President's Column

Patrick J. Mullin

As many of you know, there are often surprises that come along with an appointment to a committee or election to an office within ALA. On a committee, you may be a liaison to another committee or an interest group. Being elected LITA president has plenty of those sorts of surprises. For instance, as LITA vice president, I and the other division vice presidents served as the ALA Appointments Committee. Such "other duties" have a way of filling up my schedule at Midwinter Meeting and Annual Conference.

One such surprise (although, to be truthful, a friend told me about this a bit in advance) is that as president I receive the publications of other ALA divisions. My desk, for instance, now has copies of *Public Libraries*, *Library Resources & Technical Services*, *Reference & User Services Quarterly, ALSConnect,* and *Interface* on it. I will confess up-front that I do not read all of these cover to cover. I do skim them, but I also often read further. As I work my way through these, I am struck by how much we have in common with the other divisions.

All of us are struggling, of course, as divisions who want to retain our members, attract new members, provide opportunity, advance the profession, and help resolve the real-life problems we all face by sharing what we know and believing that others will reciprocate in our time of need.

Our jobs all have the same primary purpose, although we may serve different users or work in different functional areas. "Access" is a theme that runs through these publications and throughout our profession. "Access Services" is a part of my job title. The mission of the American Library Association is "to provide leadership for the development, promotion, and improvement of library and information services and the profession of librarianship in order to enhance learning and ensure access to information for all."

All six points of the Library Bill of Rights are about different aspects of access (www.ala.org/ala/oif/statementspols/statementsif/librarybillrights.htm), although it uses the word only once ("free access to ideas").

Patrick J. Mullin (mullin@email.unc.edu) is LITA President 2005/2006 and Associate University Librarian for Access Services and Systems, University of North Carolina at Chapel Hill.

Library Bill of Rights

The American Library Association affirms that all libraries are forums for information and ideas, and that the following basic policies should guide their services.

I. Books and other library resources should be provided for the interest, information, and enlightenment of all people of the community the library serves. Materials should not be excluded because of the origin, background, or views of those contributing to their creation.
II. Libraries should provide materials and information presenting all points of view on current and historical issues. Materials should not be proscribed or removed because of partisan or doctrinal disapproval.
III. Libraries should challenge censorship in the fulfillment of their responsibility to provide information and enlightenment.
IV. Libraries should cooperate with all persons and groups concerned with resisting abridgment of free expression and free access to ideas.
V. A person's right to use a library should not be denied or abridged because of origin, age, background, or views.
VI. Libraries which make exhibit spaces and meeting rooms available to the public they serve should make such facilities available on an equitable basis, regardless of the beliefs or affiliations of individuals or groups requesting their use.

The word "access" has slightly different meanings to us depending upon the context. You can see how it can be applied in each of the preceding six points. It is access to all kinds of ideas and resources (physical and virtual, popular and unpopular). It is also equal access regardless of condition (origin, age, background, etc.).

Back in the dark ages, when I served on LITA's Technology and Access Committee, I struggled with the word "access." It was everything and it was nothing. It was all encompassing. It had no boundaries. We struggled with access to technology and bandwidth, access to electronic resources, access for persons with disabilities, and so on. Many of today's issues and arguments are rooted in our strongly held mission to provide access.

In your work, how do you define access? How do you contribute to access for your users? It's a struggle. I believe in the power of words and ideas. So do those who seek to restrict access. Why else would they care?

Editorial: I Keep My Eyes Wide Open All the Time

John Webb

It's a Sunday morning—a little after 10 A.M. I'm on a crosscountry flight to Boston to attend the fall meeting of the ARL MetaLib Implementor's Group (playfully known as ARMPIG) chaired now by Roy Tennant. It must be time to write an *ITAL* editorial.

I have had an idea to write about the challenges and experiences I have found during my first year as editor. Of course the writing of an editorial itself is a challenge. Founding Editor Fred Kilgour never wrote editorials so why should I and some of his successors presume the necessity? This particular editorial is even more of a challenge because the key cap that contains the comma is suddenly missing from my laptop. It was there sometime yesterday or at least a couple of days ago. I've looked all around my seat—imagine how ridiculous a passenger crawling around on the floor of a nearly full Boeing 737 looks. Imagine how ominous he looks crawling when he is no further than a short jump shot—just outside the first class line—from the crew compartment. I may have to finish this from a prison cell.

This first year actually began in July 2004 when Managing Editor Marc Truitt and I spent part of a day at ALA headquarters in Chicago being oriented by Kristen McKulski of ALA Production Services. The deadline to send my first issue—March 2005—was November 23! That was eyeopener number one. Manuscripts forwarded by then-editor Dan Marmion began to arrive on my doorstep and in my e-mail within days thereafter.

I learned in orientation that ALA has no article management software to help editors track the flow of the editorial process. Each editor has to devise the systems to ingest articles; to acknowledge their receipt and provide feedback to authors about everything from time frames to changes required or suggested by the referees and editor; to accept or reject articles; to manage the flow of the refereeing process; and to package the edited manuscripts and auxiliary information that are submitted as an issue to ALA Production Services. That was the second eyeopener. And I'm now using my second homegrown system. I dumped the inadequate first and started again from scratch sometime before the third issue. I suspect I may modify it or develop a third sometime this year.

I steeled myself at the beginning for the defensiveness that authors would reasonably adopt when confronted with the often extensive changes recommended by the referees and me. As a referee I had read one published article by a well-known author that contained none of my worthy advice. Some of our authors are almost certainly much better "known" than some of our referees. But authors instead have been unanimously grateful for the feedback. Only one article published in volume 24 required no changes. The changes in one or two others were minimal. Others required multiple rewrites. Our acceptance rate has been 45 percent so far. Some authors whose articles were rejected have thanked me for considering them. Eyeopener number three.

The work performed by ALA Production Services in transforming the Word documents and any accompanying illustrations that I submit into a digital file that looks for all the world like a real journal issue astonished me. I literally jumped in my seat when I opened the PDF file that represented the *first pass* of my first issue! I can compare it only to the experience of stepping out into a cloud- and moonless night in a place unpolluted by the ambient lighting and air pollution of modern civilization and being startled by the vividness of the starry night sky. I get to experience the latter somewhat regularly because I live in the Pacific Northwest. I get to experience the former once a quarter. An appropriate fourth eyeopener. I like both.

The amount of work performed by Marc Truitt in transforming the first pass by working with authors on wording changes and footnote styles and illustration titles and placements and table clarification and grammar and punctuation and capitalization and who knows what else is enormous and was totally unanticipated by me. This is way-big eye-opener number five. (In fact with this first issue of 2006 I have tried to pay more wide-eyed attention to the copy I sent to allay some of his work.) Marc's collaboration with ALA Production Services and the authors *produces* the issues you read. The goal is that the *second pass* will require no or only minor changes. Writing this paragraph has given me another idea: Marc should write an editorial about the job of the *ITAL* managing editor. My only fear is that his description will be such an eyeopener to the membership that we'll never find a successor when he decides he has paid his dues!

I asked myself when I agreed to be nominated for this post if I were walking into it with my eyes wide shut. I asked the same question as I walked down a long hall to the LITA Publications Committee meeting at ALA Midwinter 2004 to be interviewed. I asked again after the interview. It was too late to ask the question again after I received the offer: I had already crossed too many lines in the sand. But my one-and-a-half-year-long first year has been eyeopeningly fun and rewarding. My eyes seem as open as they've been for the past few decades. I don't mind soliciting potential authors for articles. I find that I can beg shamelessly. I seem to have a good rapport with most authors. The *ITAL* board gives me no respect. You should all thank them for that.

Are your eyes open? Did you notice this editorial is comma-less? We're about an hour from Boston now. Will I soon be on my way to my hotel or my prison cell? Keep your eyes open for the next *ITAL* issue.

John Webb (jwebb@wsu.edu) is Assistant Director for Systems and Planning, Washington State University Libraries, Pullman, and Editor of *Information Technology and Libraries*.

Enriching Traditional Cataloging for Improved Access to Information: Library of Congress Tables of Contents Projects

John D. Byrum Jr. and David W. Williamson

Traditionally, standard catalog records have provided bibliographic data that mostly address the basic features of library resources. At the same time, catalogs have offered access to these records through a limited array of names, titles, series, subject headings, class numbers, and a relatively small number of keywords contained within descriptions. Today's catalog users expect access to information well beyond what can be offered by traditional approaches to bibliographic description and access. By pursuing a suite of projects, the Library of Congress (LC) has responded to the challenge of enticing patrons to continue to include the online catalog among the tools they use for information retrieval. Drawing extensively on the power of automation, staff of LC's Bibliographic Enrichment Advisory Team (BEAT) have created and implemented a variety of initiatives to link researchers, catalogs, and Web resources; increase the content of the catalog record; and link the catalog to electronic resources. BEAT's ongoing work demonstrates how, in the electronic era, it is possible to provide new and improved ways to capitalize on traditional services in the digital age. This paper will illustrate these points by focusing on BEAT's tables of contents projects to demonstrate how library automation can make significant bibliographic enhancement efforts quick, easy, and affordable to achieve.

In 1992 the Library of Congress's (LC's) Director for Cataloging established the Bibliographic Enrichment Advisory Team (BEAT) to conduct research and undertake initiatives to enhance the utility of bibliographic records. Composed of voluntary staff from a variety of service units, the team was urged to work outside the box and exempted from the restraints of many policies and practices pertaining to traditional cataloging activities. BEAT was also mandated to create and use automated methods to accomplish its work due to the impact of shrinking staff resources in the bibliographic access divisions.

Among BEAT's earliest undertakings was the development of a series of projects to focus on enriching bibliographic records to include tables of contents (TOCs) information. LC's cataloging policy had been stringent in this area because of the expense of keying such data into records. Indeed, when BEAT decided it needed a benchmark against which to gauge the cost of its TOCs projects, the team experimented with the traditional method of typing the data and concluded that the cost of adding a typical TOC would be about forty dollars per record (in 1992 dollars).

John D. Byrum Jr. (johnbyrum@earthlink.net) is Chief, Regional and Cooperative Cataloging, and **David W. Williamson** (dawi@loc.gov) is Cataloging Automation Specialist, Acquisitions and Bibliographic Access Directorate, Library of Congress.

TOC studies

The theoretical foundations for concentrating on TOCs had been established by research conducted since the early 1980s. Pappas and Herendeen have reviewed the literature and shared their findings, reporting as follows:

- A study at the University of Toronto involving two thousand books revealed that twice as many relevant items for the social sciences and three times as many for those in the humanities were retrieved when users consulted a database that had been enhanced with TOCs.
- Another study found that TOCs added 15.5 unique subject-rich words per record when included in bibliographic descriptions.
- Yet another study of thirty-one publications on the history of taxation in Great Britain found more than six hundred terms in the TOCs to be content-indicative for an average of 19.5 per publication.
- An investigation conducted in 1990 at Carnegie Mellon University using both TOCs and abstracts revealed that contents enhancements increased the number of records retrieved by 20 to 30 percent.[1]

In 1998 Winkle found that 93 percent of a sample of 648 current English language books had TOCs with an average length of 67.75 words that could be included in catalog records. However, only 1.12 percent of the bibliographic records produced by LC at that time included contents notes.[2]

Pappas and Herendeen have also distilled the major advantages of enhancing bibliographic records with TOCs to introduce subject-indicative keywords that otherwise would be excluded from descriptions of publications. Of these, three advantages are considered to be especially compelling: (1) TOCs help users to determine the relevancy of particular titles to their informational needs—a service of value, especially in a closed-stack or remote-storage environment; (2) in an online environment, words in TOCs greatly improve search effectiveness, measured by the ability to identify and retrieve relevant items; (3) by providing content-indicative information, TOCs complement subject cataloging that strives

to summarize the content of a work overall in a few carefully crafted access points per record.[3] Apropos to the latter point, according to an eleven-year longitudinal study cited by Yu and Young, "subject searching [is] being replaced by keyword searching."[4] They reference another study recommending "that subject searchers should select keyword rather than subject headings as their first access strategy."[5]

Contemporary investigations have confirmed the finding that books represented by bibliographic records with TOCs circulate more often than those with corresponding records that do not feature such data. For example, a recent case control study found that "the odds of a title being used increased by 45 percent if the titles had online tables of contents."[6] The Cataloging Enrichment Initiative (RichCat), conceived of and coordinated by Kieft (Haverford College), is being established to encourage production of TOC data for older publications—particularly, those targeted for remote storage—so that catalog users can make informed decisions before recalling particular titles for their research.[7]

Providing TOC information

As a result of such considerations, one of BEAT's earliest efforts to enhance bibliographic records focused on ways and means of providing TOC information.[8] The first application in this area centered on publications being processed through LC's Electronic Cataloging-in-Publication (E-CIP) program. In this program, publishers electronically submit texts for cataloging prior to their publication so that the printed monographs will contain appropriate cataloging information about them. Currently, 55 percent of all publications submitted for Cataloging-in-Publication (CIP) are submitted as part of the E-CIP initiative. In fiscal year 2005 (ending September 30), a total of nearly thirty-five thousand digitally formatted galleys were received.

From 1993 to 1994, an application titled Text Capture and Electronic Conversion (TCEC) was written that enabled cataloging staff to include TOC data programmatically in the bibliographic records they were creating for publications submitted for E-CIP handling. Using the TCEC software and the ASCII-text electronic manuscripts submitted by the publishers, the cataloger highlights the TOC; next, the program manipulates it and adds the result into the bibliographic record's MARC 505 field. TCEC formats the contents information to follow the *Anglo-American Cataloging Rules* specifications for recording TOCs. This includes deleting chapter, section, or part terms, and numbering; eliminating pagination, and adding International Standard Bibliographic Description (ISBD) punctuation. Because TCEC converts all words except the first word in each chapter title to lowercase, the cataloger only needs to highlight any proper nouns that need to be capitalized. The resulting transfer of information from the manuscript to the record is accomplished instantaneously, and data are recorded as accurately as they appear in the electronic manuscript, thus obviating the need for detailed proofreading. Consequently, the former cataloging policy limitation that contents could be given only for monographs that are collections was lifted for E-CIP works. Catalogers are encouraged to apply the TCEC procedure as often as possible, following four criteria:

1. Does adding the chapter titles to the record provide improved natural language keyword searching?
2. Does adding the chapter titles to the record provide a greater understanding of the contents of the item than what is conveyed in the title and statement of responsibility area?
3. Will the TOC data require extensive manual editing to prepare the notes for machine manipulation?
4. If TOC is long and contains many entries, does this dilute the value of the information once it is put into a 505 field?

Fortunately, most staff can make quick decisions in answering these questions.

An informal study suggested that about half of the E-CIP publications would qualify for TCEC-TOC treatment, but catalogers do not always elect to apply this application to TOCs when they should. However, as staff has gradually become more comfortable working with this automated tool, the percentage of catalogers using it to produce contents notes has steadily risen. In fiscal year 2005, 13,627 E-CIPs received TOC treatment, a figure that represents 38 percent of all E-CIP materials received by LC.

In a second E-CIP-TOC project, BEAT members are creating a Web-based TOC record for nearly all E-CIP records that contain TOCs. These Web TOC records are created programmatically; a hot-link in the TOC field to and from the underlying record in the LC bibliographic database is made for every item. The program has been improved recently to include most diacritical marks and to add assigned LC subject headings to the Web versions of TOCs. By the end of fiscal year 2005, approximately sixty thousand E-CIP-TOC records had been added to the Web server.

Entry to the bibliographic record

The net result of these two E-CIP approaches is entry to the bibliographic record in the online catalog through keywords indexed in the TOC field as well as access from the Web, when search engines index the HTML version

of TOCs. As of July 11, 2005, a Yahoo! search on the phrase "contents for library of congress control" produced a result set numbering 242,000 entries, all linked to BEAT's Web-based TOC records. A quick glance at some of the links reveals various uses of TOCs. Some links lead to records within the online catalogs of institutions that had downloaded TOCs. Others lead to such Web sites as "Ethical Schools of Thought," "Mongabay.com," "Hotel Marketing Associates," "Solar sites," "www.on-linenicaragua.com," and many others that cite publications cataloged by LC.

Digital tables of contents

In addition to developing a cost-effective method for enriching records for many of the important publications that are processed through the E-CIP program, BEAT has pursued two other approaches for making TOC information more widely available. The first is its Digital Tables of Contents (D-TOCs) project, which began in the late 1990s. This project has resulted in the creation of machine-readable TOC data derived from photocopied surrogates of TOCs taken from printed publications. By using scanning and optical character recognition (OCR) software as well as original programs written by BEAT's automation staff, the scanned TOCs are subsequently HTML-encoded and placed on one of LC's servers. The techniques used by the project have been modified recently to place heavier emphasis on use of imaging software and on adherence to a highly automated process to convert the TOC data to text format. The D-TOCs project has also implemented more automated and regularized quality control procedures to ensure that links work properly. In the process of HTML encoding, the underlying MARC catalog records are also automatically modified to include links to the TOC data, thus making linkage reciprocal between the two sources of information. Both the MARC catalog records and the linked TOC data may be viewed through a Web browser by accessing LC's online catalog. In addition, the pervasive availability of Web indexing and search software also makes the D-TOCs records available from almost anywhere, providing access to LC's Online Public Access Catalog (OPAC), even for the vast majority of users who are not aware of this project.

Thus, once the Web user has followed a D-TOC link back to LC's catalog, LC can then make the wealth of its collections available for structured searching in items of related interest. As with BEAT's other Web-based projects, D-TOCs serves to help bring Web users back to the library.

The following examples illustrate the various search paths and displays that might be encountered by a user in seeking information both on the Web and in LC's OPAC.

As seen in figure 1, if a Web user searches Yahoo!, for example, using the phrase "animal communication networks," because of interest in a work on this topic, the work by P. K. McGregor would appear near the top of the search results.

If the user clicks on the search result, he or she would be taken to the TOC for that book, partially illustrated in figure 2.

By clicking on "Bibliographic Record," the searcher is taken to LC's OPAC, where he or she will be shown a full description of the work for which the TOC is displayed. The display of the full record as opposed to one of the other possible views is governed by coding in the underlying link, thus providing the maximum amount of information available to the user immediately. Users searching the OPAC with the usual basic search form are initially presented with the Brief Record Display and must subsequently navigate to see more information. This step is eliminated by the link in the Web TOC display used to enter the OPAC (see figure 3).

This record provides hot links to other related works by authors, editors, or others represented by added entries through the related names link(s), as well as to other books on the same topic(s) through the subject link(s). In addition, the searcher can virtually browse the


8. Table of contents for Library of Congress control number 2004054603
... Table of contents for **Animal communication networks** / edited by P.K. ... Behaviours Specific to **Communication Networks**: 2. Eavesdropping in **communication networks** Tom M. ...
www.loc.gov/catdir/toc/cam051/2004054603.html - 4k - Cached - More from this site - Save - Block

Figure 1. Yahoo! partial search results for "animal communication networks"

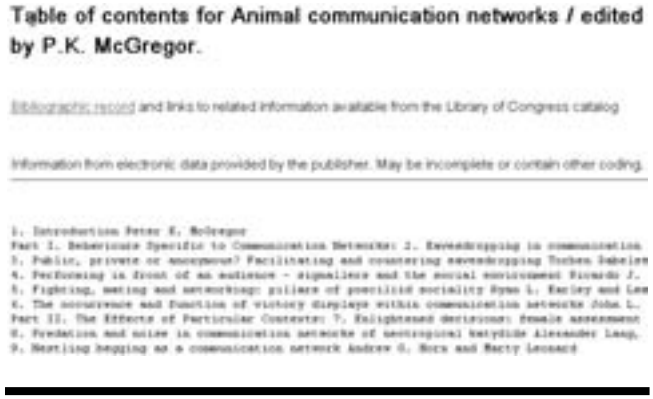


Table of contents for Animal communication networks / edited by P.K. McGregor.

Bibliographic record and links to related information available from the Library of Congress catalog

Information from electronic data provided by the publisher. May be incomplete or contain other coding.

1. Introduction Peter K. McGregor
Part I. Behaviours Specific to Communication Networks: 2. Eavesdropping in communication
3. Public, private or anonymous? Facilitating and countering eavesdropping Torben Dabelsteen
4. Performing in front of an audience - signallers and the social environment Ricardo J.
5. Fighting, mating and networking: pillars of peaceful sociality Hyun L. Earley and Lee
6. The occurrence and function of victory displays within communities networks John L.
Part II. The Effects of Particular Contexts: 7. Enlightened deterrence: female assessment
8. Predation and noise in communication networks of neotropical katydids Alexander Lang,
9. Hearing begging as a communication network Andrew G. Horn and Marty Leonard

Figure 2. HTML TOC record for "animal communication networks" (partial view)

LC shelf by using the call number link to see other books similarly classified, thereby providing entry to other resources of their interests.

Searching and retrieval are improved by various nontraditional techniques, including displaying words from the title and statement of responsibility fields of the bibliographic record, given at the beginning of the TOC display. Also, the keyword metadata tag in the TOC HTML file contains words from the subject heading fields of the bibliographic record, and the subject headings appear in the visible portion of the HTML record. This allows text-based searches on the file (as with a "find" capability resident in most Web browsing programs) while improving delivery of LC's cataloger-supplied vocabulary terms for subject content.

Figure 4 illustrates the D-TOCs project from the vantage of a catalog user. A keyword search of LC's OPAC for the terms "settlers wayne county" would produce the LC record with a hot link to the TOC.

Clicking on the hot link brings up a display of the TOC for the book (see figure 5).

Selection of titles for the D-TOCs project

By the end of fiscal year 2005, more than thirty-one thousand titles had been selected for and processed through the D-TOCs project, and the figure is growing at a rate of 250 to 350 TOCs per week. Most of the publications included are drawn from LC's current receipts, according to the following criteria: those selected should represent items of research value, including anthologies, biographies, and reference materials. In addition, the TOC should contain meaningful words and phrases and not exceed five pages in length. Titles selected are first searched in the database to eliminate those that already have been enriched as a result of other BEAT projects. To date, TOCs have been selected from English language publications. In 2005, however, coverage of the D-TOCs project was broadened to include books in German. In addition, those in Romance languages will soon be eligible. Also underway is implementation of a plan to create D-TOCs files in most of LC's overseas offices, beginning in late 2005.

As an exception to its focus on current receipts, BEAT staff have experimented with retrospective publications acquired by staff of LC's reference rooms. Upon their recommendation, the team began with genealogical works, specifically those in CS71 of the LC Classification schedules, intending to process them alphabetically by family name. (Interestingly, up to 70 percent of the titles in this collection do not have TOCs, possibly due to the fact that the majority of them are self-published.)

ONIX-TOC

The newest, largest, and cheapest of BEAT's three TOC projects is the ONline Information eXchange (ONIX)-TOC application, which was initiated in 2000. This undertaking involves extracting TOC data from publisher-supplied ONIX files. ONIX is an XML (extensible markup language) DTD (document type definition).[8] Publishers use this standardized format to provide book


Animal communication networks / edited by P.K. McGregor.

LC Control Number: 2004054603
Type of Material: Text (Book, Microform, Electronic, etc.)
Main Title: Animal communication networks / edited by P.K. McGregor.
Published/Created: New York : Cambridge University Press, 2005.

Figure 3. Bibliographic record for "animal communication networks" (partial view)


Pioneer settlers of Wayne County, (West) Virginia : including 1842 tax list...

LC Control Number: *00698783*
Type of Material: Text (Book, Microform, Electronic, etc.)
Brief Description: Massie, Evelyn Booth, 1931-
Pioneer settlers of Wayne County, (West) Virginia : including 1842 tax list with illustrations / by Evelyn Booth Massie.
Wayne, W.V. (Rt. 3, Box 3003, Wayne 25570) : E.B. Massie, c1999.
xvi, 448 p. : ill. ; 24 cm.
Links: Table of contents

Figure 4. Bibliographic record for "pioneer settlers of Wayne County, (West) Virginia"

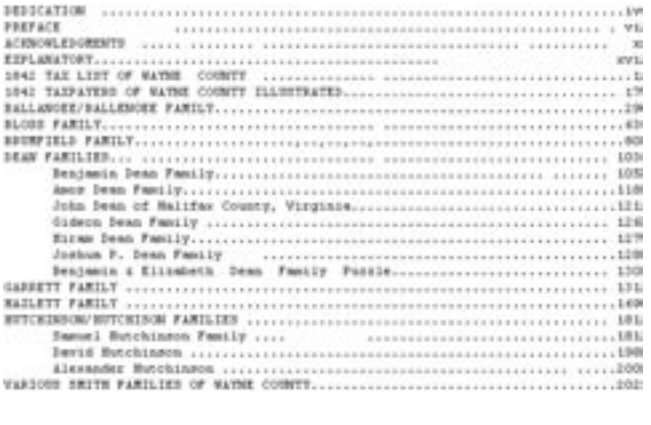


Figure 5. TOC for bibliographic record for "pioneer settlers of Wayne County, (West) Virginia" (partial view)

dealers and retailers with information about their publications; in turn, the retailers can reuse the information for promotional or other sales needs (e.g., creating Web-retailing screens). Because data used are supplied from commercial sources, BEAT's program adds the following disclaimer to each record processed on the basis of ONIX files: "Information from electronic data provided by the publisher. May be incomplete or contain other coding." In reality, such problems are quite rare.

The ONIX-TOC project is based on a Visual Basic program developed by cataloging automation specialist David Williamson, which scans ONIX files to create digital TOCs. The ONIX files are received regularly from publishers who want to make these data available to LC. The program does not validate the integrity of the ONIX file against DTD, but does sequentially seek out each ONIX record to begin processing the data in that record. Depending on the version of ONIX that was used in creating the file (as of June 2005, three versions of ONIX are being received by LC), the first element to be extracted is the ISBN for the book. This is usually the publisher's main identifier for the book. If there is no ISBN found (not yet assigned), the record is skipped and the program goes on to the next record in the file. If the ISBN is found, the ONIX record is searched for TOC information. There are three sets of tags that must be found (each tag has a mnemonic and alphanumeric equivalent):

1. <TextTypeCode> followed by a value of "04" for TOC and </TextTypeCode> to end the information;
2. <TextFormat> with a value usually indicating HTML markup or plain ASCII text followed by </TextFormat>;
3. And, then the actual <Text> tag that starts the TOC to be followed by the </Text> tag signaling the end of the TOC.

If all three sets are found, the data between the <Text> and </Text> tags are extracted. Next, the ISBN is searched against the LC database to see if there is a record for this book that also includes this ISBN.

Three problems can occur at this point:

1. The ISBN may not be unique. While ISBNs are supposed to be unique identifiers, the fact is that publishers sometimes reuse them (intentionally or not). An office outside the United States may apply for CIP for another edition being published outside the United States and may use the same ISBN as the one previously used for the U.S. edition. Tracking within the publisher's office(s) may get jumbled, and numbers may be reused. Another publisher may also accidentally put the wrong number on a publication.
2. If there are multiple records in the LC database, older versions of the program would link the TOC to the record that was entered first into the LC bibliographic database. Until LC started to receive error reports for items with incorrect TOCs linked, the idea of a nonunique ISBN had not been considered. Subsequent investigation found that less than one percent of ONIX records presented this problem. The current version of the program will skip records with duplicate entries in the LC database as manual intervention entails too much time and expense.
3. The book may be represented in the LC database, but the record for it does not contain this ISBN. Publishers create separate ONIX records for each type of binding, for each edition, for each volume in a multivolume monograph, and for associated accompanying materials. If a paperback edition is released well after the hardback edition, and the hardback edition was published before ONIX was received by LC (or was published by a division of the publishing house that does not provide ONIX to LC), then the LC record probably will not have an ISBN for the paperback version. There is no way to equate the record for the paperback edition to the LC record for the hardback edition.

Assuming there is a match in the LC database, the MARC record is further processed, extracting out the LC Control Number (LCCN), the title field, and the LC subject headings. The title field and subjects are then cleaned up for use in the header or footer and the LCCN is added to the link connecting the TOC file to the LC OPAC record.

Because publishers tend to treat their TOC data the same throughout the file (either providing HTML or ASCII), the program is told what the publisher will do. The software will then either accept the HTML coding or, in the case of ASCII text, will wrap "<pre>" and "</pre>" tags around the text in addition to adding an HTML header and footer to the TOC information. Finally, after a spot check for quality assurance, the finished file is saved on the local machine for uploading to the LC Web server. The program then moves on to the next ONIX record, and the process is repeated until the end of the file is reached.

Each of these ONIX-TOC records offers the user an option to visit the bibliographic records in the LC online catalog for further information, following the pattern of the D-TOCs project described above. Similarly, the bibliographic records for these publications are programmatically enhanced by links in the 856 field to the ONIX-TOC files. Some of these records are further enhanced through the addition of book-jacket images (see figure 6).

The ONIX approach has proven to be the most economical in that most of the processing can be started and left to run unattended. Thus, from an actual cost

perspective, the ONIX approach has proved to be very inexpensive.

Cost comparisons

The cost of adding a typical TOC is about $40 per record (in 1992 dollars) for manual keying. BEAT's early initiatives with D-TOCs were much less expensive, about ten dollars per record for the scanning and linking. With better equipment and much more powerful OCR software—BEAT is able to take advantage of LC's use of Prime OCR for performing the conversion to text—the cost-per-record for D-TOCs has fallen to approximately $2 per record. The E-CIP process where the TOC is inserted into the bibliographic record costs about $3 per record, based on guidelines that the cataloger spend no more than five minutes trying to get the TOC into the record.

In comparison, ONIX data cost $0.80 or less per record. The ONIX cost varies depending on the size of the data file received and how many new matches can be extracted from that file. The costs to set up the processing are about eight dollars (for an existing publisher) to ten dollars (for a new publisher) for each run that has to be performed. Once the program is running unattended, the number of successful new TOC files created determines the cost. If ten new TOC files are created, that's about $0.80; if one hundred are created, the cost to drops to $0.08; and if one thousand or more are processed, the cost is less than one cent per TOC for accomplishing extraction and linking.

Harvesting back files

The back files received with new sources of data usually give rise to a one-time harvest resulting in the creation of thousands of new TOC files. For example, when the firm of John Wiley and Sons sent its ONIX back file, 10,090 TOCs were extracted and linked. Wiley was the test case for ONIX; the software to process ONIX files was developed based on this back file, so the costs were a bit higher, $0.26 per record for the 10,090 TOC files. However, once the basic software was developed, it was easily adapted for new publishers, and the per-unit cost has dropped dramatically. For example, when data started to come from the Cambridge University Press DataShop, the software was able to extract and link 12,975 TOC files for $0.0008 per record. More recently received data from Cambridge has far fewer new TOC files available, but on average the cost is about $0.016 per record.

Publishers' ONIX files vary in the amount of information they contain. The information is not aimed at library use but is intended for the book trade, so information

Figure 6. TOC for *Take It From Me*, together with image of the book jacket

about such matters as print runs, availability and pricing, distribution rights, and distributors can be found in the data for each record. While some records may only contain an ISBN, title, and a projected release date (almost an equivalent of a CIP prepublication ONIX record), others are richly loaded with data, including jacket blurbs, reviews, links to the author's Web site, links to cover images, and more.

The ONIX-TOC project is just one of four BEAT-ONIX projects. It was the first, but BEAT has expanded its ONIX projects to take advantage of publisher descriptions (141,000 to date), sample texts either in HTML or PDF (twenty-four thousand), and contributor biographical information for authors, editors, illustrators, collaborators, and so on (fifty-seven thousand). In addition, there is a small test involving forty-four reading-group guides linked from the LC record to the publisher's Web site.

LC currently receives three versions of ONIX: versions 1.1, 2.0, and 2.1. New iterations tend to come out rather frequently, and publishers are not willing to reprogram for each new version, so there are many publishers still using version 1.1. A few publishers have moved up to version 2, but more waited for version 2.1. They are just now beginning to distribute data using that version, even though it has been available since June 2003. All versions through 2.1 are upwardly compatible.

EDItEUR, the group responsible for the ONIX standard, will release version 3.0 in late 2006.[9] This latest version is essentially the same as version 2.1, but all deprecated tags have been removed. Thus, there is no compatibility with the older versions, so programmers do not have to take into account any deprecated tags. Changes to the ONIX standard seem to be moving more to changes in code lists associated with the standard rather than changing the standard itself. This allows the standard to remain stable longer and requires less programming when there is a change, such as for a new type of contributor to a work or a new language code to be added. The change can be handled more efficiently in a code listing.

Today, there are nearly sixty thousand of BEAT's ONIX-TOC records available on the Web. This is a steadily expanding figure because the pool of publishers making their ONIX data available to LC continues to grow. Since February 1998, counters have monitored access to BEAT's Web-based TOC files. Hits currently range from four hundred to five hundred per hour between 8:00 A.M. and 9:00 P.M. eastern time to around two hundred per hour overnight; however, the rate is increasing rapidly. By October 2005, more than 7.5 million hits had been recorded. In addition, as with most BEAT products, the records enriched to provide links to these records are redistributed by LC's Cataloging Distribution Service (CDS), making them available to users of OCLC and RLIN, as well as to other agencies that subscribe to CDS products.

TOC Web survey

To determine whether BEAT's TOCs were, in fact, serving useful purposes, a simple Web survey was developed and a small selection of the HTML-TOC files modified to provide a link to it.[10] The survey was posted from August through October 2001 and elicited input from 360 Web users. When asked how they found the TOC file, 60 percent reported "from the bibliographic record in a library catalog," while 36 percent said "from an Internet search." Of those who responded to the question "Was this TOC information useful?" 84 percent replied in the affirmative. When asked "Did you go to the bibliographic record from the link on the TOC page?" 58 percent answered "yes." Of these, 57 percent indicated that they had "look[ed] over" the bibliographic record, and some had also clicked on the hot links within the bibliographic record to search for works by the same authors or on related subjects. Asked to describe themselves, 58 percent indicated that they were researchers or students, 23 percent were librarians, and 14 percent were casual users looking for information. Survey participants had an opportunity to comment on their Web TOC experience. Their opinions confirmed and extended the various uses and overall serviceability of TOC information summarized by Pappas and Herendeen.[11] Librarians found BEAT's TOCs helpful in making acquisition decisions, especially in cases of expensive publications, and in downloading the TOC for addition to records in their OPACs. The survey was repeated about eighteen months later with almost identical results.

505 data

Although LC's D-TOCs and ONIX-TOCs records provide access to the online catalog, they do not provide entry to bibliographic records from within the catalog, because the bibliographic records contain only links in lieu of the TOCs themselves. To counter this drawback, a program was created to add full TOCs to the bibliographic records for the Web-based TOCs. Beginning February 2005, it proved possible to use this application to enrich bibliographic records by adding information that was previously only available through 856 links.

The 505 data are automatically generated from the TOC information in the files created for the D-TOCs and ONIX-TOC records. The program scans the TOC file and extracts out each line of the TOC data, treating each line as an element in the 505 field being constructed. For many TOCs, this works perfectly well to extract the chapter titles. In the case of multiline TOC titles, this approach causes a TOC title to become two or more elements in the 505 field, potentially causing confusion. Similarly, when multiple chapter titles are on one line, some muddling of the data will occur. Each application of the program will introduce the TOC with the legend: "Machine-generated contents note."[12] Because the scanned TOCs come in a wide variety of formats and structures, some errors are to be expected in the placement and configuration of the 505 textual strings. Space, hyphen, hyphen, space will be inserted after each line break within the TOCs. In many cases, chapter and page numbers will appear as captured from the scanned TOCs images. The 505 data will not undergo review for punctuation (see figure 7).

Approximately sixty thousand LC records with existing 856 links to TOC texts are being batch-processed, modified, and redistributed until all eligible records are enhanced. Initially, after consultation with LC's public service staff, TOCs that are four thousand or less bytes in extent have been declared eligible, but larger-sized records may become eligible for processing as described above. Later, eligible ONIX-TOC records may be similarly processed.

Brief Record | Subjects/Content | Full Record | MARC Tags

Anni Albers : selected writings on design / edited and with an introduction...

LC Control Number: 00010582
Type of Material: Text (Book, Microform, Electronic, etc.)
Personal Name: Albers, Anni.
Main Title: Anni Albers : selected writings on design / edited and with an introduction by Brenda Danilowitz ; foreword by Nicholas Fox Weber.
Published/Created: Hanover : University Press of New England, c2000.
Related Names: Danilowitz, Brenda.
Description: xiii, 79 p., [24] p. of plates : ill. (some col.) ; 24 cm.
ISBN: 0819564478
Contents: Machine generated contents note: Foreword Nicholas Fox Weber vii -- Introduction Brenda Danilowitz ix -- A Start -- Weaving at the Bauhaus 3 -- Work with Material 6 -- Art-A Constant Ill -- Designing 17 -- On Jewelry 22 -- One Aspect of Art Work 25 -- Constructing Textiles 29 -- Design: Anonymous and Timeless 34 -- Review of Ben Nicholson's Paintings, Reliefs, Drawings 42 -- The Pliable Plane: Textiles in Architecture 44 -- Conversations with Artists 52 -- Weaving, Hand 55 -- Designing as Visual Organization 58 -- Tactile Sensibility 69 -- Material as Metaphor 73 -- Bibliography 77 -- Illustrations follow pages 16, 32, and 64. -- Captions for the portraits of Anni Albers appear on page 80..
Notes: Includes bibliographical references (p. [77]-79).
Subjects: Design--Philosophy.
Decoration and ornament--Philosophy.
LC Classification: NK1505 .A38 2000
Dewey Class No.: 701/.8 21
Quality Code: pcc
Electronic File Info.: Table of contents http://www.loc.gov/catdir/toc/fy034/00010582.html
Links: Table of contents

Figure 7. Sample bibliographic record with machine-generated TOC

Conclusion

BEAT's TOC projects demonstrate how, in the electronic era, LC is taking traditional services and providing new and improved ways to capitalize on them in the digital age. These projects provide a model that might be of interest to others as they ponder issues and opportunities regarding bibliographic access and retrieval in today's growing electronic environment. By responding to expanding user needs through bibliographic enrichment initiatives such as TOCs, libraries will recognize that, whether in a traditional framework or in the digital environment, researchers can and do use the catalog the way an entire library is used—not only as a source of material and information, but also as a gateway to additional information. Through adding more keyword-rich information to the catalog, libraries can serve the extended information needs of the researcher as well as offer structured pathways to their own information resources. Offering such features as standardized subject terminology and pervasive controlled headings, these catalogs are the result of more than one hundred years of intellectual effort and real capital. Considering the major investments made to create and maintain their catalogs, libraries everywhere should seek opportunities to build upon these investments to provide richer records in order to entice patrons to continue to include the online catalog as a rewarding access mechanism in their growing array of tools for information retrieval.

References and notes

1. Evan Pappas and Ann Herendeen, "Enhancing Bibliographic Records with Tables of Contents Derived from OCR Technologies at the American Museum of Natural History Library," *Cataloging and Classification Quarterly* 23, no. 4 (2000): 65–67.

2. R. Conrad Winkle, "An Analysis of Tables of Contents in Recent English-Language Books," *Library Resources and Technical Services* 43, no. 1 (1998): 14.

3. Pappas and Herendeen, "Enhancing Bibliographic Records," 63–64.

4. Holly Yu and Margo Young, "The Impact of Web Search Engines on Subject Searching in OPAC," *Information Technology and Libraries* (Dec. 2004): 168.

5. Ibid., 169.

6. Ruth C. Morris, "Online Tables of Contents for Books: Effect on Usage," *Bulletin of the Medical Library Association* 89, no. 1 (Jan. 2001): 29.

7. More information regarding RichCat is available at www.loc.gov/standards/catenrich (accessed Dec. 21, 2005).

8. A streaming videocast recorded in January 2002, containing information relating to all of BEAT's TOC initiatives as of that date, may be viewed online at http://lcweb.loc.gov/catdir/beat/eTOC/jan30-eTOC.html (accessed Dec. 21, 2005).

9. For more information on ONIX, visit the EDItEUR home page, www.editeur.org (accessed Dec. 21, 2005). EDItEUR is the agency responsible for coordinating the various national ONIX groups and distributing the ONIX standard.

10. Survey results and comments are available for review at www.loc.gov/catdir/tocsurveyresults.html (accessed Dec. 21, 2005).

11. Pappas and Herendeen, "Enhancing Bibliographic Records."

12. The 505 indicators for these machine-generated notes will be set to "8" (No display constant generated) and blank (Basic; single occurrence of subfield $a).

Index to Advertisers

Medium-sized Universities Connect to Their Libraries: Links on University Home Pages and User Group Pages

Pamela Harpel-Burke

From major tasks—such as recruitment of new students and staff—to the more mundane but equally important tasks—such as providing directions to campus—college and university Web sites perform a wide range of tasks for a varied assortment of users. Overlapping functions and user needs meld to create the need for a Web site with three major functions: promotion and marketing, access to online services, and providing a means of communication between individuals and groups.[1] In turn, college and university Web sites that provide links to their library home page can be valuable assets for recruitment, public relations, and for helping users locate online services.

Goals for Web sites may vary for different users. Colleges and universities have a number of potential user groups: current students, faculty, staff, prospective students, donors, alumni, businesses, and media.[2] Because the "fundamental organizing principle in Web site design is meeting users' needs," the organization of the home page should reflect the specific needs and interests of these user groups.[3] By segmenting their home pages into distinct user-group pages, universities may be better able to meet the needs of their various users. Segmentation into subsites for distinct user groups is an effective navigational device that manages complex Web sites with numerous pages and a variety of audiences.[4] As an added benefit to this structure, users should feel more welcome at a subsite that is designed specifically for them.[5]

Middleton, McConnell, and Davidson divided the users of a university Web site into two categories: internal (faculty, staff, and students) and external (prospective students, alumni, donors, parents, community, visitors, and news media).[6] While the needs of these groups overlap, internal users primarily need accessible, useful tools that will help them become productive and successful in their work or educational life. In contrast, external users primarily seek details on academic programs, campus environment, news and events, and contact information.

Along with the university home page in general, the representation of the library on the Web site has also diversified to meet the needs of these different user groups. Faculty, staff, and current students (internal users) may seek out the library for access to information that assists them in the completion of their work. Internal users need access to library services and resources such as course reserves, online databases, the catalog, and interlibrary loan. In all likelihood, internal users frequently will seek a wide variety of information on the library Web page. For external users, however, the library Web site's primary function is for promotion of the institution. As an example, prospective students and their families may compare the library to that of other universities under consideration. Donors, alumni, and members of the business community may be interested in visiting or contributing to the library. Consequently, external users are more concerned with such characteristics of the library as the size and age of the collection, electronic resources, and the actual facilities. Hence, a link to the library is expected to appear more often on the subsites for internal users than on subsites for external users.

Link placement is also important. Web site usability research has demonstrated that links in the upper part of the page in either corner are much more likely to be noticed than links placed elsewhere.[7] Therefore, access to the library from the university home page and associated user group pages can be affected by spatial placement of the link, the need to scroll (below the fold), drop-down menus, direct links, and terminology used for the library. A direct link to the library in a prominent position with comprehensible labeling maximizes the university's investment into the libraries' online resources.

It is important to note that the representation of the library on the university's home page and subsequent user group subsites is not an indication of the worth of the library as perceived by the institution. Although university Web site designers may group the library with general services such as food services or facilities, the absence or presence of a library link, terminology used, and link placement may not manifest the value that the university places on the library. Other factors, such as financial support and involvement in the curriculum and research, also serve to underscore the library's value within the university community. However, as noted by Astroff, the Web page is "one officially approved representation of the university's infrastructure." [8]

Literature review

Several authors have previously dealt indirectly with the association between the university home page and the library home page. In King's 1998 study of home page design of 120 Association for Research Libraries (ARL) institutions, he also looked at the placement of the library link on the associated parent university home page and counted the number of steps from the university home

Pamela Harpel-Burke (Pamela.Harpel-Burke@hofstra.edu) is Catalog Librarian and Assistant Professor, Hofstra University Libraries, Hempstead, New York.

page to the library's home page.[9] For the 111 libraries that had parent institutions, he noted that at times it was necessary to search the home page for the library link. Forty-four percent of the library links could be located in one step while 50 percent of the institutions required the user to find the library link in two or more steps. For the remaining home pages, no association between the library and its parent institution could be found.

Dewey analyzed the findability and placement of links to various services on the library home page of twelve of the thirteen member libraries of the Committee on Institutional Cooperation (CIC).[10] She also noted that a direct link to the library from the university home page affects access to these services. CIC institutions fared poorly in this regard, with only five of the twelve institutions having a link to their library home page. For the remaining seven, the link was only found by searching the institution's Web site.

One hundred and thirty-three library Web sites for medium-sized universities (six thousand to thirteen thousand students) were examined for specific features and core components in 2000 with a follow-up study in 2005.[11] Sample size and university type were similar to those in this study. Even in cases where the process was lengthy and complicated, they found in the 2000 study that they were always able to "navigate from the parent institution to the library page either by a direct link or multiple links."[12] However, 32 percent of the institutions in the initial study had no link, logo, or drop-down menu that gave access to the library home page from the university home page. It was noted that the presence of a library icon on the university home page increases the likelihood that students will easily find the library home page. In Tolppanen, Miller, Wooden, and Tolppanen's subsequent study in 2005 of the same set of universities, the percentage of the universities without access to the library home page from the university home page had only dropped to 29 percent.

In a comparison of Web sites of historically black college and university (HBCU) libraries and other public institutions in the southeastern United States, Agingu examined the usefulness of HBCU library Web sites as service providers and information disseminators.[13] Out of sixty-five library Web sites examined, only two libraries were inaccessible from the main university home page. However, she noted that 23 percent of those accessed were identified only after searching the university home page. She also advocated the presence of a direct link on the home page for visibility and ease of location.

In a more recent article, Welch surveyed 106 academic libraries (ARL and non-ARL) to evaluate the use of the library Web site for marketing library resources and services, exhibits, programs, and fundraising.[14] The placement of a link to the library from the university home page was also studied, and the increased visibility of the library through direct links was emphasized as a public relations vehicle for the library. Welch reported that 80 percent of these libraries surveyed had direct links to the library from the institutional home page.

Stover and Zink examined forty library home pages in higher education. Instead of investigating the links from the university home page, they surveyed links from the library home page.[15] Eighty percent of the library home pages in their study linked back to the university's home page. They reasoned that such a link enhanced the library's importance within the overall university community.

Astroff provided the most comprehensive research on access to libraries from university home pages in her study of 111 ARL institutions. Sixty-one percent of the university home pages had visible links to the library home page. Home pages were classified as having no visible link to the library if it was necessary for "the user to make some choice or perform some action before such a link becomes visible."[16] For example, the use of a drop-down menu box or a pathway to the library link via mouseovers was not considered a direct link to the library.

The importance of providing a link to the library from the university's home page was recognized early in the development of university Web sites.[17] Astroff noted that the trend to "organize information by the role the user plays in the university community" developed during her study. She claimed that segmentation into user groups can lead to a "very clean design but one that provides very little information and no place for a visible library link."[18] However, this author contends that a user group design may allow the library link to be more obvious on the home page. By relegating more information on various services to the specific user group, there is more white space on the home page for the library, and the user groups can be provided with information in a more usable format.

This study

This study is unique because it takes into account the trend to organize university home pages by user groups. With a focus on four-year, medium-sized universities, the accessibility of the library on pages targeted to associated user groups as well as accessibility from the university home page in general is examined.

In publications on Web site design, a variety of terms have been used to refer to these groups, such as subsites, user segments, or segmentation; in this paper, the term *user groups* will be used to represent these concepts. In addition, *university* will be used when referring to the college, university, or institution.

Method

University home pages: data set

A data set of home pages from four-year, medium-sized universities with a total student population of eight thousand to thirteen thousand students—full-time and part-time undergraduates and graduate students—was analyzed. University type and enrollment size of the author's home university (Hofstra University) was used as a baseline for comparison to other medium-sized universities. Hofstra University, a private, nonsectarian, coeducational university, is located twenty-five miles east of Manhattan on Long Island in Hempstead, New York.[19] Total enrollment at Hofstra University, including part-time undergraduate, graduate, and School of Law, is approximately thirteen thousand; full-time undergraduate enrollment is 8,067.

In order to insure that a robust data set was obtained, two resources were used to generate a list of comparable universities: the "Compare Academic Libraries" tool from the National Center for Education Statistics (NCES) Web site and *Peterson's College Bound Online*. The NCES comparison tool used data from the *Academic Libraries Survey* (*ALS*) for fiscal year 2002, which collected information from thirty-seven hundred academic libraries in the United States and outlying regions. *ALS* gathered data on libraries with "accredited degree-granting institutions of higher education and on the libraries in non-accredited institutions with a program of four years or more."[20]

The comparison group included universities within 20 percent of Hofstra's total enrollment of thirteen thousand students; the NCES generated list was comprised of 168 universities with enrollments ranging between 9,023 and 13,534 undergraduates. Three Carnegie Classification levels were retained: Doctoral/Research Universities—Extensive (Class 15; n = 16); Doctoral/Research Universities—Intensive II (Class 16; n=25); and Master's Colleges and Universities I (Class 21; n=36). Three categories were eliminated from analysis because they were not comparable to Hofstra's Carnegie Class 16: Master's (Comprehensive) Colleges and Universities II (Class 22), Associate of Arts Colleges (Carnegie Classification 33) and Schools of Engineering and Technology (Class 54). The final list from *ALS* contained 103 libraries.

Similarly, *Peterson's College Bound Online* was limited to colleges with eight thousand to thirteen thousand full-time students, excluding Canada and Puerto Rico, and community colleges. After universities that were duplicated on both *Peterson's* and *ALS* were removed, the list contained ninety universities. Because ARL universities are already studied elsewhere, they were also excluded.

A data set was formed with all universities that appeared on both lists. Finally, because Hofstra University is a private university with a Carnegie Classification 16 (Doctoral/Research Universities—Intensive II), universities with this level were retained if they appeared on either list. Eight universities were then eliminated from the data set because they were either online colleges or colleges that focused on the health sciences only. Some of the universities in the data set were part of a consortium that shared a home page, such as City University of New York (CUNY). Therefore, CUNY Brooklyn, City College, Lehman College, and Queens College were analyzed jointly and considered one home page. However, the CUNY Bernard M. Baruch College was examined separately because it had a unique home page. The final data set examined contained seventy-seven universities.

Software, hardware, and retrieval

Hardware used for the assessment of university pages was a Dell Pentium 4 (CPU 2.66GHz 384 MB of RAM) with a 16-inch, 32-bit color Intel Plug-and-Play Monitor with screen resolution of 1024 by 768 pixels. Software included the operating system (Microsoft Windows XP Professional version 2002) and browser (Microsoft Internet Explorer version 6.0.2600.000). Pages were displayed with no other toolbars than the Internet Explorer address bar. Screen shots of each page used for data collection were copied and retained for future reference.

Criteria examined

Each university home page was surveyed to determine whether there was a link to the library home page. If the university had a library link, additional data were collected. Drop-down menus or mouseovers were considered to be visible links in this study of home pages of medium-sized universities. Additional data on the use of drop-down menus for library links were collected.

If links to the library were displayed in a button bar, banner, or drop-down menu, they may have been due to page grid designs or server limitations. These navigational aids were often recorded as the second link to the library. On the university home page, a link to the library may also be represented in navigation bars or menus along the top or bottom of the university home page. When these navigational aids were present, they were also checked throughout each user group to see if they persisted as part of the structure of subsequent pages.

When there was more than one link to the library, hypertext links in the upper portion were recorded as the first link. Additional hypertext links were numbered according to their location on the page; links in the center were second, and links at the bottom were third. However, if a link was a drop-down menu, but-

ton bar, or banner menu that was the result of a page grid design, location was ignored and the link was always recorded as the last link (either second, third, or fourth).

University home pages were also examined to determine if the user needed to scroll to see the library link. If there was a link to the library from the university home page, the link was tested to see if it led directly to the library home page. Links that led users to a Web page that represented a collection or system of libraries at the parent university were recorded as direct links. University home pages were also examined to see whether there was more than one link to the library home page.

Specific characteristics of each link were examined. The exact terms, such as *Libraries* and *Library Resources,* used to label the library link were noted. Spatial placement of the library link on the university home page or specific user group was determined by dividing the screen into nine sections—upper left, upper center, upper right, center left, middle center, center right, lower left, lower center, and lower right—and noting the location of each library link.

Each site was also examined to determine whether the university had divided its home page for separate user groups. Many of the university home pages were designed to provide entry points to university information for each of these separate user groups. Data were collected for eight common user group categories; these user groups were categorized as external and internal users. Three types of users, Faculty, Staff, and Current Students, were identified as internal users. Where present, as many as five types were identified as external users: Prospective Students, Alumni and Donors, Parents, Visitors and Community, and Business and Media.

Variables examined for each category of user included the presence or absence of a link to the library on each specific user page and the number of links to the library. For the first and second library links, placement of these links and the terminology used to identify the library were recorded.

In some cases, university home pages were hybrid versions of the traditional university home page display and the segmented user group display. In these cases, some user categories were given but the home page as a whole was not segmented into specific user groups. These cases were included in the analysis. However, if the main university home page did not have divisions for user groups, the Web pages were not examined further.

Statistics

Standard two-tailed t-tests were used to compare the number of library links on university home pages, internal user groups, and external user groups. Chi-square tests and contingency tables were used to test for significant variation in link placement and the presence of at least a single library link on internal as opposed to external user group pages.

Results

An examination of home pages of seventy-seven medium-sized universities (eight thousand to thirteen thousand total student population) was undertaken in spring 2005. Thirty-two states and the District of Columbia were included in the sample. The states with the largest number of universities surveyed were New York (n=11), Missouri (n = 8), and Ohio (n = 4). Thirteen of the states were represented by only one university; ten of the states had two universities; and seven states had three universities included. The majority of universities were public (n=58); the rest were private nonsectarian (n=14) and private universities with religious affiliations.

Results of the study are arranged into three main categories: information about the library link on the main university home page, information on the library link within the Web pages for internal, and information on the library link within the Web pages for external user groups.

University home page characteristics

Seventy of the university home pages (91 percent) had an obvious, but not necessarily direct, link to the library home page (table 1). Most of the home pages of medium-sized universities had a direct link to the library home page (83 percent). Sixteen (23 percent) of the university home pages had the link to the library incorporated into a drop-down menu. In some cases, the university provided additional opportunities for the user to find the library on the home page by including a hypertext link along with the drop-down menu. Twenty-four percent of the university home pages had two or more links to the library.

University home page terminology

Terminology used to identify the library on the university home page was fairly uniform. For the first link to the library, all home pages used a phrase that contained the word *library* or its plural. Specifically, the terms Library or Libraries were used by 77 percent of the universities for the first link. Thirteen percent of the universities used a phrase containing the word Library or Libraries along with another term (i.e., Library and Media Center; Libraries and Technology; Library Resources). Nine percent of the universities used the specific name

of the library (i.e., Andruss Library at the Bloomsburg University of Pennsylvania; Booth Library at Eastern Illinois University; Maag Library at Youngstown State University). This terminology pattern was consistent for those Web pages with multiple library links—all used the word Library or its plural.

University home page link placement

Spatial placement of the first library link on the university home page was quite varied. Nearly 29 percent of the universities had a link to the library home page in the center on the left side of the page. Twenty-one percent placed the link to the library in the upper right corner. Overall, the center was the preferred place for the first link to the library; 47 percent of the universities had the first link in the center section (center left, right, or middle) of the home page. Fewer of the links were in the upper part of the university home page; 36 percent of the universities had the first link in the upper section (upper left, right, or center) of the home page. Only 17 percent of the links were in the lower part (lower left, center, or right) of the home page. Of these universities with the link in the lower part of the page, the majority of them (75 percent) required the user to scroll down the page to see this link.

Placement of the second link to the library on the university home page was quite different. For the seventeen libraries with such links, none were in the center area of the home page. Placement of the second link to the library was fairly evenly divided between the upper section (47 percent) and the lower section (53 percent) of the universities' home pages. Because this additional link was in the upper or lower parts of the page, it seems likely that the second link was attributed to a follow-through menu design.

Table 1. University home pages: library links

Criteria Examined	Responses	N= Number for subset	N= for total	%
Is there a link to the library?	Yes	70	*77*	91
	No	7		9
Is it a direct link to the library?	Yes	58	*70*	83
	No	12		17
Is it necessary to scroll to see the library link?	Yes	9	*70*	13
	No	61		87
Is the link part of a drop-down menu?	Yes	16	*70*	23
	No	54		77
Is there more than one link?	1 link	53	*70*	76
	2 links	17		24
What is the term used for the first library link?	Libraries or Library	54	*70*	77
	Library or Libraries and "other"	10		14
	Library "name"	6		9
Where is the first link?	Upper (Left, Center, Right)	25	*70*	36
	Center (Left, Center, Right)	33		47
	Lower (Left, Center, Right)	12		17
Does the menu follow through to the other Web pages?	Yes	52	*70*	74
	No	18		26
Are there user group segments?	Yes	70	*77*	91
	No	2		3
	Some	5		6

The link to the library from the university home page could also be displayed in a button bar along the top, a drop-down menu or a menu along the bottom or one side that is retained throughout the Web site. Specific characteristics, such as placement and terminology, for these links were generally recorded as the second or subsequent link. Seventy-four percent of the universities' Web sites carried the library link through and displayed it on the home page for some or all of the succeeding user groups. Page grid layout or server limitations contributed to some of the user groups' pages having up to four links to the library on their individual home pages. For example, 50 percent of the pages in the Faculty and Staff user category had two or more links to the library.

University home page user group segments

Seventy (91 percent) of the seventy-seven university

home pages had clear user group delineation. In five other cases, the university home page had indistinct user groups; these home pages presented some user groups within the traditional university home page design. These five home pages were included in the analysis of user groups. Only two of the seventy-seven university home pages had no user group segmentation and therefore were discarded from the user segmentation analysis.

All of the seven university home pages with no library link (direct or indirect) had user group segmentation. Of the twelve university home pages that had only an indirect link to the library, eleven had user group delineation. Although data were collected for eight categories of users, the majority of universities had six categories or fewer: Faculty and Staff, Current Students, Prospective Students, Alumni and Donors, Parents and Family, Visitors and Community. The Faculty and Staff categories were combined for this analysis because sixty-seven of the seventy-five universities had Faculty and Staff combined into one user group. Only two of the universities had a separate category for Staff. In one case, California State University, Hayward/East Bay (CSU), the Staff page contained the same information in the same arrangement as the Faculty page. The Staff page did not have a link to the library at the other university (Howard University). Of the seventy-five libraries with user segments, six of them had no category that included Staff.

Only six (8 percent) of the universities had a Business or Media user group represented on any of their university home pages. Three of these had links to the library home page; two of these links were only through a menu bar and one was only through a drop-down menu that appeared throughout the Web pages. Thus, in these cases the link was probably due to a follow-through menu determined by the software design. No further data for this category of user are presented here.

Internal user groups link number

As noted above, data on the user groups were divided into two primary sections: internal and external. The internal users consisted of two types: Faculty and Staff and Current Students. As shown in table 2, 94 percent of the Faculty and Staff group had a link to the library. Of these, 50 percent had two or more links to the library in their Faculty and Staff page. Pages for Current Students had results similar to the Faculty and Staff category; 97 percent provided a link to the library from the student page. For those student pages that had links to the library, 56 percent had only one link and 44 percent had two or more links to the library. Eleven of the Current Student pages had three to four links to the library.

Internal user groups terminology

The Faculty and Staff and the Current Students groups had similar terminology usage, with most of the groups using either Library or Libraries or Library name. The terms Library or Libraries were used most often by both user groups; Faculty and Staff used these two terms 46 percent of the time, Current Students used them in 52 percent of the cases. The university's library name was used on 27 percent of the Faculty and Staff pages and 22 percent of the Current Students pages. Library or Libraries and "other" were used on 18 percent of the Faculty and Staff pages and 11 percent of the Current Students pages.

Internal user groups link placement

Spatial placement of the link to the library for the Faculty and Staff group and the Current Students was fairly widely distributed. The link to the library was placed in the upper part of the page 35 percent of the time for Faculty and Staff and 27 percent for the Current Students. The center of the page was used for the library link 29 percent of the Faculty and Staff pages and 36 percent of the Current Students pages. The link to the library was placed in the lower part of the page 35 percent of the time for Faculty and Staff and 36 percent of the Current Students.

External user groups link number

The external users consisted of four types: Prospective Students; Alumni and Donors; Parents and Family; and Visitors and Community (table 3). Of external users, the Prospective Student group had the highest number of links to the library (74 percent). The other groups had similar percentages of links to the library from their page; 67 percent of the Parents and Family group had links to the library while 60 percent of the Visitors and Community group had links to the library. Only 49 percent of the Alumni and Donors category had links to the library.

External user groups terminology

Terminology used by the external groups was fairly consistent for all groups. For all of the user groups, most of the links were labeled Library or Libraries. The Alumni and Donors group used one of these two terms 72 percent of the time; the Prospective Students group used one of the terms 62 percent of the time. The Parents and Family group used either Library or Libraries for the link 54 percent of the time; Visitors and Community used one of these terms 70 percent of the time. The specific university Library name was used often by two of the groups. Twenty-four percent of the Prospective

Students group used the specific name of the library; the Parents and Family group used Library name 25 percent of the time.

External user groups link placement

Placement of the library link on the four external user group pages was varied, but most of the external user groups had the library link in the upper section of the page. Sixty-four percent of the Alumni and Donors pages, 61 percent of the Visitors and Community pages, 49 percent of the Prospective student pages, and 42 percent of the Parents and Family pages placed the link in the upper portion of the page. The library link was located in the center portion of the page least often by all of the external user groups. For example, the center position for the Prospective Students group was used 22 percent of the time; Alumni and Donors used this position only 11 percent of the time.

Discussion

Links to the library: Comparison with other studies

Comparison to other studies that examined the link to the library on the university home page shows diverse patterns. Although earlier studies that examined the university home page for a link to the library utilized a variety of data sets, criteria, and university types, simple

Table 2. Internal user groups: faculty and staff and current students

		Faculty & staff 1st occur *N*=72			Faculty & staff 2nd–4th occur *N*=43			Current students 1st occur *N*=73			Current students 2nd–4th occur *N*=45		
Question	**Responses**	**set *n*=**	**Total *N*=**	**%**	**set *n*=**	**Total *N*=**	**%**	**set *n*=**	**Total *N*=**	**%**	**set *n*=**	**Total *N*=**	**%**
Is there a link to the library?	Yes	68	72	94				73	75	97			
	No	4		6	NA			2		3	NA		
How many links are there to the library?	1	34	68	50	NA			41	73	56	NA		
	2	26		38				21		29			
	3 to 4	8		12				11		15			
What term is used for the library link?	Libraries or Library	31	68	46	20	43	46	38	73	52	23	45	51
	Library or Libraries & "other"	12		18	15		35	8		11	14		31
	Library name	18		27	3		7	16		22	4		9
	University libraries	5		7	3		7	6		8	2		4
	Other	2		3	2		5	5		7	2		4
Where is the link?	Upper (Left, Center, Right)	24	68	35	21	43	49	20	73	27	24	45	53
	Center (Left, Center, Right)	20		29	6		14	27		37	5		11
	Lower (Left, Center, Right)	24		35	15		37	26		36	16		36

comparisons remain valid. This study of medium-sized universities found that 91 percent of the universities had a link to the library (table 4). Studies by Agingu and King found similar percentages of those universities with links. In her examination of sixty-five Web sites of traditionally black colleges and other institutions in the southeastern U.S., Agingu reported that 97 percent had a link to the library home page. Ninety-three percent of the 111 ARL universities in King's study had a link.[21] Thus, the type of university surveyed seems to have little effect on the presence of a link to the library on the university home page. Studies that included ARL universities have reported 93 percent, 80 percent, and 76 percent having links to the library. These differences may be due in part to the way each link was counted. For example, in some studies, the criteria for counting a link were that it should be "obvious" in order to be counted as a link. Medium-sized universities also differed, with 91 percent, 71 percent, and 67 percent having links. Examining the data chronologically also shows no clear pattern of increasing or decreasing links. Earlier studies from 1996 and 1998 had good library representation (85 or 93 percent, respectively); while a study from 1999 had much lower library representation (58 percent). Three studies from 2000 and 2001 found links to the library on the home page from 68 to 97 percent of the time. When this study is compared with other 2005 studies, the library is on the university home page from 71 to 91 percent of the time.

User groups library link

Due to daily work and study needs, internal user groups were expected to have a link to the library more often than external user groups. Internal user groups showed a high

Table 3. External user groups: prospective students, alumni and donors, parents and family, visitors, and community

		Prospective students 1st occur *N*=61			**Alumni & donors 1st occur *N*=73**			**Parents & family 1st occur *N*=36**			**Visitors & community 1st occur *N*=38**		
Question	**Responses**	**set *n*=**	**Total *N*=**	**%**	**set *n*=**	**Total *N*=**	**%**	**set *n*=**	**Total *N*=**	**%**	**set *n*=**	**Total *N*=**	**%**
Is there a link to the library?	Yes	45	*61*	74	36	73	49	24	36	67	23	38	60
	No	16		26	37		51	12		33	15		40
How many links are there to the library?	1	34	45	76	26	36	72	17	24	71	17	23	74
	2	11		24	10		28	7		29	6		26
What term is used for the library link?	Libraries or Library	28	45	62	26	36	72	13	24	54	16	23	70
	Library or Libraries & "other"	4		9	4		11	2		8	5		22
	Library name	11		24	4		11	6		25	1		4
	University libraries	2		4	2		6	3		12	1		4
	Other	0		0	0		0	0		0	0		0
Where is the link?	Upper (Left, Center, Right)	22	45	49	23	36	64	10	24	42	14	23	61
	Center (Left, Center, Right)	10		22	4		11	5		21	3		13
	Lower (Left, Center, Right)	13		29	9		25	9		37	6		26

percentage of links to the library (Faculty and Staff: 94 percent; Students: 97 percent). External users had link-to-the-library frequencies ranging from 49 percent (Alumni and Donors) to 74 percent (Prospective Students). There was significant variation among the groups in the presence or absence of at least a single library link (X^2=20.29, d.f.=6, p<0.005); significantly more internal user group pages than external user group pages had at least one link to the library (X^2=19.48, d.f.=1, p<0.005).

In addition to the presence of a single link, internal user groups were also expected to have more additional links to the library. Additional links may appear due to follow-through menus or the deliberate addition of links because of perceived importance of these services by administrators or Web designers. In addition to the appearance of the first link, 50 percent of the Faculty and Staff and Students user groups had more than one link to the library. Internal pages had significantly more library links per page than external pages (t=8.6, d.f.=1, p<0.05).

Because of these results, comparison of the number of links on the university home page to those on internal and external user group pages was not surprising. There was no significant difference between the number of library links on university home pages and the number of links on internal user pages (p>0.05). However, there was a significant difference between the number of library links on university home pages compared to the number of library links on external user home pages (t=7.2, d.f.=1, p<0.05).

User groups: external

Little academic literature was available on college and university Web sites from the non-librarian perspective.[22] More quantitative approaches to university Web sites include literature on informetric studies. For example, a recent link metrics study examined the number and type of links on the university Web page in relation to the research at that university.[23] Internal users (Current students, Faculty and Staff) and one group of external users (Prospective students) were examined in the available research. No academic literature was available for any of the other external user groups.

In articles on prospective students and the university Web site, the library was not recommended or even mentioned as being an attribute to include on the Prospective Students Web page. In a study of fifty-five prospective students, participants were asked their opinions on content and design of the university Web site. Prospective students were most interested in information on admissions and environmental content, such as physical appearance of the campus itself, fellow students, and activities available. Poock and Lefond presented a table with the information that these prospective students expect to see when visiting a university Web site.[24] Fifty-nine answers included a variety of topics, such as information on admissions, athletics, course offerings, job opportunities, on-campus housing, majors and minors available, student social life, quick university facts, campus news and calendar, school colors, nickname and fight song, and campus location. The library was conspicuously absent from this list.

In an article about the use of search engines in a university's recruitment strategy, Whiteside and Mentz advocated the use of key departments and keywords that highlight benefits and successes that would attract students to attend the university.[25] They provided a partial list of suggested departments and terms to include for prospective students. Their list includes information on thirty items, such as accreditation, national rankings, admissions requirements, financial aid, enrollment size,

Table 4. University home page links to the library

Article author	Publication year	Institution type	*N*=	% with link
Agingu	2000	SE Black univ. & other inst.	65	97
Astroff	2001	ARL	111	76
Dewey	1999	Cmte on Institutional Cooperation	12	58
Harpel-Burke	2005	Med-sized univ (8–13K students)	77	91
King	1998	ARL	111	93
Stover & Zink*	1996	General higher education	40	85
Tolppanen, et. al	2000	Med-sized univ (6–13K students)	133	68
Tolppanen, et. al	2005	Med-sized univ (6–13K students)	133	71
Welch	2005	ARL & non-ARL	106	80

*Inverse; reports links from the library home page to the university home page

wireless capability, disability services, extracurricular activities, degree programs, advisement, athletics (and the name of the sport), housing, directions for visiting, and safety issues. Library resources and facilities were absent from this list as well.

User groups: internal

In these discussions of topics that Web-savvy prospective students look for on university Web pages, libraries may have been overlooked without good reason. In fact, Mechitov, Moshkovich, and Underwood performed an analysis of "student perceptions of academic Web sites and to determine the principal criteria that students use in forming positive and negative perceptions" for internal user groups.[26] Students evaluated aspects of the Web sites for attractiveness, information content, and entertainment. The evaluation indicated that ease of access to information and certain stylistic design issues (i.e., color coordination) were most important to the students. Analysis revealed that one of the four main reasons for visiting and revisiting university Web sites was for "additional information (e.g., on-line *library* and bookstore, research pages, writing center, job links and employment, search engines for student e-mail addresses)." Thus, the authors recommended that an effective university Web site should feature separate Web pages for all university services "(housing availability and policies, including photos of all dormitories, dining room services, *libraries*, career centers, bookstores, athletics)."[27] Therefore, it seems that links to the library should probably have been included regularly in user pages for prospective students.

The literature on faculty perceptions also included the library. Abels, White, and Hahn used focus groups and questionnaires to gather data from business faculty with Internet experience. An important component of the study was to "identify the criteria that users consider in using a Web site."[28] The researchers were interested in describing the faculty's information needs, search behaviors, and their Internet and Web use. Although the library was not specifically mentioned, most of the faculty indicated that locating literature was their greatest reason for using a Web site. This literature was described as "scholarly articles in business or professional journals."[29] Specifically, the resources that they focused on (electronic databases and full-text versions of these journals) are commonly available on the university library's Web site.

Thus, absence or presence of a link may be attributed to the universities' "dual roles of being public institutions while simultaneously having to act competitively and efficiently."[30] The university Web site's promotional functions suggest that a link to the library for recruitment purposes is not essential. However, further study in this area is warranted to determine if the presence or absence of a library link for external users has any impact on recruitment, donations, or community relations. For internal users, the importance of a link to the library was clearly recognized. Not addressed here is the issue of internal users who fail to recognize the Web site as a resource for fulfilling their information needs.

Terminology

In general, links to the library were labeled with the word Library or a phrase containing the word. Terminology was fairly uniform throughout all user groups and on the home page. In only a few cases were any other terms used to identify the library, such as Reference Desk for Faculty and Staff. Because the terminology used was not varied, no statistical analyses were performed. In general Web design principles, clever phrases and jargon are discouraged; informative and unambiguous terms are recommended. Nielsen noted that vague terminology causes users to puzzle over meanings and may serve to alienate them.[31] Use of easily understood terms also makes sense from a recruitment or marketing standpoint. Whiteside and Mentz note that in order to highlight the benefits derived from attending a particular institution, "colleges and universities must target the search engine terms . . . or keywords that Internet users are likely to use."[32]

Library link placement

Web site usability research indicates that the most sought-after information should be located in the upper part of the page in either of the top corners.[33] Findings in the Poynter Institute's Eyetracking III study, which examined how Internet users read newspapers online, also indicated that eyes tend to gravitate to the upper left corner of a Web page.[34] Because it is unclear if the library is the most sought-after information, there were no expectations associated with the placement of the library link.

Spatial placement of the first library link on the university home page and subsequent user groups was quite varied. On the home page, the first link to the library was mainly in the center of the page (47 percent). However, the upper right corner was also used often (21 percent).

For the user group pages, the internal user group pages were fairly evenly divided among the upper, center, and lower part of the page. Faculty and Staff user groups had the most links divided between the upper (35 percent) and lower part of the page (35 percent), with 29 percent in the center of the page. Student user groups had most links in the center (37 percent) and lower parts (36 percent) of the page, with 27 percent in the upper part of the page.

All of the external groups had the most links to the library in the upper part of the page. Alumni and Donors placed the link there 64 percent of the time, followed by

Visitors and Community (61 percent), Prospective students (49 percent) and Parents and Family (42 percent). There is significant variation in link placement among the user groups (X^2=34.20, d.f.=12, p<0.005).

Thus, relatively few university pages placed the link to the library in the area recommended for high-profile links. Links for internal users also had no consistently dominant position. However, links to the library for external user group pages were most often in the upper portion.

Conclusion

With 91 percent of the medium-sized universities in this study having links to the library on their home page, library representation on university home pages is quite good for the general user. Similarly, library representation on both types of internal user group pages—Faculty and Staff and Current Students—was also very high. Although a link to the library does not necessarily indicate the value that the university places on the library, these high percentages suggest that the library is important to the people responsible for university Web pages. Although library representation on external user group pages was lower (both in number of library links and prominence of those links), these groups still had a substantial number of library links.

General studies on the importance and promotion of the library to these user groups would seem valuable in determining the need, availability, and placement of links for these groups. Usability studies exploring the ways that these external user groups make use of the library via the Web site may also have many potential benefits.

Individual library Web pages targeted to a specific user group could be specially designed to showcase services for that group.

As noted by Jafari, "[m]ixing the information categories needed by both current and prospective members on a single Web site results in a compound design solution that does not serve either group well."[35] This concept could be further developed for the library Web page; each of the separate user groups could find the information about the library that they value most.

For example, prospective students and their parents may appreciate pages that would give them information that they can use in making a university choice. For these two external user groups, library pages could be designed that highlight information on collection age and size, electronic resources, assistance by library staff, and special programs such as laptop borrowing.

A library Web page designed for donors could also emphasize various services that could encourage financial support. These pages could give guidance on gifts of various amounts by suggesting specific books (with donor plates) or perhaps showcasing previous larger donations that resulted in a building, room, or a collection named in honor of the donor.

By contrast, internal user groups are more interested in Web pages that highlight specific services associated with their daily activities. Current faculty and staff may want to make suggestions for acquisitions, or place interlibrary loans. Current students may visit the library Web page to use databases, check the availability of study areas, or locate spots for wireless connections. Kvarik and Handberg studied the changes in student services in general.[36] They noted that students are now determining the flow, timing, and format of information that they need. For students, the library Web page could emphasize 24/7 access to the library, electronic formats, and remote access to databases.

Thus, further study on the overall importance of the link to the library for these various user groups is warranted. Specific content of internal and external user group pages also merits further investigation. Then the feasibility and necessity of individual Web pages or portals to the library for each user group could be considered.

References and notes

1. Roberta J. Astroff, "Searching for the Library: University Home Page Design and Missing Links," *Information Technology and Libraries* 20, no. 2 (June 2001): 93–99; Iain Middleton, Mike McConnell, and Grant Davidson, "Presenting a Model for the Structure and Content of a University World Wide Web Site," *Journal of Information Science* 25, no. 3 (1999): 219–27; Michael Poock and Dennis Lefond, "How College-bound Prospects Perceive University Web Sites: Findings, Implications, and Turning Browsers into Applicants," *College & University* (Summer 2001): 15–21.

2. Andrea Schwandt-Arbogast, "Why Most University Web Sites Suck, Part 1," *Interllectual* (Feb. 2005). Accessed June 5, 2005, http://interllectual.com/article/84/why-most-university-web-sites-suck-part-1.

3. Patrick Lynch and Sarah Horton, *Web Style Guide: Basic Design Principles for Creating Web Sites,* 2nd ed. (New Haven, Conn.: Yale Univ. Pr., 2001). (Also available online. Accessed Aug. 15, 2005, www.webstyleguide.com/index.html).

4. Lynch and Horton, *Web Style Guide*; Jakob Nielsen, *Designing Web Usability* (Indianapolis, Ind.: New Riders Pubs., 2001).

5. Nielsen, *Designing Web Usability*.

6. Middleton, McConnell, and Davidson, "Presenting a Model."

7. Jakob Nielsen and Marie Tahir, *Home Page Usability: 50 Websites Deconstructed*, (Indianapolis, Ind.: New Riders Pubs., 2002).

8. Astroff, "Searching for the Library," 95.

9. David L. King, "Library Home Page Design: A Comparison of Page Layout for Front Ends to ARL Library Web Sites," *College & Research Libraries* 59, no. 4 (1998): 457–64.

10. Barbara Dewey, "In Search of Services: Analyzing the Findability of Links on CIC University Libraries' Web Pages," *Information Technology and Libraries* 18, no.4 (1999): 210–13.

11. Bradley P. Tolppanen, Joan Miller, and Martha H. Wooden, "An Examination of Library World Wide Web Sites at Medium-Sized Universities," *Internet Reference Services Quarterly* 5, no. 2 (2000): 5–17; Bradley P. Tolppanen, Joan Miller, Martha H. Wooden, and Lori M. Tolppanen, "Library World Wide Web Sites at Medium-Sized Universities: A Re-Examination," *Internet Reference Services Quarterly* 10, no.2 (2005): 9–20.

12. Tolppanen, Miller, and Wooden, "An Examination of Library World Wide Web Sites at Medium-Sized Universities," 12.

13. Beatrice O. Agingu, "Library Web Sites at Historically Black Colleges and Universities," *College and Research Libraries* 61, no.1 (Jan. 2000): 30–37.

14. Jeanie M. Welch, "The Electronic Welcome Mat: The Academic Library Web Site as a Marketing and Public Relations Tool," *Journal of Academic Librarianship* 31, no.3 (2005): 225–28.

15. Mark Stover and Steven Zink, "World Wide Web Home Page Design: Patterns and Anomalies of Higher Education Library Home Pages," *Reference Services Review* 24, no. 3 (Fall 1996): 7–20.

16. Astroff, "Searching for the Library," 95.

17. Daniel Xiao, Pixey Anne Mosley, and Alan Cornish, "Library Services through the World Wide Web," *The Public-Access Computer Systems Review* 8, no. 4 (1997). Accessed July 31, 2005, http://info.lib.uh.edu/pr/v8/n4/xiao8n4.html.

18. Astroff, "Searching for the Library," 96.

19. Hofstra University, "At a Glance," (Jan. 2005)." Accessed Apr. 30, 2005, www.hofstra.edu/Admissions/adm_hofstra_summary.cfm.

20. National Center for Education Statistics, "Library Statistics Program: Compare Academic Libraries" (uses data from the *American Libraries Survey* for fiscal year 2002), (Washington, D.C.: Institute of Education Sciences, U.S. Dept. of Education, 2002). Accessed Apr. 29, 2005, http://nces.ed.gov/surveys/libraries/compare/index.asp?LibraryType=Academic; *Peterson's College Bound Online,* (2005). Accessed Apr. 15, 2005, www.petersons.com/ugchannel.

21. Agingu, "Library Web Sites at Historically Black Colleges and Universities"; King, "Library Home Page Design."

22. Alexander I. Mechitov, et al., "Comparative Analysis of Academic Web Sites," *Education* 121, no. 4 (2001): 652–64.

23. Mike Thelwall and Gareth Harries, "The Connection Between the Research of a University and Counts of Links to Its Web Pages: An Investigation Based Upon a Classification of the Relationships of Pages to the Research of the Host University," *Journal of the American Society for Information Science and Technology* 54, no. 7 (2003): 594–602.

24. Poock and Lefond, "How College-bound Prospects Perceive University Web Sites."

25. Richard Whiteside and George S. Mentz, "Online Admissions and Internet Recruiting: An Anatomy of Search Engine Placement," *Educause Quarterly* no. 4 (2003): 63–66.

26. Mechitov et al., "Comparative Analysis of Academic Web Sites," 652.

27. Ibid., 658. *Italics* added by author.

28. Eileen G. Abels, Marilyn Domas White, and Karla Hahn, "Identifying User-based Criteria for Web Pages," *Internet Research: Electronic Networking Applications and Policy* 7, no. 4 (1997): 252–62.

29. Ibid., 257.

30. Mechitov et al., "Comparative Analysis of Academic Web Sites," 653.

31. Nielsen, *Designing Web Usability*.

32. Whiteside and Mentz, "Online Admissions and Internet Recruiting," 65.

33. Nielsen and Tahir, *Home Page Usability: 50 Websites Deconstructed*.

34. Steve Outing and Laura Ruel, *Eyetrack III: Online News Consumer Behavior in the Age of Multimedia* (St. Petersburg, Fla.: Poytner Institute, Estlow Center for Journalism and New Media, Eyetools, 2004). Accessed Apr. 29, 2005, www.poynterextra.org/eyetrack2004/index.htm.

35. Ali Jafari, "Optimizing Campus Web Sites: Is the Portal Approach a Solution to Improving Campus Web Usability," *Educause Quarterly* no. 2 (2000): 56.

36. Robert B. Kvarik and Michael N. Handberg, "Transforming Student Services," *Educause Quarterly* no. 2, (2000): 30–37.

The State of RFID Applications in Libraries

Jay Singh, Navjit Brar, and Carmen Fong

The adoption of Radio Frequency Identification (RFID) technology by libraries promises a solution that could make it possible to inventory hundreds of thousands of items in their collections in days instead of months. In addition, it would allow patrons to check out and return library property automatically at any time of the day. Besides speeding up checkouts, keeping collections in better order, and alleviating repetitive strain injuries among librarians, RFID promises to provide a better control on theft, nonreturns, and misfiling of a library's assets. With an estimated 35 million library items tagged worldwide in more than three hundred libraries, this technology is generating ever-increasing interest. In October and November 2004, the industrial technology department and the Robert E. Kennedy Library at Cal Poly State University, San Luis Obispo, surveyed participating libraries, RFID electronic discussion groups, and Library and Information Technology Association (LITA-L) electronic discussion group subscribers to collect information with regards to the implementation of RFID systems in libraries. Opinions were gathered regarding such topics, actual or estimated, as RFID implementation costs and time; the impact of the technology on operations such as handling of volumes and security; and RFID system features adopted such as conversion stations, self-checkout units, and security systems. Information on the various RFID library components and the results from the survey are presented in this paper.

As libraries adopt the use of tiny radio frequency identification (RFID) tags for tracking their assets, it is becoming possible to inventory hundreds of thousands of items in their collections in days instead of months and to allow patrons to checkout and return library property automatically at any time of day. In 1998, RFID was proposed as a way to make possible the self-serve processing of books and media by patrons in North America. It was initially installed in the library of Rockefeller University in New York in 1999. The first public library to use the technology was the Farmington Community Library in Michigan, also in 1999.[1] Since then, more than three hundred libraries have, or are in the process of implementing, an RFID system.[2] According to the research service *RFID Knowledgebase,* United States libraries lead the world in RFID use, with the United Kingdom and Japan tied for second place. It estimates that 35 million library items have been tagged worldwide.[3] Even the high-profile Vatican library in Rome has recently started tagging 120,000 volumes of its collection.[4] It is estimated that it will take the administrators at the Vatican library only half a day to inventory these tagged volumes as compared to the one month required prior to tagging. Tagging two million of the Vatican's forty-million-piece collection during the next few years is also planned.

Besides speeding up checkouts, keeping collections in better order, and alleviating repetitive strain injuries among librarians, RFID promises to provide better control against theft, nonreturns, and misfiling of a library's assets. San Francisco Public Library, after consulting with experts and other libraries, expects RFID to help reduce workplace injuries that cost it $265,000 in workers' compensation claims over the past three years.[5] At its present stage of development, RFID technology for libraries faces a number of unresolved issues. Among the most significant are privacy, lack of standardization, and cost. The Infopeople Project and the Information Technology Section of the California Library Association (CLA) conducted a recent survey. Of the approximately 51 percent of libraries who participated in the survey and were not considering implementing RFID, 58 percent cited cost as the major issue and 9 percent listed privacy concerns.[6] Molnar and Wagner have discovered several serious vulnerabilities related to a library patron's privacy, including a lack of appropriate access control, the shortcomings in the collision-avoidance protocols in place today, and poor key management practices.[7]

From October 13, 2004, to November 1, 2004, the industrial technology department and the Robert E. Kennedy Library at California Polytechnic State University, San Luis Obispo (Cal Poly), surveyed participating libraries, RFID electronic discussion groups, and Library and Information Technology Association electronic discussion group (LITA-L) subscribers to collect information regarding the implementation of RFID systems in libraries. The need to conduct a survey arose from the lack of information regarding the present state of RFID applications in libraries, and because the authors felt it was necessary to have the opinion of those directly affected by this technology. Members and subscribers were given a link directing them to an Internet-based survey. Twenty-nine libraries participated and completed the survey; of these, ten had already converted to an RFID-based system at their location. The survey provided an insight into library RFID systems. Opinions were gathered regarding such topics as RFID implementation costs and time; the impact of the

Jay Singh (jasingh@calpoly.edu) is Packaging Program Director, Industrial Technology Department; **Navjit Brar** (nbrar@calpoly.edu) is Assistant Dean for Access and Bibliography Services, Robert E. Kennedy Library; and **Carmen Fong** (kfong@calpoly.edu) is a Graduate Assistant, Industrial Technology Department, Cal Poly State University, San Luis Obispo, California.

technology on operations such as handling of volumes and security; and RFID system features adopted (e.g., conversion stations, self-checkout units, and security systems). Results from this survey are presented at the end of this paper.

RFID basics

An RFID system (figure 1) consists of three parts:

- A scanning device
- A transceiver with a decoder to interpret the data
- A transponder—the radio-frequency tag—that has been programmed with information

A scanning antenna puts out radio-frequency (RF) signals in a relatively short range. The RF radiation does two things; it provides a means of communicating with the transponder tag (the RFID chip) and it provides the RFID device with the energy to communicate. RFID tags for library applications do not need to contain batteries and can therefore remain usable for very long periods of time. Scanning antennas can be permanently affixed to a surface or can be in handheld form. An institution could build them into a doorframe to accept data from persons or objects (or in the case of libraries, books) passing through. Various RFID system components for library use will be discussed in this paper.

When an RFID tag passes through the field of the scanning antenna, it detects the activation signal from the antenna. This "wakes up" the RFID tag, and it transmits the information on its microchip to be picked up by the scanning antenna. The RFID tag may be one of two types. Active RFID tags have their own power source, the advantage of which is that the reader can be much farther away and still receive the signal. Although some tags are built to have up to a ten-year life span, their actual life span is more limited. Passive RFID tags, however, do not require batteries, can be much smaller, and have a virtually unlimited life span. Due to their lower costs and their ability to carry large amounts of information, passive RFID tags are being used for library applications.

Figure 2 shows a typical RFID tag to be applied to the books at a library. Besides carrying a unique ID number programmed by the library, these tags also carry a security bit that is activated when an item is returned to the library, and inactivated upon proper checkout. These actions are performed during a normal checkout or return operation and do not require a separate step as with magnetic-strip-type security devices, thereby combining both inventory and security features. Figure 3 depicts a typical flow of information between the items—whether books or audio-video media—and the integrated library system (ILS).

RFID tags can be read in a wide variety of circumstances in which barcodes or other optically read technologies are useless. The tag need not be on the surface of the object (and is therefore not subject to wear), the

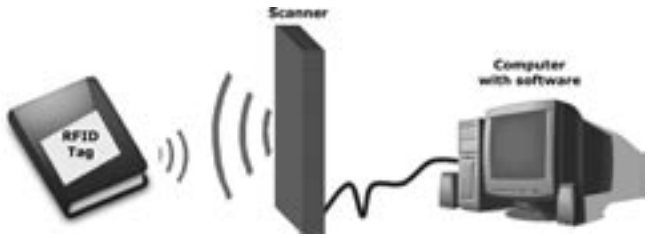


Figure 1. Components of an RFID system

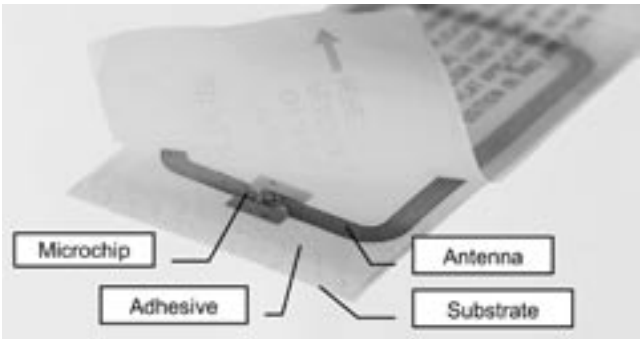


Figure 2. Typical RFID tag

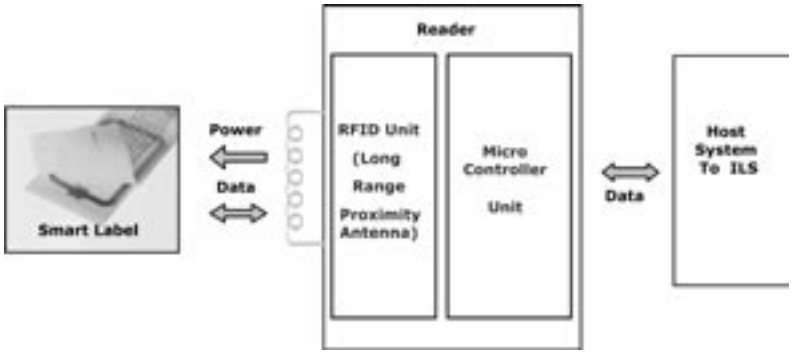


Figure 3. Information exchange for a typical RFID-based library (source: Libramation Library Systems)

read time is typically less than 100 milliseconds, and large numbers of tags can, in theory, be read at once, rather than item by item. Finally, unlike barcodes, RFID tags can be read even when they are not directly within the scanner's line of sight.[8]

RFID library system

There are various options that a library must consider before converting to RFID. Libramation Library Systems, an RFID library solutions vendor, participated in a separate study involving simulation of a complete RFID-based system at Cal Poly in May 2005. The system used 1024-bit (of which 896 bits are available for programming), 60 p/sec (gap pulse) anti-collision tags with Electronic Article Surveillance (EAS). The system operated at a frequency of 13.56 MHz and conformed to ISO 15693/18000 Standard. Various components associated with a library implementation were demonstrated.

Tag programming station

A programming (and verification) station is used to link the programmable code on the tag affixed to a volume to the existing bar code used for identifying it. This operation is very short and easy to perform. A rate of approximately 250 conversions per hour was achieved in trials at the Robert E. Kennedy Library. Figure 4 shows a typical programming station.

Self-checkout and check-in stations

These user-friendly stations are employed by patrons for checking out or returning items at the library. Easy-to-understand instructions in any language are possible to assist or prompt the patron through the process. Figure 5 depicts a typical self-serve station. These stations can be freestanding or counter-top, are adaptable to any existing form of patron membership card, and can be customized to the library's needs. Return stations can be placed at a convenient location outside the library for after-hours operations. Printed receipts confirming an interaction may also be issued. Based upon the operation, these stations are also capable of activating or deactivating the security bit in the tags. A checkout operation, for example, would require that the bit be deactivated in order for the patron to leave the premises without triggering an alarm.

Staff circulation desk

When used at library-staffed circulation desks, these ergonomically designed units have the capacity to reduce the number of movements by a staff member

Figure 4. Tag-programming station

Figure 5. Patron self-service station

Figure 6. Hybrid circulation desk

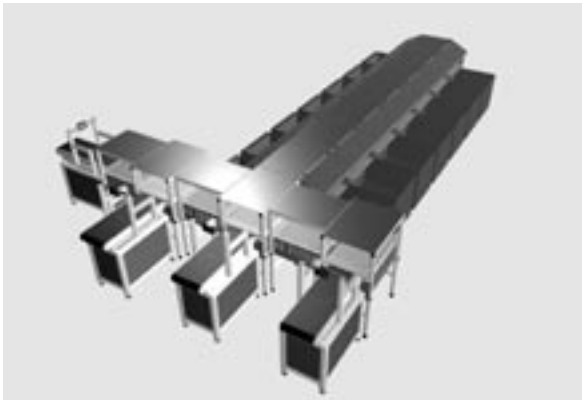


Figure 7. Sorting station

for a typical task from approximately eight to ten (barcode/magnetic strip) to two to three (RFID) per item.[9] As with the self-serve stations, the units can be customized to the specific requirements of a library. Hybrid desks such as the one shown in figure 6 may be used to perform scanning of a barcode/magnetic strip and an RFID-based operation. Hybrid circulation desks permit patrons to scan their own library cards in order to initiate the check-out process. The staff member is able to view patron status on a flat screen monitor. She or he then slides each item by hand along the desk, where it passes over a built-in or top-mounted barcode scanner. An acoustic signal indicates that the scanner has successfully identified the item, circulation records are automatically updated, and the scanned materials are desensitized and passed to the client area of the desk for retrieval by the patron. To complete the process, the staff member presses a receipt key on a mini-keyboard to produce a printed receipt with due date.

Sorting stations

Most of the usual tedious and repetitive handling of material is eliminated with these stations since the sorting unit can separate library materials into individual material bins. This system can be modular and expandable and can greatly enhance the sorting process. The security bit on the tags is automatically activated upon sorting. Figure 7 shows one such unit marketed by Libramation Library Systems.

Security gates

EAS is an anti-theft system used by libraries. Pre-RFID systems utilized metallic security strips that needed to be desensitized—a separate action—during a normal checkout in order for patrons to exit the library without setting off an alarm. This security feature is incorporated by RFID technology in the tag itself and is much more conveniently activated or deactivated during the check-in or checkout operation. Security gates, which play a crucial role in preventing loss of library properties, typically set off an alarm when an unauthorized or improperly checked-out item is detected. A customized security gate by Libramation is shown in figure 8.

Portable readers

The promise of RFID is to dramatically reduce the time and effort of taking inventory. A handheld reader makes quick and easy work of shelf-management activities like weeding or finding missing and out-of-order books. In a vendor white paper supplied to the authors, VTLS, an RFID-solution vendor, claims to be able to inventory a collection of 250,000 books in less than four hours as compared to the thirty-five hours needed by the 3M portable reader.[10] It also claims that VTLS's portable readers can find one hundred books in a ten-thousand-book collection in approximately eight minutes, as compared to eighty-three minutes by 3M. The read speed claimed by VTLS is twenty books per second as compared with two for 3M. Figure 9 shows a Libramation Library Systems portable reader.

RFID library survey

This section summarizes the responses of twenty-nine libraries that participated in the survey conducted by the authors. Out of these twenty-nine, twenty-six public libraries, two academic libraries, and a corporate library were represented. The average number of hours of operation for all participating libraries was 104 hours per month, with an average seating capacity of 370. Eighty-nine percent of all survey participants reported less than

Figure 8. Security gate

Figure 9. Libramation's Lib-Chip handheld reader

or equal to 2 million user visits every year. The average collection size for the twenty-nine libraries was 1.34 million; the average circulation was 1.24 million per year. The participating libraries had an average interlibrary loan transaction of 760 per year. Because they address both security and inventory issues, interlibrary loan transactions play an important role when considering implementation of RFID technology in a consortial environment using a shared library management system. Library assets should be able to be processed by any member of the consortium in such an application.[11] The public libraries of Fresno County, California, serve as a good example of a consortia. They have thirty-four branches and started converting two of the branches to RFID last year.[12]

Migration to an RFID-based library system

Of the twenty-nine respondents, ten libraries had migrated to an RFID-based system and nineteen libraries had not. Out of the nineteen libraries that had not migrated, thirteen were in the process of migrating to RFID, four were considering migration, and two had no current plans for migration to an RFID-based library system.

According to the previously referenced Infopeople survey, of the approximately 51 percent of libraries that participated in their survey and were not considering implementing RFID, 58 percent expressed cost as the major issue and 9 percent listed privacy concerns.[13] Forty-eight percent of the participating libraries were considering RFID; these institutions suggested that the evaluation processes would be aided by: information from other libraries that had already implemented RFID systems (ranked the highest); a checklist of key questions to consider; RFID vendor product information; and additional information about privacy, security, and health concerns.

Conversion costs and time

Twelve quantitative responses were received for the Cal Poly survey. The conversion cost (actual or estimated) ranged from $113,000 to $1.2 million, with the average conversion cost being $502,917. These costs are difficult to gauge due to the number of variables, such as partial or complete conversion of collections, paid versus volunteer employee time, and the number of RFID system components such as number of self-checkout stations.

According to Vinod Chachra, CEO of VTLS, a full RFID implementation, including tags for books, readers, software, and self-checkout stations, runs around $1 per book. In comparison, the same implementation would have cost approximately $1.60 per book a year ago.[14] According to Boss, the cost of implementing an RFID system for a collection of forty thousand items would be $70,000; for a collection of one hundred thousand items, it would be $166,000.[15] Because his estimates were made in late 2003, they may be considerably high compared to the lower costs available now. Boss also provides estimates for individual RFID components in

his paper. Smart points out several main vendors that have led the market in providing libraries with equipment and training to converting their libraries to an RFID system.[16] She also provides ballpark figures of these vendors' charges to convert a mid-sized library of two hundred thousand items including two hundred thousand tags, one self-checkout, one self-return, and one entrance gate: 3M Library Systems is estimated to charge $175,000 to $275,000; Bibliotheca approximately $180,000; Libramation Library Systems approximately $125,000 to $160,000; Vernon Library Supplies $150,000; and Checkpoint between $100,000 and $130,000.

Conversion times (concluded or estimated) for sixteen libraries that have migrated or plan to migrate to RFID vary from one library to another. Three of those libraries took less than six months, eight libraries took or estimate that it will take twelve months, and the remaining three took eighteen months to switch to RFID. Two libraries did not have a definitive answer. The average conversion time was 11.18 months.

Boss estimates a typical training time of fifteen to twenty minutes for teaching a person to tag books.[17] A phased implementation is also recommended. According to Frank Mussche, president of Libramation, a good strategy for libraries not willing to convert their entire collections at once is to first tag their most circulated items instead of slowly phasing in the entire collection to an RFID-based management system.[18] Backroom stations and even the circulation desks can be used to convert any untagged items as they are returned to the library. This strategy allows a library to move from a barcode-only system to one that also has RFID capabilities without requiring a greater amount of resources than would otherwise be needed.

Staffing since RFID migration

Most of the responding libraries showed little to no changes in the number of employees. This may be attributed to recent migrations, expansion of facilities, or staff being reassigned to other duties. Three libraries, however, showed a significant difference in the number of employees prior to and after RFID adoption. All three attributed this to a facility expansion or a move to a new facility.

Due to the automation of circulation and inventorying of a library's collection by an RFID-based system, the employees who typically perform such tasks have good reason to feel the threat of getting laid off. San Francisco and Berkeley Public Libraries in California are presently dealing with such issues.[19]

Circulation time since RFID implementation

There were ten responses for this question. Fifty percent of the participating libraries showed more than 10 percent increase in usage, as per the circulation stats, since RFID implementation. The responding libraries commented on a significant reduction in both check-in and checkout times and processes. Three of them provided quantitative responses that are presented in figure 10.

According to Checkpoint, an RFID vendor that claims that the system processes twenty items per second, libraries can realize as much as 75 percent reduction in handling time after installing their Intelligent Library System.[20] One key benefit of an RFID-based system as far as circulation is concerned is that it combines it with EAS. The security bit in a tag replaces the security strip needed with the barcode approach, and items can be moved in or out of the premises without performing two separate steps. The Mastics-Moriches Community Library in New York conducted a time and cost analysis to compare EAS and RFID systems for circulation and found that RFID gave a savings in labor time of up to 85 percent.[21] Figure 11 shows the results of the study.

Self-checkout rates tend to increase with easy-to-use circulation systems. Santa Clara Public Libraries, which in 2000 became the first system on the West Coast to adopt an RFID-based system, has achieved a self-checkout rate of 45 percent.[22] According to Margaret Hazel, Oregon Public Library's technology manager, the library has achieved a near 100 percent self-checkout rate.[23] Of note here is the fact that the Eugene Public Library has tagged all of its collection while the Santa Clara Public Libraries are still to tag such items as CDs and DVDs. Jackie Griffin,

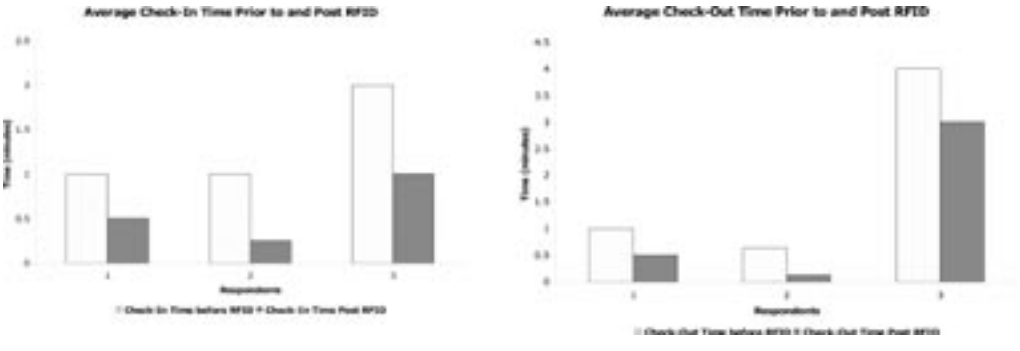


Figure 10. Circulation time since conversion

the library director at the Berkeley Public Library, estimates that self-checkouts at the library would jump from 15 to 90 percent after an RFID installation.[24] The Eugene Public Library reported "collision" problems on very thin materials and on videos as well as false readings from the RFID security gates. Collision problems mean that two or more tags are close enough to cancel the signals, making them undetectable by the RFID checkout and security systems.

RFID system features

An RFID system offers a variety of options. The various components needed to create a functional RFID library system depend on the unique requirements of individual libraries. Table 1 shows responses to the query regarding the various RFID system features, existing or sought, by the ten libraries that had already converted to RFID.

As can be seen, the patron self-checkout and conversion stations were deemed the most important RFID system feature, followed by security systems, staff service stations, inventory systems, and automatic book return systems, in that order. Of note is the fact that all of the responding libraries wanted to adopt the self-checkout units. This indicates that most of the libraries wanted to free up staff time from the repetitive task of circulating items. An automatic book return system, which can cost hundreds of thousands of dollars, was at the bottom of this list, most likely for financial reasons.

Figure 12 shows the number of self-checkout stations either existing or sought by the thirteen respondents to the question in the survey. They ranged from one to nine and averaged almost four per facility.

Security features

Two of the nine libraries that responded were using Checkpoint's RF system and seven were using 3M's Tattle Tape. The number of replacements of the pre-RFID security systems varied from twenty-five thousand to four hundred thousand.

Boss states that libraries with a properly tuned RFID-based security system reported a 50 to 75 percent reduction in false alarms at exit gates as compared to older RF or electromagnetic (EM) strip systems.[25] The RFID-based security process is automatic when an item is checked out at the circulation desk or at the self-checkout station, therefore eliminating the extra processing time of the staff or patron. Moreover, according to Bibliotheca, the accuracy rate of detecting a stolen item is very high, ranging from an average of 65 percent on conventional EM detection systems to almost 100 percent with an RFID system.[26] Bibliotheca also claims that staff cannot inadvertently damage audio-visual materials when desensitizing them. When a patron attempts to remove an item from the library without authorization, it is easily identified at the exit sensors by ID number, title, and author.

Conclusion

RFID technology has several advantages over the current barcode systems being used at libraries worldwide. For the libraries that use it, RFID promises to save time and operate more efficiently and effectively than the barcode systems. Some of the compensations of RFID over a barcode system are that RFID tags can be used for security as well as for status control, thereby eliminating the need to attach security strips to library items; RFID systems make self-checkout faster and easier for library patrons; and RFID portable scanners can take inventory by just being passed slowly along the library shelves, without having to handle each item individually. RFID vendors, however, need to resolve some issues before libraries feel confident in adopting them. In the forefront are issues such as cost, lack of standardization amongst vendors, and privacy.

This survey investigated the state of RFID applications in libraries. Of the twenty-nine survey respondents, nineteen had not migrated to an RFID-based system. Of those nineteen, thirteen were in the process of migrating to RFID. This meant that nearly 80 percent of the libraries had or were seriously contemplating adopting an RFID-based library management system. For a library

Table 1. RFID system features existing or sought

	Patron self-checkout	Staff service stations	Inventory system	Automatic book return	Security system	Conversion stations
Yes	10	8	7	4	9	10
No	0	4	5	7	2	1
Had before	1	0	0	0	1	0
In progress	1	0	0	1	0	1

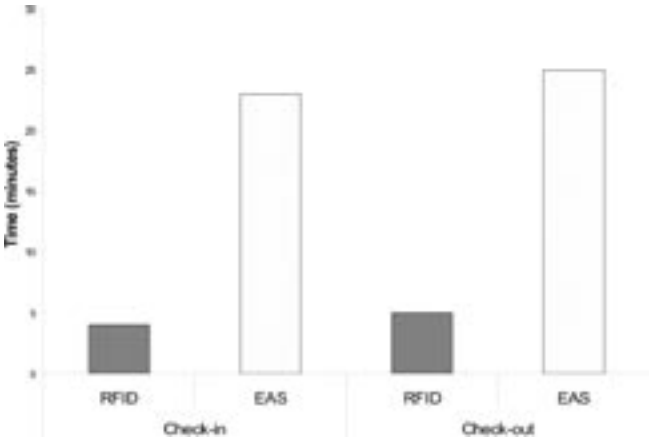


Figure 11. Time analysis of media circulation (source: www .researchinformation.info/rimayjun04radiotagged.html)

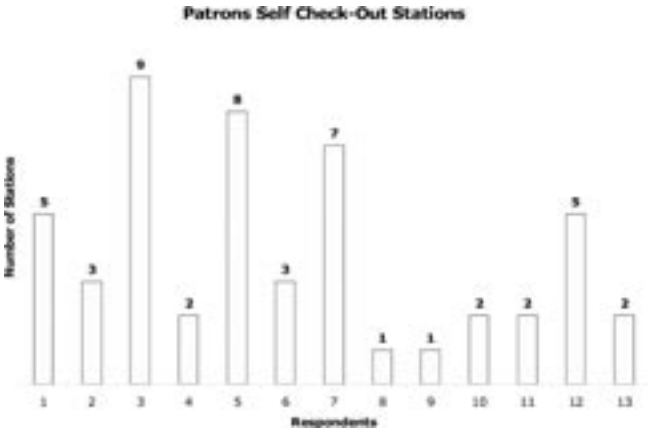


Figure 12. Patron self-checkout stations existing or sought

collection size of 1.34 million items, the average conversion cost to migrate to an RFID system was $502,917. The average conversion time from a barcode-based system to an RFID-enabled system was 11.18 months. Based on ten responses, 50 percent of the participating libraries showed a greater than 10 percent increase in usage after RFID implementation. Average times for check-in and checkout were significantly reduced after implementation. The two most-installed or sought-after features were the patron self-checkout and the conversion stations. Seventy-eight percent of the libraries that responded used 3M's Tattle Tape system as a security feature, and the number of replacements of the pre-RFID security systems varied from twenty-five thousand to four hundred thousand.

Even though RFID tags may cost anywhere from four to five times as much as a barcode and magnetic strip combined, the increase in efficiency and functionality provided by the technology is persuasive enough for an increasing number of libraries to implement or consider this technology.[27] The number of RFID-enabled library systems is constantly increasing. International collaborations in such ventures are also on the rise. In a March 11, 2005, news release, the French RFID technology firm ASK was set to deliver some four million RFID labels to U.S. and European libraries. ASK has already supplied labels and readers to more than twenty French, Italian, and German libraries during the past four years.[28]

References

1. Laura Smart, "The Market Place; Laura Smart Gives a Guide to the Leading Providers of Integrated RFID Solutions," Oct. 15, 2004. Accessed Nov. 16, 2004, http://mythic.lib.calpoly .edu:2067/universe/document?_m=aa31c69ea270d87b08.

2. Gilbert Alorie, "RFID, Coming to a Library Near You," Oct. 18, 2004, CNET News.com, http://msn-cnet.com.com/2100-1012_3-5411657.html (accessed Oct. 28, 2004).

3. Ken Young, "Using RFID Tags to Track Library Books Could Increase Security and Ease Laborious Stocktakes," *The Guardian (London)* Nov. 11, 2004. Accessed Nov. 22, 2004, www.guardian .co.uk/online/story/0,3605,1347657,00.html.

4. Jan Libbenga, "Vatican Library Adopts RFID," July 9, 2004. Accessed Nov. 29, 2004, www.theregister.co.uk/2004/07/09/vatican_library_rfid.

5. Alorie, "RFID, Coming to a Library Near You."

6. California Library RFID Survey, Aug./Sept. 2004. Accessed Feb. 7, 2005, http://infopeople.org/resources/rfid_survey.

7. D. Molnar and D. A. Wagner, "Privacy and Security in Library RFID Issues, Practices, and Architecture," Oct. 2004. Accessed Nov. 29, 2004, www.cs.berkeley.edu/~dmolnar/library .pdf.

8. "How Does RFID Work?" Accessed Apr. 26, 2005, www .technovelgy.com/ct/Technology-Article.asp?ArtNum=2.

9. Frank Mussche, president, Libramation Library Systems, personal communication, May 10, 2005.

10. VTLS, "Choosing Your RFID Library Solution," undated white paper.

11. David Dorman, "Implementing RFID Technology in a Consortial Environment Using a Shared Library Management System," Aug. 25, 2003. Accessed Oct. 25, 2004, www.lincoln trail.info/RFIDInConsortialEnvironment.html.

12. Kathleen Smith, Projects Librarian, Fresno County Public Library, personal communication, Apr. 12, 2005.

13. California Library RFID Survey.

14. Alorie, "RFID, Coming to a Library Near You."

15. Richard Boss, "RFID Technology for Libraries," May 14, 2004. Accessed July 23, 2005, www.ala.org/ala/pla/plapubs/technotes/rfidtechnology.htm.

16. Smart, "The Market Place."

17. California Library RFID Survey.
18. Frank Mussche, president, Libramation Library Systems, personal communication, Feb. 9, 2005.
19. Matthew Artz, "RFID: Library's New Technology Sparks Controversy," Berkeley Daily Planet. Accessed Mar. 4, 2005, www.berkeleydaily.org/article.cfm?issue=02-15-05&storyID=20728.
20. "Checkpoint Systems Grows RFID Library Market Share," Aug. 19, 2004, TMCnet.com. Accessed Sept. 30, 2004, www.tmcnet.com/usubmit/2004/Aug/1066833.htm.]
21. "Radio Tagged Books: Why Librarians Should Switch on Their Radios," Research Information Web site. Accessed Apr. 25, 2005, www.researchinformation.info/rimayjun04radiotagged.html.
22. Artz, "RFID: Library's New Technology Sparks Controversy."
23. Ibid.
24. Ibid.
25. Boss, "RFID Technology for Libraries."
26. "Frequently Asked Questions," BIBLIOTHECA RFID Library Systems AG Web site, www.bibliotheca-rfid.com/index.php?nav=46,47 (accessed July 23, 2005).
27. Dorman, "Implementing RFID Technology in a Consortial Environment."
28. "U.S. and European Libraries Order 4 Million RFID Labels," USINGRFID.com Web site. Accessed Mar. 11, 2005. www.usingrfid.com/news/read.asp?lc=n12239bx378zt.

WikiWikiWebs: New Ways to Communicate in a Web Environment

Brenda Chawner and Paul H. Lewis

This paper introduces WikiWikiWeb software, also known as Wiki, for use in library and information management contexts. Wikis provide an environment for Web-based collaboration and can also be used for Web site content management. The article includes an overview of the history and development of Wiki, as well as discussing basic and advanced Wiki features. It compares three Wiki engines and describes seven case studies of real-world library and library-related Wiki applications. The paper concludes with a discussion of factors that can contribute to a successful Wiki project.

Tim Berners-Lee originally saw the Web as "a system in which sharing what you knew or thought should be as easy as learning what someone else knew."[1] However, early Web browsers just provided read-only access to existing hypertext markup language (HTML) pages, and this publishing model for the Web has predominated. It was not easy for people to share what they knew on the Web; the process involved learning to add HTML to text and using network file transfer software to upload content to a Web server. WikiWikiWeb, or Wiki, software provides one option for realizing Berners-Lee's early vision by enabling authorized users to edit and create new Web content using only a Web browser.

In 1994–95, Cunningham developed the first Wiki for the Portland Pattern Repository. It can be found at http://c2.com/cgi/wiki, and is sometimes called Ward'sWiki or TheOriginalWiki. *Wikiwiki* is the Hawaiian word for fast, and a WikiWikiWeb is a quick Web site. Cunningham wrote the original Wiki in Practical Extraction and Report Language (PERL). There is now a wide choice of software for WikiWikiWebs, and Wikis are growing in popularity as people explore their potential in different contexts.

This paper, based on a presentation at the 2004 LITA National Forum, will:

- describe the history and development of Wiki;
- list typical Wiki features and concepts;
- outline factors to consider when choosing a Wiki engine;
- illustrate Wiki applications in a library and information management context using selected case studies; and
- identify issues associated with using Wiki.

What is a Wiki?

A Wiki is a server-based collaborative tool that allows any authorized user to edit Web pages and create new ones using nothing more than a Web browser and a text entry form on a Web page. Wikis free writers from the burden of mastering HTML editing and file transfer software before they can publish on the Web. Instead, Wikis use very simple text-based markup to format page text and graphic content. While the idea of letting anyone change anything they want may seem radical or naive, most Wiki engines have features to let community members monitor changes, control user-edit permissions if necessary, restore previous versions of pages, and delete unwanted pages.

Wikis make it possible for people to collaborate in a Web environment by creating, organizing, and maintaining a Web site of automatically linked pages. As such, it is one example of what is known as social software, software that makes it easy for groups of people to communicate or work together in a virtual environment.[2] Other examples of social software include Web-based discussion forums, Internet relay chat (IRC), weblogs or blogs, and instant messaging (IM). Large, successful Wikis usually have some type of constitution or philosophy that establishes goals and provides guidelines for individuals who want to participate in the group.

Some of the words in this paper have StrangeCapitalization. This is because the first Wikis used this convention to name a WikiPage. The software used to operate a Wiki is known as a WikiEngine, and is available in a wide variety of languages, and with a range of features.

Why Wiki?

When might a Wiki be an appropriate choice? Consider the following situations:

- a state library association is looking for an easy way to create and maintain its Web site;
- a professional association special interest group wants to provide an easily updated Web-based resource for its members;
- a library would like any authorized staff member to be able to update content on its intranet as necessary, without needing to use specialist software;
- an open-source software application project has participants from a range of countries who need to be able to contribute to a shared knowledge base for all software users;

Brenda Chawner (brenda.chawner@vuw.ac.nz) is Senior Lecturer, School of Information Management, Victoria University of Wellington, New Zealand. **Paul H. Lewis** (paull@usca.edu) is Government Documents Librarian, Gregg-Graniteville Library, University of South Carolina–Aiken.

- a conference planning committee needs a Web-based tool to keep track of their activities and who is doing what;
- a cross-organizational working group is preparing a collaborative report on a study trip to a range of institutions using a new type of software package; and finally,
- two people in different parts of the world are jointly writing a conference paper and would like to be able to see each other's work in context, rather than as individual word processing documents.

A Wiki could be used for each of the above scenarios. This paper will provide selected library and information Wiki examples.

History and development of Wikis

Since Cunningham developed the original Wiki in the mid-1990s, the Wiki concept has spread to many other groups, and people have written Wiki engines in a wide range of scripting/programming languages. Core Wiki features such as ease of editing, simple markup, and automatic linking of pages have been present since the beginning, but as the number of people using Wiki engines grew, extra features have evolved. These include a command to compare the current version of a page with earlier versions (PageHistory or QuickDiff), and one to browse a list of recent changes.

It is hard to determine how many Wikis exist, but SwitchWiki (www.worldwidewiki.net/wiki/SwitchWiki) lists some one thousand public Wikis, and there are many more that are restricted to specific groups (known as GatedCommunities in the Wiki world).

WikiFarms (both free and commercial) are servers that run a Wiki engine and allow people to set up their own Wikis without having to install any software. SeedWiki (www.seedwiki.com) is one example that lets individuals set up Wiki sites for personal use.

Leuf and Cunningham identify six types of Wikis, based on read- and edit-access permissions.[3] These are:

- fully open, meaning that anyone has full access to the Wiki;
- lockable, with restricted editing for some or all pages;
- gated, with some public pages (that may be locked), but other pages restricted to authorized users;
- members-only, where access is limited to registered users;
- firewalled, where access is restricted to a range of specific IP addresses; and
- personal, where access is limited to a specific computer or private site.

Wikipedia (http://en.wikipedia.org/wiki/Main_Page) is an ambitious project to build a free, open-content encyclopedia. It was begun in 2001 by Jimmy Wales and Larry Sanger, using Wiki software, and as of December 2005, included more than 880,000 articles in English. They based the idea on the open-source concept of "many eyes make all bugs shallow," and anyone can edit articles and add new ones. While the idea of an encyclopedia that anyone can edit seemed ludicrous at first, during the last four years the project has gained credibility, and the Wikipedia community has put in place mechanisms to monitor and improve the quality of its content. Observing Wikipedia pages over time generally shows that article quality initially improves, and then stabilizes. Non-English versions of Wikipedia are also available; the languages range from Arabic to Chinese to Esperanto, as well as most of the European languages. Wikipedia-related projects now include Wikiquote (http://en.wikiquote.org/wiki/Main_Page), Wikispecies (http://species.wikipedia.org/wiki/Main_Page), and Wiktionary (www.wiktionary.org).

WikiTravel (http://wikitravel.org/en/article/Main_Page), inspired by Wikipedia, is a Wiki-style travel guide begun in 2003. It uses a modified version of the Wikipedia engine, MediaWiki. By August 2004, it had 2,325 destination guides, as well as related articles.

The OpenGuides (http://openguides.org/) project, started in 2004, takes a similar approach to city travel guides. So far, Guides have been started for Glasgow, London, Northern Ireland, Nottingham, Oxford, Vegan Oxford, and Reading in the U.K.; Bologna, Oslo, and Vienna in Europe; and Orlando, Saint Paul, and San Francisco. They use the PERL OpenGuides software.

Typical Wiki features

This section describes features that are found in most Wiki engines, although in some cases the syntax may be different. It is by no means exhaustive, and the help pages for each Wiki engine will normally list the range of features and markup it supports. In contrast to HTML, Wiki markup is intended to be simple and nonintrusive.

Creating a new page

Creating a new Wiki page is a simple process. Most modern Wikis use a method known as "free linking" to create new Wiki pages. A free link is created by enclosing any word or phrase within square brackets when editing a Wiki page—i.e., [[New Page]] will create New Page. When the page is saved, the Wiki software interprets this markup and presents it as a hyperlink followed by a question mark to denote that the new Wiki page is an unedited page. Clicking on the new page hyperlink invokes

a blank editing form on which content for the new page is entered. When editing is completed, the new page is saved and becomes a part of the Wiki.

The result of adding the text [[New Page]] to an existing page is shown in figure 1.

The free-link markup can also be used to link to existing local Wiki pages, local file attachments, e-mail links, external Web site pages, and other networked resources. The approach to hyperlinking varies from one Wiki to another—for example, MediaWiki uses a single set of brackets while PmWiki employs double brackets.

Early Wikis commonly employed a convention known as CamelCase to create a new Wiki page. A typical CamelCase word consists of any two words joined together with the initial letter of both words capitalized. CamelCase is still used in some Wikis. These normally provide a way to prevent accidental WikiWords resulting from proper names like McIntosh or other situations where CamelCase occurs naturally, such as Ph.D.

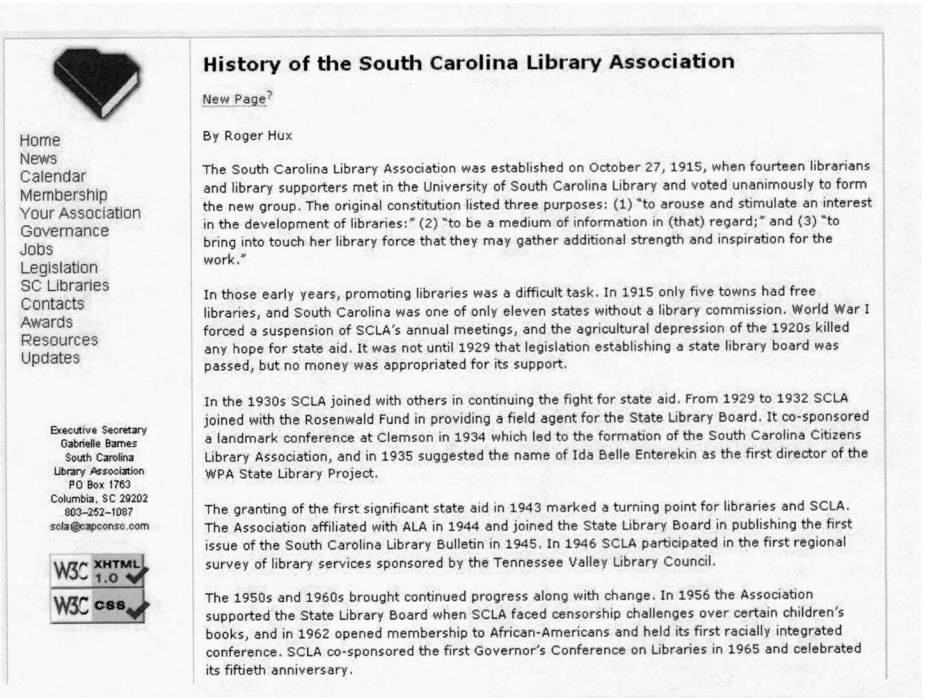


Figure 1. Saved page with the new link indicated by a "?"

Text formatting

Because Wikis use an HTML form to enter and edit content, the markup is text-based, using characters to signal special formatting. Typical formatting conventions are:

- blank lines signal new paragraphs;
- asterisks at the left margin (*) indicate a bulleted list;
- number-signs (#) at the left margin indicate a numbered list;
- two single-quotes ('), i.e., two apostrophes, indicate emphasis (usually italics);
- three single-quotes (''), i.e., three apostrophes, indicate strong emphasis (usually bold); and
- four or more hyphens (----) at the beginning of a line create a horizontal line.

In figures 2 and 3, two screenshots illustrate a range of Wiki markup and the way it is displayed.

This illustrates one of the benefits of text-based markup—the page of text displaying the Wiki markup is more readable than the associated HTML source.

Linking to an external Web page or resource

Including the text "http://," "mailto," or "ftp://" (or even "gopher://") before a URL or e-mail address will create an automatic link to the location or e-mail address. Some WikiEngines also recognize "www." as a reference to a Web resource.

Sandbox for new users

Most open Wikis have a page called SandBox or PlayGround for new users to experiment with such things as entering and formatting text, building unordered and numbered lists, and creating hyperlinks. The general rule is that anyone can edit anything on the SandBox page.

File uploads

Wikis usually provide a method for uploading images and other file types (e.g., Adobe Acrobat portable document format [PDF], Microsoft Word) to the Web site. File transfer protocol (FTP) software is the traditional approach to uploading files (such as image files or Adobe Acrobat PDFs) for posting on a Web server. One of the negative aspects of FTP is that the protocol has no way of encrypting username and password information and this is widely regarded as a major security vulnerability. FTP security issues are explained in detail at http://en.wikipedia.org/wiki/Ftp.

A key benefit of Wiki engines is that the process for uploading files is built into the program. File

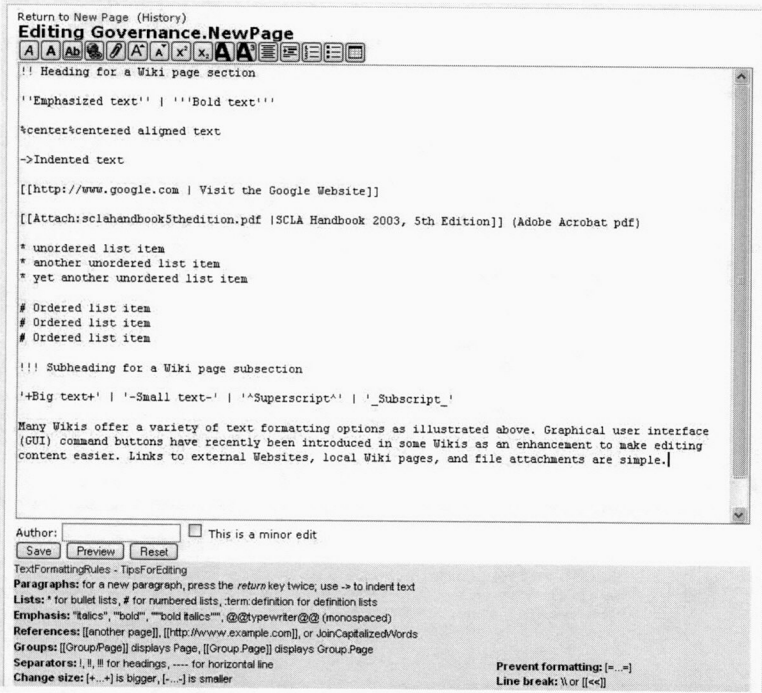


Figure 2. Range of Wiki markup as entered

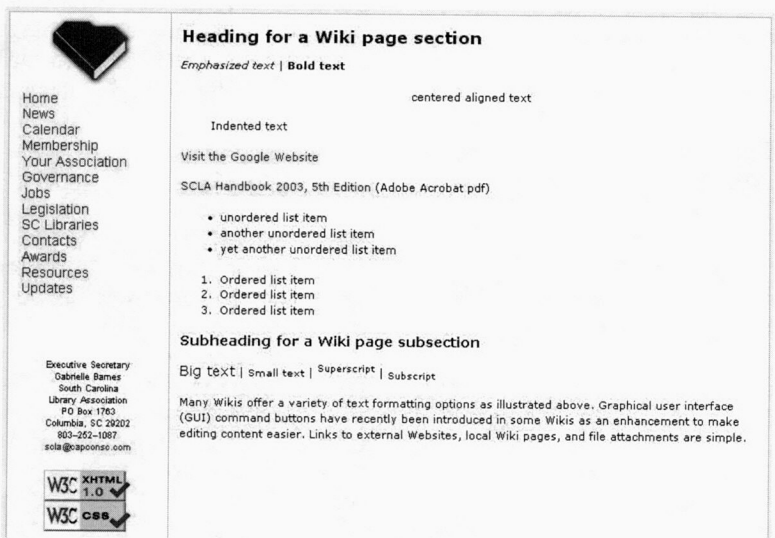


Figure 3. The way the markup is displayed

uploads are accomplished using a Web browser and an HTML form; no separate FTP program is needed. Wiki administrators can set upload parameters specifying allowed file types and file size limits in a server-based configuration file; they can also limit file upload capabilities to authorized users. This Wiki file upload process provides a more secure approach than FTP, because it does not involve sending unencrypted user account information over the Internet.

The simple syntax for hyperlink referencing of uploaded files makes it possible to use a Wiki to archive collections of document, image, and multimedia files on the Web. Leo LaPorte, the self-styled "Tech Guy" at the KFI AM-640 radio station in Los Angeles, uses the PmWiki engine to maintain a podcast archive of his radio programs at http://leoville.tv/radio/pmwiki.php/Main/AudioArchives. The South Carolina Library Association (SCLA) Governance page, www.scla.org/Governance/HomePage, is an archive of association annual reports, newsletters, and other documents in Adobe PDF, Microsoft Word, and Wiki page formats.

Recent changes

A list of recently changed pages is maintained automatically by most Wiki engines (see figure 4). As a rule, only the latest change is shown for a given page. This is a particularly useful feature, as it allows Wiki community members to easily see what has been changed. Some Wiki engines also support e-mail notification or really simple syndication (RSS) feeds for page additions, changes, and deletions.

Other features

Wiki functionality is continuing to evolve as developers improve Wiki engines. One example of a newer feature is Page History. This feature tracks previous versions of a Wiki page. With Page History, previously saved versions of a Wiki page may be recalled, reedited, and restored

if necessary. This feature is essential for a fully open Wiki, which is vulnerable to malicious damage, or WikiSpam. New Wiki content contributors often find this aspect of Wikis a comfort, as it serves as a safety net to prevent them from doing harm to the Wiki page on which they are working.

Even the simplest Wikis available today usually include a basic keyword search engine to search the contents of the entire Wiki Web site. Of the Wiki engines already discussed, only PmWiki offers full Boolean search capabilities. PmWiki's search engine can be tailored in a number of ways to limit searching to specific Wiki groups or pages.

More advanced Wiki engines, such as MediaWiki, PmWiki, and TikiWiki provide additional capabilities to organize groups of pages by hierarchical categories. Wikipedia uses hierarchical categories to organize its entries. For an example of the use of categories to manage linking to geography-related topics, see http://en.wikipedia.org/wiki/Category:Geography. The PmWiki engine offers WikiTrails as another feature for organizing groups of pages. An example of a WikiTrail for a library circulation manual is shown in figure 5.

The WikiTrail appears at the top and bottom of the Web page, allowing quick access to the preceding, home, and next pages in the manual. WikiTrail links facilitate logical movement through the pages in the circulation manual.

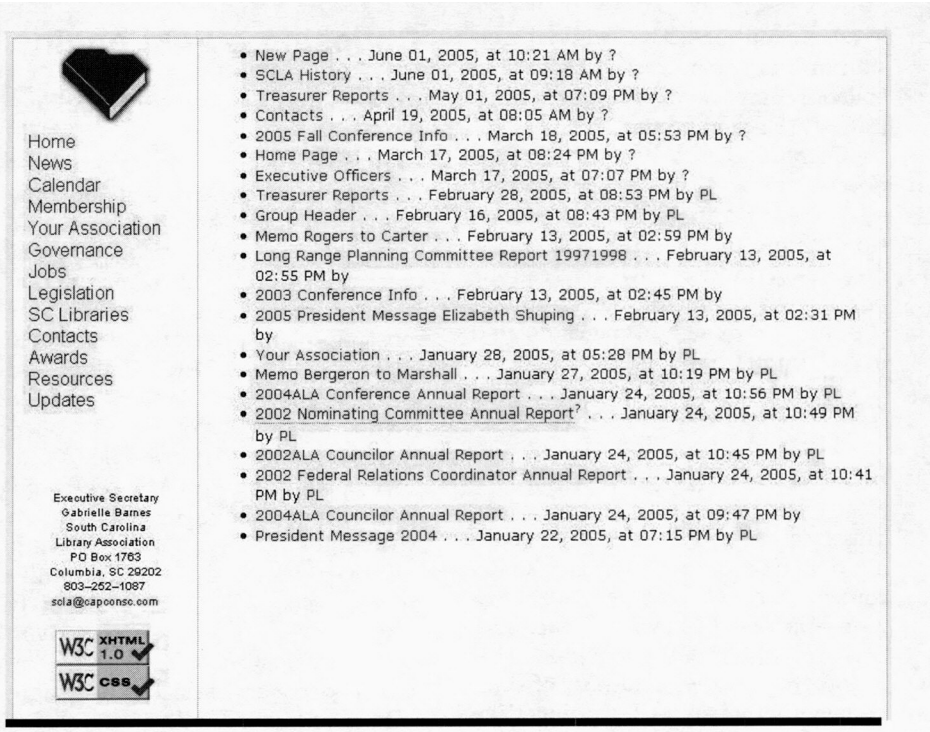


Figure 4. Sample recent changes page

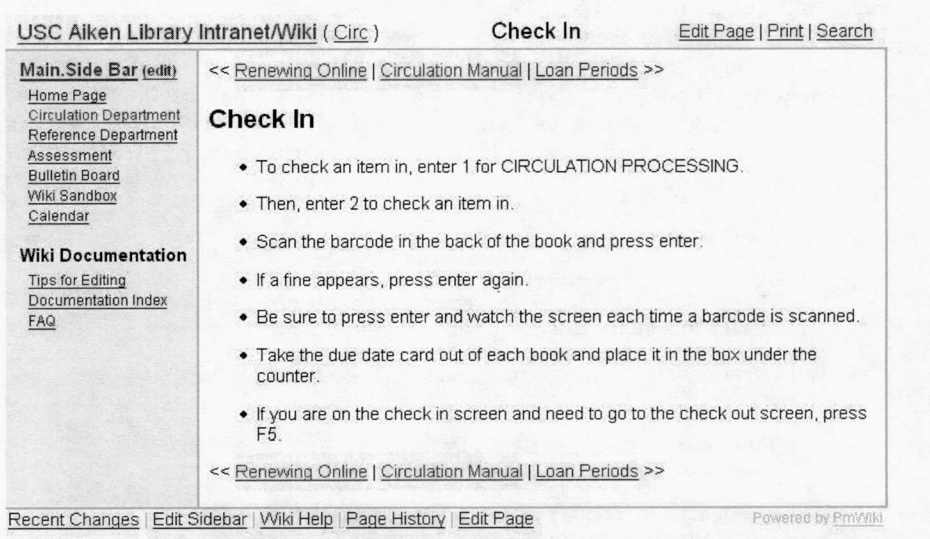


Figure 5. Sample WikiTrail page

Other Wiki concepts

Collaboration versus discussion mode

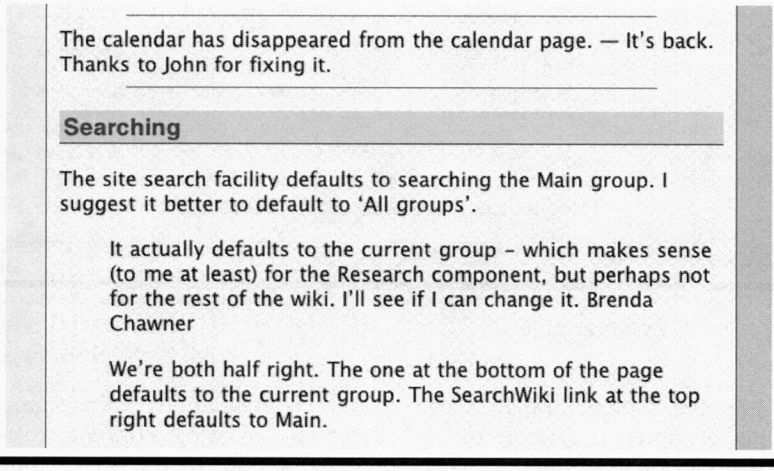
There are two main writing modes used by members of a Wiki community. In collaboration or document mode, the focus is on creating a piece of text that the community is happy with, and the content of which can be edited by anyone. Discussion, or thread mode, on the other hand, is a form of dialog in which individual contributions are kept separate (more like a conversation). Rather than editing the existing content, different authors add their own comments, and may "sign them." Horizontal lines are often used to insert a break

between authors. The original version of this paper was written using collaboration mode. Figure 6 illustrates discussion mode.

Communities whose main purpose is online discussion might find other software tools, such as discussion forums, more useful. Wiki-based discussion mode, however, can be particularly useful as a technique to support comments on pages created in collaboration mode.

Refactoring

This term is used to refer to the editing of one or more pages to make the content more coherent, or move it to a more appropriate location. Refactoring is often done when an extended discussion has created a number of similar pages, or when related points have been made on different pages. Good practice is to include a note on all edited pages to say what changes have been made, or to indicate the location to which content has been moved. Refactoring may also involve adding structure to a page—for example, by including headings or a table or contents, to allow a long page to be browsed more easily.


The calendar has disappeared from the calendar page. — It's back. Thanks to John for fixing it.

Searching

The site search facility defaults to searching the Main group. I suggest it better to default to 'All groups'.

> It actually defaults to the current group - which makes sense (to me at least) for the Research component, but perhaps not for the rest of the wiki. I'll see if I can change it. Brenda Chawner
>
> We're both half right. The one at the bottom of the page defaults to the current group. The SearchWiki link at the top right defaults to Main.

Figure 6. Selection from discussion mode page (accessed July 1, 2005, http://wiki.lianza.org.nz/index.php/ITSIG/CommentOnTheWiki)

Choosing a Wiki engine

Currently there are at least one hundred separate publicly available Wiki engines to choose from, according to an extensive list of Wiki software projects maintained at http://c2.com/cgi/wiki?WikiEngine. Wiki engines have been developed in many computer-scripting languages, though the most popular appear to be personal home page (PHP), PERL, Python, and active server pages (ASP). Some Wiki engines work only on specific operating systems. For example, Wikis developed using the Microsoft ASP programming language only run on Microsoft Windows Web servers. However, most of them are multiplatform and will function properly if their requirements are met. With so many different Wiki engines from which to choose, selecting the right Wiki can seem like a difficult task.

Wiki documentation

One can quickly begin to get a clear picture of the level of quality of a given Wiki project by checking out the Wiki developer's Web site. Is clear documentation provided on all facets of the Wiki project, including system requirements, installation, upgrading, core functions, and expanded features?

Project maturity

Every individual Wiki is a unique software project unto itself. As such it can be thought of as existing in a particular stage of development. Has the Wiki been around a while or is it just getting started? Does the project Web site have a section detailing the development history of the Wiki? In general, a more mature project is more stable and reliable with fewer software bugs.

Wiki community

Is the Wiki developed by a single individual working essentially alone, or is there an active core of developers? What might become of the Wiki project if the original sole developer loses interest or otherwise cannot continue development? Also, what methods of feedback are provided for end users and developers to communicate? A mature project Web site will often offer a combination of e-mail, bulletin board, electronic discussion list, and possibly IRC for fast responses to requests for support. Are developers friendly and helpful when new users have questions? Is there an underlying philosophy or design principle, and how closely is it aligned with the library's goals for a Wiki?

W3C standards compliance

Does the Wiki offer compliance with Extensible Hypertext Markup Language (XHTML) and cascading style sheet (CSS) guidelines recommended by the World Wide Web Consortium (www.w3c.org)? In January 2000, the W3C issued guidelines urging adoption of XHTML, the successor markup language for

HTML. Any library considering an upgrade of its main Web site or any new Web project that will be publicly available should consider XHTML compliance as a basic requirement. However, if the Wiki is to serve solely as the library's intranet or for some other nonpublic Web project, then XHTML compliance might be less of a concern. For more information about XHTML, see http://en.wikipedia.org/ wiki/ Xhtml#The_XHTML_2.0_draft_specification. A CSS is essential for developing a consistent look and feel for a Web site. Mature Wiki projects provide CSS functionality for site-wide aesthetic control over the color, layout, fonts, and so on.

Local resources and expertise

Does the library have its own Web server or does it share space on the parent institution's Web server? Although not required, it may be more convenient in the long run to have a Web server dedicated to Wiki. Every Wiki contains one or more configuration files that reside on the Web server. The Wiki administrator will need access rights to the Web server in order to edit configuration files. Most likely this person will be the current library webmaster. Is there a staff person available with the knowledge, skills, and abilities to serve this function? Also, are local staff who contribute content on board with the changes? One of the biggest obstacles to implementing new Web technology may be reluctance by content contributors to try something new.

A choice is a commitment

Moving from traditional Web site development to a Wiki or other type of content management system requires a substantial investment of time and effort. Very few systems offer batch transfer options to automatically migrate the content of an existing Web site to another format, though preliminary discussions about a WikiInterchangeFormat have begun (http://c2.com/cgi/wiki?WikiInterchangeFormat). It might be useful to identify a short-list of potential Wiki engines, and install each on a trial basis before making a final decision.

Beyond the basic Wiki

Simplicity in form and function was part of the original allure of Wiki software. Today Wikis have evolved beyond being simply a workgroup collaboration tool. They are now replacing traditional Web sites. As such, some Wikis include advanced, interactive elements that modern Web sites offer such as permissions control for content contributors and users, e-mail feedback forms, calendars, photo galleries, and RSS feeds. A library organization looking to replace its traditional Web site with a Wiki, for example, might want to consider whether a given Wiki engine offers any extra functions. Additional features are often called plugins or modules. Plugins should be stable and simple to install, and supported by an active plugin community. A list of plugins a library Wiki might consider are (to name but a few): an interactive calendar application, graphical user interface edit controls, a keyword searchable image gallery for managing a local history digital photo collection, project management issue tracking, RSS link feeds from national and international news organizations or professional associations, a Java-based interactive image editing application, and a geographic map and data server.

Disadvantages of Wikis

Though their flexibility and simplicity mean that Wikis are useful in a variety of contexts, there can be issues associated with their use. WikiSpam can be a major problem for fully open Wikis. This usually takes the form of unwanted links to commercial or pornographic sites. While active Wiki communities generally rely on members monitoring changes and reverting spam, less active ones might need to implement a form of spam protection. Techniques for dealing with WikiSpam are evolving, and currently include:

- increasing the level of security, usually by requiring a password to use the Edit function;
- blocking updates from IP addresses of known spammers;
- blocking updates that contain specific unwanted words or phrases;
- restricting the number of links that can be added in a single update;
- "hiding" the Wiki site from search engines by implementing "no index" and "nofollow" metadata tags on Wiki pages, as well as asking community members to avoid posting its URL on Web sites; and
- requiring any links to external sites to be approved by an administrator.

The lack of a standard for Wiki content markup, which makes it difficult to migrate content to another Wiki engine or Web content management, is another issue, though this applies to all Web content environments, not just Wikis. The lack of a "what you see is what you get" (WYSIWYG) editing environment can be a barrier for participants used to working in a word-processing environment.

Although all Wikis share a set of key characteristics, Wiki projects vary considerably in their underlying architecture and the feature sets they offer. The result is a rich diversity of Wiki engines designed to meet a range of needs for managing Web content. The DocuWiki site includes an extensive table comparing features of

different Wiki engines at http://wiki.splitbrain.org/wiki%3Acompare. This section examines three Wiki projects, which represent different approaches to Web content management; these range from extreme simplicity to (perhaps) feature overkill.

QwikiWiki

Project Home Page: www.qwikiwiki.com
Download: http://prdownloads.sourceforge.net/qwikiwiki
Download file size: compressed: 57 kilobytes
Real world deployment: PeacefulFuture.Net http://wiki.peacefulfuture.net/

QwikiWiki packs an impressive amount of power and flexibility in a very small amount of PHP code. It is a testament to the creativity that can be achieved by a single software developer. QwikiWiki includes an installation wizard program to walk the Wiki administrator through the process of installing the program on the Web server. The Wiki includes three template files to choose from that control the basic layout of the Wiki and four separate CSS that control such things as font types and sizes for section headings and text. The Wiki administrator is free to edit existing templates and style sheets or create completely new versions to suit personal preferences.

According to Barrett, the original developer, QwikiWiki's core design goal is simplicity. This is reflected in the limited Wiki syntax users are required to learn to begin editing and in the feature set. QwikiWiki permits HTML to work along with the Wiki syntax and this can be especially helpful—for example, when migrating Web pages that include lengthy tabular data from a traditional Web site. QwikiWiki stores its Wiki pages as data in flat text files as opposed to using a database. This is also in keeping with the focus on simplicity.

QwikiWiki includes a very basic Web site search engine, with no Boolean capabilities. An edit password feature is included to prevent page edits by unauthorized users if that option is desired. Other helpful features include an option for automatic e-mail notification when changes are made to individual Wiki pages, ability to list pages that have recently been changed, and a built-in link to a help page containing info on QwikiWiki syntax. Clear, basic documentation is provided on all features. The QwikiWiki project Web site includes a discussion forum where users can ask questions and post feature requests, and an electronic discussion list.

QwikiWiki currently does not offer W3C XHTML compliance so it would be more suited for a library intranet or other nonpublicly accessible library Web project. However, the availability of only a single-site password for this Wiki could also be a limiting factor if, for example, a site has multiple-content providers who each need separate password protection for sections they manage.

PmWiki

Project Home Page: www.pmwiki.org
Download: www.pmwiki.org/pub/pmwiki
Download file size: compressed: ~140–200 kilobytes
Real world deployment: University of Minnesota Libraries Staff Web Site http://wiki.lib.umn.edu

PmWiki, like QwikiWiki, also uses flat text files to store its data. Overall, however, PmWiki is more ambitious in its approach to Wiki design. Numerous developers collaborate with Michaud, PmWiki's creator, on the PmWiki core code, on plugins that extend the capabilities of PmWiki, and on project documentation. The main PmWiki Web site itself is an exemplary demonstration of the power of Wiki technology to support detailed project documentation. The PmWiki project Web site is an open Wiki, and users and developers regularly contribute to and refine the project Web site documentation. The basic PmWiki installation includes extensive documentation.

Updated versions of PmWiki are released on a very frequent basis. Developers and users have multiple channels to provide input and feedback, including a very active electronic discussion list, an IRC channel for real-time troubleshooting, and, of course, e-mail.

One feature that distinguishes PmWiki from other Wikis is its support of the concept of Wiki groups. A PmWiki Web site may have multiple groups that can be created according to organizational function or for individuals within an organization, or by topic, or any combination of groupings. These distinct groups may have separate stylistic elements and password settings apart from the main Wiki. Read-and-edit passwords may be set for individual pages, groups, or the entire Wiki site. This granularity of read-and-write password protections for individual pages and groups can be very useful in library or other Web site contexts where site content providers need to exercise control over who may edit their pages.

With PmWiki it is possible to establish a WikiFarm on a single server. WikiFarm is a feature somewhat similar in concept to a Wiki group. With a WikiFarm, multiple independent Wikis are maintained through a single configuration file.

PmWiki's editing syntax is more extensive than that found in QwikiWiki, and yet it is still quite simple to learn. CamelCase WikiWords can be spaced when pages are displayed, or turned off completely. This approach gives hyperlinks a more natural, grammatically correct presentation.

Like QwikiWiki, PmWiki also permits mingling of Wiki syntax and

HTML markup; this option is configurable by the site administrator, since enabling some types of HTML markup on an open Wiki can be a security risk. PmWiki has given significant attention to ensure compliance with W3C XHTML markup requirements. The look and feel of the site is controlled by templates and CSS. Its user community has contributed numerous templates to the project. As with QwikiWiki, users are free to edit existing templates and style sheets or create new ones. This design flexibility makes it possible to marry the PmWiki engine with a variety of Web site layout designs.

PmWiki offers a rich feature set. The PmWiki Web site search engine offers Boolean search capabilities. The search engine can also be configured to search selected Wiki groups within the Web site. An interactive calendar, photo galleries, graphical editing buttons, RSS feeds, and a random quote plugin are just some of the many enhancements beyond the core Wiki program. Most plugins are installed easily by uploading a single script file to the Web server and then editing a main configuration file to add a single line in the code referencing the script file. The PmWiki project Web site has a Cookbook section where all available plugins are listed along with installation details.

PmWiki is a powerful Web site content management system that can be used for a large number of applications. The ability to set read access rights for the entire Wiki makes PmWiki an excellent choice for organizational intranets. PmWiki is a good choice for small, medium, or large library Web sites and Web projects. It is one of the more widely used Wiki engines, with a Google search showing more than one million hits in mid-2005.

TikiWiki

Project Home Page: www.tikiwiki.org
Download: http://tikiwiki.org/Download
Download file size: compressed: ~4.7 to 6.9 megabytes
Real world deployment: Damosoft UK www.damosoft.co.uk/Home

TikiWiki is a large, open-source content management system (CMS) software project with more than six thousand registered users and more than 260 registered developers. TikiWiki has been called a "kitchen sink" Web CMS because of the numerous features and functions it offers. The actual Wiki component is but one of many modules included in TikiWiki. Other TikiWiki modules include public and private discussion forums, workflow management tools, interactive chat, a photo gallery, file download archives, a map server, and many others. TikiWiki, like many other traditional CMSs, is more focused on managing time-sensitive news articles, discussion forums, and blogs rather than on building Web pages. Indeed some CMS packages have created static-page plugin modules well after the initial release of the core package as a way to address user needs for traditional Web page management. By fully integrating a Wiki into its feature set, TikiWiki represents a hybrid CMS/Wiki and may well set a trend that other CMS projects emulate in the future.

TikiWiki employs a MySQL database to store all its data, including its Wiki pages. This requirement for MySQL database management skills adds a significantly higher threshold for the would-be Wiki administrator that can be somewhat intimidating. It should be noted, however, that many Wikis use a database backend and that learning how to manage a database-backed Web site is not an insurmountable task. Several excellent, free, open-source software tools are available to simplify the process. TikiWiki includes detailed installation documentation and a software wizard script to streamline initial Web site setup. Like PmWiki, the TikiWiki developers have used their Wiki to develop extensive documentation for all facets of the program.

TikiWiki includes extensive user and group read-and-edit permissions settings for its Wiki and all other modules. TikiWiki administrators pick and choose the modules they wish to deploy using a Web browser-based site control panel; other modules are turned off by default. It offers W3C XHTML and CSS compliance. There are numerous stylistic themes from which to choose. The project is under constant development with frequent updates. TikiWiki can be fine tuned to serve just a about any type of Web site need.

Selected library and information management case studies

There is an increasing number of library and information management Wikis; some are public while others are gated or closed, such as library staff intranets. Lamb identified four educational Wikis in use at the University of British Columbia, including course management, job postings, conference planning, and collaborative writing technology.[4] Frumkin describes the use of Wiki software for a reference librarian knowledge base and a Web site content editing tool, and suggests that Wikis also have potential in digital libraries—for example, as a tool for user annotations or comments.[5] Farkas set up an unofficial American Library Association 2005 conference Wiki, which was used by a number of conference attendees for conference tips, location information, and session reports (http://meredith.wolfwater.com/wiki). More recently, Farkas has set up Library Success: A Best Practices Wiki (http://libsuccess.org).

LIANZA/ITSIG Wiki

http://wiki.lianza.org.nz

A Wiki used to develop a shared resource for members of the Information Technology Special Interest Group (ITSIG) of the Library and Information Association of New Zealand/Aotearoa (LIANZA). This was launched in October 2003, using the PmWiki engine. It includes information about ITSIG and its activities, as well as a resource page with links to useful sites. It also hosts a research register for the Research Special Interest Group (SIG), which allows anyone to create a page about research projects relevant to New Zealand libraries. A template is used to format the initial edit page for a new research project page, displaying the fields and basic markup. WikiSpam (like e-mail or blog comment spam) has become an issue, and the Wiki has recently changed from being fully open to requiring a password to save changes. Known Wiki spammers are blocked.

Koha Project Wiki

www.saas.nsw.edu.au/koha_wiki

A locked Wiki set up to support members of the Koha open-source library management system project. It includes a register of Koha users, pages about installation and migration, information about reporting bugs, and documentation about Koha features. It uses the WikkiTikkiTavi engine.

New Zealand Government Tertiary e-Learning Standards Wiki

http://wiki.tertiary.govt.nz/~TertiaryELearningStandards/Main/HomePage

This is a gated Wiki set up to allow people to comment on an overview of standards for e-learning. People who wish to comment must first register, but the content is available for anyone to read. Part of a project to develop a national tertiary e-learning framework, the interim framework document was developed using a password-protected closed Wiki, with participants from a number of government departments, including the Ministry of Education and the National Library. Both Wikis used the PmWiki engine.

New Zealand Institutional Repositories Feasibility Wiki

http://wiki.tertiary.govt.nz/~InstitutionalRepositories

The National Library of New Zealand set up a project to investigate options for institutional repositories for the New Zealand research sector in May 2005. This Wiki site was set up to provide a shared resource for project members, in particular to allow them to document their findings. It uses the PmWiki engine.

South Carolina Library Association

www.scla.org

The South Carolina Library Association (SCLA) Web site serves the library community in South Carolina. It is a moderately sized Web site with more than three hundred separate pages. Page editing is password protected. Association committee, section, and round table chairs are authorized to maintain their respective pages. This Web site uses the PmWiki engine.

University of South Carolina (USC)–Aiken Library

http://library.usca.edu

The USC Aiken (USCA) Library Web site is a typical small academic library site employing PmWiki software for site content management. PmWiki is also used to manage the USCA Library's intranet. The intranet portion of the site includes the circulation department procedures manual, a section for development of the library's strategic planning and assessment, a section for collaboration for this Library and Information Technology Association (LITA) paper and presentation, and other content.

The USCA Library migrated to PmWiki from another content management system software, phpWeb site (http://phpWeb site.appstate.edu), during the summer of 2004. Librarians and some general staff contribute content to the main Web site and the intranet.

USC—Campuses Library Council

http://library.usca.edu/clc

The USC Campuses Library Council Web site is a very small site whose purpose is to support communication and planning for the eight libraries of the USC statewide campus system. This Web site uses the PmWiki engine.

Wiki success factors

Getting people to contribute to a Wiki is as much about culture as it is about technology. Even though it is very easy to add and edit Wiki content, it is not always a case of "build it and they will come." People accustomed to the WYSIWYG approach in MS Word (or even Dreamweaver) may take a while to adjust to the idea of text-based markup, for example. Some Wiki engines, such as MediaWiki and PmWiki, are now moving to a more WYSIWYG approach, with graphical edit but-

tons on the edit page to help people learn the markup.

Creating the right conditions for a Wiki involves:

- setting up an effective initial structure, so that people can see where their contribution might fit;
- monitoring new and changed content, so that inappropriate content is dealt with promptly. This might also require clear guidelines on appropriate content, and a statement about the identity of the intended audience;
- having a statement about copyright and content ownership. If people sign their contributions, then it might not be considered appropriate for someone else to edit them; a comment (i.e., discussion or thread mode) might be better;
- providing an explicit Page History link to make it obvious that content can be restored if necessary;
- having basic text formatting tips displayed when someone edits a page, to help them remember the markup; and
- having suggested page naming conventions and writing style guidelines. The basic idea here is to remove some of the uncertainty associated with a totally open environment, which might help people overcome their initial hesitation about contributing.

In the early days, it might be useful to "shoulder tap" selected community members for specified Wiki contributions, to start the habit. Some Wikis have lists of "wanted pages" to identify topics they'd like someone to write about.

In some cases having a "comment on the Wiki" section encourages people to describe their reaction to the idea of editable Web content, and it can overcome their initial fear of breaking something if they edit a Wiki page.

Conclusion

The case studies in this paper illustrate just a few situations where Wikis can be effective. Because of their flexibility and simplicity, they can be used in a wide range of contexts, and provide an environment in which Berners-Lee's early vision for the Web can be achieved. A totally open Wiki might suit a widely distributed organization like the LIANZA ITSIG, while a members-only Wiki would suit a work group collaborating on policy documents or procedure manuals (or simply wanting an easy electronic notice board). Features like Recent Changes and Page History make it easy for community members to keep up with changes. Wikis are also gaining in popularity as simple, lightweight Web site CMSs for groups and individuals. Wikis offer libraries and other organizations a tool that can be used when upgrading traditional Web sites or implementing new Web-based projects—their potential for enabling Web-based communication with staff and users is just beginning to be appreciated.

Libraries were early and enthusiastic adopters of the Web as a medium to enhance and expand access to information for the communities they serve. The Web is constantly changing, and in the next few years we expect organizations to move to more interactive e-services. New technologies offer new opportunities and evolving Web accessibility standards present new challenges for libraries. For libraries looking to take advantage of new technologies and build Web sites that comply with Web accessibility standards, Wikis offer a relatively easy path to build the next generation of library Web sites. At the same time, they raise the possibility of having more interaction with users.

Finally, Wikis illustrate a shift to an increasing ability to use a Web browser as a person's main application tool, and they foreshadow other browser-based capabilities, such as table or drawing editors, which would make it possible to create complex documents using nothing more than a standard browser.

Author's note: all URLs in the text were accessed in early July 2005.

References and notes

1. Tim Berners-Lee with Mark Fischetti, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web* (New York: HarperBusiness, 2000), 33.
2. *Wikipedia*, s.v. "Social software," June 30, 2005. Accessed July 1, 2005, http://en.wikipedia.org/wiki/social_software.
3. Bo Leuf and Ward Cunningham, *The Wiki Way: Quick Collaboration on the Web* (Boston: Addison-Wesley, 2001), 277.
4. Brian Lamb, "Wide Open Spaces: Wikis Ready or Not," *Educause Review* 39, no. 5 (Sept./Oct. 2004): 36–48.
5. Jeremy Frumkin, "The Wiki and the Digital Library," *OCLC Systems & Services: International Digital Library Perspectives* 21, no. 1 (2005): 18–22.

Graphical Table of Contents for Library Collections: The Application of Universal Decimal Classification Codes to Subject Maps

Victor Herrero-Solana, Félix Moya-Anegón, Vicente Guerrero-Bote, and Felipe Zapico-Alonso

The representation of information content by graphical maps is an extended

ongoing research topic. The objective of this article consists in verifying whether it is possible to create map displays using Universal Decimal Classification (UDC) codes (using co-classification analysis) for the purpose of creating a graphical table of contents for a library collection. The application of UDC codes was introduced to subject maps development using the following graphic representation methods: (1) multidimensional scaling; (2) cluster analysis; and (3) neural networks (self-organizing maps). Finally, the authors conclude that the different kinds of maps have slightly different degrees of viability and types of application.

Advanced techniques for information retrieval (IR) currently make up one of the most active areas of research in the field of library and information science. New models representing document content are replacing the classic systems in which search terms supplied by the user were compared against the indexing terms existing in the inverted files of a database. The objective of this article consists in verifying whether it is possible to create map displays using Universal Decimal Classification (UDC) codes, a classification system based on Dewey Decimal Classification, for the purpose of creating visualizations of a library collection.

One related topic of study in recent years is bibliographic browsing, a useful complement to querying strategies. Since the 1980s, a number of authors have dealt with this topic. For example, Ellis establishes that browsing is based on three different kinds of tasks: identification, familiarization, and differentiation.[1] Cove distinguishes three different browsing types: searching browsing, general purpose browsing, and serendipity browsing; whereas Bates presents six different types.[2] Yet most interesting is Bawden's browsing classification, which addresses similarity matching, structure-driven displays, and global vision.[3] Global-vision browsing implies the use of graphic representations, referred to in this article as *map displays*, that allow the user to grasp a global idea of the nature and structure of information in a database.

Several authors worked on this line of research throughout the 1990s, developing different types of maps. One of the most active authors was Lin, who introduced the concept of a graphical table of contents (GTOC) that is functionally analogous to the table of contents in the printed environment.[4] Lin applies the self-organizing map (SOM) algorithm to his own personal bibliography, analyzed by title and abstract fields, and represents it in a two-dimensional map.[5] The SOM algorithm is a major method for unsupervised learning, based on a grid of artificial neurons whose weights are adapted to match input vectors in a training set. It was first described by the Finnish professor Teuvo Kohonen and is thus sometimes referred to as a Kohonen map.[6] The algorithm takes a set of input objects, each represented by a vector in the matrix, and maps them onto nodes of a two-dimensional grid. Later on, Lin included such maps in the creation of GTOC Web sites based on a Java application.

Vectorization, the transformation of any information element into numerical data, using words from the title and abstract fields for co-word analysis, generates too large of a matrix, but this technique can be applied to reduced document sets. In this context, it is important to find some element that allows a less complex or "lighter" vectorization. Online public access library catalogs (OPACs) have certain elements, such as the subject codes of UDC, that can be more easily vectorized than free text in order to create GTOCs of a library collection.

Materials and methods

The OPAC selected for this study is that of the Public Library of Granada, which contains 32,700 records and 43,900 UDC codes, an average of 1.34 codes per record. These records were vectorized using the UDC codes, so as to group them into twenty-seven major subject categories, derived from the hierarchical structure of UDC. The Pearson correlation index was applied to this matrix of data (27 x 32,700) to measure the similarities among these twenty-seven major classes and to generate a new matrix (27 x 27), to which the visualization method will be applied. This correlation index approach is widely used for science mapping construction.[7]

Two basic approaches were adopted in creating the display maps: (1) statistical (based on multivariate analysis); and (2) connectionist (usually, but not exclusively, based on artificial neural networks or ANNs).

Within the techniques of multivariate statistical analysis, three basic methods deserve mention at this point: (1) cluster analysis, (2) principal component analysis (PCA), and (3) multidimensional scaling (MDS).[8] According to Kinnucan, Nelson, and Allen, "These methods are referred to as dimensionality-reduction methods because this function is to simplify what might at first appear to be a complex pattern of associations among many entities."[9]

In the following sections, we review and summarize the characteristics of the three methods:

1. **Cluster analysis.** This technique is used to create two-dimensional

Victor Herrero-Solana (victorhs@ugr.es) is Associate Professor and **Félix Moya-Anegón** (felix@ugr.es) is Professor, Department of Library and Information Science, University of Granada; **Vicente Guerrero-Bote** (vicente@alcazaba.unex.es) is Associate Professor and **Felipe Zapico-Alonso** (fzapalo@alcazaba.unex.es) is Associate Professor, Department of Computer Science, University of Extremadura, Badajoz, Spain. All of the authors are also affiliated with the SCImago Research Group (www.scimago.es/).

displays (e.g. dendrograms) of clusters of different objects whose relationships are represented by matrix values. This type of automatic classification, also known as numerical taxonomy, currently comprises more than 150 different techniques that are grouped in families according to shared procedures. Information science as a discipline generally involves polythetic clustering hierarchies, producing trees that illustrate the hierarchy of relationships among elements on the basis of individual characteristics.[10]

2. **Principal Component Analysis (PCA).** The basic premise of PCA is that the linear relation between any two variables is best summarized by a regression line. In other words, the variable that represents the regression line as a point cloud contains essential information about both variables. The two variables are thus combined into a single factor. This mechanism can be used to reduce pairs of variables to single dimensions in order to simplify the graphic display of the elements included in the matrix.[11]
3. **Multidimensional Scaling (MDS).** This multivariate analysis technique is used to identify the dimensions that best explain similarities and differences between variables. Because the purpose of MDS is to generate a map of objects, this approach can be considered an alternative to PCA.[12]

Neural networks are analytic techniques modeled after the (proposed) processes of learning in cognitive systems and the neurological functions of the brain. Neural networks use a data "training set" to build rules capable of making predictions or classifications on data sets. Neural networks can learn to assign multidimensional outputs to multidimensional inputs, and they do so while maintaining a great capacity for generalization. For this reason the better choice is the SOM algorithm. Kohonen's interest in discovering how an organization of this type might arise led him to investigate the subject.[13] The product of that research was the network model, bearing his name, that is capable of performing a topological organization of the inputs presented to it.

This type of network has recently been extrapolated to domain analysis, textual data mining, the extraction of semantic relationships among words in their contexts, and to the generation of topological maps of sets of documents, which may include labeling the zones of influence of each word or term.[14]

MDS maps

In the MDS-based display map each major subject category is placed in a certain point, depending on its relationship to other subject categories (figure 1). Also, each category is represented with a circle whose area is proportional to the volume of documents that it contains. The largest circles are located in the periphery, following the principle of center/periphery established by White and Griffith.[15]

The categories are classified in two large clusters: (1) science and technology, and (2) social science and humanities. This classification corresponds to its clustering, as in a Ward dendogram.[16] There are only two categories that do not seem to be in the expected cluster (economics and law); however, we should bear in mind that economics is related with categories of the science and technology cluster (mathematics). On the other hand, MDS places both categories at the edge of the map; in this way the dividing line can integrate them into the social sciences area.

SOM maps

The map display based on SOM is quite different from the MDS-based

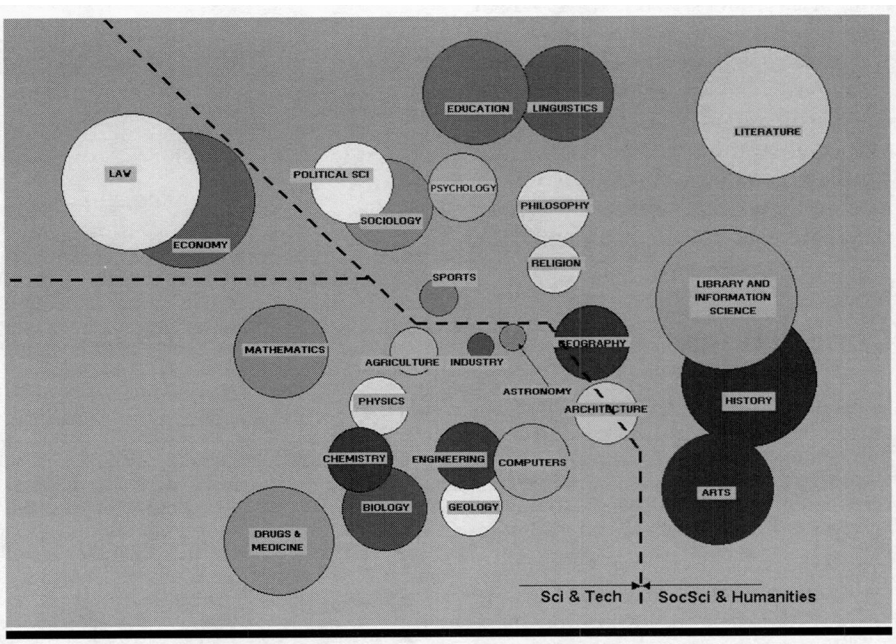


Figure 1. MDS map

map (figure 2). The SOM-based map is clearer, more schematic, and better ordered than the MDS-based map, but the size of each category is not proportional to the volume of documents, which might confuse the user. It is also very difficult to perceive the division of the categories in two big clusters as in the previous MDS-based map. The categories group according to neighboring relationships, a typical feature of Kohonen's algorithm, and the general structure of the collection that one could observe in the MDS-based map is much more difficult to discern here.

The neighboring relationships among the categories indicate the frequency of the co-occurrences of the classification codes. It is important to point out that the SOM searches out the best topology. This implies that when the representation must be reduced to two dimensions, the areas spread and the greater or lesser contact among them is indicative of the degree of interrelation. The proximity/distance among the areas is conditioned by co-classification frequency; however, this does not mean that the codes of classification of two categories that are far away from each other cannot co-occur at all.

To some extent the shapes of the areas are also determined by the co-classifications. These relationships cannot always be represented by the simplest geometric forms, for which reason their final appearance may strike the user as odd or unusual.

Conclusions

Despite the fact that user-based evaluation experience of this kind of map display is very limited, the following conclusions can be put forth:

- MDS and SOM are algorithms that can be used to generate bibliographic map displays
- An OPAC can be represented through co-classification analysis using UDC codes
- It is possible to use other decimal classifications, like DDC, but not Library of Congress
- MDS-based maps enhance viewing the structure of relations among the subject categories
- SOM-based maps are easy to use, because the view is clear, schematic, and well ordered
- SOM is easier to compute than MDS, especially when a lot of variables are involved

References and notes

1. D. Ellis, "A Behavioral Approach to Information Retrieval System Design,"*Journal of Documentation* 45, no. 3 (1989): 171–212.

2. J. F. Cove and B. C. Walsh, "Online Text Retrieval via Browsing," *Information Processing and Management* 24, no. 1 (1998): 31–37; M. Bates, "The Design of Browsing and Berrypicking Techniques for the Online Search Interface," *Online Review* 13, no. 5 (1989): 407–424.

3. D. Bawden, "Browsing: Theory and Practice," *Perspectives in Information Management* 3, no. 1 (1993): 71–85.

4. X. Lin, "Graphical Table of Contents," *Proceedings of the First ACM International Conference on Digital Libraries* (New York: ACM Pr., 1996): 45–53.

5. X. Lin, "Map Displays for Information Retrieval," *Journal of the American Society for Information Science* 48, no. 1 (1997): 40–54.

6. T. Kohonen, *Self-organizing Maps* (Berlin: Springer, 1997).

7. L. Egghe and R. Rousseau, *Introduction to Informetrics: Quantitative Methods in Library, Documentation, and Information Science* (Amsterdam: Elsevier, 1990); M. Kinnucan, M. Nelson, and B. Allen, "Statistical Methods in Information Science Research," *Annual Review of Information Science and Technology* 22 (1987): 147–78.

8. Egghe and Rousseau, *Introduction to Informetrics.*

9. Kinnucan, Nelson, and Allen, "Statistical Methods in Information Science Research."

10. H. D. Howard and K. W. McCain, "Visualizing a Discipline," *Journal of the American Society for Information Science (JASIS)* 49, no. 4 (1998): 327–55.

11. F. Moya-Anegón, E. Jiménez-Contreras, and M. D. L. Moneda-Corrochano, "Research Fronts in Library and Information Science in Spain (1985–1994)," *Scientometrics* 42, no. 2 (1998): 229–46.

12. Howard and McCain, "Visualizing a Discipline"; Moya-Anegón, Jimé-

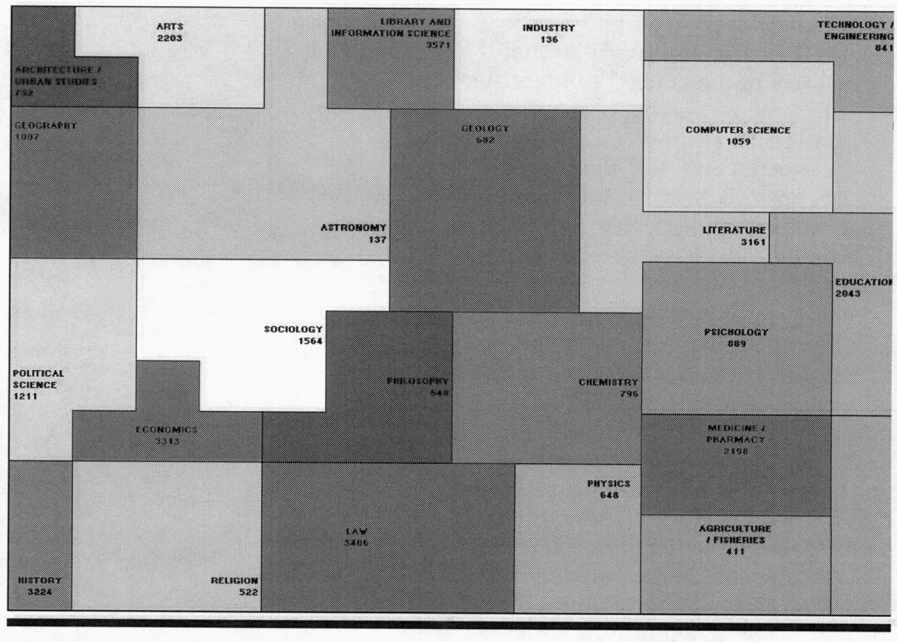


Figure 2. SOM map

nez-Contreras, and Moneda-Corrochano, "Research Fronts."

13. Kohonen, *Self-organizing Maps*; T. Kohonen, "Self-organized Formation of Topological Correct Feature Maps," *Biological Cybernetics* 43 (1982): 59–69.

14. H. D. White, X. Lin and K. W. McCain, "Two Modes of Automated Domain Analysis: Multidimensional Scaling vs. Kohonen Feature Mapping of Information Science Authors," *Structures and Relations in Knowledge Organization: Proccedings of the Fifth International ISKO Conference* (Wurzberg, Germany: Ergon Verlag, 1998); S. Kaski, K. Lagus, T. Honkela, and T. Kohonen, "Statistical Aspects of the WEBSOM System in Organizing Document Collections," *Computer Science and Statistics* (Fairfax Station, Va.: Interface Foundation of North America, 1998), 281–90; T. Honkela, V. Pulkki, and T. Kohonen, "Contextual Relations of Words in Grimm Tales, Analysed by Self-organizing Map," *Proceedings of the International Conference on Artificial Neural Networks ICANN-95* (Paris: EC2 et Cie, 1995); F. Moya-Anegón, V. Herrero-Solana, and V. Guerrero-Bote, "Virtual Reality Interface for Accessing Electronic Information," *Library and Information Research News* 22, no. 71 (1998): 34–39; F. Moya-Anegón et al., "NeuroISOC: un modelo de red neuronal para la representación del conocimiento," *Actas del IV Congreso ISKO-España EOCONSID '99* (Granada, Spain: ISKO-España, 1999), 151–56; V. Guerrero-Bote and F. Moya-Anegón, "Reduction of the Dimension of a Document Space Using the Fuzzified Output of a Kohonen Network," *Journal of the American Society for Information Science and Technology* 52, no. 14 (2001): 1234–41.

15. H. D. White and B. C. Griffith, "Author Co-citation: A Literature Measure of Intellectual Structure," *Journal of the American Society for Information Science* 32, no. 3 (1981): 163–71.

16. Ward's method, based on the minimum variance principle, is a specific method used in cluster analysis for dendogram construction. A dendogram is a branching, tree-like diagram based on one or several criteria that can be used to illustrate the relationships between elements.