



Center *for* Research Libraries

GLOBAL RESOURCES NETWORK

Shared Print Discussion Meeting

ALA Annual 2009

July 10, 2009

The meeting was hosted by the Center for Research Libraries (CRL) at the University of Chicago's Regenstein Library. CRL President Bernard Reilly opened the meeting with some introductory remarks. He noted that the shift from print to digital library collections will require that the library community build robust and efficient print archives in order to have confidence in the long-term preservation of print as well as digital archives.

Co-convenor Robert Kieft (Library Director, Occidental College) stated that the goal of the meeting was to share information about shared print initiatives already planned or underway throughout North America.

Most of the meeting consisted of reports by representatives of individual consortia or libraries involved in shared print projects, which are summarized briefly below.

University of California

- The University of California operates two shared storage facilities (Northern and Southern Regional Library Facility) which are both nearly full with 3 to 5 years capacity remaining, capital projects at UC are on hold
- A UC system-wide space planning task force recommended shared print projects with holdings mostly at RLF's including
 - Dim archive of JSTOR titles, 700 to 800 titles, 25 million pages, page-level validation, funded by JSTOR + some UC funds, learned a lot about validating materials
 - Retrospective IEEE volumes, developed standards for issue-level validation (vs page-level as on JSTOR), distributed validation work performed at the libraries
 - Prospective shared print copy of licensed e-content, a single shared print copy stored at SRLF, hoping to move to the network level with CRL
 - Shared print copy of Springer e-books, not in storage, kept at Merced campus
- Shared approval plan in process
- Beyond UC, discussing regional collaborations
- Proposal to Mellon Foundation [since approved and funded] for a shared Western Regional Storage Trust, to develop a collaborative framework for a distributed retrospective journal archives
- Proposed partnership with CRL to house print analogs of licensed journals

OCLC Research

- What is the role of regional consortia, how to articulate national print archiving strategy?
- Print journals preservation project
 - Worked with ten libraries, 200 titles distributed in print only. Looked at: what are the risks of print-only, what are the qualitative issues? Who will manage print preservation for the community?--not so much environmental risks as uncoordinated efforts lead to loss. Better to have cooperative efforts.
 - Risk has more to do with uncoordinated action (deaccessioning) rather than physical deterioration, even for top tier research institutions who would be expected to retain
- “Cloud Library” project with CLIR, Hathi Trust, NYU, ReCAP
 - How to exploit shared print and mass digital books
 - Hathi Trust (Google Books) as a potential surrogate for print collections; what is the intersection between Hathi and ReCAP stored items?
 - 20% duplication between the average research library and Hathi (based on OCLC records)
 - Some libraries may consider paying for access to digital collections so they can house the print volumes offsite.

Orbis-Cascade Alliance

- Distributed print repository for JSTOR and ACS titles, distributed throughout libraries in the consortium
- No regional facility to host materials
- Challenges were: checking to the page level is difficult for participating libraries. Time and staff needed for manual work are not available; the idea of using identifying bookplates was eliminated due to time and staff considerations. Item level information was difficult to get, lack of standardization of summary holdings statements
- MOU said in-library use only, doesn't work for monographs
- Distributed data is a major challenge – how to compare holdings

PASCAL

- Colorado plus Wyoming
- Prospector union catalog and lending system from Colorado Alliance includes non-Alliance members in state,
- PASCAL shared storage center contains about 1.2 million volumes, currently building another module, anything in PASCAL can be loaned through Prospector
- Shared serials print runs—no duplicates policy, duplicates can be discarded, not enforced at the facility.
- University of Denver is raising money for a storage facility.
- A pilot shared purchase plan (print) was abandoned in favor of an e-books shared project.

University of Illinois

- They've had a last copy policy with a shared catalog (ICCMP) for two decades, distributed, each library retains its last copies and identifies in catalog
- They have a budget for cooperative collection development of unique materials; acquisitions are housed in various Illinois facilities.

CARLI

- A new last copy project for small or mid-size libraries to donate volumes to University of Illinois, mostly monographs, which are the last copy in the state.
- Four months into the pilot project and so far 500 volumes have been offered. A web form for submitting information about individual donations has been developed.

Committee on Institutional Cooperation (CIC)

- Print demand won't go to zero so they are committed to maintaining one print copy of Google Books items (digital in Hathi Trust)
- A separate project with Google to digitize Gov Docs, expect universal accessibility for these documents, will give to GPO Access, Hathi Trust and Google.
- Sheet-fed capture, disbound, vs original Google scans of intact volumes

Ohio and OhioLink

- Five state depositories, funded separately from OhioLink, now being brought under OhioLink
- Project to dedup journals across the depositories, currently 8 million volumes stored, 3 repositories are full, the other 2 have a few years of capacity left
- Developing principles: ownership shifts to state of Ohio, not individual institutions, retention agreement, can't recall volumes to campus
- Depository materials are recorded in individual local catalogs but not in the union catalog for OhioLink, now implementing a new shared catalog for the 5 depositories, begin pilot test later this year
- Hope to open access eventually to any Ohio institutions in OhioLink
- OCLC analysis of OhioLink circulation patterns showed that 80/20 rule no longer applies, about 6% of OhioLink collections represented 80% of use

Canada (report about various projects)

- Quebec project: PREBOOK-shared agreement for subject strengths and JSTOR titles
- Council of Atlantic libraries, efforts have stalled.
- COPAL-have a distributed repository system.
- BART-uses U of ? repository for deposit of print
- Thunder Bay agreement: space and money are big issues-worse with economic downturn. Rating of Canadian institutions based in part on number of titles held in library so deaccessioning is a problem.

- OCUL print registry: titles held in various facilities, but they will run out of space. Last copy is important issue
- Risk Management/other issues: Inaccurate or incomplete holdings information, issue level information, overlapping jurisdictions, shared purchases.

PALCI

- Distributed repositories for ACS and AIP journals. Problems: getting an accurate inventory, terms for participation (currently ten years), light and dark archives (dark can be used on site), how to build collections, metadata only in local catalog, libraries validating at issue, not page level
- Cost borne by archiving library
- Metadata in local catalog and PALCI website
- Individual archive validates to issue level
- E-books project-retention of print copy-project just starting.

ASERL

- Two dozen storage facilities among ASERL membership, looking at these facilities as a “bank” which could be used by all, need agreements
- Earlier project to review monograph holdings in 9 ASERL storage facilities found unexpectedly small overlap of duplicates, only 100,000 volumes. were weeded (concept was 3 facilities should own before deduping)
- Concerns: are storage records accurate, any archive would have to be light
- Struggling with print archiving issues for government documents (there are 12 government document repository libraries in ASERL, only 3 in CIC). IMLS grant proposal to build “centers of excellence” in certain disciplines, need to catalog materials. Regionals can’t discard based on digital copy. GPO is beginning to look at this.

Washington (DC) Research Library Consortium (WRLC)

- Breaking ground for expansion of shared library storage facility (“Harvard model” high-density facility, currently holds about 1.3 million volumes).
- Recently changed deposit policy from “first come-first served” to “shared copy” (no duplicates in storage) for journals to conserve facility space and campus library space.
- Formal retention agreement: Holding library retains ownership but commits to providing intellectual content (even if they withdraw in the future).
- Disclosure through existing shared catalog (OPAC), not yet surfaced to WorldCat.
- Processing deduplication centrally at the storage facility, no validation but a general review of condition.
- American University transferred 100,000 bound journal volumes in 6 weeks, about 25% already in storage were discarded, saved the donating library 17,000 linear feet of campus shelf space

UK Research Reserve (UKRR)

- Retrospective journal archive with 3 copies of identified titles held, one at British Library storage facility and 2 in campus libraries
- British Library provides access/delivery via BL Document Supply Centre, library copies are backup

- Central funds provided by Higher Education Funding Council support BL storage, funds also distributed to individual libraries to support deaccessioning effort

Law Library Microform Consortium (LLMC)

- Members are law libraries in US and Canada
- Converting microform holdings to digital, also scanning print, a few major law libraries scan onsite
- Content comes from members deduping and weeding pressure, NBS in Minnesota scans and hosts content, print once scanned is stored in Salt lake City (it has been disbound) Metadata is done in St Louis (they have a backlog of 800 titles)
- Currently they've scanned 40,000 of 2 million volumes, they have 21 million searchable page images
- Authentication at the page level, considers it a permanent legal archive.
- Access is by subscription and the subscriber has universal access to all content.

JSTOR

- Two print archives validated at page level, one at University of California (dim) one at Harvard (Dark) both are environmentally controlled, shared costs to validate pages. Usage is small for print archive at UC. International repository being sought but not yet there.
- Challenges: Keeping up with the scanning, there are 32 million pages in JSTOR, 4-5 million added each year.
- JSTOR is expanding beyond journals literature, and don't have paper archives in place for this material.

Ithaka report: What to Withdraw

- Why to keep print at all?
- Yano study says page-verification crucial for guarantee of completeness, lack of data to know whether sufficient copies exist
- Rationales for keeping print:
 - Correct digitizing errors
 - Rescan under new standards
 - Scholarly use
- Costs are huge for these efforts.
- Limited time requirement for keeping print at all, a range of material types with different time spans for retention.
- Page level verification-need at least two copies (Berkeley study)

Group Discussion

After the project presentations, there was an open discussion of issues. The following is a brief summary of the major points that were made.

1. Shouldn't print archiving be done at the national level? PAPER report in 2003 outlined characteristics of national approach. Ithaka report (Roger Schonfeld) outlines some practical steps.
2. Government documents are a problem. They're taking up a lot of space (most hold over a million vols.)
3. National means we have to have better information about who has what and what is being used. Is the regional level an interim step?
4. It is regional because that's where the data is.
5. We don't know how many copies we need.
6. Numbers wise, a 20 year requirement for access is projected, we may need 2 page validated and 4 volume-verified worldwide (Ithaka Study)
7. Number of copies also depends on cost, more risk, less money. One problem is lack of metadata. Google and JSTOR and Hathi Trust projects generated lots of metadata down to the page level – can that be captured and used to facilitate validation?
8. The economy is going to give us permission to do a lot of things. But we need to move soon. Now is the time. OCLC needs to provide incentives for information to pull into their catalog.
9. The economy is a driver of change. We think we own our library budgets but the deans have final say and we need to lobby with them for this work.
10. We need region wide consensus
11. Chuck Henry article for CLIR addresses that University administrators have no concept of the scholarly record. This is something librarians have always done under the radar. Access through closed community models is not going to be enough.
12. Budget items may not be for a library anymore but rather for service of materials.
13. Do we need a national manifesto, a PAPER 2?

14. Are participants willing to work with CRL to get this done? Consortia can commit or work on an international effort. How can we link these various efforts together? We need a timetable, a budget, a clearinghouse for these print projects, documenting workflows, policies and procedures, a knowledge-base.
15. OCLC shared print group has done some of this.
16. Manifesto would be a nod toward this work.
17. Page validation is a difficult sell when we're moving into digitization.
18. There is a model for CRL in DC [??], membership has a fair amount of faith in CRL and this is a business model. Members will see it in their own self interest to participate while other can pay as nonmembers. A trusted network is what is needed
19. CRL's proposed 6-month planning effort is needed, meet again at ALA Midwinter in Boston..

Attendees

1. Bernie Reilly, President, Center for Research Libraries (CRL) (Co-Chair), reilly@crl.edu
2. Robert Kieft, Director, Occidental College (Co-Chair), kieft@oxy.edu
3. Kimberly Armstrong, Assistant Director, Center for Library Initiatives (CLI), CIC, karms2@staff.cic.net
4. Catherine Hamer, Interim Associate Director for User Services, University of Texas at Austin, hamer@austin.utexas.edu
5. Brian E. C. Schottlaender, University Librarian, University of California San Diego becs@ucsd.edu
6. Judith Russell, Dean of Libraries University of Florida jcrussell@ufl.edu
7. Ivy Anderson, Director, Collections, California Digital Library Ivy.Anderson@ucop.edu
8. Martha Brogan, Director, University of Pennsylvania brogan@pobox.upenn.edu
9. John Burger, Executive Director, ASERL jburger@aserl.org
10. Tom Clareson, Senior Consultant for New Initiatives, Lyris Tom.Clareson@lyris.org
11. Kimberly Douglas, University Librarian, Caltech kdouglas@library.caltech.edu
12. Sharon Farb, Associate University Librarian for Collection Development and Scholarly Communication, UCLA, farb@library.ucla.edu
13. Leslie Firth, Assistant Librarian, Systems, Carleton College leslie_firth@carleton.ca
14. Rachel Frick, Senior Program Officer, IMLS rfrick@IMLS.gov
15. Julia Gammon, Head, Acquisitions Department; Interim Marketing Manager, UA Press, University of Akron jgammon@uakron.edu
16. Michael Levine-Clark, Coordinator of Collections Management, University of Denver michael.levine-clark@du.edu
17. Constance Malpas, Program Officer, OCLC malpasc@oclc.org
18. **Emily McElroy, Head, Collection Development, Ohio State University**
mcelroye@ohsu.edu [is this correct?]

19. Lizanne Payne, Executive Director. Washington Research Library Consortium (WRLC)
payne@wrlc.org
20. Kathleen Richman, Executive. Director, Law Library Microform Consortium (LLMC)
KathleenR@llmc.com
21. Mark Sandler, Director, Center for Library Initiatives, CIC msandler@staff.cic.net
22. Karen Schmidt, University Librarian, Illinois Wesleyan University karens@iwu.edu
23. Roger Schonfeld, Manager, Research, Ithaka Roger.Schonfeld@ithaka.org
24. Peggy Seiden, Director, Swarthmore pseiden1@swarthmore.edu
25. Bryan Skib, Collection Development Officer, University Library, University of Michigan,
bskib@umich.edu
26. Michael Spinella, Executive Director, JSTOR Michael.spinella@jstor.org.
27. Dona Straley, Associate Professor, University Libraries, Ohio State University
straley.1@osu.edu
28. Tom Teper, Associate University Librarian for Information Technology Planning & Policy,
University of Illinois tteper@illinois.edu
29. Melissa Trevvett, Vice-President, CRL, mtrevvett@crl.edu
30. Marie Waltz, Special Projects Librarian, CRL, mwaltz@crl.edu
31. Catherine C Wilt, Executive Director, PALINET wilt@palinet.org