

E-book MARC records: do they make the mark?

The rapidly growing market in e-books and the options for library acquisition of these materials lead to questions regarding access. Because MARC records increase use of e-books, and some vendors provide MARC records along with e-book purchases, many libraries use their library catalogues as a conduit to e-books. This article explores the challenges associated with these records. Consideration is given to the quality of information provided by vendors, user expectations and experiences, and cataloguing workload in today's libraries. The authors seek to educate librarians about the issues associated with e-book MARC records, to empower them with questions to ask of vendors regarding this data, and to encourage them to weigh the costs and benefits associated with using this free information.



DORALYN ROSSMANN

Collection Development Librarian & Assistant Professor

AMY FOSTER

Cataloguer & Associate Professor

ELIZABETH P BABBITT

Electronic Resources Librarian & Assistant Professor

Montana State University Libraries, USA

Introduction

Given the ever increasing options in e-book offerings, today's librarians face the pressing question of how to provide access to this information. Many vendors, anticipating the library catalogue as a conduit to accessing e-books, offer MARC records as a part of e-book packages. MARC records, a long-time fixture of libraries' organization of information and a requirement for many libraries, follow commonly accepted standards developed over many decades. Vendor-produced MARC records – from home-grown records to OCLC offerings downloadable by vendor or package – may differ in quality from what has been expected of our local catalogue entries. Additionally, the purchasing models for e-books present new challenges with regard to MARC data provision and record management.

This article has three primary goals. First, it is intended to provide a fairly comprehensive overview of the issues surrounding e-book access via MARC records in an integrated library system (ILS). Second, it is hoped that this overview offers an effective set of questions for librarians to pose to e-book and ILS vendors when working with e-books and MARC. Finally, this article challenges the library community to consider whether some of

the trade-offs associated with using free MARC records are worthwhile.

Literature review

In conducting a survey of the literature, we found that many of the issues that existed as early as 2001 continue to be problematic today.^{1,2} Some of the main conclusions relating to access are:

- the catalogue or ILS is viewed as the best or most straightforward means of finding and accessing e-books^{3,4,5}
- web-based e-books which do not require special hand-held reading devices are preferred by most academic libraries, because such devices are often expensive and can be impractical for students^{6,7}
- no matter what the quality of the records, some level of cataloguer intervention is likely to be needed before the records are considered acceptable to be added to the ILS.⁸ This intervention may be as simple as adding a proxy string to the URL allowing authorized users access to materials from off campus, or as complex as making extensive edits such as filing

indicators, capitalization discrepancies, diacritics, subfield codes, authorized forms of names or subject headings, or punctuation.⁹

- it is a challenge to keep the catalogue up to date as titles are added/removed from subscription packages. Also, there is debate over which is more feasible or desirable in an ILS as regards e-books: a single record or duplicate records for the same title.¹⁰

Observations and discussion

When considering the provision of access to e-books, it is useful to understand the e-book market and what questions might be asked of vendors. Purchasing models for e-books differ considerably from those traditionally used by libraries for print books, or print/electronic serials. Usually, print books are bought outright from the publisher or through a book vendor such as YBP or Amazon. Once a print book is acquired, there is little to no communication between the purchaser and the vendor concerning that title.

Serials are generally purchased as an annual subscription, again, from either the publisher, a subscription agent, or an aggregator. For electronic serials, an annual hosting fee paid to the publisher or database provider may be required (such as with ProQuest's Historical Newspapers). E-book purchasing can follow any of these models or can take on a life of its own. E-books can be purchased outright (and, perhaps, include a small hosting fee). For example, an institution pays outright for access to a list of titles but, each year, they pay a fee to the vendor to host the materials on the vendor's server. They can also be bought with an annual subscription. Annual subscriptions allow for some flexibility in replacing older editions. Also, titles can be leased from the publisher as in ProQuest's Safari Tech Books model. In a leasing arrangement, an institution purchases a certain number of titles but the selected titles can vary as needed. For instance, a library may purchase 50 titles but they can be chosen from a list of thousands. Based on the courses taught from term to term, selected titles may be swapped out for ones that are more relevant. Thus, those 50 titles may represent 150 or more choices selected and deselected over the academic year.

Given the options for purchase, how might users access e-books? Do library users expect to

find e-books in the library's ILS, through a resource management page, or in a publisher's database, for example, as with SpringerLink¹¹? What other discovery tools for e-books are available to users? One such tool, in concept, is a fully-functioning federated search which would retrieve articles, printed books and electronic books and their chapters regardless of whether they are in the ILS, in an aggregated database, or at a publisher's website. Another discovery tool is provided through the OpenURL linking protocol which can be added to e-books at a chapter level, providing access through a library's A-Z title list. Google Book search and Google Scholar, popular choices for researchers, index full text materials across a broad spectrum of formats; but, it is unclear how often material is updated and the data are not always optimized to specify a library's holdings. More commonly accepted standards for e-book records and non-proprietary systems of sharing this bibliographic information across networks would increase discovery and reliability as well.¹²

Amid the multiple options for discovering e-books, it has been widely reported that usage goes up after records have been added to the ILS.^{13,14,15} Indeed, at the authors' institution, Montana State University, a similar increase in usage was observed. The number of times the Readex Serial Set, Knovel e-books, and Springer e-books were accessed tripled between fall 2008 and spring 2009 after MARC records were loaded into the ILS. Although there is not a clear, one-stop resource for e-book access, providing MARC records in the ILS seems to be an effective tool for reaching today's researchers.

Considerations and implications

The implications of opting to provide MARC records in the ILS will be explored in the following section.

Quality and logistical challenges of data provided by vendors

Vendor provision of e-book MARC records poses unique challenges. An awareness of these issues will help librarians work with vendors to decide what level of quality is acceptable in the MARC data provided. Some vendors do not create call numbers and detailed subject headings in their free records, so browsability is reduced (unless libraries spend time or money adding proper call numbers and additional subject access points). While many

libraries follow the field's cataloguing standards and authority control practices, do vendors providing e-book MARC records adhere to these same standards? To be fully informed about the practices being used, librarians should consider the following questions when working with vendors to acquire e-book MARC records:

- Who creates the data (e.g. librarians/non-librarians)?
- Is there authority control and does it include Library of Congress subject headings and names (in the USA) or other commonly accepted standards?
- Are call numbers provided and what schema is used?
- Are other, non-standard headings added (e.g. vendor-created subject headings)?
- When subject headings are added, are they using specific headings or very broad ones (e.g. 'Water chemistry–Mathematical models' as compared to 'Chemistry')?
- Are table of contents notes, which can enhance usability, provided?
- Do they ever provide corrected MARC records or, once created, are they considered complete (e.g. Springer provides corrected records monthly)?

In addition to questions about the quality and structure of e-book MARC records, selectors of e-books may want to consider other logistical issues. For example, are the MARC records free with the purchase of e-books? Are there records available as a subset in OCLC as well? Unlike the 'free' records provided by vendors, records from OCLC include a download cost but they are 'catalogue ready' for the most part. Can the records be shared in a consortial catalogue environment or would each library need to purchase access to these records? For instance, Readex's Serial Set records cannot be shared unless each library pays for these records. Are the records in MARC, MARC21, MARCXML, UKMARC, Dublin Core or other formats (e.g. Serial Set from Readex is offered in MARC and Dublin Core)?¹⁶ Are the records available at the time of purchase or are they forthcoming? If the records are forthcoming, when will they be available and what guarantees are there that they will be provided eventually? Are the e-book URLs stable? Is the collection OpenURL compliant? If choosing a leasing model, how does the vendor provide MARC records? For example,

ProQuest's Safari Tech books allows cataloguers to download an entire set of MARC records, or only the records for titles added for lease since the last download, or just those records no longer leased since the last download (for removal from the ILS).

ILS challenges

Once acquired, access to e-book MARC records needs to be provided to users. For those who choose to use a catalogue, what questions might be asked of the ILS vendors? Possible questions might be:

- How do they recommend handling e-book records in a shared/consortial catalogue environment?
- Should e-books be loaded in the ILS in the same manner as other records, or should they be segregated somehow for ease of searching and maintenance?
- What options exist for display of the URL: just the URL itself, a public note, the URL and a public note, a text link, or an icon link? (There may be other hyperlinked fields in the record such as subject headings and author name which may add to the visual noise of the record display. Ultimately, librarians want to avoid a 'sea of blue links' but want to be clear where the URL is and to which library it belongs.)
- How do the ILS vendors recommend exclusion of e-book records from exports for regular authority control processing (if desired by the library)?
- Unlike print books, e-book pricing models encourage purchases of large sets rather than individual titles for maximum value. In turn, more records are loaded into the catalogue than in a purely paper book environment. Are vendors creating systems that can accommodate this increase in data and that can search, re-index, run reports and retrieve data faster? Additionally, this scenario may result in increased costs from ILS vendors if pricing is based on the number of records in the system.

Also, consider how different pricing models can affect how records are handled in online catalogues. If the titles are part of a leased subscription, it is necessary to be able to easily add and remove records in batches to keep the catalogue up to date. This can be complicated further in a shared catalogue or consortial arrangement. Consortium members may use different e-book purchasing models, making the use of shared records and

batch removal of individual library holdings much more problematic. Libraries in this situation may want to consider having each library load records and holdings individually instead of sharing records. This approach allows for batch removal. This decision, however, will have other consequences, such as creating duplicate records. If sharing bibliographic records is the norm for a library system or consortium, allowing each unit to load their e-book records separately differs from how the rest of the catalogue is handled. This setup may cause confusion and waste time for users when they encounter multiple records for the same title.

Conclusions and future developments

The adage “there’s no such thing as a free puppy” should be considered relative to e-book MARC records. Vendor-supplied MARC records may need a good brushing and cleaning, some feeding of data, and booster shots before they are ready to go for a walk. In other words, receiving ‘free’ vendor-supplied MARC records when purchasing e-book packages may require some intervention by cataloguers before loading them into the catalogue. Libraries are at a crossroads when it comes to the issue of e-books. Over the past two decades, the library profession has struggled with provision of access to e-journals. Vendors such as SerialsSolutions, Innovative Interfaces Inc., and Ebsco emerged as alternatives to ILS interfaces for access to this information. But, the e-book is a different beast, with baggage of nostalgic associations with traditional books and with different challenges in comparison to serials (no volumes, issues, numbers, etc.). When libraries opt for e-book purchases, they generally provide access through their ILS and often take advantage of free MARC record offerings. This is likely to be the case at least until federated search, or other unified discovery interfaces, such as SerialsSolutions’ Summon, are widely used.^{17,18}

As has been explored, these records do incur costs and these costs should not be ignored or taken lightly. What is the cost of making vendor-supplied records ‘perfect’ before adding them to the catalogue? Is that cost justified? Is a perfectly tidy database necessary? Is easy access to e-books worth a messier catalogue database? Should we be demanding more from vendors who are not

conforming to more rigorous standards such as those used by the Library of Congress? Even with the total cost of processing such collections and loading them as a batch, vendor-supplied records represent a better deal than downloading records from OCLC individually. But, what is the cost relative to the quality of the user experience? These questions are likely to be answered differently by each library facing adding e-books to their catalogue. While there are no right or wrong answers, each library must decide what works best for their situation, based on staffing levels, budgeting implications and other workload concerns. While these questions need to be addressed at a local level, there are broader professional issues. Are we demanding enough of our vendors relative to cataloguing standards and the quality of data provided?

References

1. Joint Information Systems Committee (JISC), *Promoting the Uptake of E-books in Higher and Further Education*, Gold Leaf, 2003.
http://www.jisc.ac.uk/uploaded_documents/PromotingEBooksReportB.pdf (accessed 15 September 2009).
2. Research Information Network (RIN), *Creating Catalogues: Bibliographic Records in a Networked World*, RIN, 2009.
<http://www.rin.ac.uk/creating-catalogues> (accessed 15 September 2009).
3. Dillon, D, E-books: The University of Texas Experience, part 1 *Library Hi Tech*, 2001, 19(2), 113–124.
4. Snowhill, A, E-books and their future in Academic Libraries, *D-lib Magazine*, 2001, 7(7/8).
5. Martin, K, Cataloging eBooks: an Overview of Issues and Challenges, *Against the Grain*, 2007, 19(1), 45–47.
6. Snowhill, A, see ref. 4.
7. Dillon, D, see ref. 3.
8. Martin, K, see ref. 5.
9. Sanchez, E, et al. Cleanup of NetLibrary Cataloging Records: A Methodical Front-End Process, *Technical Services Quarterly*, 2006, 23(4), 51–71.
10. JISC, see ref. 1.
11. JISC, see ref. 1.
12. Martin, K, see ref. 5.

13. JISC, see ref. 1.
14. Dillon, D, see ref. 3.
15. Connaway, L and Snyder, C, Transaction Log Analysis of Electronic (eBook) Usage, *Against the Grain*, 2005, 17(1), 85–94.
16. Vasileiou, Hartley, and Rowley provide a helpful overview of format provisions by vendor in their e-book marketplace survey. Vasileiou, M, Hartley, R, and Rowley, J, An Overview of the E-book Marketplace, *Online Information Review*, 2009, 33(1), 173–192.
17. Hadro, J, Serials Solutions Announces Summon, a New Unified Discovery Interface, *Library Journal Blog*:
<http://www.libraryjournal.com/article/CA6630537.html> (accessed 15 September 2009).
18. Introducing Summon, *The Shifted Librarian*:
<http://theshiftedlibrarian.com/archives/2009/01/24> (accessed 15 September 2009).

Article © Doralyn Rossmann, Amy Foster and Elizabeth P Babbitt

Article received 28 July 2009; revised 12 September 2009; accepted 14 September 2009; published online November 2009

■ Doralyn Rossmann
 Collection Development Librarian &
 Assistant Professor
 Montana State University Libraries
 P.O. Box 173320
 Bozeman, MT 59717-3320, USA
 E-mail: doralyn@montana.edu

■ Amy Foster
 Cataloguer & Associate Professor
 Montana State University Libraries
 PO Box 173320
 Bozeman, MT 59717-3320, USA
 E-mail: amy.foster3@montana.edu

■ Elizabeth P Babbitt
 Electronic Resources Librarian & Assistant Professor
 Montana State University Libraries
 PO Box 173320
 Bozeman, MT 59717-3320, USA
 E-mail: lbabbitt@montana.edu

To view the original copy of this article, published in the *Serials* e-book supplement, click here:
<http://serials.uksg.org/openurl.asp?genre=issue&issn=0953-0460&volume=22&issue=3&supp=1&spage=S46>

The DOI for this article is 10.1629/22S46. Click here to access via DOI:

<http://dx.doi.org/10.1629/22S46>

For a link to the table of contents for the *Serials* e-book supplement in which this article first appeared, click here:

<http://serials.uksg.org/openurl.asp?genre=issue&issn=0953-0460&volume=22&issue=3&supp=1>

For a link to the table of contents for the main issue of *Serials*, click here:

<http://serials.uksg.org/openurl.asp?genre=issue&issn=0953-0460&volume=22&issue=3>