

Friday, 2014-06-27

Technical Services Directors of Large Research Libraries IG: "Transforming: E-Book & Collections"

ALCTS

2014-06-27 8:30-11:30

Update on 2CUL Technical Services Integration

Jim LeBlanc, Bob Wolven

<http://2cul.org/node/1> <-- info about 2CUL (Cornell, Columbia libraries integrating their tech services)

Problems/issues

- Can do work for either institution, but supervision has to be local -- ppl are employees of Cornell or of Columbia. There's no 2CUL to be an employer. Can't merge 2 physically separate institutions under 1 mgmt structure.
- Can't do some acq functions -- employee of Cornell can't commit Columbia to spend \$\$ (Not a big obstacle -- anyone can do prep work, just have to have a That Institution person pull the actual trigger -- but have to keep things like this in mind.
 - Another solution: Institution A could contract with Institution B to handle their renewals, etc.
 - Contracted w/ Serials Solutions on consortial level to have data held, managed all together
 - Jointly purchased Callisto software for e-resource problem resolution -- can monitor, manage each other's resources
- Ppl in middle levels are now starting to look on a practical level -- what can we accomplish here?
- One good time to look at doing new things is when people leave and you're replacing them (Replacing ppl? WOW MUST BE NICE -lkm). Look at what needs you have, fill accordingly

Q&A

- Shared ILS -- was this a planned efficiency?
 - It was one of the things they built around -- realizing efficiencies, advantages of the ILS itself. Efficiencies will not be realized during implementation ("That'll be a horror show") but later on.
- Employee of each institution has to place orders for that institution -- more cumbersome as things become more automated? Is it really necessary for each \$50 monograph?
 - E.g. approval plans, shelf-ready: are already delegating spending to some other entities. Might be able to work something out like that
 - Columbia uses shelf-ready more than Cornell does (Cornell can use more student labor). They thought about consolidating in some way, but each inst. is doing in way that makes sense for it.
- Shared position -- time to fill/replace. Stakeholder pool is bigger -- can't just fill?
 - They do already have one of these! Are indeed facing this situation. They're currently testing out, not forming any solid agreement. But eventually will need an agreement, shared decision-making structure.
- Gains in efficiency from combining processes?
 - More to gain in terms of scope than efficiency. Might lose some efficiency in the near future. Efficiency will come in when they start sharing collection development work.
 - "Tech Services integration" -- artificial boundary? Would need to go beyond TS?
- Contracting w/ a 3rd entity, so don't have the internal bureaucracies to negotiate?

- Setting up separate entity is so massive -- need to understand benefits, be able to make a convincing case. (One consideration: Columbia has a union, Cornell does not.)
- Shared bibliographer model?
 - E.g. agreement that shared Slavic bibliographer will keep X office hours for Cornell, Y site visits for Columbia, etc.

How do we acquire things from nontraditional resources?

Bob Wolven

Not easy mechanisms for these kinds of acquisitions -- small publishers producing ebooks (e.g. Badlands Publishers, doing book-as-art, only available through Apple or Amazon for download to 1 device). E-graphic novels. PDFs -- can't just mount on a server. Library entity "shall remain nameless" gave out a "complimentary copy" of something. Certainly there are analogues in the print world, we just don't know about them.

So anyway -- what are we doing about these?

- Perception of need/obligation -- this stuff is "free", we expect it to be there. Preservation? Things might be very ephemeral.
 - PDFs -- NLM creates catalog record, link out to it. Putting a lot of work into this! Pilot project -- putting PDFs (esp from other gov't sites) into digital repository. Kind of a one-off; not coming in in batch mode.
 - Gov Doc interest for things that might disappear. Generally don't have copyright issues to deal with; lots of interest. Would be good to be able to share this out. (Oklahoma had something kind of like this -- a pool of e-resources that we were cataloging. -lkm)
- NYU -- a couple of different models have emerged. Vendor hosting; however, this wouldn't work for some of the more vulnerable resources, so looking into local hosting. Also: musical scores, only in PDF (but not usable in that form). Currently printing them so that they'll be usable.
 - For PDFs, can work w/ commercial binder -- send them the file, they'll print & bind it.
- Things *sold* as PDF are different -- can't just preserve them.
 - Things w/ no restrictions go into the "open" folder; restricted or unknown, goes into the "restricted" folder. Catalog, provide, preserve accordingly.
- What about continuing resources like this? These things will be coming in on an ongoing basis.
 - Analog in print: zines. Use accession numbers to group the items; takes load off of cataloger b/c the title is already cataloged, just adding issues/items.
- What about one-user content? (e.g. video games, password protected things, etc.)
 - We've framed this as an acq question, but really isn't. The actual purchase is easy; it's really a question of how do we serve this stuff. Provision, preservation, etc.
 - We can serve it; solutions exist, but they're so labor-intensive that we can't scale it. (Not nec. true for Amazon/Apple purchase? Can put content on a device & then lend that device; hard to scale this, automate it.)

"We are devoting more and more time and people to moving data from one system to another, data wrangling"

Lisa German

Someone asks for data -- what do you give them? How do you know what to give them? More and more metadata is being made unshareable, or only shareable in a limited fashion. Should insist that metadata be shared freely.

- Institutions are reinventing the wheel. Also, different person at each institution; difficult to make sure that the data requests trickle to the right person.

Where's the authoritative store of metadata? Looking at how our data flows out to other institution -- "octopus on steroids"

Sometimes, ppl asking for data don't know what to put in.

- "Linked data will save everything!" Seriously though -- that's the point. Our idea of a "database of record" might have to go out the window.

Not just moving the data around to get it where you want it the first time -- is also maintaining it. Can't fix things when headings change, etc. Maintenance is harder than getting stuff in there to begin with.

- Not clear that non-cataloging ppl understand that. "Digital library people don't have that concept, in a big way." This tension has been there for a long time, but has been exacerbated with exponential expansion of the kind of info we want to collect, serve. (Tension between getting something in there rather than nothing, and then maintaining it once it's in there.)

"Database of record" -- where do you look when disaster strikes? (Like when there's an investigation or something, e.g. patent?) "It's almost like data forensics".

Transforming Technical Services: Evolving Functions in Large Research University Libraries

Jee Davis, Univ. of Texas Libraries

Title

Presenter(s), Organization

[see also handout; ppt not available yet]

How is Tech Services changing? How are our (UT's) peer institutions changing?

Research

- Contacted 25 university libraries; 19 responded. Conducted 1-hr phone interviews with them.
- Four categories of interview questions
 - Use of term: Tech Services?
 - Functions
 - Reorganization
 - Outlook (new positions, skills, TS drivers...)
- This presentation will focus on functions & drivers; Davis hopes to share the rest of the research on another occasion.

Functions (list not exhaustive, but it is hard to read from back here!)

- Includes traditional stuff like cataloging, acq, authority control; but also new/emerging functions like copyright, data creation, data curation, nonMARC metadata, many others
- Other areas performing TS include special collections; special libraries (Music, Law, etc.).
- Cross-departmental functions include resource discovery, licensing, nonMARC metadata...

Observations

- Efforts to centralize TS functions
 - Cost savings; increase consistency & efficiency
- Constantly adjusting to enhance efficiency

- Staff reallocation (TS is being cut to the bone -lkm)
 - redesign & transformation of existing job descriptions
- Shift from print to e-resources
 - Slow transition process; greater need for skilled personnel.
- Emphasizing and increasing collaboration among units, depts, divisions, institutions
 - Integrate TS fcn's with rest of the library; collaborative cost savings; increased complexity in library projects/initiatives
- New, emerging digital-centric functions
 - digital repositories; web archiving; linked open data; digital curation
- New emergence of soft skills in current job postings. Used to see "works well independently", "attention to detail". Now we're seeing more demand for interpersonal skills, flexibility, communication skills.
 - Growing need for technology skills, project mgmt skills, knowledge of intellectual property, soft skills as well as traditional TS skills (e.g. foreign languages, etc.)
 - "How much we are asking our people to be perfect"

TS Drivers & Challenges

- Asked participants: trends, challenges?

[Chart is on handout.]

Biggest challenges: technology & money.

- Challenges are "2 sides of a coin": provide great opportunities, but if opportunities are not realized, can become a threat.

Conclusions:

- Drivers of change in TS: financial constraints; technology.
- Emphasis on electronic & digital resources continues. The changing nature of academic libraries affects TS functions & org structure. Transformation of TS functions in libraries is evident.

[HEADING OUT TO TRY TO CATCH FRBR IG. -lkm]

FRBR Interest Group

ALCTS; ALCTS CaMMS

Friday 10:30-12:00

(Common problems in searching for images)

Kelley McGrath

[Got here late.]

Who is the "creator" of a moving image?

- Moving Images & Performance: Script is a starting point -- not like a music score, a static text to be interpreted. It's meant to be performed only once; is specific to the actors in the project, etc.
- Example: stand-up comedian in front of a fixed camera. Clearly the "creator" -- s/he is the only person involved in creating the content.

CEN cinematographic work -- conceptualized as a concrete entity. It's the result of a collaboration between many creators; it's inseparable from the actual product, the expression.

- Versions: language (original, soundtrack, captions, subtitles); format (16mm film, VHS, DVD)

"Mix-and-match expressions" -- can have different languages, subtitles, etc. on different publications of a film. Could have DVD/Blu-Ray, etc. Must re-verify for each manifestation.

Star Wars re-releases: devoted fans know the differences; catalogers can't be expected to. [But I bet lots of us do! -lkm] http://en.wikipedia.org/wiki/List_of_changes_in_Star_Wars_re-releases

How do we get there from here? We can't start over -- have to extract data from what we've got now. OLAC Movie & Video Credit Annotation Experiment is trying to teach computers to get data from credits -- recognize roles, etc.

<http://olac-annotator.org>

Group 1 as a Preservation Administration Framework

Jacob Nadal

Why FRBR: Version 1?

- Analytical + descriptive bibliography; similar to preservation model of intellectual content & artifactual value.

FRBR: Version 2?

- *Double Fold*. Libraries understood "newspaper" to mean a particular kind of text; Baker was thinking of the artifact.

Preservation controversy occurs when different FRBR levels are used.

How should we think about preservation? Series of activities that protect information from damage. Can be on the level of the carrier, or of the information itself. Transport problem: e.g. have a VHS, but no VHS player. "Audiovisual archivists have eBay open all the time."

M&R in FRBR: Highly Speculative

- Work: no equivalent. Preservation has to have an object.
- Expression: language. Consider it being kind of like translations.

FRBR for preservation

- Item: Preservation must address a concrete entity, which can be repaired (conservation), created (reformatting), or stabilized (preservation).
- Easy to confuse M&I in preservation. M-level & up = preservation (risk management)
- I-level is conservation (corrective/preventative) and I1M1 --> I1M2 is reformatting (digitization, etc.)
- If one manifestation has artifactual value, can't do reformatting.

Manifestation-level preservation: is a manifestation preserved when *any* of its Items exist, or is there a critical mass? (Or does it have to be *all* of them?)

Expression-level preservation: multiple copies in many manifestations: less risk, greater access. Must all of its manifestations be preserved? (Nadal says no.)

Risk Management: he has a graph! Awesome! Graphs initial # of copies needed (I) for likelihood of 1 I to still exist in 2100 CE, assuming 1% loss per year. Cool. So... once again we see that lots of copies keeps stuff safe. Also looks at adequate copies if loss is carefully controlled (less than 1%/yr).

National/Global Pan: identify dedicated repositories & support networks.

Q&A

Crowdsourcing project? (OLAC Annotator) <http://olac-annotator.org>

- Trying to parse, say, a 245 \$c. Taking that, map to any authorized name if there was one. E.g. Directed by Clint Eastwood -- so get Clint Eastwood, try to map to some controlled vocabulary. Programmer working in natural language processing on his day job (happened to be husband of someone working on the project) was appalled that they were doing QC by eyeball/spot-checking. Said they needed to invite more of the public in to help them out; wrote a program, made an interface. Weren't able to do as much as they wanted in terms of other languages, but do have some languages (Turkish, Georgian for example.)

Re: How many expressions do you have to preserve in order for a work to be preserved. What if you preserve only the translations of a work? Is the work preserved?

- This keeps Nadal up at night! [I think no. Run through Google Translate & back e.g. Snotty comparison, but not w/o validity? -lkm]
- Heddatron (sp?) -- a play where ppl are looking for someone to play Hedda in *Hedda Gabler*. Contains most of the text of *HG*; however, if you had to reconstruct *HG* from it, you'd be sunk.

More about international networks? Are all using the FRBR model?

- Nadal hopes they will. Imagine several print repositories, serving a community of thousands of libraries; it's huge. Have to prioritize the "most endangered" things.

Survival curve -- do they consider whether a copy circulates?

- Came from a Berkeley operations research researcher. They do consider it indirectly.

With TV shows on DVD, how does the individual season, episode, etc. affect FRBR considerations?

- Really have to take it into account. Called it "Larger Works" [serialists are ON THIS once again; were talking about Superworks at least a couple years ago. When will monos ppl start listening to us?:-P -lkm] Relying on info that catalogs typically don't have; but could maybe bring it in from another source (imdb, etc.)

Is there anything (article, etc.) about failings of initial FRBR model McGrath used? [I think I missed this?]

- No; she wanted to write something but hasn't yet.

Competencies & Education for a Career in Cataloging IG

ALCTS-CaMMS

Friday, 1:00 pm

Cataloger 3.0: Competencies & Education for the BIBFRAME Catalog

Allison Jai O'Dell, Univ. of Miami

- Catalog 1.0: card catalog
- Catalog 2.0: online catalog
- Catalog 3.0: looking toward more interoperability & sharing of library data.

“Authority Control Finally Gets Awesome” - switch from strings to things.

- Authority files can now be authoritative! Can link out to an auth record, then don’t have to mess with it if (i.e. when) it changes.

MARC is over 50 years old! Libraries have had to develop special software, etc. to use MARC; data entry requires knowledge of the special format. But when “everything is nonMARC”, does it really make sense to distinguish between metadata librarian & cataloger?

“Death of the out-of-the-box library catalog” -- can improve user interface when making our own designs.

- User task: “Browse” -- Links allow follow-your-nose browsing; they can go down the rabbit hole, not have to “do research”. It can be more fun to look for things.
- User task: “Repurpose” -- We can’t predict all the uses people will have for library content.

The BIBFRAME Library Cataloger

- New stuff: Variety of markup languages (not just one very specific, complex one); conversant with a handful of schemas. 1st opportunity to design a library catalog. RDF, Semantic Web. Will also need to be on top of copyright & licensing -- who owns library metadata? (Won’t have to be a copyright expert in the same way as Acq librarians, but will have to know some of it.)
- At the same time, some same old stuff: authority control; cooperative content sharing, creation, use (but now going beyond the library world); database mgmt, entity-relationship model (but now with PHP, SQL)

Education for Cataloger 3.0

- Old paradigms (librarian = original cataloger, paraprofessional = copycatter) don’t work. Now need to ask, e.g., what can a research analyst offer?

Please Send Catalogers: Metadata Staffing in the 21st Century

Jennifer A. Liss, Univ. of Indiana (Bloomington) @cursedstorm

“Competencies” = usually big, abstract things, needed to be able to perform one’s job. E.g., a performance artist: competency might be good balance. How to determine whether s/he is exhibiting good balance? Can observe. E.g. bagpiper in a Darth Vader mask & kilt, riding a unicycle. Neither the bagpipes, costume, nor unicycle is a competency, but can be considered part of a performance indicator.

How to reconsider performance indicators?

- Normalizing data -- currently, “Applies content standards (AACR2, RDA) for resource description & access”. Need to extend beyond AACR2 & RDA. Might also have DC, DACS, etc.
- Parsing resource metadata -- currently, might have “assigns descriptive metadata to textual resources”. Now have to think about how to extend to images & other non-text resources.
- Exploring info hierarchically --
- Recognizing relationships -- “provides additional access points; performs subject analysis”. Expand to: ALL

THE THINGS! E-book versions, audiobooks, etc. Authority data is important!

Human dimension of technical change -- catalogers are unfairly viewed as change-resistant. We've been changing all along! MARC has even changed in the past 30 yrs. (In fact, this is when we discover things we've been doing wrong.)

Schema.org -- summer 2011. Google, Bing, some other ppl.

- Not just headings & paragraphs are marked up! Also have ingredients & cook times in recipes; actors, directors; restaurant webpages with locations & hours of operation.
- Search results powered by semantic markup. Gmail displays too.

Google knowledge graph -- 2012

- "Graph" --> linked data. Encoded, semantic in nature.
- Must be some kind of taxonomy underneath it. (E.g. subject "dogs", narrower subject "golden retrievers".
- Can go into one area of data, then also display data from other areas. Results beyond what simple keyword search would yield.

Google Hummingbird (2014)

- "Operates like a machine-driven reference interview"
- More like a conversation than like a search -- asking questions, getting answers, refining.

Don't know if users will ever be satisfied with plain keyword search again. ("It's all about keyword searching"? NO.) Good news for catalogers, because that's exactly what kind of data we create -- either directly/manually, or by automated means.

Slides are available at <http://bit.ly/sendcatalogers>

Q&A

Will we ever be without a local catalog?

- Not completely -- need some kind of inventory of "these are what things we have, what we license, what we can access".

Practical applications? BIBFRAME is barely on ILS vendors' radar. They still struggle with RDA. Where and when would we get a BIBFRAME catalog?

- They might not come from your vendor. Maybe it's user-developed; maybe it's from another community and we tweak it. Will probably have to be community-sourced.
- LC & Zepheira already working on this -- they have an edit, "pretty terrible".
 - Shana has info! "Full-on scrum" development; iterative. Lots & lots of changes. Now is the time to get involved -- tell people what we need, what we want. Be involved. (May mean that you end up on a task force, but still. Look for the ppl asking questions, and ask questions with them.)
 - <https://zepheira.com/2014/04/viewshare-in-transition-new-look-and-feel-now-hosted-at-library-of-congress/>

Comment from an educator: MARC records aren't just going to go away. There are lots of new things people are learning in library school, but still learning the old things too. Educators' plates are full.

- Ties into performance indicators rather than competencies. The competencies may remain the same, but there are new and more performance indicators.

Cataloging is one of the first things to get cut. If we don't even know what these new things are going to look like, how can we use them to advocate? How can we explain why we're needed?

- 2 things: impact & money. Cost-effectiveness will come from moving to our own tools instead of vendor software & interfaces. Also, we talk a lot about cost -- but what about value? We need to talk more about what it would look like if we just let the catalog go?
- Shana: there are resources out there, from people who are better at explaining this kind of thing than she is. :-)

"Traditional cataloging" -- one director has said "our catalogers spend too much time on authority control. People just want to look up an author & find everything s/he's written." *FACEPALM*

- Shana: Call it "identity management", not "authority control". Then you can start pulling in ORCID's and other things beyond old-school cataloging. (It's a more modern-sounding name, but is actually more accurate IMO. As we move towards linked data, there won't be one *authority* we consult; there will be *identities* we describe, collate, and distinguish. -lkm)

LITA Open House

LITA

Friday, 3:00 pm

Membership Development Committee

Jason Battles

LITA "Guide to Participation" (committees)

#BecauseLITA hashtag -- for ppl to tell how LITA makes their lives & work better

Assessment and Research Committee

Michelle Frisque

needs ppl w/ strong statistical skills; keeper of data for the research & assessment that's done. They're also working on a membership survey for the coming year.

Education Committee

Toni Nolen

Look out for the training & professional development of members. Look for which sessions are the biggest draws; look for ppl to teach or lead those. Also Regional Institutes. They vet proposals for offering education & development; LITA Forum offerings, web courses, webinars. They have a lot of members rolling off & are looking for new members.

2014 LITA Forum Committee

- Date: November 5-8
- Place: Albuquerque, NM: Hotel Albuquerque <http://hotelabw.com>
- Theme: Transformation: From Node to Network

Accessibility IG

Ranti Junus

Universal Accessibility IG -- LITA, ACRL, ASCLA

- “Resource and discussion forum for library disability service issues” -- joint mtg. every ALA Annual & MW, jointly with corresponding IGs in ACRL, ASCLA. (Wonder if this is something else that TS ppl need to be in on? Perhaps at least with digital repositories, etc.)

Web Project: Libraries for Universal Accessibility (LUA) <http://uniaccessig.org/lua>

Authority Control IG (LITA/ALCTS)

Christina Hennessey

Auth Control for library catalogs, databases; international sharing of data

How to make our systems talk to all the other systems in the world & on the web?

They have a session on Sunday afternoon(1:00-5:30 pm)

Game Making IG

Breanne Kirsch

<http://gamemakinginterestgroup.wikispaces.com>

Also recently published a book -- will be raffling off 2 copies at their meeting, Sunday 8:30-10:00. They talk about how libraries use games for outreach & instruction.

Imagineering IG

Breanne Kirsch

Program is tomorrow, 2:30-5:30; meeting is tomorrow morning.

Heads of Library Technology IG

Meg Brown-Sica

Anyone who is interested in technology is welcome, directors or not; they talk about technology management. MW = meeting, Annual = program. This time: Evidence-Based Library Technology.

Library Code Year IG (joint group with ALCTS)

Chris Strauber

Education & coding group

“Ruthlessly, cheerfully open to all skill levels & backgrounds”

Saturday, 2014-06-28

Copy Cataloging Interest Group

ALCTS-CaMMS

Saturday, 8:30-10:00 am

Library of Congress Update

Angela Kinney

[arrived late, sorry! I'll probably copy off of Deb DeGeorge's notes for a bit. I wish I'd gotten here early enough to hear about the effect of the government shutdown on production -- it is important to remember these things and keep angry about them. -lkm]

Differentiating Ebook Packages Through Unique Identifiers

April Grey

[missed this. :(Sorry. -lkm]

Copy Cataloging Outside of WorldCat, AACR2, and RDA (title is approximate, I missed it. -lkm)

Nathan B. Putnam, Univ. of Maryland

National Trust Historic Postcards Collection

Digitized postcards from as early as the early 1900s -- many of the buildings in the pictures no longer exist. "Snapshot in time". Many of the postcards were used as postcards; have messages, postmarks w/ dates on them.

- Initial descriptions were created by digitization students; then passed the records to the Metadata Librarian for review. Metadata Librarian worked w/ Systems to create workflow.
- Metadata Librarian also worked with a graduate assistant; GA took over after Metadata Librarian left in 2010. Then GA worked w/ copy catalogers to pass knowledge, responsibility on to them so that someone permanent would be in charge (she was graduating & leaving).

Goals: work outside of RDA/AACR2; still work within copy cataloging workflow, still create high quality records. They used Fedora cataloging utility. Subjects/vocabularies: used LCSH, Getty Art & Architecture Thesaurus (AAT), and LC's Thesaurus for Graphic Materials (TGM).

Workflow: initial record was retrieved w/ status = "Pending". They'd check for errors in description; check subject headings, correcting & adding as needed; then mark records as "Complete". Very similar to WorldCat copy cataloging workflow, just in a different place with different schemas & subject vocabularies.

Many of the boxes had dropdown menus -- this helps eliminate errors/typos. Can also do things more quickly than typing stuff out.

4 copy catalogers were trained by the GA; they cataloged ~500 postcards in a month (in addition to usual work). There were some that they skipped (usu. due to technical issues, e.g. back of postcard not digitised). These went back to the GA for review.

Ebooks

Ebook titles are activated either one-by-one or in batches (full or partial collections).

- Individual activations: managed through the WorldCat Knowledge Base. Records are loaded from the vendor; vendor notifies OCLC; OCLC sets holdings.
 - Copy Cataloger, what does s/he do? Checks to make sure all this stuff actually happens. Also can fix errors in the WorldCat master record.
 - Cons: there's not always a record; links are sometimes wrong; cleanup is required.
- Batch activations: new roles for copycatters. Batch cataloging w/o MARC records; verifying title lists from

vendors. "More difficult than you might think". [Nah, I know how it is -- serials cataloging, yo. We've been to THAT rodeo before. -lkm]

- Copy catalogers can compare spreadsheets using Excel, VLOOKUP function.

RDA Enrichment: Catalog Benefit?

Roman Panchychyn, Kent State

Kent State investigated an RDA enrichment program -- had a limited window b/c Backstage had offered to do a project for free for libraries who'd done database improvements & would do ongoing authority work.

"Hybrid" records?: during RDA development, term "hybrid" = catalogs that combined both AACR & RDA records. But during the 2010 RDA test, was found that there was value in adding RDA elements (e.g. 336, 337, 338 fields) to otherwise-AACR records. So now, there was hybridization within individual records.

- Definition was formalized in Feb 2013 by PCC report.
- Not cost-effective to convert non-RDA records to full RDA. However, there were some elements that it made sense. [Going through this more quickly than I can keep up. Will try to find that PCC report.]

OCLC planned to adhere to these PCC guidelines ("OCLC RDA Policy Statement", 2013). But how will OCLC changes get into local catalogs?

RDA enrichment -- there are vendors who assist libraries with projects.

- MARCIVE - worked w/ U. Houston
- Backstage
- (others. Missed them.)

Also, MarcEdit has RDA Helper fcn. that helps libraries do their local RDA enrichment. Batch processes -- can add/delete GMDs, add 336/7/8s among others.

In late 2013 Backstage surveyed libraries to gauge interest in RDA enrichment; wasn't much interest. 90 responses; mostly academic libraries, plurality (31%) over 1M records. 93% were accepting RDA records, but overall not really moving forward much with RDA -- not even sure how they were going to proceed. [There's lots of data here that Backstage allowed Roman to share! Wish I could keep up.]

- If Backstage offered a discount, would they consider RDA? 47% said only if it were free; over 30% said not even if it were free -- too much work on their side.

Kent State was an early RDA implementer; ability to get it done for free = major motivator. They set up a profile guide -- extensive detail, options on conversion processes (as with any catalog database project contracted with a vendor.)

Issues Encountered:

- Ended up having to create 3 profiles -- main KSUL profile; special collections main profile; special collections variant profile (Sp Coll had a number of their own decisions they needed to be able to make).
- Needed public services support; data, displays had to be consistent for patrons.
- Issues with OhioLINK matching algorithm

- Record reloading -- had to get special port (\$\$\$) so that it wouldn't impact consortial catalog
- Need custom reports to identify records for cleanup

Forced KSUL to look closely at use of local abbreviations, other inconsistencies. Took a lot of staff time & involvement.

Why do it?

- Consistency of data -- ILS vendors, others can make optimal use of RDA data; improve patron display
- Only have to work with one standard. Incoming records can be enriched as need be. Once OCLC hybridizes WorldCat, they won't even have any AACR records coming in (except maybe old records from vendors).
- Unprecedented opportunity to clean up legacy data! Good to tie in with other projects (batch reclamation, ILS upgrade/migration, etc. Kent State's was tied in w/ RDA implementation.)
- BIBFRAME is in process; but we need better tools for patron experience now. MARC isn't going away yet.

Q&A (all presenters)

Catalogers (NLM) get these "massive hybrid records" & don't know what to do with them. Also, pre-AACR records -- what is Kent State doing with those? Converting through same process?

- Yes. Converting as many records as possible. Difficult to work with multiple standards.

Vendors are sending records encoded "I", but not compliant. What to do?

- Guidelines -- what to look for in an acceptable record. KSUL is considering most hybrid records acceptable, unless they're really screwed up (bad access points, etc.)

Librarians as Digital Leaders: Collaborating on the Development and Use of Digitized Collections

ACRL

Saturday, 10:30-11:30

Collaborating Globally, Scanning Locally

Martin Kalfatovic, Biodiversity Heritage Library

Encyclopedia of Life (EOL) -- 1,303,873 species pages.

Lots of stuff -- specimen collections, card files, databases, publications -- "locked up in different ways", in different museums & other places. Wanted to get it all in one place.

Charles Davies Sherbourne -- natural history library needs to have all the history.

Why not scan it all & get it over with? 2006: launched BHL.

Metadata problems! Collection assembled from different sources (museums, etc.) Scientists have reclassified, renamed species over time (e.g. sperm whale).

BHL now has 20 members and affiliates. Secretariat & Technical Staff are based at the Smithsonian; also offices at the Missouri Botanical Library (? check this)

Also have other BHL nodes around the world. BHL Europe, BHL Egypt, BHL Africa, BHL China, Majulah Singapura.

They've created the library and are now in a sustainment phase. Income comes from membership dues, donations.

- Over 44M pages of digitized text; 139K items
- 5.5M visits from 241 countries from 2007-2014
 - Over 50% of use is from other machines.
 - Mobile use is increasing -- have to make sure it's usable from mobile devices.

Recently: Mining Biodiversity grant from IMLS (\$174K for US members)

Collaborating in the Mountain West: Access to Resources and a Whole Lot More

Sandra McIntyre, Mountain West Digital Library

"DPLA on-ramp"

900K resources in 560 collections: photographs, birth & death records, music & other sounds, videos, artwork, many many others

159 partners currently – UT, NV, AZ, ID, MT, HI. Academic libraries; academic departments; publishers; lots of others!

- 31 hosting hubs in a distributed network. DPLA aggregates from MWDL; MWDL aggregates from hubs. Hubs have differing relationships w/ partners; they aggregate, but also provide support, etc.

Funding: UALC at Univ. of Utah; DPLA grants 2013-2015; Bill & Melinda Gates Foundation; IMLS; Knight Foundation; NEH

MWDL has service fee schedule (setting up repository, scanning projects, etc.) They also have a metadata model.

Open access of digital collections -- increasingly, funding agencies require that the research produced from their grants should be open. "Should not be putting licensing requirements on top of public domain [materials]." Taking cues from what Europeana has done.

Q&A

At Boston PL -- ppl said digital libraries are "not what we do". That's not true -- public libraries are about bringing information to the public.

BHL: any statistic about how many species were extant when discovered & are now extinct?

- Doesn't know, but probably could figure it out through some data analysis!

Scholarly Communications Interest Group

ALCTS

Saturday, 1:00-2:30 pm

Scholarly Communication Express

Sherri L. Barnes, UC-Santa Barbara

UCSB: public university, one of 10 Univ. of CA campuses. No professional schools (law, medical, allied health, library; not even a business school). Unusual for UC.

Library: 3M vols, 40 librarians. Library is in a time of transition w/ new administration. They've also just launched their digital repository (institutional repository/digital library/data mgmt system).

Sherri is the Scholarly Communication Program Coordinator (50% of her time). Scholarly Communication Group = coordinators of 4 collection groups; Head of CD; AUL for Collections; life sciences, chemistry, & social science data librarians; 26 faculty champions (mix of humanists, social scientists, & scientists).

The UC Academic Senate passed an OA policy 2013-07-24. They've been in a pilot period the past year; the policy goes into full force in November. This is when UCSB will be depositing their materials.

Scholarly Communication Express: 15 minute presentations on trends in academic publishing, delivered to campus departments & organizations. Topics include altmetrics, impact factors, open access, data management plans, EZID accounts, many others. The program is tied to short-term goals (create an active outreach & education structure) and long-term (become the campus authority on copyright, etc.)

The Outreach Department helped develop branding & publicity (flyer, webform, website & URL). They conducted 3 sessions between April & May. In the fall, they plan to share info at the Library Open House, in the Library newsletter, Open Access programs, and tie-in with OA Policy implementation.

Q&A

Why not make videos that faculty can stream?

- Good idea, and perhaps presentations could be enhanced by short videos; but they really do prize that face-to-face interaction. Allows librarians to gauge gaps in knowledge, ensure that ppl are understanding the topics.

Why isn't UCSB using CDL as a platform?

- Wanted to have their ETDs closer to home, in their library; didn't want the default to be open access for ETDs.

OA Policy -- faculty acceptance, willingness to do the work?

- In discussions prior to the enactment of the plan, faculty seemed cognizant of the fact that it'll be more work for the library than for them. (Still trying to make it easier for faculty as well as for librarians.)

Repository at UCSB -- is it for faculty to use?

- Repository at UCSB isn't related to the OA Policy. There's no self-deposit now; aren't sure if there'll be any in the near term. Outreach largely deals with e-scholarship.

The 3 departments that requested 15-min presentations -- how did they learn about it? What outreach effort were they responding to?

- Two responded to the email that went out (they responded immediately); the other one was due to contact with a staff member.

Public Access Policy compliance -- are faculty required to deposit their work?

- No consequences if they don't; but they're expected to. They might be contacted by library saying "we see

you have this publication; would you be interested in depositing it, or may we?"
Even if it's not publicly funded?

- Yes. Policy is a publishing agreement between university & faculty.

New England Collaborative Data Management Curriculum (NECDMC): An Educational Program and Service for Best Practices in Research Data Management (RDM)

Regina Raboin, Tufts Univ.

<http://library.umassmed.edu/necdmc/index>

- Phase 1: Course modules. Some are out-of-the-box; others can be adapted for your local purposes.
- Phase 2: Content development. Funded by grants; partner institutions joined the collaborations.

The curriculum materials are under a Creative Commons license. Lesson plans; activities; assessment, readings, etc.

Includes real research cases -- mostly biomedical, but they're trying to bring in research cases from other subject areas with the help of collaborating institutions. This helps lend context to the content -- makes it seem more real, more applicable.

It's a living site -- they're always trying to bring in more readings, keep things current.

Local piloting

- NN/LM, MAR Class (Mid-Atlantic Region, Philadelphia, April 2014)
- 15-week Simmons GSLIS course
- "Train-The-Trainer" workshops

Roll-out at Tufts: library workshops (Savvy Researcher Series); targeting specific domains; Research Data Management Software (RDMS) pilot project; Research Day at Tufts

NECDMC Pilot Student Feedback

- Now understand that research lifecycle is cyclical
- Understand that articles can be retracted if data can't be substantiated
- Technical standards -- e.g. date format affects accessibility
- Assigning access levels

They have seen improvements! Questions about specific examples of metadata preparation; details on where to go for help; RDM responsibilities (PI, students, etc.); local tools, policies available. They've had discussions about depositing, etc.

Q&A

Outreach to grad students?

- They've been doing drop-in workshops (Savvy Researcher; have also gone to Science Library & done drop-in sessions there. Are thinking of making it a requirement for RCR.

Discovering Open Access Articles: Maximum Access, Maximum Visibility!

Setting the stage for discovery and access: definitions, process, and open access at UCLA

Angela Riggio, UCLA

Early 2000s: meetings to define open access, determine what it should be: material that's free on the Internet & allows some downstream usage. (Some insist that there's no restrictions; but generally, a middle ground that most adhere to.) See Peter Suber definition.

Lots of terms flying around -- green, open, free, gratis, libre, gold. Not all open access is equal; that's why it gets confusing.

Why open access?

- Exorbitant cost of "traditional" peer reviewed journals; authors don't get paid; authors give up their copyright
- So, why not make the content free?

There are more than 200 institutional policies dictating open access (as with Bethesda meeting). Also sub-institutional, funder, and other mandates.

The UC OA Policy was passed on 2013-07-24. Became effective on LA, SF, Irvine campuses 2013-11-01; remaining campuses will join 2014-11-01. There was a review May 2014. UC has a Deposit Agreement between the faculty member and UC; if the publisher will not allow it, faculty member has to opt out on an individual basis. (Most publishers require an embargo, which is acceptable. It's usually 1 yr.)

UC's eScholarship repository (<https://escholarship.org/>)

- indexed in Google Scholar
- Submission form tries to minimize the number of keystrokes an author has to do -- lots of pulldown menus; can just drag 'n' drop an article

Implementation

- Robust scholarly communication committee; has been around for 10 yrs.
- It's a faculty policy, so info is distributed through the faculty senate
- Library is hosting open sessions; also website, presentations, email (shared email address)
- Developed a new model for the Ithaka Survey on scholarly communication (completed in June)

Discoverability -- it's in Google Scholar (which is where most researchers start). eScholarship is known to researchers too. But, need metadata (distinguish versions, indicate level of openness). Many of the journals are listed in DOAJ too.

UC's open access production is not just articles -- also course materials, digitized special collections stuff, and HathiTrust/Google stuff.

"Affordable Course Materials Initiative"

- 1st quarter: 27 applications, 23 awards; \$30,500 awarded & funds spent on acquisitions; 885 students, \$88K in student savings
- Biggest savings in one course: Theater (\$11K), Mech E (\$11K), ECON (\$4,480).

Discovery of open access articles in hybrid journals: What role does the library play?

Mary Ann Jones, Mississippi State

Hybrid journal: a regular peer-reviewed journal that allows the regular peer-reviewed submissions, but also allows an author to pay a fee to have the article published as open access. (Consists of open access and paid content together.)

- Per SHERPA-ROMEO, over 200 journals allow fee-based OA publishing; from \$60-\$5,000. Some do a per-page charge, so could be > \$5000.

It's open access, so what's the problem?

- We've drilled it into our patrons to go to journals through the library website. But then if a person discovers, say, a citation via library web structure but the library doesn't subscribe to the journal, the Find It (or whatever) will tell the researcher s/he can't have access -- even if it's an OA article in a hybrid journal.

What do we do? Rely on Google Scholar? Teach students/researchers to go outside the library? Assume researchers know where to go to find OA articles?

Possible solutions:

- Publisher's role?
 - Put all their OA articles into a searchable database?
- Aggregator's role?
 - OA indicators -- figure out a way to show that these articles are OA, if they're not allowed to replicate the actual articles w/in aggregator DB?
- Author's role?
 - Deposit articles in IRs where possible
 - Demand greater exposure of OA articles

Cross-community metadata creation and sharing: the NISO Open Access Metadata and Indicator Recommended Practice

Nettie Lagace, NISO

What's NISO? Nonprofit industry trade association; accredited by ANSI; 150+ members. Mission: develop & maintain standards related to information, etc.

When a project is brought to NISO for consideration, they look at value to be gained; feasibility; stakeholders; community/ies affected; participants required/desired; time frames.

"Wouldn't it be nice if there were an open access indicator that would show, across DBs, that something was open access?" (This idea didn't take hold in 2011; was kicked around for another couple years.) Hybrid journals are a big problem. Nettie recalls this from her link resolver days.

There's been another stakeholder come in recently: funding agencies. This is helpful for motivating people to publish OA. Also, funders want to know that their \$\$\$ has gone where they wanted. More funder mandates → more open access → more confusion over what people can get.

NISO Metadata & Open Access Indicators Working Group -- wanted to develop a specified format for bibliographic

metadata (and possibly a set of visual signals) indicating the readership rights associated with a single scholarly work (article, book chapter, etc.); recommendations for distributing info on these metadata stdds; downstream use.

- Membership includes ppl from libraries (digital & other), link resolver vendors, publishers, et al

“Open access” -- politically fraught, so decided not to use this label. (Factual info: what readership & downstream reuse rights are granted?)

- Also decided not to create/recommend a logo.
- Minimal set of metadata needed?

<free_to_read> tag -- indicates that an article is free to read (or that it will be in the future, and when).

<license_ref> tag -- content of tag would include a stable identifier expressed as a URI. Can have start dates, etc. embedded as well.

Who distributes this? Publishers, aggregators, & content providers. Could include the metadata in all standard metadata sets (intention: that this will become part of standard editorial & production workflows.) Could also include in e-TOCs, publication alerts, RSS feeds, etc.

Up next: Public review & comment period. The WG will address, potentially incorporate comments. Then, NISO approval. (It'll be NISO Publication as a Recommended Practice, not a standard.)

“It is a brave person who claims to know” where OA will be in 5 yrs.

- Simon Inger, *Workflow Implications of Open Access Publishing*, simoningerconsulting.com

Collaborative Digitization Interest Group Meeting

ASCLA

Saturday, 4:30-5:30 pm

Collaborative Digitization Interest Group

ASCLA

Collaborative Dig Listserv

- Current listserv: collaborative-dig@lyralists.lyrasis.org
 - migrated from DigiStates: BCR --> LYRASIS
- ASCLA now sponsors listservs.
 - Subscribe ascla-igcd firstname lastname
- Concerned about migration before annual – will send out migrate message following conference.

Crowdsourcing. Family History and Long Tails for Libraries

Frederick Zarndt, Secretary, IFLA Newspapers Section

(Slides will be available online.)

Crowdsourcing – put a task out to an undefined group of people (they may or may not expect to be paid). Term was coined by Jeff Howe in a June 2006 *Wired* article.

- Amazon Mechanical Turk: make money working on crowdsourcing projects.
- reCAPTCHA: now owned by Google. Verifies you're not a bot; corrects OCR text.
- Wikipedia: also an example of a crowd collaborative project.
- Gutenberg: launched Dec. 1971, longer ago than any other crowdsourced project. Digital Bentham.

Anyway, why should a library care about crowdsourcing? Refocus: think of it in terms of "user engagement". How likely is someone to come back to your website, or to stick around & use it for any length of time?

Trevor Owen: crowdsourcing is the single greatest advancement in getting ppl to interact w/ library resources.

Trove (Australia): offers a great quantity & variety of stuff, but newspaper text correcting = biggest way that ppl interact w/ the website. Australia is a very small country by population; but Trove gets 78% of web traffic for all national libraries. Avg Trove time = 8 min on site; LC = 4 min.

Calif. Digital Newspaper Project recently added crowdsourced OCR correction.

Raw accuracies could be very poor, 20%-70%. But corrected (word) accuracies rise to over 99%.

But correcting text is boring... why do it? Genealogists & family historians. Utah Digital Newspapers did a user survey (2012) & found that 72% of visitors are there for genealogy stuff. In 2013, CDNC & Cambridge PL did a survey with almost identical results (survey is available online). Ppl search for birth, death, wedding announcements; info about their families.

Trove users' report: it's a good way to pitch in while also learning. Characterizes "engaged user" population; gives their reasoning for doing text correction, etc. Available via IFLA online. CDNC report does too.

"Old, bored people: convince me there's worth."

Economics – Trove got over \$1M worth of OCR labor; CDNC, \$55K or so.

Accuracy – greatly improves search results. Raw OCR tends to have ~60% text accuracy.

Long Tails – concept discovered w.r.t. Web retailers. Popularized by Chris Anderson (2004, *Wired*) & Clay Shirky (2003). Means that a greater population lies under the "unpopular" part of the distribution curve than the popular.

Considerations: marketing/advertising? Motivation? (Don't really need to – they just need to know it's there.)

Authentication/identity of crowdsourcers? (Singapore: worried about implications of vandalism, malicious changing of text.) How to administer crowdsourced data? (Don't. Australia considered this at first; then decided to just let users at it, not curate crowdsourced data.)

Crowdsourced Manuscript Correction

Ben Brumfield, fromthepage.com

<http://manuscripttranscription.blogspot.com> (Ben's blog – slides will be available here)

OCR: Print vs Script. Example using Tesseract w/ print, script on same piece. Print was pretty good – only screwed up an apostrophe (put it as a backwards apostrophe). Script OCR was really lousy – complete gibberish.

Implications:

- We can't use machines to transcribe script.

- Goals: do we want accuracy or findability?
- Workflow?

Accuracy vs Findability – Google can't crawl images of cursive writing. (E.g. image of Ben's grandmother's diary, Julia Brumfield.) It's not characters to Google – it's pixels. No one can find it or read it.

- Diaries were distributed to her grandchildren upon her death, scattered all about. After Ben put his one online, he rec'd a package from someone he'd never heard of – another of her diaries, with a message along the lines of "You'll do more with this than I will".

Workflow – can't do as with text OCR, just have ppl do a single page/article/whatever and be done with it. People have to learn the handwriting; through extended reading they gain knowledge that lends needed context.

"Page 19": collaboration btwn Ben, Front Page, & Smith Univ. They put up a diary online in September; 2 wks later, a reader discovers it. Corrects it one morning in November; he went back & revised it later that afternoon, 6 more times within a month, and then a couple more times within the year. Very different from a one-off session!

Choosing a tool: there are lots to choose from! Many are open source, but can be pretty rough. Here's a listing of transcription tools: <http://tinyurl.com/TranscriptionToolGDoc>

Selection criteria

- Look at the kind of thing you're transcribing. Two examples: on top, a diary. On bottom, a census. The diary, you want to be able to print it out & read it. The census, you'll need it to be usable in a database.
- Markup?
 - Internal markup (looks like XML). E.g. TEI Toolbar, developed by TranscribeBentham. Very useful if you're dealing with marginalia (also crossouts, superscripts, etc.)
 - Wikilink markup (looks... well, like a wiki!) It's a different flavor of markup. Useful for linking documents within a larger transcribed work; indexing (e.g. tracking mentions of a person).
- How open is your project? (Staff? Students? On-site volunteers, or open to the web?) Might need to restrict for privacy concerns. E.g. diaries of Shelby Foote – he died fairly recently, and the diaries are still private. Don't want to put that online for the world to get at it.
- Branding – may be really important.
 - E.g. French version of Wikisource; allows people to transcribe & do OCR corrections all in the same place. Great tool, very powerful – but no place for branding! You see Wikisource all over the place, but not the name of the project or the institution.
 - E.g. Letters of 1916: Creating History. Much better branding; this is what you'd use for user engagement project.
- Integration? Have to ask what-all the tools will play nice with.
- Hosting? Host yourself, or hosted elsewhere? Most tools, it has to be local.

Yardsticks for Success: what does successful engagement look like? E.g. one guy (TX historical diaries) transcribed 100 pages all at once. Project was cancelled – they were disappointed that only one volunteer showed up! (Quality vs quantity? Number of volunteers vs deep engagement? Number of pages vs quality of transcripts?)

One such tool "Casino-Driven Design" – no exit, no windows, free drinks ;-)

Q&A

Kept mentioning history, genealogy. Does this subject matter tend to draw more ppl in than, say, science?

- FZ: kept talking about historical newspapers b/c that's what he does – doesn't know much about crowdsourcing for other topics. Wanted to stick w/ libraries.

Combining humans w/ software? Humans for first pass, then spell-check to clean it up?

- FZ: OCR can do a lot, depending on quality of image & text; but can't do everything. Most newspaper digitization projects are done by service bureaus (cost factor – cant do them on site). Many of these already do combine humans w/ software.
- BB: Science – e.g. herbarium (registries)? Labels are very heterogeneous – some are printed by computer, some have handwriting. Can look at the type of gibberish to see if it's probably got handwriting in it, then send appropriately for human review.
- FZ: Dictionaries too – name, or garden-variety gibberish?

(Missed a question. Very sorry!)

Motivation – recognizing volunteers?

- BB: Lots of research has been done into this. Number one finding: don't do anything! People are doing this for their own reasons. If you then go handing out t-shirts or something, they re-evaluate what they're doing: it quits being leisure & starts being poorly-paid work. (Substitutes external motivation for intrinsic.)
- FZ: Nat'l Library of Australia lists their top 10 volunteers (by user name) on their site; they also brought their top volunteer to the Library for a tour, gave a gift (iPad?). Other than that, nothing.

Sunday, 2014-06-29

Metadata Interest Group
ALCTS-CaMMS; ALCTS-CRS
Sunday, 8:30-10:00 am

(Slides will be available on the ALCTS Metadata IG page on ALA Connect sometime in the next few days)

Metadata Migration to Islandora: Is There An Easy Way?
Sai Deng, U of Central Florida Libraries

MODS: Metadata Object Description Schema, developed by LC <http://www.loc.gov/standards/mods/>

Islandora FLVC (Florida Virtual Campus) -- decisions made collaboratively.

FL-Islandora ingest

- can have single items, multiple items, or on a page level. Upload is through a web interface.
- Also offline ingest via FTP.

FLVC has stylesheets for DC to MODS, MARC to MODS, and ... (missed it)

Transformation from DC - MODS - MARC is challenging: mapping is never perfect, and moving from a less-granular schema to a more granular one. Introduces data ambiguity -- e.g. DC - MODS: <titleinfo><title>, subtitle is not distinguished. A;so MODS-MARC: title types (translated, other, uniform, etc.) are undistinguished.

- Element relationships (DC-MODS-MARC): e.g. can't subdivide geographically, temporally, etc.; can't show

relationships between authors; etc.

What can be done to improve the results, when LC DC-MODS and MODS-MARC stylesheets don't produce good output?

- Consortial level:
- Local: libraries can have their own stylesheets

DigiTool DC file -- I hope the slides will be available, because you really need to be able to see these screenshots. :-)

- In initial test of DC-MODS conversion, many issues: title/subtitle not distinguished; physical description is under <note>; subdivisions not available; many more. Author first/last names, etc.
- [Screen shot of MARC record produced from initially DC-MODS converted record]

Options to Improve Metadata

- XSLT customization -- will need to be able to test; and ideally a programmer. Will also need a comprehensive review of the results.
- MARC-MODS, MODS-MARC stylesheets; can do post-transformation MODS editing. FLVC recommended Notepad++ (free text & source code editor for Windows) for MODS editing. <http://notepad-plus-plus.org/>
 - Can go in & manually add the necessary granularity (separate parts of title; personal names; date patterns; etc.)
 - Can transfer data to a spreadsheet, identify patterns in data structures using search & replace spreadsheet function. E.g. find all <namePart> elements that contain "Committee", and this is a corporate author. Find all <title> with ":" in the title field, and you know this has a subtitle. (Examples of search strings given in screenshots.) Identify nonfiling characters -- search <title> for initial stop words. <topic> with "--" in it, this has a subdivision.

Metadata Migration Issues

- Is there a better way besides what we've seen here -- customising XSLT stylesheet; post-conversion pattern-based records editing?
- Should you deal with data markup based on patterns & conditions in the stylesheet or the post-transformation editing?
- Local elements to MODS: is MODS extension good practice?
- How much automation will MODS editing tool allow? How much manual or semi-manual editing work is realistic for librarians? What skills are required?

Might not be a good way to get good data when migrating from a less- to a more-granular schema. Decisions will depend on the level of the project (local, consortial, national). Still seems relevant to share collections in non-traditional DAMS in a traditional ILS.

(Very rapid-fire presentation! The slides are comprehensive though. Look for those later.)

Q/A (these questions took place after the 2nd presentation)

How many staff?

- One Digital Initiatives Librarian; working on pattern identification w/ Sai.

Documentation of procedures -- available for others?

- No, not yet. This presentation is currently the most detailed writeup.

Non-roman scripts?

- Yes, some submission packages did include special characters. (German, French.) But nothing like Cyrillic, Hebrew.

Staff on QC/QA (make sure migration was successful)?

- Digital Initiatives Librarian & Sai worked on it together; didn't have a staff person. Do have a programmer at UCF, but s/he does more high-level stuff.

What motivated choice to migrate, & why did they choose the system they did?

- Statewide system, chosen statewide through collaborative decision process.

Did Metadata have any role in this decision?

- Sai thinks so. Islandora is more flexible for metadata.

Discovering Isaac Leeser: Improving access to text collections with TEI markup

Nicole Arbuckle, Backstage Library Works; David McKnight, Univ. of Pennsylvania

U Penn's context: the Gershwint-Bennett Isaac Lesser Digital Repository is a Jesselson-Kaplan American Genizah project. (Genizah: a Torah cannot be destroyed; if no longer usable, it must be buried. So there are these buried stores of Torah fragments about, and they can be put back together & reconstructed. I hope I'm recalling this correctly.)

Isaac Leeser (1806-1868) was a Jewish author, translator, journal editor, correspondent; a major religious figure. (Was pro-South during the Civil War -- some controversy there!) He also founded the Jewish Publication Society & founded *The Occident*.

- In 2006, plans were started to produce a FT repository of Leeser's correspondence, other publications & translations, and *The Occident*.
- 2007: Judaica Curator Dr. Arthur Kiron identifies Leeser scholars actively collecting Leeser materials.
- 2007-2010: project team to scan, transcribe, edit. There was a hiatus in 2009 (NEH grant) but were able to resume w/ enthusiastic private funding.
- Spinoff: Penn Libraries in Lyrasis/Archive project
- 2010: American Genizah Project was formalized w/ private donors. Later started partnership w/ Back Stage.

Goal: integrate materials available through 3 sites into a single website. (Different file formats; different metadata structures, one might expect.)

The project has an interactive map ("pre-Google representation") for browsing the collection. Worldwide; also one for the United States. So geographic info is important.

Penn could've developed in-house TEI tagging/programming, or outsourced (had enough funding to do either). Chose to outsource.

- Opportunity to work w/ vendor interested in developing a new service
- Had plenty of money; but timeliness was a factor.

Between 2011-Jan. 2013, a Project Manager was hired; contract formalized w/ Back Stage; programmer hired; site went live.

*** **

Backstage Interest: opportunity to improve TEI skills.

Screenshot talking about some of the markup editing they had to do. What to do with different types of writing? (Subscripts, comments, addenda to letters, etc.?)

Used Validome to do actual XML validation checking. Quality assurance ran along with encoding, after an initial delivery, reworking, and redelivery of a pilot batch. At about 200-300 XML files/week, it took them about 3 mos to get through all the letters.

<http://leeser.library.edu>

Q/A:

Staffing, for that kind of records delivery?

- Dedicated Project Mgr, ~20 hrs/wk. Also had a couple/few other ppl doing full-time encoding.

No programmer around? Spent all that time copy & pasting?

- Could've been beneficial in this case; needed to get it done though.

One project manager, or one at Backstage/one at Penn?

- Each site had a PM. Backstage one was at half-time, working only on TEI; Penn one was managing other aspects of the project as well. During heavy delivery time, Penn PM was 100% on this project (managing transcription as well as TEI).

How does TEI improve access?

- Fundamental belief that transcription/full text linked to image improves ability to find it. (Just as in Ben's talk yesterday!) Can also embed other metadata into this full text transcription.

Software system for TEI delivery?

- In-house system

Business Meeting

1. Housekeeping -- review agenda, call for additional items; approve minutes from ALAMW14
2. Officer reports
 - a. Programming: lots of interest from initial call for programming. Ivy & Santi will be developing an e-form.
 - b. Blog: everything is status quo.
3. Liaison reports
 - a. LITA: (missed it. Sorry.)
 - b. CC:DA: Continued work on RDA; crossroads. Want practical instructions today, but these might not be useful in the linked-data, postMARC future.
 - c. MLA: working w/ PARS task force to update document on audio. Long-awaited clearinghouse for music metadata -- available on MLA site (Metadata for Music Resources?)
4. Elections: VC/CE: Ayla Stein; Program Co-Chair: Liz Woolcott; Blog Coordinator: Ann LePage
5. Open discussion on future programming topics

- a. Linked Data -- programs have been very heavily focused on techie side (triples, triples, triples); never seems to get beyond that. NT reframe the whole thing & actually get into what that triple *is*, how we as metadata librarians can understand & use it. Real-world examples, case studies. Also, how to allocate staff? (Human component). Tools too. And how do we implement linked data?
- b. Services on a shoestring budget, or when you have no one? How do you provide services?
- c. Authorities? Collections w/ no headings -- how do you manage that? How would you implement linked data for that?
- d. Metadata for images & text, but what about videos? What about alternative formats?
- e. Metadata assessment -- both quality (what do we need/want) and cost.

LITA Program

Sunday, 10:30-11:30 am

The role of libraries in data management and curation

Nicole Vasilevsky, Oregon Health & Science University

Ontology group www.ohsu.edu/library/ontology -- only one librarian in the group (Scholarly Comm Lib'n); also scientists, a data wrangler, grad fellows. So mostly not trained as librarians, but the function of the group is to support the research lifecycle; support biomedical researchers w/ data management.

[Illustration of 5-point research lifecycle] -- actually more of a network! It doesn't really flow in a circle -- each node is connected to the others. Completed w/ pentagon/pentagram inside the circle.

Libraries have roles in each of the stages of the lifecycle network.

Questions the group asks:

1. How can we make science more reproducible?
2. How can we educate researchers to make their data reusable & research reproducible?
3. How can we use data to generate new hypotheses & make new connections?

Ontologies -- foundation of their projects. An ontology goes beyond just a controlled vocabulary (closed, prescribed list of terms); in an ontology, the terms are defined and relationships between the terms are also defined. (Example: a list of terms about wine, shown as a simple controlled vocab and as an ontology.) "A formalized conceptualization of a specified domain of interest". Hierarchical terms are defined textually & logically; relationships btwn terms are defined; expressed in a language that can be understood by computers.

Ontologies have the strongest semantics, but also involve the most time/\$\$. Better to start with an existing ontology than to start from scratch.

1. How can we make science more reproducible?

Materials used? Protocol? It's kind of like a recipe for science. If you have a list of ingredients & a step-by-step process, you can replicate a dish to a degree. But if you don't know where the ingredients came from (e.g. farmer's market vs grocery), results might be different.

#overlyhonestmethods -- couldn't include protocol b/c only allow 200 words in description; ha to put in an

appendix

An experiment in reproducibility -- <http://biosharing.org/bsg-000532>

- Looked at 248 papers in 84 journals across domains, impact factors, and reporting guidelines; examined various resources used in the experiments (antibodies, cell lines, constructs, knockdown reagents, organisms)
- Only ~50% of the resources were identifiable (varied across resources)
 - No correlation btwn impact factor & resource identification; resources not more identifiable in journals w/ stricter reporting guidelines either. (Even though one journal recommended identifying reagents, ppl still didn't do this.)
- How to fix this?
 - Vendor names/catalog #s (Cell Signaling, cat #3576). But catalog #s change; suppliers go out of business.
 - stable, unique identifiers (Antibody Registry -- more about this later)
 - data standards (biosharing.org)

Resource Identification Initiative -- promotes use of research resource IDs in published literature. (Pilot project going on through 2014 -- initially was only supposed to be 3 mos.)

- Resources: antibodies; model organisms; software & tools
- RRIDs should be machine readable, standard & stable, easy to use.
- Sample citation: Polyclonal rabbit anti-MAPK3 antibody, Abgent, Cat# AP7251E, RRID:AB_2140114.
- Researcher submits a manuscript; editor/publisher asks researcher to include RRIDs in publication. In description, in keywords (so that RRIDs are in front of whatever paywall).

Outcomes demonstrate need for better reporting of materials/methods; cultural shift in writing/structuring papers; cultural shift in how we view literature.

www.force11.org/Resource_Identification_Initiative

What does it mean to be reproducible? Reproducibility Initiative: attempting to independently replicate research in 50 major cancer studies. <https://osf.it/e81xl/wiki/home>

- On average, approx 15% of resources are unidentifiable. Looks a lot higher than other results; but this is b/c one outlier paper used 3000 knockdown reagents, and they listed them all, Great that they did this, but it skews the results. If you take that paper out it's more similar to prev results.
- One of 1st experiments: inhibition of leishmania parasite using peptide X.
http://pt.wikipedia.org/wiki/Leishmania_infantum
 - Tried to reproduce this result; but it took a much larger amount of peptide. Turns out, in the 1st experiment the peptide had been treated beforehand, so it wasn't the same thing really at all.
 - So -- was this reproducible b/c inhibition was achieved, or not (took different amount)?

2. How can we educate researchers to make their data reusable and reproducible?

"What would you do with \$1,000 to improve data mgmt that doesn't involve another tool?"

- Workshops w/ library; individual consultation.

Held a Data Management Happy Hour w/ catering, bartender. ("Food, drink, discussion, and data". Put together a presentation; offered a \$50 Amazon GC if they scheduled a consultation..

Also went & talked to grad students. (Flyer w/ Zoidberg: Your lab notebook is bad and you should feel bad! OK, not really, but still come to the Data Wrangling Information Session.) Once again offered a \$50 Amazon GC for scheduling a one-on-one consultation. Also had sushi.

Gummy Bear: the groundbreaking paper (didn't want to use an example that was tied to actual research, because ppl in an area not the same as the example would quit paying attention).

- How to measure the springiness of a gummy bear. Had a picture of the anatomy of a gummi bear; had to manipulate the image, measure results, put together a graph & figures ("SpringOMatic")
- Results from each group varied.
 - Ppl see the same data very differently; "Detailed" means different things; Metadata???; Data preparation & management; Workflow

Initial findings: Researchers need help with finding & choosing standards. Also w/ file versioning; applying metadata to facilitate data sharing. Lack of awareness of expertise & services the library offers.

3. How can we use data to generate new hypotheses and make new connections?

Research resources as Scholarly Products -- created an ontology to help identify scholarly expertise. CTSAconnect Project (vivoweb.org)

- Merge eagle-i and VIVO ontologies into one ontology suite (VIVO-ISF)
- Extend coverage to include representation of clinical encounter
- Modularize VIVO-ISF such that it can be made available in a set of independently reusable files.
- Integrated semantic framework of people, resources, and methods. Researchers might not have overlap in institutions, where they publish, or their coauthors; but might use the same resources, machines, have same keywords, etc. VIVO-ISF shows these connections.

The Challenge: Interpretation of Disease Candidates

Undiagnosed patient: is it a known disorder that they're not recognizing, or a new disorder?

Genotype vs phenotype (genetic code vs observable characteristics).

www.monarchinitiative.org - uses typologies, compares across patients and models of genetic disease.

Q&A

No pharmaceutical companies in credits. There have been some instances where pharma cos. don't want to disclose their resources, etc. Has there been pushback?

- They're aware of the problem too; just as concerned as we are.

Gummi Bear workshop. How long does it last, how often, provide info for others to use?

- Tried to keep it to an hour; might've been an hour and a half. Haven't shared the data about it; but have written a blog post & would be happy to share.

Could NIH mandate resource identification?

- Are working w/ NIH. Currently have a mandate for data sharing; don't have one for resource identification yet. Hoping it will happen.

Pilot project to list identifiers. Any conversation about who will maintain them?

- Yes, currently working on that. Discussion whether they should be DOIs, URIs, etc; where they should live.

Success stories?

- RRI is a success story! Also Monarch Initiative.

What about connections between people established through VIVOweb?

- Don't have a way of knowing/showing that currently.

Translating BIBFRAME - What is all this #\$\$!?: Making its potential mutually intelligible to catalogers & coders alike

ALCTS-CaMMS

Sunday, 1:00-2:30 pm

Storming the Semantic Web

Dorothea Salo, Univ. of Wisconsin-Madison

(got here late. Sorry.)

Catalogers should NOT feel left out of the conversation about linked data! It's authority control. It's AUTHORITY CONTROL ALL THE THINGS.

- How to control all the things without clashes? Need global identifiers.

Don't get stuck on the idea of a URI!

Example: VIAF, the Virtual International Authority File. Also dewey.info

Getty Art & Architecture Thesaurus (AAT)

MARC was created when boundaries were clearly defined. Now, the world's looking pretty stormy. Boundaries are being smudged; it's a big, complicated world out there. We can't really control it.

Practical implications of linked data:

- The "open world". No one person (or institution) knows everything about anything -- there are always lots of sources.
- Drawing equivalences. Just because VIAF has a record for P. I. Tchaikovsky, that doesn't mean no one else can make one -- but we *do* have to be able to link the two so that we know they're the same person.
- Designing data: Don't have to make your own identifier! Borrow someone else's. Borrow someone else's structure. You'll be making the information web denser. Libraries aren't special snowflakes, the only ones who know how to organize & describe stuff. Of course we can make data structures For Us, but it makes sense to design them with an eye to the outside world.

"Show me the quick win!" ...okay, but it's not where you think it is.

Microdata -- doesn't live off on its own like triples typically do. It's right inside the HTML on your webpage; bridge between HTML world & linked data world.

- If you're on open source (Koha, Evergreen) you can do this now!
- If you're on a vendor system -- do you control your HTML? Then you can still do this now. <itemprop>,

<itemscope>, <itemtype>.

- Vendor, but don't control your HTML? Sorry, you can't do this now. But you can bug your vendor, let them know that linked data is important to you. Can write it into your RFPs.

LC's MARC to BIBFRAME Converter: Practical, Fast, and Fun!

Philip Schreur, Stanford Univ.

Linked Data for Libraries (LD4L)

Interested in this project b/c not just linking together little pieces of information; also usage data, etc. can lend context.

Identifying data sources

- Bibliographic data -- using LC, Cornell, Harvard data. Also want to look at databases, digital materials.
- Person data -- faculty as well (VIVO repository). Want to pull that into linked data store too.
- Usage data -- getting data from BioDirect, as well as circ data from Cornell & Harvard
- External data -- things that are not traditional library sources, e.g. DBpedia.

LC's MARC → BIBFRAME converter <http://bibframe.org/tools/transform/start> (2MB size limit)

- Updates issued weekly on GitHub: marc2bibframe
- Reported issues -- can see problems that have come up, discussion, resolution

Issues w/ converter:

- Works with MARC XML -- drastically reduces the # of ppl who know how to work with it.
- Working in bulk -- big batch is even harder than putting together MARC XML for a single record. Crashed the converter a couple times.
- Difficult to use repetitively -- do music, then video, then back to music, not really an efficient way to do this.

They (Stanford Libraries) added a front end to the converter. [DEMO]

Review team = 12-15 staff members

- Targeted review by format. Do their favorite kind of music, movie, whatever; then get out the records & talk about them. "Like 'RDA therapy' catalogers did."
- Targeted review by MARC field

MARC Conversion Issues

- 730 not converted (thinks it's just an oversight; reported it on GitHub, it got fixed.)
- 504 (bibliography note) repeated twice; again, report & it gets fixed.
- 300 \$e (accompanying material) -- no next-day solution for this, it'll take a little longer. There have been some other problems like this.

Broader Issues

- Edition Statements -- difference of understanding between different players. Noticed that it showed up under BIBFRAME "Instance" rather than as a different work. Reported as an error; got back a message saying that it was done on purpose, that in their eyes that's where it belonged. All have a slightly different understanding of what a BIBFRAME "work" is.
- Consistent identifiers -- e.g. William Shakespeare showed up in different places, got assigned different

URIs. "Not ideal in the linked data world."

- Not sure what they're going to do with the holdings annotations -- call numbers coming up wrong b/c converter not looking at the holdings records.

Get involved!

- BIBFRAME intended as the MARC replacement. There are others (e.g. Europeana) but BIBFRAME is yours.
- MARC data is *your* data. If you don't care about it, no one else will either.

Toy Stories: The State of Semantics in a Semiotic World

Nannette Naught, Information Management Team (IMT), Inc.

"Toy Story": not just cowboys & cowgirls & horses anymore. We've got space rangers, Chuck E Cheese aliens, it's getting messy. Not just Andy's toys anymore, and their future as Andy's toys is in question.

Request for leeway. Nannette doesn't think content & metadata are silly, but she does think we take ourselves too seriously sometimes. (2 disclaimers: #1. OVERSIMPLIFICATION, #2. Some SILLINESS)

How do we get to a shared understanding so we can move these things forward?

Back to the story...

- Long-standing leader: Woody = MARC
- New Fangled toy w/ delusions of grandeur. MARC successor = Buzz Lightyear.
- Jealousy & mayhem! Woody throws Buzz out the window; old toys contemplate their relevance. And are the new toys really all that & a Fotomat?
- Stranded & alone -- is the internet winning? Do patrons still visit libraries? How will we survive?
- Have to learn to work together. Look beyond catchphrases & hype.

Scene 2:

- Fears surface. Is the new leader just a computer game character? Toys getting their arms broken off.
- Leader is stolen -- why do marketers want MARC records? What are we missing?
- Rescue missions! MARC replacement becomes BIBFRAME. Early implementers test the water; real world "in the cracks" applications appear.
- Reunions occur -- building bridges to legacy metadata; pathways to content

Scene 3:

- The boy grows up. The Web is firmly embedded in patrons' lives, but knowledge seekers need authority amidst the chaos.
- Fear, Miscommunications -- card catalogs, strings, etc. in long term storage; normal order disrupted.
- New users get ugly: pieces are torn off & reassembled in new ways. Systems pushed; breakage occurs. truly unruly & untrained are here to stay.
- Destruction threatens. If we can't collaborate, we'll go down in flames.
- Next Gen Acceptance is achieved (Nannette believes, or else she wouldn't be here.)

Old Paradigms: how library metadata does; roles metadata fills for library systems

- Embedded punctuation, strings, complex rule-based disambiguation; MARC, AACR2, acquisitions, cataloging

Totally New Adventures: how metadata & e-content do; role(s) metadata & e-content fill for systems

- Clean data; objects; unique, linked identifiers; BIBFRAME, etc.; RDA, etc.; subject taxonomies; ask, browse, acquire, interact.

What's really changed?

- Technology modernised; scope extended beyond our library; sequence has been disrupted.

Semantics in a Semiotic World: we're connecting things to labels to ideas. We need all three -- this is the center of Library Land. BIBFRAME helps us make these connections.

Just like the toys, we will find relevance in the new world.

Q&A

How to get beyond just objects, get into relationships (say, in VIAF)?

- Of course could try that. Of concern: what's not in VIAF. Miss a lot of journal authors, b/c it's mono authors in VIAF.

Room to negotiate the definition of a BIBFRAME "work"? Do it using profiles?

- PS doesn't think there'll ever be a single definition. Instead of thinking about defining one idea of *work*, describe relationships between definitions of "work".

How are BIBFRAME & schema.org to be used together in libraries?

- PS: Lot of schemas out there; designed to do different things. schema.org is best for getting things out on the web; wouldn't nec. want to use it for our bibliographic info b/c it's not built for it. Wasn't ever meant to do that, so we shouldn't ask it to. Use different schemas to do different things -- it's what they're for.
- NN: Use the right tool at the right time.

Haven't heard anyone address RDA as a schema?

- NN: thinks RDA is becoming a schema. It's not developed like that yet; there's a group working on it.
- PS: BIBFRAME ≡ MARC. It's a way of communicating things. It'll have to communicate RDA as well as anything else; not one or the other.

Monday, 2014-07-01

ALCTS President's Program

Monday, 10:30-11:45

The Quiet Strength of Introverts

Jennifer B. Kahnweiler, author, *The Introverted Leader: Building on Your Quiet Strength* and *Quiet Influence: The Introvert's Guide to Making a Difference*

Genevieve Owens, ALCTS President

How can an introvert make budget presentations? ...be ALCTS President? Genevieve can. So can many others of us. Our speaker, Jennifer Kahnweiler (even though she's "one of those extroverts", gets us -- and she'll explain.

Erica Findley, 2012 ALA Emerging Leader

Tap into “the quiet 50%”

Dr. Jennifer Kahnweiler

[NOTE: About halfway through, I start using “I’s & E’s” for introverts & extroverts. Not RDA-compliant, but easier.]

Asked how many introverts in the crowd? Most of the room raised our hands. Any extroverts? Not many of them here. (I wonder if this is a Tech Services thing, a library thing, or what? I mean, what portion of each of those populations is extroverts?)

Introverts -- get strength from within, while extroverts are charged up by the people & places around them.

Strengths of introverts: sense of privacy; analytical; introspective; patient

David Miller: “Our outcome is not visible. Like in the theater, no one sees the technicians, but the show doesn’t happen without them.” (slight paraphrase, sorry)

Introverts are leaders in every industry -- not by acting like “us” (extroverts). We do it by inspiring discussion, provoking thought, by calling on our strengths.

1. Why do we need introverts now?
2. How do these quiet leaders/influencers do it?
3. Impact and effect that quiet people have?

1. Why quiet influence?

Time magazine: issue on Feb. 6, 2012 “The Power of (shyness)”, on introverts. (N.B.: shyness ≠ introversion)

Kahnweiler used to get her hair cut at her uncle’s barbershop; all he’d say was “shayna maideleh, shayna maideleh” (“beautiful girl”). But it was enough. She remembers that to this day, to cheer her up when needed.

[Turn to the person next to you & describe your quiet influencer. Mine is my dad. Mom’s a total extrovert; Dad is very sociable (boy oh boy is he), but introverted. The two aren’t mutually exclusive.]

Mister Rogers -- quiet influencer. (“For those too young to remember him, you missed a good one”. Oh no! There are ppl too young to know Mister Rogers. That makes me sad.)

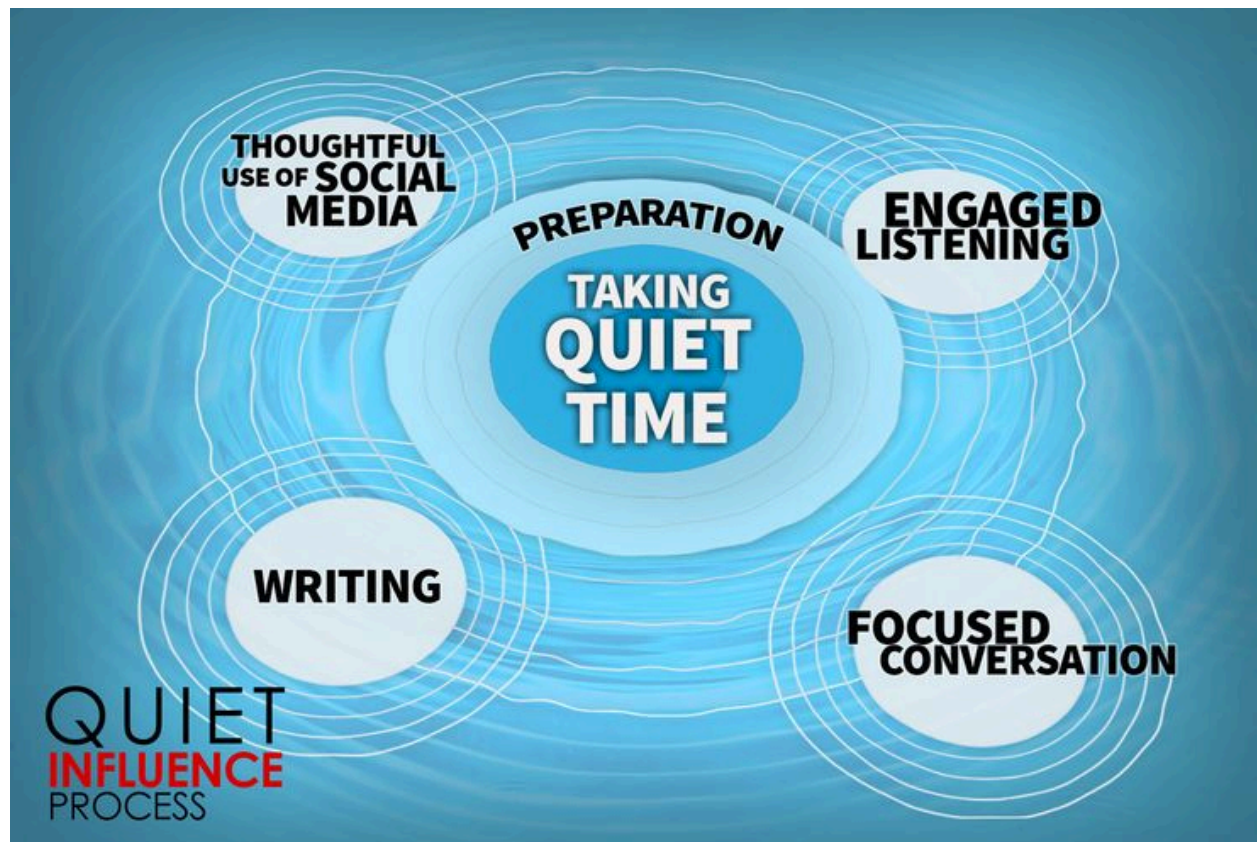
5 challenges for introverts in the workplace (value of introverts is being recognized more, but it takes time for cultures to change).

1. People exhaustion; need time alone. (“You read my t-shirt -- that’s enough social interaction for one day”)
2. Fast decisions -- more extroverted way of acting. Doesn’t accommodate introverts’ tendency to think deliberately.
3. Teams, teams, teams. I’s need time to think on their own.
4. Sell yourself (Trufax. I liked that old car commercial -- Saturn? -- where they said “we won’t try to sell you a car, we’ll help you buy one”.)
5. Put on a happy face (for those of us I’s with Resting Bitch Face, this is a doozy)

Manager, talking about an introverted person: wasn't savvy enough at the time to engage her. He later found out, after no longer president, that she had exactly the expertise he had needed. He was sitting on a gold mine & didn't know it. Won't make that mistake again.

Six strengths of introverts

"Quiet Influence Time" postcard: <http://www.pinterest.com/pin/92816442294585129/>



1. Preparation -- deliberation
2. Taking quiet time
3. Listening -- build rapport.
4. Focus -- build conversation. Address conflict; name what's going on.
5. Writing -- clarify your position in emails; hone your thoughts; make your case.
6. Thoughtful use of social media -- allows us I's to reach a large audience that we could never otherwise reach (once again, thanks for reading my notes -- I hope they're useful!)

Engaged Listening

Understand what people are saying; understand

[Write down something on our "bucket list" -- honestly, I don't have one. I just really don't!]

[Okay... after talking w/ neighbor (hers was she wants to meet Leontyne Price -- she's named after the opera

singer. That's a solid bucket list entry!) I figure that mine was learning to dance ballet en pointe. And I already did it! :-) Ten years ago, when I was 30.]

Ursula Burns, CEO of Xerox -- calls herself "Listener In Chief". Used to do "facial gymnastics" -- react to what people were saying -- and they felt intimidated. So she adopted a flat affect. People were more forthcoming.

[Conversation about bucket lists -- Daniel & Erica talking about skydiving/going to Japan. Example of engaged listening -- asking questions to get more information. "Facial expressions were interesting"; body language encouraged him to share.]

Overuses of listening can become weakness.

- Lose credibility; when we listen too much & don't talk, ppl don't know where we stand. They will project; fill in the gaps on their own.
- Don't take a stand, so ppl start overlooking us; won't tell us about new initiatives, etc. b/c they figure we'll just agree; don't know where we stand
- Avoid conflict. Ppl come to us & just vent and vent and vent; "don't know how to get them out of my office! They see me as supportive, but not good for me or the organization"

Listening tips from quiet influencers

1. Don't multitask
2. Bracket your thoughts -- "put it over here & think about it later"
3. What can I learn? Take off blinders; might get an insight about other person.
4. Move your body. On the phone a lot; sitting still a lot. Research says we should move around at least every 20 minutes!

Thoughtful use of social media

Back in the old days, had to stand in front of a crowd & bellow. Now that we can tweet & it'll reach a huge audience, new pathways to influence are open to influencers. (This is a paraphrase of a quote from the *Time* issue. I missed who said it; wasn't able to capture the exact wording.)

Giant sheep in Australia; giant peach (water tower) in Georgia -- find ways to connect. We do this online too. (I've been saying this for a while! Social media is great for I's because we can shop online for friends. Then sometimes we meet up in person!)

"Laughter is the shortest distance between two people" ~Borges

Challenge: you're missing out if you're not using it, not sharing & gathering information. Also -- have an interactive conversation. Don't just dump ideas out there.

Social media tips from quiet influencers

1. Be a giver, not a taker
2. Try it 15 min/day
3. Listen and learn
4. Use tools like Buffer and HootSuite to save time
5. Post regularly. Some of us get enthusiastic at first, but then let the ball drop later on.

What can quiet influence achieve? Ripples on a pond -- when you give it direction, it becomes a river.

What are YOU going to do to make a difference in your libraries, in your world? What's the next step you'll take? Write to Kahnweiler within next 48 hrs, and she'll send a free audio & workbook chapter from her new workshop.

Some examples:

- Slow down, try to think more "in my time"
- Quit multitasking
- Listen more
- Talk more! Speak up amidst the E's.

I'm going to start speaking up more too, or at least try to.

Going back to Kahnweiler's dad: cornerstone of the library. "Founded 1947, Expansion 1966, Expansion 1996. 'Where Dreams Endure.'" This is what he gave to the library; the library gave him so much! He was a screenwriter, and he used the library a lot. Now his papers are at the library. It took a librarian to make his work accessible to the world. Can't think of a better example of quiet influence than that.

LITA Program

Monday, 1:00-2:30

Disaster Preparedness in the 21st Century: Preserving Library Collections and Services

Catherine Tyler, Univ. of Tennessee-Knoxville

Anne Drake McClung

Branch mgr. of Goshen PL

Interest in disaster planning began with June 12 derecho & Hurricane Sandy.

"Disaster" is usually thought of as a natural occurrence or accident that causes damage & suffering. Can also be seen as any massive event with these effects.

Why "disasters in the 21st century"? What's so different now?

- Climate change -- severe weather events, billion-dollar insurance losses (<http://www.ncdc.noaa.gov/billions/>)
- Libraries have additional economic challenges -- anything that doesn't have immediate & visible payoff tends to get cut. A disaster plan
- We're (libraries and people) completely dependent on electricity in most cases. Food, water, etc. Gas pumps run on electricity.
- We (libraries) depend on Internet resources in many cases. Data is stored "in the cloud". NTBAT ("need to be able to") access outside networks.
- FEMA is recognizing the role public libraries play during & after disasters. As such, libraries may be eligible on funding. (temporary relocation funding, etc.)

Whether revising a disaster plan or starting from scratch, perfectionism is not your friend. Even starting to think about the process is better than nothing.

- Templates are available; also help get buy-in, help make your case to the people who control the library staff & resources needed to create and implement a disaster plan.

Creating a plan

1. Administrative buy-in is very important.
 - a. Documented history & likelihood of risk can help you¹ Locate your area's Hazard Mitigation Plan.
 - b. Make your case to your director/board/administration. This will help later on with fundraising, etc. (let you go ahead & buy that generator or whatever).
2. Who will write the plan?
3. Keep it feasible. Will it work for *your* library, *your* situation, *your* resources? Don't let it limit you though.
4. Staff buy-in is also important. Even if you have a plan, still need others' participation to implement it.

Step 1: Assessment of Risks (again, Hazard Mitigation Plan comes in handy)

- Flood plain? Wildfire zone? Coastal area? Earthquake zone? Tornado-prone area? Look at history of disasters in the area. (E.g. a library in Galveston, a low-lying island, is going to do much more hurricane prep than planning for an earthquake.)
- Building risks (faulty wiring, etc.)?
- Start with the cheap stuff & low-hanging fruit. Also, cost/benefit analysis.

Step 2: Inventory of Physical Assets

- Ideal: detailed lists, scans, pictures.

Step 3: Inventory of Essential Services -- think of it like packing a suitcase. What do you need most? How do you prioritize?

- Access to library catalog
- Staff/public computer use
- Staff/public printer, fax, scanner
- Lighting for stacks, reading
- Restrooms?
- Public meeting space?
- Climate control? (for equipment, collections, patron/staff comfort)
- Other?

E.g. can come up with a manual way to do checkouts & -ins, but need lighting. Restrooms perhaps less important than computer use; other places have restrooms.

Step 4: Plan for Safety of People

- Patron evacuation plan; "safe room" plan
- Staff evacuation plan & "safe room" plan
- Periodic disaster drills -- does everyone know how to secure the library in case of a disaster?
- Periodic check of detection equipment (fire alarm, carbon monoxide detector, etc) to make sure it works

Step 5: Plan for Safety of Collections

- E.g. when it's going to rain hard, make sure the buckets are out.
- Shelving bolted to walls esp. in earthquake-prone area
- Floods: have equipment & materials up off the floor.
- Surge protectors; UPS (uninterruptable power source)

Step 6: Communication Plan ("Touching Base")

- Staff call in; give supervisor cell #s to staff. Then supervisors report to higher mgmt.
- Ppl living near library may have greater responsibility if phone services is out
- Have a place to post notices in a post-disaster situation

- Build communication into your disaster plans. “It’s like teenagers -- communication won’t happen unless you make it happen”

Step 7: Re-establishment of Services (may mesh with other community service).

- Assess damage & address safety issues
- If building is OK (often happens) activate solar panels, power up generators (if you have them)

Reducing vulnerability to a LTPO (long-term power outage)

- Risks: severe weather, transformer security, blackouts/brownouts
- Solutions: alternate power sources; plan for continuing basic services; maintenance of a core emergency collection (including “how to” books, other creative solutions. Also one good set of encyclopedias, even if a bit dated; disaster recovery information). From a coll. mgmt. POV, might be good to get rid of these books; from a disaster mgmt plan, might be better to keep them.

Can be hard to get into disaster planning/recovery mode! E.g. “I’d rather you print that sign out rather than making one by hand”; but the power’s out! Can’t fax stuff either...

Generators (get one at next sale -- they’ll sell fast when needed, then can get for half-price after.)

- Pros: peace of mind; some aspects of “business as usual” can continue; pure (reliable, surge-free power) due to electrical management technology; can be hooked up to natural gas lines/tank, or diesel
- Cons: can be expensive; takes up space; has to be kept secure (vandalism/theft); fuel storage/availability. Have to check that using a generator won’t void warranties on computers, etc.

Could my library actually run in solar power? Yes! But... need space for solar panels. Also depends on the amount of sun your area gets. Also need batteries for power independence (for power at night or on cloudy days), and converters for AC --> DC and back (storing energy, reusing stored energy). Grid tie-in can also help cut costs -- feed your energy out when you’re not using it.

- www.dsireusa.org <-- State & local incentives for going solar
- Start with a demo! If you’re getting the “are you crazy?” look when you suggest going solar, start small. Maybe get a solar-powered laptop & check it out to students; have a solar panel to run a public computer (& put up a notice about it)

University of San Francisco -- small amount of solar panels. Goes to the whole campus, not just the library; but part of the energy feeds back into the library. They’ve had for at least 7 yrs.

Sierra, NV -- had a grant of \$4K to put solar panels on the library. Expected to save \$13K/mo on utilities (A/C costs here are astronomical).

Other alternative sources of power: geothermal, wind. Solar is biggest now in US, but can also investigate others.

Special Discussion II: Reducing vulnerability to loss of data

“We’re in the future, whether we like it or not” -- days of hosting our own software & data are gone.

“Going to the cloud vs. doing it in-house” -- InfoToday article by Concordia U. librarians (\$ paywall)

Talk to your service provider (perhaps via customer support) & find out about the center where your data is stored. Physical location (coastal/inland)? Redundant power? Redundant internet? Physical security? 24/7 active monitoring & customer support? “Remote hands & eyes”?

- Also want to look at the agreement you've signed or will sign. The biggest thing is to be able to withdraw your data whenever you want or need to. Also need frequent & available backups. Think about whether you want to keep a copy of your data on-site. (Might be additional \$\$ but it should be an option.) For some things, the old-school way (CDs, etc. of data) is good enough. Don't keep backups at the library though!
- Need also stable climate control, with refunds in case of data loss due to climate. Vendor should put their money where their mouth is.

Other issues

- Evaluate your ILS. Also, can you access it via cell service?
- Patriot Act issues -- federal agencies may be able to approach the service provider in order to obtain remotely-hosted data, and SP is not required to notify the library. (Some do have a policy of notification, but need not.) Storing data in Canada doesn't get around this either -- if the data is US-owned, the US gov't can have access & library need not be notified.
- Cyberattacks (need local virus/spyware protection). May not be able to let patrons download apps, etc.
- Internal systems malfunction/conflict (e.g. Windows XP vs upgrade); value of keeping system updated
- Combo of backups & open source software
- Consider DSL -- don't put all your eggs in one basket
- Again, maintain core physical collection

Special Discussion III: Community Networking

- Who is your community? Identify potential partners
- Investing in community relationships costs nothing, but can help greatly in times of crisis. Also helps build community. Get to know your first responders -- you'll have a familiar face in times of crisis.
- Think through & prepare for disasters before they happen; this way FEMA doesn't have as much to clean up.
- For public libraries willing to work w/ first responders, FEMA funding may be available. Two ways to apply for disaster mitigation funding: before (preparation) and after (recovery).

Keeping it Simple

Getting started w/o getting overwhelmed

- There are links out there that'll help you so you don't have to reinvent the wheel. Use others' work & tailor it to your own library, your own situation.

Q&A

Grants?

- You have a better chance of getting a grant if you're working with your first responders, but libs can apply for grants as individual entities

Comment: NLM website has a link to a disaster management template -- folds down & fits in your pocket.

Emergency community. There's a 10-week training course offered; also gets you on their radar.

If your library does not have a good inventory, insurance will not pay things you don't have documented. (Not just collections but also computers, furniture, etc.) If you make a claim, insurance cos will also call vendors & ask for prices on items comparable to what you're claiming, etc.

Lots of info out there, training on disaster planning via ALCTS-PARS. Also, ALA Preservation Week is always in the spring. Heritage Health Network does "May Day" (May 1st of course). American Institute for Conservation (AIC) has a lot of info on their website. Lyrrasis, NADCC (?)

Librarian Emergency Response app available for download; can be used offline.

Tuesday, 2014-07-02

Closing General Session featuring B.J. Novak
Tuesday, 9:30-11:00 am

Address: Barb Stripling, 2013 ALA President

(Got here a little late.)

We must tell our stories; empower our patrons to tell their stories, how their lives have been changed by libraries; must "build the public will" to support libraries.

(Introduction of incoming division presidents)

Ken Stewart, AASL; Ann Riley, ACRL; Mary Page, ALCTS; Ellen Rearden, ALSC; Kathleen Moeller-Piper, ASCL; Rachel Vasek, LITA; Diane , LLAMA; Mary Neil, PLA; Joseph Thompson, RUSA; ; Christopher Shoemaker, YALSA

(introduction of new Board members and President-Elect Sari Feldman)

(introduction of new ALA President Courtney L. Young)

Introduction of Keynote Speaker

Courtney Young, 2014 ALA President

Thanks Penguin Young Readers for sponsoring today's speaker, BJ Novak)

The Book With No Pictures

BJ Novak

Speaking to audience of 1000 librarians -- great experience, and "very specific sexual fantasy". (Tone note: he was poking fun at the stereotype, not being icky. He said it in a wry voice.)

1st thing he wanted to be when he was a kid was a librarian; asked parents for a date stamp for Hanukkah, so that he could make his books into a library. Never loaned out any books, but he was pleased that everything was cataloged and in order, and he was living and sleeping in a library.

"Hey librarians, call me -- 323-377-1360" with a sexxy picture of him! Hah. :-D

"Sense of controlled rebellion that humor gives him". The world has rules; when those rules are bent and broken, the results are interesting, complex, and funny. Take what we know for granted and then twist it; punchline is a new perspective. But humor for children has to provide both the setup *and* the punchline. It teaches children that there are rules, and what those rules are, and then breaks them.

Uncle Shelby's ABZ Book -- steal money from their parents' wallets and mail it to Uncle Shelby -- parody of a children's book (and BJN's favorite) that predates his actual children's books. Roald Dahl, another favorite, has dark & wry humor (probably would not have been a fan of the Berenstain Bears.)

Joined the *Harvard Lampoon*; when it came time to run for an office, he ran for Librarian and won.

(clip of BJN talking about Ben Franklin in his 1st job out of college -- stand-up comedian.)

"Only say what you like, and only keep what they like". Goes in w/ material he thinks is funny; goes back with cut of that material, the part that the audience found funny. Gave him confidence b/c material was tested & proven; but humbling b/c visceral memory of the failures.

Writer on *The Office*; credited with a certain # of episodes, but it's hard to distinguish where one person's contributions start & where another's ended.

(Clip from *The Office* where Kelly & Ryan announce their divorce -- written by BJN & Mindy Kaling in character.) Collaborative, performative nature of the writing room.

Comedy was the only thing that allowed him to write in only his own voice; but had been writing for such a rich cast of characters on *The Office*. Reserved a theater in LA; announced he would be performing live stories; went in with a pen so as to be able to edit right there on stage. This is how he created *One More Thing*, collection of short stories.

(Clip of BJN performing "Never Fall In Love" from *One More Thing*)

Waiting for book to be published, spent time with college friends & their kids. Asked the kids which books they wanted him to read; hoped it would be stuff that fit with his sense of humor; usually ended up being about trains. (FACT. Kids dig trains.) Funny power dynamic -- BJN is the one who can read, but the child is in charge b/c he's choosing the book; and the book is in charge, because it's what's written. "It's like having a little producer handing you a script, a 2-yr-old Harvey Weinstein -- 'Make it your own, but let's stick to the script. I may ask you to do it again.'"

"Book with No Pictures", a book with its secret humor: the ridiculous words it would force a grownup to say. Same writing & editing process. Write what he likes, keep what the kids like.

Had to do two things: had to take its humor from breaking the rules, but also had to introduce those rules.

(Clip of BJN reading to kids at a school in Long Island)

No more fun way for a kid to learn a rule than to learn how it can be used to their advantage.

No more important rule to introduce to a kid than the power of the written word. It can be their ally; it can make the world a better, funnier place.

Q&A

Any more public readings? Come to a book festival?

- Yes! Will the library be open on the 4th of July? ("It can be!")

Was that phone number real?

- It's his real Ralph's discount card #.

(missed question -- high school books?)

- Would like to do something where there are gaps in the stuff that's available; maybe things that would've appealed to the kind of kid he was.

Fascinated by the way BJN uses a design/prototyping method for writing; get feedback, etc. Was there a time when BJN felt really strongly that something was funny, but audience didn't? Did he keep it, or take it out?

- All the time! Used own sense of pride and shame to determine how long to keep at something he really liked. Ultimately, sees self as an entertainer rather than a solitary artist. But if he can find any shred of evidence that someday someone will like it too, he'll keep trying and honing and reworking it.

Loved BJN's work in *The Office*. Ever feel fragmented, with all the roles he played?

- Actor, writer, director -- all blended together. Wasn't much of a divide. Actors thought like writers, writers thought like directors, etc. ("If I'd had to do set construction I would've had to use a different part of my brain.")

Talk to best friend/worst enemy Mindy Kaling about writing a children's book?

- Certainly can! Thinks she would love it. There's a lot of the child in MK -- very bright child -- wants everything. Wants fun, magic, romance, calmness; has the hunger a child would.
- Very common for people to ask him "can you ask ____ of Mindy?" He has a process for it.

Shown drafts to father? Any collaboration?

- Sent him drafts; loved his notes ("after a few days for the anger to burn off"). His father's career as a ghostwriter taught BJN a lot about channeling another person's voice. Taught him a lot about writing for TV, about writing stories.

Loves that Shel Silverstein was an influence. Toyed with irreverent verse like that?

- Yes, but the toy broke. It didn't come naturally; wonders how naturally it came to SS. ("Probably pretty naturally.") (This is true! When Otto Penzler asked Shel Silverstein to write a short story for the *Murder for Love* anthology, Silverstein basically said "I don't know what I would be able to write for that. Oh wait, I have an idea". He was a firehose of creativity. -lkm)

Advocacy for libraries? Pleading w/ BJN to be a stronger advocate for libraries and librarians. Our voices are sometimes lost. Would love it if BJN would speak loudly about his passions.

- "'Speak loudly about your library' is a good phrase." Will think about how to do it best. His father is a major library advocate; would like to know more about what to say. Will be visiting a lot of libraries with his book.

Also: maybe a really hilarious show about working in a library? He'll be getting a lot of material on his upcoming visits.

- Will think about it! ("You're not taking a percentage of the show, are you?")

“Heads-up on what you just stepped into” -- so many of us here at/in ALA who are very serious, invested in advocacy.

- Ppl don’t know what’s going on in the transition into this next era in libraries.

(“Libraries: not the same old ‘SHH’.”) Question: BJN seems to measure his work by laughter alone. Why is that?

- What he knows how to do; 1st thing he felt good at as a kid, thing he feels most comfortable with, thinks it’s what ppl expect him to bring. A lot of his career is getting further and further out away from the punchline; scary for a writer, but got a lot of good feedback.

Would love to have a “READ” poster of BJN; ppl know him from *The Office*, would be great to show that he loves libraries too.

You’ve stepped into something; but haven’t stepped in so far that there’s no help. Lots of support at ALA for advocates.

- Appreciates it.

Library student, working on documentary about librarians; would love to have BJN in film.

- Can’t wait to tell agent he booked a gig while he was here! ;-)

BJN: “it was an honor and a privilege to be here.”

FIN.