



## Maine Shared Collections Strategy (MSCS) Advisory Board Update Report

---

### **OCLC WorldCat Collections Analysis tool**

OCLC's WorldCat Collections Analysis Tool proved unable to facilitate the in-depth analysis of the participating MSCS 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. As a result of repeated delays in the development of the next-generation WCA as of October 2012, the Project Team made the decision not to renew the MSCS group subscription to WCA and began investigating other collections analysis tools and services not considered initially.

In March 2013, Kathryn Harnish offered MSCS a refund for its subscription to WCA based on feedback she received at ALA Midwinter in particular regarding the reporting functionality of the current product and delays in the release of the analytics product. Instead of a refund, MSCS have agreed with OCLC to receive complimentary deferred access to the analytics product when it is available for group comparisons in late 2013/early 2014. Access to the new product may allow MSCS to report on comparisons between the services of WCA and SCS.

### **Contracting with Sustainable Collection Services**

After researching available collections analysis tools the MSCS Project Team subsequently agreed that at this time only Sustainable Collections Services (SCS) could provide the type of collections analysis services required; of particular importance is SCS's consulting support and data reporting which has allowed MSCS to make real progress on analysis and decision-making. SCS will only be assisting MSCS with the analysis of monographs; MSCS is still to address the analysis of journals.

In February 2013, SCS and MSCS agreed to work together and have worked hard to produce a quick turnaround on the preparation and processing of data to be analyzed.

### **Data pull & cleaning**

In late February MSCS provided SCS with 2.9 million bibliographic records with associated circulation and holdings data from the partner library catalogs. 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
- Reserve notes

- Internal use count
- Icode2 (Contributed to union catalog)
- Circulation Status

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

SCS matched titles to external data sources:

- OCLC WorldCat including both US and State Holdings
- HathiTrust Public Domain and In-Copyright items
- Internet Archive, which was a first for SCS

Because of the OCLC reclamation project, the data set was very clean and SCS found fewer anomalies than normal. Also, because this was the second time MSCS had performed a data pull (first time for WCA) it was a relatively smooth process. However, one issue was the late inclusion of Bangor Theological Seminary (BTS), an entire additional library collection, which is currently being subsumed into Colby's collection. BTS were not OCLC members and their data had not been through the reclamation process, which meant it was non-corrected and had missing data. Another issue is that some MSCS libraries (particularly Bates) are doing withdrawals, so the data set being used by SCS, is static, which is never representative of reality.

## SCS data reports

The results of SCS's data compiling, manipulation, and cleaning were presented in the collection summaries provided to MSCS in March. The summaries are categorical overviews of the group data set and have been used to guide retention scenario development. SCS also provided a number of graphs and charts which allowed MSCS to see and then focus on a smaller, more manageable subset of data.

The collection summary reports were in both Library of Congress and Dewey Decimal Classification, with an augmented version in both schemes. The use of both classification schemes was a new challenge for SCS. SCS devised a mapping scheme to assign broad LC and Dewey matches where needed.

## Retention scenario development

MSCS partners have decided to begin analysis and decision-making with the group of titles held by only 1-2 partners. This universe consists of just under 1.5 million title-holdings (approximately 1.6 million items) and constitutes 50% of all title-holdings in the MSCS data-set, a much 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

SCS ran a scenario based on these criteria and produced corresponding retention counts for each library. The set of resulting "Commitment to Retain" titles contained just over a million title-holdings or 73% of the eligible universe of titles. This left just under 400,000 title-holdings, or 27%, that "Needs Further Examination" for those titles that are available electronically in the HathiTrust and Internet Archive. MSCS are still evaluating whether MSCS libraries are willing to rely on digital surrogates if the item has zero circulations.

The remaining 50% where items are held by 3 or more libraries is where more in-depth collection analysis work is required. Beyond looking at items with zero circulations, MSCS have yet to decide how circulation rates will affect 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. Other factors being considered include: 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.

### Disclosing retention commitments

Having agreed retention commitments for just over one million items, MSCS is currently working on the display of those commitments in both local and union catalogs, in the MARC Subfield 583, and in OCLC WorldCat using both 583 and the OCLC Shared Print Symbol. SCS will provide MSCS spreadsheets containing lists of those items each library are committing to retain.

For the local INN-Reach systems the SCS lists will be used and retention statements will be added using global update. As a result of variances in how 583 is displayed in local catalogs, MSCS have decided to allow each member institution/group/consortium decide the how (webpub.def or OPAC message) and where to display the retention commitments themselves as long as they use common language provided by the MSCS Project Team.

MSCS have encountered issues with the display and transfer of 583 to the central union catalog, MaineCat. III have informed MSCS that an enhancement and/or individual programming is required to fix the issue. MSCS have been discussing this work with III, but have not made little progress. As a work-around Maine InfoNet staff and the MSCS System Librarian Sara Amato have managed in MaineCat to use the OCLC WorldCat API and JavaScript to perform a check of OCLC and display when an item is in shared print, which seems to solve the issue of non-transfer from local catalogs. With the API MSCS is achieving a big portion of what is required -- basically retained items will have a note saying that they are retained and by whom, with a link to MSCS retention policy information. But the API will not allow MSCS to see the retention on a brief results display list, or get the retention commitment end date -- the API will only grab and display the information on the full results page for a single item.

The MSCS Systems Librarian is currently working with OCLC on the batch loading process for Local Holding Records into OCLC for the items designated as shared print. As a result of MSCS Library Director opposition to the ILL fees associated with using the OCLC Shared Print Symbol in both ILLiad and WorldCat Resource Sharing, MSCS have decided that until a more acceptable model can be developed to use two symbols on the records in OCLC, both the main symbol which will remain requestable, and the Shared Print symbol which will be a non-supplier. MSCS have already been testing the ILL implications of the shared print symbol using Bowdoin (BBHSP) and Bangor (BYNSP), and have now ordered shared print symbols for the remaining MSCS partner libraries.

### HathiTrust membership investigations

A MSCS goal is to deliver a service model for both Print-On-Demand and E-book-On-Demand. The MSCS partner libraries are actively investigating individual and/or consortial HathiTrust membership. However, 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.

MSCS is currently exploring loading 1.6 million MARC records for the HathiTrust public domain titles into MaineCat for the purposes of making the downloadable electronic copy requestable. Loading records into local catalogs was briefly considered but discarded as an option due to the fact that a) these titles are usually accessible via a discovery layer such as Summons and b) the quality of the records is highly variable. In order to load the records into MaineCat a new agency will be created in SOLAR specifically for the purpose of centrally loading electronic resources. This allows the records to flow into MaineCat while alleviating any single library from the burden of maintaining the records, which vary in quality, in their own systems. Responsibility for HathiTrust record updates in SOLAR has yet to be assigned or distributed.

To date a small number of titles have been test loaded. MSCS are currently exploring the requesting functions and also a mechanism for keeping these records up to date.

## **Presentations/Outreach**

In 2013 MSCS Project Team members have presented at the IMLS WebWise Conference (Baltimore, MD March 6-8), Maine Larger Libraries event (Portland, ME May 5), and Timberline Conference (Mt. Hood, OR May 18-21).

MSCS Project Team members will be presenting at the: ALCTS "Shared Print Monographs" pre-conference event and the Print Archive Network forum at the ALA 2013 Conference in Chicago, June 27th & 28th respectively, IFLA Conference in Singapore on August 19<sup>th</sup>, and New England Library Association Conference in Portland, ME on October 21st.

Library Journal published an online article on MSCS titled "Major Maine Libraries, Public and Academic, Collaborate on Print Archiving Project" which resulted in a number of references to MSCS in library and partner institution news sources. MSCS received a further mention in the April 4th Library Journal editorial piece titled "Kudos for Print Archiving". MSCS Program Manager Matthew Revitt will be a panelist on the Library Journal's June 6th webcast "Data-Driven Libraries Part 1: Analyzing Data to Manage Print Collections".

MSCS also featured in the ACRL environmental scan section "Radical Collaboration in Large Regional Print Repositories" and in a sidebar article in the Maine Policy Review May edition on Maine libraries.

The MSCS website and Twitter feed are continually updated with news, project updates, meeting summaries, and reports.

## **Business model**

In December 2012, the MSCS Directors Council requested that the MOU language concerning the Executive Committee should be changed because in their opinion it was too vague and could potentially result in smaller libraries being overrepresented in the Maine Shared Collections Cooperative. The language has since been revised by Project PI Clem Guthro and included in a revised draft which will be presented to the Directors at their May 23rd meeting. Clem will present the MOU at the Maine InfoNet Board Retreat in June. If it is approved (or some modified version) it will need to go back to the Director's Council. Once approved, the MOU will be forwarded to be signed by an institutional representative (rather than the library's director), for example a Board of Trustees or Provost.

## **Budget**

MSCS submitted a Budget Change Justification Report to IMLS in December 2012 for those budget areas which met IMLS requirements for approved changes. These were using partner contributions originally intended for purchasing an Espresso Book Machine to pay for instead collections analysis tools and that one of the Advisory Board could not accept stipends. The report was submitted directly to MSCS's IMLS Officer, Chuck Thomas who approved it with the understanding that partner library contributions would not be less than originally committed.

The increases in the in-kind contributions of staff's salaries & benefits for MSCS partner library representatives meant MSCS will actually be over its originally matching amount.

Since submitting the Budget Change Justification Report, MSCS have had their requests to use grant funds to pay for foreign travel to IFLA approved by IMLS. MSCS originally intended to contract with a programmer to build the Collection Analysis System. The Project Team decided in Year 1 that this position was not required because the responsibilities were being covered by a combination of the work of the MSCS System Librarian and commercial collections analysis tools. MSCS received approval from IMLS to re-allocate programmer funds to partly fund the services of SCS and costs of bringing Jeremy York from the HathiTrust to Maine.