Quantity Meets Quality: Towards a digital library.

By Jasper Faase & Claus Gravenhorst

(Koninklijke Bibliotheek, Netherland)

Jasper Faase
*Koninklijke Bibliotheek,
Netherland*

Jasper Faase is a historian and Project Manager Digitization at the Koninklijke Bibliotheek (National Library of the Netherlands). Since 1999 Jasper has been involved in large scale digitization projects concerning historical data. In 2008 he joined the KB as coordinator of 'Heritage of the Second World War', a digitization programme that generated the following national collections: war diaries, propaganda material and illegally printed literature. He currently heads the Databank Digital Daily Newspapers project at the Koninklijke Bibliotheek, as well as several other mass-digitization projects within the KB's digitization department.

Claus Gravenhorst
CCS Content Conversion Specialists GmbH

Claus Gravenhorst joined CCS Content Conversion Specialists GmbH in 1983, holds a diploma in Electrical Engineering (TU Braunschweig, 1983). Today he is the Director of Strategic Initiatives at CCS leading business development. For 10 years Claus was in charge of the product management of CCS products. During the METAe Project, sponsored by the European Union Framework 5, from 2000 to 2003 Claus collaborated with 16 international partners (Universities, Libraries and Research Institutions) to develop a conversion engine for books and journals. Claus was responsible for the project management, exploration and dissemination. The METAe Project was successfully completed in August 2003. Since 2003 he is engaged in Business Development and promoted docWORKS as a speaker on various international conferences and exhibitions. In 2006 Claus contributed as a co-author to "Digitalization - International Projects in Libraries and Archives", published in June 2007 by BibSpider, Berlin.

Quantity meets Quality: Towards a digital library.

Jasper Faase, Project Manager, Koninklijke Bibliotheek (National Library of the Netherlands)

Claus Gravenhorst, Director Strategic Initiatives, Content Conversion Specialists GmbH

Abstract

In 2006 the Koninklijke Bibliotheek (National Library of the Netherlands) set up a production chain to digitize 8 million newspaper pages within five years. The scale of the project meant moving from state-of-the-art digitization towards a workflow where quality needs to be balanced against quantity. A completely new workflow for mass-digitization, from selecting titles to implementing scans in the KB infrastructure and publishing digital newspapers online, had to be implemented by a heritage institution that – after nearly 15 years of working with digitization – is rapidly evolving into a true digital library. As part of this process, digitization and OCR is outsourced. Newspaper bindings are delivered to and collected from the library once a week, using special transportation boxes. The location of the physical documents and the digitization process are mapped using a barcode tracking system that also enables quality and completeness checks to be performed. In preparation for article separation, the newspapers are scanned in full colour. Layout, text and structure elements are recognized and metadata collected. The documents are then issued in METS/ALTO, MPEG21, JPEG 2000 and PDF formats, which undergo extensive quality checks before they are accepted. To process the results, the KB built online services on a new infrastructure using open source software and standard protocols (SRU, OAI, Dublin Core) to ensure accessibility. The decision to develop online services in-house means that managing and maintaining the digital newspaper databank has become part of the existing organization. This presents a huge challenge, not only for the technical infrastructure but also for the KB as a whole in terms of staff and workflow management. The KB and its supplier would like share with IFLA members valuable best practice in creating a digital newspaper library.

Digitization at the KB: A short background

The National Library of the Netherlands (KB) launched its first website in 1994, in response to growing interest in the internet. Many library staff were already familiar with microfilming and photography; the next logical step was to learn how to create digital images. The launch of the Memory of the Netherlands program in 1999 was the first milestone on the way to mass digitization. This national digitization initiative was set up to digitize visually attractive objects drawn from the collections of a wide range of Dutch heritage institutions. The KB's role in the program was to coordinate and set clear technical standards. It also provided the infrastructure for publishing the results on the website www.geheugenvannederland, which is still live today.

However, libraries deal mainly with printed matter. The primary appeal of books, magazines and papers is their content, not their visual appeal. They are also bulky. It first became apparent that a different approach would be needed in 2003, when the KB started a project to digitize the Dutch Parliamentary Papers (1814-1995). This became even clearer in 2007, when the KB received funding to set up a national newspaper database. With over 8 million pages to digitize, the sheer scope of the project meant workflow had to be more standardized than ever. Digitization was no longer just about creating a state of the art image; it had to become more text oriented. Quality of metadata was paramount, as was adopting a more time-efficient approach so as to process large numbers of pages cost-effectively.

In response to this challenge the KB developed a workflow consisting of five steps: selection, preparation, digitization, quality control and presentation. Digitization is the only step that is outsourced. The workflow has proved itself by generating an output of 300,000 pages a month, and has now been implemented as the standard way of working for the KB's Digitization Department. This will allow us to increase production of the Digitization Department from 3.5 million pages in 2010 to an expected 8 million pages in 2011. With standardized work instructions and specifications in place, our aim is to ensure that choices about balancing quantity and quality are made in the same way at every stage in the production chain or, at least, according to certain criteria. How we do this, and the practical decisions we made, are discussed below.

Selection

The selection of newspapers has proved to be a time consuming process. Newspapers are nominated for digitization by a Scientific Advisory Committee, consisting of prominent historians, linguists, journalists and representatives of other potential user groups. About 1300 national, local, regional and colonial newspapers have been shortlisted.

This was by no means the end of the selection process. Listed titles had to be located by KB staff and checked for availability. To fill gaps in our collection the KB worked together with more than 30 institutions, some of them abroad, including the Vatican Library, the British Library and the National Archives of Suriname. Titles are also screened to find out if they have already been digitized elsewhere or are available on microfilm. This process was not as straightforward as it might have been; no national registry of digitized newspapers was available when the project started. A digital workflow has been developed for titles that had in fact already been scanned, involving the conversion of existing scans and metadata. If copies are available on microfilm these will be scanned, as it is more cost effective.

Another element taken into account during selection was intellectual property rights. The KB focuses on material that can be published online at no cost and without restrictions. Under Dutch law, newspaper publishers own copyright until 70 years after publication. For freelance authors and photographers copyright extends for 70 years after death. Under Dutch law users are obliged to make a 'considerable' effort to retrieve all rightful stakeholders and gain their consent in advance of publication online. In the context of mass digitization this is an extremely time-consuming, not to mention practically impossible, task. The KB entered into negotiations with publishers and freelance photographers' and journalists' associations at an early stage in the project to resolve this problem, the aim being to reach a blanket agreement with freelancers. In November 2010 the KB also struck a deal with organizations representing copyright holders. In exchange for financial compensation, the library has been relieved from its legal duty to track down all copyright holders before material is digitized. The library also successfully concluded negotiations with a number of individual commercial publishers. This year the KB expects to digitize and publish several regional and national papers cleared for publication by freelancers and publishers up to the end of 1994.

Preparation for digitization

Before the newspapers are transported to the supplier, CCS, to be scanned, every individual issue is described in a database. Newspapers are checked page by page for irregularities that demand special attention during the digitization process. Small repairs are carried out to make the newspapers ready for scanning. Where required, the supplier is given special instructions on how to handle the material.

Sometimes it is necessary to deselect titles. The biggest issue so far has been newspaper volumes that have been too tightly bound, causing loss of text. If there are too many irregularities, and no alternative copies available, these newspapers are excluded from the selection. If only a small portion of the page is affected, text loss is taken for granted.

To ensure a continuous workflow, weekly shipments are sent off to our supplier. Every week tens of thousands of newspaper pages are prepared by a highly dedicated team in the Digitization Department. It is essential that the client and the supplier keep to agreements and plan their time carefully so as to guarantee an efficient workflow.

Digitization

Workflow Design

Before each transport KB provides the supplier with metadata for each delivery of newspaper volumes via a Sharepoint platform. Accompanying these items is a transport list specifying their barcode, title and catalog ID. As original items are of great importance and must not get lost, their progress must be tracked at every point in the process. This applies to any mass digitization project. In this project a barcode-controlled tracking system is used, as part of the docWORKS mass digitization system.

All delivered material is checked against the delivery list to ensure the batch is complete. The inspection also includes manual verification of scanability and size limitations. Any items that fail the test will be rejected.

Scanning

The digitization center has been set up some 150 km from the KB. The center is operated in 3 shifts and equipped with large-format (up to A0) overhead scanners providing high-quality scans in full color. A high-performance data center with 16 blade servers and 100 TB storage capacity performs the automated processing of the scanned images.

Immediately after the check-in procedure the preparation for scanning is made. Items are grouped and assigned to suitable scan stations according to size, fragility, type, etc. High-quality overhead and microfilm scanners are used to scan the items and create images of the best possible quality. Large-format overhead scanners are used to scan outsize pages.

Each scan station runs a scan client application, an integrated software tool that performs the scanning process. A Graphical User Interface captures all relevant data in the scanning process. This will increase performance of the overall digitization workflow. The user interface helps to verify the quality of images during the scanning process. Problems will be detected immediately. In addition to the automated checks, image quality is checked manually and ingested metadata records are manually checked and edited. It is also possible for the scan operator to insert comments. The application saves this information to the workflow database.

For automated image quality checking a color target has to be scanned on every scan station once a day. Automated splitting of double-page images (2UPs) created by the overhead scanners is also supported. The resulting single-page images are automatically uploaded to the processing servers performing image pre-processing, layout analysis, OCR, structure analysis, metadata extraction, quality assurance and output.

Conversion and Verification

If necessary, automated image pre-processing can include a variety of image improvements. The layout of page images is then analyzed to build up zone information. This process separates and tags page elements like text blocks, tables, advertisements and illustrations.

OCR is then performed on the page zones containing text, as determined by the layout analysis. The OCR engine supports a wide range of languages and recognizes standard and historical fonts. The OCR results are stored together with the word bounding box information. A user interface allows manual intervention and reviewing for extra quality assurance. OCR results are then verified by a spellchecking mechanism. Unknown words can be added to user dictionaries. Once OCR is complete, the logical structure of newspapers is determined through an automated structure analysis process. Articles, headlines, author, captions, sections and associated items are tagged automatically. The related automated generation of descriptive and structural metadata is a substantial part of the conversion process.

The centralized workflow is connected via FTP with 2 near- and off-shore locations in Romania and Cambodia where remote quality assurance of digitized objects is performed at certain stages. This process, and the related data transfer, is based on a master/slave technology controlled by the workflow system. Remote operators check the digitized newspapers for correct double-page splitting, cropping, page sequence, article segmentation, tagging of certain page elements as well as headline and caption texts.

Output and Delivery

Once documents have fully completed the conversion and quality-assurance processes, related output files are generated. The digital objects are exported in different standard formats. The most significant are image (e.g. TIFF, JPEG, JPEG 2000), PDF (alternately with bookmarks and hidden text) and XML, where the international metadata standard METS hosted and supported by the Library of Congress is primarily used. The 'structural map' of a METS object defines logical document structures like articles and sections, together with other factors. In the context of the European METAe research project, the ALTO schema (Analyzed Layout and Text Object) has been developed in order to store additional information about the physical page layout.

The METS object describes the document structure and manages all links to the page image files and the related ALTO XML files. ALTO contains all page information. Depending on the document type, all XML files are validated against the appropriate schema files, ensuring a consistent high-quality output. By using XSLT transformation other required XML output based on schemas such as MPEG21 will be created as derivates of the default METS/ALTO-object. For each newspaper issue a digital object is created including master images for long-term preservation as well as a digital copy for access and presentation purposes.

Quality assessment

Once a week the digital output is delivered to the KB in complete batches stored on hard disk drives. Delivery lists track the history of each batch. Results are checked before they are accepted. File naming and the structural metadata are checked using scripts. XML files are also automatically checked to ensure they are complete and valid. Article segmentation and headline correction are checked by taking samples at random. The images – mainly the targets – are checked semi-automatically. Each batch is checked within two weeks. If any checks fail to meet the required standard the entire batch will be rejected and returned to the supplier. One of the main issues is the time it takes to copy all files from our supplier's hard disks to our own system and then – once they are accepted – to our web and long-term-preservation environment. Although copying does not require a lot of manpower it does put immense pressure on the internal storage system.

Once batches pass the KB's acceptance procedure the corresponding volumes will be returned to the library with a subsequent weekly delivery

Presentation

In May 2010 the project launched the website http://kranten.kb.nl. The aim was to build a website providing good search facilities on a solid, scalable infrastructure. About two-thirds of the overall budget for digitization is spent on metadata and OCR. As well as a full text search, users are able to perform smart searches by specifying time span, place of publication, distribution area and/or article class (advertisement, family announcement, illustration with caption or news item). Articles are divided up into classes; because, while this is relatively inexpensive, it means huge access improvement for our future users. Search results are listed on article level and highlighted in the text. The KB chose to invest in manual headline correction because of the anticipated low quality OCR text in older newspapers with Dutch, Gothic or Fraktur fonts. More advanced functionality, such as the option of OCR correction by crowd sourcing, will be added in a second release.

The website's initial launch has been well received. So far it is attracting around 50,000 unique visitors a month. For a website of this nature, seen in the Dutch context, this is a high number. It is also promising that visitors spent an average 18 minutes a time on the site. This suggests they are using it for academic purposes or their own research rather than just briefly clicking on the site and continuing. The KB's ICT infrastructure is currently being scaled up to deal with this traffic and the enormous quantities of data generated. Storage capacity has also had to be increased considerably. In total we estimate 507 Terabyte (TB) will be needed to process all results.

Where do we go from here?

The KB's ambition is to digitize 60 million pages of books, magazines and newspapers before 2014. In addition to the digitization projects run by our own Digitization Department, the KB is working with the private sector to achieve this goal. In 2010 an agreement with Google was signed to digitize 160,000 books published between 1700 and 1900. A similar agreement with ProQuest to digitize some 50,000 books printed before 1700 has recently been concluded. At the same time the library is considering new strategies such as digitization-on-demand and an in-house digitization production chain. A great deal of effort has also been put into standardization and improvement of the digitization workflow by reducing manpower costs and spending a higher percentage of the budget on cheaper digitization.

The shift from a paper-based library to a digital library presents a huge challenge for the technical infrastructure and for the KB as a whole in terms of staff and workflow management. The KB is currently under reorganization to accommodate these changes. Digitization will be less project-based and more part of the organization as a whole. This will affect all departments. For example, staff that used to collect books for the public are now selecting books to be transported for digitization. This means our customers may have to wait longer for their requests to be processed. This also requires compromise between the quality of our services and our digitization ambitions. We believe that our ultimate goal – creating a digital library that will make our written national heritage more accessible than ever before – justifies the means.

Historical development of newspapers in India: Challenges and issues

By Dr.Mohammed.Gouse Rajuddin

(Andhra Muslim College, India)

By Dr.Mohammed.Gouse Rajuddin
Andhra Muslim College
India

Dr. Mohammed Gouse Rajuddin is a Superintendent Andhra Muslim College, Guntur, Andhra Pradesh, India, since last 20years. He has been in teaching for a period of more than 10 years. He worked as Lecturer in the Department of Library and Information Science, Acharya Nagarjuna University, Nagarjunanagar from 1994-1998 During the period he taught for M.L.I.Sc. students. He is also taking classes for Acharya Nagarjuna University Distance Education Students since 2005. He is course writer of "Information Storage and Retrieval Practicals' for BLISc and 2 papers for M.L.I.Sc courses of Distance Education Acharya Nagarjuna University. He did research on "Newspaper Libraries in Vijayawada" for M.Phil. from Andhra University, Visakhapatnam. He was awarded his Ph.D. on "Newspaper Libraries in Andhra Pradesh (India): A Study on Emergence, Services and Professional Attitude towards Automation". He obtained P.G. Diploma in Library Automation and Networks from University of Hyderabad. He attended more than 15 national and International conferences. He attended International work shop on Digitalization of Libraries (Green Stone) at Indian Institute of Management, Kozikod, Kerala during 21st to 29th November 2009. He was award with Indian Library Promoter Award during 1995 by Indian Library Association, New Delhi. He is a Life Member of several professional organizations like Indian Library Association, IATLIS, APLA, PERAL.

HISTORICAL DEVELOPMENT OF NEWSPAPERS IN INDIA :
CHALLENGES AND ISSUES

Dr. Md. Gouse Riajuddin

M.A., M.L.I.Sc., PGDLAN., M.Phil., Ph.D.

Andhra Muslim College, India

Abstract

The Indian newspaper libraries are of a special type in so far as they are reference libraries. It is not easy to give a precise definition of reference work. The main function of a newspaper library is "to act as a depository of information required by the editorial and management departments of a newspaper". The important aspect is to so organize this collection, that any information can be retrieved as quickly as possible. Information is man's life blood and his need for information is as vital as the need for blood transfusion is to the sinking patient. In Indian newspaper's editorial writing will be extremely difficult if pertinent information so urgently required by the commentator, is not collected, shifted, processed and finally communicated to editorial Writers. The techniques involved in the processing and retrieving of information are designated as Information Retrieval.

INTRODUCTION

As Mahatma Gandhi said, 'One of the objects of a News newspaper is to understand the popular feelings and give expression to it, another is to arouse among the people certain desirable sentiment and third is fearless to expose popular defects.' The term "newspaper" covers surprisingly wide range of publication of comprehensive and objective information on all aspects of the country's social economic, political and cultural life. Newspaper brings every person into touch with the active world.

Newspaper, in the present era, is the most powerful media in educating, informing as well as entertaining people. Published with a huge volume and wide variety of information, it makes itself distinct from other media. It is used regularly by people from each and every strata of the society for their current awareness. Besides, academicians and professionals from diverse fields considerably use it retrospectively for their research and/ or application activities. However, it has been observed that due to absence of an efficient and effective newspaper information retrieval system in India, the retrieval or such information is being hampered, and the multidimensional development activities have been severely affected by the same.

HISTORY OF THE PRINT MEDIA IN INDIA

For about 600 years printing has been the basic tool of mass communication, education, information and transfer of knowledge. In spite of the onslaught of the electronic media from about the second half of the last century, the print words and images are not developed of their power of communication and as such their social relevance.

The contemporary newspaper has three fundamental functions and some secondary ones.

_____

The basic once are: (a) to information its objectively about what is happening locally, nationally and internationally (b) to comment editorially on the most important news of the day in order to bring these developments into focus and (c) to provide the means whereby persons with goods and services can advertise their wares.

Some other secondary functions are: (a) to campaign for desirable civic products and to help eliminate undesirable conditions; (b) to give the readers a portion of entertainment through such devices a comic strips, and cartoons; and (c) to publish informative and analytical features often written by special columnists on issues of topical interest.

Increasing importance of newspapers:

Modern electronics, which has put a television set in almost every home in the western world, has also revolutionized the newspaper publishing process, allowing many more newspapers to be born with better production quality. An increasing number of these newspapers are given away free, their production cost being borne entirely by the revenue from advertisements, which are of much grater importance. The growing demands for newspapers owing to spread of literacy and the introduction of modern printing caused the daily circulation of newspapers to rise from the thousand to hundred of thousands and eventually to millions. Newspapers continue to shape opinions in the "global village" of today. Satellites are being used for the long distance transmission of journalistic information. The advances in information technology have brought about a great increase in the speed and timeliness of journalistic activity and at the same time provide massive new outlets and audiences for the electronically distributed products.

HISTORY OF THE INDIAN PRESS

The Communication of news through oral method was used in early days in Indian Kingdoms. The news and proclamations were reached the public though massages. Even the animals like parrots, swans and deer were used to carry news or information from one point to another. So the process of communication of news by different media was there in India. The inscriptions on walls of temples and copper plates by the kings were also played a prominent role in communication of information.

The modern art of printing in India originated in Goa in 10 and the Spanish Coadjutor, Brother john de Bustamante, known as the Indian Gutenberg, was the first printer. The first material printed by him on the Indian soil was 'Conclses' or these of philosophy.

The first language printing press with vernacular types was established in 1576 at Vaipicotta. In India, communication meant primarily for the emperor, get a fillip under the Moghuls. Aurangazeb allowed great freedom in news reporting during his period.

The East India Company's news-writers were under grate control under the Moghul Emperors. There could be no free expression of views or even free expressions of views or even free communication of news within the prevailing atmosphere of grabbing, favoritism and profiteering.

FIRST NEWSPAPER

The first attempt to start a newspaper in Calcutta was made in 1760 by William Bolts, the well-known author of consideration on India Affairs. He was forced to go back to Europe and there was no one to follow the footsteps of Bolts for about fourteen years. On 29th January 1780, James Augustus Hicky brought out the Bengal Gazette of the Calcutta. In February 1784, the Calcutta Gazette and Oriental Advertiser, a quarto size paper, was published. In the following year, Bengal Journal made its appearance. In 1785, the Oriental Magazine or Calcutta Amusement was published as the first monthly publication. The Calcutta Chronicle published in 1786. After 1790 the newspapers in India grew in number.

EARLY 19TH CENTURY

Freedom of the press was considered a danger by the British East India Company and in May 1799, press regulations were issued requiring a newspaper to carry the name of the printer, the editor and the proprietor, who were to declare themselves to the 'Secretary to the Government' for prior scrutiny. As newspapers were foundnot to be submitting to pre-censorship regularly and military information was being published in the press, there were other restrictions, which led to the growth of underground press. Raja Ram Mohan Roy destined to play a bit part in the fight for the freedom of the press during 1823 has been called the Areopagitica of Indian history.

During the later half of the 19th century, Anglo-Indian press established firm foundation in India. In 1861 the Bombay Times, the Courier, the standard and the Telegraph were amalgamated under the name of the Times of India. By 1870 the press in the Indian languages was growing rapidly. There were about sixty-two Indian languages newspapers in Bombay, about sixty in North-west Indian Language newspapers in Bombay.

20th CENTURY PRESS

Mahatma Gandhi was probably the greatest journalist of all time, and the weeklies he ran and edit were probably the greatest weeklies. Under his editorship he published in English, Tamil, and Gujarati languages. Young India and Harijan become powerful vehicles of his views on the entire subject.

The Times of India, Bombay Gazette and advocate of India, Indu Prakash, Amrut Bazar Patrika were the various newspapers of the second decade of the 20th century when the country was in ferment. The Indian press, particularly the language press fought valiantly for freedom of the country, was in ferment.

ROLE OF THE PRESS IN INDIA IN THE PRE-INDEPENDENCE ERA

Before Independence, a majority of the English and almost the entire regional language press, vigorously campaigned for the freedom of the country and faced the repressive regime bravely and with tremendous sacrifices. The newspaper offices and printing presses were sealed, journalists were jailed and sever prepublication restrictions including heavy security deposits were clamped on the press by the then British regime to suppress the freedom struggle. Undaunted, the press kept the flame of freedom burning by urging the people to do their duty by the nation and challenging the government to quit. It is also important is to note with regard to the press in these days that many of the political leaders who were active in the freedom struggle like Mahatma Gandhi, Pandit Nehru, Maulana Azad, Tilak, Lala Lajpat Rai, C.R. Das and others were themselves editing journals mainly with a activation on social, educational and religious reforms propagation, inhuman practices like child marriage were banned by law and a view to advocate the cause of freedom.

AFTER INDEPENDENCE

Some measures after Independence to help the Indian press grow in the correct directions were the appointment of the First Press Commission (1952-54) and the second press Commission (1980-1982). The recommendations made by the Commission were calculated to help the development of healthy press.

STEPPED INTO 21ST CENTURY

The number of newspaper of newspapers published in India have grown to 581 and a total circulation of all the Indian 28 states/union territories nearing 4,82,85,519. One encouraging development in recent years has been the phenomenal growth of the language press, particularly the Hindi press. But quality-wise the language press, need considerable enactment.

As per the States/Union Territory language wise breakup: The Hindi language paper has been taking highest circulation 1,34,36,914 with 218 newspapers published from 17 states, and second place was taken by English with the circulation of 92,29,113 with 90 newspapers from 17 states.   Third place taken by Telugu language 34 newspapers with a circulation of 47,95,136 published from single state.   55 Marathi newspapers have taken fourth place with 44,95,790. 18 newspapers of Malayalam with the circulation of 38,26,268 stood in 5th place in annual circulation.   The lowest circulation was in the Meghalaya state language 'Khasi' with 23,780.

STATES/UNION TERRITORY & LANGUAGE-WISE BREAK-UP OF CIRCULATION OF INS MEMBER PUBLICTIONS-DAILIES

S.No	State/UT	Assamese	Bengali	English	Gujarati	Hindi	Kannada	Khasi	Malayalam	Manipuri	Marathi	Nepali	Oriya	Punjabi	Sindhi	Tamil	Telugu	Urdu	Total
1.	Andhra Pradesh	–	–	9 / 13,97,959	–	2 / 1,37,352	–	–	–	–	–	–	–	–	–	–	34 / 47,95,136	4 / 1,22,182	49 / 65,02,629
2.	Assam	8 / 4,73,329	2 / 62,514	2 / 77,422	–	2 / 21,579	–	–	–	–	–	–	–	–	–	–	–	–	16 / 6,34,844
3.	Bihar	–	–	5 / 16,750	–	5 / 6,01,801	–	–	–	–	–	–	–	–	–	–	–	–	10 / 6,18,551
4.	Chandigarh	–	–	4 / 3,40,375	–	6 / 1,20,490	–	–	–	–	–	–	–	–	–	–	–	–	10 / 5,29,237
5.	Chhatisgarh	–	–	1	–	6 / 1,90,036	–	–	–	–	–	–	–	–	–	–	–	–	7 / 1,90,036
6.	Delhi	–	12	13,76,249	–	29 / 20,90,688	–	–	–	–	–	–	–	3 / 67,872	–	–	–	8 / 1,71,919	50 / 36,80,384
7.	Goa	–	–	5 / 1,04,539	–	1	–	–	–	–	2 / 27,352	–	–	–	–	–	–	–	3 / 1,31,890
8.	Gujarat	–	–	30,962	30 / 5,38,089	1	–	–	–	–	–	–	–	–	–	–	–	–	33 / 5,69,051
9.	Haryana	–	–	–	–	3 / 1,48,335	–	–	–	–	–	–	–	–	–	–	–	–	3 / 1,48,335
10.	Himachal Pradesh	–	–	2	–	1,09,742	–	–	–	–	–	–	–	–	–	–	–	–	13 / 1,09,742
11.	Jammu & Kashmir	–	–	7 / 7,10,088	–	95,599	–	–	–	–	–	–	–	–	–	–	–	3 / 52,881	13 / 8,58,268
12.	Jharkhand	–	–	10 / 4,17,136	–	4,17,136	–	–	–	–	–	–	–	–	–	–	–	–	10 / 4,17,136
13.	Karnataka	–	–	6 / 4,07,901	–	18	18,14,389	–	–	–	1 / 1,55,828	–	–	–	–	–	–	1 / 72,646	27 / 24,50,764
14.	Kerala	–	–	2	–	–	–	–	18 / 38,26,268	–	–	–	–	–	–	–	–	–	19 / 38,26,268
15.	Madhya Pradesh	–	–	2	–	24 / 6,92,748	–	–	–	–	18	–	–	–	–	–	–	2 / 35,448	28 / 7,28,196
16.	Maharashtra	–	–	14 / 13,03,374	5 / 91,551	17 / 3,55,727	–	–	–	–	52 / 43,32,610	–	–	–	–	–	2 / 38,222	–	91 / 66,47,334
17.	Manipur	–	–	1	–	–	–	–	–	3 / 32,902	–	–	–	–	–	–	–	–	4 / 32,902
18.	Meghalaya	–	–	20,473	–	–	–	1 / 23,780	–	–	–	–	–	–	–	–	–	–	2 / 32,902
19.	Mizoram	–	–	59,218	–	–	–	–	–	–	–	–	–	–	–	–	–	–	59,218
20.	Nagaland	–	1	40,235	–	–	–	–	–	–	–	–	–	–	–	–	–	–	40,235
21.	Orissa	–	–	1	–	–	–	–	–	–	–	–	10 / 14,18,225	–	–	–	–	–	11 / 14,18,225
22.	Punjab	–	–	5 / 7,07,088	–	5	–	–	–	–	–	–	–	6 / 8,52,560	–	–	–	–	12 / 14,18,225
23.	Rajasthan	–	–	8	–	39 / 35,80,934	–	–	–	–	–	–	–	–	–	–	–	1 / 16,155	40 / 15,75,803
24.	Tamil Nadu	–	–	8 / 19,70,573	–	–	–	–	–	–	–	–	–	–	–	26 / 31,79,124	–	–	34 / 51,49,697
25.	Tripura	–	5 / 2,11,269	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	5 / 2,11,269
26.	Uttar Pradesh	–	–	51 / 4,71,033	–	29,74,073	–	–	–	–	–	–	–	–	–	–	–	5 / 2,30,061	62 / 37,02,711
27.	Uttaranchal	–	–	1 / 56,249	–	5 / 3,67,948	–	–	–	–	–	–	–	–	–	–	–	–	6 / 4,24,197
28.	West Bengal	–	13 / 27,77,858	5 / 8,44,949	–	8 / 2,75,938	–	–	–	–	–	–	–	–	–	–	–	3 / 1,24,665	31 / 40,23,410
	TOTAL	8 / 4,73,329	20 / 30,51,641	90 / 92,29,113	35 / 6,39,640	218 / 1,34,36,914	19 / 18,40,239	1 / 23,780	18 / 38,26,268	3 / 32,902	55 / 44,95,790	1 / –	10 / 14,18,225	12 / 9,89,239	–	27 / 31,79,124	34 / 47,95,136	29 / 8,64,179	581 / 4,82,85,519

CHALLENGES

Indian Newspapers have performed invaluable democratic functions somewhat better than the other news media since Independence. Print media successfully faced several challenges in the past. Radio was perceived as threat when it arrived, similar was the case when television came and arrival of internet cane with such hype that the biggest media company at the turn of country. Google was launched in 1997, skyp, 3G phones, and blogging are all products of the new millennium. Print media along with other initiated Web presence to be on safer side.   Before the web, storytelling was perform specific, Newspapers and magazines focused on text and photos, radio told stories with audio, and television dealt with moving pictures and sounds. The Web forces these platforms to integrate, Today's best media Web, sites are multimedia productions combining text, stills, audio and video.   The Web's advantages are speed and efficiency of delivery, personal selectivity, availability on demand, and the almost cost-free capability to store additional content and accept additional users. Print favors portability, the ability of reading to scan full page or portions of pages and spot headlines and stories that they would not have expected to be interested in. The merging technology and the 24-hours news cycle have only increased the amount of news available to consumers. And some believe that the abundance of information will only decrease the importance of traditional media.

According to a study released in March 2006, by the pew Internet & American Life Projects,   "more than 50 million Indians per day used the Internet as their primary news source in 2005. That's up from27million in 2002. In fact, checking the daily news is the third most popular activity on the Internet, the study found.

On August 16, 2007 the Shorenstein Center on the Press, Political and public Policy at Harvard University released a report that examines rends in Internet news. The study looked at the traffic of 160 news sites for a year and found that the sites of national, "name brand" papers were growing, while the traffic on smaller, local papers was not connected to traditional media were growing more slowly than non-traditional media, such as aggregators, bloggers, search engines and service providers. The study suggests that Web is a threat to print media, whose advantage of being a "first mover" is influx. The study also revealed that even though the web trends of favor "brand name" news sites are a brand name in their community and should use it to their advantage. News aggregators such as Google News and Yahoo News are cited as threats to traditional news organizations. Their serious and successful forays into news are taking readers away from news sites. Other online ventures, such as users-generated news and interactivity are also taking a slice of the news audience.

INFORMATION TECHNOLOGY AND THE PRINT MEDIA

Information Age: Due to electronic explosion, a digital revolution, an information super highway, or a grand convergence of media technologies, something is happening to our world that is profound and unprecedented. The information technology revolution has really arrived in India. And this has changed entire among other things, the work environment in newspaper production and distribution.

Introduction to Internet: The Internet is a computer and satellite based world-wide information network, and is composed of a large number of smaller interconnected networks called intranets. These Intranets may connect tens, hundreds, or thousands of computerized data banks, enabling them to share information with each other. The Internet thus, is a network of networks of data banks.

The Internet has made it possible for people all over the world to effectively and inexpensively communicate with each other. Unlike traditional broadcasting media, such as radio and television, the Internet is a decentralized system, each connected individual can communicate with anyone else on the Internet, can publish ideas, and can sell products with a minimum overhead cost. In the very near future, the Internet may have a dramatic impact on higher education and business as more universities offer courses and more companies offer goods and services online.

Information Technology and the Media:

The Indian media, particularly the print media has not been prompt enough to avail of the technology revolution. Though many newspapers are using computers to avail of the opportunities opened by desk-top publishing system, most of the major dailies have also opened up their web pages.

Internationally, remote publishing started with the growth of satellite communication with major Indian newspapers publishing web-edition by transmitting all over the world.

Internet and the Print Media:

Indian newspaper readership may be affected when more and more continuously updated web newspaper become available round the clock. If the convergence technology is able to combine internet and television features in a cost effective way, the print media would be compelled to find new ways of survival because they would be under attack from both advertising companies such as the internet auction houses, which are usurping the classified business, and on the editorial side from online news services.

Every time a new medium has come along, the death of newspapers has been announced yet newspapers are still with us. In most of the world's developed countries, newspaper revenues are at record level.

Circulation of Indian newspapers is on decline too. Population growth has helped newspaper sale, but circulation started falling more than a decade. According to a National Readership Survey in India:

1. 49% of **620 million** adults in India read newspapers and magazines.
2. T.V. has entered 269 million Indian homes, **237 million** Indians are hooked on the internet. While the Internet expands, the print media is strategically increasing local coverage and going closer to grass roots.

When the Internet first engaged the attention of newspaper industry, it looked like a new, cheap distribution medium. Indian newspapers began to put their copies on sites that looked like replicas of their front pages. They were worried about issues such as whether to change for access to their cities, whether to put up all their content, for fear of seeing their print sales cannibalized, whether to increase their staff to update their copy more often, how to charge for advertising. The Internet, however, will prove more than a new distribution channel for news. It is most likely going to undermine the economic that underpins the newspaper business.

INFORMATION TECHNOLOGY AND CONVERGENCE

The emergence of wireless technology has been the most talked about advance in recent times. Connecting up personal Computers, Cell phones and Personal Digital Assistants to the power of internet, it opens up the boor to immense possibilities. That brings us to convergence a multiple technology kit was demonstrated called mobile journalists workstation, the kit contains a global positioning system and a hand held device. It also packages an internet access wireless system.   These multiple technology systems were a result of convergence technologies.

NEWSPAPER NEXT

The newspaper industry is at a strategic inflection point- a period of disruptive changes that threaten its current way of doing business with no clear future path. The threaten come from many directions but are manifesting themselves in the form of declining circulation, rising costs and downward revenue pressure. These trends show no sign of reversing themselves. The industry's very survival is depend on its ability to reframe completely the way it does business, and find new way to attract and keep customers. Many industries thought history have reached strategic inflection points;   not every industry has weathered them successfully. This project's goal is to ensure that newspapers survive these disruptive times.

In exchange, these innovations introduce new benefits along dimensions such as simplicity, convenience, ease of use or low price. In the media industry, blogs, Google, eBay, Monster.com and free distributed computer papers all fit the pattern of descriptive innovation.

CONCLUSIONS

As a new information and communication medium, it is creating its own space while adapting itself to other media. In India the access of Interment for the masses would be hampered initially by the low purchasing power of people and the lower literacy rate. The greatest contribution of the Internet to the print media is that people no longer run after information, but the information is there, available on their screen at the click of the mouse, and the quantity of information is really impressive, though some time has to be spent in locating the information that would be relevant. As far as the press in India is concerned, there is no doubt that all the sections of the press, including the small and medium newspaper, will benefit from the advances in information technology. Since updated news can be received instantaneously over the internet at a very minimal cost and can be processed speedily by the newspaper using the latest information technology tools, there is bound to be a qualitative improvement of the small and medium newspapers.

REFERENCES:

Amitab Chatarjee. 1977 Libraries in the Cities' Mass Media Units in S.K. Kapoor ed. by Library Science in Calcutta. Calcutta: IASLIC, 1977, p.152.

Ashok Babu, T (2000) Vision of Future Library and information Systems,p.205-209.

Bhagavati Devi (2007) Information Security for libraries, p201-202.

Colley , Veronica (1987), The Preservation of Newspaper Clippings, Special Libraries Associations, p.41-42.

Focus on Information Technology: Mass media in India 2008., 36-38p. Publications Division New Delhi.

Gouse Riajuddin , Md. (1996). Newspaper Libraries inVijayawada – A Survey, p.4-6.

Gouse Riajuddin, Md. (2008) Newspaper Libraries in Andhra Pradesh(India) : A study on Emergence, Services and professional Attitude to words Automation, p.95-97.

Gregory, B. Newby (2000) Information Security for Libraries,p.3-5.

Hema, C. (2010). "Electronic Information Access- Role of user interface Innovation –Driven Librarianship".

Lou Chomas and Anne B. Jennings. 1983. Starting the Library in guidelines for newspaper libraries (restone American Newspaper Publishers Association Foundation). p.5

Morna ,T.P. "The Command Language Grammar: A Representation for the user Interface of Interactive Computer Systems". International Journal of Man-Machine Studies, in press hand book 2009-2010, p.58.

Neal, F. Austin, 1950. The Newspaper Library: The results of a survey completed in June 1949. Special Libraries. 41 (46).

Ram S. Kolatkar, 1987. Findings the facts. New Delhi: Press Institute of India. 1987, p.80

Ram S. Kolatkar, op. cit, p.81

Rangaswamy Parthasaradhy, 1984. Basic journalism, Dellhi: Macmillan, 1984, p.94

NewspaperSG: National Library Singapore's experience in the provision of an Online Singapore News Archive

By Judy Ng

(National Library Board, Singapore)

Judy Ng
National Library Board
Singapore

Mrs Judy Ng has been with the National Library Board for the last 25 years. She had served in different postings in the Library and has experience in running Public Libraries as well as Reference Library. Currently, she is overseeing the Lee Kong Chian Reference Library and 13 special libraries. She has led staff in the development of a suite of projects that enhances the reference services and facilitates easier access to reference content. She also play a leading role in the development of the NewspaperSG service and is closely involved in the negotiation for access to copyrighted news content.

Paper to be presented at the **IFLA INTERNATIONAL NEWSPAPER CONFERENCE 2011
(25TH – 27TH APRIL 2011)**

**NewspaperSG: National Library Singapore's experience in the provision of an Online
Singapore News Archive**

Judy Ng

Director, National Reference and Special Libraries
National Library Board, Singapore

Abstract:

The National Library Singapore (NLS) has a collection of over 200 current and historical
Singapore and Malaya newspapers. This collection resides at the Lee Kong Chian
Reference Library and serves as an invaluable resource for many users - from students to
individuals doing research on Singapore history.

These newspapers are accessible through the microfilm collection that numbers around
20,000 reels. In search for a more effective and efficient way of accessing and researching
with newspapers, as well as to extend to a wider user base, NLB began exploring
newspaper digitisation in 2005, and the development of a news portal in 2007. This
eventually led to the official launch of the NewspaperSG service in 2010.

In 2007, NLS sealed a landmark agreement with the local newspaper publisher, the
Singapore Press Holdings (SPH) on the digitisation of copyrighted local English paper, *The
Straits Times* (ST). The agreement allows NLS to digitise the ST, from the first issue of the
paper in 1845 to issues in 2006. For issues after 2007, the digital (electronic) edition of the
printed ST is deposited with the National Library. In 2009, a second agreement was signed
with SPH to allow NLS to digitise its other publications including the Chinese, Malay and
Tamil newspapers. Another agreement was also signed with Mediacorp Press Ltd to allow
access to the digital copies of the TODAY newspaper.

Digitisation of the ST for the period 1845 – 2006 has been completed and is now accessible
through the NewspaperSG portal, (http://newspapers.nl.sg), together with 20 other
newspaper titles. NLS is now working on the digitisation of the Chinese and Malay
newspapers which are scheduled for release in the second quarter of 2011.

This paper will share NLS's experience in developing the NewspaperSG service, particularly
in the following areas:

- Public-private partnership between the National Library Singapore, a public institution
  and commercial entities to deliver copyrighted content for both onsite and home
  library users.

- The innovation to bring about greater access and discovery of the newspaper content

- Latest updates on the project, including the digitisation of the non-English
  newspapers.

Introduction

1. The National Library Singapore (NLS) has a collection of over 200 current and historical Singapore and Malaya newspapers. This collection resides at the Lee Kong Chian Reference Library and serves as an invaluable resource for many users - from students to individuals doing research on Singapore history.

2. These newspapers are accessible through the microfilm collection that numbers around 20,000 reels. In search for a more effective and efficient way of accessing and researching with newspapers, as well as to extend to a wider user base, NLB began exploring newspaper digitisation in 2005, and the development of a news portal in 2007. This eventually led to the official launch of the NewspaperSG service in 2010.

3. This paper will share NLS's experience in developing the NewspaperSG service, particularly in the following areas:

    -   Public-private partnership between the National Library Singapore, a public institution and commercial entities to deliver copyrighted content for both onsite and home library users.
    -   The innovation to bring about greater access and discovery of the newspaper content
    -   Latest updates on the project, including the digitisation of the non-English newspapers.

Public-Private partnership

4. Due to copyright restrictions, most libraries in the world focus on digitizing historical newspapers in the public domain. The Singapore Press Holdings (SPH), the publisher of *The Straits Times* has a long print run of 165 years of history and is still running. NLS took a bold approach by directly negotiating for permission from SPH to digitise and to provide access to all the issues of the paper.

5. NLS started negotiation with SPH in 2005. After two years of negotiation, SPH agreed to grant the NLS the rights to digitise The Straits Time archives from its first issue in 1845 to 2006. In addition, SPH provided born-digital copies of the paper from 2007. NLS was also given the rights to provide onsite access to the digitised newspaper from the computers in all of the National Library Board's (NLB) national and public libraries. The team started to digitise the content in 2007 and at the same time, work on developing a newspaper portal called NewspaperSG. The collaboration is a win-win collaboration for both parties as SPH gets to preserve their newspaper content whilst NLS gains rights to provide access to the proprietary newspaper content.

6. Following the successful implementation of phase one of the service that allows only on-site access, NLB negotiated for a second agreement to allow the content to be accessed remotely. To obtain the agreement from SPH for content to be accessed remotely, NLS needs to balance the commercial interest of the publisher with the library's push for convenience access. After several rounds of negotiation, a second agreement was sealed in July 2009 to allow for all copyrighted content, published before 1990, to be accessed and viewed from outside the library. SPH further allows the digitization and access to 30 other newspapers and publications

that it owns, including newspapers in the Chinese, Malay and Tamil language. This was the second breakthrough.

7. Levaraging on the successful completion of phase one and using the same model of negotiation, NLS managed to seek the permission from another newspaper publisher, Mediacorp Press Ltd to allow access to digital copies of the TODAY newspaper on the NewspaperSG portal.

Service: Design or approach

8. The service was designed to allow easy discovery of content and sharing. At the same time, overcoming the challenge to protect the intellectual property rights of the publisher.

Managing Rights and Intellectual Property Protection (IPR)

9. Protecting the intellectual property rights of the newspaper publishers is paramount to the success of the service.

10. NewspaperSG has two access levels: onsite access using the libraries' computers and offsite access.

- For onsite access, users can search, view and print the full content of NewspaperSG only through the libraries' computers.

- Offsite users will also be able to search the full content, but access to the full article is limited to publishers' content published prior 1990. Users are not be allowed to print, copy or download the content thou they are free to view. To achieve this, a cost-effective combination of IP (Internet Protocol) authentication, date of publication, watermarking (offsite access) was used.
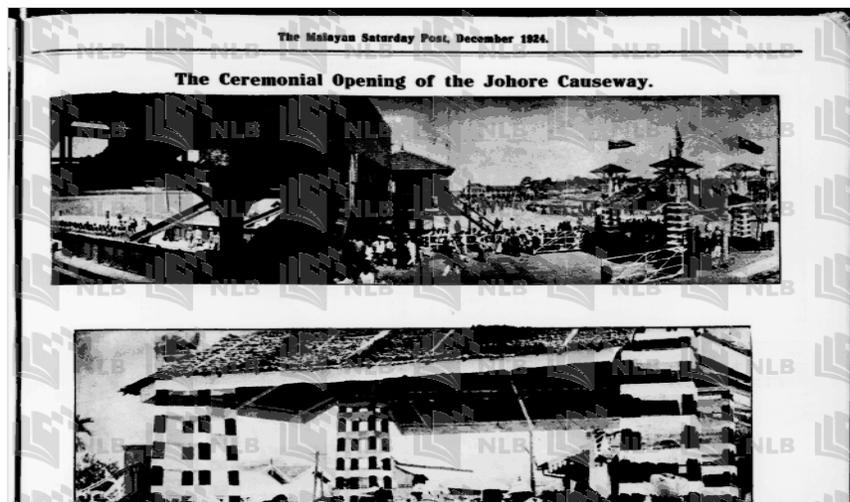
Remote access rights management (watermarking)

<u>Content Discovery</u>

11.   Most people use the search engines as the first point of search when looking for information. To allow ubiquitous discovery of the newspaper content, NLB leverages on the Internet search engines such as Google to widen the search-capability offered by the service.

12.   A "Table of Contents" of each newspaper issue, consisting of article headlines, dates of publication, page numbers and 50-word extract is generated and submitted to search engines for crawling.

13.   Offsite users are be able to see information such as the article title, date of publication, page number and a 50-word extract from the article in their search results. With this approach, users who may not visit the library website would still be able to discover newspaper content.

14.   It was found that 60% of users discovered the content in NewspaperSG through search engines, while 40% found the content by going to the URL directly or from bookmarks.

Search capabilities and effective UI

15.   NewspaperSG offers users keyword searching on the OCR text of the newspaper content. Advanced search capabilities include faceted search. With the implementation of faceted search, users are able to narrow their search in a more effective way. For example, users can narrow their search for articles by year, newspaper title and content type. To facilitate the viewing of the articles, features such as zoning, zooming, keywords highlights have been incorporated into the design of the UI.
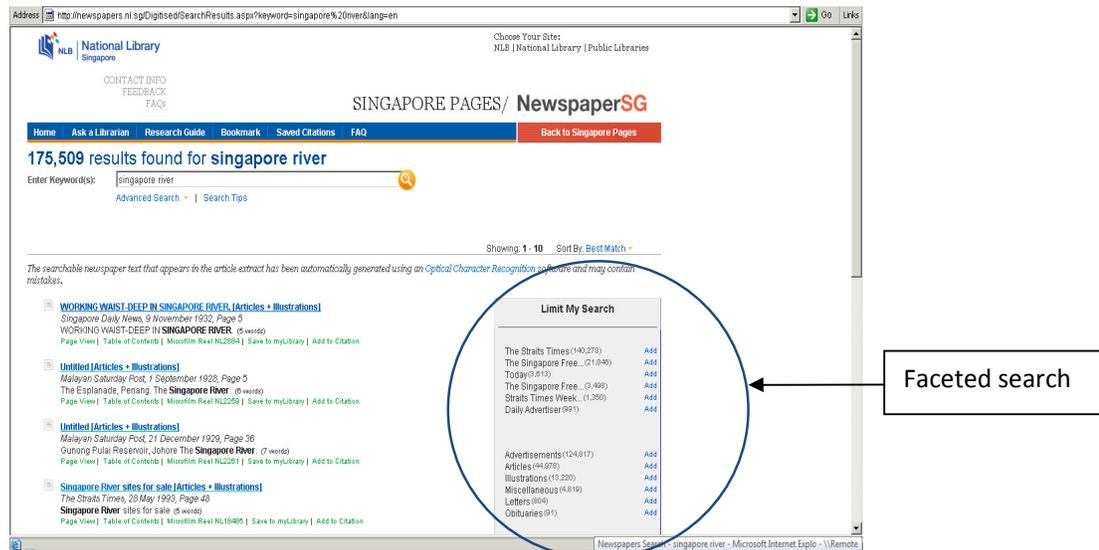
Faceted search

Saving results and sharing through social networking tools

4

16.   As users cannot download certain materials due to copyright issues, the following features were implemented:

- a citation saving feature was incorporated to enable users to email and print the citations for future reference.
- social media/bookmarking tools to facilitate knowledge sharing were also implemented. Users can share the weblinks of NewspaperSG content through social media platforms such as Facebook or MySpace.

Saving Citations

E-mailing citations

Sharing an article link

<u>myLibrary Facebook application at NewspaperSG</u>

17.   In addition to the citation feature, a complementary myLibrary Facebook application at NewspaperSG was developed. It allows users to save the links to articles on their Facebook accounts.

myLibrary Facebook application

Value of the NewspaperSG Service

18.   The NewspaperSG service has significantly improved research on newspaper content. Users can now search large extant of Singapore newspaper archive in the comfort of their home, instead of the tedious browsing of the newspaper content through the microfilms. The search feature of NewspaperSG has allowed users to discover content that they would not have done so if they use the microfilms.

19.   Usage of the library's newspaper archive has increased exponentially. Since its launch, usage of NewspaperSG has grown exponentially. The site has now an average of 50,000 unique visits and 400,000 page views a month.
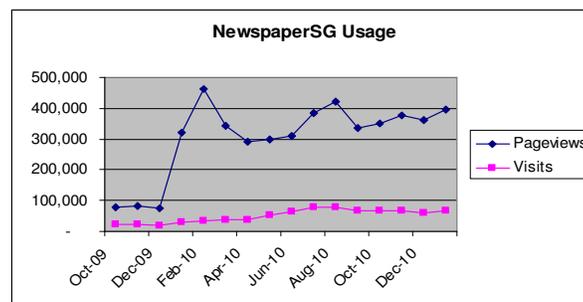
20.   The website is a boon to family history researchers and genealogists as well as academics who research on Southeast Asian history and society. Besides Singapore, a sizeable traffic comes from the United States, Australia, Netherlands and the United Kingdom, countries that Singapore has economic as well as historical links.

21.   The NewspaperSG service has received many compliments from users and some samples of the compliments are given in appendix A. [1]

---

1.   Anuar, Mazelan. (2010) *NewspaperSG: Global Access to an Online Archive of Singapore's Newspapers*. Paper presented at the The Philippine Librarians Association, Inc. (PLAI) National Congress on the theme "INTERNATIONAL LIBRARIANSHIP: A PATH AHEAD on November 17-19.

22. In recognition for its innovation, NewspaperSG had received awards at the Ministry as well as National Level. It was awarded the Gold award winner for the Ministry of Information, Communication and the Arts (MICA) Innovation Fiesta 2010 and Best PS21 Project, Gold award winner for the nation-wide Public Service 21 ExCEL Award 2010.

Next Phase of development

23. We are now moving into the next phase of NewspaperSG where digitised content of the Chinese, Malay and Tamil Papers will be available. A multi-lingual interface will be released together with the new content. A keyword translation feature will also be available to facilitate search in multi-language.

A major Chinese newspaper, Lianhe Zaobao

Multilingual interface

Key word translation

Conclusion

24. NLB is one of the few libraries in the world to make copyrighted newspaper content available to offsite users and internet search engines. The private-public partnership between NLS and Singapore newspaper publishers has revolutionised research on newspaper content. The NewspaperSG service has certainly brought value and greater convenience to our users.

Appendix 1: Users' compliments.

a) With the new National Library Straits Times search engine, doing research like this has just been completely revolutionised!

b) This is great news for us! It certainly saves the tears and the backache from sitting through hours' of scrolling through microfilms in the icy-cold rooms at the library. Kudos to the library's digitalization efforts!

c) This is pretty exciting. With the new NewspaperSG, online search can be easily done.

d) I gave it a try and the search function was quite good…this will be of great assistance when researching …as we can pinpoint the date of the article, which will also result in significant time savings when flipping through the pages.

e) The digitised newspaper service is the most fantastic advancement that the Library has done.

f) I am very excited that the Straits Times digitised newspapers are now online. I have found some matches for my ancestors and their vessels who sailed to Singapore in the mid-19th century…

g) I was reading The Straits Times at 7am on Wednesday morning when I saw the announcement of this new service. Within hours I had search through my wife's family tree and found references to many people (engagement / wedding notices). By 11:30am on that morning I was walking out of the NL with copies of all the articles. In a few hours I had been able to find a wealth of information on my wife's family tree that would have taken weeks if at all been possible before this service.

h) It is a high-quality product that enables researchers such as me to do the research very quickly and over a longer period than ever before... I have found in 2 days through your admirable searching facilities would have taken me weeks, if not months, implementing the normal methods by going through microfilms. It also prevent headaches because going through microfilms is not a particularly healthy activity for the eye muscles, with severe headaches as a normal result.

i) The article in the Malayan Saturday Post April 16 1932 titled "H.E The Governor Presents the King's Police Medals" appeared on screen and all is well... Singapore is my home City where I was born in 1926 and A.H.Dickinson shown in the picture is my father! Regrettably, as much as I would like to, I have not been able to return to Singapore. Perhaps one day if I live long enough, I will!

The selection of newspapers for digitization – a look on the multi-faceted decision process in Germany

By Ulrich Hagenah M.A.

(State and University Library Hamburg, Germany)

Ulrich Hagenah M.A.

State and University Library Hamburg

Germany

Ulrich was born on 1956 in Lüneburg (NW Germany). He studied of history, journalism and communications research, political science in Mainz and Munich, 1975-1984. His special scientific interest: history of the press (18th to early 20th century), history of communication in politics and administration. He is scientific librarian since 1988, since 1990 at Hamburg State and University Library head of the Hamburg department incl. the legal deposit of the Free and Hanseatic City of Hamburg, regional bibliography, special reading room, project management (cataloguing, exhibitions, microfilming, digitization, deacidification). He is responsible for the newspaper section of the library and for several digitization projects concerning regional periodicals, maps, books and other kinds of printed material. He is also responsible for all aspects of physical conservation of the stock and has experience in newspaper microfilming since 1992. He is a member of the executive board of Mikrofilmarchiv der Deutschsprachigen Presse e.V., Dortmund (microfilm archive of print media in German language). He is a member of the AG Zeitungen im Forum Zeitschriften GeSIG e.V. (working group newspapers in the forum newspapers GeSIG e.V.) and member of the IFLA Standing Committee on Newspapers since 2009.

The Selection of Newspapers for Digitisation – a Look at the Multi-Faceted Decision Process in Germany

Ulrich Hagenah (State and University Library Hamburg, Germany)

Abstract

In Germany no systematic funding of the digitization of historic newspapers has been established yet. Librarians, archivists and historians of the press agree that foremost criteria for selection of newspaper titles for digitization should be their information value regarding the political, economic and cultural history of the state or a region. More refined criteria have to be applied to the rich German newspaper publishing world of the past. Current digitization projects and policies of state and regional libraries follow a wide variety of determinants and practical reasons which are taken under consideration in this paper, in comparison to other European countries.From the point of view of historical communication research the paper also discusses some criteria not to be neglected, related to the typology of the European press, such as the representation of confession, political parties, local history, minorities, or the economic basis, administrative function and competitive situation of newspapers. In the selection process newspapers will have to be carefully and responsibly assessed in respect thereof, as digitization policies will inevitably exert noticeable influence on future research and on the service quality of the libraries.

Ladies and Gentlemen, dear Colleagues,

We all know: digitising historical newspapers is expensive – we cannot do all at once, we have to select and set priorities. The first problem is the mere mass of material we have to cope with. The second problem lies in the complexity of automated text conversion and indexing techniques in case we do not settle for producing millions of images, but strive for a comfortable searchability of newspaper content. This multiplies the cost of mere image production. And optical character recognition and automatic interpretation and indexing of text are additionally aggravated because – like in the Nordic countries - our German newspapers have mostly been printed in Gothic Fraktur for centuries and used Roman letters only in the 20th century.

This means, our digitising activities are heavily constrained by our budgets, especially in the field of newspapers – we cannot digitise them all, or: we cannot digitise all of them at the highest standards; or: we cannot digitise and offer them to the public under conditions of open access, but open access is one the fundamental preconditions of digitisation in German regional and academic libraries, at any rate when it was achieved with public funding.

We have to decide, according to our budget resource

- which newspaper titles should be digitised at all

- with which priority

- and to which degree, at which level all titles or certain segments of them should be indexed.

Decisions like these have to be responsibly taken after comprehensive consideration of the historical publication market. Transparency and scholarly reliability of our arguments and criteria seem to me most important in this field, because by digitising certain newspapers we privilege them for scientific use and the common public. We favour them above others that can be used only locally by microfilm or the original or possibly not at all – for many years! German scholars in historical communication research talk of 'distortion by availability' which we have to be very clear about when arguing with our users.

1 Newspaper digitisation in Germany

Before discussing several selection criteria in more detail in the second part of my talk, let me first give you a short sketch of the situation in Germany concerning newspaper digitising. We cannot discover too much progress since Klaus Ceynowa from the Bavarian State Library in Munich made his statements about the deplorable situation in 2008. There are still no reasons for optimism at the moment, because we have, on the one hand, very few good results and achievements, on the other hand there is a deplorable time-lag compared with Scandinavia, Western Europe and the rest of Central Europe, or the American and Australian programmes. Still, as we haven't really moved forward up to now, it makes sense to discuss reasonable selection criteria – and later on apply them properly.

Some figures:

I don't want to sum up a number of newspaper titles from the first German newspaper in 1605 to 1945 because we all know how misleading such figures can be – because there are too many ephemeral papers among them. Just to give you an idea of the dimension let us look at the number of newspapers existing simultaneously in certain years:

In 1648 the Holy Roman Empire of German Nation consisted of about 300 entities (principalities, free cities).

1610/20: 20

1650: 41

1700: 60-80 (Hamburg, 17th century: 11, up to 5 simultaneously)

1750/55: 90-105 (+ 60 "Intelligenzblätter" = advertising papers)

1795/1800: 210-250 (+ 170 "Intelligenzblätter" = advertising papers)

After the end of the Holy Roman Empire 1803/06 and the Congress of Vienna 1815 the Federation of German States (Deutscher Bund) consisted of 35 sovereign states and 4 free cities.

1847: 1.000

1871 Deutsches Reich consisted of 22 monarchic states, 3 free cities and Reichsland Elsass-Lothringen
1918 Republic: 24 states → 1925: 18 states

1885: 3.100

1900: 3.500

1914: 4.221

1925: 3.100

1932: 3.800

1944: 1.000

The highest total circulation is estimated for 1906: 25,5 million copies.

In the year 1914: 4.221 newspapers appeared in 2.321 cities, out of which

1.427 cities   had 1 newspaper

571 cities   had 2 newspapers

323 cities   had 3 or more newspapers

In the year 1927 out of those 1-paper-cities

197   had less than 2.000 inhabitants

532   had 2.000-4.000 inhabitants

578   had 4.000-20.000 inhabitants

41   had 20.000-50.000 inhabitants

10   had more than 50.000 inhabitants

In the years 1921/26 party papers (altogether about 1.500-1.600) appeared as follows:

320	liberal-bourgeois
600	conservative-nationalist
364-451	catholic, centristic
170	socialist, social democratic
15-30	communist
20	national-socialist

Approximately the same number of papers called themselves politically neutral, independent or non-party.[1] You see, the German newspaper landscape is very rich, its structure is polycentric, and it contains papers of diverse shapes and types, from national to local levels.

Some facts:

- Historical German newspapers are held by libraries and archives in equal shares, some archives of publishing houses added to state and municipal archives.
- Most of the newspaper holdings of the libraries – at least the academic and regional libraries – are part of the central German database for serial titles (Zeitschriftendatenbank), but not all of them.
- Newspaper collections of the archives are lacking in that central register almost completely. We know of them through old and selective printed catalogs on national level, through detailed regional inventories comprising all holdings in their area of reference and through the register of microfilms of the Microfilm Archive of the Press in German Language in Dortmund.
- The German National Library (Deutsche Bücherei, Leipzig) never got envolved in this field because newspapers weren't part of its collection profile since its foundation in 1913. The Deutsche Bibliothek (Frankfurt, founded in 1947) confined its newspaper collection to a selection of papers on microfilm.
- 45-50 % of the historical German newspapers were available on microfilm in the mid 1990s – maybe a bit more nowadays.

Regarding the digitisation of newspapers …

---

[1] Jürgen Wilke: Grundzüge der Medien und Kommunikationsgeschichte. Von den Anfängen bis ins 20. Jahrhundert. Köln / Weimar / Wien 2000, p. 52, 79, 120f., 189, 260f., 345. – Rudolf Stöber: Deutsche Pressegeschichte. Einführung, Systematik, Glossar. Konstanz 2000, p. 67ff., 75, 146ff., 155, 213ff., 220ff., 225, 230, 233ff.

- there is still no systematic funding policy of the German Research Foundation – DFG, Germany's largest research funding organisation -,
- and – what is at least as detrimental – there is no central planning board, no competence centre in terms of newspaper digitisation and history of the press in Germany.
- The German Research Foundation is funding single projects of national importance – I will get back to them later -, but hasn't set up a special programme for newspaper digitisation comparable to the one for microfilming in the 1970s to 90s. Thus there is no central coordination and strategic planning in the field of newspaper digitisation.
- I already mentioned the German National Libraries of the 20th century. You may know that Germany does not have one central historical National Library dating back before 1913. As a compensation, a financing scheme has led to the creation of a virtual national library in which six participating libraries are responsible for different historical periods. The "collection of German imprints from the 15th to the 20th centuries" has been a cooperation for more than two decades now. The new goal of this project after 20 years of collecting, cataloguing and research is the comprehensive digitisation of historical books – but: periodicals have completely been left out from the very beginning, the programme has concentrated on monographs, and this continues in respect of digitisation.
- So it's up to the state and regional libraries with historical legal deposit holdings to include historical newspapers into their digitisation programmes – what happens in about half of the 16 German states, but by far not in all regional libraries, as far as I can see.
- The basic problem is: digitisation planning for newspapers only happens inside the federal states ('laender'), and we have 16 of them. There is no accepted leadership and competence centre in terms of newspaper digitisation and history of the press.

Some good news:

I. The Berlin State Library has digitised - funded by the German Research Foundation, the Prussian Official Journals (1863-1894, Provinzial-Correspondenz 1863-1884 and Neueste Mittheilungen 1882-1894), politically most influential papers directing the contents of the smaller press – and, additionally, one of these provincial papers, the Teltower Kreisblatt, was digitised as a sort of case study. The whole project was thoroughly accompanied by historical research.

- is currently running a project to digitise three of the most important papers of the former German Democratic Republic
- and the famous Vossische Zeitung from Berlin, 1918-1934, published as a de Gruyter research database at enormous licencing fees (€ 27.390 for 127.000 pages), however, the first example of this business concept in the field of newspaper digitisation in Germany.

II. The Bavarian State Library in Munich

The library's enormous digitising efforts – you may know of their cooperation with Google books – have not included newspapers yet, but they support lots of projects in the region (Coburg, Augsburg, Passau, Regensburg, smaller cities) and present them on their platform for historic source material, the Bavarian Regional Library Online.

What I find most remarkable is their systematic preparation of digitisation activities, inside Germany a unique effort to make the whole newspaper landscape visible:

--- starting with the discovery and investigation into all titles of the regional, local and official press in Bavaria

--- cataloguing of all holdings of libraries and archives in the central German database for serial titles (Zeitschriftendatenbank)

--- followed by the identification of the most important regional publishing centres dominating the print culture of their surroundings

--- finally setting up a scheme for digitising a selection of papers on three levels of priority, and, accordingly, three funding options on national, regional and local levels.

III. Finally the Regional libraries:

Without coordination among each other and approaching the task most pragmatically about half of the German regional libraries and some university libraries as well have digitised newspapers of their area, throughout image versions without OCR or layout analysis of any kind. The selection of regional or local papers follows pragmatically their microfilm holdings, opportunities of public-private-partnerships or cooperative relations with historical societies, archives or publishers. Very often relations between regional libraries and archives are quite close. Many of them became accustomed to cooperation and complementary completion of newspaper series for preservation microfilming over the years.

12 regional libraries and 3 university libraries provide about 180 digital historical newspapers, altogether 2.300 annual volumes, if we include those regional titles partially digitised on demand by the Berlin State Library.

I cannot enter into more detail here. Instead, I would like to ask: How do we, how can we meet the expectations of our public – and of which public? And how does this influence our digitisation policies?

The National Digital Newspaper Program (NDNP) of the Library of Congress claims: „the foremost reason for selection of an individual title for digitisation is that the newspaper significantly reflects the political, economic and cultural history of the State. These include, but are not limited to, titles of statewide or regional influence and titles important as a source of information about a variety of ethnic, racial, political, economic, religious, or other special audiences or interest groups." I agree with these demands and criteria, but it is also obvious that they address very different requirements. And they do not refer to our users' needs.

I suppose, we have to be aware of quite a variety of imaginable use cases, and that we have to concentrate on different priorities according to these different expectations of users:
1. private individuals, historians, journalists, documentalists etc. who search for single data, events, persons, trying to verify facts
2. intensive studies of a certain period of time or a village or city quarter, the interest spreading wide over all kinds of topics and all what happened
3. historical communications research, historians of the press studying single newspapers, journalists or publishers, the state of the newspaper market at a certain time
4. historians of public communication, sociology or political philosophy working in the field of historical semantics or discourse analysis; they expect the searchability of topics or certain concepts through long periods of time, in large volumes of text

Group 1 and 2 and some out of group 3 might be content with image versions of digital newspapers; on the other hand, some researchers from group 1 and 3, and all out of group 4 need full text research and usability features that lighten their work considerably.

2 Selection Criteria

Let me now discuss some criteria that we are applying to our plannings for digitisation in Northern Germany. To be honest, I am not entitled to draw blueprints for projects in the whole North, because we are five states and regional libraries under quite different working conditions. But what we have in common, is the interest in digitising newspapers of highest research value.

Some figures again,
- the first only for Hamburg: If we exclude the 17th century whose newspapers are going to be completely digitised by Bremen State and University Library (350.000 pages at most for the Holy Roman Empire as a whole), there are about 35 longer running newspaper titles for Hamburg and its former satellite cities: we have to cope with an estimated amount of 3,7 million pages.
- The second for the Northeast of Lower Saxony, the directly neighbouring districts in the South of Hamburg: from 1890 to 1930 about 65 to 70 papers (main edition) appeared there, and probably even more existed in the North of Hamburg, that means: we have about 140-150 titles in the neighbouring districts of Hamburg.
- The third refers to a project that I am considering together with colleagues: to digitise the North German party press sector by sector, first of all representative leftist papers (socialist, social democratic, communist) which appeared from 1848 to 1933 representing centres of the workers' political movement and perhaps specific events such as the November revolution in 1918 which was driven by soldiers and workers in the harbour of Kiel on the Baltic Sea. The geographic scope ist broader here in the five northern states: out of 28 titles in 20 cities we could digitise 600.000–800.000 pages. (Similar follow-up-projects should cover the liberal, the conservative and the nationalist press, of course.)

Now let's pass over to the selection criteria:

I. The first bundle of criteria refers to material conditions and the digitisability of the newspapers

First of all, we consider, which titles are available on microfilm, for we know that digitisation from microfilm is remarkably cheaper than from the original, in particular if the original is to be conserved as a volume. Microfilm holdings are preferred for budgeting reasons although the digital images are to a certain degree less suitable for optical character recognition, the results slightly worse, for all we know. In the northern states of Germany librarians and archivists have continually taken care of microfilming the historical papers, and we still do complete these collections (very helpful as a clearing house is the Mikrofilmarchiv der deutschsprachigen Presse e.V. in Dortmund).

The first step is to identify complete runs of microfilms – preservation microfilming should have been a cooperative effort of all proprietors in order to provide a series of films as complete as possible. In case we discover additional original volumes we consider if primary microfilming or a combination of microfilm and digitisation from the original might be the adequate solution.

Of course we have to test if the image quality on microfilms is sufficient so that they yield satisfactory digital images and results with optical character recognition. This step may cause problems and delay in the preparation workflow as we need an evaluation by one or more of those firms which will later participate in our bidding process.

II. The second bundle of criteria: Practical reasons, pragmatically and financially important circumstances

- One could concentrate on a comprehensive newspaper collection of a library or archive and benefit from easy availability and disposability of a variety of titles.
- Sponsorship might play an important role, using the readyness of publishing houses still existing which might be interested in using the librarians' know-how in retrodigitisation to create a historical archive.
- Historical societies might subsidize newspaper digitisation in their area of interest, and more factors like this might ease our decision making.

The other way round, it is argued that orphan titles deserve special attention: "High research value titles that have ceased publication and lack active ownership likely to invest in future digital conversion should receive special consideration for NDPD", the American National Digital Newspaper Program. And they do receive it in Hamburg as well.

III. Intrinsic value

Certain single papers may be privileged by the circumstances described above, while others stay behind and do not benefit from pragmatic reasons and advantages. At this point we have to turn to an inquiry of the whole regional media landscape over time. Scrutinizing the whole system of the press in certain historical periods, we have to apply criteria to our selection process that refer to the intrinsic values of our newspapers.

We should systematically examine the region and ask: 'which newspapers had their sphere of influence here, for which segments of the public, and what kind of contents did they cover?' Doing this we are well advised to consider or have in mind the whole typology of the press of that period. This means, we have to examine both sides, supply and demand, focus our view on the typology of the press as well as on the demand of the public in past times.

In our example: we pay attention to
- broad geographic coverage in respect of information and of the public
- but we do not exclude only locally relevant titles of which we choose examples that appeared over a longer period of time
- longevity: long lasting papers are important showing the development of journalistic techniques, of the coverage and depth of reporting world news or local news
- but we do not exclude titles of short temporal coverage in case they represent specific historical phenomena, i.e.
  - famous journalists (Kleist, 'Berliner Abendblätter'; Matthias Claudius, 'Wandsbecker Bothe'),
  - political events (for example, the revolution of 1848 which brought press restrictions to an end; we are going to adapt the Bavarian programme of 'Zeitungen from 1848 to 1850'),
  - first newspapers at a certain place or of a specific type
- the representation of the typology of the press:
  - news reporting as was usual in the 17th to mid 19th centuries
  - party papers, emerging in Germany 1848 and after: liberal, conservative, socialist, communist, nationalist, regionalist, separatist, politically catholic etc.
  - non-party papers, serious or quality papers
  - yellow press, tabloid papers, 'Generalanzeiger'
  - provincial papers
  - official papers

- the representation of different types of places of publication: harbour and commercial cities, university towns, capitals of larger and of smaller territories, remote towns etc.
- the representation of the foremost socioeconomic groups of society, political sides and parties, religious groupings
- the representation of specific groups or ethnic minorities (the Danish minority at the North German border in our case)
- newspapers of indirect political influence whose content has been copied by many others voluntarily or under pressure
- There may be constellations of competitive struggle between two or more papers in a certain market which should be documented by source material.

Of course it is impossible to integrate all types and combinations of qualities into one draft scheme of newspapers to be digitised. But I tried to unfold the variety of cases and criteria in order to mention them as sort of touchstones for concrete decisions to be taken.

Thinking of the strategies of selection we can imagine different paths of proceding:
- We could first of all identify the really big titles by period and region and leave all the rest to the future
- We could try to sort out representative numbers of each type of papers and digitise type by type
- or we could, exemplarily, concentrate on a relatively small region and try to digitise a representative sample of all types of its papers

3 Instead of a conclusion

Let me try to direct your attention to a list of, in my opinion, most important criteria or points to reflect upon in the process of title selection. What we have to avoid at any rate is to cause a complete distortion of the structure of our historical press heritage on the web. It is absolutely impossible to attain a complete consensus about the selection we make – but we should keep a couple of touchstones in mind and be quite clear about the reasons for our decisions. There may be reasons to by-pass some of the criteria sometimes, but certain representations are not to be neglected if we look at our countries' press on the whole.

Material and pragmatic factors

- time-span
- microfilm
- complete microfilm run
- quality of microfilm images
- paper originals as substitute / complement to the microfilm run
- publisher still existing vs. orphan titles
- sponsoring opportunities
- practical reasons, availability

Intrinsic value

- geographic coverage and influence
- newspaper for fellow-countrymen abroad
- politics:
    - political role and influence
    - representing political party or movement
    - means of propaganda
    - connection to (a series / complex of) political or social events
- place of publication:
    - capital, administrative centre
    - special type of city (university town, harbour, commercial centre)
    - central town in sparsely populated area
- social representation:
    - ethnic minorities
    - religious orientation
    - average agrarian / town population
- economic factors:
    - competitive struggle between certain papers
    - monopolistic influence
    - specific marketing methods, business concept
- special importance for the history of the press:
    - first newspaper at the place of publication
    - (first) example of a specific type of newspaper or content
    - relation with censorship
    - important journalists or editors

Der Deutsche Bund 1815-1866

Source: Wikipedia <17.3.2011>,

http://upload.wikimedia.org/wikipedia/commons/6/6c/Deutscher_Bund.png

To Be In The News, To Remain In History: Preservation of Local Newspaper Article in Borneo.

By Imilia Ibrahim

(Universiti Teknologi MARA, Malaysia)

Imilia Ibrahim

Universiti Teknologi MARA

Malaysia

Imilia Binti Ibrahim is a lecturer at Universiti Teknologi MARA. Her academic qualification is Master of Commerce (Advanced Information Systems Management), University of New South Wales, Australia and B.Sc. (Hons.) Information Studies (Library and Information Management), Universiti Teknologi MARA. Her research are "ICT Awareness among rural communities in Sarawak (2008)" and "Information Needs and Information Seeking Behaviour among Users of UiTM Sarawak Library (2010)". Papers that she produced are:

1. Institutional Repositories: Challenges among (Proceeding International Library Conference: Sarawak 07)

2. Sarawak Committee for Information Literacy Skills (SCILS): A collaboration towards a standard information literacy skills in Sarawak (Proceeding ICOLIS 2008)

3. ICT Awareness among Rural Communities: A Case Study in Kampung Sebandi Matang, Sarawak (Proceeding KABOKA 4, 2008)

4. Library roles as towards modernizing Sarawak Communities (Proceeding KABOKA 5, 2009)

5. Librarian's Motivation Towards the Implementation of ISO9000: A Case Study in Sarawak (Proceeding ICOPS 2010)

6. Facebook as the Alternative Distribution Channel of Information: A Case Study on UiTM Vice Chancellor Facebook Page (Proceeding ICOPA 2010)

To Be In The News, To Remain In History:
Preservation of Local Newspaper Article in Borneo

Imilia Ibrahim

imiliaibrahim@sarawak.uitm.edu.my

Sabariah Abd Samad

sabariahas@sarawak.uitm.edu.my

Universiti Teknologi MARA (UiTM)

Valentino Abu Bakar

abvalentino@cais.unimas.my

Universiti Malaysia Sarawak (UNIMAS)

Abstract

In Borneo local newspaper has been an important source of information to the public. This is due to the delay in deliveries of national newspapers especially in the rural areas. Nowadays, printed newspaper has starting to be phased out by the online equivalent. Technology has improved; therefore the retrieval of latest news is no longer an issue. However, most of the local newspaper publishers do not have their news archive. As a result, it is difficult for users to capture dated news. This research analyses the act of preservation of local newspaper article by information agencies in Borneo. The issues that will be discussed are the preservation policies, the procedure and access of archival news to the public. This paper also attempts to determine the development of digital libraries and access to news archives. The findings recommend that collaboration among information agencies in Borneo in the preservation of local newspaper articles will not only benefit the public, it will also sustain the knowledge and history of the local community for the future generation.

Keywords: Preservation, Local Newspaper, Borneo

Introduction

Newspapers have been the major source of news and information to everyone around the world. Though there are many other ways to get latest news and information, the satisfaction of flipping through the pages of the thin papers are still favored by many. A lot of historical moment could be captured in newspapers. Some people said they feel more contented by reading the hardcopy of a newspaper, even though the advancement of the technology allows them to read and access to the newspaper in many other ways.

Newspapers are still one of the main sources available for many people to update themselves with latest news or to get information about things around them and throughout the world. In the early stage of newspaper production, not many people could get access to the newspaper. However, as time goes by and with the advancement of the technology, the newspaper industries manage to help more people to become literate and at the same time reduce the knowledge divide. Therefore newspaper has become one of the hidden treasures for a country. News written in the newspapers has become recorded knowledge and they are part of the history of the world.

National heritage of one country depends to their national collection. If newspapers are not treated as one of the national treasures it could affect the future generations. According to Ganguly (2007), "heritage is explained in UNESCO documents as "our legacy from the past, what we live with today, and what we pass on to future generations".

Newspapers in Borneo

The island of Borneo with emphasis on two Malaysian states, Sabah and Sarawak has a large and blooming newspaper industry. The local newspapers have played major roles in information dissemination in these two states. Historically, the access to electricity in Sabah and Sarawak were not as good as those in the peninsular Malaysia. Due to that reason, people in Borneo are lacking in the opportunity to access to the latest news on the television or radio. The sizes of the two states are also relatively larger than the Peninsular and this leads to slow distribution of the national newspaper. This is the major reason why the local newspapers are very important to the community. With the local newspapers, the society could have access to both national and local news.

Language and script

The newspapers in Borneo are majorly written in English though there are a small number of local newspapers use Malay and Mandarin. This scenario leads back to the history of these two states. The newspaper publications in Sabah and Sarawak dated way before independence and during those days, the two states were ruled by the English and they were the ones who started the newspaper industry in the states. And as the early schools in these two states were mostly missionary schools, the society are better equipped with the English language scripts.

Although there are no local newspapers using the ethnic dialects, there are some contents inside the newspapers written in the major ethnic dialects like Iban and Kadazan Dusun. The news written in local dialects are normally focusing on events and activities of the local community throughout the states.

As for the scripts used in the newspapers, most are using the English alphabets and the Chinese newspapers are using Chinese script. In the early days it is said that there are newspapers written in Jawi or Arabic Script but the publication has stop. This might be due to the reducing number of people who could read Arabic script.

Circulation

In Borneo, most newspaper publisher work independently and they cater only for the local market. Though there is a company that published Utusan Borneo and Borneo Post which are accessible to both states, the content of the newspapers are custom to the respective states. The newspapers published are mostly to be circulated in the morning. there only a small number of Chinese newspapers are distributed as afternoon newspapers. In Sabah and Sarawak, there are not many news agents as compared to the other places. Therefore newspapers are sold mostly in grocery shops, 24 hours shops and books and stationary shops. There are also small stalls being set up on the roadside of the cities. With limited distribution channels, sometimes it is difficult to identify the exact number of newspaper being published because most stalls and shops only have limited copies of newspapers that cater for the local community. Table 1 list some of the local newspapers that have been available to the public in Borneo.

Table 1: Examples of Borneo Local Newspapers

Sabah	Sarawak
The New Sabah Times	The Borneo Post
Borneo Mail	Utusan Sarawak
The Borneo Post	Utusan Borneo
Daily Express	Sarawak Tribune
	Eastern Times

Newspapers and Information Agencies

Since newspapers is one of the major information resources in Borneo, it is important for the information agencies to play their roles in storing, preserving and making sure that the community will get access to the newspapers and the news that they need. Therefore each of the different information agencies in Borneo has different roles to play.

Library

Newspaper, considered as one of the important information product would be one of the resources in most information agencies. Libraries in Borneo have played major roles not only in making the daily newspapers available to the public, they also make sure the important articles are made available and being indexed in the library collection. Though most libraries made latest and dated news available to the public, they have different ways of doing it.

Some libraries indexed and store their own newspaper articles. Pustaka Negeri Sarawak (PNS) or the Sarawak State Library for example allows users to access indexed newspaper article from their database. They indexed most of the local news about the state of Sarawak only. All the indexed articles are scanned and made available in their e-newspaper database. Similar to the PNS, the library in Universiti Malaysia Sarawak (UNIMAS) and Universiti Malaysia Sabah (UMS) have similar methods in indexing and making local newspaper articles available to the library users. Yet their policies of the news they made available would be different.

The library of Universiti Teknologi MARA, Samarahan campus makes newspaper indexing as one of the main task as well. The indexed newspaper articles are made available to the users in hard copy form as well as in CD-ROMs.

Besides going through the hassle of preparing their own newspaper index, some libraries subscribed to the e-newspaper archives. Most of these archives covered national newspapers but there are some local newspapers available for subscription like the Eastern Times.

As libraries become a reference place for news and information, some libraries also have in their possession bound newspapers for their user's needs. Libraries like the PNS and the Yayasan Sabah Library provide daily newspapers as well as bound newspaper to be used upon request.

Archive

In Borneo, there are branches of the national archive both in Sarawak and Sabah. These branches of national archive are responsible to store copies of local newspapers in the full form. In Sarawak for example, the storage and safekeeping of the local newspapers only starts in the 1970s because the national archive in Sarawak was just founded by then. The newspapers received in the national archive were bind monthly until recently when they started to use boxes to store all the newspapers according to month.

The storage of the local newspapers is done in house in the local archive but there is no additional work done to the newspapers like digitizing them or having them in index form. Therefore the accesses to local newspapers are very limited as compared to libraries.

Besides the branches of the national archive, in Sarawak the PNS is responsible to the state depository as well. The state depository includes the state archive therefore the responsibility of storing and preserving the local documents is included in their task. The depository store local newspapers in their full form starting from 2002. Similar to the national archive, the newspaper collection within their possession is kept in boxes according to month. Though they have just started collecting local newspapers since 2002, there is also a special project done by the depository center where they were given a large pile of old and dated local newspapers that they are trying to save and conserve.

In Sabah the state archive has also plays the major roles in safekeeping the newspapers collection. The state archive which used to be together with the state museum now being an independent department and being responsible fully to the collection of local newspapers. As of today, there are 33 titles of newspapers in the possession of the Sabah State Archive and the oldest newspaper is the British North Borneo Herald and Official Gazette dated way back in 1883.

Museum

The Sabah State Museum who is used to be together with the Sabah State Archive now has focused the work only to historical artifacts. Therefore they do not store any newspaper collections. Yet the case is very different in Sarawak. Though the state depository have been developed in the year 2000, the Sarawak State Museum, under their Library and Archive Departments still have in their safekeeping a large collection of the state manuscripts as well as the newspapers collection. The collection that they have is in bound form and some dated way back to 1870 as the first newspaper published in Sarawak is the Sarawak Gazette. In 1908, another newspaper, Sarawak Government Gazette emerged. Later, the first Malay newspaper, Fajar Sarawak was published in 1930 and Chinese language newspapers Shen won Kie Min Sing Pao and Xi Min Ri Bao were published in 1913 and 1927 respectively.

Preservation of Local Newspapers

As Borneo local newspapers have only being distributed in the island of Borneo, it is important for the local information agencies to take the responsibilities to preserve the newspaper collections. There are many reasons why it is very important for the newspaper collections to be preserved.

The major reason to preserve the local newspapers would be because of their historical values. As the daily newspapers record all the news and events that occur daily, this news would soon become the history and some very important events must be kept in storage and made available to the future generations as that would later become values of the society. Another reason why they need to preserve these newspapers because some of the local news highlighted is not being published in the major newspaper; therefore it is very important to make sure that those newspapers are being safely kept by the respective agencies.

The next reason is to educate the community. As the local newspapers would have very different and local approach to the news, it could be a medium to educate the community of the local ways of life. The news and event could show and highlight the local believes and their way of life. Some news was just like a lifeline to some readers as the news keeps people aware of the development and happenings in the rural areas.

As there is a some sections of local dialect included in the local newspapers, it is very important to preserve these newspapers as they are the source of reference especially to those who would like to research on the local community and language. The newspapers with local dialects should always be preserved and maintain. Newspapers in the local dialect will always be unique and interesting to be studied by researchers and young generations.

When discussing about the preservation of the local news as well as local newspapers, it seems like anybody could do it and everybody have been doing it. Libraries, archives and museums should all be aware of their responsibilities in making sure the local newspaper collections are there and available for the sake of the future generations. Besides the information agencies, the newspaper companies could also give their assistance in making sure of the availability of the local newspaper collections to information agencies as well as to the community.

The discussion on preservation of local newspapers would not be complete without discussing on how to preserve the local newspapers. As we know, most information agencies store daily newspapers in boxes and in bound. Some indexed articles from the newspaper and digitize the articles in their databases. These two methods are mostly used by all information agencies. Unfortunately these methods might not be good enough if we are planning to maintain newspapers as local heritage and to have them ready to be preserved by the young generations.

Challenges in Local Newspaper Preservation

Newspapers are known as one of the best medium used to share stories among the people. However to maintain and retain the news is a big challenge not only to the information agencies, but also to other agencies that are related to news.

In the information agencies like libraries, converting the newspaper into a digital format has created more jobs and responsibilities in the workplace. Though there are some companies that offer digital access to newspaper archive, the cost of the subscription would be added to the operational cost of the information agencies. Furthermore, the newspaper archive prepared by the newspaper companies might not be in a proper format where the access points might not be as friendly as a library catalog.

Besides the cost of preservation work, sometimes the process of preserving or digitizing local newspapers could be a redundancy. As what happen today, all the different information agencies have the local newspaper collections in their possession and they all are doing their own preservation and digitization on their own. Should they have a group or agency that does all the preservation and digitization, the resources could be shared and there will not be a redundancy of work and a waste of time and money.

Preservation, digitization and indexing of newspapers and newspaper articles could be a simple task. Unfortunately this simple task could not be done by anyone in the library. The selection of newspaper articles, the method of preservation and the keyword selection should be done by a professional staff. Therefore the preservation process might not be done without a well trained staff in the information agencies. Lack of professional and trained staff would be another obstacle that needs to be tackled by information agencies. Some information agencies might not be able to hire a professional just to do indexing of newspaper articles or preservation of newspapers. Therefore it is important for them to train the support staff so that they could do the preservation work under the supervision of a professional.

Conclusion

Preservation is very important to make sure the local heritage could be sustained for the use of future generations. In the case of local newspapers in Borneo, the needs of preservation are even bigger since the local newspapers are the evidence of events and the history of the states and the people. As both the states in Borneo have unique background and culture, there are many representation of this culture that was not captured in the national media. Therefore it is important for the local media to be well preserved to make sure all the uniqueness and the richness of the culture could be maintained for reference and research purposes.

With the knowledge of the importance of preserving the local newspapers, most information agencies have worked towards the preservation initiatives. Unfortunately, most agencies are working solo in their initiatives. This leads to waste of resources and redundancy of work. Therefore it is recommended for the agencies to work together in preserving the local news heritage.

Borneo local news would be Borneo local heritage. The locals are aware of the fact and initiatives have been done to preserve the local news. Should they have better coordination and smart partnerships between local and national organizations, the safekeeping of the resources would be materialized.

Reference

Ganguly, R.C .(ed.) (2007), **Preservation of Electronic Information**. Delhi: ISHA Books.

Mustafa K Anuar. (2007) **History of Local Media**. Retrieved on 18 Jan 2011 from
http://www.malaysianbar.org. my

Pustaka Negeri Sarawak (2011), **Pustaka Negeri Sarawak**. Retrieved on 15 Jan 2011 from
http://www.pustaka-sarawak.com.my

Sabah State Archive (2011), **List of newspaper records**. Retrieved on 22 Feb 2011 from
http://www.sabah.gov.my/ark/Newspaper.html

Yayasan Sabah Corporation (2011), **Tuan Haji Mohd Fuad Stephens Borneo Research Library.** Retrieved on 20 Feb 2011 from http://www.ysnet.org.my/library.asp?p=s&l=bi

Preserving Digital News Collection: Overview at BERNAMA and UTUSAN MELAYU which have benefited the nation

By Hasnita Ibrahim
(BERNAMA, Malaysia)

Hasnita Ibrahim
BERNAMA
Malaysia

Since June 2000, Hasnita Hj Ibrahim is managing the InfoLib of BERNAMA, which specializes in news and information portfolio. Prior to this, in May 1995, Hasnita worked as a librarian at the Judicial and Legal Training Institute (ILKAP) under the preview of the National Library of Malaysia specializing in legal documents.She graduated from the University of Technology MARA, Shah Alam with a Bachelors degree in Library and Information (1990 – 1994) and obtained her Masters degree in the similar field from the University of Malaya (1997 – 1999).Her professional experiences include in coordinating the Library System such as Micro VTLS (ILKAP) and ILMU (BERNAMA); ILKAP and InfoLib Websites; Tun Dr Mahathir/BERNAMA News on CD ROM projects; BLIS (BERNAMA Library and Infolink Service) and InfoLib Portal; Secretary to the ISO team (ILKAP) and Coordinator for Administrative Documents - *Fail Meja* (Bernama). She is also passionate about training in the weekly orientation programme at ILKAP; training for the internal users at BERNAMA and BLIS clients all over Malaysia as well as promoting BLIS and BERNAMA at exhibitions in various venues. Under her supervision, at the Ministry of Information Communications and Culture (KPKK) level, InfoLib had won the Usage of the ICT at the work place Award (20005) and The best Division Award (2006); Library Excellent Award 2010 for Small Library Category organized by the Librarians Association of Malaysia (PPM); and third place - Innovative & Creative Circle (ICC) Convention in conjunction with the Innovation Day KPKK 2010. She was conferred the Excellent Worker Award (2005 & 2010) and Promising Librarian Award (2008) organized by PPM.Hasnita has delivered papers at the Digital Library Collections Seminar (2005), International Conference of Librarian (ICOL) (2007), International Seminar - (Strategi melestarikan sumber dan perkhidmatan dalam talian Institusi penyelidikan) (2007), Costomer-focused Culture Conference (2008), Electronic Resources in Malaysia Seminar (2010) and contributed a write-up namely BERNAMA News preserved the National Heritage in the PPM Journal (2005). She is the Council member of PPM and Chairperson of the Special Library Committee. Involved in organizing conferences, seminars, workshops, social activities and visits to prominent libraries.

PRESERVING DIGITAL NEWS COLLECTION :
OVERVIEW AT BERNAMA AND UTUSAN MELAYU
WHICH HAVE BENEFITED THE NATION

Hasnita Hj Ibrahim, Masfaliza Masmuda and Fouziah Abd Rahim

Library & Information Centre (InfoLib) Bernama and

Utusan Melayu Information Centre

hasnita@bernama.com, masfaliza@bernama.com, fouziah.rahim@utusan.com.my

ABSTRACT

In recent years, technological development has had not only a profound impact on library collections but also challenged the librarians' role in preserving and maintaining digital collections. Leaving with no choice, news libraries have to digitalize the fast-growing amount of news collection. Speed and accuracy are much crucial in relations to the role of providing news on a daily basis to the users. As a result, the news is read and monitored by the nation via daily news. Thus it is important for the news collection to be processed and preserved where such news can be easily accessed by current and future generations in the realm of the country's development. In this paper we shall focus on the overview and comparison of digital activities at the Library & Information Centre (InfoLib), Bernama and Utusan Melayu Information Centre.

Keywords: News, Digital; Digital archive; Bernama; Utusan Melayu

INTRODUCTION

Sanllorenti (2005), taken from B.S. Sanchez (2006) defines digital collection can be formed by different objects of information, being originally produced in an electronic format or digitalized starting from a printed original. Digital collections need to be accessed through online or an information retrieval system.

He added, digitalizing the collections will enhance access and improve preservation. It offers opportunities to preserves and provides access to this content. Digitalization initiatives will allow users to search collections rapidly and comprehensively from anywhere at any time.

According to Smith (1999), digitalization has proven to be possible for nearly every format and medium presently held by libraries, from maps to manuscripts, and moving images to musical recordings. The use of hardware and software for capturing an item and converting it into bits and bytes, matched by a quickly developing set of practices for describing and retrieving digital objects, is giving form to the talk of a "library without walls."

Technological developments have had a huge impact on library resources and challenged librarians' roles in collecting, managing and preserving digital materials. Smith (1999) added, despite the high cost of digital conversion, many institutions are taking on ambitious projects in order to find out for themselves what the technology can do for them. They are investing large amounts of money in projects to make their collections more accessible and, too often, believing that they are also accomplishing preservation goals at the same time.

According to Moghaddam (2010), all public institutions such as archives, libraries, and museums should be involved in applying their professional skills and expertise to the long-term preservation of digital materials, just as they have taken a role in the preservation of traditional materials. Digitalization adds value to the original by providing linking, finding aids and enhanced searching options.

Rowley (1996) stated that, collection development will continue as a vital part of a library's mission for the future. It must, however, be rethought and framed in the context of the library as a developer, designer, and manager of knowledge. This will require an understanding of new information products and their delivery mechanism, as well as a commitment to support developing formats. Library leaders, empowered by new and emerging information technologies, will need to think broadly and creatively about the provision of information in this new environment. Libraries must transform their processes into new pathways to support a greater diversity of information and knowledge management. Resulting organizational structures for collection development will combine meaningful processes, knowledgeable staff, and responsive information products for new dynamic delivery of information to users.

InfoLib and Utusan Melayu Information Centre have taken the necessary actions to digitalize their collection which is growing each day. With the rapid increase of news articles, digital preservation has become increasingly important. The users do not have to know the process but most important is the result of the searches.   It is essential for libraries to preserve and make accessible all kinds of publications in digital form such as online, CD, digitalized images, and born-digital objects.

DIGITALIZATION @ BERNAMA

BACKGROUND - BERNAMA

The Malaysian National News Agency or BERNAMA is a statutory body under the auspices of the Ministry of Information Communication and Culture. Set up under the Act of Parliament (1967), BERNAMA was established on August 30, 1967 and began operations in May 20, 1968.

Bernama has over 300 journalists, editors, photojournalists and video cameramen at its headquarters at Wisma Bernama, Jalan Tun Razak, in Kuala Lumpur and bureau offices in all the states in Malaysia and capital cities of Dubai, Beijing, New Delhi, Bangkok, Singapore and Jakarta while correspondents are stationed in Washington, London, Manila, Berlin, Dhaka, Melbourne, Tokyo, Colombo and Kathmandu. A worldwide presence allows BERNAMA to cover news events of local, regional and global importance.

Bernama is producing and disseminating the news and information through BERNAMA Newswire by paying subscribers including the published newspapers, broadcast stations and websites. Many newspapers and electronic media and other international news agencies as well as the government and private companies in Malaysia are BERNAMA subscribers.

Other products provided by Bernama are Bernama.com, SMS News, Newslink, BLIS, BERNAMA Media Relations & Event Management, Bernama Images, Bernama TV (Astro channel 502), Electronic Display Board, Hyperlink, Interquotes, Multiquotes, Dragonshare and Radio 24 at 93.9 fm. (More information – http://www.bernama.com)

BACKGROUND - INFOLIB

InfoLib, formerly known as Bernama Library, was established in 1968, in conjunction with Bernama's official existence. InfoLib is a library specializing in news (Bernama news and news clipping) and information about Malaysia.

In line with Bernama's objectives of generating as well as disseminating news and information, InfoLib has been entrusted in storing, processing as well as archiving them for reference purposes which is not only accessible by Bernama reporters and journalists, but the public since April 1992.

InfoLib and BLIS users are mainly Bernama staff and bureaus all over Malaysia and foreign officers/correspondents in foreign countries.

The services are Reference & Information; Borrowing of Materials; Interlibrary borrowing; Computerized Databases; InfoLib Portal; "Le Tour de InfoLib"; Infolib Corner located at Editorial Department; Hotline "202"; and Training for BLIS & InfoLib Portal.

COLLECTION – INFOLIB

The two main collections of the so called hybrid library are books on mass media, journalism, broadcasting, politics, economy, biography / various fields, light reading materials and training notes; and the information about the collection can be accessed via ILMU (Integrated Library Management Utility), which web enable via InfoLib Portal.

InfoLib Portal, launched on the April 22, 2008 will assist users to check the WebOpac (content pages, book covers and the abstracts, if any, of the books); the borrowers status via membership; selected magazine articles; full text training notes titles; information about InfoLib, Bernama and the Bernama staff; and gateway to the important links and InfoLib on-line services. (More information – http://infolib.Bernama.com)

The full text digital collections can be accessed via BLIS (Bernama Library & Infolink Service) at http://blis.bernama.com

BLIS has 11 modules and the four main modules are:

Bernama News

The Editorial Department produces news, reports, articles, special reports, news focus/analysis in the following divisions, such as General News - social, crime, court, diplomatic, accidents, politics, election, arts, education, reading, tourism, government policies, parliament reports, development, international relationship, culture, sports, security, health, environment, science & technology, defence, disasters and other current issues; Economic News - corporate development, annual reports, company financial developments, policies, Bank Negara announcements, currency rates, market and commodity reports, and other socio-economic news; Features/Special Reports – features news, focus and news analysis pertaining to the current news or selected issues in various fields; Press Release – written statements consisting of information, explanation, formal announcement from individuals, government, private sector and society which has news value.

News sources are from personality/individual statements, media advisory, press statements, press conference, speeches, reporter's observation, interview, international news agency & local newspapers and on site visits.

Newspaper clippings

For comparison, reference and more choices of news, InfoLib is archiving the news from local published newspapers. Initially it was compiled manually, from selecting the article according to guidelines created in-house (according to the journalists needs), indexing (date and subject), cutting, pasting/stapling on the A4 size paper and lastly filing it according to the subject.

The present practice is to digitalize the article, from selecting, cutting, scanning, indexing (title, subject, author, date, newspaper category and product) and is accessible in BLIS. Since this process is time and space consuming, BLIS team is getting the news daily via FTP from The News Straits Times Groups and Utusan Malaysia (1997 – to current) in text format. With this facility arrangement, BLIS team can focus on digitalizing the millions of news cutting, which were compiled by subject files, and the copyright issue does not arise.

This was supported by Moghaddam (2010) where, in traditional preservation, some level of redundancy with multiple copies was inevitable in different repositories, but this story is different in the electronic environment.

Who's who

It is a compilation of the profiles of important persons in Malaysia from various fields. It contains details on a person's background including current post, date and place of birth, race, religion, addresses, telephone numbers, children names, educational background, career, awards they the personalities received, membership and patronage to clubs, societies and associations are listed in detail.

Databank

It is a compilation of information, which provides facts on diverse issues and events extracted from Malaysian newspapers and Bernama's wire reports. Information available includes current and archival information on various topics of special interest in Malaysia such as key personnel in the government and political arena, defense, education, election, judiciary, sports and others.

BLIS Statistics

BLIS STATISTIC

Modules	**Archival**
BERNAMA News	**2,699,332**
TOTAL= 3,882,2O4 (As at Feb 2011)	
21,039 Information Providers	**Foreign News**
Total Articles	
2,973 Calendar of Events	
	Speeches **556**
M'sian Facts & Figures	**3,155** **Who's**
	Reports **Databank** **Who**
	Directory

1,080,613	926	6,919	1,425	6,394	58,872
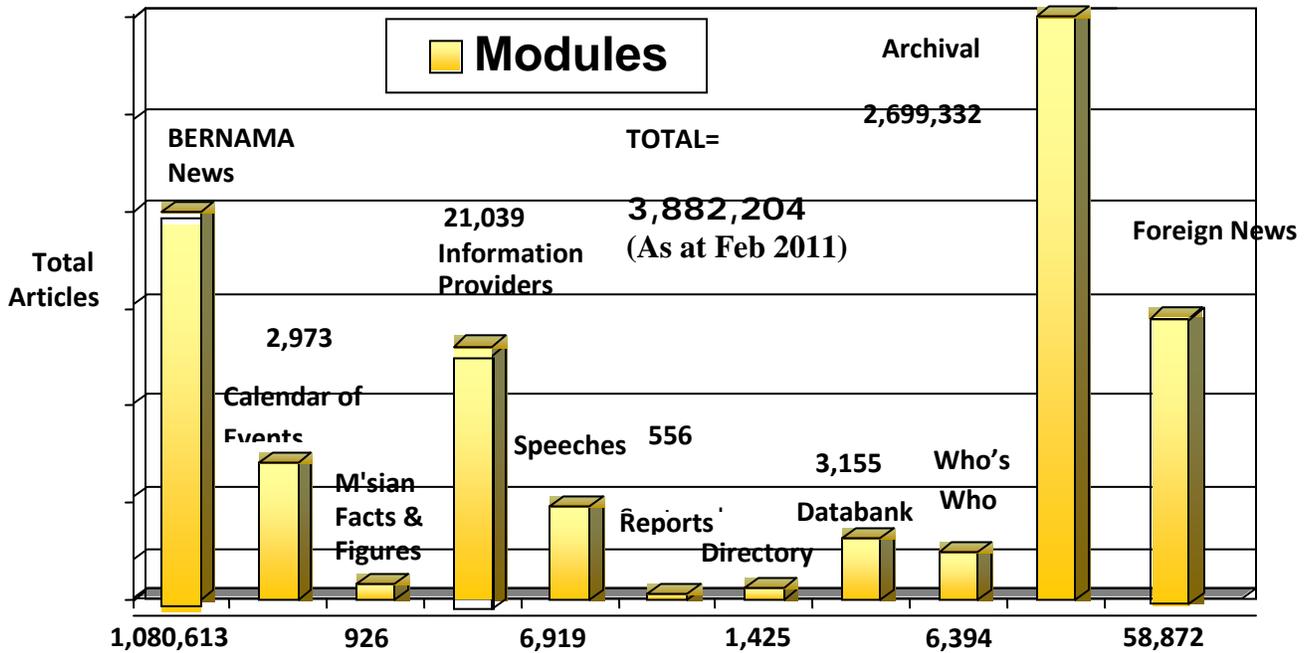
THE SYSTEM - INFOLIB

Today the news and information are for the benefit of the whole society electronically. In the early days only Bernama's journalists used them. By the end of 2001, the digital archiving system of InfoLib was in the first stage of production. Prior to that the news and information were compiled and referred manually. With the advancement of the applications and infrastructure within the digital environment, the digital archives system was up-graded and then launched in April 2008.

BLIS operates as an electronic library for easy access globally, which is ideal for anyone who needs news and information about Malaysia. Covering archival news from major newspapers, Bernama news, news reports, statistics, speeches, sectoral reports and fact files combined to offer a rich library of information for analysis and critical evaluation on a company, policy and trend.

Consists of 11 modules – Bernama News, Archival News, and Foreign news, Info Providers, Who's Who, Speeches, Sectoral Reports, Databank, Facts & Figures, Directory and calendar of Events.

Bernama researchers prepare all information by using every source of information like newspapers, magazines, Internet and sources from government agencies, private agencies, and other regional news agencies.

BLIS; jointly owned by InfoLib, which focus on the internal users and the Database and Corporate Information Division, which is focusing on promoting and selling the compiled product to the clients. BESSAR, Bernama System and Solutions Advisor Sdn. Bhd., a subsidiaries company of Bernama is executes the project on behalf of Bernama, while Karthavya Technologies is responsible in developing a web-based solution for BESSAR.

BLIS forms as an integral source of information network for all information seekers who may use the BLIS application in limitless and varied ways. Some key users such as journalists, librarians, business analysts, market researchers, lecturers, students, public relations professionals, media event practitioners, economists would get the benefits from BLIS.

BLIS is an integrated system with two sub systems that form a well-integrated system, which consists of two layers:

1. Content Management System or CMS is the software used for content management and a system to perform the workflow definition. This includes computer files, image media, audio files, electronic documents and web content. These files are available via inter-office as well as over the web. The back-end services will be performed by this system. The BLIS Admin Application is a single application in which the content management and administration of the BLIS application are integrated. The features are Content Management Module, User Management Feature, User Online Report, Subscription Management, Workflow Definition, Trial user Report, Hits Reports, Edit Hyperlink, Audit Trail and Misc Settings.

2. Web Authorization and Services (User Interface System) – The front-end system for information rendering and User Subscription Services. The BLIS User Application is the information retrieval system. The content created by the content creators is displayed in this application based on the rules defined in the BLIS Admin Application.

The system is based on the MVC Architecture. The system is using Oracle to manage both users credentials and application data. Thus data from the Alchemy/MySQL (from the previous system) database has migrated to the oracle database. All data query and search operations is directed to the Oracle database.

Bernama Library Infolink Service (BLIS) – Main page

Project Development Flow

No	Item	*Task	Date
1.	Attended workshops / seminars / conferences, visited other libraries	BLIS Team	Prior to the project
2.	Planning /Requirements Hardware's & software's	BLIS Team	Jan 2007
3.	Product demos, -Estimation budget Getting the quotations from vendors	BLIS Team	Feb 2007
4.	Prepared proposal paper	BLIS Team	Mac 2007
5.	Submited proposal paper for approval	BLIS Team	April 2007
6.	Proposal paper - Approved	BLIS Team	May 2007
7.	-User requirements -Functional requirements Indicate the Design, Functional and Database Design	Developer + BLIS team	From June 20, 2007

8.	Development of the software Oracle, Java, Jsp.	Developer	
9.	Internal Integration testing	Developer	
10.	User acceptance test BLIS ADMIN BLIS User 1.1 BLIS Admin Login 1.2 BLIS Admin Home 1.3Content Management System 1.4User Details 1.5Online users 1.6Edit subscription plans 1.7Trial users 1.8Hit reports 1.9Edit hyperlink 2.1Home page 2.2Subscription 2.3Global Search 2.5Advanced search 2.6Navigation page	BLIS Team	Jul 6, – Aug 6, 2007
11.	Integration	Developer	
12.	Testing	BLIS Team	
13.	Training	BLIS Team	Nov 11, 2007
14.	Data migration – Phase 1	Developer	From Sept 15, 2007
15.	Bug fixing	Developer	
16.	User Testing	Developer	
17.	User Testing	BLIS Team	Oct 13, 2007
18.	Checking data – Phase 1	BLIS Team	Oct 23, 2007
19.	Final acceptance (application)	BLIS Team	Nov 27, 2007
20.	Data migration– Phase 2	Developer	From Nov 30, 2007
21.	Checking data integrity – Phase 2	BLIS Team	
22.	Up (Internal users + External clients) – TESTING –Up-dating data in the present and new BLIS		Dec 3, 2007
23.	Final acceptance (data + migration)	BLIS Team	Dec 20, 2007
24.	Up – (Internal users + External clients)	BLIS Team	Jan 2008
25.	BLIS Testing	BLIS Team	Feb 15, 2008

	(Internal users + External clients)		
26.	Warranty period (6 months)	BLIS Team	May 2008
27.	BLIS Launched - Internal	BLIS Team	April 22, 2008
28.	Signed the maintenance agreement	Developer + BLIS team	Nov 2008

*BLIS Team = InfoLib and Database and Corporate Information Division

*Developer = BESSAR (Bernama System and Solutions Advisor Sdn Bhd) & Karthavya Technologies

Work Process
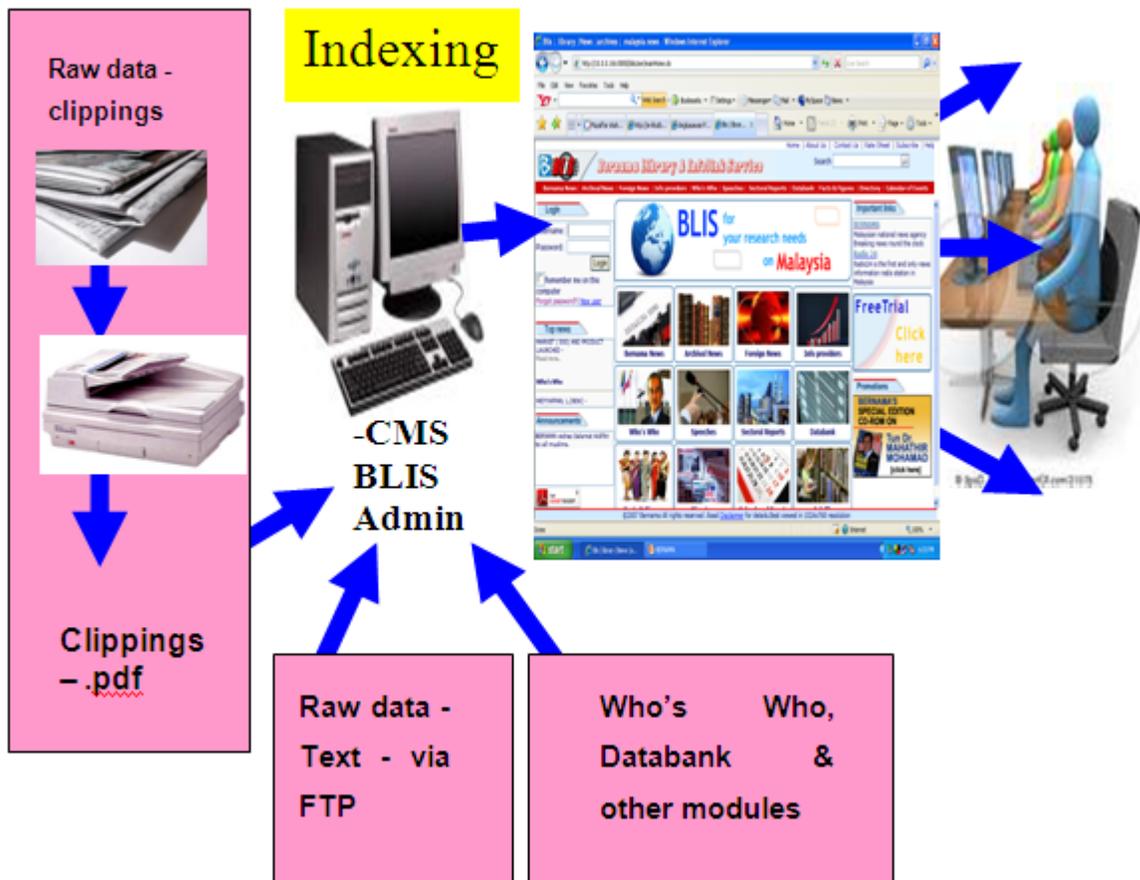
DIGITALIZATION @ UTUSAN MELAYU

BACKGROUND - UTUSAN MELAYU (M) BERHAD

Kumpulan Utusan (The Utusan Group) was incorporated in 1938 in Singapore as Utusan Melayu Press Limited. The first newspaper, Utusan Melayu, was launched on May 29, 1939. Using Jawi, the Arabic script, the newspaper was printed in Singapore and served the Jawi literate in British Malaya.

In the same year, the Sunday edition, Utusan Zaman was published. Utusan Melayu and Utusan Zaman were at the forefront of intellectual development of Malay nationalists who later provided enlightened political leadership in preparation for Independence.

In February 1958, after Independence, Utusan Melayu moved its head office to Kuala Lumpur. In Kuala Lumpur, Kumpulan Utusan (The Utusan Group) further consolidated itself and continued to play a major role and participated actively in the economic, social and political development in Malaya, which in 1963 expanded to include more territories to become Malaysia. A weekly newspaper, Mingguan Malaysia, was published in 1964. It was the country's first newspaper in romanized Malay. The first daily newspaper in romanized Malay, Utusan Malaysia was published in 1967.

1967 saw Kumpulan Utusan (The Utusan Group) being incorporated into a public limited company under the Companies Act, as Utusan Melayu (Malaysia) Berhad. Since then, Utusan or then known as Kumpulan Utusan (Utusan Group) expanded by leaps and bounds, culminating in its listing on the Main Board of the Kuala Lumpur Stock Exchange in 1994.

Even with such accomplishment, the Utusan Group did not rest on its laurels. In 1997, the Group made its entry into the world of multimedia with the launched of Utusan Malaysia On-Line, Malaysia's first on-line newspaper in full text and visuals. In 2004, a tabloid Kosmo! was published daily and in January 2005 the Sunday *Kosmo!* or *Kosmo!* Ahad were distributed around the Klang Valley area.

Today, with more than ten active companies under its wings engaging in businesses that range from publishing to computer systems management, the Utusan Group is poised to become the gateway to vast business opportunities for both local and international investors who want to take full advantage of Malaysia's rapid thrust into the information age.

BACKGROUND - UTUSAN MELAYU INFORMATION CENTRE

Utusan Melayu Information Centre is a news library, which started its services in 1965 known as Utusan Melayu library. The main function is to provide information research facility to the Editorial Department. The Utusan Information Centre is also providing services to the journalists, Utusan subsidiaries companies and outside researchers.

The Utusan Melayu library started its Utusan Info-Line database (library automation system) in early 1996. The Library was upgraded in 1998 as Utusan Melayu Information Centre. The Information Centre is one of the four supporting unit under the Editorial Department. The other three units are Photograph, Graphic and Education Unit.

COLLECTION – UTUSAN MELAYU INFORMATION CENTRE

The Information Centre collections are books on various fields such as politics, journalism and mass media, social sciences, economy, biography and autobiography, and conference notes. The information about the collection can be accessed via TARK (Tera Archiving System). The collections include books published by Utusan Publication & Distributors (UP&D) and magazines published by Utusan Karya Sdn. Bhd. Both were Utusan Group Subsidiaries.

Other collections include subscribed newspapers and magazines, newspaper clippings, microfilms, audiovisual, maps, still photos, film negatives and etc.

The Centre also has databases for newspaper daily news and articles from 1996 onwards, digital photos, biography, facts and statistics, magazines published by Kumpulan Utusan, conference and seminar papers, indexes on film negatives and newspaper pdf from 2004 onwards. New database under phase two digitization project are Electronic Information Centre Archive Repositories (E – CARE) which consist newspaper and magazine text articles and pdf from 1995 – 1939.

THE SYSTEM - UTUSAN MELAYU INFORMATION CENTRE

Utusan Melayu Information Centre started its library automation system in early 1996. The first phase of the digital project was a collaboration between Utusan Melayu Library and Utusan Melayu Information Technology Department to develop the Utusan Info-Line Database System. The system uses open sources technology, KE Texpress system purchased from Knowledge Engineering, Australia.

The Cybergrapic System used by the Editorial Department, fastens the digital process of sources whereby Library only need to copy, paste, add programming and download into the server.

Since 2004, the Information Centre used TARK (TERA Archiving System) from Milan, Italy. This new changes of system are parallel with Editorial Department where the Editorial System change to GN3 which is from the same company. Both system have been customized to fulfill the demand of the Information Centre users.

Effort for the second phase of the digital project was started in early 2005 and the project was embarked in 2008 after the Information Centre bought a high end scanner machine A1 i2S size from Excel Tech Sdn Bhd.

In June 2008, data from Utusan Database Info-Line was migrated to TARK Archiving System, which consists of E-Care Database – (Electronic Information Centre Archive Repositories) stories, pictures, facts, multimedia, magazines, biography database and etc. From January 1, 2009 the system was fully operational.

E-Care Database

Stories Database

Biography Database

Project Development Chronology

No	Item	Date
1.	Attended workshops / seminars / conferences, visited other libraries	Prior to the project
2.	Establishment of Digital team - linked closely with the establishment of the Archive Information Centre which was established in early 2005 after the Information Centre obtained their fixed location store	Early 2005
3.	As a result of the query by the Information Centre during the *Utusan Melayu* Executive Conference, the Archive Committee had their first meeting	May 16, 2006
4.	Series of meetings focused on enhancement of archive and digitalization project strategies, surveys, inviting vendors and looking at other digital project experiences such as *Yayasan Kepimpinan Perdana* (YPM) and International Islamic University of Malaysia (IIUM)	After May 2006
5.	Archive Committee approved the digital project with the purchasing of scanner machine i2s	July 20, 2007

6.	Archive Information Technology has received scanner i2s Digibook.	April 17, 2008
7.	Process of installing the scanner hardware and software	April 17 – 23, 2008
8.	System Test & Error	April 24 – May 9, 2008
9.	An official training workshop	May 12 – 14, 2008
10.	Digital project Test & Error -Analyzed momentum scanning, book restorer, pdf4books, OCR	May 15 – June 30, 2008
11.	Process of creating digital momentum – Pdf4book software – Software training pdf4book	 July 4, 2008 July 7, 2008
12.	Digital project started officially	Aug 1, 2008
13.	E-Care Database was fully operational	Jan 1, 2009
14.	Server E-care started operating. Data in text and pdf form were down loaded to the server.	Feb 11, 2009
15.	Produced the digital project report after a year	July 31, 2009
16.	Software PDF4Books Unlimited version received from Excel (vendor)	Sept 17, 2009
17.	Testing E-Care Database	Oct 1, 2009
18.	Utusan User Acceptance Test (UAT), Tark Test Scripts were signed	Jan 13, 2010
19.	Abstract file of *Utusan Malaysia* download into E-care started officially	March 4, 2010
20.	Abstract file of *Utusan Melayu* download into E-care started officially	April 27, 2010
21.	Abstract file of magazine down load into E-care started officially	May 20, 2010

Equipments

Tools	Explanations
Digibook Suprascan 8002RGB	Scanner i2s size A1 from France. Include HP work station
6 users ABBYY Fine Reader Software	Corporate version 9.0 (OCR software)
Book Restorer Software	1 unit (Editing software)
Pdf 4 Books	Limited version (Software to compress file size)
Storage	8 unit HP Eva 450GB/1TB HDD Support, 8 unit – 1TB Fata HDD
Computers	For editing process

Daily Target

Target	Quantity
Scanning	300 pages/day
OCR	12 pages/day/person
Book Restorer	Depend on scanning speed
Pdf 4 Books	Depend on scanning speed

Work Process
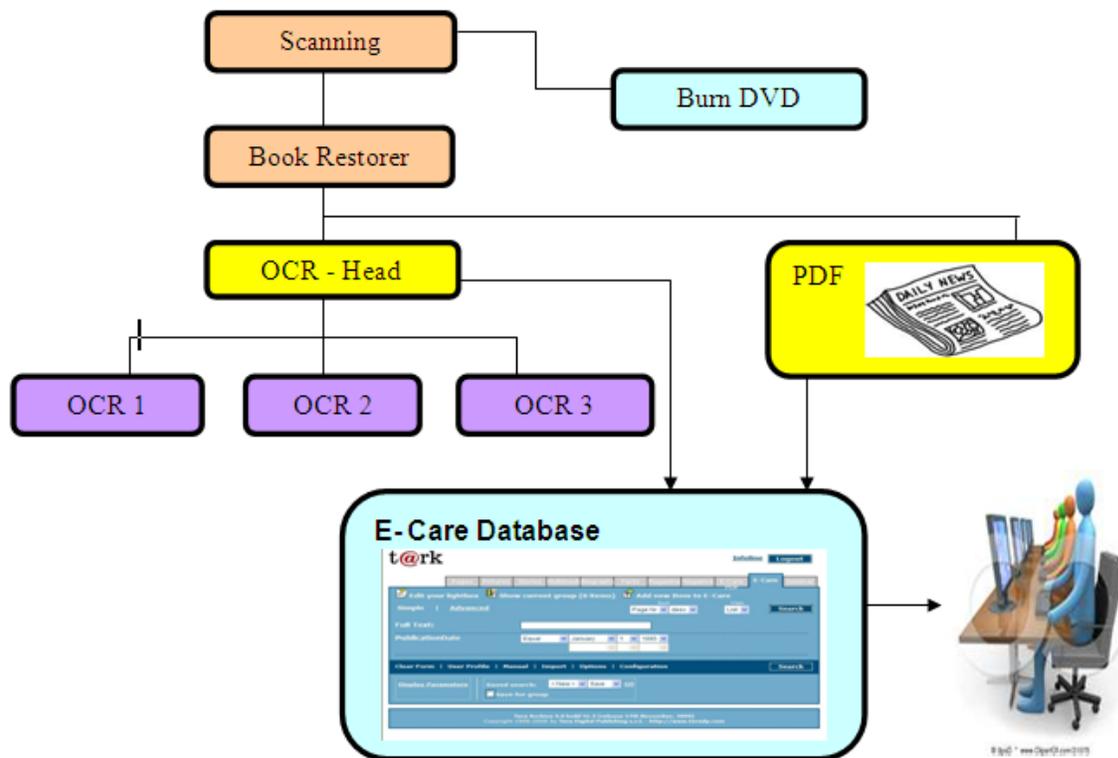
COMPARISON

Database/Modules

BLIS has 11 databases/modules, which comprises of news (Bernama news and main published newspaper namely Berita Harian, Malaysian Business, New Straits Times, Business Time, Utusan Malaysia, Bernama (1968-1995), The Edge and New Sabah Times) and information about Malaysia, which is Who's Who and Databank. Bernama news is from 1995 and news articles from 1968.

InfoLib is putting effort to digitalize millions of news articles which compiled in the manual files.

E-Care Database consists of three main databases - *Utusan/Mingguan Malaysia*, *Utusan Melayu* and magazines. They plans to include the Database of *Utusan/Mingguan* before 1995, *Jawi* Newspaper and other retrospective magazines that they have in their collections.

Utusan Melayu Information Centre has a collection of news and photos from 1960s. They are putting effort to digitalize the bound retrospective collections, which will be a great demand for research purposes about Malaysia.

As for InfoLib/BLIS the photographs are handled by the Photo Unit under the Editorial Department.

The Utusan Information Centre is also putting great effort to translate the news in Jawi to Malay to make sure that the news is preserved and can be accessed/read by the users. These collections will be very valuable to track back the history/events of our country since 1930s, which cannot be found elsewhere.

Accessibility

Under the Act of Parliament 1967, Bernama is not allowed to publish a newspaper but empowered after amendments in 1984, to generate income. Besides supporting all the national policies and development programmes of the Government of the day, BERNAMA is also a profit-oriented organization. Presently, it is producing and disseminating the news and information to many countries through BERNAMA newswire to paying subscribers including the published newspapers, broadcast stations and websites. At the end of the day, Bernama News is disseminated through our subscribers.

Our services such as Bernama.com, sms, Bernama TV and Radio24, Newslink as well as BLIS are reachable by the nation directly.

BLIS customers are the libraries from the universities, public libraries, National Library of Malaysia, government departments and private sector throughout Malaysia.

As for InfoLib, the external users on ad-hoc basis are students and individuals from government departments and private sector will be charged a minimal fee. Due to, limited facilities and physical constraints at InfoLib, only limited users can be entertained.

TARK E-CARE is mainly for internal users but they are looking to diversify the system for ROI (Return of Investment). Utusan Melayu Information Centre is also similar to InfoLib, due to the their physical constraints, the fee based services are not widely publicized. They have users from Utusan staff and subsidiaries, researchers, students and government agencies.

For BLIS, more than 1,000 daily news from Bernama, NSTP publication and Utusan Melayu will be loaded in the server. The server and the infrastructure need to be up-graded regularly according to the needs. Long term planning is important in order for the server to grow in-line with the volume. The advanced search and proper techniques will help the speed of getting the news.

It is important to ensure that the license is sufficient for the developers (up-dating) and the users (accessing), if not the imputing and accessing of information will be slow and limited.

As for Utusan Melayu, the news can be referred in a daily newspaper and via Utusan Online web address - www.utusan.com.my and Kosmo Online www.kosmo.com.my.

System infrastructure/equipments

BLIS is using the expertise from India to develop the system. The advantage of having a customized system is that we can determine the features, which are tailored to suit our requirements. But the disadvantages are that they would need a longer time and many trials and errors during the development process. Customized system will be advisable if the vendor, which develops it, is an established and skillful vendor.

As for our earlier version of BLIS, when Alchemy was bought, which was a stand alone "document management system", another working place (software) was needed which was developed by another vendor to process the news and information to the Net. As a result many problems occurred in terms of headaches and hours wasted. Hence, it is advisable to have one vendor to be responsible from A – Z of the package.

According to Green (2009), the technical arm in most of the organizations will not have the expertise to develop a proper digital library system. It is also advisable to embark a project by an established vendor and using the software, which has been established, in the market. This action can avoid circumstances where the technical personnel in-house who develop the digital archive system resign from that particular organization. The new personnel will not have the belief and understanding towards the whole architecture. Furthermore the digital archive system will not be having the up-graded version and possibly the system will be left unattended for not in line with the present technology. Green (2009) listed two aspects to ensure access:

i. Ensuring that digital items are not lost or corrupted; and

ii. Ensuring that it is still possible to read, listen to or view the items, long after the original software and hardware that was used to access the material has disappeared.

As for Utusan Melayu Information Centre, the E-care Database originated from Milan, Italy. This system is an archival system used by the Editorial Department that is GN3, which has been customized to fulfill the editorial needs. The scanning equipment is from France but maintained by the local vendor. It is crucial for the scanning equipment to be regularly maintained.

Both of us decided not to outsource the digitalization of collections task but by processing it in-house. The reasons are due to the cost factor involved to digitalize millions of news articles that we have in our collection; it is more economic; safety of the news articles in terms of loss and physical damage; less headaches liaising with the vendors; build our own expertise and for long term, the equipment bought could be used for other purposes/projects. Finally, the option to digitalize in-house according to the priority of news.

Utusan Melayu Information Centre has taken the best action by having the high-end equipment so that they can scan more news articles; the speed is 35 – 40 minutes to compress of 50 pages of news articles. This will also save the bound volume from damages in the process of scanning.

As for InfoLib we have bound volume collections of Bernama news from 1968 – 1995 which have to be processed and are trying our level best to get the budget to purchase the high-end equipment.

For Utusan Melayu Information Centre, having to scan in-house is a tedious job especially the old and sensitive volumes. The quality of the images is not according to the specification. Adding to it the tedious job involve unstitching the bound volumes which is tiedly stitched.

Due the these tedious tasks; time has to be taken in making sure that the news articles is in line with the published edition; the quality of images; partially of the images cannot be seen; and other elements such as the staff need to go on leaves, mc's, attending training and other organization functions, the team also cannot meet the target which is 720 pages a day. The actual fact is 250 – 300 pages a day only. This are the challenges that we are facing.

In InfoLib, we are facing similar challenges; the only difference is that the news articles that we digitalize are not bound volumes but articles in files. We have to make sure that the news articles are ready to be scanned.

Both information centres need to ensure that their existing staff and members can develop and continue to develop the range of competencies they need to manage the digital materials under their care.

In a paperless society it is not easy to keep up to date with one format for a long time. Moghaddam (2010) suggested to the librarians to be concerned about the fragility of digital media. The media on which digital materials are stored is inherently unstable and without suitable storage conditions and management can deteriorate very quickly, even though they may not appear to be damaged externally. In order to cope with this problem, the Digital Preservation Coalition (2001) taken from Moghaddam (2010) outlined some precautions to reduce the danger of loss:

-storage in a stable, controlled environment;

-implementing regular refreshment cycles to copy onto newer media;

-making preservation copies (assuming licensing/copyright permission);

-implementing appropriate handling procedures; and

-transferring to "standard" storage media.

Moghaddam (2010) stressed that even with these precautions the problem will not be solved because technology changes very quickly. So, even if the media is preserved in a stable condition, it may not be possible to access the information it contains yet because new technology will make the present media obsolescent. Of course, responsible care can make it easier to manage technological changes. In the volatile electronic environment, what is certain is that librarians are very much dependent on new technology for digital preservation.

Both information centres agreed that every project needed to spell out the task of the team right from the project manager to the implementing personnel and blended with the awarded vendor. Regular scheduled meetings with minutes taken are necessary to ensure that every decision will be recorded and well monitored.

Oltmans (2004) talked about the longevity of metadata. Recently, cultural heritage institutions, and the digital preservation community in general, have realized that the registration of reliable and sustainable technical information is crucial in the process of digital longevity. It has been recognized that a description of software requirements aimed at the rendering of digital publications in a current environment, will not be enough to enable viewing and editing in the long term. To use technical metadata in ten, fifty or even a hundred years from now, much more detail is needed.

BLIS and E-Care had built-up their own metadata using an in house guideline because of the diverse subjects that they are dealing with daily news. For easy access E-Care uses Abby Fine Reader to OCR their news articles and BLIS uses Adobe Acrobat Professional 8. Search engines are also very important, besides free text, searches can be made by using Title, Author, and type / title of newspaper and date of publications.

However at our end, metadata task is another tedious job after scanning, OCR, and compressing and processing to the system that we have. For the working place (BLISAdmin) we need to have a proper place to insert the metadata as well as copy and paste functions. Make sure the working place can be updated and has a back up which will avoid redundant process where we have to delete and reinsert the updated files.

According to Moghaddam (2010), the selection of media and file format is an important issue in digital preservation, as they will be subject to the influence of continued technological change. It is important to have an understanding of the various media for storage because they require different software and hardware for access, and have different storage conditions and preservation requirements. There is an essential management component for all digital media to avoid media degradation and to facilitate longer-term preservation strategies.

Open discussion and brainstorming for the requirements of the whole project including hardware, software, manpower, features needed and all related to technical issues are important. This is to avoid paying more to the vendors and to ensure speedy completion of the project. With regards budget, it is advisable to request for more revenue more than needed.

Long-term commitment to the project to sustain and maintain of its diverse and extensive range of digital assets is a must in order to avoid problems left unattended.

The system can generate reports periodically according to our needs for statistical purposes. As librarians, we need to generates library statistics, and it is important to make the developer to designed hits reports, patron report, imputer audit trail based on the databases or modules that can be printed out nicely from the system.

According to Oltmans (2004), digital publishing is causing publishers, research institutions and libraries to develop new policies, new business models and new infrastructures and techniques. A major problem is that, at the same rate at which our world is becoming digital, digital information is threatened. New types of hardware, computer applications and file formats supersede each other, making our recorded digital information inaccessible in the long term.

Hasnita (2010) recommended, while working closely with the awarded vendor is needed, but relying too much on every single thing to the vendor is not a good practice to get the best results and less hick-ups. What we have in mind may not be similar with the requested items and checked regularly to ensure the items highlighted have been implemented. Usually major changes and requirements after the application and data acceptance will involve an extra cost. So, it is best to be particular about the project before the transaction is completed.

This was supported by Cornell University Library (2002-2007) saying that technology enables the variety of formats and dissemination mechanisms change rapidly. Establishing a program that is responsive to change is a huge challenge. The advancement of ICT is changing very fast. It is sometimes difficult to catch-up with the change due to budget constraints and the present job procedure. Despite that the digital archive system must be maintained on a yearly basis and to ensure that the system is healthy and the vendors/developers need to practice the health check regularly.

To respond to evolving technological capabilities and changing user expectations, the digital archive must be revised continually. It is important to evaluate the digital content objects to determine what type and degree of formats conversion or other preservation actions should be applied, determine the appropriate metadata needed for each object type and how it is associated with the objects and the capability of the search engines.

To embark on a digital library system, it is advisable for the library to out-source to a commercial vendor to develop the project. The library employees also need to have knowledge about the whole idea of the project for them to ensure that all the requirements needed will be incorporated together after the acceptance of the system.

Recognizing and selecting a highly responsible and reliable vendor with a reliable and sustainable software/system is very important to determine the success of the project and for long term planning. Many circumstances can happen for example the company stop selling and promoting the product, the company closes shop, no/lack of commitment, getting different results/specification from what had been promised, lacking of expertise and many others.

Backup and Storage

According to Oltmans (2004) digital content is inherently vulnerable to loss or damage from hardware or software faults. Resources must therefore be allocated to the backup and recovery requirements of an organization. Initial backups should be created at the time a resource is created, with a regular routine implemented so that further backups are created during the lifetime of the resource. The recovery phase must also be considered. Procedures for data recovery should be tested periodically to ensure that data can be restored from backup media, and that the media remains compatible with changes in backup technology.

It is crucial to have a structured practice of back up such as tapes, CD or server due to the amount of information that we have. BLIS has experienced virus attack before which crippled the retrieval of the present information. After the incident the technical department deployed firewall rules to enhance security. The anti-virus is a priority and disaster planning is very important under the whole information security infrastructure.

BLIS is using two application server and two database server (one is for real-time, one is use for compress backup). As for Utusan Melayu Information Centre is using mirroring backup, which will be up if facing technical faults. Both servers can store to 4 terabyte. The volume of the articles are high, it is necessary to compress the articles so that less space is being utilized and more articles can be put in. The articles will take a longer time to be fully downloaded and it is needed to maintained the speed of getting the articles. For BLIS, we limit the resolution of each article to 200 dpi per article. As for E-Care, they can compress 1 tiff file from 85.4mb to 723 KB, which is 99.2% smaller. The capacity of the storage will be longer. Increasing the sever capacity will surely incur cost.

CONCLUSION

Providing permanent access to electronic materials is a complex problem. As has been said by Oltmans (2004), digital material is often unstable and has a brief lifespan because of the limited longevity of information carriers and the software and hardware that make the stored information accessible to users. Safeguarding the integrity and authenticity of the material is therefore a key challenge when dealing with long term preservation. Regardless of the chosen strategy, permanent access calls for continuous attention and action. The rapid pace of technological change means that the techniques and procedures for long-term storage and accessibility requirements need to be adjusted and improved constantly.

Moghaddam (2010) has indicated that, it is widely acknowledged that the most cost-effective means of ensuring continued access to important digital materials is to consider the preservation implications as early as possible, preferably at creation, and to plan actively for their management throughout their lifecycle. Therefore, the major cause of concern in the digital environment is that failure to address the long-term access requirements of digital materials at a very much earlier stage than for paper materials will almost inevitably result in their permanent loss. Establishing preservation strategies in the early stages of stepping into digital libraries will ensure the success of digital preservation.

The role of both information centres is similar to assist the journalists to produce current news to be read on a daily basis in a published newspaper or other electronic medium. It is a challenge for us being the custodian of integrity and authenticity of our collections, manually located at our premises and to make it accessible virtually. We are actually managing and organizing the digital heritage information for the nation and future generation of information users.

REFERENCES

Cornell University Library Goals and Objectives, 2002-2007. Retrieved from
http://www.library.cornell.edu/Admin/goals/objectives.html

Cornell University Library Digital Preservation Policy Framework 2004. Retrieved from
http://commondepository.library.cornell.edu/cul-dp-framework.pdf

Hasnita Ibrahim and Masfaliza Masmuda. (2010). "Sustaining and maintaining of news and information by digitalization on "Malaysiana" collections". Paper presented at Seminar Kebangsaan Sumber Elektronik held on 20 - 21 Disember 2010 at Penang Bayview Beach Resort, Pulau Pinang.

Kahin, B (1995). Institutional and policy issues in the development of the digital library. *The Journal of Electronic Publishing*. 1(1&2). Retrieved from
http://dx.doi.org/10.3998/3336451.0001.120

Karthavya Technologies (2007). Structural Design Document : BEST Module 1 – Version 1.0.0. Unpublished.

Moghaddam, G. G. (2010). Preserving digital resources : issues and concerns from a view of librarians. *Collection Building.* 29 (2) : 65-69.

Oltmans, E Diessen, R. J. van and Wijngaarden, Hilde van. (2004). Preservation functionality in a digital archive. *JCDL.* 7-11:279 – 286.

Sanchez, B. S. and Presno, I.L. (2006). Collection development in a digital environment: an imperative for information organizations in the twenty-first century. *Collection Building*. 25 (4): 139-144.

Sanllorenti, A.M. (2005), "Concepto de biblioteca, lo digital, lo electronic, lo virtual", Memorias del Curso Mercosur sobre Construccion de Bibliotecas Digitales, Unesco, Montevideo, August 29-September 2 available at www.unesco.org.uy/informatica/memoriasbibdig2005.html

Smith, Abby (1999) Why digitize? Retrieved from the Council on Library and Information Resources at http://www.clir.org/pubs/reports/pub80-smith/pub80.html

Rowley, G. and Black, W. K. (1996). Consequences of change: the evolution of collection development. *Collection Building*. 15 (2): 22-30.

Protecting your reputation with iThenticate

By Clive Wing

(iGroup (Asia Pacific) Ltd.)

Clive Wing

iGroup (Asia Pacific) Ltd

Clive Wing was until recently the Chief of Knowledge Management and Information at the UNESCO Regional Bureau for Education, Bangkok where he managed a lively print and electronic publishing programme. For the iGroup he publishes *ACCESS: Asia?s Newspaper on Electronic Information Products and Services* and commissions titles for the iGroup?s original publishing program

Watch-Read-Research the News: Marketing and Promoting News Library Services

By Debora Cheney

(The Pennsylvania State University, US)

Debora Cheney

The Pennsylvania State University

US

Debora Cheney is The Larry and Ellen Foster Communications Librarian and Head of The News and Microforms Libraries at The Pennsylvania State University Libraries. As the Larry and Ellen Foster Librarian, she works closely with the faculty and students in Penn State's College of Communications, which includes journalism and media studies departments. As head, of the News and Microforms Library her work has focused on the challenges large academic libraries face in providing access to news content for teaching and research and the role of libraries in providing access to content that is increasingly electronic and delivered via the Internet. She has co-authored articles on teaching students how to use libraries and information in Portal and Journal of Academic Librarianship and on how libraries can provide access to converging news forms in College and Research Libraries.

Watch-Read-Research the News:

Marketing and Promoting News Library Services

Debora Cheney (The Pennsylvania State University, US)

Abstract

This paper will provide an overview of the role and purpose of publicity and marketing efforts related to a library news collections and services. It will also provide an overview of publicity and marketing efforts used by the News and Microforms Library, The Pennsylvania State University Libraries. The publicity and marketing efforts have used a variety of mediums to disseminate information, including internet-based video and newspaper advertising. The News and Microforms Library also uses student interns to help develop ideas that will appeal to today's newspaper researcher.

Introduction

Libraries have long-sought to preserve newspapers and many libraries have a significant investment in and commitment to their news collections. Over time, libraries have been committed to preserving our news heritage and have adapted and adopted a variety of approaches to preserve news content. Thus, a library's news collections represent a significant library investment past, present and future. These newspaper resources provide researchers with valuable resources that are used by researchers in many disciplines.[i]

While preserving newspapers has been an important role for academic libraries, the changing nature of research, delivery of newspaper content in digital formats, and decline in newspaper/news readership results in a growing lack of awareness of the libraries' role in preserving and providing access to newspapers. Indeed, readers have an increasingly complex array of sources and delivery avenues to receive their news content. Those who read news regularly may be unaware of news research databases available only from their academic library.

Libraries contribute to this confusion by offering a complex array of news resources to their researchers and students. It is likely most academic libraries now provide access to a wide range of newspaper titles (from the first known newspaper to yesterday's blog posting), a growing variety of new "news" content, and providing varying years of coverage (even for the same newspaper). The fact libraries may license/purchase content and develop their own digitization projects of local or significant historical and copyright free titles adds to the confusion. Some sources may be free to the public, [ii] while others (*LexisNexis*, *Newsbank*, *Factiva*, and *ProQuest* databases, for example) may be licensed for specific user populations (a university's students, faculty only, for example).

The user interfaces and searching functionality to newspaper databases vary greatly. Furthermore, both researchers and libraries may also rely on *Google News Archive* search to locate and identify additional resources available the Internet. As the information environment has changed and as researchers are unfamiliar with the role of libraries in retaining and preserving news content and as more researchers expect news content to be freely delivered to their desktops, libraries must begin to re-educate our users about their library's role in providing access to news content. Publicity and marketing efforts can be used to improve awareness and use of a library's news resources and will become increasingly important.

Marketing and Promotion for News Collections and Services

Libraries have made significant progress in making their "library as place" more inviting and welcoming; they provide a great deal of classroom instruction; and they have sought to make their websites more visible, easier to navigate, and more directly related to teaching, learning and research. However, despite these changes, in truth, news collections and related services are still difficult to locate on a library's website, particularly in contrast to their visibility on the Internet. Researchers expect news content to find them and rarely expect or know how to locate yesterday's news content.[iii]

In this information environment, libraries may no longer be able to expect that if we preserve newspaper content and license it for their use--our uses will find it.[iv] Many libraries have already determined they can no longer afford to provide collections to little-used collections and have chosen to discontinue their current newspaper collections. However, libraries will need to think carefully about whether these collections still have a role in academic libraries and whether if they would market and publicize their collections and services use of these collections would increase as would visibility about the library's role in providing access to news content.

Publicity should have several purposes:

- Establish libraries as a known and reliable source for news resources and where news resources are always free;
- Emphasize services and expertise related to news resources;
- Focus on research depth and long-term preservation of news content—as compared with news websites;
- Focus on free vs. fee sources and services.[v]

Libraries can develop specific goals as they develop their marketing plan, including:

- To create collection awareness and knowledge of the library's news (in all formats) collections;
- To support newspaper readership and news literacy;
- To educate users about the functions and content provided by the library's news databases—including local historical preservation projects;
- To promote news Reference and Research Services and the library (or its website) as a support for news-based research;
- To create awareness of finding tools created by the Library to meet (often local) unique needs.

One Library's Road to Promoting and Marketing News Resources, Collections, and Services

While many libraries no longer provide a print newspaper collection, the reality is that Libraries continue to have significant news resources (databases and microfilm collections, for example) that can be used by faculty and students in teaching and research. Libraries seeking to increase the visibility of their news resources and increase the public's awareness of their role in preserving and providing long-term access to their collections should develop a marketing and promotion plan.[vi] In the case of the NML the following efforts have laid a foundation for the University Libraries, generally, and the NML, specifically, for providing access and research support to news content in every format. The following provides a summary of the approach used to increase use and readership of news resources at The Pennsylvania State University.

- Develop a vision for the news collections and services and fine tune with user input;[vii]
- Address service and collection quality:
  - Establish yourself or your department as a resource/expert for news-related reference and research services;
  - Become knowledgeable about trends and developments in the news industry;
  - Train your staff to provide strong reference and research support for news related research;
  - Work with the Digitization and Preservation Department to ensure news resources of local interest are part of their digitization efforts;
  - Develop and create finding tools to meet the most frequently asked questions of your users—*NewsCAT* and the *Centre Daily Times Obituary Database.*[viii]

- Promote and educate librarians and library staff about news resources:
  - Create a newspaper collection group—(Newspaper Advisory Group (NAG));
  - Send out e-mails and other information regularly about news resources;
  - Suggest ways news resources can be used in teaching and learning;
  - Position news content on key web pages for basic instruction ("Try These First" resources) and create News and Newspaper finding tools that seek to meet the most frequently needed types of questions and research needs;[ix]
  - Write articles for staff newsletter about specific news resources;
  - Have your staff promote collections and services during staff training and development efforts.
- Create a logo or theme that can be used in a wide variety of avenues:
  - "Watch-Read-Research the News in the News and Microforms Library"
  - "Always Free News Content in the News and Microforms Library"
- Begin to promote and market to specific user groups using appropriate media—including faculty; target faculty in disciplines which rely on news content—journalism, politics and government, etc.:
  - Partner with your Libraries/University's Marketing Department for generating press releases and quality promotional materials;
  - Use free promotional venues: International Student,[x] Honors Students, Faculty, Graduate Student and First Year Student Open House events, the Library newsletter,[xi] public displays in the library, press releases to local media, web pages, Facebook page, Twitter, etc. –make news about the news interesting and exciting;
  - Advertise in the Campus newspaper; Create eye-catching postcards, book marks, etc.
  - Create a student internship to develop promotional materials—advertising and marketing students are often interested in developing experience in non-profit settings (some may even go on to library school!);
- Position newspaper collections/resources in areas where students are "resting" or arriving and make them eye catching:
  - Create a News Wall at the entrance to the Library —front pages of newspapers from around the world each day and historical pages of newspapers in history;
  - Create news related displays—Women in Journalism; History of Newspapers; Cartoons in Newspapers, etc.;
  - Place newspapers near/with the leisure reading collection—or position displays of today's newspapers to promote your own location;

- Assess use of news "collections" and service:
  - Develop reports and summaries of database use data provided by vendors—become familiar with which databases, which news titles, etc. are being used most frequently—follow up accordingly;[xii]
  - Develop a user survey of in-house use of news collections;
  - In addition to room counts, count people who are seated in the television news viewing area, those reading newspapers or news magazines, which newspaper;
  - Capitalize on changes in the newspaper industry—once free content is still free in the Library's newspaper databases.

Conclusion

As academic libraries seek to meet the research and teaching needs of their faculty and students libraries must seek to establish their key role in providing access to (and preserving) newspapers and news content. Library users increasingly expect news content to be delivered to them, to be electronic/digital, and are rarely aware of the complexities of providing access to news content.

Clearly, it is time for libraries to step fully into the news information environment and begin to assert their role in providing content, as well as research and reference assistance for news content. Libraries serve as the primary source for access to news content from around the world. However, libraries must be prepared to promote, market, and educate their users about the availability of their news resources and begin to position their collections as an important and valuable service the Libraries provides to faculty and students to support their teaching and research.

News content will become increasingly fee-based in the coming years, and this certainly presents challenges for libraries. Yet, the fact news companies are charging for content is an opportunity to provide greater awareness of the Library's newspaper databases, their search features and their functionality (as they compare to aggregators and news search engines). The Libraries role may not be in providing access to news content to "read" the news, but certainly libraries have a strong role today and in the future for supporting the "research" needs of their faculty and students but we will have to assert our role through a variety of marketing and promotional efforts.

[1] Rettig, J., et al. (2004). *The many uses of newspapers.* Richmond, VA: University of Virginia, Boatwright Library. (Available at: http://dlxs.richmond.edu/d/ddr/docs/papers/usesofnewspapers.pdf).

[1] Such as: The Pennsylvania State University Libraries' *Historical Digital Collegian Archive 1887-1987.* (Available freely at: http://digitalnewspapers.libraries.psu.edu/Default/Skins/collegian/Client.asp?skin=collegian&AW=1302368189542&AppName=2) and the *Chronicling America: Historic American Newspapers (Library of Congress). (A*vailable at: http://chroniclingamerica.loc.gov/).

[1] "[News] is becoming less of a destination that we seek online. Instead we are expecting it to come to us in a social stream." --[Anthony Rotolo, quoted in "World's Largest Social Network: The Open Web," *The New York Tim*es, 5/16/2010]

[1] Cheney, Debora. "If You Build It, Will They Come? One Library's Experience with Converging News Content," *"On the Record": A Global Resources Forum on Electronic Media and the Preservation of News*, (Sponsored by The Center for Research Libraries), New York Public Library, October 23-24, 2008. (Summary available at: http://www.crl.edu/focus/winter-2008-09).

[1] For more guidance on developing a marketing plan see New Mexico State Library. *Library Marketing Plan Workbook.* (Available at: http://www.nmstatelibrary.org/docs/development/planning/Marketing_Plan_Workbook.pdf).

[1] Example of a marketing plan and assessment measures: *Rutgers University Libraries Marketing Plan 2007/08 with measures for success proposed by the Marketing/PR Group.* [updated March 2007]. (Available at: http://www.libraries.rutgers.edu/rul/staff/marketing/docs/marketing_plan_2007.shtml).

[1] Cheney, Debora, et al. "Convergence in the Library's News Room: Enhancing News Collections and Services in Academic Libraries." *College & Research Libraries* 67 (5): 395-417 (2006).

[1] *NewsCAT* (2004-) was developed to help researchers with the most frequently asked reference questions: where to locate a specific news source; what years and what formats are available for a specific newspaper; and what newspapers do we have for a specific time period and/or geographic area. (Available at:

http://www.libraries.psu.edu/psul/researchguides/matbytype/newscat.html) and the *Centre Daily Times* Obituary Database (2010-) an online index to obituaries from the *Centre Daily Times* newspaper from 1937-1995; 2011-. (Available at: http://www.libraries.psu.edu/psul/digital/cdtobit.html).

[1] Examples include the *Try these First* start page for many novice researchers which now includes a key newspaper database (Available at: http://www.libraries.psu.edu/psul/databases/trythesefirst.html) and the *News and Newspapers Research Guide* (Available at: http://www.libraries.psu.edu/psul/researchguides/newspapers.html).

[1] Cheney, Debora, "NML Hosts Coffee at Global Studies Lounge*," InterView (Penn State University Libraries)* 4 (41): 1-2 (Oct. 18, 2010). (Available at: http://www.libraries.psu.edu/content/dam/psul/up/news/documents/Interview10182010.pdf).

[1] Kendall Mattos, "Televisions in the Library? News and Microforms Library provides students access to news as it happens, and that's not all…It has more to offer than you might think" *Interview (Penn State University Libraries* 15(1): 1-2 (Jan/Feb. 2004) (Available at: http://www.libraries.psu.edu/content/dam/psul/up/news/documents/JanFeb04Interview_web.pdf).

[1] "Connecting the Dots: How Researchers Use Their Library's News Resources," in IFLA International Newspaper Conference, 2010 (Delhi, India). *Digital Preservation and Access to News and Views: Conference Papers*, edited by Ramesh C. Gaur, et al., pp. 68-89. (Available at: http://www.ignca.nic.in/ifla2010/IFLA_PDF/IFLA_conf_papers.pdf).

Library Services: Strategies used in marketing and promotion of news

By Nor Hazidah Awang

(Universiti Malaya, Malaysia)

Nor Hazidah Awang

Universiti Malaya

Malaysia

Nor Hazidah Awang is a librarian of the University of Malaya Library. Hazidah graduated from University of Malaya, Kuala Lumpur in B.Comp. Sc. She holds a master degree in Library and Information System from the same university. Currently, she is serving under Academic Service Division.

"UM in the News": Marketing and promotion in social media platforms

N.A, Hazidah[1]; N.N, Edzan[1];I, Y, Mahbob[1];Hasmawati[2]

[1] University of Malaya Library, Kuala Lumpur, MALAYSIA

hazidah@um.edu.my, edzan@um.edu.my, mahbobyusof@um.edu.my

[2] Universiti Kuala Lumpur, MALAYSIA

hasmawati@iprom.unikl.edu.mu

Abstract

Today's technology has made it possible for news to be disseminated in the quickest manner. Users, especially from the higher learning institutions, need to stay connected with news in order to complete some of their learning processes. The aim of this paper is to analyse the strategies used in marketing and promoting of news by the University of Malaya Library (UML). The effectiveness of these strategies will be analyzed via the number of pages viewed, shares, clicks and viral lift. The study is descriptive in nature in which it will describe the strategies used by the UML to market and promote news and also the effectiveness of these strategies. It is hoped that this study will provide some insights on the marketing and promotion of news by libraries especially at higher learning institutions.

Keywords: Marketing and promotion of news; Social media networks; Marketing strategies; News clippings

Introduction

Advancements in technology and telecommunication have affected broadcasting in a positive manner. Nowadays, e-commerce websites and online newspapers are becoming more common. Published newspapers are still very popular but on the other hand, people are also eager to check the headlines from online newspapers as well (Berna & Attila, 2010). Other factors which contribute to the selection of online newspapers or news websites include: the nature of news is changing because of the digital world's 24-hours news cycle (Robert & Kirsten, 2007); immediacy and virtually unlimited space; and the hyperlinks to the news websites which are potential doorways leading readers into the site. Studies show that link analysis is a novel and useful method that users can gauge online readership and potential impact of news sites (Yijun & Liwen, 2005). Meanwhile, Boczkowski's (2004) study reveals that online newspapers appear as one of the digital genres that have attracted the most research attention.

What is news? The traditional view is that news is information about recent events or happenings, especially reported by newspapers, periodicals, radio or television. News is the presentation of such information, as in a newspaper or on a newscast, while, an online newspaper is a newspaper that exists on the World Wide Web or the Internet.

Carlos' (2009) exploratory research discovered that newspaper reading is motivated by five key factors: knowing current news bringing forth knowledge that enables readers to interact with family and friends; searching for specific information such as stock prices and results of sport competitions; keeping informed about breaking news or current affairs; making it a leisure or entertainment activity; or being a habit where by reading a newspaper at a particular time or location has become part of many people's daily routine.

University of Malaya Library's initiative in marketing and promotion of news

University of Malaya in the News

The University of Malaya Library (UML) has made it possible for news to be disseminated within the library website. UM in the News was initiated by Dr. Nor Edzan Haji Che Nasir, UML's Chief Librarian in 2008. Among the five Research Universities, UM Library is the only University library that provides with this application. It also allows users to share the news by just clicking at the "share" button provided in each news clippings.

"UM in the News" is news coverage relating to the University of Malaya (UM) and its staff, students and alumni that have appeared in the local, regional and national newspapers. It also provides newspaper reports and pictorial coverage of events, announcements and achievements of the university. It can be accessed via http://umlib.um.edu.my/newscut.asp. The primary objective of "UM in the News" is to collect and housed all news published in the newspapers relating to UM in order to enhance the Library's services towards institutional repositories development. However, as for now, UML only uploads current news clippings.

Data generated from Google Analytics shows that the number of hits for the news clippings vary from month to month (Figure 1) but the average hits per month is 3,465.6. The highest hit is in August (4,528) and the least hit is in May (2,355). This means 2,930 pages were viewed a total of 41,587 times in the year 2010.
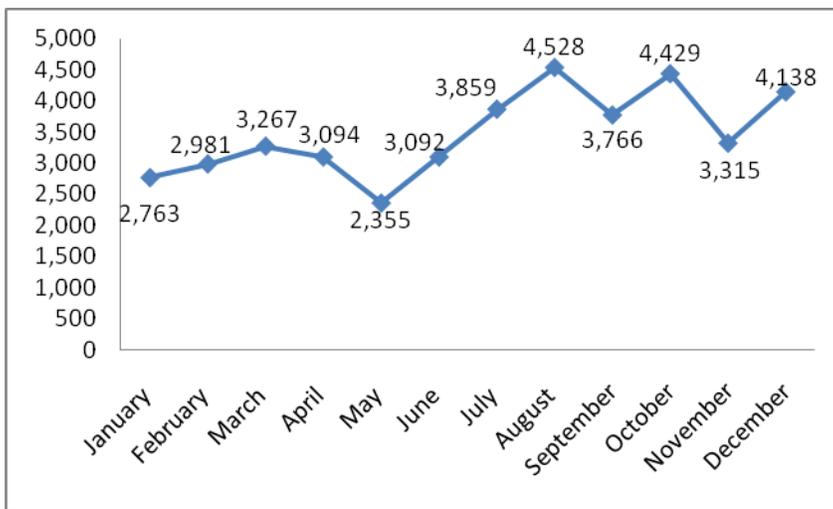
Figure 1: UM in the News: Monthly Pages Viewed (Jan – Dec 2010)

Geographic trends

In order to plan marketing and promotion strategies, it is necessary to know about geographic trends. Table 1 illustrates the top 10 ranking of countries according to their shares on the UML's news clippings. The data portrayed in this table pertains to November 2010 until February 2011.

Top Countries	Shares	Clicks	Viral Lift
Malaysia	210	354	169%
United Kingdom	3	21	700%
Indonesia	3	2	67%
United States	3	4	133%
Japan	2	19	950%
United Arab Emirates	1	0	0%
Singapore	0	14	0%
Australia	0	7	0%
Hong Kong	0	3	0%
New Zealand	0	3	0%

Table 1: Geographic trends on sharing UM Library news clippings – top 10 countries

Strategies used in marketing news clipping and the effectiveness of these strategies

Several strategies have been used to facilitate the users in accessing the news clippings. These include:

a) Highlight latest news clippings on the first page of the UML's website

With this strategy, users can view the latest news clippings uploaded without having to click on the particular site, unless they want to read more news clippings. Reports generated by AddThis Analytics show that 55% of the users access news clippings from UML's website domain (Table 2).

Top Domains	Percentages
www.umlib.um.edu.my	55%
www.google.com.my	20%
www.facebook.com	12%
www.google.com	4%
Search.yahoo.com	2%

Table 2: Top domains

b) Provide search browsing in several modes – Users may use several browsing modes such as search by subject, by month and by newspaper categories.

Sources	News Clippings	Subjects	News Clippings
Berita Harian	629	Academia - Opinion	150
Utusan Malaysia	450	Students - Activities	131
Sin Chew Daily	445	UM- Alumni	112
The Star	249	UMMC- Services	96
New Straits Times	208	Ranking	56
Sunday Star	120	UM- Lecturers	55
New Sunday Times	115	UM- Research	55
The Sun	46	UM- Activities	52
Berita Minggu	44	Educ, Higher- M'sia	46
Makkal Osai	43	UM- Commentaries	41
Mingguan Malaysia	42	UM- Convocation	36
Malay Mail	41	Crime Malaysia	35
Harian Metro	30	Students - Elections	34
Kosmo	24	UM Indian Std Dept	32
Utusan Melayu	10	Academia Award	32
New Sabah Times	5	Conference	31
Lain-lain	3	THE- QS	30
Access	2	MOE	28
Nanyang Siang Pau	2	E-Voting	28
Star Metro	1	Job Hunting- M'sia	28

Table 3: UML's news sources (as at 24th January 2011)

Table 4: UML's news clippings top 20 subjects (as at 24th January 2011)

The UML's news clipping are from different newspapers (Table 3). Most of them are from *Berita Harian* with a total of 629 news clippings. *Utusan Malaysia* ranked second in reporting news about the university (450 news clippings) and *Sin Chew Daily* ranked third with 445 news clippings. The rest of the newspapers which are significant in reporting about the university are *The Star*, *New Straits Times*, *Sunday Star* and *New Sunday Times*.

In terms of subject classifications, the news clippings are mainly related to Academia's Opinions (150), Students' Activities (131), Alumni (112) and UMMC Services (96). These are the most significant subjects that appeared in the newspapers. The rest are less significant involving subjects such as Lecturers, Researches, Activities and Commentaries (Table 4).

c) **Social sharing** – Users can share the news clippings on the social media networks by clicking on the "sharing widgets" provided by AddThis. AddThis is a social bookmarking tool that is easy to implement and is able to help in planning and provide a clear picture about the effectiveness of social medias as marketing and promotional strategies. Table 5 illustrates the majority of the users (119 shares) use "Tools - print" for sharing purpose. Facebook is the most popular social media network used to share information with 631% viral lift compared to the other social media, where by this can be considered as a successful and effective strategy. Facebook generated 404 clicks from 64 shares (Table 5). Most of this Viral Lift (called V-Lift) are from visitors who clicked on the news clippings from their friends' Facebook. Of the figure, only a few are from the email but none came from Google, Twitter, MySpace, Delicious and GMail. In summary, Table 6 shows the top services by the social media category.

Top services	Shares	Clicks	Viral Lift		Top services Category	Shares
Print	119	0	0%		Tools	119
Facebook	64	404	631%		Social Networks	67
Email	23	29	126%		Email/IM Service	24
Google	7	0	0%		Bookmarking Site	8
Twitter	4	0	0%		Blogging Platform	4
MySpace	3	0	0%			
Delicious	1	0	0%		Time period: Nov 2010 – Feb 2011	
Gmail	1	0	0%		Table 6: Top services by category	

Time period: Nov 2010 – Feb 2011

Table 5: Daily shares to top services

d) **Social medias as a marketing tools** - A part from these strategies, the UML also use social media networking such as Facebook, Delicious and Wordpress as a platform to market and promote news clippings by manually posting the link from the news clipping and providing permanent links on the social medias. These social media are also used to disseminate information on opening hours, latest services offered, notifications on services or internet disruption, information about online databases, sharing knowledge, recommendations of reading materials, announcement on the latest events in UM or UML itself. Hence, users or social media habitués can easily access the news clippings through the network services. To expand its reach, UM's website also provides permanent links to the "UM in the News". This is another way for UML and the UM to disseminate news clippings. Users can surf the site via http://www.facebook.com/pages/ The-University-of-Malaya library/ 28382147400.

As for Delicious, news clippings can be inserted directly by clicking the icon found on every news clippings. Users are free to add the news clippings into their Delicious account. User's can also tag the uploaded material. However, whatever is tagged, would require moderation from the librarians. The librarians will approve or disapprove the tags and upon approval, the tag can be added to the record.

Future Initiatives

There is a need for the UML to utilize other strategies and available social media networks tools to disseminate news about the University other than the ones mentioned above, in order to enhance user's accessibility to "UM in the News" in the near future.

a) **Search engine optimization (SEO) –** News clippings search will become easier with a search engine. Therefore, UML is planning to provide a search engine in "UM in the News". Users do not have to search the news clippings manually. All they need to do is just type-in a keyword or a title on the search engine's browser provided and they will get relevant hits.

b) **Expend content by uploading archived news clipping -** In order to expend the Library's collections towards web archiving, archived news clippings will also be uploaded into "UM in the News". Thus, users may access old news clippings online without having to come to the Library.

c) **Text version for mobile devices –** UML is also planning to provide text version for mobile devices to disseminate the latest news clippings. It is convenient for the users to read the news directly on their mobile devices in a text version.

d) **Twitter -** In the first quarter of 2010, 4 billion tweets were posted (Wikipedia, 2010). With the vast number of ongoing conversations, the UML can use Twitter to its advantage. Users can quickly share information with those who are interested in the UML's news and services, follow up with status updates and also build relationships. The viral effect it has generated over the years makes Twitter an effective social media for posting news daily headlines and update UML's followers. Users just need to find the public stream and follow the conversation of their interest.

e) **RSS (Really Simple Syndication) -** RSS is a web syndication that is a family web feed format used to publish frequently updated works such as blog entries, news headlines, audio, and video. As of 27 January 2011, the syndic8.com website has been indexing 1,404 (out of 630,990) feeds with subject tags and 4,597,192 items with subject tags. UML can apply this web syndication which will allows users to syndicate their site content besides providing an easy means to share and view headlines and contents. It also can be used to inform users that the news clippings have been uploaded and allows personalised views for different sites. In addition, it lists news with title, date and description which makes it convenient for users.

f) **Tag clouds –** tag clouds offer a more convenient and intuitive navigation for "UM in the News". They can provide users with an instant illustration of the main topics, giving a very specific and precise orientation of the news clippings' content. Since human beings tend to think in concepts and models, it is easier to get an idea of the content if the main concepts are given straight away — in digestible pieces, and prioritized by their weight. In fact, the main advantage of tag clouds lies in their ability to highlight the most important and/or popular subjects dynamically which is not the case in conventional navigation menus.

g) **MyWeboo -** MyWeboo is social news that has an interesting concept because it allows users to manage their web files from various social media networks all in one place. Since it is working with Facebook, YouTube, MySpace, Flicker, Photobucket, Picasa, Delicious, Wretch, Amazon S3 and Your PC (Windows), UML may take this opportunities to market and promote news clippings as it functions like a one stop manager that makes it easy for users to control web files from various social media networks. Since its official launch in June 2010, MyWeboo has attracted a lot of users and is now growing rapidly.

Advantages and limitations of social media networks for online marketing

Online marketing is inexpensive when one examines the ratio of cost to the attainment of the target audience. UML can reach a wide group of users for a small fraction of traditional advertising budgets. The nature of the medium allows users to read the news clippings conveniently. Therefore, UML has the advantage of appealing to users in a medium that can bring results quickly. However, the strategy and overall effectiveness of marketing campaigns depend on the library goals and users' satisfaction analysis.

In addition, the UML's management also has the advantage of measuring statistics easily and inexpensively; almost all aspects of an online marketing promotion can be traced, measured, and tested. UML can use a variety of methods, such as RSS feed and send to personal users' emails. Therefore, UML can determine which news are more appealing to the users. The results of promotion can be measured and tracked immediately because online marketing initiatives usually require users to visit a website, and to perform a targeted action. Meanwhile, there are some limitations to online marketing especially when the service providers terminate their services. Users are no longer able to access the applications. Furthermore, while social media networks can be a very useful tool, there are copyright and licensing, institutional factors, and the providers' own policies to keep in mind. Last but not least, in order to provide efficient and up-to-date service, competent staffs are required to manage the related tasks.

Conclusion

Unlike in the past where people were only depending on newspapers, television and radio for news, in recent times, news can be obtained through various media. The Internet has made it possible for people to get updated news of events anytime and anywhere they are, as long as they are connected. Social media networks create another dimension for people not only to stay in touch with their associates but it can also be use to promote products and services. The viral effect that the social media networks developed, has made it possible for services to be quickly made known faster than the old ways of advertising. In fact, a great number of people use social media networks to interact with each other. Consequently, news or news clippings should be made available through the social media networks for prompt and widespread dissemination of news by libraries. Through various social media networks, a variety of information can be shared and readers may choose the one that suits their preference whether local, regional or international. However, not all social media networks can be applied because it requires a strong commitment from the Library. Libraries are free to choose any of the suggested social media networks in order to market and promote their news.

References

Berna H. U., & Attila, A.I. (2010). A novel attribute-based dynamic content area layout for internet newspapers. *Internet Research, 20*(2),135-153. Retrieved from www.emeraldinsight.com/1066-2243.htm

Boczkowski, P.J. (2004). *Digitizing the News: Innovation in Online Newspapers*. MIT Press: Cambridge, MA.

Carlos, F., & Raquel, G. (2009). Users' motivations and attitude towards the online press. *Journal of Consumer Marketing, 26*(3), 164–174.

Robert, B. A., & Kirsten, A. (2007). Preserving digital local news. *The Electronic Library 26* (3),387-399. Retrieved from Emerald online databases.

Yijun, G. & Liwen, V. (2005). *Web hyperlink profiles of news sites: A comparison of newspapers of USA, Canada, and China.* Retrived January, 12, 2011, from www.emeraldinsight.com /0001-253X.htm

Wikipedia - Social media. (2010). Retrieved January 12, 2011, from http://en.wikipedia.org/wiki /Social_media#Examples

Digitised Newspapers – Wider Availability through Collaboration with the Copyright Societies in Finland

By Majlis Bremer-Laamanen

(National Library of Finland)

Majlis Bremer-Laamanen

National Library of Finland

Majlis Bremer-Laamanen is a Director of the Centre for Preservation and Digitisation at the National Library of Finland. She has been Project Leader and Work Package Leader in several national and international projects. Some of her ongoing memberships:

- National: National Digital Library: Availability Section, The National Library Management Group,
- Mikkeli University Consortium: Management group (vice chair) and Advisory Board, Digital Mikkeli Management group.
- International ones: IFLA Newspaper Section, EU Member States´ Expert group on Digitisation and Digital Preservation.
- LIBER Steering Committee on digitisation and resource discovery.

Some of her Publications and Conference Papers

- Bremer-Laamanen M. (2010): *Newspapers as a Digital Resource. National Infrastructure and Newspaper Publishers – in Finland.* Digital Preservation and Access to News and Views. (Ed. Ramesh C. Gaur, Frederick Zarndt, D. R. Gupta, Kavita Gaur.)
- Bremer-Laamanen M. (2008): *The Present Past – the History of Newspaper Digitisation in Finland.* Newspapers collection management: printed and digital challenges. (Ed. Hartmut Walravens.)
- Bremer-Laamanen M. (2008): *Widening Access and Legal Issues: Newspapers in Focus.* The Impact of Digital Technology on Contemporary and Historic Newspapers. (Ed. Hartmut Walravens.)
- Bremer-Lamanen M. (2006): *Connecting to the Past – Newspaper Digitisation in the Nordic Countries.* International Newspaper Liberianship for the 21st Century. (Ed. Hartmut Walravens.)

Digitised Newspapers – Wider Availability through Collaboration with the Copyright Societies in Finland

Majlis Bremer-Laamanen, the National Library of Finland

Abstract

The National Library of Finland is collaborating with the Finnish copyright society KOPIOSTO (including representatives of publishers) to achieve wider accessibility and availability for newspapers. Delegates from the Ministry of Education, KOPIOSTO and the National Library have a general interest to promote this collaboration. Promoting access to legal deposit newspapers that are in copyright through extended collective licensing: The Ministry of Education can grant the Finnish copyright societies (e.g. KOPIOSTO, TEOSTO) permission to administrate extended collective. The copyright societies are to work with the copyright holders to determine the framework for licensing. The extended licensing solution is based on legal requirements and statutory effects. This form of licensing is generally binding for copyright holders belonging to a specified group. The access to the material is specified separately.The extended collective licensing allows the copyright society to manage the copyrights also when the copyright holders cannot be reached which makes it both effective and cost effective way to achieve wider accessibility for the digitised materials.

The goal: to have the first pilot project in place in 2012.

1. The Context
    1. The Setup for Newspaper Availability
2. Extended Licensing
3. Comellus – Phase One – Funding in Place
4. Project Plan for Phase Two
5. Closing Up

---

1. THE CONTEXT

The National Library of Finland is collaborating with the Finnish copyright society KOPIOSTO and the representatives of newspaper publishers to achieve wider accessibility and availability for 20th century newspapers.

Newspapers have been and still are of great importance in the Nordic Countries, like Finland with its 5.4 million inhabitants. In terms of newspaper sales per 1000 adults, Finland is in third place following Japan and Norway.

The National Library has collected all printed materials published in Finland since 1640. The first Act on Legal Deposit from 1707, including newspapers, has been changed several times. Online newspapers, but not electronic facsimile newspapers, can be obtained only by harvesting as the Act on Legal Deposit does not cover these.

1. The Setup for Newspaper Availability

There has been a growing interest for making 20th century newspapers more freely accessible. In Finland this development has taken place in a forerunning group of newspaper publishing houses, in the Copyright society KOPIOSTO and it has been facilitated by a recent legislation change covering Extended Collective Licensing. Naturally this is also of interest for the National Library as it has already made all the copyright-free newspapers published in Finland 1771-1909 freely available on the net. We have about 550 newspaper titles digitised at the moment.

In Finland, we are able to digitise the legal deposit collections and share the digitised collections with the other six legal deposit libraries in the country based on legislation from 2008. This situation is better than in many other countries and has made it worthwhile to digitise for instance journals on a larger scale during the past few years.

2. EXTENDED LICENSING

In the Nordic countries Extended Collective Licensing is one of the means to make a wider use of digital collections possible.

Solutions based on the Finnish law of Extended Collective Licensing (16d§) are founded on legal requirements and statutory effects. The Copyright society (e.g. KOPIOSTO, TEOSTO for music composers) administrates licensing after getting permission for licensing from the Ministry of Education and Culture. This form of licensing is generally binding for copyright holders belonging to a specified group even if the copyright holders cannot be reached.

In the case of Extended Collective Licensing much effort is taken to reach the rightsholders. The rightsholders set the framework for the use and the conditions together with the Copyright Society. Naturally rightsholders can deny the use of their works. The access to the material is specified separately.

The Extended Collective Licensing agreement for KOPIOSTO covers:

- works of journals and newspapers and their rightsholders
- works of ephemera and their rightsholders
- works of photographic art and their rightsholders as well as photographs and photographers

Books are excluded from the agreement.

3. COMELLUS – THE FIRST STEP – FUNDING AND PUBLISHERS IN PLACE

To be one of the key players in the information society, the National Library has to do the right strategic choices. In order to make the 20th century newspaper collections digitally available and enhance long term preservation we have to awake the interest of the publishers to:

- digitise their collections up to this date
- receive their electronic facsimile version of the daily published newspaper
- rise an interest for long-term electronic preservation and
- make Computer Output Microfilm reproductions of the electronic versions for "eternal" storage

The name of the project describes its content:
COM – computer output microfilm
ELLU – electronic delivery
S – preservation

In December 2010, the National Library and its Centre for Preservation and Digitisation, got funding – 650.000 € for a three-year project (2011-13) from the European Social Fund. We have two regional newspapers participating: Länsi-Savo (Western region of Savo) and Etelä-Suomen Sanomat (Southern Finland). A project manager, two IT-specialists and one operator will be attached to the project.

In Helsinki, the Library has two functional units, one handling the collections and the other handling infrastructures for the library field and for the memory organizations. Our Centre for Preservation and Digitisation is in charge of preserving and digitising the library collections and for the development in these fields.

Receiving electronic legal deposit material is a hot development area in our library today. As the newspaper processes for legal deposit, collecting bibliographical information, microfilming and digitisation have been managed in the Centre, the need for taking steps to change microfilming processes to COM, and to enhance electronic delivery is seen as the next step in the development. Here all the departments in the library are working very closely together.

As a result the newspaper process will be taken into the digital age. The use of paper versions will be diminished. And the electronic delivery system will be put in place, between the newspaper houses and the National Library.

In fact the newspaper process will see a total change. In the new process the lifecycle of the newspaper will be administered from the electronic delivery to metadata questions, the preparation of electronic material to at last the "eternal" preservation on microfilm by COM.

Alongside with this project the library will digitise the newspaper holdings of Länsi-Savo and Etelä-Suomen Sanomat from 1910-2000 (2010), about 2 million pages.

4. PROJECT PLAN FOR PHASE 2

The next step is to extend the project to other publishers, address the copyright questions and make the information available.

1. *Newspaper Publishers Show an Interest*

Other interested newspaper publishers have been in contact with us in order to digitise their content and make it available. As the National Library has been high quality microfilming all newspaper collections from 1771 to this day, we have the means to digitise newspaper collections cost efficiently with quality. Perhaps astonishingly some newspapers need to digitise their publications even up to 2010.

In Finland, the Ministry of Education and Culture is in charge of the National Digital Library – project for the memory organizations. As a result, a Public Interface is put in place this year at the National Library and a Long Term Preservation (LTP) system will be put in place by 2016. These solutions are interesting for the publishers as well. Metadata questions and long term storage solutions in our digitisation process at the National Library's Centre for Preservation and Conservation attract publishers as well.

KOPIOSTO has for a long time been collaborating with newspaper and other publishers. Finnish Newspapers Association has taken an interest for the Extended Collective Licensing and its possibilities. The newspapers are interested in the possible visibility of the newspapers in the digital world alongside with a wish for some revenue. However, here the role of newspapers is primarily the one of a cultural player.

2. Goals

Building on the Comellus project we will extend the newspaper titles to five or six. The newspaper houses are chosen according to their interest to go digital and extend their influence in the cultural world. We are also looking at their status and know-how and possible other contribution to the project. We would like the pilot to be of large interest for the readers and set an example for the other newspaper publishers – and perhaps for society at large.

In this pilot plan for Phase 2, we will include:
- Two or three national newspapers: eg. Uusi Suomi
- Two or three regional newspapers: the Comellus project papers and one new paper
- We will digitise about 2-4 million pages and make them more accessible

Funding is primarily sought from the government. The newspaper houses or their background societies and the National Library might be able to participate. Other sources of income are also sought for. Other resource not to be overseen is the workforce of the participants.

3. Questions

There are many issues to be dealt with. Alongside with this project, new issues for availability and access of the 20th century collections are rising within the EU. Questions regarding copyright matters, funding and income revenue are still not solved. Some guidance for these questions is available in the EU Report of the "Comité des Sages" 10 January 2011. Discussions with eg. newspaper publishers is continuing on other arenas of copyright issues. These discussions might also reflect our project planning.

For this project we have distinguished the roles of the participants. The newspaper publishers, journalists and others are the right holders. The National Library has legal deposit to the newspapers and will digitise the newspapers from its quality microfilm collections. The Library will preserve the material and make it as available as possible. KOPISTO collects authorization from the right holders and gives them revenue.

Extended collective licensing is a form of collective rights management whereby extended collective licensing laws allow for freely negotiated copyright licensing contracts for the exclusive rights granted by copyright. These agreements by law extend to rightsowners who are not members of the collecting society agreeing the license with a user.

Agreements are made for access and use for:
- research
- teaching and study
- for the large public via public libraries
- in other ways restricted use like newspaper subscribers

For these agreements the ways of use can include:
- permission to look at and print the information from your personal computer
- to visit the information channel via a given password
- an annual permission for the organization

Agreement and technical considerations like the structural analyses of the material, distribution channels and long term preservation options including possible electronic delivery will be taken into account.

5. CLOSING UP

The pilot "access" project will begin in 2013 if funding is available, and later continue on a larger scale. We will put in place three working-groups that deal with copyright questions, the technical issues involved including digitisation, structural analyses of the material, metadata and long term preservation and interface discussions. But first we are discussing funding models to set the stage for the pilot.

We are of course not living on an island. Within EU there are many views on copyright issues, in legislation and between rightsholders and users like libraries. Discussion is continuing. We wish that we could set a good example of cooperation between the newspaper publishers, rightsholders and the copyright society. This project would share some light on the black hole of the 20th century for you and me.

Selection of newspapers for digitization and preservation as a result of present practice in the French National Library

By Else Delaunay

(Bibliothèque nationale de France)
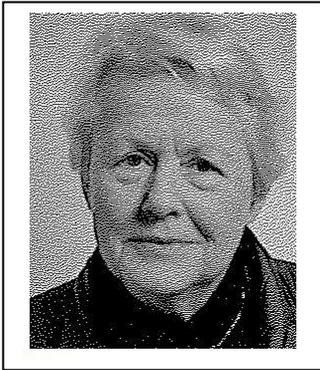
Else Delaunay

Bibliothèque nationale de France

Else is a former responsible for the General Catalogue of Periodicals, later for the Preservation and Microfilming Office of the former Periodicals Department at the Bibliothèque Nationale de France, now participating in the *Bibliography of French Local Newspapers* published from end 18th century until 1944. Member of the IFLA Round Table on Newspapers, since 2003 consultant of IFLA Newspapers Section.

Selection of newspapers for digitization and preservation as a result of present practice in the French National Library[2]

Else Delaunay (Bibliothèque nationale de France)

Abstract

The paper will give a brief overview of selection criteria and the reasons for these options as practised right now at the BNF (Bibliothèque nationale de France) for newspapers to be digitized and/or preserved in hardcopy. BNF's newspaper digitization programme which has priority will be introduced. Main selection categories are: French historical newspapers, newspapers published in French abroad, rare files only preserved by the BNF, very brittle newspapers, specialised items such as newspapers from ancient French colonies. Coloured newsprint will normally be digitized in grey because of the high cost of digitization in colours. Funding in 2011 comes to 200 000 Euros plus the digitization carried out in house, an insufficient amount compared with the millions of pages still to be digitized. At present, the BNF is negotiating with providers having huge digitization possibilities. As to current newspapers BNF's Legal Deposit Office is negotiating with several newspaper industries but it is too early to know about the conclusion. A test is running actually with 75 newspapers with online editions. Hardcopies are preserved at present but the digital format will predominate and probably take over in a near future.

I.   Introduction

In a digital era why should we keep hardcopies which are digitised and/or microfilmed and why should we preserve the original negatives when a given title has been digitised? New attitudes towards these questions seem to occur to day, all the more so, as electronic archiving still is problematic. Preservation of hardcopies and master negatives in optimum conditions is then of prime necessity. For a long time yet the hardcopy will be considered as a last resort.

---

1. *Reasons for preservation of the hardcopy after digitization*

- The hardcopy is the original item as it was published and circulated. For legal deposit libraries it is part of national heritage.
- The hardcopy may be needed for very specialized and detailed research work.
- The hardcopy is the best item to preserve as it is always possible to restore paper and it is therefore considered as a last resort.
- The digitized item may also turn out to be damaged (gaps, fuzzy print, etc.) so another digitization would be needed.

2. *Preservation conditions for hardcopies*

Newspapers need a lot of storage space, particularly expensive in big towns. They should therefore be kept in stacks off site the library (ex. a silo), meeting the best environmental conditions (air, room temperature, relative humidity, clean and dust free premises, etc.(ex. in Northern Norway at Mo I Rana, the storage of newspaper hardcopies is situated inside a mountain with optimum preservation conditions).

- Damage factors: wear and tear, self-destruction of acid machine made paper, low quality paper printed with low quality ink, pollution, fire, water, infesting with insects, etc. All these factors should be considered when building new stacks or re-equipping old stacks.

- Maintenance of the holdings

  The maintenance is essential for the good preservation of the items and their protections. The staff in the stacks needs to be particularly involved, as it is in daily contact with the holdings and can take responsibility for various checks such as:

  - detecting of all anomalies or deteriorations: supervision of the stacks and of the items sent to the reading room;
  - regular checking of thermo-hygrometers if not already checked by the Preservation Office; dust removal;
  - stocktaking, necessary moves of the holdings: estimates of the growth of the holdings are particularly tricky for newspapers

- ▪ <u>Parameters for optimum storage facilities</u>

  - cleaning and any necessary disinfection of the premises; airing
  - room temperature between +16° and +18° C (in warm countries between +20° and +22° C); relative humidity: 50 to 55 %
  - light level: not more than 50 lux
  - air conditioning, checking of pollution and dust
  - permanent checking of risks (fire, flood, infesting, robbery …)

- ▪ <u>Equipment of the stacks and transport</u>

  Traditional or compact shelving; boxes or bound volumes of large size newspapers to be stored flat on the shelves with 3 boxes or volumes per shelf; transport of the items can be manual or automatic.

3. ***Restoration of brittle or damaged issues***

Manual or mass deacidifying followed up by thermosizing/lamination or leafcasting as appropriate*.*

- • <u>Manual deacidifying</u>**:**

  Each sheet is immersed in a borax water solution, then dried in the open air or on special crates. This method is still used for newspaper sheets at the BNF. Some 400 000 sheets which will not be microfilmed nor digitized are deacidified each year.

- • <u>Mass deacidifying</u>**:**

  The Wei T'o process was developed in Canada. The base product is the methylmagnesium carbonate.The Bookeeper process developed in the USA is used by the Libtary of Congress. In France the Bookeeper process has been used to deacidify running regional dailies (1945-1970). At the end of the process an alkaline buffer is deposited directly on the fibres. This process may give newspapers in a reasonable physical condition an additional lifespan of 30 or 40 years. At present no mass deacidifying of newspapers at the BNF.

4. *BNF's Newspaper Digitization Programme*

This programme has priority within the general digitization programme of the library. Most of the 19th century national newspapers are digitized and included in Gallica[3], BNF's online library, but a great deal is still to be achieved.

- **Selection criteria**

  As a general rule the BNF will only digitize items received by legal deposit.

  - <u>Priority categories for historic newspapers</u>:

    - French national, regional or local newspapers

    - Newspapers in French published abroad (ex. Egypt, North Africa…)

    - Newspapers in foreign languages issued by ethnic groups, published in France and submitted to legal deposit

    - Rare French files only preserved by the BNF and not yet microfilmed (ex. local or regional newspapers, papers from ancient French colonies), French newspapers in bad physical condition frequently required by the users, and very brittle French newspapers needing partial restoration prior to digitization

    - French specialised newspapers such as trench papers from World War 1, underground newsprint published during the Occupation of France (1940-1944), etc.

---

[3] http://www.gallica.bnj.fr

- Heavy used magazines: news (*L'Illustration* (1843-1947)*, Le Magasin pittoresque* (1833-1952)…) ; travelling (*Le Tour du monde*, 1860-1914) ; fashion (*Journal des dames et des demoiselles* (1859-1902), *La Mode pratique* (1891-1939), *Modes et travaux* (1919→) ; children's magazines (*Le Journal de Pif* (1965→), *Le Jeudi de la jeunesse* (1904-1914), *Tarzan* (1941-1953)…

Coloured papers will normally be digitized in grey because of the cost which is three or four times higher than digitization in grey.

- ▪ <u>Priority categories for current newspapers</u>:

  - Regional newspapers with local editions

    The Legal Deposit Office receives at present 2 newspapers online with local editions but there is no guarantee that the online editions include all contents of each issue so what about preservation? The current online web newspapers available do not replicate the printed versions and therefore exclude much valuable research material including advertisements and illustrations and of course the layout/content of the printed version.
    A copy should be a true image of each issue. The BNF does not keep any hardcopy of the local editions still published in paper format. Access to some of them will soon be online on the website only

    Negotiations are going on between the Legal Deposit Office and some regional newspaper publishers but it is too early to know about the final result. A test is running actually with 75 newspapers having online editions. The BNF is trying to know if some or all of these newspapers have a PDF file which might be fit for use of the library, and how complete these files are as to the contents and look of each newspaper issue.

Newspapers with an online web edition are collected already through the web harvesting at the rate of one capture per day for national and regional dailies and one capture per week for weekly newspapers. These online editions may vary throughout the day due to regular updating. If no PDF file available, the BNF might decide to harvest the web edition only. This is possible from a legal point of view as French law allows free harvesting of the French websites. Nevertheless, technical difficulties may arise but the publishers cannot prevent the use of the online editions.

At the BNF these editions would be accessible only in the reading rooms in the ground level (Research library) via the way back machine. All other material (ex. journalists' blogs) on the French websites is captured once or twice per year and entered directly into the archival system SPAR[4].

This means that the BNF would keep only the web edition of a French regional or local newspaper and not anymore require a complete electronic copy. It would create a real change of the law on legal deposit for such items. Next step could be not to preserve the hardcopy of the main edition anymore, even though it is still published in paper format! The library would then pass on to virtual preservation only of the online editions and so trust BNF's archival system of digitized items SPAR. But we are not there yet!

The rule is still that as long as the hardcopy editions of national and all other French newspapers (except local editions) will be available, the BNF will keep them according to its duty as a national library to preserve all French printed matter. Nevertheless, though hardcopies are preserved at present and used by the readers, the digital format will predominate and probably take over in the near future.

Some legal difficulties may also occur as newspapers include many personal names which can easily be linked in an electronic edition while this is much more difficult in a hardcopy edition. So the National Commission for Information Sciences and Freedom[5] may step in.

---

[4] http://www.bnf.fr/en/professionals/preservation_spar/s.preservation_SPAR_presentation.html

[5] CNIL (Commission Nationale de l'Informatique et des Libertés)

- **French born digital newspapers:**

  They are collected through web harvesting (ex. *Rue 89*, *Bakshish*) but only once or twice a year. It is a digital legal deposit[6]. The scope of legal deposit was extended to cover the Internet by the Law of August 1st 2006. Websites belonging to the French national domain can be collected by the BNF, preserved and made available to the public.[7]

- **Preparing of the items to digitize and writing of specifications**

  - Completing of the files

    When a complete file is not available, missing issues should be sought for through interlibrary-lending (often a time-consuming task), or by photocopying. In some cases it may be necessary to use some issues from another chronological edition in order to obtain a complete file. If so, these issues should be clearly announced in the bibliographic target. Because of the cost of digitizing, it is important to prepare as comprehensive files as possible. But later completing of the run of a digitized newspaper is quite easy.

  - Brittle items:

    They must be restored if needed before any kind of reproduction. When preparing dry newspapers in bad physical condition for digitization, humidifying and ironing of the sheets are necessary. It will strengthen the paper but in many cases restoration of some issues will be needed.
    Due to lack of funding the BNF does not anymore restore newspaper sheets if not needed for reproduction (digitization or microfilm).

---

[6] http://www.bnf.fr/en/professionals/digital_legal-deposit.html
[7] See also : http://www.bnf.fr/en/professionals/digital_legal_deposit/a.digital_legal_deposit_cooperation.html

- Specifications[8]

    It is important to elaborate precise specifications for newspaper digitization. In 2010 the BNF, the Public and regional Library in Orléans and the Ministry of Culture and Communication put online a guide entitled: *How to write specifications for newspaper digitization and conversion into text mode*. It is a technical guide prepared by the Pilot Committee on Digitization, Ministry of Culture and Communication, Bibliothèque nationale de France, Bibliothèque municipale classée d'Orléans, 2010, 28 p. and a supplement of 29 p. The guide is available in Word format and in PDF.

- **Funding strategy**

    At the BNF funding for newspaper digitization in 2011 comes to 200 000 € plus the preservation digitization carried out in house, some 50 000 € per year. It is of course an insufficient amount compared with the millions of pages still to be digitized. Finances have been cut down by some 3 % this year. The reduction will be the same in 2012 and 2013.

    Digitization is in image mode plus OCR. There is still some backward ocerization to achieve from the first project of 27 national newspapers (19th century and part of 20th century) digitized in image mode only. The total of 600 000 € dedicated to the 3-year newspaper digitization programme will be partly used for the backward ocerization. But the principal goal is to digitize as many newspapers as possible for the amount of 200 000 € available each year.

    Negotiations are going on between the BNF and providers with very huge digitization possibilities (several millions of pages) who may be interested in some kind of cooperation with the BNF.

    Partnerships as well as sponsorships with the newspaper industry and with digitization providers may lead to successful results allowing to undertake and to conclude vast digitization programmes of newspapers.

---

[8] http://www.culture.gouv.fr/culture/mrt/numerisation/fr/actualit/actualit.htm

But such partnerships are very difficult to find, so much more as the BNF is not the only institution in France looking for sponsorships.

The BNF still conducts a rather vast microfilm programme: 100 000 € for preservation microfilming and 800 000 € for microfilming of 20 current regional daily newspapers with local editions, a transitional programme. From 2012 on, these newspapers will be digitized or a PDF file may be available.

5. Shared preservation

The "Associated Poles" (les pôles associés)[9]: a network of libraries and archives in France.

- It is a flexible, coherent and long-lasting system for descriptive sustainable documentary actions. It includes three-year agreement plans together with annual application plans fit for each partner. Actions such as checking, description, digitization and valorization can be linked together. The system is made up of multi-partners who have very different statutes: territorial libraries, university libraries, private association libraries, district archives.
  In 2011, the cooperating network includes 65 poles in charge of legal deposit, 79 associated documentary libraries and 8 regional poles.
  The poles share operations within acquisition, concerted digitization, conversion to computerized data files and valorization of local or national heritage.
  The yearly budget for the period 2009-2011 is 1 260 000 €

---

[9] http://www.bnf.fr/fr/professionnels/cooperation_nationale/a.creation_du_reseau.html

- The "Associated Poles" system facilitates cooperation between the National Library and university or regional libraries in various fields such as shared dynamic management and preservation of the collections, individual or shared reproduction programmes (digitization or microfilm), storage, access, union bibliographies and catalogues (ex. the French Union Catalogue online CCfr[10], BIPFPIG, a bibliography of local newspapers from the 18th century to 1944 published in paper format, should be online in 2012), etc. Thanks to the documentary map of French libraries it is easy to know which subjects are covered by each library. It is a way to avoid useless expenses and redundancy, to get information about programmes ongoing or to be undertaken in other libraries within regional partnerships, to promote interlibrary-lending so as to reproduce complete files, etc.[11]

- BNF's cooperation policy for 2009-2011 has two objects and two operating fields.
  Object 1: contribution to promote and coordinate national digitization efforts, especially for printed items.[12]
  Object 2: support dynamic efforts for developing of written heritage and accompany the action of the Ministry for Culture and Communication (MCC) in favour of the description within concerted actions (Action Plan for Written Heritage)
  Operating field 1: Thematic cooperation
  Operating field 2: Regional cooperation.

  Indeed, historic local newspapers are very much concerned by these digitization programmes. The cooperation between libraries and archives is essential in this field as French district archives preserve huge collections of local newspapers thanks to the law on legal deposit from the printers. The printer of a local newspaper must send one copy regularly to the district archives or to the main district library in charge of the preservation.

[10] http://ccfr.bnf.fr/portailccfr/
[11] http://www.bnf.fr/professionnels/cooperation_regionales/s.poles_documentaires_regionaux.html
[12] http://www.bnf.fr/fr/professionnels/cooperation_nationale/a.gallica_numerisation_partagee.html

- A new project seems to come up: a *Gallica* for newspapers with over 600 historic national and local newspapers online. Newspapers are already included in *Gallica*. But the new project should create a single common portal giving access to all digitized newspapers from BNF's collections and from other partner libraries such as regional public libraries in charge of heritage collections, district archives, etc. It would facilitate access to complete files. By overbuilding of digitized collections from various partners the user would get direct access to a complete file of the newspaper he is looking for.

  The BNF is in charge of the electronic archiving through SPAR, its digital stack based on international standards authoritative in the subject of the sustainability of digital information. SPAR respects the OAIS standard (ISO-14721: 2003). It is a reference model for an open archival information system. SPAR stands for Système de Présentation et Archivage Réparti (Distributed Archiving and Preservation System).

- Shared preservation on an international level may also be considered by national libraries. For instance Europeana[13], the European digital library financed by the European Commission and the member states, a multilingual collection of some 14 million digitized items at the moment but no newspapers yet. The Network of National Digital Libraries in French[14] (Réseau Francophone numérique) has 17 members at present among which the National Library of Cambodia[15], Library and Archives Canada[16], the Royal Library of Belgium[17], National Library of France[18], National Library of Vietnam[19]... Newspapers are included here.

  Indeed, the development of shared newspaper digitization and preservation programmes within the library community in France as well as in Europe or worldwide will hopefully continue to expand.

---

[13] http://www.bnf.fr/en/professionals/dri_partnerships/s.europe_partnerships.html ;
  http://www.europeana.eu/portal

[14] http://www.rfnum.org/

[15] http://www.bibliotheque-nationale- cambodge.com
[16] http://www.collectionscanada.gr.ca
[17] http://www.kbr.be
[18] http://www.bnf.fr
[19] http://www.nlv.gov.vn

6. *Conclusion*

Considering the high cost of preservation and digitization of newspaper hardcopies, the preservation of the national heritage of newspapers requires a close cooperation between libraries and archives, in some cases also museums and other heritage institutions.

It is important to decide which newspapers should be preserved in original format and which hardcopies may have their information preserved in a surrogate format only such as a digital copy. On a national level shared preservation in France including regional and local programmes for restoration and reproduction is organised and carried out through BNF's Department of Cooperation and the Associated Poles. To share the costs and the work seems to be the best way to preserve and save as many newspapers as possible on a national level and even worldwide.

As to born digital newspapers the web harvesting and the long-term electronic archiving are also expensive. But according to present trend the passing over to the all electronic environment may turn up in the very near future for several categories of items, even in a national legal deposit library.

Thank you for your attention

Newspapers and more in TROVE

By David Ong

(National Library of Australia)

David Ong

National Library of Australia

David Ong has been Managing the Australian Newspaper Digitisation Program (ANDP) at the National Library of Australia since December 2009, and is currently Acting Director of Digitisation and Photography Branch. David has had a long association with library systems across National and State libraries. His experience includes Inter-Library Loan and Search and Delivery systems, as well as experience in the specification and development of automated ingest and accessioning applications for digital objects. David is enthusiastic about the ANDP program, and the value it brings to providing researchers with access to fully OCR-ed and indexed content.

Newspapers and more in TROVE

David Ong (National Library of Australia)

Abstract

Via the Australian newspapers Digitisation Program, the National Library of Australia has been actively digitising newspapers for four years, and have provided unmediated access for users to find and correct OCRed text as well as tag newspaper articles. This presentation will outline the background to the project, provide an explanation of the method by which users can engage with the content and outline some of the lessons learnt and challenges faced by exposing content.

How will web 2.0 and web 3.0 technologies impact the new generation of digital libraries

By Alexandre Besson

(i2s Digibook Business Unit, France)

Alexandre Besson

i2s Digibook Business Unit

France

Mr. Alexandre Besson is an Engineer in Electronics and computer science. He has worked for 6 years as project manager in Vision system for quality control. He is a sales manager for i2S Digibook since 2007 and is in charge of USA and Asian countries.