

Document Status:

This document represents a modification of the agreement signed August 31, 2011 for the MCLS pilot project. The intent of this document is to specify the terms under which two new libraries, Oakland University and Ferris State, have been added to the pilot project group, and to incorporate those 'mid-cycle' libraries into the overall agreement with MCLS. Terms for mid-cycle libraries are outlined on pages 7 and 9-10 below.

Overview

The Midwest Collaborative for Library Services (MCLS) seeks to devise a collaborative approach to shared print collections in Michigan. By first assuring that an agreed minimum number of copies of low-use titles are held collectively, individual libraries may responsibly downsize their local print collections. MCLS will bring together a group of Michigan academic libraries in a pilot project to identify titles that are commonly-held but little-used. Participating libraries will use services and tools developed by Sustainable Collection Services (SCS) to identify such titles in their respective individual collections, and to compare results across the group. (As the analytical work progresses, it may be necessary to form multiple small groups, with the intention of uniting them into a larger comprehensive group when the initial data extracts are renewed in August/September 2013. Other approaches may also be employed to reach this end.)

Based on this information, the MCLS pilot group will develop criteria for retention commitments. The group will consider opportunities and options for shared print storage and service using a distributed model. The intent is to create a mechanism that supports a regional approach to the managed drawdown of redundant print collections, while assuring that sufficient copies of low-use titles are retained in Michigan. The comparative data produced by SCS will enable MCLS libraries to quantify overlap, identify unique titles, and to preview the size and nature of the collaborative opportunity. The data will also offer individual participants the information and latitude to act independently if a collaborative solution is not pursued. While the pilot project will initially be limited to the seven pilot libraries, provision will be made for adding new libraries in a staged process.

MCLS will bring its convening, communication, and coordination capacities to this process, and will identify participants for the pilot project. SCS will provide data-driven deselection tools, interpretation and analysis of results, and facilitation of discussions around initial findings. Most importantly, the SCS outputs will support data-driven consideration of collaborative agreements and actions, and a way to estimate their potential benefit to each participant.

Current Situation

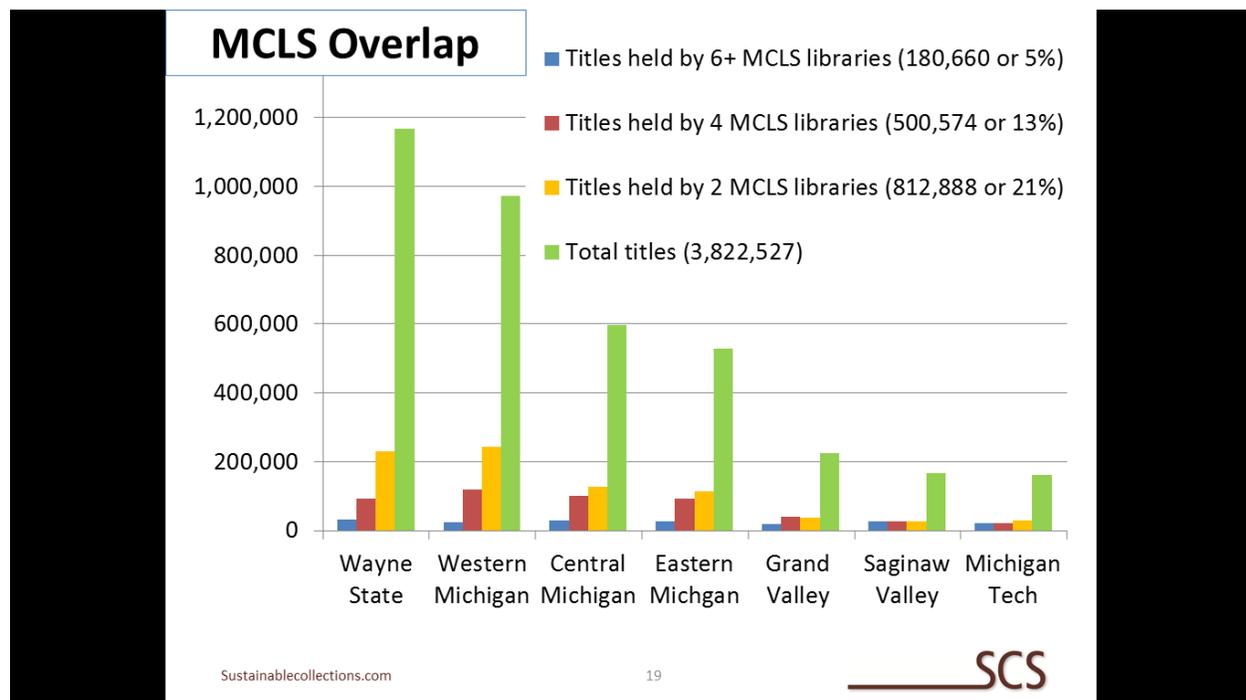
Michigan libraries have long recognized the advantages of working collaboratively, as MeLCat and other initiatives attest. It seems clear to all that a shared regional approach to print retention and storage of low-use monographs would be valuable. It would reduce overall storage needs, and would enable archiving responsibility to be distributed among participants. At minimum, it would assure that at least one copy of all titles currently held by participating libraries would remain securely and permanently archived. It would also enable little-used duplicative copies to be withdrawn, freeing significant space

and avoiding additional building costs. By working together, the MCLS pilot libraries can pursue deselection more confidently, especially once an agreement for shared retention and associated services is in place.

The proposed pilot project, which involves seven libraries, is in effect an exploration. The intent is to gauge the scale of the opportunity for collection drawdown should participants share responsibility for retention of low-use print monographs. SCS will provide the comparative intelligence and data management tools. MCLS will provide communication, coordination, and consolidated billing/payment. Although the exact methods remain to be worked out, we believe there are two main components to this project.

- First, SCS will analyze collections in participating libraries individually, to identify the effect of specific deselection criteria independently of the group collection. This analysis will be based on use (as reflected in circulation data), holdings in other MCLS libraries, and holdings in US libraries generally. Other factors, such as the presence of a Hathi Trust digital version, or the appearance of titles on authoritative lists, will be considered as necessary.
- Second, SCS will aggregate and normalize bibliographic, circulation, and item data from all pilot project participants, and determine the degree of overlap across the collections. This will help to quantify the size and nature of opportunities for collaborative action. It is expected that this information will help inform shared retention agreements, and improve each library's understanding of what can be safely deselected and what must be retained.

Data analysis completed by SCS in October 2011 shows the extent of overlap across the pilot library collections, and begins to demonstrate the size of the shared print opportunity.



Scope

This pilot project and any related ongoing processes will focus on circulating monographs in print form. Print journals, government documents, audio-visual, microforms, and other non-book materials are out of scope. Special Collections and reference works are also out of scope.

SCS is responsible for analyzing and presenting the data needed to make decisions and formulate policies. Decisions, policies, project management, communications with staff and stakeholders and implementation are the responsibility of MCLS and the participating libraries as informed by the steering committee.

Approach

Given the exploratory nature of this project, some elements of the approach will be developed or adjusted as we proceed. But we anticipate the following major components:

Building the Data Environment (*Completed: September 2011*)

SCS will obtain bibliographic, item, and circulation data from each participating library. Since participating libraries use different library automation systems and may have differing data management practices, the particulars will vary. In each case, SCS will work with the individual library to define an appropriate set of data extracts, which will be sent to our FTP server. SCS will work with each library to understand local terminology and codes such as location, item type, etc.

To the degree possible, SCS will seek to limit the extract to circulating print monographs. Once SCS has retrieved the library's data from our FTP site, we will normalize the bibliographic data to assure comparability with OCLC and other MCLS library data sets. SCS will load and normalize each participant's data sequentially. In October 2011, SCS began producing a variety of reports. We estimate that this data environment can continue to be used for approximately two years, or July 2013. At that time, a full reload of bibliographic, circulation, and item data is recommended.

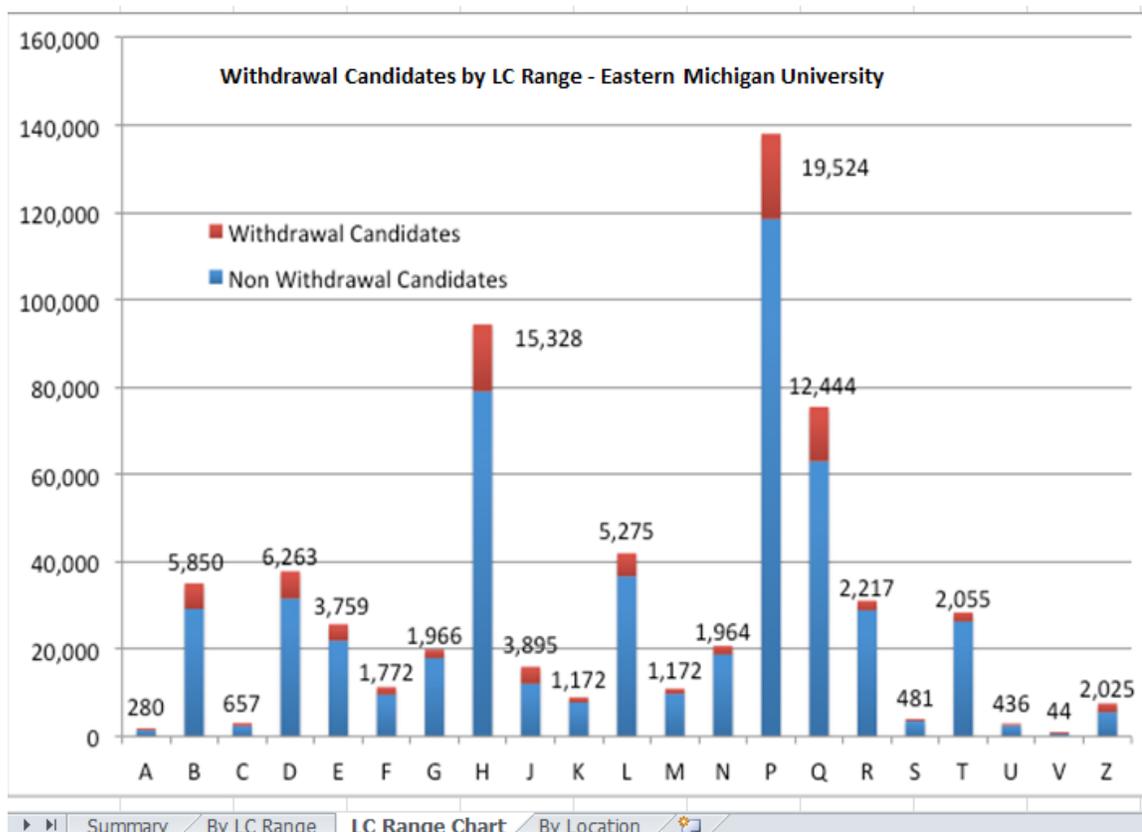
Analyzing the Aggregated Data (*Completed: November 2011*)

Once ingestion and normalization of all data sets has been completed, SCS will analyze overlap across participating library collections. During this stage we will also seek to understand the degree to which circulation data is comparable. Because there are no standards for circulation transactions, we expect significant variance in practice and in the date ranges for which use data is available. Some libraries, for instance, include in-house charges to bindery or acquisitions in their counts. Some include ILL or reserves activity, some do not. Some libraries will capture the last circulation date, some will have only total charges over time. SCS will identify and describe what the data will support for both overlap and circulation patterns.

SCS Collection Summaries for Individual Libraries (*Completed: October 2011*)

During this stage, SCS will also produce a Collection Summary for each individual library. The SCS Collection Summary provides a statistical overview of the library collection and its use. It also calculates the effect of library-defined circulation parameters and title protection rules. It provides information about the collection's overlap with target data sets and the extent to which withdrawal candidates are held in other libraries or in secure digital form. The Summary also captures the number of candidate titles by broad LC ranges and locations. A sample summary for an individual MCLS library follows:

	Title Count	Item Count	Percent of Item Records
All Records - Unfiltered	595,818	710,485	N/A
All Records - Filtered	563,999	635,668	100%
Counts for individual candidate lists			
Withdrawal Candidate List - Zero circulations and > 50 US Holdings and > 3 MCLS Holdings and Publication year < 2005 and Date added < 2005 and Not Reviewed in CHOICE	80,789	89,520	14%
Preservation Candidate List - < 5 US Holdings and no Hathi Match and no additional holdings in MCLS Pilot Group and Publication year < 2005	5,337	5,892	1%
Match rates for individual factors			
Circulation Counts			
Total Charges = 0	280,800	321,276	51%
Total Charges = 1 or less	378,298	426,172	67%
Year-to-Date Charges = 0	533,738	601,418	95%
WorldCat Counts			
> 100 holdings in USA	408,575	453,541	71%
> 50 holdings in USA	484,750	543,872	86%
> 20 holdings in USA	527,679	594,554	94%
> 10 holdings in Michigan	188,370	215,686	34%
> 5 holdings in Michigan	352,283	393,675	62%
< 10 holdings in USA	11,789	13,086	2%
< 5 holdings in USA	6,340	7,031	1%
> 5 holdings - MCLS Pilot Group	30,643	38,136	6%
> 3 holdings - MCLS Pilot Group	194,862	220,021	35%
> 1 holdings - MCLS Pilot Group	433,414	489,259	77%
Unique Titles in MCLS Pilot Group	130,584	146,408	23%
Date related counts			
Publication Year before 2005	502,128	571,994	90%
Publication Year before 2000	436,882	504,597	79%
Publication Year before 1990	314,770	375,507	59%
Added before 2005	480,007	531,140	84%
Electronic Surrogates			
Hathi Trust Public Domain Match	17,784	25,022	4%
Hathi Trust In Copyright Match	226,209	254,052	40%
Authoritative Title Lists			
Reviewed in CHOICE - match	65,227	67,897	11%



Please note that other data comparators, such as *Resources for College Libraries* and *CHOICE*, may be available to those libraries with active subscriptions to those resources.

During both the data aggregation and data analysis stages, SCS will communicate regularly with the individual libraries. Decisions will be needed regarding criteria for analysis. SCS will also provide regular status updates.

Presentation and Discussion of Preliminary Results (Completed: October 31, 2011)

Once individual Collection Summaries have been distributed and the aggregate data analyzed, SCS will visit MCLS in Lansing (or wherever designated) to present the results to date, answer questions, hear reactions, and to plan the next steps.

The individual Collection Summaries will be discussed, and modifications of those Summaries will be performed as necessary over the ensuing weeks. While a single trial 'Withdrawal Candidate List' (i.e., for one subject or one location) may be produced at this point, we do not expect to produce full Withdrawal Candidate lists until later in the project, i.e. after some decisions have been made about distributed archiving commitments.

SCS will also describe the characteristics of the aggregate MCLS data, and what we believe it will support in terms of comparison. Decisions may be needed from the MCLS group regarding the

criteria for an overall MCLS Collection Summary. Some method for allocating or claiming responsibility for specific segments of the shared MCLS collection will need to be devised.

SCS Collection Summary for MCLS Shared Collection (Expected Timeframe: October-December)

Based on decisions from the October meeting, SCS will create a Collection Summary for the group as a whole. Parameters will depend on the comparability of circulation data, the degree of collection overlap, and the publication dates and archiving thresholds (number of copies) agreed by the group. This report will be similar to the individual Collection Summaries but will also highlight shared copies, to help estimate the yield of various decisions. Some possible configurations include:

	Normalized bib records - print monographs only	Titles published prior to 2005	Pre-2005 titles with 0 recorded circs	Unique titles within this group of libraries	Titles held by two libraries	Titles held by three libraries	Titles held by four libraries	Titles held by five libraries	Titles held by six libraries
Wayne State	1,169,157	1,099,755	508,405	622,239	229,563	144,044	94,224	52,690	21,314
Western Michigan	971,995	842,941	347,542	333,475	243,804	175,812	120,583	66,794	25,710
Central Michigan	597,046	530,610	190,930	144,728	128,305	126,069	100,772	59,601	23,508
Eastern Michigan	529,502	467,682	231,046	125,401	113,717	111,083	94,200	57,609	22,480
GVSU	226,280	187,107	132,639	52,744	38,937	41,271	41,485	34,035	15,511
SVSU	167,707	151,881	61,359	40,481	27,690	27,166	26,656	24,060	16,696
Michigan Tech	161,640	150,642	76,188	43,979	30,872	26,280	22,654	19,004	13,059
Total	3,823,327	3,430,618	1,548,109	1,363,047	812,888	651,725	500,574	313,793	138,278

	Titles with 0 recorded circs	5 or fewer circs ever	Titles in Hathi Public Domain	Titles with >100 WC holdings in US	Titles with >50 WC holdings in US	Titles with < 5 WC holdings in US	Titles with 0 circs and titles with > 100 WC holdings in US
Pre-2005 titles w/6+ holdings	65,518	136,526	4,599	164,549	165,456	0	64,379
Pre-2005 titles w/5 holdings	113,019	240,885	6,617	289,161	290,259	3	110,949
Pre-2005 titles w/4 holdings	173,409	378,192	11,529	443,641	450,644	28	167,144
Pre-2005 titles w/3 holdings	225,149	486,836	17,930	523,563	566,418	354	196,968
Pre-2005 titles w/2 holdings	314,547	618,107	30,045	471,122	638,249	3,488	182,102
Pre-2005 titles w/1 holding	650,982	1,088,327	58,700	243,725	517,310	171,871	91,161

As with the individual Summaries, SCS will work with the group to produce iterations for various scenarios, to enable the group to assess the effect of different parameters.

Group-Level Withdrawal Candidate List (Expected Timeframe: January-February 2012)

The group-level Withdrawal Candidate List may be the most compelling output from this project. Based on criteria decided by the participating libraries (and on the extent of data comparability), this list will identify titles which meet the criteria – and which libraries hold copies. This is where collaboration and group decision-making finally meet the data. The group-

level list will serve as the basis for discussions regarding retention and preservation commitments.

These discussions will of necessity range beyond the data, and into which libraries have space, which libraries have deep collections in particular subjects, etc. Preservation commitments and service commitments will need to be worked out. Equity will need to be assured, so that the cost of the regional collection is shared. It is only once those decisions are made that withdrawals can begin. As of January 2012, the base data is ready and SCS has begun to work out an allocation strategy. But additional decisions from the group will also be needed.

Develop an Equitable Allocation Method for Group-Level Withdrawal Candidates (*Completed: March 2012*)

SCS has already begun work on the complex problem of equitably distributing titles from the group list to individual libraries for deselection. SCS will define a sequencing algorithm that assures 1) two copies of each title are protected; 2) withdrawal candidates are distributed proportionately across LC ranges; and 3) that the library allocated a title for deselection holds that title (based on the data supplied at the outset of the project). This allocation method will apply to any list agreed upon by the pilot libraries.

Produce Individual Library Picklists (*Completed: March 2012*)

Once it has been established which libraries may withdraw their copies, a more targeted list will be needed. The picklists will include only the individual library's withdrawal candidates, along with the necessary subset of data elements, e.g., location, barcode number, local call number. This will enable the withdrawing library to batch suppress records while items are removed from shelves and record maintenance performed.

Production of these picklists is contingent on the group's ability to finalize the decisions outlined above. Any delays in decision-making will likely affect the ability to deliver picklists in February/March. We will need to monitor this closely as the project progresses.

Produce Lists of Uniquely-Held, Zero-Circulation Titles for each library (*Completed: Dec 2011*)

This was not part of the original project plan, but the pilot group agreed to independent action among members for these titles. SCS produced the necessary lists as an added contribution to the project.

Throughout the project, SCS will provide additional iterations of the SCS Summary data as needed. This data will be used to help articulate a group retention policy which assures that at least two print copies of every title on the low-circulation overlap list will be permanently retained. For the MCLS group, but outside of the SCS purview, it will also be important to articulate how and under what conditions retained items will be serviced; i.e., lent to other libraries, or made available digitally. Similar commitments will also be articulated for items held uniquely within the group.

Finally, this data will be used to scope the size of the collective opportunity. If an MCLS-coordinated approach is adopted, how many volumes might be removed from which libraries under various policy scenarios? How many unique titles would be retained and preserved? What space might be made available for other purposes, and what corresponding dollar savings might accrue? This project will provide the data to answer those questions for a small group of libraries, and will simultaneously help define and refine the data and infrastructure needed.

Project Cycle: As of January 2012, SCS and the MCLS pilot group have collaborated for six months. The base data set has produced sufficient withdrawal candidates for approximately 2 years. It is expected that the normalized, aggregated data can serve to produce withdrawal and preservation candidate lists until July 2013. At that point, if the group decides to continue work on shared print, data for all pilot libraries will be re-loaded and other libraries can be added to the shared data set.

Mid-Cycle Libraries: Between January 2012 and July 2013, it is expected that libraries beyond the seven pilot participants will choose to join the project. This can be managed, but with some limitations. First, the data from new libraries cannot be fully integrated into the existing shared data without changing all of the withdrawal and preservation candidate lists for the pilot participants. To pursue full data integration would prevent the pilot group from acting until all recalculations had been performed.

To avoid slowing progress for the pilot libraries and to address concerns about equity, SCS and MCLS have agreed to proceed incrementally. Initially, the creation of a second group of six libraries was explored. As of August 31st, however, only two additional libraries have agreed to participate: Oakland University and Ferris State. Therefore, we have agreed to proceed as follows:

1. Participation remains open to new members, but as of August 31, 2012, the 10% discount no longer applies. Oakland and Ferris will both be accorded the discount.
2. SCS will work with Oakland and Ferris on data extract, normalization, and production of individual Collection Summary reports. This work took place in May 2012 for Oakland and will take place in August/September for Ferris.
3. SCS will not integrate any of the second group of libraries into the pilot group results. Retention commitments have already been allocated for the list of 743,000 titles generated among the original seven libraries. After discussion, the pilot group agreed to permit use of that list to the second group of libraries, and to make individual deselection decisions based on the knowledge that two MI-SPI libraries had already agreed to retain copies. This is a significant benefit to the second group. MI-SPI will revisit this issue and seek to re-balance costs and benefits when the data is refreshed in 2013-2014.
4. This approach allows everyone to move forward immediately. The pilot libraries can begin working their allocated titles from the 743K list. Oakland and Ferris can proceed immediately with the first stages of analysis. MCLS can continue to recruit additional libraries.

SCS will continue to communicate with MCLS, pilot libraries, and new libraries throughout this period. We will produce Collection Summaries in accordance with each library's specifications. Oakland, Ferris, and any other new libraries will join the MI SPI meetings in the fall of 2012, and will participate in all subsequent discussions as allowed by the group's Memorandum of Understanding.

Whatever is decided in the near term, the working premise is that when a new project cycle commences (probably in July 2013), new libraries will be integrated as full participants, and a new, expanded round of group-wide analysis can begin.

