Library Usage Data Collection and User Privacy

Emily Dust Nimsakont
University of Nebraska-Lincoln

One of the types of metrics with which librarians are often concerned is library use data. Collecting data on library usage can be very valuable to librarians, whether it is for our own use, for highlighting our importance to those who make decisions about our funding, for providing a part of accreditation documentation, or any number of other uses. These days, tracking library use can include tracking use of the library’s website, catalog, or electronic resources, using tools such as Google Analytics. My focus in this article is not the tools that are available for this, but instead pointing out that, no matter what tool you use, an important consideration to keep in mind is the privacy of our library users.

Privacy of user data should be a consideration in any scenario of data collection, but librarians especially have a duty to be concerned about this because of our professional ethics. Although our library patrons are probably used to the idea that others may access data about their online interactions, accustomed to seeing things like Amazon making recommendations based on their past purchases or ads popping up on Facebook for items they have searched for using Google, the expectations and professional standards for library data are different. The American Library Association (ALA) Code of Ethics states, “We protect each user’s right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted” (American Library Association 2008).

In a recent webinar titled “Being a True Analytics Advocate,” Tabatha Farney, Director of Web Services and Emerging Technologies for the University of Colorado-Colorado Springs Library, offered these suggestions for considering privacy when collecting library usage data:

- **Be aware of the data collected.** If you are using a tool like Google Analytics, be aware of exactly what data is collected, not just the data you are interested in collecting.
- **Purge unused data.** There is no need to retain data that you are not using for analysis.

Continued on page 27
2017-2018 Officers, Committee Chairs, and Representatives

TS-SIS
Chair:
Lauren Seney
William and Mary
Vice Chair/Chair-Elect:
Wendy Moore
University of Georgia
Secretary/Treasurer:
Diana Jaque
University of Southern California
Members-at-Large:
Carol Collins (2016-2018)
University of Tennessee
Emily Dust Nimsakont (2017-19)
University of Nebraska
Acquisitions Committee:
Marjiah Sroczynski
Milbank, Tweed, Hadley & McCloy
Awards Committee:
Kevin Carey
Ohio State University
Bylaws Committee:
Alan Keely
Wake Forest University
Cataloging & Classification:
Keiko Okuhara
University of Hawaii
Membership Committee:
Kerry Skinner
Arizona State University
Nominating Committee:
Katrina Piechnik
Jenkins Law Library
Preservation Committee:
Sharon Bradley
University of Georgia
Professional Development Committee:
Ajaye Bloomstone
Louisiana State University
Trina Holloway
Georgia State University
Serials Committee:
Patricia Roncevich
University of Pittsburgh

OBS-SIS
Chair:
Jennifer Noga
Wake Forest University
Vice Chair/Chair-Elect:
Jackie Magagnosc
Cornell University
Secretary/Treasurer:
Kevin Carey
Ohio State University
Members-at-Large:
Patty Alavyay (2016-2018)
University of Kentucky
Jason LeMay (2017-2019)
Emory University
Education Committee:
Jennifer Noga
Wake Forest University
Local Systems Committee:
Keiko Okuhara
University of Hawaii
Nominating Committee:
Scott Matheson
Yale University
OCLC Committee:
Emily Dust Nimsakont
University of Nebraska
Web Advisory Committee:
Barbara Ginzburg
Washburn University

OBS and TS-SIS Representatives/Liaisons
ALA MARC Advisory Committee
Jean Pajerek, Cornell University
ALAC Committee on Cataloging: Description and Access (CC:DA)
Robert Bratton, George Washington University
ALA Subject Analysis Committee (SAC)
Lisa Contursi, Columbia University
Funding Research Opportunities Grant (FROG)
Chair, Gypsy Moody, Belmont University (2017-2019)
Representative-at-Large, Rebecca Domm, Bass, Berry & Sims (2017-19)
OBS SIS Reps: Lisa Watson (2017-2019); Rachel Decker (2017-2019)
Online Bibliographic Services Special Interest Section

With the new year well underway and Spring on the horizon, this is a time of renewal and anticipation. I share in that feeling of expectation, not only for the changing seasons, but also for Online Bibliographic Services Special Interest Section’s (OBS-SIS) prospects and upcoming opportunities. Lately, OBS-SIS members have been busy on many different fronts. In this column, I will provide updates on those endeavors and the many things we have to look forward to this year.

**OBS-SIS Programs in Baltimore:**

First off, the OBS-SIS Education Committee met several times in 2017 to work on proposals for educational opportunities, requested by you, for the 2018 Annual Meeting. Committee members submitted four excellent and timely program proposals, and in November, we learned that two of the programs were accepted to be on the schedule for Baltimore in July. I will talk more about the specifics of those programs and other OBS-SIS related events in my next column. For now, I will just say: if you are interested in learning about APIs, OpenRefine, and system changes, among other things, start planning now to be in Baltimore.

**Repositioning Efforts:**

The Special Committee on Repositioning, which I discussed in my last column, has been hard at work and has already made some initial recommendations to the OBS-SIS Executive Board about possible updates to our section’s mission and purpose. Those recommendations will be discussed in an upcoming Executive Board meeting. I will share any decisions that come out of the repositioning efforts as soon as I can. In the meantime, let me reiterate that if you have any input about the focus and future of OBS-SIS, please communicate those to any member of the special committee or to me.

**OBS-SIS Elections:**

Finally, it is election time! Let’s take a moment to be grateful that our election cycle doesn’t involve endless campaigns, partisan bickering, and arguments with friends and family. Instead, our elections typically involve several dedicated members who are willing to put their time and energy toward supporting OBS-SIS and its goals.

This year, the OBS-SIS Nominations Committee put forward excellent candidates for Member-at-Large and Vice-Chair. We owe a huge debt of gratitude to the members of that committee: Chair Georgia Briscoe, Katie Dunn, and Angela Griffin. I also want to thank all the members who considered running. If you were not able to commit this year, no doubt there will be plenty of opportunities in the future!

Here are the candidates for 2018-2019:

**Vice Chair:**
Jessica Pasquale
Head of Scholarly Publishing & Intellectual Access
University of Michigan Law Library

**Member-at-Large:**
Barbara Szalkowski
Core Operations Librarian
South Texas College of Law Library

These candidates have already been very active, contributing members of OBS-SIS and the American Association of Law Libraries (AALL). We are fortunate to have them both as candidates for officer positions. Thank you so much to these two members for being willing to give time and effort in support of our section. Keep an eye out for more details about the election on the OBS-SIS community list.

As you can see, there are several new and exciting things going on with OBS-SIS, and we have much to look forward to in 2018 and beyond. I hope everyone will soon be enjoying warmer weather and new adventures as we approach Spring.

*Jennifer Noga*
*Wake Forest University*
A New Year, a New TS-SIS

The start of the New Year usually brings resolutions: to better oneself, to improve relationships, or to take a new outlook on life. 2018 is shaping up to be the year that the Technical Services Special Interest Section (TS-SIS) lives up to its resolutions. We spent much of 2017 taking a hard look at the Section and talking about how to improve communication, relationships, and the value of TS-SIS to its members. There was a report put together by an ad hoc committee, discussions at annual meeting, and more discussions on the listserv in the months that followed. What I heard was that the members of TS were ready for a change.

What I can tell you now is that change is coming. The TS-SIS Executive Board started the year by approving multiple changes to the structure of the Section. We will be reducing the number of standing committees from five to two—one is Description (formerly Cataloging and Classification) and the other is Resource Management (formerly Acquisitions, Serials, and Preservation). These new standing committees will have co-chairs appointed in alternating years. We will be moving the roles of the Professional Development Committee into the realm of administrative committees.

Administrative committees will continue to have a static membership for each year, but standing committees will have membership that fluctuates throughout the year. The existing advisory groups, working groups, and roundtables will continue under the auspices of the corresponding standing committee. Additional ones can be created as necessitated by the changing roles of TS librarians or specific projects that require a dedicated staffing.

The annual volunteer survey will also cease with calls for volunteers being coordinated by each committee chair. Administrative committees will make a call after the Annual Meeting, and the membership of their committees will be approved by the TS Board as these committees have an expected work product each year. Because they will be more discussion-based, standing committees will make calls for membership throughout the year. We will eliminate full membership rosters from the website, leaving the chairs as the point of contact for each committee.

All of the changes were voted on and approved by the Board during our January Conference call. The minutes to that meeting are available at https://www.aallnet.org/tssis/wp-content/uploads/sites/17/2018/02/TSSISboardminutes20180123.pdf

One outstanding item that the board will vote on this Spring is whether to continue the use of “Standing Committee” or to go with a more descriptive name such as “Forum.” If the Board approves this measure, a Bylaws update will be necessary, which will be followed by a vote from all of the membership.

I welcome your feedback during this transition as we work to make TS-SIS more fluid and adaptive.

Lauren Seney
William and Mary
A Multi-Part ILS or Management System Needs Exploration for Acquisitions, Part 2: Selection & Searching, Ordering & Order Maintenance, and Receiving & Claiming

Luz Verguizas
Columbia University

In Part 1 of this article, published in the December issue of Technical Services Law Librarian (TSLL), I discussed the need for standards for acquisitions, as well as the functional requirements which would form the basis for my investigation into the needs for an acquisitions module of an integrated library system (ILS) or library management system. The assumption here is that these are needs that should or may need to be considered for any acquisitions team, whether considering migration from an existing system or building up a homegrown system using primarily spreadsheets and/or databases (i.e., Access).

Selection & Pre-Order Searching

The first step in the acquisitions process is always selection; the acquisitions unit receives a request for a particular item. Depending on the institution, this selection can originate from a number of different sources, including but not limited to: a subject specialist in the law library, a faculty member or attorney, a student, or non-affiliate patron. How are requests currently received? Are they sent via email, a web form, paper catalogs or selection slips, or an ILS selection list? Many ILSs offer the ability to create selection lists. These can be shared between acquisitions staff and can be made visible to other staff members for collaborative input. Some also have the ability to create order records from selected items on the selection list. If your library would like to reduce its paper trail and rely primarily on electronic communications with the ILS handling this traffic, selection lists may be a way to achieve this.

Staff members reviewing any selection then need to determine whether or not the item is already held by the law library (or already on standing order for the series), or if this is a request for an additional copy. For pre-order searching, searching capability of a system is an important feature. For example, can the system easily locate duplicate titles? With any system, you should be able to easily search and verify bibliographic information.

Order & Order Maintenance

After an item has been selected and searched for in the local catalog, an order record needs to be created. For your institution, what are the crucial elements of the order record? What type of information can/should be stored in the order record? The order record should easily identify the item requested, with such information as price (in local or foreign currency), vendor, number of copies, and format. It should also show the order type, meaning whether or not this is a firm order, continuation, or subscription. The following questions should be asked:

- What fields are available? Are there local/customizable fields?
- How can records from a vendor or bibliographic utility be imported?
- Is there a standard such as MARC used?
- How does the order data interact with other data in the systems, such as vendor records and catalog records?
- Does the order data display to the public? If yes, is this optional?
- Is there an ability to create order templates, and, if so, is there a limit to how many can be created?

(Wilkinson 46)

At their most basic, there are two types of orders: firm orders and continuations. While firm orders are one-time purchases, encumbered and subsequently disencumbered after payment and involve no continued financial commitment, serials are the complete opposite. Serial order records represent ongoing financial commitments, and thus, their order records should be able to manage multiple payments over the course of time, whether that time period is one year, twenty years, or more. Some important considerations for subscriptions, as well as those libraries considering an integrated or stand-alone ERMS (Electronic Resource Management System):

- How are subscription bundles, such as packages of journals from a publisher, represented?
- How are consortial agreements represented?
- What are the options for representing deals such as electronic + print?
- Is there a knowledge base available, and which vendors populate it?
- How does the ERMS interact with the rest of the system?
- Is there a way to store license data in the system?
- How is the license manager organized?
- Is it searchable, and how?

(Wilkinson 46)
Receiving & Claiming

Whatever the acquisitions process at your institution, the receipt of ordered items will need to be tracked. For firm orders, the process is usually straightforward, with payment remitted upon receipt of the item, as most institutions will not remit payment to a vendor for a firm order until full receipt of the ordered item. It may also be important for your institution to be able to track order delays, as with items that are pre-orders, out of stock, or rare, hard to find materials.

Tracking serial subscriptions is much trickier. Although most ILS providers have individual modules dedicated to acquisitions and serials, interoperability between the two is very important. Acquisitions and/or serials staff should be able to set the frequency of publication for individual titles and corresponding expected dates for new issues or volumes. Doing so can aid librarians in assessing vendor and/or publisher performance at, or prior to, the renewal period. With regard to receiving, the following items may also be considered:

- Does the system have the ability to receive partial orders?
- When receiving items, do the staff members involved in the workflow have the ability to view order, serials, cataloging, and invoicing data all in one place?

Over the years, due to decreasing budgets, reduced staffing, and space restrictions, many libraries have abandoned binding or claiming items for non-receipt of serial subscriptions, or they have cancelled a good portion of their print subscriptions in favor of electronic equivalents. For those libraries that still send out materials for binding or that process claims, a robust serials module capable of accurately predicting receipts or claims, or when items need to be sent to the bindery, can be invaluable in terms of saving money purchasing replacements later on and staff time spent on individually reviewing titles to be bound or claimed. If your library does process claims, following are some additional considerations:

- Is there a configurable default claim cycle for unreceived items?
- Does it have the ability to configure specific claim cycles for some vendors?
- Does it offer the option for electronic claiming for vendors who support it?
- Does it have the ability to force a claim outside of the defined claim cycle?
- Does it have the ability to claim the same issue or volume multiple times?

The functional requirements detailed above, selection and pre-order searching, order and order maintenance, and receiving and claiming, are some of the most basic considerations for any library acquisitions system. While the questions suggested are in no way meant to be exhaustive in scope, they should help to start the conversation when discussing the needs for an acquisitions system. In part 3 of this article, I will discuss, among other things, payments & finance, reporting, transfer of existing data, training and communication.

Reference List


Pollution is messy

Consider a trip through time to the deep dark time before time (i.e. when there were no “K” schedules, 60+ years ago). Pollution was already an established subject of legal discourse. The laws on tanneries (a stinky business going back to antiquity) and the theory of “nuisances” in tort law were also well established. Pollution often became a “hot topic” when it resulted in people dropping dead in the street (and yes, it was much worse back then). Conservation of natural resources was a separate issue, primarily dealing with whether the economic potential of public lands should be exploited or conserved for future use (perhaps reflecting the idea that what you did on your own private property was not a proper subject for regulation). Doctors sometimes prescribed smoking tobacco, and it was often encouraged by government for its perceived health benefits. Science fiction primarily addressed climate change, and usually was about a future ice age. This is the period when the KF schedule was born.

Environmental law. The term was not included in the original drafts of the KF schedule. Eventually added to the caption for “Public health” (KF3775-3816), it now reads, somewhat awkwardly, “Environmental law. Public health. Including environmental pollution and sanitation.”

At the time, dealing with air and water pollution was a public health issue. “Conservation” is indented under “Public lands” at KF5505-5510. It made sense to the people who created the schedule (none of whom are still alive), and the same problems occur in the other older schedules, including KE (Canada) and KD (British Isles).

The more modern schedules look different. In KIE (indigenous peoples in the United States), a 21st century schedule, Environmental law (KIE1460-1511) is highest level topic, coming after (rather than indented under) Public Health and preceding Medical law. Conservation is included under Environmental law, as is pollution, though some aspects of sanitation, such as drinking water, are in public health, immediately preceding the Environmental law area. The same pattern is present in the Afro-Asian tables (the last of the geographic jurisdictions to be completed, which date from the late 20th century), KL-KWX4 3127-3135 and KL-KWX6 1507-1517.

For someone with a 21st century weltanschauung, “Environmental law” is a subject distinct from public health, and it includes protection of natural resources as well as prevention or remediation of pollution. While it may be too much to change the older schedules, we need better cross references to assist those who really don’t remember, or cannot relate to, the way things were several generations ago.

Tobacco. Today, general agreement is that tobacco is detrimental to one’s health, and the logical place to class it would be in the public health area or the medical law areas, perhaps indented under a caption for popular, unhealthy, and addictive substances, including alcohol and other “recreational” drugs. That was not the perception of tobacco a century ago. “Tobacco” and “Smoking” were not originally addressed in KF, neither under public health nor in the “Food drugs and cosmetics” areas.

In time, KF added two numbers for tobacco: one indented under “Air pollution. Control of smoke, noxious gases, etc.” at KF3812.3, and another is at KF3894.T63, indented under “Drug law.” The former number is older and made perfect sense at a time when the principal objection to smoking was that it smelled bad (it was rude and annoying, but not a health risk). Over time, the perception of tobacco as a dangerous drug has occurred. There is no justification for both numbers, and few people today are thinking of “air pollution” when they discuss tobacco usage.

In “K” general, “Tobacco smoking” is at K3593.5.T63 under “Air pollution. Transboundary air pollution regime” (consider if Canadians ever objected to second hand tobacco smoke floating over the border). In the most recent of the country schedules, the tables for Afro-Asian jurisdictions, the civil law table has a number for “Air pollution, including ... tobacco smoking” at KL-KWX4 3130.5, and both civil and common law tables have numbers for tobacco indented under drug laws (KL-KWX4 3095 for civil and KL-KWX6 1538.5.T62 in the common law table). The slightly older European civil law table has smoking under air pollution (KJ-KKZ1 3130.5). The German schedule has it under “Drug laws” at KK6196, but that is a recent addition. The KB (religious law including Jewish and Islamic law) also has smoking under the drug law numbers. The older common law schedules (KE and KD) are a similar treatment as KF, i.e. smoking as air pollutant rather than drug with public health issues.

In KIE, the most recent schedule, at KIE1539, tobacco is indented under Drug law, between “Narcotics and other intoxicating drugs” and “Alcoholic beverages.” This is where it belongs. Given the clear pattern of moving the subject to public health
and away from air pollution, I suggest that when cataloging a work on tobacco smoking, one should class it in the public health area, even if one has to make proposals to cancel the air pollution number for tobacco smoking and create a new number, indented under or adjacent to drug laws.

Summary. The real world is constantly changing, and the legal system eventually reflects changes in the real world. At present, “Environmental law” is a distinct specialty to which are devoted textbooks, law school courses, specialized journals, professional associations, etc. It includes dealing with pollution, conservation of natural resources (including biodiversity, sustainable development, etc.), as well as “climate change.” We have to be aware that paradigms can switch, which is what happened to environmental law and tobacco smoking, and we have to modify our schedules to reflect contemporary users’ expectations.

This is not an official policy statement from LC.

---

**Collection Development**

Practising Law Institute leaves
Bloomberg Law

Adrienne DeWitt
Campbell University

In December 2017, Practising Law Institute (PLI) ended its contract with Bloomberg Law. PLI’s online treatise library is now on their standalone website, PLI Plus (https://www.pli.edu/Content/Static/Online_Research_Database/_/N-1z128ex). The PLI Plus website includes 90,000 legal research documents covering 25 practice areas and is the new home of PLI treatises, coursebooks, answer books, legal forms and program transcripts. According to VP Craig Miller, PLI has no intention of licensing with any other third party publisher in the near future.¹

This could prove problematic for libraries that have either downsized or ended their looseleaf collections. A substantial number of PLI treatises are looseleafs; if those looseleafs have not been updated, libraries would need to either purchase new volumes in print or add PLI Plus to their list of online resources.

Speaking from my experience, independent legal research platforms can be a difficult sell to law students. Legal writing and research programs focus on Westlaw, Lexis, and Bloomberg Law, often leaving students unaware or uninterested in other legal research platforms. Marketing independent platforms through workshops, social media, and word of mouth is helpful, but that might not increase usage enough to justify the cost.

PLI’s recent move highlights an important question for Collection Development librarians. Should we cut a print looseleaf resource because it is accessible online? On one hand, looseleafs are bulky, hard to upkeep, and can be expensive in both updating subscriptions and manpower. At the same time, keeping looseleafs can give libraries time to evaluate an online research platform without being pressed into quickly subscribing to a licensing agreement. Plus, the timing can be problematic. Budgets are often completed by a specific date, and new subscription purchases may have to wait until the next budget cycle. Should a research change platforms at an inopportune time, patrons may have to wait a long time before the library restores access.

Weighing the cost of upkeep for a print looseleaf versus the price of an online subscription is a difficult call, especially in these austere times. However, PLI is an excellent example of the importance of keeping holdings in multiple formats. Maybe it’s not such a bad idea to keep print looseleafs, after all.

Summary of December 2017,
Q4 Legal Cataloging Forum

In this column, I am delivering my write-up of the Quarter 4 Legal Cataloging Forum, as hosted by Yan (Clara) Liao at Library of Congress’ (LC) Law Library.

There were two subjects discussed at the forum:

1. Date of promulgation of a law, etc. (presented by Robert Bratton of Jacob Burns Law Library at George Washington University Law)
2. New 7xx in Bib Records and 3xx in authority records (presented by Aaron Kuperman of Library of Congress)

Date of promulgation of law, etc.

Bratton quotes Resource Description and Access (RDA) 6.20.2.1: “date of promulgation of a law, etc.: A year a law, etc., was promulgated or brought into force.”

Referencing the 4th edition of Cataloging Legal Literature (CCL4), he pointed out there are a multitude of dates that could be used to define the date of promulgation, but that CCL4 describes the proper date as the enacted date, not necessarily the effective date.

Bratton claimed this standard for entering dates for promulgation conflicts with Black’s Law Dictionary, which defines promulgation as:

“Promulgation. The official publication of a new law or regulation by which it is put into effect. Promulgate vb. 1. To declare or announce publicly; to proclaim. 2. To put (a law or decree) into force or effect.”

Even these definitions seem to have a conflict—wherein “to proclaim” may stand in for “to enact” in the RDA definition—but this conflicts with the second part where it defines “promulgation” as “to put into force of effect.” The slipperiness of language is evident here as “force or effect” may have a technical meaning requiring further use of language to define.

The analogy would be when we talk about getting people “access to information” generally, this can oft time refer to the technical meaning of computers connecting to other computers (servers) in order to download a file to the local machine or browser—or it could be a more abstracted meaning of being “free” to get information. A case could be made along that spectrum for sure—and has.

I feel we are in a similar situation here—and the discussion that ensued as we talked this out revealed that to be true.

Bratton went through a few examples of how a law or a treaty has worked its way from approval to enactment—sometimes taking more than a year—which adds some confusion as to what the date is from the point of view and accuracy. It seems even dates are philosophical constructs, because even with the date a law was signed into being, even without year or date of enactment, RDA encourages us to think with the user in mind and use associated dates—for the purposes of retrieval.

Yes, it can get fun and less clear the more we go. But that is the good stuff.

New 7xx in Bib Records and 3xx in authority records

Kuperman spent some time discussing authority records and the input of 7xx in bib records and 3xx fields in authorities in order to increase linking potential because the technologies, along with RDA, are increasingly moving this direction—that of defining and creating relationships between works, entities, and abstractions.

Regarding 7xx fields in bib records, Kuperman spoke about the values of adding the newly updated 780, 785, and 787 fields to original MARC records (and to enhanced older records if time allows).

780 and 785 are, ideally, reciprocal relationship fields that designate a title as superseded by something else and, that the new title supersedes.

Just to glance at the complex relationships implied by 2nd indicators for 780/785, please take a look:
These indicators do not even get to the subfields in which relationship designators would be entered.

The ambiguity, and thus creativity, has yet to show itself fully with these relationship fields in (no, new) standard MARC, especially with 787 – Other relationship Entry.

787 is defined and scoped thus:

“Information concerning the work related to the target item when the relationship does not fit any of those defined in fields 760-785.

In most cases, a note is recorded in field 580 that defines the specific relationship.”

The small number of examples on LC’s Standards page for MARC 21 Bibliographic Data suggests, at least to me, that this field has yet to be explored to its potential.

[to the future…]

Kuperman also wanted to encourage authority producers (specifically NACO contributors) to make use of the range of 3xx fields in Authority data.

Some of these would be:

370 - Associated Place (R)
371 - Address (R)
372 - Field of Activity (R)
373 - Associated Group (R)
374 - Occupation (R)
375 - Gender (R)
376 - Family Information (R)
377 - Associated Language (R)
378 - Fuller Form of Personal Name (NR)

The extent of use of these fields determines much about the use of these records, but one thing in particular seems important: using these fields also means a growing number of links and references to other files (SACO, NACO, and various vocabulary thesauri such as LC Demographics Thesaurus & Genre/Form—both of which are fairly new themselves). However, the actual thesauri referenced in these fields under $2 could expand to others not hosted by LC, such as Getty’s Art and Architecture Thesaurus. All that really matters is that the correct code is added to delimiter ($) 2 so the user or machine can determine from which file the term originates.

Is it nerdy of me to suggest that this is a fun thing to think about?

It was a good forum in December 2017. Big thanks to Robert Bratton and Aaron Kuperman for their presentations to the forum.

Thank you for reading.
Long Term Financial Costs of Maintaining a Digital Collection

Wilhelmina Randtke
Florida Virtual Campus

Introduction
The point of this article is to give an idea of what costs go with long-term maintenance of digital collections. Libraries have moved most access to materials from print to electronic. Understanding maintenance costs, as opposed to initial costs to build a digital library, helps to understand long-term implications of the shift. This article is heavily academic oriented, because academic and state libraries/archives have traditionally kept long-term historic materials.

With print materials, storage costs are visceral. We can see and feel linear feet on the shelf. Rent and air conditioning are things we pay for ourselves in daily life and which must be paid on a storage facility. However, digital materials conceal costs. Home photos taken by phone may be the closest we come in daily life to understanding issues involved in long-term maintenance of digital objects.

Shift from Print to Electronic
The past 20 years—20 years!—has been the same story repeatedly in library archives. Electronic material becomes available. Libraries rent access to it but cannot own paid electronic packages. Libraries catalog, but they do not harvest open access electronic materials. Now the library has it online, so purge the books and have a nice physical space, even perhaps a nice physical space for a very different department than the library.

Acquisitions Decisions are a World Apart from Digital Preservation
Meanwhile, just-in-time purchasing has been the replacement for just-in-case hoarding of materials. The library profession has touted the logic of this with the catch phrase “patron driven acquisitions,” or PDA. The Internet has made it easier to find and buy out of print materials. You can buy an electronic resource at a faculty member’s request and have nearly immediate delivery. There is an idea hanging in the background that electronic materials never go out of print and that future price fluctuations will occur regardless of what a library does this year.

Proof that long-term preservation is divorced from the acquisitions department is right there in the buzzwords. PDA is also short for “personal digital archiving,” a huge catch phrase in the world of digital preservation. There are many conference presentations on “PDA,” with two very different meanings in two very different library worlds. A reused acronym is perfectly clear and unambiguous only as long as the two fields remain separate. Whenever you see the acronym “PDA” in use, this means that acquisitions and digital preservation have not yet met.

Here, I am using PDA to mean patron driven acquisitions, the idea of buying materials on demand rather than buying and archiving “in case” someone needs it in the unknown future.

Just-in-Time Purchasing: The Assumption that Materials Can Be Purchased
Just-in-time purchasing builds on the assumption that materials, once created, may always be acquired and purchased.

For current info, this probably works. The market can decide each year what exists. The publisher’s choice is always: Create it or not? The library’s choice is always: Buy it or not? The desire to buy happens in the same year that the availability or unavailability of the resource is apparent, so if something important disappears, there is a chance to voice concerns, know why, and even make a change.

For older info, the market cannot decide each year what exists. The choices are: Keep it available or not? and How much does it cost this year? However, once the decision is to discard it, there is not a way to go back and undo that. You need a copy to make a copy. If no one keeps a copy, then it is gone. Once an individual library decides not to retain ownership that decision is out of the library’s control. (Remember, there is a large amount of content that libraries rent rather than own.) Once everyone decides not to retain, that decision is permanent. Awareness of that decision may come long afterwards and long after there is time to voice concern.

Historically, libraries used to keep the backlog of older material. “Publishing,” at heart, means making the resource available. It used to be that a commercial publisher/printer would make something available short term. Libraries and consumers
would purchase. Then, once material was out of print, libraries made the resource available long term. Libraries for a long
time had a very concrete role in acting as the long-term archive of older materials, making those materials available for
centuries to come.

Outdated material matters for law practice. Professors studying legal history are one thing, but the Kafkaesque nature of
law throws many very relevant curves. If you have someone with a prior conviction, but court records do not state the
degree of the conviction, then you need the older statute to get the degree of the conviction. That matters for priors on a
subsequent conviction and for background checks of employment. Older material matters when a discovery rule starts the
statute of limitations running. Think of pollution—once there was no regulation of pollution, then more and more materials
have become regulated. Statutes of limitations run on the discovery rule. If someone this year discovers pollution in their
neighborhood and the dumping was long ago, then the statute of limitations starts now, but the legality of the pollution
depends on older laws in effect on the date of the pollution.

The premise of PDA is the idea that nothing is really ever unavailable.

However, the PDA model of purchasing puts long-term preservation of resources with vendors. Historically, the publisher/
vendor never could make the decision not to retain, but now can make that decision.

**Maintenance Costs and Vendors**

Online delivery of material is relatively new. The release of the first web browser was in 1993, less than 25 years ago. There
was an initial phase of adoption connecting enough people creating an audience for Internet material. Most early digital
library buzz was about digitization or producing newly available digital collections. It is far more recent that focus has
shifted from launch to long-term preservation.

Understanding costs in keeping digital material available may help to understand the issues here.

Vendors exist for profit. Profit is generally part of corporate by-laws. (Yes, many vendors are nonprofit, but copyright
concerns mean each resource stays with the owner until the sale of those rights.)

Think about cancellation decisions by libraries. (Electronic “acquisitions” decisions often instead tend to be electronic
“cancellation” decisions.) Analytics play in. It makes sense for a library to cancel less-used content. Similarly, for a publisher,
it makes sense to cancel and no longer offer for sale less-used (i.e. less profitable) content.

**Maintenance Costs in Running a Digital Collection: Computer Programming, Systems, and Testing**

This section is about costs to operate the platform, as opposed to maintenance costs of the content. In general, a software
platform has an economy of scale, where maintenance costs spread over content and more content does not have a big
impact on costs.

*Costs of Not Maintaining Code*

All digital library platforms are written in computer code. That code lives on a server. Over time, deprecated versions of
software run on servers, and, along with that, servers must be upgraded. The server configuration allows the code to run,
and a newer version may not be backwards compatible with regard to what the code allows. In fact, PHP, which is a very
popular language for web interfaces, is not backwards compatible. Code that can run on a PHP 5.4 server probably needs
some lines rewritten to be able to run on a PHP 5.6 server. You can theoretically keep running a server with an outdated
PHP or MySQL, and so keep on going with no changes to the platform, but running older server configurations opens up
maintenance issues in terms of security and eventual difficulties in sourcing hardware. For example, if you purchase a
physical server, it may have some limitations. If you purchased a physical server long ago and need a part now, that may be
harder to source and more expensive than a new server. If you rent a managed server, vendors will have rolling cut offs for
what versions of PHP, Structured Query Language (SQL), Java, etc. the vendor will support for you. Eventually, running a
very old server architecture may involve significant surcharges. Troubleshooting and tech support may become unavailable.
Going out of date costs money. Long term, not rewriting code is expensive.

Going out of date may also not be feasible as the community around you moves on. Users have expectations about how a
site should look, and those expectations change as the entire web changes. Having a site that looks a bit outdated can be off-
putting. Having a site that works in an outdated way can be a deal breaker and may prevent users from getting to material.
For example, consider sharing out metadata records. Fifteen years ago, Z39.50 was looking to be a leading standard for
sharing metadata records. Today, Z39.50 is still in use but out of favor. A fifteen-year-old platform designed to share records
exclusively in Z39.50 likely could have participated in federated search projects more so in the past than it could today.
That is because harvesters built around Z39.50 are not the norm now, and even for harvesters that support them, the skill and experience to troubleshoot or test is scarce. If you are running an archive, and the world has moved on, it is out of your control to fix the problem without changing your archive.

Costs of Maintaining Code

Keeping up-to-date costs because you have to rewrite code—there is no getting around that.

In open source, the wider community is keeping the code up, but once you use it, you are part of that wider community. You can contribute code. Adding code to a project requires a huge time commitment in order to understand the wider project fully and to make edits in an appropriate or cohesive way. It cannot be an afterthought and likely would require a dedicated position within the organization. That is a significant cost. You can contribute software testing, which is actually a huge part of debugging and huge time commitment. (Hint, hint, law libraries: this is a skill that might be easier to staff, and software testing is a skill that dovetails well with legal research. It is still a deep time commitment in order to notice the significant details that matter, but that detail oriented work to really understand how the software works for a researcher overlaps with reference and teaching activities.) Alternatively, you can contribute funds to the board managing the open source software, by a membership fee or by pledging to specific projects where multiple institutions each pledge a small amount and once achieving the goal, there is the development funding of a new feature.

In proprietary, the vendor is keeping the code up. The vendor uses salaried staff and contractors to keep software up-to-date.

There is quite a bit of overhead in keeping a whole platform running over time. Libraries can somewhat see this when getting prices for institutional repositories or digital publishing platforms. Some of the pricing is training and support, but not all of it.

Maintenance Costs in Running a Digital Collection: Metadata

Metadata maintenance comes up in two areas: migration to a new platform and keeping up with changes in the wider world. Often, this cost scales up with the amount of content. More content leads to higher costs.

Migration and Metadata

Migration is a necessary part of long-term management of digital collections. Digital content lives in software. Eventually, the software will change. Then all that content has to move over without getting all messed up. Library catalogs, which technical services is familiar with, have been in Machine-Readable-Cataloging standards (MARC) for decades, and it is glaringly obvious that MARC is the dominant standard.

However, in digital material, it is not like this. There are several widely used metadata standards for digital collections: Dublin Core, Metadata Object Description Schema (MODS), qualified Dublin Core, and others. All those standards are common, and all are very different, and it is possible to build a digital library platform around any of them. On top of that, it is also common for a platform to keep indexing information in something that is totally and completely specific to the software. For example, most metadata schemas do not have a place to represent statistical information on downloads, but some harvesters require statistics, so statistics tend to be pulled from elsewhere in the software. Also, think about a collection that an object is in. Collections can often be represented in metadata, but almost all digital library software stores, updates, and pulls collection information from something other than the metadata record. In addition, it is difficult to represent serial-specific information, such as volume, issue, etc., in any library metadata schema, so it is common for platforms oriented around journals to handle that information in ways specific to the software.

What this means is that if you change software, then you have to look closely at the metadata for each platform. Digital library platforms do not necessarily store metadata in a schema. They almost certainly will output metadata in a schema, mapped from something internal. Therefore, the material may behave as desired in the software, and you can export records in Dublin Core and in MARC, but that export may have concatenated two similar fields. The fields are stored separately, because one of them is not kept in a standard schema; the distinction is important, but to be able to export in a standard, they must be concatenated, and the distinction is lost. When you move to a new software, you will have to do a deep dive into each software and what metadata each is using to make sure you can keep certain distinctions. Even if both platforms are using a standard, there is probably some important metadata used for the search and interface not stored in that metadata schema. This metadata search and interface runs in a customized way specific to the platform. Staff time to make that deep dive is a significant cost, along with a programmer’s time to implement recommendations.

Focus and thought to migrate is often collection-specific, because the use of each collection of materials is slightly different. Think about researching statutory history in all 50 states. There is a general process of how you do it and what matters, but the details vary from state to state, and the details vary within a state over time.
There are more layers than that. Think about looking at pages of a book online. Each page is a picture. While you experience a book, the software must organize and present hundreds of individual pages. There is a standard for keeping pages in order, Metadata Encoding and Transmission Standard (METS) using the structMap elements in METS. METS is an Extensible Markup Language (XML) based standard. There is no quick query of XML, so if information is stored in XML, then it has to be pulled into something else in order for a site to operate and show pages quickly enough not to bore users. All software is really running on an index made from the XML. Software is using that index, not the XML, and the XML can seem superfluous in a self-contained system, such as your current library platform. Within a contained system, the XML is baggage and adds an extra step. Inertia cuts against writing software around XML. For page order, it is the norm to store information in something custom, so a migration to new software is extremely likely to require mapping from one standard to another. Moreover, if you are thinking about Portable Document Format (PDF) as a solution, PDF is a file wrapper and a PDF of a book is holding pictures together in a specific order while at the same time not necessarily being an open standard. The International Organization for Standardization (ISO) standard, PDFA, is tremendously limited in terms of what can be stored in PDFs in general. Happily, page order is usually a system-wide puzzle and tends to require a flat amount of time, regardless of how many items or collections exist.

**Metadata Enhancements**

Just as software updates can keep a site from looking long in the tooth, there are times when metadata needs overhauled to keep a once-current site from looking and behaving stale. For example, ten years ago, lacking faceted search was not a huge deal, but at this point, it is strange if a site does not have facets. Metadata that perfectly supports search does not necessarily support facets. Often, in adding nuance to facets, some systematic clean up is needed across a set of records.

There is no getting around touch ups, not even for perfect metadata. Think about the Virtual International Authority File (VIAF), which is Library of Congress’ controlled vocabulary and identifier system for disambiguating author names. Users expect to be able to distinguish by name—they expect to be able to separate “Tom Wolfe” versus “Tom Wolfe” by facets, by browsing from a single record and elsewhere in the interface. Twenty years ago, for a common name, you would have looked to the Library of Congress Authority record and then pulled in the demographic info to the level of detail that would allow you to differentiate. For example, you might add birth and death dates, middle name, etc., to the author in the record for the item. Then you would periodically check over all the records for an author and make sure that author was always indexed the same (i.e. if one record has a middle name, then all records have a middle name, never just a middle initial and never no middle name). In addition, of course, no misspellings. VIAF uses an author identifier to disambiguate similar names and then looks to the Library of Congress Authority record to pull the demographic info. Let us assume that in ten years, it has become standard to facet on name, to use VIAF to disambiguate similar names, and that VIAF integration has become an essential building block to a clean interface. That will require a metadata clean up and enhancement of existing records. Even a digital collection with perfectly consistent internal records would have to go over records and implement VIAF in order to integrate well into a federated search project. (After all, when you purchase electronic resources and get MARC records to boot, your catalog is the federated search project.) Internal consistency definitely saves time, but it does not eliminate the need to make changes over time and the need for a skillset to manage the change.

Metadata is as much a specialized skillset as any other tech field, and it is just as inscrutable to outsiders. Complexity in a schema is harder to maintain, and there is a balance between enabling search interfaces and still having a manageable set of metadata. There is a whole world of standards to navigate and understand, and a quick Google search is better than nothing, but not as good as deep background knowledge and understanding. It costs money to get access to that skill through staffing or through outsourcing.

**Conclusion**

As much as this article has been long and detail oriented, the key point here is that maintaining digital collections over time costs money. Understanding the nature of these costs and why they are unavoidable helps to know what you are buying beyond content when you maintain your own digital collections and when you rent platforms from vendors. It also helps in understanding the reasons why less popular and hence less profitable digital content might disappear from the marketplace.
Building strong culture with Daniel Coyle’s The Culture Code

What could the U.S. Navy SEALs, Pixar Animation Studios, the San Antonio Spurs of the NBA, and the online retailer Zappos all possibly have in common? According to Daniel Coyle in his new book The Culture Code: The Secrets of Highly Successful Groups, these extremely successful groups all have exceptional organizational cultures. In the four years he spent researching and visiting these organizations, Coyle came to believe that strong culture can be established anywhere by focusing on three essential practices: build safety, share vulnerability and establish purpose.¹

Building psychological safety in an organization addresses the trouble caused by our pesky primeval fight-or-flight response, rooted in the brain’s amygdala. According to Coyle, science has recently discovered that the amygdala, best known for constantly scanning for threats and putting us on edge, can also play an important role in building connections with others.² After sensing enough “belonging cues,” the amygdala is won over and sends out the message “You are safe here.” Psychological safety is not just a touchy-feely goal, however; research shows that having a bad apple disrupt a group’s sense of security reduces group output by 30%.³

Methods for encouraging psychological safety include:

- Communicate your listening: avoid interruptions, maintain an open posture and expressions. Make sure everyone has a voice: some organizations require that everyone share thoughts during meetings.
- Embrace the messenger (as opposed to shooting her): a vital moment for establishing safety is when someone in the group gives bad news or tough feedback. If the person giving the bad news is attacked, she will be reluctant to ever again report problems.
- Overdo thank-yous: the great San Antonio Spurs coach Greg Popovich tells his players, “Thank you for letting me coach you” at the end of each season; world-class chef and restaurateur Thomas Keller habitually thanks his dishwashers at restaurant openings, plainly stating that everyone in the organization is vital to success.⁴

Sharing vulnerability is essential for group success, and it must start at the top. If an organization is going to improve, clearly you have to know what is really happening—weaknesses cannot be hidden out of fear. While this sounds like common sense, it is extremely difficult as, in our highly competitive culture, we are “allergic to vulnerability.”⁵ Coyle believes that, in many groups, everyone has a secret second job of covering up their mistakes to maintain their status—but if you can create a space where mistakes are expected, people can stop worrying about that second job and put their energy toward actually improving.⁶

Ways to encourage sharing vulnerability include:

- Make sure leaders admit mistakes early and often. “I screwed that up” might be the most important words a leader can say, according to Dave Cooper, a U.S. Navy Seal who trained the group that raided Osama bin Laden’s Pakistani hideout.
- Deliver the negative stuff in person: avoid misunderstandings by dealing with tension face-to-face, turning a mistake into an opportunity to connect.
- Listen like a trampoline: do not be a passive sponge or immediately try to solve the problem. Instead, ask follow-up questions, perhaps even asking the same question in several different ways to get at the heart of the issue.
- Embrace the discomfort: discussing problems actually requires the group to endure two discomforts: 1) emotional pain over the mistake, and 2) a sense of inefficiency (shouldn’t we be moving on from the problem now?), but the discomfort is essential to problem-solving.⁷

² Coyle, The Culture Code, chap. 2.
⁴ Coyle, The Culture Code, chap. 6.
⁵ Hawk, “Daniel Coyle”.
⁶ Hawk, “Daniel Coyle”.
After making certain that people in your group feel safe and believe risks are shared, establishing a clear sense of purpose focuses attention on the organization’s goals. Coyle believes that building goals around stories or narratives can be a powerful way to establish purpose. For instance, an alumni fundraising call center at the University of Michigan increased its revenues by 172% after call center employees met scholarship recipients in person—the stories about how the funds raised at the call center changed these students’ lives were more inspiring to the call center staff than any financial incentives had been.8 Coyle also believes that purpose does not descend on groups from mission statements carved in stone; instead, high-purpose environments are “dug out of the ground, over and over, as a group navigates its problems together and evolves to meet the challenges of a fast-changing world.”9

Actionable ways to establish purpose include:

- Measure what really matters: in the early days of Zappos, call center workers were measured by how many calls they handled per hour. CEO Tony Hsieh decided that this measurement did not gel with the company’s purpose (even working in the wrong direction by encouraging haste and brevity), so he changed the measurement system to Personal Emotional Connections (PECs)—admittedly difficult to measure, but Hsieh believes this standard guided employees more toward customer satisfaction.

- Figure out where your group aims for proficiency and where it aims for creativity: skills of proficiency involve doing something exceptionally well repeatedly while creative skills are about building something completely different. Crystal clear directions and high-repetition, high-feedback training are suitable for proficiency skills, while creating a safe space for failure and protecting a team’s creative autonomy work better for creative skills.10

Coyle’s Culture Code could be an excellent blueprint for organizational success for groups of all sizes (from big companies to immediate families) and could be helpful to those at all levels of an organization, not just leaders. Many of the suggestions seem intuitive; however, as anyone who has ever worked in an organization knows, common sense does not always rule the day (or ever rule the day!)

10 Coyle, The Culture Code, chap. 17.
**Discussion Paper No. 2018-DP01**  
**Defining New Subfield Si in Fields 600-630 of the MARC 21 Bibliographic Format**  
Source: Program for Cooperative Cataloging (PCC), Standing Committee on Standards

This paper generated by far the longest and most animated discussion of all the papers considered at the Midwinter MAC meetings.

The PCC Standing Committee on Standards proposes adding subfield $i$ (Relationship information) to the 600-630 block of fields in the MARC 21 Bibliographic Format. The specific fields that would be affected are: 600 (Subject Added Entry-Personal Name), 610 (Subject Added Entry-Corporate Name), 611 (Subject Added Entry-Meeting Name), and 630 (Subject Added Entry-Uniform Title).

Chapter 23 of *Resource Description and Access (RDA)* “provides general guidelines and instructions on recording relationships between works and subjects.” Appendix M of *RDA* “provides general guidelines on using relationship designators to specify relationships between works and their subjects, and lists relationship designators used for that purpose.” As an example, one such relationship designator of interest to law catalogers would be “commentary on (work): A work used as a basis for a set of explanatory or critical notes.”

Appendix M relationship designators are currently recorded in subfield $i$ of 7xx related work added entries, 7xx linking entries, or incorporated into a 5xx note. As stated in the discussion paper, “the only method presently available for recording specific subject relationships between resources (and other entities) is to treat them as bibliographic relationships, not subject relationships. However, RDA specifically distinguishes between these two overall categories of relationship.” The discussion paper provides the following example of current practice:

100 $i$
   $a$ Dyck, Andrew R. $q$ (Andrew Roy), $d$ 1947- $e$ author.
245 12 $a$ A commentary on Cicero, De legibus / $c$ Andrew R. Dyck.
600 10 $a$ Cicero, Marcus Tullius. $t$ De legibus.

The paper posits that authorizing the use of relationship designators in 600-630 fields would obviate the need for the 700 related work added entry in instances like this, where the same information could be conveyed in the 600 field:

600 10 $i$
   $a$ Cicero, Marcus Tullius. $t$ De legibus.

The observation is made in the paper that “[t]he use of an RDA relationship designator … allows the field to express with greater precision the nature of the subject relationship than Library of Congress Subject Headings (LCSH) could do alone, since the subdivision ‘Commentaries’ is restricted to sacred works.”

In general, this discussion paper received little support. Some Committee members cited the unconvincing and hypothetical nature of some of the use cases presented; the German National Library suggested that RDA chapter 23 and Appendix M may end up being reorganized and cautioned against making the proposed changes to MARC 21 while RDA is still in a state of flux. Other Committee members voiced concern about the ability of current systems to implement batch changes to records based on this use of relationship designators.

LC’s reaction to the paper was particularly harsh, stating that “[w]e do not believe that it is necessary to use relationship designators in the 600-630 fields, and we intend to prohibit their use with LC subject headings if approved by MAC. … The time of the cataloger would not be saved, and the potential for confusion is real because this practice would effectively dictate the use of two different systems in the same tag block at the same time. The resulting displays in OPACs and discovery layers may also be problematic. Based on past experience, it is likely that the duplicated fields would display as separate results and could mislead users about the number of resources available on a subject.”

The sponsor of this controversial discussion paper will have to decide whether to re-work it or abandon the effort altogether.

**Discussion Paper No. 2018-DP02**  
**Subfield Coding in Field 041 for Accessibility in the MARC 21 Bibliographic Format**  
Source: Canadian Committee on Metadata Exchange (CCM)

Accessibility features in video recordings provide alternative access to both visual and audio content by employing supplementary visual description, such as described video, and captioning or signing. This discussion paper recommends
that subfield coding be added to field 041 (Language Code) in the MARC 21 Bibliographic Format to indicate:

- Language codes for written languages providing access to moving image audio content using captions
- Language codes for spoken languages providing access to moving image visual content using video description
- Language codes for signed languages that are used to provide access for the deaf and hard of hearing to moving image audio content

There was broad support for this idea among the Committee members, so this discussion paper is likely to return to MAC as a proposal.

**Discussion Paper No. 2018-DP03**

Inventory of Newer 3XX Fields that Lack Subfield $3 in the MARC 21 Bibliographic Format


Source: Music Library Association (MLA)

In this discussion paper, the Music Library Association advocates in favor of defining subfield $3 (Materials specified) for fields 377 (Associated Language), 380 (Form of Work), 381 (Other Distinguishing Characteristics of Work or Expression), and 383 (Numeric Designation of Musical Work) in the MARC 21 Bibliographic Format. Subfield $3 “allow[s] catalogers to better associate metadata with specific pieces of the bibliographic description.” It is important in music cataloging because music catalogers often create metadata in bibliographic records that applies to only one part of the resource being cataloged (e.g., a sound recording with a number of separate pieces on it). Some Committee members noted that the non-standardized, textual nature of subfield $3s limits machine actionability; the need for improved examples was also noted. After some discussion, the decision was made to accept this discussion paper as a proposal (contingent on some modifications to address the Committee’s concerns), and it was approved by unanimous vote.

**Discussion Paper No. 2018-DP04**

Multiscript Records Using Codes from ISO 15924 in the Five MARC 21 Formats


Source: German National Library

When the language of cataloging is written in a Latin script and the language of the resource being cataloged is written in a non-Latin script, a common practice of many cataloging agencies using MARC 21 is to transliterate the non-Latin portions of the bibliographic description into Latin script while linking to the same information in the vernacular script using 880 fields (Alternate Graphic Representation). This technique for handling non-Latin scripts is referred to as “Model A” (Vernacular and Transliteration: [https://www.loc.gov/marc/bibliographic/ecbdmulti.html#modela](https://www.loc.gov/marc/bibliographic/ecbdmulti.html#modela)) in Appendix D (Multiscript Records) of the MARC 21 Bibliographic Format. Subfield $6 (Linkage) is used in the 880 fields to link the transliterated field to its vernacular counterpart. One of the pieces of information encoded in the subfield $6 within an 880 field is the script identification code, “a code that identifies the first script encountered in a left-to-right scan of the field.”

The codes currently authorized for this purpose are:

<table>
<thead>
<tr>
<th>Code</th>
<th>Script</th>
</tr>
</thead>
<tbody>
<tr>
<td>(3)</td>
<td>Arabic</td>
</tr>
<tr>
<td>(B)</td>
<td>Latin</td>
</tr>
<tr>
<td>$1</td>
<td>Chinese, Japanese, Korean</td>
</tr>
<tr>
<td>(N)</td>
<td>Cyrillic</td>
</tr>
<tr>
<td>(S)</td>
<td>Greek</td>
</tr>
<tr>
<td>(2)</td>
<td>Hebrew</td>
</tr>
</tbody>
</table>

According to the discussion paper, “[t]hese values seem to be derived from the MARC 8 character set, an implementation of ISO IEC 2022.” Now that many library systems code their data in UTF8 (Unicode Transformation Format 8), the German National Library recommends that ISO 15924 (“Codes for the representation of names of scripts”: [http://www.unicode.org/iso15924/](http://www.unicode.org/iso15924/)) be used as the source for the script identification code portion of $6 subfields in 880 fields. ISO 15924 lists both alphabetic and numeric codes for dozens of scripts beyond the six currently available. This paper received general support from many Committee members, although concern was expressed as to how legacy data would be handled. The paper may return to the Committee as a proposal.

**Discussion Paper No. 2018-DP05**

Adding Institution Level Information to Subject Headings in the MARC 21 Bibliographic Format


Source: German National Library
Subfield $5 is used certain fields of the MARC 21 Bibliographic Format to indicate the “MARC code of the institution or organization that holds the copy to which the data in the field applies. Data in the field may not apply to the universal description of the item or may apply universally to the item but be of interest only to the location cited.” At Midwinter 2017, the German National Library introduced a discussion paper (https://www.loc.gov/marc/mac/2017/2017-dp05.html) that proposed extending the definition of subfield $5 to the 6XX range (i.e., subject access) of fields in the Bibliographic Format, in order to express the idea that a particular subject heading or term has been used in accordance with the policy of a specific organization. At that time, the paper generated a relatively long discussion during which Committee members expressed reluctance to conflate the kinds of information expressed in a subfield that has historically contained item-level information with work-level information associated with subject headings.

Now the German National Library has returned with a similar discussion paper. This paper offers the option for encoding institution-level information in the 6XX range of fields in field 883 (Machine-generated Metadata Provenance). The 883 field as currently defined refers to machine-generated metadata only and would have to be redefined to encompass the idea of “intellectually assigned metadata;” the authors of the paper also suggest renaming the field “Metadata Provenance.”

Some Committee members were not completely convinced by the German National Library’s use case for this change; nonetheless, the Germans were advised to pursue the 883 alternative rather than the subfield $5 option. This paper may resurface as a proposal.

Discussion Paper No. 2018-06
Versions of Resources in the MARC 21 Bibliographic Format
Source: German National Library

In this paper, the German National Library suggests several ways of designating different versions of a resource in a MARC bibliographic record. Specifically, the paper addresses the situation commonly associated with journal articles in electronic form, where preprints, postprints and the published version of an article may all be available simultaneously to readers who may be unaware of which version they are reading. The paper outlines several approaches toward handling this situation:

- Coding in an 008 fixed field Nature of contents byte
- Designation as a genre/form term in field 655
- Use of field 562 (Copy and Version Identification Note)
- A new subfield to be added to the 250 field (Edition Statement), designating version information, possibly using terms from a controlled vocabulary

The need to differentiate among various versions of an article in electronic form received wide support among the Committee members. Those expressing a preference from among the options set forth in the paper seemed to favor adding a new subfield to the 250 field. This paper may return to the Committee as a proposal.

Announcements from OCLC

OCLC Virtual AskQC office hours began January 31

On January 19, 2018, OCLC announced on the OCLC-CAT list that Virtual AskQC monthly Office Hours would begin on January 31. Starting at 1 PM Eastern Time, OCLC Metadata Quality staff are available for one hour; you can either login via WebEx or phone in to ask questions. They plan to try this for 6 months (January-June 2018) on the last Wednesday of each month. After the June office hours, OCLC will evaluate whether to continue offering this going forward.

Format: OCLC staff will offer a brief 10-15 minute presentation on a topic of interest to catalogers at the top of the hour. Following that, staff will be available to take questions. Topics will include WorldCat quality issues and cataloging questions. Questions about MARC, Resource Description and Access (RDA), Bibliographic Formats and Standards, and how to apply those standards and guidelines are welcome. Questions about OCLC product functionality (how to use Records Manager or Connexion) are out of scope and will be deferred to our product colleagues. Different Metadata Quality staff members may be available each month depending on individual schedules.
To attend the office hours, join the online meeting at https://oclc.webex.com/oclc-en/j.php?MTID=me115cb922941e78dd a5ebbe7ca6499eb the last Wednesday of each month at 1 PM Eastern Time.

**YouTube Adopts the ISNI Name Identifier**

On January 29, 2018, OCLC announced on the OCLC Abstracts email list that YouTube is beginning to use the International Standard Name Identifier (ISNI).

Already used extensively by libraries and archives to share catalog information, ISNI is being adopted by YouTube to identify music artists, composers, performers, and others who contribute works to the YouTube platform. YouTube will initially use the ISNI standard for music artists then expand use to identify contributors of other works. YouTube is the first ISNI Registration Agency that focuses on the music and audiovisual sectors. More than 10 million ISNIs have been assigned to date; some 650,000 are assigned to organizations and more than 9.4 million to persons. ISNI partners with OCLC, both as a founding member of ISNI’s board and as host and lead technology supplier for the ISNI system and database.

More information about YouTube and ISNI can be found in ISNI’s announcement at http://www.isni.org/content/youtube-adopts-isni-id-artists-songwriters.

---

**Preservation**

**Connecting the Archives to Your School’s Bottom Line (Part 2)**

In the first installment of this column, we took a look at a list of ways to enhance your archive’s visibility and prominence to your law school through several tangible steps. We listed a number of options (though the list is by no means exhaustive) which, taken together, would increase your archive’s value to the school and amplify its exposure to students, faculty, staff, and (especially) administration, often exponentially.

In this second installment, we take a look at one last step, one of the most powerful tools of all, namely oral histories and all the events that they are most often connected with: Class Reunions, Alumni Homecomings, and even Staff and Faculty Get-Togethers. Creating oral histories has the dual effect of reconnecting students and alumni (as well as staff and faculty) to their alma mater, sometimes in a tangible, long-term way (financial) AND of providing valuable new material for your archives (preservation). It is one of the most vital and dynamic trends in archiving right now.

Nowhere was the power of oral histories conveyed to me more forcefully than when I attended our New Hampshire Archives Group Spring Workshop last May at the Portsmouth Public Library. The keynote speaker was Jo Radner, whom some of you may know or have heard of. She is currently one of the premier voices in preservation work, particularly regarding the creation and preservation of oral histories. Her background is in literature, folklore, American studies, and Celtic studies, but her wisdom is applicable just as easily to recording and preserving the history of a law school.

Ms. Radner began by reminding us of the value of oral histories as a vital addition to any archives. Rather than providing us with dry dates and statistics, she reminded us that oral histories impart the MEANING of a place or an event; they turn the subject into something which is experienced—experienced through the eyes and ears of someone who was there. History comes alive in this way—it lives in perpetuity through the recording of the interview. This is something almost no other form of preserved material can do.

Ms. Radner stated that the purpose of creating a collection of oral histories was not to focus on facts; it was to capture experiences. By nature, experience is a subjective thing. Not every aspect of every recollection will necessarily match up perfectly with the facts, but that is all right. It is MEANING you are striving to record, thereby capturing a piece of history in a completely unique way.

Ms. Radner began with some preliminaries in preparation for recording oral histories:

1. **Clarity:** Be very clear about your intentions and the intentions of the law school and administration. What are you attempting to do, and what are your expectations? What are the expectations of the interviewee(s)? How are you recording the interview, and how will users access the interview? How will it be stored? Try to leave nothing out.
2. Transcription: Make sure to clarify that the interview will be transcribed. Transcription is absolutely necessary (according to Ms. Radner) and is not negotiable as a bargaining chip.

3. Video: If at all possible, try to record the interview on video. A visual record captures the maximum amount of sensory information.

4. The Interview Release Agreement: The interviewer and interviewee must sign an agreement, called the Interview Release Agreement (or something similar). This document is vital in that it lays out the expectations between the two parties. In summation, it states that the interviews will be protected by copyright and deposited in the archives for use by future researchers. The interviewee agrees to grant a non-exclusive license to use the interview at any time for exhibits, print publications, web pages, or other media throughout their lifetime. There is more, depending on the specific situation, but this is the basic boilerplate.

After these preliminaries, Ms. Radner entered into the crux of her presentation. In summation, these were her key points to keep in mind during any interview:

- Make sure one last time (at the beginning of the interview) that the goals of the interview are understood and shared. Share with the interviewee information about your oral history project, what you are interested in learning from them, and why.
- Ask brief questions, one at a time.
- Ask mostly open rather than closed questions. Keep the interview flowing and invite reflection, stories, and more meaningful answers. Begin with phrases like: Tell me about … , Can you describe … , In what ways … , What was it like to … , What stands out for you when you remember … etc.
- Ask follow-up questions, but wait until the “teller” is ready. There is a lot of flexibility here, depending on the situation, but some examples might be: Can you give me an example of that? Did that ever happen to you? Tell me more about that. How did that happen?
- Do not interrupt.
- Do not switch topics too abruptly.
- Avoid asking leading questions. This is tricky, but the key is not to force the teller into a particular position, such as “Were you angry when …”, “That must have made you feel terrible …”, “Was that during the time when …” etc.
- Show your interest. Let your enthusiasm and empathy come through in your voice, facial expression, body language, eye contact, etc.
- Rely on silence. Build the teller’s confidence that she/he will have the freedom to pause, to think, to speak slowly. This helps create a relaxed and un-self-conscious environment which is more natural and less forced.
- An interview is not a dialogue. Avoid telling your own stories during the interview or otherwise artificially inserting yourself into the conversation.
- End the interview generally. This is very broad, but some suggestions from Ms. Radner were to look over your notes to see if you have other questions or to ask the teller if she/he has any thoughts or anything else she/he would like to say.

This is a much summarized, bullet-point version of Jo Radner’s presentation due to space limitations, but it lays out most of the main tenets. What I cannot properly impart are the team experiences and personal stories and interactions. I can only advise readers to go and see her if they get the opportunity. She is wonderful.

This completes the list (my list) of ways to amplify your law school’s archives and maximize its value. Oral history is right up there at the top, hands down. It will add incalculable value to your archive, and it will create new connections to alumnae and students which may benefit the law school in unforeseen ways down the road.

So . . . give your archives a voice! Let it speak loudly and clearly of the glories of your law school and its rich history. And Happy Archiving!
Hello, I’m JoAnn Hounshell, Technical Services Librarian at Sidley Austin LLP in Chicago, the new Private Law Libraries columnist for *Technical Services Law Librarian* (TSLL). After 27 years of working in technical services departments in academic law libraries, I made the leap to Sidley in 2015. The transition was not as scary as I expected thanks to my new colleagues, but trust me when I tell you most days have gone by in a blur. Every day starts with a to-do list (go ahead, it is okay to laugh) that never seems to get shorter. No two days are the same. That is a good thing. I always suspected that firm librarians worked harder than their counter parts in other types of libraries, and I was so right!

In many respects, the technical services concerns of the firm library mirror the concerns of the academic law library: space reduction, print collection reduction, cost containment, etc. However, there is one difference between academic and firm libraries/librarians that I have experienced—firm librarians do not know what technical services librarians do. I expected attorneys to assume that all librarians are reference librarians. However, I was taken aback when I realized that my colleagues outside of technical services did not have a clue.

Last year, a new employee was being introduced around the office. When introduced to me, it went something like “This is JoAnn, and she does stuff.” I laughed it off, which was a mistake, but it triggered a memory. Months earlier, I joined a vendor presentation where another librarian introduced me, stating my title but admitting that she did not really know what a technical services librarian did in the library. Without missing a beat, I looked around the room, and said, “I solve problems.” There was that moment of “oh, yeah, problem-solver, got it.” Everyone in the room understood the concept of problem-solving. No one there needed to know the specifics of the problems I solve.

I understand why a non-librarian does not understand the distinction between a reference and technical services librarian, but I do not understand another librarian not knowing. Is it because not all firms have a dedicated or stand-alone technical services department? Regardless, it is our responsibility to take control of the narrative of the technical services team’s contribution to the firm. Almost everyone understands the idea of library operations, problem solving, project management, etc. I feel strongly that by allowing the work description of the technical services team as “they do stuff” or “I’m in charge of stuff,” I am diminishing the valuable contributions that our team makes to the library and to the firm in the eyes of others. Therefore, my new 30-second elevator pitch to new colleagues goes something like this, “Hi I’m JoAnn Hounshell, a member of the library’s technical services team, which manages library operations. Let me know if our team can be of any help; problem-solving is our specialty.”

---

**Research & Publications**

Autoethnography: Using Personal Experience in Research

**Hollie White**

*Curtin University*

In school we are taught to not use “I” when writing about anything in a serious way. Emotions, feelings, and even reactions not based on fact are all considered unscientific and not relevant or appropriate to use in formal writing or the workplace. The advantages to using this scientific approach is that it appears unbiased, well researched, and factbased. Disadvantages also exist to this approach, though. In libraries, we need to know what other people have implemented and tried in real, working libraries. What often (but not always) results when librarians report their projects or approaches in journal articles are “case studies” that are not real case studies from a research perspective. These articles mimic the fact-like, third person narrative of a scientific article by reporting a step-by-step process of project implementation and how well the project did or did not work. However, they often lack the methodological rigor required of a real case study. Also, the projects are reported in a dry manner that lacks connection to the true heart of why we conduct research or write articles at all—to share with our colleagues and to improve library services for our patrons. These colloquially named “Show and Tell” or “I Done This...
Good” articles are often mocked as pseudo-research. I believe that using more autoethnographic approaches could eliminate many of the problems that are found in current case study work which involves librarians reporting about their own projects.

Autoethnography is a sociological approach that focused on self-narrative research and writing (Anderson 2006). Ethnography is the study of a cultural group within its natural setting over a prolonged period of time. Data collection can included observational and interview methodologies (Creswell 2009). By focusing on the self instead of others in ethnographic research, autoethnography empirically relies on personal memories, perceptions, feelings, and emotions (Malsch & Tessier, 2015). Researchers are the source of the data itself, thus making data collection easy. Analysis makes researchers reflect on their own processes in order to better understand themselves and their daily lives. Because of autoethnography’s reliance on the self, it can be challenging—not in a methodological sense, but in a personal/professional reputation way. By exposing personal experiences, even at an institutional level, there are risks of judgement and misunderstanding (Malsch & Tessier, 2015).

So why is this method better than the “case studies” we already do in libraries? Yes, autoethnography seems “touchy-feely,” something we are continually warned away from embracing. However, wouldn’t it be better to acknowledge the humanity that is part of librarianship? Instead of stripping the personal out of our experiences and presenting them as not-quite-right “case studies,” we should embrace a method that from the onset declares that it will use “the personal” as a key element of study.

This article is not a how-to-guide to autoethnography nor an attempt to downplay how powerful case studies can be when done correctly. Rather I am suggesting that research in law libraries should expand to include methods that allow the profession to get the most benefit out of our collective knowledge and experiences. This may be accomplished through using methods and writing in ways that are different from what we are used to. Below I have included a list of introductory readings on autoethnography if you are interested in learning the basics of how to conduct such research.


Resources cited in the column:

The serials staff of the University of Washington School of Law, Gallagher Law Library recently identified the following serial title changes:

Harvard Latino Law Review
(OCoLC 32377824)
Changed to:
Harvard Latinx Law Review
v.20 (Spring 2017)-
(OCoLC 32377824)
Widener Law Journal
(OCoLC 52241793)
Changed to:
Widener Commonwealth Law Review
v.27 no.1 (2018)–
(OCoLC not currently available)

The serials staff of the University of Washington School of Law, Gallagher Law Library recently identified the following serial cessations:

Cardozo Public Law, Policy & Ethics Journal
CORRECTION—Ceased with: v.15 no.2 (Spring 2017)
(OCoLC 52896069)
Content merging with Cardozo Journal of International and Comparative Law to begin new journal, Cardozo International Comparative, Policy & Ethics Law Review with v.1 in Fall 2017. The publisher previously advised that
A significant number of new subject headings have appeared or been revised these last few months, and this column will focus largely on them. First, a reminder about the online tutorials in subject headings available from the Library of Congress at http://www.loc.gov/catworkshop/lcsh/. The training modules are concise and informative and serve both as a review for experienced catalogers and as a training tool for copy catalogers and new catalogers. It provides essential training for those who wish to submit subject heading proposals to the Library of Congress. As I worked my way through the modules, they answered a number of questions that have been lurking in the back of my mind since my days as a paraprofessional cataloger.

Several new or revised headings in the area of politics have appeared. The authority records for the headings State governments and Federal government were revised and should be examined. Political rights—Religious aspects and its corresponding heading Political rights—Religious aspects—Buddhism [Christianity, etc.] may now be used. The heading France—Politics and government—2012- is now France—Politics and government—2012-2017, and the heading France—Politics and government—2017- was added. Political manifestoes has changed its spelling to Political manifestos. Political activity in Russia and Eastern Europe resulted in the creation of the heading Nationalist parties. It lacks a scope note, but the sole Source Data Found note is helpful. Fake news is now a subject heading as well as a slogan of denial on the part of some politicians. Be sure to read its scope note as it provides a helpful definition. International Symbol of Access is available for use. It is commonly referred to as “the blue wheelchair symbol.” Pulse Nightclub Shooting, Orlando, Fla., 2016 may be used, and Mass shootings—Florida has been added as a broader topic. Sadly, another such proposal from Florida is now necessary.

In economics and finance, Bankruptcy—Taxation and Bankruptcy—Taxation—Law and legislation have appeared and will be useful. A somewhat amusing-sounding heading is Helicopter money (Monetary Policy). It has a scope note, thank goodness. Cooperation now takes Law and legislation. Stock exchange buildings. The heading International economic integration now has added material and should be consulted before use. The heading Paper money has been revised with new and deleted references and again should be consulted before use. Its long-time broader heading, Currency question, is no longer a broader term. Also added was African American women sports executives.

A few new terms appeared in environmental law and protection. Rights of nature encompasses the concepts of the rights of plants and animals, among other things. Vegetation management now takes Law and legislation. Additionally, Shark fisheries now takes Law and legislation.

The Library of Congress and other libraries continue work on terms of interest to researchers in the area of human sexuality. Gender-nonconforming people may be used, as well as Transgender police officers. The term Sexual predators, initially a subject proposal, was added as a UF reference to Sex offenders.

The list of available terms for Native Americans and the wars fought against them continues to expand. New terms include Cathlamet Indians (a people in Oregon), Mud Springs, Battle of, Neb., 1865, and Rush Creek, Battle of, Neb., 1865. Those concerned with the Vietnam conflict should note that the Battle of Khe Sahn has been broken into two headings.

The Library of Congress Genre Form Terms list has an important change. Casebooks is now Casebooks (Law). Multilingual dictionaries has appeared, as have Case studies and Account books.
Acquisitions

Negotiating with Vendors

As I begin to dive into acquisitions and vendor relations, a recent article in Online Searcher appears to be well timed and of potential benefit to others facing negotiations with vendors (see Michael L. Gruenberg, “Five Key Questions for Negotiators to Ask,” Online Searcher, Nov.-Dec. 2017, 44-47). In this article, Michael L. Gruenberg discusses key questions librarians repeatedly asked while he was promoting his book Buying and Selling Information. These topics are things we should, as negotiators for our institutions, be addressing with our sales representatives and vendors.

The questions Gruenberg addresses in this article:
1. Should you ask for, and expect, a price sheet from your sales rep?
2. Can the vendor defend the price?
3. Can a library request the assignment of a different representative?
4. What is the standard renewal rate?
5. Do I really need to create a negotiation plan?

If you have taken a Negotiations course in college or law school, you probably already know the answer to #5…

Gruenberg has some useful insight from his background in sales that can benefit us as negotiators for our organizations.

eBooks in the Law Library—Part 3

LLRX recently completed its look at the status of eBooks in law libraries (https://www.llrx.com/2017/12/the-state-of-law-library-ebooks-2017-18-part-three-what-law-libraries-are-doing/). The first half of this final installment is a case study of how the New York Law Institute added eBooks to their collection. It details how their original model had to evolve in response to changes driven both by publishers and by patrons.

The second half of the article is a brief survey of how other law libraries are incorporating eBooks into the collections. The main revelation of this overview is that no single approach may be appropriate for any one library. Many libraries are approaching eBooks in multiple ways, creating a custom process that is right for their users while also fitting within the budget. If you are dealing with adding eBooks to your own collection, this three-part series on LLRX is a great place to orient yourself.

Librarians to play increasing role in OER (Open Educational Resources)


A survey of more than 2700 faculty members shows that there is an increase of instructors using OER over traditional textbooks. An article in Inside Higher Ed reports on the survey findings:

“The ‘Opening the Textbook survey’ (https://www.onlinelearningsurvey.com/oer.html), published by the Babson Survey Research Group today, reports that the number of faculty members at two- and four-year institutions using OER as textbooks has nearly doubled in the last year—from 5 percent in 2015-16 to 9 percent in 2016-17. Awareness of OER—openly licensed and freely accessible teaching and learning materials—has also increased. Twenty-nine percent of faculty described themselves as ‘aware’ or ‘very aware’ of OER this year, up from 25 percent last year and 20 percent the year before. The proportion that reported they had never heard of OER fell from 66 percent in 2014-15 to 56 percent this year.”

However, as Jeff Seaman, co-director of the Babson Survey Research Group, mentioned in the article, awareness of OER is still low, and there are additional barriers such as finding and evaluating currency of materials.
Raising awareness, finding resources, and updating materials are all very familiar responsibilities for librarians. That is why Nicole Allen, the director of open education for SPARC (Scholarly Publishing and Academic Resources Coalition), a coalition that supports open policies and practices in education and research, predicts that librarians will “play an increasingly important role in helping faculty members find and evaluate OER content.”

The comments following the article further press this point. Librarians are natural resource experts to faculty. One librarian mentioned LibGuides as a tool to educate on open sources. Another librarian stated, “incentivizing for faculty is key.”

**Government Documents**

Archiving the Web in 2017

Throughout 2017, there was a renewed sense of urgency across organizations to document websites of state and government offices. In a recent update on partner program (https://archive-it.org/blog/post/notes-from-the-field-fall-winter-2017/) activities in the fall and winter of 2017, Archive-it explored highlights from the 2017 Mid-Atlantic Regional Archives Conference (MARAC), partnerships in the Community Webs Program for public libraries, and the second Documenting the Now symposium, *Digital Blackness in the Archives*.

These efforts exemplify current trends in government and community web archiving. The MARAC conference panel (https://themaracblog.wordpress.com/2017/11/15/marac-fall-2017-recap-session-12-web-archiving-democracy/), “Web Archiving Democracy,” was well represented by panelists who discussed the current trajectory of web-archiving practices. Comprehensive documentation of a democracy requires transparency from the government and includes the voice of the people. Both areas present ongoing challenges for archiving the vast array of rapidly changing or disappearing web content. On the one hand, discussions centered on evolving organizational perspectives on the nature of websites as documents and supplements to traditional government records. This includes filling in negative space by documenting peripheral websites that have been politically influential, such as grassroots and “fake news” sites. On the other hand, we also see efforts shift to a focus on people as the audience of web content, as researchers, and as community web archiving partners. These current trends in web archiving initiatives underscore the need for collaboration and partnerships, perhaps more than with any other information media.

**Information Technology**

Recommendations for Next Generation Repositories

The Confederation of Open Access Repositories (COAR) has just released a report from their Next Generation Repositories Working Group: *Behaviours and Technical Recommendations of the COAR Next Generation Repositories Working Group* (https://www.coar-repositories.org/activities/advocacy-leadership/working-group-next-generation-repositories/). The recommendations in this report provide an interesting read about the potential for a much more interlinked and standardized repository front in the future.

The report states, “many of the behaviors and recommendations for next generation repositories pertain to establishing links across repositories as a way to break down the silos and arrive at an environment characterized by interconnected networked repositories.”

We currently have a somewhat “under-regulated” system of repositories that may or may not play nice with each other. The suggested standards and protocols within this report would move towards a more standardized approach to repository content as well as the metadata backing these materials. If adopted globally, these standards would provide the foundation for more interlinked data and materials by adopting, and then adapting, resources that are already available.

Acknowledging that technology is fluid, and in some cases, the desired technology does not yet exist, the Working Group will soon be publishing their behaviors and technologies in a GitHub repository to support updates and engage the broader community.

Artificial intelligence laboratories in libraries


Libraries have often been the incubators of novel ideas and new technology. In an effort to share Artificial Intelligence (AI) development with a wider array of students and the public, the University of Rhode Island (URI) is opening an AI lab in their university library.
“The University of Rhode Island is taking a very different approach with its new AI lab (https://web.uri.edu/engineering/ai-lab-to-be-accessible-to-all-uri-students/), which may be the first in the U.S. to be located in a university library. For URI, the library location is key, as officials hope that by putting the lab in a shared central place, they can bring awareness of AI to a wider swath of the university’s faculty and student body.”

The Dean of Libraries, Karim Boughida, specifically mentions the lack of diversity of AI and its resulting issues of a biased algorithm as the reason to put an AI lab in the library, a place that values inclusivity. “Without explicit countermeasures, machine learning and AI could magnify existing patterns of inequality in our society”, says Boughida.

“Unlike a typical AI lab focused on research, the URI AI Lab will offer students and instructors the chance to learn new computing skills, and also encourage them to deepen their understanding of AI and how it might affect their lives, through a series of talks and workshops. The 600-square-foot AI lab will be located on the library’s first floor and will offer beginner- to advanced-level tutorials in areas such as robotics, natural language processing, smart cities, smart homes, the internet of things, and big data.”

Metadata

2017 DLF Forum and NDSA Digital Preservation


Jason Eiseman wrote an excellent blog post on both affiliated events for the Law Librarians of New England (LLNE) (http://llne.org/posted-on-behalf-of-llne-continuing-education-scholarship-recipient-jason-eiseman/) blog, so I will not provide another recap here. If you are also interested in viewing recordings, slides, notes, and photos from the events, in addition to keeping up with the latest Forum newsletter, the DLF has posted links to these resources in its November 22, 2017 Blog and News post, DLF Forum Recap and Working Group News (https://www.diglib.org/dlf-forum-recap-working-group-news/). The post highlights Forum activities of all active DLF working groups. I will just highlight here the work of a few of those groups potentially of interest to TS members.

At the DLF Forum, I attended the Assessment Interest Group (AIG) meeting. This group is very welcoming and encourages anyone who is interested in getting involved to take part—there is no membership requirement. Of particular interest to Technical Services Special Interest Section members may be the AIG subgroup, Metadata Assessment Working Group. The Metadata group is currently working on developing a framework for assessing descriptive metadata, building a repository of metadata assessment tools, and creating a clearinghouse of metadata application profiles. More information about the group is available on their wiki page (https://wiki.diglib.org/Assessment:Metadata) and information on their current projects is on their toolkit page (https://dlfmetadataassessment.github.io/).

Another group of potential interest is the Government Records Transparency and Accountability Interest Group. During the group’s working lunch meeting, members discussed planning for future projects and creating subgroups around issues related to sharing public information, education, advocacy, documentation, and potential special projects. The interest group will be making plans for its participation in the second annual Endangered Data Week (http://endangereddataweek.org/).

Other groups of potential interest to TS librarians include the Linked Open Data Zotero Group and Born-Digital Access Group. DLF groups of interest to librarianship in general include the Digital Library Pedagogy Group, Technologies of Surveillance Group, Labor Working Group, and Project Managers Group. You can learn more about all working groups at the DLF Groups page (https://www.diglib.org/groups/).

Continued from page 1

- **Be transparent with users.** Farney recommends having an institutional privacy policy and making sure your users are aware of the fact that their data may be collected.
- **Offer an opt-out.** Google Analytics does allow users to install a browser extension so that their web usage is not monitored.
• **Demonstrate the value of the data with users.** It is important that users are able to understand the context in which the data will be used and why the collection of the data is important.

• **If using personally identifiable information (PII), make sure that the data is secure.** It is important that the data is not vulnerable to access by others outside the library.

In the world of academic libraries, which is the law library environment I am most accustomed to, there has been a growing amount of discussion over the past few years about using data on student library usage to see how well it correlates with students’ learning outcomes. In some institutions of higher education, this is part of a larger trend toward using learning analytics to identify students who are struggling academically. Some of the data that librarians are considering looking at include: logins through proxy services to access resources from off-campus, attendance at library instruction sessions, and even how often they accessed the library building (measured by having them swipe their student identification cards). In an article on the Association of College & Research Libraries (ACRL) website titled “Keeping Up With Learning Analytics,” Steven J. Bell, Associate University Librarian for Research and Instructional Services at Temple University, writes about some of the positive aspects of collecting this kind of data, including the fact that “from an assessment perspective this can help justify library expenditures by demonstrating how academic libraries contribute to students’ retention and persistence to graduation.” However, he also points out some negative aspects of this trend, noting, “Mining data to identify students for signs of academic weakness sounds too much like surveillance to privacy proponents” (Bell 2014). In a more recent post on the Love Data Week blog, Kristin Briney, Data Services Librarian at the University of Wisconsin-Milwaukee, and her co-contributors addressed some concerns that librarians have with these data collection efforts. Briney writes, “As well-intentioned as these efforts are, and as important as it is to demonstrate that libraries aid the institutional mission, there are important issues to consider. Systematic data mining—especially of student behaviors and interactions with library resources—raises student privacy and intellectual freedom issues. Additionally, there are also practical questions regarding whether libraries have data management plans that carefully consider data anonymization, deidentification, retention, and deletion.” This blog post includes a long list of references that would be a good place to start for anyone interested in reading more about this topic.

The issues relating to what data to collect about library users and what to do with that data are very complicated. Librarians will have to think very hard about whether or not the positive outcomes of collecting and analyzing such data outweigh negative impacts on our library users and about how this type of data collection fits in with our profession’s standards regarding privacy.

**References**

Bell, Steven. 2014. “Keeping Up With…Learning Analytics.” http://www.ala.org/acrl/publications/keeping_up_with/learning_analytics