Shared Print Archiving and Formation of a National Trust

Will your campus participate in a collective print preservation strategy?

by Sam Demas and Wendy Lougee

Our nation’s libraries are nearly full and are experiencing severe cultural, space and budgetary pressures that affect how they manage their collections. American academic libraries hold about 1 billion print volumes and are growing at a rate of 25 million volumes per year. Most academic libraries, including the 80+ high-density storage facilities, are already at or are quickly approaching storage capacity.

Aggressive digital library development notwithstanding, there is simply not room in our libraries to accommodate projected future growth. At the same time, system-wide, library construction has slowed to a trickle. Universities are increasingly reluctant to build more space for books. To the contrary, libraries are being pressured to reduce their physical footprints. Growing interdisciplinary scholarship, reduced staffing, and increased production of digital content are driving the closure of branch libraries in universities. Lastly, there is increased pressure to convert high value central campus collection storage space to user services space.

As a consequence, there is intense pressure on libraries to withdraw redundant print holdings that are accessible online. How does a campus administrator help to manage these challenges? How will the nation responsibly manage its enormous print legacy collections in the years ahead?

The challenge of storage space for print collections is just one factor in an increasingly complex ecosystem of collections and content. The sidebar (see at end of article, online version only) “Rebalancing the Portfolio?” outlines how the concept and reality of “collections” has expanded and what this means in terms of institutional investments. The issue of shared print archiving must be understood in this larger context of collective and cooperative investment in infrastructure, preservation, and access.

It is increasingly clear that isolated, institution-by-institution curation of the nation’s library collections is no longer a sufficient strategy. As we address user preference for digital resources and effect a transition to a hybrid mix of collections that is increasingly digital, collective action is needed to ensure that a “national collection” of print materials is preserved in service to scholarship. Care is needed and support must be provided to ensure that libraries forced to weed their print collections do not make mistakes and withdraw materials that should be retained.

Shared and coordinated print archiving offers a strategic solution for collective action to assure preservation of sufficient print copies nationally as collections are drawn down locally. Following is a sketch of a nascent movement towards shared print archiving, highlighting some key models and programs now emerging, and suggesting how your school can support collective action and prepare to participate in emerging shared print archiving.

Coming Soon —

Environmental Sustainability of Academic Libraries
The Cost of Storing and Preserving Print Materials

Research on the cost of storing scholarly materials points to significant cost savings through cooperation, collection consolidation, and format transition. Cost data from Courant and Nielsen represent present discounted value for digital storage and for four different print shelving systems. These cost estimates include the cost of cleaning, maintenance, electricity (heating and cooling), staffing, and circulation, as well as the amortized construction cost for building space.

### Annual Average Storage Cost per Volume

<table>
<thead>
<tr>
<th>Storage Type</th>
<th>Cost per Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital Storage</td>
<td><em>$0.15 - $0.40</em></td>
</tr>
<tr>
<td>Open Stack</td>
<td>$4.26</td>
</tr>
<tr>
<td>High Density</td>
<td>$0.86</td>
</tr>
<tr>
<td>Hybrid stacks storage</td>
<td>$1.53</td>
</tr>
<tr>
<td>(first 10 years in open stacks, then moved to closed)</td>
<td></td>
</tr>
<tr>
<td>Hybrid stacks storage</td>
<td>$1.99</td>
</tr>
<tr>
<td>(first 20 years in open stacks, then moved to closed)</td>
<td></td>
</tr>
</tbody>
</table>

With a billion volumes in academic libraries, and less than 10 percent of them in storage facilities, the pressure locally and nationally to convert valuable central campus real estate to “higher yield” purposes will only increase with time.

### Coordinated Shared Print Archiving

The idea of shared print archiving is that with widespread digital access, “libraries could share their print storage, keeping only several copies nationally or regionally, rather than duplicating substantial swaths of their collections...the savings from sharing of this kind could be substantial!” The goal is to ensure the most cost-effective preservation of the print record through a coordinated system of shared responsibility.

Shared print archiving seeks to maintain and ensure the existing national breadth and depth of collections while thoughtfully “managing down” unnecessary duplication given the availability of digital surrogates. The idea of sharing responsibility for print storage and access prompts and supports the move from the traditional emphasis on local collections as free-standing silos to one of managing collections as a collective resource.

"...ensuring secure, environmentally stable storage conditions for print collections is an essential investment in protecting them from the ravages of time and ensuring enduring access to our legacy collections."

This approach is based on the notion that our research collections deliver maximum value when managed as a collective resource, i.e., their “value to the academic enterprise is less as a locally-owned asset than as a pooled resource”). Finally, ensuring secure, environmentally stable storage conditions for print collections is an essential investment in protecting them from the ravages of time and ensuring enduring access to our legacy collections.

### Shared Digital Archiving

Scanning of library collections, including the Google Books project, is producing a large corpus of digital surrogates for some portion of our print collections. For example, HathiTrust, a partnership of 52 institutional members who contribute financial support and digital content, currently holds more than 8 million volumes of digitized content, approaching the size of the largest university libraries.

Data from Hathi indicate that its holdings already substantially duplicate research library holdings (~30 percent overlap) and the corpus is growing quickly. While Hathi’s costs will grow as the resource grows, no single institution could manage nor support the scale of this service.

Despite increased investment in digitization by libraries and publishers, however, realizing the full potential of shared digital archiving is stymied by copyright law, which prevents a large corpus of in-copyright, but out-of-print materials from being fully usable. These constitute 74 percent of the Google Books corpus, for example.

### Collective Action as a Strategy

While libraries have long been activists in promoting collaboration within higher education, the current climate of economic constraint is precipitating urgency for more radical collaboration. The potential now exists to leverage local collection management and storage costs in support of a collectively managed collection designed to provide preservation and access for the nation.

The growing corpus of collectively stewarded digital surrogates makes it possible to provide digital access to and preserve digital copies of a large share of the nation’s traditional library materials. This strategy can be joined with shared print archiving: cooperative retention of print versions as backup and for their immense artifactual and intellectual value, and freeing up space by eliminating any unnecessary collection redundancies. In combination, shared digital and print archives will support consolidation and rationalization of print holdings within and across institutions. In turn, this will create economies of scale that benefit individual institutions and the community as a whole.

The potential benefits of this collective strategy are prompting early signs of a significant conceptual shift from a traditional focus on managing local collections to thinking of collection management as a national, system-level collaborative challenge.

National and regional initiatives are beginning to develop potential building blocks of a coordinated approach to “ensuring long-term survivability of the scholarly record at a cost that
is sustainable for the research library community as a whole”4. What follows are highlights several promising initiatives, or elements of a system of regional and national “trusts” to which we will all contribute and on which we will all rely.

Emerging Models of Shared Print Archiving

Centralized Archiving. The 80+ high-density storage facilities in the U.S. provide a form of centralized storage serving either single institutions or consortia/university systems. The Harvard Depository Library and Cornell University Libraries Annex Library are examples of single institution facilities. The Washington (DC) Research Library Consortium, the Research Collections and Preservation Consortium (ReCAP), and the Five Colleges Library Depository (MA) are examples of large regional facilities serving academic library consortia and/or state university systems. In most cases ownership of materials placed in storage is retained by the depositing institution, however in some cases (e.g., Five Colleges) ownership is transferred to a third party serving the collective.

Distributed Archiving. An emerging alternative to centralized storage facilities is “archiving in place” or “distributed archiving”, in which institutions in a consortium agree to retain journal runs in their campus libraries (under specified storage conditions) on behalf of the membership. When a sufficient number of schools have agreed to retain a full run of a journal, others can withdraw their own redundant runs with assurance of continued access to the shared copies.

Some of these are publisher-specific programs, focusing on readily identifiable, widely held journals that are readily available digitally. These include JSTOR, IEEE, American Institute of Physics, Elsevier, American Chemical Society, etc.

Others, however, are often more broadly based, aiming to secure retention agreements for journal titles in many disciplines. Examples of distributed archiving include the Association of Southeastern Academic and Research Libraries, Committee on Institutional Cooperation, Greater Western Library Alliance, Ontario Council of University Libraries, and Triangle Research Libraries Network.

Hybrid Models. The United Kingdom Research Reserve (UKRR) and the Western Reserve Storage Trust (WEST) employ both centralized and distributed archiving to knit together a coherent shared print archiving strategy for a region or nation. UKRR, a partnership between the UK higher education sector and the British Library, aims to preserve three copies of low-use print research journals nation-wide. One copy will reside in the British Library and two in the collections of other UKRR member libraries.

With funding from the Mellon Foundations, WEST has begun implementation of the largest and most ambitious shared print archiving program in the USA. Building on years of cooperative storage and preservation within the University of California system, WEST is expanding to incorporate 89 libraries west of the Mississippi in a carefully planned program that will likely lead the way for the nation in evolving business models, policies, operations, and coordination in shared print archiving.

Discipline-based Programs. The Center for Research Libraries is working with the U.S. Agricultural Information Network and the Law Libraries Microform Consortium to build nation-wide discipline-based approaches to shared archiving that emphasize building a national resource to enhance research in specific disciplines.

State-based Multi-type Library Programs. The University of Maine was funded by the IMLS to develop a statewide strategy of shared management for the legacy print collections of a state-wide network of public, academic, and governmental libraries. The Maine Shared Collections Strategy Project is currently working on retention policies, a service delivery model, and a shared collection management and preservation model for the state.

MINITEX, a regional, multi-type library cooperative based at the University of Minnesota, is working towards re-purposing the Minnesota Library Access Center (MLAC), its underground storage facility as a node in the nascent national shared print archiving system.

Decision Support Tools. A new class of decision support tools is being designed to normalize library catalog data and allow comparison of one’s collection with those of other libraries, with authoritative lists, and with use data. JSTOR has developed a tool to aid in decisions about withdrawing JSTOR back files.

R2 Consulting’s Sustainable Collection Services software provides tailored reports and de-selection dashboards that combine local circulation and item data with WorldCat holdings, Hathi Trust Digital Library holdings, and various authoritative title lists.

SUNY Geneseo library staff have developed an open source Gift and De-selection Manager that assists in gathering data in support of weeding decisions.

Challenges in Shaping Shared Print Archiving into a National Program

Shared print archiving initiatives will continue to cohere rapidly in widely differing contexts (e.g., small local consortia, state-wide programs, regional programs, and disciplinary programs with no geographic parameters).

In the absence of a U.S. national library or a nationally coordinated system of higher education, it is unclear how these independent efforts will be effectively knit into a coordinated national program. Optimists predict that a network level shared print archive framework will emerge within
the next 2 to 4 years. However, it may well take considerably longer.

The culture of higher education will require time to adapt to a new model for stewarding scholarly resources, and myriad practical and policy issues must be addressed. Following are some of the key challenges in shaping a federated preservation effort:

- Development of a policy-based leadership structure to shape and coordinate a national infrastructure for collective action to which all colleges and universities may decide to opt-in, or not.
- Establishment of a decision-making structure that engenders trust among partners and a mechanism for governance, commitment, investment, policy development, and risk management that ensures that legacy materials will be preserved and made accessible in perpetuity;
- Development of a sustainable budget model for shared investment in a shared collection;
- For widely held, little used materials, how many copies do we keep for the nation? More research is needed on the number of copies needed to serve regional communities.
- There will be faculty push-back, particularly in the humanities, against national collections consolidation. Some will liken it to book-burning.

How to Participate in the Formation of a National Trust

Campus administrators, working with campus library professionals, can explore shared print archiving in your region. Librarians can be the ones to recommend how best to position the college or university to participate as a node in one or more inter-connected systems of library collections. Administrators should be prepared to invest in collection action as nascent models mature by supporting librarians in building the essential local groundwork for national collective action:

- Develop a phased long-term collection management plan for your local collections based on careful analysis of collections, their strengths and weaknesses, duplication with other collections, etc. Are you making the best possible use of your library space? How can you optimize the potential for shared collection management?
- Identify the unique and little-held materials in your collections. Digitize unique materials and consider other ways in which you can contribute to the national collection.
- Engage in ongoing faculty outreach/engagement; involve faculty in discussions of your library’s space needs, collection management plan, and regional and national initiatives. Involve faculty in discussion about the costs and benefits of opting in to shared print archiving initiatives vs. “going it alone”.

Conclusion

While it will take years to fully realize the goal of a national trust comprising a collective collection for the nation, the necessity of managing down our legacy collections is a pressing reality today. Institutions concerned that large-scale weeding of collections be undertaken responsibly, as well as those interested in achieving economies of scale and benefits of cooperation, should actively position themselves to engage in shaping the future of shared print archiving – locally and nationally.

Sam Demas is College Librarian Emeritus, Carleton College & Principal, Sam Demas Collaborative Consulting; and Wendy Lounge is the McKnight Presidential Professor and University Librarian, University of Minnesota.

References

5 Both these programs, plus others are described in articles appearing in the November 2010 issue (volume 22 no. 5) of Against the Grain.
Sidebar: Rebalancing the Portfolio?

Library collections, by definition, represent a cumulating body of campus resources. Universities have long held to the traditional measures for their libraries, essentially a growth paradigm that perpetuated an insatiable appetite for annual increased funding and periodic increased commitment of new space. Further, the strategy has fed an equally insatiable publishing industry with annual inflation for publications -- in better years -- averaging 7-9 percent.

The challenge of space and storage for collections represents only one aspect of the complex ecosystem now associated with collections and content. Several forces have, over time, shaped a new paradigm that is all about access -- current and enduring access. Simply acquiring content is no longer sufficient, and the risks for sustainable access are high. The essential investment portfolio for “collections” (now broadly defined) includes:

**Digital infrastructure:** Digital content, whether licensed from publishers or created on campuses, requires robust network infrastructure for useful and stable access. In addition to the traditions of description (cataloging), libraries now invest in systems to ensure stable linkages to publisher-hosted content and also tools to enhance discovery across myriad sources of content. This evolving discovery environment serves the campus community, but also enables a global community of scholars to identify and tap unique resources, wherever they are located.

**Resource sharing:** No institution or library can acquire everything needed by the local scholarly community. Request and delivery systems to facilitate sharing of collections (including digital content) are integral to ensuring students and faculty get the information resources they need. While sharing physical published works continues to rely on traditional delivery systems, libraries have also invested in systems to share digitized or digital content via the network.

**Preservation:** Libraries have a history of attending to preservation of print collections, largely through good environmental storage and reformatting (for example, through microfilming) of at-risk publications. Opportunities now exist to digitize analog materials as a reformatting strategy, or acquire digital copies from commercial sources. With the availability of digital surrogates, collective decisions on management and storage are enabled. However, digital content carries similar risks for degradation, requiring investments in digital preservation through third party services or collaborative, shared infrastructure (e.g., HathiTrust).

**Publishing:** Libraries have also invested in opportunities to craft new models for publishing -- whether it is collective investment in alternative publishing venues (e.g., SPARC, the Scholarly Publishing and Academic Resources Coalition) or institutional investment in publishing infrastructure (e.g., Stanford University’s High-Wire Press). The goal of these investments is to support the dissemination of scholarship in cost-effective ways and to introduce new models for publishing and managing rights within the academy.

It could be argued that libraries have always supported resource sharing, access, and preservation -- and to a lesser extent, publishing. What has changed is the imperative of the associated portfolio of commitments and the emergence of new costs that are essential to ensuring and sustaining access to scholarly content.

How balanced is your institution’s portfolio? Is the return on your collections’ investment enabled by good access and a commitment to sustainable access for the long-term? Just as collective action offers promise to develop a coordinated system of print publication storage, the collective and collaborative investment in infrastructure, preservation, and access offers significant advantage for scholarship today, and in decades to come.