January 2001

Collection Development in the New Millennium–Evaluating, Selecting, Annotating, Organizing for Ease of Access, Reevaluating, and Updating Electronic Resources

Virginia A. Baldwin
University of Nebraska-Lincoln, vbaldwin2@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/libraryscience
Part of the Library and Information Science Commons

http://digitalcommons.unl.edu/libraryscience/52

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, UNL Libraries by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Collection Development in the New Millennium—Evaluating, Selecting, Annotating, Organizing for Ease of Access, Reevaluating, and Updating Electronic Resources

Virginia Baldwin

INTRODUCTION

The literature of the early 1990s suggested changes in the structure of collection management itself in order to accommodate the changes both now occurring and those foreseen. Sheila Creth wrote in 1991 “Until recently, the primary focus of collection development has been on building collections.”\(^1\) She outlined the functions of collection management in academic libraries to include: selection of materials, weeding, preservation, liaison with faculty and academic departments, reference and user education, fiscal responsibility, and policy development. Creth saw the necessity for a basic change in the library’s organization structure. “According to the principles of organizational design, the traditional structure of the university library is functional in nature . . . Change has become the common denominator for the university library.”\(^2\) Therefore, Creth called for a collection management team structure of subject librarians with subject expertise to provide the basis for accommodating these changes. In 1989 Atkin-

---

\(^1\) Baldwin, Virginia. Co-published simultaneously in Collection Management (The Haworth Information Press, an imprint of The Haworth Press, Inc.) Vol. 25, No. 1/2, 2000, pp. 67-96; and: Electronic Collection Management (ed. Suzan D. McGinnis) The Haworth Information Press, an imprint of The Haworth Press, Inc., 2000, pp. 67-96. Single or multiple copies of this article are available for a fee from The Haworth Document Delivery Service [1-800-342-9678, 9:00 a.m. - 5:00 p.m. (EST). E-mail address: getinfo@haworthpressinc.com].

© 2000 by The Haworth Press, Inc. All rights reserved.
of responsibilities by subject in favor of subdividing subjects according to functions. The functions suggested by Atkinson were notification, documentation, instructional, historical, and bibliographical. These correspond to five information sources:

1. Notification sources—journal articles and monographs written by scholars for other scholars.
2. Documentation sources—all primary materials. These will differ by discipline.
3. Instructional sources—summaries of knowledge such as textbooks or manuals.
4. Historical sources—sources that may be needed one day for historical research.
5. Bibliographical sources—those that organize and provide access to the other sources.

In the year 2000 and beyond, we can most certainly see an impact from electronic resources on all five functional areas on Atkinson’s list, with historical sources being an especially elusive one considering the ever-increasing quantity of non-archived electronic information. Creth’s team approach is one frequently used as we see more and more libraries tackle the function of providing organized access to web resources using teams consisting of subject librarians, computer specialists, educational technologists, and graduate students.

In early 1995 a group of librarians from the University of Michigan School of Information developed a virtual-library, called the Internet Public library in response to the proliferation of unorganized resources on the World Wide Web. This event marked a breakthrough team response to rapidly developing changes in the information world. During this time period OCLC created the “NetFirst” database as a way to provide bibliographic records for Web sites.

Responses to the Internet began long before these dramatic developments. Back in the days before gophers were available we telnetted to information sources and used ftp to download. When gophers were developed, we began to organize gopher sites for our patron’s use. It soon became clear, as Susan K. Martin noted in her 1996 article on organizing Internet collections, that to ignore the world of electronic materials in collection development would be to “ignore an increasingly large proportion of the knowledge of any discipline. . . . Users
should know that a particular electronic resource has been examined and identified as appropriate for their library ‘collection’ by the same person who selects their books and journals . . .”4 And, by the way, as Martin proposes, we need to keep the same statistics for these resources that we keep for traditional resources.

This chapter tracks the methods that members of the library profession have devised to deal with new electronic resources and with the continuous change in software developments to manage them. While not every individual library or librarian has managed immediate assimilation of these changes as they have occurred, as a profession we have dealt with them admirably. The question is whether or not a time will come when a sort of equilibrium is reached—a time when our libraries are organized enough in the electronic realm so that new developments will require only a modicum of accommodation.

**SHALL WE GO FOR THE GOPHER?**

The University of Minnesota developed the Gopher computer software in 1992. Soon, like their biological counterpart, Gophers began springing up in libraries everywhere. A useful tool for organizing Internet access points in a hierarchical structure, Gophers were soon used to link library patrons with external information sources such as catalogs from other libraries, as well as internal directories and databases, such as, in the case of Eastern Illinois University, to the school newspaper, the *Daily Eastern News* in electronic format. The Gopher allows the user to seamlessly telnet, download, and break connection, thereby enabling movement from one source to another with just the pressing of an arrow key or the appropriate letter key and the enter button. A “/” at the end of a line indicates that the line is a category that will lead to another hierarchical tree of sites.

The hierarchical structure of Gopher enabled Librarians to begin to organize sites by categories, and collection development librarians soon found themselves wanting to use the software to provide organized subject access to article databases, electronic journals and other electronic resources, library catalogs and local resources. In 1993 Lieberman and Rich noted, “New and innovative approaches are already redefining the way information specialists retrieve, organize and disseminate information.”5

Early on, the issue of access versus ownership became a trouble
spot. It was difficult, given the traditional librarian mindset, to provide pointers to information sources over which there was no control. Providing “access to electronic journals without downloading files and maintaining local archives . . . represents a major philosophical shift for libraries that borders on an abdication of their primary mission.”6 With the decision not to archive, however, came the obligation to regularly verify that the remote archive is being maintained and kept active. Additional problems associated with pointing rather than downloading and archiving include down servers at the site pointed to and pointers eliminated from a remote Gopher.

Draper Laboratory Library implemented a Gopher in January of 1994 through a team effort between Library and Computer Services personnel. Because of the inherent disorganization of the Internet’s resources, Draper Library customers looked to the library to provide a simple organized method of access to this Internet information. The librarians involved were reticent about launching the Gopher until the menu structure was perfected and included the most appropriate set of resources. In an article about the Draper Gopher, Rotman, Spinner, and Williams of Draper Laboratory described how the team reconciled a problem continually experienced in the realm of Internet Resource Collection Development. “By its very nature, however, gopher is always evolving. We could have kept adding new resources and adjusting the menu structure indefinitely and it still would not be perfect. We eventually realized that it was better to launch the gopher, receive feedback from end-users and make adjustments as we continued the process.”7

Another important issue that is well remembered by those of us who sought to organize resources using Gopher is that of where specifically in a remote Gopher site we should establish a pointer. Should we merely point to a remote site and then let the user find the gems that are accessible from that site, or rather, point to the individual title level, be it a document, file, newsletter, etc.? Probably we are derelict in our duties as collection development librarians if we do the former. This is expressed admirably in “The Internet and Collection Development: Mainstreaming Selection of Internet Resources” regarding Gopher development at Cornell University’s Mann Library “We believe that title by title selection of high quality resources is one of the most important values librarians can add in providing access to information
resources including those accessible via the Internet. A careful selection of resources is the touchstone of the electronic library."8

Two subject specific Gophers were produced at the National Library of Medicine (NLM) to provide organized access for both internal and external resources.9 The TEHIP (Toxicology and Environmental Health Information Program) Gopher was implemented as a result of a long range planning panel issued in September of 1993 and the AIDS Gopher was initiated in response to discussions that occurred at a conference in June of 1993. After addressing issues of stability, reliability, and currency of information, developers considered numerous organizational issues. Structures given consideration included the separation of NLM resources from non-NLM resources, separation of resources by producer, separation of resources by subject terms, such as Medical Subject Headings, and separation of resources by format (files, Gophers, listservs, etc.). Each Gopher entry is limited to 70 or less characters of text. Therefore, an explanation of the organizational philosophy of the Gopher occurred in a text file accessed from the first item on the main menu of the Gopher.

Another issue to be addressed by the NLM Gopher was the number of levels of menus that could be presented without causing the user to get lost in the Gopher. Is it better to list a large number of items on a single menu or to have short menus with many levels to get to an information source? Two things were clear, that it was best to (1) create menus that can be viewed on a single screen and (2) have multiple pathways to the same source levels. Other organizational considerations were given by Grajek and Marone. "Standardize menu features in fixed positions because consistency is important. For example, information about the gopher could be in the first position, searching in the second, etc. The most important and frequently used items should be higher up in the menu. Include a top-level menu of shortcuts to the most popular services, such as telephone directories, schedules of events and the library catalog and databases, since this facilitates access to them."10

These are the kinds of organizational concerns, which were addressed, in the early 1990s by librarians everywhere who were dedicated to providing their patrons with extensive access to this new world of electronic resources. In 1994 the University of California, Irvine “Virtual Reference Collection” of Internet resources included one of the most frequently consulted gopher resources, called the
"Virtual Reference Desk" which was designed and maintained by a librarian. This site and its organizational structure became the basis for a Web-based Virtual Reference Collection.11

Still true today, even back in the gopher age the question was being raised as to why "so many libraries and schools (were) creating gophers that all link to the same Internet information."12 "The need is not just for access to information, but for access to organized and well-maintained information"13, so why not pool resources so that each piece of information is organized and maintained by only one librarian? A case in point is the Oregon Online Project for the State of Oregon.14 The Gopher for this system is browsable at gopher://go- pher.state.or.us. Designed around an automated document management concept, the project volunteers were from 25 different state agencies. The difficulty here was obtaining agreement from the contributors as to the topical hierarchy and on a standardized document format. The outcome was an automated cooperative development of information sources on a Gopher.

Internet usage increased dramatically at Yale in October 1993 when their gopher was formally announced. The "greatest draw for external users seems to be not local information but our organization and maintenance of links to external information."15 As their portion of the contribution of several universities who are "voluntarily apportioning maintenance of global gopher resources, . . . Rice University and Michigan State University maintain a list of subject trees gleaned from searches of discipline-specific gophers at more than a dozen institutions."16

On the other hand, from a librarian’s perspective, the value of a locally established library gopher may lie more in its uniqueness rather than the collaborative efforts that may have produced it. Quite possibly, the dramatic changes that ensued may have occurred before any kind of a compromise was reached, but a locally modified gopher that had been collaboratively established may have been a workable solution and the best of both worlds.

One of the downsides of the Gopher hierarchy is that movement through the trees to the eventual location of a site of interest required either a remarkable memory for subsequent return to the site, or lengthy recording of linking information. This was soon resolved with a bookmarking capability that was available to each individual user. For the user with his or her own computer, now a personal, individual-
ized organization structure was available for more convenient access to frequently used sites.

Two other network access tools of note that were developed in the early 1990s are LIBS from Mark Resmer at Sonoma State University and HYTELENT from Peter Scott at the University of Saskatchewan, Canada. Both provide seamless telnet access to numerous OPAC’s, databases, information services, and campus-wide information systems. HYTELENT does not have the subject-based approach to accessing information resources that LIBS had. The LIBS software could be downloaded by anonymous FTP and customized for local usage. The main menu of on-line services available through the Internet had six entries. The first two entries were for library catalogs in the United States and other countries. The next three were for information systems, services, and access tools. The last was information for first time users of the program.

Most of the information in the next menu level was in alphabetical order. As an example, the second menu level under Databases and Information Services was topical in organization and included six broad areas such as Business, Education, Science and Weather, and a 7th area, Other.

HYTELENT is still available as an http protocol site. In 1997, Peter Scott announced the discontinuation of its maintenance. Figure 1 depicts the main screen of the Hytelnet Web site “HYTELENT on the WorldWideWeb” (http://www.lights.com/hytelnet/).

FIGURE 1. Portion of the HYTELENT Web Interface Main Screen as of December 1999
Librarians everywhere were realizing the value of providing access to these gophers along with access to their OPACs and electronic databases. One way to provide access was to set the client’s default “home” gopher to a designated gopher site. Another was to establish a gopher bookmark. With the advent of browsers and the proliferation of electronic indexes, abstracts, and full text databases, librarians have devised ways to direct the patron to resources appropriate to his or her needs. Methods of organizing electronic resources have become increasingly sophisticated as librarians have merged networked CD-ROM databases with telnet resources and added full text databases, electronic journals, and basic Internet resources as each of these formats has become available. Ways of presentation of these resources vary with libraries. Libraries have attempted to organize these resources by subject areas, by full text versus citation and abstract only, by monographic versus serial titles and other methods.

**BOOKMARKING**

The introduction of Mosaic, developed in 1993 at the National Center for Supercomputer Applications (NCSA), a research institute at the University of Illinois, brought about the obsolescence of gophers. At first we traced through web links with consternation and near disbelief. What were we to do now? While hyperlinking made accessing a site remarkably easy after the experience of telnetting and following gopher links, how were we to have any organized structure for accessing sites that were useful in our subject areas and for Reference work? Soon bookmarking of web browser URL’s appeared and there was optimism again, until we had an unmanageable string of bookmarks. Then, along came bookmark folders (for Netscape Navigator and Favorites for Internet Explorer). We were not certain if we should be thankful or worried that someone might be taking over our jobs.

Problems and resolutions continued to abound. In 1996 when access to electronic journals as part of digital library projects was being structured, Barber acknowledged a difficulty with using bookmarking to facilitate the monitoring of the “arrival” of new journal issues.^{19} Ideally, bookmarking could be used to bring up the site page that lists journal issues available. This would make it relatively easy to check to see if a new issue had been released. However, many publishers did
not maintain this information on static web pages, and in some cases, access to these sites were as part of a searching session, causing access by bookmark link to bring up a session expired notification.

In “Never Lose Sight of Your Site,” Terrence Young suggested using folder names that parallel the subject headings in the library catalog. Young furthers the analogy by saying that Bookmarks and Favorites are similar to the catalog records. Among the software described in this article is Smart Bookmarks which “automatically checks Web sites and proactively notifies you not only when something has changed, but also tells you what has changed. . . . Using Smart Bookmarks you can import existing bookmarks, add custom descriptions and comments to bookmarks, define categories and folders, and move bookmarks from one folder to another.”

For Mac users, there is Macuser’s Site Seer that keeps track of every site visited and creates drag-and-drop-aware windows for bookmarks. WordPerfect has a filing system named Cardfile that can be used to create discipline specific files for organizing URL’s. The files can be “filled” by copying and pasting.

Another major milestone, reached in mid-1993, grew out of a file of some 750 bookmarks that was started in 1990 as a gopher bookmark file by Carole Leita at the Berkeley Public Library (BPL). BPL mounted a web site on the City of Berkeley’s new web server as the “Berkeley Public Library Index to the Internet.” In March of 1997 the BPL index moved to the UC-Berkeley Digital Library SunSITE and the Librarians’ Index to the Internet (LII) was born. The LII is both searchable and browsable, with over 40 major categories that are modified Library of Congress (LC) subject classification categories and cross-references. Organizing the Internet began to take on a new meaning. Figure 2 shows the categories on the home page of the site (http://sunsite.berkeley.edu/InternetIndex/).

TAXONOMIES OF INTERNET RESOURCES

Because of the great variety of types and formats of Internet resources, many librarians have developed schemes especially designed for organizing these resources. In early 1993, four librarians at Cornell University’s Mann Library began a project of selecting Internet resources of potential interest to library clientele. Librarians at Cornell found that “the development of classification schemes (is) a powerful
FIGURE 2. Categories of Links on the Librarians’ Index to the Internet Main Screen as of December 1999
tool in adapting the principles of collection development to new forms of publication” (p. 281). Mann Library employed a “genre” model for selecting electronic publications to allow the grouping of resources into logical units of analysis in order to focus staff expertise on certain information types. In doing this, a taxonomy was devised to categorize Internet resources into 15 categories and 21 subcategories. Categories included reference, monographs, serials, gophers, gateways, literature and book reviews, graphic images, sound, videoconferences, and selection tools. The resulting taxonomy was strongly reflective of Mann Library’s particular subject interests. However, developers “found that overall the process of devising and working with such a classification scheme was an invaluable exercise in thinking through (their) . . . approach to Internet selection.

“Grouping titles into taxonomic categories enabled us to evaluate and compare, from a collection development perspective, a set of resources with similar characteristics.”24 Furthermore, this intellectual process served to clarify the selection criteria for collection development of Internet resources.

OCLC took the approach of cataloging Internet resources in ways similar to its traditional cataloging system. In 1996, with its FirstSearch system of databases of article indexes now enjoying widespread acceptance, OCLC added a new database, NetFirst, composed strictly of records of web resources. Each record was in traditional FirstSearch format, was retrievable through a standard search, and consisted of author, title, subject, abstract, and other fields. A second OCLC project, the Internet Cataloging project (http://www.oclc.org/oclc/man/catproj/catcall.htm) was funded by the U.S. Department of Education and it began in 1994. Librarians from 231 libraries created its MARC format records.25

Anne Callery, cataloger at Yahoo! Inc., described the Yahoo! Approach to Web site organization at the “Untangling the Web” conference in Santa Barbara, California.26 When Yahoo! created its subject directory its founders realized that as a stand-alone resource of strictly Internet sites, there was no need to integrate its sites with other longstanding resources. Therefore, Yahoo! founders devised its own classification system. Furthermore, because of the changeable nature of the Internet object being cataloged, it made no sense to create anything as structured and complex as a MARC record. Furthermore, they reasoned, access to a site was so easy compared to a trip to the stacks that
the user could just as easily jump in and make a decision about the usefulness of the contents of a given Web site. Basically, the person who submits a site to Yahoo! suggests a category. The subject hierarchy for Yahoo! is a “bottom up” approach, since it is dictated by whatever is submitted, with new subcategories often being created as a result of a submission. Yahoo! does have a structure for regional divisions, and all commercial sites are added under Business and Economy, in either Companies or Products & Services. Yahoo! accesses sites only at the top level of a document, or its significant sections. A Web search engine, on the other hand, may pull up dozens of resources within one given site. Figure 3 depicts The Yahoo! Subject Directory hierarchy.

Perhaps the obvious alternative to the MARC record of traditional library resources and the “bottom up” approach of Yahoo! is the creation of a core of metadata elements specifically created for tagging Internet and other electronic resources. Just such an effort began in March of 1995 at a Metadata Workshop sponsored by OCLC and the National Center for Supercomputing Applications, and attended by librarians and other relevant professionals from a dozen countries. The result was the Dublin Core element set of initially 13, since expanded

FIGURE 3. The Yahoo! Subject Directory Hierarchy as of December 1999
to 15 core elements: title, creator, subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage, and rights (management).  

The 15 elements or metatags, “when implemented, form the Internet’s equivalent of a bibliographic record for electronic documents. . . . The idea is that the author or creator tags electronic documents according to a set of standards. . . . The search engines then need to be able to retrieve documents according to these metatags.” Younger called for libraries to “identify incentives to encourage information creators and producers to incorporate standard metadata in their publications . . . (such as) copyright or patent registration (incentives and) revenue derived from increased access.”

The search engine INFOMINE began as a list of Web sites at the University of California, Riverside. Its entries are organized with modified LC subject headings and University of California librarians select the web sites (http://infomine.ucr.edu). There are nine basic subject areas and each subject area has five browse features, “What’s New,” “Table of Contents” (browse by subject and title), “Subject,” “Keyword,” and “Title.” The heavily funded Scout Report (http://scout.cs.wisc.edu) is housed in the Computer Science Department of the University of Wisconsin. Those reports that are cataloged are done so with LC. Since library classification doesn’t necessarily translate to the Internet, a new classification system is under investigation. Michigan Electronic Library’s Sue Davidsen is wary of traditional Dewey and LC classification since they were created for a physical object shelved in a physical location. Some of the many predictions and proposals for an individual library’s approach to organizing information sources include multi-tiered approaches, replacement of URL lists with distinct records with annotations, creation of an Internet subject thesaurus and subject gateways. One of many Multi-tiered schemes was suggested by cataloging expert Michael Gorman:

1. Full cataloging
2. Enriched Dublin core records
3. Minimal Dublin Core records, and
4. Reliance on unstructured full-text keyword searching.

For the information community as a whole, foreseen are automating functions and some melding of various methods of cataloging and access.
In 1991, the Committee on Institutional Cooperation (CIC, a consortium of the members of the Big Ten athletic conference plus the University of Chicago) created a protocol for mining the Internet for electronic journals and organizing them onto a gopher site for easy access. Journals were organized alphabetically by title and also by subject. Initially, only freely distributed electronic journals were included in the collection (now a web site at http://ejournals.cic.net/). This was a boon to librarians for it allowed them to include the CIC URL and a brief description of what was contained therein to give immediate and organized access to a wealth of free scholarly journal articles in a variety of subject areas.

In 1993, University of Michigan librarians created dozens of pathfinders on a gopher site, then organized them by subject areas on the University of Michigan server at http://www.lib.umich.edu/chhome.html. This site became the Argus Clearinghouse, a widely used subject directory and search engine for the Internet.

In November of 1995, Cyberstacks(sm), an organization of significant science and technology resources on the Internet using first level LC Classification, was formally established on the home page server at Iowa State University (http://www.public.iastate.edu/~CYBERSTACKS/homepage.html). Cyberstacks(sm) offers the following “main menu”:

- G Geography, Anthropology and Recreation
- H Social Sciences
- J Political Science
- K Law
- Q Science
- R Medicine
- S Agriculture
- T Technology
- U Military Science
- V Naval Science

Currently, J, K, U, and V are not linked to Web resources. A second tier of organization is available for the remaining LC Classification schedules, for example, for Q, which is the only class that at the time of this writing has links to all second tier classes (twelve classes, QA through QR). At the next tier under these secondary headings LC Classification subjects and number range are given for each of the
twelve levels, but contain links to Web resources for only selected ranges.

Cyberstacks(sm) also has a cross-classification index which is a subject listing in alphabetical order, and, in response to preliminary user feedback, each resource was incorporated into a newly created Title Index in 1996. The Title Index contains all resources that have been assessed including those which have not yet been described, categorized or classified, or otherwise fully integrated within the collection. An icon distinguishes between the two levels. The record for each categorized website in Cyberstacks(sm) is formatted according to the Web site information depicted in Figure 4.

Out of Great Britain came an organization of resources by Dewey Classification [http://bubl.ac.uk/link/](http://bubl.ac.uk/link/). Known as BUBL LINK, or LINK, all resources are catalogued using the Dewey Decimal Classification (DDC). On the main screen many organizational approaches are available for the catalogued sites. In addition to a link to “Dewey,” there are links to a search engine, to a hierarchical alphabetical list by

**FIGURE 4. Record Format for Cyberstacks(sm) Categorized Web Site**

<table>
<thead>
<tr>
<th>Record Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>We have attempted to provide the following categories of information about each selected resource, in the order indicated:</td>
</tr>
<tr>
<td>1) Specific Library of Congress Classification</td>
</tr>
<tr>
<td>2) Resource Title/Name</td>
</tr>
<tr>
<td>3) Full and Correct URL</td>
</tr>
<tr>
<td>4) Record Summary</td>
</tr>
<tr>
<td>a) Subject Coverage</td>
</tr>
<tr>
<td>b) Resource Size/Number of Entries</td>
</tr>
<tr>
<td>c) Record Structure</td>
</tr>
<tr>
<td>d) Special Features</td>
</tr>
<tr>
<td>e) Miscellaneous Notes</td>
</tr>
<tr>
<td>f) Source Acknowledgement</td>
</tr>
<tr>
<td>5) Search Instructions</td>
</tr>
<tr>
<td>6) Preferred Form of Contributor’s Name and Affiliation</td>
</tr>
</tbody>
</table>

For each record summary, when and where possible, we have excerpted the summary data from the original resource to provide sufficient information about its subject coverage, and other features, to enable users to judge a resource’s potential usefulness.
subject ("subject menus") which leads through the Dewey classification scheme to cataloging records of sites, and "countries" which does the same for sites relative to a given country. There is also an alphabetical link to subjects through links to letters in the alphabet. Each resulting record contains the following basic elements: Title, Description, Author, Subjects, DeweyClass, ResourceType, Location, and, occasionally, Date last checked.

Also on the initial screen is the Internet-subject-directory-appearing list of topics and subtopics that do not correspond to the structure of the DDC. However, each topic or subtopic leads hierarchically to cataloging records of sites (and these are classified according to the DDC).

CyberDewey (http://ivory.lm.com/~mundie/CyberDewey/CyberDewey.html) is another site that lists Internet sites that are organized using the Dewey Decimal Classification scheme. It consists of a listing of the 10 classes and one hundred divisions of the DDC with the number of links under each division given in parentheses after the division.

In addition to Cyberstacks(sm), Iowa State University is home to another site, "Beyond Bookmarks: Schemes for Organizing the Web," that is compiled and organized by Gerry McKiernan (http://www.iastate.edu/~CYBERSTACKS/CTW.htm). This site lists classification systems and controlled vocabularies that are used on the Web, and provides links to sites that employ them. Among the classification systems listed are Alphabetic, Numeric, Engineering Information Classification Codes, Mathematics Subject Classification, Universal Decimal Classification (UDC), Alphanumeric, Library of Congress Classification, and National Library of Medicine. Among the controlled vocabularies listed are Library of Congress Subject Headings, Medical Subject Headings, and National Library of Medicine.

More recently, in 1997, the Northern Light search engine took a completely new tact by creating a scheme to allow users to "narrow" their initial search by choosing from among a set of folders created at the time of the search to represent concepts that reoccur within the documents retrieved. Librarians developed the keywords or categories depicted by the folders. The term "custom search folder" is trademarked and NorthernLight patented the process. Figure 5 depicts the 1st and 2nd tiers of folders that result from the search "Kolb learning cycle" and the selection of "Teachers and Teaching" from the 1st tier.

The Internet Quick Reference (IQR), called "A Seamless Web-based Library" by its developer, Steve Weiss, a Document and
Reference Librarian at Utah State University, contains Internet links that are organized in a table of contents. The site is browsable in the sense that the table of contents and the annotated links to which its elements point are all downloaded upon entry into the site. Included are links to sites of particular value to Utah State University library patrons, and to some that are only available to those patrons. Figure 6 shows the Table of Contents on the main screen (http://cc.usu.edu/~stewei/hot.htm).

In late 1998, OCLC began a third Internet project when it announced plans to begin the Cooperative Online Resource Catalog (CORC) project (http://www.oclc.org/oclc/research/projects/corc).

As a response to the thousands of individual library attempts to provide lists of web links deemed useful to their patrons, OCLC’s CORC is designed to facilitate cooperative development of these individual efforts. “CORC’s objective is to provide the infrastructure support so that libraries can build the gateways automatically.”37 But, we are jumping ahead. Let us now look at some of these gateway-building and other individual organizing activities.

LIBRARIANS ORGANIZING WEB LINKS FOR PATRON ACCESS

“The Internet has changed the concept of ‘place’ in relation to both collections and collectors. In the electronic world, it has become less
important WHERE a document resides and more important to have reliable, well-organized (and presented) access to it. We want to know who produced it, who identified it as valuable, and who selected it for our use, but that person does not have to sit at the desk next to us. We no longer need to ‘own’ a physical manifestation of the information in our private institutional domain, but we must provide the appropriate technological and organizational infrastructure to access it reliably.”

Several approaches have been taken to dispense information about useful web sites to library patrons. Reference librarians together with subject specialists and often computer technology personnel have organized front-end menus that provide access to the library’s OPAC and article databases, their internal resources and CD-ROM’s, as well as Internet links and links to electronic journals available either through subscription or those with free access. Hypertext subject guides are proliferating in a variety of forms, breadth, and depth of inclusion. Most include Internet links considered valuable for the subject area. More and more of these links include annotations of contents, response time, disclaimers given, and other information about the site as well.

As lists of links grew larger, they began to require lengthy scrolling through screen after screen to view the entire list. So, librarians began to organize the links in various ways, resulting in a variety of diverse schemes almost as large in number as the number of libraries employing them. As we browsed through other web sites we found ever better
ways to organize these links, and contemplated or produced these sometimes massive changes.

In “Finding Our Way,” Kathleen Kluegel describes the all-too-frequently occurring experience of navigating through layer-upon-layer of menus, only to find we’ve gone down the wrong path and must backtrack and try another path. She suggests the tourist guide analogy, with “multiple modes of access: a spot for direct entry of a title if it is known; a list of resources for those who would recognize the name of the needed resource; a functional choice for those seeking resources of a particular type; as well as a more guided selection process.”

Very often a frequently linked site will have a very useful internal link. Annotating and providing a separate link for this internal element can make all the difference to the harried patron.

In “Netting Political Science–Finding Resources on the Web,” Lucia Snowhill gave numerous examples of Internet link organization. For sites listed in the international relations subject area, some link arrangements were by source, type of information, topics such as area studies, economic development human rights, documentary sources, major theorists vs. political theory, etc. and annotations included information about these types of arrangements.

The categories of links appear to be endless. Department links, links to links and/or gateways from other libraries, preprint sites, associations, listservs, “other” formats, newsletters, e-journals appropriate to the subject areas, links to lists of print resources for various types of reference sources, and links to databases appropriate to the subject area are some of the more common categories of links. All of the categories in the Mann Library taxonomy, mentioned earlier have undoubtedly been used at various sites.

Going back to our initial look at library organization, many issues become relevant. Should there be a standard template for organization of subject specific sites?

**Gateways**

The web-based OPAC is a candidate for the library’s gateway to Internet resources. Libraries everywhere have begun using the Electronic Location and Access (856) field of the USMARC record to provide links to electronic resources that have been purchased through standard and cooperative acquisition processes. Some libraries have begun to include free resources that have been selected by collection
development specialists. Many questions surround these practices. Links can be provided to the electronic (Internet) versions of books, and this can be done regardless of whether the book is available in print in the library. Refer to Chapter 10 in this book for Tom Peters’ discussion of the rapid development of computer networks and digital scholarly and academic information resources. In regard to electronic journals, the issue is whether the OPAC link would be to the Internet address of the journal publisher, to a separate file of online journals, or whether the record will be hotlinked to the full text. In fact, a fourth possibility is to provide a link to the publisher’s journal search screen for those publishers who provide that capability.

Furthermore, serials cataloger Amanda Xu may have the ultimate solution in the eventuality of a successful Dublin Core type tagging of Internet resources. She considers the possibility of a library OPAC serving both as a gateway to Internet metadata repositories and to its existing databases. Then she goes on to say that library systems can harvest this metadata and she discusses the potential for developing a library metadata conversion system. In fact, the Internet is approaching the size that will make it necessary to have indexing done only on document metadata rather than the document contents because of the enormous number of documents on the Web, with no end of the proliferation in sight. Xu “analyzes a library metadata conversion system that will be able to extract selected incoming external metadata from Internet resources, convert it into USMARC format, and integrate it with existing library databases automatically” (p. 195).

Regina Reynolds in “Inventory List or Information Gateway? The Role of the Catalog in the Digital Age,” concluded that the shift that is occurring from the catalog as inventory list to information gateway was both inevitable and desirable and that the cataloging record must make it clear to the patron whether or not the title is actually owned by the library. Reynolds also discussed a Library of Congress initiative to develop a cooperative archiving of electronic journals to which individual libraries could point in their OPAC’s, as opposed to pointing to the publisher’s web site.

But apart from the predictable and possibly justified aversion of librarians to catalog and place on their OPAC an Internet resource that may be here today and gone tomorrow, or changed so substantially as to require recataloging, there are other reasons to provide separate, organized, lists of links to Internet resources, in addition to those that
are provided on subject-specific Web sites. Many of these reasons can be classified under the heading of Potential Ways A Patron Might Seek Access To A Library Recognized Web Site.

Library gateways abound. Often they are embedded in a set of headline links. A whole chapter could be written on a taxonomy of library gateways. One thing a library gateway can do that an OPAC search may not (at least not yet) be able to give us is a link by subject to a list of records of electronic-only resources. This is the converse of Reynolds’ suggestion “searching strategies should allow patrons to limit searches to a base catalog of items actually owned by the library.” Both approaches are valuable organizational access methods and do not supplant each other.

The aforementioned Committee on Institutional Cooperation (CIC) site was one of the first comprehensive, organized listings of electronic journals. From its beginnings as a gopher site in 1991, it gave both subject and title browse lists. Figure 7 depicts the Electronic Journal Collection menu that has not changed since its inception. The “Topic Browse” listing was by broad subject categories that arose from the subject areas represented by the journals available. Their web page states, “Ultimately, this collection aims to be an authoritative source of electronic research and academic serial publications—incorporating all freely distributed scholarly electronic journals available online. The CIC-EJC serves as the electronic journal collection for the CIC member university libraries. The collection is fully cataloged by the CIC member libraries, and records are contributed to the international bibliographic database OCLC. Ultimately, the collection will include...”

FIGURE 7. Committee on Institutional Cooperation Electronic Journal Collection Menu as of December 1999
electronic serials licensed only to the CIC member universities, and access to licensed publications may be restricted under the terms of the applicable license agreements. . . The CIC-EJC is a collaborative initiative between the librarians of the CIC member universities, CIC-Net, and the CIC Center for Library Initiatives.”

A very useful and much referenced gateway to electronic journals comes from the Colorado Alliance of Research Libraries. Figure 8 shows the search entry and the various links on the main screen. These links include alphabetical indexes by title and LC Subject Heading, two very useful organizational schemes. The link to “Other Directories of Electronic Journals” provides an annotated list of more than twenty sites that contain listings of electronic journals, including those from CIC and BUBL. Also useful are the “Directory of major publishers of electronic journals” and “Use this form to recommend new titles for inclusion” both of which are useful for locating new titles to be added to the electronic journal list.

The most thoughtfully designed gateways and the most highly linked OPAC cannot, however, overcome the many problems that the user will find upon arriving at a site. Barber identifies several. Once the patron clicks on a link the result may be one or more of the following:

1. Service at the site is temporarily unavailable.
2. Technical capabilities at the site may be inadequate.
3. The interface at the new site most likely will be unlike any ever before encountered.
4. Full-text searching or browsing will apply only to a limited set of articles.

The aggregation of Internet resources and journals from multiple publishers at one site will alleviate some of these problems. Is this strictly a publisher function? What kind of a role can the librarian have in making this a reality? Recently, three publishers announced the establishing of an agreement to provide hot links to each other’s cited articles. These hot links would only be activated to full text at a site where the patron has established access. This is progress.
THE EMERGING ROLE
OF COLLECTION DEVELOPMENT LIBRARIANS

The “here today, gone tomorrow” phenomenon of Internet sites is somewhat mollified by software available to check links. An example is the free link checker, NetMechanic.com, which may be used at some regular frequency to ensure all hotlinks are active. However, does not relieve the webliographer of the necessity to periodically look at each site linked to make certain the contents continue to be appropriate for the use to which they have been put. For a site to be adequately maintained, new links can be added when discovered, usually through sources such as listservs, library publications, books, etc. That is, until “the evolving information technology makes it obsolete—replaces it with something better—perhaps a hand-held, artificial intelligence voice recognition system which accesses a world of information from a universal database of knowledge” (U.S. Dept of Commerce, 1999). In his article, “The Seamless, Web-Based Library: A Meta Site for the 21st Century,” Weiss describes Web design considerations, many of which take on the magnitude of new and distinct roles for the librarian. In addition to all the organizational considerations, the librarian can take on a marketing role. The Web site can be registered with a search engine like Yahoo!. Other such sites can be
found through what else, but an Internet search. Other advertising methods mentioned by Weiss are a referral service like http://Recommend-It.com, contact with other Webmasters for mutual linking, and announcing on appropriate listserves.49

All of the preceding sections of this article make it clear that collection development librarians must not only keep abreast of new technologies, they must also learn about emerging technologies and predictions of them. Although we have coped admirably with changes as a profession, individually we have been inconsistent at best in using the various organizing tools that have come along for electronic resources.

Some thoughts for the evolving future of librarians’ relationship with electronic resources:

- Eventually, we may find that all computers will act as host computers in the future evolution of resource organization. All formats and types of information will be divided among the community of users and librarians for evolving and continual addition of resources to that collective universal database of knowledge using an all-encompassing classification scheme. OCLC’s CORC effort is analogous to this. Never fully realized with gopher, now we have the possibility of global and cooperative evaluation, annotation, and classification of resources accompanied by local selection and modification to accommodate local patron needs.

- While librarians and others have been developing a multitude of classification schemes, others are working on natural language processing and voice recognition systems. If Microsoft’s Natural Language Processing (NLP) engine were “used as a Web-search engine, a user could enter a query and have a reduced number of responses, unlike current search engines.”50

- We have been amply forewarned, even back in 1996. “By the end of the decade” (the arrival of perhaps 500,000 electronic serials will raise the question of) . . . “How will this new and potentially enormous and pervasive medium be integrated into existing structures of information organization and brokering? Libraries will be the key to the success or failure of electronic scholarly publishing. With proper proactive management, the success of this new medium will benefit both libraries and the academic community. However, if the academic community and libraries
fail to take the initiative in the development of this new medium, then we will be faced with a budgetary and organizational nightmare that may lead to an increased drain on declining acquisition budgets.”51

- “For the first time in over 200 years, the paper scholarly journal can be supplanted, or at least, supplemented in a significant way by the rise of network-based electronic journals and that this may lead to a new type of scholarly discourse. . . . I think that a paradigm shift in the process of scholarly discourse is on the horizon and that this will be accompanied by a fundamental shift in the technological basis of academic communication and publishing.”52

- “Librarians are certainly destined to play a key role in disseminating the publicly funded and freely accessible electronic scholarly information when, for the first time in history, we will have a seamless interface between the university and the community. This university-community interface will take the form of a supernetwork of community Freenets that are linked to the larger academic Net and thereby have direct access to the growing wealth of electronic information resources created by and for all peoples. Hopefully the current trends in network-based electronic publishing will continue in such a way that will foster the speedy arrival of such an information system.”53

- The issue of database design and user interfaces becomes more critical as our library becomes more virtual. “Our users will demand an interface that is intuitive, easy to navigate, and dynamic; one that provides a uniform gateway for the services offered by the virtual library.

“Will our current library instruction screens and tutorials evolve into a more complex system of instructional design? Will we become instructional designers as well, or at least become part of an instructional design team, perhaps composed of both librarians and (instructional design professionals).”54

Link Personalization

Our personal computers come with their own Internet organization system in the form of browser bookmarks. If we download Netscape, we receive an initial set of bookmark folders with it. In version 4.7, those folders are “Search, Directories, Banking and Finance, Business Resources, Computers and Technology, Education, Entertainment,
General News, Hobbies and Lifestyles, Local Information, Shopping, Sports, Travel and Leisure and My Stuff.” Is this a clue to the ultimate in resource organization and access? To quote Ken Winter in “My Library Can Help Your Library,” “The promise of personalization in library settings is two-fold: First, it can help save librarians the hassle of creating countless and redundant Web pages, which are much harder to maintain than they are to create. And who can honestly say the pages they’ve created make sense to most end users? After all, a Web page designed for everyone must necessarily incorporate a least-common-denominator approach, making it simplistic for experts yet still confusing to novices. Second, personalization allows for incredibly detailed target marketing of your library’s staff, services, and resources based on such factors as patrons’ unique interests, the types of sources they use most, their academic major, or any other factor that seems relevant.”

Yes, personalization to be most effective can and will be done at the individual level, to include even the faculty and every undergraduate in a large university. In his article Winter describes many such efforts being undertaken at various universities, including North Carolina State, Calpoly San Luis Obispo, Cornell, and UCLA. At UCLA, close to 100% of a student body of over 35,000 students have created a My.UCLA (http://my.ucla.edu) Web page and the pages are heavily and regularly used. With these site personalization resources, patrons can visit their site from their dorm or any other location. Their favorite sources will be arranged in a way that is most useful to them. Alerts can be provided not only of new journal issues, but also of new books in their field.

Can we, as librarians, ultimately expand this personalization scheme to include the electronic journals and academic sites that will be needed for the research interests of every library patron whether they be students, faculty, staff, or administration? We have seen so much arrive and be supplanted in the decade of the 1990s. We have created and embraced new approaches to information organization to the benefit of our patrons. Perhaps all that we have done will become obsolete. Perhaps it will lead us to a kind of equilibrium, one that will enable us to concentrate our efforts more on changes in the flow of scholarly information and new challenges in instructional design. In so doing, we will be more able to constantly keep the most useful information sources readily available to our patrons and teach them how best to access and use them.
NOTES


Figures in Scott’s article show snapshots of files in the early HYTELNET software. The first version was released in early 1991. Screens from Version 6.2 are captured in Scott’s article “Using HYTELNET to Access Internet Resources” in *The


22. Norman Oder, Holly Hinman, and Carole Leita, “Cataloging the Net: Can We Do It?,” Library Journal 123, no. 16 (1998): 47-51. The origin and operation of the Librarians’ Index to the Internet LII is described by founder Carole Leita with Holly Hinman in a sidebar to this article. The sidebar is entitled “A Public Librarian Helps Launch an Index.”


24. Ibid. 281.


27. The Dublin Core element set, version 1.1, is defined and described at http://purl.oclc.org/dc/documents/rec-dces-19990702.htm. A formal standard for the description of metadata elements is used.


31. Norman Oder, Holly Hinman, and Carole Leita, “Cataloging the Net: Can We Do It?”


44. Ibid. 77.


49. Ibid. 51-52.


52. Ibid. 143.

53. Ibid. 144.
Ibid. 65-67. For more information on this type of resource, visit the University of Washington University Libraries’ MyLibrary Project, “My Gateway” at http://www.lib.washington.edu/resource/help/MyGateway.html.