

Browsing Library Collections: From the Shelf to the Online Catalog

The problem: Of making many books there is no end; library buildings are costly; and not many campuses have real estate to spare for the purpose of housing yet more printed books in their general libraries. *The solutions:* Put lesser-used printed materials in less-costly storage facilities; or engage in cooperative archiving agreements that retain one copy of a title in one of the partner's active stacks so that the other partners may divest of their copy.

These are indeed efficient and reasonable solutions, but a new problem arises: many library users, like bookstore customers, highly value the practice of shelf browsing. They assume that the array of items that might interest them is present on the shelf and that they can graze around an item they know meets their needs to discover and select related reading. Tales abound from library users about the happy find, the unexpected, almost magical appearance of an item whose existence they had not suspected but that turns out to be exactly the resource they need.

Since so many library users believe so strongly in the serendipitous efficacy of shelf browsing and since, for this reason, they resist plans by libraries to move items into a storage facility or to develop a consortial approach to collections, how can libraries afford users the advantages of browsing for print materials and at

the same time manage space for collection growth? Put another way, how can a library maintain the discovery and evaluation/selection function of shelf browsing for users at a time when users are requesting more and more materials from off-site storage facilities or from remote libraries through consortial or regional catalogs?

Five years ago, the Tri-College Library Consortium, a twenty-year partnership of Bryn Mawr, Haverford, and Swarthmore Colleges, embarked on a new collaborative project.¹ With a common catalog and a daily delivery service among the libraries, we decided to remove duplicate uncirculated older books and journals from our stacks in favor of retaining one copy and to make room for titles we would not otherwise have bought by working together to avoid unnecessary duplication when purchasing new titles.

How can libraries afford users the advantages of browsing for print materials and at the same time manage space for collection growth?

Enter the browsing impediment. Our users said they were skeptical of a program that, in effect, was going to distribute a single collection of printed materials over three campuses and thereby reduce the number of titles for browsing in a given library. At the same time they were telling us about their dependence on shelf browsing, however, they were also giving us other data about their preferences and habits, data that prompted us to think we could both carry out our space-management plan and satisfy our users by moving browsing from the shelf to the catalog—from the stacks, as it were, to the network.

As we discussed our project, we heard that our users loved

the way Amazon.com lets a reader “open” many books by offering sample text. They told us that it was the table of contents (TOC) and first chapters that most helped them decide whether they wanted to borrow and read a book. They told us they valued the searchable TOC we were adding to catalog records for newly published titles. We knew that they were also becoming more accustomed to working online for discovery and selection purposes through shopping and the use of search engines.

As we thought about how to solve our local problem with a physically distributed collection, we spoke with many other libraries that were trying to address this same browsing issue with book collections. Encouraged by this expression of common interest, Susan Perry, of the Council on Library and Information Resources (CLIR) and the Andrew W. Mellon Foundation, and I gathered together an interest group at the American Library Association (ALA) conference in January 2004 under the auspices of CLIR and the Digital Library Federation (DLF). With representatives from the library, library network, and bookselling communities, we asked the group to think about how we might develop a common approach to addressing the shelf-browsing problem.²

We took our cues, as in so many other cases these days in libraries, from the commercial sector and from our many users who kept asking, “Why can’t the library catalog be more like Amazon.com?” We asked ourselves: For the older titles that our libraries are moving to storage or are sharing by consortial arrangement, could we produce sample content and searchable TOCs? In other words, could we create for these older materials the kind of

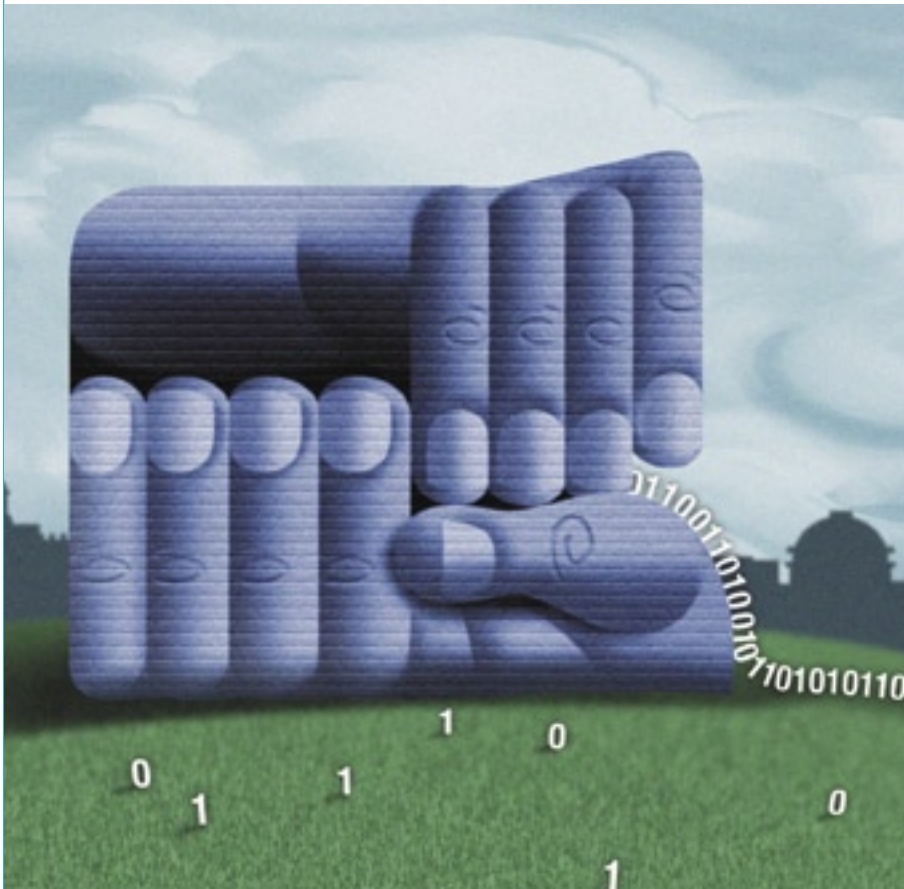


Illustration by Steve McCracken, © 2006

online discovery and evaluation resources that we have for the newer materials in our catalogs or that online booksellers make available? We also discussed the prospects for creating a national library partnership that would share the TOC and sample text that any library had produced for its local titles. To create this cooperative, we assumed that a national effort would have to be able to automate as much of the TOC data generation, verification, and distribution as possible and that the text scanning would have to be performed by low-cost, preferably student-assistant labor in order to make production and sharing of the data financially feasible for libraries.

Discussions continued at subsequent ALA meetings, and by the summer of 2005 a partnership had developed among Emory University, Haverford College, the Library of Congress, the University of Florida, and the University of Illinois to experiment with the workflows and processes that would be needed by libraries to create and share such data or, in Florida's case, to contribute already existing digitized texts. At present, our experiments are divided into two parts.

Emory, Haverford, and Illinois have been sending images of TOCs to the Library of Congress, which has software to scan and then parse into searchable data fields the TOC elements (author, title, etc.). Given the format and typographical variations of TOCs, this task is not an easy one. This experiment will conclude when we feel we have done all we can with the software to make the scanning and parsing accurate; at that point, we can decide whether that accuracy rate is sufficient for production.

The other experiment is to create a prototype catalog in which a large group of bibliographic records have all been augmented with sample text so that users can browse a sequence of books online in the same way that they would browse books on a shelf. As part of a consortial weeding project conducted in the summer of 2005, Haverford withdrew from its collection several hundred uncirculated, older economics titles that Bryn Mawr or Swarthmore will retain. Student assistants scanned a variety of page sequences, and these images were then integrated with bibliographic records. Thus in a test view

of our catalog, the image of sample pages, which can be enlarged for reading, displays alongside the bibliographic record.

We have yet to work through the means for making the TOC data and sample text images available to the library community. We have also yet to settle on the copyright implications of creating and distributing the sample text files, although we hope to work with the publishing and legal communities to make a fair-use argument for doing so. Questions also naturally arise about the relationship to, and future prospects for our project among, such mass digitization projects as the Google Book Search Library Project and the Open Content Alliance—we assume that a fully digitized, searchable text will trump images of sample text as a browsing mechanism.

Perhaps the most important consideration of all for the future of this project is the question of whether the library catalog as a discovery or browsing mechanism, focused as it is on a local collection, has a future in these days of Web searching. As libraries and the publishers who serve their market expose more of their content to indexing by search engines, as linking servers become standard equipment, and as OCLC's Open WorldCat program exposes its tens of millions of records for library items and offers a zip-code search to find a nearby or in-library copy, we must ask whether the local catalog will become less relevant. However these many questions of discovery and selection may eventually be answered, we feel that the trend is inevitably toward online methods, and we hope that our project can assist libraries and their user communities in the transition from stack to networked browsing.

Notes

1. See Judy Luther et al., "Library Buildings and the Building of a Collaborative Research Collection at the Tri-College Library Consortium," April 2003, <<http://www.clir.org/pubs/abstract/pub115abst.html>>.
2. Robert Kieft, "Collaborative Project to Enhance Library Catalog Browsing," *CLIR Issues*, no. 38 (March/April 2004), <<http://www.clir.org/pubs/issues/issues38.html#collab>>.



Robert Kieft is Director of College Information Resources and Librarian of the College, Haverford College.