Publishers, We Love You But You're Bringing Us Down: The Failure of Vendor-Supplied MARC Records

Tina Herman Buck
University of Central Florida, tina.buck@ucf.edu

Sara Duff
University of Central Florida, sara.duff@ucf.edu

Kim Montgomery
University of Central Florida, kimberly.montgomery@ucf.edu

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

Part of the [Library and Information Science 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>.

Tina Herman Buck, Sara Duff, and Kim Montgomery, "Publishers, We Love You But You're Bringing Us Down: The Failure of Vendor-Supplied MARC Records" (2018). *Proceedings of the Charleston Library Conference*.

<http://dx.doi.org/https://doi.org/10.5703/1288284317030>

Publishers, We Love You But You're Bringing Us Down: The Failure of Vendor-Supplied MARC Records

*Tina Herman Buck, Electronic Resources Librarian, University of Central Florida,
tina.buck@ucf.edu*

*Sara Duff, Acquisitions & Collection Assessment Librarian, University of Central Florida,
sara.duff@ucf.edu*

Kim Montgomery, Electronic Resources Cataloger, University of Central Florida, kimberly.montgomery@ucf.edu

Abstract

Publishers, we love your DRM-free e-books, but your MARC record options and quality need work. You're wasting librarians' time and not providing access to the products you've sold us (or are trying to). Three librarians from the University of Central Florida have a list of concerns and some bright ideas to make it better.

Introduction

The University of Central Florida (UCF) is a very large research university with approximately 68,000 students. There is one main library and several subject-specific libraries.

UCF is part of the Florida Academic Library Services Cooperative (FALSC) with 11 other state universities and 28 state colleges. All libraries in the consortia share an integrated library system (ILS), currently Aleph, which is administered by FALSC. The universities in FALSC share one bibliographic database, the colleges share another, and the public catalog display is combined per state legislation.

"Shared Bib" is the colloquial name for the cataloging rules that all state university libraries are expected to abide by. The goal is to prevent negative consequences such as inadvertent overlays and deleting each other's records. In addition to cataloging rules, each library faces restrictions in their capabilities. One of the most impactful restrictions is that we are unable to batch-delete MARC records.

In addition to the "regular" university database, the universities share a separate patron-driven acquisitions (PDA) database. This houses records for PDA and evidence-based selection (EBS) projects that the universities undertake either individually, in groups, or as a consortium. Due to the heavy amount of batch loading and deleting required in a typical PDA/EBS project, it is considered best if only FALSC have the rights to make changes there.

Local Cataloging Realities

The UCF cataloging department includes one librarian (Kim) and one staff person who work full time with electronic resources cataloging. Several others are involved at a much lower percentage of their time.

UCF currently has three local EBS/PDA programs with different vendors and one shared program with other state universities. All four sets of records must be sent, separately, to FALSC. We have purchased still-growing collections from 19 vendors, including multiple collections from many of those vendors. In sum, we continue to get new records for about 55 products. Loading large files sets temporarily displaces work on smaller ones, due to time and system requirements. All of this constitutes too many record sets to be able to baby-sit any one.

Typical bibliographic maintenance activities include notifying FALSC to remove bibliographic records at the end of a PDA/EBS program; updating records when vendor platforms undergo URL changes; adding or removing individual titles from packages; negotiating with FALSC to create a load profile for a newly acquired package; reloading record sets when a vendor makes a change that requires this; and making the needed changes when EBS/PDA titles are purchased (such as setting OCLC holdings and moving the bib from the PDA database to the local database).

All of these factors make our cataloging situation very complicated, though we think every library faces cataloging complexities, regardless of size.

Our Goal

While the title of this presentation and proceeding is somewhat provocative, we want to be clear that we are not simply venting. We want to explain why the current state is problematic and help move toward effective solutions.

The areas that we want to surface for discussion are vendor MARC websites, record quality, process issues, and issues on the users' end.

Discussion Area #1: Vendor MARC Websites

The first area of difficulties that we want to bring up is with the vendor MARC websites.

One set of challenges involves naming conventions. Package names are often inconsistent across the invoice, user end, MARC site, and sales sites. Consistently being able to identify what we have purchased is basic. Sets may be sold by the year; please identify consistently whether that means publication date, electronic release date, or another date. Sometimes package contents change but the package name stays the same (or vice versa). Updated files need different names than the original.

The second challenge is grouping conventions. We need to be able to identify and differentiate records acquired by different licensing models. As mentioned, we need to be able to download, separately, record sets for multiple PDA/EBS projects and for each individual purchase. Each of these sets must be identified differently in our databases in order for the appropriate maintenance activities to be performed. Another grouping problem is when records are only available on the publisher's site one by one or in complete sets.

Functionality problems with MARC websites often involve searching limitations. We need to be able to search for specific records and groups of records, such as by ISBN, title, invoice number, package name, and year designation. Being able to readily identify whether records have already been downloaded or not is extremely useful, as is seeing when the records were made available.

MARC file naming conventions are another sore point. Suggested improvements include: put added or updated date in the file name; put collection or package name in the file name; group records by

license and/or acquisitions method; and add the word "corrected" to the end of the file name when applicable. Finally, it is desirable to allow generation of custom MARC files by collection, license, invoice number, downloaded/not downloaded, and date range.

Discussion Area #2: Record Quality

Our second discussion area is record quality. These elements can affect the public display as well as catalog functionality. Thus, while they may seem picky, the impact can be significant.

Library of Congress subject headings (LCSH) are needed in all records. They are the default. Subject headings from other vocabularies can supplement LCSH on records for works in specialized fields but should not replace LCSH. Non-LCSH headings need to be properly coded; bad coding can interfere with catalog functionality and usability, as well as with authority control projects. Subject headings also need to be specific enough to be useful.

A second quality issue is records that don't follow the latest cataloging guidelines, Resource Description and Access (RDA). Non-RDA records impact catalog functionality. For example, RDA defines a list of terms for playback device. For e-resources, the term is "computer." The correct coding is:

```
337 $a computer $b c $2 rdamedia
```

A typical incorrect example is:

```
337 $a electronic $2 isbdmedia
```

We had planned to use this field to generate e-book icons and a search limiter, but inconsistent data makes this difficult. The result of this bad coding is less functionality in the public catalog.

Sometimes, e-book records are actually print book records with a few "e" fields. The print records were used as a basis for the electronic version of the record—which is fine—but they were incompletely converted. Some important elements that don't necessarily get converted include coding the 008 digit for "Form of item" as "o" and adding e-resource-specific fields 336, 337, 338, 347. Finally, the URL must be coded properly to indicate whether the link represents the material that the record describes (i.e., an e-book) or whether it goes to descriptive material, such as a table of contents. Correct coding

of links impacts the catalogers' ability to find appropriate fields if proxy or URL changes need to be done globally. Full-text links would be proxied while links to descriptive material would not. In addition, linking text in the public catalog relies on correct MARC coding to generate the correct message to patrons: "click to access eBook" versus "read more about it."

Harvested data needs to be cleaned up before being distributed to libraries in MARC records. Formatting from source document impacts readability and searching. For instance,

```
100 1 $aSiemie&#x144;ska,  
Teresa,$eauthor
```

is obviously wrong and prevents the author's name from being searched. It should be:

```
100 1\ $aSiemieńska, Teresa,$eauthor
```

HTML tags will display in the catalog; for example:

```
<I>History of England</I>
```

We also find indentation characters that cause long gaps in paragraphs, and end-of-file characters that prevent following records from loading.

Tables of contents need to be right-sized. Too general: "505 0 Part I – Part II – Part III . . . – Part IV." Too detailed: "Cover; Half-title; Series; Title; Copyright; Dedication; Table of Contents; Acknowledgements; List of Figures; List of Maps; Glossary; Introduction; Part One; 1: Introduction; 1.1 Getting to Work; 1.2 Starting Out Right ; 1.3 Examining the Problem; 1.4 Industry Practice; 1.5 Government Regulation;" and so on, until the record runs out of space.

The final record quality issue is with 300 field issues, that is, description of the length or pagination. People want counts of pages, not counts of PDFs. Use the standard "1 online resource" rather than specific file formats, which change over time. Remove the "height" subfield for online resources (i.e., \$c 26cm). Remove the "additional materials" subfield for online resources (such as \$e 2 CD-ROMs).

Discussion Area #3: Process Issues

Another major source of pain is what we term "process issues." These include communication and processes we attempt to enact with the records.

The first issue is inconsistent, changing, and generic match points. In order to effectively load records, catalogers need 001 and 003 fields to include both letters and numerals, and each needs to be unique for the product or the record supplier. Many vendors are now including OCLC numbers in the 035 field, but that is not enough. The 035 field is a common overlay field, meaning that in a shared bib environment it could easily be overlaid inadvertently. It may also be helpful to allow the library to choose which field to put the OCLC number in, if OCLC isn't actually the source of the records.

Another process issue is notifications from vendors for MARC pickup or deletion. At the University of Central Florida, we have at least three librarians who need to be notified when MARC records are available. Because FALSC loads our PDA records, we also need a FALSC contact to be notified of PDA record availability. We are dealing with so many packages and programs that we need to be able to set up notifications properly at the point of purchase or licensing. Yet, often we are told that vendors can only send records to one e-mail address. This creates a considerable tracking problem for us, and it is easy for records to fall through the cracks.

We also need notifications when we are losing content, preferably at least one semester in advance of the loss date. We understand that rights are complex and that vendors may inadvertently sell materials that they don't actually have the right to. If libraries aren't notified by the vendor, then we only learn of the loss when a researcher contacts the library in a panic. At that point, it may be too late to obtain the material in another manner. This makes the library look very bad. It also puts pressure on the cataloger to drop their planned workflow to perform emergency maintenance.

Communication is key. We need a contact at each vendor who understands MARC/cataloging questions. The kinds of issues discussed here require very specialized knowledge, and the regular customer services representatives aren't expected to have this knowledge. When they are forced to relay questions to a cataloging specialist—hopefully there is one—and then translate the response back to the library, it becomes a game of telephone. Often, key details get lost in translation and the problem drags on or remains unresolved.

Vendors are constantly looking to improve their products, and as such frequently migrate content

to new platforms or devise new workflows for record delivery. However, the announcements are often written by the marketing department and are devoid of technical details. There must also be a readily available source for the technical details of what these moves will entail and how to prepare for them. Too often we are left to pepper our representative with questions they are not equipped to answer.

Discussion Area #4: On the Users' End

It obviously benefits everyone to have MARC records available as soon as possible after invoicing. With so many vendor partners, our catalog is the centralized place that our community uses for resource discovery; going from one vendor platform to another is too time-consuming.

Tying into the need for advance notification of lost rights, unexpected loss of access to materials is a significant area of confusion for library patrons. Faculty members will say, "I used that in class last week. What happened? My students need it." This makes the library look unreliable.

Likewise, duplicate bibliographic records confuse everyone. A recent example: Received in January: 1 record for *Sport practice, Volume I*; and 1 record for *Sport practice, Volume II*. Received in May: 1 record for *Sport practice* (all volumes). No notification was sent to remove records for volumes I and II, resulting in three records in the catalog.

Another area of confusion is when there is a mismatch between the title in the MARC record and the title on the platform. This can be particularly problematic with streaming videos. The streaming version of an older video (perhaps one originally issued on VHS) might have a series title, an episode title, and/or the title might have been edited over the years.

Please don't use the TV show title as a main title. As with the example of the program *Frontline*, there can be hundreds of episodes, all released separately.

Positives

After this long litany of issues, we have to mention several significant things that are going right. First, most vendors supply MARC at no cost. Second, most vendors understand why libraries want and need MARC records. Third, automation, when it works, is great—such as platforms that work well and automated record feeds that deliver. Finally, our presentation was given to a roomful of people interested in working on the topic.

Going Forward

How can we improve the situation? We have two suggestions each for libraries and vendors:

- Vendor: have a cataloging contact who can be consulted directly.
- Library: explain what's meant by "Technical Services." That's usually some combination of cataloging and acquisitions. It's *not* "Tech Support" or IT.
- Vendor: better management and presentation of which titles are in which packages as well as changes over time (such as loss of rights or changes to package contents).
- Library: designate one point-person who communicates with the vendor and with cataloging. Everyone in the library shouldn't send overlapping e-mails.

A robust discussion ensued, and several vendors indicated that they now have designated cataloging contacts who actually understand technical services issues.