EAST Cohort 2 Kickoff Webinars

Monday, June 5, 2017 11:00am – 12:30pm
Tuesday, June 6, 2017 11:00am – 12:30pm
Monday, June 12, 2017 1:00pm – 2:30pm
The EAST Project Team

Susan Stearns
Project Director
sstearns@blc.org

Sara Amato
Data Librarian
samato@blc.org

Matthew Revitt
Shared Print Consultant
matthew.revitt@maine.edu

Mei Mendez
Project Manager
mmendez@blc.org
Our Agenda

Part 1
EAST Mission, Governance, Policies & Procedures
Cohort 1 Work in Review
Cohort 2 Work Outline and Timelines

Part 2
SCS Overview

- Data Extracts
- Validation WG
- Collection Analysis and Retention Modeling
- Retention Commitments and Disclosures
- 59 members from Maine to Florida
- 51 Retention Partners
- 8 Supporting Partners
- Cohort 2 libraries even more diverse than Cohort 1
EAST Members

• All Partners:
  • Pay an annual membership fee to provide program support and staffing
  • Agree to lend materials to other EAST Libraries using standard ILL
  • Have a voice in EAST governance and growing advocacy for shared print collaboration

• Retention Partners – 51 of the current 59 members
  • Monograph and/or Serials/Journals retention
  • Participate in collection analysis with OCLC SCS – subsidized by EAST grant funding and membership monies
  • Commit to retain materials per agreed upon retention model for minimum of 15 years (through June, 2031)

• Supporting Partners – the remaining 8
EAST business model

- Retention Partners contribute @30%

- One time collection analysis
- Annual member fees
- Grant funding

- Paid by both Retention and Supporting Partners
- Staffing
- Administration
- Contribution to collection analysis subsidy

- Andrew W. Mellon Foundation and Davis Educational Foundation
- Start-up and initial implementation
- Collection analysis – 70% grant funded
- Validation sampling – 100% grant funded
The EAST mission

Securing the scholarly record in support of teaching, learning and research
We aim to preserve access to printed scholarly texts, while freeing library space for new uses.
Member Resources

April 05, 2017 to May 12, 2017

May Membership Update Webinar

The EAST Project Team will provide a Members Update in May. The webinar will provide an overview of EAST's activities since January 2017, including updates on the work of the Serials & Journals Working Group, Cohort 2, and the Deeper Dive into Data validation study. Other topics include...

February 16, 2017

ILL Best Practices webinar

This webinar provides an overview of the EAST ILL Best Practices, approved by the EAST Executive Committee.

The slides are available...

January 27, 2017

January Membership Update Webinar

Member Documentation

Throughout the EAST project, documentation of a variety of types has been developed to assist the EAST member libraries. Information about EAST's work at a high level is accessible via the dropdown menus above and the following are key pieces of additional project documentation:

Collection Analysis
- Video on final retention model
- Collection analysis & allocation review resources for members
- Post allocation video tutorial
EAST governance: Your opportunity

• Executive Committee (EC)
  • Primary governing body
  • Elected by the membership
  • Representative by type and size of library, consortia, and membership type
  • Two year term (some initial EC members elected to three years for continuity)

• Operations Committee (OC)
  • Charged with addressing issues relating to operationalizing EAST policies
  • Formal OC to be elected this summer
  • To include a Cohort 2 representative

• Working Groups – More on this later
Major policies and procedures

• EAST Memorandum of Understanding – see https://eastlibraries.org/policies-mou for details of these

• General Governance Policies

• EAST Members Rights and Responsibilities

• Major Operating Policies
  • Supplemented by documentation on replacing retention commitments for lost, missing or damaged items and Best Practices for Inter-Library Loan
  • Includes an FAQ on common questions relating to retention commitments
Major milestones 2015 - 2017

- Collection Analysis: over 16 million collective collection
- Retention Commitments: more than 6 million holdings
- Validation Sampling: 97% availability
- Policy Development: MOU, Major Policies and Procedures
- Initial work on collaboration with other shared print programs – HathiTrust
- Addition of Cohort 2 libraries
- Began serials/journals work with Cohort 1
## Cohort 2 Libraries and Timeline

**Cohort 2 (13 libraries)**
- Bucknell University
- Davidson College
- DeSales University
- Florida State University
- Furman University
- Gettysburg College
- Hofstra University
- New York University
- University of Pittsburgh
- Syracuse University
- Union College
- University of the South
- Villanova – Supporting Partner

**Anticipated Timeline**
- June, 2017 – Kickoff
- Mid August, 2017 – Data extracts to SCS
- Mid-October through mid-December – Validation sample study
- Mid-October through early December – retention modeling
- January, 2018 – Retention model finalized
- February/March, 2018 – Finalized retention commitments in GreenGlass
Timelines

- **May/June**: Libraries complete SCS questionnaire
- **July**: Libraries provide SCS with bib, item, & circ. data
- **August**: CAWG explores data in GreenGlass & models retention scenarios
- **Sept**: CAWG learns about retention models, comparator groups, etc.
- **Oct**: SCS load GreenGlass with holdings data
- **Nov**: SCS allocate retention commitments to Cohort 2 Retention Partners
- **Dec**: Libraries perform validation study
- **Jan**: Libraries review retention proposals & request adjustments before signing commitments.
- **Feb**: Libraries review retention proposals & request adjustments before signing commitments.
- **Mar**: Final commitments reflected in GreenGlass

**Dates**
- 2017
  - May/June
  - July
  - August
  - Collection Analysis Working Group (CAWG)
  - Kick-off Webinars
  - CAWG learns about retention models, comparator groups, etc.
- 2018
  - Sept
  - Oct
  - Nov
  - Dec
  - Final commitments reflected in GreenGlass
Cohort 2 work - Data gathering phase

OCLC/SCS providing detailed instructions and support
- Questionnaire due to SCS: by June 30th, 2017
- Data extracts due to SCS: by August 18th, 2017

Cohort 1 experts standing by if needed!
Lessons learned from Cohort 1:
- Know your data!
- Review feedback from SCS.
Validation Sample Study

What is it?

A small **sample study** to determine an “availability metric” that provides a rough estimate of the percentage of library holdings that are likely to be available (are accounted for). This will be used by the Collection Analysis Working Group as one factor in determining how many copies to retain.

Each library will receive a list of 6,000 titles randomly selected from the in-scope records that were sent to SCS. Libraries will check shelves to determine availability, and perform a quick condition assessment of the items using a data collection tool developed by EAST.
Cohort 2 - Validation Sample Study

• Building on work processes created by Cohort 1
• Validation Working Group to begin work in September
• Training in early October; Study runs Mid-October - mid-December
• Library to provide equipment, additional local training and supervising workers in the stacks
• Cohort 1 average times:
  • Averaged 45 bph (~135 hours) / 25 hours of admin overhead
  • Plan student workers!
• $4,000 reimbursement from EAST upon completion
Timelines

May/June
- Libraries complete SCS questionnaire

July
- SCS questionnaire

August
- Collection Analysis Working Group (CAWG)

Sept
- CAWG explores data in GreenGlass & models retention scenarios
- Libraries provide SCS with bib, item, & circ. data

Oct
- SCS load GreenGlass with holdings data
- Libraries perform validation study

Nov
- SCS allocate retention commitments to Cohort 2 Retention Partners

Dec
- Libraries review retention proposals & request adjustments before signing commitments.

Jan
- Libraries review retention proposals & request adjustments before signing commitments.

Feb
- Libraries review retention proposals & request adjustments before signing commitments.

Mar
- Final commitments reflected in GreenGlass

Kick-off Webinars
- CAWG learns about retention models, comparator groups, etc.
Importance & Scope of Collection Analysis

- Collection analysis is a core activity of shared print projects like EAST. Our focus is identifying titles as a group we commit to retain.
- Partnering again with OCLC Sustainable Collection Services (SCS).
- Approximately 13 million print monograph titles across collections of 12 EAST Cohort 2 Retention Partners.
- Scope of analysis is circulating print monographs only.
Collection Analysis & Retention Modeling

• Collection Analysis Working Group will take the lead & be tasked with identifying categories of titles EAST should protect by:
  • Comparing circulating monograph holdings of EAST Retention Partners to identify uniqueness and overlap within EAST Cohort 2
  • Identifying the occurrence of these holdings elsewhere both nationally and regionally
  • Integrating circulation activity
  • Considering factors like EAST Cohort 1 retention commitments and results of validation study
Collection Analysis & Retention Modeling

• Agree on categories of material that we won’t retain:
  • Non-Scholarly titles and material that rapidly becomes outdated e.g. textbooks and guides
  • Recently published material
• Use GreenGlass Retention Model feature to develop retention rules
• Agree on how retention responsibility will be allocated across the group
Timelines

May/June
- Libraries complete SCS questionnaire
- Collection Analysis Working Group (CAWG)

July
- Libraries provide SCS with bib, item, & circ. data
- Kick-off Webinars

August
- CAWG explores data in GreenGlass & models retention scenarios

Sept
- SCS load GreenGlass with holdings data

Oct
- SCS allocate retention commitments to Cohort 2 Retention Partners

Nov
- Libraries perform validation study

Dec
- Libraries review retention proposals & request adjustments before signing commitments.

Jan

Feb
- CAWG learns about retention models, comparator groups, etc.

Mar
- Final commitments reflected in GreenGlass
Collection Analysis & Retention Modeling: Timeline

- **June 5-12 2017** - Kickoff webinars
- **June 30, 2017** - Libraries complete OCLC/SCS Cataloging & Data Questionnaire
- **June - July, 2017** - Project Team recruit and appoint an Collection Analysis Working Group
- **July - October 2017** - Collection Analysis Working Group introduced to retention concepts, review EAST Cohort 1 Retention Model, and agree on comparator groups & special category flags
- **August 18, 2017** - Libraries provide SCS with bib, item, & circ data
Collection Analysis & Retention Modeling: Timeline

- **October 20, 2017** - SCS load GreenGlass with library and EAST collective holdings data
- **October 23 - December 1, 2017** - Supported by SCS the Collection Analysis Working Group explores data in GreenGlass & develops retention model scenarios:
  - Consider retention factors
  - Develop draft retention models and seek feedback from Cohort 2 libraries
  - Revise retention model based on feedback
Collection Analysis & Retention Modeling: Timeline

- **December 22, 2017** - Project Team submit final version of Retention Model to EAST Executive Committee for approval
- **December 22, 2017 - February 4, 2018** - SCS allocate retention commitments to Cohort 2 Retention Partners
- **February 5, 2018** - SCS reload GreenGlass to reflect retention commitments
- **February 5 - March 2, 2018** - Libraries have the opportunity to review their retention proposals and request adjustments before agreeing to their final commitments
Collection Analysis & Retention Modeling:
Timeline

• **March 26, 2018** - SCS reload GreenGlass with adjusted commitments

• **January 31, 2021** - Libraries will have continued access to GreenGlass through this date
Timelines

May/June
- Libraries complete SCS questionnaire

July
- Libraries provide SCS with bib, item, & circ. data

August
- CAWG explores data in GreenGlass & models retention scenarios

Sept
- SCS load GreenGlass with holdings data
- CAWG learns about retention models, comparator groups, etc.

Oct
- SCS allocate retention commitments to Cohort 2 Retention Partners

Nov
- Libraries perform validation study

Dec
- Libraries review retention proposals & request adjustments before signing commitments.

Jan

Feb

Mar
- Final commitments reflected in GreenGlass
Disclosure

• Once retention commitments have been finalized, libraries are expected to disclose these commitments in their ILS and in OCLC when that functionality becomes available.
• EAST recommends using the 583 field, and has created best practices to facilitate this process.
• There are also ILS-specific instructions in the EAST 583 Working Group website.
Thank you.

Q & A

sstearns@blc.org

samato@blc.org

mmendez@blc.org

matthew.revitt@maine.edu
EAST – Cohort 2
Collection Analysis Kick-Off

SCS Sustainable Collection Services
Data-Driven Deselection

BOSTON LIBRARY CONSORTIUM
Today’s SCS Topics

• Overview of Sustainable Collection Services

• Project plan, strategies and schedule

• Cataloging and Data Questionnaires

• Data Extracts
WHO AND WHAT IS SCS?
A problem in academic libraries

- Stacks are overcrowded
- Use of print books is low and declining
- Library space is wanted for other purposes
- Print redundancy is significant
- The cost of keeping books on shelves is high
- Alternatives exist, but data is scattered
- Traditional approaches to deselection are costly and time-consuming
Mission
Helping Libraries
Manage and Share
Print Monographs
Sustainable Collection Services (SCS)

SCS: Helping libraries manage and share print monographs since 2011

- **2011**: SCS Founded
  - Data-driven deselection for print monographs

- **2015**: SCS Joins OCLC
  - Access to new data and resources

- **2016**: Group Functionality
  - See combined holdings as a single shared collection

- **2017**: Journal Deselection
  - (Preliminary title-level support)
SCS tools originally designed for individual library projects in the context of other library holdings

Primarily focused on responsible deselection
Group Projects

As of April 2017, retention commitments for 18.5 million title-holdings. (most of them informal)

These represent 7.4 million distinct OCLC numbers.
The challenge in working with groups of libraries is that very few participant libraries have ever done anything like this before.

SCS tools and services are designed to
- Help define the group
- Introduce new concepts and vocabulary
- Facilitate the conversation
- Ensure a shared understanding of the retention agreement
- Enable local action
Individual Library Features

- Summary info
- Visualizations
- Query builder
- Known item search
- Remediation lists
  - Lists of items

Group Features

- Summary info
- Visualizations
- Model Builder
Library's bib, item and transaction data for PRINT MONOGRAPHS
GreenGlass tutorials

Individual Library Features in GreenGlass

- What is GreenGlass? (2 minutes)
- Understanding Key Metrics (4 minutes)
- Collection Visualization (4 minutes)
- Item Lists and Item Details (4 minutes)
- GreenGlass Analysis (7 minutes)

Group Features in GreenGlass

- What is GreenGlass group functionality? (2 minutes)
- Group Retention Models (9 minutes)
- Title Sets, Title Holdings, and Items (3 minutes)
- Post-Allocation Features in GreenGlass (4 minutes)

Brief Topical Tutorials

- Establishing Your List of Comparator Libraries (4 minutes)
- Possible Duplicates (2 minutes)
- SAME and ANY Edition Holdings Tallies (3 minutes)
- Multi-Edition Titles (3 minutes)
- HathiTrust Matches and URLs (3 minutes)
- Ebook/Print book overlap data in GreenGlass (6 minutes)
- Understanding Out-of-Scope Records (5 minutes)
- Combining GreenGlass Queries (4 minutes) (NEW)
- Ruth Fischer: Project manager/analyst
- Jen John: Senior implementation program manager
- Eric Redman: Individual and group data preparation
- Rick Lugg: Market and project strategist
- Argentic Software: Developers of GreenGlass and ETL Support
EAST PROJECT PLAN
FOR COLLECTION ANALYSIS
<table>
<thead>
<tr>
<th>Task</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kick-off Meetings</td>
<td>June 5 - 12</td>
</tr>
<tr>
<td>Cataloging and data questionnaires completed</td>
<td>June 30</td>
</tr>
<tr>
<td>Comparator libraries finalized and special category defined</td>
<td>July 30</td>
</tr>
<tr>
<td>All bib and item data received from libraries</td>
<td>August 18</td>
</tr>
<tr>
<td>Random lists of items for validation study</td>
<td>October 15</td>
</tr>
<tr>
<td>GreenGlass loaded</td>
<td>October 20</td>
</tr>
<tr>
<td>Consensus reached on retention model</td>
<td>December 22</td>
</tr>
<tr>
<td>GreenGlass reloaded to reflect retention candidates</td>
<td>February 5, 2018</td>
</tr>
<tr>
<td>Retention review and rejected allocations reported to SCS</td>
<td>March 2, 2018</td>
</tr>
<tr>
<td>GreenGlass reloaded with FINAL retention commitments</td>
<td>March 26, 2018</td>
</tr>
<tr>
<td>Separate instance of GreenGlass dedicated to retentions</td>
<td>April 23, 2018</td>
</tr>
<tr>
<td>Ongoing access to GreenGlass</td>
<td>January 31, 2021</td>
</tr>
</tbody>
</table>
# EAST Cohort 2 participant libraries

<table>
<thead>
<tr>
<th>INSTITUTION</th>
<th>ILS</th>
<th>QUESTIONNAIRE CODE KEYS, ETC</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeSales University</td>
<td>Innovative, moving to WMS in June</td>
<td></td>
</tr>
<tr>
<td>Bucknell University</td>
<td>OCLC's World Share</td>
<td></td>
</tr>
<tr>
<td>Davidson College</td>
<td>OCLC's World Share</td>
<td></td>
</tr>
<tr>
<td>Hofstra University</td>
<td>III Sierra</td>
<td>✓</td>
</tr>
<tr>
<td>University of the South [Sewanee]</td>
<td>III Sierra</td>
<td></td>
</tr>
<tr>
<td>Gettysburg College</td>
<td>III Sierra</td>
<td></td>
</tr>
<tr>
<td>Florida State University</td>
<td>Aleph 20</td>
<td></td>
</tr>
<tr>
<td>Syracuse University</td>
<td>ExLibris Voyager</td>
<td></td>
</tr>
<tr>
<td>Furman University</td>
<td>III Millennium</td>
<td></td>
</tr>
<tr>
<td>New York University</td>
<td>Aleph (with Primo front end)</td>
<td></td>
</tr>
<tr>
<td>University of Pittsburgh</td>
<td>ExLibris Voyager</td>
<td></td>
</tr>
<tr>
<td>Union College</td>
<td>III Sierra</td>
<td></td>
</tr>
</tbody>
</table>
Project Scope

- **In Scope:** *Circulating Print Monographs*

- **Out of Scope:**
  - Journals (serial records)
  - Special Collections
  - Reference (non-circulating)
  - E-Books
  - Government Documents
  - Media (non-book)
  - Microforms
  - Maps
  - Scores
  - Dissertations, Theses
  - LOST, MISSING, DAMAGED ITEMS
THE QUESTIONNAIRE
# of in-scope monographs

(bib records)
OCLC holding symbol(s)
usage data

includingcirculations, browses, re-shelving counts, reading room loans, reserve charges, etc.

date of last migration
location of bib record numbers and OCLC numbers

multiple OCLC numbers
location information
classification schemas !!!
call numbers !!!!!!!!!
call number type codes
cataloging anomalies

local practices
Tell us EVERYTHING you think we should know!

We’ll get back in touch if questions arise.
DATA EXTRACTS
Step 1: MARC records for circulating print monographs

- Record type ‘a’ (print resource) in MARC leader byte 06 and
- Bib level ‘m’ or ‘a’ (monograph or monograph part) in MARC leader byte 07

(XML files are fine)
Filter out:

- Journals (serial records)
- Special Collections
- Reference (non-circulating)
- E-Books
- Government Documents
- Media (non-book)
- Microforms
- Scores
- Maps
- Theses and dissertations
Also filter out

- Lost
- Missing
- Billed
- Non-circ
- Any other status codes or type codes that indicate that an item may be unavailable
Send all corresponding item and usage data

- Item and circulation data can be delivered in 945 sub-fields of the MARC record in which case, we also need a mapping – so that SCS can know which data element will be found in each sub-field. (949 sub-fields are fine; XML files are also fine).

- Item and circulation data can also be delivered in a separate tab delimited file. In this case, we need very clear (descriptive) headers, again so that SCS can know how each column is populated. Be sure that the associated bib record number is included for each item, so that we have a dependable match point.

- Sometimes, critical item data is in a MARC Holdings record (MFHD). If so, be sure to send them!

- Multiple files are fine.
Item and circulation data
Send us all you’ve got!

- item call number
- enumeration
- last reserve date
- copy #
- in-house uses
- barcode
- last check-in date
- last check-out date
- item create date
- item record number

- location code*
- location name
- item type code*
- note field*
- opac message*
- item status code*
- total checkouts

* please send a key of these codes
SCS filters, normalizes, structures the data from each library

SCS will:

• Filter out-of scope bib records (eBooks, maps, DVDs, Gov Docs)
• Eliminate duplicate bib records
• Choose and normalize call numbers
• Eliminate trailing spaces in control numbers
• Validate OCLC numbers
• LCCN/ISBN/title-string lookups for records lacking an OCLC number
• Identify and accommodate unusual implementations of MARC
• Identify bibs without items and items with multiple bib records
• Map item-level data and interpret codes
• Assign LC (and/or Dewey) Classes to records
Data Remediation Lists Available in GreenGlass

- OCLC numbers assigned by SCS
- Records without OCLC numbers
- Holdings not set in WorldCat
- WorldCat Title/Author Risk
- Multiple OCLC numbers
- Possible duplicates
- Bib Records that Share an OCLC Record #
- Bib Records that Share an OCLC Work ID#
- Other
  - Hathi Public Domain titles
  - HathiTrust URLs
Your data is static. It’s a snapshot.

- Avoid deaccessioning in-scope monographs
- Avoid relocation projects related to monographs
- Avoid major changes to the monographs collection

Until EAST is ready to move forward in a coordinated way, avoid changes to inventory!
QUESTIONS?
For now, focus on:

1. Completing your questionnaire
2. Your extract – bib and item data (delivered via FTP)
3. Your 945 sub-field map (if appropriate)
4. Keys to your item, status, and location codes

Your primary contact for all this:

Jen John

jennifer.john@oclc.org
## Project Timeline: 2017 - 2021

<table>
<thead>
<tr>
<th>Task</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kick-off Meetings</td>
<td>June 5 -12</td>
</tr>
<tr>
<td>Cataloging and data questionnaires completed</td>
<td>June 30</td>
</tr>
<tr>
<td>Comparator libraries finalized and special category defined</td>
<td>July 30</td>
</tr>
<tr>
<td>All bib and item data received from libraries</td>
<td>August 18</td>
</tr>
<tr>
<td>Random lists of items for validation study</td>
<td>October 15</td>
</tr>
<tr>
<td>GreenGlass loaded</td>
<td>October 20</td>
</tr>
<tr>
<td>Consensus reached on retention model</td>
<td>December 22</td>
</tr>
<tr>
<td>GreenGlass reloaded to reflect retention candidates</td>
<td>February 5, 2018</td>
</tr>
<tr>
<td>Retention review and rejected allocations reported to SCS</td>
<td>March 2, 2018</td>
</tr>
<tr>
<td>GreenGlass reloaded with FINAL retention commitments</td>
<td>March 26, 2018</td>
</tr>
<tr>
<td>Separate instance of GreenGlass dedicated to retentions</td>
<td>April 23, 2018</td>
</tr>
<tr>
<td>Ongoing access to GreenGlass</td>
<td>January 31, 2021</td>
</tr>
</tbody>
</table>

We are here