Partnership for Shared Book Collections

Best Practices for Collection Analysis

Introduction
Library staff in the early stages of exploring a Shared Print Program (SPP), whether they are starting a new program, joining an existing program, or performing a program review will conduct a collection analysis. This process will help them determine whether their analyzed collections show potential benefit from a shared print program. If so, collection analysis will inform and shape their program or how their library collection will fit into an existing program.

A collection analysis will look at overlap and divergence across multiple library collections. It may also highlight scarce titles in need of protection. An SPP will then use the analysis to determine how to distribute retention commitments across libraries by assigning specific items to be retained at each library.

Note: Since collection analysis is generally the initial step in beginning or joining a Shared Print Program and there are a variety of approaches that libraries and programs can take, these are recommendations and considerations rather than good, better, best, or aspirational classes of practice.

Best Practices for Collection Analysis
Shared Print collection analysis projects require significant time and effort, with several areas that merit consideration throughout the planning and implementation process.

Program/Library Considerations
Before commencing a shared print collection analysis, the Shared Print Program should:

- Define **Project Scope**
    - Goals of the analysis, e.g., identify overlaps, gaps, or scarcity; subject (by LC Call Number) strengths; publication dates; geographic area/local subject interests; geographic area of the holding library.
    - Final outputs desired, e.g., description report, spreadsheets, record sets, or platform for tailored analysis.
    - Scope of the records to be analyzed (See the Shared Print Toolkit discussion of scope); determine if the set of records to be analyzed will be limited to this scope at the outset, or will the collection analysis tool need to be able to filter out out-of-scope materials?
    - Other data to include, e.g., circulation stats (highly circulated titles may need more than one copy retained), HathiTrust or other digital collections, holdings at peer libraries, OCLC data (uniqueness/scarcity), age/publication date, LC Call

Numbers (subject strengths), cost of items, existing retention commitments either in the target program or in other SP programs of interest.

- Identify **Project Requirements**
    - Location of records being analyzed and export process to another system, if applicable
    - Personnel responsible for gathering, preparing, processing, and analyzing data
    - Systems that match needs and budgets

- Consider **Project Obstacles**
    - Barriers to participation, including cost of the collection analysis tool, staff time required to participate, difficulty using the chosen system (i.e., does it require special training, record export and cleanup, dedicated staff?) Determine if barriers can be overcome or lowered.
    - Commitment to ongoing staff time as collection analysis and shared print are not "one and done" activities. Collections are dynamic, and retention commitments need review and care. (E.g. materials that become lost/damaged, out-of-scope, ILL considerations and/or requirements of the SSP, and onboarding new staff.)
    - Data quality of records to be analyzed. Are there any issues with the records that will be analyzed, e.g., brief records, records from legacy systems such as RLIN, records that may have been created or changed through system migrations? Are there data points that will be inconsistent across participants, e.g. if circulation data is included, was in house circulation added in some instances? How many years of circulation data are available? Does it matter to your analysis?

Selecting a Collection Analysis Tool
While selecting a collection analysis tool, the Shared Print Program should:
- Identify **Record Quality**
    - What are the potential issues with the records being analyzed, e.g., lacking OCLC number for systems that use that as a match point, brief records, bound with materials, analyzed series
    - How does the collection analysis tool handle these issues?
- Assess **Results Format Utility**
    - Will the tool provide actionable information and reports?
    - For external tools not part of a local ILS/LSP, will it give you the information you need to reflect the results in the local systems?
- Determine **Retention Allocation Functionality**
    - Can the tool allocate retention copies equitably across the members? Can it assign them based on the libraries preferences, e.g. subject strengths?

Using a Collection Analysis Tool
When using a collection analysis tool, the Shared Print Program should:
- Verify **Accuracy of analysis**

- How will you determine if the collection analysis tool is getting at what you want? Plan for iteration to hone the analysis to desired materials. This includes planning extra staff time to review results. Particular attention should be paid to the matching algorithm of the tool and if it meets your programs/libraries needs and the scope of the project, e.g., do subsequent editions of a title count as overlap or distinct entities? Check for non-print formats or other materials that may have been included by mistake.
- Do the results align with the goals stated in the program MOU?

It may also be worth reviewing the Common Pitfalls of shared print programs when embarking on a collection analysis.

Last Updated: September 2022