Community Update

Shared Print Data Infrastructure Project (AKA ‘Mellon Project’)

Matt Barnes
Director, SCS
OCLC

Amy Wood
Head of Technical Services
CRL

Marie Waltz
Special Projects Manager
CRL
Topics

1. Progress Report
2. Key Upcoming Meetings
3. Workflow Overview
4. Updates on Important Issues
5. PAPR Updates
6. Q&A / Discussion
Progress Report
Project Status

Overall Status: *On schedule*

1. Enable batch registration of serials
2. Expand workflow and metadata support
3. PAPR enhancements & syncing
4. API to retrieve commitment data
5. Record Manager & Connexion

- Development nearing completion
- Preliminary testing anticipated in August timeframe
- Multiple Formats & ICONs to Differentiate Formats
- Holdings comparison page
- Development started
- Development complete
Key Upcoming Meetings
Advisory Group Meetings  (Dates TBD)

Merge / match logic
- Latest merge / match summary to be sent in advance
- Request: Bring cataloger(s) / metadata specialists, if possible

Shared print API
- Second survey to be sent in advance
- Request: Bring developer(s), if possible

Second symbol cleanup strategy
- Questionnaire to be sent in advance
- Request: Bring cataloger(s) / metadata specialists, if possible

Change control prioritization
Workflow:

Direct Submission by Library
High Level Concept

Step 1: Create Collection Profile

New: Establish default values 1x per shared print program

Step 2: Upload Commitments

New: Use abbreviated for monographs or full csv file

<table>
<thead>
<tr>
<th>OCLC_Number</th>
<th>LSN</th>
<th>Barcode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1076037</td>
<td>.b1188423x</td>
<td>39343024935925</td>
</tr>
<tr>
<td>20629876</td>
<td>.b19695512</td>
<td>39343016183922</td>
</tr>
<tr>
<td>34887420</td>
<td>.b1981799x</td>
<td>39343016435132</td>
</tr>
<tr>
<td>2480318</td>
<td>.b25907645</td>
<td>39343006232127</td>
</tr>
<tr>
<td>5219401</td>
<td>.b10878488</td>
<td>39343016051327</td>
</tr>
<tr>
<td>11978921</td>
<td>.b19520967</td>
<td>mq1211696</td>
</tr>
<tr>
<td>62180478</td>
<td>.b3014940x</td>
<td>39343022902703</td>
</tr>
<tr>
<td>8403649</td>
<td>.b38889833</td>
<td>39343024911256</td>
</tr>
<tr>
<td>36438790</td>
<td>.b14237702</td>
<td>39343015453474</td>
</tr>
<tr>
<td>20465008</td>
<td>.b21515438</td>
<td>39343007280232</td>
</tr>
<tr>
<td>4090162</td>
<td>.b21493601</td>
<td>39343006913148</td>
</tr>
<tr>
<td>28222328</td>
<td>.b36047041</td>
<td>39343020912779</td>
</tr>
<tr>
<td>63222047</td>
<td>.b31299908</td>
<td>39343024195496</td>
</tr>
</tbody>
</table>
Detailed Workflow

**Part 1:** Establish a collection/create profile in Collection Manager

- Create profile/collection in Collection Manager for your institution
  - *One time process*

**Part 2:** Submitting Commitments

- Generate CSV file
- Make any updates to CSV file as needed
  - *OCN, LSN required*
- Receive exception/summary report
- Fix any errors from the report
- Re-submit CSV w/ corrected records with errors

**Inputs:**
- Create profile/collection in Collection Manager for your institution
  - *One time process*
- Generate CSV file
- Make any updates to CSV file as needed
  - *OCN, LSN required*
- Receive exception/summary report
- Fix any errors from the report

**Outputs:**
- Commitments registered with Shared Print flag
- Data available in WorldCat and synchronized to PAPR in separate process
3 Column .csv (Single-Part Monographs)

<table>
<thead>
<tr>
<th>OCLC_Number</th>
<th>LSN</th>
<th>Barcode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1076037</td>
<td>.b1188423x</td>
<td>39343024935925</td>
</tr>
<tr>
<td>20629876</td>
<td>.b19695512</td>
<td>39343016183922</td>
</tr>
<tr>
<td>34887420</td>
<td>.b1981799x</td>
<td>39343016435132</td>
</tr>
<tr>
<td>2480318</td>
<td>.b25907645</td>
<td>39343006232127</td>
</tr>
<tr>
<td>5219401</td>
<td>.b10878488</td>
<td>39343016051327</td>
</tr>
<tr>
<td>11978921</td>
<td>.b19520967</td>
<td>mq1211696</td>
</tr>
<tr>
<td>62180478</td>
<td>.b3014940x</td>
<td>39343022902703</td>
</tr>
<tr>
<td>8403649</td>
<td>.b38889833</td>
<td>39343024911256</td>
</tr>
<tr>
<td>36438790</td>
<td>.b14237702</td>
<td>39343015453474</td>
</tr>
<tr>
<td>20465008</td>
<td>.b21515438</td>
<td>39343007280232</td>
</tr>
<tr>
<td>4090162</td>
<td>.b21493601</td>
<td>39343006913148</td>
</tr>
<tr>
<td>28222328</td>
<td>.b36047041</td>
<td>39343020912779</td>
</tr>
<tr>
<td>63222047</td>
<td>.b31299908</td>
<td>39343024195496</td>
</tr>
</tbody>
</table>

File Naming Convention

<collectionID>.<OCLCsymbol>.sharedprint_<optionaltext>_YYYYMMDD>.<optionalfile#>.csv

Example:
• 1028071.ndclv.shareprint_20170731.csv

Future:
Potential columns added here to support serials information
### 13 Column .csv (All formats)

<table>
<thead>
<tr>
<th>OCLC_Number</th>
<th>LSN</th>
<th>Barcode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1076037</td>
<td>.b1188423x</td>
<td>39343024935925</td>
</tr>
<tr>
<td>20629876</td>
<td>.b19695512</td>
<td>39343016183922</td>
</tr>
<tr>
<td>34887420</td>
<td>.b1981799x</td>
<td>39343016435132</td>
</tr>
<tr>
<td>2480318</td>
<td>.b25907645</td>
<td>3934306232127</td>
</tr>
<tr>
<td>5219401</td>
<td>.b10878488</td>
<td>39343016051327</td>
</tr>
<tr>
<td>11978921</td>
<td>.b19520967</td>
<td>mq1211696</td>
</tr>
<tr>
<td>62180478</td>
<td>.b3014940x</td>
<td>39343022902703</td>
</tr>
<tr>
<td>8403649</td>
<td>.b38889833</td>
<td>39343024911256</td>
</tr>
<tr>
<td>36438790</td>
<td>.b14237702</td>
<td>39343015453474</td>
</tr>
<tr>
<td>20465008</td>
<td>.b21515438</td>
<td>39343007280232</td>
</tr>
<tr>
<td>4090162</td>
<td>.b21493601</td>
<td>3934306913148</td>
</tr>
<tr>
<td>28222328</td>
<td>.b36047041</td>
<td>39343020912779</td>
</tr>
<tr>
<td>63222047</td>
<td>.b31299908</td>
<td>39343024195496</td>
</tr>
</tbody>
</table>

Other columns added for data creation, overriding values, or group registration:
- Institution Symbol_852$a*
- HoldingLibrary_852$b
- Collection_ID*
- ActionNote_583$a*
- ActionDate_583$c
- ExpirationDate_583$d
- MethodofAction_583$i
- Status_583$l
- PublicNote_583$z
- HoldingSummary_583$3
Workflow:
Submission by Agent
High Level Concept

Step 1: Create Collection Profile

New: Establish default values 1x per shared print program

Step 2: Send data to agent

New: Agent will upload group data to WorldCat
Detailed Workflow

**Institution/Program actions:**

- Institution(s) creates collection(s) in Collection Manager
  
  *One time process*

- Institution(s) passes collection IDs, institution symbol, w/ other required fields (OCN, LSN) onto “admin”

  NEW: Include authorization notes in collection comments field

**“Admin” actions:**

- Create “agent” collection/profile in Collection Manager
  
  *One time process*

- Ensure collections are created for each institution you intend to register for

- Generate CSV file

- Make any updates to CSV file as needed & submit

- Receive exception/summary report
  
  Institution & group level

- Fix any errors from the report

  Re-submit CSV w/ corrected records with errors

**Outputs:**

- Commitments registered with Shared Print flag

- Data available in WorldCat and synchronized to PAPR in separate process

#ALAAC19
CRL Workflow: Behind the Scenes

Institution creates a collection in Collection Manager

Will the program store CollectionIDs?

Exploring option of OCLC providing collection IDs to CRL

What is a reasonable schedule of CollectionID updates from OCLC?

CRL registers CollectionIDs in PAPR

Institution processes collection metadata

Is change desired or needed?

Institution submits it to program manager as usual.

Is change desired or needed?

Program manager transfers the metadata to CRL

CRL validates data.

Validation reports sent to program manager

Good records are loaded into PAPR

Program manager sends validation reports to institutions

Original data set archived

CRL receives reports

CRL reviews reports to determine if PAPR or institution is origin of problem.

Fix PAPR data

Report problem to Program Manager

CRL prepares load file for ingest into WorldCat

CRL uploads CSV file

Exception and summary reports generated.
Updates on Important Issues
Change Control Process

Step 1: Create full list of requested changes & new features
Step 2: Advisory Group prioritization
Step 3: Address dev based on priority, funding, resources

Examples:
1. Program name authority control
2. Changes to merge/match rules
3. Updates/edits to existing commitments
4. Multiple $f in 13 column csv
5. Additions to discovery interfaces
6. Additional analytics/reports
Merge / Match

Rules that govern how new data interacts with existing data in WorldCat. Example questions from Advisory Group:

Q: How should I handle split locations/split holdings for the same record so that they are not overridden?

A: If separate holding locations (852$b) are provided (in two separate lines), two separate 583’s will be created and they will reside in two separate LHRs. Data will not be overridden.

Q: I have concerns about data being dropped, overridden when not intended, or duplicated when not intended. How do you plan to approach this?

A: The process is designed to never drop data from a WorldCat LHR. Matching and resolution against a pre-existing WorldCat LHR can only result in information being added to the record. Efforts are made to ensure that duplicated fields are never added to a pre-existing WorldCat LHR. We will be providing detail on the merge/match logic soon.
Second Symbol Transition/Cleanup

We intend to transition away from legacy shared print symbols (second symbols) in WorldCat.

— 92 institutions
— 2,225,765 holdings set

Example questions:
  – How are second symbols being used? Is the data used locally?
  – What data stored needs to be maintained in the new registration process?
  – Do important reports used by libraries/groups rely on this data?
Discovery – Record Manager

New shared print group search (title level)

Bulk search for commitments held locally
Discovery – Connexxion

Shared print group searching

Shared print LHR Detail in Connexion Browser
API

Follow up need on first survey to better understand:

1. Group and individual library uses cases
2. Correct search parameters are in scope
3. Ability to handle XML response format
4. Need for mass download of retention data
PAPR Updates
## Multiple Formats & ICONs to Differentiate Formats

### Search Results

1 results for ISSN: 0955-2359

<table>
<thead>
<tr>
<th>Title</th>
<th>Publisher</th>
<th>ISSN</th>
<th>Program (Institution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20th century British history</td>
<td>Oxford University Press</td>
<td>0955-2359</td>
<td>WEST (STFSP) EAST (BXM) EAST (BOS) EAST (MBB) WEST (SL3SP) PORTICO CLOCKSS</td>
</tr>
</tbody>
</table>

1 results for ISSN: 0955-2359
# Holdings Comparison Page

**Atmósfera.**

- **Publisher:** Centro de Ciencias de la Atmósfera de la Universidad Nacional Autónoma de México
- **Print ISSN:** 0187-6236
- **Electronic ISSN:**
- **OCLC Number:** 17992364
- **Subject:** Atmosphere Periodicals.
- **Subject code:**

## Print Holdings

<table>
<thead>
<tr>
<th>Archiving Institution</th>
<th>Program</th>
<th>Holdings Description</th>
<th>Retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center for Research Libraries Serials (CRL)</td>
<td>CRL-Serials</td>
<td>v8 no.1 (Jan 5, 1985)</td>
<td></td>
</tr>
<tr>
<td>University of California, San Diego Shared Print in Place (CUSSP)</td>
<td>WEST (WEST (Bronze))</td>
<td>v1 (2020-2021)</td>
<td>December 31, 2016</td>
</tr>
</tbody>
</table>

## Digital Holdings

<table>
<thead>
<tr>
<th>Program</th>
<th>Holdings Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOCKSS</td>
<td>(2014) no.27- (2017) no.30</td>
</tr>
<tr>
<td>PORTICO</td>
<td>2015-2017</td>
</tr>
</tbody>
</table>
Discussion
thank you

Matt Barnes
Director, SCS
barnesm@oclc.org

Amy Wood
Head of Technical Services
wood@crl.edu

Marie Waltz
Special Projects Manager
waltz@crl.edu