2013

In the Beginning ... There was netLibrary

Barbara Anderson
bjanders@vcu.edu

Follow this and additional works at: http://scholarscompass.vcu.edu/libraries_present

Part of the Cataloging and Metadata Commons

Downloaded from http://scholarscompass.vcu.edu/libraries_present/36

This Presentation is brought to you for free and open access by the VCU Libraries at VCU Scholars Compass. It has been accepted for inclusion in VCU Libraries Faculty and Staff Presentations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.
In the beginning ...

... there was netLibrary

Barbara Anderson
Head, Metadata and Discovery

VCU Libraries
As I thought about my experience with ebooks, I realized that my story is not about how I have “managed” ebooks in my library – it’s more about how ebooks ... in a very short time ... have taken over my life.

(one of the alternate titles I considered for this talk)

I hope that the presenters who follow will have happier stories to tell. I am beginning to regain some optimism – now and then I have seen a few little glimmers of hope on my horizon.
You have already heard enough about me and what I’ve been doing since I landed in Richmond 31 years ago.

So, I’ll just start with a few obligatory words about VCU.
In the beginning... 175 years ago!

There was MCV (which began as the Medical Department of Hampden-Sydney College).

Almost 80 years later, the Richmond Professional Institute was formed (and it actually functioned as a branch of the College of William and Mary for part of its history).

45 years ago (1968) MCV and RPI merged to form VCU.
If you follow NCAA basketball at all, you might know that the VCU men’s basketball team has had several very successful seasons.

You might also know about their slogan:
Another title that I considered for this talk (Havoc Lives Here)
A few words about the VCU Libraries

2 main buildings – medical campus and Monroe Park campus

More than 2 million volumes

Collections budget approaching $10 million

Not currently an ARL library, but we have aspirations... and ARL membership is identified specifically in the university’s strategic plan.
All of the technical processing functions are centralized in the James Branch Cabell Library on the Monroe Park Campus.

Cabell Library is an extremely busy place

It is also extremely crowded.

In less than 5 months we expect to begin construction on a major new addition to the building.
All of which helps set the stage for the rest of the story that starts in the 2000/2001 academic year.

That was the year of:
- Wikipedia
- The Apple iPod
- Windows XP
- And while election boards in Florida were busy counting chads... VCU began purchasing e-books

Our crowded conditions were among the most obvious motivating factors:
eBooks require no shelving space
- They require no labeling, binding, repairing, shelving
- No need to scan or ship to distance ed students
- Available any time (Cabell Library is open 24 hours/day now – but weren’t in 2000/2001!)
  (ARL YouTube -- LibValue: Importance of Ebooks to Academic Libraries – April 2013) [http://www.youtube.com/watch?v=XQZQE4BFpql&feature=c4-overview&list=UU81MgM4v-scw34Q6zLWYCwg](http://www.youtube.com/watch?v=XQZQE4BFpql&feature=c4-overview&list=UU81MgM4v-scw34Q6zLWYCwg)

So ... for us..
In the beginning ...

... there was netLibrary
In 2000/2001 Lyrasis was still Solinet

And Solinet began offering a series of packaged netLibrary collections.

We started out in January 2001 with a file of almost 11,000 records

And we added to that at the rate of about 10,000 records per year – usually in batches of 2-3,000 records at a time.

I had never before had to deal with so many cataloging records at a time. For catalogers who had only ever cataloged things one at a time, this was a game changer!
To the Rescue!

• MARCMaker / MARCBreaker
• NOTIS GTO record loader
• Perl scripts
• “Good enough” cataloging.

How did we manage?

For us, and for most of you, I imagine, technology came to the rescue

• MARCMaker / MARCBreaker
  (no MARCEdit yet )
• NOTIS GTO record loader - allowed us to save records to diskette and use a specially configured PC to upload the files into NOTIS.
• Perl scripts
  (just in time for the ebook onslaught I had taken a day-long library introduction to perl for librarians presented by Ed Summers, who was working at Old Dominion University at the time and, I think, is now at Library of Congress)

Plus ... I underwent a critical attitude adjustment

As I learned to do bulk editing with MARCBreaker and perl I gradually signed onto the concept of “good enough” cataloging
netLibrary process = pretty straightforward.
• Pretty good records provided via an arrangement with OCLC
• Manipulated
• Loaded
• Counted
• OCLC holdings symbol set automatically
  (established a precedent for maintaining OCLC holdings for e-resource titles)
• An adventure in cataloging. And then ...

At first, working with netLibrary records was a novelty and every time we received a new file it was an adventure.
And then there were more
Proportion of ebooks to physical books a year after netLibrary
FY2013:

- Ebooks added: 44,989
- Print books added: 24,259
- Ebooks in collection: 679,826
- Print books in collection: 2.2 million

Proportion of ebooks to physical books today.
For us, it has been a long, slippery slope into Metadata MADness (another possible title for this talk)

Why do I say that?
No two vendors/ebook packages were alike.

- Good MARC records with OCLC record numbers (but no arrangement with OCLC to set holdings automatically unless you purchased the OCLC record sets)
- Pretty good MARC records without OCLC record numbers
- Unacceptable MARC records.
- No MARC records (but title lists/spreadsheets, etc.)
- Notification / No notification

If vendor records came without OCLC record numbers, at VCU, it almost didn’t matter how good the records were — because of our commitment to synchronize our catalog records with our holdings in OCLC we just used a perl script with the vendor records to generate search keys and did batch searching in OCLC. Then, we used a second perl script to clean up the OCLC records for loading into our catalog.
Process = variations on a theme:

1. Get the records (directly from the vendor or via OCLC search keys derived from title lists)
2. Attach holdings symbol to OCLC records
3. Manipulate the records
4. Load the records
5. Count the records.

In general, though, the process was the same.
• Get the records (directly from the vendor or via OCLC search keys derived from title lists)
• Attach holdings symbol to OCLC records.
• Manipulate the records
• Load the records
• Count the records
All of this is a lot of trouble, and some might ask – why bother manipulating the metadata? Why bother loading the records at all?

For us, it’s all about the Discovery

Remember the pie charts from a few slides back… The proportion of e- to print-books in the collection is growing.

That means that the serendipity of browsing the shelves is increasingly irrelevant.

We depend on our online discovery systems to include e-books in search results and browse lists.

At the same time, it’s important to be able to distinguish easily between the physical and the electronic resources.
At first, it was the commonly accepted and expected practice to have a different bib record for each incarnation of an e-book.

Definitely a discovery downside, but easily rationalized by catalogers (different vendor platform = different version)

And relatively easy to maintain catalog records.
  Get a file of netLibrary records and load it.
  Get a file of Proquest Safari records and load it. And so on.

But we could see the mess we were creating with the multiple records in our online systems.

Discovery downside:

  Duplication of titles among packages
    \rightarrow multiple records in WorldCat
      (allowed and commonly accepted ... at first)
    \rightarrow multiple records in local catalog
And then came another game changer: the vendor-neutral standards for cataloging ebooks

This made perfect sense from the discovery side of things

But it complicates things considerably for the titles that are included in more than one vendor package. (And there are lots of them!)
So, if you get your records from ebook vendors, you might notice that their records only include 856 fields for their platform.

If you’re working with a new set of titles for one vendor, and you run OCLC search keys to obtain records and/or to set your holdings symbol, you inevitably find that your holdings symbol is already set on many of the records (because of other packages).

You will also notice that OCLC records often have many 856 fields, including urls for packages with no local access. Somehow, if you are going to use these records, you have to strip out the 856 fields you don’t want, and keep the ones you do want.

At VCU we have relied on system loaders and increasingly complicated merge routines to identify duplicates and merge into existing records.

We have also worked on the occasional search and destroy mission to identify and consolidate the legacy non-vendor-neutral records.
So far, I’ve barely scratched the surface of what it takes to support ebooks in a library.

Talked mostly about cataloging e-books so far – I’ve mentioned my favorite bugaboos with ebook cataloging:

• Different methods for retrieving records/new title information from vendors
• Duplication of content among vendor packages
• Vendor-neutral cataloging rules for ebooks

But I haven’t mentioned the numerous record maintenance challenges we face

For example, when we learn that titles have been removed from a package

And it’s usually not just 5 or 10 titles – it’s usually more like hundreds of titles
Thanks to vendor-neutral rules – it’s not necessarily a case of finding and removing/suppressing groups of records

We need to be able to identify when it’s appropriate to remove everything and cancel OCLC holdings
And when we need to remove just one vendor link and keep everything else intact.
And what about the challenges when, for example, netLibrary becomes an EBSCOHost product and the URLs of thousands of records in your catalog need to change?

And it’s not just the ebook records and vendors that change – so do local library systems.

Libraries move from one system to another (NOTIS → Aleph → Alma)

And Ebook records move along, too, but the way they are managed within each system is very different, and the way they are presented for discovery is also very different.
I don’t even remember how ebook records appeared in the NOTIS OPAC...

Aleph OPAC view

Alma/Primo view

So, in addition to managing e-book records – we need to make sure that they are discoverable and we also need to make sure that they are decipherable to users in various contexts.
Along the way, we’ve had to make some decisions about what kind of metadata we wanted to maintain for our ebooks, and how to make that metadata work best for our users – external and internal.

There are a variety of possibilities for configuring 856 fields, but consistency is usually a good thing.

Display the URL or hide it from users?
Public notes? Vendor names? (in Aleph $z was the default for any public display; in Alma $z = public note; $3 = interface name)
Indicators: Does your ILS/discovery system rely on them to recognize what constitutes an electronic resource or an electronic version – and how to reflect in 856 field (I would argue that table of contents is not an electronic version but a related resource and should *not* be coded 856 41)

Call numbers? Why?
Are classification numbers enough?
Should call number info display in the OPAC?

Maintain OCLC holdings? Why?

So, what started out as a novelty and an adventure has become somewhat of a
nightmare.
Unless you’re using WorldCat Local (UM Baltimore County?) or have come up with better methods for managing ebook records in your catalog...

... chances are you’re living in an ebook Twilight Zone...

(another possible title for this talk)
And now, 12 years after embarking on our ebook journey at VCU, I am starting to think that we are entering the twilight of the Twilight Zone ...

... and the dawn of a new era in ebook management.
Why?
I’m intrigued and encouraged by the possibilities for managing ebook collections and records with OCLC’s WorldShare Collection Manager, which allows you to

• “claim” standard packages of ebook titles or set up custom packages of your own
• arrange for delivery of locally-customized MARC records (including which URL domains to be included in your 856 fields) and receive records for new, updated, and deleted records in the collections you have identified for yourself. (wouldn’t work so well in Alma environment, but would have been very helpful in Aleph)
• manage your holdings symbol in OCLC automatically.
Nearly a year after implementing the Ex Libris Alma system, I am increasingly excited about the potential of the Alma Community Zone.

Ex Libris maintains a centrally-managed repository of bibliographic records, external links and coverage information for lots and lots of vendor electronic resource packages.

It’s sort of an expanded version of the Ex Libris SFX knowledge base - with an increasing quantity and quality of ebook records.

If records are available in the CZ for our ebook package, we “activate” them (as a collection or on a title-by-title basis), and the URLs, availability, etc. are all managed centrally. Alma generates reports of titles that have been added/dropped from various packages, which is also very helpful.

And in the Alma CZ, multiple packages are often associated with single bib records.

Minimizes confusion for users.
Minimizes maintenance for catalogers.
Not working perfectly yet, but getting there.
And finally, a shout out to Sue Polanka at Wright State University who moderates the blog *No shelf required*.

Recent entry = 68 essential resources for eBooks in Libraries

Another contributor recently taught a 1-month ALA course on ebooks (as instructor) She taught the same course in March and posted the following comments from participants, which I will end on, because I think they do a nice job of reflecting my own evolving thoughts on the subject:

“I think [the world of ebook vendors] is a mess and getting more out of control.”

“My prediction for the future is to have a set-up for e-books that is almost universal. Just as we did for cataloging with OCLC and World Cat, ebooks could be accessed and dispensed on one major platform subscribed to by many, and not fragmented into many different platforms, publishers, aggregators, etc. etc. etc.”

“The ebook world continues to be a primordial swirl of chaos. However, the maelstrom of platforms and formats and DRM headaches is spiraling down toward an eventual semblance of calm.”