Library Resources & Technical Services

ALCTS
Association for Library Collections & Technical Services

OJS Is the New Host of *LRTS*

http://journals.ala.org/lrts

Library Resources & Technical Services (*LRTS*), which went completely online in January 2015, is now available at http://journals.ala.org/lrts on the Open Journal Systems (OJS) platform hosted by the American Library Association (ALA) and is updated and maintained internally by in-house production staff. Content is no longer available on the Metapress site.

ACCESSING YOUR CONTENT

Every *LRTS* user (members and subscribers are both considered *LRTS* users) can login to journals.ala.org/lrts using the same credentials used for other ALA websites, including www.ala.org, ALA conference registration, and ALA Connect. Users will be notified in advance by ALCTS staff before the OJS login requirements take effect.

Be sure to visit http://journals.ala.org to enjoy other ALA digital journals and newsletters, including *Newsletter on Intellectual Freedom, Library Technology Reports, Reference & User Services Quarterly, Childen & Libraries*, and the *Smart Libraries Newsletter*.

For Technical Questions and Support

Please contact journals@ala.org.

Membership or Subscription Questions

To receive free access to *LRTS*, join ALA and ALCTS by visiting www.ala.org/membership/joinala or call ALA's Member and Customer Service (MACS) department at 1-800-545-2433 and press 5. Contact MACS with membership questions, too.

To subscribe to *LRTS* or to ask questions about your existing subscription, email subscriptions@ala.org or call ALA's MACS department at 1-800-545-2433 and press 5.

If you have any general questions about *LRTS*, please contact Brooke Morris (bmorris@ala.org) in the ALCTS Office.

Library Resources & Technical Services

Cover image: Reproduced with the permission of Marcia Brown. Working drawing for
Tamarindo! (1960) title page. Pen and ink, colored pencil (vermillion, ochre, cobalt blue,
charcoal). M.E. Grenander Department of Special Collections and Archives. University at
Albany, SUNY.

Editorial: Creativity and Innovation

Mary Beth Weber

One of my Rutgers colleagues recently published a book on managing creativity and fostering an innovative library culture. Creativity drives innovation in libraries, leading to novel approaches to challenging issues. Change can be daunting and working against prevailing practices and conventions can be difficult. Creativity and innovation are frequent themes in technical services work. Through collaboration and shared ideas, we as a profession can develop creative solutions.

Following on the theme of creativity and innovation, I am delighted to announce that M. Pilar Alonso Lifante and Francisco Javier Molero Madrid's paper "Enhancing OPAC Records: Evaluating and Fitting Within Cataloguing Standards a New Proposal of Description Parameters for Historical Astronomical Resources" by (*LRTS* volume 59, no. 4) is the 2016 recipient of the Edward Swanson Memorial Best of *LRTS* Award. A press release is available at www.ala.org/news/member-news/2016/02/swanson-best-lrts-award-alonso-lifante-and-molero-madrid. Alonso Lifante and Molero Madrid's paper discusses how enhancing content description of specialized resources, particularly astronomical resources, is an unresolved issue in library and information science. They provide a proposal of astronomical parameters for a better description and discuss how to accommodate these parameters using current criteria including MARC21, the ISBD and RDA. The authors take a creative approach to description to describe resources that fall beyond what typical cataloging operations handle to provide robust description that will help researchers locate astronomical resources, which typically are located outside library catalogs.

Questioning the status quo is another aspect of innovation and creativity. In "More Than a Name: A Content Analysis of Name Authority Records for Authors Who Self-Identify as Trans," Kelly J. Thompson discusses how FRAD and RDA broadened the scope of name authority records to provide a fuller description of creator. While intended to enhance discovery, these practices and the MARC 375 field for gender can be problematic for authors and creators who self-identify as trans. Thompson explores how the ability to control how information about how one's identity is shared with others is important, particularly when a person's social identity can result in discrimination or even violence. Her research includes an analysis of name authority records for authors who identify as trans that were collected during the course of her cataloging work. The analysis considers the type, breadth, and depth of information that is provided in these records.

Another aspect of innovation is customization. In "Customizing an Open Source Discovery Layer at East Carolina University Libraries: The Cataloger's Role in Developing a Replacement for a Traditional Online Catalog," Barber, Holden, and Mayo detail how they customized Blacklight to replace their traditional online catalog. Proprietary software was limiting, and they needed an interface that provided both the ability to customize and indexing flexibility to meet the different needs of their users, who are located on three separate campuses. Their paper includes a discussion of their initial work with one discovery layer, and the decision to switch to Blacklight as it provided a robust development

community. Their experience will benefit other institutions that are considering alternatives to the traditional library catalog.

Libraries increasingly face the challenge of tracking outages related to e-books and e-journals as the number of these resources in our collections continues to grow. Jennifer Wright details how her institution implemented the Foot-Prints ticketing system to track electronic resources outages and their causes. Implementing the software eliminated previous practices of using email and anecdotal information to track and resolve outages, which were time consuming and not always effective. "Electronic Outages: What Broke, Who Broke It, and How to Track It" identifies twelve different types of outages and provides problem resolution. Wright acknowledges that personnel at her library realized that changes to their workflows to better capture certain types of information related to electronic outages were necessary. Her library's experience and resolution of electronic outages

provide practical information for others who continually face similar challenges.

In addition to the papers I have highlighted, this issue also contains the following:

- "Survey of Classification and Organization of Video-recordings" by Barbara J. Bergman, Jessica Schomberg, and Dorie Kurtz;
- "Literature of Acquisitions Review, 2012–13" by Angela Dresselhaus;
- "Leveraging Author-Supplied Metadata, OAI-PMH, and XSLT to Catalog ETDs: A Case Study at a Large Research Library" by Ken Robinson, Jeff Edmunds, and Stephen C. Mattes; and
- book reviews commissioned by *LRTS* Book Review Editor Elyssa Gould.

More Than a Name

A Content Analysis of Name Authority Records for Authors Who Self-Identify as Trans

Kelly J. Thompson

With the adoption of FRAD and RDA, the scope of name authority records has broadened from a record supporting an authorized heading to a fuller description of a creator. Meant to help user discovery of resources, these practices are problematic when the record describes an author who self-identifies as trans. In this research, name authority records (NARs) for self-identified trans creators were analyzed. This analysis examined the 375 field for "gender," the contents of that field, and other representations of (trans)gender identities throughout the record. Name authority record creation practices should be examined to ensure that an author's agency to self-disclose their identities is respected.

Kelly J. Thompson (kellyt@iastate.edu) is Metadata Management and Cataloging Librarian in the Metadata and Cataloging Department at Iowa State University.

The original purpose of name authority records (NARs) was to give catalogers a framework to record the authorized form of the name of a work's creator to ensure consistency across all instances where the name was recorded in library catalogs.[1] The record can also contain information allowing a user (or a computer system) to cross-reference names creators may have used over the course of their careers under which other works of interest to the user may be found.[2] The IFLA Functional Requirements and Numbering of Authority Records (FRANAR) working group first acknowledged an expanded scope for name authority records by including that authorities should also support the user tasks of identifying (disambiguating) and contextualizing an entity.[3] Following this, with the introduction and adoption of the Functional Requirements for Authority Data (FRAD) and Resource Description and Access (RDA), the original scope of a name authority record was broadened from simply that of a carrier of an authorized heading or access point to a description of an entity with the development of an expanded list of attributes that can be included in name authority records.[4] Several fields have been added to the MARC name authority record format as part of this shift, including the optional 375 field for "gender." RDA 9.7 provides the content standard for descriptive cataloging with these new fields—the rules by which catalogers are guided to record information in these fields.[5]

On the surface, this may seem beneficial. Additional information in a name authority record could potentially help a user to find all of a creator's works regardless of the name the author used at time of publication or the name under which a user searched. Information included in a "source of information" 670 field could help a researcher to identify the cataloger's sources, such as an author's specific work, or to disambiguate similar name headings. Multiple referenced names in 400 fields could help a user to understand the author's name changes over time or redirect a search using an outdated form of the author's name. However, if NARs are considered from a transgender studies perspective, "an

emerging theoretical orientation on the nature of gender and gender identity in understanding the lived experiences of transgender and transsexual individuals," these practices can produce problematic results when the record describes an author who self-identifies as trans.[6]

Several librarians have critiqued RDA 9.7 specifically for its use of gender.[7] The author would like to extend their work to examine how NARs use gender throughout the record, both through use of the 375 field and in other practices that precede the introduction of FRAD and RDA. This research will investigate how (trans)gender identities are being represented using current NAR creation practices, specifically NARs created to describe authors who self-identify as trans. To conduct this investigation, the author analyzed the content of sixty NAR records currently available in the Library of Congress (LC) Name Authority File (NAF) that describe authors who self-identify as trans.

This research investigates whether including an expanded scope of information in a NAR is an inclusive practice, and if it serves the best interests of either authors or library users. Rich self-descriptive information about an author may be found in data stores outside of library catalogs, such as an author's personal homepage, and linked data could be employed to direct our users to this information, as controlled by the creators themselves. Although the authors whose name authority records are considered in this paper are all publicly "out," (which is how it was possible to search for and analyze their NARs), the question is related to *how* that information is shared, whether by the author or by a third party.

Background

The word *trans* is used in this paper to encompass the identities belonging to individuals whose gender identities do not correspond with society's assumptions about their gender identity based on the sex they were assigned at birth. This word is used as an umbrella term to encompass a wide array of transgender and nonbinary gender identities with which people may identify. The term is used here to communicate in a holistic way about people who, as Julia Serrano phrases it, "defy societal norms with regards to gender," and who may identify with any of the many labels for their gender identit(ies), including transgender and labels used for gender identities that fall outside of the assumed gender binary.[8]

Because many people in our society identify as *cisgender*, the prevailing societal norms of daily life do not always accommodate people who do not experience the social identities of sex and gender as congruent—people who identify with a trans identity.[9] This is sometimes referred to as *cisgender privilege*, or the specific rights and benefits that people who identify as cisgender receive, which are not

accessible to people who identify as transgender. Hill and Willoughby define *transphobia* as "the feeling of unease or even revulsion towards those who express nonnormative expressions of gender identity and expression."[10]

Because of *cisgender privilege* and *transphobia*, people who self-identify as trans are statistically significantly more likely than the general population to encounter discrimination and violence because of their gender identity or expression.[11] For individuals who self-identify as trans, queer, or with any other type of social identity that carries the possibility of discrimination or violence, the ability to control the disclosure of information about oneself is critical. When information related to an identity that may result in discrimination, harassment, or violence is shared, this is referred to as *outing*.[12] People can *out* themselves, as when people disclose their identity and *come out*. A person can also be *outed* by another person. When outing information is shared by someone else, it is unethical unless you are certain the person it is about has shared this information publicly with the intent of others knowing it; if the information was shared privately, the person needs to have explicitly given consent for others to share this information on their behalf for sharing that further to be ethical. Agentic disclosure, or a person's ability to control how information about identity is shared with others, is incredibly important.[13]

One of cataloging's core values is accurate representation of information to enable and improve access. Representations of creators' social identities can work against this goal of accurate and respectful representation, and it can lead to triggering experiences for people with marginalized identities. A poignant example can be drawn from the authority record for Ivan E. Coyote, one of the author's favorite storytellers. Coyote is publicly listed in the NAF (and thus in many library catalogs) as "Coyote, Ivan E. (Ivan E[. . .]), 19[. . .]-."[14] At the time of this writing, there is no other author by the name Ivan Coyote or Ivan E. Coyote in the NAF, yet two qualifiers are used in the heading: subfield $q for "fuller form of name" (Ivan E[. . .]) and subfield $d for "dates associated with a name" (typically the author's birth date) (19[. . .]-). The Program for Cooperative Cataloging (PCC) Task Group on the Creation and Function of Name Authorities in a Non-MARC Environment states that "authorized access points in RDA are created by combining a name with a defined set of informative qualifiers when available. There are other possible approaches to differentiating one entity name from another. . . . Authorized access points are preferably unique in RDA, but it is not required."[15] For this NAR, the level of differentiation used is not needed under any of the given guidelines. Coyote has published all of their works under the name Ivan E. Coyote, and has never published under their name assigned at birth. No other authors are listed in the authority file that could be mistaken for this author. Inclusion of this extra name information is not

respectful of the author's privacy, their self-identification, or bibliographic identity.

The author of this paper would like to cite Coyote themselves in explaining why this is problematic. In a January 2012 *Daily Xtra* column titled "'They' is me," Coyote addresses their perspective on others' inappropriate interest in their name:

> A couple of weeks ago I got an email from a young woman, a college student, who claimed that her professor had assigned her entire class a special little assignment, for extra credits, for students who could track down my legal name and bring it to class. This young woman had tried and tried, she said, to find it online, but couldn't, and she really wanted those extra marks. Would I be so kind as to just tell her?
>
> I took a deep breath. I was flabbergasted, skin crawling with chill fingers at how totally creepy this felt, an entire college English or writing or queer studies or whatever class assigned the task of violating my privacy for extra credit at school.
>
> Exactly what educational or literary purpose could my legal name serve, anyway? . . .
>
> This leaves me frustrated, and feeling violated, and worried that this will happen all over again this semester, that I will get a new crop of letters from eager students asking me about my legal name.
>
> For those people who use their legal name and have never had any dissonance in their head or life with the name given to them at birth versus the name that feels like their name, well, I am glad for you, I really am. It must be fantastic to have all your ID match your face and your gender and your tits and your birth certificate and what the border guard sees when he looks at you and decides whether or not to let you on the plane. You are lucky.
>
> I am not one of those people. I don't like my legal name, first or last; it doesn't suit me, it never fit . . .
>
> I changed my name in 1990. That was 22 years ago. Some of my family still call me by my birth name, and I let them do this only because they are my family. I cash cheques and do business and perform and publish and live and fuck and talk to my neighbours as Ivan. Because this is my name. It feels good to be called who I am. . . .
>
> Call us what we wish to be called.[16]

Coyote's words express the core issue at stake here: a person's agency to find a way of being in the world that affirms their identity and sense of self. It is essential that a service, such as cataloging, which is meant to help people to discover resources that will help them in these self-identification processes, consider this and reflect these identities, and that library systems are designed with a strong effort to reflect people's self-identities with profound respect. We have an opportunity to accurately represent authors as they choose to be known. For librarians invested in working toward gender equity, observing these kinds of misrepresentations within an information system that many trained catalogers do not have permission to edit yields an uncomfortable situation. This research considers how our practices could change to accommodate all authors.

Literature Review

New Authority for Name Authorities in RDA

Agenjo et al. wrote, "[The MARC 21] model has shifted from authority records for *personal names* to the records for *persons*, which is not a nominal but a long-range issue. In fact, some of the cataloging problems to be solved in the near future will be the combination of persons' attributes with *personal names'* attributes."[17] This shift has been influenced by the Functional Requirements for Authority Data (FRAD), which states that the function of authority control for names should support "not only the task of identification of the entity 'person,' but also the tasks of contextualization."[18] This functional requirement is the impetus for including new fields in the MARC 21 authority record format, such as the 375 field for "gender." The inclusion of a gender field is meant to support the shifting purpose of the name authority record from a documented list of authorized name headings to be used in records, to a record of characteristics of the individual authors themselves. This practice represents a fundamental shift in the theoretical framework. The Descriptive Cataloging Manual Z1 states that "an NAR does not serve as a biographical sketch of a person, nor as an account of the detailed history of a corporate body."[19] Yet RDA chapter 9.17 now provides specifically for the provision of "Biographical Information" in an NAR.[20]

The MARC 375 field was not adopted without controversy. The American Library Association (ALA) Gay, Lesbian, Bisexual, and Transgender Round Table (GLBTRT) formed an ad hoc task force, the ALA GLBTRT Task Force on RDA and Gender in Authorities, which provided a formal "comment on RDA's proposed coding of gender in authority records."[21] Their comment, issued in February 2008, included the following statements:

> The current instruction for proposed RDA rule 9.8.0.3.1,

"Record the gender with which a person iden-
tifies using an appropriate term from the list below.
    female
    male
    other
    not known"
does not acknowledge the fluidity and vari-
ety of possible gender identity or identities of an
individual over time. It also does not address cod-
ing of the variety of gender categories related to
bibliographic identities, the individuals who create
them, and the relationships among them. (How, for
example, would we code George Eliot, a woman
writing with a male name, or Barbara Michael, a
husband and wife writing novels together under
one name?) Furthermore, the limited number of
possible values, and the language used for those
values is offensive to many people.

This Task Force recommends that RDA not
prescribe any coding for gender, and that further
study be made of the issues of gender in the con-
text of resource identification and relationships
between entities.[22]

Discussion of this statement ensued on the RDA dis-
cussion list in February 2008.[23] The discussion indicated
that some voices favored the potential for this information
to advance a user's searching ability, and some argued that
we could not make exceptions for non-conforming authors
when the problems noted did not apply to the "majority" of
authors.[24] Others suggested that it was an important role of
the cataloger to respect an author's self-identity, and that this
field did not serve that purpose.[25] Potential use cases for the
information were presented, as were problematic aspects of
the field's use, such as a limited Library of Congress Subject
Headings (LCSH) vocabulary.[26]

The ALA GLBTRT Task Force's recommendation was
not accepted, and the rule was codified in RDA 9.7.1.3. As
RDA was implemented, this rule became part of current
cataloging practice, but not without continued concerns
about the usefulness of the MARC 375 field and the RDA
9.7 instructions. Billey, Drabinski, and Roberto identify how
the 375 field for "gender," while intending to provide qual-
ity bibliographic access, can be problematic when describ-
ing something as personal and fluid as gender.[27] Billey et
al. describe how the practices recommended by RDA 9.7
(which prescribes usage for the 375 or "gender" field) lead
to descriptive work that "reinforces regressive conceptions
of gender identity" and "denies the shifting and contextual
nature of gender identities."[28] Furthermore, they explain
how controlled vocabularies limited to the *gender binary* do
not leave room for the identities and experiences of nonbi-
nary gendered transfolk.[29]

Putting RDA 9.7 into Practice

The label for the MARC 375 field ("gender") conflates gender
categories and sex categories; the suggested entries of male
or female, and nonbinary categories omitted by the suggested
terms list such as intersex, are actually *sex* categories, where-
as examples of *gender* categories might be genderqueer,
agender, transgender woman, trans man, woman, man, girl,
boy, boi, etc. LC training materials regarding completion of
the 375 field also conflate sex and gender categories.[30] This
could be because of people's reluctance to use the term "sex
categories," while "gender" as a term feels much safer.

The RDA 9.7.1.3 guidelines instruct one to "record the
gender of the person using an appropriate term from the
following list: female, male, [or] not known. If none of the
terms listed is appropriate or sufficiently specific, record an
appropriate term or phrase. Example: intersex, transsexual
woman. . . . Indicate the source of information."[31] The MARC
21 documentation recommends providing a source code for
the term used in 375 $a, to be listed in 375 $2.[32] The docu-
mentation currently provides just two codes: "lcsh," referring
to LCSH, and "iso5218," which refers to the International
Standard Organization's standard Codes for the Represen-
tation of Human Sexes (ISO/IEC 5218:2004).[33] The values
provided in ISO Standard 5218 are "Not known 0 (zero),
Male 1 (one), Female 2 (two), Not applicable 9 (nine)."[34] The
documentation for this ISO standard includes the remark
"No significance is to be placed upon the fact that "Male"
is coded "1" and "Female" is coded "2." This standard was
developed on the basis of predominant practices of the
countries involved and does not convey any meaning of
importance, ranking or any other basis that could imply dis-
crimination."[35] Regardless, this standard ignores people with
trans identities and those whose sex assigned at birth may
be intersex. This standard is unable to represent the sex or
gender identities of a person who identifies with a nonbinary
identity, and also creates a hierarchy that privileges maleness
over femaleness. Even if catalogers need to work beyond the
binary of "female" and "male" to describe an author, often
appropriate labels do not exist in the controlled vocabulary
we are encouraged to use.[36]

The ALA *Library Bill of Rights* affirms that libraries
are to provide access to resources without regard to "origin,
background, or views of those contributing to their cre-
ation" and that "a person's right to use a library should not
be denied or abridged" because of these characteristics.[37]
The interpretation of this bill of rights, adopted by the ALA
Council, further explains that "Article V of the Library Bill
of Rights mandates that library services, materials, and
programs be available to all members of the community the
library serves, without regard to sex, gender identity, gender
expression, or sexual orientation."[38] Given that authors are
members of the communities that cataloging is meant to

serve, it follows that metadata and cataloging professionals should decline to adopt standards that reinforce hierarchies of privilege, and exclude already-marginalized groups in our communities. While the usefulness of controlled vocabularies and consistent metadata is invaluable to access and discovery services, it could be argued that describing identities would be better left out of this work.

Critically Engaging the Catalog

The author previously alluded to the inability of the suggested values and practices related to NARs to accommodate trans identities and experiences. Engaging the vocabulary and "spaciousness" of the catalog through a critical lens is not new. Following in the tradition of the likes of Olsen and Berman, many authors have sought to problematize vocabulary used and improve LCSH's language related to (trans) gender identities.[39] Most notably missing from LCSH are the terms *queer* and *genderqueer*. Drabinski has carried this work further by using queer theory to interrogate not only the library catalog and LCSH. While these vocabularies do change over time, Drabinski asks us to question the possibility of ever creating a truly representational and inclusive controlled vocabulary or catalog, when such an artifact is built by a relatively small number of people at a specific point in time.[40] A few people hold privilege in the creation of an authority file and catalogs: the privilege to represent many other people throughout time and place. Drabinski posits that we educate students in the work of interrogating the catalog as a cultural artifact through a queer theory lens.[41]

Bibliographical Cultural Difference

Exner's concept of *bibliographical cultural difference* furthers our understanding of why labeling authors without their input is problematic.[42] Exner sought to show how North American Indian people's names were represented by those whom he described as experiencing this bibliographical cultural difference. When those who are describing a work or person cannot access the necessary knowledge or background to make an authentic representation, it can lead to inaccuracies, misrepresentations, and silencing of identities.[43] Bibliographically it makes sense in some instances to include a creator's *sequence* of names used throughout their life in an authority record (such as if the author has published works under each name). However, especially for people who self-identify as trans, this sometimes is superfluous information that encroaches on an author's right to privacy. Because of the principle of *bibliographic cultural difference*, a cataloger may not even realize that adding a person's previous name to an authority record can be a faux pas depending on the author's personal navigation of identity.[44]

With the addition of descriptive information to NARs following FRAD and RDA, information about an author is being included in authority records in ways that were not recommended by previous cataloging rules. This information is then publicly available to those who search for name authority records (such as through LC's website), and takes on increasing visibility with the implementation of linked data and other semantic web technologies (changes for which RDA and data models such as BIBFRAME are meant to prepare catalog records.) Including a "gender" or "sex" category in a NAR is a form of *outing*. As professionals, our intention is provide information that will enable users to locate authors and to disambiguate them from others. However, there are cases when providing this information is harmful to an author and exposes information that is not bibliographically relevant.

Name Changes, Sets, and Sequences

Writers who experience gender beyond the binary are not the only people in the bibliographic world who have either a series or set of names, as illustrated in the following works. Frank Exner, Little Bear outlines clear language to use when discussing the inclusion of multiple names for an individual in a record, or what Exner deems name *sets* and name *sequences*.[45] According to Exner, a name *sequence* is a set of two or more "changing . . . names that tell an autobiographical story," or names used sequentially by a person.[46] For example, Chaz Bono has publicly shared that while his name is now Chaz Bono, his name assigned at birth was Chastity Bono (included in a 400 field in his NAR), which he no longer uses. These names have been used *sequentially* throughout Bono's life, one replacing the other. In contrast, a name *set* indicates that two or more names are or were used at the same time by a person.[47] For example, Matt Kailey's name authority record includes a 400 field (or "see from") for Matthew Kailey. Kailey used both names simultaneously and interchangeably, therefore the names constitute a name *set*.

Pellack and Kappmeyer identified several ways that name changes can create search problems for library users. Their research investigates the effect on the indexing of scholarly articles after women authors have changed their names for a variety of reasons.[48] Their technique of investigating how names are recorded in various databases, indexes, and authority files inspired the author's method of searching for name records in the national authority files.

Research Questions

This research was guided by three main questions:

1. Is the newly introduced and optional 375 "gender" field being used in NARs cataloged with RDA for

people who publicly self-identify as trans?

2. If the 375 field is present in a NAR, what content values are being used?

3. In what ways beyond the use of the 375 field has gender identity or other outing information been represented in NARs (both on records cataloged using RDA and those using AACR2)?

Method

The author began collecting sample NARs after encountering NARS for authors/creators who self-identify as trans in the course of regular cataloging work after observing the type, breadth, and depth of information included in these records. Sixty PCC NARs for individuals who are self-identified trans authors, writers, creators, and/or performers (referred to here collectively as "authors") were analyzed. To collect this set of records, the author generated a list of personal names of authors who publicly self-identify as trans to search for in the LC NAF. The names on the list were drawn from background subject knowledge, compiled from the ALA GLBT Round Table's reading lists, Goodreads, blogs, Amazon, Wikipedia's "Transgender and transsexual writers" page, and Google searches for authors' websites.[49] The author was able to generate a list of only one hundred names to search in the LC NAF. Twenty-nine personal names on the search list of one hundred were not represented by a NAR in the LC NAF so the set size was reduced to seventy-one. Eleven NARs were eliminated from the test set of seventy-one for creators who were born before 1900 or for whom a self-identity of transgender or genderqueer could not be readily confirmed, further reducing the studied record set to sixty NARs. Records for creators born before 1900 were eliminated because of the inconsistency in terminology used before the current period and to support one of the core motivations of this research—to encourage respect of the privacy and safety of living authors. Reviewers of this research have suggested that the "gender" attribute is a valuable one for researchers seeking a specific set of authors (for example, nineteenth-century women writing about civil rights and the suffragette movement), however, the author of this work feels that this potential use case falls outside of the concerns of this paper regarding privacy of living authors with marginalized gender identities currently producing works.

While sixty records may seem like a small sample size out of the entirety of the LC NAF, it was the largest number of records that could be obtained using this method. This could be the result of many factors. It could be because out trans folk continue to experience systemic oppressions that could be prevent them from becoming published authors with NARs. There could be authors in the LC NAF who identify as trans that we do not know are trans because,

regardless of profession, many people choose not to out themselves publicly as trans, for a variety of reasons, including because it is private, sensitive information and/or because it is inconsequential to their bibliographic identity and works. All sixty records were accessed between September 19, 2013, and March 13, 2014, using the OCLC Connexion Client cataloging software. After accessing these records, the author compiled selected record elements into a spreadsheet.

Table 1 provides a summary of the elements that were examined related to the research questions. The author first recorded the value in the MARC fixed field 008, position 10, for "Descriptive Cataloging Rules" used in creation of the record. To investigate research questions 1 and 2, the author examined whether the 375 field was being used in practice, and if so, what content values were used. In this analysis, the author noted whether the 375 field was used in each NAR. When records included at least one 375 field, the author noted whether there were multiple 375 fields recorded. It was noted whether the entries included "start" and "end" dates. In the content analysis, attention was paid to whether the information included in the records was *binarist*, or operated under the assumption that there are only two human genders. Notes were recorded for terminology used and whether it was represented in LCSH. The author also recorded whether these 375 fields could reasonably be considered as outing the author as trans.

The author next examined elements related to research question 3, concerning ways that (trans)gender identity is represented in elements other than the 375 in NARs. These elements are also summarized in table 1.

For records that provide multiple names for a person (such as alternate name tracings from 400 fields), the author coded the data to indicate whether these multiple names were name *sequences* or name *sets*, determined by consulting the 100, 400, and 500 fields, the bibliographic citations given in the 670 fields, and external resources such as an author's personal webpage.

When a name set or sequence was present, the author coded these instances in one of four ways:

1. alt: When the "see from" (400 field) referred to an alternate form of the same name, the author used alt (example: "Link, Aaron Raz" and "Raz Link, Aaron").

2. yes: If the names given were not alternate forms of the same name, the author coded for whether the inclusion of multiple names was bibliographically significant. For example, if the name was listed as an "earlier form of heading," or if the writer published multiple works and has used each of the names on a work, the inclusion of the "see also" (500 field) reference is bibliographically significant.

3. no: If the writer has never used one of the names given in the record for a published work, and the

Table 1. NAR Elements Examined in Relation to Research Questions 1–3, Concerning How (Trans)gender Identity is Represented in the MARC 375 Field as well as in Elements Other Than the 375 Field

Data label	MARC field source	Example data
"Descriptive Cataloging Rules"	fixed field 008 position 10	c (AACR2) z (Other) [RDA]
"gender"	375 (repeatable field, new in RDA) – uses controlled vocabulary terms/codes for a small number of sex categories, not gender identities, can include "start" and "end" dates	female $s 1969 $t 2008? male $s 2008? female Female-to-male transsexuals transsexual woman
does 375 include "start" and "end" dates?	375 $s 375 $t	female $s 1969 $t 2008? male $s 2008?
"authorized name"	100	Bornstein, Kate, $d 1948-
"see from" tracings for alternate names	400 (repeatable), 500 "see also from" tracings	Kailey, Matthew Bono, Chastity
"field of activity"	372 (new in RDA)	Gender identity
"associated group"	373 (new in RDA)	University of Sydney $a Macquarie University
"occupation"	374 (new in RDA)	Authors $2 lcsh
"source data found"	670 (repeatable) – generally cites the sources where information in the record was found	Serano, Julia. Whipping girl a transsexual woman on sexism and the scapegoating of femininity, 2007: b ECIP t.p. (Julia Serano) about the author (Julia Serano is an Oakland-based writer, spoken word performer, trans activist, and biologist; she has a PhD in Biochemistry from Columbia University and is currently a researcher at UC Berkeley in the field of Evolutionary and Developmental Biology)
use of gendered language, pronouns, gender identity labels, etc. in the record	most often in 670 field(s), but found throughout the record	Her His
any information given in the record that "justified" giving a concrete date of transition or an added 375 field through use of medicalization	most often in 670 field(s), but found throughout the record	some records cited "sex-change surgery" and/or information about a "legal name change"
date of record access	date researcher accessed record in OCLC	2/7/2014

name has never been an "earlier form of heading," it is not bibliographically significant.

4. pen: If the name in the authorized heading is a pen name, and the see also reference (500 field) is the name the person uses in daily life, the author coded the name set as "pen."

When 372, 373, and/or 374 fields were present in the record, the author recorded whether the headings could be interpreted as outing, or as placing a strong emphasis on an author's sexuality. To judge this, an examination was conducted as to whether this information would give the reader information about an author that identified them as trans or as someone with a sexual orientation other than heterosexual.

When a record included 670 field(s), the author coded whether that field contained language about gender or experiences of gender, including those which explicitly outed the writer as trans. It was recorded whether the cited source of this information was self-disclosed by the author (for example, some records cite the author's work, personal website, or an email with the author) or from an external source (such as a publisher or Wikipedia).

Each record was coded for gendered language and pronouns, including inconsistencies in name use or incidents of mispronouning (using incorrect personal pronouns of reference to refer to an author).

Some 670 fields cite both the source of the information included in the record and a quotation from that source.

The author specifically coded where the quotations in a 670 field provided information meant to support the information recorded into subfields $s and $t of the 375, meant to represent the "start" and "end" dates of the time span the creator identified with each gender. When the quotation included medical information or legal name change information, the record was coded with "medicalized" or "name change" respectively. Records that included both types of information were coded as "multiple types," and records with at least one 375 field with a subfield $s or $t but did not "justify" this information in a quotation were coded as "none." Lastly, the author recorded whether it could be perceived that the record as a whole outed the described author as trans via all of the compiled information presented.

Results

In total, thirty-nine of the sixty records analyzed contained some form of outing information in an area of the record (65 percent). Seven of these were RDA records and thirty-two were AACR2 records. This was manifested through use of multiple 375 fields, inclusion of terms in the 375 field that indicated a trans identity, and through five additional mechanisms that did not require use of the 375 field.

Inclusion of the 375 Field in RDA records (Q1)

Ten of the sixty records in the test set were coded "z," indicating that they had been created using RDA guidelines. Seven of these records used at least one MARC 375, or "gender," field. The remaining fifty records were coded "c," indicating use of AACR2 rules. Two of the fifty AACR2 records had been updated to include some of the new MARC NAR fields; specifically, they each had two 375 fields added to the record.

Content Values of the 375 fields (Q2)

Of the nine records containing at least one 375 field, three records provided a single 375 field. The other six records included more than one 375 field, which means that multiple labels for sex or gender identities were given. Among the nine records with 375 fields, there were seventeen total 375 fields (an average of 1.9 375 fields per record with 375 field(s)). Three records contained one 375 field, four records contained two 375 fields, and two records contained three 375 fields.

A 375 field may include binary "start" and "end" dates using the subfields $s and $t, per the MARC format guidelines.[50] The $s and $t subfields were used in five of the records with 375 fields (>50 percent of records with 375 fields). Of the six records with multiple 375 fields, only one

lacked "start" and "end" dates. Three of the records with $s and $t fields were cataloged using RDA and two of the records were cataloged using AACR2. These results are summarized in table 2.

Nine of the NAR records contained at least one MARC 375 field, and six of these records used only "female" or "male" as categorical terms. Three records used terms that indicated a trans identity. Two of these used three successive 375 fields to indicate that a gender identity shift had occurred by using "female" and "male" labels and "start" and "end" dates plus an LCSH term to indicate trans status. "Female-to-male transsexuals" and "male-to-female transsexuals" were the specific terms used in these two records. One record contained a single 375 field with the term "transsexual woman."

In total, from nine records with at least one 375 field, seven contained information in the 375 field(s) that could be perceived to be outing (78 percent of records with at least one 375 field). Five of these were RDA records (71 percent of RDA records with 375 fields), and two were AACR2 records (100 percent of AACR2 records with 375 fields). A summary of the use and content values of the 375 fields can be found in table 2.

Other Representations of Gender and Instances of Outing in the NAR (Q3)

There are five mechanisms through which gender and/or outing information have been represented in NARs beyond use of the 375 field: (1) inclusion of superfluous qualifiers in the authorized heading, (2) alternate names given in 400 fields or "see from" references, (3) use of gender or sexuality terms in the 37X fields (372, 373, and 374), (4) use of pronouns in the 670 or "source of information" field, and (5) disclosure of personal information unrelated to bibliographic access in the 670 field.

Unnecessary Qualifiers in the Authorized Heading (100 field)

As noted in the introduction, one record in the set (<2 percent) included a qualifier in the name heading that is not part of the author's name as currently used for publication and is not necessary to disambiguate the author from another with the same or a similar name.

Name Sets and Name Sequences (400 fields)

Of the sixty records in the test set, thirty-four records (57 percent) provided multiple names for an author. These are recorded in one or more 400 fields, which are used as "see from tracings." These fields are used to automatically refer a searcher to the authorized form of an author's name, which

is to be used on all records for their work, regardless of the name used when any given work was published.

Of the thirty-four records with multiple names given, twenty-four had name sequences (40 percent of the sixty record set, 71 percent of those with multiple names) and ten had name sets (17 percent of the sixty record set, 29 percent of those with multiple names). The author identified four general types of alternate names listed: pen names, alternate forms of the same name, an alternate name that is bibliographically significant, and inclusion of a sequence of names used by a person during various points in time that is not bibliographically significant. The frequency of these types of name sets and sequences is given in table 3 (some name sets/sequences fall under multiple of these four categories). This last type is the form identified in this work as particularly problematic: name sequences or sets that are not bibliographically significant. There were ten records with this type of multiple name listing (17 percent of the 60 record set, 29 percent of records with multiple names).

Table 2. Summary of the Use and Content Values of 375 Fields

No. of 375 Fields in Record	No. of Records with This Many 375 Fields in Sample Set	Content Values of 375 Field(s)	Cataloging Rules Used
1	3	male	RDA
		female	RDA
		*transsexual woman	RDA
2	4	*ᵇfemale $s 1969 $t 2008? male $s 2008?	AACR2
		*ᵇfemale male	RDA
		*ᵇfemale $s 1972? male $s 1926 $t 1972?	RDA
		*ᵇmale $s 1946 $t 1998 female $s 1998	AACR2
3	2	*ᵇfemale male $s 199u Female-to-male transsexuals	RDA
		*ᵇmale $s 1926 $t 1950? female $s 1950? $t 1989 male-to-female transsexuals	RDA
Total Count	9 records	17 fields	7 RDA 2 AACR2

* indicates information perceived to be outing
ᵇ indicates information perceived to be reinforcing the gender binary

37X Fields (372, 373, 374)

All MARC 37X fields are optional and repeatable. Eleven of the records analyzed included at least one 37X (including 375 fields). While these fields are optional, they have been used with a fair amount of frequency. Nine out of ten RDA records in the test set included at least one 37X field (90 percent of RDA records.) Seven records used the 372, 373, and/or 374 fields. All of these records were cataloged with RDA (70 percent of RDA records.)

Beyond the 375 or "gender" field, the author observed three other new MARC authority fields present in the NARs in the sample set used to present outing or sexualized information. These are the MARC 372 for "field of activity," 373 for "associated group," and 374 for "occupation." These fields were sometimes used to share information that could be read as outing a person's sexual orientation or gender identity, or shared information that was related to an author's sexuality in a way that did not seem relevant to the work being presented. While the outing information shared in the sample set primarily regarded sexual orientation, not gender identity as the remainder of this paper focuses on, given that these fields have been used to share sexual orientation identities, the possibility exists for them to also be used to represent gender. These additional 37X fields were each found on seven records in the sample set. All seven were cataloged using RDA. Table 4 summarizes the author's observations of these fields.

Four records contained optional 372 fields for "field of activity." Of these, two fields had information that could be seen as outing (one sexual orientation and one gender identity) and one field recorded information that focused on the author's sexuality in a way that did not hold bibliographic significance. Two records had optional 373 fields for "associated group," and one provided information that was outing (sexual orientation). Six records had optional 374 fields for "occupation," and one provided information that was outing (sexual orientation).

Pronouns and Disclosure of Personal Information Unrelated to Bibliographic Access in the 670 Field(s) or "Sources Found"

The 670 field, or "sources found," is used to cite the sources of the information recorded in the NAR. Often the information included in this repeatable field is used to "justify" the

Table 3. Records with Multiple Names Listed for an Author

	No. of Records	% of Records (*n* = 60)	% of Records with Multiple Names (*n* = 34)
Multiple names given	34	57	100
Sequence or set?			
sequences	24	40	71
sets	10	17	29
Type of multiple names			
not pen or alt but bibliographically significant	16	27	47
not bibliographically significant*	10*	17*	29*
alternative form of one name	7	12	21
pen name	1	2	3

* indicates information shared perceived to be problematic in practice.

Table 4. Use of Optional 37X fields in NARs

		No. of Records	No. of Fields
37X fields (372, 373, 374, and/or 375)		11	35
Non-375 fields (372, 373, and/or 374)		7	18
372 ("field of activity")		5	6
Contained:	outing information (sexual orientation)	1	1
	outing information (gender identity)	1	1
	reference to sexuality	1	2
373 ("associated group")		3	6
Contained:	outing information (sexual orientation)	1	1
374 ("occupation")		6	6
Contained:	outing information (sexual orientation)	1	1
375 ("gender")		9	17
Contained:	outing information (gender identity)	7	15

contents of other fields. All of the sixty records in the sample set included at least one 670 field. The 670 field is most commonly used to cite the work in which the form of the author's name was found. Typically, this is a work being cataloged for which a controlled name heading is needed, and thus the NAR was created. However, in the case of NAR records for authors who are trans-identified, this research found that information is often added to the 670 field that would not be included in the NAR for a cisgender author. Table 5 summarizes the findings related to 670 fields.

In the sample record set, thirty-six of sixty records (60 percent) contained information in the 670 field that explicitly outed the author as trans. This was done through using language such as "transgender," "transsexual," or "trans" (*n* = 23, 38 percent) by describing changes in names or pronouns used with inferences to gender identity, or by sharing information about the person's medical history using often-problematic words/phrases such as "transitioned," "gender-identity disorder," "sex-reassignment surgery," or "sex-change operation" (*n* = 6, 10 percent). These terms are outdated, generalized, and objectify transfolk, and stray from the NAR's purpose of helping catalog users find resources created by a specific individual.

The 670 fields cited information from a variety of sources, such as the author's work (*n* = 17, 28 percent), a publisher of the author's work (*n* = 9, 15 percent), and Wikipedia (*n* = 6, 10 percent). Only three records cited an author's personal website (*n* = 3, 5 percent), and just one cited direct communication with the author (*n* = 1, 2 percent). Of the thirty-six records that contained outing information in the 670 field, there were forty 670 fields in total with outing information (some records had multiple 670 fields with outing information). Some records cited the source of outing information from both the author (either their work(s), personal website, or direct communication) and another source, thus the total of forty outing instances across thirty-six records. Of the forty instances, twenty-one cited the author as the source of information (53 percent of outing 670 fields), while nineteen cited another source, such as a publisher or Wikipedia (48 percent of outing 670 fields). Table 6 summarizes the types of information sources cited. While 60 percent of records in the sample set (*n* = 36) contained information that outed the author as trans, only about half of those (*n* = 21, 58 percent of outing records, 35 percent of the total record set) cited the author as a source of information.

In addition to including outing information, many records used gendered language when a gender-agnostic term is preferable (*n* = 22, 37 percent), used "mosaic" pronouns (pronouns for the same person that correspond to multiple different genders, which can sometimes be consensual but are often seen as nonaffirming of identity) (*n* = 11, 18 percent), used outdated or incorrect pronouns (*n* = 9, 15 percent), or included nonbibliographically significant pronouns to indicate authorship of a work (a practice not required by the cataloging rules) (*n* = 6, 10 percent). See table 7 for a summary of these results.

Nine (15 percent) of the records analyzed included information related to a person's medical history or a change in name or pronouns to justify the inclusion of multiple 375 field(s) or subfields $s and $t "start" and "end" dates. This information was often recorded as one of only two sex/gender categories (male or female), which does not reflect many people/authors' lived experiences. Of these, six records defined trans-ness or gender transitions using information about surgery or medical histories (10 percent). Six records mentioned a name or gender identity change (10 percent), sometimes including the phrase "legal name change," or a change in the personal pronouns used by an author. Of the six records in the sample set with multiple 375 fields, three used binarist medical transition or name change information to "justify" this information (3 of 60 records, 5 percent; 3 of 6 records, 50 percent). These results are summarized in table 8.

Discussion

Inclusion of 375 fields (Q1)

As shown in this analysis, the new MARC 375 field for "gender" is being included in NARs cataloged using RDA. This field appeared in 70 percent of RDA NARs in the sample set (*n* = 7) and was the only new 37X field added to any updated AACR2 records that were examined (*n* = 2). Future research is needed to see how the adoption of this field fares over time, and to perform comparative analyses of NARs for

Table 5. Use of Outing 670 Fields

	No. of Records	% of Records (*n* = 60)	
Have one or more 670 fields:	60	100	
Author outed in 670:	36	60	
Method of outing in the 670:		% of records with 670(s) (*n* = 60)	% of records with outing 670 field(s) (*n* = 36)
trans identity label	23	38	64
name/pronoun change	6	10	17
medical history information	6	10	17

Table 6. Specific Sources of Outing Information in 670 Fields

Source of Outing Information (Some Cited Multiple Sources):	No. of 670 Fields	% of Records with Outing Information in the 670 Field (*n* = 36)	% of Records with 670 Fields (*n* = 60)
author's work	17	47	28
publisher	9	25	15
Wikipedia	6	17	10
personal website	3	8	5
IMDb	1	3	2
Chronicle of Higher Education	1	3	2
email from author	1	3	2
"Find a Grave" website	1	3	2
Gay & Lesbian Biography	1	3	2

authors who are transgender versus NARs for authors who are cisgender.

Content Values of 375 Fields (Q2)

The 375 field specification also includes subfields $s and $t, for "start" and "end" dates related to a particular biological sex, gender, or another identity label.[51] These subfields were used in five out of nine records with 375 fields in the test set (56 percent of records with 375 fields, 8 percent of the record set). Including a "start" and "end" date for a gender identity or sex category reinforces the problematic ideology of gender as a binary in which a person might only move from point A to point B, rather than gender as something that is constantly performed, constructed, and fluid. While the purpose for developing these subfields was not uncovered in this research, instructions in the MARC standard documentation for the 375 field and instructions provided in LC training both include examples of using these subfields for the purpose of delineating the gender of transfolk and provide no other use cases for these fields in the context of a NAR for a personal name.[52] Like many microaggressions, one can assume that these decisions came from a well-intentioned place: by representing the changes in identities

Table 7. Inconsistent Gendered Language in 670 Fields by Type

Records with Inconsistent Gendered Language:	No. of Records	% of Records that Contain Inconsistent Gendered Language (*n* = 22)	% of All Records in Sample Set (*n* = 60)
use gendered language	22	100	37
use mosaic pronouns in non-affirming way	11	50	18
use old or wrong pronoun	9	41	15
pronoun without bibliographic significance included	6	27	10

Table 8. Use of 670 Fields to Indicate Trans Status Using Medical or Name Change Information

Records with Medical or Name Change Information	No. of Records	% of All Records in Sample Set (*n* = 60)	% of Records with Medical Transition or Name Change Information (*n* = 9)
Record indicates trans status with medical transition or name change information	9	15	100
medicalization or surgery for transition date "justification"	6	10	67
mentions name change	6	10	67
have multiple 375s and justify w/ medicalization or name change	3	5	33

a person has had over the course of their life and giving the cataloger approximate time periods to use, we can represent authors as they choose to be known, rather than as they may have been known at the time a given work was published. This potentially helps users find related resources by referring them to a work under a name that they might not have known. However, gender identities are fluid and complicated, and do not instantaneously change simply with the start or stop of a prescription, the legal process of changing a name on official documents, or on the date of any kind of surgery, should an individual have enough financial and social privilege to access that type of health care or legal intervention. Identities are formed and re-formed over lifetimes, and often in an individual's life there are many transition processes, not a single monolithic transition. The author believes that this experience is erased by delineating years of identity changes, unless these dates have come directly from the author who is being described by the NAR and they have given their consent to share the information.

Limiting entries in the 375 field to values contained in a controlled vocabulary, such as LCSH or ISO 5218:2004, Information Technology Codes for the Representation of Human Sexes, constrains the record's ability to accurately and respectfully represent the identities of the authors it describes. As modern sociological research has indicated, human genders and sexes are socially constructed categories and exist outside of the binaries of simply "woman" and "man" or "female" and "male."[53] Gender identities are socially constructed, constantly performed, and tremendously fluid.[54] These are not just theories, but the real identities

and lived experiences of people across social groups and stratifications. Honoring this will require finding creative ways to incorporate authors' own words and labels, with their consent.

In discussing the creation of adequate controlled vocabularies, Drabinski argues that replacing problematic subject headings in national authority vocabularies may be useful and needed, but there are problems in imposing our view of identities and our language on people who lived in a different time or different cultures or circumstances as ourselves.[55] She also argues that a reading of the NAF through a queer theory lens could yield a conclusion that a controlled vocabulary describing people and identities can never be complete or perfect simply because of the nature of how it is constructed.[56] One way to work with these systems, she writes, is to engage with them through a queer lens, asking, "Whose voices are missing here? Whose are represented? Who has the power in this situation?"[57] If we apply this thought framework to NARs, we might conclude that if it is indeed important to include these identities in a record, it would be better left to an individual to decide and to carry out on their own terms. Further research is needed to assess whether the terms currently recorded in these records are congruent with the labels with which the authors themselves identify in a system where they can edit their own authority record, similar to the ORCID research identifier system.

A few of the NARs analyzed contained outdated information in the 375 field. Gender identity is fluid and often changes throughout an individual's lifespan, and continuing to update the NARs to reflect these changes is not easy

in the present closed system. While most current library systems lack the capacity to use this data at present, the potential future use of this data becomes problematic when maintenance is not feasible because of several constraints.

Other Representations of Gender Beyond the 375 Field (Q3)

The third research question was whether additional representations of gender were included in NARs other than the use of the 375 field. The author observed five mechanisms through which gender is represented or outing information is shared in NARs beyond use of the 375 field. This shows that regardless of whether the 375 field is used, current cataloging practice employs other mechanisms through which this private and sensitive information is being shared.

As noted in the introduction, one of the records analyzed (<2 percent) included a qualifier in the name heading that was not part of the author's name as currently used for publication and was not necessary to disambiguate the author from another with the same or similar name. Inclusion of this superfluous name information is not respectful of the author's self-identification. The PCC Task Group on the Creation and Function of Name Authorities in a Non-MARC Environment's report argues that a URI or author identifier could do a better job of disambiguating records than these subfields, while being compatible with a linked data future.[58] As the report suggests, to do this requires moving beyond the legacy parameters of the "record" and involves rethinking many of our current systems.[59] This research supports the usefulness of this suggestion.

While 60 percent of records in the test set contained information that outed the author as trans (n = 39), only about half of those (n = 21, 35 percent of the total set of records) cited the authors themselves as the source of that information. The practice of outing authors via any source but their own voice removes their personal agency to self-disclose, and the author appeals to catalogers to consider this when examining cataloging practices. Including pronouns that may not be accurate representations of the pronouns an author chooses to use can be nonaffirming to an author's gender identity and are better left out unless the pronouns of reference can be confirmed from a source that captures the author's own voice.

Some records included information that medicalized gender transition, perhaps as a form of "proof" that it had occurred. This act denies and erases the lived experiences of those transfolk who chose not to or cannot obtain medical intervention in their identity formation processes. Additionally, documenting deeply personal information that relates to medical histories or name changes is not respectful of an author's right to privacy, and violates the information sharing ethics on which laws such as HIPPA protections have been

built. Further research is needed to compare the amount of information, especially that of a personal nature, that has been provided in name authority records for authors who are trans versus those who are cisgender.

Linked Data and Self-Description as Potential Solutions

Linked data provides a solution for ensuring consistent and uniquely identifiable data. Rather than including fixed, selected data in a record, the authority record could connect a name (or series or set of names) to a URI. Links could be established to the author's works and other authoritative sources of information over which the author may have more control. Names would not need to be unique or disambiguated to create a unique heading, as the URI link would serve the purpose that the heading formerly had.

The author proposes a shift in thinking about name authorities from a system where catalogers are the authoritative voice to one where authors have the agency to self-describe their own experiences to whatever extent they wish. An example of a system with widespread adoption that does this is the ORCID unique researcher identifier system.[60] ORCID iDs have been adopted in practice by many systems, such as those that track scholarly research outputs, systems for tracking alternative research metrics, several journal article publishing platforms, and in the application processes for government granting agencies. ORCID functions on the premise that authors create their own authority record, include the information that they choose to disclose (and control levels of privacy for pieces of information on a very granular level), and link to their varied scholarly works, affiliations, and other researcher identifiers to establish their bibliographic identity as an author and researcher.

If name authorities continue as a closed system, catalogers can work to improve the system by working through the lens of transgender theory and examining the bibliographic significance of including specific information. Some questions to help guide this practice include the following:

- Is there potential for this information to harm the author through outing or violating the right to privacy?
- Is there an indication that the author consents to having this information shared publicly?
- Will including this information help a library user in the search process?

Outing, Privacy, and Safety Issues

The argument for including fields such as the 375 in NARs to enhance search and retrieval possibilities is not strong enough to justify the inclusion of such sensitive information in a publicly accessible database. While our current library systems lack the capacity to use this data, the use of this

data to search for authors or works could be quite problematic in future implementations. Not only do these new and current practices tokenize and "other" those whose gender identities fall outside of the traditional cissexist binary, but imagine the response if RDA suggested the addition of a race or religion category to the name authority record. Other legally protected classes such as race, religion, and sexual orientation are not addressed by the RDA attribute recommendations. Gender identity should be conceptualized in the same way. Gender identity is a protected class in many states, and should be given the full weight of respect as any other when it relates to talking about and potentially labeling other people. This research lends support to the recommendation to not include the 375 field in NARs unless communication and permission has been received from the author. If an authority record creator needs to justify information given in other fields in a record with a 670 field, only information pertinent to the work in hand should be included, and drawn from the author's work or personal communication.

Inability of the Suggested Values and Cataloging Practices to Accommodate Trans Identities and Experiences

Although writers who self-identify as trans in some way are not unique in the bibliographic world in having either a series of names that they use or name sets that they use, we as a community of practice should examine how bibliographic name authority descriptions, particularly when using RDA, do not necessarily treat these authors equitably as compared to their peers who identify as cisgender. It is worth considering as catalogers whether the current policies governing when to create a new name authority record are still conceptually sound when we acknowledge that identities are fluid and change over time. The practice of filing under latest-entry is only standard for personal names; serials catalogers have grappled with this concept and settled on successive entry as a better solution. Under this model, changes in the names of corporate bodies, conferences, and titles of serials lead to new authority records that link previous and/or successive names or titles. This is also true for rules governing bibliographic records for new editions of works (new editions get new bibliographic records.) How would this look in practice for personal names? What does that mean for NARs in their role as entity descriptions versus authorized lists of headings? Catalogers could consider using URIs or author IDs as part of a practice of using linked data to solve part of the issue of requiring an authorized form of name. Name changes could be recorded in an external location (an author's personal website, Wikipedia, etc.) while the URI is all that is needed to link that additional information to the bibliographic database.

Conclusion

Previous literature has critiqued RDA 9.7 and the use of the MARC Name Authority Record 375 field in relation to the problematic representation of gender.[61] This field was included in 70 percent of RDA NARs analyzed for this research, and was the only new authority field added to any of the AACR2 records in the test set. Analysis of the contents of these 375 fields confirms that its use is frequently outing in practice when used to describe authors who self-identify as trans.

Additionally, this research has shown that there are other areas of the NAR format besides the 375 field that warrant further scrutiny from the cataloging community. According to the results of this investigation, MARC 400 and 670 fields contain outing information in both NARs with 375 fields and those without. This finding raises practical and ethical concerns for catalogers, including risk of inaccurate or outdated representations of creators' identities; outing, privacy, and safety issues; and the inability of the suggested values and cataloging practices to accommodate trans identities and experiences.

Catalogers can work as allies to communities of people who experience oppression, discrimination, and often violence as a result of others' perceptions of their identities, and in particular, gender identities; this work is supported by the core principles of the Library Bill of Rights.[62] We are in a unique position as catalogers to do this work, and to open the doors for all authors to self-describe in authentic, empowered ways, and to assist users in discovering and accessing their work.

References and Notes

1. Library of Congress, Network Development and MARC Standards Office, "Introduction. Scope of the Authority Format" in *MARC 21 Format for Authority Data* (Washington, DC: Library of Congress, 1999 edition (section revised October 2009)), https://www.loc.gov/marc/authority/adintro.html.

2. Library of Congress, Network Development and MARC Standards Office, "MARC 21 Format for Authority Data" (Washington, DC: Library of Congress, 1999), https://www.loc.gov/marc/authority; Special Libraries Cataloguing, "MARC Authority Record Cheat Sheet," 2011, www.special-cataloguing.com/node/1395.

3. Bob Thomas, "Name Disambiguation—Learning From More User-Friendly Models," *Cataloging & Classification Quarterly* 49 no. 3 (2011): 223–32, http://dx.doi.org/10.1080/01639374.2011.560834.

4. Ibid.; Xavier Agenjo, Francisca Hernández, and Andrés Viedma, "Data Aggregation and Dissemination of Authority Records through Linked Open Data in a European Context," *Cataloging & Classification Quarterly* 50, no. 8 (2012): 803–29.

5. Library of Congress, Network Development and MARC Standards Office, "375—Gender (R)," in *MARC 21 Format for Authority Data*, 1999 ed., rev. 2009 (Washington, DC: Library of Congress), https://www.loc.gov/marc/authority/ad375.html.

6. Julie L. Nagoshi and Stephan/ie Brzuzy, "Transgender Theory: Embodying Research and Practice," *Affilia: Journal of Women and Social Work* 25, no. 4 (2010): 431–43.

7. Amber Billey, Emily Drabinski, and K. R. Roberto, "What's Gender Got to Do with It? A Critique of RDA 9.7," *Cataloging & Classification Quarterly* 52, no. 4 (2014): 412–21, http://dx.doi.org/10.1080/01639374.2014.882465; "GLBTRT Task Force on RDA and Gender in Authority Records," *GLBTRT News*, 2008, www.glbtrt.ala.org/news/archives/20.

8. The term *gender binary* is the assumption that there are only two human genders: women and men, even though in reality two categories do not represent the genders of all people. It includes cultural assumptions that a trans person would experience a transition only once, and in doing so, simply move from one side of an imagined (and false) gender spectrum to the other. While some people who are trans experience their gender identity and any transitions they may access in this way, many do not. Charlie McNabb explains, "Nonbinary identities are those that fall outside of the traditional binary 'man' and 'woman' gender categories. Nonbinary folks can be somewhere between man and woman, a mixture of both, or may identify completely separately from these categories. Nonbinary people can also move between genders or have no gender at all. The term 'nonbinary' is used as an umbrella term . . . individual nonbinary folks may have any number of other identity terms that they prefer." [Charlie McNabb], "Nonbinary Gender Identities in Media: An Annotated Bibiography" (annotated bibliography, 2015), www.ala.org/glbtrt/sites/ala.org.glbtrt/files/content/professionaltools/Nonbinary%20Gender%20Identities%20in%20Media%20Annotated%20Bibliography.pdf. See also Julia Serrano, "Regarding Trans* and Transgenderism," *Whipping Girl* (blog), August 27, 2015, http://juliaserano.blogspot.com/2015/08/regarding-trans-and-transgenderism.html.

9. Cisgender people experience their gender identity in ways that conform to cultural expectations about how people assigned their biological sex at birth should exist in the world. For example, a person who was assigned female at birth (biological sex), and self-identifies as a woman (gender identity), might identify as a cisgender woman. A person who was assigned female at birth (biological sex), and self-identifies as a transman, as genderqueer, or as a man (gender identities) might identify as transgender.

10. Daryl B. Hill and Brian L. B. Willoughby, "The Development and Validation of the Genderism and Transphobia Scale," *Sex Roles: A Journal of Research*, 53, no. 7–8 (2005): 531–44.

11. National Coalition of Anti-Violence Programs, "Lesbian, Gay, Bisexual, Transgender, Queer, and HIV-Affected Hate Violence in 2013: A Report from the National Coalition of Anti-Violence Programs," (New York: New York City Gay and Lesbian Anti-Violence Project, 2014), www.avp.org/storage/documents/2013_ncavp_hvreport_final.pdf.

12. "Tips for Allies of Transgender People," GLAAD, accessed August 25, 2015, www.glaad.org/transgender/allies.

13. The author has used the pronouns "they," "them," "their," "themselves," and "theirs" as third-person singular pronouns throughout this work in two types of instances. The first is to affirm the pronouns used by specific individuals consistent with their gender identities. The second is as a practice of using language uninhibited by gender when referring to a non-specified person implicitly or explicitly of any gender. For additional resources on this language construction, please see Zena Sharman, "Singular They Reading List," accessed October 4, 2013, http://ivancoyote.tumblr.com.

14. The author has redacted personal information from this name heading in this manuscript for privacy. The full name authority record cited here can be found on LC's webpage at https://lccn.loc.gov/n2001029640.

15. Program for Cooperative Cataloging, PCC Task Group on the Future of Name Authorities in a Non-MARC Environment, "Report for PCC Task Group on the Creation and Function of Name Authorities in a Non-MARC Environment," 2013, www.loc.gov/aba/pcc/rda/RDA%20Task%20groups%20and%20charges/ReportPCCTGonNameAuthInA_NonMARC_Environ_FinalReport.pdf.

16. Ivan E. Coyote, "'They' Is Me: Call Us What We Wish to Be Called," *Daily Xtra*, January 11, 2012, www.dailyxtra.com/vancouver/they-is-me-4408.

17. Agenjo, Hernández, and Viedma, "Data Aggregation and Dissemination," 824.

18. Ibid, 805.

19. Policy and Standards Division, Library of Congress, "Z1. Name and Series Authority Records—670 Source Data Found," *Descriptive Cataloging Manual*, accessed September 1, 2015, www.loc.gov/catdir/cpso/dcmz1.pdf.

20. Joint Steering Committee for Development of RDA, "9.17 Biographical Information," in *RDA: Resource Description and Access* (American Library Association, Canadian Library Association, and Chartered Institute of Library and Information Professionals (CLIP), 2009).

21. "GLBTRT Task Force on RDA and Gender in Authority Records."

22. Ibid.

23. "RDA Listserv (RDA-L) Email Discussion February 12, 2008–March 3, 2008," Joint Steering Committee for Development of RDA (JSC), Library and Archives Canada, 2008; J. McRee Elrod to RDA-L mailing list, February 12, 2008, https://www.mail-archive.com/rda-l%40listserv.lac-bac.gc.ca/msg01607.html.

24. Ibid.

25. Ibid.

26. Ibid.

27. Billey et al., "What's Gender Got to Do with It?," 412–21.

28. Ibid, 412, 414.

29. Ibid., 414, 417, 420; "GLBTRT Task Force on RDA and Gender in Authority Records," www.glbtrt.ala.org/news/archives/20.

30. Ana Lupe Cristán, "RDA Special Topics: RDA Elements in Name Authority Records (NARs): MARC 21 Fields," (presentation, Policy and Standards Division, Library of Congress, Washington, DC, 2012), www.loc.gov/aba/rda/source/special_topics_elements_NARs.ppt.

31. Joint Steering Committee for Development of RDA, "9. Identifying Persons," in *RDA: Resource Description and Access* (American Library Association, Canadian Library Association, and Chartered Institute of Library and Information Professionals (CLIP), 2009).

32. Library of Congress, Network Development and MARC Standards Office, "MARC 21 Format for Authority Data."

33. International Standards Organization/ International Electrotechnical Commission, "Information Technology—Codes for the Representation of Human Sexes (ISO/IEC 5218:2004)," (Geneva, Switzerland: ISO/IEC, 2004), www.iso.org/iso/catalogue_detail.htm?csnumber=36266.

34. Ibid.

35. Ibid.

36. It is beyond the scope of this paper to discuss the mechanisms and systems involved in changes to the LCSH vocabulary.

37. American Library Association, "Library Bill of Rights," (statement of basic principles, 1939, reaffirmed 1996), accessed August 26, 2015, www.ala.org/advocacy/intfreedom/librarybill.

38. American Library Association, "Access to Library Resources and Services Regardless of Sex, Gender Identity, Gender Expression, or Sexual Orientation: An Interpretation of the Library Bill of Rights," 1993, amended 2008, www.ala.org/advocacy/intfreedom/librarybill/interpretations/accesslibrary.

39. Melissa Adler, "Transcending Library Catalogs: A Comparative Study of Controlled Terms in Library of Congress Subject Headings and User-Generated Tags in Librarything for Transgender Books," *Journal of Web Librarianship* 3, no. 4 (2009): 309–31; Emily Drabinski, "Teaching the Radical Catalog," in *Radical Cataloging: Essays at the Front*, edited by K. R. Roberto (Jefferson, NC: McFarland, 2008), 198–205; Emily Drabinski, "Gendered S(h)elves: Body and Identity in the Library," *Women & Environments International Magazine* 78/79 (2009/2010): 16–18; Matt Johnson, "Transgender Subject Access: History and Current Practice," *Cataloging & Classification Quarterly* 48, no. 8 (2010): 661–83; K. R. Roberto, "Inflexible Bodies: Metadata for Transgender Identities," *Journal of Information Ethics* 20, no. 2 (2011): 56–64.

40. Drabinski, "Teaching the Radical Catalog," 198–205; Emily Drabinski, "Queering the Catalog: Queer Theory and the Politics of Correction," *Library Quarterly* 83, no. 2 (2013): 94–111.

41. Drabinski, "Teaching the Radical Catalog," 198–205; Drabinski, "Queering the Catalog," 94–111.

42. Frank Exner, Little Bear, "North American Indian Personal Names in National Bibliographies," in *Radical Cataloging: Essays at the Front*, edited by K. R. Roberto (Jefferson, NC: McFarland, 2008): 150–64.

43. Ibid, 162–63.

44. Ibid.

45. Ibid, 151–52.

46. Ibid, 152.

47. Ibid, 151.

48. Lorraine J. Pellack and Lori Osmus Kappmeyer, "The Ripple Effect of Women's Name Changes in Indexing, Citation, and Authority Control," *Journal of the American Society for Information Science & Technology* 62, no. 3 (2011): 440–48.

49. Nancy Silverrod et al., "TRANScending Identities: A Bibliography of Resources on Transgender and Intersex Topics," accessed November 14, 2013, www.ala.org/glbtrt/sites/ala.org.glbtrt/files/content/professionaltools/glbtrt_trans_08.pdf; Wikipedia, "Category:Transgender and Transsexual Writers," accessed November 14, 2013, http://en.wikipedia.org/wiki/Category:Transgender_and_transsexual_writers.

50. Network Development and MARC Standards Office Library of Congress, "375—Gender (R)," accessed November 14, 2013, www.loc.gov/marc/authority/ad375.html.

51. Ibid.

52. Library of Congress, Network Development and MARC Standards Office, "375—Gender (R)"; Cristán, "RDA Special Topics."

53. International Standards Organization/ International Electrotechnical Commission, "Information Technology—Codes for the Representation of Human Sexes (ISO/IEC 5218:2004)," (Geneva, Switzerland: ISO/IEC, 2004), www.iso.org/iso/catalogue_detail.htm?csnumber=36266; Judith Lorber, "Shifting Paradigms and Challenging Categories," *Social Problems* 53, no. 4 (2006): 448–53.

54. Judith Butler, *Gender Trouble: Feminism and the Subversion of Identity* (New York: Routledge, 1990).

55. Drabinski, "Queering the Catalog," 94–111.

56. Ibid, 106–8.

57. Ibid.

58. Program for Cooperative Cataloging, PCC Task Group on the Future of Name Authorities in a Non-MARC Environment, "Report for PCC Task Group on the Creation and Function of Name Authorities," 1–16.

59. Ibid, 13.

60. ORCID website, "What is ORCID?," accessed May 4, 2015, http://orcid.org/content/initiative.

61. Billey et al., "What's Gender Got to Do with It?"; "GLBTRT Task Force on RDA and Gender in Authority Records"; Elrod to RDA-L mailing list, February 12, 2008.

62. American Library Association Council, "Access to Library Resources and Services Regardless of Sex, Gender Identity, Gender Expression, or Sexual Orientation: An Interpretation of the Library Bill of Rights."

Survey of Classification and Organization of Videorecordings

Barbara Bergman, Jessica Schomberg, and Dorie Kurtz

Libraries have long struggled with the question of how to best classify and enable access to videorecordings. While giving a presentation at a state library conference, the authors observed from comments during the question and answer portion of the presentation that the libraries represented in the audience use a variety of video classification and organization practices. To better understand how local practices and librarians' attitudes regarding the efficacy of these practices vary, we conducted a survey soliciting responses from librarians representing a broad array of library types.

Classification is the act of systematically grouping similar things together. Within this study, the term "classification" refers to assigning call numbers that reflect subject content and are intended to support logical browsing of physical or virtual locations. Libraries continue to debate the question of how to best classify and provide access to video materials. While giving a presentation on the topic of video collections at a state library conference, the authors learned that libraries are using a wider variety of video classification and organization practices than we anticipated. That raised questions for us. We wanted to know whether this variation was a regional phenomenon or a general practice. Additionally, we also wanted to know whether librarians generally believed that their local practices for classifying videos adequately helped patrons. Therefore we decided to survey catalogers, media librarians (media selectors who may also do media cataloging), and generalists interested in video classification. We chose the survey method because it is an inexpensive way to obtain a broad range of responses from a wide audience nationally and, possibly, internationally. The authors anticipated that this would help identify areas requiring additional research. Because our research questions stemmed from a discussion with a diverse group of librarians, we intentionally tried to obtain responses that reflected the diversity of the profession for library type (academic, public, etc.) and organizational role (cataloging, collection development, etc.).

An additional idea that we had is that catalogers are enculturated to follow standard cataloging practices to insure interoperability between computer systems and to provide a consistent experience for patrons.[1] However, in her article on classification based on sameness and difference, Olson stated,

> I would like to build on the idea of diversity rather than universality as a way of accommodating our diverse users and collections and, thus, decreasing levels of bias. To do so requires a variety of approaches—not only can we not have a universal solution, we cannot even have a universal method for achieving solutions.[2]

Barbara J. Bergman (barbara.bergman @mnsu.edu) is Media Services Librarian at Minnesota State University Mankato. **Jessica Schomberg** (jessica .schomberg@nmsu.edu) is Media Cataloger/Assessment Coordinator at Minnesota State University Mankato. **Dorie Kurtz** (dorie.kurtz@avila.edu) is Cataloging and Acquisitions Librarian at Avila University.

Asking whether librarians generally believe that their local practices for classifying videos adequately helped patrons, what problems librarians experience with video classification, and librarians' perceptions about how their libraries handle classification enabled us to explore these ideas.

Literature Review

Discoverability

Classification is essential to making library materials discoverable. Kinney's paper on historical changes in moving image classification recounts the changes in attitude toward classifying audiovisual materials and toward open rather than closed stacks. Switching to open stacks meant that videorecordings became browsable collections that required a user-friendly classification system.[3] Present guidelines support including nonprint materials in library collections and in library catalogs, using the same classification practices as for books.[4] While classification data are not well used in current catalog systems, Kinney references Calhoun to state that its use "to support browsing by subject is among the proposed steps for revitalizing the catalog."[5] Kinney references Handman's argument that media materials are educationally significant, relatively unique, and complex in terms of bibliographic description and access requirements, all of which means that there should be more core cataloging elements for media than for books.[6]

Record Quality

The *Survey of Academic Library Cataloging Practices*, conducted in 2011 by the Primary Research Group, found that 95.71 percent of the seventy libraries surveyed performed video cataloging in-house.[7] One challenge for libraries when classifying videorecordings is that the quality of records available in OCLC varies widely. Not surprisingly, films with the widest distribution tend to have the highest-quality bibliographic records.[8] But for any given film title, catalogers may be faced with multiple choices when selecting the best quality video record.

To determine the quality of records in their OPAC and whether minimal record editing was sufficient, Myall and Chambers examined the catalogs of Eastern Washington University and Gonzaga University. They found that neither of the catalogs they studied had been able to import US national-level, full-level records of video resources without editing. Librarians at those institutions perform minimal institutional-level editing of videorecords to conform to OCLC's requirements for full cataloging (I-Level).[9] The average videorecording record required about six edits, as compared to two edits for print monographs. The fields requiring the most additions or corrections were not identified. Not surprisingly, when the encoding levels decreased in completeness, more additions and edits were required. The study also noted that most records have been modified by multiple libraries in addition to the Library of Congress (LC). This suggests that one cannot simply rely on a record-matching method of copy cataloging for videorecordings. Videorecording cataloging requires catalogers with knowledge of both the cataloging practices and the motion picture industry. Myall and Chambers recommended establishing "a national cooperative program to provide authoritative cataloging records for videos," which "could also develop a larger, more widely distributed, and more knowledgeable cohort of audiovisual catalogers as well as improve the accuracy, completeness, and consistent treatment of bibliographic records for these materials."[10]

In some ways, groups such as OLAC (Online Audiovisual Catalogers) have engaged in this type of work. However, their current video cataloging guidelines, *Best Practices for Cataloging DVD-Video and Blu-ray discs using RDA and MARC21*, addresses neither classification nor subject or genre headings.[11]

Circulation Decisions

Kinney referenced Ho's 2004 study that found that "73 percent of ARL libraries and 39 percent of all libraries surveyed shelved their video collections in closed stacks."[12] Also circa 2004, Laskowski and Bergman noted that a growing number of institutions allowed students to check out videos for use outside of the library. They stated that at one institution, policies changed because of "the discrepancy in encouraging students to make use of videos for class presentations while making it difficult for them to access these media materials."[13]

Classification Decisions Related to Serving Underrepresented Communities

Accession Numbers

Library policies evolved to treat videorecordings as full members of the library collection, compared to 1989 when Scholtz advocated including videorecordings in the library's catalog.[14] While most libraries now include videorecordings in their catalogs, many still use basic accession numbers to organize these materials, in contrast to the fully classified by subject print materials.[15] Like classification numbers, accession numbers indicate specifically where videos are located in the collection. Unlike classification numbers, accession numbers reflect only the sequence in which those materials were processed, with no other contextual information. This

practice requires little time or skill to apply, and can be helpful for inventory purposes. Using accession numbers to organize videos can cause challenges for patrons because it does not allow for shelf browsing by subject or alphabetically, and is used predominantly by libraries with closed stacks.[16] Kinney's work supports the interpretation, "The best that can be said for shelving by accession numbers is that it may make for cheaper and faster cataloging."[17]

Formal Classification Structures

Bowker and Star, in their influential book *Sorting Things Out*, noted that "information scientists work every day on the design, delegation, and choice of classification systems and standards, yet few see them as artifacts embodying moral and aesthetic choices that in turn craft people's identities, aspirations, and dignity."[18] They identified the elements of a formal classification system: consistent principles, mutually exclusive categories, and complete descriptions of the area it covers. In actual practice, classification systems do not operate according to these ideals. Local modifications to Library of Congress Classification (LCC), local shelving decisions, and classification choices based on misunderstandings or disagreements with the structure, illustrate how a library's individual decisions make classification a complex process.

Olson's research supports those ideas. She argues that no classification system is ever completely inclusive. Classification systems such as LCC are created on the basis of literary warrant, and "what gets published is also limited by powerful social discourses, it too tends to produce a corpus largely representing mainstream thought."[19]

One of these classification oversights relates specifically to video cataloging. Dewey Decimal Classification (DDC) and LCC schemes are not designed to accommodate nonprint materials. This is notably problematic for fiction television programs and films, which are often squeezed into LCC's PN schedule as examples of things that *are* drama amid nonfiction titles *about* drama. It is also a problem for other moving image materials. Kinney discussed various options to modify LCC's existing structure for music materials, suggesting possibilities such as including classifying musical performances with the same number that the print version would receive and Cuttering videorecordings of operas and other works by composer rather than title.[20] An example of a local modification from the authors' own library is to treat the director as the primary creator using the literature schedule. This works well for film studies favorites such as Hitchcock and Kurosawa. However, the Harry Potter films, which patrons would reasonably expect to be shelved together, but which had different directors, were separated both in our modified version of LCC and in strict LCC. We have therefore made selective modifications

to local practices to ensure that all the Marvel Avengers films are shelved together and that the Harry Potter series is collocated. The practices surrounding classification of videorecordings are different from those for print monographs. While it would be ideal if standards reflected those differences, they currently do not.

Problems in Classification Structures

Classification encourages browsing and discovery in a way that is not possible using accession number and closed-stacks arrangements.[21] The construction of those classification systems occurs within the confines of the culture(s) in which they are developed. Olson argued that "[people] take the classification for granted as though it were a natural landscape rather than a well-manicured lawn that is the product of intellectual labor."[22] Intellectual labor can reify the unconscious biases of its creators.[23]

Referring specifically to items physically shelved in libraries, Olson noted that those items are limited to the linear space in which each item is assigned one defining classification number/code, with distinct boundaries between classification numbers/codes.[24] Making the choice to include or exclude something from that space is never a neutral act. Another classification problem sometimes faced by video catalogers is how to assign responsibility when many people are involved in a film's creation. Classification systems based on a single-author Cutter are not optimal in this situation. While strictly following LCC and Cuttering videorecordings by title does not cause this problem, it can be challenging for libraries that modify LC rules to Cutter by director (or other creator). This problem is not unique to big budget Hollywood films. Discussing DDC's treatment of folklore, Olson notes that Western cultures prioritize the individual over the collective, even when a particular work is created by multiple people within a longstanding cultural tradition.[25]

Comparing how groups such as gays and lesbians are treated by classification systems, Olson and Ward argued that DDC's structure creates spatial gaps, or diasporas, between corresponding topics by assigning similar but distinct subclasses.[26] Similar concerns occur within LCC. For example, films about men generically and white men specifically are classed in HQ1090, while films about black men are often classed in E185. This suggests that white men is the default in the first classification area. This same process effectively racializes films about black men racialized in a manner not applied to films about white men.

Streaming Video

The adoption of streaming video in libraries has changed user expectations. Educational use of video resources on campus has accelerated across all disciplines, and even as

recently as five years ago, faculty, librarians, and administrators anticipated that use of video resources in education would grow significantly. There was the problem that "technology, legal, and other barriers continues to thwart faculty finding and accessing the segments of video they want for teaching and lectures."[27] Researchers also identified a clear need from those working in higher education for an online video repository that could be accessed as needed.[28] In this case, that need surpasses the minimal service found through vendors such as Netflix. Convenience is the primary interest reason to use streaming video, but there is also interest in the added value that can be included, such as transcripts and the ability to create a playlist of specific film clips.[29] For teaching and related purposes, users require high quality metadata about content, usage rights, format, and platform specifications, in addition to streaming video resources being reliable enough for classroom use.

Classification of free streaming videos of lectures and from major research institutions is important.[30] While these lectures are usually posted on video sites, author-provided metadata are not always as user-oriented as that provided by professionals. Creating records for these videos in OCLC will allow us to facilitate patrons' access to these important resources.

The use of streaming video has already reached the tipping point in the majority of academic libraries licensing at least some streaming video. Data from the 2010 Primary Research Group survey indicated that across all Carnegie classifications of Institutions of Higher Education, approximately 33 percent of academic libraries provided streaming video.[31] Whereas results of the 2013 Survey of Academic Streaming Video (SASV) conducted by Farrelly and Hutchinson showed an increase to 70 percent of all academic libraries providing some streaming video, and most of the remaining libraries planned to do so within three years.[32]

SASV results also found that libraries use a wide variety of tools, including the OPAC, discovery layers, and LibGuides, to provide access to streaming videos. Librarians who responded to the survey expressed their preference for using the online catalog to search for streaming videos because of title-specific searching. However, 25 percent of respondents did not provide catalog access to streaming videos, and only 41 percent used the catalog as the primary access point. Licensing restrictions were cited by respondents as one reason streaming video was not being included in their catalog, but the major reason appears to be inadequate staff time being devoted to managing streaming video. As stated by Farrelly and Hutchinson,

> Librarians are largely unaware of many factors related to streaming video. Many librarians are unfamiliar with models, practices, systems, and other

factors related to the acquisition and support of streaming videos in their collections. Significantly, librarians overwhelmingly report relatively low level of staff time to support streaming video. In the aggregate, for selection, licensing, encoding and uploading, and metadata, respondents report a staff commitment of less than one full-time equivalent. More than a quarter of respondents, however, report not knowing what that commitment is.[33]

Method

During a presentation the authors made at a state conference, librarians and library staff in attendance disclosed a variety of local practices for videorecordings. This inspired us to ask a larger audience of librarians whether most believed that their local practices for classifying videos adequately helped their patrons; what problems librarians experience with video classification, such as inconsistent categories or bias; and librarians' perceptions about how their libraries handle classification.

To increase our understanding of the existing variety of classification and shelving practices for videos, we developed a survey. Because our intention was to capture diverse perspectives that we might not otherwise have considered, we intentionally solicited responses from a broad array of librarians. Our reason for implementing this survey was to have a diverse sample of responses. We knew that surveys using convenience sampling are often ungeneralizable, but diverse convenience samples increase their usefulness. Diverse sampling ensure that a broad range of potential participants are invited, and provide similar attitudinal results as one finds from probabilistic samples.[34] Catalogers and public services librarians bring different insights, as do public librarians and academic librarians. Since our goal was to learn about participants' practices and opinions, we permitted open textual responses for many of the questions in addition to listed response options. A copy of the survey is in the appendix.

The questionnaire consisted of two sections. The first section was on demographics: the kind of libraries the respondents represented and the location of these libraries. The second section was the decisions that respondents made about video classification: the kinds of videos classified, where videos are shelved, what classification standards are used, and how satisfied respondents are with the classification standards they use.

We used Qualtrics (www.qualtrics.com) to disseminate and analyze the survey results. To solicit the broad response we desired, links to the survey were posted to email discussion lists for catalogers, media librarians and generalists, including OCLC-CAT, OLAC-L, VideoLib and colllib-l. The researchers also shared survey links on their personal

Table 1. Are Your Videos on Open Stacks? (i.e., patrons can go to the shelves and select their own videos)

	Academic Libraries	Public Library	Special Library	School (K-12) Library	Other	Total
Yes	148	81	13	3	9	254
No	89	0	9	1	6	105
Some	65	6	1	2	1	75
Total	302	87	23	6	16	434

Table 2. Are Your Videos Integrated with Other Formats?

	Academic Libraries	Public Library	Special Library	School (K-12) Library	Other	Total
Yes	28	5	9	1	4	47
No	230	64	9	4	9	316
Some	44	18	5	1	3	71
Total	302	87	23	6	16	434

Facebook or Twitter pages. In the two-week response period (January 12–23, 2015), 412 surveys were completed.

Results

Respondents were primarily from academic libraries (7 percent from two-year colleges, 17 percent from four-year colleges, and 45 percent from universities, totaling 69 percent of responses). Because there were no real differences in percentages for how the academic libraries responded, they are grouped as "academic" for discussion of the survey results. Twenty percent of responses were from public libraries, 1 percent were from school libraries, 5 percent were from special libraries, and 4 percent were from other types of libraries. Several of the "other library" responses were from law librarians. Regarding location, 98 percent of respondents were from North America, 1 percent were from Europe, and 1 percent were from elsewhere.

The authors based the survey questions on ideas discussed during our conference presentation and identified through a literature review. Questions were oriented around decisions various library types make about physical access, streaming access, and classification. Because the questionnaire was distributed to individuals and participants were not required to list their institutional affiliation, it is possible that multiple librarians from the same institution answered the survey. The goal was to capture a diverse array of perspectives, so we did not consider this a barrier. While the survey was open to all library types, the number of responses from the school, special, and other library categories were too small to be meaningful when separated from the whole. Therefore we have chosen to present only separate responses from public and academic libraries.

Physical Access

Our previous research indicated that shelving and checkout policies often affect whether collections are classified with call numbers.[35] Therefore we included questions about these practices in the survey.

Open or Closed Stacks

Not surprisingly, all public librarians that responded indicated that their libraries shelve videos in open stacks. Half of the academic librarian responses indicated that their video collections are in open stacks, with an additional 22 percent using a mixed open/closed stack arrangement for portions of the collection. Comments added by respondents indicating mixed open/closed stack shelving included a variety of configurations, such as popular fiction titles are in open stacks, while more expensive educational DVDs are kept in closed stacks. In other cases, VHS tapes have been moved to open stacks. Some libraries display DVD cases in open stacks while retaining the actual DVD behind a service desk. Table 1 shows numerical results.

Format Integration

Most librarians (73 percent) indicated that their libraries do not integrate videos with other formats. But 7 percent of public library respondents and 12 percent of academic library respondents do integrate videos with other formats, with the integrated shelving generally being used for nonfiction videos (see table 2). Comments showed processing variations in cases when DVDs are treated as accompanying materials to books.

Table 3. Who Can Checkout Videos?

	Academic Libraries	Public Library	Special Library	School (K–12) Library	Other	Total
Students	261	9	8	4	5	287
Faculty	271	8	6	6	6	297
Community Members	131	72	3	2	5	213
In-Library Use Only	9	1	2	0	3	15
Not Applicable	0	0	0	0	2	2
Other	60	18	9	1	3	91
Some Mix Depending on Material Type and User Type	46	4	3	0	0	53

Table 4. How Do You Classify Videos at Your Library?

	Academic	Public	Special	School (K–12)	Other	Total
Library of Congress (LCC)	183	5	9	1	7	205
Decimal (DDC, UDC)	27	60	2	4	2	95
Other	22	12	5	0	2	41
Total	232	77	16	5	11	341

Check-out Permissions

The majority of libraries have less restrictive lending policies than revealed in past surveys.[36] Responses showed that the majority of libraries lend to multiple patron types, including some interlibrary loans. A few still restrict patrons to on-site viewing or faculty-only checkouts. Unfortunately, there appears to have been some confusion about how to answer the question, which resulted in many open-ended descriptions about specific policies (see table 3).

Reserve Shelving

The survey did not specifically ask about videos in reserve collections. It was assumed that reserve and/or booked videos would be held in restrictive closed stacks. Comments provided by several respondents confirmed this assumption.

Classification

Most respondents use a formal classification scheme to arrange videos, with almost equal numbers strictly following cataloging rules or making local modifications. The predominant classification schemes are LCC (60 percent) or DDC (28 percent), with percentages similar to the percent of responses from academic libraries and public libraries respectively. Other formal schemes referenced include

Superintendent of Documents (SuDoc) and Medical Subject Headings (MeSH). Two libraries use BISAC subject headings as their classification system. BISAC stands for Book Industry Standards and Communication. Based on open-ended comments added to this question, it appears that BISAC subject descriptions are often used as shelving guides to facilitate browsing within public libraries (see table 4).[37]

As previously noted, we allowed respondents to comment on their activities and perceptions to provide a richer understanding of actual practices. In the next section, much of what we describe came from free text responses.

Some respondents reported that the format shift from videotapes to DVDs enabled them to improve local cataloging practices. By shelving DVDs separately from videotapes, and therefore avoiding the laborious process of reclassifying all of their videotapes from accession numbers, they fully classified DVDs as they were added to the collection. Even libraries that still assign accession numbers appear to otherwise fully include DVDs in the catalog.

Unlike nonfiction films (documentaries, etc.), which were almost universally classified using DDC or LCC, there was more variety in how libraries handle fiction films (see tables 5 and 6).

From the survey data, we found that there is not a single way that a majority of libraries classify their fiction videos. However, it is clear that librarians make choices on the basis of their patrons' needs and to maintain consistency

Table 5. What Kinds of Videos do You Classify?

	Academic Libraries	Public Library	Special Library	School (K–12) Library	Other	Total
Fiction Only	0	0	0	0	0	0
Nonfiction Only	10	20	3	0	4	37
All of Them	245	62	15	6	8	336
None of Them	34	2	5	0	2	43
Other	18	4	0	0	2	24
Total	307	88	23	6	16	440

Table 6. Classification Variations

		Academic	Public library	Special library	School (K–12) library	Other	Total
Do you strictly follow a library classification standard for videorecordings? i.e., there are no local modifications	Yes	143	23	10	1	4	181
	No	157	64	13	5	11	250
Do you loosely follow a library classification standard for videos—one with local modifications?	Yes	100	58	7	3	8	176
	No	53	6	6	0	2	67
Do you arrange videos alphabetically by title?	Yes	13	5	2	1	0	21
	No	45	1	4	1	3	54
Do you arrange videos by accession number?	Yes	39	0	3	0	3	45
	No	5	1	1	1	0	8
Do you arrange videos by subject or genre?	Yes	3	1	0	0	0	4
	No	3	0	0	1	0	4

in the stacks. A few respondents noted that their current system was inherited from previous catalogers or administrations and that they would change it if possible, but were constrained by lack of time, money, and staff. Many of the respondents addressed which section of LCC they use to classify their videos. Librarians identified if they were using PN1995.9 to classify their videos by genre or PN1997-1997.2 to class films by their title. PN1997 is used to class films that were created before 2001 and PN1997.2 is used to class films from 2001 to the present. More than one librarian chose to classify all films under PN1997 to have the entire collection in alphabetical order without the separation of years. This type of classification can present challenges when Cuttering for the title as longer Cutter numbers are more difficult to locate and shelve. Multiple librarians mentioned the length of the Cutter as a cause of frustration in their library.

The majority of participants indicated that they are happy with their video classification scheme. Through a five-point Likert scale that used smiley faces to indicate level satisfaction, 83 percent of 325 responses received for this question indicated that they were happy to very happy (4–5) with their classification scheme (see table 7).

This overall satisfaction level was followed with questions regarding what classification problems concerned participants, as shown in table 8.

Again, using a five-point scale, the average value for the prompt "We're pretty happy with how this works for us" was 4.13 ("Sometimes/Always"). On the other end of the scale, 9 percent indicated at least some dissatisfaction with classification. An additional 8 percent responded neutrally. These percentages were consistent across all library types. We also asked whether classifications are too specific or too broad. Seventy percent of respondents indicated that they are "never" or "rarely" too specific. Respondents seemed divided about whether they are too broad. Approximately 30 percent of respondents answered that they are Sometimes too broad and 29 percent of respondents indicated that classifications are "rarely" too broad.

Considering the possibility of providing multiple classification numbers in a streaming environment, we asked whether librarians ever assign multiple classifications.

Table 7. Satisfaction with Classification

#	Answer	Bar	Response	%
1	😦	▌	4	1.23%
2	😦	▬	25	7.69%
3	😐	▬	27	8.31%
4	🙂	▬▬▬▬▬▬	184	56.62%
5	😄	▬▬▬	85	26.15%
	Total		325	100.00%

Seventy-four percent of respondents said that was "never" or "rarely" the case, while only 17 percent indicate that they "sometimes" or "always" want to do so.

The prompt with the least support, with a 1.98 average ("never/rarely"), was "The whole classification scheme seems kind of racist, sexist, ableist, etc.," which 204 of 376 respondents said was never a problem. While one could easily question the wording of this question, as it was intended to elicit an instinctive response. Considering the literature review above, some of which is deeply critical of existing classification schemes because of systemic issues of racism and sexism, we thought it was important to include this perspective.[38]

Streaming Access

In response to the question, "Does your library purchase access to streaming video?," 64 percent indicated "yes." Of those librarians who responded with "yes," 56 percent were public librarians and 74 percent were academic librarians. Streaming videos are licensed from the vendor in various options for wide ranging lengths of time. Among librarians who indicate that their libraries acquire streaming video, only 67 percent include streaming video titles in their catalog. Comparing library types, we found that only 50 percent of public librarians and 70 percent of academic librarians who offer streaming video also say that they include the records in their catalogs. Compared to the SASV survey, which indicated that 75 percent of libraries offer access to streaming videos through the catalog, we infer that not including all streaming video titles in the catalog is still unfortunately a common situation.

For libraries that do not include streaming videos in the catalog, the primary reasons cited by twenty-seven librarians was that the library did not consider the streaming videos a priority or that they lacked the time and money necessary for this work. Similarly, an additional eleven responses indicated a reliance on vendor databases to provide access to streaming videos rather than including title-level access points to those databases from within the library catalog. Four of the librarians surveyed responded that streaming videos were available only to off-campus students for short-term use and therefore it was unnecessary to add the titles to the catalog.

Discussion

One of the questions that prompted this project was whether librarians believed that local video classification practices adequately help patrons. This study has revealed that the concerns about local video classification practices do not necessarily focus on the lack of uniformity, but the feelings of uncertainty that catalogers and media specialists have about their decisions. While many respondents noted that their video cataloging systems work for the most part, there are still areas that require more consideration. Areas that are particularly in need of further study include how to better represent genre for patron discovery and how to classify video adaptations of film and other formats.

Whether the library chooses to class using PN1995.9 or PN1997, the main issue is that they must constantly create their own Cutters for the film to fit into their collection, which can be time consuming and requires a professional cataloger. As the Primary Research Group noted in their report, only 4.29 percent of academic libraries in 2011 were outsourcing their video cataloging, while 44.29 percent of libraries outsourced their e-book cataloging.[39]

For the questions that asked about bias in classification, responses were as anticipated. Because of the sensitive nature of this area, we were also unsure whether participants would answer on the basis of their conscious thought processes rather than their automatic valuations, which Hofmann et al. found to be more reliably represented when participants answer spontaneously.[40] After seeing the strongly negative response to this prompt, we wondered if there were differences in perspective by library types. However, this did not seem to be the case: the "never" responses was selected by 54 percent of academic librarians, and 49

Table 8. Classification Opinions

Are these issues ever a problem with your video classification?		Academic	Public Library	Special Library	School (K–12) Library	Other	Total
a. Inconsistent classification (for example: some documentaries are in subject areas and some are mixed in with fiction films	Never	88	23	10	1	4	126
	Rarely	78	24	4	1	0	107
	Maybe?	28	11	1	2	4	46
	Sometimes	68	21	3	2	3	97
	Always	4	1	0	0	0	5
	Total	266	80	18	6	11	381
b. Classifications are too specific	Never	89	21	8	0	4	122
	Rarely	102	34	9	2	5	152
	Maybe?	23	11	2	2	1	39
	Sometimes	44	13	1	0	1	59
	Always	3	0	0	1	0	4
	Total	261	79	20	5	11	376
c. The whole classification scheme seems kind of racist, sexist, ableist, etc.	Never	140	38	16	4	6	204
	Rarely	71	25	3	2	2	103
	Maybe?	33	11	0	0	1	45
	Sometimes	16	4	1	0	2	23
	Always	1	0	0	0	0	1
	Total	261	78	20	6	11	376
d. I want to put multiple classifications on things. One number is too limiting!	Never	126	29	12	2	4	173
	Rarely	69	24	5	2	4	104
	Maybe?	22	7	2	1	2	34
	Sometimes	41	16	1	1	1	60
	Always	3	0	0	0	0	3
	Total	261	76	20	6	11	374
e. We don't classify videos so it's hard to find specific titles	Never	179	50	13	4	8	254
	Rarely	24	10	4	0	1	39
	Maybe?	13	3	0	0	1	17
	Sometimes	13	2	1	0	1	17
	Always	4	1	0	0	0	5
	Total	233	66	18	4	11	332
f. We're pretty happy with how this works for us	Never	9	3	1	0	2	15
	Rarely	9	0	0	0	1	10
	Maybe?	31	6	1	0	1	39
	Sometimes	123	36	8	5	4	176
	Always	100	34	10	1	5	150
	Total	272	79	20	6	13	390
g. Classifications are too broad	Never	68	13	7	1	3	92
	Rarely	68	28	5	4	2	107
	Maybe?	41	14	1	0	1	57
	Sometimes	80	21	7	1	5	114
	Always	5	0	0	0	0	5
	Total	262	76	20	6	11	375

Table 9. Streaming Video

		Academic	Public Library	Special Library	School (K–12) Library	Other	Total
Does your library purchase access to streaming videos?	Yes	225	49	2	2	2	280
	No	81	39	21	4	14	159
	Total	306	88	23	6	16	439
Do all streaming video titles appear in the catalog?	Yes	158	24	1	1	2	186
	No	67	24	1	1	0	93
	Total	225	48	2	2	2	279

percent of public librarians. Based on the comments, including one participant who was appalled at the other suggestion and one participant who expressed resignation about it, this appears to be a good area for further investigation through a more interactive method.

Demand for streaming video has grown quickly, requiring vendors and libraries to change their business models to include licensing access. Survey responses confirm that many libraries are licensing access to streaming video. The rapid changes in streaming video offerings and license options, especially in the availability of large subscription packages of films without concurrent staffing additions, have caused libraries to lag behind in including complete holdings in the catalog. From these results, we can make certain conclusions about the hierarchy of streaming videos compared to other library materials. It also raises questions about the purpose of streaming videos in libraries being used as support materials for a course. Survey respondents appear to recognize that streaming video resources should be in the catalog, but library processes and procedures need time to catch up, especially staff time assigned to this work. In 2015, Hutchison and Farrelly conducted a follow-up to their 2013 survey, the results of which we anticipate will show considerable changes in just two years.[41] For example, the licensing platform Kanopy has made huge inroads into hosted video content with a large number of video vendors, but was not yet available to US libraries in 2013. The recent addition of on-demand and evidence-based licensing options by several vendors greatly increases streaming licensing options for libraries.

Conclusion

The only clear conclusion reached from this survey is that there is not one consistent way to handle videorecordings in libraries and certainly no one right way. Libraries have generally increased their comfort level with managing videos as a more equal part of their collections, but in some ways the processes for handling these materials have not advanced greatly from the time when libraries began acquiring

videotapes twenty-five years ago. Many academic libraries still use closed stacks and accession numbers for physical items, although that number is decreasing for both. The increasing availability of streaming video is changing how films are acquired and made discoverable and accessible. As libraries have seen with e-journals and e-books, we are now increasingly providing licensed access as opposed to owning a physical item. As libraries shift to a focus on virtual browsing, they will need to reevaluate how those films are classified. At this point, a substantial number of both academic and public libraries do not include bibliographic records for streaming video in their catalogs.

One topic that emerged from the questions related to problems librarians experience with videorecording classification is how systemic discrimination is manifested within classification schedules. While this survey suggests that it is not a major concern for most librarians, cataloging research suggests that it is an important issue for users. To overcome limitations within the survey method when asking potentially sensitive questions, we suggest using qualitative methods including a discourse analysis of classification decisions, reflexive case studies, interactive interviewing, or focus groups with library patrons from marginalized populations.

In a related note, there may be differences in patron perceptions of video classification processes when compared to library employee perceptions. User studies examining patron efficacy and comfort with library terminology, organizational structures, and spaces are an important area that needs considerable investment in the future.

References

1. Lois Mai Chan and Marcia Lei Zen, "Metadata Interoperability and Standardization: A Study of Methods Part I, Achieving Interoperability at the Schema Level," *D-Lib Magazine* 12, no. 6 (2006), www.dlib.org/dlib/june06/chan/06chan .html.
2. Hope A. Olson, "Sameness and Difference: A Cultural Foundation of Classification," *Library Resources & Technical Services* 43, no. 3 (2001): 121.

3. Daniel W. Kinney, "The Classification of Music Moving Image Materials: Historical Perspectives, Problems, and Practical Solutions," *Cataloging & Classification Quarterly* 47, no. 1 (2009): 2–22, http://dx.doi.org/10.1080/01639370802451975.

4. Association of College and Research Libraries, American Library Association, *Guidelines for Media Resources in Academic Libraries* (Chicago: American Library Association, 2012), accessed June 2, 2015, www.ala.org/acrl/standards/mediaresources.

5. Kinney, "The Classification of Music Moving Image Materials," 8.

6. Gary Handman, "I Lost It (or Found It) at the Movies: Public Service Implications of Minimal-Level Cataloging of Audiovisual Materials," *Technical Services Quarterly* 9, no. 2 (1991): 39–47.

7. Primary Research Group, *The Survey of Academic Library Cataloging Practices* (New York: Primary Research Group, 2011).

8. Barb Bergman, Jessica Schomberg, and Victoria Peters, "Video Collecting for the Sometimes Media Librarian: Tips and Tricks for Selecting, Purchasing, and Cataloging Videos for an Academic Library," *College & Undergraduate Libraries* 14, no. 1 (2007): 57–77.

9. Carolynne Myall and Sydney Chambers, "Copy Cataloging for Print and Video Monographs in Two Academic Libraries: A Case Study of Editing Required for Accuracy and Completeness," *Cataloging & Classification Quarterly* 44, no. 3–4 (2007): 233–57, http://dx.doi.org/10.1300/J104v44n03_06.

10. Ibid., 252.

11. "Best Practices for Cataloging DVD-Video and Blu-ray Discs using RDA and MARC21," OLAC, accessed October 18, 2015, http://olacinc.org/drupal/capc_files/DVD_RDA_Guide.pdf.

12. Kinney, "The Classification of Music Moving Image Materials," 6.

13. Mary Laskowski and Barb Bergman, "Academic Media Center Collection Development and Circulation Policies: A Comparative Analysis," *College & University Media Review* 10, no. 2 (2004): 86.

14. James Scholtz., *Developing and Maintaining Video Collections in Libraries* (Santa Barbara, CA: ABC-CLIO, 1989).

15. Bergman, "Making the Most of Your Video Collection," 335–48.

16. Bergman, Schomberg, and Peters, "Video Collecting for the Sometimes Media Librarian," 57–77.

17. Kinney, "The Classification of Music Moving Image Materials," 10.

18. Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things Out: Classification and Its Consequences* (Cambridge, MA: The MIT Press, 1999): 4.

19. Hope A. Olson, "Mapping Beyond Dewey's Boundaries: Constructing Classificatory Space for Marginalized Knowledge Domains," *Library Trends* 47, no. 2 (1998): 235–36.

20. Kinney, "The Classification of Music Moving Image Materials," 11.

21. Ibid., 10.

22. Olson, "Sameness and Difference," 115.

23. Olson, "Mapping Beyond Dewey's Boundaries," 233–34.

24. Ibid., 245.

25. Olson, "Sameness and Difference," 118.

26. Hope A. Olson and Dennis B. Ward, "Ghettoes and Diaspora in Classification: Communicating Across the Limits," in Bernd Frohmann, ed., *Communication and Information in Context: Society, Technology, and the Professions* (Toronto, Ontario, Canada: Canadian Association for Information Science, 1997): 27, www.cais-acsi.ca/ojs/index.php/cais/article/download/185/359.

27. Peter B. Kaufman and Jen Mohan, "Video Use and Higher Education: Options for the Future," (New York: Intelligent Television and Copyright Clearance Center, 2009): 2, https://library.nyu.edu/about/Video_Use_in_Higher_Education.pdf.

28. Ibid., 11.

29. Matt Enis, "On Demand Academic Streaming Media," *Library Journal,* October 15, 2015, http://lj.libraryjournal.com/2015/10/books-media/on-demand-academic-media.

30. George Wrenn, "Hidden in Plain Sight? Records for on-Demand Academic Public Lectures in OCLC WorldCat: A Survey," *Cataloging & Classification Quarterly* 48, no. 8 (2010), http://dx.doi.org/10.1080/01639374.2010.492465.

31. Primary Research Group, *Library Use of Video and Audio* (New York: Primary Research Group, 2011).

32. Deg Farrelly and Jane Hutchison, "Academic Library Streaming Video: Key Findings from the National Survey," *Against The Grain* 26, no. 5 (2014): 73.

33. Ibid., 75.

34. Samuel J. Best et al., "An Assessment of the Generalizability of Internet Surveys," *Social Science Computer Review* 19, no. 2 (2001): 131–45.

35. Barb Bergman, "Making the Most of Your Video Collection: Trends in Patron Access and Resource Sharing," *Library Trends* 58, no. 3 (2010): 335–48.

36. Bergman. "Making the Most of Your Video Collection," 335–48.

37. "Tutorial and FAQ," Book Industry Study Group, accessed October 18, 2015, https://www.bisg.org/tutorial-and-faq.

38. Olson, "Sameness and Difference," 115.

39. Primary Research Group, *The Survey of Academic Library Cataloging Practices*, 2011–12 ed. (New York: Primary Research Group, 2011), 108–9.

40. Wilhelm Hofmann et al., "A Meta-Analysis on the Correlation between the Implicit Association Test and Explicit Self-Report Measures," *Personality & Social Psychology Bulletin* 31, no. 10 (2005): 1369–85, http://dx.doi.org/10.1177/0146167205275613.

41. Enis, "On Demand Academic Streaming Media."

Appendix. Video Classification Survey

1. Do you want to take this survey?

Demographic Information

2. What type of library organization do you work for?
    ○ 2-Year College or equivalent
    ○ 4-Year College
    ○ University
    ○ Public library
    ○ Special library
    ○ School (K–12) library
    ○ Vendor
    ○ Other
3. Location?

And now to the real questions . . .

We are going to use the words video and videorecording to include all formats of this type: VHS, DVD, film, streaming, etc.

4. The way that your organization is structured, where does the cataloging occur?
    ○ Off-site cataloging department
    ○ In-house cataloging department
    ○ We purchase catalog records, we don't have a cataloging department
    ○ Other
5. Who can check out videos? (Check all that apply)
    ○ Students
    ○ Faculty
    ○ Community members
    ○ No one—in library use only
    ○ Not applicable
    ○ Some mix depending on material type and user type that I'll describe here
    ○ Other
6. What kinds of videos do you classify?
    ○ Fiction only
    ○ Nonfiction only
    ○ All of them
    ○ None of them
    ○ Other
7. Does your library purchase access to streaming videos?
    ○ Yes
    ○ No
8. Do all streaming video titles appear in the catalog?
    ○ Yes
    ○ No

9. You indicated that not all streaming titles are displayed in your catalog. Why not?
10. Are your videos on open stacks? (i.e., patrons can go to the shelves and select their own videos without mediation)
    ○ Yes
    ○ No
    ○ Some of them
11. Are your videos integrated with other formats?
    ○ Yes
    ○ No
    ○ Some of them
12. Do you strictly follow a library classification standard for videorecordings? i.e., there are no local modifications
    ○ Yes
    ○ No
13. Do you loosely follow a library classification standard for videos - one with local modifications?
    ○ Yes
    ○ No
14. Do you arrange videos alphabetically by title?
    ○ Yes
    ○ No
15. Do you arrange videos by accession number?
    ○ Yes
    ○ No
16. Do you arrange videos by subject or genre?
    ○ Yes
    ○ No
17. Do you put videos on the shelves in no particular order?
    ○ Yes
    ○ No
18. How do you classify videos at your library?
    ○ Library of Congress (LCC)
    ○ Decimal (DDC, UDC)
    ○ BISAC
    ○ Other
19. What kinds of local modifications do you use? (for example: you follow LCC for non-fiction films but modify how you treat fiction films)
20. How do you choose what subjects/genres to organize videos into?
21. How happy are you with the classification standard that your library uses for videorecordings?

22.   Are these issues ever a problem with your video classification?

	Never	Rarely	Maybe?	Sometimes	Always
Classifications are too broad	○ Classifications are too broad Never	○ Classifications are too broad Rarely	○ Classifications are too broad Maybe?	○ Classifications are too broad Sometimes	○ Classifications are too broad Always
We're pretty happy with how this works for us	○ We're pretty happy with how this works for us Never	○ We're pretty happy with how this works for us Rarely	○ We're pretty happy with how this works for us Maybe?	○ We're pretty happy with how this works for us Sometimes	○ We're pretty happy with how this works for us Always
I want to put multiple classifications on things. One number is too limiting!	○ I want to put multiple classifications on things. One number is too limiting! Never	○ I want to put multiple classifications on things. One number is too limiting! Rarely	○ I want to put multiple classifications on things. One number is too limiting! Maybe?	○ I want to put multiple classifications on things. One number is too limiting! Sometimes	○ I want to put multiple classifications on things. One number is too limiting! Always
We don't classify videos so it's hard to find specific titles	○ We don't classify videos so it's hard to find specific titles Never	○ We don't classify videos so it's hard to find specific titles Rarely	○ We don't classify videos so it's hard to find specific titles Maybe?	○ We don't classify videos so it's hard to find specific titles Sometimes	○ We don't classify videos so it's hard to find specific titles Always
Classifications are too specific	○ Classifications are too specific Never	○ Classifications are too specific Rarely	○ Classifications are too specific Maybe?	○ Classifications are too specific Sometimes	○ Classifications are too specific Always
The whole classification scheme seems kind of racist, sexist, ableist, etc.	○ The whole classification scheme seems kind of racist, sexist, ableist, etc. Never	○ The whole classification scheme seems kind of racist, sexist, ableist, etc. Rarely	○ The whole classification scheme seems kind of racist, sexist, ableist, etc. Maybe?	○ The whole classification scheme seems kind of racist, sexist, ableist, etc. Sometimes	○ The whole classification scheme seems kind of racist, sexist, ableist, etc. Always
Inconsistent classification (for example: some documentaries are in subject areas and some are mixed in with fiction films)	○ Inconsistent classification (for example: some documentaries are in subject areas and some are mixed in with fiction films) Never	○ Inconsistent classification (for example: some documentaries are in subject areas and some are mixed in with fiction films) Rarely	○ Inconsistent classification (for example: some documentaries are in subject areas and some are mixed in with fiction films) Maybe?	○ Inconsistent classification (for example: some documentaries are in subject areas and some are mixed in with fiction films) Sometimes	○ Inconsistent classification (for example: some documentaries are in subject areas and some are mixed in with fiction films) Always
Other	○ Other Never	○ Other Rarely	○ Other Maybe?	○ Other Sometimes	○ Other Always

23.   What additional concerns do you encounter?
24.   Is there anything else you would like the researchers
      or the library community to know?

Literature of Acquisitions in Review, 2012–13

Angela Dresselhaus

The acquisitions literature published in 2012–13 shows a strong focus on non-traditional purchasing models, especially for electronic books (e-books). Patron Driven Acquisition (PDA) is one method that helps librarians cope with budget constraints that continue to plague library budgets. The expense of Big Deals has some libraries seeking more efficient alternatives such as Pay-Per-View or Evidence-Based Selection, however, many libraries are still reliant on the depth of coverage and perceived value of Big Deals. This review will cover these trends along with developments in Electronic Resources Management Systems (ERMSs), workflow efficiencies, and negotiation and licensing techniques.

*L*ibrary Resources & Technical Services has published four prior reviews of acquisitions literature covering the period between 1996 and 2011.[1] This installment covers literature published in 2012–13. Dunham and Davis noted that the literature published from 1996 to 2003 reflects an environment seriously disrupted by the implementation of automated systems and the Internet, requiring acquisitions departments to move from paper systems to automated library systems. Print acquisitions job responsibilities were restructured to accommodate electronic resource management. In 2004–7, Dunham and Davis observed that electronic journal (e-journal) purchasing was well underway and becoming increasingly complex, necessitating the development of electronic resource management tools. Harrell characterized 2008–9 as driven by budget reductions with the transition of subscriptions from print to electronic collections. The role of the Big Deal was questioned as library budgets decreased and employee skill sets continued to shift to accommodate electronic publications, especially in the area of negotiation and licensing. Moeller noted that budget constraints continued to shape the literature, which is evident in the interest in Open Access and Patron-Driven Acquisitions (PDA), and the increasing scrutiny of Big Deals. For the current review period, budget constraints continue to shape the conversation, causing librarians to question the long-term sustainability of big journal deals, seek consortial-level PDA plans to share costs, and implement open source electronic resource management tools instead of subscription based solutions.

Method

The author replicated the research method described in Moeller's 2010–11 literature review.[2] A search in Library and Information Science Abstracts (LISA), Web of Science (WOS), and Worldcat.org was conducted using the keywords: library and acquisitions. The date range for each search was limited to materials published in 2012–13. Materials were further filtered to only include scholarly materials published in English. The LISA search yielded 544 results, the WOS search yielded 57 results, and the Worldcat.org search yielded 134 results. Selected

Angela Dresselhaus (dresselhausa15@ecu.edu) is Head of Electronic Resources at East Carolina University.

journals and conference proceedings from the Charleston Conference and the North American Serials Interest Group (NASIG) meetings were systematically reviewed for inclusion in this literature review. The author struggled to find a clear distinction between acquisitions literature and literature covering collection development and e-resources. Despite this challenge, papers focused on purchasing models, supporting workflows, and acquisitions management tools were selected. Papers covering public libraries, secondary school libraries, collection management studies, and gift acquisitions were excluded from this literature review.

Literature Review

The materials considered for this review are primarily peer-reviewed papers, along with relevant columns, interviews, trade publications, conference proceedings, and monographs.

Purchasing Models

Firm Orders

Few studies pertaining to monographic firm orders were published during 2012–13. Tony G. Horava, associate university librarian, Collections, at the University of Ottawa wrote a case study documenting the implementation of a new firm order monographic process using the vendor YBP.[3] Horava noted that a user-centered approach to monographic acquisitions requires that the library develop a workflow to quickly deliver requested books. Economic pressures require that the library use financial and human resources efficiently, reducing the amount of staff involved in fulfillment. In response to these pressures, the librarians at the University of Ottawa decided to implement newly released YBP features to streamline order fulfillment. Advancements in technology allowed the library to partner with YBP to create a more efficient monograph firm order process. Implementing a direct ordering method in YBP reduced fulfillment time for monograph purchases and freed librarians to focus on more complex electronic resource purchases. Horava concluded that the case study was successful, as the library was able to realize a one-week reduction to shelf time for books, acquisitions budget balances were readily available, and time saved by streamlining firm order processing was invested into the management of e-resources.

Rita Cauce, head of the Resource Development Department, Florida International University, discussed the development of an electronic monograph request system that replaced a paper request system.[4] The system was named Online Library Acquisitions System (OLAS), and was designed to replace a paper-based tracking system used to manage firm order purchase requests. The electronic system was designed to allow librarians to quickly determine the status of an order before it was entered into the ILS. Cauce determined that OLAS improved efficiencies and increased accountability within the acquisitions unit.

Approval Plans

In previous years, approval plan management was challenged by publication delays that often existed between print and electronic monographs, which complicated combining print and electronic acquisition profiles. Forzetting, Wiersma, and and Eager demonstrated that a partnership between library vendors and librarians was essential for developing comprehensive approval plans built on careful profiling and meticulous tracking of all library monographic purchases.[5] The most compelling points detail how the librarians and vendor representatives incorporated the established PDA models into the print approval plan workflow and how they accommodated the sometimes disparate electronic and print publication dates. In a separate publication, Wiersma developed a study to determine publication date differences between electronic and print format.[6] Wiersma found that the publication gap has closed dramatically to the point that in 2011 many publishers had moved to simultaneous print and electronic publication. Wiersma also analyzed publisher and subject trends. Ultimately, the data she collected was used to develop a highly tuned approval plan to meet patron needs.

Consortial Purchases

The literature yielded several studies on consortial purchasing in the United States, from state-based groups such as OhioLINK, the Colorado University System, the California State University Library Consortium, the Arizona Universities Library Consortium, and the Consortium of Academic and Research Libraries in Illinois (CARLI), to large regional consortia such as the Orbis Cascade Alliance (OCA) and Association of Southeastern Research Libraries (ASERL).[7] This literature review included only English language materials, limiting the diversity of international consortia; however, studies from Canadian and Chinese library consortia were discovered.[8] Whether large regional or statewide, or small groups such as the Triangle Research Library Network (TRLN) or the Five College Consortium, many groups share similar goals, such as increasing buying power and providing broad access to research materials.[9] Machovec's paper provides a comprehensive look into consortia, their histories, and goals.[10]

Pilot projects for Demand Driven Acquisition (DDA) and Patron Driven Acquisition (PDA) were prevalent, including examples from the OCA and Ontario Council of University Libraries (OCUL). OCA's DDA was discussed in an interview conducted by Jill Emery, collection

development librarian, Portland State University.[11] She interviewed members of OCA's Demand Driven Acquisitions (DDA) implementation team and their two vendor partners. The interview format provides a firsthand look into how various OCA members and their partners perceived the planning, implementation, and success of the project. While this paper provides a big picture perspective, Arch et al. describe in detail the process and factors in place that led to a successful consortial DDA plan with OCA members.[12] They determined through a review of literature that very little has been published regarding the evaluation of large DDA programs, and the contribution of their work to the literature filled a gap with a detailed treatment of how the evaluation was conducted.

Davis et al. published a paper on OCUL's shared PDA experience.[13] The authors reported that OCUL, a twenty-one-member organization serving 420,000 users, attempted to develop a shared PDA program to meet the needs of individual schools, while maintaining a group shared cost benefit. Each perspective is presented; especially helpful are the vignettes from individual institutions detailing implementation concerns, usage, and rewards of the program.

Big Deals

Frazier, Bergstrom, and Nicholas et al. extensively addressed the Big Deal in prior years.[14] Best, Gatten and Sanville, and Gibbs laid the foundation as early as 2004 for a shift away from Big Deals and have refocused the debate in 2012–13 on whether the Big Deal is dying.[15] The University of California Libraries, the California Digital Library, Southern Illinois University-Carbondale, and the University of Oregon have cancelled Big Deals.[16] McGrath discussed the efforts of a steering committee tasked with finding alternatives to the Big Deal by Research Libraries UK (RLUK), a consortium of research libraries in the United Kingdom.[17] Despite efforts at some institutions to show less dependence on large journal packages, Big Deals have deeply penetrated the library market; Strieb and Blixrud reported on data collected from Association of Research Libraries (ARL) libraries and found that "three of four publishers covered in the two most recent surveys (Elsevier, Springer, and Wiley) are now licensed as bundles by 90% or more of libraries for which data were collected."[18] The survey also found that Big Deal purchases are still commonly made through consortia, a factor which no doubt led authors Ashmore, Grogg, and Weddle to state "rumors of the Big Deal's death have been exaggerated."[19]

An analysis of the literature suggests that the question should be further refined to "are serial Big Deals dying and big e-book deals flourishing?" Proctor notes that e-book Big Deals are mirroring what was seen in the e-journal market, namely desirable content packaged with low use materials, high costs, and little ability to choose titles included in a Big Deal.[20] Big Deals are not dying, and are evolving and may need to continue to evolve as alternatives are developed.

The Balance Point column, published in *Serials Review*, is a long-standing venue for experienced librarians to share individual perspectives on key issues in librarianship. One installment, edited by Dyas-Correia, consisted of interviews with representatives from small and large publishers, vendors, libraries, and consortia and asked if the Big Deal was on the way out.[21] The panel agreed that there is a great deal of talk about cancelling Big Deals, but very few packages are actually canceled. In fact, the interview participants predicted that a large scale cancellation of Big Deal packages could permanently alter the role of consortia or cause their demise. Dooley noted, "If consortia members opt out, it could well increase cost for the remaining members, not to mention the effects on trust and good will."[22] However, increasing costs and declining budgets are causing librarians to question the sustainability of purchasing all or a substantial portion of one publisher's content. Pay-Per-View (PPV), interlibrary loan (ILL), individual subscriptions, and Open Access models were suggested as alternatives to the Big Deal. Stanford University was cited as a model for meeting patron needs without subscribing to Big Deals. Van Rennes noted that "a little further down the line, I suspect that articles, rather than journals, will become the main unit of information commerce, and new models will be based upon that development."[23] This could produce additional revenue streams for publishers in the area of PPV article models, small packages, and as Bucknell found at University of Liverpool, evidence-based selection.[24]

Patron Driven Acquisitions (PDA)/Demand Driven Acquisitions (DDA)

Interest in PDA continues to grow and implementations are widely documented, creating a robust body of published literature covering implementation issues, workflows, and evaluation. PDA plans are often executed as a way to realize cost savings, yet England and Anderson acknowledge that "PDA models do not necessarily lead to cost savings, but they do by their very nature forge a strong connection between acquisition and real-world usage, making PDA a potentially attractive model for libraries concerned about maximizing the effective use of their acquisition fund."[25] This finding is consistent with other researchers, such as Dinkins and Schroeder, who have found that PDA selected materials enjoy excellent circulation.[26]

Implementation

Ward's book, *Guide to Implementing and Managing Patron-Driven Acquisitions*, provides a complete how-to guide for implementing and managing a PDA model.[27] Her coverage

of PDA is inclusive as she describes ILL requests and patron catalog discoveries as points of triggered purchases. Chapter three covers issues to consider before implementing a PDA program, such as timing, PDA parameters, and determining who will be authorized to trigger purchases. Ward acknowledges criticisms and some of the challenges leveled against PDA in chapter six. Finally, she concludes with a discussion of future directions that includes the possibility of digitization on demand, print on demand, and the potential for collection development to be completely directed by patrons, potentially a new era in collection development. Allison's monograph, *The Patron-Driven Library: A Practical Guide for Managing Collections and Services in the Digital Age,* complements Ward's advice, providing information on the technological environment, challenges in library acquisitions, patron preferences, and other big picture topics.[28]

A case study by Fischer et al. documented the implementation of a completely unmediated e-book PDA plan at the University of Iowa Libraries that began in September 2009.[29] The launch of the PDA plan was silent, the public was not informed to reduce bias or unreliable results during the evaluation of the pilot program. The authors found that during the trial, the allocation was quickly spent and they had to implement ways to control spending. Limits were not set to prevent format duplication, i.e., purchase of both print and electronic formats. Fischer et al. evaluated circulation rates for print titles that were duplicated by an e-book purchase from the PDA plan and discovered that the circulation rate of print titles fell drastically once electronic equivalents were acquired.

McCaslin described a PDA program built on the Amazon Kindle platform at the California Institute of Technology.[30] The library decided to circulate Kindles and provide patrons the authority to make $25 in e-book purchases. The Kindle was chosen because the library could associate up to six devices with the library's controlled account, thereby providing access to the same selection of e-books on all Kindles. The library used Amazon gift cards linked to a library account to eliminate the financial risk involved with circulating a device linked to an open credit line. The paper covered details such as circulation, acquisitions policies, and accounting issues. A patron satisfaction questionnaire was distributed as part of the Kindle check-out process, and results indicated that most of the users were undergraduate students interested in reading previously selected material rather than purchasing new content, and there were many repeat users. The six month trial period, including staff time, startup costs, and book purchases cost $10,243.14, a figure considered affordable by the library administration.[31]

Evaluation of PDA/DDA

Ongoing evaluation of the overall performance of PDA plans is needed to ensure value of this selection and purchasing

method. Tyler et al. designed a study consisting of non-parametric statistical tests to evaluate PDA effectiveness, particularly investigating the factors that show favorable circulation over traditionally selected materials in prior studies.[32] The authors found statistically significant differences in circulation rates between librarian-selected, vendor-selected, and patron-selected materials. The statistically significant result held even when the variables of book price, years available, and subject area were added to the model. The authors recommended continuing patron and librarian selection, while reducing the focus on vendor selection.

Walter's paper criticizes the value of PDA and considers it a risk to the long-term quality and ability of the library collection to meet the institution's educational mission.[33] The author believes that librarians play an important role in the selection of materials and advocates mediating patron requests. This paper includes a chart comparing major PDA implementations in the literature. The chart includes data such as year started, vendor, number of titles made available, triggering thresholds, number of titles purchased annually and the price per title.

A case study conducted by McLure and Hoseth at Colorado State University examined a user survey and use statistics to produce a snapshot of patrons' attitudes regarding PDA.[34] The library acquires a large portion of their e-books from the electronic PDA program, which is modified to include only a few select subjects. The authors discussed limitations of the survey, such as the brief online survey period and technological barriers that impacted the sample. The survey, set to pop up when a catalog record was viewed, asked if patrons had used e-books and whether they had a format preference, among other questions. The survey found an even split between those with no experience and experience using e-books. The survey showed "29.22 percent preferring an e-book, 32.80 percent preferring a print book, and 37.98 percent indicating no preference."[35] The authors used Dewey and LC subject headings analysis to gain an understanding of how faculty from various subject disciplines used the PDA model. The study concluded that most subjects were well served with PDA and the approval plans were reduced to cover a few select subjects.

Evidence-Based Selection

Bucknell described how the University of Liverpool (UOL) evaluated e-book packages and used modeling techniques to determine whether alternative acquisitions methods, such as PDA or DDA, were more cost effective than purchasing large e-book packages.[36] UOL librarians typically did not engage in speculative book purchases, but a change in user demand for e-books prompted a modification in practice that allowed the purchase of e-book packages. Since this was a departure from routine practices, the librarians were interested in

determining if there was an alternative way to meet patron demand while avoiding "just in case" collections. Bucknell modeled usage reports from purchased e-book packages to see if expenditures would have been higher with a PDA plan in place and found that the e-book package was more cost effective. Given this result, the next question is clearly how to identify what collections are most likely to be used. This is where evidence-based selection offers an opportunity to provide access to a large catalog of books and allows the library to buy books on the basis of consistent patron demand.

Pay Per View

In response to rapidly escalating prices and looming budget cuts, Hosburgh sought to mitigate the effect of journal cancelations, especially in the STEM area, by implementing both mediated and unmediated PPV models.[37] PPV was established with three publishers and the new "Get It Now" program from the Copyright Clearance Center. Challenges such as expired tokens, duplicate purchases, incompatible vendor systems and link resolvers, and conflicting collection development philosophy, were addressed.

ReadCube, a new service providing journal article PDA has been evaluated by England and Anderson. They address ReadCube's business model and its trial at the University of Utah's Marriott Library during a one-year case study.[38] While some feedback included negative experiences with DRM and the ReadCube client, they considered ReadCube to be a cost effective alternative to ILL. The novel approach discussed in this article is the acceptance of stringent digital rights management limitations in return for reduced costs.

An email thread on LIBLICENSE-L turned into a viable PPV model and the process was documented by Sowards.[39] The conversation was geared toward solving frustrations with Big Deal pricing using a PPV model. The author noted that continued access to published research materials is a goal shared by librarians and publishers. Over the course of the year, the author worked with Multi-Science Publishing to create a PPV plan. The article meticulously details conversations that shaped the deal, providing an interesting look at how partnerships can be formed to solve common problems. As a final note, Sowards's literature review was well constructed and covers the depth of articles related to PPV, including four articles that merit more than a mere mention by Fisher, Kurt, and Gardner; Brown; Weicher and Zhang; and Powell.[40]

Workflows

E-Books

As e-book publishing grows, librarians have developed and refined workflows to meet patron demand. E-book workflows frequently involve multiple steps and intradepartmental collaboration, and are often not linear. Walter's bibliographic essay is an excellent review of published literature covering various impacts on e-book workflows including: availability of scholarly e-books, electronic publication delays, licensing, e-book file formats, pricing models, and preservation.[41] While Walter's article focuses on challenges, the review is not negative and does not recommend abandoning e-book acquisition.

Building and Managing eBook Collections, edited by Kaplan, is the 184th installment in Neal-Schuman's How-To-Do-It Manuals series. This resource is an excellent addition to the literature because it provides a comprehensive collection of articles covering a wide range of topics in the management of e-book workflows.[42] The first of three sections, "E-Books in Context," includes three chapters dedicated to outlining the history and development of e-books and e-book publishing from industry and library perspectives. The second section, "E-Books in Detail," covers practical areas of e-book management such as selection, licensing, budgeting, cataloging, and assessment. Six case studies of e-book implementations in high school, public, and academic libraries are included in the third section, "E-Books in Practice," covering topics including marketing, e-book readers, staff workflow for managing e-books at a university, and managing e-books in an ERM.

Roncevic's installment of the *Library Technology Reports* series, "E-book Platforms for Libraries," evaluated e-book platforms and provided a comprehensive resource for public, academic, and school librarians seeking information on where to purchase e-books.[43] The author covered marketplace issues including corporate structure and business models and tied each business model to how it applies in public, academic, and school libraries. The bulk of the technology report is a comprehensive directory of e-book platforms available to libraries; the directory includes information such as target market, type of platform, type of e-book, subject, backgrounds, business model, and vendor websites. The last chapter provided comparative tables to give librarians a quick view so they can compare products against one another in such areas as scope, technical aspects, and business model.

Geller and Roscoe share the experience of selecting, processing, and circulating e-books on e-readers at Lesley University.[44] After a selection process, the Apple iPad and Sony PRS-600 were chosen because a library policy prohibited purchasing materials from Amazon, eliminating the Kindle. The selection of e-book titles conformed to the library's collection development policies for two of their collections, New and Noteworthy and Casual Collections. While patron input was solicited, librarians were ultimately responsible for selection. Details regarding physical processing, circulation, and marketing were discussed, and

there was an informative section titled "Lessons Learned." Among the lessons learned, the library discovered that the iPad was viewed as a versatile piece of equipment beyond its capability as an e-book reader. Patrons preferred the iPad, but some found the convenient size of the Sony reader to be better suited to e-book reading. Librarians at Lesley intend to purchase additional devices to provide library patrons with a broader selection of e-readers.

The complicated e-book workflows at the University of Tennessee library span six library units, requiring collaboration and standardization. Hodge, Manoff, and Watson discuss the issues and challenges faced in scaling up e-book purchasing to accommodate an e-preferred purchasing environment.[45] Of particular complexity is the nonlinear nature of e-resource workflows with respect to licensing, invoicing, establishing access, and ensuring discoverability. The e-book workflow was prone to mistakes that could potentially cause purchased content to go undiscovered. To mitigate these problems, an E-Book Study Group was created to design clear and consistent workflows. Among the study group's recommendations was the implementation of a PDA program for e-books, addition of standardized language to record access terms and restrictions in MARC records, establishment and application of a minimum standard for record quality to MARC record batch-load processes, and elimination of redundant workflows. Workflows were designed, implemented, evaluated, and slowly evolved to ensure that e-books were available to library patrons.

Beisler and Kurt's case study conducted at the University of Nevada, Reno revealed that the complexity of e-book purchase models had grown, that there was an increased demand for e-books, and the workflow spanned multiple departments and was prone to communication breakdowns.[46] In response, the University of Nevada, Reno assembled an interdepartmental task force to develop a new workflow and tracking system to address discoverability and technical service workflow inefficiencies. Significant challenges were communication breakdowns leading to undesirable outcomes, such as uncompleted licenses and improperly cataloged e-books. Workgroups were formed to focus on workflows, discovery, and ERM management issues. The authors outlined a series of questions to be addressed, which could be a helpful resource for any library developing or revising their e-book workflows. A request tracking system was created to provide key functions such as action item alerts, status update alerts, and a general communication tool. The system was used to gather supporting documents used by selection committees, licensing terms, and access information to ensure smooth customized workflows and efficient movement from ordering to discovery. The authors provided a detailed flowchart, visualizing the entire process from request to access. The task force's work enabled the University of Nevada, Reno libraries to create the necessary

workflows for ensuring timely delivery of e-books purchased through a variety of acquisitions methods.

Duan and Grace outlined the e-book acquisitions workflow used at the Open University, a large online university based in the United Kingdom.[47] With more than 200,000 distance education students, the e-book purchasing workflows must be efficient and serve the needs of the large student population. For Duan and Grace, a simpler method for buying e-books was through Big Deals, but they found a "long tail" of unused e-books, and excessive expenditures on "just in case books." The library service developed three tools to facilitate individual e-book purchases: (1) a series of questions to determine how a requested resource will be used, (2) a checklist of preferred options on vendor platforms, and (3) a database with a web form front end to track the purchase requests received from patrons. These tools work together to ensure that the library purchases e-books that match patron expectations.

Streaming Video

Duncan and Peterson describe their process for building a collection of streaming video by licensing content and employing fair use methods at the James Madison University library.[48] The authors detail, with an emphasis on unique and complex issues, all the stages in a streaming video life cycle, including acquisition, access, administration, support, and evaluation. The acquisition stage mirrors other formats; however, due to the cost of streaming video titles, additional care is needed to make justifiable choices. The authors recommend careful evaluation of usage statistics and reliance on faculty requests to inform purchase decisions. Licensing can be complex because of the many options available, such as archival rights, performance rights, etc. In contrast, some publishers include simple use terms on an invoice in lieu of a license. The authors emphasize the importance of knowing what is being purchased and what type of licensing is involved, and tracking data in an ERM.

Accounting and Budgeting

Kirk's book, *Balancing the Books: Accounting for Librarians*, is a much needed guide on accounting topics for librarians. This monograph fills a gap in the literature, in library science education, and in continuing education.[49] The first section discusses responsibilities of acquisition librarians and library directors as typically described in position descriptions. Librarians often find themselves responsible for accounting and budgeting, a need that has grown more acute with declining budgets that require librarians to project costs and track expenditures to prevent overspending. In this environment, training in accounting principles would enable librarians to effectively manage and balance library

budgets. Kirk emphasized that outside of libraries, it is not common to assign responsibility for maintaining budgets to employees who lack formal training in accounting techniques. The author reviewed courses offered by sixty-three ALA-accredited library science programs and suggested that students interested in acquisitions and budgeting enroll in technical services and management courses.

The second section of Kirk's book provides detailed information about principles, philosophy, and practices followed in the field of accounting. Kirk introduces Generally Accepted Accounting Principles (GAAP), a framework of rules, adapted for different organizations, that ensures adherence to the basic elements of financial accounting. Also included is a highly informative discussion on accounting regulations, sample balance sheets, and general guidelines for federal, private, nonprofit, and publicly supported libraries. The chapter focuses on four themes: reporting by program, accounting for collections, accounting for investments, and the financial statement's relationship to budgets.[50] The author introduced the *Statement of Financial Accounting Concepts no. 8*, a framework produced to document the context in which accounting takes place, define concepts, and outline the purpose of accounting.[51] The final chapter of this section draws connections between library services and accounting practices. Financial accounting records provide proof of an organization's financial activity and need to be created and maintained with utmost accuracy and integrity. In-depth coverage of acquisitions ledgers created in ILSs is presented with an emphasis on planning the structure to match reporting requirements.

The final section of Kirk's book discusses budgeting from both a philosophical and practical perspective. The author asks readers to consider that budgeting is a plan for reaching goals, not a restrictive tool for monitoring the library's basic expenses. When creating a budget, librarians should develop a consistent narrative that will concisely justify budget requests in conjunction with a contingency plan for when requests are not fulfilled. Kirk also provides practical guidance on the principles of good budgeting: clarity, accuracy, consistency, and comprehensiveness. She suggests cultivating trend analyses and research to inform current and future budget projections and planning. Kirk includes common library budget challenges in her examples, offers suggestions for managing inflation, and poses considerations for cancellations.

Albitz and Brennan's book chapter, "Budgeting for E-Books," published in *Building and Managing E-Book Collections: A How-To-Do-It Manual for Librarians,* addresses the complex nature of purchasing models and the challenge of allocating funds to support PDA, subscriptions, one-time purchases, Big Deal e-book packages, and single firm order acquisitions.[52] The authors make an interesting observation that unlike e-journal options, which started out as a free add-on feature to print subscriptions, e-books were always a separately sold product. Like e-journals, e-books are often sold with some type of continuing fee, as in the case of a subscription Big Deal for e-books, or a maintenance/platform fee for e-books purchased outright. Clearly, this diverges from print monograph acquisitions budgeting techniques, where an expense occurs once and the transaction is complete. The chapter discusses budgeting for PDA and ways to control costs, such as limiting the number of bibliographic records loaded into to the library catalog, placing cost limits, and implementing some form of librarian mediation of purchases. The chapter concludes with a brief discussion centered on strategies for funding e-book acquisitions, noting that while e-books have higher initial costs and represent a larger drain on collection budgets, studies have found that processing costs and staff time are reduced when moving to e-book acquisitions.

Licensing

Ashmore, Grogg, and Weddle's book, *The Librarian's Guide to Negotiation,* is a practical guide to developing negotiation skills.[53] This book can help bridge the gap between what is covered in library school and the skills expected of an acquisitions professional. The first chapter introduces librarians to language used in business negotiations; techniques from well-known books such as *The Power of Nice, Getting to Yes,* and *Start with No* are reviewed and translated into terms relevant to librarians. The authors draw from the expertise of librarians, vendors, publishers, and consortia managers to compile practical advice, helpful checklists, and pointers. The experts urge librarians to address any anxieties, which likely emanate from the unknown, by reading library literature and seeking continuing education and mentorships.

Negotiation strategies to use in an economic downturn are outlined, including when it is appropriate to play hardball and how to do so in a respectful and productive manner. The concept of ordered flexibility, that is, knowing when and how to make a decisive action, is discussed as a tool for improving negotiation skills. Chapter 5 is written for librarians who negotiate with government officials for budget allocations, but many of the concepts can be applied broadly. Negotiating from a position of strength is important, and the authors note that strength is developed through comprehensive preparation. Strength is also built by becoming an active member in the community and being responsive to the community's needs.

The consolidation of publishers and vendors around a few large companies further exacerbates the challenges to successful negotiation for library materials. The authors remind librarians that they are not necessarily in weak negotiating positions simply because they are acquiring unique items, as academic libraries are the core customer base for

many publishers. Chapter 7 provides strategies for negotiating in this environment, where Big Deals and monopoly-like competitive structures dominate. The authors point out that Big Deals offer an opportunity for strong negotiation, especially when done in collaboration with a consortium. While Big Deals may seem like a fixture, the authors cite the rejection of Big Deal packages by Cornell libraries and the Triangle Research Library Network consortium, and note that in a Big Deal negotiation, one of the most powerful tools is the ability to walk away.

Chapter 8 introduces the idea that negotiation does not end when a deal is finalized, and continues through into the management of acquired e-resources. This negotiation occurs internally between library staff members. Collins advises, "negotiation often manifests itself as management of political nuance, perception, and expectation with the organization."[54] In a similar vein to chapter 8, chapter 9 reminds readers that negotiations occur even in situations that do not involve financial resources. When using open source products, local expertise is often needed for successful implementation. This chapter covers negotiations for the costs associated with free resources such as implementation, maintenance, processing, and marketing. The authors also define open source and Open Access, and launch into an interesting discussion on negotiation in open communities.

Albitz and Brennan's book chapter, "Licensing of E-books," published in *Building and Managing E-Book Collections: A How-To-Do-It Manual for Librarians*, provides an overview of licensing issues of particular concern for e-books, such as platform choice, license types (i.e., unlimited user, nonlinear lending, single user), ILL terms, use terms, and preservation.[55] The section on permitted uses, which differ from those for print books, is particularly helpful in understanding how copyright law applies to e-books. For example, the first sale doctrine does not apply equally to e-books, and may often be compromised by the terms of license agreements.[56] The authors argue that in addition to legal issues, contract law limiting permitted uses is common because of the publisher concerns about piracy and a fundamental shift toward defining book chapters as the commodity.

The Primary Research Group regularly publishes surveys of topics relevant to acquisitions librarianship. The *Survey of Scholarly Journal Licensing and Acquisition Practices* published in 2012 presents a thorough account of questions and responses covering topics ranging from Big Deals to pricing and Open Access.[57] The report details the research method and questionnaire and discusses the characteristics of the sample. The *Survey of Library Database Licensing Practices*, 2012 and 2013 editions, follows the same pattern of presenting detailed information on the research method, questionnaire and sample characteristics as well as providing valuable insight into database licensing practices.[58] Additionally, the report provides information on mobile device use, staff time dedicated to database management, and many other relevant topics.

Acquisitions and Electronic Resource Management Tools

ERM Systems

ERMs and how they can improve the e-resources workflow continue to be popular topics at the annual Charleston Conference. Appleton and Reagan investigated methods for streamlining workflows, with particular attention to using an automated alert system to improve communication and transparency in e-resources management.[59] England, Fu, and Miller described their use of Six Sigma, a business process management solution.[60] Six Sigma is based on a statistical method used to evaluate a process and make data driven improvements in quality by reducing the number of errors.[61] By applying concepts from Six Sigma, the authors enabled small workflow adjustments resulting in improved patron satisfaction. Similarly, Brett, Castro, and Vacek discussed the importance of improving communication between various stakeholders in the e-resource acquisitions workflow through the use of a web portal.[62] Langhurst, Marien, and Schmidt discussed the important work that occurs after an ERM implementation.[63] The authors reported how the University of Notre Dame handled the post-development steps needed to align workflows with CORAL, an open source product created by Notre Dame's Hesburgh Libraries.[64] For example, cataloging workflows fell outside of CORAL, prompting the development of an area of CORAL for tracking MARC record loads.

NASIG annual conferences featured several presentations discussing ERM implementations.[65] England conducted a survey and found that librarians were managing local administrative data such as FTEs, internal contacts, and IP addresses with shared drives on computer networks. The presentation continued with a description of how England adapted an existing ERM record to store administrative data. McQuillan reported on the ERM Data Standards in Best Practices Review Steering Committee, a subgroup of the National Information Standards Organization (NISO). The author noted several areas that were under review, such as link resolvers and knowledge bases, work, manifestations, and access points; cost and usage related data license terms; and data exchange using institutional identifiers.[66] McQuillan then discussed standards and other issues for each of the aforementioned categories; a paper outlining final recommendations can be obtained on the NISO webpage.[67] Imre, Hartnett, and Hiatt presented CORAL implementations from three different university libraries. Each author discussed the processes undertaken at their library, the selection, implementation, and future developments. What is most

interesting about this paper is that the three implementations followed different paths, yet accomplished similar goals.

Hartnett et al. presented their experiences with implementing e-resource management systems; they detail two failed attempts and their success with CORAL.[68] Texas A&M (TAMU) is the flagship institution in a group of fourteen campuses, which required a system that could handle consortial purchases in addition to their own acquisitions. The library had intended to implement a commercial system, but the implementation team failed to make measurable progress and the ERM subscription was canceled. TAMU's next choice was the Gold Rush system created by the Colorado Alliance of Research Libraries, which lacked the robust features necessary to manage large collections. At this point, the library decided to follow a thorough selection process to avoid investing additional time in systems that failed to meet expectations. As the selection team reviewed the literature, implementation case studies, and vendor supplied data, CORAL quickly rose to the top of the pool. The authors then discuss the implementation of each module. The implementation team organized training sessions for everyone who would be required to use CORAL. Lastly, the authors sent a short survey to training session attendees and of eleven responses, five of the six respondents who have used CORAL since the training session are involved with e-resource management. The authors discovered that the intended target audience outside of e-resource management staff was not reached.

Jensen described the creation of an ERM system based on Google Sites implemented at the University of Alaska Fairbanks (UAF).[69] The library used many technologies to manage their e-resources collections, including a commercial ERM, a trouble ticket system, a web-based database, spreadsheets, a network computer system, and email. The UAF library lacked a dedicated position for e-resources management and required a management system that was easy to access and capable of serving multiple library departments. The library first attempted to implement a system created using Drupal and found that the searching capabilities were not as robust as desired and various pieces of data generated during the e-resources life cycle were not integrated. The library then chose to move to a major vendor system, but discovered that the ERM was too difficult to use and lacked some of the necessary features. This led to the development of an ERM on Google Sites that met their e-resources management needs. The author discussed Google Sites' helpful features, including renewal notifications via Google Calendar, the ability to see site changes by looking at the revision history and create websites to share administrative information and document workflows, and most importantly, the presence of a Google search box. The Google Sites solution is also useful because it does not require specialized knowledge of computer programming to manage the ERM.

Purchase Request Management Systems

Many papers in the acquisitions field focus on the early stages of the acquisitions process, such as licensing and purchase models. Studies that address the final stages of the acquisitions process typically focus on usage statistics as a means of analyzing the effectiveness of purchasing models such as PDA or Big Deal packages. McMullen and Gray focus on the implementation of a service designed to inform requestors when the requested items are available in the library.[70] The library had complicated, time-consuming, and antiquated protocols for generating PDF lists of new acquisitions from the ILS. Through a process of iterative changes, the library developed a current awareness service on an RSS feed platform. While this advancement was helpful, its major pitfall was the technology barrier that it presented to those unfamiliar with RSS feeds. The finalized service would allow requesters to track their requests through a department website dynamically generated by feeds from the ILS. This system requires catalogers to enter a four letter code from an existing system used to tie fund codes to academic departments in the ILS into a 945 MARC field, which is a field that can be locally customized. When placed in the 945 field, the codes created a unique text string used by a script designed to pull information from the ILS through the Z39.50 connections and publish it to a website. This system allows requesters to track their books after the order was placed even if a book falls outside the discipline normally requested. This is accomplished by using the 945 field to note the requester's department, rather than the book's subject. The authors found that the current awareness service was well used, as indicated by Google Analytics and anecdotal evidence.

Downey described Kent State University Libraries' implementation of a locally designed system, named Pre-ILS, to manage the steps of an e-resource workflow that occurs before a record is created in an ILS.[71] One of the complaints Pre-ILS addressed was the avalanche of emails required to move an e-resource request through the trial to access process. The library had used many disparate systems including the ERM, ILS, spreadsheets, and email, and cobbled them together to manage the workflow in a way that was prone to communication breakdowns. Several meetings were convened to develop working requirements for the system and to plan the development and implementation of Pre-ILS. The system is intended to be used by selectors as a place to request pricing information, review resources, store trial feedback, and approve purchases. As a resource moves through the acquisition stages, each stakeholder is notified of progress. An interesting feature of the system is the ability to offer five different levels of user access, from level one (view-only access open to the public) through level five (access reserved for decision-makers). Pre-ILS was not developed on an open platform and cannot be easily shared.

Yang and Hung, affiliated with the Department of Information Management at Nanhua University, Taiwan, developed an innovative approach to gathering patrons' book recommendations through data mining and algorithm development.[72] The authors noted in their technical paper that book recommendation workflows are often complicated and require significant staff intervention causing long waits for books, and are often not well marketed to patrons. The authors created an innovative electronic system that uses text mining of user search input from failed searches in the library catalog and generates a purchase list from the mined data. The recommended prototype was well tested and appeared to be an innovative method for mining patron search queries to create a recommend for purchase list.

Conclusion

Purchasing models saw rapid changes in the period from 2012–13, generating several publications. Interest in firm order purchasing and approval plans waned during this review period, and was refocused on consortial purchases, PDA, and DDA. Consortial purchasing is a popular topic in the literature covering Big Deal packages and PDA/DDA pilot programs. While some libraries, such as the University of California Libraries and Southern Illinois University-Carbondale have cancelled Big Deals, ARL surveys have revealed the deep and persistent presence of Big Deals in the library market, especially in the consortial arena. Interest in PDA and DDA continues to grow and implementation reports are widely shared, creating a robust body of published literature. As this literature review has shown, studies have been conducted to determine the utility of PDA. Many have found that PDA offers a way for libraries to allow direct input from patrons in the selection of materials, which often has a positive impact on circulation rates and has been used as a tool to control spending. PDA models have evolved to include e-book readers such as the Amazon Kindle. The literature has also demonstrated a growing interest in Evidence-Based Selection as an alternative to PDA, and PPV programs such as ReadCube and "Get it Now" as alternatives to Big Deal packages.

Advancing technology has proven to be a challenge and opportunity in the acquisitions field. The challenges include: increasingly complicated workflows, licensing challenges, rising interest in e-books on dedicated e-reader devices, growing reliance on streaming media, and an expanding array of purchase models for electronic content. These challenges have opened opportunities to leverage technology to improve workflows and ultimately patron satisfaction through the acquisition and delivery of needed research materials. Innovative solutions include the pre-ILS system developed at Kent State University, an ERM system built on Google Sites, and a book acquisitions recommendation model developed at the University of Taiwan.[73] CORAL implementations are commonly discussed in the literature and 2012–13 saw growing interest in improving CORAL and improving workflows with the goal of increasing efficiency and in turn improving patron satisfaction. An underlying theme of library technology is the basic need to improve communications, whether it be between library employees, vendors, or patrons. New purchasing models require new workflows and new workflows necessitate the development of workflow management tools. This literature review has shown how these areas are interconnected and growing.

References and Notes

1. Barbara S. Dunham and Trisha L. Davis, "Literature of Acquisitions Review, 1996–2003," *Library Resources & Technical Services* 52, no. 4 (2008): 238–53, http://dx.doi.org/10.5860/lrts.52n4.238; Barbara S. Dunham and Trisha L. Davis, "Literature of Acquisitions Review, 2004–7," *Library Resources & Technical Services* 53, no. 4 (2011): 231–42, https://journals.ala.org/lrts/article/view/5181/6287; Jeanne Harrell, "Literature of Acquisitions Review, 2008-9," *Library Resources & Technical Services* 56, no. 1 (2012): 4–13, https://journals.ala.org/lrts/article/view/5493/6740; Paul D. Moeller, "Literature of Acquisitions Review, 2010–11," *Library Resources & Technical Services* 57, no. 2 (2013): 87–99, http://dx.doi.org/10.5860/lrts.57n1.87.

2. Moeller, "Literature of Acquisitions Review, 2010–11," 87–99.

3. Tony G. Horava, "Direct Ordering by Subject Librarians in a Vendor System: A Case Study," *Partnership: The Canadian Journal of Library & Information Practice & Research* 7, no. 1 (2012): 1–11, https://journal.lib.uoguelph.ca/index.php/perj/article/view/1723/2472#.VgAO5t9VhBc.

4. Rita M. Cauce, "Where's Professor Watt's Request? Streamlining to a Paperless Acquisitions Workflow," in *Something's Gotta Give: Charleston Conference Proceedings, 2011,* edited by Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012), 446–49.

5. Sarah Forzetting, Gabrielle Wiersma, and Leslie Eager, "Managing E-Book Acquisition: The Coordination of 'P' and 'E' Publication Dates," *Serials Librarian* 62, no. 1–4 (2012): 200–205, http://dx.doi.org/10.1080/0361526X.2012.652921.

6. Gabrielle Wiersma, "What's the Delay? Managing E and P Publication Dates," *Library Collections Acquisitions, & Technical Services* 37, no. 1 (2013): 13–18, http://dx.doi.org/10.1016/j.lcats.2013.03.001.

7. Rob Kairis, "A Subject-Based Shared Approval Plan for Consortia Purchasing of US University Press Books," *Library Collections, Acquisitions, & Technical Services* 36, no. 1 (2012): 30–38, http://dx.doi.org/10.1016/j.lcats.2012.03.002; Wen-Ying Lu and Mary Beth Chambers, "PDA Consortium Style," *Library Resources & Technical Services* 57,

no. 3 (2013): 164–78, http://dx.doi.org/10.5860/lrts.57n3.164; Jodi Shepherd and Marc Langston, "Shared Patron Driven Acquisition of E-Books in the California State University Library Consortium," *Library Collections Acquisitions & Technical Services* 37, no. 1–2 (2013): 34–41, http://dx.doi.org/10.1016/j.lcats.2013.08.001; Jeanne Richardson, "The Arizona Universities Library Consortium Patron-Driven E-Book Model," *Insights: The UKSG Journal* 26, no. 4 (2013): 66–69, http://dx.doi.org/10.1629/2048-7754.26.1.66; Lynn Wiley and Elizabeth Clarage, "Building on Success: Evolving Local and Consortium Purchase-on-Demand Programs," *Interlending & Document Supply* 40, no. 2 (2012): 105–10, http://dx.doi.org/10.1108/02641611211239588; Jill Emery, "The Demand Driven Acquisitions Pilot Project by the Orbis Cascade Alliance: An Interview with Members of the Demand Driven Acquisitions Implementation Team," *Serials Review* 38, no. 2 (2012): 132–36, http://dx.doi.org/10.1016/j.serrev.2012.04.008; Xan Arch et al., "By Popular Demand: Building a Consortial Demand-Driven Program," in *Something's Gotta Give: Charleston Conference Proceedings, 2011,* edited by Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012); Connie Stovall, Leslie O'Brien, and Edward Lener, "Virginia Tech's Participation in ASERL's Cooperative Print Journal Retention Project" in *Something's Gotta Give: Charleston Conference Proceedings, 2011,* edited by Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012).

8. Kate Davis et al., "Shared Patron-Driven Acquisition within a Consortium: The OCUL PDA Pilot," *Serials Review* 38, no. 3 (2012): 183–87, http://dx.doi.org/10.1016/j.serrev.2012.08.002; Tony Horava, "Today and in Perpetuity: A Canadian Consortial Strategy for Owning and Hosting Ebooks," *Journal of Academic Librarianship* 39, no. 5 (2013): 423–28, http://dx.doi.org/10.1016/j.acalib.2013.04.001; Xiaoxia Yao, "Evaluation and Promotion: The Cooperative Purchase Experience of Academic Libraries in China," *Library Collections, Acquisitions, & Technical Services* 36, no. 3 (2012): 97–106, http://dx.doi.org/10.1016/j.lcats.2012.05.003.

9. Sarah Kalikman Lippincott et al., "Librarian, Publisher, and Vendor Perspectives on Consortial E-Book Purchasing: The Experience of the TRLN Beyond Print Summit," *Serials Review* 38, no. 1 (2012): 3–11, http://dx.doi.org/10.1016/j.serrev.2011.12.003; Leslie Button et al., "Reducing Unintentional Duplication: Adventures and Opportunities in Cooperative Collection Development," in *Something's Gotta Give: Charleston Conference Proceedings, 2011,* edited by Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012).

10. George Machovec, "Library Consortia: The Big Picture," *Journal of Library Administration* 53, no. 2–3 (February 2013): 199–208, http://dx.doi.org/10.1080/01930826.2013.853504.

11. Jill Emery and Bonnie Parks, "The Demand Driven Acquisitions Pilot Project by the Orbis Cascade Alliance: An Interview with Members of the Demand Driven Acquisitions Implementation Team," *Serials Review* 38, no. 2 (June 2012): 132–36, http://dx.doi.org/10.1016/j.serrev.2012.04.008.

12. Xan Arch et al., "By Popular Demand."

13. Davis et al., "Shared Patron-Driven Acquisition within a Consortium," 183–87.

14. Kenneth Frazier, "What's the Big Deal?," *Serials Librarian* 48, no. 1 (2005): 49–59, http://dx.doi.org/10.1300/J123v48n01_06; Theodore C. Bergstrom, "Librarians and the Terrible Fix: Economics of the Big Deal," *Serials* 23, no. 2 (2010): 77–82, http://search.proquest.com.jproxy.lib.ecu.edu/docview/753822310?accountid=10639; David Nicholas et al., "What Deep Log Analysis Tells Us about the Impact of Big Deals: Case Study OhioLINK," *Journal of Documentation* 62, no. 4 (2006): 482–508, http://dx.doi.org/10.1108/00220410610673864.

15. Jeffrey N. Gatten and Tom Sanville, "An Orderly Retreat from the Big Deal: Is it Possible for Consortia?," *D-Lib Magazine* 10, no. 10 (2004), www.dlib.org; Rickey D. Best, "Is the "Big Deal" Dead?," *Serials Librarian* 57, no. 4 (2009): 353–63, http://search.proquest.com.jproxy.lib.ecu.edu/docview/57691425?accountid=10639; Nancy J. Gibbs, "Walking Away from the 'Big Deal': Consequences and Achievements," *Serials* 18, no. 2 (2005): 89–93, http://search.proquest.com.jproxy.lib.ecu.edu/docview/57614754?accountid=10639.

16. "UC Libraries Discontinue Taylor & Francis Systemwide Journals License," California Digital Library website, July 3, 2013, accessed December 27, 2015, www.cdlib.org/cdlinfo/2013/07/03/uc-libraries-discontinue-taylor-francis-systemwide-journals-license/; Jonathan Nabe and David C. Fowler, "Leaving the 'Big Deal': Consequences and Next Steps," *Serials Librarian* 62, no. 1–4 (2012): 59–72, http://dx.doi.org/10.1080/0361526X.2012.652524.

17. Mike McGrath, "Fighting Back Against the Big Deals: A Success Story from the UK," *Interlending & Document Supply* 40, no. 4 (2012): 178–86, http://dx.doi.org/10.1108/02641611211283831.

18. Karla L. Strieb and Julia C. Blixrud, "The State of Large-Publisher Bundles in 2012," *Research Library Issues: A Report from ARL, CNI, and SPARC* no. 282 (2013): 13–20, http://publications.arl.org/rli282.

19. Beth Ashmore, Jill E. Grogg, and Jeff Weddle, *The Librarian's Guide to Negotiation Winning Strategies for the Digital Age* (Medford: Information Today, 2012), 144.

20. Julia Proctor, "Avoiding Ebook 'Big Deals': Alternatives to Ebook Backlists," *New Library World* 114, no. 7–8 (2013): 301–7, http://dx.doi.org/10.1108/NLW-02-2013-0018.

21. Robert W Boissy et al., "Is the 'Big Deal' Dying?," *Serials Review* 38, no. 1 (2012): 36–45, http://dx.doi.org/10.1016/j.serrev.2011.12.012.

22. Ibid., 42.

23. Ibid., 41.

24. Terry Bucknell, "Buying by the Bucketful: A Comparative Study of E-Book Acquisition Strategies," *Insights: The UKSG Journal* 25, no. 1 (2012): 51–60, http://dx.doi.org/10.1629/2048-7754.25.1.51.

25. Mark England and Rick Anderson, "Patron-Driven Acquisition of Journal Articles Using ReadCube at the University of Utah," *Insights: The UKSG Journal* 26, no. 3 (2013): 267–71, http://dx.doi.org/10.1629/2048-7754.77.

26. Debbi Dinkins, "Individual Title Requests in PDA Collections: A Small University Library's Experience," *College & Research Libraries News* 73, no. 5 (2012): 249, http://crln.acrl.org/content/73/5/249.full.pdf+html; Rebecca Schroeder, "When Patrons Call the Shots: Patron-Driven Acquisition at Brigham Young University," *Collection Building* 31, no. 1 (2012): 11–14, http://dx.doi.org/10.1108/01604951211199128.

27. Suzanne M. Ward, *Guide to Implementing and Managing Patron-Driven Acquisitions* (Chicago: ALCTS, 2012).

28. Dee Ann Allison, *The Patron-Driven Library: A Practical Guide for Managing Collections and Services in the Digital Age* (Oxford: Chandos, 2013).

29. Karen S. Fischer et al., "Give 'Em what they Want: A One-Year Study of Unmediated Patron-Driven Acquisition of E-Books," *College & Research Libraries* 73, no. 5 (2012): 469-492, http://dx.doi.org/10.5860/crl-297.

30. David McCaslin, "Collection Building through Patrons: Caltech Library's Kindle Program," *Collection Management* 38, no. 3 (2013): 172–91, http://dx.doi.org/10.1080/01462679.2013.792308.

31. Ibid., 83.

32. David C. Tyler et al., "Patron-Driven Acquisition and Circulation at an Academic Library: Interaction Effects and Circulation Performance of Print Books Acquired Via Librarians' Orders, Approval Plans, and Patrons' Interlibrary Loan Requests," *Collection Management* 38, no. 1 (2013): 3–32, http://dx.doi.org/10.1080/01462679.2012.730494.

33. William H. Walters, "Patron-Driven Acquisition and the Educational Mission of the Academic Library," *Library Resources & Technical Services* 56, no. 3 (2012): 199–213, http://dx.doi.org/10.5860/lrts.56n3.199.

34. Merinda McLure and Amy Hoseth, "Patron-Driven E-Book use and Users' E-Book Perceptions: A Snapshot," *Collection Building* 31, no. 4 (2012): 136–47, http://dx.doi.org/10.1108/01604951211274043.

35. Ibid., 139.

36. Bucknell, "Buying by the Bucketful," 51-60.

37. Nathan Hosburgh, "Getting the Most Out of Pay-Per-View: A Feasibility Study and Discussion of Mediated and Unmediated Options," *Journal of Electronic Resources Librarianship* 24, no. 3 (2012): 204, http://dx.doi.org/10.1080/1941126X.2012.706112.

38. England and Anderson, "Patron-Driven Acquisition of Journal Articles," 267–71.

39. Steven W. Sowards, "Library-Publisher Experimentation and Partnership in Alternative Models for Journal Content," *The Serials Librarian* 65, no. 3–4 (2013): 309, http://dx.doi.org/10.1080/0361526X.2013.837858.

40. Erin S. Fisher, Lisa Kurt, and Sarah Gardner, "Exploring Patron-Driven Access Models for E-journals and E-books," *Serials Librarian* 62, no. 1–4 (2012): 164–68, http://dx.doi.org/10.1080/0361526X.2012.652913; Heather L. Brown, "Pay-Per-View in Interlibrary Loan: A Case Study," *Journal of the Medical Library Association: JMLA* 100, no. 2 (2012): 98–103, http://dx.doi.org/10.3163/1536-5050.100.2.007; Maureen Weicher and Tia Xiao Zhang, "Unbundling the 'Big Deal' with Pay-Per-View of E-Journal Articles," *Serials Librarian* 63, no. 1 (2012): 28, http://dx.doi.org/10.1080/0361526X.2012.688167; Allen Powell, "Navigating the New Norm: Vendor, Publisher, and Librarian Strategies to Cope with the Changing Information Industry," *Journal of Library Administration* 52, no. 5 (2012): 370, http://dx.doi.org/10.1080/01930826.2012.701112.

41. William H. Walters, "E-Books in Academic Libraries: Challenges for Acquisition and Collection Management," *portal: Libraries & the Academy* 13, no. 2 (2013): 187–211, http://dx.doi.org/10.1353/pla.2013.0012.

42. Richard Kaplan, ed., *Building and Maintaining E-Book Collections: A How-To-Do-It Manual for Librarians* (Chicago: Neal-Schuman, 2012).

43. Mirela Roncevic, "E-book Platforms for Libraries," *Library Technology Reports* 49, no. 3 (2013): 5–43, https://journals.ala.org/ltr/issue/view/130.

44. Marilyn Geller and Linda Roscoe, "Accessing and Circulating E-Books with E-Readers—Lesley University," in *Building and Managing E-Book Collections: A How-To-Do-It Manual for Librarians*, edited by Richard Kaplan (Chicago: Neal-Schuman, 2012), 171–76.

45. Valeria Hodge, Maribeth Manoff, and Gail Watson, "Providing Access to E-Books and E-Book Collections: Struggles and Solutions," *Serials Librarian* 64, no. 1–4 (2013): 200–205, http://dx.doi.org/10.1080/0361526X.2013.760411.

46. Amalia Beisler and Lisa Kurt, "E-Book Workflow from Inquiry to Access: Facing the Challenges to Implementing E-Book Access at the University of Nevada, Reno," *Collaborative Librarianship* 4, no. 3 (2012): 96–116, http://collaborativelibrarianship.org/index.php/jocl/article/viewArticle/188.

47. Sharon Duan and Claire Grace, "E-Book Purchasing: The Open University's Experience," *Insights: The UKSG Journal* 26, no. 3 (2013): 250–55, http://dx.doi.org/10.1629/2048-7754.100.

48. Cheri Duncan and Erika Peterson, "You Ought to be in Pictures: Bringing Streaming Video to Your Library," in *Something's Gotta Give: Charleston Conference Proceedings, 2011*, ed. Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012).

49. Rachel A. Kirk, *Balancing the Books: Accounting for*

Librarians (Santa Barbara: Libraries Unlimited, 2013).

50. Ibid., 45.

51. Financial Accounting Standards Board, "Statement of Financial Accounting Concepts No. 8," accessed September 19, 2015, www.fasb.org/jsp/FASB/Document_C/DocumentPage ?cid=1176157498129&acceptedDisclaimer=true.

52. Becky Albitz and David Brennan, "Budgeting for E-Books," in *Building and Managing E-Book Collections: A How-To-Do-It Manual for Librarians*, edited by Richard Kaplan (Chicago: Neal-Schuman, 2012), 85–94.

53. Beth Ashmore, Jill E. Grogg, and Jeff Weddle, *The Librarian's Guide to Negotiation Winning Strategies for the Digital Age* (Medford, NJ: Information Today, 2012).

54. Ibid., 168.

55. Becky Albitz and David Brennan, "Licensing of E-Books," in *Building and Managing E-Book Collections: A How-To-Do-It Manual for Librarians*, edited by Richard Kaplan (Chicago: Neal-Schuman, 2012), 75–84.

56. Rachel Ann Geist, "A 'License to Read': The Effect of E-books on Publishers, Libraries, and the First Sale Doctrine," *Idea* 52, no. 1 (2012): 63–100, http://ipmall.info/hosted _resources/IDEA/idea-vol52-no1-geist.pdf.

57. Primary Research Group, *Survey of Scholarly Journal Licensing and Acquisition Practices* (New York: Primary Research Group, 2012).

58. Primary Research Group, *Survey of Library Database Licensing Practices* (New York: Primary Research Group, 2012); Primary Research Group, *Survey of Library Database Licensing Practices* (New York: Primary Research Group, 2013).

59. Betsy Appleton and Shannon Regan, "Going Beyond Electronic Resource Management System (ERMS) Implementation: ERMS-Focused Work flows and Communications," in *Something's Gotta Give: Charleston Conference Proceedings, 2011*, edited by Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012).

60. Lenore England, Stephen Miller, and Li Fu, "Optimizing Your ERM: Application of Business Process Management to Operations," in *Something's Gotta Give: Charleston Conference Proceedings, 2011*, edited by Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012).

61. More information about Six Sigma can be found at http://asq .org/learn-about-quality/six-sigma/tools.html.

62. Kelsey Brett, Jeanne M. Castro, and Rachel Vacek, "How to Build a Better Mousetrap: Developing an Easy, Functional ERM" in *Accentuate the Positive!: Charleston Conference Proceedings, 2012*, edited by Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012).

63. Andrea Langhurst, Stacey Marien, and Kari Schmidt, "Beyond Implementation: Making Your ERMS Work for You" in *Accentuate the Positive!: Charleston Conference*

Proceedings, 2012, edited by Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012).

64. More information about CORAL can be found at http://coral -erm.org/about.

65. Bob McQuillan, "Gateway to Improving ERM System Deliverables: NISO ERM Data Standards and Best Practices Review," *Serials Librarian* 62, no. 1–4 (2012): 112–24, http:// dx.doi.org/10.1080/0361526X.2012.652482; Deborah England, "We have our ERM System, it's Implemented: Why Am I Still Going here and there to Get the Information I Need?" *Serials Librarian* 64, no. 1–4 (2013): 111, http://dx.doi.org /10.1080/0361526X.2013.760148; Andrea Imre, Eric Hartnett, and C. Derrik Hiatt, "CORAL: Implementing an Open-Source ERM System," *Serials Librarian* 64, no. 1–4 (2013): 224, http://dx.doi.org/10.1080/0361526X.2013.760414.

66. McQuillan, "Gateway to Improving ERM System Deliverables,"114.

67. NISO ERM Data Standards and Best Practices Review Steering Committee, *Making Good on the Promise of ERM: A Standards and Best Practices Discussion Paper* (Baltimore: NISO, 2012), www.niso.org/apps/group_public/download .php/7946/Making_Good_on_the_Promise_of_ERM.pdf.

68. Eric Hartnett et al., "Charting a Course Through CORAL: Texas A&M University Libraries' Experience Implementing an Open-Source Electronic Resources Management System," *Journal of Electronic Resources Librarianship* 25, no. 1 (2013): 16–38, http://dx.doi.org/10.1080/194112 6X.2013.760402.

69. Karen Jensen, "Managing Library Electronic Resources Using Google Sites," *Journal of Electronic Resources Librarianship* 25, no. 2 (2013): 115–23, http://dx.doi.org/10.1080/19 41126X.2013.785289.

70. Anthony McMullen and Barry Gray, "From Static to Dynamic: Using the OPAC to Generate Real-Time Lists of Departmental Acquisitions for Library Current Awareness Service," *Library Hi Tech* 30, no. 4 (2012): 673–82, http://dx.doi .org/10.1108/07378831211285121.

71. Kay Downey, "Kent State University Libraries Develops a New System for Resource Selection," in *Something's Gotta Give: Charleston Conference Proceedings, 2011*, edited by Beth R. Bernhardt, Leah H. Hinds, and Katina P. Strauch (West Lafayette: Purdue University Press, 2012).

72. Shih-Ting Yang and Ming-Chien Hung, "A Model for Book Inquiry History Analysis and Book-Acquisition Recommendation of Libraries," *Library Collections, Acquisitions, & Technical Services* 36, no. 3 (2012): 127–42, http://dx.doi .org/10.1016/j.lcats.2012.05.001.

73. Downey, "Kent State University Libraries Develops a New System for Resource Selection," 157–62; Shih-Ting Yang and Ming-Chien Hung, "A Model for Book Inquiry History Analysis and Book-Acquisition Recommendation of Libraries," 127–42.

Notes on Operations
Customizing an Open Source Discovery Layer at East Carolina University Libraries

The Cataloger's Role in Developing a Replacement for a Traditional Online Catalog

Marlena Barber, Christopher Holden, and Janet L. Mayo

Open source discovery layers offer the ability to extensively customize every aspect of the search experience for a local user population. However, discovery layers have primarily been discussed in the professional literature in terms of the installation or configuration process. In this paper, three catalogers present a case study of an open source discovery layer implementation with a focus on the problems and solutions from the cataloging perspective.

East Carolina University (ECU) is the third largest school in the North Carolina system, with almost 27,000 students.[1] As of January 27, 2015, the university "employs nearly 2,050 full-time, part-time, and temporary faculty."[2] These faculty and students are supported by more than 3,500 staff members.[3] In fall of 2014, 27,511 students were served by ECU Libraries, a system that consists of a main library (Joyner Library), Laupus Health Sciences Library (Laupus), and the music library.[4] These three very different libraries have a wide range of discovery challenges, from multiple classification schemes and subject vocabularies to differences in desired MARC fields in a brief record display to varying requirements for metadata granularity depending on the subject areas or collections. Such a wide array of needs requires a sophisticated and robust discovery tool that offers maximum potential for users to find and access the information they seek.

ECU Libraries has been using the e-Library software from SirsiDynix's Symphony line of products for its OPAC since 2009. A 2013 internal assessment of e-Library identified problems that fell into three broad areas: poor relevancy ranking, an inflexible user interface, and cumbersome functionality.[5] Because of the proprietary software's limitations, local customization and improvement of the relevancy of search results ranged from difficult to impossible to implement. Desired features, such as customized bibliographic displays for each library or highlighting search terms in record displays, could not be accommodated. The e-Library interface also failed to consistently index certain MARC elements; for example, the music library found that uniform titles in the 240 field were indexed differently from the same uniform titles found in the 700 field.

To supplement the traditional OPAC, the libraries were tasked in 2009 with reviewing and recommending a discovery tool.[6] The decision was made to implement ProQuest's Summon product. While e-Library is used for traditional OPAC tasks such as title or call number browsing, placing holds, and tracking course

Marlena Barber (barberm@ecu.edu) is Collection Services & Metadata Librarian at Laupus Health Sciences Library, East Carolina University. **Christopher Holden** (holdenc@ecu.edu) is Assistant Music Librarian, Joyner Library, East Carolina University. **Janet L. Mayo** (mayoj@ecu.edu) is Head, General Cataloging, Joyner Library, East Carolina University.

reserves, Summon serves as a broader, web-scale discovery tool, allowing users to receive results from all of the ECU Libraries' resources through a single search.

However, because of the product's proprietary nature, it was often difficult to understand why certain search results were elevated to the top. In particular, Summon's interface tended to rank electronic resources higher than physical items, a trait that did not work very well for certain types of searches. Additionally, while Summon's MARC mapping and faceting was customizable to a certain extent, some desired features remained unavailable. For example, because Summon allowed only a one-to-one mapping between an item and a format facet, librarians were frustrated that certain items, such as digital audio files, could be mapped only as an "electronic resource" or as a "sound recording," but not both. Summon's display included several undesirable features that were impossible to suppress, such as the display of foreign language subject headings for faceted searches.

Because both e-Library and Summon are proprietary software, customization was limited. ECU Libraries needed an interface with fewer constraints and increased indexing flexibility to meet user needs for all three campus libraries. The libraries decided to pursue open source discovery layer options as a solution to this problem. Initially, planning and development for a new discovery layer began with VuFind, one such open source option, but the libraries later opted to develop another open source catalog, Blacklight.[7]

Literature Review

Because discovery layers are a relatively recent phenomenon, the literature about them focuses on certain definitions of what constitutes a "discovery layer," and discuss libraries' experiences in setting them up. Moore and Greene point out that many terms have been used interchangeably to discuss these new catalog interfaces, including "next-generation catalog," or NGC, "discovery layer," and "web-scale discovery tool." They differentiate between the three terms as follows:

1. Next-generation catalogs are public search interfaces that integrate Web 2.0 technologies such as RSS feeds and social media;
2. Discovery layers are search interfaces that specifically exist apart from the traditional ILS, and may incorporate other discoverable content beyond MARC records, such as digital collections and institutional repositories;
3. Web-scale discovery tools draw from a central index of vendor and publisher databases of scholarly articles, and the local institution's MARC records. This allows a user to search across practically every resource to which the library has access.[8]

A few papers have discussed the failings of the traditional OPAC interface and the problems that discovery layers are intended to solve. Ballard and Blaine note that most OPAC interfaces are "often not intuitive and are inconsistent with well-established user interface conventions" and lack good relevancy rankings.[9] Sadler points out that relying on an OPAC interface for discovery purposes removes the development of the interface from the feedback of local users, leading to a lack of customization for individual libraries.[10] Ho and Horne-Poppe posit that traditional OPACs frustrate users with their "un-intuitive library catalog interfaces that can't handle searches that start with articles, that don't enable easy discovery of similar items, and that don't allow for interaction with the library records."[11]

Open source discovery layers have been developed partially as a way to resolve these problems with the OPAC, and partially as a means to provide an alternative to the proprietary commercial products that are so difficult to customize. Open source products allow a library to extensively customize the search experience while avoiding upfront costs, though programming such software does involve an extensive amount of manpower and technical skills. Two such products, VuFind and Blacklight, are open source discovery interfaces meant to overlay a traditional OPAC. VuFind was developed by Villanova University as an open source alternative to the Endeca discovery layer, a commercial product implemented at North Carolina State University.[12] Blacklight development was initiated by the University of Virginia. While both VuFind and Blacklight use SolrMarc, a version of the search utility Apache Solr that specifically works with MARC, to index MARC fields, VuFind uses the PHP scripting language to structure and present the web interface, and Blacklight uses an application framework known as Ruby on Rails for the same purpose.[13] Nagy and Katz wrote that VuFind emphasizes easy installation and an "out-of-the-box" approach, while Blacklight allows for more extensive customization for those working in a Ruby on Rails environment, giving libraries the ability to provide a diversity of displays for different items, facets, and branch libraries.[14] Nagy and Katz assert that Google and other search engines have changed the discovery landscape and user expectations, and the relational database structure of the OPAC does not allow for the speed or relevancy that Solr indexes can provide.[15] They make the case that VuFind's capability to offer post-search faceting, synonym matching, stemmed searches, and spelling correction position it as a marked improvement over the traditional OPAC.

Ballard and Blaine also believe that faceting is a major advantage of discovery layers, pointing out that users were fifteen times more likely to refine their search in a faceted environment than a traditional OPAC in which all refinements had to be entered at the time of the search.[16] Sadler appreciated the "tremendous advantages in terms

of flexibility in defining what is indexed and how it is searched" within Blacklight and points out that discovery layers can be customized even for branch libraries within a single institution.[17] She mentions a customized interface for the University of Virginia music library, which included an index of instrumentation, and the ability to search for musical works by "date of composition." La Barre mentions the importance of analyzing and customizing facets before the implementation of a discovery layer.[18] Leebaw reports that feedback from users on their VuFind implementation revealed that users are very interested in accessing specific collections such as films, sound recordings, and reference materials, using facets.[19]

Many authors praised the flexibility of open source discovery layers, which provide the ability to devise customized solutions to local problems. Leebaw writes, "Being able to immediately incorporate feedback is one of the best advantages of implementing an open source overlay."[20] This kind of customization can create local solutions for specific groups of researchers and library users. For example, Sadler points out that the University of Virginia created a music controller in Blacklight that incorporates slightly different relevancy rankings and facets than the standard search controller.[21] She writes that this is a significant departure from the traditional OPAC, "which assumes that there can be a single interface that will be good enough for most users, and that this interface must be managed centrally."[22]

Several papers discussed the impact that discovery layers can have on music materials in particular. Snyder conducted a survey of School of Music faculty and graduate students at the University of Chicago following that institution's implementation of the AquaBrowser faceted catalog. Survey respondents appreciated the format facet in particular but that faceting for categories such as "genre," "geographic region," and "era" was inconsistent and confusing.[23] The Music Library Association's report on discovery requirements indicated that some data, such as the "date of composition," have been used inconsistently and would be difficult to take advantage of in a discovery environment.[24] The report touches on several important requirements for the discovery of musical materials, such as the ability to map multiple formats to a single item, the continued use of text strings for subject headings and uniform titles, and the ability to implement authority control and take advantage of the data in authority records.

There are many comparisons of discovery layers to the Google search engine, with writers pointing out that Google's ubiquity in the discovery landscape has led to a major shift in expectations among users. Katz and Nagy note that users now prefer a "self-service-oriented approach to searching."[25] However, Ballard and Blaine worry that the "Google-esque" single keyword search query box may not be enough for many academic libraries, and a link to a more advanced search page may still be necessary.[26] Leebaw agrees that the "Googlization" of library resources does not always advance the library's discovery goals and points out that reference librarians raised some concerns about discovery overlays.[27] This relates to the trend noted by Moore and Greene that public services librarians are often involved in the process of setting up the discovery layer, which makes them more likely to buy into the completed product.[28]

A recurring caveat throughout the literature is that developing a local open source discovery layer requires significant amounts of time and in-house programming knowledge. While both VuFind and Blacklight have basic default configuration settings, taking advantage of local customizations and other flexibilities requires a great deal of technical skill. Skinner summarizes GIL-Find, a Georgia-wide implementation of VuFind, and notes that "it would be unlikely for more than a few of the largest universities in the Georgia system to have the staffing or the expertise to adapt an open source product such as VuFind at the local level."[29] Emanuel goes so far as to wonder if next-generation catalogs will create a new "digital divide" between those libraries that can afford to implement an open source product and those that cannot.[30]

As detailed later in this paper, the university's Blacklight implementation was contingent on the presence of strong, knowledgeable programmers working for the libraries. Throughout the literature, there is extensive scholarship on the history of next-generation catalogs and aspects to consider when choosing one. However, there is little to no scholarship on the process of setting up local customizations for a next-generation catalog. The intellectual effort to set up the indexing and display of various MARC fields is a significant investment; additionally, the programming knowledge needed to implement these customizations is also important. Each library's open source installation will differ according to local needs, and perhaps this factor has led to a lack of scholarship describing the process.

Mapping

From the beginning of the project, it was agreed by all parties at the ECU Libraries that intense customization of the discovery layer was needed to specifically address research issues unique to ECU and to fix problems that had been identified with previous search tools. Initially, planning and development for a new discovery layer began with VuFind, an open source option. Throughout the development of VuFind, comments from the university libraries' Discovery Advisory Board (DAB) were considered and addressed by the catalogers and members of the Application and Discovery Services (ADS) department. The DAB is a library-wide committee that considers all the libraries' discovery

initiatives and includes both public services and technical services faculty and staff from all three libraries. The catalogers and members of the ADS department began to reconfigure the out-of-the-box VuFind product to meet the needs of all the libraries' user groups, and a working group was assembled to customize the public interface. Using a wiki, they assembled a list of "canned search suggestions" for search testing, and provided ongoing suggestions and feedback. Meanwhile, a cataloging subcommittee consisting of all catalogers and the ILS administrator, met to discuss which MARC fields needed to be indexed and displayed to meet the needs of the larger working group. Despite the default MARC mappings included with the initial VuFind download, the catalogers thought it best to spend a considerable amount of time refining these defaults to maximize discoverability and address the unique needs of specialized researchers at the university. Over the course of their meetings, they discussed indexing for keyword and browse capabilities of various searches (title, author, subject, etc.) and the display of specific fields and subfields in the item-level view. The library staff tasked with customizing the discovery layer used the PBworks wiki, a collaborative online editing system already being used by the library for other projects, as a working space.[31] Meeting minutes and documentation including test search terms, mapping decisions, and comments on the current OPAC, were all included.

During the customization process, Joyner's Discovery Services Librarian, in consultation with her team, made the decision to switch from VuFind to Blacklight. She shared with the DAB that there were performance issues with VuFind for which she and her team were unable to secure satisfactory solutions from other adopters or the VuFind development community. Other reasons cited for the switch were stronger partnerships with other universities that had implemented Blacklight, a more robust development community, the complex customizations offered by Blacklight's Ruby on Rails environment, and the University of Virginia, also a SirsiDynix Symphony institution, was a Blacklight development partner. All of the MARC mapping work initially done for VuFind was directly transferrable to Blacklight.

The process of mapping MARC fields for a discovery layer was not unfamiliar to the libraries, as the catalogers had previously engaged in reviewing and replacing the default MARC mappings for the libraries' instance of ProQuest's Summon. Initially, the catalogers reviewed the default VuFind mappings, and then customized them to meet the needs of differing user populations across all the libraries, drawing on knowledge of these populations gained from direct interactions with the users or from discussions with public services personnel. Specific user communities, such as Laupus's users, the music library's users, and the Teaching Resources Center's users, were discussed in detail,

along with specific material formats (musical scores, media, kits, and government documents).

While the default MARC mapping settings provided a good starting point for discussion, it was agreed that these settings were inadequate for the libraries' catalog. Many important subfields were initially not mapped or completely ignored in the defaults, and the catalogers had to identify these and add them as necessary. For example, the default "author" keyword search did not consider such fields as the 245 $c (the transcribed statement of responsibility), or any entities in 8xx fields (which include author and title information of series). These were deemed extremely important to include in the author index, and promptly added; without them, any author keyword search would be missing vital information such as the name of the author as transcribed from the title page.

As another example, many subfields for uniform titles were ignored; VuFind's default settings included only the 240 $a and 700 $t subfields in the title index. This ignored the multiple subfields used extensively by the music library for musical uniform titles, and their use in other formats. The 240 field was customized so that the discovery layer took into account subfields $a, $d, $f, $j, $k, $l, $m, $n, $o, $p, $r, and $s for a title search, plus subfield 700 $k, $l, $m, $n, $o, $p, $r, $s, and $t. The addition of these fields allowed searches to take into account more than simply the first few words in a musical uniform title. For example, the uniform title for Mozart's Clarinet Concerto is *Mozart, Wolfgang Amadeus, $o 1756–1791. $t Concertos, $m clarinet, orchestra, $n K. 622, $r A major*. Adding these subfields now allowed keywords like "clarinet," "A major," or the thematic catalog number "K. 622" to come up in searches, allowing keyword searches such as "Mozart clarinet concerto" or "Mozart concerto K. 622" to function. These new fields were given a specific relevancy "weight," so that the 245 field would still function as the title field with the highest relevancy, but other title information in the 240 and 700 fields would be considered. These omissions from the default settings indicate the importance of a thorough review by trained catalogers before implementing any open source discovery layer.

The MARC mapping for e-Library's alphabetical browse search was compared to the default alphabetical browse settings for Blacklight. Changes were made to Blacklight's browse search to correct deficiencies in e-Library's browse indexes and to ensure that each index corresponded with its search function. The general material designation (GMD) was considered outdated and removed from display. While this may be considered a radical step, the catalogers thought that with RDA's elimination of the GMD and Blacklight's ability to facet based on material format, it was no longer needed. Other changes included tweaking mapping so that uniform titles in the 7XX fields would be treated in the

same way as their counterparts in the 240 field. Care was taken to separate the subfields specific to personal names from the subfields specific to uniform titles, with the idea that uniform titles could be displayed in a dropdown menu next to the author's name in the index. Subfields considered important to identifying titles in the 245 field, such as $n and $p (number and name of a part or section of a work), were added to display.

The classification facet was another aspect of the discovery layer that provided additional functionality lacking in the older OPAC. Users could choose a facet that represented a specific section of call numbers, allowing post-search browsing for specific subject areas. The default set up was only for Library of Congress Classification (LCC); however, for the health science library users,

24	pattern_map.callnum.pattern_23 = ^R.*=>R - Medicine
25	pattern_map.callnum.pattern_24 = ^S.*=>S - Agriculture
26	pattern_map.callnum.pattern_25 = ^T.*=>T - Technology
27	pattern_map.callnum.pattern_26 = ^U.*=>U - Military Science
28	pattern_map.callnum.pattern_27 = ^V.*=>V - Naval Science
29	pattern_map.callnum.pattern_28 = ^WA.*=>WA/ZWA - Public Health
30	pattern_map.callnum.pattern_29 = ^WB.*=>WB/ZWB - Practice of Medicine
31	pattern_map.callnum.pattern_30 = ^WC.*=>WC/ZWC - Communicable Diseases

Figure 1. Example of Classification Facet Mapping

Facet Display	MARC
Atlas	007/00-01=ad
Globe	007/00=d
Map	007/00=a AND 007/01 ≠ d
Map	Leader/06=e,f

Figure 2. Example of Format Fact Mapping

National Library of Medicine (NLM) classification was included (highlighted in figure 1) to interweave seamlessly with LCC. Joyner's Teaching Resources Center includes books classed using Dewey Decimal Classification, and the music library classifies all media holdings using a non-subject-based alphanumeric model. These were not included in the call number facet; however, all call numbers remain searchable using a call number keyword index.

With previous search tools, many users and librarians at the university expressed frustration with the "format" search limiter in e-Library. This format function was limited to a one-to-one mapping between a system-specific format type and a bibliographic record, meaning that some types of items had to be filed under "format" labels that did not adequately describe them. For example, streaming audio recordings could be filed under the "e-resource" format or the "audio recording" format, but not both. Because of this, a decision was made to use only MARC bibliographic data to map the format facet in the discovery layer, rather than relying on the OPAC's proprietary "item type" information. The libraries began with the format values that existed within e-Library before adding their own additional formats that were not previously available. For example, under the new list, one could differentiate between "maps," "globes," and "atlases," whereas before each of these was available only under "maps" (see figure 2). Additional facets for "print" and "electronic" were developed to allow users to limit their searches according these terms. Each format was assigned a

mapping using MARC fixed field data. By mapping MARC values to the format facet in this manner, the catalogers could now ensure that a resource encompassing multiple formats was discoverable under all applicable formats. However, this required a great deal of time and effort refining the format facet to account for all the fixed field data in MARC records, and led to the discovery of certain errors within the university libraries' bibliographic data.

Errors

Errors were found in many records at the beginning of the Blacklight testing phase. The aforementioned "format" function proved to be one of the most difficult to develop. Because the university libraries had relied on proprietary item types for format mapping, rather than using MARC bibliographic data, there were many gaps and issues in format fields such as the 006 field, the 007 field, and numerous fixed fields.

Blacklight did not initially recognize the hand puppets held by Joyner, categorizing them as "unknown"; the format mapping was updated to correlate the material type "r" (for realia) with the format facet "physical object." The music library found that preliminary format mappings had marked long-playing (LP) records as globes. This was not due to poor mapping, but rather due to extensive cataloging errors in the bibliographic records themselves. Graduate assistants were trained to edit the 007 fields of over 600 LPs to make sure

they were assigned to the correct format. A similar problem occurred for several hundred monographs. The presence of "t" in the Leader/06 byte is primarily for manuscripts; however, the Libraries uncovered many non-manuscript books with the "t" in this byte. Many were manually investigated and changed to prevent their inclusion in the "Archival and Manuscript Materials" facet. Through updating the faulty data in hundreds of catalog records and refining the format mapping based on a more intimate knowledge of the fixed fields, catalogers resolved all "unknown" and incorrect formats in the Blacklight results. Another example of format-related errors were items that were included in both the "print" and "electronic" facets. While some of these were correct (such as a book with an accompanying CD-ROM), a large number of records had elements of both the print and electronic manifestation in their fixed fields. Many of these were vendor records that had been improperly coded; others were records with fixed fields that had not been appropriately updated when they were derived from existing records and contained faulty data that were generating inaccurate mapping.

Though labor-intensive, fine-tuning the format facet eventually proved to be beneficial to the libraries. By incorporating such fields as the 006 and 007 fields into format mapping, the discovery layer could now take into account records representing multiple formats, whereas previously only the "dominant" format could be mapped to the record. Rather than mapping items to an "electronic resources" format, users can now select "electronic" *and* "book," or "electronic" *and* "video" to fine-tune their search by both medium and physical carrier.

An additional class of non-format errors involved previously undetected irregularities in subfield assignment. For instance, Blacklight flagged many instances of repetitive or incorrect usages of the subfield b in the 245 in many records with AACR cataloging, which required correction. These errors appeared in reports generated by Blacklight as files of records were imported because the MARC 21 standards are integrated in the Blacklight software.

Another problem that became evident soon after the beta phase began was that items with copyright dates for years beyond the current one were not populating to the top of the results list when a user chose to sort by most recent publication date. A customization was quickly added so that the dates up to 9999 in the 260 or 264 date fields would display at the top of the "most recent" results list; this solved the problem for those books with future publication and copyright dates.

The errors uncovered showed the importance of clean, correct, and consistent MARC data in all fields. Many were the result of varying cataloging practices over several decades. The fixed fields in particular received a lot of attention; while many of these had not been used by the previous OPAC, they quickly became important with the new discovery layer. The tremendous workload involved in fixing problems was a useful reminder to curate good MARC data, regardless of which fields are used by the current software.

Relevance

The Libraries found that the ability to customize the relevance ranking was crucial. Blacklight allows librarians to customize the relevancy "weight" given to each MARC subfield, ensuring that certain fields (e.g., 100, 700, 245 $c) are weighted highly for certain methods of searching (e.g., author search). Unlike proprietary OPACs, with a "black box" relevancy system that remains mysterious and unknown, librarians can see the nuts and bolts in Blacklight and customize accordingly. For example, the music library ensured that the specialized music MARC fields were given proper weight in the relevancy engine; it is unclear and difficult to tell if this was the case in the previous proprietary OPAC.

Using e-Library, a keyword search for the famous nursing theorist "Virginia Henderson" failed to retrieve any of the works written by her or with her name in the title until the second page of results, even with the "sort by relevance" option selected. In contrast, a search for "Virginia Henderson" using Blacklight's search box provides truly relevant results with works written by her and with her name in the title populating within the first results. The university libraries fine-tuned the relevancy to the point that a search for the Beatles' album *Help!* and Kathryn Stockett's novel *The Help* each yielded the correct result in the first hit, even with only a one word difference between the two generic-sounding titles. This was done by creating a separate index with stopwords such as "the," "and," and "a" that was given a slightly lower relevancy weight than the main keyword index with all non-stopwords. By manipulating the relevancy of stopwords in this way, it ensured that they were taken into account when there were two records with a close match, separated by only a stopword (such as *Help!* and *The Help*) but also that a search without stopwords included would still bring up the relevant records.

Unique Customization

The libraries used fixed field data to represent a particular format held in the e-Library catalog: the Electronic Theses and Dissertations (ETD) collection. In e-Library, the decision had been made to create a specific format for these that could be used to find institution-specific theses and dissertations. However, these ETDs provided nothing within the MARC bibliographic record to indicate their status, so their format functionality was lost within both Summon and the default Blacklight discovery layer. To solve this problem, the

libraries developed a facet that searches for the words "East Carolina University" in the 502 field, and looks for local item data from the ILS output in the 999 field. This allowed for the creation of a "Local theses and dissertations" format.

In an open source discovery layer, it is possible to customize the indexing of fields that are displayed on the basis of MARC indicators. This function had been problematic in e-Library. The ability to refine by indicator is very helpful to differentiate between each 856 field that displays. In ECU's current e-Library setup, if a record contains an 856 field, a customized "Online Content" button is generated. However, this button can be misleading, as "Online Content" can refer to any manner of supplemental content, such as an online table of contents for a print book. In their Blacklight discovery layer, however, the libraries were ensured that such a button appeared only for those items with full-text content available online. This was done by using the 856 field's second indicator, which indicates the kind of online content available. An 856 with a second indicator 2, for example, indicates a "related resource" is available online; these did not receive an "Online Content" button, but instead were displayed as "Supplemental Content."

In some cases, uneven cataloging practices limited what could be done with indexing customizations. An 856 with a second indicator 1 means a version of the resource is available. This often includes tables of contents and not the full text, which can be problematic for users. However, identifying all 856 41 fields as "Supplemental Content" would have excluded thousands of PURLs (Persistent Uniform Resource Locators) on federal document records that link to full text from displaying as "Online Content." Because PURLs are not the actual resource, but rather a version of the resource, many have been appropriately coded as 856 41, making it difficult to correctly identify which 856 41 links feature full-text content, and which feature supplemental material.

The Libraries also customized the MARC fields that are displayed when a results list is generated after a search. There had been problems with e-Library, which allowed only a finite number of MARC fields to display in a search results list. The various campus libraries each valued different fields in this brief display, depending on their users' needs. Laupus, for example, felt the display of the 250 field (edition statement) was of great importance for a user selecting the proper item (see figure 3) In contrast, the music library preferred the display of the 300 field (physical description), so that users could distinguish between different kinds of published scores (see figure 4). While e-Library could not accommodate multiple displays for items from different libraries, it was relatively easy to program an open-source discovery layer to do this. Now, items display in a search results list with an edition statement if housed in Laupus, and with a physical description statement if housed in
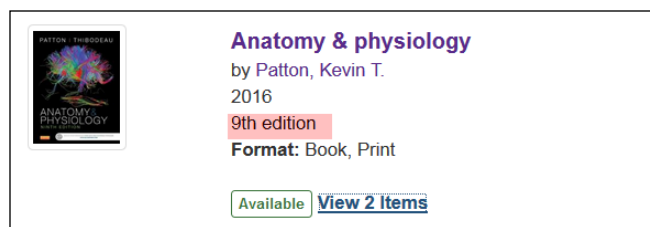

Figure 3. Example of Search Result for Health Sciences Library Item
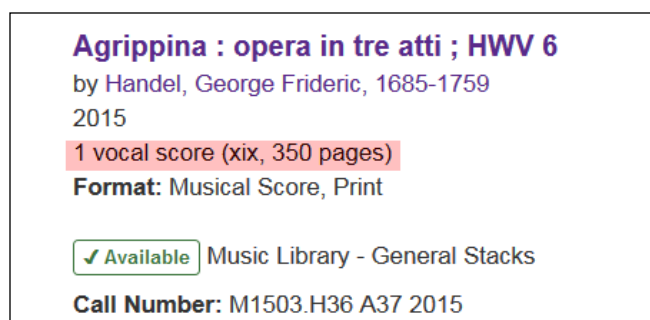

Figure 4. Example of Search Result for Music Library Item

the music library. This customization allowed the discovery layer to specifically serve each library's user base.

Future Considerations

While many of the developments of open source discovery layers focus on enhanced relevancy and accuracy for keyword searching, long-standing practices in information organization, along with the looming future of BIBFRAME and linked data models, mean that authority control practices remain relevant for facilitating discovery. From the beginning, librarians at the university have been committed to retaining all functionalities of the older OPAC, including browse lists, authority links, and utilization of see and see also references in authority records. Other libraries have expressed the desire to implement these features, such as the plan to implement cross-references in the University of Georgia's GIL-FIND.[32] In the short term, the university libraries plan to create browse lists that take advantage of authority records in the same way as most OPACs. While Blacklight lacks a default function that allows this, programmers remain optimistic about its implementation. VuFind documentation has recently been updated with instructions for implementing alphabetical browse lists, showing the possibility of creating such lists within a Solr system.[33]

The future holds bigger aspirations for using authority records in discovery layers. Katz and his colleagues make the case that "the value of rich authority data should be shared with everyone, not just catalogers."[34] One idea is to harvest

Figure 5. Example of MARC Authority Record with 373 Field

the cross-references in subject thesauri such as LCSH and MeSH for keyword searching, with a keyword term simultaneously querying all related terms. This has been proposed by Pace and other developers, and PubMed has created a system like this using MeSH terminology.[35] Another possibility for using authority data is to use the recently added RDA-compliant MARC fields in authority records. The 37X fields, for example, contain a wealth of robust information; to cite one specific example, the 373 field could perhaps be used to allow one to search for works by faculty members associated with the university (see figure 5).

It remains to be seen what the future holds for authority records, and how authorities will be administered in new schema such as BIBFRAME. Regardless of future developments, open source discovery layers will allow libraries to customize their setup and experiment with different methods of taking advantage of this data.

Conclusion

Through customizing an open source catalog, the university libraries gained an understanding of the inner workings of how the data could be manipulated in a way that best meets the users' needs. An important lesson learned was not simply to accept an open source product's default settings. While VuFind or Blacklight can be a powerful tool, implementation is not simple. Many hours of work were necessary to clean up errors and set up the indexing and display functions that were present in the older OPAC, and additional hours were needed to establish added features. The implementation also revealed previously unaddressed cataloging errors that required significant clean-up. This process requires the time and staffing to perform a systematic review of MARC mappings for indexing and display, and a working knowledge of the features desired by users and other stakeholders.

Such an understanding should help catalogers to more easily comprehend how data will be manipulated in other item description environments. Additionally, the collaboration between public and technical service departments from the three university libraries has led to continued cooperation and increased efficiencies. Finally, all three libraries, with their unique materials, classification and subject schemas, and user groups were able to customize a product that effectively takes all of those areas into account.

References and Notes

1. East Carolina University, "Measures of Success," last modified February 2, 2015, www.ecu.edu/cs-admin/news/measuresofsuccess2014.cfm.
2. East Carolina University, "ECU at a Glance," last modified September 22, 2014, www.ecu.edu/cs-admin/mktg/points_east_facts.cfm.
3. Ibid.
4. East Carolina University, "Fact Book 2014-2015," accessed January 27, 2015, www.ecu.edu/cs-acad/ipar/customcf/DL/FB/FactBook14-15.pdf.
5. *An Assessment of the Integrated Library System Currently Shared by the Libraries of East Carolina University and Elizabeth City State University* (working paper, East Carolina University, Greenville, NC, 2013).
6. Elizabeth Ketterman, Megan E. Besaw, and Michael Tucker, "Rethinking How Patrons Discover Information: Implementing a Discovery Tool," accessed July 14, 2015, http://thescholarship.ecu.edu/handle/10342/3506.
7. SirsiDynix offers next-generation catalog products such as Enterprise or BlueCloud PAC; however, these were still in the early stages of development when ECU Libraries began to look for other discovery options, and were not considered a viable replacement at the time.
8. Kate B. Moore and Courtney Greene, "Choosing Discovery: A Literature Review on the Selection and Evaluation of Discovery Layers," *Journal of Web Librarianship* 6, no. 3 (2012): 146.
9. Terry Ballard and Anna Blaine, "User Search-Limiting Behavior in Online Catalogs: Comparing Classic Catalog Use to Search Behavior in Next-Generation Catalogs," *New Library World* 112, no. 5–6 (2011): 262.
10. Elizabeth Sadler, "Project Blacklight: A Next Generation Library Catalog at a First Generation University," *Library Hi Tech* 27, no. 1 (2009): 58.
11. Birong Ho and Laura Horne-Poppe, "VuFind—An OPAC 2.0?" in *New Directions in Information Organization*, edited by Jung-Ran Park and Lynne C. Howarth (Bingley, UK: Emerald, 2013), 159–71.
12. John Houser, "The VuFind Implementation at Villanova University," *Library Hi-Tech* 27, no. 1 (2009): 95.
13. Demian Katz and Andrew Nagy, "VuFind: Solr Power in the Library" in *Library Automation & OPAC 2.0: Information Access and Services in the 2.0 Landscape,* edited by Jesus Tramullas and Piedad Garrido (Hershey, PA: Information Science Reference, 2012), 93.
14. Ibid.; Kate B. Moore and Courtney Greene, "The Search for a New OPAC: Selecting an Open Source Discovery Layer," *Serials Review* 38, no. 1 (2012): 28.

15. Katz and Nagy, 74.

16. Ballard, 267.

17. Sadler, 58.

18. Kathryn La Barre, "Faceted Navigation and Browsing Features in New OPACs: Robust Support for Scholarly Information Seeking?," *Knowledge Organization* 34, no. 2 (2007): 86.

19. Dayna Leebaw et al., "Improving Library Resource Discovery: Exploring the Possibilities of VuFind and Web-Scale Discovery," *Journal of Web Librarianship* 7, no. 2 (2013): 178.

20. Leebaw, 175.

21. Sadler, 59.

22. Ibid., 58

23. Tracey Snyder, "Music Materials in a Faceted Catalog: Interviews with Faculty and Graduate Students," *Music Reference Services Quarterly* 13, no. 3-4 (2010): 71–72.

24. Nara L. Newcomer et al., "Music Discovery Requirements: A Guide to Optimizing Interfaces," *Notes* 69, no. 3 (2013): 503.

25. Katz and Nagy, 73.

26. Ballard, 265.

27. Leebaw, 165.

28. Moore and Greene, "Choosing Discovery," 155.

29. Debra G. Skinner, "A Comparison of Searching Functionality of a VuFind Catalogue Implementation and a Traditional Catalogue," *Library Trends* 61, no. 1 (2012): 209.

30. Jenny Emanuel, "Next Generation Catalogs: What Do They Do and Why Should We Care?," *Reference & User Services Quarterly* 49, no. 2 (2009): 118.

31. "PBworks Overview," CrunchBase website, accessed December 1, 2015, https://www.crunchbase.com/organization/pbworks#/entity.

32. Skinner.

33. Demian Katz, "Alphabetical Heading Browse," VuFind website, last modified January 26, 2015, https://vufind.org/wiki/alphabetical_heading_browse.

34. Demian Katz, Ralph LeVan, and Ya'aqov Ziso, "Using Authority Data in VuFind," *Code4lib Journal* 14 (2011), http://journal.code4lib.org/articles/5354.

35. Kristen Antelman, Emily Lynema and Andrew K. Pace, "Toward a Twenty-First Century Library Catalog," *Information Technology & Libraries* 25, no. 3 (2006): 128; National Center for Biotechnology Information, "Entrez Help," last modified April 9, 2014, https://www.ncbi.nlm.nih.gov/books/NBK3837/.

Notes on Operations

Leveraging Author-Supplied Metadata, OAI-PMH, and XSLT to Catalog ETDs

A Case Study at a Large Research Library

Ken Robinson, Jeff Edmunds, and Stephen C. Mattes

Most academic theses and dissertations are now born-digital assets (i.e., electronic theses and dissertations). As such, they often coexist with author-supplied metadata that has the potential for being repurposed and enhanced to facilitate discovery and access in an online environment. The authors describe the evolution of the electronic thesis and dissertation (ETD) cataloging workflow at a large research library, from the era of print to the present day, with emphasis on the challenges and opportunities of harvesting author-supplied metadata for cataloging ETDs. The authors provide detailed explanations of the harvesting process, creating code for the metadata transformations, loading records, and quality assurance procedures.

I n August 2013, the Cataloging and Metadata Services Department of the Pennsylvania State University Libraries created the Digital Access Team in response to the need to devote more resources to the management of metadata for digital assets.[1] One of the team's primary activities is repurposing metadata from existing MARC records in Penn State's online catalog, The CAT, for digital collections in CONTENTdm and other platforms.[2] The team also works closely with the Library Technologies Department to repurpose MARC records in The CAT for mass digitization partnerships, such as HathiTrust and the Internet Archive.

The team began looking at repurposing metadata from other platforms for use in The CAT in October 2013. One promising source of metadata was Penn State's electronic theses and dissertations (ETDs) server.[3] Metadata for each ETD is available in unqualified Dublin Core (DC) format and can be harvested using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).[4] An important tool for harvesting this data is MarcEdit, a freely available metadata editing suite designed by Terry Reese.[5] MarcEdit provides many default crosswalks for mapping between multiple metadata schemes. These schemes can be customized for local metadata harvesting. MarcEdit also includes a tool for harvesting metadata from sites that have implemented or use OAI-PMH.

This paper describes the Digital Access Team's efforts to design an ETD cataloging workflow by harvesting author-supplied metadata using a customized DC-to-MARCXML Extensible Stylesheet Language Transformation (XSLT) crosswalk in MarcEdit to create a file of Resource Description and Access (RDA) MARC records for batch loading into The CAT.[6] The history of thesis cataloging at Penn State is described, including the transition to cataloging ETDs, and how the new harvesting method has improved access to ETDs while simultaneously freeing up staff time. Examples of MARC records for ETDs before and after the new procedure was implemented are provided, and time savings are quantified

Ken Robinson (kjr106@psu.edu) is a Digital Access and Metadata Specialist in Cataloging and Metadata Services at Penn State University Libraries. **Jeff Edmunds** (jhe2@psu.edu) is Digital Access Coordinator in Cataloging and Metadata Services at Penn State University Libraries. **Stephen C. Mattes** (smattes@uoregon.edu) is an Information Specialist in the Oregon Career Information System at the University of Oregon.

on the basis of studies conducted over a twelve-month period (three semesters). The paper also describes in detail the mappings created to harvest the metadata, the customizations made to the XSLT crosswalk, and the steps taken to ensure that the metadata batchloaded into The CAT is of sufficiently high quality.

Literature Review

Literature addressing the harvesting of ETD author-supplied metadata for creating MARC records for online catalogs is somewhat sparse, although efforts date back as far as 1999. Early harvesting strategies used Perl scripts. Sharretts, Shieh, and French described how the University of Virginia Library's pilot project using the Unix command-line utility Grep to extract bibliographic data from thesis PDF title pages and how it evolved into a series of Perl scripts that ran when a student submitted an ETD online.[7] Surratt and Hill described a similar process at Texas A&M University using a Perl script called ETD2MARC that took advantage of the open source MARC::Record Perl module.[8]

As OAI-PMH became more common, libraries began using this protocol to harvest ETD author-supplied metadata. Reeves described a process that Library and Archives Canada (LAC) used to harvest metadata with OAI-PMH queries that retrieved ETD Metadata Standard (ETD-MS) records for ETDs submitted from various Canadian universities in the Thesis Canada Portal.[9] Using this method, LAC had a cost savings of $95,000 in the 2006–7 fiscal year and expected progressively larger savings as more Canadian universities implemented ETD submission programs. McCutcheon et al. described an elaborate process at Kent State University in which a Perl script called ETDcat ran when it received an automatically generated notification from the OhioLINK ETD Center that an ETD had been submitted.[10] The script constructed an OAI-PMH query and retrieved the metadata as an ETD-MS record, which was converted into a MARC record using the MARC-Perl library.[11] Reese documented efforts made by Oregon State University (OSU) to harvest ETD metadata using MarcEdit's Metadata Harvester, which sent an OAI-PMH query that retrieved unqualified DC records.[12] A specialized XSLT crosswalk derived from a default DC to MARC crosswalk that is part of the MarcEdit installation was used to convert the records into MARCXML. Boock and Kunda also described the OSU experience, but focused more on workflow changes and cost savings.[13] They noted a time savings of seventeen minutes for cataloging each thesis using the new method described by Reese. Deng and Reese described further attempts of XSLT crosswalk customization at both OSU and Wichita State University for OAI-PMH ETD metadata harvests.[14] Bower, Courtois, and Turvey-Welch presented a similar OAI-PMH harvesting process for ETDs at Kansas State University.[15] Walsh provided an overview of metadata repurposing using XSLT and gave a user case showing a step-by-step process for harvesting author-generated metadata for ETDs using MarcEdit.[16]

Another avenue for acquiring ETD author-supplied metadata was to repurpose data supplied by ProQuest. Averkamp and Lee documented how the University of Iowa Libraries transformed ProQuest XML files using XSLT to create metadata that could be loaded into their online repository and was used to create MARC records for their online catalog.[17] Middleton, Dean, and Gilbertson described how the University of Arkansas Libraries used ETD author-submitted metadata supplied by ProQuest in MARC format.[18]

Although the literature addressed multiple ways to acquire ETD author-supplied metadata, the variable and often substandard quality of this metadata arose as a common theme. McCutcheon gave a good summary of the issues and noted that "the descriptive record created by automatic harvesting is only as good as the quality of the author-supplied metadata, which varies from author to author."[19] Metadata quality issues included representation of scientific symbols and diacritics, separation of titles from subtitles, nonfiling characters in the title proper, capitalization, management of whitespace, spelling, and other data entry errors.

History of Thesis and Dissertation Cataloging at Penn State

Cataloging of print theses and dissertations (TDs) at Penn State has historically been minimal level and formulaic. Catalog records generally consisted of the full title, author, date of issuance, a pagination count, degree type, and graduate degree program (in a local MARC 699 field). Library of Congress Subject Headings (LCSH) were assigned until 1964, though the headings were generally broad in scope. From 1965 until 1974, LCSH were added only when a personal name, corporate name, or title of a work were present in the TD title. Beginning in 1975, full subject analysis was performed and LCSH was assigned only for TDs containing the term *Pennsylvania* or a local Pennsylvania name (such as a town or county) in the title. This practice has continued to the present. With this workflow, the average thesis required ten to fifteen minutes to catalog, with an additional five to ten minutes per thesis if referred for subject analysis.

Such a relatively minimalist approach was designed primarily as a balance between providing sufficient access for TDs while minimizing the amount of time spent on complicated subject analysis for what are generally very narrow and specialized subject areas. Special Collections Team catalogers perform full subject analysis for any TDs added to Penn State's Special Collections Library.

Penn State University Libraries initiated a pilot project in collaboration with the Graduate School, Information Technology Services, and Digital Library Technologies in the fall of 1998 to investigate the possibility of allowing theses and dissertations to be submitted and archived electronically. The Graduate School began accepting ETDs in 2000. Penn State originally used ETD-db, an open-source ETD database developed at Virginia Tech.[20] The current ETD application is based on ETD-db and developed with Django, an open source web application framework and MySQL, an open source relational database management system.[21] Records provided in OAI-PMH feeds are currently in unqualified DC format.

Cataloging of ETDs began in 2004. The existing minimalist approach to cataloging TDs was used as a foundation for cataloging ETDs. Electronic aspects were added to the catalog records (MARC 006, 007, 538, and 856), added entries for thesis advisors were included for the first time (in MARC 700), and author-supplied keywords were added to MARC 653. Because ETDs were accessible online, catalogers began copying much of the data from the online record found in Penn State's ETD database and from the title page of the ETD's PDF file. This data was pasted into a MARC record in The CAT. To save time, ETD cataloging was supplemented with a series of Macro Express macros for data repeated in every MARC record.[22] Repeatable data included fixed field data elements in MARC 006, 007, nd 008, MARC 260 (Publication, Distribution, etc.), and MARC 538 (System Details Note). Because thesis titles in PDF files were sometimes entirely capitalized, copying and pasting proved to be as time-consuming as typing the title from scratch. The Digital Access Team's programmer created a script using AutoIt (a freeware automation scripting language) that adjusted the capitalization for pasting into The CAT.[23] Finally, an additional script was created to convert a bulleted list of keywords into a single MARC 653 field.[24] Using this approach, the average ETD required between five and ten minutes for cataloging. Cataloging an ETD generally took about half as long as cataloging a print thesis, a time savings due primarily to the efficiency gained through the copying and pasting of data.

Old Workflow, 2004–14

After receiving a list of ETDs from the Graduate School thesis office each semester, the thesis cataloger cataloged each ETD individually. Starting with a blank template in the local SirsiDynix Symphony ILS, the cataloger used macros line-by-line to fill in constant fields (fixed fields, 006, 007, 040, 260, 538). The cataloger transcribed or copied the title, author, degree type, advisor(s), and thesis department as they appeared on the ETD document. The cataloger took metadata from the ETD server page when it did not appear

in the document, such as keywords for the 653 field. The URL provided in the 856 field led to the splash page for the individual thesis.

The cataloger provided local authority control for ETD authors and advisors by searching the local catalog for any previous works by the author or advisor and using the form of name found. If no previous works were found, the name was entered in MARC 100 and 700 using the usage found on the ETD. The cataloger added MARC 246 fields for title variations, such as an alternate form for hyphenated words, or discrepancies between the title on the PDF and that on the ETD server page. The cataloger also added pagination information in the 300 Physical Description field, and checked for additional files to list in 300 subfield $e, such as audio or video files. Typically, the cataloger would spend the bulk of a month (100–160 person-hours) cataloging 300–400 ETDs after each semester.

With a shrinking staff, competing demands for time, and new priorities (such as the creation of metadata for digital projects), Cataloging and Metadata Services felt the time was right to transition from a largely manual, title-by-title process for cataloging ETDs to a more automated, batch approach that leveraged the power of harvesting author-supplied metadata.

Harvesting Metadata versus Records from ProQuest

All doctoral dissertations at Penn State are submitted to ProQuest/UMI Dissertation Publishing for microfilming.[25] This arrangement, initiated for the purpose of preservation, has been in place for over fifty years. Penn State does not submit master's theses to ProQuest, but authors may submit their master's thesis abstracts to ProQuest's Master's Abstract program. Undergraduates in Penn State's Schreyer Honors College are required to complete an undergraduate honors thesis. These are not submitted to ProQuest.

All doctoral dissertations and master's theses are currently submitted to Penn State's ETDs site and all undergraduate honors theses are currently submitted to Penn State's Electronic Honors Theses (EHTs) site.[26] Both sites are OAI-PMH compliant. Because metadata are readily available and can be harvested for all ETDs and EHTs and ProQuest only has metadata for doctoral dissertations, the Digital Access Team chose to harvest this data locally.

Harvesting Metadata: The Process

Metadata are harvested from Penn State's ETD server (etda.libraries.psu.edu) using MarcEdit's Metadata Harvester. The harvest process includes an XSLT crosswalk that transforms the DC data output by the OAI-PMH feed into

RDA-compliant MARC records.

The server name and query date parameters are entered in the Server box, for example,

> https://etda.libraries.psu.edu/oai/?verb=ListRecords&from=2014-04-01&until=2014-09-15&metadataPrefix=oai_dc

The Metadata type is set to "Dublin Core." Other options include OAI MARC, MODS, and MARC21XML.

The Crosswalk is set to the XSLT file locally customized to provide RDA-compliant MARC data in conformance with local standards for ETD metadata:

> C:\Program Files\MarcEdit 6\xslt\ETD_OAIDCtoMARCXML-rev5.xsl

Clicking on "OK" initiates the harvest, which usually takes only a few seconds for batchloads containing hundreds of records. The file of harvested records then automatically opens in MarcEdit's editor window. At this point, the MARC data can be further manipulated as needed for quality assurance. The edited .mrk file (a MarcEdit file format that is readable and easily editable by a human) is ultimately compiled into a MARC file in .mrc format for loading into the local ILS where the records are again spot-checked for quality to verify that aspects of the records more readily noticeable in the public Webcat interface are in fact correct and display as expected.

Tweaking the Dublin Core Mappings

The metadata available from the ETD server via the OAI-PMH harvest are largely author-supplied (i.e., input by authors at the time they upload their ETDs to the site). These data elements are internally mapped to DC elements. Since DC is a much less precise framework than AACR2 or RDA, the first hurdle faced was mapping the vagaries of DC to the precision of RDA expressed in MARC.

The data elements available on the ETD site include "Graduate Program" and "Keywords." Both were originally mapped to the DC element *subject*, which meant that the out-of-the-box XSLT transformation Penn State used as a test (OAIDCtoMARCXML.xsl) transformed both elements to the MARC 690 field, a local subject access field. Penn State's practice for ETDs has been to distinguish between keywords, which were manually input in MARC 653 (Index
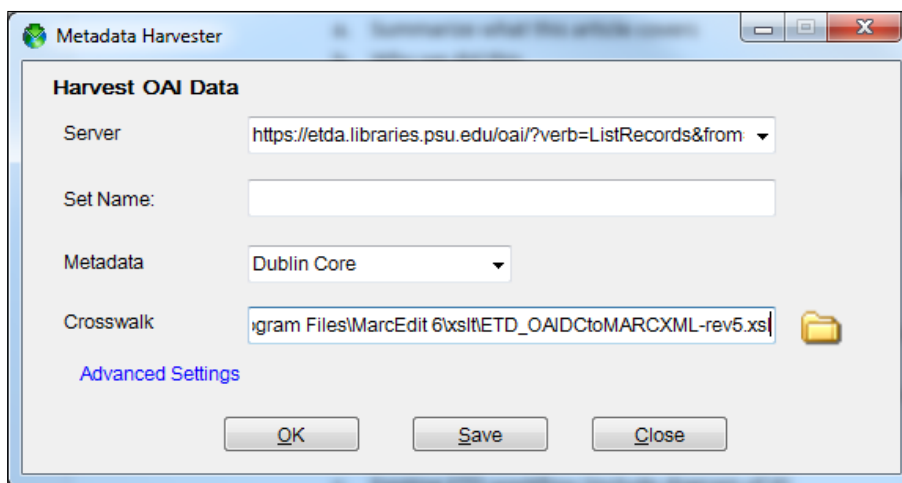

Figure 1. MarcEdit's Metadata Harvester

Term—Uncontrolled) fields, and graduate program data, which was input using MARC 699, a local subject access field specifically for collating theses by graduate program in The CAT. The solution was to change the mapping of the Graduate Program ETD element to a DC element not used elsewhere in the data, *coverage*, and edit the .xsl file to output the DC element *coverage* as MARC 699. The .xsl file was also edited to output *subject* not as the default 690 but as MARC 653 instead.

A similar conundrum existed for ETD data elements originally not mapped to DC. Neither "Degree" nor "Committee" data was mapped to DC, and therefore not output in the harvest. Degree information needed to be mapped to the MARC 502 (Dissertation Note) and committee member data (i.e., personal names) mapped to MARC 700 fields (Personal Name Added Entries) with relationship designators. The solution was to map degree to a DC element not used elsewhere in the mappings, *relation*, and then edit the .xsl file to map DC element *relation* to MARC 502. Because this mapping is not standard, Penn State Libraries' Digital Content Strategist (who remains in close contact with stakeholders at the Graduate School, the Honors College, and other institutions) was consulted to ensure that mapping the data in this way would not disrupt existing harvest workflows or negatively impact other potential harvesters of Penn State's data via the OAI-PMH feed. The DC element *contributor* had not been used in the mappings, so Committee was mapped to *contributor* and the XSLT file was edited to output this data in 700 fields.

Finally, there was the issue of MARC data completely absent from the ETD author-supplied data and handled imprecisely or not at all by DC: 006, 007, 008 (Fixed Length Data Element fields), 040 (Cataloging Source), etc. Much of this could be added to the records with MarcEdit following harvesting, but customizing the XSLT transformation

allowed us to add this data as part of the harvest itself.

Table 1 shows the ETD data elements available for harvest and their corresponding DC and MARC mappings.

DC data output from the OAI-PMH harvest was transformed and correctly formatted to RDA/MARC using a customized XSLT file.

XSLT Customization

Following initial testing of MarcEdit's *OAIDCtoMAR-CXML.xsl* crosswalk, it became evident that further customization would be necessary to make use of all the data available from the harvest of Penn State's ETD website, particularly regarding new mappings of degree type, graduate degree program, and access restrictions.[27] After ten separate revisions of the original crosswalk, members of the Digital Access Team working with the Authority Control Librarian of the Cataloging and Metadata Services Department reviewed samples of several hundred MARC records. Through each iteration, feedback was provided, errors were noted, and corrections and modifications were made to the crosswalk until its output met the department's quality standards.

One of the authors had prior programming experience, but no experience in XSLT coding. An online tutorial and a standard reference book were used to acquire a basic understanding of XSLT before making changings to the XSLT crosswalk.[28] Further information used in helping to debug the crosswalk was obtained by searching forums at Stack Overflow.[29] In total, 92.5 hours spread out over a seven-month timeframe were used to learn XSLT, code the crosswalk, debugging, testing, getting feedback, and writing documentation. While this amount of time is considerable, the initial investment paid dividends almost immediately, as the time required to process a semester's worth of ETDs plummeted from 100–160 hours to fewer than 8 hours. Return on investment (92.5 hours) occurred as soon as the new procedure was implemented in addition to paying dividends: 10–70 hours of newly available staff time.

The first customizations made on the XSLT crosswalk handled local non-standard assignments to the DC elements *coverage*, *relation*, and *rights* that are discussed in the previous section. *Coverage* contained the name of the graduate degree program, such as Architecture, Aerospace Engineering, and Kinesiology. The original *OAIDCtoMAR-CXML.xsl* crosswalk mapped this to the MARC 500 field. This was changed to MARC 699, a local subject access field specifically for collating theses by graduate degree program in The CAT. In initial tests, the MARC records output by the harvest showed MARC 699 positioned between MARC 300 and MARC 520. To correct this, the code for mapping MARC 699 in the XSLT crosswalk was moved between mappings of MARC 653 and MARC 700.

Table 1. ETD, Dublin Core, and MARC Mappings

ETD Data Element	Dublin Core	MARC
Author	creator	100
Email	N/A	N/A
Graduate Program	coverage (previously mapped to "subject")	699
Degree	relation (previously not mapped)	502
Document Type	type	N/A
Date of Defense	date	264 $c, 008 Date 1
Committee	contributor (originally not mapped)	700
Availability	rights (originally not mapped)	506
Title	title	245
Abstract	description	520
Keywords	subject	653
Files	identifier	856

Similar changes were made to other DC elements. The DC element *relation* contained the degree type, such as PhD or MS. The original *OAIDCtoMARCXML.xsl* crosswalk had mapped it to the MARC 787 field. Because degree types belong in MARC 502, the code in the crosswalk was changed to map it to that field. The DC element *rights* contained access restrictions. There were three possible values for this DC element: "Open Access," "Restricted," and "Restricted (Penn State Only)." In the original crosswalk, *rights* was mapped to MARC 540 (Terms Governing Use and Reproduction Note). This was changed to MARC 506 (Restriction on Access Note). The DC element *subject* was remapped from MARC 690 to MARC 653, since this DC element only contained author-supplied keywords. These three remaps were also re-positioned in the crosswalk so that their output displayed in the correct positions within a MARC record.

The second stage of customization involved adding MARC fields to the crosswalk that contained constant data appearing in every thesis MARC record. Examples included the MARC fields 006, 007, 008, 040, 260/264 (Production, Publication, Manufacture and Copyright Notice), 300 (Physical Description), and 538 (System Details Note). The MARC 008 was a special case: positions 00–05 required a computer-generated, six-character numeric string indicating the date the record was created in the format yymmdd. A function that could retrieve the current date was required. XSLT uses a language called XML Path Language, or XPath, that addresses parts of an XML document and performs calculations on it.[30] XPath provides several functions that XSLT can use. One of these is a function that retrieves the current date, *current-date()*, but because this function is from XPath
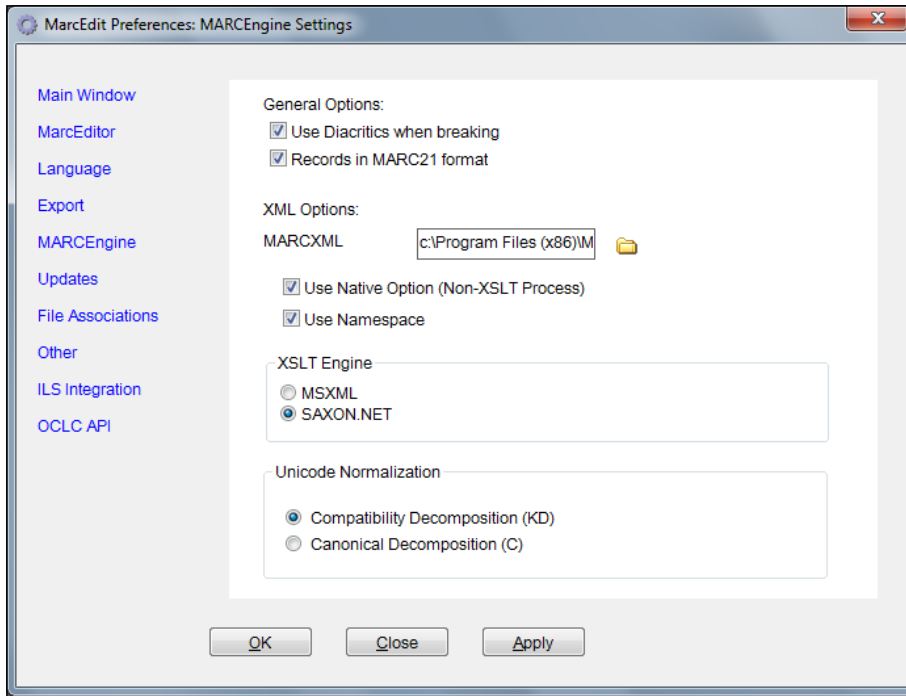
Figure 2. MARCEngine Settings in MarcEdit

2.0 (with the current crosswalk in XSLT 1.0), the XSLT Engine in MarcEdit needed to be set to SAXON.NET, an XSLT processor designed to run using the Microsoft .NET Framework, in the MARCEngine section of MarcEdit's preferences. By default, MarcEdit is set to the MSXML XSLT Engine, which does not support XPath 2.0 functions. Changing the XSLT Engine allowed us to add version 2.0 functions to a version 1.0 XSLT crosswalk without having to upgrade the entire crosswalk to version 2.0. This saved coding time, but in the future it may be desirable to convert the entire crosswalk to XSLT 2.0. The output of this function yielded the date in the format yyyy-mm-dd-hh:mm. To convert this date into the format needed for MARC 008, the output was concatenated using three separate *substring()* functions together.[31] Figure 2 shows the MARCEngine settings in MarcEdit used for the customized crosswalk.

In addition to retrieving the record creation date, the publication date is also required for MARC 008 positions 7–10 (Date1) and MARC 260/264 subfield $c. This was obtained during the harvest from DC element *date*. The date was in yyyy-mm-dd format and with the use of a *substring()* function, the first four characters were mapped to all of these MARC21 positions.

During the early stages of testing, harvested data was output into MARC records using AACR2. The department began the transition to RDA in early 2013, but thesis cataloging had not yet made the transition at the time of testing. This was an opportune time to convert the XSLT crosswalk

to output RDA-compliant MARC data. Following Program for Cooperative Cataloging (PCC) guidelines for RDA records, MARC 260 was changed to MARC 264.[32] The General Material Designation (GMD) in MARC 245 subfield $h was replaced with three new MARC fields: 336 (Content Type), 337 (Media Type), and 338 (Carrier Type). Instead of using MARC 502 subfield $a (Dissertation Note), dissertation information was parsed into separate subfields: $b (Degree Type), $c (Name of Granting Institution), and $d (Year Degree Granted). Relationship designators were added to MARC 100 and 700 in subfield $e. In other areas of the MARC record, abbreviations were spelled out, such as converting "Pa." to "Pennsylvania" in MARC 264 subfield $a (Place of Production, Publication, Distribution, Manufacture).

The next area customized was the display of degree type in MARC 502 $b. ETD authors submitted this information via a dropdown box, but the format in which this information was stored in DC element *relation* did not coincide with the punctuation conventions currently used in 502 $b, for example, Ph.D. was stored as "PHD" and M.Agr. as "M AGR." XSLT provides a method for expressing multiple conditional tests, by using an *<xsl:choose>* element in conjunction with multiple *<xsl:when>* elements.[33] An *<xsl:otherwise>* element is used in conjunction with *<xsl:choose>* when none of the *<xsl:when>* elements matches the conditions being tested. To use this method, the degree type harvested from DC element *relation* was stored in an XSLT variable called *degree*. Ten different *<xsl:when>* tests were performed on the value in *degree* to see if it matched one of ten different degree types offered by Penn State. If it matched one of them (i.e., PHD), the corrected form (i.e., Ph.D.) was stored in another variable called *degree_output*. The value in *degree_output* was mapped to 502 $b. If none of the *<xsl:when>* tests resulted in a match, then the value in *<xsl:otherwise>* was used. The term "Unknown" was assigned for this case. This will be helpful for detecting any future new degree types. Each time a harvest is performed, a visual scan of the records is sufficient to catch these for manual correction and future updating of the crosswalk. Figure 3 shows the XSLT coding for mapping the 502 field.

A similar *<xsl:choose>* structure was created for a single subfield in a local MARC 949 field. At Penn State, the 949 field is used to create holdings information for each record

during batchload into the catalog. The field contains nine subfields of which eight are constant data, set by local policy and coded directly into the XSLT crosswalk:

- $a (Call Number) = Electronic thesis
- $w (Class Scheme) = ASIS
- $m (Library) = ONLINE
- $k (Current Location) = ONLINE
- $l (Home Location) = ONLINE
- $o (Notes) = no value assigned
- $r (Circulate Flag) = Y
- $s (Permanent Flag) = Y

The ninth, subfield $t contains the item type. For ETDs, only two values are valid: THESIS-D for doctoral dissertations and THESIS-M for master's theses. These values are used in The CAT as limits for searching on doctoral dissertations or master's theses. The *<xsl:choose>* coding for this

```xml
<xsl:variable name="degree" select="dc:relation"/>
<xsl:variable name="degree_output">
  <xsl:choose>
      <xsl:when test="$degree='PHD'">Ph.D.</xsl:when>
      <xsl:when test="$degree='DED'">D.Ed.</xsl:when>
      <xsl:when test="$degree='DMA'">D.M.A.</xsl:when>
      <xsl:when test="$degree='MS'">M.S.</xsl:when>
      <xsl:when test="$degree='M AGR'">M.Agr.</xsl:when>
      <xsl:when test="$degree='MArch'">M.Arch.</xsl:when>
      <xsl:when test="$degree='MA'">M.A.</xsl:when>
      <xsl:when test="$degree='M Ed'">M.Ed.</xsl:when>
      <xsl:when test="$degree='ME'">M.Eng.</xsl:when>
      <xsl:when test="$degree='MLA'">M.L.A.</xsl:when>
      <xsl:otherwise>Unknown</xsl:otherwise>
  </xsl:choose>
</xsl:variable>
<datafield tag="502" ind2=" " ind1=" ">
  <subfield code="b">
      <xsl:value-of select="$degree_output"/>
  </subfield>
  <subfield code="c">Pennsylvania State University</subfield>
  <subfield code="d">
      <xsl:value-of select="concat(substring(dc:date,1,4), '.')"/>
  </subfield>
</datafield>
```

Figure 3. XSLT Coding for Mapping Degree Type in MARC 502 Field

was almost the same as that for mapping to 502 $b, except it mapped an item type based on the value found in DC element *relation*. For example, "PHD" maps to THESIS-D and "MS" maps to THESIS-M. The *<xsl:otherwise>* value was set to THESIS-M because there were larger numbers of master's theses than doctoral dissertations during testing. This will prevent this subfield from being blank and causing a batchload to fail. During a visual scan following a harvest, any instances of "Unknown" found in 502 $b requires that the cataloger check and correct 949 $t. Figure 4 shows the XSLT coding for using the degree type to determine the item type in Penn State's local 949 field.

Coding was added to the XSLT crosswalk to handle initial articles in thesis titles. Because the majority of Penn State theses are written in English, the crosswalk handles only the initial articles "a," "an," and "the." Respectively MARC 245 indicator position two is set to 2, 3, and 4. In all other cases, it is set to 0.

Another challenge was determining where a title ends and a subtitle begins. Sharretts, Shieh, and French noted that they considered anything following a colon as a subtitle.[34] Penn State took a similar approach, but expanded it to include the space following the colon. This decision was made in anticipation of unusual usage of colons in acronyms or for artistic or typographic effects. Our samples showed that the space following the colon was used in all cases and future testing will determine whether more elaborate coding is warranted.

Author names were already stored in inverted order on Penn State's ETD website and did not contain fuller forms or birth dates. Consequently, the original code in the *OAID-CtoMARCXML.xsl* crosswalk was simplified by removing the "persname_template," an XSLT template designed to construct the MARC 100 and 700 fields for names with fuller forms and/or birth dates. Functionality for mapping additional authors to MARC 700 was retained even though co-authors were not found among any of the samples tested.

Unlike author names, thesis advisor and committee member names were stored in the DC element *contributor* in direct order. The form that ETD authors used to submit their thesis advisors and committee members is in free format, though there are separate areas for the advisors and committee members. In addition to the name, DC element *contributor* contains the role the individual played following the name and separated by a semicolon and space character. Roles include Thesis Advisor, Dissertation Advisor, Committee Chair, and Committee Member. Examples include "Jane Doe; Thesis Advisor" or "John Doe; Committee Member." There can be multiple thesis advisors and multiple committee members for each thesis. When a thesis advisor's name is not present, the committee chair is assumed to be the thesis advisor. Adding to the complexity, names as entered by the ETD author sometimes include prefixes (Dr., Professor), suffixes (Jr., III), and the degree of the advisor or committee member (PhD, DEd). In some rare cases, several members' names were entered consecutively in the same field. The

goal was to get all thesis advisors associated with a thesis mapped to MARC 700 fields with their names in indirect order. This was a particularly challenging and complicated coding task.

An Open Archives Initiative harvest of 773 theses was used as a sample to determine the variations found in the DC *contributor* element. Each variation was noted and an algorithm was developed to address the most common forms and some of the more prevalent problematic forms. While processing the *contributor* element, any unusual findings were mapped to the MARC 720 (Added Entry— Uncontrolled Name) for evaluation after the harvest. As a backup, anything the algorithm missed will be detected by our authority control vendor and reported as errors that can be cleaned up later. In future harvests, the algorithm may need to be adjusted to address new issues that may arise.

A simplified version of the algorithm to convert direct-order personal names as harvested into name added entries in the MARC bibliographic record in inverted order (i.e., Last Name, First Name, other data):

```
<xsl:variable name="degree" select="dc:relation"/>
<xsl:variable name="itemtype">
  <xsl:choose>
    <xsl:when test="$degree='PHD'">THESIS-D</xsl:when>
    <xsl:when test="$degree='DED'">THESIS-D</xsl:when>
    <xsl:when test="$degree='DMA'">THESIS-D</xsl:when>
    <xsl:when test="$degree='MS'">THESIS-M</xsl:when>
    <xsl:when test="$degree='M AGR'">THESIS-M</xsl:when>
    <xsl:when test="$degree='MArch'">THESIS-M</xsl:when>
    <xsl:when test="$degree='MA'">THESIS-M</xsl:when>
    <xsl:when test="$degree='M Ed'">THESIS-M</xsl:when>
    <xsl:when test="$degree='ME'">THESIS-M</xsl:when>
    <xsl:when test="$degree='MLA'">THESIS-M</xsl:when>
    <xsl:otherwise>THESIS-M</xsl:otherwise>
  </xsl:choose>
</xsl:variable>
<datafield tag="949" ind2=" " ind1=" ">
  <subfield code="a">Electronic thesis</subfield>
  <subfield code="w">ASIS</subfield>
  <subfield code="m">ONLINE</subfield>
  <subfield code="k">ONLINE</subfield>
  <subfield code="l">ONLINE</subfield>
  <subfield code="r">Y</subfield>
  <subfield code="s">Y</subfield>
  <subfield code="t">
    <xsl:value-of select="$itemtype"/>
  </subfield>
</datafield>
```

Figure 4. XSLT Coding for Setting Item Type Based on Degree Type

1. Gather the roles of the first nine contributors (an arbitrary value intended to exceed the typical number of possible individuals).
2. Separate the name from the role using the position of the semicolon.
3. If the role is "Committee Chair" and no other role contains the term "Advisor," then set the role for that individual as "Thesis Advisor."
4. Because the name in DC element *contributor* is in direct order, assume the presence of a comma to mean the name contains a suffix or degree following it. Separate the data following the comma from the name. An example would be Martha Evans, PhD.
5. If the role is "Dissertation Advisor" or "Thesis Advisor," continue to the next step. Otherwise, ignore this name, exit the algorithm, and then start the whole process over with the next name.
6. Remove any titles from the beginning of the name (Dr., Prof., etc.).
7. Tokenize the name (i.e., split the name into individual elements).

8. Check the last token. If it contains a suffix (Jr., Sr., III, etc.), choose the second to last token as the surname. Otherwise, the last token is the surname. This check for suffixes is because sometimes they appear before the comma.
9. Output the surname into a MARC 700 subfield $a, followed by a comma, and then the rest of the tokens preceding the last token (first and middle names or initials).
10. If the name contains a suffix (Jr., Sr., III, etc.), then output the suffix into MARC 700 subfield $c.
11. Output "thesis advisor" into MARC 700 subfield $e.
12. Discard any titles or degree information following the comma (PhD, MS, Prof., etc.). After discarding, if any remaining data are found, output into MARC 720.

This algorithm may not appear to follow a logical pattern. The apparent illogic is due to having to write code within the limitations of variable assignments in XSLT. Unlike many traditional programming languages, XSLT does not allow a variable's value to be modified once it has been set.[35]

In testing the customized XSLT crosswalk, there was concern about how to handle unusual characters that might encode or display incorrectly in Penn State's online catalog

or in OCLC WorldCat. This occurred when a thesis was found to contain a Unicode line separator character, which caused half of a summary note (MARC 520) to appear at the end of the record in MarcEdit's MarcEditor. It became apparent that thesis authors often copied and pasted much of their information from whatever word processing software they used into Penn State's thesis submission forms. This practice introduced a large number of characters not generally compatible with online library catalogs. Penn State's Digital Library Technologies staff partially addressed the issue by applying a filter to strip control characters from the OAI-PMH feed. During additional troubleshooting, a sample harvest was imported into a local save file in the OCLC Connexion client. A considerable number of these records did not validate because of incompatible characters.

A two-pronged approach was used to address these characters. First, a script was written using AutoIt that reads in a MarcEdit .mrk file and writes the record numbers and incompatible characters found in those records into a spreadsheet.[36] Second, after reviewing the resulting spreadsheet, an XSLT template was developed for the crosswalk to convert all of the incompatible characters found in the 773-record sample harvest noted above.

The first process the template performed was a Unicode normalization using the Normalization Form Compatibility Decomposition (NKFD).[37] This converted single characters (a letter and a diacritic as a single character) into their decomposed forms of a letter and a combining mark. It also separated ligatures (such as "fi") into two separate characters. There are four different Unicode normalization forms, and through testing, NKFD produced the best results. For this to work, the XSLT Engine settings in MarcEdit required that Unicode Normalization be set to Compatibility Decomposition (KD).

The remainder of the template converts individual noncompatible characters into their compatible equivalents or as bracketed interpolations. These included both lowercase and uppercase Greek letters used as mathematical variables, right and left quotes and double quotes, a variety of dash and hyphen symbols, a large number of characters found in the Unicode Private Use Area for Microsoft symbol fonts, and other assorted mathematical symbols (such as the infinity symbol).

Because there is no way to predict what kinds of noncompatible characters thesis authors might include in their metadata, and developing a template to handle thousands of such characters is time-consuming, the AutoIt script used to detect them will be part of the workflow for future harvests. Henceforth, any non-compatible characters found will be manually corrected before loading into Penn State's online catalog. During future reviews, non-compatible characters that appear in large numbers may warrant additions to the crosswalk's template.

Throughout the development of this customized template, the Digital Access Team invested a significant amount of effort to minimize the amount of work needed in future harvests. The team expects that further tweaks will be necessary; the decision to implement a given enhancement will be based on whether the time required to implement the enhancement will save time or resources in the future.

New Workflow, 2014–

As noted above, the old workflow was largely manual. The new workflow was designed to free up a copy cataloger's time by leveraging the power of harvesting author-supplied data and batch loading records into the catalog. The time required to develop and test the process was not insignificant, but the immediate and long-term savings in time and gains in efficiency and quality warranted the decision to invest resources up front. The new process is outlined below.

After receiving the .mrk file of ETD metadata, the thesis cataloger begins quality assurance procedures. Using MarcEdit's Export Tab Delimited function, the cataloger exports a set of fields into a tab-delimited text file which is then opened in Excel to allow for sorting by fields. The included MARC fields are 001 (control number), 100$a (author name), 245 (title), 502 (degree type), 506 (access note), 699 (academic program), 700 (advisor), 856$u (URL), 720 (used for advisor fields that could not be properly parsed), 949 (holdings note, which generates call number and item information), and 520$a (summary).

The cataloger first sorts by author's name (MARC 100) and compares the file with a list of ETDs provided by the Graduate School to ensure that all ETDs were included in the extract. It may be necessary to either manually catalog ETDs that are missing from the extract, or to delete ETDs from the .mrk file that are not on the Graduate School's list for that semester. Some discrepancy in names may be caused by name changes, different parsing of compound surnames, or misspellings. The cataloger then compares the 502 field (degree type) to ensure that it matches the item type contained in the MARC 949, either THESIS-D for doctoral degrees (PhD, DEd, etc.) or THESIS-M for master's degrees (MS, MA, MEng, etc.). The cataloger checks for the presence of any 720 fields, which indicate that manual intervention is required on the 700 advisor note, and resolves as necessary. Next, the cataloger scans the 506 field and makes a list of all "Restricted" files, to shadow these records in The CAT after load. In SirsiDynix parlance, shadowing a record leaves it intact in the catalog but removes it from public view.

The cataloger checks for any ETDs that require assignment of LCSH based on departmental standards. These records are referred to original catalogers after the file has been loaded into The CAT. The cataloger uses the Find function in Excel on the column containing MARC 245 data

to find references to Pennsylvania or Penn State in ETD titles. The cataloger sorts by Graduate Degree Program (in the 699 field) and manually scans the titles of all ETDs in departments (e.g., humanities, arts, languages) most likely to be associated with such data to check for references to authors and works. At this stage, the cataloger also scans for ETD titles in non-English languages, and corrects the language fixed field for those records.

Next, the cataloger uses the AutoIt script, which checks all fields in the .mrk file for text characters not compliant with OCLC load requirements. After manually correcting these in the .mrk file, the ETD cataloger compiles the file into an .mrc file that is ready to load into SirsiDynix's staff module WorkFlows. When the load is complete, the cataloger shadows all records for "Restricted" ETDs to hide them from public view, and refers any records requiring subject headings to an original cataloger. Access restrictions last up to two years, after which the cataloger receives notification from the Graduate School and makes the records for shadowed ETDs visible to the public. For examples of thesis MARC records from 2002 (when many theses were still produced in print), 2004–14 (when ETDs were cataloged by hand), and from 2014– (after the adoption of the new workflow), see the Appendix.

Conclusion

Ensuring discovery of and access to materials in low-barrier self-deposit services, such as ETD databases, requires an enormous investment of time and resources when approached with a traditional cataloging mindset, i.e., cataloging items one-by-one. By leveraging metadata supplied by authors at the time of deposit, OAI-PMH harvests, and the transformations of data possible with XSLT, the authors devised tools and a workflow that greatly improved the efficiency of the cataloging process with minimal impact on metadata quality. Development and testing of the new procedure required a considerable investment of time, but with the scripts now in place and a redesigned workflow, a procedure that previously required months of staff time annually now takes hours. As cataloging and metadata departments are being asked to provide new services while still keeping up with traditional workflows, it is imperative to make every effort to streamline procedures that can be simplified. Cataloging ETDs is one such procedure. By extension, variants of the tools and processes described above could be applied to similar cases, such as institutional repositories or, in fact, any database in which metadata resides and is harvestable via OAI-PMH.[38]

References and Notes

1. Penn State University Libraries, Cataloging and Metadata Services, Digital Access Team, "Digital Access Team" (University Park, PA: Penn State University Libraries, 2015), www.libraries.psu.edu/psul/cataloging/digital.html.

2. Penn State University Libraries, "The CAT" (University Park, PA: Penn State University Libraries, 2014), http://cat.libraries.psu.edu.

3. Pennsylvania State University, "ETD, Electronic Theses and Dissertations" (University Park, PA: Pennsylvania State University, 2011), https://etda.libraries.psu.edu.

4. Open Archives Initiative, "Open Archives Initiative Protocol for Metadata Harvesting" (n.p.: Open Archives Initiative, n.d.), accessed November 23, 2015, www.openarchives.org/pmh/.

5. MarcEdit Development, "MarcEdit—Your complete free MARC editing utility" (n.p.: MarcEdit Development, 2013), http://marcedit.reeset.net.

6. American Library Association, "RDA Toolkit" (Chicago: American Library Association, 2010), www.rdatoolkit.org.

7. Cristina W. Sharretts, Jackie Shieh, and James C. French, "Electronic Theses and Dissertations at the University of Virginia," in Proceedings of the *ASIS Annual Meeting*, 240–55 (Medford, NJ: Information Today, 1999).

8. Brian E. Surratt and Dustin Hill, "ETD2MARC: A Semi-Automated Workflow for Cataloging Electronic Theses and Dissertations," *Library Collections, Acquisitions, & Technical Services* 28, no. 2 (2004): 205–23, http://dx.doi.org/10.1016/j.lcats.2004.02.014; Galen Charlton, "MARC-Record-2.0.6" (n.p.: CPAN, 2013), http://search.cpan.org/~gmcharlt/MARC-Record-2.0.6/.

9. Sharon Reeves, "User-Generated Metadata for ETDs: Added Value for Libraries," in Proceedings of the Tenth *International Symposium on Electronic Theses and Dissertations* (Uppsala: Uppsala University Library, 2007), http://epc.ub.uu.se/ETD2007/files/papers/paper-40.pdf.

10. Sevim McCutcheon et al., "Morphing Metadata: Maximizing Access to Electronic Theses and Dissertations," *Library Hi Tech* 26, no. 1 (2008): 41–57, http://dx.doi.org/10.1108/07378830810857799.

11. "MARC/Perl" (n.p. : SourceForge, n.d.), accessed November 23, 2015, http://marcpm.sourceforge.net.

12. Terry Reese, "Automated Metadata Harvesting: Low-Barrier MARC Record Generation from OAI-PMH Repository Stores Using MarcEdit," *Library Resources & Technical Services* 53, no. 2 (April 2009): 121–34, www.ala.org/alcts/sites/ala.org.alcts/files/content/resources/lrts/archive/53n2.pdf.

13. Michael Boock and Sue Kunda, "Electronic Thesis and Dissertation Metadata Workflow at Oregon State University Libraries," *Cataloging & Classification Quarterly* 47, no. 3–4 (2009):

297–308, http://dx.doi.org/10.1080/01639370902737323.

14. Sai Deng and Terry Reese, "Customized Mapping and Metadata Transfer from DSpace to OCLC to Improve ETD Work Flow," *New Library World* 110, no. 5–6 (2009): 249–64, http://dx.doi.org/10.1108/03074800910954271.

15. Merry Bower, Martin Courtois, and Michelle Turvey-Welch, "ETDs Transformed: Maximizing Cataloging Efficiencies and Open Access," in *International Symposium on Electronic Theses and Dissertations* (ETD 2009) (Pittsburgh: University of Pittsburgh, 2009), http://docs.ndltd.org/dspace/handle/2340/1129.

16. Maureen P. Walsh, "Metadata Repurposing Using XSLT," *More Technology For the Rest of Us: A Second Primer on Computing for the Non-IT Librarian* (Santa Barbara, CA: Libraries Unlimited, 2010), 125–39.

17. Shawn Averkamp and Joanna Lee, "Repurposing ProQuest Metadata for Batch Ingesting ETDs into an Institutional Repository," *Code4Lib Journal* 7 (2009), http://journal.code4lib.org/articles/1647.

18. Cedar C. Middleton, Jason W. Dean, and Mary A. Gilbertson, "A Process for the Original Cataloging of Theses and Dissertations," *Cataloging & Classification Quarterly* 53, no. 2 (2015): 234–46, http://dx.doi.org/10.1080/01639374.2014.971997.

19. Sevim McCutcheon, "Basic, Fuller, Fullest: Treatment Options for Electronic Theses and Dissertations," *Library Collections, Acquisitions & Technical Services* 35 (2011): 64–68, http://dx.doi.org/10.1016/j.lcats.2011.03.019.

20. Virginia Tech, "etds@vt" (Blacksburg: Virginia Tech, 2016), https://theses.lib.vt.edu/ETD-db/index.shtml.

21. Django Software Foundation, "Django" (n.p.: Django Software Foundation, 2016), accessed February 5, 2016, https://www.djangoproject.com; Oracle Corporation, "MySQL" (n.p.: Oracle Corporation, 2016), www.mysql.com.

22. Insight Software Solutions, "Macro Express: The Windows Automation Tool" (Kaysville, UT: Insight Software Solutions, 2015), www.macros.com.

23. AutoIt Consulting, "AutoIt: Automation and Scripting Language" (n.p.: AutoIt Consulting, 2014), www.autoitscript.com/site/autoit/; Ken Robinson, "PSU Thesis PDF Title to WorkFlows Conversion" (University Park: Penn State University Libraries, 2015), www.libraries.psu.edu/psul/cataloging/techresources/macros/autoit_docs/thesis_pdf.html.

24. Ken Robinson, "653 Creation from PSU Thesis Keywords" (University Park: Penn State University Libraries, 2015), www.libraries.psu.edu/psul/cataloging/techresources/macros/autoit_docs/create653.html.

25. Penn State, The Graduate School, "Thesis and Dissertation Guide: Requirements and Guidelines for the Preparation of Master's Theses and Doctoral Dissertations" (University Park, PA: Office of Theses and Dissertations, n.d.), accessed November 23, 2015, www.gradschool.psu.edu/current-students/etd/thesisdissertationguidepdf/.

26. Pennsylvania State University, "EHT, Electronic Honors Theses" (University Park, PA: Pennsylvania State University, 2011), https://honors.libraries.psu.edu.

27. The crosswalk and a description of how it works is available for download. See Ken Robinson, "Penn State eTD Dublin Core-to-MARCXML Crosswalk" (University Park: Penn State ScholarSphere, 2015), https://scholarsphere.psu.edu/collections/x346dj68d.

28. W3Schools, "XSLT Tutorial" (n.p.: W3Schools, 2015), accessed November 23, 2015, www.w3schools.com/xsl/; Jeni Tennison, "Beginning XSLT 2.0: From Novice to Professional" (Berkeley, CA: Apress, 2005).

29. "Stack Overflow" (n.p.: Stack Exchange, 2015), http://stackoverflow.com.

30. James Clark and Steve DeRose, "XML Path Language (XPath), Version 1.0" (n.p.: W3C, 1999), http://www.w3.org/TR/xpath/.

31. W3Schools, "XSLT, XPath, and XQuery Functions," *XSLT Tutorial* (n.p.: W3Schools, 2015), www.w3schools.com/xsl/xsl_functions.asp.

32. Program for Cooperative Cataloging, "RDA PCC Proposed Guidelines and Standards" (Washington, DC: Program for Cooperative Cataloging, 2015), www.loc.gov/aba/pcc/rda/RDA%20PCC%20Proposed%20Guidelines%20and%20Standards.html.

33. W3Schools, "XSLT <xsl:choose> Element," *XSLT Tutorial* (n.p.: W3Schools, 2015), www.w3schools.com/xsl/xsl_choose.asp.

34. Sharretts, "Electronic Theses," 251.

35. W3Schools, "XSLT <xsl:variable> Element," *XSLT Tutorial* (n.p.: W3Schools, 2015), www.w3schools.com/xsl/el_variable.asp.

36. The program code for this AutoIt script is available for download. See Ken Robinson, "AutoIt Script for Identifying OCLC Non-Compatible Characters in a MarcEdit .mrk File" (University Park: Penn State ScholarSphere, 2015), https://scholarsphere.psu.edu/files/x346dv29k.

37. Mark Davis and Ken Whistler, eds., "Unicode Standard Annex #15: Unicode Normalization Forms (Mountain View, CA: Unicode Consortium, 2014), http://unicode.org/reports/tr15.

38. It should be noted that subsequent to the adoption of the new ETD workflow, the Digital Access Team devised similar workflows, based on OAI-PMH and batch loading records, to catalog items available through the NASA Scientific and Technical Information Program (the NASA Technical Reports Server, www.sti.nasa.gov) and selected collections with good metadata in the Internet Archive.

Appendix

A typical print thesis MARC record from 2002:

LDR 00638 tm a2200169 a 4500
008 030902s2002\\\\pau\\\\\\\\\\000\0\eng\d
040 __ $a PSt $c PSt
100 1_ $a Halperin, Gregory S.
245 10 $a Therapist personality and tactics of interpersonal influence /$c by Gregory S. Halperin.
260 __ $c 2002.
300 __ $a 83 leaves.
502 __ $a Thesis (Ph.D.)--Pennsylvania State University, 2002.
533 __ $a Microfilm (positive). $e 1 reel 35 mm. $c (University Microfilms 30-64931)
699 __ $a Psychology.

A typical ETD MARC record from the Old Workflow period, 2004–14:

LDR 01428cam a22003137 4500
006 m\\\\\\\\d\\\\\\
007 cr\n|||||||||n
008 121205s2012\\\\pau\\\\\\\\\\000\0\eng\u
040 __ $a PSt $c PSt
100 1_ $a Shipley, Peggy Zoe.
245 10 $a Life patterns of family caregivers of patients with amyotrophic lateral sclerosis $h [electronic resource] / $c by Peggy Zoe Shipley.
246 30 $a Family caregivers of patients with amyotrophic lateral sclerosis
260 __ $a [University Park, Pa.] : $b Pennsylvania State University, $c 2012.
300 __ $a 1 electronic document (443 p.)
500 __ $a Mode of access: World Wide Web.
500 __ $a Thesis advisor: Janice Penrod.
502 __ $a Thesis (Ph.D.)--Pennsylvania State University, 2012.
533 __ $aMicrofilm (positive). $e 1 reel ; 35 mm. $c (University Microfilms 35-34763)
538 __ $a The full text of the dissertation is available as a Adobe Acrobat .pdf file ; Adobe Acrobat Reader required to view the file.
653 0_ $a Family caregiver; health as expanding consciousness; ALS; amyotrophic lateral sclerosis; life pattern
699 __ $a Nursing.
700 1_ $a Penrod, Janice L., $e thesis advisor.
856 40 $u https://etda.libraries.psu.edu/paper/15131/

A typical ETD MARC record harvested using the New Workflow, 2014–:
LDR 03310nam a22003737i 4500

006 m\\\\\l|\\d\\\\\\
007 cr\|n|||||||n||
008 141027s2014\\\\pau\\\\\sm\\\\000\0\eng\d
040 __ $a PSt $b eng $e rda $c PSt
042 __ $a dc
100 1_ $aTownsend, Andrew, $e author.
245 10 $aRevisiting Mesopredator Release : $b Carnivore Dynamics Along A Gradient of Landscape Disturbance.
264 _1 $a [University Park, Pennsylvania] : $b Pennsylvania State University, $c 2014.
300 __ $a 1 electronic document.
336 __ $a text $b txt $2 rdacontent
337 __ $a computer $b c $2 rdamedia
338 __ $a online resource $b cr $2 rdacarrier
502 __ $b M.S. $c Pennsylvania State University $d 2014.
506 __ $a Open Access.
520 __ $a Human induced habitat loss and predator persecution caused severe declines in apex carnivores throughout the North American continent. Removal of apex predators allowed smaller, lower rank predators from the Order Carnivora to become prominent. These "mesopredators" flourished, destabilizing ecosystems by driving many prey species toward extinction. However, some suggest that mesopredators still benefit from contemporary vegetation changes and fragmentation by thriving in disturbed areas. Many worry the versatility of these mesopredators could further threaten their prey species by leading to increased predation in anthropogenically-disturbed areas. This study seasonally sampled predator distributions along land cover gradients in forested, riparian corridors in Appalachia to identify whether landscape modification results in changes in carnivore community structure in the region. The study area consisted of randomly generated sites along streams in central Pennsylvania. I gathered data from camera traps and field surveys to catalogue the spatial ecology of mesopredators. I analyzed these data with landscape metrics to test the hypothesis that as forest contiguity decreases, both the abundance and richness of the predator community increases, possibly adding pressure on vulnerable prey populations. Through the analysis of these habitat metrics and carnivore occurrence data, this study found that carnivore species richness and abundance do generally increase with human disturbance in rural settings. However, this pattern is not due to the behavior of every species as many mesopredators are present across these rural landscapes and exhibit different responses to disturbance. Nevertheless, a few important generalists, namely the canids and raccoons, do show preferences toward more human disturbed areas and thus, are most accountable for this observed pattern.
538 __ $a The full text of the dissertation is available as an Adobe Acrobat .pdf file ; Adobe Acrobat Reader required

to view the file.
653 _0 $a spatial ecology
653 _0 $a mesopredator
653 _0 $a predation
653 _0 $a landscape gradients
653 _0 $a riparian corridors
653 _0 $a Appalachia
653 _0 $a ecological cascades
699 __ $a Geography
700 1_ $a Brooks, Robert P., $e thesis advisor.
700 1_ $a Bishop, Joseph A., $e thesis advisor.
700 1_ $a Serfass, Thomas L., $e thesis advisor.

856 40 $u https://etda.libraries.psu.edu/paper/22618 $z Connect to this object online.
949 __ $a Electronic thesis $w ASIS $m ONLINE $k ONLINE $l ONLINE $r Y $s Y $t THESIS-M

Notes on Operations
Electronic Outages

What Broke, Who Broke It, and How to Track It

Jennifer Wright

As electronic books and electronic journals have become more prevalent, so too do the number of electronic resources outages related to those resources. This paper, distilled from a presentation delivered at the 2015 American Library Association Midwinter Meeting, describes the implementation of a new tracking system for electronic resources outages at the University of Michigan (UM). It elaborates on the decisions that went into building the system and the insights gleaned from analyzing a year's worth of outages. It is hoped that such data might better inform decisions related to electronic resources at UM, and that its collection might inspire similar data-driving tracking elsewhere.

The landscape of electronic resources available to institutions is large, and ever-growing. As the number of resources increase, so do the kinds of technologies used to access those resource and the number of things that can go wrong with that technology.

After years of relying on email and anecdotal information to track electronic outages, the University of Michigan's (UM) Library's Electronic Access Unit (EAU) took a more concerted approach in 2013, working with the collaboration of their paraprofessional colleagues in the Electronic Acquisitions Unit and the Electronic Cataloging Unit. Managed by an electronic resources librarian, the paraprofessional staff in EAU developed an outage framework for the FootPrints ticketing system, which enables the team to track when electronic resources (e-resources) fail and under what conditions, with far greater precision than was previously possible. Additionally, FootPrints' extensive reporting capabilities enable more detailed analysis of e-resources outages and their causes, which analysis we hope to use to inform future purchasing decisions.

This study examines outages that occurred between June 2013 and June 2014. EAU sought to determine patterns based on outage type and vendor. The team also sought to highlight areas where it might improve the ticketing system to better capture types of outages missed in the initial implementation. Although Footprints was implemented spring 2013, this study begins in June of that year to allow for time to become familiar with the system.

Literature Review

Libraries are aware of the problems endemic in electronic resources e-resources (e-resources), particularly journals. Donlan notes that "even when e-journals are part of an aggregation (a standard package of titles as opposed to an individual menu of titles from one provider)) there can be problems with establishing correct holdings data. Content providers drop journals or lose them to competitors, so the end date for a journal run must be established in the link resolver's knowledge

Jennifer Wright (jennywri@umich.edu) is an Information Resources Assistant at the Hatcher Graduate Library, University of Michigan.

base for the OpenURL to 'know' where a particular issue is available."[1] Such issues are hardly minutiae of concern only to staff member on the backend of the system, for as Donlan notes, "Nothing is more frustrating to a user than to click a full-text icon that leads to 'web page not found' or 'sorry, no full-text was found for your article.'"[2] These problems have a direct and measurable impact on user experience, and the failure of those resources to perform reflects poorly on the library and the institution.

This effect is understood on the user's end only in the most basic terms, irrespective of the various workarounds devised to offer content. As Trainor and Price observed, "Put another way, users generally do not care whether the item is in the library's collection: they clicked the resolver button because they want to know whether the item is immediately accessible to them."[3] This points to the nature of the end user's problem ("Do I have access?"), which differ greatly from the complexities seen by staff attempting to troubleshoot the issue. While troubleshooters may ask where in the chain of communication between publisher, content provider, link resolver, and institution the access failed, the user ultimately has a different view of things. It is good to be reminded that despite the time and energy we devote to these troubleshooting efforts, the end result is still a struggle to answer a yes-or-no question.

In their survey of e-resources librarians, Rathmel et al. point to the difficulty not only in managing the relationships between libraries and vendors, but in managing relationships between different divisions of the library. Tools public services librarians have been utilizing for years as part of their reference work can be applied successfully to technical services interactions. According to Rathmel et al., "Customer relations management (CRM) and ticketing systems, both underused according to the survey, are one of the ways reference desks have managed handoffs, tracked statistics on common questions and resolutions, and gathered user feedback."[4] Though typically thought of in its outward-facing functions, CRM when harnessed internally within the library can greatly help communication between the many parties called upon over the course of outage resolution. At Oakland University, though BMC's FootPrints product was under consideration, ultimately staff went with a combination of Trello and Zapier's CRM software to improve internal communication on e-resource outages, and they saw that Trello's label system was useful in tracking trends in reported errors.[5] In their implementation of CRM software, Borchert noticed a marked increase in the expedience with which troubleshooters could organize reports of electronic outages (and, more useful still for long-term reporting, the ability to search and sort logs of previous outages greatly increased.)[6] In a presentation about the same implementation, Borchert and Graves received "lively discussion from attendees" on the topic of harnessing the power of CRM software internally to better keep disparate parts of the library aware of each other's efforts to solve the complex issues attendant upon e-resources. Borchert and Graves reported that "one audience member pondered why electronic resource management systems or integrated library systems were not providing this service. Another wondered if this expectation of integrated library systems was valid and whether libraries would be able to wait for vendor development in this area."[7] While the first question may remain unanswered, the answer to the second, in the intervening decade between 2006 when it was asked and now, appears to be "no." Libraries are done waiting.

In addition to CRM, cross-training or an introduction to other units' workflows may also assist in improving internal communication, as Malinowski observed at California State University, Fullerton (CSUF). At CSUF, "acquisitions staff met with staff responsible for managing the SFX server to discuss e-journal problem solving. As part of this discussion, the SFX staff developed an understanding of acquisitions workflows; this effort created a common understanding of answers to questions such as 'why is it taking so long to get vendors to respond?'"[8] Without exposure to the realities of acquisitions work (including the sometimes slow pace of library-vendor communications), troubleshooters might not understand why answers are so slow in coming.

Various organizations have attempted to wade into the shifting morass of electronic access and create best practices which should decrease the amount of failed access attempts. The Counting Online Usage of Networked Electronic Resources (COUNTER) initiative, for example, has attempted to leverage journal usage statistics to inform libraries' acquisition decisions, but is unable to account for the kind of discovery services most in vogue now. The International Federation of Library Associations and Institutions (IFLA) notes that "the availability of quality statistical data is important in understanding how well resources are used and how cost effective they are compared to other products. This is particularly important in supporting renewal and de-selection decisions."[9] However, "index-based discovery services may or may not directly embed full-text content, and therefore a number of the prescribed reports that deal with full text are not relevant."[10] With better data about how e-resources are or are not being used (for example, turn-away statistics), acquisitions specialists would be better able to avoid purchasing the kind of resources whose performance and stability turns out to be substandard. However, while prescribing best practices is useful to vendors already committed to or in the midst of change, they are guidelines given in the fervent hope that vendors might take notice of them and act accordingly. Without repercussions for a failure to abide by these guidelines, vendors are able to ignore them.

The implementation of the OpenURL standard maintains a great influence over whether users access

content. OpenURL linking, when supported by (for example) abstracting and indexing services, allows for the construction of links to full-text content that is subscribed to by the institution utilizing the OpenURL. Using bibliographic metadata to form, "OpenURLs are then sent to link resolvers run by individual institutions with which users are affiliated, which check the bibliographic information about the located resource against a local database of licensed and open access resource. The user is then presented with a list of options for how to access different versions of the resource in print and in licensed databases."[11] In seeking to map the degree to which different levels of error affected students' willingness to search out content despite outages encountered, Mann and Sutton examined what they termed "severe system error" and found that "failure to display an OpenURL link for some item types (e.g., gray literature) and refusing to accept the OpenURL link were serious problems which overpowered multiple students during the usability tests. . . . System errors add additional complexity to student interactions with e-resources—perhaps just enough complexity to overwhelm them."[12] While data suggested that users are willing to determine metadata differences like mismatched titles or journal issues with a quick Google search, encountering severe system errors like error codes, blank screens, and absent links may cause students to cease a search completely. The level of error is so great that users assume the desired content lies behind an insurmountable barrier of system malfunction. Such studies highlight the need both for explanatory error messages and for always providing "a way out" rather than resolving to a blank screen with no link to alternative paths to content.

Various examples exist of scholarly institutions' attempts to manage the response to e-resources outages. At Illinois State University, librarians polled both public and technical services staff and found a need both for increased cross-training between these two groups, and for public services staff to be more included in the conversations around electronic access. Foster and Williams note that "much of the literature that emphasizes the importance of collaboration rarely mentioned public services staff and limited their participation in electronic resources management to traditional roles. The public services interest in this study indicates that an important group of potential collaborators has been frequently overlooked."[13] Such studies exemplify the notion that e-resources outage resolution is a team effort extending across all branches of the library that come into contact with these resources. At Colorado State University, staff have been able to leverage LibGuides to expose some of the technical knowledge unique to those who acquire and maintain e-resources to all staff. They even go so far as to provide audience-specific e-resource LibGuides, stating that, "e-books might present a challenge for help desk staff who neither work with e-books regularly nor are familiar

with the general routines of ER access. A separate, abbreviated e-books guide was designed to directly target the library information desk environment with wording tailored for library staff and students working at the information desk rather than a general audience."[14] Texas A&M University has found screencasting to be of assistance in troubleshooting the errors of remote colleagues and documenting a problem for vendors.[15]

Lacking from existing research on this topic is a detailed examination, not just into how careful documentation of e-resources troubleshooting is done, but into what the results of such troubleshooting tracking show about the quality of vendor products. Such analysis might prove useful both in present and future negotiations for better services from the vendors libraries deal with on a daily basis. It is this absence that the author seeks to address in this study.

Selecting a Tracking System

In 2012, an internal analysis of library workflows at the University of Michigan (UM) Library revealed challenges in library communication. As a result, each unit looked more carefully at their methods of communication, including the EAU outages team. The use of electronic resources had been on the rise, creating new access problems and new needs for staff who can troubleshoot. While this helped to justify the hiring of two additional team members, the problem of communication remained, and in seeking to address it, the FootPrints ticketing system came up as an option.

Widely implemented across various units in the UM library system, FootPrints processes incoming emails to set an email address into ticket form in an online, browser-based workspace. All team members can access this workspace, which can be customized to track everything from ticket arrival time, status and description to time spent on a ticket. Tickets can be transferred across workspaces to other teams (such as acquisitions, cataloging and public services), and the fact that each of these teams were already using FootPrints as a ticketing system made its implementation all the more attractive to EAU.

Previous communication occurred primarily via emails, resulting in pertinent information becoming siloed in individuals' email boxes if respondents neglected to "reply all" in their correspondence. This increased the amount of time needed to address issues, as members of the troubleshooting team sought to determine who knew what about which outage. This situation was further complicated when the person who had received the pertinent information was out of the office. While we considered other ticketing systems including Trello and Jira, the ease of transferring tickets between units, the availability of a programmer who was familiar with FootPrints and how to customize it, and the fact that the

library had already purchased it drove the team's decision to use FootPrints.

Because of the high degree of customization the product offers, EAU customized dropdown menus, enabled time tracking, and provided free-text fields as necessary. The team established a group email address to forward emails received into the FootPrints system, where the newly formed tickets would await troubleshooting by one of the four members of EAU. The aim is to respond within 24 hours to any incident, with the exception of weekends. All team members receive notifications any time the group email address receives an outage report.

In addition to the group email, other units can cross-copy tickets into our FootPrints workspace, and by far the unit contributing the most of these kinds of outages is the Ask a Librarian reference service, which on average send about half of their problems per month to EAU. The issues referred from Ask a Librarian into our workspace tend to be complicated technical issues that cannot be solved by changing a search method or ensuring that a user has logged into the system before attempting to access resources.

Outage Types

A dropdown menu in Footprints provides a controlled vocabulary, eliminating the possibility of typographical errors or variants in naming when assigning an outage type. FootPrints allows for the classification of outages, which EAU has separated into the following twelve categories:

- *Bundled Content:* When several articles are bundled into a single document, making it difficult for both link resolvers and patrons to recognize that their desired content is inside
- *Configuration:* Including but not limited to concurrent user limit errors, missing IP ranges, improperly formatted SICI (Serial Item and Contribution Identifier) errors, poor site navigation, reset passwords
- *Proxy:* Where off-campus or wireless access through EZ Proxy is not working
- *Violation/Breach*: The amount and speed of accessed content exceeds limits set by the vendor and access is cut off as a precaution
- *Holdings:* Coverage in our catalog or knowledgebase does not match the vendor's coverage
- *Metadata:* Where incorrect metadata (author last name, volume/issue number, etc.) is causing linkage to break down
- *OpenURL:* Where the link resolver's linking syntax is insufficient for the resource and is failing to reach the article
- *Scheduled Maintenance:* Where the vendor has informed us of upcoming maintenance (this can also

occur on our end)
- *Target Content Lacking:* The article, e-book, volume or issue is not available on the vendor's site
- *Target Site Down:* Vendor's site is down
- *Subscription:* Where vendor does not recognize that we have a subscription to their product, likely due to renewal, licensing, or payment changes
- Other

Another option is to leave the problem type blank. Upon review, EAU determined that this typically occurred during a time crunch, when the troubleshooter filled out the form in a rush to attend to the surge of outages. Since the description field in Footprints contained a log of both the initial report and all the subsequent updates and final resolutions to these outages, the team felt comfortable assigning problem types retroactively, based on the wealth of information available on each outage and preserved in the ticket. In the future, EAU will take greater care to ensure that the Problem Type field is always populated with an outage type before a ticket is closed.

Problem and Resolution

To minimize the possibility of typographical errors muddling statistics, the "Who Caused Problem" and "Who Resolved Problem" fields are populated by dropdown menus, which contain the following options:

- Acquisitions
- Cataloging
- ER (Electronic Resources) Staff
- Knowledgebase Vendor
- Link Resolver Vendor
- Metadata Vendor
- Selector
- Subscription Agent
- User
- Vendor
- Web Systems

The system distinguishes between Knowledgebase Vendor, Link Resolver Vendor, and Metadata Vendor (the vendor of the discovery service's metadata), even though at present the same company (ProQuest) is used for all three services. This distinction assists both in knowing which part of the company to contact in case of future issues (repeat issues can be resolved faster if the team is able to point to previous occurrences and how they were resolved), and in maintaining a separation of functions in case the same company is not always in charge of all three services.

When the team attributes the cause of the problem to "Vendor," the intent is to indicate the content provider. For

example, if a user discovered an issue of a journal missing from ProQuest Research Library, and the cause of the "Target Content Lacking" outage was determined to be the vendor's fault, the "Who Caused Problem" field would be populated with "Vendor," and the "Vendor" field would be populated with ProQuest.

EAU lacked serviceable statistics regarding vendors before implementation of FootPrints, and relied on anecdotal memory to determine which vendors to add as dropdown menu options. The list in the following graphs does not encompass all of the vendors, but instead the large, easily recognizable names associated with enough outages at the time of implementation to warrant addition to the list. Using a dropdown list made choosing the correct version of a vendor name a more efficient process. The team knew it could not account for all possible vendors, however, even if it were not trying to keep the dropdown menu items to a manageable number on a single computer screen without scrolling. The team added the option of "Other" in the Vendor field, with a free-text entry box to enter other vendor names. While the option for free-text potentially raises the possibility of discrepancies in naming between individual troubleshooters (one troubleshooter may list American Medical Association, for example, and another may put AMA), EAU has discussed a controlled vocabulary and ensured that the same name is being used across troubleshooters for the same entity.

Life Cycle of an Electronic Resource Outage

To better grasp the troubleshooting workflow, the life cycle of a possible e-resource outage from beginning to end is illustrated. On Tuesday evening, a patron attempts to access an article in EBSCO's Omnifile Full-text Select via the library's OpenURL link resolver. The page the link resolves to is an EBSCOhost error page that states "No Results Found." The patron clicks the "Full-Text Not Working? Please let us know!" link below the 360 Link OpenURL button. This provides a Qualtrics form (www.qualtrics.com) for users to specify the problem (see figure 1). Users can also enter their email address for follow-up from reference staff, and any notes they wish to add to the report. Once submitted, this form generates a ticket in FootPrints' Ask a Librarian workspace, where a reference staff member seeks to re-create the outage. The staff member is able to re-create the error, suggesting that it was not a temporary glitch or user error. When the reference staff member can confirm that the problem is of a technical nature not noted in the catalog as a known issue, he forward it to EAU.

On Wednesday morning, the EAU staff member who handles that week's outage rotation looks at the outage, which is now in FootPrints' Electronic Outages workspace. The staff member drills down to the journal-level in the Omnifile

Figure 1. ArticlesPlus Qualtrics Form

Full-Text Select database on EBSCO and successfully locates the article in question. The staff member then tabs back and forth between the 360 Link page detailing the information searched upon unsuccessfully by ProQuest's link resolver, and the page containing EBSCO's metadata on the article. She identifies what she believes to be the problem: ProQuest calls the issue of the journal "7–8," whereas EBSCO calls it "7/8." Suspecting that this is causing the error, she opens up a ticket with ProQuest's Summon team, requesting that they change their metadata to match EBSCO's since that is how the article can be successfully reached on EBSCO. Later that day, a member of ProQuest's support team responds, indicating that ProQuest is unable to make this change since it would cause sites successfully using that metadata to link out from 360 Link to then fail in making that transfer. ProQuest suggests that the EAU staff member contact EBSCO, which the staff member then does.

EBSCO responds on Thursday, requesting screenshots of the error message, which the EAU staff member provides. EBSCO notes that the link resolver in use is not an EBSCO product, and that they cannot see what ProQuest is doing to cause the linking to fail. When the EAU staff member points again to the discrepancy in the metadata between the two companies, EBSCO responds that they cannot change their metadata for the same reason as ProQuest. Because of this impasse, the outage is marked as "can't resolve," with the problem type listed as "metadata." The outage is closed— all the vendor correspondence having been copied and pasted into FootPrints, allowing reference staff, whose Ask a Librarian ticket was linked to the Outages ticket, to see the resolution. The data are preserved in the event that it can be used to provide examples to vendors of products in need of improvement. Public services staff see EAU's updates in their own FootPrints workspace, which is linked to the
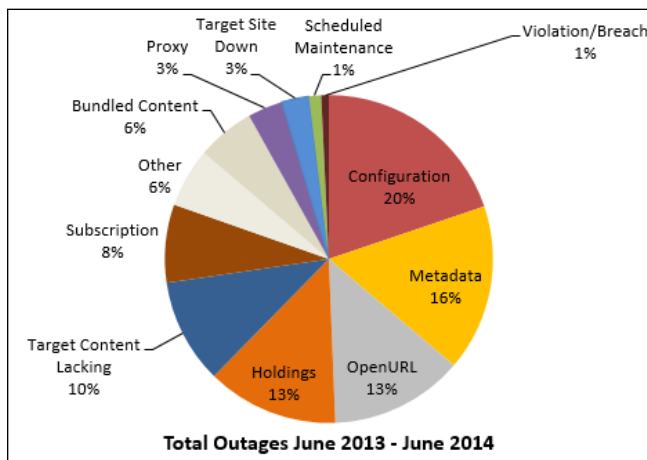
Table 1. Problem Type, Cause, and Number of Outages

Problem Type	Who Caused Problem	No. of Outages
Bundled Content	Link Resolver Vender	4
	Metadata Vendor	2
Configuration	Cataloging	1
	Knowledge Base Vendor	5
	Link Resolver Vendor	15
	Metadata Vendor	32
	User	5
	Web Systems	7
Holding	Cataloging	28
	ER Staff	8
	Knowledge Base Vendor	16
	Link Resolver Vendor	2
	Metadata Vendor	7
	User	1
Metadata	Knowledge Base Vendor	3
	Link Resolver Vendor	2
	Metadata Vendor	189
Upen URL	Knowledge Base Vendor	5
	Link Resolver Vendor	151
	Metadata Vendor	9
Other	Knowledge Base Vendor	1
	Link Resolver Vendor	4
	Metadata Vendor	2
	User	2
Proxy	Knowledge Base Vendor	2
	Metadata Vendor	1
	User	1
	Web Systems	7
Subscription	Subscription Agent	7
Violation/Breach	User	5

Figure 2. Total Outages June 2013–June 2014

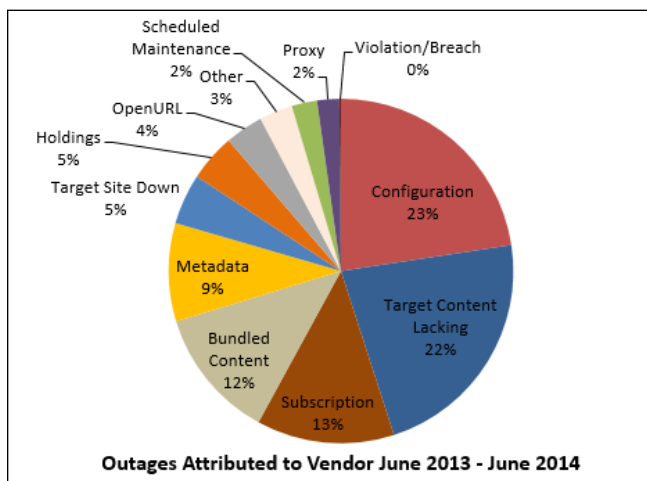**Figure 3.** Outages Attributed to Vendor June 2013–June 2014

troubleshooting workspace. By this point, public services have typically suggested Interlibrary Loan or other alternative routes of access to the patron. If there was a specific request for an update, the Ask a Librarian staff member will inform the patron of EAU's findings.

Findings

During the period studied, EAU received 1,586 tickets. Of those, the percentage breakdown by outage type is shown in figure 2, with most of the outage falling into the Configuration, Metadata, Holdings and OpenURL categories. Compare this to the breakdown of outages determined to have

been caused by the vendor (47 percent of total outages; see figure 3), where the numbers are slightly different. Here, a greater number of outages appear to fall into the Configuration, Target Content Lacking, and Subscription categories.

The reason for this discrepancy may be that most of the issues that fell into the Configuration category tend to involve information provided to the vendor that the vendor subsequently misplaced—for example, IP ranges, which the library sends vendors when we acquire resources. When vendors change platform or initiate site redesigns, full IP ranges are not always transferred, and EAU may need to remind the vendors of the full range the library has provided to them. The same holds true for "Target Content Lacking." If a vendor advertises that they can provide a given article, journal issue or e-book, and that content is missing, EAU populates the "Who Caused Problem" field with "Vendor."

Subscription issues may also be attributed to vendors who failed to register the library's active subscriptions to their content. When EAU provides a paper trail indicating that the library has paid for content, and when the resolution of the issue involves telling this to various customer help representatives, it is easy to attribute this problem to the vendor. The library had a subscription, the vendor failed to recognize it, but when they did, access was restored.

An analysis of outages not caused by vendor indicates room for improvement from our Metadata Vendor and our Link Resolver Vendor (see table 1), specifically for metadata and OpenURL outages. At the same time, we also see a need for our own staff to improve workflows regarding holdings and their maintenance to keep the catalog up to date with current coverage dates and access points. If our catalog is not up to date, the users seeking to go directly to journals or e-books rather than entering via the discovery service portal will never reach their chosen resources. EAU must also ensure that the e-resources we activate in the Knowledge Base are activated in a timely manner, and in the right locations.

It is important to keep the previous contexts in mind when assessing outages attributed to vendors versus those which are not (see figure 4). Often, because of issues like those previously described, "not attributed to vendor" does not mean that there is a different known contributor to the problem Rather, it means that there are so many breakdowns in the system that it is difficult to ascribe responsibility to any one party. The problems elucidated by such issues are greater than anyone single company or institution, and will need to be addressed at an industry-wide level beyond the scope of this article.

In some instances, however, the cause of an e-resource outage and the expected source of its solution is much more concrete. Take, for example, the numbers for Gale, ProQuest, and EBSCO, in figures 5, 6, and 7. Each of these vendors were unable to provide access to content that they said they would provide. This happened thirty-six times for Gale, twenty-one times for ProQuest, and twenty-two times for EBSCO.

Discussion

Metadata and holdings outages present problems greater and more endemic than a single paper can address. Because there is no established hierarchy for the maintenance of metadata, there is often not one place to resolve problems. When sources of responsibility can be identified, there are times when the multiple parties with a stake in the metadata reach an impasse regarding what metadata should be provided. For example, during one outage, it was discovered that the OpenURL link resolver was failing to arrive successfully
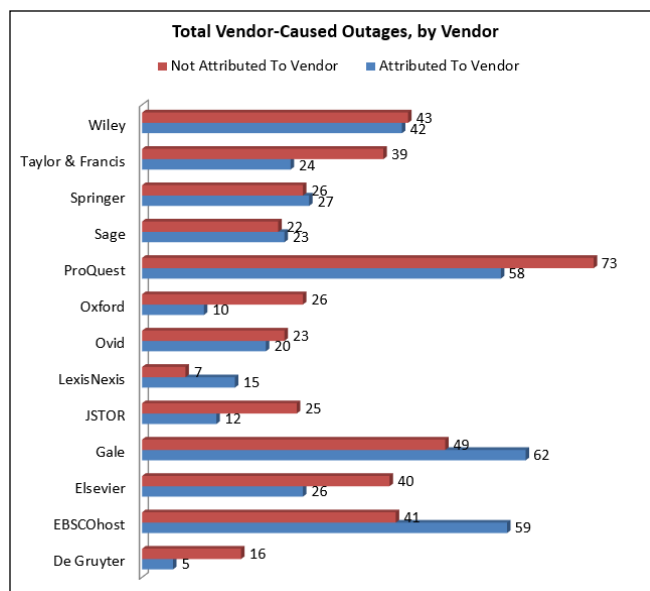

Figure 4. Total Vendor-Caused Outages, By Vendor

at a given article because JSTOR had indexed it as issue 1/2 instead of issue 1. The metadata vendor, ProQuest, indexed the article as appearing in issue 1. The end result was that the issue was resolved, and the link continued to fail. On another occasion, both the metadata vendor and the content provider acknowledged that the metadata was "not clear at all," and that because of this, they would not change their indexing, despite explicit examples indicating that leaving it as it appeared would perpetuate a lack of access to the article in question through OpenURL linking. In such instances, while the OpenURL was failing, the failure was due to discrepancies in the metadata, which was apportioned into the metadata category. However, in the absence of a chain of responsibility or the authority to encourage the party (or parties) to change their metadata, the problem remains. It also remains difficult to track. This raises the question as to who is responsible—the vendor or the metadata provider? In the face of such thus far unanswered questions, the metadata and holdings categories (subject to a similar dearth of responsibility) appear much less represented in the Outages Caused by Vendor data. The low percentages of these problems should not be taken as indication of stellar service from vendors. Were there a clearer chain of responsibility, the numbers for metadata and holdings issues would be much higher, given that combined they represent roughly a third of the total outages received.

Bundled content also raises questions of attribution. Usually involving large numbers of abstracts, reviews or proceedings, bundling tends to be beneficial for the content provider. Rather than create a separate web page for every article in quarterly collections of hundreds of
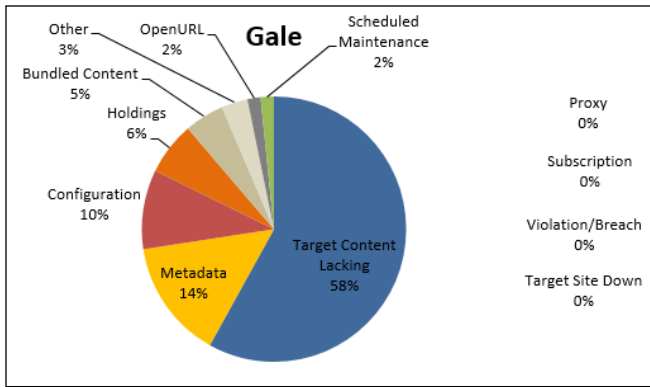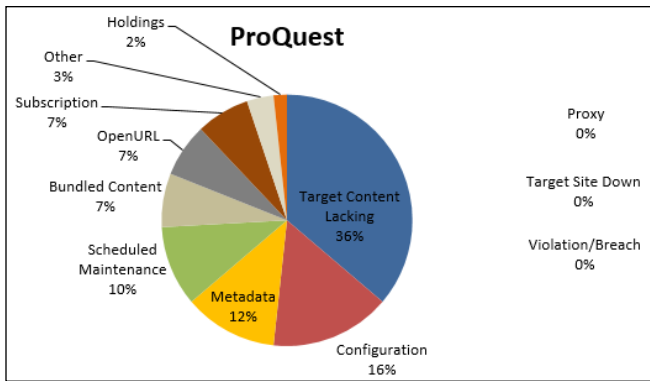
Figure 5. Outages Attributed to Vendor, Gale

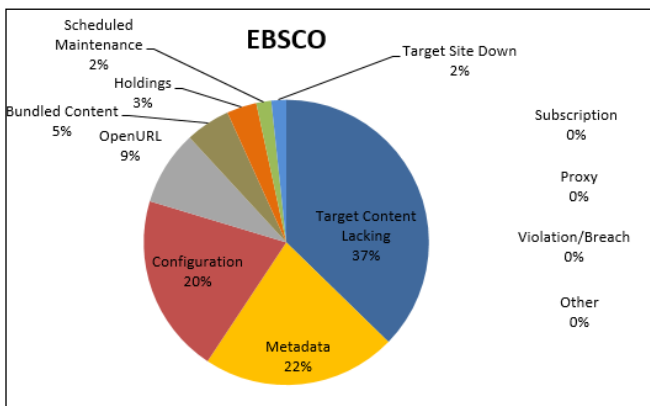
Figure 6. Outages Attributed to Vendor, ProQuest

Figure 7. Outages Attributed to Vendor, EBSCO

articles, vendors can combine articles into one PDF and make it available as a single file. While this is convenient for content providers, it is inconvenient for link resolvers since the metadata treats these abstracts and reviews as separate articles, not as parts of a larger PDF. It is as separate articles, with their own individual files, that the link resolver seeks to find them. The content provider

does not attach the metadata of every article contained within the PDF to the web page hosting the PDF, resulting in an error message and a flawed linking system. Multiple vendors have indicated that they have no plans to separate out such bundled content into individual files on individual pages, and for this reason we tend to attribute bundled content problems to the vendor in our FootPrints ticketing system. The immediate reaction might be to suggest abandoning OpenURL linking for a more reliable linking solution. While OpenURL constructs a link based on metadata, Direct Linking relies on an identifier available in the discovery tool's records to use as a marker to indicate where the link is intended to go. This identifier is a link to a particular full-text provider such as Gale or ProQuest or a Digital Object Identifier (DOI), ideally obtained through registration of the article with a DOI registration agency upon its publication. Because the onus is on the vendor to register the DOI, we attribute metadata issues with unregistered DOIs to vendors. However, the index, such as Summon, that contains those markers and controls which of those markers to use to direct patrons to the desired resource is proprietary material.

Broadly speaking, as Stuart, Varnum and Ahronheim point out in their analysis of direct-linking versus OpenURL linking, "From the Library's perspective, the trend to direct linking creates the risk of a vendor lock-in because the vendor-created direct links will not work after the library's business relationship with the vendor ends."[16] More narrowly, using direct-linking to circumvent the errors encountered in using OpenURL linking to access bundled content would still require a tremendous amount of maintenance. Some entity would need to track which vendors bundled individual abstracts and reviews into PDFs and which separated them into their own files. In addition, those vendors that separated bundled content into several smaller combined files (for example, turning abstracts from a conference into three separate PDFs, one for each day of the conference) would need to be tracked. The same entity would then need to manage direct links to the various providers offering this content in all the numerous ways they offer it. Such an undertaking would be prohibitively costly even in the short-term, let alone in the long-term when maintenance is inevitably required.

Changes

As a result of this study, EAU's personnel realized that they needed to make many changes to their workflows to better capture certain kinds of information related to electronic outages. A systematic review of the miscellaneous "Other" category of outages revealed that there were repeat outage types common enough within that category to warrant their own category. These types included the following:

- Concurrent User Limits: Where a resource was inaccessible because the maximum number of concurrent users had been reached
- User Error: Where the resource was thought to be inaccessible due to a user's misunderstanding of linking or access buttons, etc.
- Temporary Glitch: Where the resource was determined to be inaccessible at the time of reporting, but by the time of troubleshooting the problem had been resolved

Separating these issues from the "Other" category so that they have their own statistics will better enable EAU to track these problems and resolve them. For example, if concurrent user limits continue to rise for certain titles, selectors might choose to act upon this information and pay to increase the limit on concurrent users for affected resources. Or, if user errors keep arising for specific resources, EAU might investigate them and see if the vendor's display or the catalog's explanation of it could be improved to reduce the number of access issues. Temporary glitches, while ephemeral in nature, might still prompt a missive to the vendor, should they occur repeatedly and in such numbers as to become a concern. All of these issues will be reflected more reliably in future statistics, since they have been drawn out into separate categories.

EAU also realized that it needed to be more proactive about deleting correspondence tickets. These are tickets generated by emails erroneously sent to the group email address that migrates content into Footprints. Such emails tend to be about outages that have been addressed, usually from people other than those who originally reported the outage, and are sent to the ticket-generating email in ignorance of the preexisting ticket. Responding through the existing ticket is no trouble—and EAU endeavors to do that to keep a record in the ticket of all correspondence related to the outage. However, when the team does this, members need to make sure that they delete the erroneously generated outage tickets. While such tickets have been removed from this data, EAU has determined that it needs to be more aware in the future of the need to address such tickets as soon as they appear.

Conclusion

The landscape of electronic resources is anything but easy to articulate, even within one's own institution. Issues attendant upon communication, technical knowledge, and organization make capturing problems with e-resources difficult. Even knowing how to group outages into categories that can then be reported may require many discussions and adjustments before reaching consensus. Even then, with a reliable list compiled, attributing outages to the correct source of the problem swiftly becomes a point of contention, especially with vendors less than willing to work with one another on a given issue.

The roadblocks presented by e-resources are not easily surmounted. No one institution can systematically rid itself of the kinds of errors seen repeatedly, across platforms, vendors and content delivery services. However, the multifaceted nature of e-resources troubleshooting should not force libraries to accept the status quo as incontrovertible. Deciding where to place responsibility when it comes to thorny issues like metadata does not need to be a process that librarians as a profession avoid. With enough data gathered through systems like FootPrints and shared with both vendors and fellow institutions, libraries stand poised to improve the functionality of e-resources, not just for their own patrons, but for patrons everywhere. Improving our ability to describe errors, to capture examples of them and the attempts made to fix them, is the first part of what is sure to be an arduous but ultimately worthwhile process.

References

1. Rebecca Donlan, "Boulevard of Broken Links: Keeping Users Connected to E-Journal Content," *Reference Librarian* 48, no. 1 (2007): 99–104, http://dx.doi.org/10.1300/J120v48n99_08.

2. Ibid., 102.

3. Cindi Trainor and Jason Price, *Rethinking Library Linking: Breathing New Life into OpenURL* (Chicago: ALA TechSource, 2010).

4. Angela Rathmel et al., "Tools, Techniques and Training: Results of an E-Resources Troubleshooting Survey," *Journal of Electronic Resources Librarianship* 27, no. 2 (2015): 88–107.

5. Megan Finch, "Using Zapier with Trello for Electronic Resources Troubleshooting Workflow," *Code4lib Journal* 26 (2014), http://journal.code4lib.org/articles/10034.

6. Carol Ann Borchert, "Untangling the Jungle of e-journal Access Issues using CRM Software," *Library Collections, Acquisitions, & Technical Services* 30, no. 3-4 (2006): 224–37.

7. Carol Ann Borchert and Tonia Graves, "Using Customer-Service Software to Manage Serials Online Access Issues," *Serials Librarian* 50, no. 3-4 (2006): 213–15.

8. Susan Davis et al., "Who Ya Gonna Call? Troubleshooting Strategies for E-resources Access Problems," *Serials Librarian* 62, no.1–4 (2012): 24–32.

9. Sharon Johnson et al., "Key Issues for Resource Collection Development: A Guide for Libraries" (The Hague, Netherlands: International Federation of Library Associations and Institutions, 2012), www.ifla.org/files/assets/acquisition-collection-development/publications/electronic-resource-guide-en.pdf.

10. *Open Discovery Initiative: Promoting Transparency in Discovery* (Baltimore: National Information Standards

Organization, Open Discovery Initiative, 2014), www.niso.org/apps/group_public/download.php/14820/rp-19-2014_ODI.pdf.

11. *ANSI/NISO Z39.88 - The OpenURL Framework for Context-Sensitive Services (OpenURL)* (Online Computer Library Center (OCLC)), accessed January 20, 2016, www2.archivists.org/groups/standards-committee/ansiniso-z3988-the-openurl-framework-for-context-sensitive-services-openurl.

12. Sanjeet Mann and Sarah Sutton, "Why Can't Students Get the Sources They Need? Results from a Real Electronic Resources Availability Study," *Serials Librarian* 68, no. 1–4 (2015): 180–90.

13. Anita Foster and Sarah C. Williams, "We're All in this Together: Library Faculty and Staff and their Reporting of Electronic Resource Problems," *Journal of Electronic Resources Librarianship* 22, no. 3–4 (2010): 124–43.

14. Rachel A. Erb and Brian Erb, "Leveraging the Libguides platform for Electronic Resources Access Assistance," *Journal of Electronic Resources Librarianship* 26, no. 3 (2014): 170–89.

15. Eric Hartnett and Carole Thompson, "From Tedious to Timely: Screencasting to Troubleshoot Electronic Resource Issues," *Journal of Electronic Resources Librarianship* 22, no. 3–4 (2010): 102–12, http://dx.doi.org/10.1080/1941126x.2010.535736.

16. Kenyon Stuart, Ken Varnum, and Judith Ahronheim, "Measuring Journal Linking Success from a Discovery Service," *Information Technology & Libraries* 34, no. 1 (2015): 51–76.

Book Reviews

Elyssa M. Gould

Self-Publishing and Collection Development: Opportunities and Challenges for Libraries. Edited by Robert Holley. West Lafayette, IN: Purdue University Press, 2015. 185p. $29.95 softcover (ISBN: 978-1-55753-721-8).

Self-publishing has exploded over the last ten years, resulting in an entirely new ecosystem of self-publishing platforms, marketing options, and collection development tools. Major companies such as Amazon provide a relatively easy way for individuals to format and upload writing for public consumption, and the public's demand for these materials is increasing. This demand has created a new and challenging set of problems for librarians who would like to leverage the growth of self-publishing to improve library collections and services. The essays collected in *Self-Publishing and Collection Development: Opportunities and Challenges for Libraries* make a convincing case that academic and public libraries have significant opportunity to participate in this new publishing landscape while making clear that, particularly for academic libraries, the acceptance and active collection of self-published titles is still something of a work in progress.

As editor Robert Holley lays out in the introduction, this volume may be "the first monograph to deal with self-publication and its present and potential impact on libraries" (1). As such, the essays contained within cover a lot of ground, but cluster around three ways for librarians to interact with self-publishing: as collection development/acquisitions specialists, as facilitators for patron publishing, and as publishers themselves. Chapter contributions from both public and academic librarians as well as vendors and authors provide a broad spectrum of perspectives on the challenges and opportunities of self-publishing.

Several chapters contain a brief history or overview of self-publishing, noting its recent growth and the traditional reluctance of librarians to collect self-published work. Multiple authors note that public libraries tend to collect more self-published works than academic libraries and that the term "vanity press" has gone out of favor. References to the "stigma" of self-publishing appear frequently. It is clear that each author is trying to provide context for his or her contribution, but tighter editing of these sections may have reduced some of the repetition, allowing more space for discussing the issues and experiences that were unique.

Within the cluster of essays dealing with self-publishing as a collection development issue, Bob Nardini of ProQuest Books presents a discussion of the challenge of applying traditional vendor services to independent publishers,

highlighting the idea that many libraries may avoid collecting self-published books because there is no easy way to discover and acquire them. He imagines a future when vendors are able to create library profiles and bundle self-published titles for purchase in a way that is cost-effective and provides acceptable metadata for institutions. This is an interesting thought experiment, highlighting the fact that there are issues of quality control, scalability, and economy throughout the self-publishing and distribution process, and libraries are not the only institutions struggling to come up with workable solutions. This chapter serves as a useful glimpse into vendor priorities.

In contrast, the following chapter focuses on Ingram Content Group and reads much more like promotional material, including explicit directions for formatting and preparing material for upload to the IngramSpark self-publishing platform. This chapter sticks out in both content and tone, being more of a company overview and guide to getting started using the company's services than an academic exploration of self-publishing's impact on libraries. A few paragraphs outlining some of Ingram's partnerships with libraries eventually connect this section with the larger themes of the book.

Several chapters touch on the role libraries can play to facilitate patron self-publishing. Author Henry Bankhead outlines a successful partnership with e-book publisher Smashwords that allowed the Los Gatos Library to guide local authors through the publishing process and provide library access to their books via OverDrive. Public libraries likely already have a subgroup of patrons eager to write a novel, local history, memoir or family biography, and libraries can help patrons navigate self-publishing platforms and highlight the work of their patrons in the library. The chapters detailing individual author experiences with self-publishing reveal just how many options there are for those looking to self-publish and perhaps unintentionally underscore how much opportunity there is for libraries to guide potential authors through the dizzying maze of platform selection, book design, marketing and distribution.

The book is capped by a brief, yet wide-ranging and thought-provoking, bibliographic essay highlighting some of the recent research on the self-publishing landscape and its impact on libraries. Chapter author Joseph D. Grobelny writes that libraries have lagged behind book publishers and the public when it comes to interest in self-publishing but ends the volume on a positive note, writing that "it is worth taking the longer view that libraries will

most likely successfully adapt to the changed publishing environment" (177).

What is missing from this volume is in-depth discussion of academic libraries as publishers and the special considerations that might apply when publishing scholarly content via an open access journal or institutional repository. Two chapters deal with self-publishing as an acquisitions issue in academic libraries, but there are no chapters outlining university library publishing programs or the academic library's increasing role in the scholarly communication process. While both Donald Beagle and Grobelny mention the growing prominence of institutional repositories in academic libraries, the focus of the book as a whole is squarely on self-publishing through third-party vendors such as Smashwords, IngramSpark, and Amazon's Kindle Direct Publishing. Several authors rightly point out that public libraries are much more active in this area.

Taken in its entirety, *Self-Publishing and Collection Development* is a wide-angle view of the ways that self-publishing can impact libraries. Chapters vary from resource-rich guides containing practical advice and descriptions of self-publishing experiences to more philosophical explorations of the challenges of discovering and acquiring self-published works. At times this breadth can be a bit disorienting, as chapters jump from collection development to programming development to vendor partnerships. However, this eclecticism means that there is in some sense "something for everyone," from librarians struggling to locate, acquire and properly catalog self-published materials, to those who are considering self-publishing their own writing.—*Rebecca Brody (rbrody@westfield.ma.edu), Westfield State University, Westfield, Massachusetts*

FRBR, Before and After: A Look at Our Bibliographic Models. By Karen Coyle. Chicago: ALA Editions, 2016. 179 p. $50.00 softcover (ISBN: 978-0-8389-1345-1). *Editor's Note: as of January 2016, the full contents of the book are available for Open Access with a CC-BY license. See Coyle's website (kcoyle.net) for details.*

While many in the metadata creation community are familiar with the Group 1–3 entities described in the *Functional Requirements for Bibliographic Records* (FRBR), the historical context for FRBR as a bibliographic model is less familiar.[1] In 1990, the International Federation of Library Associations (IFLA) sponsored the Stockholm Seminar on Cataloguing. One of the outcomes of this Seminar was the creation of the FRBR Study Group whose purpose was to identify a minimum set of data elements necessary to satisfy the needs of users. Using this element set in the creation of records would both further facilitate the sharing of bibliographic records and reduce the cost of cataloging for participating institutions. The FRBR Study Group's final report has far reaching influence, including serving as the

conceptual model upon which *RDA: Resource Description and Access*—the successor to the *Anglo-American Cataloging Rules* as the content standard used by many libraries across the world.[2]

In *FRBR, Before and After*, Karen Coyle puts FRBR into context, both historically and with regard to earlier bibliographic models. But Coyle's book is not merely an analysis of FRBR and whether the FRBR Study Group built a model that successfully meets the objectives set forth by IFLA. Coyle also asserts that bibliographic models inform, and are informed by, the technology being used most prevalently at the time of the model's creation. In the introduction Coyle lays out her argument, writing, "This book looks at the ways that we define the things in the bibliographic world, and in particular how our bibliographic models reflect our technology and the assumed goal of libraries" (xv). Coyle's book succeeds as an analysis of the relationship between bibliographic models and technology and as an analysis of the effectiveness of FRBR as a bibliographic model.

In part one, "Work, Model, and Technology," Coyle lays the foundation upon which her claims regarding FRBR as a bibliographic model are built. She begins by introducing readers to the concept of work; drawing from the fields of philosophy, semiotics, and information science and even advancing her own theory on the topic. Coyle moves on to discuss modeling, including both a general discussion on data modeling and a more specific conversation about library data modeling. Finally, Coyle addresses advances in library technology from printed library catalog cards through the rise of the Semantic Web.

In part two, "FRBR and other solutions," Coyle builds upon the groundwork laid in the first three chapters as she addresses FRBR as a bibliographic model. She begins by recounting the history that led to the development of the FRBR Study Group and their report. Coyle follows that up with a detailed explanation of the entity-relationship model, which is the model used in the development of FRBR. She then offers a brief explanation of what is being modeled in FRBR. In Chapter 8, Coyle uses the objectives that guided the work of the FRBR Study Group as a measure of the effectiveness of the bibliographic model they created. She then addresses a few of the fundamental problems she has identified with FRBR, including the concepts of inheritance and disjoint. Coyle concludes the book by discussing the future of bibliographic description and the application of the FRBR model in the Semantic Web environment.

In chapter 3 of *FRBR, Before and After* Coyle addresses the parallel development of cataloging standards and library technology standards, writing "there is no interaction between technology standards development and cataloging standards" (44). Prospective readers might imagine that Coyle's book could serve as a bridge upon which catalogers and library technologists could stand while building

standards in closer alignment. However, synthesizing Coyle's deep and thoughtful analyses on the topic of bibliographic and data models may be a challenge for those without a background in data modeling. While Coyle addresses both the historical context of FRBR and the model's component parts, *FRBR, Before and After* should not be mistaken for a primer on the bibliographic model. Those looking for a basic introduction to FRBR will find Coyle's discussion of the both the bibliographic model and the data model upon which it is based too esoteric as to be instructive.

While the focus of Coyle's book is the development of FRBR, Coyle also attempts to contextualize bibliographic models and the technology that influence their development more generally. Each topic on its own is a significant undertaking, worthy of its own tome. Putting them together in a single book, especially one of moderate length, means that certain topics may be underdeveloped while other topics are repeated in multiple places. Coyle focuses most of part 1 on the general analysis of bibliographic models and data models, while part 2 focuses almost exclusively on FRBR. The reader may feel disjointedness between the two parts and wish that Coyle had focused exclusively on one topic or the other—especially when bits of information from part 1 are repeated in part 2.

Coyle shines brightest when she writes about library users. Throughout *FRBR, Before and After* Coyle returns her readers back to what should be the fundamental question of both cataloging standards and library technology standards: *Where is the user in all of this?* Coyle's most scathing critique of the work of the FRBR Study Group is the absence of users in both the development of the bibliographic model and in the final report that documents it. In chapter 8, Coyle writes, "For a study that was purported to be user-centric, the user's absence is notable" and goes on to state that "the FRBR Final report reads as a study *by* catalogers *for* catalogers" (106). Throughout the book, Coyle continually challenges readers to consider whether "find, identify, select, obtain" are the tasks that users most want to do when interacting with library technology. And, if not, what might the preferred tasks be.

In the afterword, Coyle writes "It is taken for granted by many that future library systems will carry data organized around the FRBR groups of entities. I hope that the analysis that I have provided here encourages critical thinking about some of our assumptions" (159). As an analysis of past and current developments in bibliographic modeling, *FRBR, Before and After* can be seen as an important contribution to the conversation about the development of bibliographic models and data models. If the reader applies to Coyle's text the same test that Coyle applies to FRBR—does this book meet its objectives?—the answer is a resounding yes. Coyle makes a cogent argument about the relationship between bibliographic models and the technology that both informs them and is informed by them.—*Erin Leach (eleach@uga .edu), University of Georgia, Athens, Georgia*

References

1. International Federation of Library Associations and Institutions Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records: Final Report* (The Hague: International Federation of Library Associations and Institutions, 2009), http://archive .ifla.org/VII/s13/frbr/frbr_2008.pdf.
2. *RDA: Resource Description & Access* (Chicago: ALA; Ottawa: Canadian Library Association; London: Chartered Institute of Library and Information Professionals, 2015); *Anglo-American Cataloguing Rules*, 2nd ed., 2002 rev., 2005 update (Chicago: ALA; Ottawa: Canadian Library Association; London: Chartered Institute of Library and Information Professionals, 2002).