Chapter 8

Information access and user interfaces

Outline

The central theme of this chapter is information access with particular reference to the user interfaces of digital libraries. It begins with a general discussion of information access and information-seeking models. It then examines the various stages of an information search process and user interface issues in the context of information seeking in a digital library environment. Sample digital library interfaces are shown to describe the various types of interfaces available to support the browsing, searching and display of digital information. Finally, the chapter mentions some areas of research into improvement of information access in digital libraries in order to support the creative activities of the end-users.

Introduction

One of the major objectives of a digital library is to provide improved access to information. In fact, as mentioned in Chapter 1, developing the enabling technologies to support improved access to digital information has been the primary goal of many digital library researchers. Digital library collections have several characteristic features that make information searching difficult. These collections are typically very large. They may involve many different kinds of objects, including all kinds of electronic publications – books, journals, conference proceedings, reports, images, audio, video and databases. Even within a single category, these objects may have widely different formats and internal structures. Furthermore, they have complex relationships with each other, and are required by different categories of user to meet different types of information needs.

Since digital libraries do not have any physical presence, users do not get a view of the collection and the contents as they do in a printed library. In fact, users of a digital library may not know much about its contents, and thus may find it difficult to formulate a query to get access to the required information. Users may not know how exhaustive the collection is in their area of interest or the size of the entire collection. Hence, it may become difficult for them to assess how exhaustive the retrieved information set at the end of a search is, and how precise the search results are. Determining how much information to show the users of a digital library is a major design choice in information access interfaces (Hearst, 1999).

Borgman (2000b, 79) comments that the concept of access to information has its roots in different areas such as library services, telecommunications policies, and so on. She defines access to information (Borgman, 2000b, 57) as connectivity to a computer network and to the available content, given that the technology is usable; the user has the requisite skills and knowledge; and the content is in usable and useful form. Thus, the three major factors influencing access to information are the technology, content, and above all the users.

In this chapter we discuss the basic issues of information access to digital libraries and the concept of user interfaces and interface design issues. A close look at the interfaces of some digital libraries reveals the different practices followed by designers to support the browsing, searching and display of digital information.

Information users

Theoretically speaking, a user of a digital library may be anyone living anywhere in the world. This is especially true for those digital libraries that aim to offer information access to a global audience, for example, the Networked Digital Library of Theses and Dissertations (NDLTD), the Networked Computer Science Technical Reference Library (NCSTRL) or the Greenstone Digital Library. Users of these digital libraries may vary in terms of their information need, characteristics, capabilities, and so on. Some digital libraries aim to cater for a rather homogeneous group of users though they may not be bound by any geographical boundary. Examples of such digital libraries include the Association of Computing Machinery (ACM) digital library and Institute of Electrical and Electronic Engineers (IEL) digital library. Yet there are other digital libraries that are designed specifically to support a defined group of users.

Most university digital libraries, for example the California Digital Library and HeadLine, belong to this category.

However, as Nicholas and Dobrowolski (2001) argue, the characteristics and role of users in a digital environment vary significantly from those in a traditional library environment. They propose a new term, 'information players', to replace the word users or end-users in order to reflect the changing nature of the users in a digital environment.

Users in a digital library environment need not only have adequate information literacy skills, but also some ICT skills to help them make optimum use of the digital environment. The SCONUL position paper (SCONUL, 1999) identified seven information skills that are considered necessary in a digital library environment:

- the ability to recognize a need for information
- the ability to distinguish ways in which the information need may be addressed
- the ability to construct appropriate information strategies
- the ability to locate and access appropriate information
- the ability to compare and evaluate information obtained from various sources
- the ability to organize, communicate and apply information in problem-solving tasks or in decision making
- the ability to synthesize and build upon existing information, and to contribute to the creation of new information.

Information needs

Information needs and how to provide access to information has remained a central theme of study among information and computer science researchers for many years. It is commonly agreed that the provision of information begins when a user has an information need. A significant amount of research on such matters as information needs and users' information-seeking behaviour has been conducted, and a large volume of literature has been produced over the past decades (for example, Bates, 1977, 1979, 1981, 1984; Belkin, Oddy and Brooks, 1982a, 1982b; Case, Borgman and Meadow, 1986; Dervin, 1977; Ellis, 1989; Kuhlthau, 1988a, 1988b; Marchionini and Komlodi, 1998; Spink et al., 2002a, 2002b, 2002c, 2002d; Wilson, T., 1981, 1994, 1999). Information need is a stage where the user senses that it may be useful to know something that they

do not know at that particular point in time. Marchionini (1995, 6) comments that information seeking is a fundamental human process that is closely related to learning and problem solving. According to Borgman (2000b, 109) a need is a psychological construct, and it cannot be observed by a researcher, a librarian or an intelligent agent; only indicators or manifestations of needs can be observed.

The information needed by a user to accomplish a goal – to resolve a problem, to answer a specific question or to satisfy a curiosity – may be quick and brief factual data or exhaustive and detailed. Figure 8.1 shows a simple model of information access. Although it appears to be very basic, in essence several complex processes take place throughout. Some of these are technological and related to the information retrieval system, users interfaces, and so on. Other

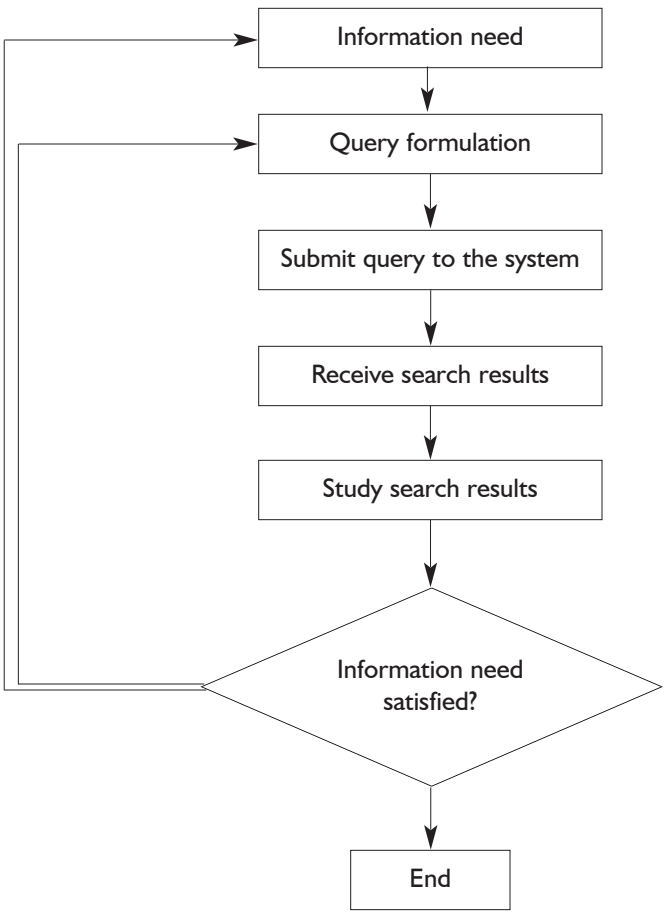


Fig. 8.1 *Basic information access model*

processes relate to the nature and characteristics of the content as well as the specific user. The process may take more or less time, and may become simple or complex depending on the nature of the users – their cognitive abilities and background, the specific nature of the information need, and so on.

Information seeking is an interactive process that depends on initiative on the part of the user, feedback from the information system, and the user's decisions about subsequent actions based on the feedback (Marchionini, 1995, 17). The user's initial information need may often change on receipt of some information. Hence, the information search process continues till the user gets the information required to satisfy the revised information need. Appropriate technology, such as a suitable information retrieval system and user interface, may facilitate the process, but is not the ultimate answer, because the process depends largely on an individual user and their information need, as well as the nature, volume and variety of the content.

Users often learn during the information search process. For example, they may come across some information that influences their information need. The user may also acquire new knowledge about the system, and thus be able to formulate queries more skillfully and appropriately to retrieve better output. The berry-picking model of information seeking proposed by Bates (1989) suggests that as a result of reading and learning from the information retrieved through the search process, users' information needs and the queries continually shift. The berry-picking model also suggests that the user's information needs are satisfied by a series of selections and bits of information found along the entire information search process, as opposed to the output of any particular search set. The ASK (Anomalous States of Knowledge) model proposed by Belkin and his associates (Belkin, 1980; Belkin, Oddy and Brooks, 1982a, 1982b) suggests that an information-seeking process begins with a problem, but initially that problem and the information needed to resolve it are not clearly understood. Hence, the information seekers need to go through an iterative process to articulate a search request, and the information system should support interactive searching.

The sense-making approach of Dervin (1977) posits that users go through different phases in making sense of the world. The first phase establishes the context for the information need, which she calls a situation. People find a gap between what they understand and what they need to know in order to make sense of the current situation. These gaps are manifested by the formulation of questions. The answers to these questions are then used to move to the next situation. Marchionini (1995, 29) comments that Dervin's model applies more to

general human conditions than to information seeking, but the model has been adopted by researchers in information science and communications as a framework for studying the information-seeking process. Kuhlthau (1988a, 1988b; Kuhlthau et al., 1990) proposed a model of how students search for information as part of their writing process. The process involves seven stages: task initiation, topic selection, pre-focus exploration, focus formulation, information collection, search closure and the starting of writing.

Bates (2002) suggests that each layer in an information system interacts with every other design layer, and this cascade of interactions culminates in the interface, where all the prior interactions have either worked to produce effective information retrieval or to produce system elements that work at cross-purposes. She thus proposes a design model, called the Cascade Model, for operational information retrieval systems. The basic proposition of this model is that without the design of all the constituent layers of an information retrieval system being integrated, the resulting system is likely to be poor. There are four layers in this model:

- The first layer comprises the infrastructure – network, hardware, software and databases.
- The second layer comprises the information or content combined with the metadata structure.
- The third layer represents the information retrieval system – from information in searchable form to the interface design supported by the technical infrastructure.
- The fourth layer is the human part of the system, comprising user searching activities and user understanding and motivation.

Nicholas (1996) stresses the need for qualitative assessment of user needs. He identifies 11 major characteristics of information need: subject, level, quality, place of publication or origin, function, viewpoint, date, processing and packaging, nature, quality and speed of delivery. A number of interesting issues related to the information seeking and searching in digital libraries have been discussed in a DELOS workshop (Boehm, Croft and Schek, 2000).

The four-phase framework for information search

Information searching is a complex process. It involves a number of stages and at each stage a number of actions are taken and decisions made. The

information retrieval system and the user interface may provide support in performing these actions and in making appropriate decisions. Shneiderman, Byrd and Croft (1997) divide the major activities in an information search process into four major phases: formulation, action, review of results, and refinement. They propose that this four-phase framework for interface design will provide a common structure and terminology for information searching while preserving the distinct features of individual digital library collections and search mechanisms (Shneiderman, Byrd and Croft, 1998).

Phase 1: Formulation

The formulation of a search is triggered by an information need, and several decisions are made regarding sources, fields, what to search for, and the search variants.

The selection of sources (collections and/or databases) is an important step in a search process. In a digital library environment, users may have access to many collections, and each collection may have one or more databases. Users need to have some idea about the nature and content of the collections and databases and use this to make a selection. Some digital library interfaces show a list of the available collections and allow users to select one particular collection; for example the Greenstone Digital Library allows users to select one particular collection to browse or search, and NDLTD allows users to search for the theses of a selected university or universities. In the case of the California Digital Library, users have many choices. For example, they may select a particular collection or use the Searchlight option to select by broad subject grouping like science and engineering, or social science and humanities. However, the selection of sources to search is not always an easy task, especially for new and novice users. Some systems provide support for this. In this case, users are asked to enter a search expression, and then the system searches across the databases and produces an output of best-matching databases instead of best-matching records. This gives the user an idea of the content of the collections and thus facilitates the selection of sources. In Dialog, for example, the DialIndex option allows users to search across a range of databases to get a list of those that best match their chosen search topic (Chowdhury and Chowdhury, 2001a, 141).

A search may be conducted against one or more selected fields in a database. A search on specific fields produces more specific search results than one on a complete record. However, it is sometimes difficult for the user to decide which field to search. This calls for a familiarity with the structure of the chosen data-

base and also with the nature and content of the fields. Users may go to the help files, or to some other source, for example to the blue sheets (pages that contain information on a database including its content, coverage, structure and indexing) in the case of a Dialog database search. Some systems provide search interfaces (usually in the advanced search mode) with structured fields. For example, in the ACM digital library's advanced search interface, users can select a number of fields to which to restrict a search (see Figure 8.2). Similarly, in THOMAS (a digital library service of the Library of Congress), users can select a number of fields with which to conduct and/or restrict a search (see Figure 8.3) for Congressional records.

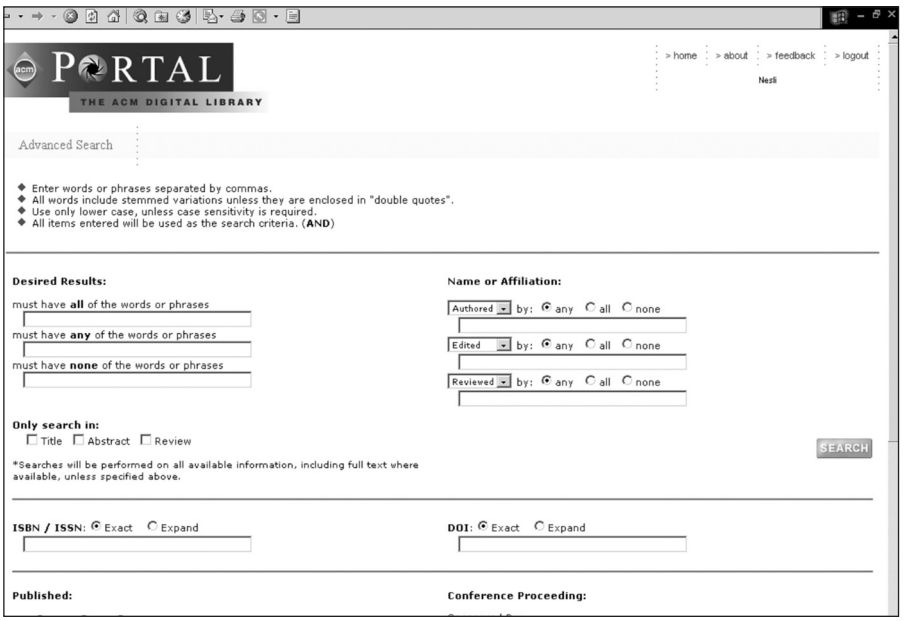


Fig. 8.2 *Advanced search interface of the ACM digital library*

A major challenge for users comes in writing the actual search statement. A search statement tells the system what to search for in the chosen database(s). Various techniques are available for specifying how the constituent search terms are to be looked for, for example by using appropriate search operators. The search operators are not always intuitive and are purely dependent on the chosen system. Users need to be familiar with the various operators and the conventions appropriate for the chosen search system. Examples of various search operators are provided in Chapter 9.

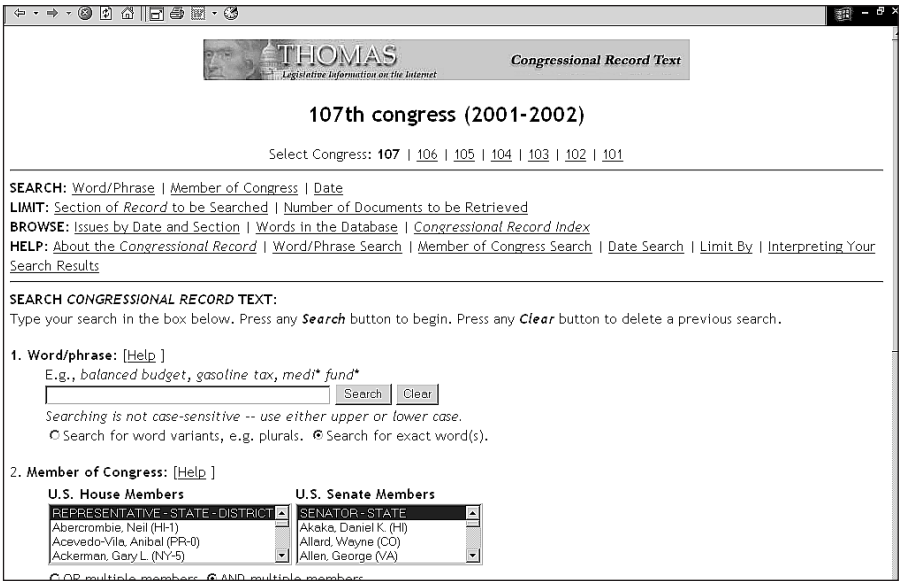


Fig. 8.3 *Search interface for text search in THOMAS*

A search term may be represented in a variety of ways. Users may want to choose a given search term or phrase that appears in the database records in a variety of ways, for example in singular plural forms, in various synonymous forms, with variant spellings, and so on. In such cases it can be very difficult for the user to decide which form or variant of a search term to use. The user interfaces of digital libraries often help by allowing for case sensitivity, stemming, phonetic variants, synonyms, abbreviations, broader and narrower terms, and stop words. Some of these issues are discussed in Chapter 9.

Phase 2: Action

Usually a search button needs to be pressed to conduct a search, although in some cases the user just needs to press <CR>. Once the search begins, the user is usually expected to wait till the search process end. Sometimes, this may take a long time and thus be frustrating. In some cases, the interface prompts the user that the search is being processed; it may also tell the user about the progress of the search (for example, Figure 8.4 shows the progress of a search in the California Digital Library search interface). A very appealing method of information searching is 'dynamic queries', a system where there is no search button and the result set is continuously displayed and updated as phases of the search are changed (Shneiderman, Byrd and Croft, 1997).

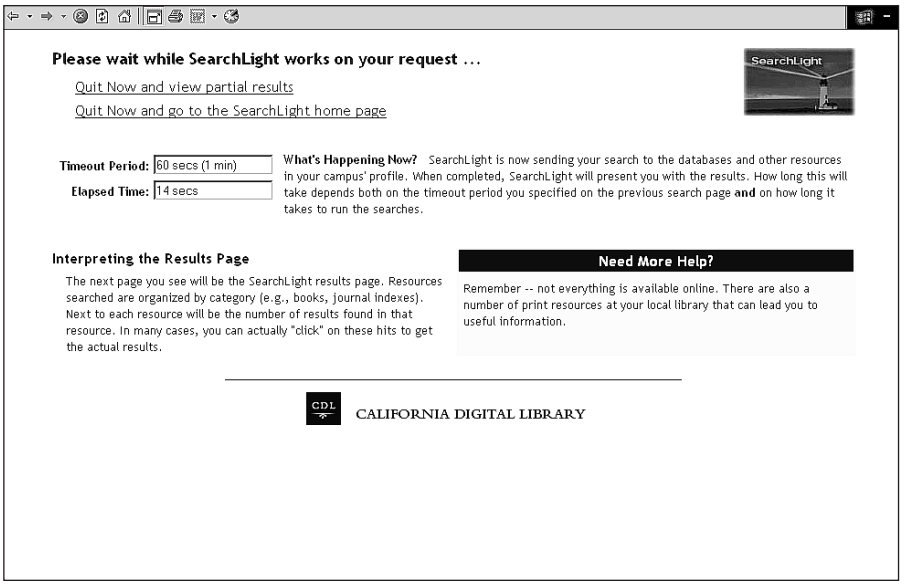


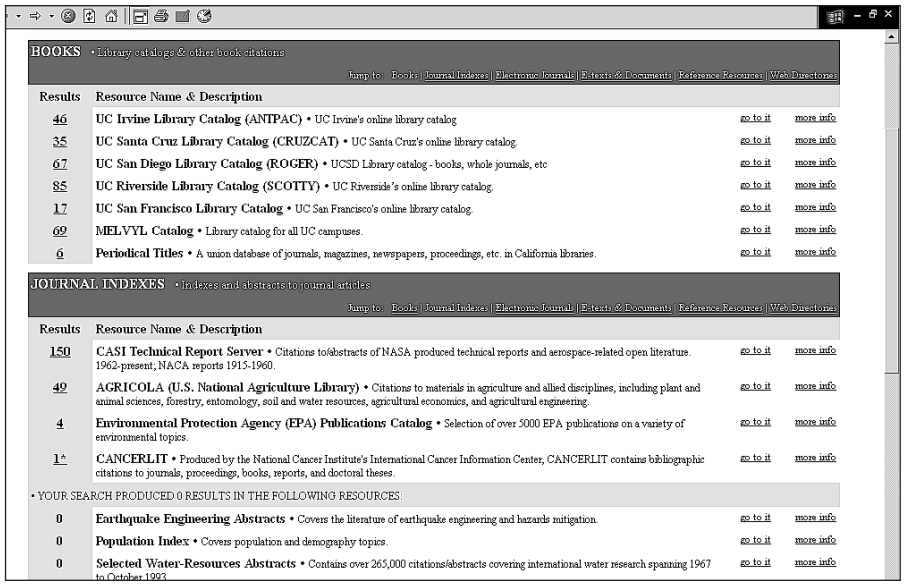
Fig. 8.4 *Progress of a search session in the California Digital Library*

Phase 3: Review of results

Information retrieval interfaces usually offer the user various choices for viewing results such as the size of display, display format and the order of the retrieved items (sorted by author, date, and so on). Some interfaces use different visualization techniques for the display of search results. Some interfaces also use helpful messages to explain the results, for example the degree of relevance. Some digital libraries, for example the California Digital Library, shows results from different collections separately (for example, Figure 8.5).

Phase 4: Refinement

Different search interfaces provide different facilities for modifying and refining queries. In some cases, users need to reformulate the search statement and conduct a new search, while in others users can refine a search and conduct a new search on the retrieved set. For example, in Dialog search, each search is automatically given a set number, and the user can specify any search set on which to conduct a refined search. Many digital library interfaces also support relevance feedback, a system where the user selects some retrieved output as relevant and the system conducts a new search based on the characteristics of the items identified as relevant (discussed further in Chapter 9).



BOOKS • Library catalogs & other book citations

Jump to: Books | Journal Indexes | Electronic Journals | Etexts & Documents | Reference Resources | Web Directories

Results	Resource Name & Description		
46	**UC Irvine Library Catalog (ANTPAC)** • UC Irvine's online library catalog	go to it	more info
35	**UC Santa Cruz Library Catalog (CRUZCAT)** • UC Santa Cruz's online library catalog.	go to it	more info
67	**UC San Diego Library Catalog (ROGER)** • UCSD Library catalog - books, whole journals, etc	go to it	more info
85	**UC Riverside Library Catalog (SCOTTY)** • UC Riverside's online library catalog.	go to it	more info
17	**UC San Francisco Library Catalog** • UC San Francisco's online library catalog.	go to it	more info
69	**MELVYL Catalog** • Library catalog for all UC campuses.	go to it	more info
6	**Periodical Titles** • A union database of journals, magazines, newspapers, proceedings, etc. in California libraries.	go to it	more info

JOURNAL INDEXES • Indexes and abstracts to journal articles

Jump to: Books | Journal Indexes | Electronic Journals | Etexts & Documents | Reference Resources | Web Directories

Results	Resource Name & Description		
150	**CASI Technical Report Server** • Citations to/abstracts of NASA produced technical reports and aerospace-related open literature. 1962-present; NACA reports 1915-1960.	go to it	more info
49	**AGRICOLA (U.S. National Agriculture Library)** • Citations to materials in agriculture and allied disciplines, including plant and animal sciences, forestry, entomology, soil and water resources, agricultural economics, and agricultural engineering.	go to it	more info
4	**Environmental Protection Agency (EPA) Publications Catalog** • Selection of over 5000 EPA publications on a variety of environmental topics.	go to it	more info
1^	**CANCERLIT** • Produced by the National Cancer Institute's International Cancer Information Center, CANCERLIT contains bibliographic citations to journals, proceedings, books, reports, and doctoral theses.	go to it	more info
• YOUR SEARCH PRODUCED 0 RESULTS IN THE FOLLOWING RESOURCES:			
0	**Earthquake Engineering Abstracts** • Covers the literature of earthquake engineering and hazards mitigation.	go to it	more info
0	**Population Index** • Covers population and demography topics.	go to it	more info
0	**Selected Water-Resources Abstracts** • Contains over 265,000 citations/abstracts covering international water research spanning 1967 to October 1993.	go to it	more info

Fig. 8.5 *Results of a search in the California Digital Library showing output from various categories of sources*

Information seeking and user interfaces

User interfaces to information retrieval systems that support information-seeking processes have been widely discussed in the literature. Mitchell (1999) comments that the user interface is the means by which information is transferred between the user and the computer and vice versa. Mitchell further comments that well-designed user interfaces should allow users to find and use the information that the information system provides access to more efficiently. In fact a good user interface greatly enhances the quality of interactions with information systems.

Interface design encompasses what appears on the user's screen, how they view it, and how they manipulate it. Functional design specifies the functions that are offered to the user such as selecting parts of a digital object, searching a list or sorting retrieved output, obtaining help, and manipulating objects that appear on the screen. Almost all present-day PCs have a user interface that is based on the style derived at Xerox PARC and made popular on Apple Macs, and uses the metaphors of files and folders on a desktop (Arms, 2000b, 45).

Shneiderman, the guru of HCI (human–computer interaction) and user interface design, proposes a number of guiding principles for the design of user interfaces (Shneiderman, 1998b; Shneiderman, Byrd and Croft, 1997):

- Strive for consistency: terminology, layout, instructions, fonts and colour should be used consistently throughout the interface.
- Provide shortcuts for skilled users.
- Provide informative feedback: the system should provide users with appropriate feedback about the sources and what is being searched for.
- Design for closure: users should know when they have completed searching the entire collection or have viewed every item in a browse list.
- Permit reversal of actions: users should be able to undo or modify actions, for example they should be able to modify their queries, or should be able to go back to the previous state in a search session.
- Support user control: the user should be able to monitor the progress of a search and should be able to specify the parameters to control a search.
- Reduce short-term memory load: the system should keep track of some important actions performed by the users and should allow them to jump to a formerly performed action easily, for example to a former query or to a specific result set.
- Make error handling facilities simple: users should be able to rectify errors easily, and all error messages should be clear and specific.
- Provide plenty of space: a lot of room should be made available for entering text in search boxes.
- Provide alternative interfaces: separate interfaces should be available for expert and novice users.

Bates (2002) stresses that interface design is pivotal to the effective use of an information system, and that the application environment of information retrieval systems has its own distinctive needs and characteristics that need to be understood and addressed in design. Hearst (1999) comments that an interface designer must make decisions about how to arrange various kinds of information on the screen and how to structure the possible sequences of user-system interactions.

Marchionini (1992) provides a description of the essential features of interfaces to support end-user information seeking and suggests five information-seeking functions, namely problem definition, source selection, problem articulation, result examination and information extraction. He argues that much interface work has focused on problem articulation (including query formulation) and that the other functions need to be investigated in designing information-seeking interfaces. Marchionini and Komlodi (1998) discuss the evolution of interfaces

and trace research and development in three areas, namely information seeking, interface design and computer technology. They provide a brief review of interfaces to online information retrieval systems as well as to online public access catalogues (OPACs). They also discuss the new generation of user interfaces influenced by the emergence of the web. They conclude that interface design has become more user-centred and that the trend is toward more mature interfaces that support a range of information-seeking strategies.

Savage-Knepshield and Belkin (1999) discuss the trends related to interface design challenges within the context of information retrieval interaction over the last three decades. They divide the period into three major eras, which they refer to simply as the early years, the middle years and the later years, and provide a description of the types of interfaces designed in each. Command language interfaces provided the main approach in the early years. In the middle years menu-driven and form-fill-in interfaces, which were more appropriate for novices and casual searchers, became the dominant interface type. In the later years, users and their information needs became the focus of the most complex interface design challenges. This period is characterized by use of the natural language and direct manipulation user interfaces. The authors note that the degree of interaction between the searcher and the IR system has dramatically increased but that much research is still required to meet the challenges in interface design for IR interaction.

Hearst (1999) discusses user interface support for the information-seeking process and describes the features of these interfaces that aid such processes as query formulation and specification, viewing results and interactive relevance feedback. She describes a number of graphical user interfaces that provide information seekers with a wide range of approaches to specify, view, analyse and evaluate queries and documents within the context of information retrieval systems. Interfaces that support the formulation of Boolean and natural language queries as well as those providing categorical and subject support are examples of those reviewed. Hearst points out that there is an increasing interest in taking the behaviour of individuals into account when designing interfaces.

User interfaces and visualization

Since human beings are highly attuned to images, and since visual representation facilitates rapid and easy communication, several visualization techniques have now been applied to the design of user interfaces. Various graphical

representation and manipulation methods are used to represent information on the user screens though the visualization of textually represented information is challenging (Hearst, 1999).

Users of popular operating systems and common software packages use a number of visual tools and techniques for day-to-day operations. These include the icons, colour highlighting, windows and boxes, and so on. The most commonly used visualization techniques used in user interfaces for information access include the following (Hearst, 1999; Kakimoto and Kambayashi, 1999; Rao et al., 1995; Robertson, Card and Mackinlay, 1993):

- *Perspective wall*. This resembles a grey wall folded into three parts and provides a sort of a fish-eye view (a three-dimensional picture taken with a special lens called the fish-eye lens). The centre panel provides a detailed view, and the two wings provide a contextual view. It is suitable for information that has a linear structure.
- *Cone tree*. This provides a fish-eye view by displaying the nodes that are closer larger and more brightly than the ones that are further away. It is suitable for information that has a hierarchical structure.
- *Document lenses*. These are used to focus on one page in a document.
- *Hyperbolic tree browser*. This is used to show the hierarchical structure of a collection as a hyperbolic tree (there is an example in Figure 8.9).
- *Brushing and linking*. This involves connecting two or more views of the same data in such a way that a change to the representation of one view affects the representation of the other.
- *Panning and zooming*. This mimics the actions of a movie camera that can scan sideways across a scene, panning, and can move in for a close-up or back away to get a wider view, zooming.
- *Focus plus context*. One portion of the collection is made the focus of attention by making it larger while shrinking the surrounding objects that form the context.

User interfaces of digital libraries

Digital libraries vary in terms of design, objectives, characteristics, content and users. Consequently, many different types of digital library user interface can be found. While some of these user interfaces are simple, others are sophisticated in terms of design features as well as visualization techniques. In this section we shall brief discuss the user interfaces of some selected digital libraries and

related services. The examples are by no means exhaustive, but give an idea of the various design techniques used in digital library interfaces for supporting various activities, such as query formulation, display of search results, and so on.

The digital library user interface is an exciting area of research, and researchers from various fields are now working in this area. The first International Workshop on Visual Interfaces to Digital Libraries was held at the first Joint ACM and IEEE conference on digital libraries in Roanoke, Virginia, on 28 June 2001. This one-day workshop drew an international audience of 37 researchers, practitioners and graduate students in the areas of information visualization, digital libraries, human–computer interaction, library and information science, computer science and geography. The primary aim of the workshop was to raise the awareness of several interconnected fields of research related to the design and use of visual interfaces to digital libraries, especially in information visualization, human–computer interaction and cognitive psychology (Börner, 2001).

The eighth DELOS Workshop on User Interfaces in Digital Libraries, held in Stockholm, Sweden, 21–23 October 1998, discussed the following issues: (Hansen, 1999; Hansen and Karlgren, 1999):

- information seeking and retrieval as embedded activities within digital libraries
- techniques and methods to analyse and evaluate different systems as well as different users, their behaviour, tasks and the ideas behind the systems developed
- interactions with information, such as texts and multimedia, and access to multilingual information in information-seeking activities
- alternative modalities for representations of information-seeking activities in digital libraries.

The major problem for designers of a digital library user interface is their lack of knowledge about the users, their infrastructure, needs and characteristics, since virtually anyone anywhere in the world can be a potential user. A good design must be effective in a range of computing environments. DLITE (the Digital Library Integrated Task Environment) is an example of a user interface model (developed at Stanford University as part of the Digital Library Initiative), which takes into account the tasks that the users of a digital library

typically carry out (Cousins et al., 1997). It describes digital libraries in terms of four major components: documents, queries, collections and services. DLITE uses the InfoBus technology, which provides a unifying framework for bringing together services provided on the web as well as other traditional information retrieval services. Baldonado (2000) and associates (Baldonado and Winograd, 1996) describe the design and implementation of SenseMaker, an interface for information exploration across heterogeneous sources. The basic tenet of this approach is that a user's interests go through a process of development and that SenseMaker supports their context-driven evolution. SenseMaker allows users to first examine the current context and, secondly, progress from one context to the next.

Information access in digital libraries

The two basic modes of access to information in digital libraries are browsing and searching. Most digital libraries provide facilities for browsing as well as searching. In this chapter we shall take a look at the user interfaces of some digital libraries highlighting their features for supporting browsing, query formulation and display of search results.

A number of portals and subject gateways have been developed over the past few years to facilitate access to digital information. These gateways are not digital libraries per se; some prefer to call them virtual libraries since they provide organized and systematic access to the information resources on the web. Examples of such gateways include BUBL, EEVL, SOSIG, Biz/ed and BIOME. Chowdhury and Chowdhury (2001b, Chapter 5) provide useful discussions on the characteristics of these subject gateways.

Digital library interfaces and browsing

Digital library interfaces offer different types of browsing facilities. In this section we shall look at some such interfaces to get an idea of the different browsing options and features. Figure 8.6 and 8.7 show the browsing interfaces of the Virginia Tech ETD (Electronic Theses and Dissertations) digital library. When users open the VT ETD web page, they are offered the choice of either the search or the browse by author option. If the browsing option is chosen, a page appears which divides up the theses according to the letter of the alphabet with which their author's surname begins, gives a total number of theses for each letter, and offers a choice between them (Figure 8.6). This screen also tells the user which kind of access facilities are available for those theses. Once the user

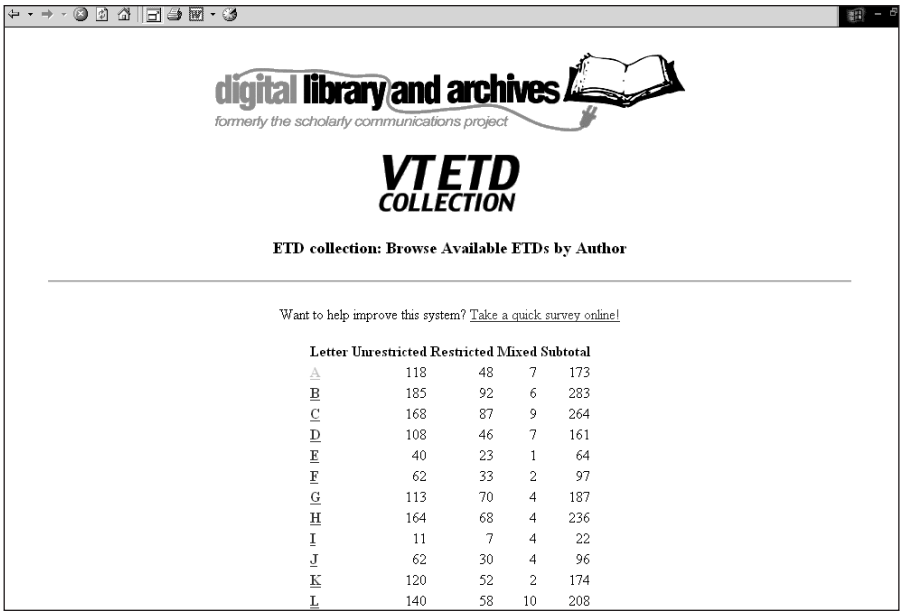


Fig. 8.6 *Virginia Tech ETD browse screen (1)*

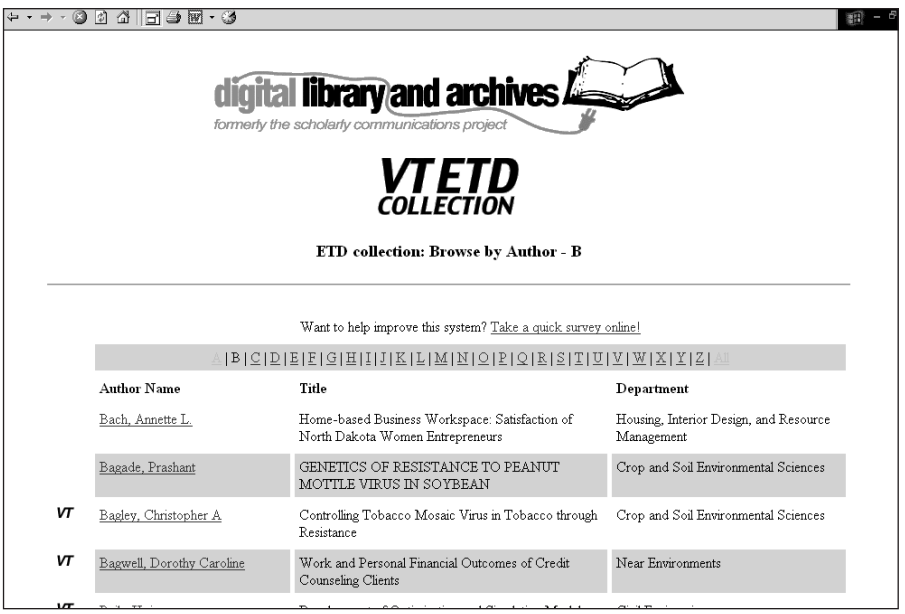


Fig. 8.7 *Virginia Tech ETD browse screen (2)*

chooses a specific letter of the alphabet, another screen (Figure 8.7) appears showing the actual titles available arranged alphabetically by author. The list also shows the name of the department. A click on the hyperlink on the author name takes the user to another page showing the metadata of the concerned dissertation, from where the user gets a link to the full text. The metadata page also shows the required download time for different types of access – through different types of modem, ISDN, and so on. Thus, a user has to follow four steps to get access to the full text, if available, of a thesis.

Figure 8.8 shows the browse screen of NCSTRL. Here, users can browse the documents by archives. The screen has two parts: the left part shows the list of archives and the total number of items available in each archive, and if the user selects any archive, the right window shows the list of documents available in it. A click on the title of a selected item takes the user to another screen which shows the abstract and other details, including a document identifier that takes the user to the full text of the item. The Universal Library, hosted by Carnegie Mellon University, provides a demonstration of a browsing screen using the hyperbolic tree from Inxight Software (www.inxight.com). The various topics for each subject collection appear as nodes on the hyperbolic tree which take the user to the different parts of the collection (Figure 8.9).

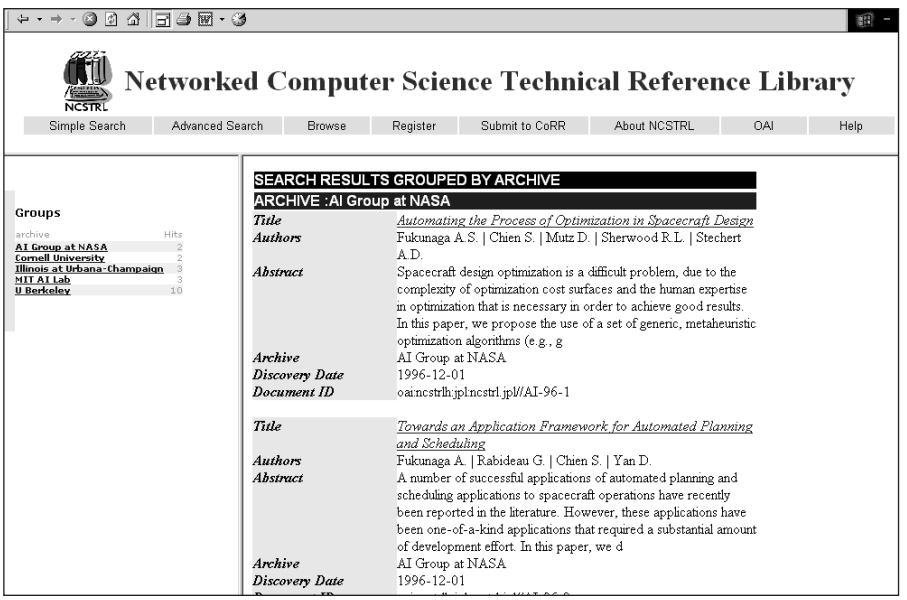


Fig. 8.8 *NCSTRL browse interface*

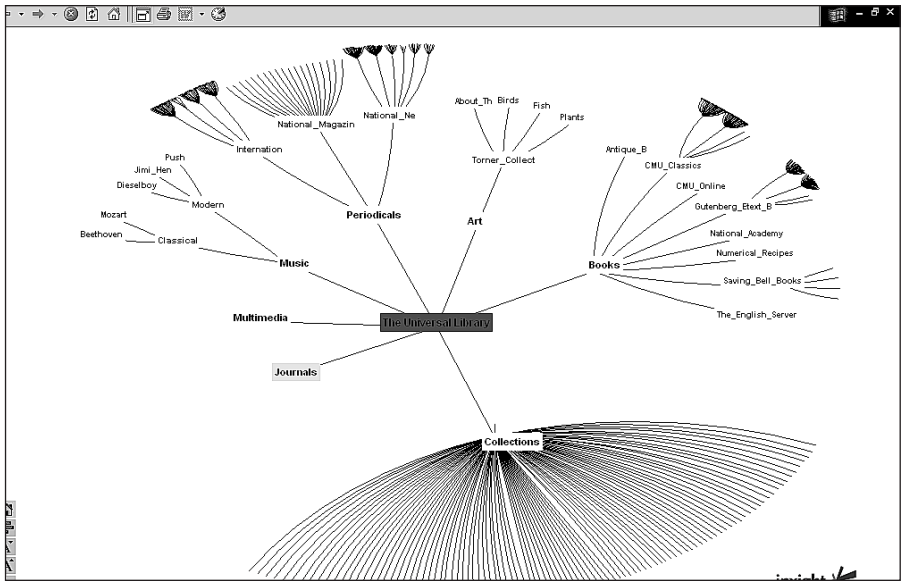


Fig. 8.9 *Universal Library browse interface*

Digital library interfaces and search and display

Digital library interfaces vary significantly in terms of the search features and facilities provided. They also differ significantly in terms of such matters as design, layout, conventions, fonts and colours. In this section, we shall look at the search interfaces of some selected digital libraries with a view to getting an idea of the different varieties that are available. These are by no means exhaustive, and indeed there are many more.

We shall begin with a very simple search interface, that of the NCSTRL digital library shown in Figure 8.10. This search screen allows users to enter a search or phrase, group the results by either archive or year, and sort the results by relevance or date. The simple search interface of the Virginia Tech ETD (Figure 8.11) provides a box for entering search term(s), and allows users to select one of the two collections. Here, users can also choose the advanced search option. Hence, even if users want to use the advanced search screen, they have to come first to the simple search screen from the main VT ETD page.

The advanced search screen (Figure 8.12) provides a number of options to formulate complex queries. Here users can select a field to search and specify whether the search terms should be treated as individual search terms or a phrase. They can also specify Boolean search conditions, but the conventions

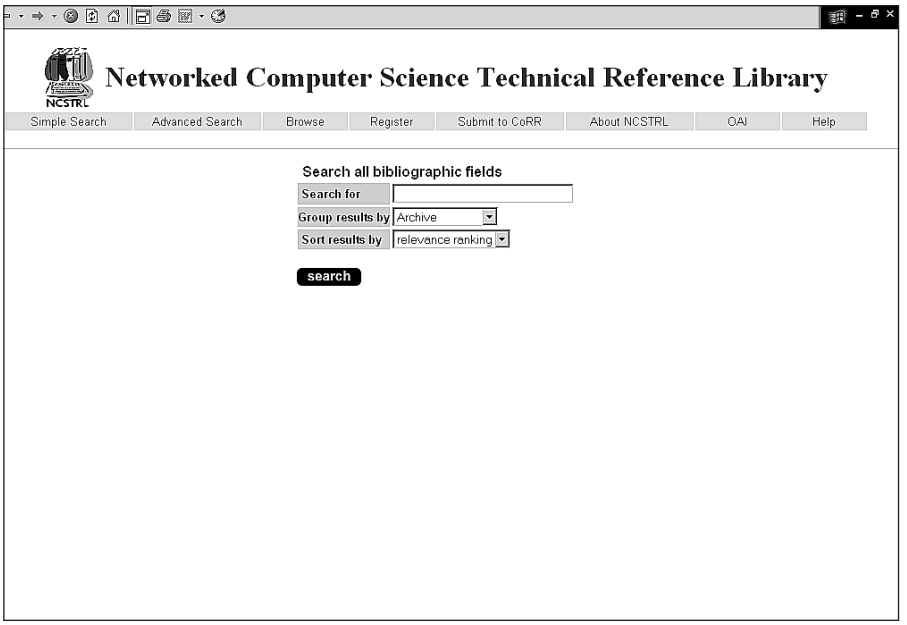


Fig. 8.10 *Simple search interface of NCSTRL*

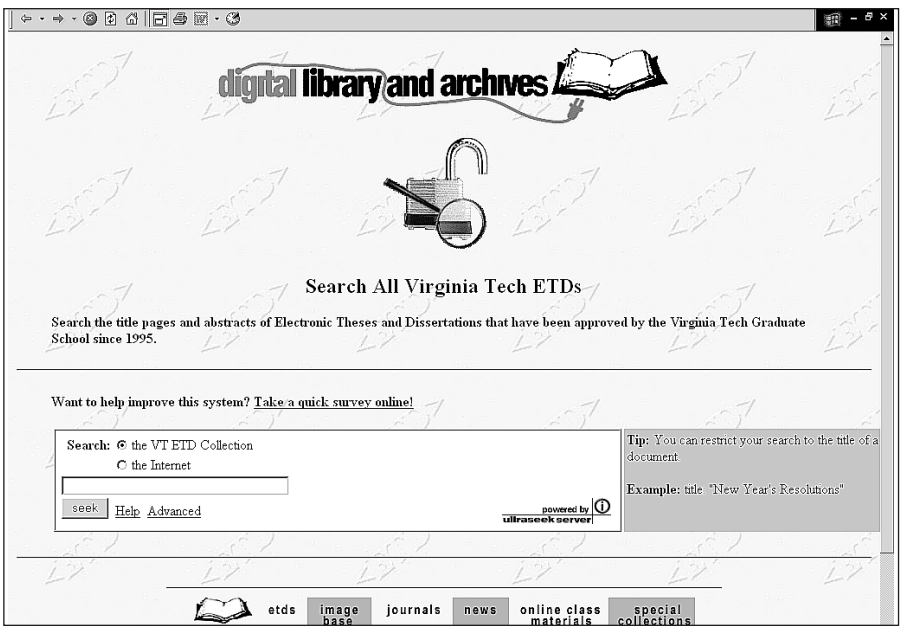


Fig. 8.11 *Simple search interface of VT ETD*

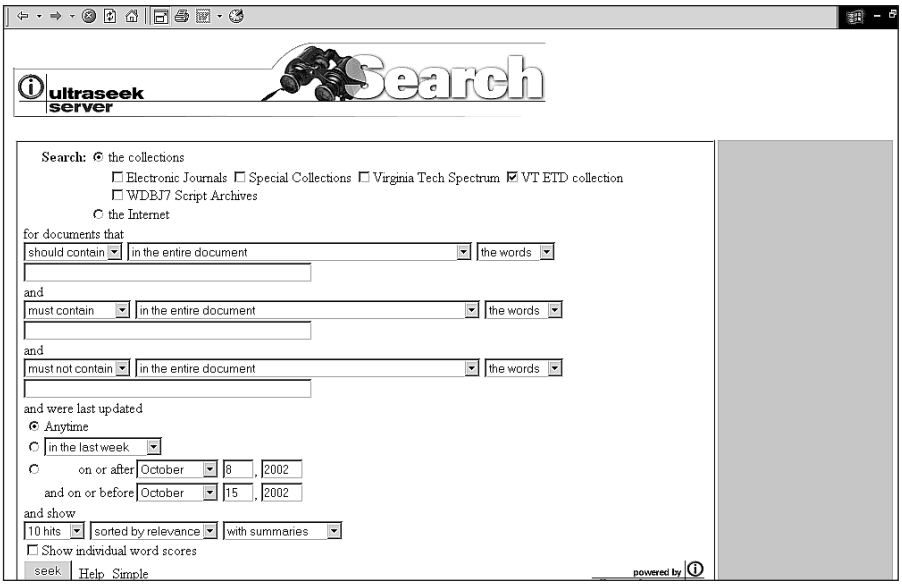


Fig. 8.12 *Advanced search interface of VT ETD*

are different here – 'should contain', 'must contain' and 'must not contain'. Users can also select display conditions, such as the number of hits per page, the criteria for sorting and presentation (with or without summaries). All these options can be chosen from one page.

However, when users conduct a search using the advanced search screen, the same search screen re-appears with the search results at the bottom. At times, it may be confusing for users, because unless they scroll down the screen they do not get any clue as to what has happened. However, once a search is conducted, they can select the option 'Search these results' to conduct a new search on the retrieved set, or can choose to 'search entire web'; if users do not choose either of these options, the ETD collection will be searched. Users also have to alter the display format by choosing 'hide summaries'.

If the user has used the simple search interface, then the results display interface appears differently – it is the same results display as for the advanced screen, but in this case it appears on top. Figure 8.13 shows the output of a search using the simple search interface of VT ETD. An interesting display screen appears in the California Digital Library interface where search results in each collection are shown separately (see Figure 8.5).

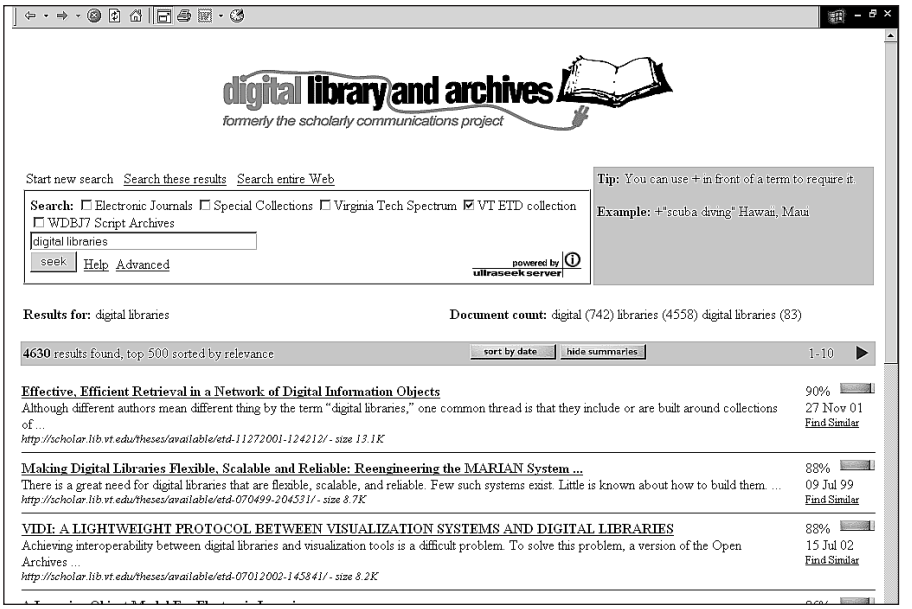


Fig. 8.13 *Search results from VT ETD*

User-centred digital libraries

While many of its users may be local, as in the case of a university library, a digital library by definition should provide global access both in terms of content and users. In other words, the users of a digital library may be located anywhere in the world. Since digital libraries are meant for end-users and eliminate the physical presence of human intermediaries, the only link between a digital library and its users is the user interface. Ideally speaking, therefore, the user interface of a digital library should be so simple and intuitive that it can be used easily and comfortably by anyone from anywhere at any time. This is practically an impossible proposition since a single interface cannot serve all the potential users, who may vary in terms of age, subject interests, profession, social and economic background, linguistic and cognitive abilities, and so on. Also the same interface cannot be suitable for all types of information content. Many digital library researchers have proposed user-centred design of digital libraries. Some of these user-centred models have been discussed in Chapter 4. Brophy (2001) argues that in the future libraries will need to keep detailed and accurate data about their users in order to provide personalized services. Such user information may include user activities, subject interests, display and reading preferences, and so on.

User-centred design digital library interfaces have been proposed by many researchers (Baldonado, 2000; Fox and Urs, 2002; Marchionini and Komlodi, 1998; Meyyappan, Chowdhury and Foo, 2001a, 2001b, 2001c; Sutcliffe, 1999 and Theng et al. 1999), and several researchers have proposed information access models to support creativity (for example, Ford, 1999; Shneiderman, 1998a, 1999).

Web browsers have significantly improved and standardized the features of user interfaces; thus it is now possible to have more uniform access to digital information. Before the emergence of web browsers, developers had to provide a separate interface for each type of computer and system. Now, the designers of digital libraries can to a large extent rely on the web browsers for a standard interface for the users, leaving aside the complexities of hardware and software variations, and can concentrate on how to organize the flow of information to the user (Arms, 2000b, 147).

Shneiderman (1998a, 1999) proposes the *genex framework*, which supports creativity through four phases:

- *Collect*. Learn from previous works stored in digital libraries.
- *Relate*. Consult with peers and mentors.
- *Create*. Explore, compose and evaluate possible solutions.
- *Donate*. Disseminate the results and contribute to the digital libraries.

He further describes eight activities that need powerful interfaces to support creative work. In other words, he proposes eight areas that need attention from researchers to make future digital libraries useful for creative work. These activities are (Shneiderman, 1999) as follows.

1 *Searching and browsing digital libraries*. Users should have more control over searching and browsing so that they can make use of their prior knowledge and retrieve information that supports their creative activities. Since searching is a part of the entire creative process, users should be able to save the search results into appropriate systems or software for future use, for example on a spreadsheet for further manipulation, as a file that can later be used for consultation with peers, or on a personal notebook for later referral.
2 *Consulting with peers*. Users may often consult their peers about new findings or research ideas. Information is collected at different stages in the consultation process. Different tools and techniques are also used for

consultation. This concerns the design and use of technologies for the interface design, since the appropriate balance of privacy with rights to, and ease of, access to information is very important.

3 *Visualizing data and processes*. Interfaces that support visualization of a digital library's contents are very useful and further work is necessary for the smooth integration of the technologies. For example, the interface should allow the user to view the results of a search using appropriate visualization tools that would help them select the most appropriate results. It should then be possible to export those results to appropriate packages, for example to a spreadsheet or to a database, and eventually to include the processed information in a report or presentation.

4 *Thinking by free association*. The association of ideas and using related concepts is a useful method of thinking and creativity. Thesauri are sometimes used in information retrieval systems to support the association of search terms. Various online tools are now available to associate concepts, for example IdeaFisher (www.ideafisher.com) and MindManager (www.mindman.com). Search interfaces should allow users to use these tools appropriately throughout the search process.

5 *Exploring solutions*. Digital libraries can help users make a decision by making use of the appropriate information and software. Various generic software packages allow users to explore solutions; one prominent example is the spreadsheet software that has 'what-if' tools to help users explore the various alternatives. Digital library interfaces should help users run a simulation, save the whole session, and/or send the session to someone for further discussion and exploration. The digital library interface should also allow users to store any successful sessions to help future researchers build on the best work.

6 *Composing artefacts and performances*. A number of software packages are available that allow users to compose artefacts and performances. Digital library interfaces should enable the integration of the results of a search with such interfaces, for example templates and macros for business operations, Adobe Photoshop Macros for redoing images, or music composition programmes in the case of music digital libraries to help users become more creative.

7 *Reviewing and replaying session histories*. Digital library users may like to replay previous sessions to get some new information, or to begin from there for a new search session. However, as Shneiderman (1999) comments,

success in this requires careful user interface and software design to ensure that the results are compact, comprehensible and useful.

8 *Disseminating results*. New information may be disseminated to different types of user. One possible group would be previous and current researchers in the field. Digital libraries should allow users to find workers in a field of study easily. Shneiderman (1999) recommends that digital libraries should be conceived of as digital library communities by the extensive use of online community software to turn every object into the focus of a discussion group.

Thus research and evaluation are necessary to build systems that can support users in all their activities related to creativity. Other researchers propose the need for further research in information access and user interfaces so that digital libraries can support (1) more collaborative activities (Nichols et al., 2000), (2) better access to digital video collections (Lee et al., 2000), (3) geographic and spatial information (Oliveira, Goncalves and Medeiros, 1999), and (4) complex interactions (Kovács et al., 2000). Abdulla, Liu and Fox (1998) suggest that comprehensive logs of digital library use should be kept for analysing user behaviour to provide insights for future developments in the mechanisms of information access and user interfaces.

Summary

In this chapter we have considered the basic process of information access. Information access and information seeking have been the major areas of research in information science for many years and consequently many theories and models have been proposed by researchers. We have discussed some basic models that are also applicable to a digital library environment.

Since digital libraries are designed for end-users without the involvement of intermediaries; the interface is the only gateway and tool that the user can use to get access to and retrieve required information. User interfaces and computer–human interactions have been major areas of research in the study of both computer science and information science. In this chapter we have discussed the basic issues of user interface design with particular reference to the interfaces for digital information access.

Information access also depends on the information retrieval systems used. Discussions on information retrieval system models, techniques and tools are included in Chapter 9. Although digital libraries are at an early stage of

development, there have been some evaluation studies of them. Issues related to the evaluations, particularly on usability issues, appear in Chapter 13.

The development and proliferation of web browsers have largely standardized the design features of user interfaces, but digital libraries still vary significantly in terms of interface features. The examples given of user interfaces for supporting information browsing, searching and display in different digital libraries show some of the variations that exist. Advanced visualization techniques will greatly enhance access to, and display of, information in digital libraries.

In future digital library interfaces should integrate with a number of software packages and utility tools to increase the creativity of users.