Together we are Stronger: A Cooperative Approach to Managing Print Collections

Matthew Revitt  
Program Manager, Maine Shared Collections Strategy, University of Maine, Orono, ME, USA  
matthew.revitt@maine.edu

Clem Guthro  
Director, Colby College Libraries & Project Investigator, Maine Shared Collections Strategy, Waterville, ME, USA  
cpguthro@colby.edu

Copyright © 2013 by Matthew Revitt & Clem Guthro. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License: http://creativecommons.org/licenses/by/3.0/

Abstract:

Maine Shared Collections Strategy (MSCS) is a collaborative library project seeking to create a model for the long-term preservation and management of legacy print collections. Amongst the many challenges facing libraries is the increasing cost of housing legacy print collections together with the lack of funding to build new stacks. Users expect libraries to devote increased room to social space for more active use, but also expect that they will still have the same access to information resources. Libraries also feel pressure to responsibly steward sizable, historic print collections. There are also growing opportunities to provide E-book-On-Demand and Print-On-Demand services via large scale e-book collections such as the HathiTrust, which allow libraries to rethink the management and delivery of their collections. These tasks exceed the capacity of any single library or organization. The MSCS multi-type library partners have taken a cooperative approach to facing the challenges and opportunities of managing legacy print collections. This paper will focus on the cooperation required in:

- Extracting and using data for collections analysis. Which data are available and useful and the challenges of comparing multi-type libraries, particularly when attempting subject analysis.
- Establishing the selection criteria for print titles to be retained, managed, and preserved, as well as titles that could be de-accessioned due to availability and/or preservation in a large-scale digital collection.
- Displaying shared retention decisions in shared catalogs.
- Implementing a centralized, shared solution to EOD and POD, incorporating large-scale digital collections (HathiTrust and Internet Archive).
- Developing a business model and Memorandum Of Understanding for sharing the responsibility long-term stewardship of print collections.
**Keywords:** Collection Development; Cooperation; Cooperative Collection Development; Maine Shared Collections Strategy; Managing Print; Shared Collections; Shared Print

---

**Introduction**

In the U.S. state of Maine, we have long believed that the future of libraries lies in a cooperative approach to library collections and services. In 2010 the two largest public libraries, Portland and Bangor; the Maine State Library; the two largest public universities, University of Maine and University of Southern Maine; three private colleges, Colby, Bates and Bowdoin; and the statewide consortium Maine InfoNet formed the Maine Shared Collections Strategy (MSCS) to create a strategy for the long-term preservation and management of legacy print collections. MSCS is supported by a US Government Institute of Museum and Library Services National Leadership Demonstration Grant.

MSCS aims to provide an example of statewide, multi-type collaboration that can serve as a model for others wishing to collaboratively manage their print collections. MSCS is developing a state-wide print collection management strategy and Print-On-Demand (POD) and E-book-On-Demand (EOD) service which includes data-driven collaborative models for collection management, de-accessioning, and collection storage; agreements regarding ownership and management of joint collections; protocols for requesting stored materials; protocols and policies for requesting and delivering POD services; and a cost model for delivering POD and EOD.

MSCS is unique in its approach to shared print with an emphasis on retention and preservation rather than weeding, a primary focus on monographs rather than journals, collaboration between public and academic libraries, and utilization of large-scale digital collections such as the HathiTrust in shared collections analysis. This paper will explore the drivers and growth of shared print initiatives, why shared print is a good fit for Maine, MSCS project objectives, and the activities undertaken to meet these objectives.

**Drivers & growth of shared print initiatives**

Print collections represent an essential component of the cultural, scientific, and scholarly record and remain a vital asset for libraries. No library has the resources to manage its own print collections indefinitely without regard for shelf space or long-term preservation, while at the same time accounting for the rise of digital collections that may supplant or augment portions of a library’s print collection.

In the last ten years, a number of shared print projects and forums have been established. Most projects have approached shared print from the point of view of a university library system, a collaborative of academic libraries, or independent groups that focus on single formats (e.g. print journals). Some major shared print projects are:

- **Western Regional Storage Trust (WEST):** a distributed print journal repository serving more than 100 research libraries, college and university libraries, and two library consortia in the western U.S.
- **UK Research Reserve (UKRR):** a partnership between the British Library and 29 UK higher education institutions to provide continued access to low use print journals.
- **Copac Collections Management (CCM):** a project with the Research Libraries UK and the White Rose Consortium to develop and test a web-based collections analysis tool.
• The Five Colleges consortium in Massachusetts (Amherst, Smith, Mount Holyoke, Hampshire, and University of Massachusetts): a shared repository for printed materials with ownership of volumes transferred to Five Colleges, Inc.
• The Center for Research Libraries (CRL) established the Print Archives Network (PAN) forum to shape a national approach to long-term preservation of and access to print collections. The most recent PAN meeting held at American Library Association Midwinter Conference highlighted the growing number of shared print projects being established and the variance in approaches.

A number of major studies and pilot projects aimed at understanding the issues around shared print management and digital collections have been completed by OCLC, the Council of Library Information Resources (CLIR), Ithaka, and the Center for Research Libraries (CRL). Reports of these initiatives shed light on the issues surrounding print management and have been used to shape best practice. Payne (2007) in a study of library storage facilities identified key trends, issues, and recommendations for future management. Malpas (2009) reviewed policies for existing shared print collections and identified policy elements necessary for shared print collections. Lavoie, Malpas and Shipengrover (2012) considered systems of consolidated regional collections, rather than local collections.

The OCLC Print Archives Disclosure Pilot Final Report (2012) details the findings of the OCLC Print Archives Pilot Project which explored ways in which libraries could use OCLC features and services to disclose retention commitments and support resource sharing for shared print resources. Out of the pilot project has grown the OCLC’s Shared Print Management Program which seeks to help libraries register shared print collections holdings in WorldCat.

Why here in Maine?

The fact that Maine is rural and economically challenged has contributed to its progressive approach to library collaboration. MSCS partners are members of Maine InfoNet, one of the few state-wide multi-type library consortia in the US. Maine InfoNet provides a robust digital technology infrastructure, a shared state-wide library catalog (MaineCat), and services for electronic resources that deliver collections in physical and electronic formats and link the libraries organizationally. The partners share costs of a state-wide delivery system that connects the majority of libraries in the state.

A history of collaboration across all library types dating back almost 100 years, a robust service delivery model, and the fact that the collections of the MSCS partners comprise more than three quarters of the print collections in Maine makes Maine an ideal location for a shared print initiative. MSCS is addressing demands for space and service expansion by sharing collection storage expenses while at the same time integrating collaborative practices in Maine into larger national efforts.

Maine Shared Collections Strategy project management

The day-to-day coordination of MSCS is directed by the MSCS Project Team consisting of a full-time Program Manager, Technology Director, and three Co-Principal Investigators. The Program Manager is based at the University of Maine and administers the grant on behalf of the partner libraries. The Project Team meets monthly and publishes its progress on the MSCS website (http://www.maineinfonet.net/mscs/progress/).

MSCS includes stakeholder groups with each partner library represented in every group. The Director’s Council provides leadership with respect to governance, strategic direction
and the resolution of issues that are referred to it by the Project Team. The Collection Management Committee assists in the development of collections analysis, retention and preservation criteria. The Technical Services Committee consults on the cleanup and extraction of library catalog data, testing the use of MARC Subfield 583 to document retention and preservation decisions in our library catalogs, and implementation of an OCLC Shared Print Symbol to expose those decisions in OCLC WorldCat.

The MSCS Project Team has also contracted with an independent Systems Librarian to extract data from library systems, plan for the integration of metadata for large-scale digital collections into the libraries’ catalogs, and for the implementation of the Print-On-Demand and E-book-On-Demand functionality.

MSCS is working with an Advisory Board of nationally known experts with recognized expertise in shared collection development and print storage and preservation. The Board provides evaluative oversight and external review and feedback of the project, advises on best practices and standards, and offers guidance in the creation of an infrastructure that ensures the long term viability of the collaborative initiative, and will ensure that MSCS interfaces with other related national initiatives (for more information regarding the different MSCS groups see http://www.maineinfonet.net/mscs/about/people/).

**Objective #1: Develop a strategy for a state-wide, multi-type library program for managing, storing and preserving print collections among public and private institutions to achieve greater efficiencies and extend the power of every dollar invested in collections and library facilities.**

**Collection and use analysis of print collection**

The MSCS partners want to answer the following high-level questions:

- What monographs should be designated for long-term retention?
- What is an equitable and/or common-sense distribution of retention responsibilities?
- What monographs are candidates for incorporating into POD/EOD services (via local or shared catalogs) by virtue of HathiTrust or Internet Archive public domain material?
- What monograph copies (by library) could optionally be deselected, once retention decisions have been finalized?

In order to provide the in-depth collection analysis necessary for the project, the following data elements were identified as being necessary:

- Item record number
- Created date
- Barcode
- Itype (value in the item that defines how it circulates)
- Volume and copy
- Item call number
- Location
- Total checkout and total renewal
- Year to date circulation
- Last year circulation
- Last checkin
- Out date
- Last out date
An initial analysis of the MARC records from MSCS partner libraries revealed issues of data quality and accuracy. In order to accurately compare holding and circulation data with OCLC WorldCat, an OCLC reclamation project was needed to clean up the data and verify accuracy of holdings in OCLC. The reclamation, originally projected to take three months, took 12 months to complete and required significant work on part of each library's technical services staff.

Circulation data comparisons brought to light a number of data issues. Not all libraries have the same number of years of circulation data. Academic libraries have reserve and in-house use statistics, but there is no comparable data for public libraries. The public libraries bill items more quickly and requested that recently billed materials be included in the analysis, while the academic libraries excluded these items. Treatment of such data variations has yet to be addressed.

The academic partner libraries classify their library collections using Library of Congress Classification Scheme (LC) and the public libraries use Dewey Decimal Classification (DDC). To facilitate the collections analysis process, in particular looking at subject strengths, the libraries requested that any data reports provided to them are in the classification they use. However, there also needs to be an augmented version of data reports in order to make group comparisons across both LC & DDC. It was with this type of issue in mind that MSCS sought the outside technical support of a library collections analysis tool that provides data manipulation and reporting functionality. Technical support is especially important due to the scale of the analysis, involving approximately 3 million items across the partner libraries.

In November 2011, a group subscription to OCLC WorldCat Collection Analysis (WCA) was chosen, but WCA proved unable to facilitate the in-depth analysis of the participating library collections. Particular issues were the inability to: provide both a combined group and individual analysis; import item and/or title level circulation data; and manipulate and report out large batch files of title and other data for the eight libraries individually and as a group.

While using WCA, the Project Team also began discussions with OCLC regarding the development of OCLC's new analytics product. The persistent delays in the development of the new analytics product led the Project Team to investigate other collections analysis tools and services, including: Center for Research Libraries (CRL) Print Archives Preservation Registry, Bowker Book Analysis System (BBAS), Ulrich’s Serials Analysis System, Portico book and journal comparisons, CUFTS Resource Comparison and Capstone’s CollectionWiz. These products compare a library’s holdings to their proprietary lists of published materials, for the most part journals, which is not the primary focus of MSCS. None of these products proved feasible because they do not have the ability to ingest massive loads of data from multiple libraries and output reports and data for use in a group and cross-library analysis and batch loading retention statements in multiple systems.

MSCS subsequently contracted with Sustainable Collection Services (SCS) to provide collections analysis services. SCS are unique in their ability to provide tailored reports combining local circulation and item data with OCLC WorldCat library catalog holdings and HathiTrust Digital Library and Internet Archive holdings. MSCS provided SCS with 2.9 million bibliographic records with associated circulation and holdings data from the partner library catalogs. Using this record set SCS:
- Filtered out-of-scope bibliographic records including: Government Documents, non-monographic material, Non-language material, non-print resources, records missing OCLC numbers, Bibliographic/author mismatches with OCLC and multiple OCLC numbers per record
- Eliminated duplicate bibliographic records
- Normalized call numbers
- Eliminated trailing spaces in control numbers
- Validated OCLC numbers
- Matched bibliographic records on OCLC numbers (with title string check)
- LCCN/title-string lookups for records lacking OCLC numbers
- Identified and accommodated unusual implementations of MARC
- Mapped item-level data and interpret codes
- Provided Dewey Decimal numbers for records that lacked them

Because of the OCLC reclamation project the data set was very clean and SCS found fewer anomalies than normal. SCS matched titles to external data sources:

- OCLC WorldCat including both US and State Holdings
- HathiTrust Public Domain and In-Copyright items
- Internet Archive

SCS provided partner libraries with collection summary reports in Library of Congress and Dewey Decimal Classification and an augmented version in both schemes. They also included summary comparisons against HathiTrust and Internet Archive. The use of Dewey and Internet Archive were new challenges for SCS. SCS devised a mapping scheme to assign LC and Dewey matches where needed. Some additional refinement of the Internet Archive data and subsequent comparisons is underway.

Collection management, stewardship and preservation model

The majority of shared print projects to date have focused on de-accessioning materials. In order to gain political approval from library stakeholders and avoid the negative publicity seen in high profile weeding projects (Demas, 2013) MSCS emphasized the role of libraries as stewards of legacy print collections and focused on retention and preservation of material to ensure it remains accessible to library patrons.

Through analysis of the collections data, partners will identify:

- Number of copies of a particular work owned by partner libraries
- Number of circulating copies
- Number of times a title circulated and date of last circulation
- Number of titles/copies uniquely held in the group/Maine/OCLC WorldCat (U.S. only)
- Subject strengths across the group and the state
- Titles represented in HathiTrust and Internet Archive
- Overlap between general collections and special collections

MSCS project groups are developing criteria for print titles that should be retained, managed and preserved, as well as titles where the print title may be de-accessioned and replaced with the digital copy from one of the large-scale digital collections. These criteria will be based on risk management principles and use models that help libraries determine the optimal number of copies needed at the network/regional level. Risk management principles developed within the larger library community will assist the partner libraries to position their
efforts within the broader regional, national, and international efforts and help create a
decision-making framework based on environmental criteria (duplication rates, publication
date, electronic availability) as well as use.

MSCS contracted with SCS for consulting services to help interpret the data reports and
refine retention criteria. MSCS requested some guided analysis because of the scale of the
data and time constraints of a three-year grant period. With the help of SCS, the MSCS
partners decided to begin with titles held by only 1-2 partners, for which there is a higher
than expected proportion. The following criteria for making decisions on these titles were
developed:

- Analyze and take action only on pre-2003 copies
- Retain the copies if any circulation or internal use
- Retain material that falls into local protection categories (Specific Maine items) even
  if no circulation
- Retain Special Collections/Archives copies even if no circulation
- Retain materials on course reserves even if no circulation
- Retain unique in OCLC (only 0-9 copies in OCLC) even if no circulation
- Compare remaining 0 circulation copies with both HathiTrust and Internet Archive

Beyond looking at items with 0 circulations, MSCS have yet to decide how circulation rates
will effect retention decisions. The average circulation rates for items were higher because of
the public libraries. This will have to be factored in when looking at thresholds for circulation
rates.

Retention commitments will be allocated to the partner libraries based on a number of
factors including available storage space, subject strengths, and existing preferential loan
periods. Colby has built a new storage facility which may allow them to keep a larger share
of materials. Colby, Bates and Bowdoin have common preferential loan periods so they may
choose to retain an additional copy if these loan periods cannot be guaranteed by other
partners. The public libraries are more likely to build on their strength in fiction and may
choose to ingest fiction from other partners.

MSCS partner libraries will disclose their retention commitments in both local and union
catalogs and in OCLC WorldCat. A result of MSCS decisions will be that libraries both
MSCS partner and non-partner can weed their collections safe in the knowledge that the
material will be preserved within Maine and remain accessible to their patrons.

MSCS will follow the recommendations of the OCLC Print Archives Disclosure Pilot Final
Report (2012) and define separate OCLC Institution Symbols to identify items selected for
retention. Recommendations include holdings-level data in MARC Holdings records (OCLC
Local Holdings Records, LHRs) and using the MARC 583 Action Note to describe specific
action(s) for each set of holdings. MSCS has tested the use of the 583 field in local system’s
holdings records and identified issues with the display and transfer of the data from local
catalogs to the central union catalog, and also variances in how 583 will display depending
on the discovery layer used. MSCS is currently exploring various options with OCLC and our
system vendor, Innovative Interfaces, Inc. MSCS hopes to have clear procedures in place by
the time retention decisions are ready for implementation. MSCS has started to test the
Interlibrary Loan implications of the Shared Print Symbol.

In contrast to many shared print projects which have implemented centralized storage
facilities for their shared print items, MSCS is developing a distributed management model
where ownership and storage of the retained material will remain with the partner libraries.
Shared print items will not have different Interlibrary Loan workflows and therefore remain requestable in both state and national catalogs.

**Objective #2: Expand access to existing digital book collections by developing Print-On-Demand (POD) and E-book-On-Demand (EOD) services to support long-term management of a shared print collection, and the integration of digital resources with print collections.**

**Print/digital management model**

The growing availability of e-books in large-scale digital collections like the HathiTrust and Internet Archive provides an opportunity for rethinking management of print collections.

MSCS will develop criteria for when and how a digital copy of a title would substitute for retaining print copies or for when the delivery format of choice would be digital even when print copies were extant. Some MSCS partners have expressed a willingness to rely on digital copies as surrogates. However for some humanistic disciplines, faculty at the academic libraries are likely to prefer a physical copy.

The MSCS partner libraries are actively investigating individual and/or consortial HathiTrust membership. HathiTrust membership will allow patrons to view and download electronic copies of public domain materials. It will also allow member libraries to provide a digital copy of a public domain book via Interlibrary Loan. HathiTrust’s requirement of Shibboleth for authentication makes it impossible for public library participation and difficult for academics. Shibboleth implementation by partner academic libraries needs to be completed prior to becoming members. ([http://www.hathitrust.org/shibboleth](http://www.hathitrust.org/shibboleth)).

SCS compared MSCS partner print holdings (including item-level data) against HathiTrust and Internet Archives to see where there is overlap. There is surprisingly low 6% overlap for all public domain titles across all partner collections, but a 37% overlap for in-copyright titles.

**Service delivery model including POD and EOD**

MSCS partner libraries plan to implement a service delivery model for Print On-Demand (POD) and E-book-On-Demand (EOD) titles to complement the physical book delivery service. The ability to provide POD and EOD are based on having HathiTrust and Internet Archive records discoverable in the MaineCat union catalog.

MSCS have investigated a number of different options for implementing POD including the Espresso Book machine and commercial POD vendors such as Booksurge / CreateSpace (Amazon), TextStream (Baker and Taylor) and Lightning Source (Ingram). MSCS intends to test the demand for POD using the University of Maine’s Printing Services. The MSCS Systems Librarian is investigating different options for discovery of HathiTrust records in MaineCat as well as methods for providing EOD and POD links.

**Objective #3: Formalize organizational agreements, establish a budget, and develop policies essential to the management of shared print and digital collections.**

**Sustainable business model**

At the end of the three year grant, Maine Shared Collections Strategy must move from a project to a long-term sustainable model and will operate under the name of Maine Shared
MSCS is currently working on a number of areas that will be essential for long-term sustainability. MSCS developed a draft Memorandum of Understanding (MOU) that received the preliminary approval of the Directors Council. It will need to be vetted and signed by campus administrators and/or Boards of Trustees. The MOU specifies that an Executive Committee will provide overall governance for the Cooperative, because of the distributed nature of the shared print collections, daily management of retained items remain the responsibility of the retaining library. The five-member Executive Committee will have mandatory representation from the private non-profit academic libraries, the public university libraries, the public libraries, and state library in order to maintain a balance representation of the major constituencies. A Collections and Operations Committee will oversee issues related to the selection of materials for retention, as well as issues related to holdings disclosure and access/delivery.

Maine InfoNet will serve as the administrative host for the Cooperative. Members agree to be responsible for all of the costs and expenses associated with maintaining their own retained materials, disclosing holdings to MaineCat and WorldCat, lending materials to other libraries, and deselecting materials from their own collections.

Academic members, in whole or in part, will join HathiTrust and pay the associated yearly membership fees. Academic members who become members of HathiTrust will also cover the staffing costs of EOD requests for HathiTrust titles. Print on Demand requests will be routed through the University of Maine print shop and costs will be borne by the requestor or the requestor's library.

The MOU specifies that partner libraries are making a 15-year commitment to retain titles with a 5-year review of the agreement to ensure that all partners are on board and any issues are addressed. By the end of the grant period a mechanism for adding other partners to the Cooperative will be delineated.

**Conclusion**

MSCS is important in the emerging international effort to manage legacy print collections. MSCS significantly advances the understanding of how to manage the relationship among large scale digital collections and established print collections.

**References**


