146 LRTS 51(2)

Notes on Operations

Linking Print and Electronic Books

One Approach

By Betsy Simpson, Jimmie Lundgren, and Tatiana Barr

Library catalog searchers expect to retrieve information for all resources in the catalog that matches their search strategy. They expect keyword searching to retrieve a rich array of resources. In an effort to enhance service to users, the University of Florida Smathers Libraries acquired table of contents data to enrich bibliographic records for print books with publication dates from 1990 to the present. Many of these books have also been acquired in electronic format. Because the record for the same book in electronic format did not include the enhancements, catalog users were likely to retrieve the catalog record for the print version only and remain unaware of the availability of the electronic version. The authors, using insights from discussions surrounding the Functional Requirements for Bibliographic Records (FRBR) initiative, developed a method for serving users more effectively by linking these records to leverage the enhancements for both versions (two manifestations) of the same title.\(^1\)

Introduction

The proliferation of electronic resources and increasing expenditures on electronic resources have had a profound effect on library services. Librarians have been forced to rethink assumptions about basic library operations as well as long-held notions about user needs and behavior. Within the cataloging community, the Functional Requirements for Bibliographic Records (FRBR) has fostered a renewed commitment to creating library catalogs that allow users to find, identify, select, and obtain library material, and to navigate through the catalog database more effectively. FRBR inspired the authors to look for ways to improve the link between library catalog records for corresponding print and electronic book (ebook) titles. The impetus for creating these links was the presence of table of contents (TOC) data only in the records for print materials. Searches by keyword in the online catalog that matched data in the TOC retrieved only the records for print books. As e-books become more accepted as alternatives and supplements to their print equivalents, users will benefit from efforts to enhance access to them through the library catalog.

Background

A major thrust of current national cataloging initiatives is toward improving the display of connections and relationships among bibliographic entities. The FRBR conceptual model promotes a framework that highlights the interrelatedness of works and allows users to navigate easily among expressions, manifestations, and items.2 Embracing the underlying tenets of FRBR, libraries have been motivated to explore changes that leverage bibliographic data in new ways. The Research Libraries Group's (RLG) RedLightGreen service was an early, large-scale, innovative application of FRBR principles. Launched in 2003 (and ended November 1, 2006), RedLightGreen sought to mine RLG's union catalog for "conceptu-

Betsy Simpson (betsys@uflib.ufl.edu) is Chair, Cataloging and Metadata Department, and Jimmie Lundgren (jimlund@uflib.ufl.edu) is Head, Science and Social Science Cataloging Unit, George A. Smathers Libraries, University of Florida, Gainesville. Tatiana Barr (tatiana.barr@yale.edu) is Catalog Librarian, History and Social Science Team, Sterling Memorial Library, Yale University; she was a librarian at the University of Florida when this paper was written.

Submitted January 16, 2006, and returned for revision and resubmission; revision submitted June 9, 2006; accepted for publication July 9, 2006.

al relationships and holdings data."3 OCLC's Fiction Finder also employs automated means to find and collocate material in order to present users with a clear summary of various editions of fictional works and which libraries own them.4 Advances in data harvesting as well as semantic interoperability hold promise for dramatically improving user interfaces.

Three obstacles stand in the way of libraries fully FRBRizing catalogs. First, typical library management systems do not adequately manipulate the links that currently exist among bibliographic records. Referring to the inherent problems associated with record linking within the library environment, Gradmann speaks of the "library automation applications and the data architecture underlying these, which strangle libraries, creating a structural lack of technical flexibility" and discusses the need to "free librarian bibliographic data from its golden catalogue-cage."5 Yee also decries the state of library catalog software, stating that libraries are forced to "choose among undesirable alternatives" when seeking a system that has an adequate search engine with helpful displays.⁶ Strong words, but justified given that much data created by librarians sits essentially unused due to system limitations.

Second, catalogs lack the data necessary to reflect relationships because catalogers frequently have not provided it. Catalogers have not entered all data necessary to reflect relationships due, in part, to increasing pressure for catalog departments to economize operations and improve throughput. This has resulted in workflows in which downloaded copy is accepted as is, and original records are created with less detail.

The third deterrent to more thorough recording of relationship data is that library management systems are not able to make use of all data. Yee notes that the trend toward deprofessionalization has reduced the number of highly trained, knowledgeable

catalogers, and suggests that this void leaves the profession without a voice that understands the nuances of bibliographic description and can advocate successfully for change.⁷

Links between records may be missing entirely, not consistently applied, or entered in a way that is difficult to extract. The labor that would be required to create links manually on a record by record basis seems unthinkable in an age of shrinking budgets and staffs. Bowen discusses the possibilities for cataloger-created collocation, but acknowledges that the additional effort is likely prohibitive and will necessitate selective adoption.8 Catalogers, however, using the technological tools at hand and working closely with systems professionals, public service librarians, and vendors, can develop more ways to provide and utilize relationship data, and play a key role in making library collections more accessible to users.

What should catalogers do during this time of transition when developing a FRBR-like catalog portends a significant outlay of time, money, and technical expertise? The authors suggest that catalogers can begin by taking whatever small steps are possible while also collaborating with vendors and colleagues to institute more sweeping changes. Librarians can follow Bowen's recommendation to "look for opportunities to implement some aspects of the FRBR model within other activities that are more under the library's immediate control."9 In this spirit, the authors approached a local experiment to link e-book records to their print equivalents. A literature search did not produce evidence of similar projects elsewhere.

NetLibrary, a division of OCLC, is a leading provider of e-content. It is one of a growing number of companies offering access to e-books and, often, the corresponding files of MARC bibliographic records for downloading to local library catalogs.10 Since 2001, the University of Florida Smathers Libraries (UFL) has

batchloaded approximately 250,000 ebook records from NetLibrary, Early English Books Online, Eighteenth Century Collections Online, History e-Books, Past Masters, Gale Virtual Reference Library, and Books24x7. With this level of activity, UFL wanted to maximize its investments by improving the ease with which users can find these resources. Other libraries are likely experiencing similar needs. The library catalog can provide a solution by alerting users when the electronic version of a title is available along with the print version.

UFL, like many other libraries, has loaded separate catalog records for e-books rather than attempting to utilize a single-record method of access. As a result, users often are presented with multiple entries for the same title, which they must examine individually to discover the alternative format option. Since MARC e-book records usually replicate their corresponding print records in key access points, "browse" catalog searches retrieve both formats next to each other in the index. The TOC enhancements, which also provide access through browse searches to chapter titles for UFL's print books, do not retrieve the corresponding e-books because they lack the same TOC data. If a user, through any search option, retrieves a record for the print version, the burden is on him or her to go back to the catalog index to note the existence of a record for the e-book, and vice versa. Often that index (because of the search strategy used) does not include the other record at all and calls for a new search (not based on TOC enhancement data) to determine whether or not another version is available. Figure 1 illustrates the problem. UFL owns both the print and e-book versions of Social Cognition: Making Sense of People. When the term "hot cognition" is searched, only the print version record is retrieved because that term is present in the TOC.

To better serve UFL users, the authors sought a means to improve

searching. Although the authors' project led to the use of a system-specific field not widely available at other institutions, this paper seeks to highlight a creative and collaborative approach to a linking problem that resulted in a solution—an approach that might be adopted by others coping with imperfect user interfaces.

Linking Project

Because the authors recognized that UFL holds many titles in both print and electronic forms and, through FRBR, had a heightened awareness of the value of relationships in the catalog, they wanted to find a good way to connect users to the two manifestations represented by the print and e-book versions of the same title. In describing future practice in "Draft Interim Guidelines for Cataloging Electronic Resources," the Library of Congress (LC) made a distinction between the collocating function and the linking function. 11 LC advised using the added entry technique for

collocating and the linking function when appropriate. The 776 linking field (additional physical form entry) was designated to represent horizontal relationships. In 1999 Florida's State University Libraries, including UFL, required the use of the 776 field in its "Access and Cataloging Guidelines for Electronic Resources" for its cooperative digitization program. 12 This proved advantageous when LC wanted to know which of UFL's holdings had been digitized. These were identifiable through the 776 field in catalog records. Critical to the success of the project described here is the fact that NetLibrary distributes catalog records for e-books that include 776 linking fields with both the Library of Congress Control Number (LCCN) and the OCLC number for the print version. Unfortunately, not all e-book vendors do.

Initially, the authors explored making the 776 field a functional linking field in the catalog as are 780 (preceding entry) and 785 (succeeding entry) fields in the UFL catalog. For those fields, a hidden search is trig-

gered by clicking on the linking field, which leads the user to a results list that includes the related serial record. Such indirect linking or "pseudohyperlinking" is described in detail in a 2005 report issued by the Task Group on Linking Entries of the Program for Cooperative Cataloging (PCC) Standing Committee on Automation.¹³ This option, while definitely valuable, is less than ideal because it might lead to no results or force the user to select the related record from an often ambiguous display list. Many libraries use this kind of solution for connecting serial records based on either title or ISSN (International Standard Serial Number) searches, although not as many as the authors expected.

The authors conducted a small, informal analysis in spring 2006 to determine how large academic libraries make use of the earlier and later serials titles linking fields. The authors searched ten serial titles in the catalogs of twenty-one libraries randomly selected from those represented in the Association for Library Collections & Technical Services' Technical Services

Directors of Large Research Libraries discussion group. They observed the options provided to facilitate users' ability to connect between the earlier title (780 field) and later title (785 field) records. Results fell into four categories (see table 1). In some cases, the presence of multiple 78x fields apparently prevented the clicking function of both fields. This exploratory survey, while limited in scope, did uncover typical patterns of service for this function in at least five integrated library management systems. More extensive research along these lines could be both interesting and useful. Recording earlier and later titles of bibliographic records for serials appears to be the most consistent practice; other methods for connecting users to related



Figure 1. Search results for "hot cognition"

| Table 1. Use of linking fields in online catalogs (<i>N</i> =21) | | | |
|--|--|---|------------|
| Presence of Link | Description of Functionality | Libraries | % of total |
| No clickable link | Display of journal record does not include clickable link to earlier or later title. | Duke University, Harvard University, Indiana University, Princeton University, University of Chicago, University of Minnesota, University of Pennsylvania, University of Texas at Austin | 38 |
| Clickable link to search screen | Display of journal record includes clickable link for earlier or later title, and clicking leads to search screen with browse and keyword search options for related title. Clicking there leads to results list for chosen search. | University of California at Berkeley, University of Michigan | 10 |
| Clickable link to results list only | Display of journal record includes clickable link for earlier or later title, and clicking leads to results list for journal title search for related title. | Cornell University, New York University, Pennsylvania State University, Stanford University, University of California at Los Angeles, University of Illinois at Urbana-Champaign, University of Virginia, University of Wisconsin at Madison, Yale University | 42 |
| Clickable link to related record or | Display of journal record includes a clickable link that leads either directly to the related record or to a | Ohio State University, University of Washington | 10 |

titles appear lacking. This falls short of the ideal of providing complete information for catalog searchers as they seek to identify and locate relevant resources.

results lists

results list when the entry is not unique.

The Florida Center for Library Automation (FCLA), which provides automation services to the libraries of Florida's publicly funded universities, was unable to identify any type of similar capability for 776 fields in UFL's ALEPH integrated library management system (implemented in May 2004). ALEPH, however, offers a non-MARC field that can be used to connect records directly. This ALEPH system-specific field allows direct functional connections among bibliographic records, holdings records, and item records. It is useful for connecting bound-together titles, analytics for collection or set level records, and (as the authors discovered) has other interesting possibilities. The PCC Task Group on Linking Entries discussed the benefits of creating such logical links, even suggesting a possible 7XX subfield utilizing local system

or standardized numbers, although the group voiced concern about the limited availability of data in catalogs to support linking and the lack of cataloging staff to enter it.14 The group focused primarily on complex linking among serial records rather than on the straightforward one-to-one relationship between equivalent print and e-book records.

Use of a local non-MARC field could be called "guerilla cataloging" for several reasons. While one can enter and save the field in the local system, this cannot be done when cataloging in OCLC. It may or not be retained upon migration to another system. However, the authors hope that the greater awareness of the value of bibliographic relationships, which has been highlighted by FRBR discussions, will result in improved online catalog systems that will continue to use this data.

Because this field is system-specific to ALEPH, it is not addressed in AACR2, MARC 21, or OCLC cataloging and coding rules. This leaves

local catalogers without guidance as to how and when to utilize new tools such as this, and (at UFL) is leading to open-ended discussions among public and technical services staff. The UFL Cataloging and Metadata Department charged a committee to evaluate local use of this field. The Task Force focused chiefly on serials and special collections materials, but staff are encouraged to explore other possibilities. When used for parallel bibliographic records, this field generates a reciprocal note about the other form available in the public displays of both matched records. These can be clicked to directly connect a user to a matched record. No intervening index displays as it does with the 780 and 785 field linking.

Worthy of note is the way in which UFL leveraged investment in the TOC records. The TOC enhancements were acquired, in effect, as a "two for the price of one" bargain because catalog users retrieve both records (for the print and electronic versions) in searches even though the

TOC are loaded only in the records for the print version.

Implementation

Project implementation was possible without costly investment of staff time to implement the changes and sustain the new steps for linking records. UFL's technical coordinator designed a highly automated, multi-step process. The first step was to identify the pairs of records for alternate versions. This was done by generating system reports of the NetLibrary records with the LCCN in 776 subfield w (record control number). A system-specific loader software (GenLoad) available from FCLA was run to find any matching print version records using an LCCN search. The system numbers were extracted and merged into an Excel spreadsheet. A macro was then used to insert the local field in one record of each pair. While either record could contain the local field, the authors decided to place it in the NetLibrary record because the library was still in the process of adding the tables of contents to the print version records and did not want to risk overlaying the linking data. The flowchart in figure 2 shows the steps involved.

After the field is inserted in one of a pair of records, a user who retrieves one of the records is both informed about the related item and supported in easily connecting to the other record by a simple click. Thus, a remote user who retrieves the record for the print version will easily benefit from connecting to the electronic version, which can be viewed without leaving the computer. Similarly, a user who first retrieves the e-book record but desires to borrow and use a print book will be connected to the record giving shelf location and availability. Figures 3 and 4 show the OPAC views for the linked records.

This process, in addition to the advantage of using little staff time,

be applied retrospectively to records that may have been in the catalog for years. While UFL has already begun connecting new boundtogether items and analyzed sets of various kinds, there is also interest in identifying other methods for better presenting other related materials to users. The process of identifying the record pairs using the LCCN in the 776 field may be replaced, at some point, by a more sophisticated process that could use author-title combinations to relate many of the varied formats and editions held by the library. Such methods already are being used else-

where by the larger-scale FRBR projects and library catalog systems to process existing catalog data to enable new indexing and display options. The authors observed that Duke University makes use of a feature in their catalog that creates an author-title entry for each record. When clicked, the authortitle entry opens a window populated with matching records, although versions are not differentiated. While not quite as intuitive as a clickable note on the record that says "Available in other form: E-book" and links directly to the corresponding record, the author-title link allows the user to navigate among different versions. The authors encourage others to explore creative solutions that will overcome the absence of data (for example, uniform titles)

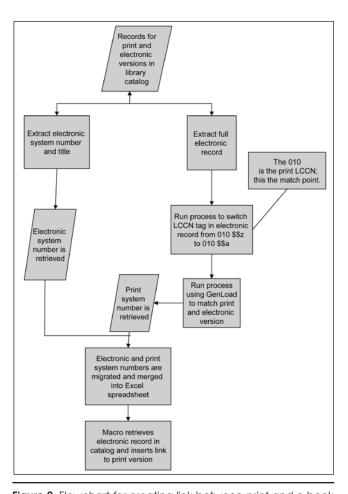


Figure 2. Flowchart for creating link between print and e-book

that might have facilitated navigation among different versions, but which were not added to records for cost reasons in the past.

Conclusion

Linking between NetLibrary and print version records has provided an exciting new way to connect users to materials they need in the format they prefer. It extends the benefits of TOC enhancements that were only in the records for print versions to the records for matching electronic versions. It does so in a way that facilitates better user awareness of and connection between the two versions. Additional access thus gained includes browse searches for authors

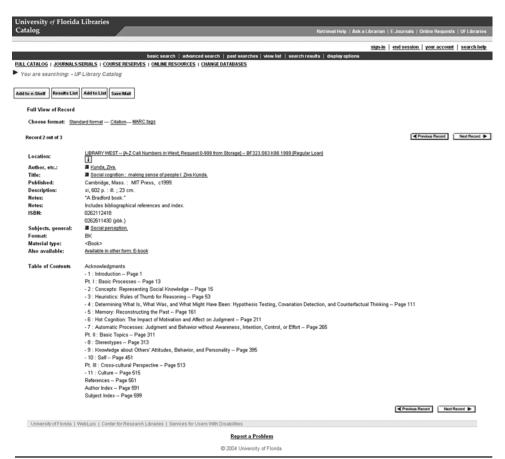


Figure 3. OPAC view of print version



Figure 4. OPAC view of e-book version

and titles of chapters, and keyword access to terms included in the TOC fields of those records. This project enables users to access their preferred format for a given content regardless of the retrieval method and, in so doing, promotes what Tillett refers to as the fifth function of the catalog—to support navigation.15 Better navigation options are necessary if libraries are to satisfy users' growing needs and expectations. While the authors' preferred method for linking was, and continues to be, through the established MARC standard linking fields, in the absence of that possibility, an alternative process served the immediate need.

This exploration of systemspecific functionality had many other beneficial outcomes. The authors identified a function of the bibliographic record, thought in terms of the user's tasks, and then translated it into reality. This process led to a deeper understanding of the uses of linking fields and how they can be incorporated into local workflow and thinking. The process itself expanded awareness and knowledge of the many issues that arise when librarians today try to offer users specific improvements to navigational and display capabilities in their catalogs. It was a learning experience that brought a refined understanding of FRBR terms. The catalogers' collaboration with FCLA and UFL's technical coordinator that made this project possible benefited from using the FRBR model as a conceptual tool—one that enables various partners in the world of online information to communicate with each other effectively. The project made clear that today's catalogers must go beyond their traditional functions, explore new options in

technology, and communicate their ideas to those who can implement them and to those who benefit from the outcome. It highlighted the need for more consistent and coordinated practices by the cataloging community, system vendors, and suppliers of cataloging services. The next stage of the project will be to speak with public services staff at UFL about moving the linking note to a higher level of visibility near the top of the record and in the brief view. Collaborating across library divisions and with systems designers to improve the navigability of the catalog for the users is in the best tradition of cataloging and represents the greatest hope for realizing the spirit of *FRBR*.

References

- 1. Standing Committee of the IFLA Section on Cataloguing, Functional Requirements for Bibliographic Records: Final Report (Munich: K. G. Saur, 1998).
- 2. Barbara B. Tillett, "FRBR and Cataloging for the Future," Cataloging & Classification Quarterly 39, no. 3/4 (2005): 197-205.
- 3. "RedLightGreen Will Enhance Stud-

- ent Research," RLG News 57 (Fall 2003): 1.
- 4. OCLC, "Fiction Finder: A FRBRbased Prototype for Fiction in WorldCat." www.oclc.org/research/ projects/frbr/fictionfinder.htm (accessed Oct. 21, 2005).
- Stefan Gradmann, "rdfs:frbr-Towards an Implemenation Model for Library Catalogs Using Semantic Web Technology," Cataloging & Classification Quarterly 39, no. 3/4 (2005): 64, 66.
- 6. Martha M. Yee, "FRBRization: A Method for Turning Online Public Finding Lists into Online Public Catalogs," Information Technology and Libraries 24, no. 2 (June 2005): 78.
- 7. Ibid.
- 8. Jennifer Bowen, "FRBR Coming Soon to Your Library?" Library Resources & Technical Services 49, no. 3 (July 2005): 175-88.
- 9. Ibid., 181.
- "Membership Reports: 10. OCLC, Information Format Trends: Content, Not Containers" (2004). www.oclc .org/reports/2004format.htm (accessed Oct. 23, 2005).
- 11. Library of Congress, Cataloging and Support Office, "Descriptive Cataloging Manual: Draft Interim

- Guidelines for Cataloging Electronic Resources" (Dec. 18, 1997). www.loc .gov/catdir/cpso/dcmb19.pdf (accessed Dec. 19, 2005).
- Cataloging and Access Guidelines 12. for Electronic Resources Subcommittee, Technical Services Planning Committee, Council of State University Libraries, "Access and Cataloging Guidelines for Electronic Resources." www.lib.usf.edu/tech services/CAGER/CAGERGuidelines Contents.html (accessed May 5,
- 13. Program for Cooperative Cataloging, Standing Committee on Automation, "Task Group on Linking Entries Final Report, Feb. 2005." www.loc.gov/ catdir/pcc/archive/tglnkentr-rpt05.pdf (accessed May 1, 2006).
- 14. Ibid.
- 15. Barbara Tillett, "FRBR and Cataloging Rules: Impact on IFLA's Statement of Principles and AACR/RDA" (paper presented at FRBR in 21st Century Catalogues: An Invitational Workshop, Dublin, Ohio, May 2-4, 2005). www.oclc.org/research/events/ frbr-workshop/presentations/tillett/ FRBR_and_cat_rules.ppt (accessed Dec. 21, 2005).