Shelflessness as a Virtue: Preserving serendipity in an electronic reference collection

Lyle Ford, Lisa Hanson O'Hara, Jared Whiklo

Abstract: The reference collection traditionally rewards serendipity by presenting key resources in a compact group. At the University of Manitoba Libraries, reference collection space is at a premium and, increasingly, electronic versions of reference materials are selected for purchase. However, our space saving comes at a cost: our patrons can't browse electronic reference materials across various online platforms and they miss out on potential serendipitous discoveries. This presentation will outline the problem, look at what ARL libraries offer as online reference collections, and offer possible solutions, including what we hope to do at the University of Manitoba Libraries (UML).

Introduction

> "Serendipity, *n.*" The Oxford English Dictionary. 2nd ed. 1989. OED Online. Oxford University Press.
> "The faculty of making happy and unexpected discoveries by accident. Also, the fact or an instance of such a discovery."

Electronic reference (ereference) collections have been provided in academic libraries since the early 1990's when people began to make resources available on the Web. Many traditional reference sources, especially ready-reference sources such as dictionaries, directories and encyclopedias were ideal for converting to Web-based services and were made freely available. Libraries, developing their own web-pages at the same time, took advantage of these free resources by creating lists of reference web sites. As electronic publishing matured and grew in influence and scale, academic libraries began to lose ground in their attempts to provide useful and comprehensive guides. Moreover, as publishers began to apply strategies and models adopted from electronic journals, librarians encountered the complexities of competing interfaces, bundled collections, authentication, usage restrictions and other imposed limitations.

Despite these limitiations of course, electronic books are commonplace. They often enter library catalogues as collections of materials not freely available on the Internet, and equate to the sort of books which would be found in a traditional print reference collection in an academic library.

Libraries have been providing electronic journals and electronic books for a number of years. Studies have shown that electronic books are discovered the same way as traditional print collections, through the library catalogue (Connaway, 344). Electronic journals are discovered through the catalogue, but also through databases that are OpenURL enabled, and through A-Z lists, making discovery of journals improved in the electronic age (O'Hara, 127). However, reference collections are approached and used differently, by patrons and by librarians. The library catalogue serves as a representation of a given collection, using metadata to facilitate discovery. Many reference books, though, function as collections in their own right. Dictionaries, encyclopedias, handbooks and bibliographies all bring together resources on particular topics, organizing them by subject, time period, area and myriad other systems. As discrete items, reference books provide value by concentrating information at varying levels of depth and breadth. That concentration, though, complicates effective representation of the item in the catalogue.

Serendipity in the discovery of information in the reference collection is key, since many reference books are not discovered through the library's catalogue, as Karen Sendi has written. "Approximately 50 percent of the respondents [to a survey on the reference collection] learned about the collection from a librarian and/or a professor. Other

methods of discovering the materials in the reference collection included browsing, location specified in the online catalog, mention of the collection in a library instruction class, referral by other classmates or other units on campus" (Sendi, p.20). In most academic libraries separate reference collections are the norm precisely because they are recognized as distinct from everything else. The separation allows for browsing and obviates the necessity of representing the individual items in the catalogue. The reference collection remains an intellectual construct, organized by call numbers, but the representational construct is easier to grasp and navigate because it can be approached in the concrete rather than in the abstract, via the catalogue.

Sendi's article, and the established practice of distinct reference collections, raised questions about the content and usefulness of electronic reference collections. We wondered if, as print resources change over to electronic format, was the notion of a reference collection still necessary and, if so, has the idea evolved to take full advantage of electronic resources? It was also interesting to discover whether or not librarians had made any efforts toward creating anything analagous to the print collection. We also hoped to find out how collections of electronic reference books were being used and how they were being made available, expecting to find some patterns overall. And finally, the central question was, can libraries preserve the browsability of the print reference collection in an electronic reference collection?

The authors undertook to answer these questions in two ways. Our first objective was to survey the web sites of members of the Association for Research Libraries to look for electronic reference collections and determine what types of materials were included in

those collections and how they were made available. Secondly, we undertook to develop a virtual collection of electronic reference books for the University of Manitoba Libraries' website, and gauge the usage of this collection against the usage of the print reference books in the Libraries' arts and humanities library. While undertaking the first objective, we also looked for examples in library literature and on the web that we could use to develop our own virtual electronic reference book collection.

E-Reference Collections at ARL Libraries

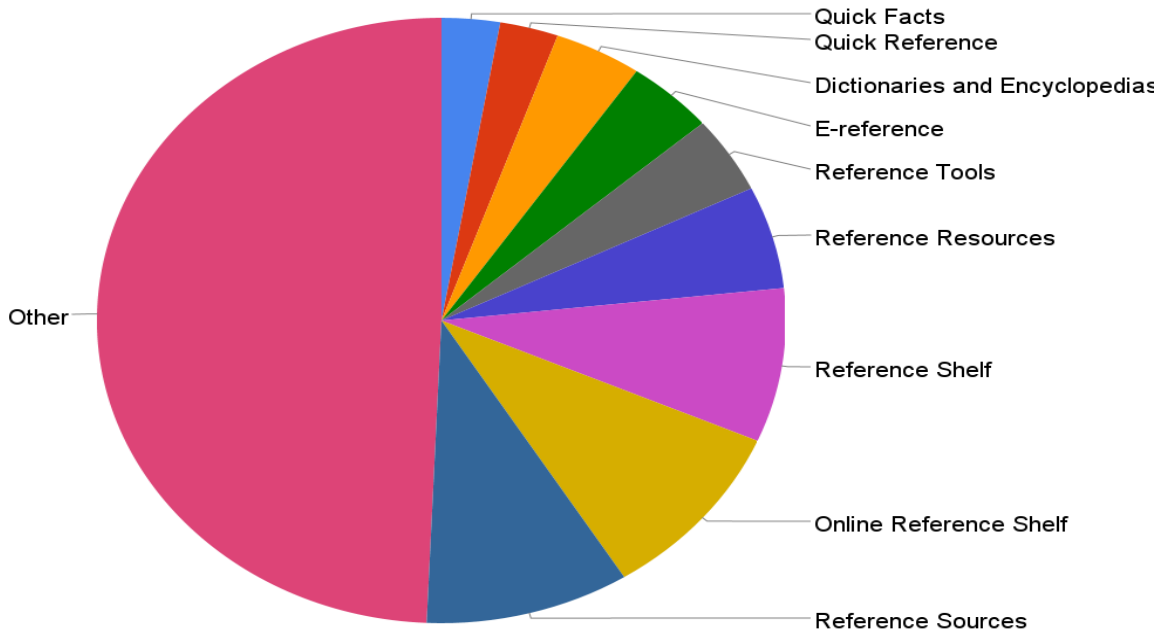
We surveyed the academic library members' of ARL and CARL (Canadian Association of Research Libraries) websites in June of 2008. Consulting the Survey Random Sample Calculator from CustomInsight.com (www.custominsight.com/articles/random-sample-calculator.asp), we learned we would need to sample 93 of the ARL members in order to achieve a 5% error rate and a 95% confidence level. We decided to do an environmental scan of all 123 members. Data was collected from 123 different libraries and included information on whether or not the library had an ereference collection, whether it was accessed through a subject portal or not, how many clicks from the homepage it was located, what it was called, whether it was dynamically generated or not and finally, whether it included links to electronic books paid for by the library (not just free resources). When looking for electronic reference collections, we decided to spend a maximum of three minutes looking on each website and if the ereference collection was not found, we assumed that it did not exist. Nielsen and Loranger say people users generally spend less than 30 seconds looking for something on a website, and about 10% of users will spend about 5 minutes; we felt that three minutes was a fair test given these statistics.

We found that 51 (41%) of the libraries surveyed did not have electronic reference collections. Of those libraries that did provide electronic reference collections, 32 (63%) included electronic reference books in the collection as well as, or instead of, free resources. None of the libraries' websites included "Web 2.0" features such as book covers, browsing, tagging et cetera. Twenty-five percent of the electronic reference collections were generated dynamically, as determined by the authors from evidence in the webpage. The others relied on manually updated subject guides, title lists and the like.

Fifteen of the 73 electronic reference collections were accessed through a subject portal. This means that the ereference collection was found by navigating to a subject listing page, and and then locating either a subject category labelled as reference materials, or by finding a subcategory labelled as reference materials within the subject portals themselves (i.e. "Dictionaries and Encyclopedias" found under a "Anthropology" subject portal).

Of those libraries whose websites contained an electronic reference collection, the names varied widely. Fourteen websites contained an "Online Reference" collection, followed by eight sites with collections found under "Reference Sources". "Electronic Reference Sources/Shelf," "Reference Resources," "Reference Shelf" and "Find (varied as: Databases/Dictionary/Digital Resources/Facts)" were used at six libraries each. Various other names were used such as "Quick Facts," "Quick Reference," "Virtual

Reference" and "Digital Reference." The term "Reference" was used in 52 of the electronic reference collections.



Links to electronic reference collections were relatively easy to find on the libraries' websites. Sixty-six percent of the websites had electronic reference collections available on the homepage itself or within one click of the homepage. Only 9% of the websites required 3 or more clicks.

Analysis

It appears that electronic reference collections in academic libraries are slowly evolving to match the availability of electronic reference books. Although 47% of libraries are still using an old model of providing electronic reference collections, more than half of the libraries are providing access to electronic reference books through an online collection. Browsability of these collections is being maintained through the listing of

the books together in one place on the website, but it was disappointing that the sites were basically just that, a listing. The authors expected to find that some libraries were taking advantage of Web 2.0 technologies to make these resources more "Amazon-like." Perhaps part of the problem is that libraries are still working to make their main catalogues "Amazon-like" and will turn their attention to electronic reference collections once that has been accomplished. We did look at the main library catalogues to see if there was any way to distinguish electronic reference books in cases where we knew a "next-generation catalogue" was employed, but this was not done systematically and no way of grouping electronic reference books as distinct from print was found.

The naming of the electronic reference collections across ARL & CARL libraries was also an issue. There is no best practice for naming the collection and it was interesting to note that not all of the libaries even included the word "Reference" in the name of the online collection, although we will make an assumption that the print collection bears that name, as it does in most academic, public and school libraries.

Literature Review

The authors searched two databases, seeking Library, Information Science & Technology Abstracts (LISTA) and Library Literature & Information Science (LLIS) articles published in academic journals since 1999, using descriptors such as "reference sources," "electronic reference sources," "electronic book," "ebook" and keywords such as "reference collection" and "collection development." Largely, the articles considered selection criteria and licensing issues, or assessed platforms and the publishers' practice of bundling collections (see Works Consulted). What we did not find were any articles

that specifically addressed the presentation of such collections in a way that related to our project.

University of Manitoba Libraries

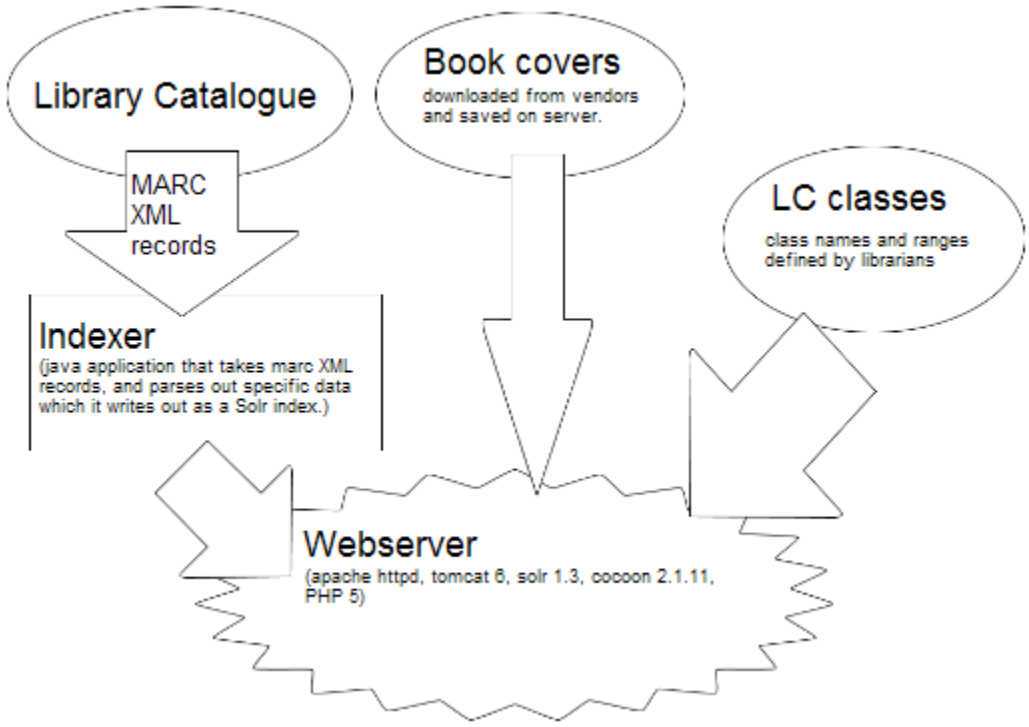
The University of Manitoba Libraries (UML) has been collecting electronic reference books since its acquisition of the Credo Reference collection in 1995 (then called xrefer plus). Since that time, the electronic reference book collection has grown, both through the purchase of collections such as Gale Virtual Reference Libraries, and through the purchase of reference books in other electronic collections, such as the Springer electronic books collection. Whenever possible, the Libraries purchases MARC records along with the ebooks, and they are added to the catalogue along with a local subject field (MARC 650_4) "Ebooks -- UML". Electronic reference books also get the local subject field "Ereference -- UML". Access to these materials is dependent upon finding the MARC record in the UML's catalogue (not a preferred discovery method for reference books as demonstrated in the literature) or by accessing the vendor's platform. Usage statistics indicate that the electronic reference books are being used fairly consistently throughout the year, and that usage is increasing. In 2008 there were 3382 searches run in 799 sessions in Credo Reference which has 180 titles in the collection; in 2007 there were 5938 searches run in 746 sessions. This is not an extensive increase in the number of sessions, and it is not possible to determine where these sessions originated (library staff, faculty, students, etc.)

The literature indicated that most patrons find and use reference books through browsing the reference collection or through referral by a professor or librarian. In

order to make our electronic reference books more accessible to library clients, we wanted to create a browseable collection across various vendor platforms and we wanted to incorporate some Web 2.0 technologies into our model to make it more user-friendly.

Creating the University of Manitoba Model

To create our model electronic reference book collection, we used MARC records extracted from the Libraries' catalogue which we converted to MARCXML, book covers obtained from the vendors, and a table of LC class numbers linked to subject titles.

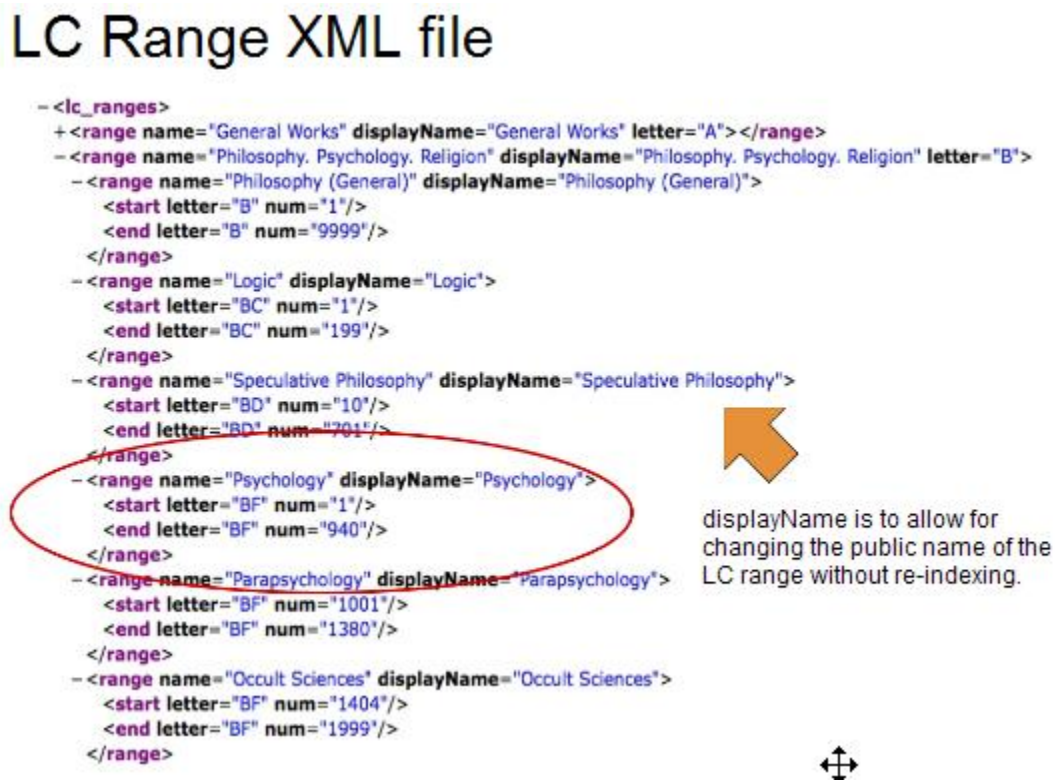


The MARC records were obtained by searching our catalogue for all records with the local subject heading (MARC 650_4) "Ereference -- UML". These records were then converted to MARCXML using MarcEdit software. We then wrote an indexer in Java to loop through a MARCXML file and push the records into a Solr index. The indexer was based on one that had been previously built for the Manitobia.ca project (this is a collection of historical newspapers from Manitoba). The indexer uses the date in the MARC 005 field in order to determine whether a record is new or has been previously indexed. If it is a new record, it is automatically indexed; if not new, the record is ignored. We chose to use Solr because it is a very fast read-only index which takes a query and returns XML. It also has the advantages of built-in spell-checking, syntax-highlighting and faceting features.

Solr record


```
- <response>
  + <lst name="responseHeader"></lst>
  - <result name="response" numFound="1" start="0">
    - <doc>
        <str name="ISBN">9780631222618</str>
        <str name="id">8181455</str>
      - <arr name="lc_class">
          <str>BF713.5</str>
        </arr>
      - <arr name="lc_subject">
          <str>Psychology</str>
        </arr>
        <date name="modified">2008-06-07T16:45:37Z</date>
        <str name="pub_date">2008</str>
        <str name="publisher">Blackwell Publishers,</str>
      - <str name="title">
          Handbook of research methods in developmental science
        </str>
      - <arr name="url">
        - <str>
            http://www.blackwellreference.com/subscriber/uid=411/book?id=g9780631222618_9780631222618
          </str>
        </arr>
        <str name="xmlType">marcXML</str>
      </doc>
    </result>
</response>
```

A table of LC class numbers was created which listed LC ranges and associated display names (the "subjects" that clients could use as starting points in a browse). The association of class numbers with display names allows us to change the display names in the LC Range XML file instead of having to change each individual record. Using the LC class number as the filing point allowed us to create a mdoel that browses in the same way that the items would browse on the shelf.



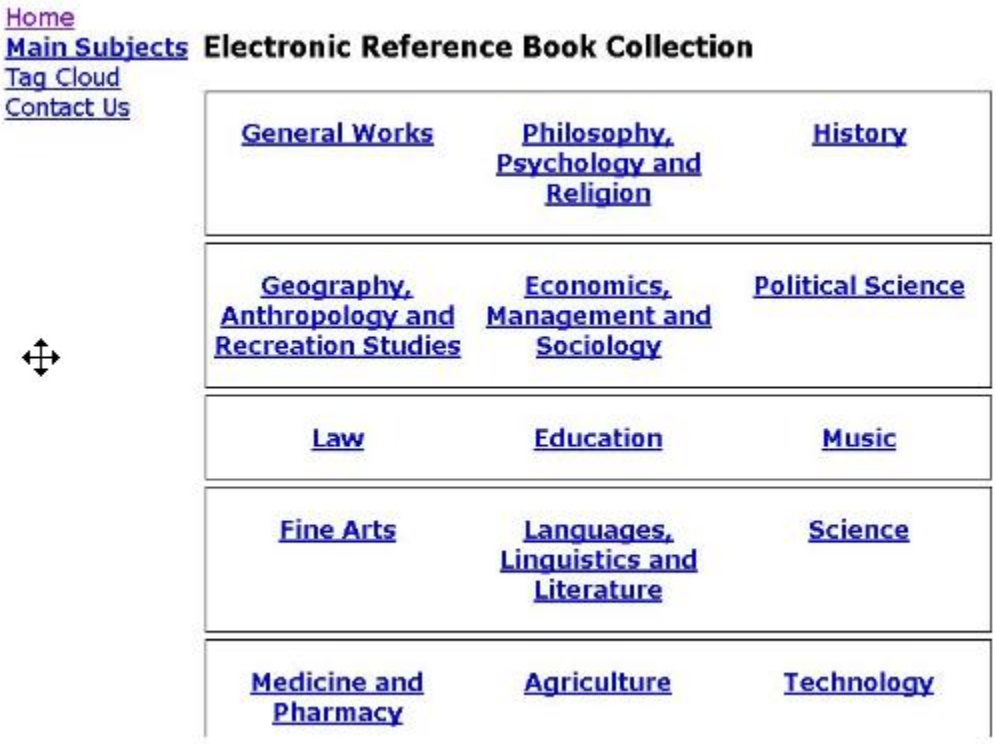
Cocoon allowed us to map between the standard LC ranges and the LC class that we indexed for each record in order to give us a hierarchy of browseable subjects. Cocoon is a robust framework that allows for on-the-fly XSLT (eXtensible StyLesheet

Translations). We used Cocoon v.2.1 because we were familiar with it and v.2.2 was too complicated to install and get working for a demonstration project.

We also obtained book covers for all of the electronic reference books in order to make our model more user-friendly, and to maintain the similarity to browsing a shelf of print books. Because Amazon no longer allows libraries to use their book covers without providing a link back to purchase the book from Amazon, and because LibraryThing doesn't have very many covers for reference books, we asked the vendors themselves for the book covers, all of whom were very willing to allow us to download them. Some vendors made them available in a batch and in some cases we created a little program to download them rather than going into the vendor database and downloading them individually. The Solr records were then made more user friendly.

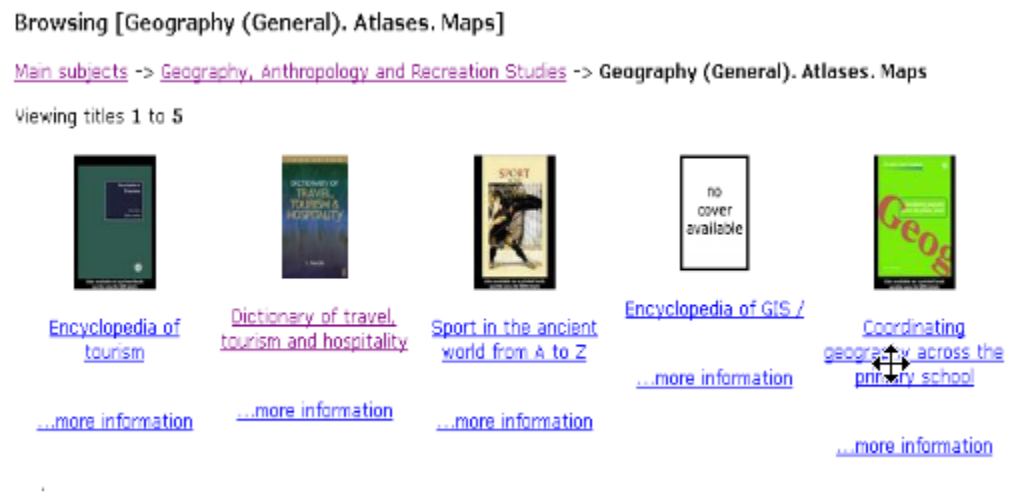
The final piece for creating the model was allowing for tagging. We felt that this would be a useful Web 2.0 feature for clients but did not want to create login and passwords or allow free tagging with the possibility of inappropriate use of the tagging feature. Instead we decided to create a blacklist which would not allow the use of certain words as tags.

The University of Manitoba Model



Click on "Geography, Anthropology and Recreation Studies" with the resulting screen:

Click on "Geography (General), Atlases. Maps" with the resulting screen:



Clicking on one of the book covers or titles takes the client to the book at the vendor platform.

Evaluating the Model

The model we created was purposely simple, intended to serve as a “test of concept.” Primarily, we wanted to know if the model would ease discovery and access. Our first

goal for evaluation was to gather input from our colleagues, before developing a more sophisticated product for the public. Accordingly, aesthetics did not factor into our prototype design. In fact, we deliberately sidestepped design issues in order to limit discussion about colours and screen layout, and to maximize our colleagues' focus on the navigation scheme, subject access, link visibility, and other fundamentals.

We chose to poll a select group of six librarians rather than the entire complement of librarians employed by the UML. We chose the librarians providing public service to specific faculties and departments, in hopes of gaining a variety of opinions. The librarians were invited, by email, to visit the test site and then tell us their first impressions. We presented the site as an adjunct to the library catalogue, a way of making electronic reference materials browsable so that patrons could discover books related to their topics, arranged by "aboutness." We also explained that we had gone forward on the assumption that people tend not to search for reference materials in the catalogue but tend to rely on serendipity to find material on the shelf.

Following these introductory comments, what we asked our colleagues four direct questions:

1. Where do you think this ought to go on the UML site?
2. What do you think about the navigation scheme? Does it work? What could be clearer?
3. Will this ease or complicate discovery of electronic reference books?

4. An OCL report showed that people associate libraries with books, but don't define e-books as books. They don't view our e-resources as trusted info-sources on a par with books, but rather compare them to other e-resources. Do you think a service like this will help promote e-reference books as something as trustworthy and valuable, but more convenient, than print books?

In response, our colleagues commented on several aspects of the prototype. Subject access to the collection, though, was the most popular topic for our respondents. Four of the six commented on the terms we used to organize the collection. One commentator told us that there were too many subdivisions, and that simpler, broader divisions would be better. Another person thought it would be preferable to use the names of faculties and departments (ie. Arts then English).

The departmental arrangement is one that the UML has long used to organize its collection of databases. Based on that precedent, and on the feedback provided by our catalogues, we decided that the next step is to reuse that schema and map the LC headings to it. Once that mapping is completed, the e-reference collection will be added into the larger UML website as a beta version. Once the resource is made available, our intent is to conduct three informal focus groups: one of faculty, one of graduate students, and one of undergraduates.

Conclusion

While it remains to be seen if our approach will serve all of our purposes, we have successfully demonstrated that it is possible to design a sophisticated apparatus, in-house. Using local expertise within a task-specific group, comprised of members from

reference, technical and cataloguing departments, allowed us to rapidly prototype a workable system. We began the process by outlining our most optimistic goals for the system. After defining our ideal, we began to assess our current workflows and systems to see what was achievable. That process, although fairly complex, was eased by the collegial attitude of the group members. If nothing else, the exercise brought home to the participants, and to the rest of the library system, that all of the staff play a significant role in public service, regardless of departmental affiliation.

Bibliography

Aked, Michael J., John C. Phillips, and Karen A. Reiman Sendi. "Faculty use of an Academic Library Reference Collection." Collection Building 17.2 (1998): 56-64.

Bradford, Jane T. "What's Coming Off the Shelves? A Reference use Study Analyzing Print Reference Sources used in a University Library." The Journal of Academic Librarianship 31.6 (2005): 546-58.

Buckland, Michael K. "The Digital Difference in Reference Collections." Journal of Library Administration 46.2 (2007): 87-100.

---. "The Digital Difference in Reference Collections." Journal of Library Administration 46.2 (2007): 87-100.

Colson, Jeannie. "Determining use of an Academic Library Reference Collection: Report of a Study." Reference and User Services Quarterly 47.2 (2007): 168-75.

Connaway, Lynn Silipigni. "A Web-Based Electronic Book (eBook) Library: The netLibrary Model." Library Hi Tech 19.4 (2001): 340-349.

Fishman, Diane L., and Megan DelBaglivo. "Rich in resources/deficient in Dollars! which Titles do Reference Departments really Need?" Bulletin of the Medical Library Association 86.4 (1998): 545-50.

Frost, William J. The Reference Collection: From the Shelf to the Web. Haworth Information Press, 2005.

---. The Reference Collection: From the Shelf to the Web. Haworth Information Press, 2005.

Landesman, Margaret. "Getting it Right-the Evolution of Reference Collections." The Reference Librarian.91/92 (2005): 5-22.

---. "Getting it Right-the Evolution of Reference Collections." The Reference Librarian.91/92 (2005): 5-22.

Nielsen, Jakob and Hoa Loranger. Prioritizing Web Usability. Berkeley, CA: New Riders, 2006.

O'Hara, Lisa. "Providing Access to Electronic Journals in Academic Libraries: A General Survey." The Serials Librarian. 51.3/4 (2007): 119-128.

Reiman Sendi, Karen A. "Assessing the Functionality of the Reference Collection." Collection Building 15.3 (1996): 17-21.

Welch, Jeanie M., Lynn A. Cauble, and Lara B. Little. "The Evolution of Technology in the Management of Noncirculating Library Collections." Technical Services Quarterly 17.4 (2000): 1-11.

---. "Automated Reshelving Statistics as a Tool in Reference Collection Management." Reference Services Review 25.3/4 (1997): 79-85.