

**The Canebsco-OCULA
Award-Winning Student Essay**

**SUBJECT
ACCESS
for the World
Wide Web**

by Candice Dahl

The growth and popularity of the World Wide Web has made it a widely-used research tool among the staff, faculty and students of universities.

Even though these Web users often cannot find what they want on the Web, they still frequently turn to it for information. Consequently, the Web is a popular topic of discussion among those concerned with indexing and cataloguing information. Providing subject access to documents and other resources on the Web is not a simple task and it poses both familiar and new problems for cataloguers and indexers. One way to help people use the Web efficiently is to provide subject access to its content. While such a project can be implemented on a large scale, it can also be facilitated through the creation of subject guides (often mounted on a web page) for a specific group of users, such as a university community. In order to create both appropriate and effective subject guides, several factors must be considered. After discussing guidelines for deciding when it is appropriate for librarians to create their own subject guides, it will be argued here that the provision of subject access to the Web

One way to help people use the Web efficiently is to provide subject access to its content.

through such guides is influenced by the nature of the Web itself, the concept of collection development (particularly in relation to the issue of filtering), and structural and design considerations. All of these points should be kept in mind by those who wish to create subject guides, in order to increase their understanding and skill in this area of librarianship.

WHEN TO CREATE YOUR OWN GUIDE

Before taking on any project, librarians must determine both if and how it should be done. There is clearly a need to facilitate subject access to relevant Web resources, particularly in light of the frequent inadequacy and inefficiency of keyword searching, the simplicity of cer-

tain search engines, and the rate at which the Web is growing. However librarians should consider whether it is necessary for different librarians at different universities to create subject guides for the same subject area. They should assess what is available before undertaking the creation of new subject guides. Because it is very time consuming to create quality guides, librarians should strive to avoid duplication if an adequate guide already exists (Calcari and Wells, 28). Such an assessment will, in part, help librarians decide in which subject areas to develop their Web collections and to produce subject guides.

Another influencing factor in this decision is that, for a particular library's users, subject guides in certain areas will be more helpful and well-used than others. For example, it would only make sense to create a forensics subject guide if there was a forensics program at the university. Also, it may be the case that,

because of a user group's unique needs, a guide that would be useful has not been produced anywhere else. When deciding whether or not to make a subject-specific

guide, first determine if a need exists (or is likely to exist in the future) and whether or not the Web resources available for your guide add something unique to your collection (Guarnino, 97). If they do not, or if the topic is one in which the best information is available in books, then the guide may not be necessary. But if the provision of subject access to a Web collection adds something more to the print collection or has the potential to produce a guide that is better than existing guides, time should be taken to make the best subject guide available for users (Toub, 150).

THE NATURE OF THE WORLD WIDE WEB

The Web is dynamic and ever-

growing, in a way and at a rate that differs from most print collections in libraries. Therefore, while some issues relating to subject access remain relevant, other considerations must also be made. First of all, since the Web is "egalitarian and diverse" (Harris, 36) it grows in an uncontrolled and uncontrollable manner, unlike print collections in our libraries. It is, perhaps, exactly because of this egalitarianism, where traditional structures surrounding publication and peer review can be ignored by Web page authors, that the Web is somewhat resistant to attempts of imposing order; many people participate in Web authoring precisely because of the lack of such structures. This freedom, and the fluid nature of the Web that allows innumerable links between sites, permits pages to be less subject-specific and more varied than reviewed and published works. Consequently, providing subject access to the Web can be difficult.

This egalitarianism also allows for the exponential growth and changeability of the Web, which has made it impossible for those who wish to create structures for subject access to keep up. Although most cataloguing departments in libraries do fall behind, the size of the backlog is unlikely to be anywhere near the backlog for those cataloguing the Web according to subject. Another consequence of the magnitude of the Web is that it contains many things that library users do not even need subject access to, such as individual messages from online discussion groups (Harris 39). In contrast, our physical libraries do not usually contain items to which no one needs subject access; our collection development policies tend to prevent this from happening. Providing subject access to Web-based documents is also unique in that the actual content of the documents sometimes changes, unlike documents kept in physical libraries. On the Web these changes can occur without prior warning or notification after the fact. So, besides the fact that the document may get lost (something which does frequently happen in libraries) due to a change in location, its con-

tents may also expand or decrease to make previous attempts at subject description inadequate.

Another important point regarding the nature of the Web is that subject access provides an intellectual entry point into documents, and is different than mere physical access. Simply being connected to the internet, then, does not ensure adequate accessibility since "an intellectual organization is the only way in which a user can interact with [a] collection" (both print and otherwise) (Harris, 36). For this reason it is vital to engage in some degree of subject cataloguing. Just being able to get to a web page, without having a way to determine beforehand what may or may not be relevant, is not sufficient. This is particularly so when documents are included in search results based on the appearance of keywords within them, and when the number of items retrieved by a search engine is in the thousands.

There are now many ways in which to gain subject access to the Web, both through search engines with subject hierarchies, such as Yahoo!, and through subject guides created by people, where the Web sites listed are often both evaluated and summarised. Such endeavors are concerned with the 'big picture' — the entire Web — while subject guides created by academic librarians are often geared to their particular groups of users and their specific needs. When individual librarians wish to exercise their influence, the 'big picture' of the whole Web must necessarily be whittled down to a manageable 'little picture.' In physical libraries, where the acquisitions and size of the library reflect its limited budget and physical capacity, these parameters exist automatically and it is easier for librarians to have a handle on their collections. The need for this process in the electronic environment demonstrates another way in which providing subject access to the Web differs from doing so in traditional collections.

Clearly, when making subject guides for Web resources, certain issues that do not arise when dealing with print collections must be

acknowledged and accommodated because of the nature of the Web. Even so, the standard skills needed for accurate description and evaluation are still essential, since documents that are inappropriately classified might remain unknown to the users or be found only by those who are looking for something else. Whether a local subject guide is being made, or a clearinghouse is being assembled, librarians must be aware of both the differences and the similarities between providing subject access to print and Web-based texts.

COLLECTION DEVELOPMENT

Creating subject guides to resources on the Web is a form of collection development because librarians are, in a sense, incorporating into the collection the sites that they decide to link to the guide. Although Web resources are not often incorporated into a library's collection in the traditional way (by listing them and their URLs in the library catalogue) (Guarnino, 95-96), librarians intentionally direct users to these resources by linking them to a library publication. Therefore the process should be



done with care and according to appropriate selection criteria.

The first step of building a subject guide to Web resources is to find them. This can be done by using search engines, listserves, and virtual libraries, for example. Clearinghouses such as the Scout Report Signpost <<http://www.signpost.org/signpost/>> and BUBL Link <<http://bubl.ac.uk/link/>> have relied on librarians and others to uncover, evaluate, and arrange by subject thousands of Web pages. Such clearinghouses can be extremely helpful when looking for valuable sites. These directories provide subject access through the use of subject headings (Library of Congress

or Dewey, for example), classified structures, or metadata such as that proposed by the Dublin Core initiative to describe all Web resources (Burton, 5). Such directories are described as "interdisciplinary and selective indexes with resource descriptors" (Calcari and Wells, 43). Their overall goal and scope makes them different than subject- and institution-specific guides under consideration here, but their potential role in developing these specific guides should be recognized.

Because the amount of information that is accessible to library users on the Web is so great, it is not feasible for them to sit down and sort through it all themselves. In fact, when users

are looking for something very specific the prospect of unstructured 'surfing' will likely be quite unappealing (Harris, 34). Further, they may not even have the skills to find what they need in many cases. Therefore, after librarians have found a variety of subject-specific Web resources the filtering process must begin. In order to filter adequately it is important to first define the subject area that the guide will cover and create a profile of the library's users in order to determine what is needed (Fidel and Crandall, 16-17). Also, a selection policy that defines appropriate boundaries, "such as the selection of sites tailored to the primary clientele," should be followed in order to successfully gear the guides to the intended user group (Diamond, 9). In other words, filtering should be carried out according to set objective and quantitative measures that reflect the needs of the library community that the subject guide is to serve (Toub, 13). Questions to consider are, "Who are the end-users, what do they want to know, and how can they find it? (Dean, 83)" So, while internet directories such as the Scout project can serve as "a first layer filter for content and authority control," further filtering must be done in order for librarians to tailor their guides to the needs and level of their users (Calcari and Wells, 34).

Filtering according to the profile of intended users can only happen when a reasonably well-defined group of users can be identified. While such groups can often be defined in a university setting, for example, identifying a specific user group for the Web in general is an impossible task. Thus, librarians who filter Web resources with the users of their specific library in mind have different guiding principles and a more significant filtering role than librarians creating guides for the entire Web community. This filtering process is also known as gatekeeping since the tasks of gatekeepers are to monitor resources, provide and repackage information for known needs, and to analyze information use (Burton, 5). Since such activities cannot be carried out without some knowledge of the user population for

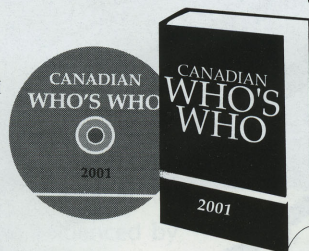
CANADIAN WHO'S WHO 2001

BOOK
Cloth 0-8020-4958-3 \$185

CD-ROM
0-8020-4960-5 \$195

BOOK AND CD-ROM
0-8020-4959-1 \$275

PUBLICATION
JUNE 2001



'The spirit of Canadians.'

Roch Carrier, National Librarian

- For ninety-one years, *Canadian Who's Who* has recorded the diversity of Canadian culture.
- Still one of the most consulted reference works in Canada, this annual publication documents the lives of 15,000 notable living Canadians.
- Useful to researchers, recruiters, students, media, business people, and many others, *Canadian Who's Who* is a must for every library.

Visit our Website at www.utpress.utoronto.ca/cww



UNIVERSITY OF TORONTO PRESS CALL 1-800-565-9523 FAX 1-800-221-9985

A black and white photograph of a desktop laminator machine, showing its top and front panels. The machine is white with a black control panel on the right side.

TCC

LAMINATORS

• Business Cards • Luggage Tags •
• Licenses • Membership Cards • Others •

Other Products: Display & Merchandising Products, Brochure Holders, Print Protectors

Make Your Best Impression!

HOP

HAVE OUR PLASTIC INC.
www.hop.ca

Please call for more information
1-800-263-5995

HAVE OUR PLASTIC INC.
6535 Millcreek Drive, Unit 67
Mississauga, ON L5N 2M2

Tel.: (905) 821-7550 • Fax: (905) 821-7553

which the gate is being tended, the issues of filtering and gatekeeping pertain most significantly to situations in which user-specific subject guides are being created.

Clearly, gatekeeping requires that irrelevant sites be filtered out according to users' needs, but another important component is that they be evaluated. It is only through evaluation that informed filtering can occur. Further, when these evaluations are included in the subject guides themselves, users will not be left merely "to follow the connections that are provided and analyze the materials themselves" (Harris, 38). Criteria for evaluation can include the intellectual level at which the resource is written, the type of information it contains (i.e. statistics, study results, etc.), the frequency of updates if applicable, the source of the page, and how closely the information contained relates to the information needs of the user group for which the subject guide was created. Only by taking such factors into account can accurate subject access be provided. Filtering is clearly an important part of providing subject access to Web documents because unsophisticated and keyword searches can lead users to a tremendous number of inappropriate documents. Users

are best served when they are led to the sites that pertain directly to the subjects of their inquiries.

While filtering is an important part of collection development, there are other decisions to be made that relate to the collection development of Web resources. For example, librarians should set guidelines for the depth and breadth to be represented in their subject guides. They must also decide what types of resources to include. Depending on their own print collection or the usefulness of the particular resources, they may or may not decide to include Web-based dictionaries and other reference sources, for example. One extremely important decision to make is whether or not to list primarily gateways (the equivalent of print bibliographies) to information about the chosen subject or to list mostly sites that contain the actual information users seek. In other words, will the subject guide be a "list of lists" or a list of individual sites that contain relevant information (Dean, 83)? The problem with compiling a list of lists is that the links they provide can be followed endlessly by users who, even after moving through several of them, may still not be in a position to retrieve the actual information they seek. Although such gateways

can be useful in certain situations, they are sometimes very large and are not usually tailored meet the needs of a particular group of users. Although these guidelines pertain mainly to issues of scope, other collection development criteria that apply when looking at Web resources are authority, format, and audience. Although cost is the fifth criterion when collecting for traditional collections, most web sites added to subject guides are freely accessible.

Once a subject-specific Web collection has been 'developed' librarians can concentrate on adding more value. The acts of filtering and drawing together appropriate and relevant sites contribute to the creation of value-added products, such as subject guides. However, when the focus is on subject access further steps must be taken. Briefly, the contents of each resource listed should be described and subject words or headings should be listed to enable users to quickly hone in on the information most appropriate for their needs. These are precisely the sorts of features contained in Web directories such as Scout Report Signpost. Although LC or Dewey headings may be used, it is not necessary to follow such classification structures if a simple combination



of terms will suffice, and if those who create subject guides are careful to maintain consistency in assigning descriptors. This 'cataloguing' of resources is an important function that helps make manageable a broad range of resources and types of information on a variety of topics within a particular subject area.

STRUCTURE AND DESIGN

Almost all subject guides look different from one another and their structure depends on a variety of decisions made by their creators. While it is not necessary for all guides to look alike, creators must carefully consider the structure they adopt since it influences the kinds of subject access allowed and the overall design of the guide. Subject access can be facilitated in different ways and it is helpful to present users with choices. Morville and Wickhorst advise the use of a topical scheme in which the guide is divided according to topics within the larger subject area

(Morville and Wickhorst, 31). For example, a subject guide for urban forestry could be divided into sections on planning, management, community involvement, etc. This type of breakdown gives users quite specific subject access at a glance. In some subject areas it is helpful to provide "additional access paths": a history subject guide may include a chronological organizational scheme, while an international relations guide might allow access based on geography. Guides may also include arrangements according to format, such as FTP or .jpg files, especially since it is relatively easy to create such a variety of organizational schemes on the web (Morville and Wickhorst, 31).

Hierarchical arrangements can also be useful and can even be incorporated into topical arrangements. In order to decide whether or not to use a hierarchical structure librarians must consider how extensive the subject guide will be and how deeply it will treat the subject. Hierarchical structures are

popular among large Internet directories but are not always necessary in significantly smaller, institution-specific subject guides. Both LC and Dewey structures can be used, allowing users to move from a broad subject such as 'health' to the specific topic 'health and medicine' and then to the subtopic 'health, preventive medicine,' for example (Vizine-Goetz, 14). As users move down the hierarchy the number of relevant Web resources is reduced, which makes the search for information more manageable. However, even if a hierarchical structure is adopted, users should still be able to jump right to 'health, preventive medicine' if they wish, without having to determine how to get there through the hierarchy. Providing users with choices is more likely to facilitate adequate subject access for all.

One issue that is raised by the use of Dewey or LC structures and their subject headings is that of natural language versus controlled vocabulary. The controlled vocabulary offered by these systems can aid precision and consistency, which is particularly useful when librarians are trying to describe the resources included in their guides. However, if users are not familiar with the controlled vocabulary librarians must consider carefully if "subject terms that reflect the normal speech and usage patterns of the defined patrons" should be used instead (Toub, 152). If natural language is used, for the sake of consistency decisions must be made about which terms to use and to what they refer. The goal in such a case would be to create a controlled vocabulary that is made up of natural language terms.

Although decisions about the above-mentioned methods for providing subject access clearly relate to design as well, there are also other design issues to consider. One very important component is the description of and metadata for each resource, which allows users to determine its particular topic and the exact type of information it contains. Although subject headings are useful for this purpose they do not always provide enough information. Descrip-

WHAT LIBRARIES NEED

Service • Selection • Speed

S&B BOOKS LTD. is THE National Book Wholesaler featuring:



Automatic Release Plans
Children's Bestseller Program
Large Showroom Inventory
Large Print, Video, Music

- Support to Canadian Publishers
- U.S. Ordering Capability
- U.K. Direct Ordering
- Non-Print Materials
- Audio/CD-ROM
- Three Bestseller Lists a Year
- MARC Records on Disc
- Electronic Ordering
- Fast and Efficient Delivery
- High Fulfillment Ratio
- Monthly Reports on Orders
- Fund Accounting
- Competitive Discounts
- RFP Quotations

S&B BOOKS DELIVERS

S&B BOOKS LTD.

3085 Universal Drive, Mississauga, ON L4X 2E2
(905) 629-5055 1 (800) 997-7099
(905) 629-5054 1 (800) 826-7702
(905) 629-1455 1 (800) 997-2597
root@sbbooks.com www.sbbooks.com



tions are more specific and should include the "scope, aims, and goals of the resource" (Toub, 154). They should also include metadata such as the resource's title, URL, size, source (and the authority of that source), and indicate what audience the resource is intended for. A set of 15 such elements has been enumerated by the Dublin Core initiative in order to specify what information (metadata) should be collected from Web resources (Burton, 5). Although this level of detail may not be necessary on all subject guides, certain basic components, such as those listed above, can be helpful to people using the subject guide.

Usability is also essential and can be facilitated by the design of the subject guide. Scope notes should be included to indicate exactly what is being covered, terminology should be simple but precise, and the labels or headings for each section must accurately and adequately describe that to which they refer (Dean, 85). Mis-labeled and poorly organized information is unlikely to be found by those who seek it and jargon is both unclear and unhelpful to those unfamiliar with the terms. A straightforward layout with a simple guide to the whole, or a table of contents, will do much to facilitate successful subject searching. Further, in order to ensure usability, form should follow function. For example, large graphics are seldom necessary and because they can take a long time to load they sometimes have a negative impact on usability (Morville and Wickhorst, 31).

Other design features have been recommended by users of subject guides. Users suggest that "duplication of resource links across multiple pages" should be minimized and that they should have "the capability to search through all the subject guide's pages by keyword or phrase" (Dean, 87). It can also be helpful to have a glossary of key terms (both library-related and subject-specific) available somewhere within the guide if such terms are not explained in context. All of these features increase the clarity of subject guides and although some users might benefit from other specif-

ic design features, these basic principles should guide the design. Before designing a guide, then, it is helpful to review others in order to both get ideas and see what designs work best.

CONCLUSION

Creating Web guides is an important task for librarians, particularly in light of the Web's increasing popularity as a source for information. Because of the magnitude of the Web there is no doubt about the need for organized and reliable subject access to its contents. However, because a significant amount of time and effort are needed to produce quality subject guides, librarians should assess what is already available and then decide if it is necessary to create their own. If librarians do decide to proceed, it is important that they be aware of the challenges that the subject cataloguing of Web resources may pose so that special considerations can be made when necessary. It has been shown here that providing subject access to Web resources is closely related to collection development because of the need to determine what small part of the Web an individual librarian can catalogue for his or her library's particular user groups. Clearly, then, librarians must recognize that a small part of the whole Web becomes part of the library's 'collection' when a subject guide is created. Once the Web collection has been assembled, subject access can be facilitated through the use of subject headings, hierarchical and topical arrangements, and written descriptions in subject-specific guides. In order to facilitate easy use of such guides, several structural and design features must be taken into account. When a subject guide is developed through a sound process and created with care it can provide helpful and relevant access to Web resources according to subject, which saves the time of library users and begins to organize a small part of the Web that is relevant to them.

Candice Dahl has successfully completed her Masters' program at the Faculty of Information Studies of the University of Toronto.

BIBLIOGRAPHY

- Burton, Paul F. "Issues and Challenges of Subject Access." *Catalogue & Index* 128 (1998): 1-7.
- Calcari, Susan and Wells, Amy Tracy. "New Tools for Collection Development: The Internet Scout Project." *Collection Management* 23, 3 (1998): 33-48.
- Dean, Charles W. "The Public Electronic Library: Web-Based Subject Guides." *Library Hi-Tech* 16, 3-4 (1998): 80-88.
- Diamond, Tom "Librarians and the Digital Frontier: Drivers, Passengers, or Hitchhikers?" *LLA Bulletin* 59, 1 (1996): 6-12.
- Fidel, Raya and Crandall, Michael. "The Role of Subject Access in Information Filtering," in *Visualizing Subject Access for 21st Century Information Resources*, ed. Pauline Atherton Cochrane and Eric H. Johnson. Urbana-Champaign: University of Illinois at Urbana-Champaign, 1998: 16-27.
- Guarnino, Patricia H. "Making the Internet a Part of the Library's Collection." *The Acquisitions Librarian* 20 (1998): 91-100.
- Harris, Steven R. "Webliography: The Process of Building Internet Subject Access." *The Acquisitions Librarian* 17/18 (1997): 29-43.
- Morville, Peter and Wickhorst, Susan. "Building Subject-Specific Guides to Internet Resources." *Internet Research: Electronic Networking Applications and Policy*, 6, 4 (1996): 27-32.
- Toub, Stephen E. "Adding Value to Internet Collections." *Library Hi-Tech* 15, 3-4 (1997): 148-154.
- Vizine-Goetz, Diane. "OCLC Investigates Using Classification Tools to Organize Internet Data." *OCLC Newsletter* 226 (1997): 14-18.