

Purdue University
Purdue e-Pubs

Charleston Library Conference

The E-Book Story: The Key to a Happy Ending

Denise Branch
Virginia Commonwealth University, dbranch@vcu.edu

Katy Aronoff
Ex Libris, Katy.Aronoff@exlibrisgroup.com

Evelyn Elias
Taylor & Francis, Evelyn.Elias@taylorandfrancis.com

Emma Waecker
EBSCO, ewaecker@ebSCO.com

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

 Part of the [Higher Education Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Denise Branch, Katy Aronoff, Evelyn Elias, and Emma Waecker, "The E-Book Story: The Key to a Happy Ending" (2018). *Proceedings of the Charleston Library Conference*.
<https://dx.doi.org/https://doi.org/10.5703/1288284317025>

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.

The E-Book Story: The Key to a Happy Ending

Denise Branch, Virginia Commonwealth University, dmbranch@vcu.edu

Katy Aronoff, Ex Libris, Katy.Aronoff@exlibrisgroup.com

Evelyn Elias, Taylor & Francis, Evelyn.Elias@taylorandfrancis.com

Emma Waecker, EBSCO, ewaecker@ebSCO.com

Introduction

Both advantages and challenges are associated with e-books. Advantages are easy to identify (24/7 access, no shelving required). However, the presenters had some difficulty deciding which other numerous challenges to address since many are intertwined. For example, the level of maintenance for e-books is dependent upon the quality of the bibliographic metadata. This paper will focus on three core challenges. The e-book story will be introduced from a librarian's perspective and three other stakeholders will each share their stories.

VCU Libraries' E-Books Story

VCU Libraries (VCUL) recognizes the importance of creating a learning environment where scholarly communication is easily disseminated and users can connect to information that is high quality and barrier free. E-books are essential in accomplishing this goal. E-books allow the Libraries to provide discovery and especially access that is not possible with print books. Users like the convenience, flexibility, and functionality of an online environment and expect this for books. Sitting in class, being stationed in the hospital, or away on vacation should not stop users from discovering or accessing books from their devices.

E-books have been a part of the VCUL collections since 2001 when the first file of 10,973 e-book records was received from netLibrary. Today there are over one million e-books. After the first big influx of e-books into the collection, the acquisition of e-books has become a regular collection practice. The largest number of e-books was first acquired through aggregators. E-books are now acquired through a variety of methods, including PDA, subscriptions, collections, e-preferred, and individually. As members of the Virtual Library of Virginia (VIVA) consortium, the VCUL have access to four major e-book collections (Brill, Taylor & Francis, University Scholarship Online, Wiley).

There are three major factors that catalyzed VCUL's purchase of e-books. They are space issues, support for distance education programs, and 24/7 discovery and access. Space issues plague the Libraries. Three million print volumes can be found on the shelves and print continues to be purchased since some books are not available electronically and some users still prefer print. There is overcrowding and not enough space for newly purchased books. E-books are an excellent choice for distance education programs. VCUL tries to support the curriculum needs of distance education students by acquiring electronic copies. Authorized users from remote locations are able to use their devices to locate thousands of e-book titles, save citations, link to other resources, and copy and paste, something they cannot do with print books. With today's technology, users can access an e-book any time, any place.

The popularity of e-books has grown and the VCUL has seen a high demand in the areas of STEM, business, and social sciences disciplines. As a library with over one million e-books, advantages can be easily identified. The 24/7 accessibility is a great advantage. No labeling, binding, barcoding, scanning, repairing, or shelving are required. E-books cannot be lost. Users can access and check out e-books from practically any device without intervention from library staff. Every word and phrase can be found by searching within text. Downloading of some portion or in some cases the entire book is available. Multiple search engines can be used to discover e-books.

Despite their many advantages, e-books are complex and untidy. They exist in a complicated ecosystem with many different stakeholders, including libraries, publishers, content providers, vendors, students, and faculty. Different business models, formats, and licenses are found in the ecosystem and each presents challenges. We will focus on three core challenges. We call them the three M's: (1) metadata, (2) management, and (3) models.

Metadata is one of the biggest challenges. Bibliographic metadata is necessary for discovery and holdings data is needed for access. They work hand-in-hand. However, this data can be inaccurate, incomplete, and lack content and quality, causing inconsistencies and irregularities in the knowledge base. When users perform searches in the library's online discovery and access systems that contain this poor-quality data, they are faced with significant barriers.

Management of e-books is time-consuming. Data can be received in multiple formats and require action to be taken by the library. Staff are needed to catalog, activate, and check links. Ongoing maintenance is needed as data and formats change. Troubleshooting may need to be performed, but the library is not always aware of this and might end up with users reporting a problem before the library is able to act on its own. Workflows can be established to help manage the process, but may change since so much is dependent upon evolving circumstances. Staff must stay alert to changes in the quickly evolving ecosystem. Manual updating is not sustainable for large e-book collections because there is not enough staff to successfully perform this task.

Different models are available and new ways of acquiring e-books constantly emerge. E-books can be acquired directly or from a publisher, aggregator, or vendor. Libraries can choose to select e-book titles through an approval plan, individually, in DDA or PDA plans, or through packages. Access rights vary depending on the provider and the license that is signed by the library. DRM restrictions can exist, along with a limit on the number of users. Standards for downloading, cutting and pasting, and discovery tools can differ among providers. Libraries may be confused about what e-books are available to them and if perpetual access is assured.

All of these challenges cause frustration for users and libraries. Sometimes frustrated users will seek out other resources such as Google to find what they need. Libraries might cancel e-books if usage is found to be low or nonexistent. These electronic resources are acquired to support scholarly communication, but what good are they if they cannot be discovered?

EBSCO's E-Books Story

EBSCO relies on publisher partners for the majority of our e-book metadata. Working with over 1,000 publisher partners worldwide enables us to sell and

provide access to over 1,000,000 e-books. Depending on how publishers send metadata, EBSCO may not receive data elements required for consistent representation in downstream systems, which impacts librarians managing ebooks and end users trying to use those e-books. Publishers have options when sending metadata; today, EBSCO's primary sources of publisher metadata are ONIX or a standardized Excel template. We strongly prefer ONIX because it enables transmission of rich metadata in a standardized and automated fashion. Among other benefits, ONIX's machine-created XML avoids subjectivity and human error; it's designed for international audiences, so it accounts for worldwide metadata variances; and it uses controlled vocabularies and over 200 code lists, which allow for a wide variety of data points about a book. EBSCO receives a good amount of ONIX, but 55% of publishers still only send Excel metadata, which only captures about 30 metadata elements, so it doesn't provide a deep representation of a book. EBSCO doesn't create metadata because we don't own the e-book content we sell, but we offset information gaps with a team of librarians who choose which metadata sources populate elements on our products. For example, we use the publisher's representation of the title as it's more natural and matches what's in the e-book itself, whereas the MARC record's representation of the title may be slightly different. Conversely, we prefer the MARC publication date as publisher-provided dates often reflect their workflow, not the book itself. For example, some publishers send PDF and EPUB publication dates that are different from each other, and which differ from the print publication date as well. EBSCO's normalization efforts help provide a better end user experience, but ultimately e-books are better represented when publishers send rich and consistent metadata up front.

E-book management is highly dependent on e-book metadata, as management is only possible when timely, accurate, and rich metadata is available where e-books are managed. EBSCO wants our customers to be able to manage their EBSCO e-books wherever they manage their other resources. The strength of knowledge bases underlying management systems is that they receive and aggregate data from so many sources. In most cases, knowledge bases already have the metadata EBSCO gets, plus more, so presuming an EBSCO e-book is matched in the knowledge base underlying a management tool, it should reflect the same quality of metadata as any other instance of that e-book. However, we hear from customers that EBSCO e-books often lack

complete metadata in management systems, and since this process happens outside our systems, we aren't positioned to know exactly why a disconnect has occurred. When records are missing things like descriptions and subject info, librarians will augment records themselves rather than risk lack of discovery. Since libraries manage a huge quantity of e-books across different systems, we want to consistently reflect the books in a library's EBSCO e-book collection, with associated rich metadata, to relieve that manual burden.

The third significant challenge with e-books is also one of its main strengths: the variety of models. Models are constantly evolving as new library and end user needs emerge. This is a benefit as it provides choices for libraries to meet their budget, collection development, and academic goals, but it introduces complexity in tracking and effectively managing e-books, and end users are confused when e-books work differently. EBSCO's end user research confirms that 74% of end users believe there should never be restrictions on library e-books, and most end users are confused by the terminology and limitations for e-books. Most will move on to other resources if they struggle with e-books. We don't expect these challenges to disappear on their own, but we strongly believe that by working together as a community, we'll get our happily ever after.

Ex Libris's E-Books Story

When working with e-books, libraries need robust and reliable software platforms that cover the entire e-book lifecycle, from purchase or licensing to delivery to the patron. As a software vendor catering to academic and research libraries, Ex Libris is invested in solving problems related to e-book management and access, particularly as libraries devote a larger percentage of collection budgets to electronic resources.

The main challenges for a software company supporting e-books align with the three areas of focus for this paper: metadata, management, and models. As neither a publisher nor an aggregator, Ex Libris must approach these issues from a different perspective than the other vendors represented in our paper (while Ex Libris is a ProQuest company and provides tight integration between the Alma platform and ProQuest platforms including OASIS and EBook Central, we will consider Ex Libris as a software provider separately for the purposes of this discussion). We

do not have direct control of the content itself, and Ex Libris's primary concerns in the e-book lifecycle are related to the library's ability to manage acquisition and access and to provide sophisticated interfaces for patrons to discover and use e-books across a range of providers and management models. These challenges overlap, and they can be difficult to truly separate from the Ex Libris perspective.

Ex Libris works with publishers and aggregators to gather and curate metadata for e-books and other electronic resources in our knowledge bases and indexes. Working with a large number of publishers in a variety of fields and industries means that achieving uniformity and meeting consistently high standards are major concerns not only on a day-to-day operational level but also as big-picture goals. Maintaining accurate and up-to-date knowledge bases is a challenge of both scale and precision.

Without high-quality metadata, e-books cannot be discovered by patrons, and they are harder for the library to manage. Metadata is also an element of tension in libraries today across formats as more and more records for print as well as electronic materials are supplied by vendors, frequently in large batches that can be loaded using automated tools. Batch data manipulation tools must be robust enough to provide specific and sophisticated transformation of metadata, but if records do not meet certain minimal standards for content, completeness, and quality, then all the data manipulation in the world will do no good without extensive enrichment, which can be prohibitively time-consuming for librarians, especially when it is performed manually.

On the discovery and access front, libraries need tools for troubleshooting, and they need to be certain that all titles that should be available are available and that they can be delivered in simple and predictable ways. Link resolution is a critical component of electronic resource management, as is the ability to provide direct linking where appropriate and desirable. Library staff should also be able to access and edit information about license data and agreements, embargoes, usage restrictions, and authentication requirements. Records must be published or transformed for discovery and decisions made about how resources should be displayed and made available in discovery. Which metadata elements should be indexed or displayed? Should certain providers be privileged or excluded in the delivery options displayed to patrons? How should buttons and links in the discovery interface

be designed, and what wording should be used? Our role as a vendor is to make sure that our systems are not only meeting libraries' needs but also continuing to evolve as these needs change.

Because Ex Libris is neither a publisher nor an aggregator, we do not have a role in designing or implementing acquisition or access models, but it is incumbent upon us to work with other vendors and libraries to make sure that our software can support models in use today and in the future. This is one of many reasons that Ex Libris strongly supports standards and is an active participant in discussions both formal and informal surrounding standards and projects such as KBART, the NISO Open Discovery Initiative, BIBFRAME, and more. Libraries should be able to provide access to e-books acquired or delivered under different models and from different vendors. The infrastructure and tools must be in place in the system to manage varying models without workarounds or unnecessarily complicated processes. Patrons should also be able to access e-books in consistent and straightforward ways without having to understand—as far as is reasonably possible—the differences between different models for acquisition or access.

Taylor & Francis's E-Books Story

At Taylor & Francis, we publish about 7,000 new books a year and are actively digitizing our backlist books as well. As of November 2018, we have approximately 138K DRM-free e-books on our own platform, and an additional 12K DRM-protected titles that are only sold on aggregator platforms. Many of these titles were signed and published by T&F's editorial staff, and some of these titles came to T&F when we acquired other imprints, such as Ashgate.

We as publishers face very similar challenges that librarians and vendors do. Our metadata is supplied to market and to our own platform through a variety of automated feeds and manipulated spreadsheets.

To be as customer-centric as we can, T&F offers many purchasing models for e-book sales. The most straightforward model is title-by-title purchase, which is very explicit as to what the customer should have and supplying MARC and other metadata is very clear. Other models, such as frontlist purchasing or evidence-based models, are less clear. For frontlist titles, we only truly know a book will be published once we have a manuscript in hand, about

4–6 months before publication. It takes time after publication to have a robust MARC record available, as the book must be cataloged properly after publication. These dynamic and ever-changing collections are reliant on our staff keeping all records up-to-date with the proper metadata.

As my co-authors mentioned, publication dates could also be tricky. The definition of publication date is the day the book, in that format, becomes available to consumers. For publishers, this means the day we digitized the book. We understand this has become problematic for libraries who expect that a 2018 title be published for the first time in that year. We have recently added a "First Published On" date to our systems, which allows us to deliver what librarians and users alike expect. This date reflects the first publication of the content in any format. This new field is one example where we have found ambiguity as booksellers. Different customers have different needs, and we must find a way to serve each market sector in the way that suits them.

In addition to the publication dates, we are challenged by how to send metadata to our partners for out-of-print content. In a print world, when a book goes out of print, it disappears from the market: it is removed from all feeds, inventory is pulped, and rights may be reverted. In an e-book world, we cannot ignore out-of-print titles. We must maintain those titles on our platform and continue to support the metadata needs of our customers. What happens when Library A bought a collection five years ago, and Library B bought a collection yesterday, and 20 titles in that collection have gone out-of-print in the meantime? This presents unique challenges, and we rely on market feedback to guide us to a solution.

Over time, we have also learned that we need to place strict controls over who can change fields that impact collections, and when they can be changed. We have recently locked down DRM status, subject collection assignments, and date fields so as not to have titles moving in and out of collections. The more automation we create, the more consistent our metadata is. Controlled vocabularies and standardized formats such as KBART and ONIX help us deliver that metadata to our partners, such as EBSCO and ProQuest, and to our customers. Publishers will continue to invest in systems and work with libraries and organizations such as OCLC and NISO to ensure that we develop and comply with industry standards that help us manage metadata and evaluate models.

Vision for the Future

Each stakeholder wants to support the success of their communities and institutions by addressing e-book challenges. Here are some suggestions for the future from the presenters.

- Providing standardized updates for both scheduled and unscheduled change data so that there is an audit trail of changes that can be easily tracked.
- Encouraging publishers to send ONIX metadata to ensure rich, consistent metadata is available across the e-book workflow.
- Building for interoperability so e-book customers and users are successful wherever they find an e-book. EBSCO recently began sending an industry-standard KBART file for use in Alma to improve representation and currency of their e-book subscription collections. They intend to continue increasing interoperability for EBSCO e-books.
- Collaborating on the models that best serve the library community. In early 2018, EBSCO heard the library community's call for an e-book model that provides end users with the most access with the fewest restrictions and released DRM-free EBSCO e-books. The Charlotte Initiative was instrumental in setting a goal for the industry to meet, and we see great opportunity for future evolution of models through similar collaboration.