

Concerted Thought, Collective Action, and a Collaborative Future for Access to Print Monographs (v.2)¹

Dan Cohen, Digital Public Library of America; Seth Denbo, American Historical Association; Kathleen Fitzpatrick, Modern Language Association, Convener; Mike Furlough, HathiTrust; James Grossman, American Historical Association; Charles Henry, Council on Library and Information Resources; Geneva Henry, George Washington University; Robert H. Kieft, Modern Language Association; Deanna Marcum, Ithaka S+R; Judy Ruttenberg, Association of Research Libraries; Elliott Shore, Association of Research Libraries; Andrew M. Stauffer, University of Virginia; Elizabeth A. Waraksa, Association of Research Libraries; Steven Wheatley, American Council of Learned Societies

Summary

“What is the place of print in the digital age?” That is the fundamental question informing this white paper, and one for which many and varied answers have been registered over the last 20 years as publishing has engaged with digital technologies and as libraries have addressed collaborative possibilities for the collection, housing, and preservation of serials, books, analog and digital media, government publications, manuscripts, ephemera and gray literature, etc. Our work focuses on monographs but builds on research, experimentation, and policy development among library consortia and organizations for a number of formats and on a thoughtful sequence of publications and conversations conducted over the last twenty years that persuasively articulates two correlate conclusions: 1) the collective cultural heritage embodied in the vast print record held in libraries’ general, circulating collections must be preserved and made accessible for future generations, and 2) that no single institution or existing organization can provide a satisfactory solution for sustaining this cultural legacy. A coherent, collective action is required on the part of the entire higher education community.

¹ The version of this paper posted for comment on *Humanities Commons* in December 2016 bore the capacious title “Concerted Thought, Collaborative Action, and the Future of the Print Record.” Early comments on that working draft called our attention to a disjunction between the scope suggested by our title and the focus of our work, which is on monographs. In this second version of our paper, we have revised the text to more clearly define our focus on the corpus of 19th- and 20th-Century printed books that libraries make available for general circulation. We assume that printed books from the hand press era and many other rare and unusual printed books published since are held in special collections or governed by arrangements that tend to secure them. When we refer, therefore, to “library collections” or “library materials,” we are referring to general collection monographs from the last two centuries of industrial-scale book production. Our work is prompted by and depends conceptually on the several studies released by OCLC Research in the last decade on the “collective collection,” and we have learned much from corollary work done on serials by the Center for Research Libraries, Western Regional Storage Trust (WEST), and Scholars Trust as well as by Eastern Academic Scholars’ Trust (EAST) and other consortia on shared print agreements for monographs.

In this paper we propose a research agenda for, and suggest a path toward, a national system for print book management that derives from core values informing academic research and teaching and will provide over time a salient public good. Recognizing the rational imperative of the 20th-century model of large-scale acquisition of print book collections by single institutions, we recognize as well that contemporary robust networks, sophisticated communication, and infinitely reproducible digitized content offer us unprecedented means to augment and extend this traditional model of collection and access. Our proposal combines suggestions for a policy and governance structure and the rationalization of existing, high-density book facilities with strategically built and managed new facilities that would be operated as a coherent system. Fundamental to this built environment is the ongoing digitization of print book content. Our proposal seeks to create a commonwealth of analog and digital resources in service to scholarly productivity along with new means for discovery that together expand human capacity in ways that are more efficient, cost effective, and elegant than is possible within the academy today. Our work thus affirms the value of and many uses for physical monograph collections; enables libraries to design rich on-shelf browseable collections in concert with other forms of access, including digital surrogates and remote storage; and calls on the library community to commit through a national collaboration among librarians, scholars, and institutions to defending and maintaining the value of print books, preserving them in rational ways, and making them more accessible and useful to scholarship.

The title, “Concerted Thought, Collaborative Action,” refers to the logical consolidation and management of currently diffuse and distributed collections, and also describes the collaborative effort necessary to make such an ambitious and necessary project work. This is a rich amalgam of technical, behavioral, and cultural considerations. Implicit in successfully addressing this grand challenge is an elusive but foundational condition of engagement: mutual trust among institutions, their libraries, scholarly communities, information specialists, and administrators, a trust based on a shared commitment to the highest quality of teaching, learning, and scholarship.

I. Libraries Today, the Place of Print, and a New National Context

As noted in the introduction, the framing issue of this paper is the place of print monographs in the digital age. While ‘the place of print’ as an often heated topic dates back many decades to the appearance of microform as a medium for reproducing and distributing paper based content, the question is posed ever more vigorously as digital technologies have made mass reproduction and

Internet distribution possible. As print runs for academic books and library budgets for books have declined over the years,² digital resources, tools, and organizations devoted to maintaining and making accessible the electronic surrogates of the print corpus have burgeoned, and new forms of scholarly publishing have emerged. Despite this vigorous flourishing of digital resources, the number of books maintained in thousands of college and university libraries across North America is still enormous; these volumes and those that continue to be added to library collections remain as integral to scholarship as they have been for centuries.

Assessing contemporary libraries and scholarship we note the following:

- As stewards of the print record, libraries must judiciously consider the place of print materials within a dominant digital ecology;
- The amount and types of research data being created by scholars are continually growing and evolving;
- Readers' relationship to print is evolving, and their needs for it are changing; it is now possible to compile use cases for reading in one way or another and in one format or another and to align technological services with such multiple modes of reading;
- While scholars continue to work in longstanding modes of research (e.g. close reading and interpretation of primary sources, writing journal articles and monographs), changes in the techniques of *reading* are also becoming clear: a) machines are increasingly used to compute humanistic content, and such machine reading could drastically change the nature of discovery; b) large scale solutions like textual analysis and data mining tools facilitate microscopically close reading, as well as distant reading of massive content; c) distant or digitally-enabled reading entails new methodologies, new intellectual strategies, and new discovery tools for collection description, management and preservation;
- New technologies facilitate access for many who could not obtain or use the physically printed form

In response to this complex, hybrid nature of library collections, new investigative methodologies, and shifting reader response, we believe a comprehensive collaboration at the national level is both feasible and urgently needed for the following reasons.

² *Academic Library Book Purchasing Trends*, Proquest, 2016. Available at http://contentz.mkt5049.com/lp/43888/438659/D187_Ebooks_Aquisition_whitepaper_v5.pdf

1) *A robust, ubiquitous infrastructure.* Library and scholarly communities enjoy the availability of a robust, credible, and developing information technology infrastructure (provided by OCLC, local networked catalogs, other discovery services, and physical or digital delivery services) that enables not only the sharing of physical items and the metadata about them, but also the integration of print with digitized collections and fully born-digital resources. That this infrastructure for collection access, management, and preservation is essential to the conduct and progress of education is a strong guarantor of its sustainability and thoughtful development over time.

2) *Erratic budgets.* Budgets are constrained at almost every institution of higher education; developing fiscal strength through collaboration is intuitive. As our blog post, “Preserving the Print Record in the Digital Age,” noted,³ current budgetary and resource pressures add urgency to any question pertaining to the disposition of the print record. The numbers are stunning: in 2012, academic libraries alone in the United States house and care for 1.1 billion print volumes⁴ on shelving that, at 10 volumes per foot, would extend almost 21,000 miles. Despite the proliferation of digital materials, the acquisition of print materials continues, which not only consumes campus space—itsself an increasingly scarce commodity—but also competes for space with services and facilities that support new teaching and research methodologies at a times when new money for fulfillment of materials stewardship is not increasing. Although off-site storage has been controversial among scholars, some of whom object to the loss of open-stack browsing, libraries turn to off-site shelving facilities to relieve space pressure in on-campus buildings, placing little-used materials in remote warehouses for on-demand retrieval. These off-campus book-housing centers themselves entail construction and maintenance costs, require staffing, and consume energy, albeit at lower total cost per square foot than open-stack campus libraries; collectively, this storage investment is not insignificant, despite its efficiencies.

3) *The costs of local decision making.* Our call to national action is not just about stressed budget numbers but focuses importantly on the ways by which these constraints, and the

³ Available at <https://printrecord.commons.mla.org/>.

⁴ Tai Phan, Laura Hardesty, and Jamie Hug. “Academic Libraries: 2012 First Look.” U.S. Department of Education, National Center for Education Statistics, 2014. Table 4, p.7. NCES-2014-38. Available at <http://nces.ed.gov/pubs2014/2014038.pdf>. ARL member libraries held roughly 687,000,000 as of the 2013/14 fiscal year: Martha Kyrillidou, Shaneka Morris, and Gary Roebuck. *ARL Statistics 2013-14*. Washington, DC: Association of Research Libraries, 2015. 13. In her 2007 report for OCLC, “Library Storage Facilities and the Future of Print Collections in North America” (<http://www.oclc.org/content/dam/research/publications/library/2007/2007-01.pdf>), Lizanne Payne estimated (p. 6) that “As of July 2007, at least 68 such [storage] facilities exist in the United States and Canada, holding more than 70 million volumes in total.”

responses to them, affect scholarship and intellectual advancement. Given impacted spaces and strained budgets, as well as the continued acquisition of print materials, most libraries are cognizant of the necessity of reshaping their collections, not to mention their relationships with each other, consortial organizations, publishers, and their students and faculty. Such reshaping often results in decisions about retention and deaccessioning that are frequently, and understandably, made at the local level. Records of titles removed are rarely shared outside of the local or consortial context, and no scalable infrastructure now exists to record and publicly query commitments to retain printed works. Whether volumes are housed in a remote facility or removed permanently from the active shelves, thousands of individual, local decisions strongly influence the future of the print record and thereby determine in part the future of scholarship⁵.

Given the situation described above, a comprehensive change in both mindset and procedure is needed in order to allow local decisions to be made within the larger national context. More funding at the institutional level is not a solution; a greatly increased investment to support the traditional model of the disposition of print monographs might alleviate local pressure but would not conduce to an efficient and effective new conceptual solution. Consensus is growing that the higher education community needs a national collection management system because such a system can ameliorate the inhibiting costs of academic information organization and enhance scholarly productivity—a sizable investment certainly, but libraries have gained experience in the last 10 years with the design and implementation of shared print collections and creating technologies and policies to manage them. The investment for we advocate has the potential to unburden current constraints and strengthen the mission of higher education.

II. The Benefits to Scholarship

Our work on this paper began as a reconsideration of three seminal reports - *Statement on the Significance of Primary Records*,⁶ *Preserving Research Collections: A Collaboration Between*

⁵ We do not intend in making this point to detract from or criticize the procedures that have accrued to the management of print or the people making them. The decisions made are difficult, even agonizing, and executed with acumen and diligence possible under current circumstances.

⁶ Modern Language Association of America, *Statement on the Significance of Primary Records*, 1995. <https://www.mla.org/Resources/Research/Surveys-Reports-and-Other-Documents/Publishing-and-Scholarship/Significance-of-Primary-Records/Read-the-Report-Online/Statement-on-the-Significance-of-Primary-Records-Modern-Language-Association>.

Librarians and Scholars,⁷ and *The Evidence in Hand: Report of The Task Force on the Artifact in Library Collections*.⁸ While it is instructive to note the changes between 1995 and 2016, it is also important to state that some principles from the last century (and this can be legitimately extended to the past millennium) remain unquestionably true today: genuine guiding principles. “For centuries the library has been the repository of the written record and a powerful symbol of human intellectual achievement,” argued a Mellon Foundation study on university libraries published in 1992.⁹ The purpose of the research library “is, and has been throughout its long history...to collect the record of human knowledge, to organize and make it available for use, and to conserve it for future use.”¹⁰ Collecting, organizing, conserving, and making accessible our common intellectual achievement are incontestable responsibilities of great libraries.

The digital revolution has changed libraries and users’ perceptions of them. Local ownership is pervasive; accumulated academic information promotes a distinct pride of institutional place: libraries continue to compete for rankings based on arithmetic, counting annually the number of books, journals, staff, and funding of local resources.

We recognize these elements of competition in academic and institutional culture, but we maintain that the hierarchy and competition they entail are no longer productive for the effective support of scholarship. A national collection management system for print books would replace this paradigm with a collaborative model, and we believe the benefits for librarians and scholars alike are considerable:

- Better documentation of the books on the shelf of both circulating and non-circulating repositories, and better metadata about those books, including their condition, history, and archival status;
- Increased attention to book history and reading culture;

⁷ Jutta Reed-Scott, *Preserving Research Collections: A Collaboration Between Librarians and Scholars*. Washington, DC: Association of Research Libraries, 1999. Available at <https://www.mla.org/content/download/2935/80094/prc.pdf>.

⁸ Stephen G. Nichols and Abby Smith, *The Evidence in Hand: Report of The Task Force on the Artifact in Library Collections*. Washington, DC: Council on Library and Information Resources, 2001. Available at <http://www.clir.org/pubs/reports/pub103/pub103.pdf>

⁹ Anthony M. Cummings, Marcia L. Witte, William G. Bowen, Laura O. Lazarus, and Richard H. Ekman, *University Libraries and Scholarly Communication: A Study Prepared for the Andrew W. Mellon Foundation*. Association of Research Libraries: Washington, DC, 1992, 1. Available at <http://files.eric.ed.gov/fulltext/ED371758.pdf>.

¹⁰ Reed-Scott, 1999, 3, quoting Sheila Dowd.

- A more mindful systematic digitization of print journals and books than has heretofore been the case;
- Increased attention to preserving unique, scarce, or unusual materials;
- Potential for libraries to gain flexibility in staff to more deeply support the needs of scholars and students in the areas of - for example - special collections; data management; or scholarly communication.
- More careful collecting with the increased ability to tie new acquisitions to older materials;
- Widespread collaborative models for the curation of print materials - especially monographs - in libraries' general collections;
- Democratized access through collective management and the sense of common purpose it requires, target and plan digitization.

A well-managed collection of book collections, with item level metadata subject to advanced search strategies and data mining techniques, would significantly enhance discovery and long term curation of this vital record while preserving the print record for future generations.

Creating this item-level data for tens of millions of volumes will entail considerable and complex effort, but it will be necessary in order to support collection management and scholarly work.

In re-imagining their physical environment, librarians, university administrators, and scholars can also re-imagine their roles as stewards of the print record. A national collection management system for books would secure and keep safe millions of print volumes with an acceptable, cost effective redundancy that would be audited annually; it would make this collection visible and easily discovered by cataloging using universal linked data schema that improves one's understanding of the content of the print record. Volumes would be retrieved upon request, and would be easily available for digitization. Software exists to facilitate virtual browsing across vast arrays of titles and content by discipline or by chance, the results organized, annotated, and integrated at the scholar's discretion: a mutually reinforcing combination of the physical with the digital.

This solution cannot be accomplished by a few institutions working in tandem; it cannot be achieved by regional efforts that employ different schema of metadata, are separately governed, and do not integrate. It is also beyond our collective reach if competitive acquisition, independent action, and disciplinary silos persist. This proposal entreats the intersecting

communities of librarians, scholars, administrators, and information technologists to think concertedly in support of a larger scale endeavor and a more encompassing and productive reach of professional expertise and responsibility. If we do not proceed in a collaborative, collective fashion, we will certainly be subject to the prevailing inconsistencies and local disruptions that compel us to think within parochial fences, inhibiting our higher agency in support of a compelling public good.

III. Initial Thoughts toward the Envisioned System

The following sections present areas of focus requisite for effective planning and execution of a national program.

A. Planning and research agenda

Numerous perspectives are required for effectively planning a national system to steward the print (book) record, including those of scholarly societies, research libraries, higher education administrators, and information technology specialists, including experts in network design, electronic storage and preservation, input of user experience, and contributions of scientists working in various fields associated with knowledge organization. Some of the more salient areas for research include:

- Review and analysis of current practices and needs in both scholarly research and teaching.
- Identification of existing projects and programs that could serve as cost, policy, technology, and governance models: potential partners and collaborative organizations, focused on either print storage or digital resources. These can include the Big Ten Academic Alliance, CRL, DPLA, EAST, HathiTrust, Scholars Trust, WEST, and others.
- Examination of selection criteria, including publication dates, disciplines, formats, and/or at risk areas based on research/data analysis for immediate decision-making. Questions to consider might include: Where is the need for a national print retention and/or digitization policy for books most pressing? For example, is it late 20th/early 21st century books that do not exist in digital editions; English-language literature that is at risk of being over-weeded on the assumption that other libraries would retain copies of books for long-term preservation; area studies/foreign-language printed books that have not been a digitization priority either in the US or abroad?

- Assessment and evaluation of the necessary and prudent level of duplication in a national-scale system.
- Exploration and demarcation of the boundaries where variations among print copies become meaningful in terms of the number of copies to be retained in a national system; definition of the descriptive information needed about these copies in order to make them useful for scholarship.
- Development of guidelines for what should be digitized and why; definition of priorities to guide the digitization of the corpus of print monographs; articulation of standards for digitization so that such materials are machine discoverable, machine actionable without excessive human cataloging.
- Development of systems/protocols/workflows for tracking/describing/identifying items and for discovery and use (access mechanisms, delivery etc.).
- Development of policies for maintaining and preserving materials and collections; organizations for oversight, research, and decision-making; access mechanisms (including copyright, physical and digital delivery).
- Investigation of ways to increase participation of publishers in preservation.
- Exploration of possibilities for consortial acquisitions of electronic and print monographs and broader sharing in given groups.
- Development of a plan for libraries to move budgets to the collaborative acquisition and preservation of books over the years.
- Review and analysis of interpretations of copyright laws that would encourage monograph digitization and wide distribution.

B. A schematic description of a national system

A national plan will consist of a policy and governance structure that specifies roles, responsibilities, the composition and management of the collective corpus, rules for access and use, and other elements as suggested by the agenda above. It will also include facilities for housing those parts of the shared print book collection that do not remain in local stacks. The physical plant of a national collection management system would likely include some existing facilities (e.g. ReCAP), with the possibility of creating large-scale, super high-density facilities geographically distributed in North America. Sites would be chosen for a combination of relatively unimpeded physical access with a minimum susceptibility to geologic and meteorological disruption based on historical record and current predictability algorithms.

The integration of digital technology, essential to this system, has several major components:

- a robust infrastructure connecting libraries and repositories;
- sophisticated tools and applications that allow users to search across very large arrays of titles and content descriptions and request items;
- digital methods of aggregation of resources, data mining, pattern recognition, semantic analysis, and related features to help readers navigate the corpus;
- the ongoing digitization of print monograph content based on need and relevance to research and teaching; and
- suitable redundancy of print holdings.

C. Governance

An acceptable model of governance for the national system is essential. While many calls for the establishment of large scale preservation programs have been issued over the years, so far, none of these calls has been realized in a concrete, sustained project.¹¹ For all practical purposes, then, there is little if any precedent for an academic program of this scale, influence, and longer term potential. However, we are encouraged to think that a national-level system for collection management can now succeed because of the examples set by initiatives such as HathiTrust, Eastern Academic Scholars' Trust (EAST), the Digital Public Library of America (DPLA), and the many other consortial projects underway among academic libraries. We infer that a governing body would be representative of the various constituencies that contribute to the design, execution, and sustainability of the project. Annual audits, oversight, and other external methods of routine assessment and checks and balances will be required. Not-for-profit organizational models might be reviewed for structuring governance.

D. Costs

The ultimate ROI of our plan will be understood in the less tangible or unquantifiable ease of scholars' work and the continued advancement of human knowledge. We will nonetheless need to create a detailed business plan that delineates the costs of implementing the plan. Since the

¹¹ Deanna Marcum, "Due Diligence and Stewardship in a Time of Change and Uncertainty." New York, NY: Ithaka S+R. Available at <http://dx.doi.org/10.18665/sr.278232>.

actual work of developing a national plan will rest with libraries, library consortia, national library organizations, and funders should develop a business plan that accounts for the costs of services and logistics, the costs incurred for digital infrastructure installation and maintenance, and the costs of re-use, construction, and maintenance of physical repositories, along with an analysis of return on investment (ROI) for higher education institutions, with the expected cost recovery over the years calculated to recoup the initial outlay. The ROI will likely include cost avoidance relating to curtailing local storage sites and library-archival building additions. Cost avoidance savings are less common in university budgeting, and will probably require more itemized descriptions of accounting practices and useful formulas.

E. Planning for a National System for Preservation of Printed Books: Longer term

This paper tightly correlates with the recommendations made in *The Evidence in Hand: Report on The Task Force on the Artifact in Library Collections*¹² written by scholars, librarians, and other academic professionals: digitally connected, skillfully managed repositories, located in safe, distributed geographic regions, which house an acceptable redundancy of titles that are easily accessed, maintained over time, annually audited, and collectively evolved, would safeguard this extraordinary legacy, assuring that these collections are searchable and exponentially reusable. A digitization program, coupled with the appropriate copyright reform, could create electronic copies of the most requested volumes over time, rendering those volumes even more accessible and reusable.

If successful, this project will have a thoroughly researched, adequately funded, and rigorously governed beginning. It will not, if successful, have an endpoint in sight, its middle a long, thoughtful iteration of decisions that contribute to its evolution, strengthening the provision of services and resources. Like any well-designed project, many of its benefits will accrue to future generations, less so to those who lay its foundation. How it will be used over time is, if designed flexibly and intuitively, unknowable; the hope is that users would be astonished at this system's utility and integrity as a public good decades hence.

For the capacity to undertake this, scholars, librarians, and university administrators will need to both function exactly within the context of our era, and be able to frequently step out of their skin. Over time, the national collection management system for the print book may more

¹² Council on Library and Information Resources, 2001.

accurately be understood as a process, a pivotal bridge that allows for the transposition of a significant and invaluable collection of our cultural heritage into emerging and future manner of human expression: a process vital to a more open, fluid, and generous organization of knowledge.

Suggested Next Steps

We will take on the following responsibilities to broaden the conversation we have started in this paper and to further test its assumptions and recommendations:

- December 2016: propagate a working draft of the paper by posting it to Humanities Commons and simultaneously to mailing lists and committees of the sponsoring organizations as well as ALA and other relevant groups and organizations;
- January 2017: hold session at MLA Convention;
- First quarter 2017: assuming the ongoing availability of the communication channels used above, develop plan to provide routine updates and other kinds of communications to the constituencies touched by this proposal:
- First quarter 2017: once the paper has been released, initiate a process for gathering and incorporating reactions from mailing lists and meetings into further iterations of the plan outlined in the paper;
- January 2017–May 2017: plan and subsequently convene meetings and sessions at conferences of the sponsoring organizations and in other professional venues of interested parties to discuss these assumptions and recommendations; meetings and sessions at conferences and discussions for executive bodies of organization to occur throughout 2017;
- January 2017–July 2017: based on the assumptions and plan described in the paper, and on early responses to it, outline the components of a rigorous analysis of the costs and benefits of the proposed strategy to scholarly productivity, access, sustainability, and discovery; prepare a detailed description of how the higher education community can create, govern, and finance a productive future for the print record;
- January–December 2017; early in the year and with the leadership of ARL and CLIR, begin to discuss the paper with potential funders; based on the analysis described above, develop proposals for funding planning grants and research to test assumptions and provide more rigorous assessment and budget models of the proposed plan.