

Assessing Bibliographic Inaccuracy as a Contributing Factor for Unintended Loss in Shared Print Monograph Programs

Helen N. Levenson, Sara Amato, Ian Bogus, Fern E. Brody, Mary Miller, and Jacob Nadal

Shared print programs are helping their member libraries right-size their collections. As they do, there are concerns about the adverse impact of bibliographic inaccuracies. This paper studies bibliographic record inaccuracies and the resulting frequency of mismatches between an item owned and the record representing ownership. Through analysis of interlibrary loan (ILL) survey data, the authors found an overall low rate of bibliographic inaccuracy, and that inaccuracies rarely prevented library staff from verifying the ownership of the item. The authors review how this finding contributes to confidence in library holding records, which is necessary for evaluating the optimal number of copies in shared print networks and for minimizing the risk in withdrawal decisions.

Introduction

As the shared print community continues to grow and mature it has been finding creative ways for libraries to work together, not only to make resources more easily accessible, but also to ensure that the printed word is adequately secured for the future. During this growth, practitioners often ask about how many copies of a title need to be retained to ensure long-term access. While it appears to be a simple question, it inevitably becomes more complicated as one considers the variables that contribute to the determination of the minimal number of copies required. The types of risks and variables involved in being able to make this determination were explored in the C&RL paper “Everything Not Saved Will Be Lost”¹ in which Maiorana et al. stated, “The variables are known, but we can only identify the values of those variables through study and research.”² The Partnership for Shared Book Collections³ (AKA The Part-

* Helen N. Levenson is Associate Professor, Collection Development Librarian at University Libraries, Oakland University, email: hlevenson@oakland.edu; Sara Amato is Program Manager at Eastern Academic Scholars' Trust, email: samato@eastlibraries.org; Ian Bogus is Executive Director at Research Collections and Preservation Consortium (ReCAP), email: ibogus@princeton.edu; Fern E. Brody is Associate University Librarian for Collections and Technical Services at University Library System, University of Pittsburgh, email: feb@pitt.edu; Mary Miller is Director of Collection Management and Preservation, University of Minnesota Libraries - Twin Cities Campus, email: memiller@umn.edu; Jacob Nadal is President of the Center for Research Libraries, email: president@crl.edu. ©2024 Helen N. Levenson, Sara Amato, Ian Bogus, Fern E. Brody, Mary Miller, and Jacob Nadal, Attribution-NonCommercial (<https://creativecommons.org/licenses/by-nc/4.0/>) CC BY-NC.

nership), a federation of 17 shared print monograph programs, convened a Risk Research Working Group* to further explore the issue of optimal number of copies. In doing so, the group has taken up the work of delving deeper into the unanswered questions raised in the Maiorana article, including the rate of inaccurate bibliographic records. In our study, we explored two topics arising from bibliographic inaccuracies: first, the frequency of occurrences of bibliographic mismatches and, secondly, how detrimental bibliographic inaccuracy might be to long-term retention in the context of shared print programs.

Bibliographic inaccuracy can be perceived to be a significant risk factor in shared print management. Librarians often use the number of copies listed in OCLC as a stand-in not only to identify how many copies are held in other libraries, but also as the basis for deciding to withdraw their own copies. This decision assumes that the number reported in OCLC is above their critical threshold for the number of copies that need to be retained. When evaluating how many copies need to be retained, print archiving programs also need to have confidence that bibliographic records match the objects they are supposed to represent. In meetings of the Partnership and its Risk Research Working Group, participants frequently expressed concern that there were significant problems related to bibliographic inaccuracy, especially with details such as edition statements. These may be attributable, for example, to matching algorithms used in retrospective conversion projects that transformed card catalogs to electronic records in bulk.

The risks inherent in bibliographic inaccuracy can play out in two ways. First, catalogers may have created a new record unnecessarily leading to a proliferation of separate records for items that are, in reality, the same work or edition. Each of these records would then have a smaller number of holdings, which would create the appearance of numerous, scarcely held works. This, in turn, would lead libraries to determine that particular title as being at risk and, thus, as a candidate for retention. This false scarcity is a bibliographic problem, but not a preservation risk since a larger number of copies are retained, albeit accidentally. However, it is a concern for the overarching goals of shared print archiving because it works against the efficiency and accuracy of these programs.

Second, and of more concern to shared print preservation programs, catalogers may erroneously attach holdings to an incorrect record. If different works or editions are erroneously attached to a given record bibliographic inaccuracies may cause an overestimation of the number of copies in existence. Such inaccuracies between the bibliographic record and the actual physical item can lead to libraries unintentionally dropping the number of retained copies below the acceptable thresholds.

Attaining complete certainty about the level of accuracy across the hundreds of millions of records in OCLC is, in practical terms, an impossible goal. Instead, our research sought to determine if the risks posed by bibliographic inaccuracy could be managed in the course of normal shared print activities, or if they presented a risk of sufficient scale and complexity that it might endanger the enterprise. Our research does not, therefore, seek to determine the absolute level of bibliographic inaccuracy across the collective collection. Instead, it evaluates whether bibliographic inaccuracy is frequent enough or severe enough to compromise shared print efforts. Our research used interlibrary loan data to assess where bibliographic

* The Partnership for Shared Book Collections formed in 2019 to support and promote shared print programs, and formed a Risk Research Working Group to help the Partnership make informed recommendations on creating responsible commitments. <https://sharedprint.org/>

inaccuracy falls on a spectrum of risk. We chose to analyze interlibrary loan data for our research for several reasons:

1. Data on fulfillment rates of ILL are already collected by many libraries, so our study could benefit from a large volume of data without requiring a high level of additional effort from participants.
2. ILL is a core library function. Therefore, participants could gain an immediate benefit from participation not only through assessing their own data, but also via comparison and benchmarking opportunities with other participants, regardless of the outcomes of the research project.
3. Evaluation of ILL requests against the item in hand resembles the activity that shared print participants perform when evaluating items for transfer to an archive facility or for withdrawal, thus making this a good proxy for the operational risk that needed to be evaluated.

Literature Review

The number of collectively owned copies is an integral variable in the calculation of determining the necessary number of copies to retain within both individual library collections and larger collective collections, as part of establishing sufficient and dependable access in conjunction with necessary preservation strategies.⁴ The frequency of this topic in the literature suggests that the shared print community fully recognizes the importance of being able to make this determination. As shared print programs establish retention scenarios among participants, they strive to guarantee a minimum number of retained copies, typically dependent on criteria for both widely held titles, and scarce or unique holdings.⁵ As one of the fundamental goals of ensuring continued access to print resources, overall sufficient numbers of copies need to be retained among the partners of shared print programs.⁶ However, there is a prevailing element of uncertainty as to what, exactly, is a sufficient number—or optimal number—of copies that must be retained.⁷ This uncertainty regarding what quantity would constitute the optimal number of retained copies was recognized in the early phases of the development of shared print programs⁸ and has remained as a significant concern.

Several studies have emphasized that reliable data is needed to determine how many copies are necessary to retain and preserve to ensure enduring access and overall responsible stewardship of the print scholarly record.⁹ Areas for which there can be insufficient data include: reliable discovery of print retention commitments or digital surrogates; levels of validation conducted verifying item ownership; types of storage environments used by shared print partners; and physical condition of the monographs partners have committed to retain. Although not commonly explored in the library literature, the margin of error in the bibliographic and holding records of shared print program participants is an essential element of the data required for proper retention quantity determination.¹⁰

Do We Own What We Think We Own?

Accurate bibliographic record data regarding what other libraries own is critical for libraries to have confidence about their own withdrawal decisions.¹¹ Teper included bibliographic inaccuracy, among other risk factors, as part of her study of seemingly identical monographs among a survey sample of 625 books. Teper found a 3.4% cataloging error rate exclusively due to edition inaccuracies through the use of an incorrect OCLC record.¹² Teper's findings may

skew higher for errors in editions because her study focused on books published between 1851 and 1922. During this earlier time, cataloging rules and practices were not standardized; efforts to determine what constituted quality cataloging and benchmarks to measure quality cataloging did not begin to emerge until the 1990's.¹³ Teper found that factors other than bibliographic inaccuracy—such as item degradation—were more likely to be responsible for risk compromising access. It is important to note that, in contrast to our study, Teper's study of catalog record inaccuracy was based on examining the local catalog record alone, compared to the OCLC record. Our study furthers Teper's research by not limiting the publication year of the monographs studied and by directly comparing the item in hand to the corresponding catalog record. This methodology resulted in a bibliographic inaccuracy test more closely related to the processes used in validation and resource sharing request fulfillments, both germane activities to achieving important shared print program goals.

Accuracy of Bibliographic Records

The bibliographic database is core to any library's operations. Petrucciani stated that libraries are the only social institutions that are responsible for the "control, organization, communication, and preservation of information about the published output of human knowledge and expression."¹⁴ Although Nero & He noted that academic libraries' cataloging departments are responsible for taking "ownership of the bibliographic integrity of the collection" in support of bibliographic record inspection,¹⁵ it is unrealistic and unreasonable to think human errors do not occur. Additionally, there can be variances of cataloging records based on individualized library cataloging practices. However, most human errors and library cataloging variances do not hinder identification and retrieval. Therefore, the necessary quality and quantity of monograph bibliographic access points for a book in question to be accurately identified is maintained.¹⁶ Bade noted that it is not necessary to have perfect catalog records but ones that provide sufficiently accurate identification information.¹⁷

The study of record errors conducted by Michaels and Neel found that their processes for relocating items to a closed stack location or facility functioned very effectively as verification of ownership. In their study, the bibliographic records were verified against each individually handled item resulting in "a 100 percent retrieval rate" for over 640,000 requested items over a 16 year period.¹⁸ However, Michaels and Neel did find record error occurrences for items located in open stacks, although they applied a much broader interpretation for such errors than what we used for this ILL-based bibliographic inaccuracy study. For example, Michaels and Neel included circulation status errors as well as catalog record mismatches to physical items in their study, thus resulting in an overall 2.16% error rate. Taking into account only cataloging record errors, which included labeling and barcode errors, Michaels and Neel found a 1.42% error rate. Of all the various error types that Michaels and Neel studied, they acknowledged that not all of the issues "encountered would have equal implications for shared print initiatives."¹⁹ In some cases, these record errors resulted in an item being more widely held than previously known, which leads to over-retention and reduces the risk of retaining a lower than acceptable quantity of a monograph title.

Bibliographic inaccuracy is but one of several variables, many others being more impactful, in determining and mitigating risk within shared print programs.²⁰ A recent model of how to determine the optimal number of copies includes a variable for bibliographic inaccuracy²¹ and requires a default value for this variable to be entered. The dearth of published research

on bibliographic inaccuracy applicable to the shared print endeavor prompted us to undertake our own study to more clearly understand which instances of bibliographic error would lead to an incorrect conclusion of a library's ownership of an individual monograph. We intended that this research would provide a reasoned quantitative number to apply in the risk model tool for shared print programs mentioned above.

Methods

Through this study, we gathered data to further our understanding of two areas:

1. The frequency of occurrence, or percentage, of bibliographic mismatches for monographs.
2. Whether bibliographic mismatches, or inaccuracy, presents a significant risk in a collective collection, specifically the risk that copies may be incorrectly withdrawn.

To the best of our knowledge, interlibrary loan transactions have not previously been used to measure or study bibliographic inaccuracy. As noted in the introduction, we chose to use ILL data for this project for several related reasons:

- Ease of data collection: In the course of daily work, resource sharing staff look closely at the item in hand and compare it to the request. Through many resource sharing systems, like the service run by OCLC, the requests are normalized so that when staff are comparing books to the request, they are checking it against a record. Because we requested limited data that could be gathered and recorded relatively easily while carrying out existing daily library functions, the study did not require a high level of additional effort from participants.
- Data quantity: Because data on fulfillment rates of ILL are already collected by many libraries, our study benefited from including a large volume of data collected from diverse and geographically distributed participating institutions in a relatively short period of time.
- Applicability to shared print program management: The evaluation of ILL requests against the item in hand is similar to the activity that shared print program participants perform when evaluating items for transfer to an archive facility or for withdrawal. Because of this, we were able to leverage this data to assess risk of bibliographic inaccuracy and to apply it in a shared print context.

In short, by leveraging existing resource sharing workflows and keeping data collection efforts to a minimum, it was possible to maximize participation, to create a large data set of 29,630 items within a relatively short period of time, and to use this pool of data as a litmus test for the severity of mismatches.

The intention of our study's scope is to capture data related to issues that affect identification of duplicate monograph copies, rather than overall cataloging accuracy. For example, it is less important that subject headings be correctly listed or author names normalized. The deciding factor in our study was whether the difference in bibliographic data either caused confusion about the copy in hand or prevented a successful retrieval for fulfillment of a loan to a user. In other words, in practice, do the book and the information in the record sufficiently match?

Data Collection

Prior to beginning our research, we drafted a simple data collection form. We used this form to solicit feedback from eight resource sharing departments; several institutions also undertook

a three-day pilot to evaluate the proposed process. We initially considered an approach that required participants to capture categories of errors and source of request (such as WorldCat or a union catalog). After receiving feedback, we narrowed the scope to focus on capturing data that was most relevant to a bad retention outcome.

We then distributed a call for interest in participation to the Print Archive Network and ILL-L listservs, in addition to direct messages to eight libraries and consortia already engaged in shared print efforts. Respondents received an email with a brief project description and an invitation to a webinar on October 1, 2019. They were invited to submit suggestions for refining the study and questions at or in advance of the video conference, which also addressed the intent, process, and expectations of the study.

Participants then completed a brief registration form that included institution name, contact name, contact email address, and dates of study participation. They were asked to gather data for four to six weeks, completing their data collection by November 28, 2019. Fifteen libraries participated.

Participating libraries were asked to use a paper form to record two categories of mismatches: major (i.e., title or author, which may indicate an entirely different FRBR Work) and minor (i.e., edition, language, or form, which indicates a different FRBR Expression). Only requests that used full bibliographic records were included. Participants were also asked to record other issues they discovered along with a description. For mismatches, photos of the title and verso pages were encouraged, although not required (see Appendix for the description of the project provided to the participants, instructions to the participating libraries, and the form to use to record results).

Once all data was submitted, the research team reviewed each dataset for completeness and compiled the results. It bears emphasis that this study focuses on the ability of resource sharing staff to fulfill requests based on bibliographic data they receive and uses that information as a proxy for what shared print programs look for when comparing records. This is not a study of cataloging practice directly, but of the effects created by those practices and the way that environment shapes shared print retention decision-making.

Results

Out of the fifteen libraries that submitted datasets, thirteen were complete, valid, and comparable. From the thirteen sets of valid data, there were a total of 29,630 items reviewed during the study period.

Data from two libraries were not included in the results. One library's data contained useful information on matching accuracy but could not be used due to incomplete counts of total ILL transactions. It is impossible to calculate the percentage of mismatches between the OCLC records and the physical items without the total count of items analyzed. Another library captured data as books were sent to storage and not through their ILL process. While the data itself is valuable, it cannot be compared with data collected through ILL because the demographics of the books involved may be different from those requested through ILL transactions. Therefore, this library's data was not used in this study.

The initial interest of the research team was determining a rate of bibliographic inaccuracy. By all counts, bibliographic inaccuracies for items requested through ILL were low. Nine libraries (69%) out of a total of thirteen, reported no major or minor mismatches. These nine libraries reviewed 23,224 (78%), the majority of items in the study overall. The remaining

four libraries found 15 (0.05%) major and 14 minor (0.05%) errors, accounting for 0.10% of all items reviewed. It is possible that the libraries that found errors had more thorough review practices. To account for the worst-case scenario, we also calculated the error rates only within datasets that found at least one error. For the libraries that did find errors, major and minor errors accounted for less than half of one percent (0.45%) of the items that they reviewed.

TABLE 1			
Results of Comparable Data from Participating Libraries			
	# Errors	% of All Items (29,630)	% From Libs Reporting Errors (6,406)
Major Errors	15	0.05%	0.23%
Minor Errors	14	0.05%	0.22%
Total Errors	29	0.10%	0.45%

The four libraries that reported errors included varying degrees of details in their submissions. Our participating libraries had some differences of opinion as to what to include as a major versus a minor error; however examining the error descriptions show that the errors themselves rarely cause confusion about the item in hand. Even title or author differences may be more accurately described as variations rather than wholly different titles or authors.

Examples of Mismatches

Library #4 reported the following:

- Major mismatches
 - One error appears to have a variant title in the local 245 field. The OCLC record reports “The man who invented instant replay” in the 245 with an additional title in the 246 of “Instant replay : the day that changed sports forever.” The local record has “Instant replay : the day that changed sports forever” in the 245 without a 246 at all.
 - There was an item with the opposite problem where OCLC’s record appears to have a minor typo. The OCLC 245 contains “All hall to the archpriest” rather than “All hail to the archpriest” which is in the local record and the title page of the book.
 - Similarly, a different book appears to have an omitted word in the record, or a standardized title. Both the OCLC record and local record list “Prima che te lo dicano gli altri” (English translation: “Before the others tell you”) in the 245 but the title page is written as “Prima che te lo dicano altri” (English translation: “Before others tell you”).
 - One item appears to have a standardized author in the record. The 100 field is listed as “Hawdon, Sarah Elizabeth, 1851-1921” but the title page lists “New Zealander.”
 - The last major error reported was a photocopied book that should have been noted in the record. While the title is correct, it understandably may have been confused for an original.
- Minor mismatches
 - Dates were a common issue for minor mismatches and were involved with five minor mismatches.
 - Four books were listed to have a variation of one or two years in the

date field.

- » One of those four listed New York as the place of publication instead of Cambridge.
- For the fifth, 2018 was listed in OCLC, 2014 in the local catalog, but the book states that the edition was first published in 2015. The library reported that the record was “pulled in 2018 for copyright information but [they] can’t find any proof that a 2018 edition exists.”

Library #7 reported two major mismatches.

- Major mismatches
 - The first was that the OCLC record appears to have two ISBN numbers (905349040x and 9789053490402) though the book lists only one which has eleven digits instead of the required ten (90-53940-040-x). This ISBN has a typo, potentially confusing the staff member.
 - The other mismatch appears to be an author normalization. The record states “Johnson, A.B.” as an author although in the book it is written as “Burt Johnson.” It was submitted as a major mismatch because it was viewed as listing the first name of the author incorrectly.
- Minor mismatches
 - Both minor mismatches appear to be minor normalizations using initials for the first and middle names. The first case is A.V. Seaton in the record but Tony Seaton in the book (one of five editors). The other is written as Gurevich, Aron IAKovlevich in the record but A.J. Gurevich in the book.

Library #2 reported the following:

- Major mismatches
 - Seven major mismatches with no more detail than five were for title and three for author errors.
- Minor mismatches
 - They also reported five minor mismatches, two for edition and three for year errors.

Library #17 reported the following:

- Major mismatches
 - Responding library reported an incorrect ISBN in the record as a major mismatch.
- Minor mismatches
 - One item was reported as having an incorrect edition but the respondent wondered if it was due to an internal logic error with their Information Delivery Services instance.

Discussion

In analyzing the data sets received, we were surprised at how infrequently errors that would affect the shared print endeavor were found. Only four out of the thirteen libraries with valid results found either major or minor errors in their records. Additionally, among these four libraries they only found 29 major or minor errors. One can infer that the occurrence of such errors is rather low and the specific data supports this; the combined errors only accounted for 0.10% of all the items reviewed by all the participating libraries. The number of errors found is too small to break down further with any accuracy.

Of the four libraries that reported mismatches, the total number of items that were processed by each library during the survey period varied widely, therefore the percentages of

major and minor errors had corresponding variances. The number of items reviewed as part of the ILL process was, from lowest to highest, a total of 220 (library #7), 1257 (library #4), 1873 (library #13), and finally 3056 (library #17). By percentage, those libraries had error rates of 2.3% (library #7), 0.8% (library #4), 0.6% (library #13), and finally <.1% (library #17). However, the significant comparison is against the larger group as a whole, comprising a total of 29,630 records reviewed. This is in keeping with the evaluation of the occurrence of bibliographic inaccuracy, as it exists within the collective collection in the context of shared print collections.

The impetus of this study was to assess how often a bibliographic record would point to a different object than the one described in the record, causing confusion about duplicates held across institutions. Correctly identifying the number of duplicates is important for libraries that are trying to assess the risk of withdrawing copies, as well as for the shared print community because they assess the minimum number of copies that need retention commitments. Interpreting the results through the lens of this intended purpose may shed some more light and further refine the findings.

Only four of the thirteen libraries with viable results found mismatches. At first, it may appear odd that errors were concentrated in 30% of the participants; however, examining the errors found more closely may suggest an explanation. Two of the libraries, accounting for fifteen of the errors, reported details for their findings. Of these fifteen errors, eight of the records clearly are describing the item in hand without potentially pointing to another item or edition. Four of the remaining eight errors were reported for date discrepancies, all of which were within a year or two of the book in hand and none of which showed edition variations. These, more likely than not, are because of differences in copyright, publication, and distribution date rather than actually different editions. Overall, this leaves three errors out of the original fifteen errors that lead to practical differences between the item in hand and the bibliographic record.

It is possible that nine libraries did not find any errors because errors are exceedingly rare. Two of the libraries did not provide specific documentation about the errors found except for the base cause (e.g., date, edition etc.). If we were to remove the twelve errors that do not cause confusion about the book in hand the overall error rate reduces to 0.06%, or less than six instances in every 10,000 cases.

As noted previously, results of Teper's study²² showed significantly higher error rates than what our ILL-based study demonstrates. An important difference between this study and Teper's is the age of the monographs analyzed. The books Teper studied were older and from a time when cataloging practices were less standardized; these books, therefore, do not reflect the bulk of ILL transactions. In contrast to a review of nineteenth century and early twentieth century monographs, our review of ILL data constitutes a study of far more current materials. This claim is substantiated by studies that show the preponderance of items requested through ILL have publication dates within the most recent one to two decades from the date of the ILL request.²³

The differences in the findings of these two studies are significant in a shared print environment. The number of books published grew steadily from the early nineteenth century to the mid-twentieth century, at which point growth accelerated rapidly.²⁴ Therefore, the number of newer publications far outstrips the number of older ones in library collections.²⁵ Since the availability of resource sharing is core to shared print programs, the lower number of bibliographic mismatches in this population of newer publications is noteworthy.

An analysis of a large data set of bibliographic records further confirms the ILL study results. OCLC's Sustainable Collection Services (SCS) provided data from approximately 430 separate libraries. The SCS dataset represented approximately 205.5 million title counts. SCS found that a mismatch of either title or author of a book only occurred in an average of 0.12% (or median of 0.03%) of the 205.5 million title counts and that 0.6% of the records could not be validated with an OCLC control number. While the SCS reports offer insights into the occurrences of mismatches it lacks the component of a physical comparison. As ILL staff handles each requested item, a more precise validation process occurs to verify if the lending library indeed owns the exact book in question.

Finally, a further impetus for this study was the question of whether bibliographic inaccuracy impacts retention decisions. We considered the possibility that bibliographic errors create a false impression of scarcity. This is arguably good for preservation, since those apparently scarce titles may be retained. However, this is contrary to the goals of shared print programs to effectively and efficiently manage access to resources. ILL data helps to put this hypothetical scenario into context.

Conclusion

In general, the review of the data sets results in an overall low rate of bibliographic inaccuracies—specifically 0.10%, and no more than 0.47% in the most cautious reading of the data—of the total 29,630 items reviewed. Moreover, these instances of bibliographic inaccuracy rarely prevented identification of the correct item and fulfillment of the ILL request. The error rate number drops to less than six instances out of every 10,000 items, and most likely even less, when the errors are limited to only verifiable errors that point to a different item than what is actually in hand. Even limiting the data to only libraries that found any error, the error rate is still below 0.5%. These are relatively small error rates, although in the context of millions of items they can amount to a significant number of affected items overall.

Areas for Future Research

This paper examines instances of bibliographic inaccuracies from a collections perspective, specifically in the context of shared print retention programs. However, it does point to an area of additional research that would be of potential value, namely that of the examination of cataloging practices and current trends and how that relates to the overall issue of accuracy of bibliographic records. Although that discussion is outside of the scope of this paper, it presents the opportunity for future meaningful research. Additionally, a follow-up study that is not limited to the ILL-based study conducted for this paper could constitute another area of future valuable research. This would allow for further exploration of bibliographic mismatches to items in hand.

As shared print programs look to mitigate risk as part of the application of retention commitments, other related areas hold value for future research. As mentioned previously, we did not include in our analysis the information from the library that submitted data for books being sent to storage, as opposed to data from ILL transactions, due to the potential of substantially different book demographics between the two types of data sets. However, data from books sent to storage offers a compelling direction for continuing research on bibliographic inaccuracy. Exploration of instances of bibliographic inaccuracy discovered by shifting items to a storage repository could uncover some useful data as it relates to shared

print programs. Namely, it serves as a validation process that the library has confirmed that it indeed owns the item in hand because bibliographic records are often verified and corrected in the process of sending items offsite. This could result in a more comprehensive review of a much larger quantity of possible mismatches between record and book since total numbers of items relocated to storage facilities would typically outstrip the numbers of ILL requests within the finite time period our study covered.

The implementation of shared print programs may need to account for different bands of risk management based on various collections demographics. Although some subsets of the collective collection may require more rigorous validation than others, such as the nominal occurrences of bibliographic inaccuracies as uncovered in this study, these subset groups may ultimately allow for individualized targeted and effective retention efforts. We hope that the data presented here can assist shared print programs assess the risk of bibliographic inaccuracies, in particular in conjunction with the use of the optimal number of copies model.²⁶

Appendix. Instructions to Participating Libraries

Monograph Bibliographic Accuracy Study

The Partnership for Shared Book Collections is a fledgling community that has formed to support and promote shared print programs. One of its initiatives is to make informed recommendations on creating responsible commitments. Shared print programs exist in an ecosystem and few operate totally independently. Libraries are often members of more than one shared print program.

A working group of the Partnership is trying to determine how many copies of a particular title should have commitments, either in a single program or in all programs, through an evidence-based process. There are several variables that are necessary in determining an adequate number of copies. One issue that has been raised several times is how much confidence can we have that bibliographic records match the objects they are supposed to represent. There is some speculation, especially with details like editions, that there may be a significant number of errors, attributable, for example, to recon project matching algorithms.

We have not found existing research in this area. In trying to find data, the prospect of partnering with resource sharing staff was raised, since they look closely at the records and items during their work. We have contacted a few resource sharing departments and from them have heard that staff do not find bibliographic errors at a high rate, which makes us optimistic that a few resource sharing departments may be willing to help us capture data about errors. If you are willing to help us out with a six-week study by capturing data as you lend items, please let me know. We hope to start the study in early to mid October and finish by Thanksgiving.

You can use the form on the reverse side of this sheet for books lent. Please provide:

Library name

Date you started recording errors

Date you stopped recording errors

The total number of requests for monographs you searched during this time period

Tick marks for each record/book discrepancy in the designated categories; total of tick marks

Photographs of the request and the title page/verso for books with mismatches are greatly appreciated but not necessary.

Please return completed forms and any images you create, along with any comments, to me at:

Library Name: _____

Date Start: _____

Contact Name: _____

Date End: _____

Contact email: _____

Total number of requests for monographs processed during this time span: _____

Place a tick mark in the box for a mismatch between the bib info on the request and the book in hand.*Only record one error per item, preference given in the order listed below:

Major Mismatch (title, author):

Minor Mismatch (edition, year, place of publication, publisher)

Other, tally and type (e.g. III - not on shelf,)

N/A

* Photographs of the request and TP/Verso are greatly appreciated

Acknowledgement

We are indebted to the fifteen anonymous libraries that contributed data to the project. This work would not have been possible without their generosity in collecting and submitting data.

Notes

1. Maiorana, Z., Bogus, I., Miller, M., Nadal, J., Risseuw, K., and Teper, J. H. (2019). Everything not saved will be lost: Preservation in the age of shared print and withdrawal projects. *College & Research Libraries*, 80(7), 945-972. <https://doi.org/10.5860/crl.80.7.945>.
2. Maiorana et al., 2019.
3. Partnership for Shared Book Collections (2022). <https://sharedprint.org> [accessed 21 August 2023].
4. Martin, J., Kamada, H., and Feeney, M. (2013). A systematic plan for managing physical collections at the University of Arizona libraries. *Collection Management*, 38(3), 226-242. <https://doi.org/10.1080/01462679.2013.797376>; Schleicher, M. C. (2010). Assembling selection criteria and writing a collection development policy for a variety of older medical books. *Journal of Hospital Librarianship*, 10(3), 251-264. <https://doi.org/10.1080/15323269.2010.491424>; Snyder, C. E. (2014). Data-driven deselection: Multiple point data using a decision support tool in an academic library. *Collection Management*, 39(1), 17-31. <https://doi.org/10.1080/01462679.2013.866607>; Soma, A. K., and Sjoberg, L. M. (2010). More than just low-hanging fruit: A collaborative approach to weeding in academic libraries. *Collection Management*, 36(1), 17-28. <https://doi.org/10.1080/01462679.2011.529241>.
5. Acadia, S. (2016). Books be gone! Reducing an academic library's print collection by half to meet strategic planning initiatives and participate in a joint library resource-sharing facility. *Journal of Library Administration*, 56(2), 144-157. <https://doi.org/10.1080/01930826.2015.1105668>; Gillies, S., and Stephenson, C. (2012). Three libraries, three weeding projects: Collaborative weeding projects within a shared print repository. *Collection Management*, 37(3-4), 205-222. <https://doi.org/10.1080/01462679.2012.685420>; Kieft, R. H., and Payne, L. (2012). Collective collection,

- collective action. *Collection Management*, 37(3-4), 137-152. <https://doi.org/10.1080/01462679.2012.68541>; Levine-Clark, M., Jobe, M., and Holliday, S. (2009). Uniqueness and collection overlap in academic libraries. *Proceedings of the Charleston Library Conference*, 233-237. <https://doi.org/10.5703/1288284314751>.
6. Ward, S. M. (2015). *Rightsizing the academic library collection*. American Library Association.
 7. Kieft, R. H. (2014). Beyond my people and thy people, or the shared print collections imperative. In B. Albitz, C. Avery, & D. Zabel (Eds.), *Rethinking collection development and management* (pp. 297-319). Libraries Unlimited.
 8. Demas, S. (2011). Curating collective collections—Prospectus for a new ATG column. *Against the Grain*, 23(6), 76-78. <https://doi.org/10.7771/2380-176X.6066>; Kieft, R. H., and Payne, L. (2010). A nation-wide planning framework for large-scale collaboration on legacy print monograph collections. *Collaborative Librarianship*, 2(4), 229-233. <https://digitalcommons.du.edu/collaborativelibrarianship/vol2/iss4/8>.
 9. Clement, S. K. (2012). From collaborative purchasing towards collaborative discarding: The evolution of the shared print repository. *Collection Management*, 37(3-4), 153-167. <https://doi.org/10.1080/01462679.2012.685413>; Crist, R. D., & Michaels, S. L. (2013). Shared print on the move: Collocating collections. *Proceedings of the Charleston Library Conference*, 162-166. <https://doi.org/10.5703/1288284315252>; Maiorana et al., 2019; Palazzolo, C., and Meyer, L. (2018). Shared print initiatives. *Proceedings of the Charleston Library Conference*, 163-168. <https://doi.org/10.5703/1288284317033>; Stearns, S., Revitt, M., and Leonard, K. (2020). Taking shared print to the next level: The partnership for shared book collections. *Journal of Library Administration*, 60(7), 812-820. <https://doi.org/10.1080/01930826.2020.1803020>; Stearns, S., and Wohlers, A. (2020). Shared print on the threshold: Looking back and forging the future. *Collaborative Librarianship*, 12(2), 159-169. <https://digitalcommons.du.edu/collaborativelibrarianship/vol12/iss2/5>.
 10. Michaels, S. and Neel, B. (2021). Conducting an inventory with shared print in mind. *Collection Management*, 46(2), 142-156. <https://doi.org/10.1080/01462679.2020.1818343>.
 11. Baich, T. & Michaels, S. (2017). Facilitating information sharing through library collection maintenance and preservation. In B. Posner (Ed.), *Library information and resource sharing: Transforming services and collections* (pp. 93-110). Libraries Unlimited.
 12. Teper, J. H. (2019). Considering “sameness” of monographic holdings in shared print retention decisions. *Library Resources & Technical Services* 63(1), 29-45. <https://doi.org/10.5860/lrts.63n1.29>.
 13. Snow, K. (2017). Defining, assessing, and rethinking quality cataloging. *Cataloging & Classification Quarterly*, 55(7-8), 438-455. <https://doi.org/10.1080/01639374.2017.1350774>.
 14. Petrucciani, A. (2015). Quality of library catalogs and value of (good) catalogs. *Cataloging & Classification Quarterly*, 53(3-4), 303-313. <https://doi.org/10.1080/01639374.2014.1003669>.
 15. Nero, M. D., & He, J. (2018). Is it necessary: Quality control in cataloging. *International Journal of Librarianship*, 3(2), 85-95. <https://doi.org/10.23974/ijol.2018.vol.3.2.96>.
 16. Intner, S. S. (1989, February 1). Much ado about nothing: OCLC and RLIN cataloging quality. *Library Journal*, 38-40; Snow, 2017.
 17. Bade, D. (2008). The perfect bibliographic record: Platonic ideal, rhetorical strategy, or nonsense? *Cataloging & Classification Quarterly*, 46(1), 109-133. <https://doi.org/10.1080/01639370802183081>.
 18. Michaels and Neel, 2021.
 19. Michaels and Neel, 2021.
 20. Maiorana et al., 2019.
 21. Bogus, I., Yano, C. A., Zachary, S., Nadal, J., Miller, M., Levenson, H. N., Brody, F., and Amato, S. (2023). A model to determine optimal numbers of monograph copies for preservation in shared print collections. *College & Research Libraries* 84(5), 767-801. <https://doi.org/10.5860/crl.84.5.767>.
 22. Teper, 2019.
 23. Reighart, R., and Oberlander, C. (2008). Exploring the future of interlibrary loan: Generalizing the experience of the University of Virginia, USA. *Interlending & Document Supply*, 36(4), 184-190. <https://doi.org/10.1108/02641610810919516>; Ruppel, M. (2006). Tying collection development's loose ends with interlibrary loan. *Collection Building*, 25(3), 72-77. <https://doi.org/10.1108/01604950610677530>; Tolppanen, B. P., and Derr, J. (2010). Interlibrary loan patron use patterns: An examination of borrowing requests at a midsized academic library. *Journal of Interlibrary Loan, Document Delivery & Electronic Reserve*, 20(5), 303-317. <https://doi.org/10.1080/1072303X.2010.522941>.
 24. Lavoie, B. F. and Schonfeld, R. C. (2006). Books without boundaries: A brief tour of the system-wide print book collection. *The Journal of Electronic Publishing*, 9(2). <https://doi.org/10.3998/3336451.0009.208>
 25. Lavoie and Schonfeld, 2006.
 26. Bogus et al.