Design and Development of a User-Centred Digital Library System: Some Basic Guidelines

N. Meyyappan (*), G.G.Chowdhury (#) and S. Foo(**)

PN839200@ntu.edu.sg (*), assfoo@ntu.edu.sg (**)
Division of Information Studies, School of Computer Engineering,
Nanyang Technological University, Singapore 639798

gobinda@dis.strath.ac.uk (#)
Department of Computer and Information Sciences, University of Strathclyde, Glasgow, United Kingdom

Abstract

Beginning with an introduction and justifications for the need of user-centred digital library development, this paper outlines the architecture, the basic components and functionalities of a Digital Work Environment (DWE) that forms the basis of a user-centred digital library development at Nanyang Technological University in Singapore. The information resources include full-text articles, online and other forms of databases, collections of theses and dissertations, library OPAC, union catalogues, and so on, that are scattered across the Internet, university's intranet, NTU Library's Gateway to Electronic Media Services (GEMS) and Online Public Access Catalogue (OPAC) system. The user-centred DWE design aims to provide a systematic access to these resources by organising them using a task-based approach. With this, users are able to browse through a hierarchy of tasks and sub-tasks and identify appropriate resources. In order to support added search and information retrieval variety, DWE also supports browsing of an alphabetical list of all resources, searching of individual information resource attributes of intranet documents, GEMS and OPAC, user-selected Internet search engines, and a combined search of these various options through a single unified query. Based on the experience of a continued research on the design and development of the DWE, a set of generic guidelines for the design of a user-centred digital library system is provided.

Introduction

Digital library research and development across the world over the past few years have significantly improved the facilities for accessing and retrieving digital information resources. While researchers are constantly trying to come up with mechanisms that aim to provide users access to timely, accurate, comprehensive and relevant information from a variety of sources and systems, most digital library systems to date expect that users are aware of where the information is located or the information systems that it resides, and that they are proficient in search and retrieval techniques of different information systems. A review of twenty working digital libraries from different parts of the world representing academic, special and public libraries attest to the diversified collection of information resources and different forms of digital libraries that are in existence today (Meyyappan, Chowdhury and Foo, 2000b). The various information resources accessible to users include those of full-text articles, proceedings papers, CD-ROM databases, theses and dissertations, e-journals, e-books, examination papers, photographs, images, maps, audio, video, multimedia databases, collections of manuscripts, sound recordings, music collections, OPACs, union catalogues, and so on. Many digital libraries host special collections of a sole homogeneous resource type such as maps or sound recordings. Other digital libraries take the form of hybrid libraries with a combination of general and special collections of multiple resource types in both digital and print form.

The academic community is possibly the largest and the most important user group of digital libraries. The nature of work of academics, students, researchers and other users vary according to the tasks that they perform. Apart from the various information resources that are currently managed, provided and access by digital libraries, there are many other information resources that are needed by users in an academic environment. These resources include course calendars, university statutes, various course offerings, course registration, thesis and dissertation guidelines, style guides, laboratory facilities, availability of software, hardware, equipment, course materials, reserve book/handout collections, local publication databases, locally produced theses and dissertations and so on.

Current digital libraries and the Web environment are mainly designed based on a systems approach. Less or little emphasis has been put on the users in such settings. Users are expected to be able to address the "what", "where" and "how" questions in their quest to seek information. They are expected to be able to formulate good queries to represent their information needs, or to map their queries onto often unknown knowledge structure such as subject directories. This is not an easy job. Although users information needs are often related to some form of their immediate task in one way or another, current digital libraries or information systems do not generally organize information according to the various user tasks. As a result, users usually employ the trial and error method to move from one Web page to another or from one information resource or system to another.

Another common problem faced by users is that the required information resources are generally scattered all over the place with the quality, content and coverage differing at different locations. When this happens, users have to traverse from one server to another or from one information resource to another in order to find the desired information. The user-centred approach to digital library design is therefore desirable as it aims to shift the

focus from a system-oriented to a user-oriented design in an attempt to users' real needs and facilitating means and ways to support their information seeking and use behaviour.

Continued research at Nanyang Technological University in Singapore, have attempted to resolve some of the above problems. The central focus of this research is the design and development of a Digital Work Environment (DWE), a task-based information retrieval environment, that is designed to serve the needs of the Nanyang Technological University's (NTU) academic community. Design features of DWE have recently been reported in the literature (Meyyappan, Chowdhury and Foo, 2001a, 2001b). In this paper, based on our experience with the current research, we shall provide some basic guidelines for the design and development of a digital work environment that is the building block of a user-centred digital library. Beginning with a brief review of the related literature, we shall briefly describe the features of the DWE, which will follow the design guidelines.

Review of Related Literature

As early as 1979, Garvey, et. al. noted that the success of information services was more likely to be achieved through adjusting the services to meet the specific needs of an individual rather than trying to adapt the individual user to match the wholesale output of an information system. Mick *et. al*. (1980) recognised that the effective transition into the information age will require switching from information systems that are technology and content driven to information systems that are user driven. Despite such recognition and early calls, the majority of existing digital libraries are still designed in the more traditional way. However, there is also evidence to show that some digital library design had heeded these calls and attempted to create more user-centred designs. The MyLibrary (Cohen et. al., 2000) and HeadLine (Gambles, 2000) digital library projects have focussed on the delivery of a personalised information environment and attempted to cater to the differing needs of users. Michelle and Wang (2000) designed a user interface in SenseMaker for information exploration in a heterogeneous digital library environment. Leong (1998) proposed a Web-based architecture for searching distributed heterogeneous multi-Asian language bibliographic sources, and described a pilot implementation of the system. Gates and Lawhead (1998) and Moussawi *et. al.* (1999) designed information systems to provide information from heterogeneous sources through a user-customized view on a Web interface.

Various ways can and have been used to organise the information of digital libraries. These include the use of an alphabetical listing, subject categories, broad groupings, by tasks, and so on (Meyyappan, 2000b). For instance, the ACM digital library uses its own Computing Classification System (CCS) for organising information, while the American Memory digital library organises information in the form of broad groups, format, time, place, original format, digital format and library classification. A number of researchers have recommended the task-based approach (see for example, Lesk, 1991; Fox *et. al.*, 1993 and Cousins, 1996). In these cases, a task-based approach is shown to be a useful design for effective and efficient of information retrieval. In particular, it aims to filter out irrelevant information resources and presenting to the users those relevant and needed information resources in a suitable format.

Basic Architecture of DWE System

The basic architecture of DWE is shown in Figure 1 (Meyyappan, Chowdhury and Foo, 2001a). MS Access was used at the backend for creating, maintaining and updating of user categories, user tasks, personal space and information resources tables. Additionally, the Java Servlets, JavaScript scripting language, HyperText Markup Language (HTML) and Open Database Connectivity (ODBC) were used for developing the various routines of the system**.** The Dublin Core (DC) metadata element set was chosen to describe the resource entities together with other data attributes such as the last date accessed, count of resource usage, authentication code and password that are necessary to access external resources such as the subscription-based online databases and e-journals.

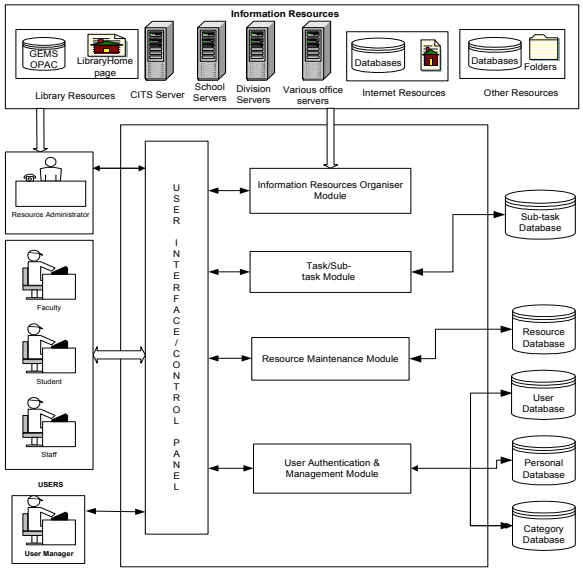

Figure 1: Basic Architecture of the System

The environment is used by three main user groups, namely, the *Information Resource Administrator, User Manager* and *the General Academic Users.* The *Information Resource Administrator* is responsible for the overall management of all the information resources that are interfaced to DWE. This includes ensuring resource relevancy, integrity and currency. The *User Manager* has the task to manage the user base of DWE. The end users of DWE are *General Academic Users* that include the various Faculty, Student and Staff of the University.

The system consists of a *User Interface Module* that is linked to four other main modules. The *User Authentication and Management Module* is used for the login process to ascertain the type of user by interacting with the user category database. With this information, the system will automatically present the list of tasks that are associated with the user type for subsequent selection by the user. The *Task/Sub-task Module* is used by the Resource Administrator to update the respective tasks and sub-tasks in the tasks database. This is used to define the relationships between the main task, sub-task and sub-sub-tasks, thereby forming the hierarchy of parent-child tasks relationships. The *Resource Maintenance Module* is maintained by the Resource Administrator to update

the resource information that includes the Uniform Resource Locator (URL) and resource password (if any) in the resource database.

The *Information Resources Organiser Module* plays an important role by processing users requests and interacting with various servers in the Intranet, library home page, databases, folders, Internet resources, and so on, to bring back the needed information resources to the users.

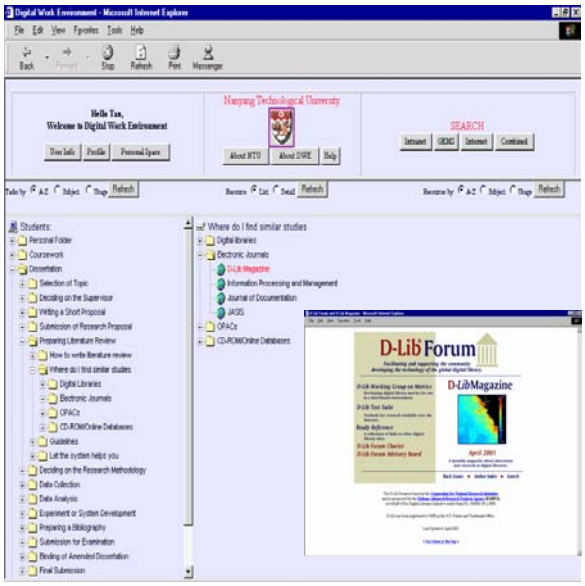

Figure 2: DWE Interface

User Interface/Control Panel

DWE employs a frame-based Web interface as shown in Figure 2. The interface is divided into three frames: *Welcome, Navigation Tree and Resource* frames. The *Welcome* frame is used to display greeting messages, University-related information and the DWE search menu. In the *Navigation Tree* frame, various tasks associated with each user category are organised in the form of a hierarchical tree structure starting with the general task and ending with the specific lowest level of sub-task. This part of the interface has a similar approach and look-and-feel of the widely accepted MS *Windows Explorer* model. Different tree structures are presented to users upon successful logins since it is tailored to support different users of the system. Upon selecting an appropriate resource in the *Resource* frame, a pop-up window is correspondingly activated to display the content of the selected information resource. In the figure, this window is scaled down to highlight the various parts of the interface.

DWE Workflow

The DWE databases of the current DWE prototype have been populated with data pertaining to the information needs of the Division of Information Studies (DIS), at Nanyang Technological University in Singapore. These information needs were elicited through a survey questionnaire for all the users groups in the Division, and a follow-up

focus group with the faculty (Meyyappan, 2000a). The data collected during these processes were used to define the different tasks that are undertaken by the different user groups of the Division, and the list of information resources that are needed to accomplish these tasks. These basically represent a set of "best practices" data for the Division.

The DWE workflow is simple and intuitive. This is illustrated through an example of a DIS graduate student carrying out a literature review as part of the "Dissertation" task. Figure 2 shows the DWE interface upon successful user login. The selected task comprises a number of sub-tasks, which in turn is divided into further sub-tasks that are arranged in a proper hierarchical manner in the interface. The student selects and expands the desired folders to show a set of resources that includes Digital Libraries, Electronic Journals, OPACs and CD-ROM/Online Databases. Upon selecting the Electronic Journals option, the various DIS e-journal titles are displayed in the Resource frame. The student selects the desired title that is in turn brought back by DWE and displayed in the separate pop up window.

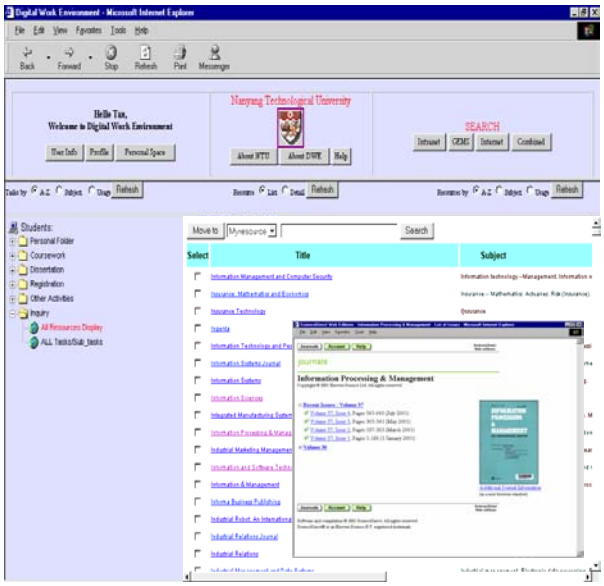

Figure 3: Alphabetical List of Resources

Browsing and Searching Alternatives in DWE

In addition to finding and using resources through the task-oriented manner, DWE supports other forms of browse and search strategies since users might sometimes want to search for other forms of external resources (e.g. documents on the Internet) or multiple systems or services (e.g. NTU Library OPAC or GEMS) that are not presented to them by default. DWE therefore supports browsing of an alphabetical list of all resources that are defined in the resource database (Figure 3). Users work with a list of all resources and make use of their own personal space to organise into folders and manage their own set of information resources that are extracted from the master list. Users can create any number of folders and put any number of resources in each folder. Users can also search the university Intranet, GEMS or Internet individually, as well as a combination of these last three

search options through a single query. Depending on the search option selected, different search features are supported. This includes simple keyword search, phrase search, or the more advanced complex search through the use of Boolean, truncation and proximity operators.

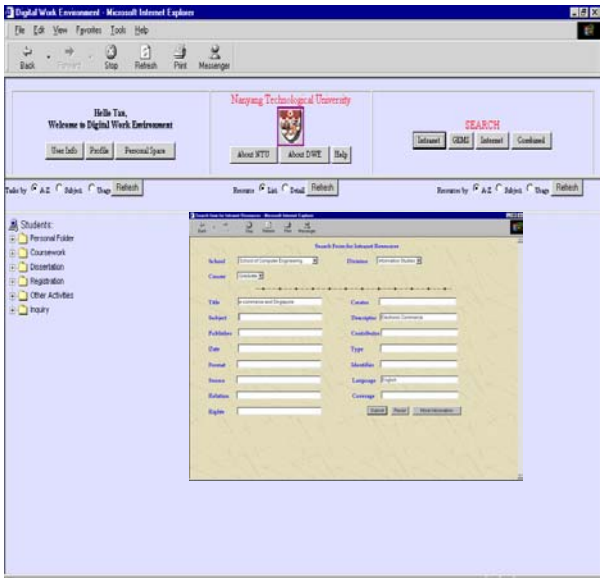

Figure 4: Intranet Resource Search Form

Intranet Resources

The indexed information resource descriptor fields of the DC metadata set and other attributes used for defining the university's intranet resources can be individually or collectively queried (using AND between fields) as shown in Figure 4. These intranet resources include those that are available in the NTU administration servers, library server, Centre for IT Services servers, and various schools, divisions, laboratories, office and other servers that are connected by the campus wide network.

GEMS

NTU's Gateway to Electronic Media Service's (GEMS) is basically the university's digital library system that is maintained by the NTU Library. It provides access to a range of CD-ROM databases, Chinese CD-ROM titles, online search services, e-journals, audio-visual sources, and OPACs. Users can search the system through the GEMS search interface as shown in Figure 5. Searchable fields include those of title, author, subject, keyword, call number, ISBN, ISSN, year, language and format.

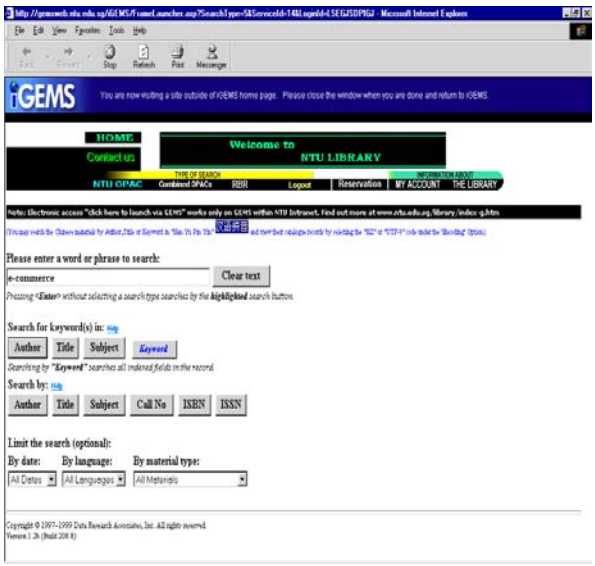

Figure 5: GEMS OPAC Search Form

Internet Resources

DWE adopts the concept of the NetScape Net Search Page to support the searching of Internet resources as shown in Figure 6. Users can choose any of the nine search engines, namely, Netscape, AltaVista, AskJeeves, GoTo, Hotbot, LookSmart, Lycos, NBCi and Google to carry out the search using simple keyword search. The selected search engine remains as the default search engine in subsequent search sessions. This default search engine information is captured by DWE automatically and stored as part of the user profile information. The majority of users are expected to use the simple search features as opposed to the advanced search features.

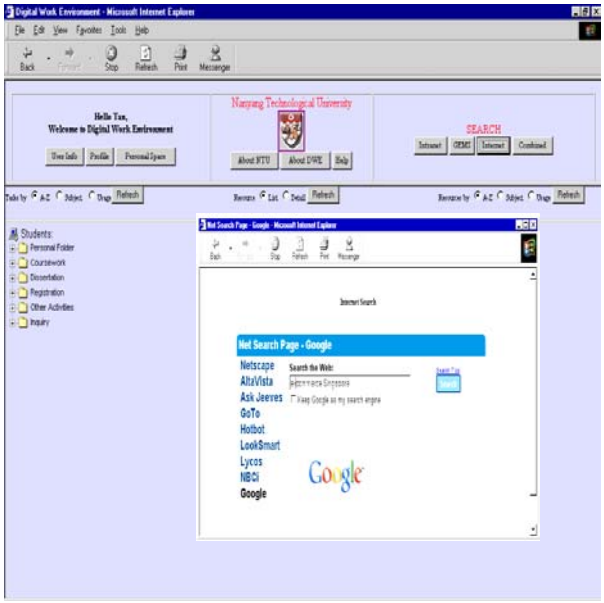

Figure 6: Internet Search Form

The simple search is basically a free-text search of all indexed fields of the search engines. As different search capabilities are supported by different search engines, it become important for users to consult the help documentation and be familiar with it prior to carrying out a search.

Combined Search

Searching different information systems and services individually, such as those supported by DWE, for the same search query obviously takes time and requires the user to be familiar with the search features that are supported by different search mechanisms. As such, DWE also supports a combined search facility to allow the user to carry out a keyword, subject and/or author search using a unified query that is processed and automatically directed to the different chosen search systems as shown in Figure 7. It translates the user query to an appropriate format that is supported by individual systems.

The combined search allows the user to select a combination of individual or all search systems that are supported by DWE. The Intranet resources contain the matching author, subject and keyword fields in the form of the adopted DC metadata elements, viz creator, subject and description respectively. GEMS have matching author, subject and keyword fields for searching. For Internet search, keyword and subject are matched with the title or general search string field. There is usually no separate author field to search author name in the Internet.

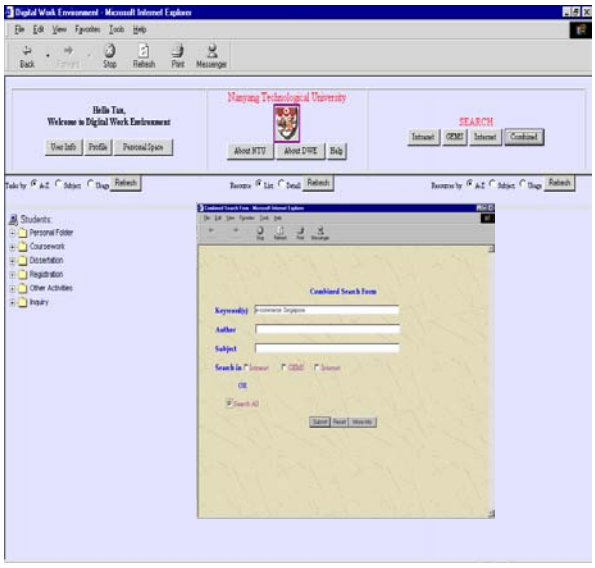

Figure 7: Combined Search Form

The query translation into suitable search strings for searching different system requires a detailed understanding the search facility available in individual systems. This translation and conversion of Boolean operators are written using HTML, Java Script Java and Java Servlets with the data taken from the appropriate Resource database table. This is in turn thrown to the different systems using the Common Gateway Interface (CGI) protocol to

facilitate searching. Figure 8 shows an example of the display output using the combined search facility. The use of the combined search facility reduces the burden of the user visiting different sites and relieves the task of forming individual queries for each system albeit losing some of the distinct search support that are not common across systems. It is especially useful when a wide coverage search is desirable.

Figure 8: Results of Combined Search

Summary of DWE Search Support

DWE supports both the browsing and searching mode for locating information resources that are interfaced to the system. These differing modes are meant to be used as complimentary alternatives to support the information seeking process. Most importantly, the task-based approach envisages the natural way to seek and use information through the accomplishments of tasks so that unwanted and irrelevant information are filtered out, leaving the most useful and relevant information for the user. These resources, however, have to be maintained and updated constantly by the Resource Administrator to make it useful for users. Apart from the task-based and "all-resources" browsing, DWE supports individual and combined direct search facilities to help users gain access to other information resources that are available within and outside the DWE environment. Preferred Internet search engines are selected by users for Internet searching. The combined search facility reduces query formation time, searches heterogeneous information systems/sources all at one time from a single query, and avoids the need to visit multiple search systems. As such, DWE has the potential to address the needs of both novice and expert users in different information seeking situations.

The Design Guidelines

In this section, we shall provide some generic guidelines that may be useful in designing a users-centred digital library system. Although these guidelines are based on our

experience in the context of an academic community, they may be equally useful in any other institution or user environment, with little modifications.

Gathering Information about the various Information Systems and Resources

As mentioned above, the DWE, and hence a user-centred digital library system, provides two major facilities: (1) it allows users to choose a particular task and then automatically brings up, or links users to, the appropriate digital information resources – library resources, internet and intranet resources; and (2) it allows users to conduct a search in a chosen or in a range of information resources and/or systems. In order to be able to perform the first set of tasks, the system should know the:

1. various categories of users and the tasks performed by them;
2. various information resources, their nature and characteristics, such as availability, organization, how to access, etc.;
3. information resources, and/or systems, that are necessary to perform a particular task, and so on.

The DWE handles these issues through the *Information Resource Administrator*, and the *User Manager*. The role of the *Information Resource Administrator* is to capture and organize information about all the different types of information systems and resources. This is a continuous process since new information resources and/or systems appear every now and then, and the nature, characteristics, location, etc., of the resources change from time to time. For a large digital library system, the job of the Information Resource Administration has to be performed by a group of appropriately trained information professionals. Critical study of the available library resources, Internet and intranet resources on different subjects have to be made to capture information about the resource/system itself as well as its usefulness and suitability to accomplish various user tasks.

Gathering Information about Users

The role of the *User Manager* is equally important and critical. This involves identification of the various categories of users (and this means identifying as many distinct user groups as possible so that they can be categorized according to a set of specific tasks), and then identifying the various distinct tasks performed by them. This is again a continuous process, since in an institution, especially in a large University, the type and category of users change, and so do their tasks, with changes in the University policies, curricula, etc. This is again a massive job for a large academic institution, and hence has to be performed by a group of appropriately trained information professionals.

Design Guidelines and Standards

As shown in the architecture of the DWE, there is a web interface that allows the end users to interact with the system. However, at the back end there are also databases that store information about the various information resources, users, their tasks, etc. The interface interacts with the databases through the appropriate programming interfaces. Although a designer may choose his/her own programming language, database software, etc., for this, it is important to make sure that the interface is not tied to any specific hardware or software from the end-user perspectives. Similarly, since the information

about the information resources as well as users and their tasks change quite often, the databases should allow frequent changes and updates.

There are also the question of standards, and they relate to the ways information about the various information resources are stored in the databases, ways in which the various screens, user tasks, and other forms of information are displayed on the screen, and the overall design features of the user interfaces. Again these are critical decisions, and should be decided by the system designer based on the user as well as the institutional needs and characteristics, available resources, frequency and speed of access, and so on. The DWE, described in this paper uses simple design features and generic standards like the Dublin Core, but one may change them to suit the specific user and institutional needs.

IR from Heterogeneous Information Systems and Resources

One of the most difficult problems of a user-centred digital library system is to allow users to conduct search on one or more information systems and resources through a single query and thereby reducing the hassles of learning the information retrieval features, user interfaces, etc., of the various information systems and resources. As mentioned earlier in this paper, the information systems and resources may range from a an intranet resource sitting on a local website, to an internet resource searchable through a specific URL or by using a web search tool, a library resource that can be searched through a specific search interface such as an OPAC, an online journal, a CD-ROM or an online database accessible through a vendor, and so on. The first and the most important thing to do in this case is to design a simple, and yet useful, search interface that can be used by the users to enter a query. This may be a simple search screen allowing users to enter just the URL, keywords, etc., or a detailed search screen allowing users to search by various keys, such as keywords, title, URL, subject descriptors, etc. The difficult part is to translate the query, automatically behind the scene, into appropriate search statements for every information system/resource and then to throw the queries to each chosen information system/resource. Again when the search results come back, the interface has to make sure that they are displayed for the user in a sensible way. Again all these technical considerations should be left with the system designer, and one should make use of the appropriate web programming language such as Java Applets, Java Servlets, Java Script, VB Script, ASP, etc., and browser features, to make these happen. At present, the combined search facility has been provided for selected information systems that interact with DWE. Designing a common search interface is a difficult task as it involves information systems that are developed by many vendors and whose proprietary search mechanisms and functionality can change anytime. This is an issue of ongoing research to explore how best this can be made portable and flexible to support such imminent changes.

System Evaluation

This is the last, but definitely not the least, part of any system design job. A detailed usability study has to be carried out with the digital library system to make sure that it meets its stated objectives. In our ongoing project, we shall be carrying out and concluding the evaluation of the DWE that will be reported in a separate paper. We

envisage conducting a comparative study against additional new versions of interfaces to support different means of interacting with information resources. For example, the concept of a personalised information environment such as MyLibrary can be emulated so that users work with a list of all resources and make use of their own personal space to organise and manage their own set of information resources that are extracted from the master list.

Conclusion

In this paper, the architecture and workflow of a digital work environment, that forms the basis of a user-centred digital library system has been described. This follows a set of general guidelines that may be used to design and develop a user-centred digital library system. Although our experience, and consequently the guidelines, are based on a University scenario, we believe that these may be useful for designing user-centred digital library for any user community.

References

Cohen, S., Fereira, H., Horne, Kibbee, A., Mistlebauer, H., Smith, A. (2000). MyLibrary: Personalised Electronic Services in the Cornell University Library. *D-Lib Magazine,* 6(4). Available at: http://www.dlib.org/dlib/pri100/ mistlebauer/ 04mistlebauer.html.

Cousins, S. (1996). A Task-oriented Interface to a Digital Library. *Proceedings of the CHI '96 Conference Companion on Human Factors in Computing Systems: Common Ground* (1996), 103 – 104. Available at: http://www-diglib.stanford.edu/diglib/WP/ PUBLIC/DOC44.html.

EveryWare Development Inc. (1999). Tango Enterprises 3:User's Guide (1999). Mississauga, Canada.

Fox, E.A., Hix, D., Nowell, L.T., Brunei, D.J., Wake, W.C., Heath, L.S., and Rao, D. (1993). Users, User Interfaces, and Objects: Envison, a Digital Library. *Journal of the American Society for Information Science*, 44(8), 480-91.

Gambles, A. (2001). The HeadLine Personal Information Environment, *D-Lib Magazine,* 7(3). Available at: http://www.dlib.org/dlib/march01/gambles/03gambles.html.

Garvey, W.D., Tomitaa, K. and Woolf, P. (1979). The Dynamic Scientific-Information User. In: Garvey, W.D. Communication: The Essence of Science. Elmsford, NY: Pergamon Press; 1979, 256-79.

Gates, K.F. and Lawhead, P.B. (1998). Towards and Adaptive WWW: a Case Study in Customised Hypermedia. *New Review of Hypermedia and Multimedia*, 4, 89-113.

Leong, C.K. (1998). Searching Heterogeneous Multi-lingual Bibliographic Sources. *Computer Networks and ISDN Systems,* 30(1/7), 612-15.

Lesk, M. (1991). The CORE Electronic Chemistry Library. *Proceedings of the 14th Annual International ACM/SIGIR Conference on R&D in Information Retrieval*, Chicago, IL, 92-112.

Meyyappan, N. (2000a). Design and Implementation of Digital Work Environment (DWE). (Unpublished Technical Report TR-IS (SCE) #07/2000, Nanyang Technological University, Division of Information Studies, September 2000).

Meyyappan, N., Chowdhury, G. G. and Foo, S. (2000b). A Review of the Status of Twenty Digital Libraries. *Journal of Information Science*, 26(5), 337-55.

Meyyappan, N., Chowdhury, G.G. and Foo, S. (2001a). Use of a Digital Work Environment Prototype to Create a User-centered University Digital Library. *Journal of Information Science*, 27(4), 2001, 249-264.

Meyyappan, N., Chowdhury, G.G. and Foo, S. (2001b). An Architecture of a User-Centred Digital Library for the Academic Community. In: NIT 2001, *Conference Proceedings on Global Digital Library Development in the New Millennium, Beijing,* May 29-31, 175-82.

Mick, C.K., Lindsey, G.N. and Callahan, D. (1980). Towards Usable User Studies. *Journal of the American Society for Information Science,* 31(5), 347-56.

Moussawi, W., Pierre, S. and Hoang, H.H. (1999). Modelling Intelligent Information Resource Agents for Dynamic Heterogeneous Sources. Engineering Solutions for the Next Millennium. *1999 IEEE Canadian Conference on Electrical and Computer Engineering. IEEE.,* Part Vol.2, 1999, 990-5.