University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

E-JASL 1999-2009 (volumes 1-10)

E-JASL: The Electronic Journal of Academic and Special Librarianship

Summer 2002

Challenges, Strategies, and Tools for Research Scientists: Using Web-Based Information Resources

Danielle Bodrero Hoggan
San Jose State University

ELECTRONIC JOURNAL OF ACADEMIC AND SPECIAL LIBRARIANSHIP

v.3 no.3 (Spring 2002)

Challenges, Strategies, and Tools for Research Scientists: Using Web-Based Information Resources

Danielle Bodrero Hoggan
School of Library and Information Science, San Jose State University
danielle@scripps.edu

Thanks to Paula King, Michaeleen Trimarchi, and Rachelle Georger for their valuable suggestions.

Abstract

Scientists face many challenges in harnessing web-based resources. Information overload, misinformation, fees, poorly designed navigation, and loss of browsability all hamper the scientist searcher. In addition, many scientists rely on only one or two databases and often miss unique information that is available through other sources. Librarians can team up with scientists to develop strategies to overcome the challenges of web-based information. For example, a librarian can teach effective information seeking techniques, including how to use controlled vocabularies, how to evaluate information on the web, and how to complement web-based resources with print resources. Librarians can also help scientists to identify multiple electronic resources to more thoroughly cover a given topic. In addition, librarians can provide many tools to help scientists better utilize web-based resources. Subject directories can help scientists to cut through the information overload on the web and more quickly find relevant and high-quality information. Database selection tools can help scientists to find relevant databases for their research. Navigation of e-journal web sites can be eased or eliminated by linking software, and preprints can be more easily located using a preprint server. Finally, cost issues can be helped somewhat by making the most of the free web-based resources that are available from the U.S. government. Web-based resources present possibilities as well as problems. Librarians can provide the strategies and tools to help scientists make the most of web-based information.

Introduction

Scientists consume massive amounts of information from many sources: journal articles, books, theses, databases, conferences, and collaborators, to name a few. Journal articles are one of the most prominent sources of information for scientists. For example, physicists read an average of 204 articles per year, while chemists read approximately 276 articles each year (Tenopir & King, 2001). In addition, scientific researchers and medical workers read three times as many journal articles as legal, management, and sales professionals (Belefant-Miller & King, 2001). In order to find these articles, scientists browse through journals or search bibliographic databases or indexes. PubMed (the National Library of Medicine's free web-based interface to the Medline database of biomedical literature) serves up answers to 500,000 to 1,000,000 searches every day (Tenopir & King).

With such a large portion of their time spent finding and reading journal articles and other information resources, scientists are turning to the World Wide Web as a means to more efficiently locate the information they are seeking. A web browser can be used to access facsimiles of journal articles and theses, to search databases on many subjects, and to obtain data from other scientists. A survey of scientists at the Oak Ridge National Laboratory (ORNL) found that 35% of their readings came from electronic (web-based) journals (Tenopir & King, 2001). Web-based resources are widely viewed as timesaving because individual copies of articles can be produced at the scientist's desktop, eliminating the need for a trip to the library. Yet a study at the ORNL found that scientists take twice as much time to identify and locate electronic journal articles as they do for print articles (Tenopir, et al., 2001). This suggests that scientists are not using web-based resources as efficiently as possible. This paper will address the challenges that scientists face in locating and using web-based information. Problems in scientists' current practices will also be addressed, along with strategies and tools that librarians can provide to help scientists to make better use of web-based resources.

In this paper, web-based information will be defined as information that is available over the Internet via a browser using the Hypertext Transfer Protocol (HTTP). This includes information that is available for free as well as fee-based information. Many of the scholarly journals that scientists access over the web are fee-based resources. The fees for web-based journals are usually paid by the library at the scientists' institution (Tenopir & King, 2001). Libraries also subscribe to scholarly databases for their scientist patrons to access on the web. Journals that are accessed via the web will be referred to in this paper as *electronic journals* or *e-journals*. These terms include journals that are only available online as well as electronic journals that are based on a print source.

Challenges to Harnessing Web-Based Information

While accessing web-based information, research scientists face major problems including an overabundance of information, poor web site design, misinformation, and monetary concerns. These and other challenges are discussed below.

Information Overload

The volume of available information is more than any one information retrieval system can index, more than any library can purchase, and more than any scientist can read (Belefant-Miller & King, 2001). The number of journal publications per scientist per year has more than doubled since 1977 (Belefant-Miller & King). In addition to e-journal articles, much scholarly information is available on other web sites and web-based databases and databanks. The lack of central organization and indexing on the web makes the information overload even more confusing because the precision of search engine results is often low (i.e. many irrelevant documents are retrieved). At the same time, no search engine can index the entire contents of the web, so searches also have poor recall with respect to all relevant information on the web.

Web Site Navigation

To retrieve electronic journal articles, users must navigate a vast number of web sites, and each publisher structures their web site differently than the next. This can be very challenging because the web sites are not always designed for easy navigation. Sometimes it is very difficult to determine what link on the crowded homepage leads to article content. Often the links to "subscribe today" are much more prominent.

Preprints

Scientific research is very competitive, and researchers must keep up with the latest developments in their field. Consequently, most of the articles read by scientists are less than one year old (Belefant-Miller & King, 2001). In fact, researchers often want to know the results of a study before they are published. Preprints, copies of articles that have not yet been published, are sometimes available on the web. However, finding preprints on the web can be very difficult because they are poorly organized and because most of them are in portable document format (PDF) or PostScript format and are therefore not indexed by most search engines (He & Hui, 2001).

Financial Concerns

Peer-reviewed articles are available on the web, but they are usually published in fee-based e-journals. Scientists access these resources via personal online subscriptions or, more commonly, institutional online subscriptions paid for by their libraries (Tenopir & King, 2001). For researchers not associated with an institution, such as retirees or independent consultants, these fees are barriers to information access (Line, 2001). Even researchers with access to a small private library may be discouraged from requesting journal articles because of the cost involved. Some information resources are available at no charge, but they tend to be unreliable. For example, PubSCIENCE, a free web-based database for physical scientists maintained by the U.S. Department of Energy, may be discontinued because the U.S. Congress feels that it is an "undesirable duplication of private-sector activities" (Peek, 2001, para.2). Some private-sector online services offer free content, but they are much more likely to go out of business than the fee-based services (Tenopir, 2001). Thus, financial concerns can limit scientists' access to scholarly information.

Loss of Browsability

A 1993 study at the University of Tennessee, Knoxville (UTK) found that science faculty and staff found most of the articles they read through browsing (Belefant-Miller & King, 2001). In fact, browsing accounted for twice as many documents (53%) as database searching (27%, Belefant-Miller & King, p. 109). Other studies in 1962 and 1977 also found that browsing accounted for most of the readings of scientists (Martin, 1962, as cited in Belefant-Miller & King, 2001; King, McDonald, & Roderer, 1981, as cited in Belefant-Miller & King, 2001). Professor Barry Trost of Stanford's chemistry department feels that browsing print materials is an important source of innovative and creative ideas (Schevitz, 2002).

The challenge to scientists is the loss of browsability of today's electronic journals. At the time of the 1993 study, electronic journals were not as popular as they are today, and most of the UTK science faculty and staff preferred print articles to electronic ones, largely because of the poor printing quality of figures in the electronic versions (Belefant-Miller & King, 2001). However, scientists' interest in electronic journals has increased dramatically in recent years, largely because of their convenience and the improved quality of figures that are now available in PDF format. The emphasis that today's scientists place on electronic journals may come with a corresponding loss of serendipitous readings (Schevitz, 2002). Electronic journals are not as browsable as print journals, and this may have a marked impact on the range of articles read by scientists, who have traditionally depended on browsing to find most of the articles they read.

Reliability

Print journals do not experience technical difficulties. However, when a publisher's server goes down, access to electronic content is temporarily lost. Compounding the problem is the fact that access to many electronic resources is rented rather than owned, so if the institution cancels a subscription, access to backfiles is lost. With a print journal, on the other hand, the previous volumes of the journal would still be available after a subscription was cancelled. Because of these issues, librarians and scientists cannot rely completely on electronic resources.

Scholarly Misconduct and Misinformation

Scholarly misconduct includes plagiarism, fabrication of results, and manipulation of data (Calvert, 2001). Researchers have been known to manipulate numbers and even publish lies (Calvert; Woolston, 2002). In addition, many publications include incorrect bibliographic citations and other errors. It is already difficult for editors of print journals to detect scholarly misconduct and other errors, and the chance for misinformation to slip past an e-journal editor is even more likely given the faster turnaround time of online-only journals (Calvert). Online journals may actually attract dishonest scholars who want to pad their resumes with quick, fabricated publications (Calvert). In addition, scholars can publish results and papers on their personal web sites without peer review. Online discussion groups often contain biased information or information taken out of context (Calvert). Thus, the potential for misinformation on the web is a real concern for research scientists.

Problems in Scientists' Current Practices of Using Web-Based Information

Before the advent of CD-ROMs and the web, when online searching was charged by the minute, doctors and scientists were most likely to ask a librarian to perform their literature searches for them. Today, however, research scientists are using web-based resources to perform most of their own literature searches. In fact, "most patrons prefer performing their own searches—even if the librarians can do a better job" (Tenopir, 1992, p. 98). Scientists bring important subject knowledge to a literature search, but they often lack sophisticated searching skills. Problems in scientists' current searching practices are discussed below, followed by tactics and tools that scientists can use to become better searchers.

Overlooking Alternative Web-Based Resources

Librarians at Texas A&M have found that many knowledgeable scientists limit themselves to one or two databases and ignore the rest (Quigley, Church, & Peterson, 2001; Schaffer, 2001). Often scientists are unaware of other databases that may be relevant to their research, and they need database instruction at the most basic levels (Hall, 1999; Quigley, Church, & Peterson; Schaffer). For example, Michaeleen Trimarchi (personal communication, March 16, 2002) related an experience where a science professor was surprised to discover that the Science Citation Index database allowed for cited reference searching in addition to general subject and author searches. Many subject-specific databases are complicated to use, and even common database interfaces like PubMed have many features that most users never utilize. Scientists need to take the time to learn to use all of the resources available to them that may be relevant to their research, either by attending a class, talking with a librarian, or completing an online tutorial.

Overlooking Print Resources

E-journals are a convenient way for scientists to access scholarly journal articles without leaving their offices. Unfortunately, these web-based resources are more expensive than their print counterparts, so libraries are limited in the number of electronic subscriptions that will fit in their budgets (Schevitz, 2002). In addition, only six to eight percent of the scholarly journals in existence are available online (Schevitz), so even a very rich library could not provide web-based access to all scholarly journals. This means that a thorough search of the literature requires scientists to make a trip to the physical library or to place document delivery requests that may take days or weeks to fill. However, in an era of instant gratification, researchers are tempted to read only the information that is available online because it is the easiest to obtain. Print-only journals, books, and older volumes of journals that are not archived electronically may be ignored.

The ultimate (although tragic) example of ignoring alternative web-based resources and print resources is the summer 2001 clinical trial at Johns Hopkins University that resulted in the death of a 24-year-old woman after she inhaled hexamethonium (Schevitz, 2002). Prior to the clinical trial, Dr. Alkis Togias searched Medline and the web for literature about hexamethonium and reported that it had been used in four studies without adverse effects (Ogilvie, 2001; Perkins, 2001). However, after the tragedy, a more thorough search of resources revealed 11 articles published between 1953 and 1962 reporting the toxicity of the substance (Ogilvie). It is dangerously easy for scientists to assume that all medical literature is in the Medline database (McLellan, 2001). Many medical students and residents do not know that Medline searches only cover literature from 1966 to the present (Huffman, 2001). The hexamethonium articles mentioned above could have been located using the web-based OLDMEDLINE database, which covers literature from 1958 to 1965 (McLellan), and the printed *Index Medicus*, which covers medical literature from 1879 to the present (see footnote j of Table 6). In addition, searches of Toxline (http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE; National Library of Medicine [NLM], 2002) and Micromedex's PoisonIndex Toxicologic Management (a fee-based online database) turned up information on the toxicity of hexamethonium (Perkins, 2001). Thus, if Dr. Togias had included these web-based resources or print resources in his search, he would have been more aware of the dangers of this drug. When researchers need a thorough literature search, they should collaborate with a librarian who can recommend appropriate databases and perform searches when needed.

Not Evaluating Resources

Misinformation on the web was discussed above as challenge for research scientists. A problem in current practice that relates to this issue is the lack of critical thinking that users apply to web-based information. Readers give more cognitive authority to (i.e. they put more trust in) information that they receive in an electronic form (Calvert, 2001). This is a scary fact given that electronic resources are likely to attract higher levels of scholarly misconduct than print publications (Calvert). When surfing the web, users are exposed to large amounts of information at high speeds, and they rarely take the time to evaluate what they are reading (Calvert). Even university faculty can be duped. Calvert conducted a focus group discussion in which he showed library faculty of the Nanyang Technological University in Singapore a fictitious journal article in a fictitious electronic journal. Some of the faculty admitted to being fooled by the fictitious journal and article. In a day when anyone can put anything on the web, the need to remain vigilant is essential.

Ignoring Copyright Terms

Scientists in academia share their discoveries and knowledge with colleagues by publishing journal articles for which they receive no royalties. In fact, the author often must pay the publisher to offset the cost of printing the article. Because the authors do not receive compensation for the intellectual property presented in the article, scientists often do not understand why journal subscriptions cost so much. They are particularly baffled by the prices of online subscriptions, which seem to avoid even the cost of paper and ink. Some scientists think that electronic articles cost the publisher nothing. Therefore, researchers feel justified in sharing passwords to electronic journal subscriptions with colleagues, even though this is a violation of the terms of the subscription and of copyright law.

In reality, one journal article costs a publisher $4000 for recruiting authors, refereeing, editing, preparing illustrations, marketing, and maintaining subscriptions (Belefant-Miller & King, 2001). Electronic articles do not require the additional cost of paper, ink, or shipping, but they do require hardware, software, and staff. Thus, the cost of an electronic journal article is very similar to that of a print article (Belefant-Miller & King), and publishers are entitled to charge for their services. If researchers do not abide by the terms of subscription agreements and database licenses, they risk being sued and losing the privilege to use those resources in the future. Scientists should also be reminded that violations are much easier to detect in the electronic environment because the publisher’s web log records every transaction that occurs on the server. Librarians need to take an active role in educating scientists about the terms of access to web-based resources.

Tactics and Strategies for Dealing with Web-Based Information

As librarians we need to encourage scientists to consult with us about literature searches, particularly if a comprehensive search is needed (McLellan, 2001; Perkins, 2001). Librarians may not have the subject knowledge of the scientist, but we have expertise in searching many resources. "The best searches would involve researchers and their librarians working closely together" (Perkins, para. 21). Consultations with librarians will teach researchers about the different databases that are available and how to use them, thus making the scientist a better searcher. All of the strategies mentioned below can be taught at a bibliographic instruction class or during a consultation.

Heuristic Search Strategies

Scientists, like all searchers, need to learn basic searching strategies that will be applicable to a variety of databases. A basic searching class can teach techniques such as truncation, Boolean logic, and use of controlled vocabularies. Libraries should offer hands-on searching courses targeted to scientists. Written instructions and follow-up assistance after class are often needed to help searchers become confident in using the complex online searching tools that are available (Nahl, 1999).

Controlled Vocabularies

Scientists can improve their searches by learning to use controlled vocabularies or thesauri when they are available. For example, Medical Subject Headings (MeSH) used in Medline can be accessed online from the National Library of Medicine's MeSH Browser (http://www.nlm.nih.gov/mesh/MBrowser.html; NLM, 2001a). Some interfaces to Medline, such as PubMed and OVID, provide automatic mapping of search terms to MeSH headings. The MeSH Browser is helpful in searches of older literature. For example, the MeSH Browser entry for a term includes the date the term was added to the thesaurus, previous indexing terms, and tree structures. It is sometimes necessary to choose a broader term from the tree structure when searching older literature because the more specific terms may not have been indexed as such (Perkins, 2001). Other controlled vocabularies that may be useful to scientists are the Library of Congress (LC) subject headings, National Agricultural Library subject headings, and the Mathematics Subject Classification (MSC) scheme. MSC can be browsed, searched, or downloaded from the Zentralblatt MATH MSC site (http://www.zblmath.fiz-karlsruhe.de/MATH/msc/index; European Mathematical Society, FIZ Karlsruhe, & Heidelberg Academy of Sciences, 2001). LC subject headings are not available online, but can be browsed in print (Library of Congress, 2001). Often, however, the controlled vocabulary for a database is not available. In these situations, pearl-growing searching techniques are especially useful. Librarians should discuss controlled vocabularies and searching techniques with scientists.

Web Site Evaluation

The overall quality of information on the web is not likely to increase with time. In fact, an evaluation of consumer health web sites showed a decline in reliability from 53% to 47% between 1996 and 1997 (Calvert, 2001, p. 233). The nature of the web makes any sort of quality control infeasible (Calvert). Thus, scientists need to learn to evaluate web-based resources and to detect misinformation (Calvert). If a library does not offer an information evaluation class, librarians can refer scientists to one of the following online information literacy sites, each of which has a specific section on evaluation of web-based resources: Comprehensive Online Research Education (http://core.lib.purdue.edu/; Purdue University, n.d.), Texas Information Literacy Tutorial (http://tilt.lib.utsystem.edu; University of Texas System Digital Library, 2002), or Research Instruction Online (http://dizzy.library.arizona.edu/rio; University of Arizona Library, 2000).

Alternative Web-Based Resources

A thorough literature search can save a scientist weeks in the laboratory. Librarians and scientists should discuss what databases are available that may be relevant to the research. A study at Texas A&M University found that many science faculty members are unaware of electronic article databases appropriate to their discipline (Quigley, Church, & Peterson, 2001). For example, many biologists rely primarily on Medline for their literature searches. However, the Biological Abstracts database contains many biological and biomedical citations not found in Medline; there is only 33% overlap in the journal coverage of the two databases (Miller, 2002). Thus, biologists would do well to search both. Similarly, scientists searching the web should be encouraged to use multiple search engines because different search engines show little overlap in their results (Notess, 2000).

Older Resources

The Johns Hopkins clinical trial discussed above drives home the need to thoroughly search the literature, including the older resources. Many of today's popular databases do not cover literature before the 1970s. With a librarian's assistance, scientists can complement their usual web-based searches with databases and print indexes that cover earlier literature. For example, scientists can search the Web of Science®, the web-based version of Science Citation Index, for scholarly science journal articles from 1945 to the present (Institute for Scientific Information, n.d.). In most cases, the web-based indexes do not cover literature as far back as their print counterparts, so scientists should be prepared to include print indexes as part of their search strategy. Tools for finding older resources will be discussed below.

Print Resources

Librarians should encourage researchers to complement web-based resources with print resources. For example, scientists should search their library's catalog for books, which may provide a historical analysis that they will not find in a journal article (Schevitz, 2002). Also, I recommend that scientists browse the stacks and recent journal arrivals at the library. There is no electronic substitute for the browsability of a physical book or journal. Often researchers browsing through journals read articles that catch the eye, but that they would never have searched for online (Schevitz). This cross-fertilization process leads to an interdisciplinary approach to research that fosters innovation and creativity (Schevitz). In addition, a trip to the library facilitates social interaction with colleagues, an important source of information for scientists (Belefant-Miller & King, 2001). A visit to the library also enables researchers to use print indexes that may cover time periods outside the scope of the usual electronic databases. Thus, scientists have much to gain from a trip to the library. Librarians should promote the physical library as an important part of a research strategy.

If scientists cannot make frequent trips to the library to browse, they should consider browsing the electronic tables of contents of the journals in their field, which can be accessed for free from most journal web sites (Poss, 2000, p. 181). In addition, many journals provide electronic table of contents services that send users email of the tables of contents as new issues are published. Many journals have a link to this free service from the journal's web site. In addition, librarians can inform scientists about institutional subscriptions to alerting services such as Ingenta that allow a scientist to receive emails of tables of contents for multiple journals and updates of citations with a specified author or keyword (Ingenta, 2002). This electronic form of browsing can at least partially substitute for print journal browsing.

License Terms for Fee-Based Resources

Scientists need to take the time to learn about copyright and the license terms that they agree to when they subscribe to an electronic resource. These terms often limit the sharing of information with others who have not paid for a subscription. The terms may also limit the number of records that can be downloaded in a given time period. Librarians need to inform scientists about these stipulations.

Tools for Dealing with Web-Based Information

This section describes tools that can help scientists to find quality information on the web quickly and easily. Librarians need to make scientists aware of these tools through marketing and bibliographic instruction, which can take many forms. Science faculty tend to prefer workshops to formal classes, and they are especially fond of instructional handouts and online tutorials that allow them to learn at their own pace (Quigley, Church, & Peterson, 2001). Science faculty also prefer targeted instruction, so a workshop on a particular database or subject area should be received well (Quigley, Church, & Peterson; Hall, 1999).

Subject Directories

When searching for very specific information on the web, it is often most effective to use a search engine. However, for information about a field of study, a science subject directory will lead to quality information more quickly (Beekink, 2000). Science libraries, universities, and other non-profit organizations compile directories of high-quality science web sites, often including annotations and ratings. (Beekink). Table 1 shows subject directories that would be useful to scientists on the web.

Database Selection Tools

DialIndex is a ranking tool for databases offered by Dialog. A searcher enters a term and DialIndex searches all or a subset of Dialog's databases and returns a report of the number of hits in each database. Thus, DialIndex is a valuable tool for discovering databases that are relevant to a particular search. Scientists can search DialIndex (http://www.dialogweb.com/cgi/dwframe?context=databases&href=topics/All/; Dialog Corporation, 2002) for free on the web, but their institutions may not subscribe to all of the databases in the results list. The University of California, San Diego (UCSD) libraries have overcome this problem by creating a tool called Database Advisor, which is like DialIndex, but it is customized to search the databases available at UCSD (Hightower, Reiswig, & Berteaux, 1998). Librarians can encourage their institutions to adopt the Database Advisor software, which UCSD will freely share. However, the software requires significant customization and maintenance, so librarians will need to assess their technical resources before implementing the software.

Similarly, a chemist seeking thermodynamic data for a particular compound will find the ThermoDex web site (http://thermodex.lib.utexas.edu/; Mallet Chemistry Library, University of Texas at Austin, 2001) to be a useful resource. The searcher chooses the compound of interest and the thermodynamic property desired (choices including enthalpy, entropy, surface tension, boiling point, and over 100 other properties). ThermoDex returns a list of print handbooks and web-based compilations of thermodynamic data that may contain the desired information.

Free Information Sources

The financial issues faced by scientists (and their librarians) have been discussed above. Of course, information is not free. However, some of it is paid for by tax dollars and is available to the public at no charge. In addition, journal publishers will often provide citations and abstracts of their content to database producers because this increases demand for the articles. The federal government compiles these citations into databases for biomedical and physical scientists. Some of the tools available to research scientists at no charge are described in Tables 2 through 5.

Table 2 describes some of the databases from the U.S. National Library of Medicine (NLM) for biologists and medical researchers. The NLM's (2001c) complete list of electronic information resources can be viewed at http://www.nlm.nih.gov/databases/databases.html and includes information on specific topics such as bioethics, cancer, AIDS, or malaria.

The National Center for Biotechnology Information (NCBI) at the NLM provides the Entrez databases (http://www.ncbi.nlm.nih.gov/Entrez/; NCBI, 2002), which provide biotechnology researchers with important structural and sequence data for biological molecules. Table 3 shows some of the databases provided by the NCBI. The NCBI (2002) also maintains databases about population studies, human genetics, and other topics that can be accessed from http://www.ncbi.nlm.nih.gov/Entrez/.

Biologists that are focused more on agricultural and environmental studies can use the databases shown in Table 4. These resources are sponsored by the U.S. National Agricultural Library, the Department of Agriculture, the National Oceanic and Atmospheric Administration (NOAA), and the Department of Energy (DOE). Table 5 shows resources for chemists, physicists, and mathematicians provided by the U.S. National Science Foundation, National Aeronautics and Space Administration (NASA), and the DOE.

Tables 2 through 5 show information resources available to the public at no charge. Librarians should also let scientists know about fee-based resources that are available to them via institutional subscriptions.

Tools for Finding Older Resources

Most of today's science databases do not cover literature before the 1970s or 1980s, and some, like Current Contents®, index only the latest 12 months. Finding citations to earlier literature can be a challenge. Table 6 summarizes databases and indexes that can be used to find citations to early scientific literature. The table includes many print resources because print indexes tend to cover literature much farther back than their web-based counterparts. A notable exception is the Beilstein Crossfire database, which indexes chemical structures, reactions, and publications back to 1771 (MDL Information Systems, 2002). Most of the web-based resources in Table 6 are fee-based, and therefore may not be accessible; print resources may be preferable in this situation. Two very notable print resources, *Catalogue of Scientific Papers (1800-*

1900) and *Catalogue of Scientific Papers, 1800-1900: Subject Index*, both compiled by the Royal Society of London (1965, 1968), provide author and subject access to biochemistry, chemistry, physics, and mathematics papers from the nineteenth century. These print indexes were continued by the *International Catalogue of Scientific Literature*, which indexed scientific literature from 1901 to 1914, when the cooperative project was halted by World War I (Parrott, 2001).

Tools for Finding Funding

Information about federal grants can be found at the National Institutes of Health (NIH) Grants and Funding Opportunities site (http://grants1.nih.gov/grants/index.cfm; NIH, n.d.). The CRISP database, which is accessible from the NIH Grants and Funding Opportunities homepage, lists funding opportunities from the NIH, the Centers for Disease Control and Prevention, the Food and Drug Administration, the Health Resources and Services Administration, and the Agency for Health Care Policy and Research (Reavie, 2000). Additional grants and funding opportunities can be found at the National Science Foundation's Funding site (http://www.nsf.gov/home/grants.htm; National Science Foundation, n.d.). In addition to these free resources, librarians can make scientists aware of institutional subscriptions to the Illinois Researcher Information Service (IRIS; http://www.library.uiuc.edu/iris/; University of Illinois Library, 2001) or the Community of Science (http://www.cos.com/; Community of Science, 2002), databases of government and private funding opportunities.

Linking Software

Open uniform resource locator (URL) technology is making it possible to bypass the navigation of a journal publisher's web site and go straight to the desired full text article. PubMed and PubSCIENCE have linking features that will take a searcher directly from a citation to the full text article on the publisher's web site if the searcher or the searcher's institution has a subscription to the journal (DOE, Office of Scientific and Technical Information, n.d.a; National Center for Biotechnology Information, n.d.). Commercial databases such as Dialog and EbscoHost offer similar linking features (Ebsco Information Services, n.d.; K. Quinn, Dialog Information Consultant, personal communication, March 22, 2002). Commercial linking software is also available, e.g. SFX from Ex Libris. SFX will provide linking from many commercial databases to full text articles (Ex Libris, 2002). SFX can also be customized for the library's holdings such that the searcher is directed to an interlibrary loan (ILL) request form in the event that the library does not have a subscription. SFX inserts the bibliographic information into the ILL form to minimize typing for the scientist. Librarians should inform scientists about linking software and encourage their libraries to adopt this new timesaving technology.

Preprint Servers

Use of preprint servers like arXiv (http://arxiv.org/; Cornell University, 2001), the PrePRINT Network (http://www.osti.gov/preprints/; DOE, Office of Scientific and Technical Information, n.d.b), and the Chemistry Preprint Server (http://www.chemweb.com/preprint; ChemWeb, 2002) will direct scientists more quickly to preprints than a general search of the WWW. If the desired article cannot be found using a preprint server, it is recommended that the scientist or librarian use the Google search engine to search the web. Google indexes PDF files and therefore will be more likely to index the preprint (He & Hui, 2001).

Conclusion

Scientists face many challenges in harnessing web-based information. The web contains a plethora of information, some of which is not accurate. High-quality information may be difficult to access because of poorly designed web sites or associated fees. In addition, each web-based resource seems to hold such a vast quantity of information that it is easy to assume that one has conducted a comprehensive search, when, in fact, other resources contain unique, relevant information. It is too easy to violate copyright on the web. Finally, users of web-based resources lose the browsability of their print counterparts.

Librarians can help scientists to overcome these challenges by becoming partners in literature searches and by teaching effective searching strategies. While consulting with a scientist or at a bibliographic instruction class, librarians can teach better information seeking techniques, including how to use controlled vocabularies and how to evaluate information on the web. Librarians can teach researchers to make a habit of reading and understanding the terms of license agreements. Librarians can also guide scientists to search multiple web-based resources to more thoroughly cover a given topic and to complement web-based resources with print resources, which contain much scholarly information that is not duplicated on the web. In addition, print resources tend to be more browsable than electronic ones, thus helping to provide a scientist with an interdisciplinary perspective that contributes to innovation.

Librarians can also make scientists aware of tools that are available to assist the use of web-based resources. For example, subject directories can help scientists to cut through the information overload on the web and more quickly find relevant and high-quality information. Database selection tools can help scientists to find relevant databases for their research. Navigation of e-journal web sites can be eased or eliminated by linking software. Preprints can be more easily located using a preprint server. Finally, cost issues can be helped somewhat by making the most of the free web-based resources that are available from the U.S. government.

Web-based resources present possibilities as well as problems. However, with the right strategies and tools, scientists and librarians can make the most of web-based information.

References

Beekink, M. (2000). Training end-users using scientific internet-subject directories. *EContent*, *23*, 57-60. Retrieved February 8, 2002, from Library Literature and Information Science Full Text database.

Belefant-Miller, H., & King, D. W. (2001). How, what, and why science faculty read. *Science and Technology Libraries, 19,* 91-112.

Biosis. (n.d.). *Biosis timeline.* Retrieved March 16, 2002, from http://www.biosis.org/75/timeline.html

Calvert, P. J. (2001). Scholarly misconduct and misinformation on the World Wide Web [Electronic version]. *The Electronic Library, 19,* 232-240.

Chemical Abstracts Service. (n.d.). *Printed chemical abstracts.* Retrieved March 16, 2002, from http://www.cas.org/PRINTED/printca.html

Chemical Abstracts Service. (2002, February). *Explore general chemical information with SciFinder Scholar.* Retrieved March 16, 2002, from http://www.cas.org/SCIFINDER/SCHOLAR/scholchemistry.pdf

ChemWeb. (2002). *The chemistry preprint server.* Retrieved April 6, 2002, from http://www.chemweb.com/preprint

Cinahl Information Systems. (n.d.) *Cinahl history.* Retrieved March 16, 2002, from http://www.cinahl.com/about/abt-history.htm

Community of Science. (2002). *Community of science: Research funding, expertise, abstract publishing for academia and industry.* Retrieved March 30, 2002, from http://www.cos.com/

Cornell University. (2001, December 13). *ArXiv.org e-print archive.* Retrieved April 6, 2002, from http://arxiv.org/

Department of Energy, Office of Scientific and Technical Information. (n.d.a). *About PubSCIENCE.* Retrieved March 13, 2002, from http://pubsci.osti.gov/aboutfrm.html

Department of Energy, Office of Scientific and Technical Information. (n.d.b). *PrePRINT network: Energy, science and technology preprints.* Retrieved April 6, 2002, from http://www.osti.gov/preprints/

Dialog Corporation. (2002). *Dialog® web.* Retrieved March 28, 2002, from http://www.dialogweb.com/cgi/dwframe?context=databases&href=topics/All/

Ebsco Information Services. (n.d.). *Ebsco expands linking capabilities.* Retrieved March 23, 2002, from http://www.ebsco.com/ess/ays/56linking.stm

European Mathematical Society, FIZ Karlsruhe, & Heidelberg Academy of Sciences. (2001, September 13). *MSC2000*. Retrieved March 28, 2002, from

http://www.zblmath.fiz-karlsruhe.de/MATH/msc/index

Ex Libris. (2002, February 25). *Welcome to SFX*. Retrieved March 13, 2002, from http://www.sfxit.com

Hall, L. (1999). A home-grown program for raising faculty information competence. *Computers in Libraries, 19*(8), 28-30, 32, 34.

He, Y., & Hui, S. C. (2001). PubSearch: A web citation-based retrieval system. *Library Hi Tech, 19,* 274-285. Retrieved February 9, 2002, from Emerald Fulltext database.

Hightower, C., Reiswig, J., & Berteaux, S. S. (1998). Introducing Database Advisor: A new service that will make your research easier. *College and Research Libraries News, 59,* 409-412. Retrieved February 8, 2002, from Library Literature and Information Science Full Text database.

Huffman, D. G. (2001). When is a literature search done? [Electronic version]. *Lancet, 358*, 2088.

Ingenta. (2002). *Ingenta: About ingenta*. Retrieved March 16, 2002, from http://www.ingenta.com/isis/general/Jsp/ingenta?target=/about_ingenta/uncover_users.jsp&WebLogicSession=PJN9u2sIQGsaihybLnwy|-5659241230871430413/-1052814329/6/7051/7051/7052/7052/7051/-1

Institute for Scientific Information. (n.d.). *Science citation index*. Retrieved March 16, 2002, from http://www.isinet.com/isi/products/citation/sci/

Johnson, W., Golden, M., & Norrisey, S. (2000). Federal science and technology web sites: An annotated bibliography. *Science and Technology Libraries, 19*(1), 103-121.

Library of Congress. (2001). *Library of Congress subject headings* (24th ed., Vols. 1-5). Washington, DC: Library of Congress.

Line, M. B. (2001). Opinion paper: Access to documents by the independent researcher [Electronic version]. *Interlending and Document Supply, 29,* 175-176.

Mallet Chemistry Library, University of Texas at Austin. (2001). *ThermoDex search page.* Retrieved April 6, 2002, from http://thermodex.lib.utexas.edu/

McLellan, F. (2001). 1966 and all that—When is a literature search done? [Electronic version]. *Lancet, 358,* 646.

MDL Information Systems. (2002). *CrossFire Beilstein information*. Retrieved March 20, 2002, from http://www.beilstein.com/products/xfire

Miller, J. M. (2002, January 3). *Medical College of Ohio Library: Materials for Biological Abstracts*. Retrieved April 15, 2002, from http://www.mco.edu/lib/education/ba.html

Nahl, D. (1999). Creating user-centered instructions for novice end-users [Electronic version]. *Reference Services Review, 27,* 280-286.

National Agricultural Library. (2001, February 14). *About AGRICOLA.* Retrieved March 16, 2002, from http://www.nal.usda.gov/general_info/agricola/agricola.html

National Center for Biotechnology Information. (n.d.). *LinkOut*. Retrieved March 13, 2002, from http://www.ncbi.nlm.nih.gov/entrez/linkout/

National Center for Biotechnology Information. (2001, January 29). *NCBI BLAST home page.* Retrieved April 6, 2002, from http://www.ncbi.nlm.nih.gov/BLAST/

National Center for Biotechnology Information. (2002). *Entrez search and retrieval system.* Retrieved March 14, 2002, from http://www.ncbi.nlm.nih.gov/Entrez/

National Institutes of Health. (n.d.) *NIH: Grants and funding opportunities.* Retrieved March 28, 2002, from http://grants1.nih.gov/grants/index.cfm

National Library of Medicine. (2001a, October 31). *MeSH browser.* Retrieved March 28, 2002, from http://www.nlm.nih.gov/mesh/MBrowser.html

National Library of Medicine. (2001b, August 17). *NLM online databases and databanks*. Retrieved March 16, 2002, from http://www.nlm.nih.gov/databases/databases_summaries.html

National Library of Medicine. (2001c, November 1). *NLM's databases and electronic information sources*. Retrieved March 14, 2002, from http://www.nlm.nih.gov/databases/databases.html

National Library of Medicine. (2001d, July 26). *PubMed tutorial.* Retrieved March 19, 2002, from http://www.nlm.nih.gov/bsd/pubmed_tutorial/m1001.html

National Library of Medicine. (2002, February 13). *TOXLINE basic search.* Retrieved March 28, 2002, from http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE

National Science Foundation. (n.d.) *NSF: Funding: Overview of grants and awards.* Retrieved March 28, 2002, from http://www.nsf.gov/home/grants.htm

Notess, G. R. (2000). *Search engine statistics: Database overlap*. Retrieved March 19, 2002, from http://www.searchengineshowdown.com/stats/overlap.shtml

Ogilvie, R. I. (2001). The death of a volunteer research subject: Lessons to be learned. *Canadian Medical Association Journal, 165*, 1335-1337. Retrieved March 11, 2002, from http://www.cma.ca/cmaj/vol-165/issue-10/1335.asp

Parrott, J. (2001, October 16). *Bibliographic projects and bibliographic standards.* Retrieved March 23, 2002, from http://www.scholarly-societies.org/bibliog.html

Peek, R. (2001). What if PubSCIENCE is shut down. *Information Today*, *18*, 42. Retrieved February 8, 2002, from Library Literature and Information Science Full Text database.

Perkins, E. (2001, August 7). *Johns Hopkins' tragedy: Could librarians have prevented a death?* Retrieved March 14, 2002, from http://www.infotoday.com/newsbreaks/nb010806-1.htm

Poss, A. J. (2000). *Library handbook for organic chemists.* New York: Chemical Publishing Co., Inc.

Purdue University. (n.d.) *CORE: Comprehensive online research education.* Retrieved April 6, 2002, from http://core.lib.purdue.edu/

Quigley, B. D., Church, G. M., & Peterson, A. (2001). Defining the need for information technology instruction among science faculty. *Science and Technology Libraries*, *20*, 5-42.

Reavie, K. (2000). The virtual library III: Electronic journals, grants, and funding information. In S. Misener & S. A. Krawetz (Eds.), *Methods in molecular biology: Vol. 132. Bioinformatics methods and protocols* (pp. 481-490). Totowa, NJ: Humana Press.

Royal Society of London. (1965). ***Catalogue of scientific papers (1800-1900)*** **(Vols. 1-19).** New York: Johnson Reprint Corp.

Royal Society of London. (1968). ***Catalogue of scientific papers, 1800-1900 : Subject index*****(Vols. 1-3).** Metuchen, N.J.: Scarecrow Reprint Corp.

Schaffer, T. (2001). Databases and political science research. *Online Information Review, 25,* 47-54. Retrieved February 9, 2002, from Emerald Fulltext database.

Schevitz, T. (2002, February 10). A new chapter for libraries: Preference for online research has its price. *San Francisco Chronicle.* Retrieved February 11, 2002, from http://www.sfgate.com/cgi-bin/article.cgi?file=/chronicle/archive/2002/02/10/MN169969.DTL

Science citation index. (1961-2002). Philadelphia: Institute for Scientific Information.

Shooshan, S. (2001). How to search OLDMEDLINE using the NLM Gateway. *NLM Technical Bulletin*, *320*, e7. Retrieved March 13, 2002, from http://www.nlm.nih.gov/pubs/techbull/mj01/mj01_gw_hands_on.html

Tenopir, C. (1992, April 1). Is it any of our business? When end users are happy with poor search results. *Library Journal*, *117,* 96-98.

Tenopir, C. (2001, June 1). The web and content companies. *Library Journal, 126*, 47-50.

Tenopir, C., & King, D. W. (2001, June). *What do we know about scientists' use of information?* Retrieved February 18, 2002, from http://web.utk.edu/~tenopir/NFAIS6-21final.ppt

Tenopir, C., King, D. W., Hoffman, R., McSween, E., Ryland, C., & Smith, E. (2001, May). *Scientists' use of journals: Differences (and similarities) between print and* *electronic*. Retrieved February 18, 2002, from http://web.utk.edu/~tenopir/NOM4-20-01.ppt

University of Arizona Library. (2000, September 15). *Research instruction online*. Retrieved April 6, 2002, from http://dizzy.library.arizona.edu/rio

University of Illinois Library. (2001, December 14). *The Illinois Researcher Information Service (IRIS).* Retrieved March 28, 2002, from http://www.library.uiuc.edu/iris/

University of Texas System Digital Library. (2002). *TILT: Texas information literacy tutorial.* Retrieved April 6, 2002, from http://tilt.lib.utsystem.edu

Woolston, C. (2002, March 4). Scientific misconduct: Keeping your lab clean. *The Chronicle of Higher Education*, Article 2002030401c. Retrieved March 8, 2002, from http://chronicle.com/jobs/2002/03/2002030401c.htm

Table 1

Subject Directories of Web-based Resources

Directory	URL	Description
Librarians' Index to the Internet	http://www.lii.org	Subject directory covering almost every imaginable topic, including the sciences and how to search the Internet, created by California librarians.
SAGE	http://libraries.ucsd.edu/sage/	Subject directory of biology, medicine, physical sciences, engineering, social sciences, and humanities web sites created by librarians at UC San Diego.
BUBL	http://bubl.ac.uk/link/	A subject directory of a variety of academic subject areas arranged by Dewey Decimal Classification and created by the University of Strathclyde, Scotland
PINAKES	http://www.hw.ac.uk/libWW/irn/pinakes/pinakes.html	A directory of subject directories from the Heriot-Watt University, Scotland.
Scout Report Archives	http://scout.cs.wisc.edu/archives/	Directory of web sites categorized by Library of Congress Subject Headings, created by the University of Wisconsin-Madison, and funded by the National Science Foundation.

Table 2

Selected Free Databases from the NLM for Biologists and Medical Researchers

Database	URL	Description
PubMed	www.pubmed.gov	The NLM's free interface to Medline, a bibliographic database of medical and biological journal articles from 1966 to the present.
OldMedline	Can be searched through the NLM Gateway website (http://gateway.nlm.nih.gov/gw/Cmd)	Bibliographic database of biomedical literature from 1958-1965 (NLM, 2001b). In addition, one to two years' worth of older citations are being added each year (McLellan, 2001). Instructions can be found at http://www.nlm.nih.gov/pubs/techbull/mj01/ mj01_gw_hands_on.html(Shooshan, 2001).
PubMed Central	http://www.pubmedcentral.nih.gov/	Provides full text access to biomedical journal content.
ClinicalTrials. gov	http://www.clinicaltrials.gov/	A search interface for information about current clinical research.
TOXNET	http://toxnet.nlm.nih.gov/	A group of databases (including TOXLINE) about toxicology and hazardous chemicals.

Table 3

Selected Free Databases from the NCBI for Biotechnology Researchers

Database	URL	Description
BLAST	http://www.ncbi.nlm.nih.gov/BLAST/	Basic Local Alignment Search Tool. BLAST allows scientist to search all available sequence databases to find genes similar to any given DNA or protein sequence.
Entrez Nucleotide	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide	Database of genetic sequences from GenBank, RefSeq, and the Protein Databank.
Entrez Protein	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein	Database of protein sequences compiled from several databases.

Entrez Structure	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Structure	Database of three-dimensional structures of proteins, DNA, and RNA.
Entrez Genome	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Genome	Database containing the sequences of all of the genes of over 800 organisms.

Table 4

Free Resources for Agricultural and Environmental Studies

Database	URL	Description
AGRICOLA	http://www.nal.usda.gov/ag98/	Database of journal articles about plant biology, animal sciences, nutrition, agriculture, and environmental studies provided by the National Agricultural Library (2001).
Plants National Database	http://plants.usda.gov/	Database of information about plant research, classification, phylogeny, as well as many plant images from the U.S. Department of Agriculture (Johnson, Golden, & Norrisey, 2000).
National Climatic Data Center	http://lwf.ncdc.noaa.gov/oa/ncdc.html	A repository of global and historical climatic information provided by the U.S. Department of Commerce, National Oceanic and Atmospheric Administration.
Energy Citations Database	http://www.osti.gov/energycitations/	Bibliographic citations for reports, conference proceedings, journal articles, patents, books, and dissertations of interest to the Department of Energy (DOE) as well as scientific and technical information from the DOE and its predecessor agencies from 1948 to present. The database covers environmental science, geology, climatology, oceanography, mathematics, chemistry, physics, materials science, engineering, and computer science.

Table 5

Free Information Tools for Chemists, Physicists, and Mathematicians

Tool	URL	Description
PubSCIENCE	http://pubsci.osti.gov/	Bibliographic database with abstracts (and some full text) from peer-reviewed journal articles in physics and chemistry. Provided by the DOE.
PrePRINT Network	http://www.osti.gov/preprints/	Gateway to hundreds of preprint sites and servers, most of which are located on university web sites. The PrePRINT Network allows the user to search for full text math, chemistry, physics, biology, and environmental and materials science articles before they have been published. Provided by the DOE.
DOE Information Bridge	http://www.osti.gov/bridge/	Full-text access to reports from the DOE from 1995 forward.
Energy Citations Database	http://www.osti.gov/energycitations/	In addition to environmental science (see Table 4), this database covers mathematics, chemistry, physics, materials science, engineering, and computer science. Provided by the DOE.
arXiv	http://arxiv.org/	A preprint archive that resides at Cornell University and is funded by the U.S. National Science Foundation. ArXiv contains pre-publication full text of articles in the areas of physics, mathematics, and computer science.
NASA Technical Report Servers	http://techreports.larc.nasa.gov/cgi-bin/NTRS	Citations, abstracts, and in some cases full text of unclassified technical reports from NASA programs (Johnson, Golden, & Norrisey, 2000).

Table 6

Tools for Finding Early Scientific Resources

Database/Index	Format	Coverage	Subject
Catalogue of Scientific Papers (1800-1900)[a]	print	1800-1900	Biochemistry, chemistry, physics and mathematics papers (author index).

Catalogue of Scientific Papers, 1800-1900: Subject Index[b]	print	1800-1900	Biochemistry, chemistry, physics and mathematics papers (subject index).
International Catalogue of Scientific Literature	print	1901-1914[c]	Mathematics, mechanics, physics, chemistry, astronomy, meteorology, mineralogy, geology, geography, paleontology, general biology, botany, zoology, anatomy, anthropology, physiology, and bacteriology.
Science Citation Index: Web of Science[d]	web-based, fee-based (http://www.isinet.com/isi/products/citation/sci/)	1945-present	More than 100 disciplines, including biology, chemistry, physics, mathematics, medicine, geology, computer science, engineering, and behavioral science.
Science Citation Index[e]	print	1961-present	More than 100 disciplines, including biology, chemistry, physics, mathematics, medicine, geology, computer science, engineering, and behavioral science.
OldMedline	web-based (http://gateway.nlm.nih gov/gw/Cmd)[f]	1958-1965[g]	Biomedical literature.
Biological Abstracts[h]	print	1926-present	Biological literature.

MethodsBASE	web-based, fee-based(http://www.methodsbase.com)	1900-present	Methods and protocols for the life sciences.
Cumulative Index to Nursing Literature & Cumulative Index to Nursing and Allied Health Literature[i]	print	1956-present	Nursing and allied health literature.
Index Medicus indexes[j]	print	1879-present	Biomedical literature.
Beilstein Crossfire[k]	web-based, fee-based(http://www.beilstein/com/products/xfire/)	1771-present	Chemical structures, reactions, & publications.
SciFinder Scholar[l] from Chemical Abstracts Service	web-based, fee-based (http://www.cas.org/SCIFINDER/SCHOLAR)	1907-present	Chemical literature.
Chemical Abstracts[m]	print, microform, or CD-ROM	1907-present	Chemical literature.
CompactMATH	web-based, fee-based (http://zb.msri.org/ ZMATH/)	1931-present	Mathematics literature.
MathSciNet	web-based, fee-based (http://www.ams.org/ mathscinet/)	1940-present	Mathematics literature.

[a]Royal Society of London, 1965.
[b]Royal Society of London, 1968.
[c]Parrott, 2001.
[d]Institute for Scientific Information, n.d.
[e]*Science citation index*, 1961-2002.
[f]Instructions for searching OldMedline through the NLM Gateway website can be found at http://www.nlm.nih.gov/pubs/techbull/mj01/mj01_gw_hands_on.html (Shooshan, 2001).
[g]One to two years' worth of older citations are being added each year (McLellan, 2001). [h]Biosis, n.d.
[i]Cinahl Information Systems, n.d.
[j]*Index Medicus* (1879-1927), *Quarterly Cumulative Index to Current Medical*

Literature (1916-1926), *Quarterly Cumulative Index Medicus* (1927-1956), *Current List of Medical Literature* (1941-1959), *Cumulated Index Medicus* (1960-2000), and *Index Medicus* (1960-present). Retrieved by searching the NLM's LocatorPlus database (http://locatorplus.gov/) for the phrase "index medicus."
[k]MDL Information Systems, 2002.
[l]Chemical Abstracts Service, 2002.
[m]Chemical Abstracts Service, n.d.

Back to Contents