Sustaining and Maintaining of News and Information by Digitalization on "MALAYSIANA" Collections

Hasnita Hj Ibrahim
Library & Information Centre (InfoLib) BERNAMA
Emel: hasnita@bernama.com

Masfaliza Masmuda
Library & Information Centre (InfoLib) BERNAMA
Emel: masfaliza@bernama.com

Abstract

The up-graded version of the digital archiving system of InfoLib, known as BLIS (Bernama Library & InfoLink Service), meant to sustain and maintain for easy access the news and information pertaining Malaysia. The development of BLIS is dedicated to the long-term storage and access to large quantities of documents that are in digital form (born digital) and those that were converted to digital form through a digitization process, which can be accessed via BLIS internal users and outside clients. In this paper we shall highlight on the development, design, implementation and some issues pertaining to digital activities.

Keywords: Digital; Digital archive; Content Management System

Introduction

 "It is a change and behavior management issue. It is not going to happen overnight. You have to rewrite policies and procedures and audit for compliance." Anything that comes in on physical paper has to be scanned into the system. Oregon is not alone in trying to address how to deal with the explosion of electronically stored information in a more efficient and structured manner. (Raths, 2009).

BERNAMA has dealt with the collections of news and information which considered as the heritage of BERNAMA, and that as Malaysia.

Library & Information Centre (InfoLib) formerly known as BERNAMA Library, was established in 1968, in conjunction with BERNAMA's official existence. In line with BERNAMA objectives of generating as well as disseminating news and information, InfoLib entrusted in storing, processing as well as archiving them for reference purposes.

Today the news and information are for the benefit of the whole society electronically. In the early days only BERNAMA's journalists used them. But the end of 2001, the digital archiving system of InfoLib was in the 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 has been up-graded and then launched in April 2008.

BLIS (BERNAMA Library & Infolink Service)

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

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.

Jointly owned by InfoLib, which is focusing on the internal users and Database and Corporate Information Division which is focusing on promoting and selling the compiled product to the clients. BESSAR, Bernama System and Solutions, a group company of Bernama is executing 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.

The information from various sources are categorized, sequenced and stored as Information modules as described below (Table 1):

TABLE 1: Modules and description in BLIS

No.	Modules	Description
1.	Bernama News	Covers news ranging from business, politics, sports, event calendar and courts to human-interest issues. These are archived and indexed ready to be retrieved in a search on the BLIS site 24 hours after they appear on BERNAMA's news wires. Collection of news items as reported by BERNAMA from 1996 to current.
2.	Archival News	Collection of news clippings and full text from the major local newspapers and magazines. News cover all the politics, business, human-interest issues and major events in sports.
3.	Foreign News	Collection of foreign news from various news agencies from around the world like Kyodo of Japan, Yonhap of Korea, Xinhua from China, MAP from Morocco, PTI from India, International Islamic News Agency(IINA) under the organization of the Islamic Conference(OIC),

		Antara of Indonesia and many others from the African nations.
4.	Info. Providers	Malaysian statistics on industry performance, commodity production and economic forecast provided by government research agencies.
5.	Who's Who	Compilation of profiles of prominent persons in Malaysia.
6.	Speeches	Full text of speeches made at local and international events like world forums, launches, official openings, announcements and other major events delivered by key government and corporate officials in Malaysia.
7.	Sectoral Reports	Compilation of various industries in Malaysia. Sourced from various organizations.
8.	Databank	Provides facts on diverse issues and events extracted from Malaysian Newspapers and BERNAMA's reports.
9.	Facts & Figures	Profile of Malaysia on its economy, government industries, labor and business environment. State profiles cover projects, governments and industries. Details on primary commodities.
10.	Directory	Covers banks, finance companies, unit trusts, money brokers and insurance institutions.
11.	Calendar of Events	Details of events such as what it is about, venue, time, cost and contact persons.

System Specification

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

1. Content Management System or CMS is a 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 single application in which the content management and administration of the BLIS application are integrated. The features built in are shown below (Table 2):

TABLE 2: System Features and Specifications

Content Management Module	The standard features of content management systems like content creation, content modification and content deletion based on user access privileges is provided in this module.
User Management Feature	Create users for the BLIS system. This module allows the creation of the users of the BLIS system. In this module the Admin can create internal users, external users, bulk

	subscribers, trial users for the BLIS System.
User Online Report	Generates the reports of the users currently online on the BLIS system. The online users will include both the internal and external user reports. The Admin can view list of all the users logged in both the BLIS Admin Application as well as the BLIS User Application.
Subscription Management	Allows for definition of the subscription plan and changes to the subscription plan.
Workflow Definition	Defines the workflow of the Content Management group. The Administrators can categorize and define user and user type.
Trial user Report	The details are displayed in this module. The Admin can extend the trial period and perform other settings like de-activate/activate the trial access features for the trial users.
Hits Reports	Gives the details of the user hits reports which can be edited by the Admin.
Edit Hyperlink	The Admin can define links to external application in the Edit Hyperlink module. The Admin can attach an image and description for the links. The hyper links will be placed in the BLIS User Application page. The Admin can add, update or delete any number of external links.
Audit Trail	The Admin can monitor the activities of users of the BLIS Admin Application as well as the BLIS User Application by viewing the Audit Trail report which is separated for both the internal and external users.
Misc Settings	Additional application settings like addition of Categories, Product Types, Content tree structure to BLIS User Site mapping and other settings can be performed by the Admin in this section. The Admin can define the BLIS User Application display in this module.

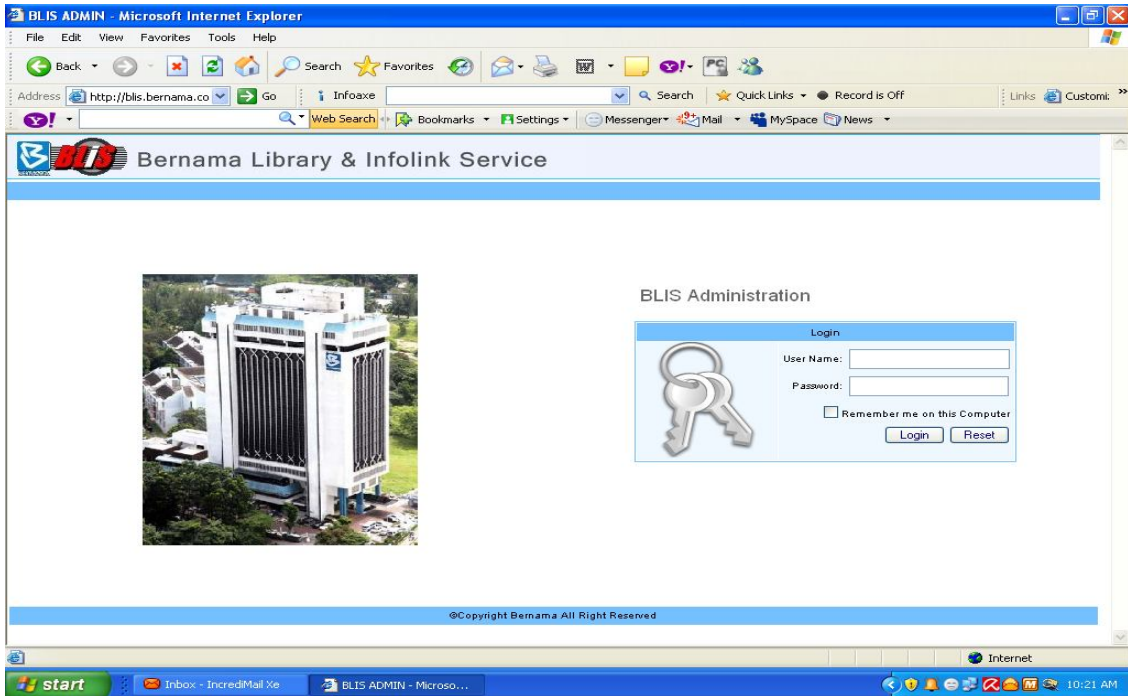FIGURE 1: BlisAdmin Application

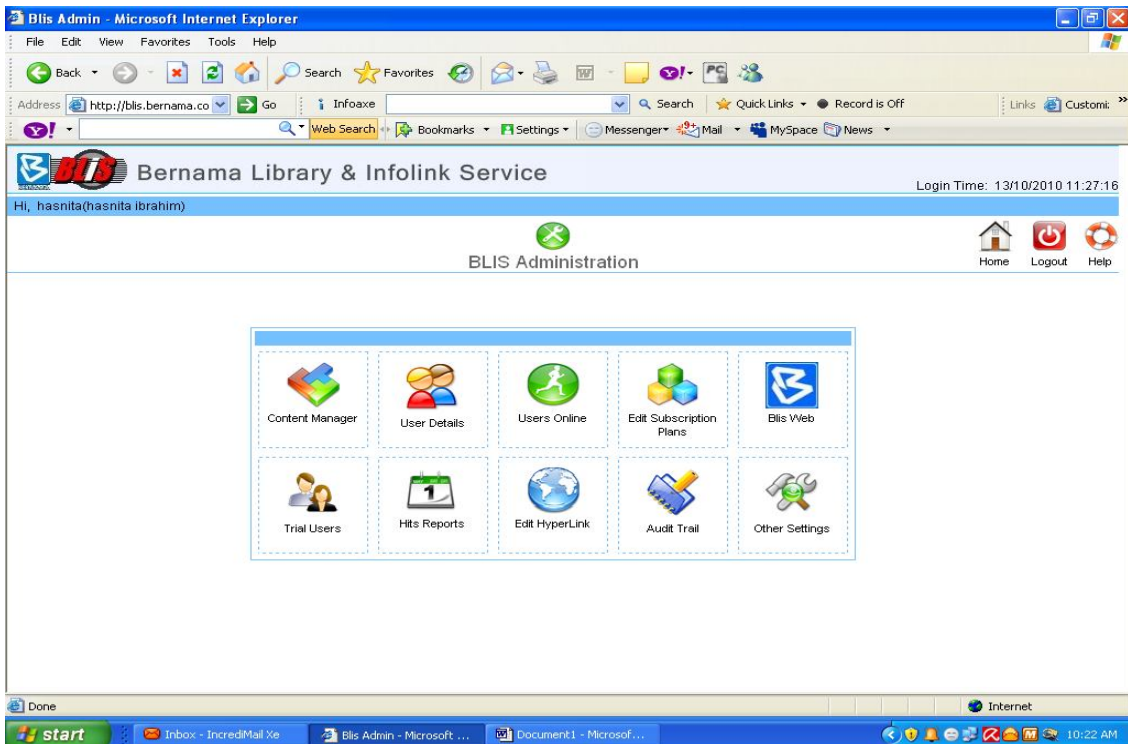FIGURE 2: Blis Administration Features
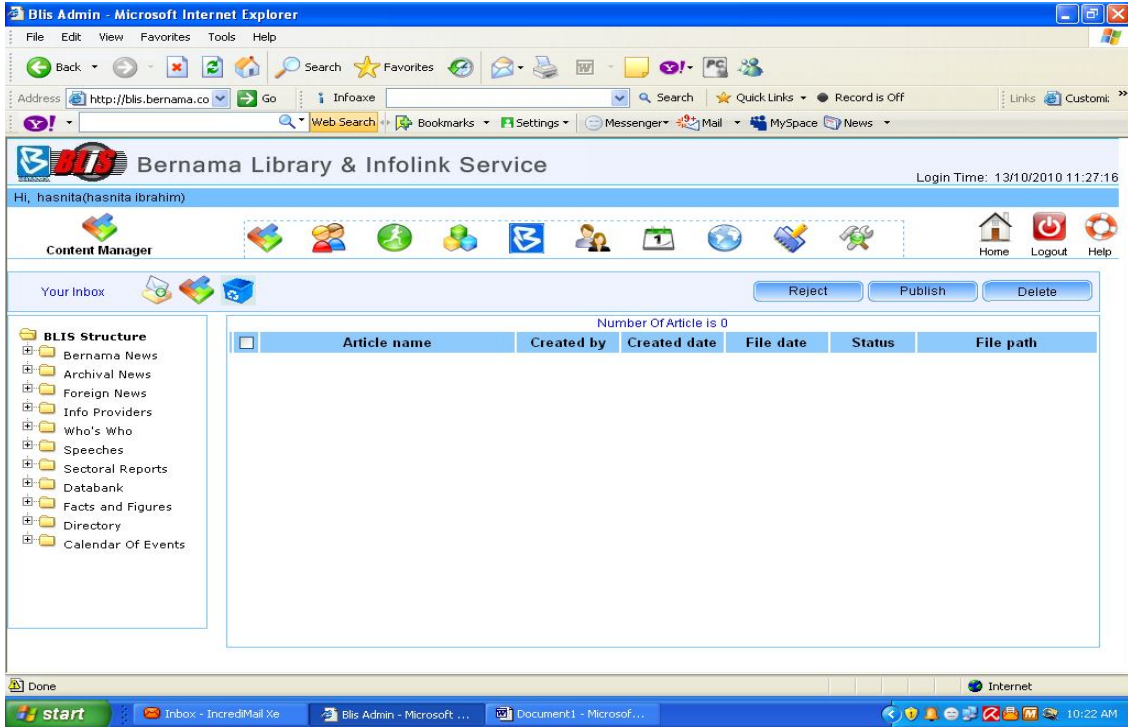
FIGURE 3: Module – Listing
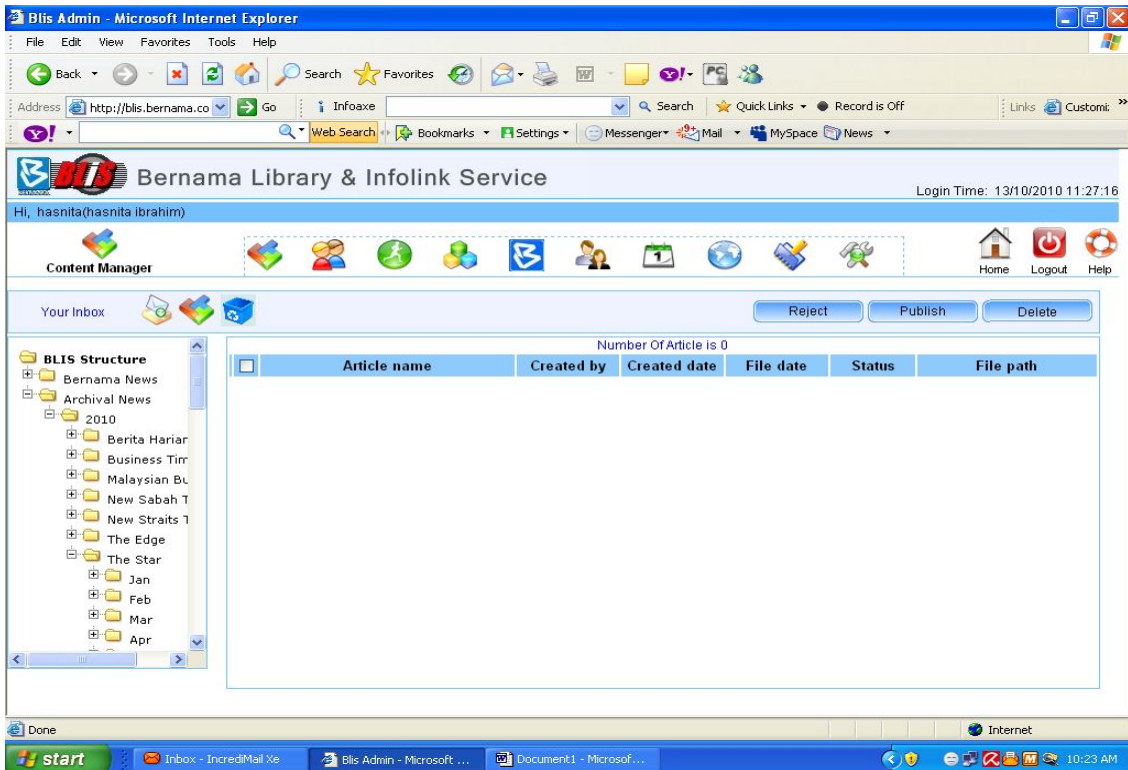

FIGURE 4: Sub Module – Listing

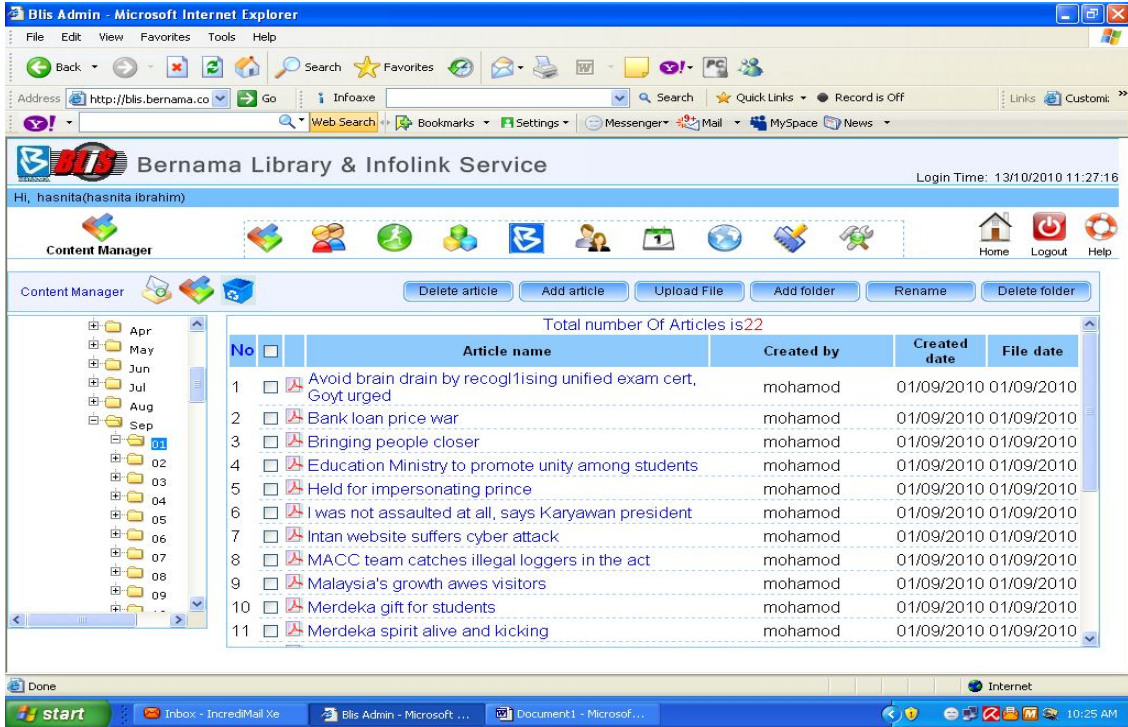FIGURE 5: Articles listing - Archival News Module
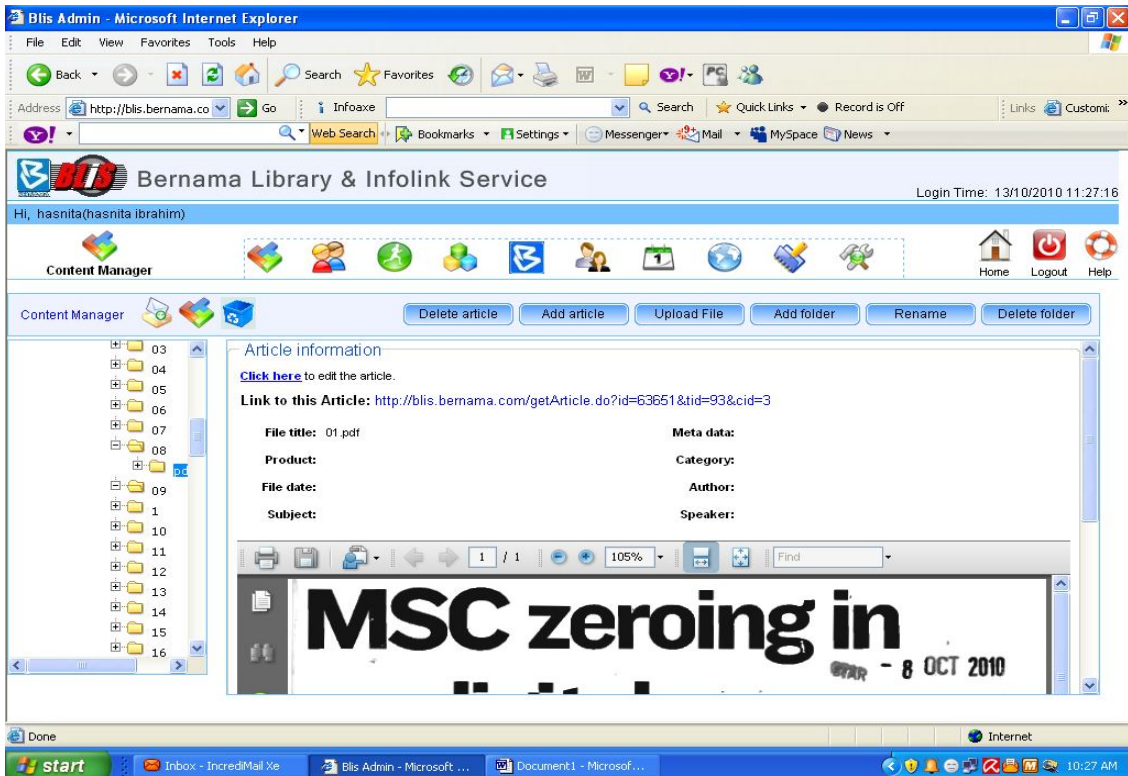

FIGURE 6: Digital article – Archival News Module

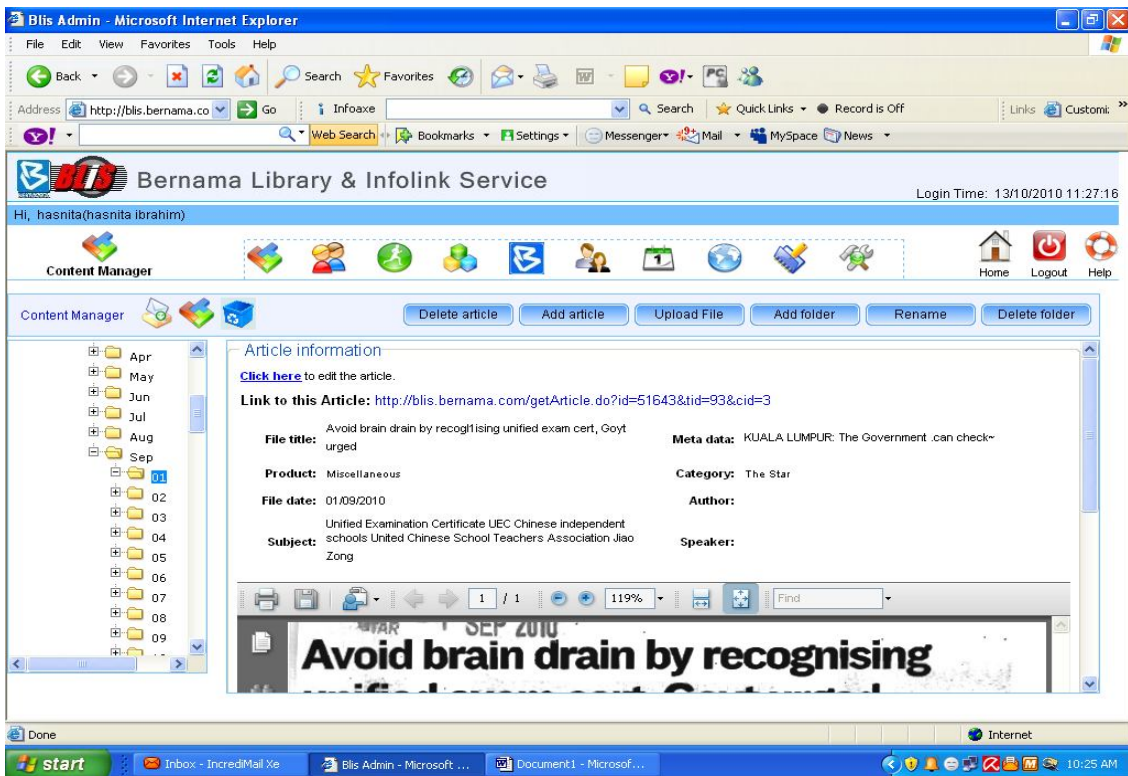FIGURE 7: Metadata - Digital article – Archival News Module

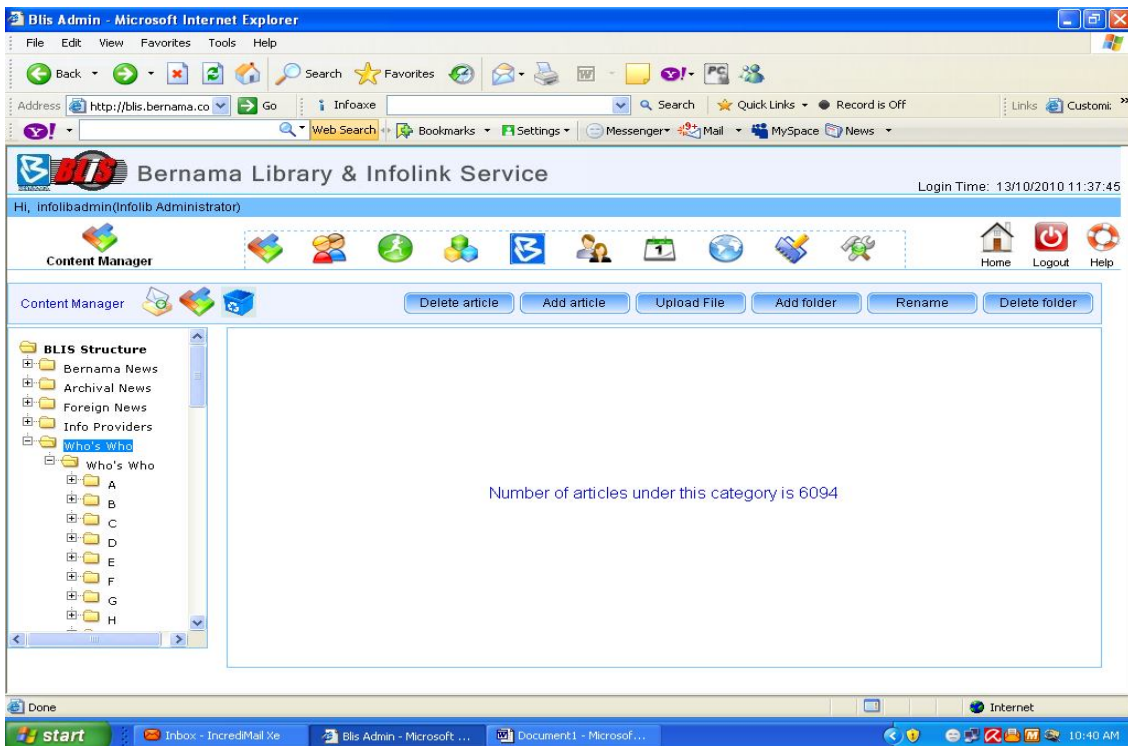FIGURE 8: Articles listing – Who's Who Module
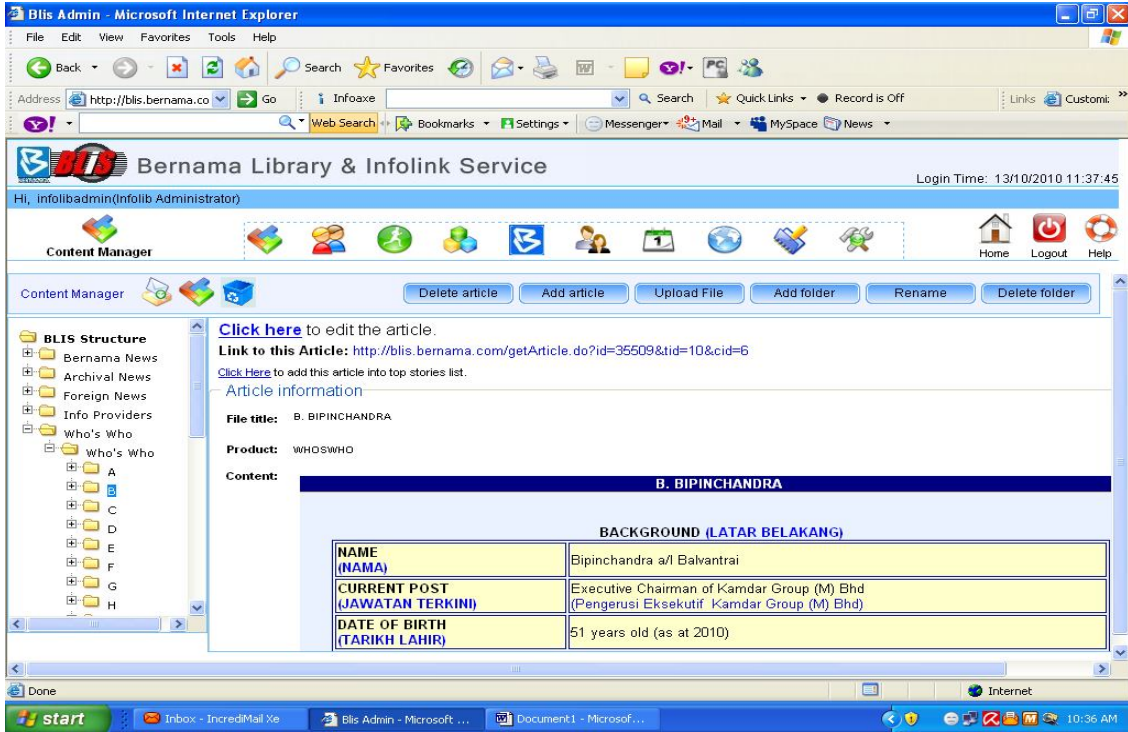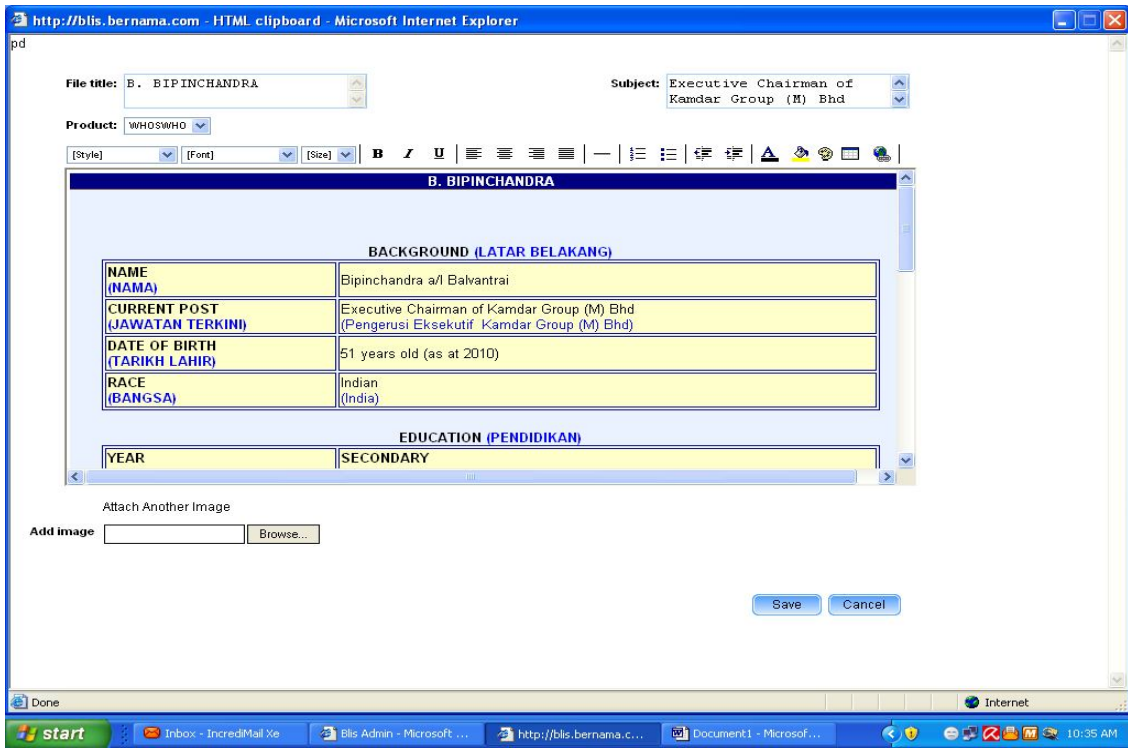
FIGURE 9: Profile – Who's Who module

FIGURE 10: Metadata – Profile - Who's Who module

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 features provided are shown below (Table 3) :

TABLE 3: Features for User Application

Any user of the BLIS User Application can browse through all the information modules without logging in to the system. However to access the contents of the articles the user must login in to the BLIS User Application. If the internal Bernama users access the BLIS User Application within the local intranet they need not login in to the BLIS User Application. They will be authenticated based on the IP address of the request.
The unsubscribe users of users not logged in to the system will be able to view the articles titles and cannot view the articles until the users are logged on.
The User Application Site will have bulk registration facility. Organizations the wish to subscribe on bulk will be provided with a single user id and password and can have a configurable session count. The BLIS Admin can alter the session count in BLIS Admin Application.
The BLIS Site will have a default time-out of 10 minutes. The users who remain inactive for than 10 minutes are logged off from the system and the session count is released for additional users.
BLIS User application will provide 2 kinds of search facilities.
Global search is an option on the BLIS user Application where the articles matching the key words across all the information modules are fetched. This search is available in the header section of the BLIS User Application. An optimized collective index will be built for a faster search.
Advanced Search is an option on the BLIS User Application where the users can search a subset of information modules and drill down the search results. The User can search based on the date or can drill down on specific information module all provide other search criteria like search on a specific category or product or on a specific author. The search mechanism in this level will be a local search. In the advanced search limits the search or the desired articles at the information module level.

| |
|---|
| Prospective users who wish to register into the BLIS User Application can register in the BLIS User Application. |
| Subscribed users can view/modify their subscriptions details in the BLIS User Application. News users can also subscribe to the BLIS User Application through the subscription link available on the BLIS User Application Site. |
| Links to external related sites will be available on the BLIS User Application. Users who wish to view these can click on the hyperlinks. The user will be displayed the selected page in a separate window. |
| The design of the BLIS User Application is built to easily integrated with any available payment gateways. |

FIGURE 11: Bernama Library Infolink Service (BLIS) – Main page

The system is based on the MVC Architecture. The figure below depicts the overall functional architecture of the system. 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.

FIGURE 12: Architecture of BLIS
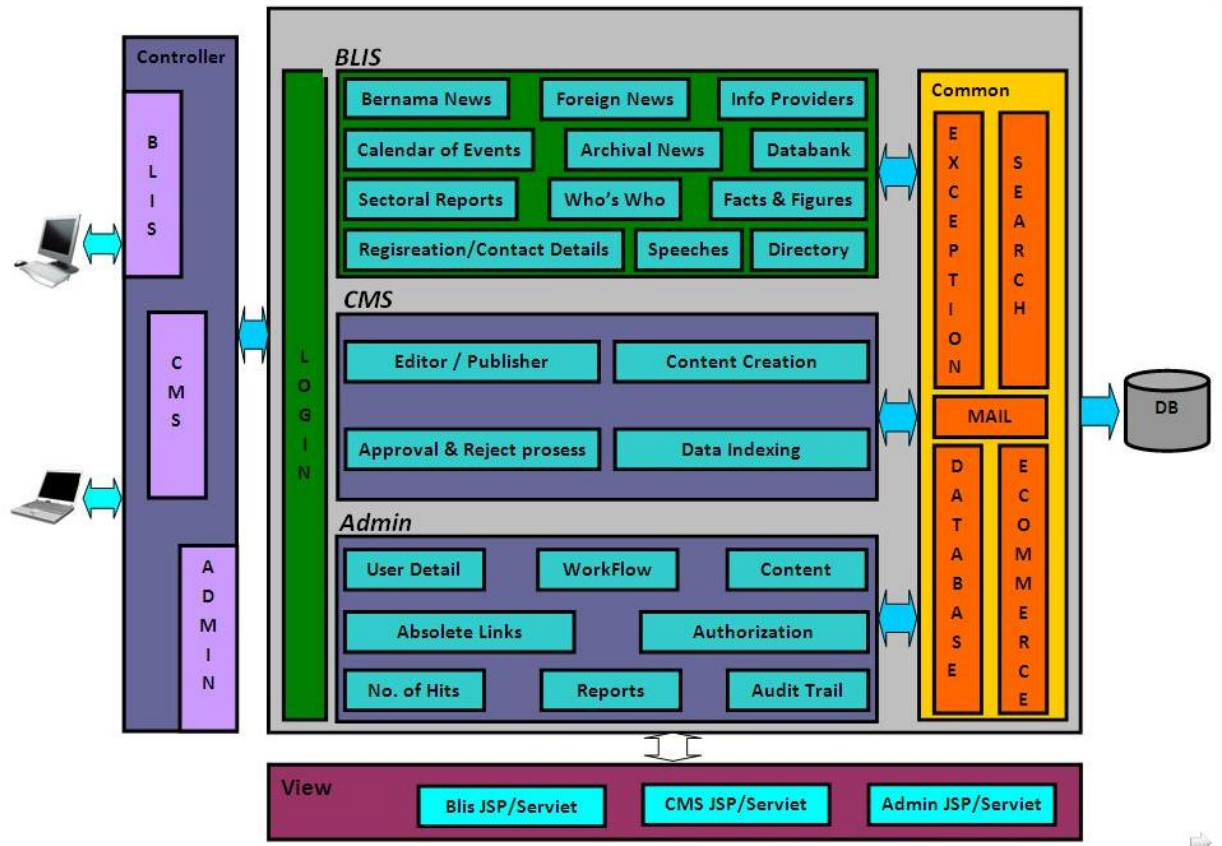
Project Work Flow

*BLIS Team = InfoLib and Database and Corporate Information Division
*Developer = BESSAR (Bernama System and Solutions) & Karthavya Technologies

TABLE 4: BLIS Work Flow

No	Item	*Task	Date
1.	Attending workshops / seminars / conferences, visits to other libraries	BLIS Team	Prior to the project
2.	Planning /Requirements Hardware's & software's	BLIS Team	Jan 2007
3.	Product demos, -estimation budget	BLIS Team	Feb 2007

	Getting the quotations from vendors		
4.	Prepare proposal paper	BLIS Team	Mac 2007
5.	Submit 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	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

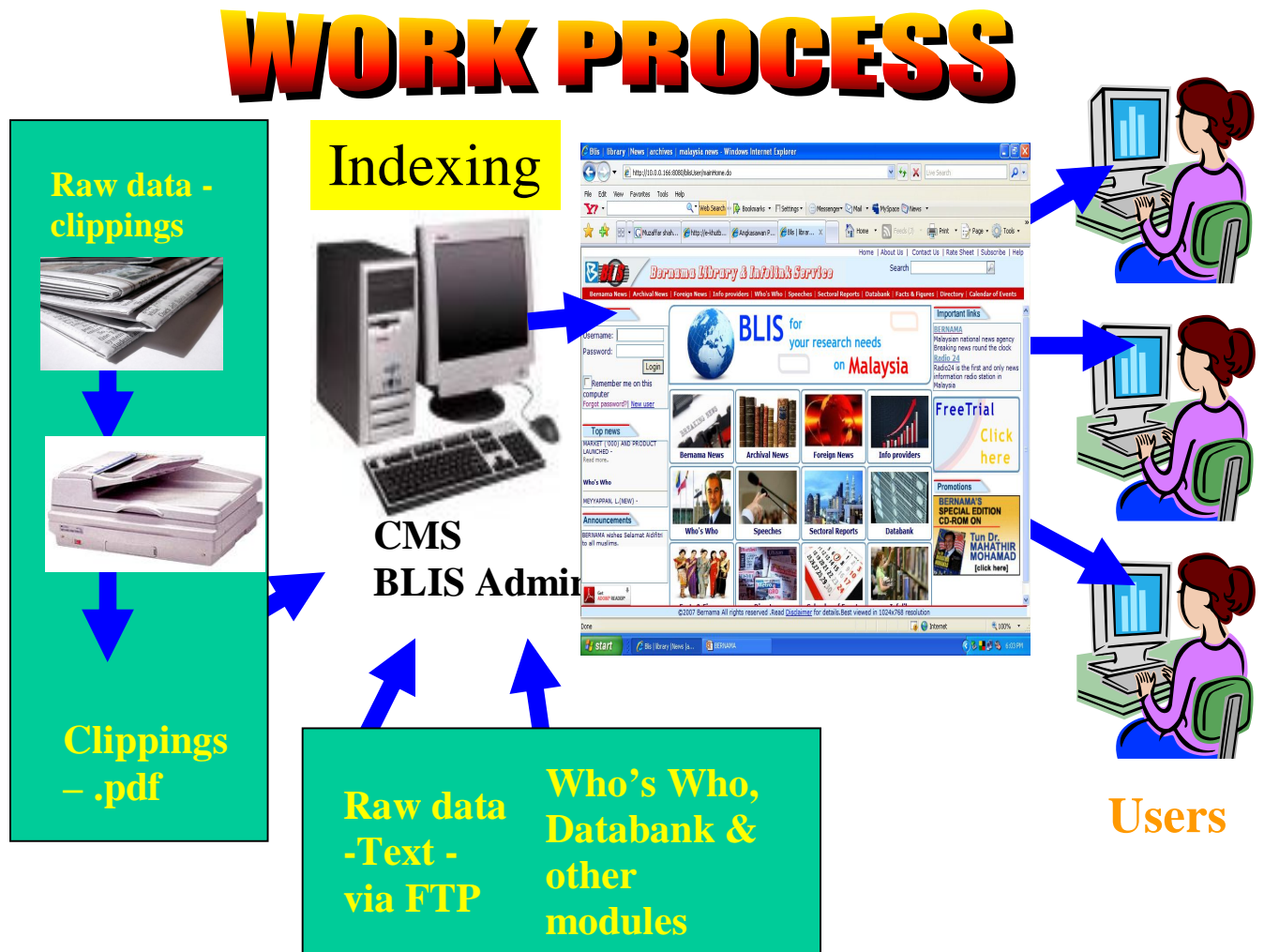| 24.	Up – (Internal users + External clients)	BLIS Team	Jan 2008
25.	BLIS Testing (Internal users + External clients)	BLIS Team	Feb, 15 2008
26.	Warranty period (6 months)	BLIS Team	May 2008
27.	BLIS Launching - Internal	BLIS Team	April, 22 2008
28.	Signing the maintenance agreement.	Developer + BLIS team	Nov 2008

FIGURE 12: BLIS Work Process
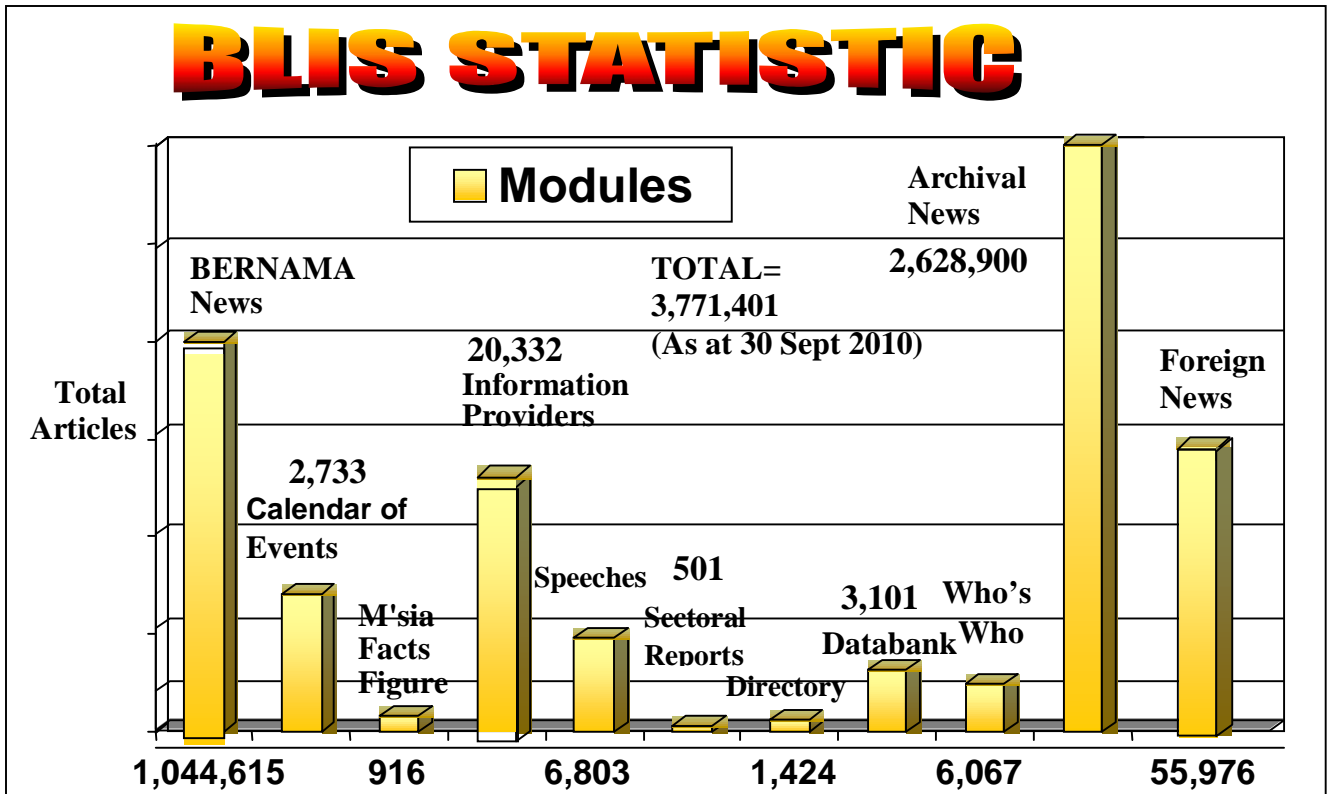
FIGURE 13: BLIS Statistics

TABLE 5: Searching

Global Search	Global search is used to search articles based on keywords. This search provides quick results based on the keywords entered.
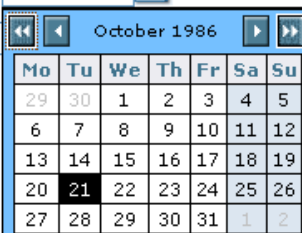
Advanced search	Advanced search helps to easily drill down to the article we are searching for. This search will generally be faster and more precise than the global search. **Bernama News** *Advanced Search* **Select Category** ☐ All ☐ Berita Dalam Negeri ☐ Berita Ekonomi ☐ Economic News ☐ Features/Special Reports ☐ General News ☐ NewsPaper Summary (Publication Creased - March 2005) ☐ Siaran Akhbar ☐ Press Release Eng ☐ Rencana/Laporan Khas **Search Criteria** Title: [_____] Author: [_____] From Date: [____] 🔲 ↺ To Date: [____] 🔲 ↺ Key Words: [_____] [Search] [Reset] October 1986 Mo Tu We Th Fr Sa Su 29 30 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 **21** 22 23 24 25 26 27 28 29 30 31 1 2

Understand the Issues

Planning stage

Prior to embark a digital archive project, the personnel should acquire the knowledge and skill by attending workshops/seminars/conferences and similar activities which are related to technologies and managing a project.

A visit to the other libraries, which has embarked similar project, is equally important. The learning part from the visit is that libraries can avoid facing similar problems and mistakes by identifying vendors and needed digital archive software in the market so that best results could be achieved.

For as, appoint project manager plus committed team members who are supposed to be working closely with the appointed vendors. List all the requirements with the whole team.

Next, arrange demos by the identified related vendors. This can identify and assist selection of digital archive preservation across new generations of technologies. Short listing the vendors and get quotations. These inputs are to be including in the budget requirements, and the proposal paper for management approval.

As soon as the vendor is appointed, we must have proper schedule is required to achieve the target date. The vendors then will mange the whole

project with our cooperation. With proper planning and disciplines we can meet the challenges and attain good results.

Rapid growth and evolution of ICT

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 (Oltmans, 2004)

Technology that enables the variety of formats and dissemination mechanisms changes rapidly. Establishing a program that is responsive to change is a huge challenge (Cornell University Library, 2002-2007).

The advancement of ICT is changing very fast. It is sometimes difficult to catch-up with the change due to budget constrains 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.

Though BLIS is an up-graded version, it is a totally new entity starting from a scratch, in which the process is very tedious and time consuming. It is best to have a properly maintained digital archive system so that the system can be up-graded regularly in-line with the new technology, not to totally replace with the new infrastructure and to ensure the system in good condition and expandable. Develop the digital archive to maximize, scalability, flexibility and reliability.

Ensuring that the digital archive is as interoperable as possible is by utilizing open source options whenever feasible.

The technical arm in most of the organizations will not have the expertise to develop a proper digital library system. It is 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 whom develop the digital archive system resign from that particular organization. The new personnel will not have the believed and understanding towards the whole architecture. Adding to it the digital archive system will not be having the up-graded version and possibly the system will be left unattended for not inline with the present technology.

There are 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. (Green, 2009)

Vendor

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

Recognizing and selecting a highly responsibility and reliability 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 such as the company stop selling and promoting the product, the company closes shop, no/lack of commitments, getting different results/specification from what had been promised, lacking of expertise and many others.

It is equally important for the vendor to be well equipped with the expertise, structured documents and manual along with the project. They need to have proper planning at the early stage until the project is accomplished and the maintaining aspect must also be included.

The up-graded BLIS was awarded to the BERNAMA's subsidiary which is BESSAR and assisted by Karthavya Technologies, from India. It is a customized system using Oracle, Java and Jsp. The arrangement with BESSAR will ensure a long term relations for the future of BLIS.

The advantage of having customized system is that we can determine the features, which are tailored suit to our requirements. But the disadvantages are that they would need a longer time and to 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", thereby another working place (software) was needed which 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 package.

Project team

Every project needs 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 meeting as with minutes taken are necessary to ensure that every decision will be recorded and well monitored.

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. On budget, it is advisable to propose revenue than needed.

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

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 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 completed.

Training for the staff is very important to ensure that the system is properly updated and managed. The challenge will be the resistance to change at the daily job and the time for the staff to adjust to the new environment of work. Technically everyone needs time to implement the adjustment.

Storage/accessing

For BLIS, more than 1,000 daily news from BERNAMA, NSTP publication and Utusan Melayu will load 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.

Out-sourcing cannot be implemented due to cost implication to digitize millions of articles in the collection. Processing in-house is very time consuming and more manpower is needed. Then, the option digitize in-house according to the priority of news.

Data Migration

For up-grading projects, data migration would be a tedious job due to the volume and should be sorted wisely. At one end, checking data is needed to ensure that the whole information/articles will not be missing or corrupted in the up-graded version.

From our past experienced, the vendor had minor problem with the migrating of data due to the errors and different formats in the present systems and this slowed down the project. The possibility of losing the data is there if it is not properly checked.

Other capabilities

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 (Oltmans, 2004).

If the volume of the information/articles are high, request software that can compress the articles so that less space being utilize and more articles can be put in. The server will take longer time to be full and will maintain the speed of getting the articles. Make sure that the system can generate reports periodically according to our needs for statistical purposes.

For the working place (BLISAdmin) we need to have a proper place to insert the metadata as well as copy and past capabilities. Make sure the working place can be up-dated and having to have back up which to avoid double process where we have to delete and reinsert the up-dated files.

Back-ups

It is crucial to have a structured practice of back-ups such as tapes, on cd's or server due to the amount of information in BLIS. Sometimes BLIS experienced virus attack before which crippled the retrieval of the present information. For a client-based system this cannot happen. The anti-virus is a priority and a disaster planning is very important under the whole information security infrastructure.

Training

To increase the utilization of the digital archive system, to introduce, to highlight searching techniques users need to be trained. Whoever is using the system should undergo training so as to simplify work.

Conclusion

The new world of digital information requires a new way of providing access to that information. Since BLIS is a growing in terms of the volume of information materials. It has to be properly maintained by competent team. The system and supporting processes are continuously being improved, based on BLIS working experience with volumes of born digital and digitized collections.

Under Digital Archives, we have developed a secured and resilient store for about 4 millions of digital items. Over the years, we foresee a massive expansion in the size of the store both in terms of the number of items and its size capacity. At the same time, we will add facilities to make it easier for users and clients to find and access the relevant content.

We believe in long-term commitment to a digital future. For preservation of news and information, we therefore need higher end equipment to digitalize the rest of Bernama News and newspaper clippings into our collection.

Acknowledgement

1.Infolib Staff
2.Mr Izham Shuhaimi bin Ahmad – Editor, Bernama.com
   (Prove read this paper)
3. Organizers – UniMap Library, Malaysian Online E-Resources and Librarians Association of Malayisa (North)
4. Mr Visvakumar a/l Sangiah – Senior Manager
   Human Resources and Administration Department
5. Our family

References

Beebe, L. and Meyers, B. (2000) Digital workflow: managing the process electronically. *The Journal of Electronic Publishing.* 5(4). Retrieved from http://dx.doi.org/10.3998/3336451.0005.403

Green, S. (2009). The digital library programme at the British Library: goals and priorities. *Emerald database.* 37(3) :136–139.

Conway, P. (1996) Preservation in the Digital World. Retrieved from http://www.clir.org/pubs/reports/conway2/

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

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

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

Raths, D. (2009). States step up to the electronic challenge. Retrieved from http://www.kmworld.com

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