February 2007

ATG Special Report -- Cataloging eBooks: an Overview of Issues and Challenges

Kristin E. Martin

University of North Carolina at Chapel Hill, kmarti@email.unc.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

Part of the [Library and Information Science Commons](#)

Recommended Citation

Martin, Kristin E. (2007) "ATG Special Report -- Cataloging eBooks: an Overview of Issues and Challenges," *Against the Grain*: Vol. 19: Iss. 1, Article 12.

DOI: <https://doi.org/10.7771/2380-176X.5233>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

of money for public relations campaigns by all stakeholders, including, for example, the PLoS advertising and marketing budget that was close to a half a million dollars in 2004. PLoS is not the only group that has launched such campaigns; SPARC has been very aggressive, and now the Association of American Publishers has retained a public relations guru.

Thousands and thousands of dollars are being expended on the pro-con open access debate, and yet it has not been fully examined from a fiduciary point of view.

Without a sound fiduciary model that is sustainable, all the rest is an exercise of eloquent (and very repetitious) prose. And wasted money.

We do not know if the money for sustainability and affordability is assured. Who is going to demand that answer? Until we have long standing evidence of sustainable and affordable models, we have to be absolutely sure that ideological fervor does not overtake the realities of what all this will really cost, and, please.... Repeat after me, where will this money come from? And for how long?

Does the subscription system have flaws? Indeed it does. Should publishers and librarians still try to create a better system together while we grapple with the unknown? Indeed we should.

Beware of unintended consequences. It is well to keep in mind the phrase "Don't Throw the Baby out with the Bath Water." Credited to the first written occurrence in the satirical book, *Narrenbeschwörung* (1512), by Thomas Murner (1475-1537), a chapter is entitled such: it is a treatise on fools who by trying to rid themselves of a bad thing succeed in destroying whatever good there was as well.

Well said. And very good advice.

And remember to send for your application to law school. 🐼

Balancing the Needs ... from page 43

A longer term benefit of this approach is the development of understanding and familiarity between MASU staff and content producers. It is hoped these relationships will increase their comfort with approaching MASU for future assistance or advice regarding metadata or cataloging. Moreover, it provides a tested model for working with content providers outside the library, say the engineering faculty, who want to contribute materials to the DAMS for safeguarding.

MASU is confident our extensible normalization approach meets the needs of aggregating legacy data while remaining flexible enough to evolve along with the changing needs of the DAMS and the UCSD Libraries. 🐼

ATG Special Report — Cataloging eBooks: an Overview of Issues and Challenges

by Kristin E. Martin (Electronic Resources Cataloger, Davis Library, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514-8890; Phone: 919-962-0153; Fax: 919-962-4450) <kmarti@email.unc.edu>

Web-based eBooks have become popular with a wide variety of library users and are an increasingly important part of libraries' collections. eBook content now encompasses databases of retrospective eBooks (such as *Early English Books Online* or *Literature Online*), aggregated packages of relatively current content from multiple publishers provided by an eBook vendor (such as NetLibrary or ebrary), and titles offered directly from the publishers (such as Springer and Elsevier). As the volume of eBook content grows, libraries are grappling with how to integrate this content into their online catalogs. Librarians trying to provide title-level catalog access to their eBook collections must answer multiple questions to determine optimal workflow. Questions include:

- Where will the record come from?
- Can the eBook records be processed in batch?
- Should electronic holdings be placed on the same record as print holdings?
- What changes will need to be made to vendor-supplied records?
- How can the records remain accurate as titles are added and subtracted to eBook collections?
- Should holdings be added to OCLC? Why or why not?

At the University of North Carolina at Chapel Hill University Libraries we have been analyzing the issues raised by these questions to figure out how to provide the best access to our growing number of eBook collections. This article does not purport to be able to answer all of those questions, but rather introduces them as a series of topics that librarians will need to address when adding eBook records to their catalogs.

Although many eBook collections offer their own search mechanisms, having individual title records for eBooks in the OPAC provides library users with a single discovery tool for eBook titles across all collections and allows users to simultaneously view the library's print and electronic holdings. Initial studies of eBook use, mainly looking at NetLibrary content, have demonstrated the importance of catalog records in enhancing use to electronic books (for example see Dillon 2001; Gibbs 2001; Langston 2003). In a particularly dramatic example at the University of Rochester, the use of the NetLibrary eBooks increased by 755 percent when comparing use in the five

months before and after loading the catalog records (University of Rochester Libraries 2001). Later studies of eBook usage have taken title-level catalog records for granted, when comparing usage of print and electronic counterparts (Christianson and Aucoin 2005; Littman and Connaway 2004).

Despite the preponderance of evidence supporting the need for access to eBooks through the catalog, many libraries have been quicker to purchase eBooks than to provide title-level access through the OPAC. Several issues have contributed to this delay in cataloging. Acquisitions and cataloging workflows have been developed around the processing of physical items, generally on a title-by-title basis, while eBooks are intangible objects that have frequently been made available in large collections that could overwhelm a cataloging department. Staff may still have a "print is primary" mindset, and view electronic resources as supplementary, rather than as a core part of the library's collection. Additionally, eBooks may only be available on subscription, rather than owned, and titles may be swapped in and out as new material becomes available in large collections. Finally, cataloging standards for electronic resources have been subject to multiple revisions, making libraries reluctant to spend time and resources creating catalog records that will need to be updated.

Fortunately, as eBooks have become more widespread, so has the availability of MARC records for individual titles, frequently from the vendor. One of the first questions librarians must consider is whether to use vendor-supplied records for eBook collections. Records may be free with the purchase of the resource, available for a fee through OCLC's Collection Sets, or available for purchase separately from the vendor, with price and quality of records varying widely. These vendor-supplied records free the library from having to provide title-by-title cataloging, and may be loaded quickly into the catalog; however, there is still work to be done at the library's end.

Librarians must scrutinize the records carefully for quality and ensure the records correctly represent the titles the library purchased. Given the size of some eBook collections, it may not be possible to examine each record, but it is important to at least spot check records or to examine a selective sample for quality and accuracy. To date, vendor records have typically treated eBooks as electronic

continued on page 46

reproductions, frequently basing the record for the electronic version on a record for the print version of the book. This "electronic reproduction" model follows the **Library of Congress Rule Interpretation 1.11A**, which allows all of the original data of the eBook to remain in the standard **MARC** fields (such as the 300 note for physical description and 260 for the publication, distribution, etc.), while information about the reproduction is listed separately in a 533 note. This method of handling reproductions had already been in use for microforms by the **Library of Congress**, although it does not follow *The Anglo-American Cataloguing Rules*. New guidelines from the *Program for Cooperative Cataloging (PCC)* for monograph aggregator vendors also recommend the electronic reproduction model for PDF files (*Program for Cooperative Cataloging* 2006). By having vendors base their records on the print version, the quality of the records can be quite high, depending on the level of the original cataloging. While this model works particularly well for retrospective digitization projects, it has some limitations.

First, the reproduction model becomes increasingly untenable for those titles that are actually distributed in electronic format simultaneously with or prior to the print edition. Although most eBook collections to date have been retrospective collections or published after the print version is available, future publications from large publishers, such as **Springer**, promise to begin supplying libraries with eBooks before the print publication is produced, so the eBook will need to be cataloged independently of any print version. Second, following **LCRI 1.11A** is not consistent with **AACR2** and will most likely not be consistent with **Resource Description and Access**, the successor to **AACR2**. With the implementation of RDA and online catalogs that follow the **Functional Requirements for Bibliographic Records (FRBR)**, in which both eBooks and print books could be identified as manifestations of the same work, the reproduction model may fall out of favor.

Regardless of whether eBooks are cataloged as reproductions or as born-digital, because the resulting records are vendor-specific, multiple records for the same title, as provided by different aggregators may result, particularly when purchasing sets of eBooks without title-by-title selection. While **Cooperative Serials Online**, the body overseeing cooperative serials cataloging, has implemented aggregator-neutral records for serials and is considering them for integrating resources (**Cooperative Serials Online** 2006), there has not been consensus on how to handle monograph titles offered by multiple aggregators. As a result, users in the OPAC may retrieve multiple hits for the same title, one from **NetLibrary**, one from **ebrary**, one from **Safari Tech Books Online**, etc. This problem too may be solved in the future with a good **FRBR** implementation that allows hierarchical linking of records. The record for the work could provide general information about

the title, while information about each aggregator could be presented at the manifestation level and link to the record for the work.

eBook records, regardless of quality, frequently will need adjustments before being "catalog ready." The PCC guidelines provide a list of required and recommended fields for electronic reproductions, and librarians should use this as a checklist when examining vendor records. Some common errors include leaving the original **OCLC** number or ISBN for the print record and book, which can create confusion for local catalog systems that use these numbers for duplicate record detection; leaving supplementary materials notes in the 300 field which no longer apply to the electronic version, or forgetting to add the general material designation "[electronic resource]." The library may also need to make individualized customizations, such as altering URLs to have them go through a proxy server or adding special fields to identify the record set. As much as possible, this work should be done at a batch level, either by performing global changes through an integrated library system or using a program like **MarcEdit**. **Sanchez** (2006) provides an excellent description of using **MarcEdit**, **Microsoft Word**, and **Microsoft Excel** to customize **NetLibrary** records in batch prior to loading to the catalog.

Once records are in the catalog, a library's work is not done. In a small survey done in 2005 on eBook adoption in libraries, all respondents noted that keeping the catalog up to date was the greatest maintenance issue related to eBooks (**Blummer** 2006). This is particularly true for subscription eBook packages where titles may be added and subtracted from a library's subscription. Records may not be available at the same time as the eBooks, and updates may only be placed on a vendor Website without notification. For those titles for which the library owns the print version or an additional electronic version, librarians must decide whether to try to consolidate the information onto a single record. Consolidating records when working with loads of eBook records for large collections requires a substantial amount of additional work after the load. Additionally, maintaining a single record is more difficult because of the potential need to remove titles from the catalog, should an eBook subscription be dropped.

If vendor-supplied records turn out to be of inadequate quality, too expensive, or unavailable, libraries still have the option to catalog eBooks title-by-title. While obviously more feasible for smaller collections, libraries may also decide to go this route when purchasing perpetual access to expensive titles. It may also be possible to use **OCLC** to search for groups of titles from eBook collections and batch process the records, using some of the same techniques as on vendor records. Expect staff used to managing print resources to need additional training and potentially new equipment, such as dual monitor systems, to facilitate the processing of the eBooks. If libraries need to perform original cataloging on their eBooks, (**Bothmann** 2004) provides a detailed description and examples of eBook records,

although the article is based on the 2002 revision of **AACR2**.

The final issue for eBook cataloging relates to the availability of holdings information in national union catalogs, i.e., **OCLC**. Unless using records from **OCLC's Collection Sets**, libraries loading records directly from vendors will not have their corresponding holdings information (or possibly even the record) in **OCLC**. Reducing the information in **OCLC** creates several problems: less reliable information for interlibrary loan, less ability to mine **OCLC** data for accurate information regarding holdings, and less sharing of cataloging records. Additionally, for those libraries that might be considering using **OCLC** as access to their library holdings, rather than maintaining a separate system (as suggested in *Bibliographic Services Task Force* 2005), such a discrepancy will not be an option. The inability to use eBooks for ILL may not be important for certain eBook licenses but many eBook licenses do allow for at least limited lending. The sharing of records through **OCLC** also provides a way for libraries to share access to records that are enhanced or improved. Libraries will need to weigh the advantages of having their holdings in **OCLC** against the extra work required to add that information to the database.

As this paper illustrates, the cataloging of electronic books, while vital to enhancing access, is not a simple task and requires careful analysis and thoughtful decisions. Because eBooks are relatively new, cataloging principles and workflows are still in flux. Already, though, some standards are being established, such as the recommendations set forth by the *Program for Cooperative Cataloging*, which librarians can use as guidelines when determining the best way to handle new collections. Decisions for how to catalog, perform quality control, maintain records, and update holdings in **OCLC** will vary from library to library, depending on individual circumstances. Making these decisions and planning for the eBook cataloging workflow should be an important part of the whole process of purchasing sets of eBooks, rather than just an afterthought. Regardless of individual decisions, the future for cataloging eBooks should include the ability for libraries to accurately provide access to electronic books and for cooperative cataloging and sharing of information, so well established for print materials, to continue.

References

Bibliographic Services Task Force. 2005. Rethinking how we provide bibliographic services for the University of California. Oakland, Cal.: University of California Libraries. Available: <http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf> (accessed December 19, 2006).

Blummer, Barbara. 2006. eBooks revisited: the adoption of electronic books by special, academic, and public libraries. *Internet Reference Services Quarterly* 11, (2): 1-13.

Bothmann, Robert. 2004. Cataloging electronic books. *Library Resources & Technical Services* 48, (1) (January): 12-19.

continued on page 47

Cataloging eBooks from page 46

Christianson, Marilyn, and Marsha Aucoin. 2005. Electronic or print books: Which are used? *Library Collections, Acquisitions, and Technical Services* 29, (1): 71-81.

Cooperative Serials Online. Homepage of the CONSER Program. Washington, D.C.: Library of Congress. 2006 [cited December 19 2006]. Available: <http://www.loc.gov/acq/conser/> (accessed December 18, 2006).

Dillon, D. 2001. eBooks: the University of Texas experience, part 2. *Library Hi Tech* 19, (4): 350-362.

Gibbs, Nancy J. 2001. eBooks two years later: the North Carolina State University perspective. *Against the Grain* 13, (6) (December 2001/January 2002): 22-26.

Langston, Marc. 2003. The California State University eBook Pilot Project: implications for cooperative collection development. *Library Collections, Acquisitions, and Technical Services* 27, (1) (Spring): 19-32.

Littman, Justin, and Lynn Silipigni Conaway. 2004. A circulation analysis of print books and eBooks in an academic research library. *Library Resources & Technical Services* 48, (4) (October): 256-262.

Program for Cooperative Cataloging. 2006. MARC record guide for monograph aggregator vendors. Washington, D.C.: Program for Cooperative Cataloging. Available: <http://www.loc.gov/catdir/pcc/sca/FinalVendorGuide.pdf> (accessed December 19, 2006).

Sanchez, Elaine, Leslie Fatout, and Aleene Howser. 2006. Cleanup of NetLibrary cataloging records: a methodical front-end process. *Technical Services Quarterly* 23, (4): 51-71.

University of Rochester Libraries. 2001. netLibrary eBook usage at the University of Rochester Libraries. Rochester, N.Y.: University of Rochester Libraries. Available: <http://www.lib.rochester.edu/main/studies/analysis.pdf>.

Column Editor's Note: In early January 2007 YBP Library Services called me to let me know that Bob Nardini had left YBP and joined a competitor, Coutts Library Services. I spoke with Bob two weeks later. — PC

Upgrade Your Print Subscription & Add Online Full-Text!

Electronic access to the American Economic Association journals for libraries, institutions, and firms!

Institutional site licenses to issues of the *American Economic Review*, *Journal of Economic Literature*, and *Journal of Economic Perspectives* are now online!

Institutional access includes reference links to full-text articles through CrossRef and JSTOR, links to full-text articles citing AEA articles, table of contents alerts, and

searchable full-text. Site licenses include all years from 2002 forward so long as a subscription is active!

Contact the AEA at www.aeaweb.org/subscribe.html or your subscription agent to add an institutional site license to the AEA journals today!

**Site License price as of 1/1/2007
(includes all three AEA online journals).....\$345.00**

Ask about . . .

EconLit – The foremost bibliographic database in economics is linked to full-text now more than ever!

Full-text linking options are available from all EconLit vendors: CSA – Dialog – Ebsco – OCLC – OVID/SilverPlatter

American Economic Association
2403 Sidney Street #260
Pittsburgh, PA 15203-2167, USA
info@econlit.org

ATG Interviews Bob Nardini

Group Director of Client Integration, Coutts Library Services

by Pam Censer (Science Librarian, University of Florida) <pam.censer@gmail.com>

ATG: In the press release from Coutts about your appointment your title includes the phrase "group director of client integration" was does that mean?

BN: I think it is more or less a British term. It means that my job will focus on making sure

that Coutts' services mesh with the processes of academic libraries, so that it is easy to do business with Coutts.

ATG: You joined Yankee Book Peddler (now YBP) in July 1985, many academic
continued on page 48