Leveraging the Library Series

LEVERAGING **DATA DISCOVERY** IN THE LIBRARY

AND DEVELOPING PARTICIPATORY INTRANETS

Routledge
Taylor & Francis Group

TABLE OF CONTENTS

INTRODUCTION

Discovery tools are becoming increasingly popular in academic libraries, yet there is little information available concerning the tools used, and their implementation, maintenance, and evaluation. This FreeBook thus provides library practitioners and students of Library and Information Science (LIS) with guidance on how to best enable data discovery, while creating agile participatory intranets, in order to best serve patrons, by facilitating better internal communication, knowledge, and data.

This FreeBook features contributions from experts in their field, including:

John S. Spencer is Coordinator of Reference and Chair of Public Services at Foley Center Library, Gonzaga University, USA.

Christopher Millson-Martula is Director of the Library at Lynchburg College, USA.

Nina McHale is Assistant Systems Administrator at the Arapahoe Library District, USA.

William Miller, PHD, MLS, is Dean of University Libraries at Florida Atlantic University. He formerly served as Head of Reference at Michigan State University, and as Associate Dean of Libraries at Bowling Green State University, Ohio. Presently, he teaches courses in English Literature and Library Science.

Rita Pellen is the Associate Director of Libraries at Florida Atlantic University. Previously, she was Assistant Director of Public Services and Head of the Reference Department.

Note to readers: As you read through this FreeBook, you will notice that some excerpts reference other chapters in the book – please note that these are references to the original text and not the FreeBook. Footnotes and other references are not included. For a fully referenced version of each text, please see the published title.
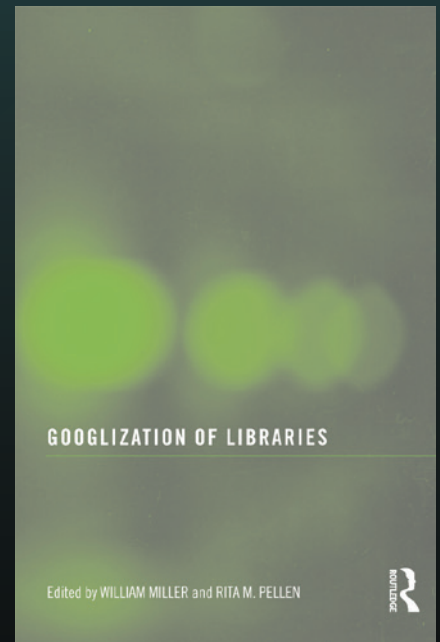
1

ACADEMIC LIBRARIES AND DISCOVERY TOOLS
A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

This chapter is excerpted from

Discovery Tools: The Next Generation of Library Research

Edited by John S. Spencer
and Christopher Millson-Martula

LEARN MORE ❯

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

Discovery tools are becoming increasingly popular in academic libraries, yet there is little information available concerning the tools used regarding their implementation, maintenance, and evaluation, or the tools' effects on public service librarians. Key questions considered through a literature review include the most common tools used, how their effectiveness is evaluated, and user satisfaction data. Readers will learn how public service librarians are involved in discovery tool processes; the usability of discovery tools; and how to incorporate discovery tools into information literacy endeavors.

INTRODUCTION

Many libraries want to identify "hidden collections" in order to make them available to users. In order to accomplish this goal, libraries primarily needed to enter these items in their catalog or create some sort of finding aid. Of course, this approach does lead to discussions about "findability" of items in the catalog in general, particularly as libraries have moved to include a wide array of digital and electronic content in their collections.

Substantial work on metadata had been done to make these collections and other items Web accessible, but it seemed very cumbersome when users were required to perform a catalog search as well as a Google search to locate library resources. Library catalogs tend to be designed using outdated ways of thinking, and they are well known to be difficult for users. Most librarians would agree that, out of all library resources, users have the most difficult time using the catalog. Catalogs are not searchable by Google, and, therefore, they leave a huge resource of information unavailable to users. Users know the library has things they need but they can't find the right pathway. Combining a "Google-like" search box with the wealth of a library's information resources may help us recover users who left our "walls" for the ease of Web searching. Hence, discovery tools!

S. R. Ranganathan's fourth law is "save the time of the reader," and libraries would do well to place this at the top of the priority list when designing products. Howard and Wiebrands (2011, 1) state that " ... it is the behavior of our information seekers that should drive our services." The library should be built into user workflows rather than trying to fit the user into the traditional library (Liu and Liao 2009). The focus should be on the users, and our systems should fit the information seeking behavior of our patrons (Howard and Wiebrands). Lougee (2009, 612) adds that the " ... library is no

ACADEMIC LIBRARIES AND DISCOVERY TOOLS
A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

longer the archive that deals primarily with the products of scholarship, rather there is a role to be played in all aspects, in all stages, of the processes of scholarship."

Allison (2010) emphasizes also that library tools need to follow (or lead) innovations on Internet search engines. She adds, "Librarians need to compromise between powerful search interfaces that are too complex   for students and featureless interfaces that are too basic to be useful when drill-down is needed for large results sets" (377). Gross and Sheridan (2011) note that libraries need to make themselves more a part of the search pro- cess again. Since digital media has become so accessible, fewer and fewer searches start at the library; thus, we need to maximize resource use by reducing user frustration. Howard and Wiebrands suggest that the focus on information literacy is due to the need to teach students how to use complicated resources that cause unnecessary frustration due to poor system design. Anderson (2006, 33) writes that " ... tutoring is a far less effective and efficient solution to the problem than fixing the search interface, which is exactly what Google has done."

In a 2004 discussion panel, Kenney notes that libraries need a Google or Google-like system rather than nonintuitive databases in order to provide the value of all the best information versus all that can be found in a shorter time. Library resources need to be made available through Google so that users can find it, especially since not all people have access to the powerful library databases. Users need to find items in order to use them. It is well known that users will not use a Web site if it is not user friendly. Now that options are available, this can be applied to library resources, especially the library catalog. Perrett (2010) reports that even adding additional copies of textbooks does not increase user satisfaction since patrons are still unable to find items in the catalog. Extra copies are only useful if students can find them. Providing a discovery tool that may make it easier for students to find items will have an impact on their overall satisfaction with the library.

McDonald and Thomas (2006) note that there are disconnects between current users' wants and needs and the traditional values of libraries. They classify the disconnects into three areas: technology, policy, and unexploited opportunities. Patrons become frustrated when having to use library systems that are limited by library culture rather than by technology; hence, they   seek systems that are easier to use and freely available. Michael Gersch, senior vice president and general manager of Serials Solutions, states, " ... we've all seen the statistics that show users want the quality of the library but prefer the simplicity of Google." (Luther and Kelly 2011, 4) Wilkes and

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

Gurney (2009) report that, based on two surveys of freshmen, more than two- thirds of respondents preferred using the Internet to find information. The majority of respondents used Google scholar as their preference for a known interface, but they were frustrated by the lack of full text and the inability to limit to peer review. Students using a database reported that easily limiting to full text and peer-reviewed sources was a significant benefit. There was a significant increase in comfort using Google Scholar over two semesters and only a very small increase in comfort using library databases. Discovery tools take a "Google" approach and include limiting features to help define the topic and get the best of the best (Kenney 2011). Way (2010) reports that users found databases and catalogs somewhat confusing, particularly the aggregated content and embargoed items, together with the problematic direct linking to many full text articles. In addition, there are many choices of databases, and each database has a variety of limits. Often the limits, as well as catalogs and databases in general, do not utilize terminology familiar to patrons.

Liu and Liao note that 89 percent of students use search engines first, with only 2 percent starting at a library site. However, patrons will use library resources if they are found through the Web. The results of the survey indicated that users want a single point of discovery for all resources, speed, an easy to use interface, relevant results (facet, natural language, topic clusters), and everything electronic (full text, multimedia download, e-Books) and seamless (and accurate). Users also desired widely available means to con- tact library staff, self-service, personalized services based on needs, librarians with technological skills to allow the user to exploit functions of resources, and staff with the " ... interpersonal skills to enable them to communicate effectively and sympathetically with users ... " (Liu and Liao 2009, 303). The authors added that "nowadays the number of visits to the library Websites is no longer as important as it was, the most important thing is that your users find your library and its services in their favorable environment" (304). The authors recommend focusing on promoting information reliability and special collections but that these must be easily available and findable. The authors also advocate for the promotion of librarians as one of the strengths of a library, whether through in-person or virtual means of contact.

LITERATURE REVIEW

A concept called the "next generation catalog" appeared when librarians came to the conclusion that the library catalog was no longer working for users (Wisniewski 2010). Yang and Wagner (2010) compare proprietary and open source tools according

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

to the checklist of twelve features desirable in the next generation catalog according to Breeding (2010). This is a very detailed report and presents several charts that were among the very first to compare the different types of discovery tools available.

McHale (2009) notes that metadata is vital to retrieval in federated searching and next gen catalogs. This also applies to discovery tools. If metadata is not present in sufficient quantity and quality, retrieval is jeopardized regardless of the system. Additionally, functionality is limited depending on the items included in the search system. Reasons for items not being included are vendor prohibition, limited number of seats, prices based on pay per view, and no current "translator" for the item.

FEDERATED SEARCH

The federated search engine represents an attempt to help users in their discovery process (Keene 2011). This technology allows users to select the databases that they found useful with a single search box query that is sent to each of the chosen databases separately. However, the average user is often unaware of the best databases to use as discovered through evaluation of federated search systems. Williams and Foster (2011) investigate how federated search engines were incorporated into Association of Research Library Member Websites in 2010. They conclude that more and more libraries were implementing federated search engines and that on-going research should explore librarian and user opinions. McHale (2009) reports that, after a usability test of their federated search system, graduate students and faculty would recommend it to undergraduates in their class, but not use the tool themselves. Several respondents did note that the availability of previously unknown resources was helpful. These results show the value of the one box search method despite a variety of concerns.

While federated search systems have issues, such as response time, they do serve as a precursor of discovery tools (Bhatnagar et al. 2010). Pradhan, Trivedi, and Arora (2011) provide a good review of the pros and cons of federated searching. Slow response time and time outs can be quite frustrating to users. The results can lead to information overload since limiting is not generally performed. It can be difficult to analyze results from all of the different databases. This article offers a very good explanation of all of the components that go into resource discovery, and it introduces the major players in the game. Stone (2010) has an interesting discussion on federated search engines and their lower than expected success. R. F. Woods (2010) shares her concerns with using federated searching in a law library. Her findings show that subject specific databases are vital to specialized areas of research and

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

practice. Thus, regardless of the technology, specialized tools need to be in the forefront of advanced research.

DISCOVERY TOOLS

After discovering that federated searching did not satisfy many of our users' needs, notably speed and thoroughness, we may have a better prospect in the discovery tools that have been developed from the next generation catalogs. Generally, these products offer a single point of access utilizing a centralized consolidated index that combines library catalogs, e-journals, databases, and Web-based resources, as well as digital archives. Since the content is pre-indexed, response time is significantly improved. The development of Google Scholar in 2005 was definitely another impetus for librarians to move from federated searching to something faster and more comprehensive.

Breeding (2010) discusses his idea of what discovery tools should be and the progress that libraries have made in developing an interface for today's library users. Vaughan (2011, 8) describes Web-scale discovery and states, "The primary goal of this work is to provide a valuable foundation to libraries that wish to know more about library-focused web scale discovery services and to aid libraries contemplating a marketplace review for their local environment." He does an excellent job describing concepts that librarians should know about, and he provides important information on all of the major players on the field. Luther and Kelly (2011) discuss what discovery tools are and the areas that need to be considered before purchase. These areas or factors include content, the search process, the fit, and the cost. Discovery tools are changing and expanding continuously. Some libraries have combined different services in various combinations, customizing the tools to meet their individual institutional needs (Vaughan 2011). Bhatnagar et al. (2010) note that more vendors are being more cooperative by allowing coordination among products; this allows for the full potential of Web-scale discovery services. It is possible that vendors have become aware that their products are no longer meeting the needs of libraries, and, thus, they see their inclusion in discovery tools as essential to their survival.

The most popular discovery tools are Summon by Serials Solutions, WorldCat Local by OCLC, EBSCOhost Discovery Service by EBSCO, Primo Central by Ex Libris, and Encore Synergy by Innovative Interfaces. Open source discovery tools, such as VuFind, Blacklight, and eXtensible Catalog allow institutions to tailor their tools for different users and collections (Parry 2009).

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

IMPACT OF DISCOVERY TOOLS

This section deals with the impact of discovery tools on usage of resources and any changes to departments other than public services. Stevenson et al. (2009) report that the availability of facets in Encore enabled the library to increase the number of formats employed in the cataloging process. Tosaka and Weng (2011) report that records with richer content tend to be found more frequently and lead to more circulation. Since discovery tools enhance "findability," there should be even more use with additional metadata. How- ever, a study by Stone (2010) notes that discovery tools highlighted major inconsistences in cataloging practices and the difficulties in providing consistent journal holdings and titles. Allison (2010) reports that the use of Google analytics with an Encore-based tool increased the use of databases and facets. She concludes that Encore showed the breadth of the collection to users, thereby increasing overall use.

Way (2010) indicates that Summon provided for less usage of traditional databases and a dramatic increase in article downloads through the discovery tool, indicating that more users find articles with Summon. How- ever, the author expresses concern about the reduction in use of specialized databases. It is generally assumed that users will move to the subject databases when needed; however, the authors of this article wonder how users will know when to move beyond the discovery tool search and use specialized databases to their full capacity. Way also notes that the large database aggregators showed reduced use, with a large increase in newspaper and book use. He cautions, though, that it is unknown whether students actually search the results for different formats or just look at the first hits regardless of format. Luther and Kelly (2011) find an overall increase in downloads of items while the use of databases decreased. They caution about findability and question whether all publisher and in-house records are equally findable.

Several Primo-based vendors stated that the new discovery systems seem to increase the number of search sessions, shorten search session times, introduce users to new content types and increase database usage overall (Harris 2011). Ballard and Blaine (2011), however, note that users spent more time in Encore than the classic catalog, which is thought to be predictive of a valuable search. The authors add that Google Analytics showed that users are more likely to use facets rather than limiting in a traditional catalog by about 15 percent. The authors share information from two other Websites that showed similar results. It is possible that students are more likely to use limits or facets since they are presented to users on the first screen rather than requiring another screen. Fagan et al. (2012) report that

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

the Quick Search (EDS) implementation reduced use of the Research Databases link by 55 percent in one year, although the authors did not find that this was correlated to an actual decrease in use during a spot check analysis.

USER ACCEPTANCE AND EASE OF USE

Generally, it is accepted that today's students are familiar with and will use technologies. However, library resources don't necessarily follow this trend, as proponents of virtual reference technologies know. Kim (2010) discusses the technology acceptance model (TAM) with a research sample from four universities in the United States. There is a positive correlation between perceived usefulness and perceived ease of use. Resources must be seen as useable and adding value. Interestingly, all three groups (masters and doctoral students, and faculty) showed a need to have guidance as they experience difficulties while searching. This last item indicates that discovery tools result in a need for instruction. Similarly, Pouratashi and Rezvanfar (2010) report that, as the relevance or need to use library resources increases, so does motivation. Additionally, as perceived ease of use increases so does the likelihood that patrons would use the tool. Jeong (2011) performed a study of student teachers and their acceptance using TAM. Interface characteristics influenced perceived usefulness indirectly through perceived ease of use. System characteristics (relevance and system quality) had a direct influence on the perceived usefulness. System quality positively affected perceived usefulness and perceived ease of use. Perceived ease of use is the overall dominant variable in use of the system. Thus, tools need to be easy to use for students to make effective use of them.

Shachak and Fine (2008) also note that hands-on training aided informed decision-making as to usefulness of the tools and how best to use them, as well as the value to their work during a study of a group of researchers. Results indicated that both faculty and students needed to find the tools easy to use and valuable for their research and education. Hands-on training was recommended by the participants to help them develop a realistic value of the tool (know its features and functionality) and improve ease of use.

USABILITY OF DISCOVERY TOOLS

The preceding section provides evidence that the easier and more relevant users find a tool to be, the more they will use it. Thus, libraries need to find useable products and provide training on those products. This section will review some recent usability studies of discovery tools and provide some conclusions about user preferences. This is not a completely exhaustive list, nor is it meant to evaluate the tools.

ACADEMIC LIBRARIES AND DISCOVERY TOOLS
A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

Joo and Lee (2011) performed a usability measurement of digital libraries. Their results are applicable to discovery tools since these represent the portal to digital content. Usability tests measure both system characteristics (speed and interface design) and user-centered qualities (needs and satisfaction). The authors developed a survey methodology that gathers large amounts of data easily from users rather than having to use the traditional formal usability tests with small numbers of participants. In her literature review, Allison (2010) discusses researcher expectations and skills along with the need for libraries to develop discovery interfaces as a result of the increasing lack of personal contact due to the accessibility of research materials on the Web. She discusses how users have been unhappy that they are unable to place limits post search with the current catalogs during a review of several discovery tools.

Although Olson (2007) used an early version of a faceted catalog with an Endeca layer, nine of the twelve doctoral candidates in the study discovered items they had not found previously despite having done "complete" searches of the classic catalog. All participants noted that the facets provided easy refinement of the searches. Majors (2011), using five current discovery tools across a number of participants, reports favorable responses on the ability to find a variety of resources in one search, ease of narrowing, saving items, and e-mailing items. Participants did recommend a spell checker and locating the refine options near the search box along with an explanation   of their purpose. Facets were appreciated, although multiple options within each facet were proposed and participants wanted the results to update automatically. Tam and Bussey (2009) report that international students favored relevance ranking, tag clouds, facets, and borrowing suggestions.

Bhatnagar et al. (2010) employed usability tests with three systems. Using personas, the authors developed tasks and features needed in a discovery tool. After creating a review matrix to compare all three tools, the authors utilized an online user survey to prioritize the items discovered through the persona process. Several face-to-face interviews were conducted as well. Top priorities for users included the following: "I can find articles from the top journals in my subject area"; "I can get to the full text of an article in one mouse click"; "I can perform advanced searches"; "I can easily narrow down my search results … "; and "I can limit my search so I only get articles from scholarly (peer-reviewed) journals." The authors summarized the priorities as being able to find useful content in the shortest amount of time. Participants also prioritized features such as e-mailing, printing, citation analysis, and similar features.

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

There are numerous usability studies for Summon. Howard and Wiebrands (2011) report on usability studies involving undergraduate students. Students reported difficulty finding reserve items due, in part, to their being classed as Web resources rather than course readings. Addition- ally, patrons seemed even more disappointed than usual that not everything is available online. The authors note that this has always been an issue, but   it is even worse now that one-stop shopping is expected. It was reported that students would reject print books even if they are available in the same library as the user. However, a number of participants commented that they may use the library more now that searching is easier. Although a slightly older study, DeFelice et al. (2009) also report that students, especially undergraduates, were more frustrated by the lack of full text than with the traditional catalog and database systems. Gross and Sheridan (2011) recount a task-based usability study with five students using Summon. Participants experienced difficulty with terminology. The issue of greatest concern was the use of "Web resource" rather than "reserve reading," which is partially   a terminology problem but also related to a lack of understanding about   the search results. The authors also noted that students used simplistic key- words rather than taking time to analyze the question. All participants were first-year students, which may have affected the results.

Philip (2010) reports that students found the single search option in Summon helpful as a central search point. However, they still desired instruction or enhanced documentation on the advanced features. There was high use of the article preview option and keyword searching. Another study from North Carolina State University (2010) revealed that participants liked the single search option but that there were problems with finding known items along with some concerns with facet searching and the relevance ranking of results. A study by Sharman and Hiller (2011) from the University of Huddersfield notes that students preferred Summon over the current system due to the quick and efficient retrieval of items. The participants also appreciated that the specialized databases are prominently available from the Summon results page. Ford (2010) reports on a usability study of Sum- mon at the University of Manitoba. As with several studies in this section, students have a higher expectation to search and access books and articles at the same time and immediately. The lack of full text was a major concern for participants. Users did not notice limits until prompted; however, participants did comment on the ease and simplicity of Summon in terms of retrieving useful items. Stone (2010) provides a comprehensive report on the selection and implementation process of the Summon tool at the University of Huddersfield. He discusses some of the issues that were

ACADEMIC LIBRARIES AND DISCOVERY TOOLS
A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

found, such as major inconsistencies in cataloging practices and difficulties locating specific journal titles, together with how these issues were addressed.

DeFelice et al. (2009) report that undergraduate students found Summon easy to work with and that it returned good results. As mentioned above, there was significant frustration that the full-text icon did not always lead to the full text. There was some use of facets. Graduate students had less favorable responses for Summon compared to their traditional subject- specific databases. Faculty and graduate students both desired a cited reference feature. Faculty members noted that they would recommend the system to their students. While Summon was sufficient in many ways for their own searches, faculty wanted improved relevancy. Summon did not perform well for single author searches or for limited discipline searches. Faculty also missed the search history option. The report team had concerns that Sum- mon does not search all items and thus emphasized that libraries cannot do away with specific databases or resources that have key features for particular disciplines. They added that the lack of full-text items is exacerbated with Summon (the Google effect), and they had concerns about not knowing precisely what is included in Summon.

Williams and Foster (2011) researched usability studies before performing an extensive usability study with EBSCO Discovery Service (EDS), where they found that EDS still had some major issues. Overall, the review results indicated that users do not want advanced search techniques, although graduate students and faculty desire easily available limiting options. The authors used usability task-based testing. Similar to Denton and Coysh (2011), students did not progress past the first page of results (the default was thirty items), and they tended to select only results with images. Students frequently selected limits prior to beginning a search, which may be due to familiarity   with searching EBSCO products. Post-search limits were also used. Books, academic/peer reviewed journals, and periodicals were generally clear to students, while other formats were less clear. There was less use of sharing options than expected, as participants tended to use only the e-mail function. The use of "additional results," which links to results from the EBSCOhost Integrated Search but are not included in EDS, was poorly used. Participants were generally unaware of the purpose of the function. Format icons were   not employed frequently, but rather participants used the limiting options that were available. Users did not explore non-opened facets unless struggling with a task. Special features were not well used. The authors suggest making these more prominent. Post-test questionnaire results indicated that participants wanted more instruction and suggested pop-up help

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

options or a one-page reference sheet for advanced features. Participants noted that these might help their search, but they wanted to know more about search options before searching. Powers (Thornton-Verma 2011) report that EDS worked well for undergraduates needing a few articles and for graduate students and faculty needing to find obscure information. Otherwise, there were some issues. Ultimately, they found that freshmen, PhD students, and faculty were happiest with the product. Upper classmen and master's students found that the specialized databases were more helpful for their purposes.

Fagan et al. (2012) utilized task-based testing with a mix of ten fresh-men, sophomores, juniors, one senior, one graduate student, and two faculty members for a usability study of EDS. There was overall confusion about what is being searched. Most used limits were "just books," "just articles," "title," and "scholarly articles," and there was high use of these items, similar to results found by Williams and Foster (2011). Participants had significant trouble with changing the sort order. User comments included easy and speedy, with disadvantages focusing around the overwhelming number of results. The authors recommend increased concentration on highlighting special collections and databases not included in the discovery tool.

Despite popularity in academic organizations, Primo had few published usability studies. A study by Sadeh (2008) involving eight participants reported some terminology issues, but overall the Primo interface was easy to understand. Students expressed value in the faceted system, appreciated on-   line versus print sorting options, and found usefulness in the tagging options. Fernandez (Thornton-Verma 2011) reports that students loved the ability to use their mobile devices with Primo.

Several studies looked at WorldCat Local. Becher and Schmidt (2011) discuss a study in which students explored preferences using Aquabrowser and WorldCat Local. Undergraduates preferred WorldCat Local due to the simplicity of the interface. Graduate students favored Aquabrowser due to the more professional look. All students liked the links to full text. Sixty percent of respondents across both categories found more resources than with the traditional systems. No participants preferred an older classic catalog. Facets were appreciated by all participants, while, interestingly, those students who noticed the Web 2.0 tools did not view them as being academically relevant. Fahey, Gordon, and Rose (2011) list confusion in interpreting holding statements of books and other items, as well as locations and call numbers in a usability study with WorldCat Local. Participants still preferred WorldCat Local over the catalog; however, the authors note that this may have been due to the inclusion of book covers, which had significant appeal to the participants.

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

VuFind, seemingly the most popular open-source option, has received a variety of usability studies. Emmanuel (2011) reports on Yales's VuFind usability testing and qualitative task-based testing. Participants preferred key- word search to be the default search type. They also enjoyed the ability to search broadly and then narrow, although there was a low understanding of facets. As found in other studies of discovery tools, users experienced confusion about format types and frustration with the continued use of library-ese. Out of a group of fifteen participants, only one preferred the current catalog due to its functionality. Participants liked the "cite this" feature but wanted a link to RefWorks. An earlier study by Bauer (2008a) involving medical and nursing students found that these students preferred to search rather than browse topic lists. They also found that narrow topics were difficult to search due to the large number of results retrieved and also because of the inability to limit to medical and nursing items. A second study (Bauer 2008b) with undergraduate students showed that participants wanted a browse feature similar to browsing a bookshelf, more help with spell checking, and ways to limit to the more relevant results. Additionally, students had no interest in the tagging option.

Denton and Coysh (2011) performed an extensive study of multiple groups of users of VuFind. All users preferred a unified interface to the traditional catalog and electronic resources rather than separate search options. Terminology was a major issue. As an example, Library of Congress Subject Headings were initially called "subject keyword," which participants assumed was a normal keyword search. As a result of the study, this search was changed to the more usual "subject heading" search. Respondents to the survey noticed inconsistencies in the catalog, such as the language field sometimes saying English and other times being left blank to mean English. Journal holding information was poorly recognized. Respondents to both the survey and task-based testing appreciated the wealth of results although they looked only at first page. Unlike some studies, facets seemed to be   well understood by graduate students and faculty. In summary, users were able to find items related to their searches faster, although there were some issues.

DISCOVERY TOOLS AND INFORMATION LITERACY

The Technology Acceptance Model proposes that technology users have a requirement for some training, whether mediated or unmediated (Shachak and Fine 2008). Results of several usability studies as presented above also indicated that students desire some instruction in order to make use of the functionality of discovery tools.

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

Fagan (2011) provides an excellent editorial examining how discovery tools assist with the information literacy of users based on Standards 1 and 2 of the Information Literacy Competency Standards for Higher Education (Association of College and Research Libraries 2000). As expected, discovery tools support some competencies while not supporting others. Librarians need to determine how to overcome any deficits and promote the use of features that support information literacy.

While not discussing discovery tools specifically, Goh (2011) provides an interesting study on SMS (Short Message Service) adoption differences for females versus males, and finds them similar to other technology adoption articles. Gender differences do relate to instruction, and, as instruction librarians know, instruction content and techniques need to take attendee differences into account. Goh recommends promoting the usefulness of the system for both and concentrates on training for females to increase their confidence with using new technologies. Given the historical evidence provided by a multitude of articles, this result can also be applied to training for discovery tools.

Ansari and Zuberi (2010) studied researchers at the University of Karachi and reported that the participants cited lack of knowledge for not employing library information resources. Similarly, master's and doctoral graduate students and also faculty showed a need to have guidance when searching (Kim 2010).

Due to the lack of studies on librarians and discovery tools, we included Lampert and Dabbour (2007), who look at librarian attitudes toward federated and meta search systems. A survey of librarians attending the 2005 LITA Forum (thirty-three respondents) noted that 61 percent of the respondents did not teach federated searching. Several reasons existed for this: not all content was included, loss of controlled vocabularies and specialized features found in specialized databases, too time-consuming for a one-shot session, and a dislike of promoting the use of a Google-like system. Twenty- one percent of respondents who taught federated searching appreciated the ability to save citations and search strategies, as these were seen to help with information management. Most participants outlined the value of the native databases and the federated search system to students. The most frequent way of teaching was to introduce the federated search as a starting point to identify more useful items, including which specialized databases to search further. Similar responses were received for those who teach federated searching during reference transactions. Librarian confidence in using and teaching federated search was mixed. Some respondents expressed concerns about confidence in search

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

results due to a lack of knowledge about how the results were generated. Other respondents found it easy to search and teach. Perceived value is most often seen when searching in multidisciplinary areas or when identifying useful databases.

The authors performed a similar survey during a 2005 mini-conference of the California Academic and Research Libraries South. Out of ten tables of participants, three tables agreed that federated search should be taught as a starting point since it shows the range of information available and permits focus on search strategy rather than the specifics of each database. Three tables gave a negative response due to lack of limiting and too many varied results, loss of subject headings and thesauri, and concern that not all databases are included. Four tables responded that it would depend on student needs and experiences and generally only if obscure topics were being searched or as an option to determine possible native sources. Lam- pert and Babbour (2007) note that librarians preferred to instruct in native products— familiarity and ease of use of specific techniques. They also point out the need to show what is freely available since users will not always have library resources.

Buck and Mellinger (2011) also performed a survey of librarians using a popular librarian listserv. Although the response rate was low, the results were helpful. Ten percent of respondents demonstrated Summon during instruction sessions, reference desk transactions (14 percent), or both

(62 percent). Forty-two percent included Summon in instruction sessions regardless of discipline or course level. Others considered including Sum- mon based on population or content of session. Summon was taught during reference transactions as a good starting point in order to emphasize the availability of databases, or when only a few articles were needed. Most instruction librarians taught Summon, the catalog, and databases, and 29 percent considered Summon to be just another tool to teach. Four percent taught Summon only, while another 4 percent taught only Summon and the catalog. Both groups taught Summon to lower division undergraduates only. Reference and instruction librarians were asked how discovery tools would affect information literacy. Lampert and Dabbour (2007) report a negative or neutral effect from their survey. Generally, this reaction is due to students not being able to recognize the value of databases, or not knowing what they are searching. Ten percent of positive responses expressed the belief that students should be given a choice and that federated search gives students a choice in that it shows the variety of resources available. Buck and Mellinger (2011) find similar results with Summon librarians. Twenty- three percent felt that Summon had

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

negative effects on information literacy due to the broad results that are returned, with 38 percent of the responses being neutral or mixed due to users being able to access a greater variety of information and its ease of use. Yet, there were some concerns with students knowing how to search effectively and taking the time to limit appropriately and identify formats. Eight percent of librarians believed that Summon has a positive effect on information literacy since it allowed students to focus on selecting and evaluating results rather than choosing the best resource.

In terms of teaching discovery tools to groups of users, Lampert and Babbour (2007) report that 30 percent of the respondents limited instruction on federated search products to upper-level undergraduates, graduate students, and faculty. Respondents felt that the search system was too specialized or complicated for lower division undergraduates. Buck and Mellinger (2011) report that 12 percent of respondents targeted undergraduates for using Summon during reference transactions. Of the respondents who included Summon in instruction sessions, the majority taught Summon to undergraduates, with only 5 percent teaching it to graduate students. Several respondents clarified that they include Summon instruction only when students need to search a variety of topics, such as for writing or debate classes. Librarians from the survey felt that students had difficulty limiting and using the advanced features. Most undergraduates were satisfied with Summon's performance, while graduate students seemed less satisfied.

Discovery resources also were identified as having some effect on instruction. There was a variety of responses to Summon from across the spectrum in the study by Buck and Mellinger (2011). Fifty-eight percent thought that Summon had changed instruction by reducing time spent on selecting databases and teaching the catalog, leaving more time to refine the search and develop higher level search skills. Forty percent spent the extra time on narrowing and broadening searches, explaining scholarly and general articles, evaluation of results, citing sources, reading citations, and other more "high level" literacy skills. Thornton-Verma (2011) also reports that librarians found bibliographic instruction to be easier to provide because they can spend the time on the finer points of searching rather than just the basics. Librarians responded that they save time demonstrating various article databases and online catalogs; since there is just one product, the extra time can be used to emphasize higher-level search skills (Sharman and Hiller 2011). Luther and Kelly (2011, par. 4 under "User Opinions") note that "librarians have observed how they have gone from explaining the mechanics of search to focusing on evaluating search results." A few librarians expressed feelings of being forced to teach Summon due to its prominence

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

on the home page, and that it added to instruction instead of replacing current content (Buck and Mellinger 2011).

Librarians had opinions about the value of discovery tools. Buck and Mellinger (2011) note that librarians appreciated Summon for ease of use, one-stop-shopping, and enabling students to always find something of value. Content or scope was good for some yet also a reason not to include in sessions due to the lack of inclusion of desired databases, overly large result sets, and lack of relevant results. Librarians considered ease of searching, comprehensive results, greater use of resources, and facets to be advantages for users (Howards and Wiebrands 2011).

Difficulty in understanding as a result of too many options was a major flaw (Buck and Mellinger 2011). Howard and Wiebrands (2011) report concerns over the change in how patrons access resources as it is very different from traditional systems. Additionally, librarians thought that Summon reduced user ability to know when a specialized database is necessary, especially for those that are not included in the Summon index. Other major disadvantages were not knowing what was indexed and how results were generated (Buck and Mellinger; Howard and Wiebrands). This leads librarians to distrust the results (Howard and Wiebrands). Librarians are well-versed in the workings of databases and catalogs. Discovery tools are less easy to understand, and, therefore, librarians have a lack of trust in the results. Luther and Kelly (2011) report that librarian concerns included loss of original con- tent from databases, the potential demise of the catalog, reduced analysis of use in that discovery tools may not track the same statistics currently collected, and knowing when content is added or removed or features updated. Howard and Wiebrands (2011) find that a pre- and postquestionnaire after Summon implementation resulted in a 25 percent decrease in full text searching. There were some concerns about collections still being hidden, with the frequently used resources having greater opportunity to be dis- played than the specialized materials. Buck and Mellinger (2011) note similar results, as the variety of items and depth of items searched was a positive, yet it could also be a negative, especially for disciplines with general or common terminology. Not knowing how enhancements were implemented was another stressor.

Howard and Wiebrands (2011) add that librarians were moved beyond their comfort zones, and, therefore, they tried to "break" Summon by finding every mistake, missing citations, indexing changes, etc. Librarians had to make a mental shift to realize that discovery tools are for the general user, not primarily for librarians or library professionals who may require a different way of using the system. Reference

ACADEMIC LIBRARIES AND DISCOVERY TOOLS
A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

and instruction librarians learned to move away from complex searches and use a broad search initially and then limit the results with facets. There is still a push/pull between the "correct" way to find information and the fact that users want quick and simple results. Users do not usually expect to find all the results or necessarily find the best results, concepts which do not fit easily with traditional librarian thinking.

Using a pre- and postquestionnaire for Summon implementation, Howard and Wiebrands (2011) also note that librarian perceptions changed over time, and that there was an increase in the awareness of Summon functionality. The authors recommend that libraries should create positive first impressions of a new system and allow time for trust to be established. As a result, librarians are now more involved on integration issues, especially in how favored databases and the library catalog work within Summon. The authors also suggest that librarians be included in the process earlier in order to avoid sending negative impressions to users and to ensure best use of the tool.

SUMMARY

Discovery tools offer mixed blessings to librarians and library users. Valuable characteristics are ease of use, one stop shopping, limiting by facets, and citation information. Concerns center around too many results, lack of relevancy, not all content included, loss of catalog specificity, loss of database functionality, and user knowledge. Additional issues may be that users look only at the first page of results and user willingness to move to the specialized database when needed.

Usability studies indicate that typical library things are still problematic for users. Call numbers, multiple locations, item formats, and holding statements, among others, are not well understood; they are possibly more difficult to find in discovery tools. Search and limiting options should appear on the main page as users tend not to explore the tools. Several usability studies identified issues with which facets should be available and "open" to show users the options. Usability studies are still a test situation, and observing users in action will provide further information.

In general, librarians as a group are not comfortable with change, especially change where they have little involvement or input. Including public service librarians in early and ongoing decisions will help to build trust and confidence in the system as shown from the literature review. For optimal use of discovery tools, public service librarians need to be more involved in the selection, implementation, maintenance, and evaluation of the tools. Librarians need to be comfortable with a resource in order to teach and advocate for it.

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

Lampert and Babbour (2007) recommend including discovery tools in instruction. Based on the literature, librarians who include the tool in instruction consider it to be just another option for users. Given the wealth of options available, it is vital to explain the value and potential use of each item. Librarians should focus on when it is appropriate to move to a specialized database or to the catalog.

An unexpected result from the literature review was the request for instruction and/or documentation on advanced features. However, based on best practices, instruction needs to focus on the user and what she/he needs. In other words, market the value of the resource and how it benefits the user. Jubb, Look, and Sparks (2007) emphasize that researchers rely heavily on recommendations from colleagues. Thus, if we want to market a product, we need to gain "advocates" to help spread the word. Green and Macauley (2007) emphasize that librarians need to understand student needs and engagement in order to understand how to develop information literacy, programs, and services that captivate the users.

Libraries and vendors need to use the system-related terminology (Google and Facebook, for example) to which our users relate. Tools should also be designed to fit into the user's way of thinking and operating, rather than trying to "force" them into the traditional library system. Best practice—involve librarians early. Let them develop comfort with a discovery system prior to rollout so that they will be able to share its positive features with users. If we can attain these goals, we will have advanced considerably in line with the major advancement that discovery tools represent.

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

REFERENCES

Allison, D. A. 2010. Information portals: The next generation catalog. Journal of Web Librarianship 4:375–89.

Anderson, R. 2006. The uncertain future of libraries in a google world—sounding an alarm. Internet Reference Services Quarterly 10:29–36.

Ansari, M. N., and B. A. Zuberi. 2010. Use of electronic resources among academics at the University of Karachi. Library Philosophy and Practice June: 1–7 http://www. webpages.uidaho.edu/mbolin/ansari-zuberi.htm.

Association of College and Research Libraries. 2000. Information Literacy Competency Standards for Higher Education. http://www.ala.org/acrl/sites/ala.org. acrl/files/content/standards/standards.pdf.

Ballard, T., and A. Blaine. 2011. User search-limiting behavior in online catalogs: Comparing classic catalog use to search behavior in next-generation catalogs. New Library World 112:261–73.

Bauer, K. 2008a. Usability test of VuFind as a subject-based display of e-Books. https://collaborate.library.yale.edu/usability/reports/Forms/public.aspx.

———. 2008b. Yale University Library VuFind Test—Undergraduates. https:// collaborate.library.yale.edu/usability/reports/Forms/public.aspx.

Becher, M., and K. Schmidt. 2011. Taking discovery systems for a test drive. Journal of Web Librarianship 5:199–219.

Bhatnagar, G., S. Dennis, G. Duque, S. Henry, M. MacEachern, S. Teasley, K. Varnum. 2010. University of Michigan Article Discovery Working Group Final Report. http:// www.lib.umich.edu/files/adwg/final-report.pdf.

Breeding, M. 2010. The state of the art in library discovery 2010. Computers in Libraries 30:31–34.

Brown-Sica, M., J. Beall, and N. McHale. 2010. Next-generation library catalogs and the problem of slow response time. Information Technology and Libraries 28:214–23. http://www.ala.org/lita/ital/sites/ala.org.lita.ital/files/ content/29/4/ brownsica.pdf.

Buck, S., and M. Mellinger. 2011. The impact of serial solutions' summon on information literacy instruction: Librarian perceptions. Reference Services Quarterly 16:159–81.

DeFelice, B. et al. 2009. An Evaluation of Serials Solution Summon as a Discovery Service for the Dartmouth College Library. Report submitted to the Dean of Libraries and Librarian of the College, November 10, 2009. http://www.dartmouth.edu/library/ admin/docs/Summon_Report.pdf.

Denton, W., and S. J. Coysh. 2011. Usability testing of VuFind at an academic library. Library Hi Tech 29:301–19.

Emde, J. Z., S. E. Morris, and M. Claassen-Wilson. 2009. Testing an academic library website for usability with faculty and graduate students. Evidence Based Library and Information Practice 4:24–36.

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

Emmanuel, J. 2011. Usability of the VuFind next-generation online catalog. Information Technology and Libraries 30(1). http://www.ala.org/lita/ital/sites/ala. org. lita.ital/files/content/30/1/pdf/mar11.pdf.

Fagan, J. 2011. Editorial: Discovery tools and information literacy. Journal of Web Librarianship 5:171–78.

Fagan, J. C., M. Mandernach, C. S. Nelson, J. R. Paulo, and G. Saunders. 2012. Usability test result for a discovery tool in an academic library. Information Technology and Libraries 31:83–112.

Fahey, S., S. Gordon, and C. Rose. 2011. Seeing double at memorial university: Two WorldCat local usability studies. Partnership: The Canadian Journal of Library and Information Practice and Research 6(2). http://journal.lib. uoguelph.ca/index.php/ perj/article/view/1552/2282.

Ford, L. 2010. Better than Google Scholar? Presented at Internet Librarian 2010 Monterey, California. October 25, 2010. http://conferences.infotoday. com/ documents/111/a105_ford.pdf.

Goh, T. 2011. Exploring gender differences in SMS-based mobile library search system adoption. Educational Technology and Society 14:192–206.

Green, R., and P. Macauley. 2007. Doctoral students' engagement with information: An American-Australian perspective. portal: Libraries and the Academy 7:317–32.

Gross, J., and L. Sheridan. 2011. Web scale discovery: The user experience. New Library World 112:236–47.

Harris, S. 2011. Discovery services sift through expert resources. Research Information, April/May, 18–21.

Howard, D., and C. Wiebrands. 2011. Culture shock: Librarians' response to web scale search. Edith Cowan University Research Online. http://ro.ecu.edu. au/cgi/ viewcontent.cgi?article 7208&context ecuworks.

Jeong, H. 2011. An investigation of user perceptions and behavioral intentions to- wards the e-Library. Library Collections, Acquisitions, and Technical Services 35:45–60.

Joo, S., and J. Y. Lee. 2011. Measuring the usability of academic digital libraries: Instrument development and validation. The Electronic Library 29: 523–37.

Jubb, M., H. Look, and S. Sparks. 2007. Researchers' use and perceptions of discovery services. Learned Publishing 20:147–53.

Keene, C. 2011. Discovery services: Next generation of searching scholarly information. Serials 24:193–95.

Kenney, B. 2004. Googlizers vs. Resistors. Library Journal, December 15, 44–46. http://www.libraryjournal.com/article/CA485756.html.

———. 2011. Liverpool's discovery. Library Journal, February 15, 24–27.

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

Kim, Y. 2010. The adoption of university library web site resources: A multigroup analysis. Journal of the American Society for Information Science and Technology 61:978–1003.

Lampert, L. D., and K. S. Dabbour. 2007. Librarian perspectives on teaching metasearch and federated search technologies. Internet Reference Services Quarterly 12:253–78.

Liu, S., and S. Liao. 2009. Surviving in the digital age by utilizing libraries' distinct advantages. Electronic Library 27:298–307.

Lougee, W. 2009. The diffuse library revisited: Aligning the library as strategic asset. Library Hi Tech 27:610–23.

Luther, J., and M. Kelly. 2011. The next generation of discovery. Library Journal, March 15. http://www.libraryjournal.com/lj/home/889250-264/ the_next_generation_of_discovery.html.csp.

Majors, R. 2011. Usability of next-gen interfaces and discovery tools. Presentation at the Innovative Users Group, San Francisco, April 2011.

McDonald, R. H., and C. Thomas. 2006. Disconnects between library cultures and millennial generation values. Educause Quarterly 29:4–6.

McHale, N. 2009. Why reference and instruction librarians hate federated searching and nextgen catalogs. MileHighLibrarian (blog). http://milehighbrarian. net/node/1.

North Carolina State University. 2010. Summon usability testing. http://www. lib.ncsu. edu/userstudies/studies/2010_summon/.

Olson, T. A. 2007. Utility of a faceted catalog for scholarly research. Library Hi Tech 25:550–61.

Parry, M. 2009. After losing users in catalogs, libraries find better search software. Chronicle of Higher Education, September, 288.

Perrett, V. 2010. Information literacy skills training: A factor in student satisfaction with access to high demand material. The Australian Library Journal 59:23–29. Philip, M. 2010. Do Students Want a One-Stop-Shop to Help Them Navigate Their Way Around the Maze of Library Resources? A Usability Study Looking at the Best Version of Summon, the New Library Search Engine at the University of Huddersfield. Master's thesis, University of Sheffield. http://eprints.hud.ac.uk/9824/1/Final_ version_dissertation.pdf.

Pouratashi, M., and A. Rezvanfar. 2010. Analysis of factors influencing application of ICT by agriculture graduate students. Journal of the American Society for Information Science and Technology 61:81–87.

Pradhan, D., K. Trivedi, and J. Arora. 2011. Searching Online Resources in New Discovery Environment: A State-of-the-Art Review. Paper presented at the 8th International CALIBER-2011, Goa University, Goa, March 2–4. http:// ir.inflibnet.ac.in/ dxml/bitstream/handle/1944/1623/14.pdf%3Fsequence%3D1.

Rowe, R. 2010. Web-scale discovery: A review of summon, EBSCO discovery service, and world cat local. Charleston Advisor 12:5–10.

ACADEMIC LIBRARIES AND DISCOVERY TOOLS

A SURVEY OF THE LITERATURE

Beth Thomsett-Scott and Patricia E. Reese

Excerpted from *Discovery Tools: The Next Generation of Library Research*

Sadeh, T. 2008. User experience in the library: A case study. New Library World 109:7–24.

Shachak, A., and S. Fine. 2008. The effect of training on biologists' acceptance of bioinformatics tools: A field experiment. Journal of the American Society for Information Science and Technology 59:719–30.

Sharman, A., and E. Hiller. 2011. Implementation of SUMMON at the University of Huddersfield. SCONUL Focus 51:50–52.

Stevenson, K., S. Elsegood, D. Seaman, and C. Pawluk. 2009. Next-generation catalogues: Reviews of encore, primo, summon and summa. Serials 22:68–77.

Stone, G. 2010. Searching life, the universe and everything? The implementation of summon at the University of Huddersfield. Liber Quarterly 20:25–52.

Tam, C., and Bussey. 2009. Student user preferences for features of next-generation OPACs: A case study of University of Sheffield international students. Program: Electronic Library and Information Systems 43:349–74.

Thornton-Verma, H. 2011. Discovering what works: Librarians compare discovery interface experiences. Reviews by Library Journal (blog), December 7. http://reviews.libraryjournal.com/2011/12/reference/discovering-what- works-librarians-compare-discovery-interface-experiences/.

Tosaka, Y., and C. Weng. 2011. Re-examining content-enriched access: Its effect on usage and discovery. College and Research Libraries 72:412–27.

Vaughan, J. 2011. Web scale discovery. Library Technology Reports, January, 47:15–59.

Way, D. 2010. The impact of web-scale discovery on the use of a library collection. Serials Review 36:214–20.

Wilkes, J., and L. J. Gurney. 2009. Perceptions and applications of information literacy by first year applied science students. Australian Academic and Research Libraries 40:159–71.

Williams, S. C., and A. K. Foster. 2011. Promise fulfilled? An EBSCO discovery service usability study. Journal of Web Librarianship 5:179–98.

Wisniewski, J. 2010. Web scale discovery: The future's so bright, I gotta wear shades. Online, July/August, 55–57.

Wood, M. S. 2011. Discovery tools and local metadata requirements in academic libraries. SLIS Student Research Journal 1:1–15.

Woods, R. F. 2010. From federated search to the universal search solution. The Serials Librarian 58:141–48.

Yang, S. Q., and K. Wagner. 2010. Evaluating and comparing discovery tools: How close are we towards next generation catalog? Library Hi Tech 28:690–709.

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY

A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

This chapter is excerpted from

Discovery Tools: The Next Generation of Library Research

Edited by John S. Spencer
and Christopher Millson-Martula

LEARN MORE ›

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY

A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

This study observes user behavior within the Web-scale discovery tool, Summon by Serials Solutions, at Montana State University. The most common search query within Summon at the time of this study was facebook.com. Two hypotheses are explored: a majority of search queries performed within Summon are of low quality; and search query quality improved during the first two semesters of implementation. Transaction log analysis was performed on a random sample of search queries. Search query quality was consistently high over the first two semesters of implementation, but it did not increase thereafter. Further analysis revealed that users neglected to use Boolean operators.

INTRODUCTION

The widespread use of Internet search engines has made searching library-provided resources seem difficult and cumbersome in comparison, even among university students (Griffiths and Brophy 2005). Google is often the first-place university students visit for scholarly information, and the resources they acquire via Internet search engines often only "satisfice" their information needs. "Satisfice" refers to being satisfied with resources that only suffice for an information need (Brophy 2004, 145). Using a single search box that yields relevance-ranked results has created expectations when searching for library-provided resources as well (Griffiths and Brophy). This sort of behavior embodies the Principle of Least Effort, in which humans habitually choose the most convenient path with the least amount of work in order to accomplish a task (Zipf 1949, 5 6). It is believed that libraries need to compete with ease of use and create their own simple interfaces that allow their expensive and high-quality content to be more discoverable (University College London CIBER Group 2008).

Since approximately 2000, libraries have made greater efforts toward ease of searching library materials, especially since the increased use of electronic resources. Federated searching, first put into practice in 2003, has come close to realizing the single search model. Federated searching allows for a cross-database search utilizing one search box. The databases are usually grouped by subject area (Caswell and Wynstra 2010). Federated searching provides for an improvement in users' search experiences; how- ever, the results include article content only (Wisniewski 2010). The time it takes for federated search tools to search multiple databases in real-time and de-duplicate the search results can seem too long for users accustomed to search engines as well (Wrubel and Schmidt 2007).

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY
A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

Web-scale discovery tools, a relatively new technology first implemented in 2007, bring libraries a step closer to making their collections easier for users to search. Vaughan (2011) states that bibliographic information from library resources is pre-harvested and centrally indexed, streamlining the search process for users, and, thereby, ultimately providing improvements in content, discovery, deliverability, and flexibility. Through central indexing to the article level and brokered agreements with publishers, resources are available in one location via a single search box, similar to a search engine. Relevance-ranked results are displayed in an intuitive interface, creating a space where results and content are deliverable to the user. Flexibility in the system is achieved through the interoperability of the interface with content hosted in multiple areas, as well as customizable features for user preferences. For the Montana State University (MSU) Library, Web-scale discovery was an appealing, low maintenance way to improve the usability of its resources.

After the processes of collaboration, trouble-shooting, and testing be- hind the scenes, MSU Library provided Summon by Serials Solutions for its users in August of 2010. The public interface was rebranded as CatSearch, named after the MSU Bobcats. CatSearch was prominently displayed on the library Website, and it has been heavily utilized since it went live, with over 100,000 search queries in the first two semesters of implementation. Ac- cording to Google Analytics, there were over 375,000 visits to MSU Library's Website during that time. This gives an idea of the amount of user traffic directed to CatSearch. Although the search interface is being heavily utilized, it is important to evaluate how this is being employed as well. In fact, when Serials Solutions made Summon usage data available to MSU, the most com- mon search query performed within the interface, other than a blank, was facebook.com. The frequency of facebook.com was problematic, due to the fact that CatSearch does not support Web addresses the way search engines do. With the use of transaction log analysis, this study aims to understand how the single search box model is being utilized. With the prolific query facebook.com in mind, two hypotheses were formed in order to understand how Web-scale discovery was being used at MSU. The first hypothesis was that search queries performed within CatSearch are of a low quality. Due to the fact that this was the first academic year of implementation at MSU, the second hypothesis of this study is that search query quality improved during the first two semesters of implementation.

RELATED STUDIES

There are numerous articles concerning Web-scale discovery that focus on the products offered and descriptions of their implementation. Wisniewski (2010)

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY
A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

provides an introductory overview of Web-scale discovery. Vaughan (2011) offers a series in Library Technology Reports that defines Web-scale discovery as well as describes the products offered. Dartmouth College Library (2009) details the evaluation process of these tools. Newcomer (2011) outlines how to configure Web-scale discovery tools to better work with music information retrieval.

Two studies have approached Web-scale discovery tools from the perspective of librarians. Howard and Wiebrands (2011) surveyed the librarians at Edith Cowen University, and they found some general concerns about Web-scale discovery, such as the "dumbing down" of students, content coverage, and trust in the interface. Buck and Mellinger (2011, 171 72) surveyed instruction and reference librarians across institutions implementing Summon to determine if and how they are teaching Summon to students. They found a general ambivalence toward Summon among librarians, and that it has   not yet been fully integrated into instruction. Way (2010) studied the impact that Summon had upon collection usage at Grand Valley State University Libraries. He found that full-text use increased with the implementation of Summon.

Transaction log analysis (TLA) has been used in a number of studies to learn more about human computer interaction and information seeking in the library and information science field. Jansen (2006, 408) defines a transaction log as "a file (i.e., log) of the communications (i.e., transactions) between a system and the users of that system." Peters (1993) provides a comprehensive history and literature review of TLA until 1993. Initially TLA was used to evaluate information retrieval systems, but the benefits of using TLA to unobtrusively observe users' interactions with information retrieval systems was quickly recognized among researchers (42). Jansen (2006) outlines three steps in the TLA methodology: collection, preparation, and analysis. TLA has been used to better understand the use of OPACs and how to improve OPAC interfaces (Peters 1989; Wallace 1993; Blecic et al. 1998; Ciliberti et al. 1998; Moulaison 2008; Liu 2010). Peters (1989, 270) examined the number of hits within an OPAC, and he found an overall failure rate of roughly 30–40 percent. Wallace argues for better OPAC system design to meet users where they are, rather than where they should be. She found variable use of system search help, high scanning of results lists, and that the majority of keyword searches produced twenty-five or fewer results within an OPAC. Blecic et al. (1998) tracked changes in transaction logs from an OPAC after usability changes in the interface were made, resulting in improved user performance within the interface. Ciliberti et al. (1998) combined demographic survey data with transaction logs, resulting in a number of collection decisions, interface design changes, and service changes within a specific

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY
A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

library. Moulaison (2008) found low Boolean operator and limiter use, and a failure rate of one in three within an OPAC at a medium sized academic library. She recommends better system help and more customizable interface design based on location specifications. Liu (2010) measured the use of default search settings in two different OPACs and discovered that users are inclined to use the default settings regardless of the types of searches set as the default. Liu stresses the role of the default settings in users' satisfaction of an OPAC. Other information retrieval systems have been observed with TLA as well. Shiri (2011) observed patterns of search behavior within a nascent and interdisciplinary field among users of a nanoscience and technology digital library and databases. Asunka et al. (2009) used TLA to discover how academic library users employed different features of a library Website. Not surprisingly, they found high use of the library resources and services, high use of the upper left portion of the Web page, and low use of library announcements and advertisements. Xie and Wolfram (2009) performed a longitudinal study of database usage within a public digital library over a six-year period. They found growing academic use of the databases, as well as a growing focus on the social sciences and humanities along with decreased use in business, finance, and entertainment. Wolfram (2008) compared transaction logs from four types of information retrieval systems, and he found that users change their search query patterns depending upon the type of system being used. Jansen and Spink (2006) compared transaction logs from nine Web search engines and outlined the methodology employed in the use of TLA on the Web. Cothey (2002) observed the browsing and searching patterns among novice and expert users on the Web to understand changes in user behavior as experience with the Web is gained. She also found that users typically assume a more passive browsing role as they gain more experience using the Web. Huntington, Nicholas, and Jamali (2007) compared a number of metrics from transaction logs within a specific search engine, arguing that effectiveness and success of a search engine could be measured with this method.

The literature does provide some information about the strengths and weaknesses of TLA. Kurth (1993) outlines a number of limitations in the use of TLA. He notes that although TLA is an unobtrusive way to collect large data sets, there are limitations to the method, such as system constraints in the collection of log data, unknown demographic information about users, ineffective analysis, and privacy. Peters (1993) focuses on the inability to understand the actual intentions, emotions, and information needs of users during TLA. Combining TLA with other methods, such as surveys, has been proposed as a way to overcome some of the limitations of TLA (54).

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY

A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

The patterns of search query formulation among academic users within a Web-scale discovery tool are poorly understood. As described above, recent research reveals some information about users' and librarians' perceptions of Web-scale discovery tools (Buck and Mellinger 2011; Howard and Wiebrands 2011). Usage of materials after Web-scale discovery implementation has been evaluated as well (Way 2010). As described above, TLA has been applied to understand user behavior in other types of information retrieval systems, but transaction logs from web-scale discovery tools have not yet been analyzed. This study uses TLA to fill in the knowledge gaps about query formulation in a Web-scale discovery interface. This study addresses the hypotheses that the quality of search queries performed within Summon are low, and that the quality of search queries performed within Summon improved during the first two semesters of implementation.

METHODS

This study followed Jansen's (2006) suggested methodology for transaction log analysis, utilizing collection, preparation, and analysis.

DATA COLLECTION

Serials Solutions stores all of the search queries that are performed within the single search box on the MSU Library home page in a feature called the administration console. The total number of search queries from each month during the period August 2010 through April 2011 were downloaded from the Summon administration console in May 2011. Table 1 shows the total search queries performed for each month within each semester (see Table 1).

Semester	Month	Number of Search Queries
Semester 1	Aug	4193
	Sep	11488
	Oct	16199
	Nov	13000
	Dec	6887
Semester 2	Jan	6418
	Feb	9870
	Mar	11425
	Apr	14726

TABLE 1 • Total Number of Unique Search Queries by Month
Note. The table displays the total number of search queries, excluding blanks, for each month of the study. The data is divided into semesters 1 and 2.

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY

A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

All search query processing and statistical analysis were performed in R (R Core Development Team 2010). The data set was constructed from a stratified random sample of one hundred search queries for each month (Connaway and Powell 2010, 123). Since Serials Solutions lists the search query and its frequency, search queries were multiplied by how many times they were performed before the random sample was extracted. The month of origin for each search query was retained for analysis of results. Thus, a total of 900 search queries was used for the study.

DATA PREPARATION

The majority of search queries listed in the Serials Solutions administration console were blanks, meaning nothing was entered into the search box. A blank search box could mean that users are skipping the search box and navigating to other areas in the interface, such as the advanced search; thus, blank search queries were removed from the data set. In order to input the data into R, some special characters (such as *, {}, "", and ~) were also removed. The removal of these characters from the data set does eliminate some potentially useful information, especially since SUMMON supports the use of * and "" symbols. The removal of some special characters is addressed further in the discussion.

Each of the nine hundred search queries was assigned to one of seven search query types, and the subjectivity of search query quality and type was addressed by creating a set of rules to determine the following search query types:

- URL (URL) search queries were determined based upon whether there was a ".com," complete URL, or if the search showed an obvious intention of accessing a URL, such as facebook.com.

- Invalid (Inv.) search queries were determined based upon whether the in- formation could have been contained anywhere in the library's collection. These types of search queries revealed that the user did not understand the purpose of CatSearch or had confused CatSearch with a site search. Search queries such as CHEM 141, reserves, and peer reviewed were placed in this category.

- Natural language (NL) search queries were determined by the use of complete sentences or words that are normally considered stop words, such as the, if, or an. Any search query with three or more words in it was copied and pasted into a title search in WorldCat in order to determine if the search was for a known item. Natural Language search queries with no results listed in WorldCat were coded as such.

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY

A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

- Database/journal (DB/journal) search queries were generally recognized by the author, who has knowledge of the databases and journal subscriptions at MSU Library. If there was a question about a publication being a journal, it was copied and pasted into WorldCat. For example, Montana Stockgrower was entered into a WorldCat title search and was determined to be a journal.

- Subject (Subj.) search queries refer to queries in which topics, names, or keywords were searched. As stated above, if there were three or more words in a search query, a title search was performed in WorldCat. If the search query yielded no results with a matching title, it was coded as a subject search query. Any author's names listed without a preceding "Author:" or "AuthorCombined:" were determined to be subject search queries as well. Search queries that contained two or fewer words were determined to be subject search queries unless they were obvious titles or journal/database titles. Shining was considered a known item search, and a u was determined to be a journal.

- Known item (KI) search queries were easier to determine. These were search queries in which the user is searching for a particular item in which the author or title is known. As discussed above, three or more words resulted in a title search in WorldCat and determined if the search query was for a known item. Conserving energy in the home was coded as a known item search query. It is worth noting that search queries with two terms or less required somewhat of a judgment call from the author. For example, photography is part of a number of book titles, but since the topic is so broad, it was determined to be a subject search.

- Boolean (Bool.) search queries were determined based upon the use of the operators: AND, NOT, OR. An example of a Boolean search is no impact AND environment.

After the search queries were coded, the seven search query types were grouped into high and low quality. The concept of quality for this study was dependent upon whether or not the system would support the methods implemented within the search query. For example, a URL is an effective way to locate information within a search engine; however, URLs are not supported by Web-scale discovery, and, thus, they were labelled as a low-quality search query. The low-quality grouping consisted of the URL together with invalid and natural language search query types. The high-quality grouping consisted of subject, Boolean, known item, and database/journal search query types.

STATISTICAL ANALYSIS

Two-sample t-tests were used to compare the proportions of high- versus low-quality search queries for the entire academic year, as well as for both semesters

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY
A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

individually; the proportions of high-quality searches by semester were also compared to detect a change in search quality between the semesters. Semester 1 consisted of months August through December, and semester 2 consisted of months January through April. Simple linear regression was performed on each search type, as well as for the combined high-quality search queries, through the academic year to test for a change in any of the categories over time.

RESULTS

Figure 1 shows the percentages of each search query type by month, with high-quality categories as the lightest shades of gray and low-quality as the darkest shades. Subjects were clearly the most common type of search query, followed by known items. It is also apparent that low-quality search queries are a small proportion of the overall search queries coded in the study. This also shows that the quality of search queries did not appear to change over time. The high-quality search percentage is similar between semesters with no significant change ($p = .536$, from a two-sample t-test). The mean for high-quality search queries in semester 1 is 91 percent, and the mean for high-quality search queries in semester 2 is 92.25 percent. The standard deviation for semesters 1 and 2 is 2.83 and 2.87, respectively. There was a significant difference between high- and low-quality search queries over the entire academic year ($p < .001$, from a two-sample t-test). The mean of the high-quality search queries was 91.56 percent, while the mean of low-quality search queries was 8.44 percent. Finally, there were no significant trends in either high-quality search queries through the year or individual search query categories through the year ($p > .1$ for the slopes of all linear models).

The results put the hypotheses of this study into perspective. The first hypothesis of this study stated that the quality of search queries within Summon is low. This hypothesis is rejected given the large and consistent differences between high- and low-quality search queries. The second hypothesis of this study stated that the quality of search queries in Summon improved during the first two semesters of implementation. This hypothesis is also rejected since there was no difference in high-quality searches between the two semesters; we also found no significant trends through the year for any of the search query categories.

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY

A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow
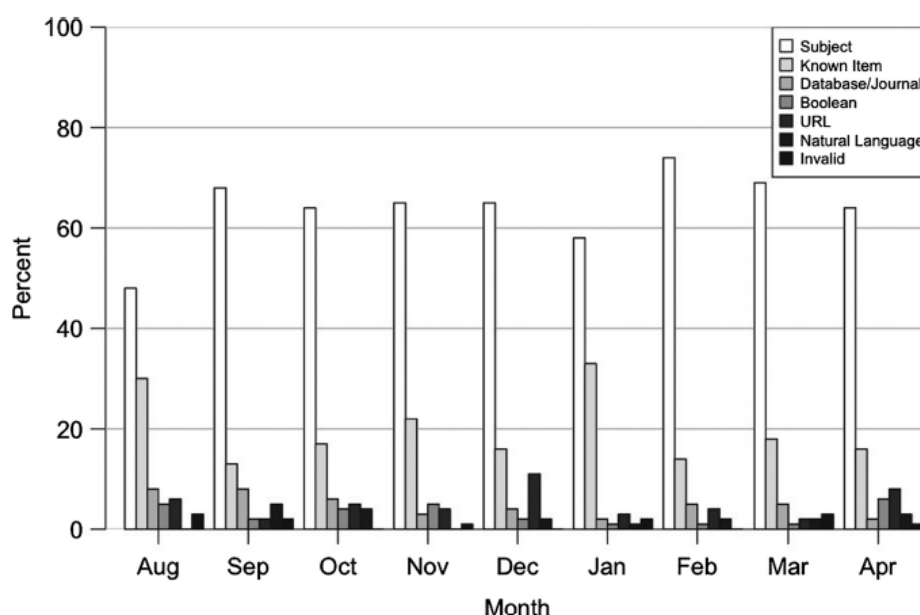
FIGURE 1 • Search queries by category. This bar graph displays the frequency of search types by month with high quality search queries in four light shades and low-quality search queries in three dark shades.

DISCUSSION

These results could signify that the single search box model of discovery is sufficient for most students. However, there are some who do not understand the most effective ways to search and others who do not even understand the meaning of searching library resources. The majority of students are already accustomed to the single search box method through the use of Internet search engines, such as Google, which has set the standard for information seeking (Way 2010). The fact that most search queries are legitimate could mean that library resources are becoming easier to access through Web-scale discovery tools.

There are some implications for further instruction when only the high-quality search queries are considered. Table 2 shows the number and percentage of search query types for the high-quality search queries performed.

Since the use of Boolean operators yields more effective results, a higher percentage of this search query type would be an indication of effective searching behavior. As noted in the methods above, since special characters were removed from the data

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY
A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

Search Type	Number of Search Queries	Percentage of Search Queries
Boolean	27	3.30%
Database/Journal	43	5%
Known Item	179	21.70%
Subject	575	70%
Totals:	824	100.00%

TABLE 2 • Number and Percentage of Search Types within the High-Quality Search Queries Performed

set, a greater percentage of Boolean search queries could have been possible. The low percentage of Boolean operator utilization found in this analysis was consistent with other studies (Moulaison, 231; Shiri, 141; Peters 1989, 271). There is also a wide variety of qualities found within the search queries when the mere validity or invalidity of the queries is disregarded. For example, the search queries wolf management and peter were both considered high quality subject search queries for the purposes of this study. However, wolf management would likely yield much more effective results than peter, which is perhaps too broad to be meaningful and would require further specification to yield effective results. Unfortunately, with the limitations of TLA, it is impossible to determine whether or not the users considered their search sessions to be successful. Overall, the concept of quality is relative, and the ability to form an effective search query appears to be out of reach for many users of the system.

Some librarians believe that Summon has afforded them more time to focus upon concepts like forming effective search queries rather than the specifics of certain databases (Buck and Mellinger 2011, 173). This study indicates that although users are likely getting by using CatSearch, information literacy instruction focusing on the use of Boolean operator searching and keyword selection would be helpful in the academic environment.

This study also opens a communication pathway between libraries and vendors. For example, although the most common search query per- formed within CatSearch is facebook.com, when the queries are divided into types of search queries the URL searches are quite low in comparison to subject searches. This type of information is useful for vendors so they can better understand how their products are being utilized in real-world situations.

Implications for interface design, especially at MSU, can be gleaned from this study. For example, discussion about enabling an MSU Library site search feature within

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY
A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

CatSearch has ensued based upon the results of this study. In the future, it may be possible for the users who searched reserves or CHEM 141 to arrive at their expected destination.

Further studies applying TLA with more metrics to Web-scale discovery tools would be beneficial for understanding not only what search queries are being used but also the number of results listed, number of search results visited by the user, time spent within the interface, and turnaways. As discussed above, the intentions of the users can never be known, but analyzing the entire search session can allow for better understanding of what, if anything, discovery layers are delivering for library users. The usage of Web-scale discovery tools could also be compared to the usage of other information retrieval systems in order to determine whether users treat Web-scale discovery tools more like Web search engines, databases, or OPACs. The impact of discovery layers could also be determined by com- paring how much Web-scale discovery tools are used in relation to other features or information retrieval systems within a library Website. Web-scale discovery is a nascent technology with much potential for future research from both user and system perspectives.

CONCLUSION

The implementation of CatSearch as MSU Library's Web-scale discovery tool has been relatively successful; however, there may be further implications for the instruction of university students in the formation of an effective search query. Additional instruction on how to form an effective query may be necessary in order to distil the information library users need from the volume of resources available via discovery layers. The various search query types that emerged from the study may leave room for a response from an interface design perspective as well. This study opens a pathway of communication between libraries and vendors to better understand how Summon is being used in a real-world academic environment. This simple study merely scratches the surface of what can be accomplished in understanding Web- scale discovery tools and how users are interacting with them. By better understanding users, their expectations, successes, failures, and skill levels, libraries can offer more effective services to meet the needs of information seekers with the quality of library-provided resources.

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY

A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

Excerpted from *Discovery Tools: The Next Generation of Library Research*

REFERENCES

Asunka, S., H. S. Chae, B. Hughes, and G. Natriello. 2009. "Understanding Academic Information Seeking Habits through Analysis of Web Server Log Files: The Case of the Teachers of College Library Website." The Journal of Academic Librarianship 35:33–45. http://search.ebscohost.com/login.aspx? direct true&db lxh&AN 36911604&site ehost-live.

Blecic, D. D., N. S. Bangalore, J. L. Dorsch, C. L. Henderson, M. H. Koenig, and A. C. Weller. 1998. "Using Transaction Log Analysis to Improve OPAC Retrieval Results." College & Research Libraries 59:39–50. http://crl.acrl.org/content/59/< 5848   >/39.full.pdf html.

Brophy, P. 2004. Evaluating the joint information systems committee's information environment: The EDNER and EDNER projects. VIN 43:143–47. doi: 10.1108/ 03055720410699928.

Buck, S. and M. Mellinger. 2011. The impact of Serials Solutions' Summon on in- formation literacy instruction: Librarian perceptions. Internet Reference Services Quarterly 16:159–81. doi: 10.1080/10875301.2011.621864.

Caswell, J. V., and J. D. Wynstra. 2010. "Improving the Search Experience: Federated Searching and the Library Gateway." Library HI Tech 28:391–401. doi: 10.1108/0737883101107664.

Ciliberti, A., M. L. Radford, G. P. Radford, and T. Ballard. "Empty Handed? A Material Availability Study and Transaction Log Analysis Verification." 1998. The Journal of Academic Librarianship 24:282–89. doi: 10.1016/S0099-1333(98)90104-5

Connaway, L. S., and R. R. Powell. 2010. Basic Research Methods for Librarians. 5th ed. Santa Barbara, CA: Libraries Unlimited an Imprint of ABC-CLIO, LLC.

Cothey, V. 2002. "A Longitudinal Study of World Wide Web Users' Information-Searching Behavior." Journal of the American Society for Information Science and Technology 53:67–78. doi: IO.lOO2/asi.lO011.

Dartmouth College Library. 2009. An Evaluation of Serials Solutions Summon as a Discovery Service for the Dartmouth College Library. http://www.dartmouth.edu/ library/admin/docs/Summon_Report.pdf.

Griffiths, J., and P. Brophy. 2005. "Student Searching Behavior and the Web: Use of Academic Resources and Google." Library Trends 53:539–54. http://search.ebsc ohost.com/login.aspx?direct true&db afh&AN 16811207&site ehost-live.

Howard, D., and C. Wiebrands. 2011. "Culture Shock: Librarians' Response to Web Scale Search." Edith Cowen University Research Online. http://ro.ecu.edu.au/cgi/ viewcontent.cgi?article 7208&context ecuworks.

Huntington, P., D. Nicholas, and H. R. Jamali. 2007. "Employing Log Metrics to Evaluate Search Behaviour and Success: Case Study BBC Search Engine." Journal of Information Science 33:584–97. doi: 10.1177/0165551506077407.

Jansen, B. 2006. "Search Log Analysis: What It Is, What's Been Done, How to Do It." Library & Information Sciences Research 28:407–32. doi: 10.1016/j.lisr. 2006.06.005.

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY

A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

Jansen, B., and A. Spink. 2006. "How Are We Searching the World Wide Web? A Comparison of Nine Search Engine Transaction Logs." Information Processing and Management 42:248–63. doi: 10.1016/j.ipm.2004.10.007.

Kurth, M. 1993. "The Limits and Limitations of Transaction Log Analysis." Library Hi Tech 11:98–104.

Liu, W. 2010. "Remote Users' OPAC Searching Habits: A Comparative Case Study through Web Transaction Log Analysis." Kentucky Libraries, Summer, 6–13. http://search.ebscohost.com/login.aspx?direct true&db lxh&AN 52597526& site ehost-live.

Moulaison, H. L. 2008. "OPAC Queries at a Medium-Sized Academic Library: A Transaction Log Analysis." Library Resources & Technical Services 52:230–37. http://search.ebscohost.com/login.aspx?direct true&db afh&AN 34692535& site ehost-live.

Newcomer, N. L. 2011. "The Details Behind the Web-Scale: Selecting and Configuring Web-Scale Discovery Tools to Meet Music Information Retrieval Needs." Music Reference Services Quarterly 14:131–45. doi: 10.1080/10588167.2011.596098.

Peters, T. A. 1989. "When Smart People Fail: An Analysis of the Transaction Log of an Online Public Access Catalog." The Journal of Academic Librarianship 15:267–73. http://search.ebscohost.com/login.aspx?direct=true&db=afh&AN=7825790&site=ehost-live.

———. 1993. "The History and Development of Transaction Log Analysis." Library Hi Tech 11:41–66.

R Development Core Team. 2010. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org.

Shiri, A. 2011. "Revealing Interdisciplinarity in Nanoscience and Technology Queries: A Transaction Log Analysis Approach." Knowledge Organization 38:135–53. http://search.ebscohost.com/login.aspx?direct true&db lxh&AN 59621133& site ehost-live.

University College London (UCL) CIBER Group. 2008. CIBER Briefing Paper 9: In- formation Behaviour of the Researcher of the Future. http://www.jisc.ac.uk/media/documents/programmes/reppres/gg_final_keynote_11012008.pdf.

Vaughan, J. 2011. "Chapter 1: Web-Scale Discovery: What and Why?" Library Technology Reports, January, 5–11. http://go.galegroup.com/ps/i.do?id GALE%7CA246955415&v 2.1&u mtlib_1_1123&it r&p AONE&sw w.

Wallace, P. M. 1993. "How do Patrons Search the Online Catalog When No One's Looking? Transaction Log Analysis and Implications for Bibliographic Instruction and System Design." RQ 33:239–52. http://go.galegroup.com/ps/i.do? id GALE%7CA14769930&v 2.1&u mtlib_1_1123&it r&p AONE&sw w.

Way, D. 2010. "The Impacts of Web-Scale Discovery on the Use of a Library Collection." Serials Review 36:214 20.doi: 10.1016/j.serrev.2010.07.002.

Wisniewski, J. 2010. "Web-Scale Discovery: The Future's So Bright, I Gotta Wear Shades." Online, July August, 55–57. http://go.galegroup.com/ps/i.do? id GALE%7CA230959395&v 2.1&u mtlib_1_1123&it r&p AONE&sw w.

SEARCH QUERY QUALITY AND WEB-SCALE DISCOVERY

A QUALITATIVE AND QUANTITATIVE ANALYSIS

Kelly Meadow and James Meadow

Wolfram, D. 2008. "Search Characteristics in Different Types of Web-Based IR Environments: Are They the Same?" Information Processing and Management 44: 1279–92. doi: 10.1016/j.ipm.2007.07.010.

Wrubel, L., and K. Schmidt. 2007. "Usability Testing of a Metasearch Interface: A Case Study. College & Research Libraries 68:292–311. http://crl.acrl.org/ content/68/4/292.full.pdf html.

Xie, I., and D. Wolfram. 2009. "A Longitudinal Study of Database Usage within a General Audience Digital Library." Journal of Digital Information 10:1–19. http:// search.ebscohost.com/login.aspx?direct true&db lxh&AN 48438957&site ehost-live.

Zipf, G. K. 1949. Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology. Cambridge, MA: Addison-Wesley Press.

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

This chapter is excerpted from

Designing and Developing LIbrary Intranets

Edited by Nina McHale

LEARN MORE ›

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

Intranets should provide quick and easy access to organizational information. The University of Alabama Libraries' intranet was only partially satisfying this basic expectation. Librarians could use it to find forms, policies, committee assignments, and meeting minutes, but navigating the libraries' intranet was neither quick nor easy, and it was only one of multiple sources for essential internal information. The Web Services Department of the University of Alabama Libraries was responsible for directing the redesign of the intranet. Moving to the open-source Drupal content management system (http://drupal.org), Web Services launched a revamped public Web site in January 2009. The intranet was slated for a similar redesign and conversion to Drupal by the end of the same year. The goal was to build a site that served as a center for information for library faculty and staff, provided a stream of information to keep librarians throughout the system connected, contained personalized features based on an individual's group memberships, and created a collaborative environment for all library personnel. The new intranet is a one-stop source for internal information and includes features to promote communication, professional development, and collegiality. The database-driven Drupal framework provided greater flexibility in organizing and presenting information and allowed inclusion of additional data sources. By wrapping together existing disparate information sources, adding new functionality, and giving users a platform for adding content, the new intranet is designed to be an integral part of librarians' daily workflow.

Intranets serve a basic function as storehouses of organizational information. In the case of the University of Alabama Libraries, the intranet has historically consisted of little more than forms, policies, and committee minutes. Is that enough? Some researchers recognize the potential of intranets as conduits of information that help "management and group members obtain a clearer picture of what is really going on" throughout an organization (Denton 2006, 6). Aligning the approach of this project with this broader concept of the intranet as a way to move information, the University Libraries undertook a redesign that expanded functionality by centralizing all important internal information, facilitating and promoting broad participation, and leveraging existing technologies to make the intranet easy to use and manage.

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

The new intranet ties library faculty and staff to the daily activities and events of the organization and all of its parts while inviting them to be active contributors. Serving the research needs of faculty and more than 28,000 students, the University Libraries are a vital part of Alabama's oldest public university. The University Libraries employ more than 120 faculty and staff in 5 libraries across campus. Within the University of Alabama Libraries, the Web Services Department is responsible for developing and maintaining Web interfaces, applications, and services for faculty and students as well as library personnel. The department consists of two library faculty and two professional staff. While directing the libraries' Web presence, the Web Services Department works diligently to include all applicable stakeholders in their projects. This consultation occurs formally in the form of the libraries' Public Interface Working Group. This group consists of library faculty and staff from units, departments, and libraries throughout the organization and provides counsel to the head of the Web Services Department, who serves as the group's chair. During the intranet redesign, Web Services frequently briefed the Public Interface Working Group and periodically updated all library personnel on the status of development. Web Services also made certain that all library faculty and staff had an opportunity to offer feedback prior to launch.

BACKGROUND

Libraries have taken a variety of approaches to intranet design, from infrastructure and functionality to content and presentation. Using a third-party application versus in-house development is one of those differences. Some librarians have had great success in building their own intranet (Engard and Park 2006). Others have relied on proven commercial and open source solutions. As it pertains to functionality, the University Libraries are moving from a narrow focus of the intranet as an information repository to an approach that enables and emphasizes collaboration among library personnel. Facilitating library faculty and staff interaction opens the potential of the intranet to serve as a decision-making and strategic tool (Denton 2006). With this in mind, the Web Services Department sought to "deliver the goods" by building a site that gave library personnel a "compelling reason to use the intranet" (Fichter 2006, 52).

DESIGN DECISIONS

The decision to use a content management system (CMS) for the new University Libraries' intranet was made in early 2008 when the Web Services Department was

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

developing a plan for redesigning the libraries' public Web site. At that time, both the public Web site and intranet consisted almost entirely of static Web pages. Web Services researched several infrastructure options for the redesign. Flexibility in design, ease of site management, and the ability to precisely control access were all important. The department quickly recognized the capacity of CMSs to address all these needs, with   the greatest advantage of CMSs being their separation of the site's data from the presentation layer. This infrastructure change would make future aesthetic site redesigns much easier, since changes to style elements would not have an impact on the database housing library content. The CMSs considered also enabled better control over access to particular content areas. This alleviated the reliance on UNIX file permissions or .htaccess directory restrictions on the server. Overall, CMSs were a much more efficient mechanism for managing the Web sites for this library.

In deciding which CMS to use, functionality, the size of the user base, and security were considered. To assist in the selection process, a Web site called "The CMS Matrix" (http://www.cmsmatrix.org) was used, which com- pares the attributes of more than 100 open-source and commercial CMSs. Several were reviewed, with the focus on Drupal, Joomla, and Microsoft's SharePoint. Drupal and Joomla are open source solutions. As an organization that relies on Microsoft's Exchange and Active Directory, SharePoint was also a viable candidate. However, Drupal was chosen because of its powerful access-control settings, large user community, easily modify themes for presentation, and capability to create custom modules.

The deliberation over the infrastructure of the University Libraries' re- designed public Web site and intranet and the selection of Drupal were important steps. However, as with any CMS, knowledge and expertise to configure and implement Drupal had to be gained to meet the goals for the Web environment. Proceeding with the redesign of the public Web site first may seem counterintuitive when learning a new piece of software, but the need to deliver an improved user experience and a manageable administrative environment for the public Web site was much more immediate and had an impact upon thousands of users. By the time the Drupal-based University Libraries' public Web site launched in January 2009, Web Services had learned much about how to adapt this particular CMS to the needs of this project. This knowledge provided a great head start in the work to redesign the libraries' intranet.

Having resolved the infrastructure decisions for the new intranet, the Web Services Department outlined and addressed key goals for the redesign. The desire was for the

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

new intranet to centralize as much information as possible. The existing intranet of forms, committee minutes, and statistics was a destination that library personnel visited as needed rather than daily. The intranet omitted key pockets of information that were contained in wikis, blogs, and network drives. Understanding the organizational culture provides in- sight as to why this disparate information existed. All University Libraries' faculty and staff are involved in multiple groups with intranet presences. From their organizational position in units or departments to numerous faculty and staff committees, working groups, and task forces, library personnel are never tied to fewer than two of these entities. Many groups have some type of intranet content. Bringing the information for an individual's affiliations into a central location without them having to check multiple sites or pages was a primary goal. Also desired was the addition of new functionality that encouraged participation, fostered collaboration, and promoted collegiality among librarians from across the organization.

A ONE-STOP SHOP

The goal of providing an intranet that centralized all important internal in- formation was not just about putting everything in one place but also about serving this information to users in an easily accessible and personalized way. The most significant part of the redesign that helped accomplish this goal was the incorporation of existing internal wikis into the CMS by using the third-party Organic Groups module for Drupal (http://drupal.org/project/og). Over the course of the last few years, numerous wikis were created for a variety of committees, departments, task forces, etc. Using the open source MediaWiki (http://www.mediawiki.org) software, Web Services installed and maintained these wikis. In almost every case, usage dramatically declined after a few months. Some of the wikis that remained active did so only because Web Services updated them.

The wikis fell into disuse for a few reasons. First, each wiki had a unique set of usernames and passwords. Since this information did not coincide with any existing user authentication system, wiki members quickly forgot how to log in. Initially, this led to requests for password resets, but eventually many users stopped trying to log in. Although one of the strengths of wikis is their ease of use, they introduce yet one more markup style for users to learn (Notess 2006). Most library personnel at the University of Alabama found wiki syntax frustrating when trying to format their content.

 Organic Groups (see Figure 1) enabled us to preserve the collaborative aspects of wikis while resolving the specific problems that impeded wiki usage in the

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles
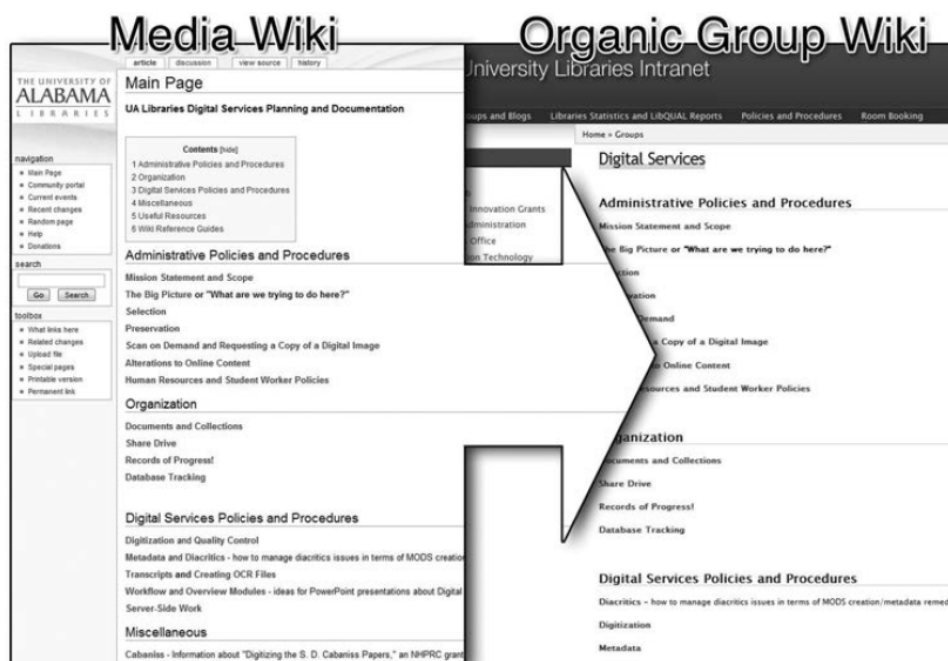
FIGURE 1 • Migrating MediaWiki pages to Drupal using Organic Groups.

environment. This module allowed the creation of both access control and content divisions based on any parameters chosen. This permitted the recreation of wiki content and easily set wiki content editors in the Organic Group environment. With this change, library personnel could make posts unique to the groups to which they belonged. Groups can be viewable to the public or viewable only to users who are members of the group, and each individual post in a group can be public or private as well. Posts to the groups can also function as a "wiki post," allowing all members of the group to edit the post, or as normal posts, where only the creator and administrators can edit. Automatic versioning of each edit meant that, as with wikis, users could easily revert to a previous version in case of error. Group posts are edited using the CKEditor (http://drupal.org/project/ckeditor), which library faculty and staff also use for creating and editing content on the public Web site. This editor accepts plain text and uses familiar word processing application icons for formatting. Creating or updating wiki posts does not require users to make the usual selections of pathname, alias, revision, etc. that regular Drupal page or node creation entails. The Pathauto module (http://drupal.org/project/pathauto) makes it possible to set the pathname and alias.

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

The ability to replicate, but simplify, the wiki environment within Drupal was the largest, but not the only, usage of Organic Groups in the libraries' intranet. As mentioned, the existing intranet had content pages for commit- tees, working groups, task forces, and a smattering of units and departments. These existed as both static and dynamic pages, but access was restricted to the head or chair of the organizational entity. Moving this content into the Organic Groups environment opened access to all members of the respective group but still enabled the traditional administrators of that content control over an individual's level of access.

With wikis and group pages in the CMS, Web Services sought to further encourage the organizational movement of information by serving up each individual's group information in a personalized feed directly to the front page of their intranet view. Using the power of the third-party Drupal Views module (http://drupal.org/project/views), users get a custom "My Groups" tab on the intranet front page (**see Figure 2**). This tab contains all the latest posts made to any of the groups to which they belong. Library faculty and staff no longer have to go to different sites or pages to stay current with all the information pertinent to their memberships.

**FIGURE 2 • ** The "Main Stream" and "My Groups" tabs and the new intranet.

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

NEW FEATURES

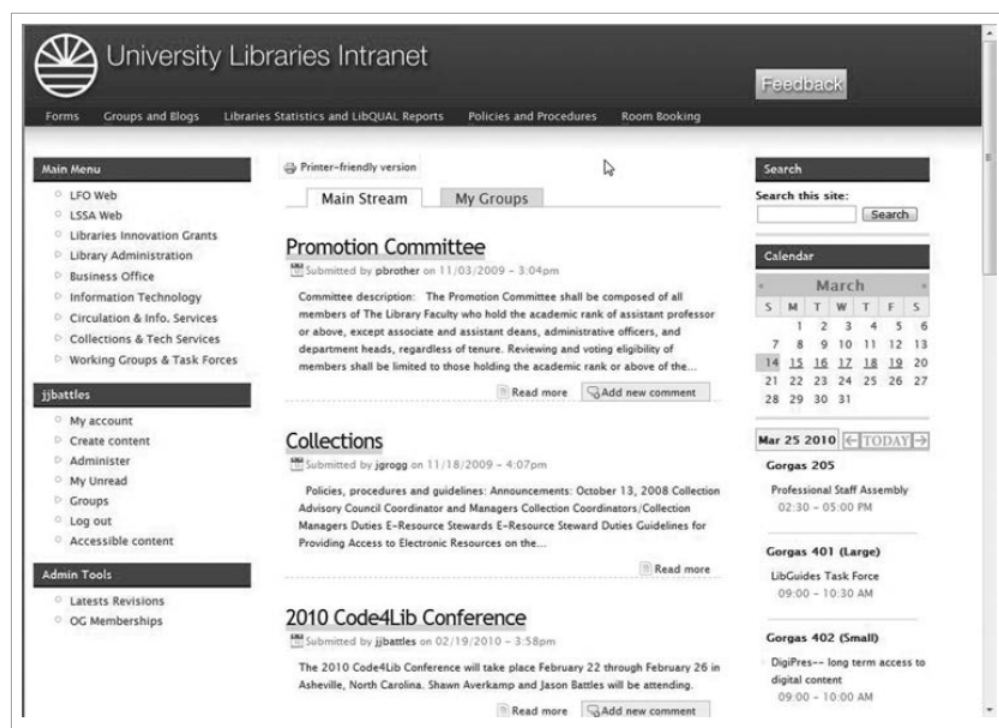Room-booking assignments for key venues within the libraries were added to the intranet's front page. The University Libraries have long relied on the open source MRBS room-booking software (http://mrbs.sourceforge.net/) for scheduling space for meetings, events, etc. During intranet development, other solutions were sought that would fit better within the CMS, but, in the end, MRBS functioned well in this environment, and its use of a MySQL database made it easy to pull into a custom Drupal module written by Web Services staff. As implemented, the right-hand column of the intranet top page includes the current day's bookings for select meeting space. This serves as much to inform library faculty and staff of what is taking place as to alert individuals of meetings they may have overlooked in their personal calendar. The custom-built MRBS module allows users to page forward a day at a time. Included alongside the room bookings listing is Drupal's Calendar module. All users are allowed to add items to the calendar. The purpose of the calendar is to display conferences, Webinars, and other events of organizational and professional interest. The idea is that librarians from various areas throughout the University Libraries will have a shared knowledge of their colleagues' professional events. As modern librarianship is very diverse, colleagues may not be aware of the distinctiveness of specialized areas of librarianship. Systems librarians, for example, attend very different events from catalogers or serials librarians. A simple calendar does not lead to an instant understanding of those unfamiliar areas, but it does expose the entirety of the library faculty and staff to the existence of those events, and it may lead the curious to gain a better understanding of their colleagues' work.

Three other new features were added to improve communication and build collegiality among library faculty and staff. The core panel of the Web site is the "Main Stream," a chronological list of the latest page updates and additions. While the "My Groups" tab personalizes updates for the user, the "Main Stream" section gives library faculty and staff a broader view of what is going on throughout the organization. Also, a professional development section was included so users could post information on conference presentations, published articles, or reports from conferences attended. Previously, this information was disseminated only within the users' respective departments or units—or not at all. Again, this gives the entire organization information about the professional activities of their colleagues. Web Services also included a more casual "water cooler" section to the site. The intent of this area was to allow people to post about topics not necessarily related to work. For example, perhaps someone is interested in quilting or gardening and wants to find like-minded

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

colleagues. Granted, there is no direct professional gain from such a forum, but it does build connections that translate into a richer work atmosphere.

While moving forward from the launch of the intranet, additional features will hopefully be included. The most significant addition considered is integrated chat. There are several different ways chat can be added to Drupal, including chat modules. Third-party applications like Meebo (http://www.meebo.com) can also be incorporated. Hosted chat applications, such as the open source Jabber server, are also an option. Tying chat into the Drupal profile system means online/offline status could easily be provided for users. Chat can be a useful internal tool for communication, as it is an easy way to ask a quick question with little disruption compared to a phone call or in-person visit, and it can be quicker to compose or reply to than e-mail.

AN INTRANET FOR ALL

Bringing all of the internal information important to library faculty and staff into a single Web space was key to building a highly effective intranet. Making sure all library employees could easily access, contribute, and collaborate within the intranet from anywhere was also vital to success. The University of Alabama Libraries' previous intranet relied upon unique accounts for out-of- library or off-campus access. Likewise, Web Services faculty and staff were the only library personnel able to upload files. Page editing was available to fewer than ten faculty and staff via the client-based Adobe Contribute software (http://www.adobe.com/products/contribute/). Any Web updates submitted via Contribute were reviewed by Web Services prior to publication, which could delay information getting to users. A better solution to both problems was needed.

Since intranets contain organizational information that may not always be fit for public consumption, some authentication method is usually used to validate users. For the University of Alabama Libraries' previous intranet, this method was .htaccess usernames and passwords. This method of directory access control is available via the Apache Web server (http://www.apache.org). The .htaccess usernames were not always the same as those used by the libraries' Active Directory or the campus LDAP. These differences led to frequent requests from library personnel for pass- word changes or general inquiries about how to log in to the intranet. This frustrated users and decreased use of the intranet. In the last year, IP ranges were added to the Apache configuration so that in-library faculty and staff workstations were automatically allowed into the intranet without a login. IP restriction was possible because the University of Alabama Libraries use a fixed range of IP addresses for

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

library employee workstations that does not share network subnet addresses with public workstations. While this did help viewing access from within the libraries, off-campus and editing access still required the user to log in.

The intranet redesign sought to address this limitation by relying upon existing authentication systems. Drupal's LDAP integration module (http://drupal.org/project/ldap integration) allowed the use of an outside authentication source for intranet logins. Since the intranet was internal to the University Libraries, Active Directory was a more appropriate source than the campus LDAP directory against which to authenticate. This meant intranet users could use their existing daily workstation login credentials to edit and upload files to the intranet from wherever they were. This simple change greatly improved intranet accessibility while also saving IT staff time in addressing username/password issues.

Having improved intranet access via Drupal's use of our existing Active Directory, the next barrier for faculty and staff usage was the ease of editing and uploading files. Coupled with a need to make intranet content editing easier was the desire to expand higher-level access to more library personnel. Historically, intranet editing privileges did not differ greatly from the public Web site. Prior to the public Web site's move into Drupal in January 2009, updates were also made via Adobe Contribute. Web Services was then responsible for reviewing and publishing the public Web site changes. The University Libraries' intranet worked the same way. Using a per-seat licensed product like Contribute limited those with access to make Web updates on both the intranet and the former public Web site. Even for those with the ability to make changes to the intranet, the publication of those changes was not instant.

The redesigned intranet's CMS made possible the inclusion of a built-in WYSIWYG editor and eliminated reliance on Contribute. This change was an essential part of opening the intranet to broader participation from library faculty and staff. The redesigned University Libraries' intranet uses the CKEditor module for Drupal (see Figure 3). This provides a Web-based, graphical editor similar to popular word-processing applications. No software is installed on any workstation. The editor has a familiar feel. The advanced user still has the ability to edit HTML code and add style elements, but the novice can avoid those complications. Even with the easy-to-use editor, the Web Services Department conducted multiple sessions with library faculty and staff to instruct them on its use. These sessions provided an opportunity to provide both assistance in usage and also explain overall changes to the editing and approval process.

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles
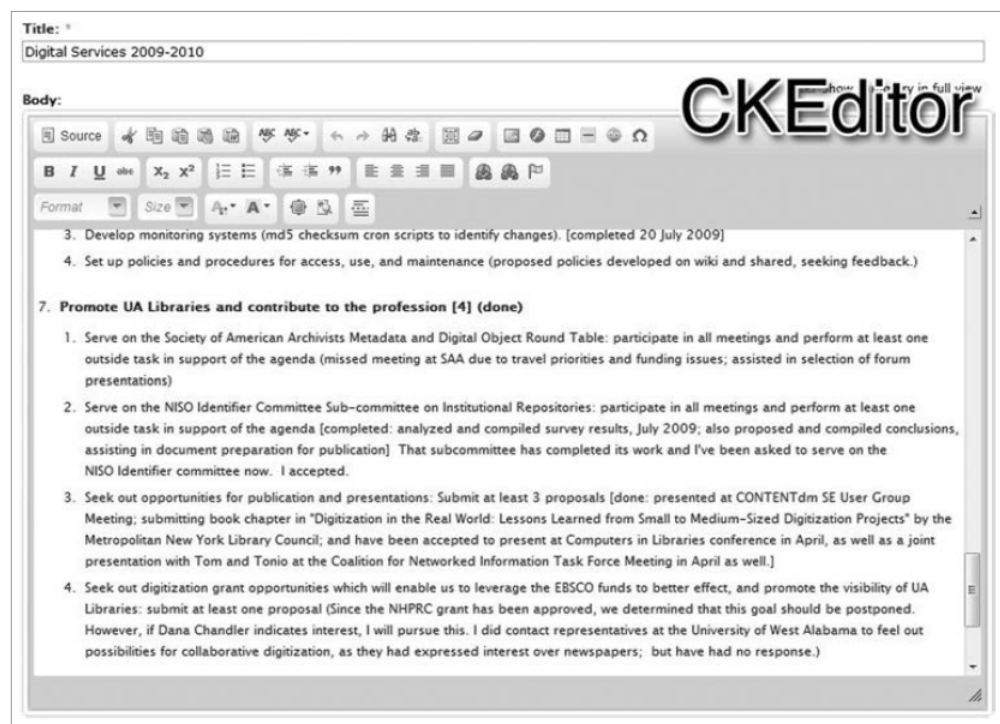
FIGURE 3 • The WYSIWYG CKEditor.

One of the many goals of the redesigned intranet was expanding access and building a collaborative space for library personnel. Intranets are different from public Web sites. With public sites, there is a level of oversight needed to control editing privileges and ensure content is reviewed prior to publication. The intranet, however, is designed for the organization. Minor errors made on the intranet will get noticed by a handful of people, but errors on the public site are noticed by thousands. That is not to make light of the importance of always providing good information whether the consumer is the public, campus, or staff; it is a balance between restraining intranet participation to protect content and encouraging participation with the risk of errors. Moving to a truly collaborative intranet means conceding some control. In the redesigned intranet, most edits are automatically published. However, there are technological solutions to make it easier to overcome the resulting problems. For example, Drupal was enabled to save each revision of a page, allowing the roll back to previous versions should a problem occur. As mentioned earlier, this feature is also one reason all of the intranet wikis were moved into Drupal.

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

IMPLEMENTATION PLAN

Putting the new features within the Drupal framework allowed a move for- ward with the implementation plan that would lead to the intranet's launch. While additional research and development time was spent mastering the modules necessary for this effort (Table 1), the bulk of the work to build the framework and move the content took place within a three-month period. In parallel with the Web Services Department's development work, all library faculty and staff responsible for the existing intranet content were contacted and a review of the policies, committee information, forms, and any other information they managed was requested. In most cases, these contributors were department, unit, committee, or organizational heads. It was asked whether their intranet content could be moved into the new intranet as-is, needed updating first, or should be deleted. As Sarah Robbins, Debra Engel, and James Bierman (2006) said, "Low contribution and lack of ownership can lead to dated and inaccurate information" (262). Giving ample time for content review and updating meant the content that was moved onto the new site was current and therefore useful to library faculty and staff. Coupled with the new content being brought into the intranet, this further improved the overall quality of the end product. By launch, 20 percent of the old intranet's content was marked obsolete and deleted, and 30 percent was updated by the content providers and moved into Drupal by Web Services. The remaining 50 percent of the content was moved as-is. However, two months after launch, roughly one-quarter of that content had been updated in Drupal by the content providers themselves.

By combining the content review with the intranet framework development, a beta phase could be more quickly moved into, where feedback, usability testing, and instruction sessions for library faculty and staff could begin. During the beta phase, the existing intranet remained accessible. This gave users time to try out the redesign without being forced to switch before becoming more comfortable with the changes. A similar approach was used with the public Web site launch but with a much longer beta period because the audience was much larger and more diverse. For the intranet, feedback was collected informally through e-mail and phone calls. A feedback link was also included on the site that stored user comments in a Web-accessible MySQL database. After launch, laptops with webcams and TechSmith's Morae (http://www. techsmith.com/morae.asp) software were used to conduct one- on-one usability tests to get a more detailed look at usage of the new site. Morae is a commercial usability software application that allows one to set tasks, track mouse movement and clicks, record video, create custom surveys, and compose it all into presentations and reports.

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

Function	Modules	Description
Access	Access Control List Content Access **LDAP Integration** Login Destination **Override Node Options** **Secure Pages**	The LDAP Integration module allows login from Active Directory of another LDAP source. Secure Pages allows the use of SSL. Override Node Options enables non-admin users to promote content to the front page.
Blogs, wikis, and committees	Automatic Nodetitles **Organic Groups** Pathauto Talk **Views**	One of the unique approaches to the Drupal intranet is that former stand-alone blogs and wikis have been incorporated into Drupal while preserving their traditional functionality. Organic Groups is the linchpin of this organizational structure, with Views providing the user interface framework.
Calendar	Calendar Date	These modules make the events calendar work.
Content moderation	Module Grants Revisioning	While Organic Group posts for blogs, wikies, and committees do not go through moderation, normal pages do.
Customized content types	Content Creation Kit (CCK)	CCK allows the use of custom content types, giving added flexibility in creating pages.
Editing	CKEditor WYSIWYG Filter	The CKEditor was released the end of 2009 and includes numerous editing options through a user-friendly graphical interface.
File Management	IMCE	IMCE is used predominantly for file uploading but also provides a browse and delete capability for uploaded content.
Printer friendly pages	Printer, e-mail, and PDF versions	Printing policies, meeting minutes, and other intranet content is highly requested.
Statistics	Google Analytics	Google Analytics is widely used throughout this organization. This module allows integration with Drupal.

TABLE 1 • Intranet Functionality and Corresponding Third-Party Drupal Modules Function Modules Description

Morae was used extensively for usability tests for the public Web site redesign. Since the new intranet was designed to not only allow but promote usage by all library faculty and staff, the Web Services Department conducted multiple sessions to instruct library personnel on how to use Drupal's CKEditor implementation and walked them through posting comments, uploading files, and creating pages. The beta phase lasted several weeks before we officially switching to the new Drupal-based intranet.

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

A few months after launch, the effect of the redesign began to be seen. The initial move of content from the old intranet into Drupal was completed by Web Services. Once the intranet launched and library employees began receiving training on using the new framework, they took responsibility for contributing and updating intranet content. This was new ground for many library faculty and staff. Prior to the move, the intranet averaged less than five page updates a month. In the new Drupal-based intranet, there were more than 50 updates a month. The previous intranet was open to editing by fewer than ten library faculty members. Now, all library employees have the ability to add content to the intranet. Currently, almost 50 staff members have taken advantage of this capability. More than one-third of those have created content or updated pages in the last 30 days. The new intranet includes functionality previously fragmented across various wikis, blogs, and Web sites or not editable by most University Libraries' employees. Not only has the new content led to more people using the intranet, but by opening permissions to allow broad participation those pages are being visited, updated and new content have been added at a considerable rate compared to the old site.

CONCLUSION

More feedback and usability information on the redesign are still being gathered, but an initial review has shown a significant change in intranet usage. Web Services will continue to assess the work thus far, making adjustments to better meet the needs of library faculty and staff. As updates to existing applications and new technologies emerge, constant re-evaluation of how the intranet functions and whether this approach remains the best solution for users is expected. For libraries using Drupal, 2010 is expected to bring the release of Drupal 7. The Drupal 7 core will integrate some of the most used third-party modules and include significant improvements to file and image handling. Will the changes in the new release result in extensive work for those libraries using earlier Drupal versions? Is there another CMS on the horizon that will better fill the needs of this library?

There are far too many unanswered questions to know the future effect of specific products, but it is safe to say that libraries undertaking intranet redesigns have much to consider. For now, the new University Libraries' intranet fulfils the initial goals. The robust and feature-rich redesign was produced by using a specific piece of software, but there is a need to stress that this approach was not determined by the technology. Drupal obviously was an important part, but it remains merely the vehicle that allowed for the achievement of goals. The new intranet marks a dramatic leap forward for the University Libraries, not only by putting so much information at users' fingertips but also by giving them new opportunities to participate and collaborate with colleagues.

DESIGNING AND BUILDING A COLLABORATIVE LIBRARY INTRANET FOR ALL

Jason J. Battles

REFERENCES

Denton, Keith. 2006. Strategic intranets: The next big thing? Corporate Communications: An International Journal 11(1): 5–12.

Engard, Nicole C., and RayAna M. Park. 2006. Intranet 2.0: Fostering collaboration. Online 30(3): 16–23.

Fichter, Darlene. 2006. Making your intranet live up to its potential. Online 30(1): 51–53.

Notess, Greg R. 2006. Diverging Web markup choices. Online 30(6): 43–45.

Robbins, Sarah, Debra Engel, and James Bierman. 2006. Using the library intranet to manage Web content. Library Hi Tech 24(2): 261–272.

This chapter is excerpted from

Designing and Developing LIbrary Intranets

Edited by Nina McHale

LEARN MORE ›

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

The days of top-down communication and controlled internal messages at a library organization are—or should be—behind us. Modern libraries must be fluid and flexible organizations with equally nimble internal communication infrastructures in place to keep up with the fast-paced environments that have been created in these organizations. As is the case at many institutions, McMaster University Library (about 100 employees) put a great deal of effort into public-facing resources and content, while the library intranet languished as an afterthought. Static Web pages were haphazardly created and linked to from the site's index page. As the site grew, the lack of global navigation, search functionality, and clarity about content ownership led to a large, confusing collection of pages that was increasingly difficult to maintain. In 2009, a project was undertaken to redesign the staff intranet and implement Drupal, an open-source content management system, to power the new site. This case study outlines the issues faced with the former intranet, requirements gathering, staff feedback, and usability tests performed to inform the redesign, site architecture, and Drupal modules implemented, features and benefits of the redesigned intranet, the use of the new intranet to corral existing Web 2.0/social media channels, governance, evaluation, and lessons learned from the project. Future phases of the project will focus on integrating other internal communication tools used by staff in their day-to-day work, including internal file-sharing drives, staff e-mail and instant messaging platforms, meeting scheduling software, and external document sharing tools such as Google Docs.

Consider this scenario: Michelle, a liaison librarian at a large academic library, completes a successful shift on the reference desk, during which she assisted two faculty members with finding research materials using the library's highly functional and recently redesigned Web site and catalog. During her shift, Michelle used the library's first-year experience wiki and new instructional videos on the library's YouTube channel to orient first-year students to library services and resources. Finally, she assisted a graduate teaching assistant with setting up a wiki on the library's wiki server for her second-year multimedia class.

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

Satisfied with her success on the reference desk, Michelle retreats to her office to work on a library committee. Searching for the committee page, she fruitlessly clicks around the library intranet, only to be frustrated by confusing navigation and repeated dead ends. Turning to the library blog directory, she finds the committee blog, where she becomes trapped in a loop that bounces her to the committee wiki, back to the intranet, then right back to the blog. Deciding not to allow her frustration to thwart her, she decides to edit the committee's page on the intranet to include links to all important online spaces and documents. When she remembers she needs   to complete an online request form to have the IT department make the changes she requires, she decides to set the task aside and move on to other priorities that can be completed with fewer hurdles and roadblocks.

In Intranets for Info Pros, Mary Lee Kennedy states, "In highly dynamic, flattening, and porous organizations, many information professionals are struggling with what information to manage, how much effort to focus on internal or external information sharing, and when to integrate information into natural workflows" (2007, 5). This statement perfectly sums up Michelle's plight in the above scenario. While Michelle is an entirely fictional character, her experience mirrors what many staff members faced at McMaster University Library prior to its intranet redesign. As is the case at many institutions, at McMaster University Library, a great deal of effort was put into public-facing resources and content, such as the library Web site, the catalog, and various Web 2.0 channels, while the library intranet languished as an afterthought. Through three reorganizations in a two-year period, the intranet remained a dispersed collection of rarely updated, static HTML pages that did little to mirror the evolving organizational structure. And with content updates being relegated to a few individuals in the library's IT department, every aspect of the intranet needed to be rethought, from the back end to governance and ongoing maintenance.

FROM WHENCE WE CAME: THE OLD INTRANET

The former intranet, referred to internally as Libstaff, got its inauspicious start a number of years ago when library staff realized they needed an online space to record internal communications such as procedures, forms, and contact information. Static Web pages were haphazardly created and linked to from the site's index page. As the site grew, the lack of global navigation, search functionality, and clarity around content ownership led to a large, confusing collection of pages that was increasingly difficult to maintain. The extent of organization on the intranet's index page was an alphabetical list of links to content that was loosely grouped by library department.

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

Knowledge of HTML and access to the server to publish pages were technical barriers in keeping content up to date. One common workaround strategy that developed was for library staff to create documents in a number of file formats (PDF, .doc, .xls, etc.) and have them uploaded to the intranet instead of creating well-indexed Web pages to house that content. This led to a jarring navigation experience, where clicking on a link in a navigation area would suddenly open a PDF file instead of leading to another Web page, as one might expect. Compounding this problem was a lack of standardization of file types.

In addition to site organization, the former intranet had security issues. Access was controlled by restricting IP ranges, which meant anyone with a campus IP was able to view pages on the intranet. As a result, sensitive information such as passwords and staff contact information could not be stored on the intranet. Furthermore, if a library staff member was working off-campus, use of a Virtual Private Network client was required in order to view intranet pages.

FROM THEN TO NOW: CONCEIVING THE NEW INTRANET

The system-wide frustration with the old intranet allowed quick and easy development of a list of basic requirements for the new site. These requirements included global navigation, site search, site authentication using existing staff log-in credentials, customizable permissions with different levels of access rights, simplification of content creation and maintenance, dynamic content, and a flexible underlying architecture that would allow the site to grow and evolve as the organization changed.

It was also acknowledged early on that one of the redesign's most important aspects was staff engagement and involvement in the process. This would play a role in the level of staff buy-in for the new site when it came time to launch. In response to this issue, an in-house, Web-based survey was performed prior to starting the redesign to find out more about library staff's use of and experience with the existing intranet. The following five questions were posed:

- Why do you visit the Libstaff Web site?
- What do you like about the current Libstaff Web site? What don't you like about the current Libstaff Web site?
- What features would you like to see on a redesigned Libstaff?
- Do you have any other comments you would like to share about Libstaff?

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

Excerpted from *Designing and Developing Library Intranets*

The survey, which was completed by 20 percent of staff, indicated:

- Staff visited the site for information on various library committees (minutes, meetings), for human resources needs (time sheets, performance documentation), to fill out forms (statistics recording, problem reporting), for policy/procedure information and help materials (staff FAQ, help sheets), and for news and current events.

- Staff liked to have one place to look for information and access to quick links.

- Some staff said they did not like anything about the site.

- Staff did not like the disorganization of the current site, the difficulty in editing and keeping it up to date, the site's general user unfriendliness and the absence of a search function, the lack of security, and the inability to easily access the site from off-campus.

- On a newly designed site, staff wanted the above items addressed, shared document drives incorporated into the site, and fun and interactive features such as photos, feeds, and comments added.

To further the level of staff engagement, about halfway through the re-design, the new intranet was presented at two all-staff forum events while it was still under construction. The intranet redesign was one of eight library projects showcased at the events and provided staff with an informal opportunity to comment on the proposed changes.

PLUMBING CONSIDERATIONS: BUILDING THE BACK END

It was realized early on that a database-driven content management system (CMS) would be the only option to fulfil many of the desired technical requirements, which included authentication, good indexing for site search, and author permissions. Fortunately, the authors had prior experience working with Drupal (http://drupal.org), an open source CMS, as part of a library Web site redesign just a few months earlier. The valuable experience gained from customizing Drupal for the library Web site redesign meant it instantly became the number-one choice for the intranet implementation. Armed with the requirements list and valuable staff feedback, requirements were mapped with Drupal's core functionality and added modules (see Table 1).

One of Drupal's biggest selling points is that the core application can suitably serve any Web site, and with the addition of user-developed modules that plug into the framework, just about any additional functionality that was needed could easily be added. Table 2 provides a list of modules implemented on the intranet and what each module accomplished.

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

Excerpted from *Designing and Developing Library Intranets*

Requirements	Drupal functionality
Global navigation	Drupal Themes allowed this to be accomplished easily
Site search	All content is housed in a database ("nodes") and indexed, allowing for excellent content indexing and search
Site authentication using existing staff logins	Decided on using Drupal's built-in authentication
Customizable permissions with different levels of access rights	User management within Drupal core allows for granular access and editing permissions
Simplification of content creation and maintenance	Content templates coupled with user permissions streamlines the creation and editing of content
Dynamic content	Drupal's database-driven back end allows the display of content in different ways and places on the site and easy inclusion of content from other channels using RSS syndication
Flexible underlying architecture	Drupal's database back end is extensible; also the core Taxonomy module allows for standardized architecture across the site

TABLE 1 • Mapping Requirements to Drupal Functionality

INFORMATION ARCHITECTURE AND NAVIGATION

Kim Guenther (2003) and Allison J. Head (2000) both noted that one of the major pitfalls of intranet design is organizing sites around a library's organizational structure. Being sensitive to this, the project began by creating a content inventory of the old intranet, allowing the effective categorization of that content. After some contemplation over how the site was likely to grow, the following "types" of content were decided on: department, service, project, committee, staff training and development, and policy and procedure. Working with the taxonomy and Content Creation Kit (CCK) modules in Drupal, content types and templates were created for each type. With these content types, discrete items such as a department could be easily added, deleted, or moved to reflect organizational changes. Furthermore, having "committee" and "project" content types allowed a link to be made from committees and projects to certain departments and for those linked relationships to just as easily be changed in the event of departmental reorganization.

The global navigation, which appears across the top of all intranet pages, consisted of the six content types and a seventh label for "Old Stuff." Three different taxonomies were used on the site to assist with generating dynamic views, and all new pages must be identified as a certain content type (department, committee, service, project,

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

Drupal module	Features/output
Content Creation Kit (CCK)	Allows site administrators to create templates for different types of content; e.g., committee template contains the following fields: committee name, chair/lead, links (optional), mandate, membership, start/end date, minutes (optional), additional information (optional)
Comment	Allows logged-in users to comment on any page or node
Contact	Provides contact forms for every user on the site, allowing staff to contact each other easily
Development	For site administrators, allows for easy troubleshooting
FCKEditor	WYSIWYG editor that allows site authors to format content using recognizable formatting icons (no knowledge of HTML required to add/edit site content)
Filters: Headings Anchors and Table of Contents	Provides authors with an easy way to create internal page links and tables of contents for pages based on headings
Forum	Threaded discussion boards for improved communication and discussion among staff
Menu	Allows site administrators to easily customize and maintain global navigation
Path & Pathauto	Allows users to rename URLs and set up predetermined paths for certain types of content
Poll	Useful for getting staff feedback quickly and engaging staff in fun ways
Profile	Staff can edit their own profiles, add avatars, etc.
Search	Essential for site-wide searching
Statistics & Google Analytics	Usage statistics and analytics to help with continuously improving the site
Taxonomy	Powers the organization of the site based on category terms
Taxonomy Breadcrumb	Allows site administrators to build page breadcrumbs based on taxonomy terms in use
Upload	Allows individual site authors to upload files to pages; particularly useful for committee minutes
Views	Provides site administrators with various options for displaying the database-driven content

TABLE 2 • Drupal Modules Implemented

etc.). Additionally, the views module in Drupal was used to dynamically update the landing pages for each of the items in the global navigation. Thus, the landing page for "Department" lists all the library departments in an alphabetical browse-all list. To move an item to "Old Stuff," it simply needs to be tagged as such. Site authors could also give this item a secondary tag, such as "committee" or "project," to indicate an old committee or completed project. Some items, such as policies and procedures, are also tagged as belonging to a certain library department.

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

CONSIDERING GOVERNANCE

The core ability in Drupal was used to customize site permissions and designated five different user roles with varying degrees of editing access to content. Compared to the former staff intranet, this greatly increased the number of people who were able to maintain and update certain portions of the Web site. With the new site, 48 staff members—nearly half of the staffing complement—hold one of the five editing roles (see Table 3). Permission to create new pages that were not a designated content type was not granted widely. To maintain the integrity of the site architecture, only site administrators were given the right to perform this task.

Role on the intranet	Level of permission	Role in the library
Administrator	Create and edit any content type; administer all other functionality on the site (taxonomies, menus, blocks, comments); create new pages	User Experience Librarian (public website manager), Marketing, Communications and Outreach Librarian
Committee/service/project editor	Create and edit committees, services, and projects	All staff who chair/lead a project, service, or committee
Department editor	Edit department landing page content; create and edit policies and procedures; edit any page	Department heads, designated department contacts
Staff training editor	Create and edit staff training development pages	Teaching and Learning Librarian, other staff as designated
General page editor	Edit page content	Miscellaneous staff as designated

TABLE 3 • Site Governance and Permissions

FEATURES AND BENEFITS

Peter Griffiths outlines the following useful features for intranets: collaborative working, communities of interest, discussion groups, electronic forms, internal newsletters, internal trade, search facilities, and training materials (2000, 165–166). In redesigning this intranet, these features were kept in the forefront of the development plans and were incorporated wherever possible.

Real-time content creation, allowing for easy collaboration on the fly, is easily accommodated by the new intranet. Mandatory standardized fields for new intranet

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS
Amanda Etches-Johnson and Catherine Baird

pages, such as new projects or new committees, bring consistency to these staple categories of content. Small-scale ventures, such as a small weeding project, can just as easily be represented on the intranet as large-scale initiatives, such as the creation of a new integrated library system. Previously, information about smaller projects may not have been captured except on local hard drives and through e-mail exchanges among certain staff. As a result, there may have been little to no access to this information by other library staff. Now, those working on smaller projects can easily document their work on a single intranet page. Conversely, large-scale initiatives that use a project wiki for day-to-day workflow and documentation are often dense with information. Staff members who may not be involved with the project but require some basic piece of information about the project are able to avoid sifting through an entire wiki and access key project information from the intranet's project page, from which a link to the project wiki is provided.

Drupal's poll module has provided some much-needed fun and informality to the site. An initial poll was used to gauge the reaction to the new site, asking staff to vote on how much they liked or disliked the new design and organization, the results of which indicated the initial reaction was overwhelmingly positive. Subsequent polls allowed staff to vote on their favorite campus eatery, summer vacation activities, and new TV shows.

As indicated in Table 2, the comment module in Drupal was implemented to allow staff members to post comments on any page. Allowing comments on the site has proven to be very useful, particularly after the launch of the new site. The staff use comments to request changes to pages and suggest improvements. Since site administrators are notified of new comments, the ability exists to respond to comments quickly and make iterative changes on the site. More recently, staff members have been observed addressing problems themselves by responding to each other using comments rather than waiting for a response from site administrators.

Finally, one of the single biggest feature enhancements to the intranet is the addition of search functionality. Whereas on the previous staff intranet library, staff were forced to browse and stumble through page after page to find the information they were looking for, the new site has a search box on every page. Usability testing has proven that many staff members use the search function when looking for information, some choosing to use it exclusively over browsing. As usability best practices have taught, it is crucial to provide users with multiple avenues to access the information they are seeking, and the addition of a search box on all pages fulfils that principle on the new intranet.

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

Excerpted from *Designing and Developing Library Intranets*

CORRALLING WEB 2.0 CHANNELS

Two years prior to the intranet redesign, McMaster University Library staff took part in a very successful Learning 2.0/23 Things (http://macetg.blog.lib.mcmaster.ca/about-learning-20-mac/) staff program. The success of the program, coupled with the launch of internally hosted blog and wiki software (the highly dysfunctional intranet at the time), led staff to enthusiastically create blogs and wikis to organize projects and commit- tees and handle internal communication within service areas (and effectively circumvent the existing intranet). These tools have satisfied their purposes very well, and while the goal of the intranet redesign was not to replace them, the intention was to use the new site as a single jumping-off point for all internal library communication.

In practical terms, this means that all library departments, committees, and projects have at least a single page on the new intranet, with links to   all other relevant documentation. For example, a library committee has a committee page on the new intranet and is listed with all other existing committees. If the committee uses a blog or wiki to record meeting minutes or share documentation, then the committee intranet page simply links to that space. This structure allows committee chairs to continue using blogs and wikis for communication/documentation purposes, but it also enables library staff members, who might not be familiar with a particular committee, to easily access all committee information in a central location.

In addition to bringing together the blog and wiki channels already in use, the library intranet also needed to provide access to the staff newsletter, which was converted to a blog at the same time as the intranet redesign. Drupal's aggregator module was used to syndicate an RSS feed of the staff newsletter blog posts and display the most recent headlines on the main page, making the intranet the number-one referring site to the staff newsletter blog. This crucial connection fulfils two important requirements: providing dynamic content on the intranet's homepage and connecting the staff newsletter to the intranet, which allows staff to easily discover new newsletter posts instead of forcing them to bookmark the newsletter site and visit the site for updates.

To further satisfy the dynamic content requirement, a slideshow of recent images added to the library's Flickr account was included on the intranet homepage. As one of many social media channels in use at the library, the redesigned intranet was once again chosen as a hub for our social media activities rather than to supplant those activities. In addition to providing the homepage with some much-needed visual interest, the image stream also allows library staff members to easily view images of

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS
Amanda Etches-Johnson and Catherine Baird

recent library events and staff celebrations rather than forcing them to visit our Flickr page to view this content, as was the case prior to the redesign.

THE BIG REVEAL: ROLL OUT, TRAINING, AND POST-LAUNCH EVALUATION

As with any major project, success often hinges on good communication. In an effort to keep staff informed about the redesign throughout the project, existing communication channels, such as e-mail and all-staff meetings, were used to remind staff of the project at regular intervals. A few weeks prior to the official launch, staff were invited to participate in one of six hour-long training sessions that were designed to familiarize them with the new Web site prior to flipping the switch on the new intranet and removing access to the old one.

The training sessions were customized to the groups in attendance. For example, those staff who were likely to only view the new site and not be involved with behind-the-scenes changes and updates were given training on logging in, changing log-in passwords, a general orientation to the new site, and basic information about editing permissions on the new Web site so they would know the procedure for requesting a change or update. Other groups of staff who would be involved with editing and updating were also trained on attaching files, creating new content, and adding and deleting related links. Step-by-step documentation of the training was recorded in a wiki, including annotated screenshots and written instructions, so staff would be able to refer back to directions on how to perform various tasks.

A few months after the initial launch, a second Web-based staff survey was conducted, modelled closely on the first survey about the previous intranet. The following questions were asked:

- Why do you visit the Libstaff Web site?

- What do you like about the Libstaff Web site?

- What don't you like about the Libstaff Web site?

- What other features would you like to see on Libstaff?

- Do you have any other comments you would like to share about Libstaff?

The same response rate as the first survey was maintained, with 20 percent of staff responding. Results included the following insights:

- Staff continue to visit the site for information on committees and projects, human resources needs, to fill out forms, and for policies and procedures and library news.

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

- Staff reported liking the site's ease of use, new design, easy home access, security, access to the staff newsletter, and good content and new items such as polls, comments, photos, and a search function.

- Staff did not like that they still had problems finding information, although, in contrast to the last survey, no one said the site lacked organization. Some did not like the new layout, the labels, or the multiple navigation areas; some did not like the default font size and colors.

- One outstanding feature staff requested but was not rolled into the redesign was access to documents on shared staff drives, a feature currently under investigation for phase two of the redesign.

In response to lingering findability issues, a round of usability testing was conducted. Eight staff members from departments across the library were observed while they attempted to complete nine tasks on the redesigned site. Findings from the usability testing include some accessibility and design issues and some confusion over labels in use in global navigation. While changes are made to the site based on feedback and test results, it is planned that a 90-day cycle of conducting usability tests and making iterative changes be maintained.

CONTEMPLATING THE TAKEAWAYS: LESSONS LEARNED FROM THE REDESIGN

As any good project management text will say, it is crucial to plan the transition from project to maintenance. In this case, it was learned that the intranet cannot simply be redesigned as a discrete project and then left to sort itself out. Even with distributed editing, the intranet needs central ownership to realize complete adoption by individuals in the organization and become an invaluable tool. Since Drupal allows for distributed editing and content management, the project coordinators naively underestimated the amount of post-launch governance necessary for the new staff intranet.

Staff feedback also indicated that the importance of providing a unified experience was overlooked. Redesigning an entire Web site, implementing a CMS, and revising all content is a large project with a broad mandate. In this single-minded effort to improve the intranet, there was a failure to scope out other tools that affect library staff's ability to carry out their day-to-day tasks, including internal file sharing drives, communication tools (e-mail, IM), meeting scheduling software, and external document sharing tools (like Google Docs). Plans for phase two of the redesign include these tools within scope to provide a more unified experience for library staff.

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

Finally, despite everything that is known about the importance of resisting the urge to design an intranet around an organization's structure, there was a frequent temptation to turn to the institution's organizational chart to inform navigation and labelling decisions. Ongoing user testing and feed- back from staff helps the project's coordinators to resist this temptation, but it serves as a repeated lesson on a daily basis.

CONCLUSION

While phase two of the intranet redesign project aims to target some of the outstanding issues described, staff feedback and testing have already proven that the new site is a vast improvement. Content is more findable, thanks to good indexing, site search, and consistent navigation; staff workflows have been streamlined as a result of easier site access and simplified content creation and maintenance; the scalability of the site architecture ensures the intranet will grow and evolve more comfortably with the organization; and the integration of the library's existing 2.0 channels provides a more seamless experience for library staff. Thinking back to fictional librarian "Michelle" and the hurdles and roadblocks she faced with the former site, it is certain she would have a much better experience with the library's intranet today.

CORRALLING WEB 2.0
BUILDING AN INTRANET THAT ENABLES INDIVIDUALS

Amanda Etches-Johnson and Catherine Baird

REFERENCES

Griffiths, Peter. 2000. Your intranet. In Managing your internet and intranet services: The information and library professional's guide to strategy, 161–173. London: Library Association.

Guenther, Kim. 2003. Ten steps to intranet success. Online 27(1): 66–69. Head, Alison J. 2000. Demystifying intranet design. Online 24(4): 36–42.

Kennedy, Mary Lee. 2007. Introduction. In Intranets for info pros, ed. Mary Lee Kennedy and Jane Dysart, 1–13. Medford, NJ: Information Today.

FOOL'S GOLD

WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

This chapter is excerpted from

Googlization of Libraries

Edited by William Miller and Rita Pellen

LEARN MORE ›

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

Naysayers are a peculiar fish, always trying to swim upstream, go against current trends. Look at salmon, for instance. Upstream for all their lives and then what: shotten herring as it were. Surely those who presume to swim upstream should take that morality play for what it is: it may seem necessary at the time but it sure ends badly. Naysayers are a lot like scryers who complain that the proverbial crystal ball everyone else sees so clearly isn't that clear at all. What is it that George Eliot said? "Among all the forms of mistake," she wrote, "prophecy is the most gratuitous."[1] Shouldn't restraint be the order of the day, especially since so many naysayers have been wrong?

You'd think. But in the case of the Web, someone has to speak out, has to say aloud that the emperor doesn't have on any clothes. Too many people–especially many intelligent people–keep repeating that same inane bromide that libraries are obsolete, or are being made so by the Web, the Internet, the "Net."[2] This is particularly bedeviling, especially when so many librarians nod in agreement, but the library-is-dead- crowd is everywhere.

It's rare to open a professional library magazine and not find an article on why library search engines should be more like Google, or why the entire library shouldn't imitate the Web in some significant manner. For all other magazines, from business ones to ones on theology or the humanities, it's Google über alles. To hear the claque rehearse their lines, you'd think the Web had cured the common cold. I say enough al- ready. Google will not replace anything, and the Web will not make libraries obsolete. Both are useful and important but neither is a panacea for anything, least of all research. Although this belief may place me in the company of the few, perhaps only the handful, it's company I prefer to keep.

Herewith are reasons why the Internet will not replace the library. But before going too far into that discussion, two disclaimers must be lodged. I have used the Web extensively since 1992, and even created Web pages when doing so required knowing elementary code. I have used the Web extensively to write the articles and books I've written, including this one. So my assertions about the Internet vis-à-vis the library do not qualify me as either a Luddite or a hypocrite.

The Luddite charge is one I know I cannot avoid but I'll make the case now that I know a fair amount about the Web and have used it enough to know both its merits and its defects (yes, there are some). I would never claim to be an expert on the Web, but as one who uses it daily, and has since its early, clunky inception, I know enough about it at least to know whereof I speak. I also know that we Americans love our technology almost as a much as we love our football (both have become ersatz spiritual obsessions). And, just like football, I know if you criticize technology you'll

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

like be branded as an outcast, even a nutcase, for declaiming against the semi-official religion. Besides, wasn't the Unabomber one of the early critics of technology? Who wants to be thrown into that class!?

I, for one, don't, but someone must step forward to stop all this codswallop about the Web replacing libraries, or Google making them obsolete. Someone had to say something, so I did, first in an article and then a poster. But criticizing the Web is a dangerous area and one where I do not tread lightly. One respected librarian at a major research institution in the northeast saw my poster and wrote to tell me that the library "was nothing until the Web."[3] The assertion proved an astonishing one to me, especially when one considers that we've been without the Web for about 99.9% of our civilization! Alas, it's not the first such comment I've heard from colleagues. Anyway, the point is, I'm not a Luddite, though I'm pretty sure some who read this article will dismiss me after the first sentence as one because I dared criticize the God of Google or the Tao of Technology.

Although all those who criticize the Web are "straightway handled with a chain," not all of them should be. The coiner of the phrase, "paper- less society," F. W. Lancaster, has come forward regretting saying that for many reasons, not the least of which is the one I contend with here: it has not made libraries obsolete.[4] Many doubts have now surfaced about the so-called social networking of MySpace.com and Facebook.com, which are discussed later. Thomas Friedman, author of the rightly famous The World Is Flat, now opines that he gets into a cab and he and the cabbies do not speak for miles, both of them multitasking like madmen.[5] Perhaps the world isn't really flat so much as we're just all flat-headed when it comes to the "wow" of the Web. Now comes word from the creator of the Web itself, not Al Gore, but the real inventor, Sir Timothy Berners-Lee, wondering aloud what HTML hath wrought.[6] Apparently, this grand experiment in democracy turns out to be a failed experiment in kakistocracy. So, perhaps I'm in pretty good company after all.

As for the hypocrite charge, it runs something like this. If you cite from the Web and use it, how can you criticize it? But I view this charge much as I would view a charge of a similar nature from one who, say, criticized traffic deaths and yet drove to work every day. Or take another example. Is a person who criticizes healthcare but holds health insurance and visits his physician annually a hypocrite? Would anyone be willing to call Hillary Clinton who held (and holds) possibly the best health insurance in the world as the First Lady (and now as a Senator) a hypocrite because she spearheaded a group to dismantle that healthcare while her husband served as

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY
Mark Y. Herring

President? Regardless of whether one thought her suggestions excellent, or just socialized medicine, I don't think one could say that because she used healthcare but also criticized it, she's hypocrite. Just because I cite Web-based articles here and else- where while claiming that it is a most inferior library-like substitute doesn't qualify me as a hypocrite either.

Okay then, so why is the Web no substitute for a library? Why do I level the charge that the Web is the new fool's gold?

FORGET THE NEEDLE! CAN YOU JUST TELL ME WHICH HAYSTACK?

We've all experienced Web-euphoria. You input a search term into Google (or Lycos, Hotbot, Yahoo) and get back 1,234,456,345 hits in 3.652 seconds. Wow! That's what I'm talking about! How can this not be better than a dumb old library where you have to think and actually look for your answers? Then you begin to scroll through the first ten hits, then the second ten, then the third ten and slowly your euphoria turns first to disappointment then to outright indignation. How dare the Web trick you like that!

All of this assumes you really look past the first ten hits. I teach search classes where I work and for the last seven years I've asked students (over 300 so far) how many go past the first screen. To date, only one has admitted to it, a non-traditional student over the age of forty. Now this is very anecdotal and unscientific, but also very descriptive. While the Web is great at supplying discrete data–Bogotá, 3.14, 5,280 feet–it isn't very good at providing you the right bit of knowledge. Be- cause most search engines use relevancy ranking, the first ten or twenty or one hundred hits may or may not be what you want. Moreover, those same returns are not always the best: the Web is jammed-packed with misinformation, disinformation, fraud, and more.

Many examples abound. In 2002, the highly regarded Associated Press reported that PETA (People for the Ethical Treatment of Animals) had outfitted more than 400 deer in orange vests so hunters would mistake them for other hunters and not shoot them.[7] Guy Lockey, a sporting goods store owner offered a reward for each vested deer bagged. Al- though the story is just crazy enough to fit the usual PETA modus operandi, trapping a wild deer and fitting it with a vest is next to impossible. Moreover, Guy Lockey is a fictitious name. But if it came off the web, it had to be true, right?

Then there is the well-known but tragic one that occurred at Johns Hopkins and represents the limits of the Web, falling under the heading of what you don't know can hurt, perhaps even kill you. Ellen Roche, a seemingly healthy 24-year-old, volunteered

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

for an asthma study at the well-known and highly regarded institute. After entering the study, she was given a drug protocol that had been vetted, or so it seemed, by the researchers. She inhaled a chemical treatment, hexamethonium, which led to the progressive failure of her lungs and her kidneys, and her eventual death.[8] This sad case is made all the more tragic because it could easily have been averted. The supervising physician, Dr. Alkis Togias, made what appeared to be a thorough search of the literature. He approved the drug as did the ethics panel who reviewed his research methodology. But Dr. Togias relied heavily on PubMed, an electronic source that only goes back to 1966. A traditional print search would have found articles published in the 1950s that warned against using this drug in such experiments for the very reason that Ms. Roche died: rapid, progressive lung and kidney failure. Togias relied on PubMed and even did a Google search on hexamethonium. But his search did not reveal the inherent danger with his drug protocol, though the warning existed in the citations section of PoisonIndex Toxicologic Management. These citations are on the Web but are not easily found by inexperienced searchers.

Because the Web is so easy to use as a publishing medium, because there is no gate-keeping, because there are no fact-checkers, misinformation is not only inevitable but also abundant. Some may argue that these misprisions are small and on balance unworthy of our attention. But in a less Web-based delivery of information, bad or incorrect information would never make it through all the hoops before being found as bogus. The Web's instant gratification attribute makes certain many bogus claims will be made public long before we learn that they are, in fact, specious.

Space does not allow a longer rehearsal of the millions of business scams, health frauds, and more. In the case of the latter, physicians are particularly worried because patients come into their offices having diagnosed their illness. These patients are hard to convince they are wrong. Of course there are excellent sites on the Web, such as WebMD, but there are also thousands of others that are not only wrong (such as the coughing-to-prevent-a-heart-attack), not only stupid, but also dangerous.

WEARE18.COM

We've always been a nation awash in pornography but now, because of the Web, we're drowning in an ocean of smut. In my own lifetime spanning barely a half century, pornography has gone from its once in- famous brown paper wrappers to today's most revered and protected forms of "free speech." Sadly, the American Library Association (along with the usual suspects like the ACLU and First

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

Amendment absolutists) have gone to great lengths to protect everyone's right to view any kind of intercourse with any kind of partner (human or not) anywhere. In its infancy, television did not hint at sexual misconduct (most married couples slept in separate beds) and mainstream movies had to pass the rigorous Hayes Commission standards. Today, gratuitous sex is rampant in mainstream movies, even when those R and NC-17 movies lose millions.[9] Meanwhile, television sitcom characters snigger about everything from veiled sexual innuendo to blatant declarations of every kind of sex known (and some unknown) to humankind. Sub rosa storylines, even in commercials, titter like junior high-schoolers over erectile dysfunction or vaginitis. Oh, we have come a long way, baby! Sex is everywhere.

So why pick on the Web? Is Web-based pornography really that big a deal? Are we a nation awash in voyeurism because the Web has made it so readily available? If statistics are any indication, it would appear a forgone conclusion. The argument here isn't that we have only recently discovered pornography. As almost any book or article is quick to point out, pornography has been around quite a long time. Livy records that certain Roman emperors spent their leisure hours viewing or reading it. Diogenes of Sinope, that quintessential liberal figure, fought against the customs of his day by eschewing manners, dressing shabbily and masturbating in public just to make a point.[10] The argument here isn't that pornography is new but that its nonstop ubiquity is courtesy of the Web.[11]

About 25% of all websites are pornographic, a seemingly small amount; but its influence is far-reaching. The Web boasts of nearly 400 million pornographic pages. Nearly a third of all daily searches (close to 70 million) are for pornographic retrievals. About 10% of all e-mails are pornographic (nearly 3 billion). Almost 40% of all downloads are pornographic in nature (nearly 2 billion), and close to 120,000 daily requests to Gnutella, a file-sharing network, are of the unquestionable constitutionally illegal child pornography kind. More than 100,000 sites on the Web are of children in pornographic repose.[12] KaZaA, another file- sharing network, has been downloaded more than two hundred million times, with more than four million users sharing and searching files at any given time, many of which are pornographic in content.[13] We are a nation that does not discriminate, either, as about three-fourths of all men and a quarter of all women admit to visiting pornographic websites regularly. Not to put too fine a point on it, more than 50% of all men who are members of the religious group "Promise Keepers" (and have pledged to stay away from such things) drink from the pornographic cesspool often.

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

So, is pornography a problem, and is it specifically a Web problem? To coin a phrase a former President once made famous, it depends on what how you define "problem." Among our cahier de doleances against the Web, pornography, and all its twisted sisters, ranks chief among them. While the Web did not invent pornography, it certainly made its longevity and ubiquity a certainty. Child predators have found a safe haven in the Web. Clearly the Web has made it more ubiquitous, more readily available, and far more likely to be part of everyone's Web existence, whether they like it or not.

This rich, supposedly unprecedented educational and information resource that will eventually replace libraries routinely delivers the worse that humankind can offer. It's important to remember, however, that pornography is not only available in millions of Web pages, but also in chat rooms, e-mails, alternative groups, FaceBook, MySpace, Friendster, and other so-called social communities. Moreover, the local child molester who might not have shown his face twenty years ago "shows" it in every community in America, thanks to the Web. The anonymity that the Web provides makes it easy to prey on the young and unsuspecting without the risk of capture, thus emboldening those who might other- wise not take the risk.[14] MySpace, especially, has become a haven for pedophiles and porn.[15]

In fact MySpace has become so bad ("skanky" as the New York Times puts it) that even corporate sponsors such as Weight Watchers and T-mobile have already pulled their ads or are threatening to do so. Some of these social community users are 14 years of age or even younger with such graphic profiles they cannot be repeated here. Some of the "more presentable" profiles include young girls "practicing to be porn stars" or the sharing of pictures of the anything-goes porn star Tera Patrick, and others. If you think pedophiles have missed this opportunity, think again. They are trolling for the wary and the not-so-wary.

The point is, the Web has made pornography ubiquitous. While the American Library Association opines against filters, it has yet to say one word about fellatio, cunnilingus and more that can be found on the Web by any eight-year old with Web access. A few years ago we shut down the apple juice industry because if one ate 70 alar-treated apples every day for 70 years, cancer might ensue. We shut down the apple juice industry "for the children." But the Web can produce both known and unknown sex acts in Technicolor and we all defend it as the newest and best library in the world. Readers need to remember that only fifteen years ago, none of this was available because the Web had not been invented yet. Moreover, none of this was available in any library. We are living proof of what one wag said: "As information doubles, knowledge halves and wisdom quarters."

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

FOOTNOTES? WHO NEEDS THEM?

"You may observe," said Sir Arthur in Sir Walter Scott's The Antiquary, "that he never has any advantage of me in dispute, unless when he avails himself of a sort of pettifogging intimacy with dates, names, and trifling matters of fact–a tiresome and frivolous accuracy of memory, which is entirely owing to his mechanical descent."

"He must find it convenient in historical investigation, I should think, sir?" said the young lady.[16]

Ah yes, that "pettifogging intimacy with dates"! In our politically correct, let's-not-make-anyone-memorize-anything school of thought, those who are intimate with dates, names, facts and so on are constantly having to defend themselves against the modern-day know-nothings who wish everyone were as little informed as the least informed among us. The "snatch and grab" mentality of the Web encourages all of us to let something else do everything for you, including the thinking part. Part of our brave, new Web-world is to do everything in 3.25 seconds; we no longer have time for silly things like footnotes.

The Web, per se, is guilty of missing, omitting, or ignoring footnotes altogether. But this complaint is of a far more serious kind. While the dotcoms and the dotnets are certainly the most egregious offenders when it comes to vanishing footnotes, even the electronic proprietary databases–databases like Lexis-Nexis, Academic Search Premier, Infoseek and so on–offend insensibly and unnecessarily. Not only must librarians worry about the retention of electronic titles in aggregate databases, but also fret over whether to keep both print and electronic collections, for the sake of footnotes. Now, while wringing hands over what to do about lost titles within database collections, they now must also worry about a new phenomenon: linkrot.[17] Linkrot, or what some refer to as the half-life phenomenon, has become an increasing worry since the late nineties when it was first noticed. The Web had not been around long enough before then to notice the growing problem. Moreover, no one perhaps gave it any real thought.

In a study of some 416 citations over four years, researchers Bugeja and Dimitrova found that only 61% of them were still accessible.[18] Moreover, 19% contained an error in the URL while 63% did not cite the date they were accessed. This is an unprecedented turn of events, especially when one considers that only a decade and half ago such careless scholarship would have been unthinkable, not to mention unacceptable. While it is not a matter that is slipping by unnoticed, it is worse: it is slipping by with the full knowledge of many scholars.[19]

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

One might excuse it if the journals in question were what might be considered (it is silly to say it) in the academic world as "back-ups" or "second string" sources, assuming there is such a thing. But the sad state of affairs is that the journals in question are among the most prestigious in a specified discipline. Given that academe is a discipline wherein we teach others the importance of scholarly, honesty, and integrity, it is most odd to find that we, too, when pushed by the Web, simply accept the state of digital life in which we now find ourselves. We accept, like so many braying sheep, that this is the state of progressive affairs, and so we must learn to accept and adapt to it. Many already have. Baby-Boomers once chanted that no one would push them around; GenXers pride themselves on being independent, solitary. Yet when it comes to Web-based information, both groups take their marching orders from the Internet regardless of how inferior it may be.

But the linkrot story worsens. When researchers studied citations moving from .pdf to html formatting, a 17% failure was noted.[20] This means, at its simplest level, that not only it is likely that citations only four years old may not only not be on the Web at all (there is 39% chance they will be gone), but that common formatting changes will increase the likelihood of disappearing citations. What does this mean for Google's grand plans to digitize some 15 million volumes? Will these volumes (assuming–and this is a grand assumption–that Google can surmount the copyright challenges) be complete texts, or texts with vanishing footnotes? Yes, of course, these will be mere images of pages, or so we are assured. But is the mass digitization plan accounting for changes in the Web over time so that these materials are here year after year, decade after decade? Moreover, it is common knowledge (or at least it once was) that re-mastering of digitized texts must take place at least twice a decade.[21] If foot- notes are vanishing, and original texts are lost or discarded, what about the future re-digitizing of these texts? Will it be possible to create a stable text, much less a definitive one? And while the subject of definitive texts is on the table for discussion, what about them? With this state of affairs can any digitized text be considered definitive?

Electronic law journals and medical journals are also coming to the foreground as linkrot-infested information resources. "Over the past de- cade," writes one researcher, "the use of Internet citations in the foot- notes of law review articles has grown from a trickle to a flood. But it is well documented that Uniform Resource Locators (URLs) experience linkrot, that is, over time, the URL is more and more likely to become a dead link, making footnote citations worthless or nearly so."[22] In a discipline (law) where footnotes are, in large part, the "meat" of the article, this becomes much more a

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY
Mark Y. Herring

serious matter than just a handful of journals in an esoteric discipline. According to Lyons, there were four instances of Web citations in three law reviews in the early nineties. By 2003 there were more than 96,000.[23] This would not be a concern if these citations could be relied upon for a reasonable period of time but they cannot.

According to Lyons and others, within one year, almost 18% of Websites and nearly 32% of Webpages had disappeared. In one study cited by Lyons of 31 academic titles, almost half the links were dead at the end of a three-year period. In a study by JAMA, also cited by Lyons, of 515 scientific links, nearly 28% were useless after only twenty-four months.[24] What is perhaps most disconcerting is that law review journals have a standard–the Bluebook (Rule 18.2)–that specifically discourages the use of Internet-only citations. While proprietary databases do fare better, they are not immune to linkrot. This is because even proprietary databases like Lexis-Nexis, InfoSeek, and Academic Source Premier do not maintain the same core journals from year to year. Sometimes this occurs because publishers of those journals pull them. At other times, the vendor decides, for whatever reason, to replace a journal with another more frequently used and/or cited one. And so the linkrot saga continues, in law, in medicine, in academic and science journals.

Some will mock this as so much ado about nothing and will quote, approvingly, Noel Coward's assertion that looking up footnotes is like answering the doorbell while making love. While humorous, Coward's assertion is surely wrong for those who wish to substantiate a claim, or to look back at a claim to see from whence it came. Oddly, there was an Ophrah-sized hand-wringing over James Frey's fictional autobiography. Frey's book was eventually withdrawn. But let the Web sell you a pig in a poke and we all oink approvingly.

GOOGLE ÜBER ALLES

Google's grand plan to digitize 10+ million unique titles from the so-called G7: University of Michigan, Harvard University, Stanford University, Oxford University, the University of California, the University of Virginia, and the New York Public Library, seemed like a good, even a great idea. These 10+ million unique titles were most likely not being read regularly, few people had access to them, and no one had access to them all. What could possibly be wrong about it? When the news was announced in 2004, almost everyone jumped on board. The giant Google with its vast resources would undertake the project, the now G7 libraries (more by the time this article appears) would partner with Google, little read books would be showcased in

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY
Mark Y. Herring

those 100+ million daily hits, and everyone would be served. Or so we thought. Digital libraries do not appear all that difficult to build, do not require very many employees (easily the largest expenditure item in conventional libraries), and are, we are told, able to adapt quickly to market and technological demands. In the twilight of the new digital day, everything appeared possible. Then the radar screen began blipping with numerous intruders threatening to spoil our mad rush to the mass digitization party. But the trouble began even long before Google's plan. The symptoms were there, but few wanted to acknowledge them.

Like a frat party that begins well enough only to end in too much drinking, the next day hangovers ensued. Some began to question how all of this would be done. Others asked about copyright. The French asked why the titles Google chose to digitize were all English ones. A few librarians began thinking seriously about the implications of a giant Google library. Some scholars wondered about how Google would be reimbursed (would Hamlet's soliloquy appear alongside an advertisement for Viagra?). Not a few publishers asked why they were not consulted, and by late 2006, the Google plan began to flag.

It's not likely it will fall to half-staff forever. Google has always prided itself on "built-in innovation" and so will surely find ways to make some- thing on this order (though, perhaps, not of this magnitude) eventuate. Still, it brings to mind an important question: is the rush to mass digitization the right approach right now for scarce library dollars?

Obviously my answer is in the negative and here's why. We're rushing vast resources in this direction with no real standards, no real plans other than we will digitize these millions of books and then … what? We seem to think that once in cyberspace everyone will have them. But this is patently untrue. Only those will have them, assuming Google surmounts the copyright challenges, who also have high speed Internet ac- cess, a stable Internet Provider, good, regularly updated equipment, the ability to download, and unlimited free printing (where's a tree-hugger when you really need one?).[25]

We are nearly always given the scenario of the poor, underserved soul in the darkest of Dafur who will now have access to millions of titles. Really? Is this right before or right after he has died of malnutrition, scurvy, pox, and a million other hunger-deprived diseases? The fact of the matter is only in the U.S. could this be assumed, and even here, of, at most, only about 50% of the population. For example, the only way that I am able to navigate the Web with great ease is because I work for an institution that spends, literally, hundreds of thousands of dollars on access annually. If left to my own paltry pocketbook, I would not be accessing the Web every week, much less every hour of every day.

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

But let's assume that Google really is the new Guru when it comes to web-based information. What of those other tens of thousands of lesser (much lesser) digitization projects? Are those being held to high standards, are there funds for re-digitization, are they guilty of "cherry picking" their collections (picking the best and leaving the rest for the dustbin), and is anyone using digitization as a replacement for preservation? In the case of the latter, using digitization as a replacement for preservation, there is real concern. Digitization cannot, at least for now, ever be used as a substitute. Yet the problem remains that many are using it as such even though they know it's a bad replacement. Why? Because that's where the funding is. Years from now we'll know what Google et al. hath wrought in this mass hysteria but by then it may be too late to rectify any askance matters.

EBOOKS TO THE RESCUE?

Surely eBooks will come to our rescue in this grand rush to all things Web-based. Sure, print books have been easy to handle, come in all shapes and sizes, can generally be carried everywhere and anywhere, but it's only a matter of time before eBooks replace them. Surely eBooks will prove all my assertions about the Web wrong!

The inevitability of eBooks is a safe bet because so much time, effort, and dollars have already been sunk into them although so little has been seen in return. eBooks have, in one shape or another, been in the reading conversation for more than 15 years.[26] In all that time, we are still far away from a consuetudinary acceptance of them. But their inevitability is secure because printed books are the new bête noir.[27]

eBooks are today still on the evolutionary chain's first link, at the point of the once two-shoebox-size mobile phone: clunky, cumbersome, frustrating, yet with some promise. What makes them less hopeful than mobile-phones-turned-ubiquitous-cell-phones is that they are not anywhere like an improvement over what they purport to replace. And therein lies the problem.

So far, all we have are bad imitations of print books, not grand improvements over an "outdated" medium. Formats are not interchangeable. eBook readers are not interchangeable and again, when one goes to invest, the initial start-ups are steep, beginning at about $400 and ending around $2,000. While the texts themselves may be cheaper than print counterparts (but not by much), there is still the problem of the reader.[28] But this is only the beginning. They also change the reading experience whether one reads them on a reader, or downloads an HTML text for reading on a

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

computer. The argument is always that these changes enhance the reading experience. In fact, they radically change it.

For example, a text will be embedded with hyperlinks, colors, perhaps even a digital cicerone who guides the reader through the text. Whatever else one may say about these added electronic ingredients, they make the reading experience a spectator sport, not an intellectual encounter. The reader is led through the text; he/she does not read it. Rather than allowing the reader to think for himself/herself, the reader is led on a wild goose-chase to a myriad of other "helps" that serve only to exhaust (if they do not first confuse) the reader. Even if we suppose a strict e-text that precisely duplicates a printed text, the reader is still struggling through the new physiological experience of reading with a light in her eyes, not over her shoulder.[29] We may be able to "evolve" away from this process but it will not happen over the next 100 years. Even after 15 years, eBooks still only hold 5% of the reading population. Moreover, in Web-like "snatch and grab" fashion, "reading" is measured on eBooks in minutes, not in hours. So far we do not have anything approximating cover-to-cover reading. And yes, while this may change, surely after 15 years we might have something more than a whisper warehoused as a shout? When eBook users are not complaining about the readers, they are complaining about the lack of content.[30] Again, is this really a good way to use up very scarce library funds?

THE PAPERLESS REVOLUTION IS COMPLETE!

The din of grand claims about the Web often drowns out some of its inglorious failures. None is more colossal or more inglorious than the paperless society schlepped for the last thirty years as the coming brave, new world. In fact, everything about the Web, the Prime Mover of the paperless world, screams "Ultimate Significance" when we cannot possibly know that at this early stage. "The Internet is too young," writes Robert J. Samuelson, "for anyone to foretell its ultimate significance.... But some present claims aren't true."[31] Samuelson goes on to point out that it did not spread faster than any other innovation, that it is a work in progress, and that many other innovations have far outstripped it in seismic afterclap (take indoor plumbing, for example, or electricity). We forget that we are, as Robert Burton reminds us in the Anatomy of Melancholy, dwarves sitting on the shoulders of giants.[32] We moderns hate the thought that, not only did many things and many greater people come before us, but also that they invented many things far better than we have, including the Web.

Over the last twenty-five years, automation's prognosticators made increasingly bold claims about the coming paperless age. Offices, businesses, industries, libraries, and

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

Excerpted from *Googlization of Libraries*

universities would eventually all be- come paperless and the New Jerusalem would be ours at last. Some made claims by redefining the library as no longer a single entity but "a range of services" that would be (somehow) seamlessly connected and would reach beyond any one campus or laboratory.[33] Others suggested that the library would be freed, as if once a slave, from paper and occupy "infinite space" and be "interconnected in a transparent way."[34] (All predictions must apparently include either the word "seamless" or "transparent.") The library as a place would disappear; it would be replaced by a network of every imaginable file or database, there would be no walls, buildings would vanish, communities of every size would provide its citizens laptops, and electronic files would never, ever be lost.[35] Of course, there would have to be paradigm shifts for where would we be without them? Librarians and users would have to learn to accept those changeling paradigms or be lost in the dust of the parabolic arch, or some such matrix-like vacuum.

A funny thing happened on the way to the new paperless world: society didn't get the memo, maybe because it was e-mailed and ended up deleted as spam. Somewhere along the way the New Paperless Jerusalem ended in the pages of Samuel Butler's old Erewhon. The only paperless society so far successful occurs in the heady environment inhabited by George Jetson, his wife Jane, his daughter Judy, his son Elroy—well, you get the picture. The paperless society expired even before that new world began. We are no closer to the paperless world than we were thirty-five years ago when it was touted as the next new thing. Surely this helps explain why in the last ten years almost no one has been talking about it but that has not stopped the forcible move in its direction. What is disturbing to some is the continued march, lockstep really, toward this goal, al- though most, if not all the predictions of the coming paperless society ended early, or simply failed outright.

"We are drowning in information," writes John Naisbitt, "but starved for knowledge."[36] Naisbitt hits upon the very reason for this embarrassing failure. Along the way to paper-freedom, someone forgot to distinguish between an answer and the right answer, between information and the right information, between information (raw data), and knowledge (raw data turned into something usable). In other words, having an ocean of fish doesn't mean you've caught anything yet, and weltering in information does not mean you're awash in wisdom. Paper has allowed us to tame the information revolution that threatens to undo us. No one seemed to notice that having an industry–computers–that could churn out with lightning quick speed a disorganized morass of data did not mean we would be any better off than when we had to distill it into wisdom as we once slowly ferreted it out on paper. Take, for

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

example, the new datum that 75% of all e-mails today are spam.[37] In the end, this grand new educational resource that's better than libraries threaten to undo us with "You May Be A Winner!" notices.

Why does so much of the world remain agog over the putative but little seen "vast" resources of the Web? Sure, eight billion pages (the web's current size) is vast. But when two-thirds or more are the equivalent of a crank phone call, "vast" suddenly seems small. Yet in our rush to make everything electronic, we are called upon to ignore most of these crank phone calls and stand amazed in the presence of emoticons, video clips of stupid animal tricks, Internet Solitaire, a cesspool of pornography–and all for the grand benefit of getting tomorrow's weather forecast today, or being able to e-mail the latest puns to your best friend ... who happens to work down the hall. But such talk will only land us in the Office of Luddite Control. We must face up to the fact that the paperless society is winning even when it's losing. Sadly, many of its wins come largely from those in my own profession, librarianship.

The modern version of this "treason of the clerks" goes like this. Recently OCLC, the national database of library-owned materials, released a report that proves the treasonous clerks are not only winning battles, they may have already won the war.[38] Not all the news is inexorably bad. Students are using libraries, have library cards, and have used a library Web site.[39] But students go on to point out that library usage in the future will remain flat, that the Web is out-vying libraries for "information," and that search engines are the preferred (first and last) approach to re- search.[40] Moreover, almost 70% are more likely to begin (and, sadly end) with Google, regardless of the fact that vast sums of money are being spent by libraries on proprietary databases chock full of scholarly re- search.[41] The first choice of students (college students, mind you) is a search engine (Google, Yahoo, Ask.com, MSN Search, etc.)–72%–over a physical library (14%). In other words, they are choosing free search engines over multimillion dollar libraries five times as often.[42]

Some may be shaking their heads ruefully at this point. Poor fellow, the information highway built its right-of-way to zoom past his epigone library. He didn't adapt his "business" to the growing demands of the electronic world and his customers are leaving him.[43] This charge would resonate more loudly if the inventor of the phase, "the paperless society," F. W. Lancaster, were not himself having second thoughts.[44] Lancaster recalls when he coined the phrase at a conference in Finland as describing "a largely paperless, network-based communications system having many of the characteristics to today's Internet-based environment."[45] He saw the confluence of

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

events moving us quickly along to a time when most everything would grow out of this putative electronic medium. "As the transition actually occurred, however, I became less enthusiastic about the developments and implications and, in the past few years, downright hostile to them."[46]

Lancaster cites dehumanization as much to blame, arguing that, at least in the United States, it has replaced the human element. But he goes on to blame librarians because they have become "completely uncritical of information technologies."[47] While this is largely true, it cannot be left unsaid that those of us who have been critical have been branded as Luddites or worse.[48] He also goes on to point out "wild assertions" about technology that cannot possibly be true, such as providing access to in- formation anywhere, anytime, any place, because they assume if it is in a database, it can be found easily, when "nothing could be further from the truth."[49] He argues that many scholars and experienced researchers are happy with results found until they discovered how many important items they missed (Individual librarians, as well as a team of librarians, missed fewer important citations but more than they should have.).

If we proceed along these lines, searching for materials to answer pertinent questions of inquiry will be all too painfully familiar: "Press one, if your answer is yes; press two, if you need more help. Press three, if …" After a long line of such questions the poor user will be referred back to the menu, and the endless, useless loop will begin anew. And for what? For the paperless chase that leads to nowhere.

A MILE WIDE AND A MIND-NUMBING INCH DEEP

Throughout this article we have examined the hidden costs of the Web, such as its very content, or the lack thereof, and dispelled many myths regarding the size of the Web. We also addressed issues such as fraud, pornography, scams, hate sites, identity theft, and the vast numbers of young people providing TMI–too much information–about themselves and others in places like MySpace.com and Facebook.com, without respect to future consequences (when did a library ever do this to anyone, much less to young people?). While we have been quick to point out that anyone of these obvious or hidden costs would hardly be enough by themselves to question the Web's utility, cumulatively they do present a strong argument against the Web's stance as a possible substitute for a full-service library.

Reading comprehension now presents itself as a possible casualty of the Web.[50] We know for example that when we read on the Web we do not read in detail but in snatches (sometimes referred to as the "snatch and grab" process) and not in the

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

same manner as we read printed books, cover-to-cover. Yet very few are examining this problem in great detail with an eye to possible contraindications. Even so, preliminary studies do not provide promising results for the future. Indeed, if anything, they point to just the opposite, that Web-based "reading" contributes to an even more cursory reading comprehension, along with a decline in literacy. Very little is being done to point out how this problem is a negative consequence of making all our access to information digital. As with most of the problems we have encountered in connection to the Web, our solution appears to be that we need more of the Web, not less; and if it turns out to be a bad thing, well, it will somehow manage to right itself in the long run.

Reading proficiency scores for students have been declining for a number of years, and dramatically since 1990.[51] For example, reading proficiency for 9-year olds from 1971 to 1999 has remained almost static, rising from 208 to 212 (scores that are just above the barely literate rates) and then dropping only slightly and levelling off entirely. The 212 level, the highest one reached, is barely above the "partial skills and understanding" level, and far removed from "learns from specialized reading materials," a level we hope they will achieve before entering college. Although the National Assessment of Educational Progress (NEAP) has been conducting reading studies for a number of years, some would argue that the proficiency levels noted there (in the low 200s) may not even rise to the level of what many would call literate.

While these students may well be able to read, they still exhibit what is sometimes referred to as a "knowledge deficit." Such young people are unlikely to overcome this lost or weak ability once it is entrenched at the age of 9, the last age at which any positive gains in reading can be measured. It could be argued that the stagnant levels of this proficiency roughly correspond to the rise of computers in the home and in the school as merely coincidental, but it strikes some as more than that. If the skills are weak to begin with, and activities and interests pull readers possessing those weak skills into yet other reading proficiency-attenuating activities, strengthening those reading skills later will be highly unlikely.[52]

In more than twenty years, the skill could not be improved, even marginally (a period, it should be pointed out that mimics the rise of the computer age). It appears as if literacy simply stopped growing. Furthermore, while the levels increase somewhat at age 13 (in the middle 250s, barely a literate reading ability), at age 17 those skills remain too weak to understand complicated information. Again, reading levels at each age do not improve so that the level of reading proficiency at ages 13 and 17 in students in 1971 was equal to those measured in students more than twenty years later, in 1999.

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

In other words, today's students are no better in spite of all the money, time, energy, and programs (on- line and otherwise) we have applied to reading improvement.[53]

Reading proficiency should at least require at a minimum some basic literacy, something these figures do not show. In fact, by 2009 students in grades 4, 8, and 12 must be able "to read both literary (fiction, nonfiction, and poetry) and information texts (exposition, argumentation, persuasive and procedural texts or documents)" in order to declare one reading with proficiency.[54] Current scores indicate that we are light years from achieving that goal. Why would we add anything to the mix that weakens our chances of getting there?

In other words, the abysmal scores previously reported will only grow worse as standards tighten. The 2009 NEAP standards define reading as "an active and complex process that involves understanding written text; developing and interpreting meaning; and using meaning as appropriate to type of text, purpose and situation."[55] If students are encouraged to "snatch and grab," if they are losing vocabulary, if they are rapidly declining in reading interests the older they get (and they are), the Web should be the last place we focus their attention. It behoves us all the more to make certain we are not saddling students with a skill acquisition (the digital context) that will only deepen the pool of potential illiterates. The Web has us on a fast-track collision course with shared ignorance. It encourages spectators to knowledge, not active participants. Students view watching television or a computer screen as equivalent activities. Moreover, online material has also increased plagiarism to epidemic proportions.[56] Having more of the same will be unlikely to decrease these offenses. Moreover, today's students do no read these texts, but merely snatch and grab what they want, not the best of scholarly practices or a way to improve reading proficiency. Having millions more books to pilfer does not seem like the best way of attenuating our current and growing illiteracy problem. It is not that students today are so much less virtuous than students of former years, but that having such an arsenal at one's fingertips only encourages more snatching and grabbing, more reading in bits and pieces, and more cutting and pasting. It is too risible to think that students will read these texts online from digital-cover-to-digital-cover when all evidence to date indicates they are spending minutes with online books, not hours. We believe that students must be learning better because the "text" in question not only has words, but also has pictures, sounds, video clips, and more. We think that because students are forced to numerous other important and tangential concerns, the learning experience is thereby deeper and made more powerful. What we do not have to support these velleties are empirical data substantiating that these are

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY
Mark Y. Herring

transferable skills. We just assume it must be so, and that wishing will make it right. Even when we know that "many Web-based environments also introduce a new set of cognitive barriers that can cause competent readers of conventional text to be cognitively overloaded and frustrated, we damn those torpedoes and fund another round of mass digitization projects.[57]

Furthermore, a new study now indicates students cannot multitask. While they may well be "ultra-communicators," they cannot read, study, use the cell phone, text message, and listen to their iPods at the same time, though most try. We also know that they are rapidly growing allergic to text-based materials. This should alarm us. Whether we like it or not, literacy is still measured largely by the written (i.e., printed) word.[58] It does not seem either right or wise that we should encourage students to become independent of text-based learning. Moreover, we should make absolutely certain that any Web-based learning is equivalent to text-based learning before we send a generation of students to their academic peril. It's no secret that elementary children are fascinated by what they see on the Web, especially if pictures, video, and audio clips are present. But if these readers become too distracted by the technology, are they still learning?

The so-called transferable skills do not appear to be as transferable as we once thought. Web-based learning not only proves to be more difficult, but also capturing all the right elements–hardware, software, analysis of contents and so on–is very elusive. Even on college campuses we find that a variety of approaches meant to meet the so-called various kinds of learning are just as likely to be as bad as it is good. For every young person who feels emboldened to speak up in the anonymous context of Web-based or online environments, a dozen more are frustrated or annoyed by it. This should not come as any surprise to those who have tried to force audio-visual learning into the classroom. While it does not impede learning per se, many have found regardless of the approach used that such "teachable moments" turn quickly into disruptive ones and class management imbroglios.

We are progressing too rapidly down a descent from which we may never recover. When young people are poorly educated, they spend a lifetime regretting it. So, too, does the society which allows it to happen. As the "can-do" nation, we are slow to admit defeat about any invention but especially about one that appears as successful to us as the Web. However, we cannot turn a blind eye to its defects. Mary Shelly's Frankenstein began innocently enough. We need not let our modern-day Frankenstein, the Web, end with the same fate as hers. Now comes the news that many students are using the same moral code in instant messaging that they use in

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

formal research papers and cannot understand why it's unacceptable.

Finally, there is the Web content itself. No matter what one chooses to look at, what the Web proffers does not inspire us with its educational content. Whether one chooses the "best" of the Web from Yahoo! Google, or even Web-savvy experts, what young people and old rate as the best is far from the true educational content that even a small library proffers. For example, in October 2006, Google purchased YouTube, a site offering thousands of videos. Before the purchase, YouTube could boast hundreds of thousands of hits, apparently the only measure of a "good" Website. Today, the issue of copyright emerges again and threatens to constrain the site. Already numerous clips available in one month are the next month or the month after removed.

In the fall of 2006, I looked at the top 100 videos. The headers proved amusing, showing everything from "Young girl with large hooters" to a young man trying to set fire to his flatulence with a cigarette lighter, to some white guy trying to dance. This is the new educational site that threatens libraries?

Of course there are valuable videos on this site (though not much I could find in the top 100 that one would want to spend much time examining). When this site is coupled with the fact that 75% of e-mails are now spam, perhaps all our "work" on the Web is so much wheel-spinning. Sure, for advertisers seeking to reach millions of people, the Web is a bonanza. But should this be the work of a library? In November 2006, I heard a presentation at a library conference about Google Print. During the question and answer period, one participant in the audience asked about Google's plan to make money off all this. The representative said that Google "was not yet sure how they were going to monetize all this" but perhaps it might be something like the ads that appear so familiar now. Or, perhaps worse, the full text will be pay-per-view?

And this is just what troubles me and others. I'm certain I do not want Dante with an ad for Virago, or one for erectile dysfunction. Or even for clothes or anything in between. There may still be time to make the Web what it should be, a tool, like many other tools, that can aid and abet our pursuit of turning information into knowledge, and that knowledge into wisdom. But the present state of affairs puts us exactly light years from this goal. Are librarians paying any attention to these things? Is anyone? A few more years down this road and the question will no longer matter. We will have, not the future we want, but the future we allowed. We have arrived on the Information Superhighway, all right, but are we rushing all too fast to make libraries, and library services, that high- way's first roadkill?

FOOL'S GOLD

WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

NOTES

1. George Eliot. Middlemarch 1950 ed., London: Zodiac Press, 83.

2. Yes, yes, I know that the Web and the Internet are not synonymous terms though they have become so in common parlance. Rather that try to swim upstream entirely by keeping the two separate, I have chosen to use them interchangeably. The Internet is, of course, made up of protocols or rules that let computers "talk" to each other. The Web or the World Wide Web, on the other hand, is made up of software protocols that run on top of the Internet and that let users see and access files stored on other computers. For more on this see Chris Sherman and Gary Price. The Invisible Web: Uncovering Information Sources Search Engines Can't See. CyberAge Books., Medford, New Jersey, 2001, specifically their chapter 1. I also use Google as eponymous for the Web.

3. For the poster, see <www.winthrop.edu/dacus> and click on the "ten reasons link." All the proceeds go to the library's faculty development fund. Look on the bright side. If you disagree, buy a poster anyway and suggest that your dean go to a few technology conferences.

4. F. Wilfrid Lancaster. "Second Thoughts on the Paperless Society." Library Journal (September 15, 1999) Vol. 124 (15), 48. See also his "The Paperless Society Revisited." American Libraries (September 1985) Vol. 16 (8), 553-555.

5. Friedman, Thomas L. "The Taxi Driver." New York Times (Wednesday, November 1, 2006) LE, A23.

6. Ghosh, Pallab. (November 02, 2006) "Web inventor Fears for the Future." BBC News accessed via <http://news.bbc.co.uk/go/pr/fr/-/1/hi/technology/6108578.stm> on November 7, 2006.

7. Paul S. Piper. "Web Hoaxes, Counterfeit Sites, and Other Spurious Information on the Internet." In Web of Deception: Misinformation on the Internet. Anne P. Mintz, ed. Medford, NJ: CyberAge Books, 2002, 7-8. Fox News also reported the story in January 2002.

8. Eva Perkins. "Johns Hopkins's Tragedy: Could Librarians Have Prevented a Death?" Information Today, Inc. (August 7, 2001). Accessed via http: <www.infotoday.com/newsbreaks/nb010806-1.htm>, March 2006. The description which follows comes from this article. See also Susan Levine. "Hopkins Researcher Faulted in Death." The Washington Post (July 17, 2001) accessed via Lexis-Nexis February 2006.

9. Movie critic Michael Medved has discussed this oddity. G-rated movies do two and three times better at the box office than PG, PG-13, R, or NC-17 movies. Yet

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

producers make five or even six times as many of the PG, PG-13, R, and NC-17 movies than the G-rated ones that earn so much more money.

10. The story is repeated in Siva Vaidhyanathan. The Anarchist in the Library: How the Clash Between Freedom and Control Is Hacking the Real World and Crashing the System. New York, Basic Books, 2004, 25.

11. Some of the best books for discussing the philosophical underpinnings of pornography and its deleterious effects on all of society are as follows: Robert P. George. The Clash of Orthodoxies: Law, Religion, and Morality in Crisis. Wilmington, Del.: ISI Books, 2001; Pamela Paul. Pornified: How Pornography Is Transforming Our Lives, Our Relationships, and Our Families. New York: Times Books, 2005; Roger Shattuck. Forbidden Knowledge: From Prometheus to Pornography. San Diego: Harcourt Brace, 1997; United States House of Representatives. Committee on Energy and Commerce. Subcommittee on Commerce Trade, and Consumer Protection. Online Pornography: Closing the Door on Pervasive Smut: Hearing Before the Subcommittee on Commerce, Trade, and Consumer Protection of the Committee on Energy and Commerce, House of Representatives, One Hundred Eight Congress, Second Session. May 6, 2004. Washington: U.S. Government Printing Office, 2004. Paul's book is good because Paul is herself a First Amendment die-hard and Times reporter. Shattuck's thesis is that there are things that we can do that we should not and no better example exists than the widespread availably of pornography. That is, just because we can do some- thing doesn't mean we should.

12. Other sources indicate that in 2005, almost 75% of searches worldwide were for pornographic material.

13. House of Representatives. Stumbling onto Smut: The Alarming Ease of Access to Pornography on Peer-to-Peer Networks: Hearing Before the Committee on Government Reform. House of Representatives, 108th Congress, March 13, 2003, 1. KaZaA and Gnutella are two of many file sharing sites. Others include Morpheus, BearShare, and Grokster.

14. Dick Thornburgh and Herbert S. Lin, eds. Youth, Pornography and the Internet. National Academy Press, Washington, DC, 2002, ES-3.

15. Rebecca Hagelin. "Porn, Pedophiles, Our Kids and MySpace." Townhall.com. Available at <ww.townhall.com/opinion/columns/ rebeccahagelin/2006/05/30/199114. html>. Accessed June 2006.

16. Sir Walter Scott. The Antiquary, New York: Macmillan & Co., 1895, 43.

17. For the so-called hybrid collections, see Mark Rowse. "The Hybrid Environment: Electronic-Only versus Print Retention." Against the Grain (April 2003), 24-28.

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

The problem of lost titles is this: an aggregate database can contain 100 titles in September and 100 in May. The trouble is they are not always the same 100 titles. Many libraries come to know about these lost or vanishing titles only after patrons come   to them complaining that what they saw last month isn't there this month. For more, see the book, Fool's Gold: Why the Internet Is No Substitute for a Library. McFarland, 2007.

18. Michael Bugeja and Daniela V. Dimitrova. "The Half-Life Phenomenon: Eroding Citations in Journals." The Serials Librarian Vol. 49 (3), 117.

19. Scott Carlson. "Scholars Note the Decay of Citation to Online References." The Chronicle of Higher Education. (March 18, 2005), A30. Also cited in Bujega and Dimitrova.

20. Bugeja and Dimitrova, 121-122.

21. Some experts claim re-mastering is required every three years.

22. Susan Lyons. "Persistent Identification of Electronic Documents and the Future of Footnotes." Law Library Journal (Fall 2005) Vol. 97 (4), 681. Accessed via Lexis-Nexis, viewed, May 2006.

23. Ibid., 681.

24. Ibid., 683.

25. No one but no one reads anything on the Web after the third screen, and this is being generous!

26. To be honest, eBooks have been promised for the last fifty years. See Jennifer Adamec. "EBooks in the College Classroom." Unpublished Thesis. N.p.: MS Publishing, 2006, 2.

27. See Scott Carlson. "Library Renovation Leads to Soul Searching at Cal Poly." Chronicle of Higher Education (September 1, 2006) Vol. LIII (2), A59-61, where professors and librarians complain about the shift from print to online exclusively, and where print is considered so passé.

28. In some cases, readers are charged for books that are in public domain. While the charge may only be a nominal one (say $3 or $4), one still must pay for what is freely available elsewhere on the Web. This fee is ostensibly to defray the cost for loading in the Sony Connect eBook Reader format.

29. Bear in mind that we have been reading with the light (whether candle or electric) over our shoulders for the last 500 or so years.

30. Content is pre-chosen so one is electing what someone else thinks one should read, not what you, the reader, have decided to read. Some critics will complain

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

that lack of choice is always the case in any library one enters. Perhaps, but I like my odds in a library of half a million books over one with only 5,000 choices, most of which are the usual half-baked best-sellers.

31. Robert J. Samuelson. Untruth: Why Conventional Wisdom Is (Almost Always) Wrong. AtRandom.com: New York, 2001, 218.

32. My favorite edition is the Vintage, published in 1977, a reprint of the 1932 Everyman edition. The book was of course published in 1621 and is a towering monument to scholarship of the highest, best, and most humorous kind, though not many critics see the latter. I cannot help but point out he, in addition to spending a lifetime on this magnificent book, was also a college librarian.

33. Mark Kirby and Nancy H. Evans. "The Network Is the Library." EDUCOM Review (Fall 1989) Vol. 24 (3), 16. Also quoted in Karen M. Drabenstott. Analytical Review of the Library of the Future. Council on Library Resources, Washington, 1994, 9.

34. Landoni Monica, Nadia Catenazzi and Forbes Gibb. "Hyper-Books and Visual-Books in an Electronic Library." Electronic Library (June 1993) Vol. 11 (3), 176. Also quoted in Karen M. Drabenstott. Analytical Review of the Library of the Future. Council on Library Resources, Washington, 1994, 10.

35. Karen M. Drabenstott. Analytical Review of the Library of the Future. Council on Library Resources, Washington, 1994, 11, 15, 20, 27.

36. Also quoted in Walt Crawford and Michael Gorman. Future Libraries: Dreams, Madness, & Reality. Chicago: American Library Association, 1995, 4.

37. "Ipswich, Inc., Warns that Spam Continues to Rise." Accessed via <http://www.ipswitch.com/company/press_releases/060831_spamometer.asp>, (August 31, 2006) viewed September 2006.

38. OCLC Online Computer Library Center. College Students' Perceptions of Libraries and Information Resources. OCLC Online Computer Library Center: Dublin, Ohio, 2006. This is a companion piece to an earlier report mentioned in this book, Perceptions of Libraries and Information Resources, 2005.

39. Ibid., 1-1.

40. Ibid., 1-2, 1-4, 1-7.

41. Ibid., 1-8.

42. Ibid., 1-11.

43. I should add that while I cannot extrapolate for all libraries, the one where I work has increased in usage over the last seven years as measured by our door counts.

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

44. F. Wilfrid Lancaster. "Second Thoughts on the Paperless Society." Library Journal (September 15, 1999) Vol. 124 (15), 48. See also his "The Paperless Society Revisited." American Libraries (September 1985) Vol. 16 (8), 553-555.

45. Ibid.

46. Ibid. (Emphases mine.)

47. Ibid.

48. I would not want to suggest his comment here contra electronic everything is the only reason, or even a primary one, that Lancaster himself has had troubles of late, but the timing is surely right. Only recently he had to step down as the editor of Library Trends, a move that was not only not of his making, but also not of his liking. It strikes me that the "trends" portion of that journal's name may well not like the idea that its editor is not fully on board with the current electronic uber alles trend? See "Lancaster Steps Down from Library Trends but Not by Choice." Library Hotline (August 14, 2006), 3.

49. Lancaster, 49.

50. I am not the first to raise this issue. See Sven Birkerts. The Gutenberg Elegies: The Fate of Reading in an Electronic Age. Boston: Faber and Faber, 1994. Birkerts' book was prophetic, foreseeing more than a decade before it happened most of what I record here.

51. Figures for what follows have come from Trends in the Well-Being of America's Children, 2003, specifically "Achievement/Proficiency." See also Tom Loveless. How Well Are American Students Learning? The 2006 Brown Center Report on American Education. Washington: The Brooking Institution, 2006, 8-11.

52. It isn't part of this argument about libraries and literacy but it is interesting to note that the Brown Report also points out that not only do students not do better if they are happy (i.e., have strong self esteem), but that trying to making learning relevant not only does not improve student learning but may also very well impede it. See Tom Loveless. How Well Aare American Students Learning? The 2006 Brown Center Report on American Education Washington: The Brooking Institution, 2006, 13-18.

53. Loveless, 327-330.

54. National Accessible Reading Assessment Projects. Defining Reading Proficiency for Accessible Large-Scale Assessments: Some Guiding Principles. February 17, 2006. Minneapolis, MN, 1. Available at <www.narap.info>. Viewed, September, 2006. 3. The changes are in response to standards made tighter by the NCLB Act. While flexibility is stressed, it is flexibility not in what defines proficiency but in what a child is able to read.

FOOL'S GOLD
WHY THE INTERNET IS NO SUBSTITUTE FOR A LIBRARY

Mark Y. Herring

55. Ibid., 6.

56. Center for Academic Integrity, at <http://www.academicintegrity.org/>.

57. Julie Coiro. "Reading Comprehension on the Internet: Expanding Our Under-standing of Reading Comprehension to Encompass New Literacies." Reading Teacher (February 2003) Vol. 56 (5) 462.

58. Robin Goodfellow. "Online Literacies and Learning: Operational, Cultural and Critical Dimensions." Language and Education (2004) Vol. 18 (5), 380. While some of what Goodfellow argues here agrees with my conclusions, he, too, is more hopeful about the future of Web-based learning than am I.

Emily F. Blankenship

LEARN MORE ›

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

INTRODUCTION

The vast majority of knowledge now lies outside the realm of a physical library. The general public and many librarians now rely upon mega search engines to locate, in a timely manner, the most obscure data. Libraries could still play vital roles in these transactions because libraries can provide access to more scholarly resources, but the mega search engines, in reality, serve as Internet guideposts for most people and our challenge is to bring people back to their library holdings and services.

To paraphrase Dr. Vartan Gregorian during his speech at the American Library Association Annual Conference in June 2007, "Librarians are functionaries of our society. They have not changed the course of library service, but they have been changed by outside powers."[1] Google, Yahoo!, and Microsoft comprise the top three powers or mega search engines that are slowly and systematically changing the face of reference work. Librarians must use their advanced search skills in partnership with Google, Yahoo!, and Microsoft in order to provide timely, high quality, and relevant reference work to the new generation of users.

BACKGROUND

Seemingly just a few years ago, communities turned to the knowledgeable reference librarian who could place his/her finger on the correct page of information. Then, in 1996, came the wake-up call found within the Benton Foundation's report "Buildings, Books and Bytes: Libraries and Communities in the Digital Age." Benton's report focused on frequent library users and their perception of the library of the future.

The youngest Americans polled, those between the ages of 18 and 24, were found to be the least enthusiastic boosters of maintaining and building library buildings. This age group, commonly known as Generation Y, was also the least enthusiastic of any age group about the importance of libraries in a digital future. The young poll participants voted to spend their money on personal computer systems rather than contribute the same amount in tax dollars to the library. When asked to think about the role of libraries in the future, they placed libraries firmly in the past. "In 30 years," the Generation Y group noted, libraries would be relegated to a "kind of museum where people can go and look up stuff from way back when."[2]

The library of the future, according to the young Benton Foundation poll participants, far from being a technology leader, would function as an historical archive. The young participants presented an equally diminished view of the future role of librarians. The

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

participants noted that librarians could perform a useful role as navigators in the as-yet difficult-to-navigate universe of the Internet. Yet, poll participants just as easily sanctioned the notion that trained library professionals could be re- placed with community volunteers, such as retirees. For these unsophisticated library users, the concept of "librarians as trained professionals" was nebulous at best.[3]

Predictions from the Benton Foundation report proved to be on track. From 1997 to 2007, the public perception of the need for actual reference librarian work changed so much that the relevancy of the existence of libraries and reference desks is now questioned. The concept of a reference search, itself, has been transformed in the popular mind from perusing printed multivolume sets of works and printed serials indexes to an Internet search and the mania with all things Internet. Reference librarians are continuing to find increasing competition from the massive growth of the mega search engines like Google, Yahoo!, and MSN Search. Google, with its academic search engine, Google Scholar, provides especially alarming search statistics, competition, and annoyance to conventionally trained reference librarians.

TRADITIONAL REFERENCE SERVICES FACE COMPETITION

Maurice York, in his recent article regarding Google Scholar, notes that librarians are divided into two camps regarding the use of Google Scholar: the purists, those who wish to protect their users from the gap-ridden, non-scholarly content; and those who are in some way trying to come to terms with the new 800-pound gorilla in their living room by introducing it to users in a cautious manner.[4]

Librarians possess legitimate concerns about users embracing acceptance of the mega search engines. Shared is the central fear that users will abandon expensive library databases and the library catalog and proceed to use the mega search engines exclusively for their research.

Support for reference librarian concern is shown in the results from East Carolina University's 2007 LibQUAL+ survey. Of the respondents, 58.23% said they used non-library gateways daily to retrieve information; 12.85% of the respondents stated they accessed library resources through the library gateway daily. Finally, only 2.81% of the respondents reported actually using resources housed on the library premises on a daily basis.[5]

Nancy Becker, in her article "Google in Perspective: Understanding and Enhancing Student Search Skills" notes that during a study of undergraduate student search behavior, "many college students continue to assign unwarranted primacy and

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

authority to information found through Google and on the Web."[6] Ms. Becker further states that many of the students were able to articulate the importance of source evaluation for assessing the authority and reliability of Web-based resources. In practice, however, the same students usually abandoned source evaluation and followed the path of least resistance, relying exclusively on basic Google searching. Librarians now share the experience of being relegated to the brink of irrelevancy. Librarians are daily encountering users eliciting exclamations such as "I have looked all over the Internet for articles, but didn't find any. Now I'll ask you!"

Recently, the Search Engine Users report issued by the Pew Internet & Life Project found that most Internet users were naïve about doing re- search via the Internet. Overall, 18% of searchers said they could consistently distinguish between paid and unpaid search results and 70% of the searchers stated they were comfortable with the concept of paid or sponsored results.[7] This finding is particularly ironic, as nearly 50% of the users stated they would stop using search engines if they thought engines were not clearly designating paid results. Such outcomes demonstrate that search engine users are unaware of the analogous sets of content commonly presented by search engines.

PROSPECTIVE PATRONS AND NEW EXPECTATIONS

In the summer of 2004, the Pew Internet & American Life Project released its report on American Internet use. The Pew findings answer the very important question: "Who searches the Internet?" The answer is nearly everyone. According to Pew, 84% of adult Internet users, about 108 million United States citizens, have used search engines to help them find information on the Web.[8] Only the use of e-mail, with approximately 120 million users, eclipses searching in popularity as an Internet activity. The amazing part, at the core of these phenomena, is one per- son in front of a screen, typing in a query. Prospective library users, now also known as searchers, expect the mega search engines to bring back perfect results within seconds.

People want results now, and they desire perfect answers without learning an unwieldy programming language or drilling through copious layers of Web pages on an academic library Web site to reach the desired article. As the Benton Foundation report summarizes, the existence of the Internet, home computers, super-bookstores, and private in- formation brokers require libraries, of all types, to change their means of providing the materials the users want and need.[9] Richard Wayne, writing about technology trends in Texas libraries notes, "Google has become the bane of many

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

librarians and library school professors. Yet, the public–and even many librarians–love Google, and Google continues to grow like gangbusters."[10] People have become accustomed to almost instantaneous answers; now it is up to librarians to adopt new technologies and means of providing reference services or to face extinction.

In further evidence of the changing tide of librarian opinion, Jan Lewis, Associate Director for Library Services at East Carolina University, notes:

> *It is clear that most of the world, including our students, are "Googling." Google Scholar looks and works like the Google they are familiar with, yet gets them to the content of some (not nearly all) of our licensed scholarly content materials that they would not otherwise be able to access through Internet search engines. We want to enable access to our licensed resources however we can to maximize use by our students. We hope that once they get into a database at the article level through a Google Scholar search, they may decide to use the database's search functions for more in-depth information retrieval.[11]*

THE COMPETITION IN NUMBERS

Where are the students and general populous finding the answers to their questions while searching on the Internet? Many answers are found in the 2007 Piper Jaffray Research Report entitled The User Revolution. Released in February 2007, the Piper Jaffray Research Report states:

> *Google dominates the United States search market with 46.3% of users performing their searches via Google. Yahoo! is a distant second with 28.5% of the United States search market share. MSN- Microsoft appears in third place with 10.5%, Ask follows with 5.4%, AOL holds 4.9%. Various other search engines comprise the remaining 3.4% of the United States search market. Piper Jaffray expects Google to continue to make gains in the volume of searches until levelling off at around 70% of the total search volume.[12]*

Frighteningly for some librarians, John Battelle, founder and chairman of Federated Media Publishing notes, "the search engine of the future isn't really a search engine as we know it. It's more like an intelligent agent–or as Larry Page told me, a reference librarian with a complete mastery of the entire corpus of human

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

knowledge."[13] Libraries indeed may soon be in peril if they continue to lag behind in adoption of new technology. Such lag causes the public to assume that libraries are old-fashioned. In stark contrast, the mega search engines are issuing press releases for each minuscule update or change.

Larry Page has not missed a step in his path to knowledge domination, yet. Google comes to us offering harmony, but on its own terms, which libraries have not attempted to challenge or even question. The challenge we face is one of understanding both the existing power and the potential of such a world knowledge force. As Google moves into for-fee services, libraries must work out the business models which suit our positions and discontinue being functionaries.

In June 2007, searchers in the United States conducted 8.0 billion searches online. This figure rose 26% from June 2006, as noted by digital survey company ComScore in its June 2007 United States Search Engine Rankings Report. ComScore also reports Google further increasing its search market share compared to the other mega search engines.[14] These statistics should serve as startling evidence for librarians to collaborate with the mega search engines, and especially with Google.

THE FOCUS ON GOOGLE

Since 2002, Google has been synonymous in the public mind with Internet Search. Google's recent inclusion as a verb in the "Oxford English Dictionary" confirms what all competitors, including reference librarians, feared: Google means Search to the majority of Internet users. The company has relied on word-of-mouth marketing and the ingenuity of its founders, Larry Page and Sergey Brin, to become a common house- hold verb. To Google stands even more popular than the use of Xerox as a synonym for photocopying, and Kleenex as a synonym for tissues. Unlike Google, both older, long ago established companies actively con- done the use of their trade marks as verbs, given that such use places the trademark in danger of being declared a generic term.

Google does not intend to rely on its past accomplishments or allow its lead against other Internet search engine competitors to wane. Ever striving to be all indexing, Google, in May 2007, announced its critical first steps toward a universal search model. This new search model promises users more integrated and comprehensive ways to search for and view information online. Marissa Mayer, Vice President of Search Products and User Experience at Google states, "The ultimate goal of universal search is to break down the silos of information that exist on the Web and

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

provide the very best answer every time a user enters a query. While we still have a long way to go, today's announcements are a big step in that direction."[15]

Google's vision for universal search is ultimately to search across all its content sources, compare and rank all the information in real time, and deliver a single, integrated set of search results that offers users precisely what they are seeking. Google plans to subtly incorporate information from a variety of previously separate sources–including videos, images, news, maps, books, and Websites–into a single set of results. Over time users are promised recognition of additional types of content integrated into their search results as the company attempts to deliver a further comprehensive search experience.[16]

To further ease the rigors of Internet searching across countries and continents, Google, also in May 2007, unveiled a new cross-language search feature, Google Translate. Google Translate allows users across the world to find and view search results on foreign language Web pages in their own native language. By doing so, additional content on the Web is accessible to more users, regardless of what languages they speak. Reference librarians, working especially in multicultural environments, will find Google Translate beneficial. Google Translate is available in: English, Arabic, French, Italian, German, Spanish, Portuguese, Russian, Japanese, Korean, Chinese (Traditional), and Chinese (Simplified).[17]

To this end, Google is pouring the bulk of its resources into enhancing the breadth and quality of the search experience. Recent innovations involve wholly different methods of searching that may make today's Google seem primitive. As these projects evolve, the search mechanisms of the future will produce better answers to queries, much as the current Google is superior to the search engines that preceded it. "The ultimate search engine," says Page, "would understand exactly what you mean and give back exactly what you want."[18]

FALLING BY THE WAYSIDE

Yahoo!, the second largest search provider, at 26.8%, commands a market share more than double that of the next largest search engine, MSN. Outside of the top two search market shareholders, MSN, Ask, and Time Warner Network (AOL) round out the top five search engines used in the United States. Yahoo's market share in terms of query volumes has declined over the last year as Google's brand draws away users. Yahoo! hopes to bring back users through social search, known as Yahoo! Answers.

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

Yahoo! also recently unveiled a new integrated brand campaign that demonstrates how Yahoo! can help people be better at whatever they're into. Whether it's being a better shopper, Frisbee player, or salesman, Yahoo! advertises to empower users to connect to their interests and passions and evolve that part of their lives through Yahoo!'s information, tools, and services. Much to the reference librarian's chagrin, Yahoo!'s multi-million-dollar advertising campaign "Be a Better …" spotlights Yahoo!'s trademark irreverence about the academic and scholarly re- sources.

The "Be a Better …" campaign focuses on Yahoo! Answers and Yahoo! oneSearch, a new search service available on users' mobile phones. oneSearch showcases how people can enhance their lives by using Yahoo! to access the world's knowledge and make informed decisions from anywhere, anytime. The "Be a Better …" messaging will integrate into Yahoo! fans' lives through interactive ads and promotions, for example "Be a Better Globetrotter" (Yahoo! Travel) and "Be a Better Roadie" (Yahoo! Autos).[19]

Very little is appearing on the forefront of development at MSN, Ask. com, and the host of other search engines available on the Internet. MSN, in 2006, announced a number of new features and enhancements including an updated version of Windows Live and a Windows Live Toolbar. Windows Live Search includes Web search, as well as an image search, news search, RSS feeds, mail, local search, shopping, and MSN spaces, Microsoft's blogging service. New and updated search engine components were also available, including search history, clock, notepad, Live Favorites, stock quotes, weather, and MSN Infopane, to access MSN content. Unfortunately for MSN, similar features have been available in the past via numerous other sites, including Google.

FUTURE DECISIONS

Gazing into the crystal ball of the future, what resources and responsibilities will tomorrow's reference librarian command? How long will it be before Google and/or the mega search engines require fees for access? Many concerns have already presented themselves in the reference librarian community.

One concern arises as librarians include Google Scholar on library Web pages alongside subscription resources. Librarians are concerned they may be seen as promoting Google Scholar's use for subject searching of scholarly material as well as suggesting that it is equivalent to the other indexes and databases listed.[20] Will the placement of Google Scholar search boxes on library Web pages affect the use of other resources? Are librarians purposely or unintentionally putting themselves out of business by doing so? Many librarians would agree in the affirmative to both questions.

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

Among the many libraries with no intention of putting themselves out of business are Baruch College Library in New York City, the University of Georgia Libraries, and Joyner Library at East Carolina University. At Baruch, electronic reference is at the forefront of student and faculty research. Reliance on electronic search is evident not only in the remarkable number of computer stations spread throughout the reference section, but also in the quantity of empty shelving units that formerly housed print sets and periodicals. The University of Georgia Libraries report cancellation of all print indexes, in favor of Web-based versions of the same materials. Both Baruch and Georgia universities state they can meet most of their undergraduates' needs without a print collection.[21] Joyner Library at East Carolina University has also switched to electronic versions for most of its serial subscriptions. In addition, Joyner Library has begun an ongoing series of Google drop-in sessions. These 15-minute sessions, held twice per month, provide in-depth training on the many tricks and short cuts embedded in Google and Google Scholar's search features. All three libraries are now actively marketing their services to their communities, as they have learned to do from their commercial counterparts.

CONCLUSION

As Paul Courant notes in his article "Scholarship and Academic Libraries (and their kin) in the World of Google," except for the most arcane materials and users, that which is not available online will simply not be read. Libraries must make sure true scholarly material is available online, or users will only see the inferior material. In addition, collaboration is the fundamental method of scholarship, and without it libraries can do nothing of value. If libraries focus on the purposes and mechanisms of scholarship, the new technologies should be our friends. In order to remain relevant to their communities, libraries will need to work together and decide which of the mega search engine offerings they wish to promote, which they are prepared to pay for, and which to reject. Libraries will be required to stand up for the principle of free and equal access to content, and for the principle of high-quality index provision, because without those principles we are no longer running libraries.[22] It is important not to take sides, but to expend energy and resources on useful and valuable innovations within each library's grasp. Certainly a middle road exists between the old and the new reference services and it is highly conceivable for libraries to prosper in both worlds.

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

NOTES

1.   Vartan Gregorian. American Library Association Annual Conference. Washington Convention Center, Washington, DC. June 23, 2007.

2.   Benton Foundation. Buildings, Books, and Bytes. Washington, D.C.: Benton Foundation, 1996. <http://www.benton.org/publibrary/kellogg/buildings.html> (accessed June 3, 2007).

3.   Ibid.

4.   Maurice C. York. "Calling the Scholars Home: Google Scholar as a Tool for Re-discovering the Academic Library," Internet Reference Services Quarterly, 10:3/4 (2005): 118.

5.   LibQUAL+ Survey, Greenville, NC: Joyner Library, East Carolina University, unpublished report, 2007, p. 38.

6.   Nancy J. Becker. "Google in Perspective: understanding and enhancing student search skills." The New Review of Academic Librarianship, 9:1 (2003): 84.

7.   Deborah Fallows, Search Engine User, Washington, D.C.: Pew Internet & American Life Project, 2005. <http://www.pewinternet.org/PPF/r/146/report_ display. asp> (accessed April 27, 2007).

8.   Ibid.

9.   Benton Foundation. Buildings, Books, and Bytes. Washington, D.C.: Benton Foundation, 1996. <http://www.benton.org/publibrary/kellogg/buildings.html> (accessed June 4, 2007).

10. Richard Wayne. "T4: Top Texas Technology Trends," Texas Library Journal, Spring (2005): 13.

11. Janice S. Lewis. "Re: Google and Libraries." E-mail to the author. April 27, 2007.

12. Safa Rashtchy. The User Revolution: The New Advertising Ecosytem and The Rise of the Internet as a Mass Medium. East Palo Alto, CA: Piper Jaffray, (2007), p. 169.

13. John Battele. The Search. New York, NY: Portfolio, 2005.

14. Andrew Lipsman. ComScore Releases may U.S. Search Engine Rankings. Reston, VA: ComScore, Inc., 2007. <www.comscore.com/press/release.aps?press=1525> (accessed July 25, 2007).

15. "Google Begins Move to Universal Search." May 16, 2006. <http://www.google. com/intl/en/press/pressrel/univeralsearch_20070516.html> (accessed June 5, 2007).

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

16. "Google and Four US States Improve Public Access to Government Web- sites." Google. (April 30, 2007) <http://www.google.com/intl/en/press/pressrel/ govt_access.html.> (accessed June 5, 2007).

17. "Google Leaps Over Language Barriers." May 23, 2007. <http://www.google. com/intl/en/press/annc/translate_20070523.html.> (accessed June 5, 2007).

18. David A. Vise and Mark Malseed. The Google Story. New York, NY: Portfolio, 2005.

19. "Yahoo! Empowers Users To 'Be a Better …' " Through the Most Integrated Brand Campaign in Company History. April 30, 2007. <http://yhoo.client. shareholder. com/press/ReleaseDetail.cfm?ReleaseID=239962> (accessed June 4, 2007).

20. Laura Bowering Mullen and Karen A. Hartman. "Google Scholar and the Library Web Site: The Early Response by ARL Libraries." College & Research Libraries, 67:2 (March 2006):117.

21. Mirela Roncevic. "The e-Ref Invasion." Library Journal, 130:19 (2005):8.

22. John MacColl. "Google Challenges for Academic Libraries." Ariadne, 46:64 (February2006) <http://www.ariadne.ac.uk/issue46/maccoll/> (accessed June 5, 2007).

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

REFERENCES

"Google and Four US States Improve Public Access to Government Websites."

Google, (April 30, 2007) <http://www.google.com/intl/en/press/pressrel/govt_access.html> (accessed June 5, 2007).

"Google Begins Move to Universal Search." Google, (May 16, 2007). <http://www. google.com/intl/en/press/pressrel/univeralsearch_20070516.html> (accessed June 5, 2007).

"Google Leaps Over Language Barriers." Google, (May 23, 2007). <http://www. google.com/intl/en/press/annc/translate_20070523.html> (accessed June 5, 2007).

Battele, John. The Search. New York, NY: Portfolio, 2005.

Battele, John. "Web 2.0 Principles." (May 20, 2007). <http://www.federatedmedia.

net/blog/> (accessed June 5, 2007).

Becker, Nancy J. "Google in Perspective: Understanding and Enhancing Student Search Skills." New Review of Academic Librarianship, 9 (2003): 84-100.

Benton Foundation. Buildings, Books, and Bytes. Washington, D.C.: Benton Foundation, 1996. <http://www.benton.org/publibrary/kellogg/buildings.html> (accessed June 4, 2007).

Brophy, Jan, and David Bawden. "Is Google Enough? Comparison of an Internet Search Engine with Academic Library Resources." Aslib Proceedings, 57:6 (2005): 498-512.

Courant, Paul N. "Scholarship and Academic Libraries (and their Kin) in the World of Google." First Monday, 11:8 (2006):1-14.

Fallows, Deborah. Search Engine User. Washington, D.C.: Pew Internet & American Life Project, 2005. <http://www.pewinternet.org/PPF/r/146/report_display.asp> (accessed April 27, 2007).

Gregorian, Vartan. American Library Association Annual Conference. Washington Convention Center, Washington, DC. June 23, 2007.

Lewis, Janice S. "Re: Google and Libraries." E-mail to the author. April 27, 2007. LibQUAL+ Survey, Greenville, NC: Joyner Library, East Carolina University, unpublished report, 2007.

Lipsman, Andrew. ComScore Releases may U.S. Search Engine Rankings. Reston, VA: ComScore, Inc., 2007. <www.comscore.com/press/release.aps?press=1525> (accessed July 25, 2007).

MacColl, John. "Google Challenges for Academic Libraries." Ariadne, 46:64 (February 2006) <http://www.ariadne.ac.uk/issue46/maccoll/> (accessed June 5, 2007).

Mullen, Laura Bowering, and Karen A. Hartman. "Google Scholar and the Library Web Site: The Early Response by ARL Libraries." College & Research Libraries, 67 (2006): 106-122.

Rashtchy, Safa. The User Revolution: The New Advertising Ecosytem and The Rise of the Internet as a Mass Medium. East Palo Alto, CA: Piper Jaffray, 2007.

WHO HOLDS THE KEYS TO THE WEB FOR LIBRARIES?

Emily F. Blankenship

Roncevic, Mirela. "The e-Ref Invasion." Library Journal, 130:19 (2005): 8-10, 13. Vise, David A., and Mark Marlseed. The Google Story. New York, NY: Delacorte Press, 2005.

Wayne, Richard. "T4: Top Texas Technology Trends for Libraries." Texas Library Journal, 81:1 (2005): 12-17.

"Yahoo! Empowers Users To 'Be a Better … '" Through the Most Integrated Brand Campaign in Company History. (April 30, 2007). <http://yhoo.client.shareholder. com/press/ReleaseDetail.cfm?ReleaseID=239962> (accessed June 4, 2007).

York, Maurice C. "Calling the Scholars Home: Google Scholar as a Tool for Rediscovering the Academic Library." Internet Reference Services Quarterly, 10:3/4 (2005): 117-123.

ASSESSING INFORMATION NEEDS

- Why Undertake Information Needs Assessments?
- What Are Information Needs?
- A Framework for Evaluating Information Needs

The following is excerpted from

Assessing Information Needs in the Age of the Digital Consumer

Edited by David Nicholas and Eti Herman

LEARN MORE >

1   Why undertake information needs assessments?

The user-oriented holistic approach to the development of information systems and services has been strongly advocated by the vast majority of information studies researchers for quite some time now (for extensive reviews of the literature on the subject see Dervin and Nilan, 1986; Hewins, 1990; Pettigrew et al., 2001). Indeed, according to Stefl-Mabry et al. (2003, 441) 'user-centred design has become an established goal of much of the work in information science'. Practice, however, seems to lag (far) behind theory where the actual set up, evaluation and auditing of information provision systems are concerned. As Pettigrew et al. (2001) point out, the realisation that information systems and services should be designed to support information behaviour and that the design of such systems should be based on our understanding of this behaviour, has not often led to the forging of a direct link from the study of information needs and behaviour to information provision specifications or practice. Thus, information professionals may be forever paying what is really not much more than lip service to the 'user', but while space-age information systems grace our desk tops, information centres and libraries, we still do not use suitably modern and effective management methods to ensure that these systems are providing their customers with what they need and want. To say that information systems are largely free from consumer evaluation and are seldom challenged with user needs or usage data, would be to exaggerate, but not by very much. Rarely are high-quality data fed into the design, evaluation and running of information systems, like intranets, libraries and websites.

It is hardly surprising, then, to find that people in the information professions, alert as they are to the technical changes that have taken place in the virtual information space, are nevertheless going about their business as if nothing really fundamental has happened to their clients, the users-cum-consumers. Indeed, contrary to what is plainly there for all to see, they seem to believe that the digital information consumers of today are no different from the 'library patrons' or 'readers' of yore. However, as we have noted elsewhere (Nicholas et al., 2008b), these days information professionals are confronted with an entirely different breed of information seekers, looking for information, yes, but also for goods, services, new experiences, titillation,

excitement and entertainment. Moreover, they can do their information seeking unbelievably easily and expediently through a plethora of devices and platforms at their disposal 24 hours a day, seven days a week. Indeed, as our whole experience of the virtual scholar clearly indicates, for today's information consumer convenience and user satisfaction will triumph, even over content, any day of the week (Nicholas et al., 2008b). What is more, as CIBER's 'Google Generation' project amply proves, today's digital consumers benchmark their online experiences against more immersive environments like Amazon or Facebook (Williams et al., 2008). It is a foregone conclusion, then, that virtual library spaces need to be involving, easy to use and simple. Why, then, are so many of the sites produced by information professionals and publishers as complex and austere as they are? Why can't library catalogues be like Amazon, with sample pages, trust metrics, referral metrics, user feedback and colour? Why do they not speak to the user? We need to remember: if the traditional purveyor of information is not there (at best) or gets in the way of communication (at worst), consumers will simply abandon what to them is a sinking ship – after all, they can go their own way! This key aspect of the digital revolution applies to all members of the communication and information food chain, and ignoring it brings about a real risk of libraries becoming decoupled from the user. We will take up this point again further on, but first, why this patent disregard of the user?

The whys and wherefores of the ongoing neglect of information needs

Would anyone actually doubt that libraries, archives, information units and databases are there solely to service the needs of their clients? Would anyone really accuse information professionals of not being wholly aware that this is the gist of their mission? There are, however, six factors that add up to a general neglect of the user:

1 Believe it or not, there are still many (quite well-known) information professionals who feel that it is not necessary to consult the client in what they consider to be professional matters.
2 Information professionals tend to be preoccupied with information systems and not the users of these systems.
3 The professions, especially librarianship, are plagued by insular attitudes and poor communication skills, something that does not lead to a close relationship with the consumer.
4 It is by no means easy to get hold of the necessary needs data.
5 The expenditure of resources involved in the obtaining of needs data is not thought to be justified in these hard budgetary times.
6 There is an absence of a standard, commonly understood framework for the assessment of information needs – something, which, it has to be said, lets information professionals off the hook.

There is little point in conducting information needs assessments (trust us)

It is hard to credit, but there is a school of thought that believes there is very little point in consulting the users: people do not know what their needs are; they do not know what they are talking about; why ask them; far better to trust professional judgement. Just listen to this: 'there is something rather absurd in being constantly enjoined to meet the needs of the user, when needs have been probed the outcomes have been worse rather than better' (Shinebourne, 1980). Shinebourne is not alone, either: some quite eminent (more recent) Library and Information Science (LIS) researchers have also been questioning the value of user-centred design, development and implementation at the individual and system levels (Rosenbaum et al., 2003; Stefl-Mabry et al., 2003). Thus, Rosenbaum et al. (2003, 429) may freely admit that 'much has been gained intellectually and culturally from the emphasis on the user, from improvements in the design and use of information and communication technologies (ITCs) to enhancements in library services for patrons', but they still wonder as to the 'cash value' of the concept of the user (not something Google wonders about, of course). Has it reached a point of marginal or diminishing returns in LIS research, to the point that we may be actually witnessing 'the death of the user in LIS'? After all, they go on to say, the outcome of user-centred design can hardly be tooted as a success story: 'many ITCs do not work well for their intended audiences and there is a rich history of costly information systems failures in corporate, educational and governmental organizations.'

Instead, proponents of this approach often argue that since we are the experts, we should simply trust our professional judgement to get it right, and then it is just a question of convincing the users of it. True, people no longer dare say 'we know better what is good for you' in so many words (it would be politically incorrect), just as today's physicians refrain from voicing such patronising attitudes (at least in the presence of their patients), but these stances, perhaps worded more carefully, are still endemic in our profession. Thus, for example, Stefl-Mabry et al. (2003, 441), exploring 'the extent to which there is real substance behind the rhetoric of user-centeredness in our research and practice', set out to answer questions such as 'What do you fail to learn when you rely solely on users for input on decisions?', or 'When do users not know what is best for them?' They even speculate if 'we have uncovered significant knowledge of users since the advent of the user-centred approach'. Cronin (1981, 38; 46), too, despite claiming that 'user studies are on the whole a jolly good idea', has his reservations as to their practical worth: 'a fine sentiment no doubt but sentiment is not always at home in the world of commerce'. Admittedly, he might have changed his outlook in the 20-something years since he put forward this view of his, for, clearly, today's commerce is precisely all about getting closer to the user. With very good reason, too: sophisticated market research and demographic profiling are behind the success of many leading retailers and service providers all over the world. Actually, as Kujala

(2003) suggests on the basis of a careful analysis of the LIS literature, user involvement in information systems development has generally positive effects, especially on user satisfaction, and taking users as a primary information source is an effective means of requirements capture.

A systems-driven profession

The information profession is, by and large, systems driven; it shows an enormous interest in the processing and storing of information, to the detriment of the consumer. Indeed, the profession manifests a marked fixation with powerful, innovative information systems, sometimes irrespective of their direct suitability to users. A prime example of this mindset is libraries' recent attempts to halt the massive desertion of their patrons by succumbing to any digital fashion that comes their way: institutional repositories, portals, internet cafés, learning spaces, open access and social networks. What is more, once such facilities have been obtained it is on to the next powerful, innovative one, without any user evaluation of the former system or facility. The sheer pace of technological change provides the ideal opportunity and excuse. The qualities that are appreciated tend to be systems characteristics, such as speed of response, storage size, or number of network stations. If needs are considered at all, it is by and large in terms of how the system might meet these needs, rather than be built, developed or changed to meet them. No wonder, therefore, as Adams et al. (2005) note, that librarians are held to be possessive of resources rather than supporting and understanding users' needs. The words of the academic they quote do indeed seem to echo faithfully enough the popularly held sentiments on the topic: ' … the librarians are not user-centred, they're information resource centred … they want to protect their resources, not to gain access to them'.

Admittedly, in their deliberations and writings, the systems-driven of this world pay due allegiance to the idea of the user (generally mentioning the worn-out slogan 'user-friendly', which is, of course, a wholly systems phrase). It is almost as if by citing the term or having a conference on the topic, their guilt is assuaged. However, that seems to be the extent of it; hardly any libraries (or publishers for that matter) actually investigate users' information needs, follow their information behaviour and then relate their findings directly to outcomes and impacts. The following tale truly drives the point home: at a multimedia conference one of the authors attended a number of years ago, the delegates, from the top media companies in the country, were asked two questions: firstly, how many of them had online access, to which, with some alacrity, all put up their hands; secondly, how many of them conducted annual user needs assessments, to which just one admitted so doing. When confronted by the sheer disparity, most felt uncomfortable at first, but soon gathered their wits and asserted that this did not mean that they did not 'know' their clients – they did, of course! However, their main methodology for obtaining user data appeared to be osmosis. It truly looks as if information

systems are seen as omnipotent, while users are too often perceived as the supplicants. This, when users are the ones who, after all, drive all the major changes in the digital information environment. How can we possibly see to it that we are delivering the right information services, unless we think about people's needs, rather than the allures of our systems?

Poor communication skills and insular and antagonistic attitudes

Information professionals tend to demonstrate a marked reluctance to keep in touch with their clients and, in result, they do not always know their clients as they patently should – though they would inevitably protest otherwise. Take, for example, Weintraub's (1980) gentle reproach of librarians, who, according to him, soldier on in brave isolation, as if they were a beleaguered community of martyrs, instead of dealing with the problems of the library by establishing a more effective interaction between them and the communities of which they are a part. However, a continuing dialogue between client and information professional is often a rarity – has always been a rarity – so that we cannot even blame it on the recent massive advent of disintermediation (loosely defined and understood as 'cutting out the middleman'). Why, then?

In defence of their customary ways, information professionals often cite shortage of time and work pressures. However, more frequently the real reason seems to be a characteristic insularity and four-wall mentality. Information professionals are typically inward-looking and tribal (something which seems to be particularly true of librarians): anything that happens outside their strictly defined discipline and/or job boundaries is not their prime concern, even though the user and the internet are dissolving these boundaries. In direct consequence, they are insufficiently concerned with information problems that occur outside the information unit – in the digital world, for instance. However, it is in the office, coffee bar and home that information needs are hatched and only occasionally, one would have thought, in the physical information space. This would necessitate that the information professionals go to their clients, which these days means both in the literal and the virtual sense of the word. Thus, it is important to interact with potential customers in their natural surroundings, say in the office or the laboratory, but it is no less crucial to follow closely the digital footprints of those who will never be encountered in the flesh. Luckily, we are in the enviable position that our predecessors could not even begin to dream about, of being able to monitor closely information seekers' many activities, as these take place anonymously in the virtual space.

Poor lines of communication can also be attributed to the low status of information professionals in the organisation, making it difficult for them to initiate contact and maintain ongoing dialogues. Generally seen as fulfilling a supporting role, information professionals are often overlooked and disregarded, and certainly not considered peers of the 'real doers' in the establishment. This perception of the information professional is especially prevalent in academe, as Biggs (1981) points out in her seminal work on the topic. Quoting a librarian,

according to whom she and her ilk are treated as 'poor relations' and 'super clerks', Biggs traces this to faculty myths about librarians' alleged rigidity, meticulousness and lack of intellectual knowledge. More recently Adams et al. (2005, 183) also reported that librarians do not see themselves as being on equal footing to the academics with whom they work; as one interviewee of theirs puts it: 'There is also that whole professional thing as well that you have to be wary of that we're not seen as maybe their equal in knowledge or whatever and that maybe they feel that they're not going to learn anything gainful from us'. Regrettably, Biggs' (1981) conclusion is that the information professionals are the ones who, by their reluctance to confront their clientele, create a nearly insuperable barrier between them and their scholarly customers (although faculty members undoubtedly need to help things along by making an effort to listen and co-operate). Perhaps not very surprisingly, her diagnosis still seems to hold true today, not only in academe, either.

Further to this, it is not an exaggeration to say that a good number of information workers are antagonistic towards their users – such antagonism is bred of long and close proximity. It is an old cliché, but for battle-weary information professionals, people get in the way of the systems they are so busily building and defending. It is hardly surprising, then, that students complain of negative experiences with librarians, of encountering preoccupied, rude, inconsiderate and discourteous service, in result of which they apparently often think twice before resorting to professional help in resolving their information problems (Hernon and Pastine, 1977; Swope and Katzer, 1972). In fact, the phenomenon of library anxiety (Mellon, 1986) has been traced back, among other factors, to students' negative perceptions of library staff (Jiao and Onwuegbuzie, 1997; Onwuegbuzie et al., 2004), the result of previous off-putting experiences. In all probability not independently of this state of affairs, interpersonal skills, especially oral and written communication skills, have emerged as primary requirements for library positions, as content analysis of job advertisements appropriate for LIS graduates indicate (Kennan et al., 2006). Concurrently, theoreticians of the profession have been advocating the need to pay much more attention to communication and interpersonal competencies and skills in LIS education (see, for example, Gorman and Corbitt, 2002). Unfortunately, it is quite plain to all those who teach such courses that for many students acquiring communication skills-based oral competencies presents quite some challenge.

No single or easy method of collecting the data

Plainly, the method of analysis is only going to be as good as the data that are fed into it. Therefore, the data collection methods have to be effective and robust. There is, of course, a rich choice of data collection methods, to the point of too much choice, which, in itself, may put people off. To complicate matters further, there is also the vital need to use different methodological approaches (methods triangulation). The prevalent view in the literature may very well be that it is use or information-seeking behaviour data which point most directly to the needs

experienced by people (see, for example, Cronin, 1981; Wilson, 1981). However, if we aim for effective information provision, we need to probe beneath the visible surface of people's actual behaviour, important as it is, into their needs, too (be they expressed or dormant, a point we will examine later).

Typically, since reality is subjective as well as culturally derived and historically situated (Crotty, 1998), phenomena need to be studied through the eyes of people in their lived situations (Hjorland, 2005), in an attempt 'to know what the actors [in a particular social world] know, see what they see, understand what they understand' (Schwartz and Jacobs, 1979, 7, quoted in: Wildemuth, 1993). This calls for qualitative research methods, which aim at studying people in situ, without constrictions of preconceived notions, so as to capture what their lives, experiences and interactions mean to them in their own terms and in their natural settings. However, not only are such methods very time consuming, but the practicalities of using them for obtaining data are such that the samples studied are by necessity relatively small and unrepresentative. Thus, crucial as the utilisation of qualitative techniques for determining people's needs is, it must be complemented by monitoring large (as large as possible) populations through the use of quantitative techniques. Only then can the veracity of the insights derived via qualitative methods be judged and put to a reality check, and the extent to which they are generalisable to a wider population established. By the same token, whilst quantitative techniques are invaluable for highlighting patterns of use and identifying broad sweeps of information-seeking behaviour, the data thus derived need to be further investigated for their validity and significance via quantitative methods. After all, quantitative methods tend to generate data of a more superficial nature: big numbers are there, but what it all means is not always clear. Therefore, it is truly vital that information need investigations combine quantitative and qualitative methods, for, as Hammersley (1981) and Greene and Caracelli (1997) suggest, the use of different methodological approaches (methods triangulation) serves to enhance the validity and reliability of the study by counterbalancing the flaws or the weaknesses of each method with the strengths of the others.

Expensive to collect the data

Another reason for neglecting information needs is that the data do not come cheap, at least not if you set out to attain an accurate and comprehensive picture, which, obviously, is the only way to do it. The problem is that providing the just noted necessary triangulation and reach necessitates the use of as wide a range of methods as possible, which comes at a (steep) price. The qualitative methods most suitable for running to earth information needs data (interview and observation) are typically very expensive indeed, although, luckily, the quantitative ones, offering the best insight into use, are often relatively cheaper.

In any case, as both qualitative and quantitative methods have to be used, information needs studies are a costly manoeuvre, especially as they have to be repeated regularly in these fast-changing times and, to make things even more

complicated, the financial trade-off is not immediately obvious. Can the conducting of information needs investigations really justify taking money away from the Book Fund (the information professionals' original ultimate threat) or the website (the new ultimate threat)? Can it excuse removing staff from critical front-line routines, like shelving books, managing the short-loan collection or updating the website? With these comforting thoughts information professionals renounce the information needs survey. However, what is the point of efficiently overseeing a collection or website that amounts to no more than a dim and distant reflection of the needs of its users? It is bound to be so in a world that has probably seen more change in information needs and information behaviour in the last five years than in the previous 50! How can we be sure that the information unit or system is heading along the right track?

It seems, then, that the correct and only view must surely be that it is too costly *not* to collect needs data. The stance must be that, in these dynamic and fast-changing times, it is wholly *economic* to collect these data on a regular and ongoing basis. Not that it is going to be easy to convince people of the wisdom of this. On a recent short course, when one of the authors explained that needs data was best collected by interview and that the interview might take 45 minutes, a number of participants audibly choked. The implication was obvious: that is a lot of time (which it indeed is). Still, is it too long a time to ensure that people get the service and information they need and deserve? Would anybody choke if they were told it would take ten hours to design a website?

Lack of a commonly understood and agreed framework of analysis

There are few easily understood and practical frameworks available with which to explore people's needs. Many of the works looking into the topic have tended to be too personal, as well as too theoretical and academic. Rather than clarify the situation, more often than not they have muddied the waters; rather than encouraging people to conduct user needs surveys, in fact they have provided an excuse not to do so. Maurice Line (1969; 1974) did propose a suitable method many years ago, but it has not been widely adopted. This guide embraces and expands upon Line's original model in aiming to rectify the situation. Therefore, it offers up an analytical, off-the-shelf method that can operate in the hurly-burly of today's high-tech (digital) information centres and units. It is a method for the systematic collection of information needs data to facilitate the design, assessment and auditing of information systems, which, although highly practical, is firmly entrenched in sound theoretical outlook and principles. The main intention is to provide a practical, usable and reusable method of analysing and evaluating information needs, which can be employed by information managers, information system designers, publishers, records managers and digital information consumers alike.

There are, then, a host of factors that bring about the ongoing neglect of information needs. However, none excuse it. Indeed, this state of affairs has to be put right. There are good and pressing reasons why it simply cannot continue.

So why indeed undertake information needs assessments?

1 Increased competition and competitive tendering from within and without the profession force information services to get closer to their customers (or go to the wall).
2 The huge growth in end-use and migration to the virtual space, which so unnerves information professionals worried about their jobs, should compel them to monitor closely what their customers are thinking, doing and wanting.
3 The challenge facing today's information professionals, the provision of custom-made, personalised information infrastructures, tailored to the distinctive needs of individuals, can only be met through a nuanced understanding of information seekers' idiosyncratic needs.
4 The huge and rising costs of introducing IT-based innovative systems demand that these systems be evaluated in the light of people's actual information needs.
5 The increased rigour with which information units are being audited and questioned about how they spend money require that information services collect data on their clients and outcomes to justify their expenditure.

Competition and deregulation

We are ten years into an information consumer revolution occasioned by the arrival of the internet (and, increasingly, the mobile phone), which is changing society, education and commerce on a massive and global scale. Digital consumers worldwide, numbered in their billions, are rapidly transforming today's ubiquitous virtual information environment through a preference for search engine exploration, a dislike of paying for information and a short attention span. Moreover, they use their new-found freedom to relocate themselves and their activities in this altered and incessantly changing information landscape, taking on many of the tasks previously reserved for information professionals. Indeed, information professionals and their systems are no longer the first choice of supplier for many people looking for information. As a direct result, the traditional information flagships are heading towards the reefs, blown there by the hurricane-force winds that have been generated by the digital transition. Public libraries appear to be in real trouble, but academic libraries, too, clearly grow increasingly decoupled from their user base as information consumers continue to flee the physical space (Martell, 2008).

Complicating things even further, libraries, the traditional agents for the preservation and provision of knowledge, are facing probably the fiercest competition in the history of their existence. As part and parcel of the inter-communal strife among the core content providers – librarians, publishers, journalists and television broadcasters (Gunter, 2008) – librarians and publishers, especially in the scholarly communications field, are on the brink of coming to blows over the possession of the keys to information. Their territorial disputes

and skirmishes, a direct consequence of the introduction of open access and institutional repositories, are bound to become even more ferocious with the inexorable growth of the popularity of e-books (Nicholas et al., 2007b). Thus, publishers and librarians already find themselves fighting for the spoils, with the threat of the loser being mortally wounded ever more looming over the horizons of both. At least for now, there seems a real possibility that librarians will be the ones to suffer defeat. Take, for example, the case of scholars, surely one of the most prominent information consumer communities: increasingly deserting the physical space for the virtual space, they move closer to the publisher and further from the library. As they become more and more anonymous and removed from the scene of scholarly information consumption, the librarians attempting to meet their needs grow correspondingly less and less informed about them (Nicholas et al., 2008b).

The conclusion to be derived from all this is truly inescapable: if information professionals do not get close to the consumer, others will and the information professionals will become an irrelevance (out of sight, out of mind). Whatever the reasons for the profession's neglect of information needs, it is surely the dearth of a robust and appropriate evidence base on information seekers that is responsible for the palpable danger of libraries becoming increasingly rudderless and estranged from their users and paymasters. It is clearly the reason why in this so-called information society, librarians, in particular, are increasingly marginalised, frequently under attack, in danger of losing their jobs and having their salaries depressed, with the organisations they so lovingly tend losing market share by the minute. The profession should look to successful businesses for guidance on how to survive the new climate. The advice of a past chairman of Kingfisher, the owner of B&Q, ScrewFix and other DIY brands, is certainly worth taking on board: 'We are concerned with meeting end-consumer needs. If you look at the success or failure of many organisations, the root of the problem often comes back to not anticipating how customer's needs have changed, and therefore not adapting to those changes' (Kay, 1994). Sound advice, indeed, and in the digital information world even sounder.

The end-user cometh and cometh again and again

Related to the above point about changes in the marketplace, there is also the huge number of connected end-users (digital consumers) now populating the information space. During the 1980s the profession constantly debated the outcome of end-use – some doubting whether it would ever happen and others forecasting the Apocalypse (how wrong could they be!). Since then, a veritable armoury of computers and modems have been built up in people's homes and offices, and, concurrently, the ultimate user-friendly tool – the internet – has become effortlessly accessible for all and sundry, so much so that, as Russell (2008) notes, access to the internet has become more the norm in a household than not. Indeed, the information landscape has been totally transformed:

millions and millions of people are connected directly to the information they need, courtesy of the ubiquitous search engine, on a scale that dwarfs any library, publishing or newspaper effort. The days of information seekers as supplicants are pretty much over. Who needs the traditional 'gatekeepers'? After all, people can now easily and expediently meet all aspects of their information needs on their own, and do so at any time of the day or night, too! The end-user is now king, truly the dominant player in today's information scene. No wonder disintermediation seems to be moving at such a rate.

The reaction of too many people in the profession to this upheaval taking place all around them is rather like that of a frightened rabbit in a car's headlights. They are paralysed by fear, possibly in denial. The appropriate reaction is surely to recognise that there is now much common ground, a common vocabulary, a willingness to discuss information problems; and the opportunity should be grabbed with both hands. The potential terrain for the information professional has increased enormously, although most of it lies outside the boundaries of the physical space. The key to mastering this terrain, of course, is information needs assessments. When you talk about end-users and digital consumers, you are really talking about information needs.

The challenge of custom-made, personalised information services

For many people the internet has resulted in the overnight transformation of an information-poor world into an information-rich one. We have moved from a situation in which information needs were rarely ever effectively met – certainly not without a tremendous expenditure of effort and cost, to one in which they are only too easily met, in theory anyway. To be sure, the internet fulfils information needs, triggers information needs and attracts people – information voyeurs and navigators, who have no needs at all. It has turned information seeking into a global (and fun) pastime. However, just broadcasting ever-greater amounts of information is not what it is about. There is a mistaken belief amongst the profession that the future is all about sharing information knowledge-management style, or storing and distributing information digital-library style; it is, in fact, about getting closer to what people need in the way of (instant) information and producing it in a processed, packaged form for individuals to consume at a particular point in time that they choose. Plainly, customisation, individualisation and segmentation in the information market are the next stage of the information revolution. Indeed, the Information Society will never become a reality until we can genuinely meet people's individual and special needs. There is still a long way to go.

However, although we talk glibly enough about the commodification of information, we seem to stop there: essentially, we – and the systems we provide – remain crude batch processors. At least for the time being, information products remain incredibly raw and general; even in the case of the web search, success depends largely on the lucky dip of single keywords or the input of a hieroglyph (URLs). This, when the future of information provision is surely

personalised information flows; it would be an extremely brave (or foolish?) person who would argue against that. How else do you get personal information other than from information needs assessments? Undoubtedly, in the information-rich environment in which we find ourselves, with undreamed-of quantities of information just a couple of mouse clicks away, we need to be even clearer about consumers' information needs than ever before. What else is going to help steer a path through the information jungle? Certainly not any of the current search engines. What else is going to ensure the precision of searching and the filters for the push technology that is coming our way?

Cost of IT-based innovations

With vast amounts of money being spent on novel communication and retrieval systems (intranets, websites, blogs, RSS feeds), mistaken judgements have increasingly serious and perilous consequences. Also, the more rapid the pace of change, the greater the risk of investing considerable sums in passing fads or enhancements divorced from real needs. It is only through an understanding of what information people need, how they prefer to set about finding that information and whether they achieve positive outcomes from their information seeking, that professionals can ensure that scarce budgets are used for the provision of suitable information systems; professional judgement alone cannot be relied upon, certainly not in these fast-changing times. A case in point is libraries' recent experiments with Web 2.0 interactive facilities. In an attempt to ensnare today's allegedly new brand of information consumers, used to involving, dynamic and personalised content experiences that can compete with the likes of Facebook, many libraries now have profiles on social networking sites. However, as we will discuss in more detail later on, at least for now, libraries' presence on social networks, as well as the blogs proudly sported by many of them, attract very little positive use. Obviously, costly innovative endeavours need to be based on people's actual information needs, otherwise there is a real risk of wasting money on systems and services that fail to meet their intended purpose.

Accountability and auditing

Traditionally, the quest for knowledge and learning was seen as an end in itself and, in consequence, its major (if not sole) supporting agencies, libraries, were generously provided for from public funds to enable their fostering progress and education. However, with knowledge and information becoming commodities of major value, the state's financial resources have become reserved for producers and propagators of 'knowledge for use', rather than 'knowledge for its own sake' (Calas and Smircich, 2001; Delanty, 1998; Duderstadt, 1997), with a subsequent decline in the budgets of the latter. Furthermore, driven by the rhetoric of 'quality', 'efficiency' and 'value for money' (Harvie, 2000), whatever government funding remains is made more targeted by allocation mechanisms

mimicking the market. Hence, public sector organisations are called upon to adopt 'new' management processes and systems for greater efficiency, even more marketisation and, especially, greater accountability.

Libraries, information centres and archives are thus part and parcel of today's value-driven environment, from which they are unlikely ever to escape. They are now on the same cost-conscious footing as any other business and, in result, they are subject to the same concerns, such as customer care, customer charters, cost benefit and the like. Whilst this is generally recognised as a *fait accompli* by the profession, few seem to realise that these concerns are nothing but a sham unless they are underpinned by the systematic collection of data on consumer needs and behaviour. How else is performance to be determined? Certainly not by the traditional measure of the number of documents on the shelves or new titles bought per year, or by the item that has superseded it: the number and power of computers on display. The yardstick unquestionably has to be changed, for it is only through customer satisfaction and outcome that success or effectiveness can be truly measured, and satisfaction and positive outcomes can only be obtained by meeting user need. It is customers who we should be proudly showing off, not computers, websites or shelves. Placing the customer care charter on show in the library, scattering a few complaints forms around and employing a number of staff to monitor a service is woefully inadequate, but it is too often the common response. No library that we know has a department devoted to the monitoring and evaluating of customer needs and usage of information, but they are taking on institutional repositories with some alacrity. These days, with value-for-money exercises having become the norm, there is really no other way to face up to the challenge they present, but through the careful collection and appraisal of information needs and use data.

All this seems to boil down to one point: the main reason for undertaking needs assessments must be that the information profession has neglected doing so in the past, a practice that has not got either the profession or the systems and services they provide very far; indeed, it continues to frustrate the progress of them both. Thus, unless information workers, from all walks of life, are reconnected with their user base, they are not doing the job they are meant to be doing: playing their part in supporting the information/knowledge society and economy. The message is loud and clear: the information community must stop thinking it knows best. The information consumer knows best. It is consumers who are calling the shots, so why keep them at arm's length? Bridging this gap necessitates, however, that we clarify to ourselves what information needs really are, which is the next question on our agenda.

2 What are information needs?

The concept of 'information need' is hardly unequivocal, although people talk about it as if it were, without ever bothering to define what they mean when they use the phrase. All the same, as Shenton and Dixon (2004) point out, citing a host of articles in ample proof of their assertion, there is a lack of a common understanding of the term 'information need'; indeed, the absence of universal agreement on the topic is a recurrent theme in library and information science (LIS) writing. True, there are very good reasons for refraining from the attempt to clarify the expression, for, when definitions are provided, they are often vague or highly complex in nature. Actually, they tend to cloud further some already muddy water and really serve very little practical use. Unfortunately, despite years of academic debate and much intellectual borrowing from other disciplines, like management and psychology, not too much has emerged that would aid information practitioners in their information needs deliberations.

Furthermore, people often talk about information needs when, in fact, they are referring to wants or use. Indeed, as Elayyan (1988), Green (1990) or Hewins (1990) contend, many studies that claim to be studies of information need are really studies of information use. However, while it is true enough that wants or use are both manifestations of need – and, as such, undoubtedly should be considered, they are neither identical to need, nor fully or accurately describe it. Thus, in order to attain a correct and comprehensive picture, we should be evaluating the need people have for information, the wants and demands they express for it and the use they make of it. Needs analysis may very well be fraught with difficulties and complexities, but still, we cannot always employ use indicators, such as website 'hits', as proxies. Therefore, as part and parcel of any definition, information needs have to be distinguished from some closely associated, but distinct information concepts, like want, demand and use, which are frequently (and sometimes deliberately) confused with information needs – to the general detriment of information provision and system design. We will attempt to do so in this section, which follows much of Maurice Line's original thinking, as set out in his 1974 article 'Draft Definitions: Information and Library Needs, Wants, Demands and Uses'.

Information needs: a working definition

For Line (1974), information needs were seen as the call for 'information [that] would further this job or this research, and would be recognised as doing so by the recipient'. Belkin and Vickery (1989) add that information needs arise when people recognise a gap in their state of knowledge, that is, when they experience 'an anomalous state of knowledge' and wish to resolve that anomaly. One can build upon these definitions by adding that it is the need for information that individuals *ought* to have to do their job effectively, solve a problem satisfactorily or pursue a hobby or interest happily. The operative word here is surely 'ought', the assumption being that for people to perform efficiently, effectively, safely and happily, they need to be well informed, that is, their information needs should be met. There is an implied value judgement in this – the meeting of need is beneficial or necessary to the person – and would be recognised as such.

Of course, people do not usually have information needs *per se*; rather, when they experience a problem or difficulty or are under some pressure, these cognitive and emotional needs of theirs may be met, or at least partially met, by obtaining and then applying some appropriate information. Indeed, information needs arise out of a desire to meet one or other of the three basic human needs: physiological needs (need for food, shelter, etc.); psychological needs (need for domination, security, etc.); and cognitive needs (need to plan, learn a skill, etc.). Thinking very much along the same lines, Norwood (cited in: Huitt, 2004) proposes that Maslow's (1954) well-known hierarchy of needs can be used to describe the kinds of information that individuals seek. Thus, individuals at the lowest level of the pyramid of needs, focused on their basic physiological needs, such as hunger, thirst, bodily comforts, etc., require *coping information*; individuals at the safety level, intent upon avoiding danger and ensuring their personal security, need *helping information*; individuals higher up on the pyramid of needs, at a stage where they are looking to belong, to affiliate with others, to be accepted, need *enlightening information* of the kind to be found in books on relationship development. Individuals on an even higher level of the hierarchy of needs, that of esteem, seeking to achieve, to be competent, to gain approval and recognition, need *empowering information*: information on how their ego can be developed. Finally, people who have reached the highest level of needs, that of a need for self-actualisation, for self-fulfillment and the realisation of their potential, seek sources (whether human or documentary) of *edifying information*.

All this, however, can in no way be taken to mean that information needs are any less important than the primary needs they serve; rather to the contrary, because success in meeting the one (the primary need) is dependent on meeting the other (the information need). True, the latter might be classified as secondary to the former, but in this increasingly information-dependent age, where information has obviously moved to centre stage, lack of information could certainly have serious, or even perilous consequences for the

individual. Yet, despite the pivotal role accorded to information in all spheres of present-day social and economic life, people's information needs may go unmet, either because they are unaware of having a need for information, or because, for various reasons of their own, they do not set out to meet a recognised (but unexpressed) need.

Unrecognised and recognised (but unexpressed) information needs

People do not always know what their information needs are. They do not know they have an information gap, for they are not aware that there is information out there that could be of help to them. They do not know that new information has rendered obsolete what they know and, in result, has given rise to another information need. It is only when exposed to the relevant information that the need is recognised. This might be called *dormant need* or *unrecognised need.* Take this case as an example: a person goes down to the photocopying machine to copy a letter and in the queue overhears a conversation about a television programme, screened the previous night, about globe artichokes. The person in question is rather fond of this vegetable, so listens with keen interest. Apparently, globe artichokes contain a lot of chemicals because of the infrequent rain (in Israel, where they are grown), and because the washing process fails to penetrate their tightly closed petals. Now, the person did not come to the photocopier with an information need, but goes away having obtained a needed piece of information.

Conversely, users may be well aware of their information needs, that is, their needs are by no means dormant/unrecognised, but, nevertheless, they do nothing about meeting them, either because they cannot or will not. A case in point is the all too familiar phenomenon of people refraining from pursuing their information needs for lack of time. Such non-use of relevant information, as Wilson (1993a, 1995, 1996) points out, may not happen by accident or by mistake even in academe. Rather, it often reflects a routine and normal approach for coping with the prevalent situation, in which the concurrent pressures of the constant dearth of time, on the one hand, and the huge quantities of available information, on the other, combine to instigate a policy of deliberate disregard of one's information need. In fact, even at the best of times today's information seekers tend to be satisficers (a term resulting from the blend of the two words 'sufficing' and 'satisfying'). That is, they stop information seeking after finding material that is good enough (Savolainen, 2007), so that they can juggle the need for comprehensive information with the constraints placed upon them.

Clearly, in today's internet-based information world, in which information is being generated in ever-increasing volumes and people are connected to information sources of unparalleled power and reach, taking a conscious decision not to attempt to meet one's information needs, at least not fully, is commonplace and will increasingly become more so. At the same time, the huge popularity of the internet must be at least partly due to the fact that it

has an unlimited potential both to uncover dormant information needs in the searcher and to solve recognised information problems expediently. However, turning to the internet, with or without a particular purpose in mind, frequently means relying on happy accident. This, in its turn, may come at some considerable cost: missing out on a vital piece of information. Thus, users cannot possibly count on this serendipitous method for obtaining all their information; it is too much of a lottery in these information-dependent times. The uncovering of dormant need, just the same as the efficient meeting of recognised need, has to be put on firmer and surer ground, but in this disintermediated age it is not clear who is going to do this: librarian, publisher or academic.

There is an opportunity for information professionals in all this: they are in many respects the best positioned to run to earth individuals' primary-needs-contingent information needs and to help them meet these needs effectively. They know what is available, on the one hand, and are able to control the information filters, on the other, so that exposure to information can be balanced with the problems that the availability of masses of information may bring about. This, of course, creates for them a positive and proactive role: where information needs are concerned, they are the experts. However, this is also where the big challenge lies for them: so far, information professionals have, at best, only really concerned themselves with satisfying the direct and specific articulation of information needs, but they need to go further if they are to win back a strategic place in the information chain. To attain the holistic understanding of their customers' information needs, wants, demands and uses, which alone can ensure that these needs are successfully met, they need to get very close to the information consumers and to remain steadfastly close. A clearer understanding of how information needs differ from information wants, demands and use, the issue we are about to tackle, may be a good place to start.

Information wants

Information wants are what an individual would like to have – *like* being the operative word here. Of course, in a perfect world information needs and information wants would be one and the same. However, we live in a far from perfect world, in which, for a variety of reasons, stemming from idiosyncratic factors of personality, time and resources, not all that is needed is wanted, and not all that is wanted is actually needed. Thus, for example, individuals may not attempt to meet their information needs fully, that is, may not strive to obtain all that they in fact need, for lack of time, skills or finances, or, alternatively, individuals may be tempted to obtain information that they do not in fact need (a prime example of this is the way people surf the web).

Nevertheless, people tend to equate wants with needs. Indeed, questionnaires, aiming to explore needs typically turn out to be want studies – with questions taking a 'would you like more information, more journals, enhanced facilities' line. Results can prove to be very misleading because users unreservedly tick

all the want choices on offer, happy in the knowledge that they will probably never be required to exercise their options (but just in case!). The following story amply illustrates to what an extent 'real life' can differ from wish-lists. One of the authors, at the time a practising librarian in charge of reader services, was requested over and over again to see to it that her library's opening hours were extended until 10pm. A bit hesitant to expend the considerable sums the move required, especially in view of the limited number of patrons on the premises by 7pm or so, she decided to ask the potential beneficiaries whether they were interested in longer opening hours. The results left little room for doubt: people were overwhelmingly in favour of the option. However, when the plan was actually realised, the author and her colleagues were more or less the only ones around in the evenings.

In any case, where information wants are concerned, we are moving into an almost wholly subjective domain, where, as it has already been noted, personal characteristics, available time and affordable assets make themselves felt. It is, of course, taken for granted that a price has to be paid if information needs are to be met – time, effort and possibly money have to be expended. Still, individuals may not be motivated to chase information, perhaps because these days information so often comes to them unsolicited and at no cost, or because they may not have the time to look for it or the skills to locate it, or, maybe, just do not have access to the necessary information resources (through lack of finance, perhaps). Job satisfaction must be a big determinant of whether individuals go ahead and attempt to meet their information needs fully. If you like your job, you will want to do it well, improve it and keep yourself up to date. This will inevitably mean going out of your way to meet your information needs.

Be it as it may, one point is hardly arguable: society sends us confusing signals about whether we should want information. On the one hand, we are enjoined to sample the joys of the internet (surely, *the* information-enabling mechanism), and on the other hand, in today's atmosphere of relentlessly increasing demands for accountability in everything we do, we have less and less time to enjoy the fruits of this easily come-by information. Moreover, when we succumb to the temptation to pursue our information wants, we may end up burying ourselves with information, thus digging our own information graves. So, should we, or should we not follow up our information heart desires; that is, should we actually demand the information we want?

Information demands

An information demand is a request for an item of information believed to be wanted. This is where information seeking starts, where the potential consumer first encounters the information system, source (human or documentary) or intermediary. However, people may demand information they do not really need, perhaps because their initial perception of its value does not match with reality: someone tells them it is a good site, but on arrival it turns out to be a

disappointment (lots of information seeking must lead to blind alleys, especially courtesy of search engines). By the same token, they certainly need or want information they do not demand, for instance, because they are not aware that it is there. Certainly the internet stokes up demand and leads inevitably to (very) large amounts of material that is demanded but not needed. Indeed, as the findings of the CIBER study into the use of scholarly journals (Nicholas et al., 2008b) clearly indicate, people download huge quantities of material, but not everything (possibly not much) that is downloaded is actually read or used. Much material is just squirreled away for another day, though that day may never come because of a shortage of time and the amount of squirreling that already has been undertaken.

Also, demand is at least partly dependent on expectation, which, in its turn, depends upon existing information provision. Indeed, as long as traditional libraries were the only game in town, customer expectations were notoriously low; after all, libraries only ever offered a limited window on information and could never respond within the tight time frames expected by most busy individuals. Also, as it has already been noted, there is a legacy of poor service in the profession. Add to this that many people are altogether unaware of what the information service can do for them, and the picture in its dismal entirety becomes all too clear. Seen in this light, the rush towards disintermediation (or doing it yourself) is less astonishing than it seems at first glance, especially as the advent of computers has raised user expectations enormously. Information seekers naively believe that these 'black boxes' can deliver anything and quickly, to boot. The web has raised people's expectations even more, to sky-high levels, in fact: plainly, it is seen as an inexhaustible source of (mostly) reliable information on anything and everything, which is effortlessly available 24/7. Indeed, web logs provide an awesome indicator of global demand for information, which, though, as we have already seen, is not necessarily synonymous with actual consumption or satisfaction.

Information use

Here we arrive at the more visible end of the information-seeking process – the information the individual actually uses or consumes. This is an area about which information professionals know most, but even here, not enough. So what exactly do we mean when we talk about information use?

First of all, use is both intended use and unintended use; that is, it may be the direct outcome of a satisfied demand, but, just as much, the result of browsing or serendipitous discovery, while not looking purposively for anything or when looking for something else. Not, it must be noted here, that browsing and accidental unearthing of information are invariably akin to unintended use, for browsing can be quite directed and structured. Some people browse because they have no choice; they cannot recognise and articulate their need until something they see reminds them of it. Indeed, humanities scholars are famously fond of browsing precisely because some newly encountered information may

uncover for them a dormant need, bringing about the fortuitous discovery of connections between ideas and words (Saule, 1992). People also browse because they are forced to do so: the manner, whereby the information system – the web, for instance – displays the information, leaves little room for locating information any other way – and logs point to this being the dominant form of navigating the vast virtual space. At any rate, the difference between intended use and unintended use is an important one in terms of information system design. For this very reason, usage studies should really make an effort to distinguish between the two.

All in all, use is a word that comes with a lot of baggage, beyond the fact, which has already been pointed out, that it is certainly not the purported clean, hard, direct manifestation of need, which it is so readily assumed to be. To begin with, it is not all that easy to determine when the use being witnessed can be counted as 'real' consumption of information, for 'use' can and does refer to, at the very least, two clearly distinguishable levels of use. The first level of use simply involves determining whether something is worth using in the first place, for, obviously, use and satisfaction do not always go hand-in-hand. Thus, information seekers clearly need to establish at the outset whether a given item of information amongst a vast sea of data will satisfy or fail to satisfy need; but can we consider their actions, typically measured by transaction log analyses, as constituting 'use'? Probably better considered as power browsing. Still, this type of use might lead to other people being alerted to the potential worth of the information consumed, so perhaps it is use, after all! The second level of use is the actual consumption of information, subsequent to its having been found relevant, that is, the actual putting of some information to purpose-relevant use. This type of use, generally more effectively measured by citation studies, probably corresponds more directly to the popular concept of 'real' use. However, even use data of this kind can really tell us very little about many of the key needs characteristics: thus, for example, the need for information presented from a particular viewpoint, approach or angle certainly cannot be gleaned from usage data.

If the problem of what can be considered 'real' use and what cannot is not complicated enough, there is also the dilemma of how use is to be measured. After all, use has many recorded manifestations. What can be taken then to be an indicator or record of use? Citations and logs have already been mentioned, and there are also issue statistics, library loan studies, book sales records and the ubiquitous tick boxes in questionnaires, which seek to find out whether people used a particular information system daily, weekly, monthly, etc., during a given period of time (mysteriously, always assuming that information use is a rhythmic or periodic activity). Plainly, many of these use indicators are measuring different phenomena, each of which leaves plenty of room for various interpretations.

Consider, for example, what constitutes use on the web and what can be read into it. Putting aside for the time being the problems of actually determining use on the web (we will come to that presently), let us try to establish

what 'hits' (pages viewed) really signify. For a start, with the loose and idio-syncratic method of searching on the web and the shotgun approach of most search engines to retrieval, the chances that you actually want to see the specific page you end up with have to be relatively low. However, you 'used' it and you are recorded on the logs as having done so, and action will be taken by others – advertisers, sponsors and web managers, on the basis of these data. Furthermore, how many times do you navigate through a site, going down numerous pathways, to get what you really want? Each page you go through on your way to the page you really need is another page 'used', but not actually needed.

Obviously then, use data are really very problematic, and need to be handled with great care, but are generally not. True, bean (use) counting does come stripped of the wish-list or fantasy factor that is so endemic of questionnaire surveys, which is an undoubted advantage. Also, data on use are generally to hand and plentiful, two very attractive features indeed where there is a chronic shortage of time and funding (and where is there not?). However, as we have just seen, use data have to be treated with caution. Also, perhaps even more importantly, use data are too crude indicators of need to serve as a comprehensive enough foundation for services aiming at the meeting of information needs: whilst use can be a manifestation of need, an information need is, in theory, greater than demand plus use. Moreover, as people can only use what is available, use is very heavily dependent upon provision and access, albeit this is less of a problem these digital days. Perhaps most disturbingly, where use studies are concerned, non-users – whose number can often amount to quite a significant percentage of the population, are not taken into the equation. Non-users may loom larger in certain fields and users may form a small, self-selecting group, but they may also prove to be a more financially attractive and/or influential group. We know, for instance, that senior managers have traditionally been the ones that have shied away from using information systems, as a senior partner of an accountancy firm, trying to explain why he never searched for information himself, told one of the authors who came to interview him: 'I know how to make coffee, but I don't'.

Evidently then, use data can only offer a partial view of need. Even when augmented by demand data, use data can only help an information system improve on what it is already doing, but since there is no guarantee that it was on the right lines to begin with, this is of limited value only; use data will not help build a system which will provide new services and solutions. Thus, use data may be very welcome for measuring the usage of what is provided, but it is no substitute for needs data in establishing whether what is provided is what is best. Plainly, the case for basing effective information services and systems on a holistic view of the need people have for information, the wants and demands they express for it, and the use they make of it is very strong. Before we leave the topic of information use, though, let us consider the term 'user' itself.

The digital consumer (yesteryear's user, reader, customer, client, patron …)

User, along with its various synonyms in the professional literature – reader, costumer, client, patron – has traditionally been used to denote people who might avail themselves of an information service. The expression has never been too fortunate a choice of words (for one, it was employed to describe non-users, too), but, in any case, it is surely *passé* by now: user (and users) has lost much of its meaning. In many respects it is a tired, cheap, over-used and misused word, which provides the information profession with a debased currency. The word 'users' conjures up a picture of a featureless mass, a homogenous body – people who are accustomed to being fed (print-based) information in batch-processing mode, 1950s-style. It fails to reflect the close, complex and virtually incessant engagement that takes place between people and information in today's digital world, where knowledge, perceived as the key to success in all walks of life, is a major democratic right and leveller. Basically, then, 'user' is the wrong word, in the wrong place, at the wrong time.

What we really need is a more accurate term, much richer in meaning, which acknowledges the multi-dimensional relationship between an individual and the internet-redefined, vastly widened and 'viewing' information environment. 'Digital consumer', the term suggested by Nicholas and his colleagues in their recent book of the same title, could fit the bill, and 'digital consumer', rather than the more specific descriptor 'digital information consumer' it should be. This, because, as Nicholas et al. (2008a) contend, in today's information realities a digital consumer is, to all intents and purposes, the equivalent of an information consumer. True, people visit the multi-purpose, encyclopaedic virtual space that is the internet for many different objectives, much the same as they would go to a bricks-and-mortar superstore looking for goods, services, new experiences, titillation, excitement and amusement. However, as their pursuits on the internet invariably involve choosing or buying e-documents or information services, it is now almost impossible to say what is information and what is not; what is information seeking and what is not. Take the example of e-shopping: as Russell (2008) explains, first a person is a digital information consumer and then an e-buyer. Thus, people shopping at an e-store will be using the internal search engine to find what they want, navigating through the site, employing browsing menus and opening another window on a cross-comparison site, to make sure they are getting value for money. Only when the information-seeking component of the shopping process is successfully completed will they actually purchase the item they need. Thinking very much along the same lines, in a recent *New York Times* article on the growing reliance on the internet for health information Schwartz (2008) argues as follows: 'As patients go online to share information and discuss their care, they are becoming something more: consumers. [For instance,] Amy Tenderich, the creator of Diabetes Mine has turned her site into a

community for diabetes patients and an information clearinghouse for treatments and gadgets … '

'Digital consumer', not 'digital information consumer' it should be, then, but why this specific term? Well, to begin with, because 'digital consumer' is very much an internet-type word, which the word 'user' most certainly is not. The internet, with its own rich and picturesque language, is so very much a part of everybody's life today (at least in the so-called 'developed' countries), that there is a pressing need to get our professional jargon in-line with its vocabulary. We really cannot ignore the call for employing the 'right' language, for it is only by doing so that information professionals can address the much larger information audience that the internet commands.

Also, perhaps more importantly, the term 'digital consumer' is a truer representation of things as they really stand, namely, that it is the individual, rather than the system that now holds centre stage. Thus, while users of a bygone era used information systems, today's digital consumers explore the information space (cyberspace); while users were supplicants, standing outside the system, looking (beseechingly) in, digital consumers are part of the system. Indeed, the digital consumer has come to play a much more dynamic, complicated, creative and engaged role in the evolution of the information domain than the user of yore could ever have imagined. In fact, the digital consumer is now King, actually driving the changes in the virtual space with a wholly novel style of information seeking: frenetic, promiscuous, volatile and intent on the pursuit of quick wins. Exchanging the old term for the new one thus acknowledges this shift in power from information producer to information consumer.

Perhaps it is the adoption of this more accurate terminology that will help to convince librarians, publishers and media moguls alike – indeed, anyone who manages large, centralised, inflexible, batch-processing-style information factories – that their users/customers/patrons/readers have really flown the coop. Having become digital consumers, rather than passive users, they are no longer the captive audience of the past, wholly dependent on the information providers' goodwill; these days they are the ones who call the shots. After all, they have a huge digital choice and can quickly vote with their feet (mouse?). In fact, as Gunter (2008) points out, they will have even more choices in terms of sources of information about commodities and services. Thus, thinking about people as consumers is a key step towards delivering the right services to meet their needs. Information professionals had better take these relatively new developments to heart, challenging as they may be, or ignore them at their peril. If they opt for the latter, they run the risk of information seekers completely abandoning them and going their own way; there are plenty of convenient enough alternatives to the traditional information services. With disintermediation gaining momentum by the minute, the writing is there on the wall, for all to see …

3   A framework for evaluating information needs

Having delineated in some detail the whys and wherefores of the imperative to undertake routine collection and analysis of information needs data, we now come to the thorny question of how to go about it. The secret to it all seems to place the slippery concept of information need in a comprehensive, clear-cut and understandable analytical framework, which is precisely the form of scrutiny offered here. Not that the parameters of such a framework too readily come to mind, for it is far more difficult to describe the characteristics of information needs than, say, those of housing needs. This is probably to do with the fact that information needs arise from other needs: as such, they are more likely to be accorded less individual thought and consideration, and, in result, their characteristics are not so easily remembered or disentangled. Still, despite the fact that information needs are perceived as less concrete and more diffuse, just as you can describe the key characteristics of housing need as being: building material, site location, type (apartment, semi-detached), number of rooms, architectural design/character, and age/period of the property, so too can the characteristics of information need be described.

Thus, it is possible to identify 11 major characteristics of information need: subject, function, nature, intellectual level, viewpoint, quantity, quality/authority, date/currency, speed of delivery, place of publication/origin, and processing and packaging. These characteristics combine to form a comprehensive evaluatory framework, for it takes the holistic consideration of the different attributes of an information need to provide a truly fitting answer to a problem encountered. Suffice to cite the example of the unsuitability to most UK-based high-school students of some information, which may be right on target subject-wise, but, say, highly scientific in its level, 20 years old and in Chinese, to demonstrate the point. It is important to mention at this juncture: although the portrayal of the various facets of need may take a lot of words, this does not mean that using the framework is a long and laborious task. It is the understanding that takes the time. Once the form of analysis is mastered, then the various sections outlined below shrink into a headings checklist.

The 11-pronged framework proposed here can, thus, ensure that information delivery is consumer-centred, targeted, personalised and relevant. Indeed, it

can be profitably used at both a macro level – for effective strategic information management planning, and at a micro level – for the efficient carrying out of routine enquiry work and consumer online searching. More specifically, the framework can be put to use for the following purposes:

1 Laying the foundations for the design of personalised information services by benchmarking the needs of different information communities and making comparisons between them.
2 Monitoring and evaluating the effectiveness and appropriateness of existing information systems from a consumer perspective.
3 Detecting gaps in information service provision and remaining vigilant to changes in need necessitating modifications, adjustments and fine-tuning.
4 Aiding the assessment of the never-ending tide of new information products.
5 Ensuring that one-to-one information service encounters are set on a firm footing and conducted in a systematic and comprehensive manner.
7 Bringing the information consumer and the information professional closer together (something that is inherent to the information needs assessment process put forward here).
8 Providing an information literacy training tool for the enfranchised, but untrained, digital consumer.
9 Offering a self-help guide for the e-citizen who wants to maximise the benefits of the information cornucopia. Students of all subjects would particularly benefit.

Subject

Subject is probably the most obvious characteristic of information need, central to nearly all information need statements. Indeed, it is probably the one feature most readily coming to mind for describing an information need. Libraries arranging their document collections by subject and search engines providing access to the world's information resources by the means of keywords is a testament to the importance accorded to this aspect of information need. This is not to say, though, that describing information needs in subject terms alone will lead to wholly satisfactory outcomes. As has already been noted, only if all relevant attributes of the information need are considered, can a truly appropriate solution to the problem be found. Still, while subject is not the sole aspect of significance in the portrayal of an information need, it is plainly a very important one (hence Google's interest). Unfortunately, it is also one that often defies our attempts to get it right, for the successful matching of a person's subject requirements with the 'right' information is far easier said than done, and no wonder: subject requirements vary so with the idiosyncratic circumstances of each and every person, or even with those of the very same person at different times.

Inter- and intra-individual variations in subject requirements

In today's multidisciplinary and multitasking world most people undertake an ever-increasing variety of roles, each with its own subject requirements. On the job front alone an individual might function in a number of capacities. Thus a university lecturer might have teaching, counselling, consultancy, administrative, professional and union responsibilities, and, probably above all, research obligations. In consequence, people need to concern themselves with quite a few subjects: mastering some, learning a bit here and there about others; keeping in touch with some, occasionally (if at all) revisiting others. It is not necessarily the individual's lead role(s) either, where the most pressing need for information is. Indeed, it could be argued that the individual would already have ensured that arrangements were satisfactory in this department, for example by putting to good use one or other of the widely available current awareness/alerting/RSS services. It is elsewhere – outside the mainstream interest – where initiating 'jewel hunting expenditures' in a library or on the internet may be deemed necessary. Thus, for example, a focus group of academics attended by one of the authors, convened to discuss and reflect upon scholarly research behaviour, said they followed developments in their specific areas of interest mostly via subscriptions to alerting services, which sent them journal tables of contents (TOCs) on a regular basis; however, outside their immediate fields they considered Google searches the better option for the purpose.

What makes the whole matter rather tricky, though, is the fact that every role people undertake requires them to have detailed knowledge on some things and a broader understanding of many other things. Obviously, people have different information needs and different methods of meeting these needs in areas outside their fields of expertise, in which their knowledge is limited or even altogether non-existent, and in which they are certainly not as well versed in the literature, the methodologies and the jargon. Take again the example of the information needs of university lecturers. As Menzel (1964) points out, in their particular role of researcher alone, each scholar's area of attention comprises several fields or sub-fields arranged in concentric circles: the primary field of attention, at the centre, is to be kept up with in full detail; the secondary fields, at varying distances from the centre, are also to be kept up with, if not in the same detail; and fields towards the periphery merely warrant knowing about progress made. A similar differentiation holds true where university lecturers' other roles are concerned, too: in some areas they follow the advances made pertaining to their teaching in minute detail, while in others they dip into the material every now and then, just to gain a smattering of understanding of current professional issues. Thus, for instance, one of the prime teaching interests of Library School lecturers – on which they spend a lot of time and have done so for many years – is online searching. It follows that they would be grateful for everything newly written about, say, Dialog, a system they usually teach in depth. However, the same lecturers must also keep in touch with developments in the broader Library world, for they are training students to

work in libraries. The need here, though, is for general, contextual data only: a general item on the financial problems faced by libraries would be acceptable, but perhaps only the one, not too detailed, item would be sufficient.

Furthermore, not only do different roles and endeavours call for material in diverse subjects and in a range of detail, but they may very well require information that also varies as to the extent to which it delves into a subject. It is very much a question of how deep the interest lies. True, these days people's 'concentric circles of interest' are getting smaller and smaller in span, as one academic explains: 'You know that you are unable to cover the whole field, and in consequence you concentrate on specific issues … you deal with the trees, rather than with the whole forest'. However, at the same time, these circles of interest grow ever-more specific. In fact, the developments in this direction seem to form a self-perpetuating circle, for the need to cope with the huge quantities of information being constantly generated seems to dictate, as much as to originate in this ever-growing specialisation. Thus, a salient fact of contemporary life seems to be a focusing of interests and, as its direct derivative, a focusing of information needs. Another academic, also testifying to a focusing of his research interests and his information needs, elaborates: 'You have to specialise, otherwise you won't be able to cover all the knowledge in a given field, [and] because of the specialisation you need to inquire more deeply into your subject, you have to know more about it'.

However, just how specific a request is may greatly vary from person to person, or even from situation to situation for the very same person. An interest in, say, organic gardening might entail a need for some quite broad-spectrum information, if the person is a newcomer to the subject area – 'The Manual of Organic Gardening', for instance – or for some very specific, in-depth information, if it is for the use of an expert – such as 'The No-digging Approach to Potato Cultivation'. By the same token, a visitor to a health website may sometimes search the site for a specific purpose, looking for exhaustive information on a given subject, whilst at other times the same visitor turns to the site simply to browse the general health news. Of course, information seekers often have a very good idea of just how specific and detailed the material they are looking for must be, but does the intermediary or system know? If this is not untangled at the outset, either a flood of dense information is unleashed on the unwary, or the supply of information is choked off to the needy.

To complicate matters even further, people's interests and responsibilities change – nothing is set in concrete. Rather to the contrary, as a casual comment of a physicist, musing aloud on his information needs, seems to indicate: 'My area changes so much, that I constantly have to spread out to more and more domains, and in consequence I need to know increasingly more. It's not the way it used to be; in the past you accumulated the information you needed in the first few years of work in a field and from then on you only needed to keep track of further progress made. Now it's the other way round: with the passing of time I need to know more and more of things I have never

needed to know before'. It seems, then, that routine and regular monitoring of the subject premises upon which the increasingly more widely available current awareness/alerting services are built must be carried out if they are to maintain their effectiveness. Virus-checking programmes today often update their scanning lists daily and information systems should take their lead from them. Updating needs profiles should be conducted at the very least once every six months to maintain their effectiveness. Once a quarter is better.

Having seen how differences in personal circumstances entail discretionary information needs and uses both on the inter-individual and the intra-individual level, we now come to the biggest challenge of all: the matching of a person's subject requirements with the 'right' information.

Locating pertinent information on a subject

The root of the difficulties encountered when attempting to find appropriate information on any given topic seems to lie in the problems associated with effective subject description. The conversion of the need for information into terms that adequately clothe its subject sounds easy, but it is not: people are unlikely to furnish all the terms the information system needs for it to produce the goods, or, even when prodded (by pop-up boxes, for instance), to provide the most productive terms. It all boils down to the problem of translating user-generated keywords into the retrieval language of the information system, which purportedly has been solved with the introduction of structured access to information. However, as the huge popularity of keyword-based information seeking irrefutably proves by now, bibliographic description and controlled subject access are not held to be as crucial to information work as librarians would have us believe. In fact, information seekers have long been 'voting with their feet' in manifesting quite some reluctance to locate the information they need on a given topic in the methodical, bibliographic-tools-based fashion wistfully recommended by information professionals. Thus, for example, Palmer and Neumann (2002) observe that academic researchers in the humanities are renowned for their propensity for serendipitous locating of information despite the plethora of secondary searching tools at their disposal, to which King and Tenopir (1999) add that science and social science researchers too have been found to prefer less systematic methods of information retrieval.

Perhaps not surprisingly: first of all, as Stoan (1984) points out, no subject heading or descriptor can adequately analyse a book or an article for the reader, as bibliographic access tools to the literature introduce another layer of human minds through which information must be filtered, evaluated, classified and labelled. Also, locating information through the use of the controlled vocabulary of a catalogue, an abstract or an index necessitates lighting on the 'right' subject terms, which for all practical purposes depends on the seeker's ability to second-guess correctly the indexer's choices. Happening on the 'correct' term is clearly difficult enough even in knowledge areas, most

notably the sciences, where the information content of a publication is definable in concrete and universally accepted terms. It is obviously a far more formidable task in fields of a less predictable terminology, such as the humanities and social sciences, in which vocabulary is conventionally assumed to be fuzzy and hard to pin down. It is for this very reason that a university-based political scientist argues so hotly against unconditional reliance on alerting services in his efforts to keep up with new developments in his areas of interest:

> I'm constantly on the lookout lest I fail to spot relevant material, as I can never be sure that I can correctly predict an indexers' choice of terms! Just last week I almost missed an article of importance because it was under "prejudice" rather than "xenophobia"!.

There is, of course, a time-honoured (but increasingly less-used) solution to these problems of relevance and precision in defining the subject of an information need: professional assistance. Take, for example, the attempt to solve an information problem on a one-to-one basis. An information specialist will spend some time, ideally in the presence of the customer – not so easy in the virtual world – scanning thesauri and sample issues of secondary services, often coming up with quite satisfactory results: the broad-narrow and related-term networks will provide an excellent word map in which to place a topic, resulting in sufficient terms to effectively cover a subject. Yet, even enlisting the help of an intermediary does not guarantee problem-free information retrieval, for clients seem to find it very difficult indeed to pinpoint the subjects of their information needs.

One frequently encountered problem is that people, in their attempt to communicate to the information professional or system the subject of their concern, generalise the query. Mostly they do this in order to ease the way of the intermediary into what they consider a complicated and intractable problem that they have spent a long time considering, but to which the intermediary comes cold. Interestingly, people appear to do the same with remote information systems, too. However, there are other reasons for generalising requests: to provide browsing room to allow for the inadequacy of the keywords; to simplify things for the intermediary, who is not necessarily a subject expert; to minimise the risk of early rejection and to provide space for negotiation, in case the intermediary can only offer limited assistance; and to get a prompt reply by means of a short, perceived then as simple, question. Information workers, indeed telephone helpline operatives, all have their own pet examples of hopelessly general questions, such as the following gem: a request for books on fish in a public library. Now, in a public library there are books on catching fish, cooking fish, the biology of fish and fish as pets, to name just the most obvious possibilities. Take, too, this real-life query: 'I wonder if there is any information on new cars?' The actual requirement was for dealers' prices for the Honda Civic.

Closely related to general question-framing is vague question-framing – sometimes the two are indistinguishable. Confidentiality concerns can lead to

people cloaking or camouflaging their interest from the intermediary or system (in case it is monitored), so that others are not alerted to their particular line of enquiry. This can happen in the information centres of newspapers, especially in those that serve journalists from a number of papers, as is the case at News International and the UK Mirror Group. More often, vague subject specification mirrors the users' own confusion and uncertainty as to what they want: it is difficult for them to verbalise their problem, although they will recognise what they want when they see it. After all, users are asking for information to fill a gap in their own knowledge: this must inevitably lead to some imprecision in the formulation of the query.

Not that all these are insurmountable problems; far from it. In point of fact, information professionals learn early in their careers to identify such obstacles to subject specification and deal with them effectively in the reference interview. Unfortunately, though, in these disintermediated days, where most searching is conducted remotely and anonymously, they have fewer and fewer opportunities to put their expertise to the test. Even before the digital information world became for many people a far superior alternative to the print-based traditional library, turning to the information professional was perceived as time consuming and labour-intensive compared to researching information independently. So much so, that even researchers, for whom attaining the right information at the right time is absolutely crucial, were found as long as some 35 years ago to shun professional librarians (Meadows, 1974), and this to such an extent that, apparently, they were 'prepared to consult almost anyone, except a librarian' (Line, 1973, 33). By now, this reluctance to use the services of an intermediary seems to have become an overpowering trend: in our age of ubiquitous desktop access to massive quantities of information on any and every subject, help yourself is very much the name of the game. After all, who needs information professionals performing their feats of conjecture to alight on the right subject descriptors when all it takes is typing in a keyword or two? Searching is easy, is it not?

Well, as it happens, it appears not. Thus, as Bates (1998) contends, in study after study, across a vast variety of environments, it has been found that for any target topic, people will use a wide range of different terms. Two examples help to illustrate the nature of the problem. A research academic, searching for material published on the topic of people doing work 'on the side', i.e. without the various government authorities knowing, uses the term 'moonlighting'. However, a comprehensive trawl of the literature would soon uncover more terms: second economy, underground economy, black economy, black market. All of these terms will have to be employed if the search is not to become a lottery, which much searching is. Similarly, a search for material on the elderly – a relatively simple concept one might have thought – is, in fact, even more problematical with the following possible alternatives: retired people, old age, the aged, senior citizens, pensioners, old people, old persons. It gets much more complicated than this when two or more concepts are involved.

Still, people seem to favour greatly what is in effect the epitome of the shallow thinking characterising 'trial and error' behaviours: the portrayal of information need through keywords alone. 'Google is doing a great job for us these days … you only have to know how to search, how to choose the key-words', says a computer scientist, and his psycho-oncologist colleague joins him in extolling the wonders of the technique: 'I search for information by trying various word combinations which I think will get me to the information I need, all sorts of word combinations, until I find the ideal combination'. However, she seems to be unique at least in one respect: these days not even scholars construct searches with many terms in them. Typically one-third of users enter one word in their search statements, about the same proportion two words, and only the remaining third enter three words or more (Nicholas et al., 2008b). To be sure, today's information seeking is very different from that prevailing in the hard-copy environment of the late 20th century. Searching is no longer a serious activity in terms of thought, preparation and execution. No doubt, subject keywords are perceived as providing the easy and quick fix, although, of course, they fix a little as every user of a web search engine is only too aware.

Indeed, as the use of search engines is fast becoming the first-line option for tracking down pertinent information, there is a proportionate increase in the retrieval of irrelevant documents. It is easy to see how this comes about: today's information seekers, steering clear of 'superfluous' bibliographic access tools or professional support because 'it is so easy to conduct an information search on the web', forgo the benefits of accurate and comprehensive analysis of need, which alone can guarantee relevant and precise results. Also, the ease of use characterising search engines comes at a price: much of the material procured by this form of wide-angled (shotgun) searching will inevitably be irrelevant.

Apparently, then, the awareness that much of the material served up in spades by a search engine will probably be found irrelevant does not deter people from its use; they are quite happy to trade failure for convenience. This dovetails neatly with the waning of the hue and cry characterising the early days of the web concerning the problematic nature of locating information of relevance via search engines. Thus, for example, gone are the (not so long ago) days, noted in study after study (Kibirige and DePalo, 2000; Massey-Burzio, 1999; Voorbij, 1999; Wang and Cohen, 1998; Zhang, 1999), when academics consistently clamoured for librarians' professional intervention to remedy the problems seen as emanating from the lack of bibliographic description and controlled subject indexing of the information to be found on the web. Indeed, the CIBER studies into the use of various e-information platforms, most notably in the areas of health (Nicholas et al., 2007a), scholarly journals (Nicholas et al., 2008b) and scholarly books (Rowlands et al., 2007) find time and time again that whatever the specific audience, users tend to shun on-site menus, complicated interfaces and myriad search options, opting instead for search engines. Thus, the majority of users (the proportion normally varies

between about two-fifths and three-quarters) find a relevant site through the use of Google or other search engines. Anecdotal evidence gleaned in the aforementioned focus group, discussing the information seeking of academics, indicates that even among researchers, with very specialist, sometimes esoteric fields of inquiry, the notion of discipline-specific databases seems to have bitten the dust in favour of convenient but incomplete generic services like Web of Science and Science Direct. The attitude appears to be that 'if the information isn't found there, it's not worth looking for'. This mirrors the behaviour of their undergraduates; it is just that they look for something a bit more 'select' than Google. No wonder then that, as Russell (2008) points out, there is a marked increase in search engine-based information retrieval: in 2007 approximately one-third of internet users utilised a search engine to find a site, even if they had visited the site before, whereas in 2003 this was just one-quarter! Furthermore, as the evidence amassed in CIBER's Virtual Scholar research programme (2001–08) indicates, this is just the beginning, for the younger the information seekers, the more likely they are to tackle an information need by keyword-based retrieval via the use of search engines (Nicholas et al., 2008b).

Not that this growing popularity of search engines, bringing about the habitual reduction of an information need to a few haphazardly chosen keywords, is all that surprising. Keyword searching may not be the most efficient information retrieval method, yielding, as it usually does, a considerable amount of 'noise', but it does provide the information seeker with a much wider and disparate view of what is on offer – more titles, older material, from more subjects. Also, it is certainly a far more convenient method of information retrieval than using an arbitrarily chosen descriptor or subject heading assigned by a third party. Finally, perhaps most importantly in today's hurried times, when 'fast, easy and trouble-free' is so often the overriding consideration in everything we set out to do, the use of a search engine serves the ultimate goal of the information consumer, the simplifying or short-circuiting of the information-seeking process. In addition, using a search engine is costless, except for the time spent – and it only takes a typical internet user 10 seconds to check out a page (Nielsen, 2000), and, courtesy of associated advertising, you might even find something to buy.

Thus, search engines offer the prospect of trouble-free, targeted and direct access to meet an information need, providing as they do massive choice in response to a query of a word or two. No wonder they are perceived as offering a relatively sophisticated search facility for people with limited knowledge about either information retrieval or the content sought. In this respect they might have become the digital equivalent of the 'returned book shelves', the place where the tried, if not necessarily proven items are to be found. The quest for effortless searching is so pervasive that young people even exhibit a strong preference for expressing themselves in natural language rather than analysing which keywords might be more effective (Williams et al., 2008).

The ramifications of this inexorably growing preference for unmediated and uncomplicated information activity may be quite far reaching. On the most basic level, as it unmistakably emerges from the substantial evidence base amassed by the CIBER research group over the years, search skills and levels of digital literacy are (largely unacknowledged) problems for a considerable number of people. There seems to be some disturbing data which indicate that much use appears to be passing and/or ineffectual, and could possibly constitute a 'dumbing down' in information-seeking behaviour. Thus, lots of hits are just searchers passing through; they put in the wrong word and got to the wrong place. Take the example of an investigation of the BBC website (Nicholas and Huntington, 2005), which found that a considerable number of users made input errors when entering their search queries, and often did not notice what they had done until the search results were displayed. When participants noticed spelling suggestions (i.e. 'were you looking for … ') they often welcomed them, although several users scrolled directly down to the results list. The recurrent indications of poor/limited searching do make one reflect on the effectiveness of information literacy strategy and programmes. It would indeed be ironic if the web that provided for the enfranchisement of the user in information terms was also guilty for the disenfranchisement of whole swathes of the population unable to take advantage of the information deluge.

Furthermore, this may be the first inkling of a major change underway in today's information consumption dictates. The fact that search engines, unlike browsing mechanisms like content, subject and alphabetic lists, do not require information seekers to have any prior knowledge of the formal literature, its structure and hierarchy might well mean that future users will bring with them less knowledge of the scholarly system, of which libraries are currently an important component. Seeing that most libraries were originally – and still are – designed so that their contents could be browsed first, this is quite a fundamental shift, which is not yet reflected in information provision.

Function (use to which the information is put)

People frequently need information for achieving the vast variety of their role-, task- or interest-dictated goals. The inevitable outcome of this state of affairs is that each individual puts information to work in diverse ways, contingent on the specific circumstances in which the need for information arises. Take, for example, the different functions information fulfils for individuals in their professional capacity alone. It begins with the organisation or professional community to which they belong: since the end products of each such organisation/community are distinctive, so are, as you might expect, their uses for information. In the case of journalists information is used to write stories; in the case of social workers it is often used to answer resource questions concerning their clients; and in the case of academics it will be used to root a new inquiry in its context, help compile a lecture or update a reading list. Further to that, within each profession (and organisation) the prime function to which

information is put will vary according to the role and specialism of the individual. Thus, managers in social work departments would be using information to monitor the progress of the organisation, rather than to answer the resource questions of the client group. Nevertheless, there are some generalisations that can be made about the functions to which people put information. Essentially, people need information for six broad functions or purposes, and it is very important to distinguish between them, for they require very different information solutions. They are: (1) getting hold of answers to specific questions (fact-finding function); (2) keeping up to date (current awareness function); (3) investigating a new field in depth (research function); (4) obtaining a background understanding of an issue/topic (briefing function); (5) procuring ideas or stimuli (stimulus function); and (6) looking for interesting titbits of information just for the fun of it (recreational browsing function).

The fact-finding function

Very often indeed, people need information simply to obtain answers to specific questions. These questions are familiar to all reference librarians: they are of the 'who, why, what, where, when and how' kind. Such questions may be straightforward, like the address of an organisation or individual, a biographical portrait; or complex, like the number of aircraft near-misses that occurred in 1987. The huge popularity of such fact-finding tools as Wikipedia and the now electronically accessible 'The Statesman's Yearbook', and the enduring high regard in which the still print-only 'Whitaker's Almanack' is held, bear testament to the strength of the need amongst end-users.

Indeed, everybody has this fact-finding need and for most of us it is a recurrent, perhaps everyday need; many of the queries of e-shoppers, for instance, fall into this camp. Even practised and experienced researchers, who know their literature very well, frequently seem to encounter those gaps in their knowledge (Bernal, 1959) that send them in pursuit of a piece of necessary information: a bit of data, a method, the construction of a piece of apparatus, an equation … Luckily, the need is relatively precise and well-defined, generally met by facts, names, addresses, statistics and the like. Not a lot of information is involved in meeting this type of need and the interchange between user and information system/intermediary is consequently brief: therefore, such needs are by and large easily and cheaply met. They are also easily delegated.

There is a lot of evidence to suggest that nowadays fact-finding needs are met almost universally via the web; it seems to be everybody's handy encyclopaedia and telephone directory, and then, of course, there is Wikipedia. Hardly surprisingly, of course: this is simply part and parcel of the by now truly widespread tendency to regard the internet as the first-line source for meeting all information needs. Thus, for example, in the aforementioned comprehensive survey of the use and impact of key digital health platforms and services in the UK (Nicholas et al., 2007a) it has been repeatedly stated that the internet is now the first source consulted, with people saying that other sources were consulted

only 'when I can't get what I need from the internet'. This was true even for one respondent working in a location where there was a medical library:

> I usually first try to find relevant info [sic] on the net, because it is easier than getting hold of hard copies of the same or similar info. If the net can't offer enough, then I will try to get the information from medical library at work.

Obviously, where easily solvable problems, necessitating only some fact finding are concerned, the convenient accessibility of the truly wide-range information on the web does indeed render it the prime option for resolving painlessly and quickly the occasional disruptions to workflow or thought processes caused by problems of this sort. Yet, interestingly, it is precisely this need for specific information which can at times still bring people to the physical library, despite its plainly diminishing importance in the eyes of many (Martell, 2008). Thus, for example, as it has been shown by Herman (2005), humanities scholars still regard the library as their primary option for fact-finding purposes, even though turning to a colleague or searching the internet could leave them comfortably seated at their desks. Apparently, as Brockman et al. (2001) point out, their investigations often raise questions pertaining to details which can be found only in lesser-known primary documents or secondary sources. These stand a better chance of being located in a traditional library, for, in many of the commercial full-text and indexing products available in the humanities the marginal and the esoteric are ignored in favour of the canonical and the influential.

The current awareness function

Moving on to another use to which information is put, we now come to people's need to keep up to date, to follow the new developments in their areas of interest. This is also a generally widely felt need, especially in today's knowledge- and information-based society, but in some fields and professions the concern is much more pressing. In fact, as Wilson (1993b) suggests in his essay on maintaining currency, for the large class of knowledge workers – i.e. knowledge producers (those active in research and development) and members of the professions – the requirement to keep up with one's field is an ethical requirement, sometimes even dictated by the law. However, social pressure strongly reinforces the demands of ethics and law: people do not want to appear to their peers to be behind the times, because that is likely to expose them to contempt. This is how a biologist puts it: 'I need to know what's going on so that I don't turn out to be the laughingstock of the field, proposing a 'new' project five other labs are already working on'. Further to that, Wilson (1993b) adds, there is a kind of 'logical' pressure involved, too: keeping up with professional advances made is, of course, plain common sense, as people's principal assets are likely to be their stock of specialised knowledge, which, for them, makes keeping up to date nothing less than a form of self-

preservation. Literally so, apparently, as a neuro-biologist, mincing no words in getting across the message, explains:

> Keeping up is one of the measures you take in order to safeguard yourself … otherwise you've no way of knowing what goes on in the world, and in science it's truly critical and essential that you do. If you want to survive, you've got to do it, if you don't know other people's work you're as good as dead.

Maintaining a hold on what is going on is a particularly essential part of professionalism in dynamic fields, characterised as they are by sudden, frequent and widespread change. This is plainly the case with all journalists, and enormous amounts of money are spent assembling complex information systems to enable them to keep track of events in real time, although they have been helped considerably in this by social networks, like Twitter and Facebook. In academia, too, it is the fast-moving areas that require constant vigilance over developments, although in the case of scholars, it is the research advances made that need to be followed, rather than events. Indeed, keeping up to date is considered a must in all scholarly fields, as an academic, an expert in philosophy, elucidates:

> Keeping current is very important. You can't conduct research … if you are isolated from other people's work; that is, you can, but it will be much less efficient … You can't discuss a problem without acknowledging that somebody wrote something on it recently, you can't write an article on a problem which somebody else has already solved … Only if all researchers determinedly keep pace with the developments in their fields can scholarly progress be guaranteed, otherwise they'll re-invent the wheel over and over again.

Still, in matters of keeping up with the developments in one's field it has long been demonstrated that all researchers are definitely not created equal. Not that they differ as to the theoretical importance accorded to keeping abreast of new developments in their areas of interest; it is rather the definition of keeping up which differs from discipline to discipline and, in result, so does the pace of the activities aimed at attaining currency. Indeed, there can be little doubt that 'making every effort to keep current' does not mean the same for scientists, social scientists and humanists.

Scientists consider keeping current an ongoing task of great urgency and no wonder: as their fast-moving fields are characterised by rapid changes, unless they are invariably up to date, they are liable to put in jeopardy the successful outcome of their research efforts. Not realising that some research has already been done, they may repeat it and, as a computer scientist puts it 'you turn out to be an idiot … because you have wasted your time, and because people will tell you: what, didn't you know?' Moreover, in addition to the waste of energy and resources and the loss of face involved, ignorance of new research developments

may also slow down their progress to the point of thwarting a claim to priority of a discovery (Becher, 1989; Garvey et al., 1970; Garvey, 1979; Price, 1986).

In comparison, social scientists can afford to adopt a more relaxed attitude to the need to maintain currency, since, as Line (1973) contends, the penalty for refraining from doing so is not as severe as in the case of the science researchers; with the circumstances of the typically empirical research projects of the social sciences differing from place to place and from time to time, it is not too likely that any findings would be discarded altogether because somebody else 'got there first'.

As for the humanist researchers, they are altogether complacent about the whole issue of keeping current, far less concerned than their scientist and even their social scientist colleagues with making sure that that no new contribution in their subjects escapes their attention immediately upon publication. Thus, although they too are mindful of the need to keep pace with new developments, they are believed to contentedly adopt an 'if not sooner, then later (or even much later)' frame of mind to the whole issue (Fulton, 1991; Stone, 1982; Wiberley and Jones, 1989). Stone (1982) links this relative tranquillity attributed to the humanities researchers, with regard to following the progress made in their areas of interest, to the nature of scholarship in the humanities: since the humanist researcher's innovative contribution to knowledge can consist of different perspectives or different understandings of the same work and might not present any new 'facts', awareness that others have worked or are working in the same field is less important; there is small chance of actual duplication occurring and it may not matter much if it does, so long as each presents an original interpretation. Indeed, at first glance, this danger of reinventing the wheel does not seem very relevant where humanities research is concerned. After all, who dare claim that the ultimate word on, say, Hamlet has already been said? However, the following story, recounted by a professor of literature, clearly demonstrates that in the humanities, too, the danger is there:

> Having re-read Jane Eyre I realised that there was a recurring motif of opening and closing of doors and windows in it. Rather happy with my insight, I wrote a long article on the subject, only to discover that somebody else had come up with the idea long before I did and there was a very good article on the subject, far better, unfortunately, than mine. There was nothing to be done about it; I literally threw the article into the bin.

Evidently then, humanists, too, have very good reasons for following the scholarly advances made in their fields.

Not very surprisingly, then, seeing to it that current awareness needs were adequately met was a major concern of the information community up until recently, and with good reason. The huge amounts of information generated, often referred to as the information explosion (a phrase nobody uses any more), did make the systematic monitoring of advances achieved quite complicated, even with close support forthcoming from computerised information

systems. Thus, despite the ready availability of Selective Dissemination of Information services – regular, pushed information services based upon user-supplied keywords that represent ongoing interests, now increasingly being replaced by RSS (Really Simple Syndication) feeds – maintaining currency does necessitate a great expenditure of time and effort on the part of the individual to be spent on vetting and digesting the data. Still, in the above-mentioned survey of the roll-out of digital consumer health services in the UK during the period 2000–05 (Nicholas et al., 2007a), one user type identified was people who rated medical news and research highly and wanted to keep up to date. In fact, almost one-quarter of the users were on the site simply to browse the general health news, rather than to search it for a specific purpose. By the same token, the ways of a psycho-oncologist, who five years ago still reported seeing to her needs for up-to-date information in a fairly regular manner 'because information becomes obsolete in a fraction of a second', were quite the norm. No longer, though.

Logs, as well as interview and focus group data tell us that people are not keeping up to date as they once did: with increasing time and resource pressures at most workplaces and in many households, current awareness is no longer a discrete or regular activity. Thus, for example, in the just-quoted survey of digital health services in the UK (Nicholas et al., 2007a) doctors testified that at least some of their professional updating was done in response to a specific need: on the comparatively rare occasions when they were faced with complaints that required information beyond their personal stock of knowledge, they turned to the internet to aid them in forming a diagnosis and advising the patient. That it should be so is perhaps not wholly unpredictable: after all, the need for keeping up is inevitably vaguer than the need for facts; therefore, it is an information activity that may more easily be put on the back burner, dropped or conveniently overlooked. Also, whereas fact-finding is usually associated with immediate and often urgent problem-solving – and hence has to be dealt with speedily – this is not the case with current awareness. There is often no direct pay-off; the effect is much more long term.

Indeed, even academic researchers, who, as it has just been noted, are very conscious of the need to keep up, do not, as a rule, routinely invest time and effort in proactive information seeking aimed at learning of new developments. True, as Herman (2005) points out, the range, variety and frequency of researchers' activities aimed at keeping current are determined by the level of awareness deemed necessary in their disciplinary milieu to the work being done by others. Thus, for example, in fast-moving areas, where the disciplinary culture dictated norm is the ongoing exchange of pre-print based information among researchers, users seem to appreciate greatly the benefits of obtaining early intelligence via e-print repositories, as an academic, specialising in high-energy physics, describes:

> First thing in the morning, I check the new articles posted overnight. I'm addicted to this, I spend on average between half an hour to an hour each

morning checking if there's something new and interesting, or something which may link up with what I've been working on … Since everybody, from students to the most valued researchers, sends the results of their work first to this archive, and only later to some journal, this is all I need to keep up.

However, this is really quite the exception to the rule, for nowadays, contrary to widely held notions, academics seem to follow the progress made in their respective fields with what can be termed as 'serene interest', but no more than that. Thus, with the notable exception of perusing e-journal TOCs when these land on their desktops, they tend to update themselves when the need arises, tacitly relying on search engines to deliver 'current awareness on demand'. Indeed, libraries' attempts to interest their patrons in innovative alerting services, such as RSS feeds and the like, meet with very little success, if any (Nicholas et al., 2008b).

It seems, then, that these days keeping current is a tactical, problem-driven activity rather than the strategic, time-driven concern that it used to be. Presumably this is because keeping a finger on the pulse of the developments presents less difficulty in an electronic environment, in which the information tap is always in the on-position in any case.

The research function

Researching a new field in depth is a far less frequent and widespread concern. Most people encounter the need to review the existing knowledge on a topic only occasionally and irregularly, but, given the realities of contemporary life, it does not come as much of a surprise to find that this is on the increase. Indeed, the insatiable demand for a multi-skilled and mobile workforce in a world based on knowledge has turned us all into permanent students, if not amateur scholars. There can be little doubt that these days, when life-long study, as well as training and retraining have become customary and taken for granted by large segments of the population, assembling a solid information base in preparation for embarking on a new undertaking (and not necessarily in work-related circumstances, either) is fast becoming routine for many people. It is certainly no longer the prerogative of those in research and academe! Indeed, the survey of digital health platforms and services in the UK (Nicholas et al., 2007a) lent further support to the notion, held by various health information providers and researchers (London, 1999; Eysenbach and Diepgen, 1999, to name but two) that the internet is exploited at a deeper level than that which might be expected of lay users, with websites intended for medical practitioners accessed by non-professional consumers alongside healthcare workers. Thus, for example, one respondent commented: 'I prefer sites, which are written by Doctors for Doctors [sic]. I also like to see references to the research papers that back up articles'.

This increasingly more prevalent need for anchoring a topic in its information context often necessitates extensive coverage of the knowledge existing on it, a state of affairs which is not invariably problem-free. True, the computer-aided accumulation of data allows for the ongoing creation and easy management of a dynamic, growing knowledge base, which is so wide-ranging, that seekers of information have at their disposal an unprecedented array of information resources. Also, computer applications can be of immeasurable help in locating, accessing and retrieving relevant information. Still, the time required to absorb and put to good use the data resulting from a simple search for the information foundations of a given subject can reach such enormous proportions, that a seemingly insurmountable problem is entailed. Fortunately, though, the need to research a new field in depth does not equal a need to assemble all the existing information on a subject, not even as an ideal goal to which to aspire. This holds true even for academics, who have to base their investigations on steadfast information underpinnings just as much as anybody else, perhaps more, but, in addition, also need to identify the lacunae in it for further investigation. Indeed, among the academics interviewed by Herman (2005), while a few did testify to aiming for obtaining the maximum information coverage possible on the eve of a new research project, others were more in agreement with a neuro-biologist, who said that she only aimed at 'getting the general drift of things':

> I had quite a few arguments with colleagues [on the thoroughness with which the literature needs to be reviewed]. Each time I mentioned considering a new topic for investigation, I was told: 'Go read the literature first'. And I keep insisting that I'm not interested in the literature when I can see with my own two eyes what's going on … I don't hold with performing a full review of the literature … it should suffice that you know in general terms what has been done, and what hasn't, for you to go ahead.

It seems, then, that whilst it is all too easy to mix up a good search with a big search – and none come bigger than those associated with the research need, the two are by no means synonymous. The novice or naive user (and intermediary) may be particularly prone to pursuing such information needs much too exhaustively, but this really need not be so. Establishing at the outset of an information-seeking expedition the scope of comprehensiveness and thoroughness required can simplify matters considerably, whether it is end-users searching on their own or an intermediary searching on behalf of somebody else. Moreover, as Wilson (1995) suggests, evaluating the search results and then intentionally ignoring some of the potentially relevant information to be found on a subject, which is really a routine and normal strategy of research work, can be a very effective way to cope with the problem. Obviously, this tactic can only work if the decisions taken as to which items to use (and which to skip) are based on knowledgeable assessment of the quality, authority and value of the information found, rather than resorting to

the popular practice of picking the first item(s) produced by Google, which, of course, only proves the vital importance of the much-discussed need for information literacy training.

One last point: it is important not to confuse a fact-requiring question for a research one, because the information outcomes are as different as they can be. The trouble is that the information need, which seems to occur most readily to people, is the need for finding all relevant information on a topic. As it has already been noted, even academics no longer consider it essential to locate 'everything' known on a subject under investigation, but, nevertheless, their instantaneous association of a need for information often seems to be with ostensibly gathering literally everything published on a given topic (Herman, 2005). The root of the problem seems to be the fact that so much Information Science research concerns academics. They really are a prominent group of information users, but they also commend themselves because they are a pliant and orderly population, who, to boot, are easily accessible for a university-based Information Science researcher. The consequence is a stereotype of information need, which, although no longer accurate even for academics, is carried over to the population as a whole. No wonder, then, that when graduates of Information Science programmes come to work in a practitioner environment, they tend to opt for information solutions more appropriate to academic information problems, both in the planning of information services and day-by-day work, much to the detriment of their adequately meeting their clients' information needs.

The briefing/background function

As previously mentioned, not everyone has the time or need to research a field in depth. However, many people need a briefing on topics with which they are broadly familiar, but perhaps insufficiently acquainted with the detail, and sketchily and fleetingly need to be. For most of them this need is probably met by the newspapers, which thus perform a key briefing function.

Generally, the broader the subject interest (which, inevitably, equals more information), and the less the time available, the greater the need for the background brief. Journalists, for one, caught between the need to say something authoritative about almost anything and with very little notice and time to do this, are great practitioners of the background search – cuttings traditionally fulfilled this need until the advent of digital newspaper archives. Another group of users frequently in need of briefing are politicians. They are expected to have a view on anything, so it is not surprising to discover that one of their preferred information forms is the background papers produced by The House of Commons Library. Indeed, this is a prime example of information professionals anticipating information needs, and an area where they can really bring their skills to bear.

Ironically, the web – the chief culprit for the information deluge we are all experiencing, is a great briefing source. Thus, for example, the overwhelming

majority of the seekers of health information in the survey of digital health services in the UK (Nicholas et al., 2007a) said that the information found on the web had helped them in understanding more about an illness or injury, and more than a quarter of them reported that the information found was even sufficient, in their judgement, to meet their health query and substitute for a visit to the doctor. Coupled with the ability 'to search on a wide range of topics – much more than a local library, and to consult a range of sources and perspectives', it is hardly surprising that for many of them the internet fulfilled effectively enough the briefing function, providing as it did what one user described as 'information that just wasn't available before to normal [i.e. non-medical] people'. However, patients bringing in internet printouts or discussing information they had acquired through this medium is not without its drawbacks and dangers, as one health professional explained: 'we are inundated with half heard or understood fragments of information from the television or half read magazine articles'. In result, added his colleague, 'usually the doctor has to allay fears or iron out misconceptions'.

Perhaps somewhat unexpectedly, today's academic researchers have briefing needs, too, especially when they participate in a scholarly endeavour attempting to coalesce the kind of multifaceted, and often inter- or multidisciplinary expertise that one researcher working alone cannot always provide. Thus, they may on occasion look only for the basic level information needed to aid them in understanding the wider picture within which the specific point being investigated is embedded, as a psycho-pharmacologist explains:

> If I need some information in a subject … I'm no expert on, I'll look for a source, which can provide me with the information on a sufficient level to answer my question. I won't delve ad infinitum into each subject, because I won't see the end of it.

However, they have their own strategies for fulfilling this kind of information need (Herman, 2005), although they, too, utilise the web for the purpose.

Apparently, more often than not a brief discussion with a colleague suffices to point the researcher in the right direction. Obviously, the best option is asking a renowned expert on the subject, which is indeed the course taken by the more senior people. An archaeologist, for example, a prominent authority in his field, has no qualms whatsoever as to the right way to proceed when he finds himself up against a need for some background information: 'I'll contact a friend of mine who is sure to know the answer, I'll just send him a quick e-mail, he'll get back to me in no time, and that's that'. His economist colleague, also of a standing in his field, takes much the same course of action, but he will insist on talking to the colleague who can provide the information: ' … it'll cost the university for half an hour or an hour of a transatlantic phone call, but I'll have my answer on the spot'. However, life is not as simple for the novice academics, as a young philosopher explains:

> Sometimes, when you don't know the answer to a question, you're well aware that there are others who do. However, then it becomes a question of … the people you are in contact with in your day to day activities: if you work in a central place, where the action is, you just ask your colleague down the hall, but if you're not …

Clearly, the problem boils down to a researcher having the right professional contacts, for in this day and age, courtesy of the ubiquitous e-mail, technically everybody is 'down the hall' from everybody else.

Well, if having the answer straight from the horse's mouth, so to speak, is not feasible, there is always recorded knowledge, which is indeed an often-cited solution among academics. However, these days this is not invariably the preferred option, as a neuro-biologist explains:

> If in the midst of work I need some information, I leave everything and dash to the library to get it. Though nowadays it's best to search the internet, it saves going to the library, looking for the specific volume, which may not be on the shelf, and even if it is, you have to go to the photocopying service, where there is a queue … Also, often a little e-mail, which you write in a second, can save you all that trouble, you just ask a colleague to send to you the information.

Still, when academics require a briefing on a topic, turning to the literature is a prime option, for they have to hand review articles and extensive literature surveys designed for such purposes. These handy summaries of the state of the art are obviously of considerable time-saving capabilities and, as such, are very popular indeed among researchers. True, as Herman (2005) found, the practice of making do with a synopsis of the achievements on a topic in lieu of assembling the original contributions comprising the information base of a new research project is held, at least notionally, in very low esteem indeed. Even in the sciences, where it should be definitely feasible to do so, for in the 'harder' knowledge domains one reads to discover the outcome of somebody else's research (as opposed to reading for retracing the discovery and analysis at the core of the research in the 'softer' areas), a summary of advances made is rarely taken to suffice for laying the information foundations of a new exploration. Still, the more pressured the researchers are, the greater is the likelihood that they will rely on literature reviews. Indeed, in these days of Research Assessment Exercises and comparable institutional evaluation measures, which exert ever-increasing pressure on academics to produce as much research output in as short a time as possible, literature reviews are becoming much more important for scholars; they are considered a particularly useful strategy for getting up to date and avoiding too much reading.

Having seen how information is put to work to fulfil the need for acquiring some basic knowledge on a given topic, we now proceed to another of its uses: procuring ideas or stimuli.

The stimulus function

The creative aspect of many a new undertaking feeds on information, which can often serve as a stimulating agent, the source whence initiatives for an original venture may hopefully ignite. This role of information as a fund of inspiration is nowhere as pronounced as in the case of academic researchers, intent as they are upon contributing new knowledge and understanding. To be sure, it is their relentless search for the next problem to pursue which so frequently accounts for their ongoing concern with access to information, and with very good reason, too. As Palmer and Neumann (2002) contend, scholars prime for future discoveries: by working at maintaining a high level of interaction with a wide variety of information, they develop a state of preparedness for new discoveries, a conclusion, which, of course, echoes Pasteur's well-known saying, 'chance favours the prepared mind'. This is how an archaeologist puts it:

> You read the literature, you know what you know, and all of a sudden you realise that something looks at odds with what's held to be true, something doesn't fit into the existing body of knowledge. So you think about it, and then think about it some more, and perhaps come up with an alternative solution, which provides a better explanation. The existing knowledge is the trigger, yes, indeed, always, there's no such thing as not relying on previous information. Even if the idea occurred to Newton because the apple hit him on the head, it was his previous knowledge, which enabled him to come up with it.

Hardly debatably then, information serves a particularly vital purpose in sparking off fresh ideas and spurring new projects or investigations; the only question is how exactly it happens. However, the processes that lead people to light on new insights and/or formulate problems of consequence seem to defy attempts to describe it (Schwartz, 1992; Bath University Library, 1971a). What we do know is that these processes, including the utilisation of information for the purpose, are obscure and intensely personal. Indeed, whilst information seekers generally know what they are looking for, it is not so when they are on the lookout for inspiration. Here they have only the vaguest idea of what they are trying to find – and sometimes no idea at all. In fact, they interact with information sources in the hope that this will result in their discovering just what it is that they need; so much so, that quite frequently it is by seeing something that they do not want that they are perversely alerted to something they do want. No wonder, then, that surfing the web is a prime example of the stimulating effect information is capable of affording. The full-text documents the web liberally offers, with their hyperlinks and idiosyncratic natural language indexing, can in fact uncover all kinds of dormant information needs by providing unusual and unexpected associations of ideas.

Consequent to this state of affairs, the information seeking associated with this particular need is unavoidably unfocused and unstructured, which often leads intermediaries who observe it to the mistaken belief that what they are witnessing is poor searching. Much of the rubbishing of end-users' searching skills found in the professional literature results from a poor understanding of the characteristic information behaviour that results from individuals trying to meet this kind of need. However, it is only through this kind of vague and wandering exposure to great amounts of information that users can discover what they want and uncover their dormant information needs. This is why, as we have already noted, humanities scholars, whose manner of conducting research is highly individualistic and subjective, are so fond of browsing. As Saule (1992) points out, their seemingly aimless examining of the catalogue, scanning titles of books in the stacks or skimming a document sometimes results in the fortuitous discovery of connections between ideas and words. Of course, what is meant to stimulate can also irritate – and it can all too easily happen that the individual is overloaded and/or led on a wild goose chase. Still, it seems to be a small price to pay for fulfilling an information need that can hardly be resolved in any other way.

Evidently, if people seeking novel insights and fresh ideas are to fulfil their rather vaguely defined need for information, they have to rely on serendipity. It is perhaps not very surprising to find, then, that people do not often deliberately set out to discover stimulating information; rather, they seem to be on the alert while they browse around on the web or on the library shelves. Still, scientists, with their ongoing awareness of what the truly significant issues at the frontier of the developments in their fields are, do report that from time to time they scour the literature with the express purpose of locating topics for new research (Herman, 2005).

So now we come to the last of the uses to which information is put: enjoyment, pure and simple.

The recreational browsing function

Information seeking is seldom seen as an end in itself; rather, it is viewed instrumentally, as a means towards achieving one's role-, task- or interest-associated goals. Still, unearthing interesting titbits of information just for the fun of it is a much-loved pastime: people while away many a happy hour rummaging among the displays in a bookshop, leafing through books, journals and magazines in a library and, perhaps most notably, surfing the web. Obviously, as Das et al. (2003) maintain, they regard such a non-goal-directed information-seeking activity as its own reward, a form of entertainment, relaxation and escapism.

There is, of course, nothing new about this leisurely use of information, but the ubiquitous and effortless availability of the web, with its enormous information reach and its ability to provide undreamed-of quantities of information from one's desktop, have made it more widespread than ever. Thus, for example,

Fallows (2006) notes that almost two-thirds of American users (78 million people) testify to having gone online to browse the internet for no particular reason, just for fun or to pass the time, a percentage that has held about steady since the Pew Internet & American Life Project, which is a daily tracking survey on Americans' use of the internet, began asking the question in 2000. Indeed, she says, surfing the web has become one of the most popular activities that American internet users will do: nearly one-third of them go online on a typical day just to hang out. In fact, compared to other online pursuits, browsing for fun now stands only behind sending or receiving e-mail (52% of American internet users do this on a typical day) and using a search engine (38% of American internet users do this on a typical day), and it is in a virtual tie for third with getting news online (31% of American internet users do this on a typical day). Not that the huge popularity of recreational information seeking on the web comes as a surprise for anybody living in our times (at least in the industrialised world); we all do it, all the time. We have all become information voyeurs, who set out to look for information when we have no needs at all. Of course we do: these online sessions of wandering around on the web just for the fun of it offer a smorgasbord of information, attractively spread out for the consumption of a tasty bite here and there, wherever the fancy takes us.

Interestingly, this manner of information seeking is by no means reserved for recreational purposes; rather to the contrary, as we are about to see in the course of our discussion here. However, first we need to see how people obtain the information they put to use for the variety of purposes we have just enumerated.

Coping with the call for information in an era of abundant choice

Having examined in some detail the different functions information fulfils for individuals, as well as the distinctive information solutions each requires, all that remains now is the question of 'how'. How do people go about meeting the considerable range of their role-, task-, interest- or entertainment-associated needs in today's plentiful, if not over-abundant information world? The answer emerges loud and clear from the vast evidence base that the CIBER research group have assembled over the past few years. It is a very, very different form of behaviour than the one that might be expected on the basis of reading the classic information-seeking texts of Ellis (Ellis et al., 1993; Ellis and Haugan, 1997) and Wilson (1999). This is partly because we have undergone a massive paradigm shift in information-seeking behaviour since they developed their ideas, and partly because it has only recently been possible to observe information-seeking behaviour on a huge scale and in minute detail. Indeed, the digital consumer revolution requires us to consign to the bin much of what we have been holding to be true about people's ways and means of meeting their information needs.

In result of being given digital choice, present-day information consumers manifest a widespread, pronounced, endemic form of digital information-seeking behaviour, best described as 'bouncing', although the terms 'flicking' or

'hopping' would equally do. This is a form of behaviour where users view only one or two web pages from the vast numbers available to them and a substantial proportion (usually the same ones) generally do not return to the same website very often, if at all. Thus, for example, all the CIBER studies showed that around 55%–65% of e-journal users typically viewed no more than three pages in a visit and then left; the studies also showed that around half of all users did not return or only returned after a prolonged gap. This suggests a promiscuous, checking-comparing, dipping sort of behaviour that is a result of being provided with huge digital choice, search engines constantly refreshing that choice and a shortage of time that results from so much to look at. In this respect the behaviour is best seen as being akin to television channel hopping using the remote control – you flick around alighting on things of interest and when the interest fails or wanes you flick to something else.

In addition, information seekers seldom penetrate a site to any depth; on average, most people spend only a few minutes on a visit to a website, insufficient time to do much reading or obtain much understanding. This, as the extensive literature review undertaken in preparation for the Google Generation project (Williams et al., 2008) clearly indicates, is very much in line with other research based on observations or surveys. Thus, for example, observational studies have shown that young people (boys especially) scan online pages very rapidly and click extensively on hyperlinks – rather than reading sequentially. Users make very little use of advanced search facilities, assuming that search engines 'understand' their queries. They tend to move quickly from page to page, spending little time reading or digesting information and they have difficulty making relevance judgements about the pages they retrieve. When all this is put together with the bouncing data it would appear that we are witnessing the emergence of a new form of 'reading', with users 'feeding for information' or 'power browsing' horizontally through sites, titles, contents pages and abstracts in their pursuit of quick wins.

Plainly, this novel style of information seeking, frenetic, promiscuous, volatile and viewing in nature, is eminently suited to browsing for recreational purposes. Indeed, it bears a close resemblance to the characteristic behaviour of an e-shopper confronted by the cornucopia of shopping opportunities offered by the web. The problem is that when the search is goal-directed, whether it is for getting hold of answers to specific questions, for keeping up to date, for investigating a new field in depth, for obtaining a background understanding of an issue/topic, or for procuring ideas or stimuli, it can be construed to point to inadequate information skills resulting in negative outcomes (not finding what you need/want). However, in today's new information realities people expect instant gratification at a click, looking for 'the answer' rather than for a particular format or source. At the same time, they seem to await entertainment even when they set out to meet their formal, work- or study-associated needs, looking for involving, dynamic and personalised content experiences that can compete with the likes of Facebook. No wonder they scan, flick and 'power browse' their way through digital content!

This state of affairs clearly poses an enormous challenge for librarians. Forced by circumstances to compete for the attention of information seekers, who seem to be turning their backs on the library as a physical space at an alarming rate (Martell, 2008), they are increasingly conscious of the need to find innovative ways to enhance interest without impeding the effective meeting of people's information needs. Hence, in an attempt to adapt to the change they perceive in information-seeking behaviour, many librarians have started to experiment with Web 2.0 interactive facilities. Paying homage to the astonishing success of Facebook, MySpace and YouTube with the young (and not so young), and cognizant that if you are not certain of your brand or presence, it might be possible to obtain this by association, libraries now have profiles on social networking sites. It is too early to tell whether this kind of initiative will eventually bear fruit, but for now library presence on social networks as well as the blogs proudly sported by many of them seem to account for a small proportion of use only; most people concentrate on mainstream, traditional bibliographic activities. Thus, for example, Williams et al. (2008) cite a 2007 OCLC survey, which asked both college students and members of the general public how likely they would be to participate in activities on a social networking or community site, if built by their library. The responses left little room for doubt: clearly, neither group was much interested, with the numbers of those who said they would be 'extremely likely' or 'very likely' to do so coming to about 6% or 7% in most cases.

By the same token, at the above-mentioned focus group of academic researchers, the idea that an e-journal database should adopt Web 2.0 facilities went down like a lead balloon. They could not see the need for this or how people had the time to indulge in what they clearly thought were side-show activities. 'There is enough to do without engaging in blogs, wikis, RSS and the like', was the general message coming from the group. It seems, then, that at least for the time being social software does not have much to contribute to the rebuilding of relationships with users in an increasingly disintermediated environment. Perhaps not surprisingly: as Nic Howell (cited in: Sherwin, 2008) points out, 'social networking is as much about who isn't on the site as who is'; thus, when libraries and museums start profiles on a social network, its 'cool' brand is devalued. Anyhow, it is not very clear that any such attempt can be the answer to libraries' current plight: in 2008 the number of (British) Facebook users fell for the first time, and MySpace and Bebo visits were down too (Sherwin, 2008).

In any case, concurrent to these valiant attempts to harness social networking to the reinstating of libraries, librarians also concentrate a lot of effort on what is considered home ground for them: information retrieval. Having picked up on the popularity (and problems) of search engine searching, they are taking Google and the Google generation head-on by developing their own engines to stop user flight. The latest initiative and, maybe, the last hurrah, is federated searching, also known as meta-searching, broadcast searching, or cross-searching. The various federated search tools allow users to search simultaneously multiple library databases, catalogues, multimedia sources and other

collections via one common Google-like interface, and to get the collected results in a succinct and unified format devoid of duplication (Cox, 2006). In doing this librarians aim to create the kind of simple and straightforward (not to say 'dumbed down') search environment that today's information seekers want, but with a difference: their search tools feature the components missing from the big search engines and Google Scholar: authority and quality. The early signs are quite promising.

Nature

This information need characteristic seems to defy any attempt at straightforward definition, although a clear enough idea of what is involved emerges from a look at the various types of information in existence. Thus, on one level it is possible to differentiate between conceptual or theoretical, historical, descriptive, statistical or methodological information. On another level, there is the distinction between primary information (that is, uninterpreted evidence) and secondary information (books and articles reporting the results of research). Plainly, information of different types is available on any subject, in consequence of which, the nature of the information being sought is a crucial factor in ensuring that the answer found is truly relevant to the question asked. After all, when people specifically need one type of information, say, some statistics to prove a theoretical or conceptual point, their finding any other type of information on the subject most emphatically will not do. Yet, the type of information does not play the pivotal role in information seeking that it might conceivably do.

In fact, it seems that information seekers do not pay much attention to the nature aspect of the material they need; not even academic researchers, for whom the judicious use of information is truly vital. Thus, for example, in a questionnaire survey, canvassing the entire population of researchers at an Israeli university, more than a third of the participants rated the notion of setting out to look for a specific type of information as altogether irrelevant to them (Herman, 2005). Apparently, the kind of information required only comes to the fore when it is a truly prominent feature of the information need. Thus, for example, an academic researcher, musing aloud on the subject, remarks that she really should look specifically for methodological information, as she does not always know the new statistical techniques in vogue. Her colleague goes even farther than that, remarking that his ways are quite typical of biomedical and physics researchers: he actually sets out to look for particular types of information, for example, if he wants to learn how a certain procedure is done or what the best equipment is to be used for a specific purpose. When he does need such information, searching the internet is his favourite solution to the problem, 'as it has sources, which are practically handbooks, telling you do this and this, the procedure goes like this'. However, on the whole the nature of information sought is not a matter with which people customarily concern themselves; indeed, it is not very often consciously formulated in the context of information seeking.

In view of the scant attention given to this aspect of information need, it is not very surprising to find that designers of information systems have traditionally neglected it. Systems seldom provide for retrieval along these lines, although most, offering a wide degree of access through their subject or word indices, do enable access to specific types of items, such as statistical accounts.

In any case, whilst there are different kinds of information on any subject, not all of them are equally suited to the needs of different individuals. As a matter of fact, some of these information types will prove to be highly unpalatable to some people, as a computer scientist interviewed by one of the authors insists: 'Stories are not for me; I leave them to my humanities colleagues'. Arguing much along the same lines, another interviewee, an economist, draws a clear demarcating line between information that he terms 'opinions', as opposed to what he calls 'facts':

> In the areas we deal with in economics we don't concern ourselves with opinions … When it's a matter of opinion, you collect the facts … , and then form an opinion … , [but then] somebody else can have a different opinion about the same facts … However, in economics it's not what I think, it's what's actually going on … if I say that you need 5 percent to reach your goal, and somebody else says 4 percent will suffice, reality will soon put one of us right … so it's not a matter of opinion; the concept that will be accepted is the one which reflects reality more accurately, and reality is usually not disputable, it's easily verified … if I tell you the price of tomatoes, and you don't agree, the two of us go to the market … and see what tomatoes cost.

Plainly, the nature aspect of some information depends in no small measure on its intended readership/audience. Thus, for instance, social science practitioners will hardly ever require their information produced in a theoretical manner or from a historical point of view – though their academic counterparts most probably would.

Indeed, it is scholarly research, which, intent as it is on querying every aspect of life on earth, perforce involves the use of the whole range of information types in existence: theoretical, conceptual, empirical, historical, descriptive, factual, statistical and methodological. Needless to say, though, that the nature of the information required in scholarly work is subject-contingent, that is, it varies first and foremost with the subject matter of the research underway, inclusive of the disciplinary-conventions dictated approach taken to it. Hence, if theoretical, conceptual, factual or methodological information is probably indispensable in any academic endeavour, regardless of its topic and the knowledge domain in which it is embedded, other types of information are, on the whole, reserved for specific disciplines. Thus, for example, historical information, so central for much of humanities research, will rarely, if ever be needed in science enquiries. By the same token, statistical information, the bread and butter of most social sciences research, is more often than not quite uncalled for in humanities scholarship.

Actually, even when researchers of different subject areas need the same kind of information, the extent of its utilisation may vary from one discipline to another, as Garvey et al. (1974) found in their extensive studies of over 2,000 sciences and social sciences researchers. Apparently, when physical scientists set out to look for methodological information, it is more frequently needed to formulate technical solutions and to design equipment or apparatus, whereas for social scientists the greater need is for information to select a design or strategy for data collection and to choose a data-analysis technique.

Further to that, since the activities associated with the successive steps of the 'typical' research process involve different kinds of mental processing, different kinds of information are required at the various stages. Thus, in the initial stages, the perception of the research problem involves heavy use of theoretical and conceptual information; at the stage of reviewing the existing knowledge on the subject being queried historical and/or descriptive information is needed; the formulation of procedures appropriate to the inquiry necessitate methodological information; in the intermediate stages that follow, when information is required to solve problems as they come up, specific information is usually the answer, along the lines of statistics or details of techniques and methods; and in the final stages, when researchers seek to fully interpret their data and integrate their findings into the existing body of knowledge, the need for information is focused yet again on theoretical and conceptual, as well as descriptive and/or historical information (Egan and Henkle, 1956; Garvey et al., 1974; Menzel, 1964). Very much in agreement with the conclusions arrived at in these studies into the matter, an academic researcher, describing the way he usually works, links the different stages of the research process with different types of information:

> At the initial stage of a research project, when I'm at the conceptualisation stage, I go to the literature in order to see what questions to ask, so I read theoretical works. [Then], when I'm at the stage of research where I'm looking for tools, I read works on applications of concepts and tools, that is to say works which investigate empirical parameters I am not familiar with or works which investigate issues in a new way, one I haven't tried yet. And when I analyse [my] data, I need theoretical information, which helps me to interpret my data, because in my field theory and data interact. I try to understand the data in light of what I've read, and I try to answer the question, what do I know now that I didn't know before, that hasn't been already covered in the literature. That is to say, at this stage I turn to the literature in an attempt to find out how to interpret the answers.

It seems, then, that even academics, who undoubtedly have very wide-ranging information needs indeed, present a rather nuanced picture of the requirements for the different types of information in existence. For other populations it must be 10 times so, with many people needing mostly descriptive and

historical material. Theoretical information, for instance, seems to have a very limited circulation. It is descriptive, methodological and statistical data that are much more the province of the practitioners in all fields. The need for methodological information, however, cuts across the academic/practitioner divide: it is required by practitioners such as teachers, engineers and social workers, as well as by research scientists of all kinds.

Finally, if we interpret the 'nature' aspect of an information need widely to include 'how to do' information, then hobbyists of all kinds are certainly among the most ardent seekers of specific kinds of information. In fact, as a Pew Internet & American Life Project survey proves, looking for information about hobbies is among the most popular online activities in America, on par with shopping, surfing the web for fun and getting news. Fully 83% of online Americans say they have used the internet to seek information about their hobbies and 29% do so on a typical day (Griffith and Fox, 2007). Incidentally, the need for such 'how to do' information is not reserved for hobbyists alone, as a neuro-biologist, sounding very appreciative indeed of the capabilities of today's information services to steer her to this particular type of information, reports: 'I wanted to investigate some animal behaviour using a labyrinth, but I had no idea what the labyrinth should look like or anything [else about it]. So, I searched in a database, found an article which described the kind of labyrinth I wanted to use … and asked our technician to build for me a replica of the labyrinth described'.

Having explored the nature aspect of an information need from one angle, we now come to another way of grouping the different types of material to be found: primary information versus secondary information. Stoan (1984) makes the following useful distinction between the two types of information: the term 'primary information' refers to essentially uninterpreted data, which may be gathered in the field, in a laboratory, in a library or in an archive, whereas the term 'secondary information' is used to denote the books, articles and papers in which the results of research are reported. Obviously, each discipline has its own kind of primary data and its own techniques for gathering and testing that data, as Wiberley and Jones (1989; 1994) point out: humanists use as their primary evidence existing sources created by the subjects of their research; in comparison, scientists and social scientists initiate and participate in the creation of their sources, the former in their laboratories and the latter in the field. Therefore, while in the sciences and the social sciences no primary evidence exists until the researchers begin to work, as the primary evidence is the product of the scientific quest, in the humanities the primary evidence is there first, for the researchers to reconstruct, describe and interpret. Moreover, as the subjects of humanistic research create the primary evidence of the humanities, these sources are the products of a specific place and time and shaped by the distinctive personalities of their creators, with primarily qualitative and aesthetic dimensions. No wonder, then, that humanists, unlike their scientist or social scientist counterparts, are wary of accepting even true-to-life replicas as substitutes for the original source they are about to analyse.

The fundamental nature of primary information may thus vary from field to field, but its vital importance for expanding human knowledge is indubitable; indeed, it is the genuine core of any research undertaken. Happily, in the course of the past few years, access to primary sources of information, especially current information, has been greatly enhanced through the ubiquitous availability of internet-based resources, as an academic explains:

> We download the latest 'Human Development Report' in a jiffy, whereas in the past we had to wait two months for it to arrive … today you can download everything so easily, reports of the World Bank, of the Palestinian Central Bureau of Statistics, the Israeli Central Bureau of Statistics … you can check the price of oil every few days, the exchange rates, the reports of the banks in the Middle East, everything, all the time.

The vast improvement in the ability to obtain primary material is nowhere more conspicuous than in the case of information deemed unmarketable by commercial presses. Indeed, the elusive character of the primary sources of information dubbed 'grey' or 'fugitive' (non-commercial literature, which is produced by government, academies, business and industries) has been radically changing. As Cronin and McKim (1996, 165) sum it up: ' … grey literature is no longer the step-child of primary publishing. The web invites and envelops semi-published, unpublished and vanity items, blissfully unmindful of provenance or pedigree – grey becomes black and white on the World Wide Web'.

As we have just noted, somewhat in passing, much of the primary evidence so readily accessible on the internet is of the kind that is particularly useful for seekers of current or recent information on human behaviour (both on the group and the individual level), and the social and environmental factors affecting it, such as data sets, original archival material, and legislative, governmental and demographic information. However, there is also a wide availability on the internet of science primary information, such as raw data and technical reports. Perhaps somewhat more surprisingly, there is also an abundance of humanities primary information to be found on the internet, from the facsimiles of manuscripts and artefacts, through films and music recordings, to archival material. True, as Stone (1982) clarifies, it is really vital that people, especially scholars, who criticise or study the history of paintings, sculptures or buildings, see the genuine article, for information absorbed from an original artefact cannot be gained in any other way; however, even in the humanities the need for access to original research reports is more questionable. Apparently, then, the oft-heard argument that humanists would use electronic alternatives to the primary material they need, perhaps not invariably, but to a greater extent, if only their need for old, if not antique material could be met, that is, if more primary sources of theirs were digitised (Brockman et al., 2001), seems to present only part of the picture.

All the same, where access to primary material is concerned, there are very good reasons for singing the praises of the internet. Of course, it is the

realm of 'current history', truly the forte of the web, which is probably the best proof of the vastly improved access to primary material in today's world. Still, the by now routine practice of allowing public access to research data 'milked dry' in a given project is also quite some improvement over past practices. The resulting databases of primary information surely would have been impossible to come by in the days before the electronic era! There can be little doubt then that the likelihood of locating primary information has risen quite dramatically thanks to the internet, to the great benefit first and foremost of academics, but not only them. People pursuing topics of interest or study can be quite pleased, too, at the ease with which they can now come by original letters, photographs, audio recordings, moving pictures or video recordings.

Intellectual level

This characteristic refers to the minimum extent of knowledge and sometimes the level of intelligence an individual might need in order to understand the information available. However, it is not simply a question of matching the intellectual powers of individuals with documents of appropriate academic level, because intellectually advanced individuals might require elementary knowledge in a related or marginal field. Thus, a research scientist coming to grips with spreadsheets for the first time could be on the same information footing as a school secretary. In fact, matching people's information needs with the suitable level of information dictated by their requirements and abilities has more to do with the intelligibility of information (or, all too often, its dearth …). Thus, information is made complex not just by how much knowledge and education it assumes, but also by how abstract or compressed it is. Writing styles and skills do count here too; after all, some broadsheet newspapers deal with some very complex ideas, but their articles are made comprehensible by excellent presentation of the data. Indeed, journalists play an important role in making the contents of academic reports and research accessible to a much wider public. Thus, for example, the medical correspondents of broadsheet newspapers, like *The Guardian* and *The Independent*, regularly repackage articles from *The Lancet* or *British Journal of Medicine* for an essentially lay audience.

Unfortunately and, perhaps, somewhat paradoxically in our age of omnipresent information, locating a piece of information fitting to the knowledge and intelligence level of the person requiring it seems to have become more problematic than ever. Obviously, with so much information everywhere, rounding up something on any given topic rarely, if ever, poses a problem; however, finding an answer matched to an individual's needs, with all the intricate combination of subjective and objective factors that enter into it, is something else again. This is nowhere more apparent than in the case of meeting an information need on an appropriate level: clearly, for a child who wants information on the amount of food his newly acquired dog will consume daily, to be presented with a

scholarly article in a learned journal on the topic is hardly the right solution for his need, and a zoologist would likewise find little joy in a children's book on pets. Each to his own, of course, which until recently was a widely accepted and easily adhered to information behaviour norm. The academic went to his research library and the child to the neighbourhood public library or the school library, content in the knowledge that their information needs were sure to be met on an appropriate level. No longer, though. With the easy accessibility of myriad items on the web, the effortlessly discernible dividing lines of yore between popular/scientific, elementary/advanced, detailed/superficial information are apparently not as visible as before. More amazingly, distinguishing among them is not deemed very necessary at all, regrettable as it may be.

In point of fact, regrettable it is: getting the pitch right can be of utmost importance, as there is a huge spread of values between the advanced user and the elementary user, and between advanced information and elementary information. This seems to be true even where relatively homogenous populations are concerned. Thus, for example, academic researchers may all want scholarly level information, but, as Menzel (1964) notes, scholarly communication channels are overwhelmingly tailored for the specialist in a given field. Opining that seekers of information from a disciplinary area that is not their main concern could benefit greatly from a different sort of selection, editing, grouping and packaging of information than that which is most suitable for insiders, he suggests the establishment of special channels of communication that would be dedicated to 'information from field A for researchers in field B'. Thinking much along the same lines, Kircz (1998) proposes the breaking up of the classical scientific or scholarly article into modules. This, in order to cater to the different levels of information needed during the various stages of a research project, in the course of which an academic alternates roles between an uninformed reader, who is out to learn something entirely new, in a field that is either unknown to him or her, or of which he/she has only a rudimentary knowledge; a partially informed reader, who is not conversant with the specific research as such, but is interested in the general aspects that might be of use for his or her own investigations; or an informed reader, who is well-versed in the literature of the field.

Despite their obvious merits, neither Menzel's nor Kircz's proposed solutions to the problem of different levels of information required have ever been realised, although probably not for a lack of a very real need for them. After all, information systems are cognisant enough of the requirement for suitably written material, which is why they index documents according to their intellectual level. Thus, for instance, ERIC, an educational database, allocates academic/research, practitioner and consumer codes to documents to assist with their digestion. However, the construction of a truly finely-tuned modularised system was hardly feasible before the widespread utilisation of information technology, whereas today, when the technical know-how could easily allow for such an undertaking, nobody seems to be clamouring for it. Given the scant attention paid these days to information-seeking strategies – as it has already been

noted, searching for information is no longer considered a task requiring serious investment in terms of thought, preparation and execution – this comes as no surprise.

Furthermore, nowadays the sharply delineated demarcation between the need for popular as opposed to scientific, elementary as opposed to advanced, detailed as opposed to superficial information is truly crumbling down. In direct consequence of the vital role played by knowledge and information in contemporary life, which has turned us all into life-long learners, practitioners and even amateurs will read academic/research literature. Indeed, findings of the Virtual Scholar research programme (2001–08) clearly indicate that users from outside academe are heavy consumers of the Open Access scholarly material on the internet (Nicholas et al., 2008b). By the same token, as it has already been noted, lay persons use extensively the health sites on the internet primarily intended for healthcare professionals and researchers. No wonder, then, that many academic health websites, mindful of the obvious demand for their services, now often provide pages targeted at the general public as well. Developers of the academic cancer website, Kimmel Cancer Centre (at www.kcc.tju.edu), for example, having found that their database of currently open clinical trials, designed for physicians, was frequently accessed by non-professionals, began to include lay descriptions in their trial listings (London, 1999).

Interestingly, there is a move in the opposite direction as well: perhaps not very often, but at least every now and then internet-based popular databases are utilised for research purposes, too. Apparently, when academics need information in areas outside their chosen spheres of expertise for gaining some essential understanding of an issue at hand, they can and do compromise on the level of the material required, making do with consulting an information resource of a more elementary level (Herman, 2005). Thus, for example, a psycho-oncologist, who looks on occasion only for the basic-level information needed to aid her in understanding the wider picture within which the specific point she is investigating is embedded, says: 'for the fundamental medical information necessary for my investigations at times I turn to those popular health sites on the web'. However, her psycho-pharmacologist colleague, who joins her in testifying to sporadic use of non-scientific material or material not published via scientific channels for research purposes, clarifies: ' … of course I know enough to sift out the wheat from the chaff'.

Having said all that, much as today's information seekers tend to ignore any considerations pertaining to the level aspect of their information requirements, their competencies to understand and handle information, be it in electronic or hardcopy form, do vary quite considerably. This, obviously, can be quite problematic for information providers: their potential customers, coming as they do from different educational backgrounds and possessing diverse levels of computer literacy, reading skills and reading comprehension capabilities, may find the information presented to them unacceptable level-wise. Either too complicated, deep and detailed, and as such uninteresting and unappealing, or else too simplistic and superficial, and therefore not only just as uninteresting and

unappealing, but also patronising and lacking in authority. Seeking to address this problem of people's varied abilities to comprehend information, Williams et al. (2003) came up with the idea of organising information into 'vertical' layers of pages. These would offer information on each topic at different depths or levels of detail, in addition to the 'lateral' arrangement of material organised by topic. Clearly, the designers of any such system would have to be cognisant of digital visibility issues and consider how the information could be displayed in the fewest hierarchical levels, perhaps with the informational levels 'side-by-side' on the menu option, e.g. 'treatment of kidney disease: basic information; more detailed information; advanced information' – where each option was an active link.

One further point of considerable relevance to the intellectual aspect of information needs: the much debated possibility that in the new digital environment information-seeking behaviour is increasingly being 'dumbed down'. Indeed, the CIBER studies exploring the information work of various communities, most notably those associated with news (Nicholas et al., 2000), health (Nicholas et al., 2007a), voluntary and charitable work (Nicholas et al., 2004a) and scholarly publishing (Nicholas et al., 2008b) point to characteristic patterns of information behaviour, which, taken together, could possibly amount to such 'dumbing down'. Indeed, there is some disturbing evidence that much of today's information use, shaped by massive digital choice, unbelievable (24/7) access to, if not bombardment from a vast array of channels, disintermediation, and hugely powerful and influential search engines, constitutes passing and/or ineffectual activity. Everyone exhibits bouncing/flicking behaviour, which sees them searching horizontally, rather than vertically. Power browsing and viewing are the norm for all; reading appears to be undertaken only occasionally online, probably undertaken offline and possibly not done at all. Promiscuity is endemic as a direct result of a combination of massive choice and the constant refreshing of choice by search engines. Much digital activity involves navigating and not using. The quality and reliability of some information is judged by its popularity; the wisdom of the crowd is the key measure of worth. All this arouses a real fear of a world full of 'information malnutrition', where people consume vast amounts of the information equivalent of fast food: easily obtainable, flattering to the undiscerning palate, but of little actual value at its best, harmful at its worst.

Viewpoint

The particular viewpoint, approach or angle from which (overtly or covertly) some information is written up, is obviously very important for its potential consumers. Indeed, this is one needs characteristic that really gets information seekers worked up. More often than not people require information sympathetic to the views that they hold, to the point that information on the same topic, but tackled from a different point of view or slant would be unpalatable to them. This is probably most commonly seen in newspaper readership, where people like to subscribe to those newspapers that present news and views from the

same political and social standpoint as their own. Still, as we will see further on, information perceived as representing a different viewpoint from one's own, or even as biased and/or one-sided can nevertheless have its welcome uses.

In any case, this aspect of an information need is very much subject field-contingent: the nearer the subject of a piece of information is to the hard end on Storer's (1967) 'hard' to 'soft' continuum of knowledge domains, the less applicable becomes the notion of its being reported from a specific viewpoint. After all, Newton's laws are Newton's laws, regardless of how they are put forward; they are hardly open to different understandings or interpretations, are they? However, the 'softer' disciplinary fields, the social sciences and the humanities plainly allow, if not outright call for different approaches to writing up information, so that topics are frequently treated from a certain angle or with a particular perspective. So much so, that social sciences information, for one, is almost by definition 'information presented from a specific perspective', as an academic, an expert on social welfare, explains: 'I proceed from the more or less always given notion that information is subjective; it is based on the subjective theoretical approach and presented through the subjective perspective of its originator. I never even expect anything else'. Indeed, this disparity among the different disciplinary areas is so widely accepted that it is considered among the most fundamental axioms of information work. No wonder the aforementioned psycho-oncologist has no doubts whatsoever when she ponders the topic:

> I always want facts and data, and although it's not mathematics, it's not open to different interpretations, either. It's in history that they mess about with interpreting the data this way or that way; in my field, if the statistics say that 80% of the patients prefer a certain surgical procedure, neither you nor I can say that it's not the majority of women who feel this way … it's not debatable, is it?

Indeed so, which is why in those areas where the subject matter is open to different interpretations, any piece of information encountered may vary by one or more of several determinants of viewpoint: school of thought, political orientation, positive or negative lines of attack and, in interdisciplinary fields, discipline orientation.

School of thought

Schools of thought, those informal bodies of people united by a general similarity of principles, opinions, points of behaviour and practices, are most evident in the social sciences, although clearly the humanities have their own schools, too. These schools are large, widely known and have handy labels to describe them, such as feminism, Marxism, monetarism, neo-classicism/modernism and structuralism, to mention just the first ones coming to mind. In addition, there are countless mini-schools or 'departments', which inhabit nearly all

disciplines. In the field of Information Science, for instance, we have the systems-driven and user-driven schools of thought. It is at this rather more specific level that it is possible to discern schools of thought even in science. In fact, Thomas Kuhn (1963), the renowned American philosopher and historian of science suggests that progress in science is based on the existence of precisely such schools of thought, for it is through the breakdown of old paradigms (i.e. theories or ways of looking at the world) and the emergence of new ones, which 'attract an enduring group of adherents away from competing modes of scientific activity', that world views change. Thinking much along the same lines, Max Planck (1968) even contends that ' … a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it'. In addition, in scientific fields, too, ethical considerations bring about the fragmentation of information so characteristic of social science disciplines. Thus, for example, the ongoing intellectual battles over abortion, animal-testing, genetically modified food and organic farming techniques are every bit as heated and polarised as they are in the social sciences.

It is hardly surprising, then, to find that in many academic disciplines values may be highly charged and the incidence of major ideological disagreements quite prevalent. However, since the scientific ethos stipulates for scholars to be impartial and unbiased in their communications even as the existence of different schools of thought entails the possibility that a topic would be treated from a particular approach or with a particular perspective, they must routinely evaluate the material they handle for any tell-tale signs of a specific angle of reporting. This, as Ellis (1989) and Ellis et al. (1993) point out, can greatly aid researchers in establishing the worth of some material for their investigations, for information appearing in sources which take a similar perspective to their own is more likely to be pertinent to their needs. However, as Meho and Tibbo (2003) add, in this process, which Ellis (1989) calls differentiating, researchers also aim to detect biased information, the product of scholarship which has a hidden (or not so hidden) agenda. Obviously, identifying such slanted information for what it is forms an important part of research work, striving as it does to present a disinterested and open-minded picture.

Still, if the lack of objectivity is openly acknowledged, or at least once it is detected, there seem to be important benefits to be derived from information that is not presented objectively, or, in point of fact, because it is not presented objectively, as a historian suggests:

> I may definitely set out to look for an article, which has been written from a specific point of view, because in many instances you can deal with a topic only if you are well aware of the points of contention involved. If you're not familiar with the controversy concerning the subject, if you don't know who's against whom, and what each contender has to say for himself, you can make mistakes, you can misunderstand the situation.

His philosopher colleague apparently is of much the same opinion:

> As long as you realise that some information has been written from a particular point of view, there's no problem whatsoever with using it. Quite the opposite, actually: this way you learn that there is such a viewpoint, that it's possible to think … and to answer the question … differently; it can be of tremendous help even, because otherwise such an approach may not have occurred to you.

Indeed, the one-sided or non-objective nature of some information can render it all the more useful, but, probably just as often, also altogether unusable. A telling example of how this might happen came up in the aforementioned survey of the use and impact of key digital health platforms and services in the UK (Nicholas et al., 2007a). One health information seeker, a staunch supporter of homeopathic medicine, professed to a general mistrust of the information to be found on the NHS website because its philosophy of health and treatment was at odds with hers:

> I guess I never just trust what the NHS says. I always research their diagnosis as well as their proposed remedy before using it because Western medicine in general is myopic in its approach to the matter of health. Basically I feel there is too much of the notion that current medical practices can 'heal' you and too little on the concept of responsibility for one's own state of health; mentally, physically and spiritually. I guess you could say that I fall into the category of people who really do subscribe to the notion of a holistic approach to life. What the NHS offers is just one part of the whole.

The problem is that recognising the school of thought a piece of information represents can be quite tricky, seeing that authors – founding fathers and key disciples aside – do not normally identify which, if any, school they are writing from. True, for academics, familiar as they are with a field, discerning the school of thought guiding the writing up of some information is a manifestly undemanding task (Herman, 2005), and no wonder. After all, they develop such an intimate knowledge of the literature of their areas of interest, that they are easily able to match the features of the source they are examining with those long known to represent various schools of thought, viewpoints, approaches and perspectives. Thus, they can effortlessly recognise the point of departure of the information under consideration by the sources cited and the terminology used. Also, apparently they look for certain 'codes', which, for them, clearly signify the author's ideological affiliation. As a senior historian, well-versed in the ways of the research enterprise, elucidates:

> There are these little symbols … for example, in an article [written by somebody] from the post-modernist school of thought, the opening lines

send you to a footnote citing Foucault or Derrida; now, this is a code, a sign, which says 'I'm a post-modernist, I accept the main tenets of post-modernism', even though he is actually writing on, say, the relations between Britain and Kuwait in the 17th century. There's no connection whatsoever? True, but this is how he pays allegiance, how he tells you 'classify me correctly'.

Obviously, not everybody is familiar enough with a given field to be able to identify in this manner the tacit evidence pointing to the school of thought; thus, expert help can come in very handy here. However, given the afore-mentioned trend towards a widespread marginalisation of intermediaries in information seeking, it is not too often that the viewpoint aspect of an infor-mation need is deemed important enough to dispatch people in search of a librarian. True, some information systems do help by providing viewpoint indexing – the British National Bibliography did so for a time with its PRECIS system – but, as we have already noted, the use of search engines, rather than information systems, is the first-line option for unearthing pertinent informa-tion. Still, today's practised information users seem to have their ways of dealing with the problem, if and when they are aware of it at all. Thus, they may simply search for an author representing the 'right' school of thought, possibly supplementing their findings with the names of the authors who subsequently cited the original publication (on the assumption that there is a good chance that they all follow the same intellectual traditions). After all, tracking down a list of subsequent citations of some information no longer poses too much of a problem: no need to resort to complicated bibliographic tools, or, for that matter, to the help of information specialists; suffice that you are aware of the existence of Google Scholar, which routinely offers citation data. Also, many people are experienced enough information users to realise that the organisation the author works for is the best guide. Thus, when it comes to organic farming, we know clearly where the Soil Association and the National Union of Farmers stand.

Political orientation

If there is one area more susceptible to subjective reporting of information than others it must be that of politics, as a university lecturer, an expert in mass communications, notes: 'I always teach my students to try and find out who the author of some information is; if he is, say, a politician, think of his party; if he is a journalist, you have to be cognisant of the political points of view represented in his newspaper'. Indeed, information can be, and often is written up from a political point of view: there are right-wing, left-wing, conservative, socialist stances – to name the most obvious. Of course, school of thought and political orientation may coincide. It would be difficult, for instance, to characterise the pro- and anti-European standpoint that drives much of today's politics in the UK. While people feel more comfortable with,

and more disposed towards reading information sympathetic to their own political allegiances, it would be wrong to assume that they would not be interested in information emanating from a different political persuasion. Thus, politicians will often scan the opposition parties' newspapers, hoping that they can use something there as ammunition to throw at the opposition in debate or interview. A quote from *The Guardian* (a left-of-centre newspaper) used by a Conservative politician in the debating chamber can prove most effective.

As with school of thought, the political perspective of information is not always immediately recognisable, except when it comes from a mainstream political party, of course. Nevertheless, here again, if you are privy to the terminology typically used by holders of a specific political stance, discerning the point of view, approach, or angle from which some information is presented becomes much easier. Thus, for example, where the heavily politics-laced issues of the contemporary Middle East are concerned, if the term used is 'the Zionist state' instead of 'Israel', it is clearly written from a particular point of view, effortlessly recognised by all those in the know. However, what if you are not among the chosen few who are wise to such subtle representations of political orientation? How is your need to be met then?

Unfortunately, little help can be expected from the available information systems, but there are shortcuts, because so many organisations are associated with a political point of view. We are not just talking about political parties, either; think tanks, research centres/organisations, associations, unions, voluntary/pressure groups, governments, newspapers and even university departments can be associated with a political point of view. A document on the environment by the Automobile Association is likely to differ fundamentally in tone from one issued by Friends of the Earth, for instance. Therefore, compiling lists of suitably-minded organisations can help a lot in the meeting the need for information shaped by political considerations.

Somewhat surprisingly, information professionals, especially those in the public service, sometimes fight shy of providing information presented from a political point of view, believing that this might compromise them in some way. However, there can be nothing wrong with providing information corresponding to the political stand being avowed. Indeed, at the House of Commons, where the Library has to walk a political tightrope, librarians frequently prompt the inquirer as to for what political end the data is required. That way the information can be used to its best advantage. If they, of all people, have no qualms about this, why should anyone else?

Positive/negative approaches

Sometimes there is a need for information to be presented in a positive or negative form – as the existence of 'spin doctors', those political press agents or publicists employed to promote a favourable interpretation of events to journalists, confirms. The demand for information presenting the facts in a favourable/non-favourable light is greatest amongst those in politics, the media

and business. Thus, for example, both the Labour and Conservative parties in the UK have what they call 'dirt' databases on individual MPs, in which they keep unsavoury stories, injudicious quotes, incidences of poor behaviour, voting inconsistencies – all data to be leaked or exposed at a suitably telling time, such as a media interview with an opposing politician. The need for this type of data is best shown in an enquiry one of the authors conducted for a national newspaper. A prominent politician was suing the newspaper over being quoted as having said something, which made him look bad, although, according to him, he had not. For the newspaper the best defence was to get the 'dirt' on this politician in order to show that in the past he had said things, only to deny them later, proving that his memory/word could not be relied upon.

Apparently, then, it is not all that rare to find that there is a definite need for information that presents the issues being considered in a positive/negative light. No doubt, if more information systems catered for this approach, the demand would be even greater. Yet, there are not many people who would or could easily field this important and valid information need, and even information professionals consider such requests just as controversial and challenging as the above-noted need for information put forward from a specific political point of view. Information science students are particularly aghast when told that this is a legitimate characteristic of information need. Certainly 'dirt' in this context is unlikely to feature as a database keyword. Indeed, needs of this kind are best met through oral and informal channels of communication, but to have its greatest impact, the information obtained has to be recorded. Still, the problem is not unsolvable. Thus, for instance, there are publications that are only ever going to show some politician or personality in a bad light: *Private Eye* and some biographies, for instance. Also, a politically motivated organisation that is opposed to the views held by the person under consideration can be a great source of negative information.

Today the web is the best peddler of 'documented dirt' information, as The Drudge Report political site, for one, amply proves. In fact the web, especially the blogs, seems to meet much of the strong demand for the controversial, gossip and plain dirt. For many people, but especially journalists, the fact that the internet contains information of uncertain, but interesting quality is a plus. In particular, features journalists and those charged with producing articles of unusual 'human interest' or of a generally lighter nature tend, as might be expected, to be most interested. So much so, that they may even use information of dubious authenticity.

Subject orientation

In interdisciplinary fields, where authors and information providers possess a number of different subject backgrounds, there is the need to consider the subject orientation of an information requirement. Criminology provides a good illustration of this. University criminology departments are staffed by academics who approach the subject from a wide range of disciplines. Thus, for instance,

the criminology department of Middlesex University studies criminology from the perspective of sociology, whereas the academics at the University of Manchester study criminology from a psychological perspective and those of the University of Cambridge from the standpoint of law. Any assessment of the information needs of criminologists must take this into account.

It seems, then, that diverse interpretations of facts and data are possible, if not expected in certain knowledge areas, although even there users may at times actually want information that is wholly objective. Given this state of affairs, detecting whether a particular piece of information on hand is written from a certain point of view, approach, or angle becomes an important component of information work. Indeed, as the above quoted philosopher suggests, every piece of information should be treated as potentially representing a specific point of view:

> You have to take into account that people often write with hidden layers of meaning … it's important to be aware of the possibility … [and] to recognise the concealed elements in the text. In fact, part of the understanding of a text is identifying whatever the author assumes to be self-evident and therefore leaves unsaid, or whatever he chooses to conceal.

No wonder the vast literature on information literacy unanimously acknowledges the need to keep a wary eye on authors' possibly slanted presentation. Concurrently, as we have seen, people may at times look for and happily use information representing specific viewpoints. However, perhaps somewhat surprisingly, nothing much seems to be happening in result of these parallel trends in information work; information seekers certainly do not clamour for a solution, either because the problem is not all that pressing, or because they do not even realise that they have a problem (of course, it is hardly a rare occurrence that information needs go unrecognised, or at least unvoiced).

Quantity

Life in our present-day postmodernist society, at least in theory, requires us to be better informed than ever before, emphasising as it does plurality of values and diversity, tolerance of ambiguity, acceptance (indeed celebration) of innovation and change, on the one hand, and challenging of convention and authority on the other (Buschman and Brosio, 2006). Now that we no longer rely on tradition or past experience (for the circumstances of yesterday have surely changed by today), or on the decrees of some authority figure (for who knows what is suitable or good for us as well as we ourselves do), how else can we weigh up the possibilities, the pros and cons of proceeding one way or the other? Plainly, nowadays we are only able to cope in all walks of life by getting hold of information to serve as the basis for competent decision making.

Not that attaining the necessary information poses any problem in today's information-saturated world; rather to the contrary. We may no longer live in

fear of drowning in a vast sea of information, a point we will examine in more detail shortly, but we do have (literally) at our fingertips a dynamic, constantly growing and changing knowledge base, which is so wide-ranging, that it is truly all-encompassing. However, while all people require information to do a job or solve a problem, the size of their information needs (though, as we will see presently, not necessarily that of their information appetites) varies greatly, not only between individuals and groups, but also according to the nature of the need. Motivation, diligence and the amount of time available to take in information are all influential factors in determining the amount of information actually consumed, although the very presence of huge quantities of data in every form encourages excessive demand. Indeed, by now we have all moved from a situation where the main information problem was getting hold of information, to a situation where the chief difficulty is digesting (or avoiding) the information that all too easily flows our way. So the quantity aspect of an information need is unusual in that it can also be an information constraint.

It is not that information seekers never, or even rarely, want all the information on which they can lay their hands. Someone starting a new job, for example, might feel that the more information to be had on the organisation he/she is about to join, the better. An investigative journalist embarking on a new story might think similarly. Thus, the request to 'give me everything you've got on … ' is far from uncommon in journalist–information worker exchanges. Indeed, quite a few of the journalists interviewed by one of the authors actually said 'you can't have too much information', although admittedly this was a few years ago, and the web might have changed that. The prime example, though, of people with truly voracious information appetites are academic researchers, and no wonder. As it has already been pointed out, they have to make sure that they see almost every new publication in their respective fields in order to keep up with the developments at the research front of their specialisations. Moreover, for each new research endeavour they undertake, they need to amass, as comprehensively and exhaustively as possible, the knowledge accumulated on their subject, first to identify the gaps necessitating further investigation and then to anchor the topic to be investigated in its information context. However, given today's unprecedented availability of vast quantities of information, 'everything on the topic' may be too tall an order even for academics. Have not we all heard ad nauseam about the problem of information overload, at least as a threat, if not a reality, to quote Wilson (1996)? However, is it a threat, or a reality, or perhaps neither?

The past few decades saw an ever-increasing concern with the possibility that the growth of knowledge had surpassed the growth of the knowledge of how to manage it (Gaines, 1995), to the point of its having assumed the stature of a widely accepted truism. Any mention of the subject seems to have served to unleash a flood of woeful prophecies concerning the difficulties to be encountered in a world plagued with excess quantities of information. Not any more, though. Plainly, the gloomy forecasts as to the huge availability of data resulting in 'information overload', or its more dramatically put counterpart

'information explosion' simply never came true (in fact, both terms may very well be on their way to becoming misnomers). Today's information consumers, far from being beleaguered by problems of information overload, consider the current state of almost unlimited access to information unproblematic, if not eminently satisfactory. Indeed, it looks as if somewhere along the way many of them have fallen in love with the information affluence characterising our times, actually revelling in the abundance of information unabatedly accumulating all around them: 'There is a fantastic amount of information available … you can access wires, you access libraries, you access information all around the world … in an office … or at home … it is quite fantastic'.

Even academic researchers, with all their patent awareness that an attempt to read everything that may be relevant to their scholarly interests is perforce doomed to failure, seem to be at least resigned to this state of affairs, regarding it as one element of the academic assault course, and a small price to pay for the unbelievable level of access obtained (Nicholas et al., 2008b). However, many go even further than that, considering the large quantities of information flooding them not merely an acceptable trade-off for the ease with which they obtain it, but a veritable blessing (Herman, 2005). Thus, for example, an academic communications expert explains: ' … there truly is an inordinate amount of information all over the place, though I'm glad that it is so … ', and his colleague, a professor of social welfare, apparently feels the same. True, he does point out the evident hopelessness of knowing everything, of reading every piece of possibly relevant information, which, according to him, entails an insatiable thirst for new material, along with some stress and guilt, but still he insists: 'I would never go back to the time there was less information around, I'm overjoyed with this explosion of information'. He is not alone, either, in happily taking avail of the wealth of information at his disposal, as the CIBER findings on the information behaviour of the digital information consumer amply prove. Take, for example, the levels of activity associated with scholarly sites. The volume of their use is very impressive indeed and seems to be rising inexorably. We are witnessing not only constantly escalating use on the part of the core audience of these sites, although that too, thanks to their ability to access the site anytime and anywhere via broadband, wireless, the Blackberry, mobile phone and the like, but also a surge of huge masses of non-subscribers, coming in via search engines and making enthusiastic use of the scholarly net (Nicholas et al., 2008b).

However, it is a mistake to assume that more (and faster) information increases knowledge: rather like food consumption, after a while it does you more harm than good. Similarly, it would be wrong to believe that there is always something essentially important about getting (more) information. Take, for instance, this quote from an academic researcher:

> It's important that I keep current, but … it is my thinking, which is the core. All the rest [i.e. the information generated by others] just testifies to my having made my investigations, that I know what's going on, but it's

> marginal to the heart of the matter, which is your own opinion and your reasons for forming that specific opinion … Additional reading will not matter much for your thesis; in fact … it'll be redundant, it'll obstruct your train of thought, it'll impede your ability to say to yourself: 'O.K., what do you make of this, where are you going from here'.

All this is very much in line with what a respondent in the Information Requirements of Social Scientists (INFROSS) research project (Bath University Library, 1971b) had to say on the subject more than 30 years earlier: 'The importance of information can be overrated. More information does not always result in increased knowledge and probably seldom produces increased wisdom'.

Unfortunately, it is precisely at this juncture that information professionals so often get their customers' requirements wrong. Too many judge their own information prowess by the amount of information that they can provide in response to an enquiry, considering a long list or bibliography the physical proof that they can do the job. The following incident amply illustrates information professionals' mindset where the quantity aspect of information need is concerned: a class of information science students were set an online exercise to discover what the winter of 1989 was like. After about 20 minutes a very unhappy student put up a hand for help. The problem? Well, they had only found one newspaper article on the subject. The fact that the article gave all the details mattered not; something was driving them on to find more of the same!

Actually, people do not always have the time, inclination and, perhaps most importantly, real need to wade through large volumes of information. The dearth of time necessary to make effective use of the information resources available can be especially problematic in this context, as people may find it difficult to come to terms with their inability to deal with the piles of documents accumulating on their desks or the long lists of unread e-mail messages awaiting their attention, to the point where they feel actually overwhelmed by the situation (Bawden et al., 1999). There again, people today are, by necessity, fairly efficient managers of time and they have certainly mastered the art of coping with the profusion of information characterising our world: indeed, they tend to be satisficers (a term, which, as it has already been noted, results from the blend of the two words 'sufficing' and 'satisfying'). That is, they stop information seeking after finding material that is good enough (Savolainen, 2007), so that they can juggle the need for comprehensive information with the constraints placed upon them. Thus, if need be, they are quite content to have sufficient, but small quantities of information, preferring limited information that meets deadlines rather than complete information that does not, which is why senior managers, for example, often insist that all written communication fit on one-side of an A4 sheet. Indeed, people ignore what they perceive as unnecessary or irrelevant; they sample and select, choosing the best/most suitable/most interesting; or they even information gamble, taking a chance on what comes to hand (say, the first item among thousands of search engine-generated items).

These tactics are by now such a customary element of contemporary information behaviour that even academics opt for them (well, one hopes for the first two only) in order to deal with the profusion of scholarly and scientific publications at their disposal. Thus, as Wilson (1993a, 1995, 1996) contends in a series of studies on communication efficiency in research, contrary to conventional understanding, non-use of relevant information in a research enterprise may not happen by accident or by mistake, but rather reflect a routine and normal approach to coping with the huge quantities of information in each and every field. Indeed, he suggests, rather than overdosing on information, researchers consistently and deliberately ignore material of which they are aware, even though it may be pertinent to their enquiries, as part of their individual research strategies. True, they do so only after more or less careful perusal and prioritisation of incoming information, much along the lines of the 'differentiating' component of Ellis's (1989) model of researchers' information-seeking patterns, through utilising the differences between sources as filters on the nature and quality of the material examined. Herman (2005) ties the obvious prevalence of this information management strategy, which, in effect, amounts to no more than the age-old policy of selective reading, to the lowering of academic standards associated with the present-day profusion of scholarly publications (the 'publish or perish syndrome'). This, she maintains, has brought about a change in attitude to information in academe: perceived to be declining in quality, information is no longer treated with deference bordering on reverence; rather, it is customarily appraised for its merits, just like any other commodity, and of the more easily available and plentiful variety too. This is why the key to contemporary researchers' effective information consumption is selective reading. This is why their strategy for coping with the time pressures typical of today's scholarly endeavour, on the one hand, and the vast quantities of information incessantly flowing to them, on the other, is screening, evaluating and filtering, not just to distinguish relevant from irrelevant, but to separate dispensable from indispensable relevant material.

To be sure, by now the real predicament for most people is not that they do not have effective enough ways and means of dealing with the inundation of information; they certainly do. It is rather that they cannot resist the temptations of the information affluence surrounding them. In fact, the behaviour of today's information seekers may very well exhibit the characteristics of 'The Sweet Shop Syndrome' (Ball, 2004): just like children, who, suddenly given the freedom of a sweet shop, will gorge initially far beyond the true limits of their actual needs, they too may get overly excited when they encounter the (as yet still) novelty of easy access to a seemingly unlimited array of information. This, coupled with the behaviour characteristic of the e-shoppers that they are, which is fashioned by the dictates of what could only be described as a sales mentality, results in a tendency to accumulate far more material than they actually read or use.

In the way that shoppers are easily swayed in their choice by 'offers', so too are present day information seekers. This point could not have been made

clearer than by the findings of a CIBER study of Emerald (www.emeraldinsight. com), a scholarly journal database, the policy of which is to offer for free the articles from two journals once a week. It transpired that for these two journals – whatever they were – use jumped immediately by a factor of 10, only to drop down again to pre-offer levels once the promotion was over. Clearly, users hastened to download the articles while they were free, rightly surmising that this state was a temporary one. An analysis of download times before and after the free week suggested that a squirreling (or access expectation) behaviour was being witnessed: download times during the free week were much shorter, an indication that people were simply storing for a later day, rather than 'reading' at the time (Nicholas et al., 2008b). Indeed, the same behavioural trait was displayed by off-campus e-book users, noted in another CIBER project, Superbook (Nicholas et al., 2007b). Much material is just squirreled away for another day and that day never comes because of a shortage of time and the amount of squirreling that has already been undertaken. However, the increase in use for the two journals, on offer free for a week, was so remarkable, that it must have had something to do with their enhanced digital visibility at the time as well, with the very fact that they were part of a promotion.

It seems, then, that in our information-rich and information-driven world the quantity aspect of an information need has certainly come to the fore. However, flying in the face of popularly held notions, it is no longer the threat of information overload which is a cause for worry. Having masterminded the intricacies of coping with the ever-growing abundance of information at their disposal, today's information seekers see no problem whatsoever with the vast quantities of information available on almost any topic imaginable. However, whilst people perceive the situation all in all as a happy state of information affluence, they also find it hard to curb their information appetites, demonstrating a propensity to bite off far more information chunks than necessary, or even consumable.

Quality/authority

Assessments of the value of information are not easily made, especially in the digital environment, with the ever-more conveniently accessible internet bringing to the desktop of the contemporary information consumer a truly alluring array of material on every conceivable subject. Knowing that the information needed may be only a mouse click away seems to pose an almost irresistible enticement to opt for the instant information gratification so easily provided by the web, especially these days, when speed is so often the paramount consideration in whatever we set out to do. Inevitably, then, quality and authority concerns loom particularly large in our era of ever-present information of unprecedented volume: the extremely vocal, ongoing debate about the questionable reliability and worth of information to be found on the web is a testament to this. Hardly surprisingly, of course, for the more information there

is around, the more tempting it is for people to power browse, to scan huge lists rapidly, grabbing an item here and an item there, without giving much thought to the relative merits of each. Often it is simply a question of fancy-ing this rather than that, but even when the choice is more considered, it is likely to be made on the basis of where an item ranks on Google. The higher the item is on a Google-generated list, the more it is invested with quality, but of course quality is hardly the prime determinant in Google's ranking. This when, given today's information affluence, the key to effective information man-agement can only be well thought-out selection: locating first the relevant and then, from among the relevant, the indispensable in terms of value.

Selection thus being the choice strategy for coping with the vast quantities of information surrounding us, obviously far better that it is conducted along logical grounds than arbitrary ones. Surely quality and reliability are better criteria for knowledgeable information selection than congeniality: after all, people do not really want to base a decision or a course of action on unstable foundations or be sent on wild goose chases, do they? Indeed, the quality of information, its veracity, trustworthiness and accuracy, are held to be, at least in theory, critical considerations, and for some fields and occupations particularly so.

Health is obviously a case in point, for poor information provision on ailments could have serious, even fatal consequences; but journalism, science, finance and business also come to mind. Take, for example, the following quotation from a leading crime correspondent:

> There is a problem with inaccurate information and particularly in my field that could be very dangerous because we run into problems of libel and we run into problems of contempt. Reporting crime if you get bogus information or inaccurate information about people's convictions or about crimes – then you are in trouble and I am wary of the internet for that reason.

Indeed, journalists are extremely concerned with authority and accuracy. Much of the unsolicited material that comes to them looks suspiciously like propaganda, public relations or advertising: it is difficult to distinguish fact from hype. In consequence, a good deal of cross-checking is done. At *The Economist*, for instance, where none of the articles are signed – and, in result, the reputation of the whole magazine is at stake, unchecked facts and unver-ified sources are simply not used. An advertisement for the journal once stated: '*The Economist* believes in collective responsibility. It commits its own reputation to every sentence it writes, good or bad'. To be sure, *Economist* journalists take the authority invested in them extremely seriously, as the fol-lowing story amply proves: at one editorial board the following week's edition was being considered and the conversation got around to an article about Mozambique. The question was: should the article appear the week before the forthcoming elections or during election week? One journalist argued that the

article should appear the week before, because, that way, the powers that be in Mozambique would read it and that would stop them making the same mistake as they did the time before!

Academics are another group with stringent requirements for excellence and dependability of information. They are very much aware of the importance of paying attention to the quality/authority aspect of their information work, and no wonder: they simply have too much at stake to neglect doing so, given that the inevitable prerequisite of every new scholarly progress made is its firm anchoring in previous knowledge and peer review. It is hardly surprising, then, that they are rather wary of information that is not endorsed by the formal scholarly communication system:

> … for the purposes of my research work … I use a good journal, something written by somebody from a good university. I can't base my hypothesis on findings posted on the web, which haven't been published in a journal, it's just not reliable, that's why, even if the idea is terrific, I'll still take it for something unreliable.

By necessity, then, academic researchers are considerably more skilled and competent at the task of evaluating the information on hand than most people – and it does take quite some expertise.

For one, determining the quality of information is rather more difficult these days than it used to be: the digital environment is a complicated one in which to make quality and reliability judgements. It is a relatively new (for some) and fast-changing environment, with new sites appearing all the time. Also, there are so many parties associated with the production of a digital information service – experts, governmental, scholarly or commercial content providers, broadcasters and publishers (and Joe-public in the likes of blogs), to name just the major ones. It is clearly a situation in which authority is 'up for grabs' and worth is by no means assured.

Further to that, in today's information realities the task of sifting out the wheat from the chaff in the information of potential relevance often falls to the user; the value and trustworthiness of information is no longer a priori established for its potential consumers. The traditional library-driven user of the not so distant past relied on the library for (limited) choice, and for a stamp of quality or authority. Seeing that ascertainment of quality was part and parcel of the authoritative professional preparation of information for its central storage and provision in a library, the not wholly unfounded assumption was that if it was in the library, it was good. In any case, the choice was largely made for the consumer because the intermediary conducted the search. Today most people search for themselves, often from non-library or -evaluated information environments. To stay afloat in the ever-expanding, mostly digital information environment they need to evaluate, and evaluate well. Presented with massive and increasing choice via internet-based, often unvetted channels, they are forced to make the evaluations once made by librarians; in this

context the phrase 'we are all librarians now' is an especially apt one. With so much choice and new products coming on stream, they have to make many, many evaluations, and quickly. No mean feat, that, as for a user to select information on quality grounds involves multiple layers of interpretation derived from their experiences, perceptions and private knowledge related to the particular information need at hand (Park, 1993).

Seeing that the perceived quality and reliability of a document thus rests on a mix of highly personal knowledge, past experience and accumulation of information, the best way to assess its value is simply reading it 'cover-to-cover' to see what it says. However, the convenient availability of huge amounts of information on any and every conceivable topic under the sun renders the attempt to do so unrealistic; so much so that nowadays even academics are far less inclined to take their quality decisions on the basis of a straightforward perusal of the information presented to them. Instead, people look beyond the inherent appropriateness of documents in information need situations, making judgements about the *source* rather than the *content* of the information. In point of fact, they resort to long-standing aids or 'props' for picking out as efficiently as possible the worthwhile items from among the huge amounts of information at their disposal.

Thus, knowledgeable users aim to single out from the abundance of obtainable information the items of appropriate authority and quality, employing one or more indicators of value, often using all available indicators in combination. Indeed, establishing the authority and/or determining the quality of information is for many people a two-tiered process of first selection made on the basis of both information giver/sender and source (authorship and channel of publication/dissemination), followed by a more in-depth scrutiny of the items, which have been found to merit further consideration. Hence, the value and trustworthiness of some information is often assessed by first noting who recommended it, who its author is (with what organisation the author is affiliated, the author's academic background, etc.), and the journal in which it is published or the site on which it is posted. It is only if by this stage the information looks worthwhile that a more in-depth examination of its contents is deemed to be in order.

Starting out the evaluation process of some information with its giver/sender makes eminent sense in our era, when more and more information comes to us, rather than vice versa. With social networking on the web having become almost normative, friends and colleagues seem to be behind much of the new information that flows to us. Indeed, Schwartz (2008) cites Ted Eytan, medical director for delivery systems operations improvement at the Permanente Federation, who says that 'patients aren't learning from websites – they're learning from each other'. Given this state of affairs, the following quotation, which captures the popularly held sentiments concerning the information giver/sender, sounds more pertinent than ever: 'One of the filters on what stuff you read is who sends it to you; the weight given to the information will depend on the source providing that information and more weight will be

given to information provided by a source if that source has a high position in an organisation'. Indeed, the aforementioned study into the use and impact of key digital health platforms and services in the UK (Nicholas et al., 2007a) found a relationship between how users heard about health websites and the sites' rating. Those users who were recommended the site were least likely to say that site trustworthiness was poor or okay. Users who arrived at the site via an advertisement or a search engine were most likely to rate trustworthiness as either poor or okay.

Beyond the perceived authority and trustworthiness of the information giver/sender, it is the source of the information that is often the basis for determining its quality; therefore, a very good appreciation of the information producers in the subject field is quite important. This is how one academic put it, pondering the rationale behind it all: 'If I know the author, I know exactly what to expect: I know his areas of interest, and more importantly, I know the worth of the information he'll give me'.

Indeed, another important factor in determining the reliability or quality of some information is the reputation of its author and that of the institution with which he/she is affiliated, as these are perceived by the reader, with the latter, as Liu (2004) notes, considered even more significant than the former. This presents no problem for those in the know, for, as Becher (1989) points out, one of the striking features of academic life is that nearly everything is graded in more or less subtle ways, institutions, departments and, possibly above all, the scholars in a given field. Thus, to quote an economist, well-versed in the ways of academe, academics:

> Walk around with invisible ranks … they all look the same, not even wearing suits or ties, or perceptible ranks … and yet, whoever has to know, knows without any doubt who's leading and who's led, who makes the decisions and who follows.

It is not very surprising, then, that academics can form a well-founded first impression as to how good and solid a given piece of work is almost at a glance: if the information being considered originates with somebody who holds a senior rank in a 'first line' university with a 'good' department in the relevant area of specialisation, whose name is known in professional circles, it passes muster. So much so, that, as Park (1993) notes, a prominent scholar in the field tends to become an independent quality/authority parameter, regardless of the subject matter of the publication being considered (along the lines of 'anything he writes is bound to be good'). In fact, as Kling and McKim (1999) assert, a non peer-reviewed posting on a website by a high-status and well-respected scholar may well be trusted and valued more than a peer-reviewed journal article by someone not as well-known in the community.

Yet, with all the importance accorded to the author's status in the evaluation process, basing 'a vote of confidence' solely on the author's standing is obviously inadvisable, as an experienced academic researcher explains:

> Clearly I treat differently the work of somebody, whose contribution to the field is indisputable, than that of a fledgling, who still sports a few bits of the eggshell on his head … Of course I do, and so does everybody else … Still, it is always possible that in some of his articles the author just reiterates previous sayings, and if so, prominent or not, these articles go straightaway to the pile of 'not worth reading'.

Moreover, basing quality/authority judgements on the reputation of the author is pertinent to no more than a small segment of information consumers; it would be hardly reasonable to expect that outside academe information seekers always or even often be aware of the scholarly weight of the authors, the actual producers of some document on hand. Still, things are very different indeed when it comes to the standing and dependability of the organisations with which these producers are affiliated.

People are well aware that certain organisations, because of their economic or political power, command particular authority or respect and, as a consequence, their publications or websites are in effect brand names, which carry a lot of clout. Thus, for example, BBC Radio 4's Today programme is widely recognised throughout the media and political worlds as being agenda setting – largely because of its flagship status and because important and authoritative people are drawn to its studios. Perhaps as a consequence of people's concerns over the quality of the information on the web ('I think that most people realise that there isn't any particularly effective restrictions in place as yet on the internet, therefore, anybody can set up a website containing false, misleading, incorrect or offensive content'), the websites of government, the European Union and some academic establishments, for instance, are particularly popular. Indeed, they are frequently used by journalists. A case in point is the science editor at *The Guardian*, who regularly consults sites such as The NASA Pathfinder mission, The American Association for the Advancement of Science, or The Global Seismology Unit, in pursuit of quality information. However, if not so long ago people unquestioningly accepted that information from such governmental bodies would be accurate and reliable (whether conveyed via the internet, hard copy publication or any other means) and were prepared to use it straight off, this no longer holds invariably true.

The findings of the above-mentioned health survey (Nicholas et al., 2007a), which, having canvassed hundreds of thousands of users on a national scale, can be taken to be very indicative indeed of generally held views, serve as ample proof of this point. Apparently, people may be by and large supportive of the British National Health Service (NHS), declaring that they trust the NHS 'to a large extent' or 'on the whole', but nevertheless, they also look elsewhere for information. They largely do this checking on the basis of long experience of searching the web, a lot of practice in making constant comparisons and through a process of trial and error. Plainly, people apparently take full advantage of the above-noted capability provided by the web to 'suck it and see'. Here is how one survey respondent put it:

> I trust the info [sic] provided by the NHS. However, like treating any other … sources, I usually check up a few other professional websites for the same info, so I get a more complete picture of the subject, instead of relying completely on … one source.

This seems to suggest that not even so-called 'authoritative' publications and sites are believed to be always correct. Rightly so, it seems: as the literature testifies, even 'official' information published by the NHS and other government bodies can be of dubious quality. Coulter et al. (1999), for example, point to a multitude of problems: much of the information at the time of their research was inaccurate and out of date, technical terms were not explained, and few materials provided 'adequate' information about treatment risks and side-effects. Huntington et al. (2004) also found that the NHS Direct site was in fact poor at sourcing its information, possibly, or so they surmise, as the NHS considers itself an authority in its own right and hence does not see the need to quote its information sources. The site was also found wanting in terms of date-stamping its information.

Further to users' awareness of the possibility that even organisations formally qualified to furnish authoritative information do not always live up to expectations, there is also the problem of government bodies being suspected of having an official 'axe to grind'. This, obviously, may greatly impact on the perceived quality and credibility of the information these establishments offer. In the words of one participant in the just-cited health survey, the NHS is considered:

> Trustworthy in the sense that the information present will be accurate on the whole but biased in the sense that there is an economic consideration with healthcare provision and the NHS is representing the Department of Health (DoH) policies and may not promote treatments that are not available widely on the NHS or new research.

Another instance where users can and do put brand indicators to good use in accrediting the information found is when their source is a peer-reviewed scholarly publication. Each discipline has its fairly well-defined, if unwritten hierarchies of scholarly journals and publishing houses, with some publication venues considered of a higher quality than others. This pecking order of scholarly information dissemination channels is so well-known, so much part of the disciplinary culture of any given field that active researchers can rank publishing venues on impact or quality grounds, and the level of agreement amongst them would be exceedingly high. They usually acquire this intimate familiarity with the relative standing of the different 'brands' of information dissemination routes in their respective fields as an essential part of their socialisation to their chosen profession, but also the hard way, through the process of submitting manuscripts (first-line publications are in such high author demand, that they can afford to be very choosy indeed). Accordingly,

although more often than not they know exactly what the quality of their research is and send the manuscript to a journal that reflects that quality, if they cannot get into the journal of appropriate ranking chosen, they go to the next one down the ladder. Quality thus tiers down and, in result, the archiving of scientific ideas and findings is in fact on a continuum, with varying degrees of value, reliability and peer validation, from the most prestigious and rigorously reviewed at the top, to what is virtually a vanity press at the bottom (Harnad, 1990; 1999). Indeed, it is by no accident that a relatively small number of the titles available account for a relatively high proportion of use, a long-known characteristic of information-seeking behaviour, which has been proven to hold true in the electronic environment, too (Nicholas and Huntington, 2006).

This state of affairs is well-known to potential information consumers, certainly within academe, but often without it, too: 'If I read an article in *Nature* or *The Journal of the British Medical Association* then it is not an unreasonable assumption that the guy knows what he's talking about and that his colleagues believe he knows what he is talking about'. Furthermore, if in the past only the leading information dissemination 'brands' were known to the general public (*Nature* and *Science* may always have been household words, but not many other examples readily come to mind), it is hardly the case these days. The ethos of the knowledge society, with its emphasis on knowledge and information as the key to success has brought about a huge demand for scholarly information (Nicholas et al., 2003a). We are all amateur scholars now, and thanks to the greater access afforded by Open Access and institutional or disciplinary repositories, we have a library of immense power at our beck and call. Indeed, as Nicholas et al. (2008b) found in their Virtual Scholar research programme (2001–08), the so called disenfranchised users of Open Access material (that is, users from outside academe) are large in number and are fuelling a lot of the growth in scholarly information consumption. With many of the general public thus having joined the army of virtual scholars, they, too, are quite familiar with the literature in their areas of interest, inclusive of the aforementioned 'pecking order' of the scholarly information dissemination channels. In result, they are almost as adept at judging some information by its brand, that is, by the reputation of its venue of dissemination, as their academic counterparts.

Let there be no mistake, however; scholars we all may be by now, but amateurs, not professionals: translators of research into actionable outcomes – practitioners, research users, policy-makers and opinion-formers. Most people are, then, readers of scholarly information rather than its producers. Thus, amateur scholars may recognise brand names of publication outlets as readily as professional researchers, but they have little need or use for the time-honoured tool widely utilised in academe for identifying qualitative and authoritative information: citation analysis data. Actually, citations are very important, offering insights into which journals are, on average, the most highly cited (ISI impact factor), which journals researchers turn to first (immediacy index),

and the long-term value that academics ascribe to particular titles (cited half-life). However, all these are indicators of value from an author perspective, providing as they do reliable information about the preferred publication outlets in a given field. For the vast constituencies of readers who do not write papers, citation data cannot be of much help. After all, the fact that an article has been highly cited does not necessarily testify to its reliability or commendable quality – the very opposite may well be true, with the citations reflecting refutations of its content! Indeed, Rowlands and Nicholas (2007) propose that we systematically measure, at the individual article level, journal use ('votes by readers') as well as journal citation ('votes by authors'), among other purposes in order to aid all scholarly information consumers, and not just the authors among them, in acquiring a good understanding of the usefulness of some research.

In any case, at least for academics, who need trustworthy and high-quality information for most of their work-related purposes, it seems to be a foregone conclusion that they will give much more weight in their choices to the brand names in scholarly publications, the ones considered to be as near as possible to the top end of Harnad's continuum. After all, many people set great store by a brand precisely because it is perceived as providing special benefits and added values. Indeed, as it has already been mentioned, the esteem in which a journal and/or a publishing house are held is often well-known and frequently used as an indication of the quality and authority of information, as one academic, summing it all up, says:

> The author, and in fact the whole journal, serve as parameters for determining the quality and authority of a publication. As it happens, some very, very interesting things appear in negligible journals, written by authors you don't know, but when you have such an avalanche of material, that's the way you work.

However, contrary to popularly held notions, this building on brand names in quality and authority assessments of scholarly material cannot be taken as a sure-fire defence against alighting on information that is not really up to par. In fact, these days a degree of wariness seems to be called for in our dealing with scholarly information, the consequence of the ever-more pervasive doctrine of 'publish or perish' in academe, which, according to the editor of a history journal amounts to 'far too much publishing and not enough perishing' (quoted in Garner et al., 2001, 253).

Originally coined by Logan Wilson in 1940, the term 'publish or perish' has become the universally accepted shorthand for the onus on academics to publish copiously, as Schauder (1994, 82), summarising Wilson's thesis, describes:

> … because publication is one of the main measures of academic productiveness, there is pressure on authors to fractionise their research projects into as many separate articles as possible to 'add yardage to the author's

bibliography' … [Since the] publishing of articles … [is] a principal means by which academics can achieve visibility, and therefore, advancement … , [these] situational imperatives dictate a 'publish or perish' credo within the ranks.

Unfortunately, the 'publish or perish' system, focusing as it does on the quest for quantity of publication, may lead to inconsequential publications: 'salami publications', wherein material adequate for a single paper is sliced into several 'least publishable units', or 'meat extenders', i.e. papers re-issued with no new data or papers that are in fact two previously published papers merged into a new one. The result of this perceived need in academe to skew research toward what is 'acceptable' rather than toward what may be important is, at least at times, a decline in the quality of the information, as Ziman (1970) puts it:

> Not only is there too much scientific work being published; there is much too much of it … the need to get recognition by publication forces each of us to shout a little longer and louder so as to be noticed at all in the gathering, swelling crowd of voices … The result has been a proliferation of semi-literate, semi-scientific, half-baked and trivial material which threatens to swamp the system.

It seems, then, that these days even governmental/official or scholarly establishments can no longer be unhesitatingly considered purveyors of innately valuable and authoritative information. Rather predictably, commercial organisations are held even more suspect in this regard. So much so, that in the UK survey of health platforms and services most respondents, 63%, said that being too commercial would be a reason not to visit a site (Huntington et al., 2004). Indeed, over half of the people polled said that they had not returned to a site because it was too commercial (Nicholas et al., 2007a). In the same vein, when a site carried 'obtrusive' or too many advertisements, people were much more likely to rate its trustworthiness as poor or only okay. No wonder, then, that clicking on a banner advert was found to be the least popular way of finding health information.

True, some commercial establishments do carry a lot of respect in result of the quality of their work and excellent track record. A particularly telling example of this came up in CIBER's above-mentioned health survey (Nicholas et al., 2007a). People who used the touch-sensitive screen health information kiosk at a Safeway (supermarket) Pharmacy were asked about the authority of the information provided by the kiosk and the trust they exhibited in it. In reply, users tended to say that they had 'every trust in the information', but also to assume that it was a Safeway service (it was, in fact, an NHS-provided one). When told that this was not the case, many were disappointed, and the younger the users, the more disappointed they were. This, because Safeway was seen as a very successful business, whereas the NHS was always being criticised in the press for not coming up to scratch. Moreover, even when it was or became

clear that Safeway was just the host organisation, its reputation carried over to serve as a guarantee for the provision of quality information only on its premises: 'Well, they wouldn't let any cowboy stick a kiosk here, would they?'

By the same token, newspapers, with their wide and large readership – including lots of decision-makers, have great impact. Indeed, articles in newspapers like the *Financial Times* and the *Wall Street Journal* can actually move financial markets and bring down economies (in one of its advertising campaigns the *Financial Times* played upon the crucial role it has in meeting the information needs of the business world: 'no FT, no comment' was the advertising slogan). Undoubtedly, for-profit ventures of this stature are invested with special authority by many people, but they are the exceptions to the rule, certainly not the rule.

All this seems to lead to the conclusion that by now the time-honoured reliance on the source of some information in the initial assessment of its authority and quality certainly has undergone fundamental changes. Not that it has become any the less important; rather the contrary. Given the surfeit of choice characterising the mostly online information environment of today, singling out the items of intrinsic merit on the basis of a preliminary judgement of their extrinsic characteristics is an inescapable imperative. Indeed, the evidence that the CIBER research group has amassed in several projects points to the fact that on average most people spend only a few minutes on a visit to a website, insufficient time to do much reading, but perfectly adequate enough to note some indicators of its worth as a source. However, if this preliminary assessment of the information on hand is as crucial a component of information seeking as ever, if not more crucial, the way people go about it is much changed in today's web-based world. People simply shop around for information and on the basis of comparisons take decisions for themselves on what they perceive to be 'good' information. Present-day digital consumers are what can be termed 'promiscuous users'.

Promiscuity results from consumers' uncertainty in the face of massive access and choice, coupled with short attention spans and a tendency to leave their memories in cyberspace. In information-seeking terms it manifests itself in two ways. Firstly, people visit a (large) number of websites to find what they want. Secondly, and this is related, they do not often return to sites they have once visited. This form of information behaviour may sound like a dumbed-down form of information searching and retrieval: people are seemingly unable to make up their minds and, in any case, succeed in obtaining just a thin veneer of information. However, it is simply a different way of tackling an information need. One is minded of the father watching his young daughter using the remote control to flick from one television channel to another. A slightly irritated father asks his daughter why she cannot make up her mind and she answers that she is not attempting to do so; she is actually watching all the channels! She, like today's information consumer, is gathering information horizontally, not vertically. The single, high-quality and authoritative source, which is always consulted and deeply mined, seems to be a thing of the past.

Indeed, a recurrent finding in the various CIBER projects evaluating the information-seeking behaviour of a number of information communities, most notably those associated with news (Nicholas et al., 2000), health (Nicholas et al., 2007a), voluntary and charitable work (Nicholas et al., 2004a) and scholarly publishing (Nicholas et al., 2008b), is that appraisal of the quality and trustworthiness of information is largely undertaken by making comparisons, which seems to be a key element of digital literacy. People will look at several internet sites (and, therefore, consult several organisations) for information. This appears to be true even when people 'trust' one site, or when they go to 'brand name' sites. Perhaps not surprisingly: as Simon (2001) observes, brand recognition has taken on a new dimension on the web, with the surfeit of choice online producing 'a concomitant change in consumer attitudes', moving them from what he describes as 'receptive space' to 'sceptical space'. Users can afford to be sceptical about the attributes of an individual site, as these attributes can be maximised by visiting a number of sites. Indeed, in the aforementioned health services survey (Nicholas et al., 2007a), a relationship was found between a respondent's rating of the content and the number of health sites visited. As the number of health websites visited increased, so the user's appraisal of content depth, breadth and trust declined. The ease of information access has made internet users into information connoisseurs.

Hardly surprisingly, then, loyalty might be a thing of the past, too, although we have loyalty to Google, Facebook, BBC and ebay in bucketloads. Coming back to a site constitutes conscious and directed use, which clearly suggests that it is a tried and trusted source of information. This 'site stickiness' (as the industry calls it) makes return visits indicative of the perceived quality and authority of the site. Indeed, a service with a high percentage of returnees can be regarded as having a brand following, the goal of all service providers. However, loyalty or repeat behaviour generally is not a trait of the digital information consumer. Thus, for example, a study of the SurgeryDoor website (Nicholas et al., 2007a) found that over a relatively long period of 12 months, two-thirds of visitors never returned and the remaining 33% visited the site only two-to-five times. By the same token, in the 'Google Generation' project (CIBER, 2008), the data on the return visits to ScienceDirect, a scholarly journal database, show that over a five-month period 40% of users just visited once, 24% visited two-to-five times, 15% visited six-to-fifteen times and 21% visited over fifteen times.

It seems, then, that in today's information setting people have become consumer 'checkers' or 'evaluators'. They may happily avail themselves of innovative information services, with some of them actually describing their routine usage of websites and internet-based repositories as being 'hooked', but often they are also very cognisant indeed of the perils entailed by indiscriminate use of information distributed this way:

> … these electronic archives are definitely no substitute to a regular journal. For example, if you log in today and find a pertinent article, tomorrow

you may discover that it is no longer there, in all probability because the author has decided to retract it for corrections … an item posted there has not been refereed, so you've got to be very careful, you can't trust the information too much.

Indeed, the information seeking of the present-day consumer has more in common with the behaviour of shoppers than with that of traditional library users. This is because there is now a huge, rich market for information, and obtaining information is part of the shopping experience. As a shopper it is our duty to be a smart shopper, who plays the market. Nobody wants to spend time and money on obtaining information of inadequate quality and authority, and with good reason, too, as Nicholas et al. (2007a) found in their evaluation of digital health platforms and services in the UK. A person's trust in a site was a significant factor on health outcome, with those users demonstrating the greatest trust being more likely to claim a positive health outcome. Thus, those rating the site's trustworthiness as either good or excellent were more likely to say that they had been helped a lot and were less likely to say that the site was of no help.

There is a possible role here for the information professional in providing quality assessments, especially in the construction of information-filtering mechanisms that take account of quality and authority criteria. Plainly, web search engines and their relevance-ranking methods are not always sufficient. However, neither are the commonly cited quality criteria for online content, as a 2008 study into breast cancer information online clearly indicates: apparently, no quality criteria or website characteristic, singly or in combination, reliably identified inaccurate information (Bernstam et al., 2008). Librarians seem to be facing up to the challenge by developing their own search engines. These are meant to offer the trusted and effective searching environment that is missing from the big search engines and Google Scholar.

Date/currency

Everyone works with a mix of new and old information: even stockbrokers, preoccupied as they are with the telephone and real-time services, need to place data into a context. Thus, two closely related questions have to be asked regarding this information needs characteristic: firstly, how up-to-date does the information need to be; and secondly, how far back in time the information is required to go.

Information seekers are always likely to require the most up-to-date material, even if they do search for data back in time, too; by definition – the new captures the most interest. This seems to hold all the more true in today's information society, in which people consider keeping informed of new developments in their diverse areas of interest and attention an essential part of their customary pursuits. Thus, currency may be only one aspect of the date-range requirement, but these days the pressing need to always have the very latest

information puts it very much in the spotlight. However, 'current' may have quite different meanings for different people and even for the same person in different circumstances, from 'this very moment' to 'nowadays', although the web has raised everybody's expectations. Thus, as Wilson (1993b) points out, one may claim to be current without claiming to have information about what is happening in 'real time' – i.e. right now. A market maker in a leading stockbroking firm would probably consider the last few minutes' information to be current, but for the historian the definition of current might well extend to a year or more.

Be an individual's precise definition of 'current' what it may, conventional wisdom holds that an information source is at its peak of use just after release, becoming less and less frequently used with time. This truism might not have gone uncontested – for instance, Rothenberg (1993), reviewing the body of literature testing the notion, argued that studies had failed to show the expected measurable decline in use. Still, whether there are demonstrably fewer readings of older materials or not, they are rated more important than new articles, more time is spent reading them, and a higher proportion of them are consulted to prepare a formal publication, such as an article or a book (King and Tenopir, 1999; Tenopir and King, 2000). Indeed, early retrieval systems traditionally proceeded from the notion that the most current information was the most important for users. This is the reason why they ordered their output in reverse chronological order. In fact, the provision of current information was long held to be the hallmark of a good information system – and for many people it was the performance indicator by which they measured its efficiency. Clearly, users are, and have always been interested in the new, but, as Odlyzko (2000) prophesied almost a decade ago, and as the data gleaned in CIBER's Virtual Scholar research programme (2001–08) prove beyond doubt, it would be a mistake to go on believing that this is the picture in its entirety. Apparently, once the visibility of older material increased, thanks to the massive improvements in access to back files, on the one hand, and the ubiquitous use of search engines that prioritise relevance over age, on the other, a much wider usage of older materials followed (Nicholas et al., 2008b).

Indeed so, but what do we mean when we say 'older material'? Undeniably, if people's perceptions as to their currency requirements vary greatly, so do their notions of how far back in time information may still be relevant to their needs. Hardly surprisingly, though, for obsolescence, the decline in the use of published material over its lifetime, is the result of change in general, and in the individual circumstances of the information seeker, in particular. Indeed, new discoveries, new equipment, new technologies, new legislation, as well as political and economic factors can render valueless – even dangerous – what we know and do. For instance, consider the value today of a book published on central heating systems for the home in the 1990s – books that can still be found in some public libraries. The relative prices of fuels have changed; much of the equipment featured would be obsolete; and new energy-saving features would not be mentioned. No wonder that it is the shelf-life of

technical information that seems to be the shortest, on par with that of news. However, information decays in all fields, as the following account of an archaeologist amply illustrates:

> Without a doubt, in archaeology information does become obsolete. A book published, say, in 1932 is worthless by now … you can't use the information in it, because it is no longer correct … the dates given are wrong, the facts are incorrect … I have in my private collection some older books, but not one line therein is still valid.

Given that the decay of information is so much contingent on change, it comes as very little of a surprise to find that obsolescence is clearly a discipline-specific phenomenon. After all, as it has already been noted, the various areas along Storer's (1967) continuum of 'hard' to 'soft' domains are characterised by a very different pace of change and development. Thus, the knowledge domains towards the 'hard' end of the continuum are fast-moving fields, characterised by rapid and linear changes. Since, as Meadows (1974) puts it, scientific knowledge grows in the orderly fashion of a skyscraper being built, with each new floor depending on the previously constructed floor for support, steps forward can be made only when the current problem at the frontier of research is resolved. This state of affairs obviously dictates a dynamic tempo of information creation, bringing about frequent change and a correspondingly high rate of obsolescence. Take, for example, the above-quoted psycho-oncologist's description of her field:

> … the world of medicine is one of the most rapidly changing, one of the most dynamic … whatever held true five years ago is more or less rubbish by now … just like my PhD dissertation, which is all but ready for the bin by now, since the huge progress in the prevention of the side effects of chemotherapy for cancer has rendered all I had to say on the subject of patients' coping with the treatments just a few years back very much outdated.

In comparison, the humanities do not normally evolve in a linear fashion; one discovery is not necessarily the result of a prior one and will not necessarily lead to a later one. As such, to use again Meadows' (1974) picturesque simile, growth in the humanities might more reasonably be compared with the construction of a rambling country house. Therefore, new developments in the knowledge domains towards the 'soft' end of the continuum occur at a much more leisurely pace, in consequence of which information decay occurs much more slowly. Still, if in the past obsolescence in humanities research was an almost unheard of phenomenon, with new research usually supplementing rather than superseding previous knowledge, nowadays there seem to be instances of information becoming obsolete. This is what a historian has to say on the subject:

… in the humanities too, in history, philosophy, theology, there is so much research going on that now, and it was not so in the past, material does become obsolete … A significant part of this innovative research truly sheds new light on the issues being considered, provides us with further understandings and different approaches … some humanities research done ten, fifteen years ago has simply become obsolete and it never used to be the case! So you can't say today 'I don't get to the recent work in my field', it won't hold water, because in most every area there are some new, very central, very important works, which have changed in one way or another the concept, the outlook, the understanding, the whole information infrastructure.

As a rule, the very nature of information use in the different knowledge domains (particularly for research work purposes) dictates these differences in the levels of reliance on past information. As it has been previously noted, at the 'hard' end of the continuum information consumers must have the results of previous research pertaining to their own work, though not the specific writings reporting it, for where the findings of previous generations of scientists are still relevant, they are part of the building blocks of science and, therefore, readily available in textbooks, treatises, handbooks, etc. In stark contrast, in disciplines at the opposite, 'soft' end of the continuum, seekers of information cannot incorporate prior knowledge in their own undertakings unless they get hold of the specific documents that convey it, for the unique insights of the author form a vital part of the breakthrough reported (Bates, 1996; Stoan, 1984, 1991). No wonder, then, that there are distinct, discipline-specific patterns of the time-depth needed in information work.

Obsolescence of primary ('raw') material occurs (if it occurs at all) more slowly towards the 'softer' end of the continuum of knowledge domains: thus, for example, nobody could ever substitute an article summarising, say, Plato's thinking for the original works of Plato, and not only because the reviewer may have misunderstood the great philosopher or made a mistake in interpreting his work, but also because new insights can only arise from the original text. In fact, such primary information can actually gain value with age. Something written by Churchill during the Second World War would be valued for its age – it would be a source document; by its very nature it cannot be superseded. However, secondary works, which interpret the primary material, may very well age and become obsolete as new scholarly advances are made in the 'soft' disciplines too, although, again, the process is likely to take much more time. To be sure, citation studies show that the shelf-life of information in the 'softer' areas is a great deal longer than in the 'harder' domains (see, for example, Line and Sandison, 1974).

Even in the sciences not all information ages rapidly: the theory and fundamentals of many subjects are fairly constant and, in result, long-established (but revised) textbooks and manuals are still well-thumbed. Moreover, in every field there are key publications, which, having shaped the course of

research, left an indelible imprint on its development: ' … these are the most important works on the subject, the corner-stones of a knowledge area', says a biologist, 'which everybody seems to remember and cite, no matter when they were written. Therefore, I cannot say that I only need current material; sometimes I may go back as much as fifty or sixty years for the basics'. In the same vein, medicine often needs long-term retrospective information in considering the development of a disease. Also, apparently, there are instances when going back to the original publication is important, if not essential, in the sciences too, as a mathematician explains: 'I've often found that if I encounter some difficulties in understanding something, then the original article, the one which first reported the breakthrough on the issue I'm trying to understand … clarifies things for me. Since the original article explains what the author really wants to do and how he goes about it, it gives you a different perspective. In the books subsequently written on the subject you find all sorts of things he never thought about when he wrote his article, and the improvements made on his initial notion, but you are better able to understand the original idea, the message nobody had come up with till then, when it is being described for the first time'.

However, as people often intuitively know what the obsolescence factor is in their field, they customarily set out to look primarily for materials from a given age spectrum (from the almost unlimited time-depth in the humanities to less than five years in many areas of the sciences and social sciences). In direct consequence, they frequently use date of origin as a means of selection – sometimes, as a substitute for quality. Date can indeed serve as a useful cut-off point for a search that produces a lot of documents. Information professionals have long understood this, which is why the aforementioned reverse chronological display is such a well-established practice. Still, it is important to remember that the shelf life of the information is not the only factor that determines the date range required – the amount of time that a person has available to read and digest information also comes into it.

Of course, information units and systems themselves have a need to weed, discard and archive information, because of the cost and space involved in storing information. This is often done using date criteria. However, where information professionals often have got it wrong is in holding on to information long after it has become deceased (the collector syndrome), and in failing to understand the complex relationship between occupation/job role and currency requirements. Thus, for example, just because journalists deal with news as it is breaking, it does not mean that they do not need archival information. In point of fact they do – to put the breaking news into some kind of perspective.

For information to retain its currency it must be distributed quickly – something that is taken up more fully in the next section. Some information channels are more conducive to the rapid transmission of information than others: hence the tremendous popularity of e-mail and the mobile telephone. Traditional information systems have never been geared to providing access

to really current information – typically, abstracting services still serve up information that is three-to-six months old, although they no longer claim they are providing a current awareness function. Indeed not, for online technology and the internet enable the provision of services that are much closer to what people regard as acceptably current information. Thus, for instance, in the Western world same-day access to the full text digital edition of the major newspapers is quite the norm by now, and in many cases the digital edition is available before the hard copy.

Still, as the participants of the above-mentioned focus group of researchers reported, innovative alerting services, such as RSS feeds and the like, meet with very little success, if any. For example, on the eve of the 2008 elections in the USA, when online news consumption tripled (fully one-third of Americans reported getting most of their election news online, up from the 10% who did so four years earlier), still only 4% of American internet users subscribed to receive campaign or political information through an RSS feed. Even among voters who used the internet in one way or another for political purposes, who did take somewhat greater advantage of the ability to customise their news and get the latest updates on the campaign, only 5% set up a politics-related RSS feed (Pew Center for the People & the Press, 2008; Smith, 2009). Generally, people much prefer to rely on search engines to deliver 'current awareness on demand'. Mistakenly, as it happens, for the internet may be doing much to raise the currency performance, but too often it flatters to deceive. Sites are sometimes not even as up-to-date as their hard-copy equivalents, and when they are, their 'date stamp' is frequently either not very prominent or missing altogether.

Yet, currency is popularly held to be a prominent attribute of internet-based information. True, in the aforementioned survey of digital health platforms and services in the UK (Nicholas et al., 2007a) few people cited currency as an advantage of the internet over other sources, but that because their assumption must have been that digital information, almost by definition, was up to date; a dangerous assumption, of course. Nevertheless, people did say that the 'fact' that it was up-to-date information was instrumental for them in replacing other sources. Thus, for example, according to a company director 'books are too out of date relating to medical matters'. Similarly, another respondent felt that 'you will always be able to access the most up to date information on the internet, whereas a library may not have it available'. Information professionals have long taken it upon themselves to alert their customers to the importance of evaluating websites not only on the basis of content, but also on the basis of other key elements of digital literacy: authority, access, design, currency and interactivity, to name only the most important. Indeed, this is what Friedewald (2000) must have had in mind when he urged medical practitioners to cultivate professional websites of their own, because doing so could help them steer their patients towards current and authoritative information on the internet.

Speed of delivery

Speed of delivery is all about getting information to people quickly – as quickly as the need for it. Obviously, information should not go 'off' in transit or transmission. This is ever-more true these days, when time is such a rare commodity. However, the centrality of this aspect of information need goes beyond the unprecedented time constraints characterising life in the 21st century. Clearly, we also seem to have set ourselves new standards for the speed with which information is counted on to reach us. Hardly surprisingly, of course, for the easy availability and effortless accessibility of the host of resources, channels and facilities truly enable the transferring of information from one end of the world to the other in a matter of seconds. Indeed, the host of internet-based information services, with their live broadcast qualities, on the one hand, and the omnipresent electronic information communication tools, the e-mail, the fax and, more recently, even the mobile phone, on the other hand, add up to a seemingly boundless capacity to satisfy every information appetite on the spot. Inevitably, then, where speed of delivery is concerned, people's expectations are by now sky-high. Total access, as quickly as possible, appear to be the present-day information seeker's key information needs requirements. Nobody wants to wait; nobody wants to queue – even if they could. We are all impatient and have zero tolerance for delay, as the CIBER studies (Nicholas et al., 2003b; Nicholas et al., 2004b; Nicholas et al., 2005; Jamali et al., 2005; Nicholas et al., 2006a) have shown time after time. Needs must be fulfilled immediately and information needs are no exception. A computer scientist, for example, admittedly of the ilk constantly preoccupied with 'getting there' faster than his colleagues, thinks nothing of setting the standards of speedy information delivery quite high when he says: 'who has the patience to sit and wait for an hour until a piece of information arrives?' Real-time information, once the exclusive and treasured domain of the journalist/stockbroker, is open to all and is now what everyone wants, it is the benchmark.

Indeed so: with IT-enabled increased speed of information delivery, a more instant response is demanded of everyone; knee-jerk reactions become the norm. Take the case of the stockbroker. Some 25 years ago human messengers would bring price information from the Stock Exchange twice a day and maybe this information would arrive 20 minutes late. On arrival of the prices a frenetic period of activity would begin. This would subside after a while and then the stockbrokers would prepare themselves (perhaps by reading) for the next price announcements that came much later in the day. Now, though, thanks to real-time online systems, they watch the prices change on the screen in front of them, seconds after they have been posted. This goes on all day long, there is no relief or quiet time; they watch the screens all the time. Stockbrokers are not the only ones who work under such time pressure; rather the contrary. Today's working environments are characterised by the urgency and immediacy with which tasks have to be done, which is of course the reason why rapid information delivery is so highly prized – almost above

everything else. In newspapers, for instance, if information cannot be obtained within half an hour or so, it simply will not be used. By the same token, a physicist readily admits that he will give up on a worthwhile idea if he is too short of time to do the information work involved:

> I may very well have an idea which necessitates that I go to the library. Now if I don't have the time or if it's too much trouble, it's not that I'll write the article without information, but I may decide that I won't try to solve the problem at all.

Thus, even academic researchers look forward to as swift a solution as possible to their information needs, and it is not merely because they want to avail themselves of what there is for the taking. Actually, obtaining information quickly is a very real need of theirs, as a historian explains:

> When you are in the midst of this process of investigation and analysis, and you get to a certain link which seems to be missing, you go to your information reservoir … if you find your answer there, then you can continue with your work, but if it's not there, and you can only get what you need later on, you're stuck at the point you've reached until you do … So the velocity of the information flow is indeed immensely important for people engaged in intensive research work.

Another academic, an expert on social welfare, goes even further than that. Contending that working under pressure is inherent to research work, he believes that in consequence the speedy meeting of information needs is an enduring prerequisite in academe:

> Even when you already have a reputation in your field, you still feel that you're under pressure to publish quickly, and therefore, you also feel under pressure to obtain information quickly … First of all, I think that in academe you're socialised to work under pressure, so when you no longer need to do so, you're already 'infected', so you keep working in the same manner. When you work on something, you always want it either validated or refuted so that you can get going, and this is not contingent on your academic rank, it's just how our work is. Also, when you're in the midst of developing something, you don't want to defer gratification for lack of information.

If researchers' perceived need for speedy access to information is the direct consequence of their feeling compelled (for extrinsic or intrinsic reasons) to produce and announce the results of their work quickly, it is particularly so in the highly competitive disciplinary culture of the sciences. Indeed, their circumstances are not conducive to tranquillity in any aspect of work, inclusive of its information-gathering component: 'When somebody comes up with an idea', says the above-quoted physicist,

It may very well happen that simultaneously five others in the world come up with the same idea. You want the idea to be chalked up to you, but if somebody precedes you, you can't very well say that you've done it too. If you look at it from this angle, there is a definite need in my field to obtain information quickly.

So much so, in fact, that according to Herman (2005), researchers' ability to find the time necessary to handle a certain quantity of information may occasionally even impact on their choices of topics for investigation. Firstly, the rigorous dictates of the 'publish or perish' mentality in academe, coupled with the externally imposed norms of gauging faculty productivity, enforce these days a brisk pace in every scholarly field. In addition, though, in the sciences and the social sciences the greater potential for profitable research findings, and the aforementioned much greater danger of those revenue-generating discoveries being 'scooped' by somebody else prescribe a more intensive – occasionally frenetic – rate of activity. In result, scientists and social scientists, especially the young among them, who still have to prove their abilities, are more inclined to weigh among the pros and cons of a planned project the amount of information needed, as it can have an impact on the time investment required.

No wonder, then, that many of the visitors to scholarly sites seem to be in a great hurry: they 'power browse', skimming titles, contents pages and abstracts, view only a few pages and do not stop long enough to do any real reading (Nicholas et al., 2008b). Seekers of scholarly information are not alone in manifesting this manner of behaviour, either. Indeed, the current interpretation of how quickly is 'quickly' has clearly become much more stringent in the universally felt pressured atmosphere we all inhabit now. In these times of electronic access to information, 'quickly' seems to be no less than 'immediately'; not 'in a few minutes', not 'soon', but now! Plainly, it is not that we all have to have our information needs met in real time, or with a high degree of urgency, but that even people who do not need information that quickly are impressed by rapid response – after all, it is another performance measure. It is synonymous with efficiency and efficiency at all times and in all respects is one of the hallmarks of life in our time-starved realities. Computer experts' conduct is perhaps the quintessence of this attitude, as one of them explains:

I strive for efficiency; anything slow is inefficient and as such, irritating. My wife likes to recount how in the cafeteria, where there is a choice of two soups, one chock-full of vegetables and noodles, but more expensive, and the other watery and unappetising, but cheap, computer people unfailingly opt for the low-priced alternative, because it is the more cost-efficient choice. What about the pleasure factor? Well, for computer experts that doesn't enter into it … That's how we are, that's the way things are in our field; we try for maximum efficiency. We relentlessly struggle to improve our solutions … constantly seek to find ever-more efficient ways to solve the very same problem … so anything inefficient gets on our nerves, really offends our sensibilities.

Unfortunately, traditional information services generally respond relatively slowly, though perhaps not as slowly as they used to (a prompt information service, according to librarians of the old school, was putting a book on the shelf within three months or so of the request made for its acquisition). Still, it was not so long ago that The British Library Document Supply Centre's inter-library loan performance of five-to-seven days from query to receipt of document was being trumpeted as a success story. Given the standards we are used to by now as to the speed with which information can be delivered to our doorsteps, it becomes more understandable why information seekers growingly desert the physical library (Martell, 2008). Thus, even academics, whose work is ostensibly done at a relatively leisurely pace, increasingly consider taking even a short journey to the library building the least preferable option for fulfilling their information needs, if an option at all, for:

> If I don't have the information on-line, I have to go to the elevator, wait a long time, go down to the library, only to discover when I finally get there that somebody has just checked out the item without which I'm unable to continue working. It can mean a waste of three, four, five days of work.

Indeed, none of the participants in the aforementioned focus group of practised and experienced academic researchers used libraries in any regular or strategic sense. 'Libraries are empty nowadays', said one of them. Another volunteered that he had not been in the place for three years and he said this with no sense of guilt. Many researchers, especially scientific ones, really seem to have fled the library (Nicholas et al., 2008b).

Yet, the regrettable fact remains that many information units still trundle along in a rather unhurried fashion, ignoring the huge changes that have occurred around them, which makes them look almost prehistoric by comparison. Thus Amazon's ability to get a book to a user within a day has shone the spotlight on the performance of most libraries in getting books to their customers. No wonder, then, as Nicholas and his colleagues argue (Nicholas et al., 2008b), that the consumer is moving closer to the publisher or distributor and away from the library as a provider of information. True, full-text online services, electronic document delivery, the fax and, above all, the internet have all come to the aid of the information centre to make for – in theory, anyway – a much more responsive service. However, as Moss (2008) contends, making a strong case for bringing the library back to its rightful place at the forefront of the information chain, the technology-driven access and service focus of the present-day library is hardly the way to go ahead, in any case. Rather, he suggests, it is the time-honoured curatorial actions of the library which should be emphasised, the selection/appraisal/privileging of content for user communities. Quoting Levy (2001, 197), he thus calls for reinstating the library as a warehouse of organised and ordered knowledge, a move that, alone, can ensure that it fulfils its true role as a space for reflection and contemplation:

> For some of us, books and libraries symbolize some of the very qualities and modes of being that are threatened in our fast-paced instrumented lives. Books speak of time and depth and attention. They speak of a slower rhythm of life … Libraries are places not just where books can be found, but where people can temporarily remove themselves from the speed and busyness of life, where they can read and write and reflect.

It is speed, speed and more speed that appear to be what most people want, whether they wish to get from place to place or have their information needs met. It is in cognisance of this fact that web designers and computer manufacturers continue to see the reduction of response times as their main goal. There can be little doubt about it – everybody wants quick wins, pure and simple. So much so, that, as Russell (2008) points out, speed is often preferred to accuracy (or authority), which is why, according to Lippincott (2005), students usually prefer the global searching of Google over the more sophisticated, but more time-consuming, searching system provided by the library, where they must make separate searches of the online catalogue and every database of potential interest. In a similar vein, as the findings of a benchmarking survey of e-book usage and perceptions in more than 120 UK universities indicated, when searching (rather than browsing) was used as the means for locating e-book content, the most popular search, by a considerable margin, was the quick search (Nicholas et al., 2008c). Indeed, as Russell (2008) goes on to say, website usability is not just about making sure that everything on the site works and relatively easy to navigate and use; it looks at speed of use, as well. Thus, when it was still common for sites to exhibit advanced technological facilities that had the effect of slowing down page loading, visitors were more likely to try their luck elsewhere. By the same token, those sites that offered a range of search methods were preferred, for then users could find and select the fastest method to suit them. Therefore, he concludes, where internet users are concerned, registration or application needs to be as simple and quick as possible, and any feature that speeds up the process is an advantage.

Finally, one last word about speed of delivery and its affect on currency. The faster people can get information, the more current it is, and that drives up their currency expectations further, with people wanting ever-more current data. Thus, for example, news has a shorter and shorter shelf life. Should, say, an abstracting service provide abstracts of news items months later, as sometimes is the case (*British Education Index* for instance), then its value as a publicist of news information is severely circumscribed. Because of this it is not wholly implausible to imagine a time when news will migrate totally from hard-copy newspapers to newspaper websites, possibly reducing newspapers to features magazines.

Place of publication/origin

The term 'Global Village', by now a rather worn-out cliché, nevertheless captures particularly well the essence of the contemporary information scene:

information production, communication and use certainly seem to have moved into a worldwide, borderless arena, unimpeded by technical barriers. Still, the place or country from which some information hails may not be invariably inconsequential to its potential consumers, and not because of any racist, discriminatory or derogatory attitude on their part. In ample proof of this point suffice to remember that unless information seekers can understand the language in which some material is written, it is of no use to them, especially since the automatic translation services available on various websites are still far from providing adequate conversions from one language to another. It is in recognition of this that search engines routinely offer the opportunity to restrict a search by language. The place or country of origin of some information can thus be very significant; just how significant, is contingent largely on three factors: (1) subject; (2) whether the user is a practitioner or academic; and (3) language proficiency.

Subject

The profiles of internationalisation differ between fields: the subject matter of some is truly international, whilst that of others is less so. In direct consequence, the importance accorded to the geographical origins of information is decidedly discipline-specific. The social sciences transcend national boundaries much less well than either the sciences or the humanities.

The sciences cross national frontiers readily by virtue of their universality ('an atom in New York is an atom in Moscow', says Brittain, 1984, 11), as well as their highly codified language. Therefore, academics studying say, cancer, or salmonella poisoning, are likely to be interested in research emanating from anywhere in the world. The humanities, too, 'travel well', concerned as they are with unique topics that have universal relevance, such as an event, a person or a work of art, literature, or music (Tibbo, 1994).

Not so, though, the social sciences, which, as Line (1973, 29–30) asserts, are characterised by an inherent instability, the result of their concern with human beings, and particularly their interactions with one another, whether this interaction be social, political, or economic. Thus, he maintains,

> … however carefully a particular study or experiment is carried out, and however valid the data that may come from it, a similar study of a different population – in a different town or country, or at a different time – will almost certainly give different results.

In addition, the lack of universally accepted methodologies and definitions, and the existence of uniquely national social institutions (like the UK's NHS, for instance) further contribute to the locality-specific nature of social science information. All this adds up to an overall tendency among seekers of scholarly information in the different fields of social science to be much more parochial in their information needs and information behaviour.

Not that this is always the case: some subjects in the social sciences – economics and psychology, for instance, do have broad international horizons. Still, in subjects like law and social welfare, communication is country-bound, although the European Union is increasingly drawing even law out of its traditional insularity. In the case of social welfare the concerns are even more local – at the regional rather than national level. Also, in the cases of some social science and humanities subjects – history, political science and geography come most immediately to mind – countries are frequently the subjects of study and, in result, place of origin is of special importance to them. Thus, for example, if you wished to study, say, the politics of Cyprus, then you would be well advised to examine the publications emanating from that island.

However, even in fields characterised by world-spanning scholarly communication, users will place a higher priority on the literature of some countries than others. Citation studies have long borne testimony to the fact that international scholarly communication is not a two-way or reciprocal process; rather, as Arunachalam and Singh (1992) point out, the actual distribution of scholarly and scientific research among different nations is rather skewed. Indeed, a small number of countries produce much of the mainstream research, whilst a very large number of countries contribute very little to the generation of knowledge. Naturally enough, the literature of the countries renowned because of the quality and size of their research, most notably the USA, the UK, Switzerland, Sweden, Canada, the Netherlands, Denmark, Australia, Finland and France, is universally held in high esteem, whilst the publications of the developing countries, because of the poverty of theirs, is generally ignored. Thus, for example, a professor of literature professes to a pronounced partiality to information originating in the Western world (and sounds rather apologetic about it, too):

> In my field, there's quite a lot of information coming from India, and somehow I have this somewhat derogatory attitude toward it; I keep expecting that it won't be all that significant. I don't know how justified I am for thinking so; I can't say I've read the piece and found it superficial, but my expectation is not the same it would have been were the author from the UK or the States or Holland. In fact, in India they have great English, we don't come near their level, and yet I have the feeling that whatever they write in India can't be serious. Part of it is the quality of the book: you see how cheap it is, both the paper and the print, and on the spot your expectations drop. It's not right, there's no justification to it, but that's how it is. You see something from Princeton or Oxford, and immediately think it is God knows what, although that's not always the case, not at all, but …

In defence, scholars point out that if someone from a Third World country had something really worthwhile to say, he/she would say it in a Western journal, on account of the recognition and prestige to be gained in so doing. True

enough: in scientific journals, no matter what the geographic origin of the publisher, the authorship is likely to be international. Thus, it is mostly through international conferences that academics (qualitatively) sample the non-Western literature, especially since, as a mathematician points out, 'if a journal is published in, say, a third-world country, there's a good chance that you've never heard of it and it will not be in the library either'.

True, as Russell (2001) observes, the shift to electronic scholarship could, at least in principle, bring about a change in this state of affairs, seeing that these days, researchers in developing countries can interact with their colleagues in any part of the world unhindered by geographical constraints. Also, Open Access is clearly hugely beneficial to people in poorer countries, where users cannot afford access to expensive journals (Nicholas et al., 2007e). However, even if information poverty is no longer the barrier that it used to be, academics from the periphery may still find it difficult to assume a more central position in international scholarship, for it is still questionable to what extent today's information consumers are prepared to take note of possibly valuable work done and published outside the Western world.

The problem does not seem to be a dearth of aptitude, knowledge or skills on the level of the individual researcher. After all, international research collaboration is especially prevalent between scholars hailing from small or developing countries and their colleagues in the wider academic world (Thorsteinsdottir, 2000). The phenomenon of 'brain drain', the significantly increased mobility of academics in the global knowledge society from low to high knowledge intensity places also speaks against such a possibility (Meyer et al., 2001). Yet, as Arunachalam (1999) contends, developing country scientists are not easily accepted into mainstream science: they can rarely get their research published in well-known journals and, even when they do, their work may not be quoted in subsequent work as often as papers published in the same journal by scientists from the advanced countries. Thus, the genuine inequities in opportunities, from less-developed infrastructures for electronically mediated research, through a shortage of research funds, to inadequate statutory and organisational environments both at the government and institutional level probably do not tell the whole story. Rather, as Russell (2001) suggests, there are also social and cultural barriers that prevent academics from the developing world taking their rightful place in the international community of scholars, due to a measure of subjectivity in scientific evaluations. Thus, until neither the objective conditions for the conduct of the scientific work, nor the subjective perceptions of its quality measure up to Western standards, the scholarship of the developing countries is bound to be approached warily.

Having seen the impact of the country of origin on information production and consumption, we need to probe deeper than that in order to see a more finely-grained picture. Thus, taking the place of publication aspect of information to its logical extreme, to the immediate work environment, it is important to note that it is in-house information, that is, formal and informal

information produced within the organisation, which is the more valued. Hardly surprisingly, of course: inside information is immediately relevant and directly touches upon the individual. Indeed, internal information flows are perceived as most critical and, in result, fellow workers seem to be the most frequently utilised information sources for work purposes (Baldwin and Rice, 1997; Huotari and Wilson, 2001).

Practitioner/academic divide

Academics, because they tend to be more interested in ideas, theories and comparative approaches, adopt a more international approach to information gathering than practitioners or, for that matter, amateurs; although, as it has already been noted, with scholarly information having become available at everybody's desktop, a much wider range of people have been drawn to its products. Nevertheless, information seeking across national boundaries is particularly typical of the scholarly enterprise (Nicholas et al., 2008b), and one which has been gaining considerable momentum throughout the 20th century and into the 21st. This, obviously, happened as part and parcel of general globalisation trends, although policy initiatives on national and supranational levels also contributed their share (Smeby and Trondal, 2005). In addition, the increasingly widespread availability of information and communication technologies not only enabled and underpinned the developments in this direction, but further accelerated them, too.

Practitioners and consumers of general-interest material also feed off a more international diet of information, though not necessarily intentionally or even knowingly. Thus, people may still obtain much of their foreign information second-hand from the national news services, but the ever-increasing popularity of the internet and the growth of satellite television have certainly brought the world to almost everybody's doorstep. Here again, the web has had a major impact. The impact is not a straightforward one, though. By making it much easier to get hold of information from any country on earth it is promoting the use of 'foreign' material. However, because the vast majority of information on the web is from the USA (although this is being challenged now by the economic growth of China), use is even more concentrated.

In any case, the world is indeed getting smaller (hence the term 'the Global Village'). Even Americans, long held to be quite self-sufficient in their use of information, take considerably more interest in the literatures of other countries. This is best illustrated by reference to two CIBER studies (Nicholas et al., 2008b): one, of the British Library's learning site for young scholars, the other, of Intute, a Joint Information Systems Committee (JISC)-sponsored scholarly gateway site. In both cases the UK audience was a minority one and US scholars were the majority user group, as was the case with *The Times* five years earlier (Nicholas et al., 2000). This might be explained by the perceived high quality of UK education, which, gratifying as it may be, plainly has big implications for decision-makers. After all, what will the tax payers say, in

this specific case in the UK, if they learn that government money is going to help the Americans become better and more informed searchers?

It seems, then, that on the national level the advent of the Global Village may not be wholly devoid of problems. Another instance of this is the ever-more heavy slant of today's globalisation-induced realities towards the so-called Anglo-Saxon culture. Suffice to cite, in ample proof of this, Google's initiative 'to organise the world's information and make it universally accessible', which, to all intents and purposes, is tantamount to a universal library of mostly Anglo-Saxon origins: a digital file of 15 million English-language books available on the web. No wonder the president of the French National Library has launched a counterattack, aimed at redressing the situation, which, he believes, will result in an unbalanced treatment of the literature of other countries (Jeanneney, 2007). Such a development, in its turn, will, as Gerald Grunberg, Senior Consultant to the Project Bibliotheca Alexandrina argues, run counter to the need for assembling and conserving the collective memory of a community or of a country, a need which has clearly become especially important now that globalisation has become an all-pervasive fact of life (cited in: Moss, 2008).

Language proficiency

Another problem associated with the geographic origins of information is what seems to be the most obvious of them all: the language barrier, which may thwart the attempt to read the literature of another country. Linguistic ability thus clearly enters into the question of whether information from foreign countries is consumed, although it seems that even when people can read literature written in another language, they are not highly motivated to do so. Mindful of this, the European Union publishes all its significant papers in all the languages of the Community. In any case, with the universal trends of globalisation and internationalisation, English has truly become the Esperanto of our times (only far more thriving than the original has ever been). Indeed, the dominance of the English language in contemporary international communication is indubitable, as Jorna (2002, 158) asserts:

> English is the geographically most widespread language of the world … [It] is the official language of relatively affluent and influential countries in North America, the British Isles and Australia, and has special status as a second language in over 70 countries … Across the world there are about 350 million native speakers and 250 to 350 million people who speak English as a second language … [although] if the most basic level of English is included, one might count up to 1.5 billion English speakers … Also, three of the most important international organisations communicate primarily in English: the United Nations, the World Bank, and the European Community.

If rallying behind English as a *de facto* lingua franca of international discourse is the way to overcome language barriers, then the web has certainly been

instrumental in helping things along. Firstly, it is encouraging people of all languages to disseminate information in English to obtain the largest audience. Indeed, well over 80% of websites are in English. Secondly, as has already been noted, a number of search engines provide a translating facility.

In point of fact, the use of English as the one, commonly agreed-upon language of dialogue between people from all over the world looks as if it is fast becoming quite the norm. Academe is a prime example of a milieu where this has actually happened. Indeed, by now proficiency in English is absolutely vital for academics; so much so, that the above-quoted Israeli-based computer scientist considers not having English a disability and his mathematician colleague actually likens it to not being able to breathe! With good reason, too, as another colleague of theirs explains:

> The whole of the Western world these days is centred on the U.S. and U.K … When somebody publishes in his native tongue, in local journals … it is inevitably less of a contribution, just like in our case with Hebrew … What you really want is to present your work to the international scientific community, to measure up to the standards set by the scientific community, which today begins with publishing in English … If an author did not publish his work in English it's probably not because he hadn't wanted to, but because his work was rejected.

Yet, interestingly, at least in the 'softer' knowledge domains, academics consider not having other languages detrimental to their work; perhaps not always seriously damaging, but certainly disadvantageous (Herman, 2005). First of all, as we have already seen, at least the social sciences are more insular in their scholarly activities, in result of which considerable quantities of information are published domestically and in the local language, too. Indeed, in an interview with an academic researcher, whose multidisciplinary interests frequently send him in pursuit of information in various fields of the social sciences, he greatly lamented the problem incurred by his linguistic limitations:

> I only know two languages, Hebrew and English, and it's an obstacle in my research, a real obstacle … I'm very interested in Germany … their constitutional organisation, both from the legal and the political point of view is probably the best achievable … [However,] I don't read German, and they hardly ever translate themselves … two, three, four, five publications in my field have been translated, and they are the most important ones, but you can't understand a legal system without knowing the language … To me my ignorance is a great hindrance, I feel like an illiterate person, an ignoramus … it's a real obstacle, because knowing English doesn't always suffice.

Further to that, researchers' language preferences are not only information availability-dictated, but need-driven. Indeed, as is more often than not the

case where scholarly information behaviour is concerned, needs inherent to the nature of the research endeavour shape the requirements for information, which, in result, tend to be very much discipline-specific. In fact, here again the needs of researchers in the various disciplines along the 'hard' to 'soft' continuum seem to be associated with the previously noted differences in the utilisation of information for generating new knowledge. Thus, at the 'soft' end of the continuum, where the primary evidence used is the product of a specific place and time, shaped by the distinctive personality of its creator (Wiberley and Jones, 1994), and where the new contribution to the corpus of knowledge actually 'happens' in the research article, inheres in the way the scholar analyses, extracts and develops insights about the material (Bates, 1996), reading the original-language publication can be absolutely crucial. For the social scientists, with their aforementioned parochial information needs, it must be even more so than for their humanities colleagues.

It is very different indeed at the 'hard' end of the continuum, where, as has already been noted, the new discovery is reported in the research article, not contained in it (Stoan, 1984, 1991). As researchers in the 'harder' knowledge domains only need to learn the results of the progress made by their peers, without attempting to get inside their thought processes, they can more easily afford to save the time, effort and money costs of mastering any other language, bar English. Anyhow, precisely because they only need to know 'the bottom line' of new contributions to the corpus of knowledge in their fields, they can put translations to very good use indeed. Thus, their need for information exchange across linguistic barriers can be met these days via a host of convenient translation journals at their disposal. Another possible solution are the many English-language international state-of-the-art reviews in existence, which, as we have already seen, may not be considered adequate substitutes for reading the original publications, but can definitely help at a pinch.

Processing and packaging

These two aspects of an information need, concerned with the different ways and formats in which the same ideas and data can be represented and presented to potential users, are intertwined and overlapping to such an extent that they are best treated together. Indeed, the vast literature of recent years on the subject rarely separates the two. Vast literature it is, too, for the processing and packaging of information have been the focus of untold studies, articles and books ever since the first attempts to harness novel electronic technologies to information management wrought irrevocable changes to our information environments. Unavoidably so, of course: with information figuring higher and higher on our inventories of 'bare necessities', its processing and packaging could not but come to the fore; after all, the way a piece of information is geared up for consumption is a crucial factor in its accessibility and usability, although with so much talk centring on the mechanisms of information use and management, it does look from time to time as if it is the tail

wagging the dog … In any case, let us take a closer look at this two-pronged characteristic of an information need.

Processing refers to the different ways that can be utilised to convey the very same information. Thus, for the same topic a researcher might want raw data with as little manipulation and interpretation as possible (unprocessed data), whereas a practitioner or somebody from outside the field might want the bare bones of data, with really only the significance of the information being spelt out (highly processed data). In fact, and this happens all the time, a single scientific discovery, social survey, government inquiry can be processed for a whole range of audiences and purposes. Take a piece of research undertaken on the effect of increased lighting on crime in a housing estate in East London. The work was originally published as a Home Office research report. As a research report it was typically densely and closely argued, full of data, descriptions of research methods and statistical appendices and, as a matter of fact, only accessible to other researchers. However, the topic itself was of interest to a much wider audience. Consequently, it was then condensed and fashioned for an article in a professional journal and, after that, it was picked up by the newspapers before, finally, being featured as a one-minute item on the local television news. At every stage in this chain detail was removed, interpretation featured more strongly and the information content was reduced. Thus, at each stage of processing the information was further compacted and simplified, resulting in a progressive reduction in the quantity, as well as a lowering of the intellectual level of the information. Processing does indeed often aim at achieving both these aims, although this need not be so, for condensation may not involve simplification and vice versa.

The reason processing, nevertheless, frequently does end up doing both is because it is quite typically all about popularisation. The newspapers, the radio, the TV and the internet are all purveyors of heavily processed information. Indeed, specialist correspondents spend a good deal of their time simplifying, popularising and explaining government reports, research studies and major surveys – and they are generally very good at this, making all kinds of difficult topics accessible to the uninitiated. By the same token, professionally written book reviews – together with the 'blurbs' – provide those concise summaries that, allegedly, are often the only part of a book that is ever read. Unfortunately, some of these popular reports can be so highly processed that they probably pass through the system without ever being absorbed by the brain. So, maybe the process can be taken too far?

There are, of course, other forms of presenting highly processed information, and not necessarily at a popular level, either. These include, most notably, abstracts, state-of-the-art reviews, executive summaries and interpretations. The best-known among them (and probably the one dearest to the information professionals' heart) is the abstract. Still, it has been long held that with the exception of academics, users do not much like the abstract, short and pithy as it is, possibly because, in many cases, abstracting results in too much loss of information. Even students did not use them, we were told (Keene,

2004). Thus, it was somewhat unexpected to find in a CIBER study of an e-journal database (Blackwell Synergy), reported in Nicholas et al. (2005), that undergraduates were the biggest users of abstracts (by a 5% margin over all other groups), and the population surveyed were all subscribers, so it cannot be simply put down to poorer access to full-text versions!

Perhaps more predictably, the abstract has remained popular among academics in the digital scholarly environment, too, as findings of the Virtual Scholar Research Programme clearly demonstrate (Nicholas et al., 2007d). Still, it is more of a surprise to see just how popular abstracts are, which is all the more interesting in today's information environment, rich as it is in full-text documents. The popularity of abstracts stems in part from technical reasons: search engines and gateways tend to point seekers of information to abstracts in the first instance and, obviously, viewing the abstract, typically free for all users, is the only option for the non-subscribers or the 'disenfranchised'. Also, as Pinto and Lancaster (1999) point out, abstracts may still be most advantageous for retrieval purposes, because the searching of full text will frequently cause an unacceptable level of irrelevancy.

However, there seems to be much more to it: abstracts are very important indeed in helping scholars deal with the information flood. First of all, abstracts allow for determining fairly quickly whether an article is of interest and of the appropriate quality and level, as one researcher explains: 'When I scan the literature, I do a very quick screen on the titles first. For the rare articles (from many journals) that make that cut, I then read through the abstracts to see if I wish to read more. I then pull down the full article only for a select few'. Further to that, present-day academics, working under unprecedented time pressures, may use the abstract as a substitute for the article itself: 'In some cases the abstract even provides me enough information and I don't need to read the full paper. Given the amount of papers published, good scientists only rarely can afford reading a full paper'. Apparently, at times the abstract is quite sufficient to fulfil an information need, as, for example, in the case of setting out to learn of new developments in one's field. True, this is very much field-dependent: as it has already been noted, at the 'hard' end of the continuum of disciplines, where the new discovery is reported in the research article, not contained in it (Stoan, 1984, 1991), it is probably fairly clear from the abstract how the results or conclusions of a paper fit with one's own research programme. However, at the 'soft' end, where the new contribution to the corpus of knowledge actually 'happens' in the research article, inheres in the way the scholar analyses, extracts and develops insights about the material (Bates, 1996), achieving the same end may often necessitate reading (or more likely skimming) the whole paper. Indeed, the conclusions of Nicholas et al. (2009) in their Research Information Network (RIN)-funded investigation of the impact of scholarly e-journals on the UK research community underscore this suggestion. The evaluation of the usage logs of the Oxford Journals in regard to use by 10 major UK research institutions and three representative subjects (life sciences, economics and history) indicated that

despite the fact that a good proportion of history journals had abstracts, very few historians viewed an abstract during their visit. Three times as many life science and economics sessions viewed an abstract.

In any case, using the abstracts in lieu of reading the article in its entirety fits in very well with the 'power browsing' form of information behaviour, which, as it has repeatedly been noted, is endemic in the digital environment. Today's information consumers Hoover through titles, contents pages and abstracts at a huge rate of knots to help them stride across the digital information universe. They feed for information horizontally rather than vertically, looking for 'bite-size' information chunks and, in result, seldom delve deeply into a website or even return to it. It is not difficult to see how abstracts are suited to this style of behaviour; in fact, they may even encourage it: after all, abstracts and contents pages are made for that, they are the motorways by which users drive through content.

Another way to process information in order to make it more palatable is in the form of a review article. Such subject-specific synopses of the recent major research advances made ideally add up to a coherent view of the 'state of the art' in a given knowledge area. When such a review article bears the signature of an influential scholar, which, in point of fact, it often does, for the authority figures in the various knowledge areas frequently function as gatekeepers, the information to be found therein is greatly enhanced by the expert interpretation of its compiler. Therefore, review articles understandably carry a lot of weight. Actually, the mere inclusion of a work in a review article serves as warranty of its excellence. Indeed, this highly processed form of scholarly information can serve as a great starting point for people who want to master the up-to-date basics of a topic.

Carrying the notion of the subject-specific synopsis of the knowledge attained on a topic to its very extreme, managerial information is often condensed into one page listing bullet points. With very little time to spend on any one problem from among the many awaiting their attention, senior management and busy professionals generally have support staff to provide them with the highly processed information they require: brief and focused pointers on possible solutions and actions, as well as warnings concerning the obstacles that may crop up.

Proceeding to the second component in the processing and packaging aspect of an information need, we now come to the external presentation or physical form of the information – the form in which it is stored and communicated. However, first, a word on the relationship between processing and packaging. It is, obviously, a very close one, because certain information packages are designed for the storage and dissemination of specific levels of processed data. Thus, dissertations and theses are packages that convey a good deal of data and detail, as do research reports and statistical series. Almost inevitably then, there is a limited audience for these information packages. Conversely, the internet, newspapers, television and leaflets – all purveyors of highly processed information – have vast and popular audiences. Of course, it is not always as simple as that, for theses have abstracts and

broadsheet newspapers have their heavy articles, and it is almost impossible to typecast books.

However, it is not the level of processing alone that attracts users to various forms of information package. There is a lot more to it than that, for, as we have already seen, some packages are more current than others (the web, news wires); some are far more exclusive (oral sources); some demand much less of the individual in digesting their messages (television); some are simply more 'in', and as such, have much greater appeal (mobile phones); others are very accessible (newspapers); and some are just plainly more familiar (books). The personality of the individual comes into it too – for instance, there are those who are more comfortable with oral sources, say, with having something explained to them, as opposed to fans of reading, who need to absorb information through their eyes.

In the same vein, the purpose for which information is sought can have quite some effect on the packaging required. That is, people quite consistently match the information task they are facing with the appropriate source. Factual information may be obtained easily enough via written sources, whether computer-mediated or hard copy, but when the need is for information that conveys complex ideas and thoughts, it is best attained via face-to-face interaction with human experts. The following account of an academic amply illustrates the point:

> If you write an article with a colleague, first you have to solve the problem you're working on … Now … this solution finding usually involves a face-to-face encounter, since you have to explain yourself, you have to use abstract arguments; it's not entirely trivial getting all that across to somebody, so it has got to be done verbally, it is truly essential to do it face-to-face.

Indeed, people unanimously appreciate the need for face-to-face communication for some purposes, as a historian contends:

> There's no substitute for the human touch, no substitute whatsoever … The electronic devices can help to decrease the need for human touch, but they are no substitute for it … After all, you wouldn't consider e-mailing your kids or even talking to them over the phone the equivalent of hugging them … You can't join forces with somebody you don't know, haven't met, haven't had coffee with.

No wonder, then, that the aforementioned survey of the use and impact of key digital health platforms and services in the UK (Nicholas et al., 2007a) found that where information was needed for solving medical problems, which, of course, often necessitates some give and take of possibly sensitive and/or vital information, the two most important sources were a person's own doctor and the practice nurse. Similarly, Garvey et al. (1974), pointing out the intra-individual variations in scholarly information use, came to the conclusion that journals

may be the most useful for providing information needed to place a scientist's work in proper context and to integrate his or her findings into current scientific knowledge, but informal channels, such as local colleagues were essential as sources of ideas, opinions and creative solutions to technical problems.

It seems then that a person's preference for a certain package is likely to be a result of an amalgamation of factors. Indeed, different user populations consistently want their information presented in specific packages, often to the exclusion of others. Thus, scientists have a love affair with journals and humanities researchers with monographs, students cannot get enough of the web, newspaper cuttings similarly smite journalists, and community workers revel in grey literature. It goes without saying, then, that information consumers are very much inclined to opt for the forms in which they regularly communicate, the forms to which they are accustomed. This is, of course, such a well-known phenomenon that people rarely stop to consider that it can lead to tunnel vision, as the media correspondent of the *Evening Standard* points out: 'We [journalists] do not know enough about the world. We rely too much on other newspapers [for information] … ' (Glover, 1994). Not that anybody seems to be overly bothered about the possibility; people happily remain tradition-bound where the packaging of their information is concerned.

Still, nothing is set in concrete, as the massive uptake of electronic packaging of information amply proves. To be sure, the foundation-laying years of the digital revolution passed to the tune of heartfelt laments regarding 'the sluggishness of human nature and its superstitious cleavage to old habits' (Harnad, 1999). Yet, with all that not so long ago IT-based sources and services were a huge novelty for everybody, now we all seem to have joined the vast ranks of consumers of digital information. Generally speaking, today's information users seem to be quite at peace with the novel technologies, although some are undoubtedly more enthusiastic about it all than others. A philosopher's earnest diatribe against electronic texts, which, he says, 'lack the vitality of the printed word just the same as the canned music accompanying your purchase of a pair of underpants lacks the vitality of a live concert', undoubtedly attests to the soft spot many people still have for the erudite tradition of the book and the library. His colleague, a professor of literature, who readily admits to being a 'dinosaur' of sorts where anything electronic is concerned ('I have finally mastered the art of searching the computerised catalogue, but it has taken me fifteen years', he says), blames the ephemeral qualities of IT-based sources for his lingering wariness of them:

> There's a transient feeling about it all; when you hold in your hand a piece of paper, it has presence … however, when you encounter the information on the computer screen, it is not only that your eyes and your brain are not accustomed to it, but it is also somehow of a temporary, insubstantial nature … When you find your information in a book, it's something tangible, standing on a shelf, but if it's on the internet, today it's here, tomorrow it's gone, so how can you trust it?

However, most people seem to be of a rather more pragmatic mindset: the often reserved, sometimes outright reluctant attitude towards electronic information, so frequently encountered in the past, has now mostly been replaced by the matter-of-fact approach usually reserved for the rudimentary conditions and routine practices of life and living.

Interestingly, though, whether a specific electronic information form or practice is adopted at all, and if it is, to what extent and how fast, may vary greatly among different populations. Originally, it was widely held that all people would eventually flock as one to all of the promising-looking solutions to the need for information. After all, the novel technologies did afford easy and quick access to more and better information! Indeed, or so the reasoning went, it was only a matter of time before all information seekers would 'see the light', perhaps simply a matter of waiting for children, who were born into the realities of a digital world, to grow up. The technologising force sweeping over society was bound to culminate in an ultimate, unreserved conversion to a wholly electronic way of life; it was an inescapable imperative, even, and information work was no exception. Arguing strongly against this technological determinism, subsequent thinking, taking into consideration the many idiosyncratic factors governing the needs of different populations, posited that as all technological changes are weighed against a normative order, new technologies are either not adopted by some groups of people, or modified to fit in with the existing social structure of these groups.

A prime example, amply proving the validity of this way of reasoning, is the integration of electronic media into academic work. Many of the studies into the impact of IT-based resources on scholarly work practices (see, for example, Bruce, 1998; Erens, 1996; Lazinger et al., 1997; Liebscher et al., 1997; Pullinger, 1999; Starkweather and Wallin, 1999) proceed from the notion that the move to electronic scholarship is indeed just a matter of time across all disciplines. Proponents of this view, as Kling and McKim (2000) explicate, typically conceptualise their vision in either one of two ways. Those who focus on the technical features of the various media maintain that all the novel electronic channels are essentially equally valuable in all disciplines; they all are said to reduce the costs of communication, expand the range of people and locations from which materials are accessible, and generally speed communications. As scholars in all scientific fields work with data, and communicate both formally and informally with other scholars, all of the electronic media should be adopted and used fairly uniformly. Others of the same mindset employ an evolutionary approach: since various fields, through somewhat random experimentation, have developed a series of electronic communication forums, soon we should expect scholars of all fields to adapt these successful discoveries to enhance their communications. Thus, it is simply a matter of time – perhaps simply a matter of waiting for today's internet-savvy students to become working scientists – before academics of various fields will catch up with those among their colleagues who are already on the leading edge of an inexorable trend. True, so the logic of such analyses goes, first some basic problems need

to be resolved, from lack of access, lack of awareness to the existence of electronic sources, lack of computer skills, lack of user friendliness of some IT-based systems, to the especially knotty issue of academics' conservative attitudes. Also, humanists, popularly assumed to be technophobes, might take longer before they, too, are persuaded that it is 'good for them'. Still, for those proceeding from this standpoint there seems to be little doubt as to the final outcome: all are bound to realise sooner or later that the advantages of electronic information work (ease, speed, convenience, etc.) are well-worth the effort of converting to IT-based practices.

However, other experts studying the phenomenon put forward an alternative scenario, maintaining that it is more likely that we will see field-specific or even sub-field-specific variations in the adoption of electronic research work practices (Covi, 1999; Covi, 2000; Fry, 2004; Kling and Covi, 1997; Kling and McKim, 2000; Mahe, 2003; Mahe et al., 2000; Talja and Maula, 2003; Walsh and Bayma, 1996). They suggest that it is the idiosyncratic nature of the scholarly undertaking in the different knowledge domains which determines the extent to which electronic resources are utilised and the rate of their adoption, remonstrating that the move to novel information work practices is not just a matter of time. Thus, the shaping of technology is highly specific to and emerges in reaction to the dynamic needs of particular communities. Therefore, as Kling and McKim (2000) maintain, field differences in the willingness to convert to electronic scholarship stem from the social practices that support trustworthy communications in each field. Take, for example, the much-debated uptake of e-print repositories. In some fields productive scholars are more aware of the work of their fellow researchers than in others. If the ongoing work is thus relatively transparent, the risks associated with sharing reports prior to their formal publications are fewer, and the willingness to base the scholarly communication on e-prints will be correspondingly higher.

Having thus looked in some detail at the specific point of electronic packaging of information, we now return to another of the more general aspects of the topic of our discussion here, focusing this time on the information professional's point of view. Plainly, information professionals, most notably librarians, are very well acquainted with this characteristic of need: after all, they do spend a good deal of their time organising and storing the physical embodiment of information – hard-copy and digital. So much so, that they are probably guilty of giving the processing and packaging aspect of an information need too much prominence. All too often it is a package and not the information that is given in answer to a question, an attitude that might explain why librarians tend to present to their clients the electronic version of some material, even when the hard-copy alternative is available and, if it were only given some consideration, more fitting to the enquirer's circumstances.

In addition, librarians seem to show marked bias towards some packages, in result of which the traditional library is full of books and journals. Newspapers, leaflets, CDs, unpublished information and personal contacts are generally neglected, even when they would appear to be more appropriate in

dealing with an enquiry. However, in the case of the academic librarian, the manifest disregard for some sources could be a case of responding to the wishes of their users. As has already been noted, scholars may hold their colleagues to be excellent sources for news about pertinent work underway and for detailed information about apparatus and procedures, but they certainly would rather not rely on them for obtaining the knowledge base accumulated on a subject (Herman, 2005). Thus, they have a preference for authoritative, published sources – sources that in turn can be cited as acceptable evidence (e.g. Einstein *wrote* ranks ahead of Einstein *said*).

Information professionals' many practical concerns with form tend to result in their building information systems for certain publication packages – and so fragment and complicate the search for information. Thus the online public access catalogues (OPACs) in most of our college libraries provide access to content to be found in books only (something of which students are often totally unaware). The periodicals – often a far more suitable and certainly a much more expensive resource – are largely left to a scattered and mismatched set of abstracting/indexing services. No wonder 'Google Scholar' enjoys the degree of popularity it does – it offers a rather more convenient, one-stop option for the retrieval of scholarly literature: 'from one place, you can search across many disciplines and sources: peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, preprint repositories, universities and other scholarly organizations'. This is, of course, much more in line with our expectations these days: as Featherstone and Venn (2006) point out, thanks to the internet we can now think beyond the desktop covered with piles of opened books, journals and photocopies. Our resource base is now the screen with its own virtual desktop, on which we are offered new layout and graphics for text along with images, pop-ups and video-clips, as well as mobility between sites and, perhaps most importantly, the possibility of to-ing and fro-ing between the two modalities. Undoubtedly, there is a lesson here for information professionals, which should send them scurrying to rethink the still prevalent ghettoisation of information packages. The recent efforts made towards the implementation of federated searching at university libraries certainly represent a step in the right direction.

This is all the more important for, as Williams and Rowlands (CIBER, 2008, I, 22–3) find in their examination of the literature on the myths surrounding the digital information behaviour of the 'Google Generation', the prevalent belief that today's young people are format agnostic and have little interest in the containers that provide the context and wrapping for information, is yet to be proven and grounded in reality. Rather, they contend, future developments may turn out to be quite different: presentation will still be important, for it is expected in an online environment, in which the technology offers potentially more diverse and interesting styles of content presentation. Thus,

> … the idea, often referred to in the online news context as shovelware, of migrating offline content online in a form as close as possible to its

original offline design will not generally work. The online world promises more and must deliver to be successful.

   With this we conclude our review of the 11 characteristics of information need and the holistic approach to information needs analysis it represents. We hope that the implications from all this are crystal clear by now: only if most, preferably all relevant aspects of a need situation are considered, be it on the collective or on the individual level, can the call for effective information provision be appropriately met. However, for attaining a full understanding of an information need, it is also important that we look at the host of factors that may come into play when people set out to look for data in response to a problem perceived as calling for additional information.