

Collective Collections

Final Report

September 25, 2015

revised November 15, 2015

Introduction

The term “collective collections” refers to **shared** content and content that can potentially be shared across institutions. Collective collections can be in any format, are local, regional, national, and global. They provide more substantive and useful research experiences for scholars, students, and other interested parties than collections held by any single institution. Collective collections can encompass legacy print materials, rare and distinct (special) collections, research data, subscription-based and open digital content, and emerging forms of digital scholarship, all of which will require different tools, strategies, and partnerships to ensure current and future access and use. In all cases, the work of supporting the most effective access, retention, and preservation will take place through a collective investment that respects and supports local interest. As ARL leverages individual collections into a collective, we anticipate increasing institutional specialization in both content and expertise.¹

ARL will stimulate the creation of deep and wide platforms for ensuring that knowledge resources essential to the academy and to support research and scholarship are accessible and sustained through federated networks of print, digital, data, and artifactual repositories, created and managed by collectives of institutions (e.g., HathiTrust) in North America and beyond. The SHARE initiative is a key part of this strategy, operating at the network level and unifying and making visible distributed resources.

ARL activities should have high impact across the academy, not just for individual campuses or institutions. ARL’s work will not only guide the creation of governance, shared protocols, best practices, trusted relationships, and financial models, but will in some cases extend to convening parties to pursue the creation of new entities that conduct work in this space. Some existing multi-institutional entities, including but not limited to HathiTrust, DPN, and DPLA, are not managed by ARL but are mission critical to ARL Collective Collections. The Collective Collections Design Team will explore ARL engagement with these initiatives, including strategies for making investment and participation in them by individual ARL institutions more efficient. ARL’s membership also includes national collections and repositories, such as the Library of Congress, the National Agricultural Library, the National Archives and Records Administration, and the Center for Research Libraries, which positions ARL to coordinate projects which either leverage or contribute to such collections.

ARL can also accelerate successful collections activities that begin on a local or consortial level, and is in a unique position to help avoid duplicate efforts or to initiate new projects at scale. This Design Team will need to explore the balance of local branding of projects (which comes with responsibility, accountability, investment, and credit) with collective stewardship and funding.

Framework and Context for Collective Collections

¹ The CC Design Team can leverage the Mellon-funded Center of Excellence (CoE) model research by ARL Leadership Fellows in 2013-14. <http://www.clir.org/pubs/reports/pub163>

The 2012 ARL Issue Brief “[21st-Century Collections: Calibration of Investment and Collaborative Action](#)” set the stage for this Design Team and outlined the future of research library collections as collaborative and multi-institutional. The brief noted that “rather than focusing on acquiring the *products* of scholarship, the library is now an engaged agent supporting and embedded within the *processes* of scholarship,” and that “collections management is now a juggling act” among local needs, shared investment, and coordinated decisions. The report acknowledges both constraints (shrinking budgets, subscription prices, and inflation) and incentives (e.g. community tool development, aggregating content) fueling increased collaboration.

The Collective Collections Design Team has identified the need for collective agendas and corresponding projects in the following five broad categories in order to address the central challenge of the 21st Century Collections report: preserving and making accessible the scholarly record.

I. Shared print: A coordinated North American strategy to ensure the preservation of the print record. This area should include engagement with HathiTrust as the digital library and partial basis for a distributed print archive for books and government documents. ARL’s shared print strategy should also engage with other partners (for example, historically black colleges and universities) to ensure diversity and inclusion of the print record. Though North American in focus, it is critical to monitor related international activities and potentially participate in them if benefits of scale can be realized. Finally, part of ARL’s shared print strategy should be to maintain the Association’s long-standing conversation with scholarly societies, such as the Modern Language Association (MLA), who work in this area.

II. Licensed electronic resources: With more than half of ARL’s aggregate collections budget devoted to e-resources, opportunities for a collective agenda in this area include addressing the financial sustainability of the scholarly publishing environment and the terms under which electronic resources are acquired, preserved, and shared. This category also includes developing new, open models of access, ownership, and distribution of content; and attention to the scale and efficiency in associated workflows for both commercial and open content. Common infrastructure could reduce redundant investment in operations that sustain the current scholarly communications system.²

III. Cultural heritage and primary source material: Preservation and discovery, especially across institutions, and coordinated expertise of content held by libraries, museums, and communities, or created on the web. This area should include coordinated engagement with, for example, DPLA as a critical piece of the discovery infrastructure. Like shared print, ARL’s strategy in this area should be focused on diversity of materials and inclusion of stakeholders, including institutions and communities beyond ARL libraries. This area may explore collective solutions related to community-based collections,³ including innovative custodial and governance arrangements.

IV. Institutional and/or disciplinary research assets: Track, preserve, and make visible and accessible the complex scholarly output of our institutions, including records, publications, research data, and other materials associated with the scholarly research process. Within this area, the “library as publisher” and the Institutional Repository is where the Scholarly Dissemination Engine and Collective Collections Design Teams overlap.

² Potential development models include Quali GOKb, DPLA service hubs

³ [SPEC Kit 347](#), published July 2015

V. Preservation: Beyond projects specific to types of information resources, ARL needs a collective agenda for technologies and platforms specifically related to preservation. This area should include coordinated engagement with DPN, APTrust, and others.

Within these five broad categories of focus, Collective Collections projects might support existing infrastructure or build where it's needed; develop or implement policy; coordinate specialization of content and/or expertise; or develop or support common tools. Creating new kinds of research tools to make scholarship and source materials for scholarship more available and easy to find is a creative and forward looking way of showing the value of research library collections. (This may also be a critical intersection with the Scholarly Dissemination Engine team.) Collective Collections should reflect the diversity of the human experience and be constructed according to principles of accessibility and inclusivity.

Current ARL activities that fit within the scope of Collective Collections include:

- ARL participation in the [MLA Future of the Print Record](#) Working Group, an important librarian-scholar partnership (Shared Print)
- ARL-LYRASIS e-book licensing model (Licensed Electronic Resources)
- [Model license terms](#): Maintain in partnership with Advocacy and Public Policy Committee (Licensed Electronic Resources)
- SHARE (Institutional and/or disciplinary research assets)
- ARL-SAA partnership to host Digital Archives Specialist courses (Institutional and/or disciplinary research assets; Cultural heritage and primary source material). Potential exists here for an intersection between Collective Collections and the ARL Academy.

New projects that the Collective Collections Design Team sees emerging out of the conversations thus far include:

I. Shared Print

- ARL analysis of standards and strategies for disclosure of print retention agreements and holdings, including those of regional efforts such as WEST and ASERL, and support for greater disclosure to national registries such as CRL's PAPR database and OCLC WorldCat.
- Development of model policy agreements that enable shared access, retention, preservation, archiving, and deduplication efforts between and among institutions.
- Analysis of overlap and redundancy of ARL print collections. Explore potential partnership with OCLC. Potential OCLC partnerships include issues around print monographs or international efforts such as OCLC's existing collaboration with British universities.⁴ The [COPPUL Shared Print Archive Network \(SPAN\)](#) and Jisc's [National Monograph Strategy](#) are other regional models to watch.
- Continued engagement with the MLA Future of the Print Record Working Group around coordinated stewardship of the print and digital record by ARL libraries. This Working Group might include direct engagement with the HathiTrust (Shared Print Monograph Program).
- ARL should continue to hold strategy discussions around US federal documents holdings in print, in conversation with the HathiTrust.

II. Licensed Electronic Resources

⁴ "RLUK Collective Collection in Context – the OCLC Research Perspective." <http://www.rluk.ac.uk/about-us/blog/rluk-collective-collection-in-context-the-oclc-research-perspective/>

- Work with ARL assessment committee to collect data about the sustainability of the current scholarly publishing environment and ARL investments in open access content and infrastructure.
- Evaluate ARL's existing e-book partnership with LYRASIS paying particular attending to new and emerging publisher models.
- Collaborate with the Scholarly Dissemination Engine on the larger issues of e-book and other licenses.

III. Cultural heritage and primary source materials

- The web is too big for any one institution to archive and will therefore require coordination and specialization. Project to coordinate institutional specialization.

IV. Research Assets:

- Analysis of institutional repositories, including investment, shared metadata issues (standards and rights, e.g. [DPLA Standardized International Rights](#)), relationships with disciplinary repositories such as arXiv and SSRN, consolidation of regional or consortial repositories, and use of standards and identifiers for interoperability, especially with respect to SHARE. SHARE is a collective tool for making research assets more discoverable. Through this project, the Collective Collections Design Team will actively engage with SHARE to ensure that its development supports local and collective collecting.
- Development of new visions, strategies, and tools for stewardship and discovery for digital archiving (e.g. email, manuscripts, institutional records).

V. Preservation

- Conduct an analysis of ARL investment in large-scale preservation initiatives such as DPN, AP Trust, Portico, and LOCKSS. Are they overlapping, redundant, or complementary? How relevant are these tools to collections? What is the return on investment?

Top three to five projects within Collective Collections to present to the membership

- ARL analysis of standards tools and strategies for disclosure of print retention agreements and holdings, including those of regional efforts such as WEST and ASERL, and support for greater disclosure to national registries efforts such as CRL's PAPR database and OCLC WorldCat. Include cost-benefit analysis of different approaches.
- Work with ARL assessment committee to collect data about the sustainability of the current scholarly publishing environment and ARL investments in open access content and infrastructure.
- Development of new visions, strategies, and tools for stewardship and discovery of born-digital archival materials (e.g. email, manuscripts, institutional records).

The Collective Collections Design Team notes that coordinating shared print and print retention did not rise to the top of member priorities in the open session at the Association's Fall Meeting. This discrepancy may be due to the amount of activity in the shared print space, making the focal area appear less urgent than economic sustainability of scholarly communications and born-digital archiving, both of which garnered more enthusiasm from the membership.

However, the ARL Board, in its discussion on October 5, did see merit in continuing the partnership with the Modern Language Association (MLA) on the future of the print record, which does address this focal area from a broader perspective.

The ARL Board also noted an interest in the analysis of existing digital preservation strategies.

Incorporates Board feedback from July, 2015; edited week of August 24; finalized September 2. Additional Board feedback received September 2015; edited and finalized September 16, 2015. Minor edits following Design Team Meeting, Sep 25.

Collective Collections Design Team

John Culshaw, chair (Iowa)
Nancy Gwinn (Smithsonian)
Sara Lowman (Rice)
Sarah Michalak (North Carolina)
Bernie Reilly (CRL)
Judy Russell (Florida)
Scott Seaman (Ohio)
Bonnie MacEwan, Board Liaison (Auburn)
Staff Liaison: Judy Ruttenberg