July 20, 2015

EAST
Collection Analysis Kick-Off

SCS Sustainable Collection Services
Data-Driven Deselection

BLC
BOSTON LIBRARY CONSORTIUM

OCLC

Welcome!

Let's introduce ourselves.

OCLC

Today's Agenda

- Overview of the EAST Project – Susan/Lizanne
- Brief environmental scan - Rick
- Overview of SCS – Rick
- Project plan, strategies and schedule – Rick
- Lunch
- GreenGlass and GreenGlass for Groups - Ruth
- Data Extracts - Ruth
- Cataloging and Data Questionnaires - Ruth

OCLC

EAST – EASTERN ACADEMIC SCHOLARS' TRUST

EAST Goals

- Share stewardship of library print holdings to ensure more effective preservation of the scholarly record and recognize that some institutions may need to alleviate space pressures

- Provide access and delivery of the shared materials to meet the needs of scholars, researchers, teachers and their students

- Provide information to support the needs of libraries for separate contracted offsite shelving for local collections.

BLC
BOSTON LIBRARY CONSORTIUM

EAST Policies

- Recommended by EAST working groups, approved by the Steering Committee, and endorsed by the plenary EAST meeting in July 2014.

- Will be reviewed and potentially revised during EAST implementation

- Will be embodied in a formal Memorandum of Understanding (MOU) as the basis for ongoing membership in EAST

Retention

- Distributed retention model, libraries retain specific titles in their local collections

- <u>Minimum</u> of 15 year retention commitment for both monographs and serials. Libraries can commit to longer retention.

- Retention commitments to be revisited at least 2 years before they end and at least every 5 years after that

Access and delivery

- EAST is a light archive

- Holdings discoverable using existing systems

- Fulfill requests for materials through ILL using existing systems

- Current recommendation is for no charge for ILL to other EAST members

EAST Membership

- Two forms of membership
  - Retention Partners
  - Supporting Partners

- 47 current members
  - 38 monograph Retention Partners (some also for journals)
  - 1 journal Retention Partner only
  - 8 Supporting Partners

EAST Business Model

- EAST expenses include program management and collection analysis

- Staffing = full time project manager, half-time data librarian, planning consultant, and statistical consultant

- Financial support
  - Mellon grant for Years 1 and 2 ($995,000 total)
  - Davis grant for Years 1 and 3 ($400,000 total)
  - Membership fees for Years 1 and 2 (~$500,000 total) support program management and a portion of collection analysis costs

BLC
BOSTON LIBRARY CONSORTIUM

Major Project Activities

1. Analyze circulating monograph collections in order to propose commitments to be made by retention partners
2. Design, test and analyze a sample-based validation study
3. Secure retention commitments
4. Finalize EAST policies
5. Plan for future EAST membership
6. Explore relationships with other regional and national shared print programs

BLC
BOSTON LIBRARY CONSORTIUM

Timeline Year 1 (2015-2016)

- Establish EAST organization (July – October)
  - Elect Executive Committee, establish Working Groups
  - Hire project staff
- Conduct monograph collection analysis
  - Ingest and aggregate data (July – December)
  - Model scenarios and prepare retention proposals (January – June 2016)
- Design & conduct validation study #1 (October 2015 – April 2016)
- Reconfirm or revise policies (throughout year)

BLC
BOSTON LIBRARY CONSORTIUM

Timeline Year 2 (2016-2017)

- Design & conduct validation study #2 (July – November)
- Finalize retention agreements (December 2016 – March 2017)
- Finalize and execute MOU and retention agreements (April – June 2017)
- Define plans for future EAST membership (October 2016 – June 2017)
- Plan for collaboration with other regional/national shared print initiatives (September 2016 – April 2017)

BLC
BOSTON LIBRARY CONSORTIUM

EAST Validation Studies

- Design and execution will be supervised by the Validation Working Group, in consultation with other working groups and the Steering Committee, as needed.

- Goal is to build trust in retention commitments, by answering these questions:
  - Will confidence levels vary by library collection?
  - Will confidence levels vary across tiers of redundancy?
  - Will first-round sampling indicate other criteria for offsetting risk?

- Validation will focus on inventory confirmation, not item condition.

BLC
BOSTON LIBRARY CONSORTIUM

Two-stage sampling process

- Study #1: Each Retention Partner's collection will be sampled to determine the local likelihood that an item in the catalogue is on the shelf.

- Study #2: In order to offset local uncertainties, specific subsets will be sampled globally, across consortial collections. The selection criteria for these subsets will be based on the findings in the first round of sampling. Redundancy will be one criterion.

BLC
BOSTON LIBRARY CONSORTIUM

EAST Governance

Executive Committee

- Approve EAST operating policies

- Approve the initial MOU and future changes to the MOU, including changes to the business model

- Approve future changes to the business model, operating budget, and member fees

- Approve EAST membership criteria and new EAST members

Working Groups

- Monographs

- Journals

- Validation

- Governance

BLC
BOSTON LIBRARY CONSORTIUM

THE ACADEMIC LIBRARY ENVIRONMENT

Evolution of the Library Paradigm

<u>Reader-centered</u>: from monastic scriptorium and library; dominated by light and reading tables

<u>Book-centered</u>: collection growth; unrelenting need for more shelving

<u>Learning-centered:</u> digital content; information commons; learning spaces; information literacy

Source: Scott Bennett, *Libraries and Learning:*

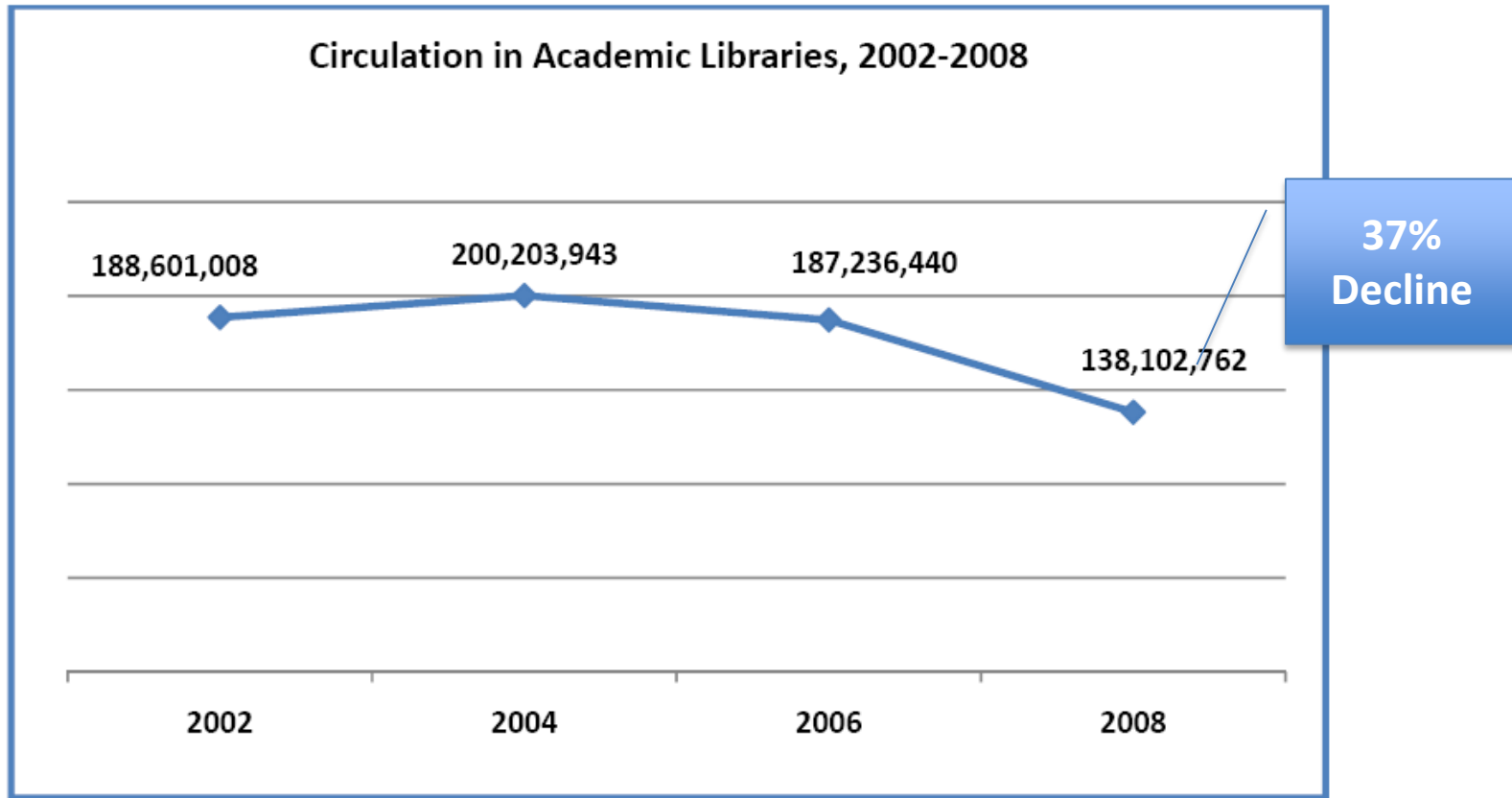*A History of Paradigm Change (2003)*

SCS

The Problem

- 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

SCS

Stacks are crowded *and* empty

SCS

Circulation in Academic Libraries Continues to Decline

Circulation in Academic Libraries, 2002-2008

188,601,008   200,203,943   187,236,440

138,102,762

37% Decline

2002   2004   2006   2008

Source: National Center for Education Statistics. *Academic Libraries*, various years. Table 1 (http://nces.ed.gov/pubsearch/getpubcats.asp?sid=041#)

SCS

Space Requirements: Monographs

Volumes	Square Feet
100,000	20,000
250,000	45,000
500,000	80,000
1,000,000	150,000
2,700,000	405,000
	Source: Stephen R. Lawrence, Lynn Silipigni Connaway, and Keith H. Brigham, "Life Cycle Costs of Library Collections" *College & Research Libraries,* November 2001, p. 546.

SCS

Library space is wanted for other purposes...

"The crowding out of readers by reading materials is one of the most common and disturbing ironies in library space planning."

--Scott Bennett

Lifecycle Costs: Monographs

The Idea of Order
Transforming Research Collections for 21st Century Scholarship
June 2010

COUNCIL ON LIBRARY AND INFORMATION RESOURCES

On the Cost of Keeping a Book

Paul N. Courant and Matthew "Buzzy" Nielsen*

1. Introduction

Among the critical functions of research and academic libraries is preservation—keeping both the scholarly record and much of the associated cultural record fit for use over time. Until recently, maintaining this record entailed managing primarily printed works, of ink on paper, bound in book form. The advent of electronic texts poses a novel and expensive set of preservation problems for academic libraries that have been addressed by many current and recent studies on the cost of digital preservation.[1] The topic of this report is the cost of storing and using print in old-fashioned codex form.[2]

We have two motivations for doing this work. The first is something of a straw man: as librarians and their funders become increasingly aware of the daunting technical and economic problems associated with digital preservation, there is often a certain wistfulness for

- CLIR, June 2010

- Courant & Nielsen

- Estimated Annual Costs

  ➤ $4.26/ volume annually in central stacks

  ➤ $0.86/volume in high-density facility

SCS
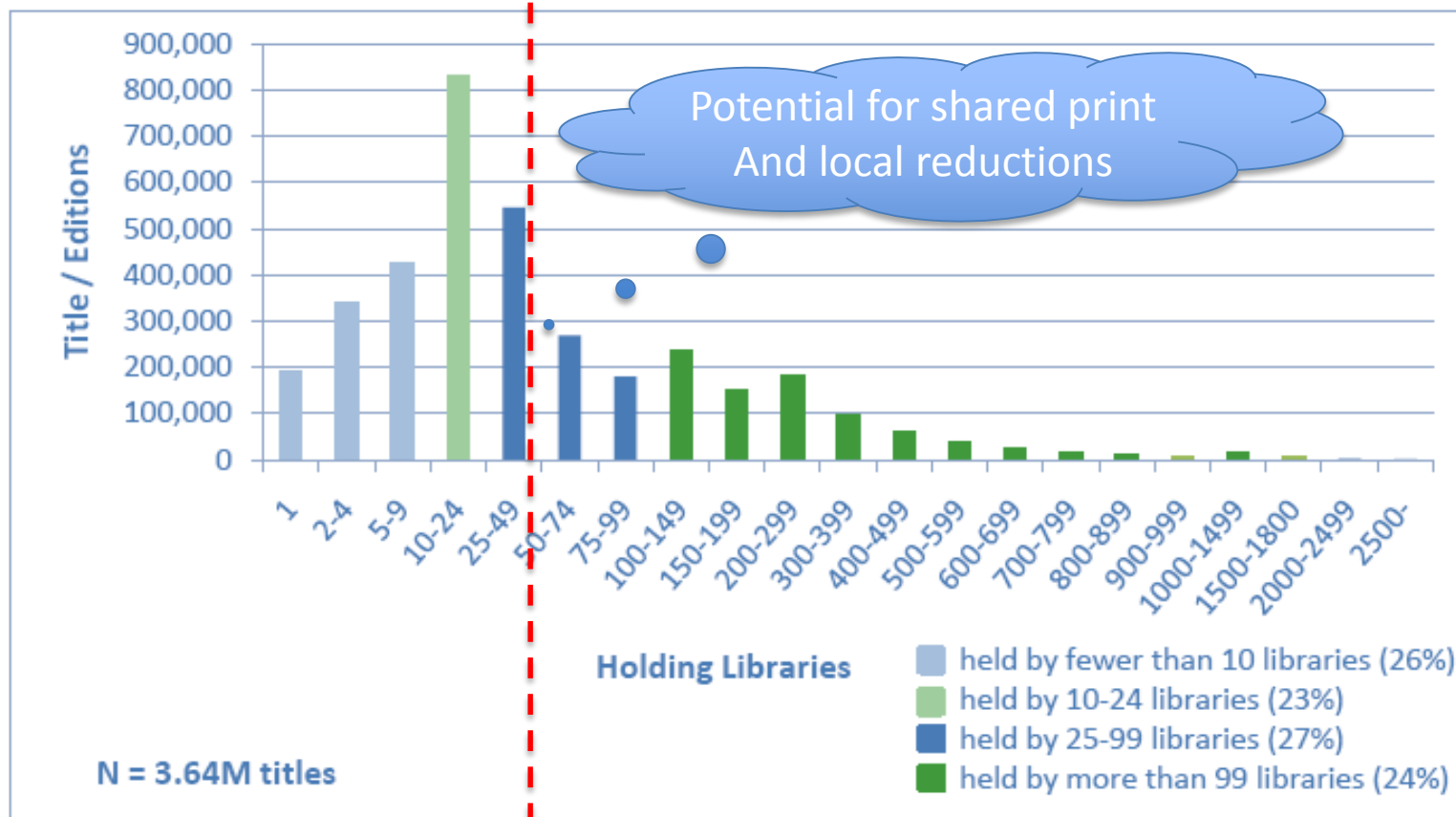
Print redundancy is significant...

Figure 8. System-wide distribution of library holdings for titles in HathiTrust Digital Library (June 2010)

Two functions of library print collections

- <u>Preserving</u> function

- <u>Dispensing</u> function

Source: Michael Buckland, Redesigning Library Services: A Manifesto (Chicago: American Library Association, 1992).

SCS

'Archive' copies

- Print Archives
  - Failsafe for technological or natural disaster
  - New digital surrogates or re-digitization
  - Dark, dim, or light?
  - People trust print

- Digital Archives
  - Secure, high-quality
  - Hathi Trust, Portico
  - CRL certification

SCS

'Service' copies

- Once content is securely archived, 'dispensing' function can be managed with fewer copies

- Focus on distribution, convenience, speed of delivery

- Borrow or re-purchase; print, electronic (including PDA, DDA, Short-term Loan); POD

SCS

'Surplus' copies

- Archiving requirements satisfied

- Sufficient service copies to meet anticipated demand

- How many holdings/copies remain?

- Are all of them needed?

- Share? Store? Withdraw?

SCS

WHO AND WHAT IS SCS?

Davidson College, January 2008

Recommendation: Adopt a Rules-based approach to weeding, storage, and sharing...

OCLC

SCS Sustainable Collection Services
Data-Driven Deselection

Sustainable Collection Services (SCS) offers deselection decision-support tools to academic libraries.

What is GreenGlass

Sustainable Collection Services

Upcoming Events

- **August 29, 2013** *Maine Shared Collection Strategy (MSCS):* Meeting with Collection Development committee.

Current Projects

Boston College
Washington Research Library Consortium (WRLC)
San Jose State University
College of the Holy Cross
Virtual Library of Virginia (VIVA)
US Naval Academy
Wake Forest University
Richard Stockton College of New Jersey
Tri-University Group (Waterloo, Guelph, Wilfrid Laurier)
Trinity University (TX)
Central Iowa–Collaborative Collection Initiative (CI–CCI)
Maine Shared Collection Strategy (MSCS)
Gonzaga University
Quinnipiac University
Regis College (MA)
Northwest Nazarene University
McGill University
Macalester College
Drake University
College of Brockport (NY)

Sustainablecollections.com

Deselection: A Range of Options

- **'Deselection' can encompass several goals:**

  – Transfer to offsite storage, automated storage & retrieval systems (ASRS) or compact shelving

  – Shared Print Archiving

  – Retention and Preservation

  – Digitization

  – Weeding or Withdrawal

- **Connotes the care with which print collections must be managed**

SCS

Mission

Helping Libraries **Manage** and **Share** Print **Monographs**

SCS

January 8, 2015

SCS: Vendor Support for Monographs Analysis

Consulting

Project Management

Tools to interact & visualize

Value added to data

Data

THE SCS TEAM & EXPERIENCE

SCS People & Roles in EAST Project

- Ruth Fischer   Project manager/analyst
- Rick Lugg   General oversight/strategy
- Andy Breeding   Project analyst as needed
- Eric Redman   Individual and group data preparation

- Argentic Software   Developers of GreenGlass

OCLC

Individual Library Projects

- Johns Hopkins
- Boston College
- MIT Libraries
- Amherst College
- University of Vermont
- Middlebury College
- Rhode Island College
- Roger Williams University
- SUNY/Brockport
- SUNY/Oneonta

OCLC

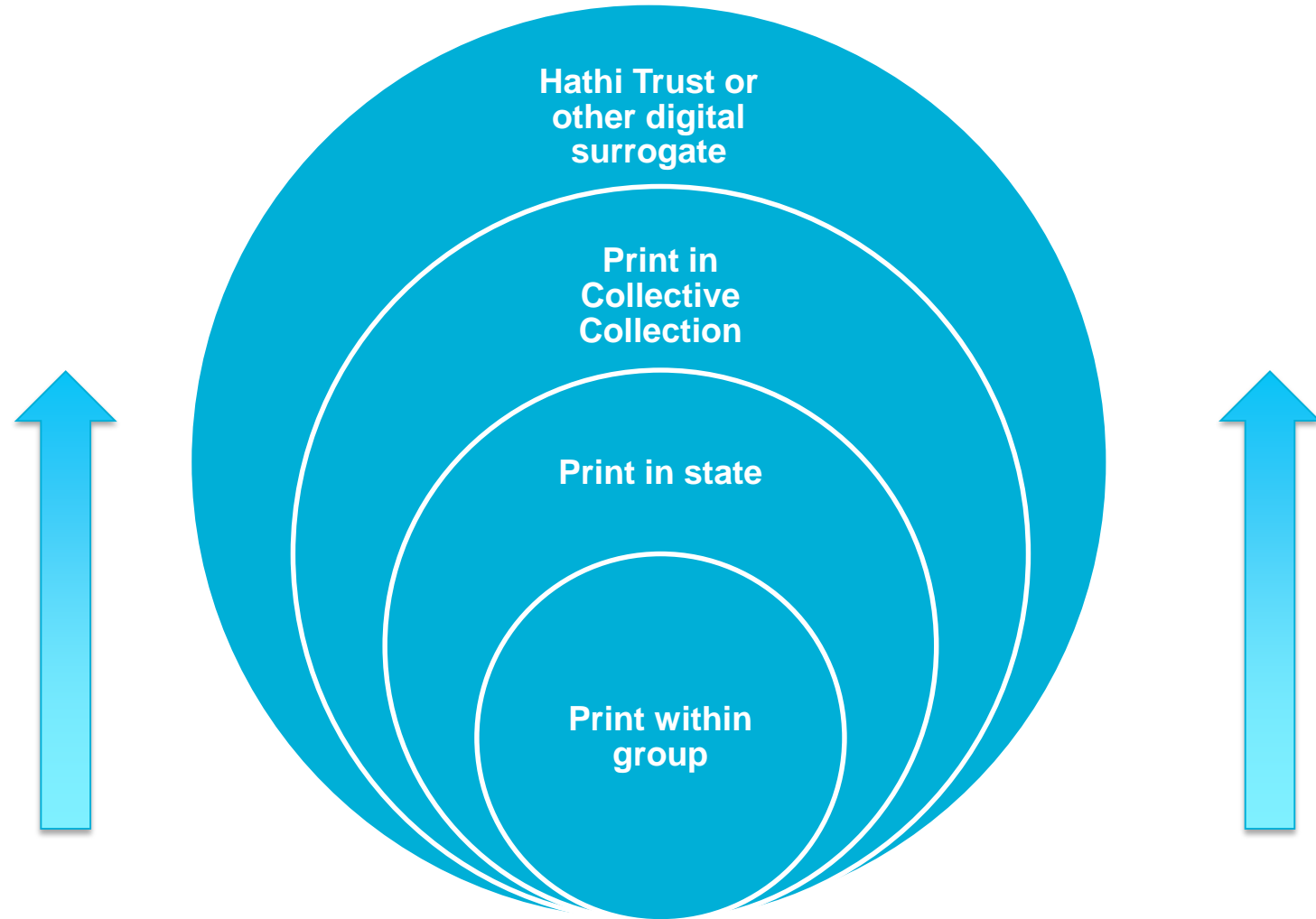# Independent action in a collective context
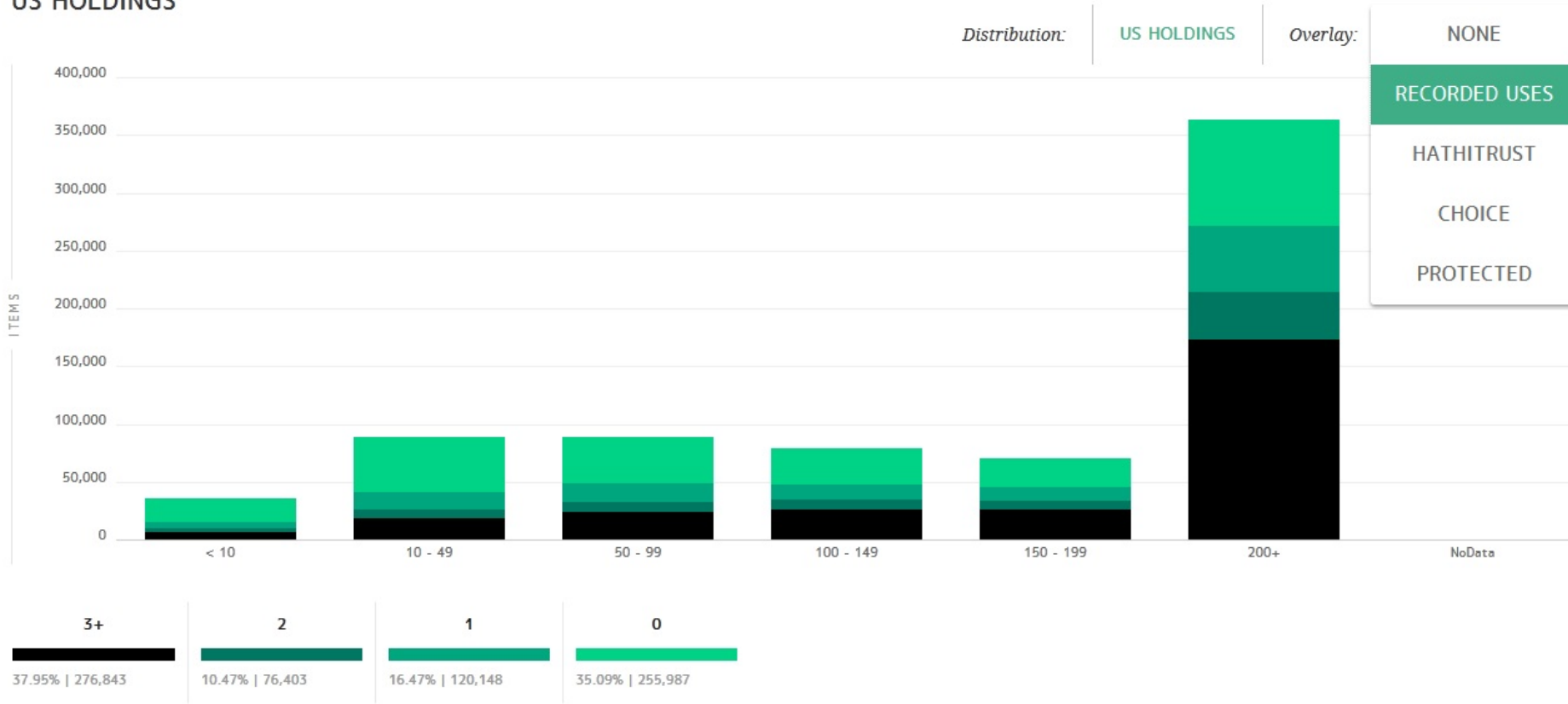
Group Projects

- MI-SPI (x2)
- CNY
- Maine
- VIVA Video
- VIVA Books
- WRLC
- CI-CCI
- TUG
- ALI/PALNI
- CSUs

As of February 2015,
retention commitments for
8 million title-holdings.

(most of them informal)

These represent
3.8 million distinct OCLC numbers.

Strong Preferences: Print, Self-Sufficiency, Proximity

Hathi Trust or other digital surrogate

Print in Collective Collection

Print in state

Print within group

Shared Print Projects

- 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
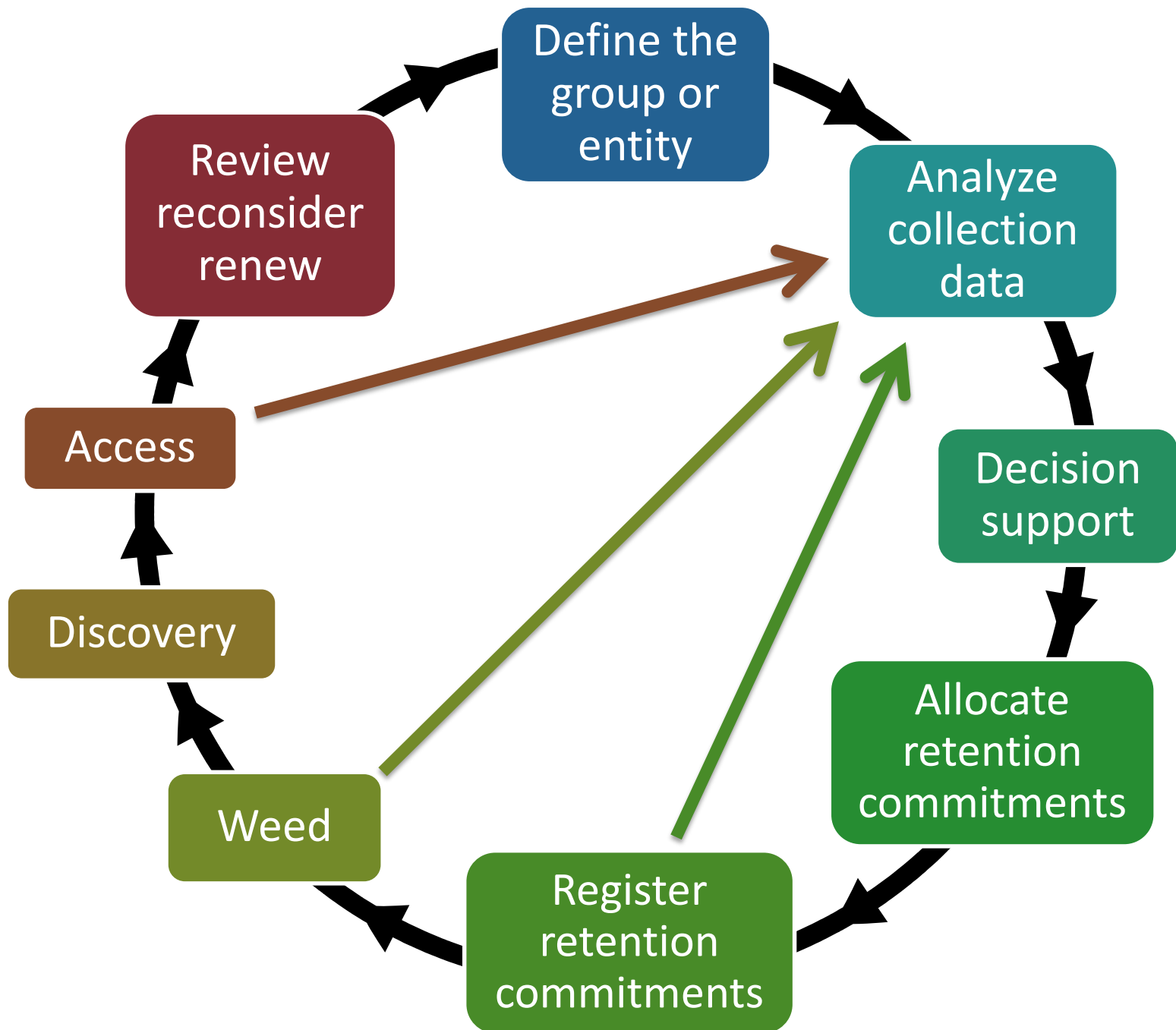
EAST PROJECT PLAN FOR COLLECTION ANALYSIS

Collection Analysis Participants

1. New collection analysis participants (~30 libraries)
   - Data work will start from scratch
   - Bib, item, and circ extracts

2. Recent SCS clients (5 libraries)
   - Data is less than two years old
   - Original data will be used, but re-scoped

3. Older SCS Group projects (MCSS, ConnectNY)
   - Data is more than two years old
   - Retention commitments will be used as targets

OCLC

Data work varies by category

1. New data sets: extracts of MARC records for in-scope items, plus associated item and circulation data

2. Re-worked data sets: SCS will use the library's original data, filtered and scoped to EAST group specifications

3. Retention commitments as targets: SCS will reformat CNY Pod 1 retention lists, and incorporate those commitments into the EAST group data set. MSCC will be used as a WorldCat comparator group. No other data from these projects will be used.

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
  - Scores

Project Timeline: 2015

Task	Timeframe
Kick-off Meeting: Organizational	June
Kick-off Meetings: Collection Analysis	July
Confirm Project Leadership/Communication	June/July
Confirm Scope and Specifications	July/August
Category 1 Libraries: Data Questionnaire	July-September
Category 1 Libraries: Supply Bib, Item, Circ Data	July-October
Category 2 Libraries: Modify Existing Data	July-September
Category 3 Libraries: Configure Retention List for Matching	July-August
Conduct Validation Study #1	October-April
Category 1 & 2 Libraries: WorldCat validation & matching; Hathi Trust matching; match against Cat retention lists	November-December

We are here

Project Timeline: 2016

Task	Timeframe
Category 1 & 2: Compile and validate individual data roll-ups	November-January
GreenGlass access for Individual Libraries [optional]	November-January
Compile the EAST group-wide database and provide GreenGlass for Groups	January 2016
Experiment with Retention Scenarios	February - June
Discussion/Decisions on EAST retention commitments	February - June
Allocation of retention proposals by SCS	June - July
Re-load of GreenGlass to reflect retention proposals and establish a safety net	July
Conduct validation study #2	July - November
Finalize retention commitments incorporating validation study #2	December 2016 – March 2017
Individual libraries validate as needed, and declare retention commitments in catalogs, WorldCat	April 2017 and future
Individual libraries deselect as needed for titles they have not committed to retain	April 2017 and future

GREENGLASS AND GREENGLASS FOR GROUPS (G3)

OCLC

SCS

OCLC WorldCat®

HATHI TRUST

CHOICE · OUTSTANDING ACADEMIC TITLE

GreenGlass

Library's bib, item and transaction data for PRINT MONOGRAPHS

OCLC

GREENGLASS DEMO

For a comprehensive series of GreenGlass
video tutorials go to:

http://www.oclc.org/sustainable-
collections/resources.en.html#tutorials

Retention Allocations and Beyond

- SCS allocates titles to participant libraries – based on agreed parameters and thresholds – to be considered for retention

- SCS reloads GreenGlass to reflect retention allocations

- Participant libraries validate and secure their retained items

- Participant libraries weed according to local needs

OCLC

DATA EXTRACTS

Send ALL in-scope MARC records

- Filter out:

  - Journals (serial records)

  - Special Collections

  - Reference (non-circulating)

  - E-Books

  - Government Documents

  - Media (non-book)   (XML files are fine)

  - Microforms

  - Scores

Also filter out

- Lost

- Missing

- Billed

- Non-circ

- Any other status codes or type codes that indicate that an item may be unavailable

OCLC

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 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

OCLC

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

OCLC

Example of bib records filtered out

Bib Records filtered out	21,675
Government docs	*1,880*
Non-language materials	*2,821*
Non-monographic materials	*1,880*
Non-print resources	*13,725*
Unable to obtain OCLC number	*3,461*
Bib Title/Author mismatch with OCLC	*279*
Multiple OCLC numbers per record	*47*

OCLC

Data Summary

Once SCS has taken these steps, we will send you a Data Summary, so that you can see exactly what we've done.

We will ask you to validate our work with regard to:

- Bib record counts, filtered and unfiltered
- Bib records filtered out by cause
- Circulation / internal use counts
- Title/item counts by location
- Title/item counts by subject
- Transaction dates

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
- Mixed Multi-Volume Sets*
- 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.

THE QUESTIONNAIRE

of in-scope monographs

(bib records)

OCLC holding symbols(s)

date of last migration

OPAC URLs

location of bib record numbers and OCLC numbers

multiple OCLC numbers

classification schemas !!!

call numbers !!!!!!!!!

call number type codes

copy numbers and enumeration

WorldCat holdings

usage data

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

OCLC

cataloging anomalies

local practices

**Tell us EVERYTHING
you think we should know!**

I'll get back in touch if questions arise.

QUESTIONS?

OCLC

For now, focus on:

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

Send all these things to Ruth Fischer

Sustainable Collections

fischerr@oclc.org

603-746-5991

Project Timeline: 2015

Task	Timeframe
Kick-off Meeting: Organizational	June
Kick-off Meetings: Collection Analysis	July
Confirm Project Leadership/Communication	June/July
Confirm Scope and Specifications	July/August
Category 1 Libraries: Data Questionnaire	July-September
Category 1 Libraries: Supply Bib, Item, Circ Data	July-October
Category 2 Libraries: Modify Existing Data	July-September
Category 3 Libraries: Configure Retention List for Matching	July-August
Conduct Validation Study #1	October-April
Category 1 & 2 Libraries: WorldCat validation & matching; Hathi Trust matching; match against Cat retention lists	November-December

We are here

Contact Info

oc.lc/scs

fischerr@oclc.org

luggr@oclc.org

https://www.blc.org/east-project

sstearns@blc.org

lizannepayne03@gmail.com