

The ReShare Paradigm — What can we achieve?

Karin Gilje
University of Pennsylvania Libraries, USA

Abstract

Purpose & Goals

The Partnership for Academic Library Collaboration & Innovation (PALCI) and Ivy Plus Library Confederation (IPLC) have moved to the ReShare platform by Index Data (in Fall 2021 and Winter 2022 respectively). Through close collaboration, Index Data has provided the University of Pennsylvania access to the data for import into our MetriDoc Data Warehouse where deeper analyses can be performed through tools such as Microsoft PowerBI. Rather than relying on simple counting statistics, we can evaluate the uniqueness of our collections and the efficacy of sharing within consortia. One of the challenges with the new system is to change the assumptions stakeholders have in calculating essential metrics. However, the transparent nature of the data provides more opportunities to examine data in new and unique ways.

Design & Methodology

The ReShare interlibrary loan platform relies on a centralized PostgreSQL data store hosted by Index Data. Index Data has made their underlying code base open source which increases the transparency of how resource sharing works. The Penn Libraries host a Ruby on Rails User Interface based on a PostgreSQL database called MetriDoc. Index Data provides a selection of their data tables to MetriDoc for Access Services to have direct access to it. Dashboards were created within the User Interface for quick lookups of key datapoints and further analyses can be performed by our Resource Sharing colleagues through an interactive query system. Additionally, tools such as Power BI or Tableau can connect to the PostgreSQL database directly to create fully interactive Data Visualizations and Reports.

Findings

As with any migration, there have been issues, but the collaboration between the Penn Libraries and the vendor, Index Data, has improved all of our understanding of the data. Through our close interactions, we have developed a complete data model to track loans through the system. The data model is rich in collection, operational, and management information. We can use the data to analyze which parts of our collections are in high demand at other institutions and where our patrons may need to rely on our partners' collections. The two consortia contract with shipping vendors to ensure items have minimal transit time. The exit and entrance scan information helps us evaluate the shipping performance. ReShare data also provides detailed information on the rota, or ordered list, of institutions asked for an item. This rota is developed for each item based on the current share of an institution's resource sharing load. This information has never been available before and has opened new venues for understanding the depth and breadth of the joint collections and for gauging the balance of workload and operational effects on participating libraries.

Action & Impact

The Penn Libraries will continue to maintain and improve the MetriDoc service. Although the main part of this project is complete, we will be developing a series of workshops for our

stakeholders to help them understand and leverage the data. We will also be sharing the research questions with the vendor, Index Data, to help them understand how the data is being utilized and continue our collaborative effort to provide answers to new questions. It is important to note the uniqueness of this data collaboration with a library vendor. While library systems like Alma generally provide reporting modules, those tools comprise a set of independent data silos that restrict useful integration of disparate transactional source data, limit the resolution of statistical analysis to canned results, and restrict opportunities for visualizing data or incorporating emerging tools in the AI frontier into data analysis. The Penn Libraries' workflow model with Index Data presents a forward-thinking approach to library/vendor collaboration in library business intelligence.

Practical Implications & Value

Although the University of Pennsylvania hosts and maintains MetriDoc, it also houses the Resource Sharing Assessment Tool (RSAT) developed for the IPLC. Completing the migration to the Index Data ReShare platform fulfills our compact with the IPLC to continue to provide RSAT access to BorrowDirect resource sharing program data. However, we have gone beyond simply providing data to also providing in situ dashboards and querying capability. In addition, we have developed an excellent cooperative relationship with the vendor, Index Data.

Characters

Index Data

Index Data serves as the certified service provider for the ReShare resource sharing platform. They build open platforms to support libraries and their communities. Their goal is to build products based on librarian input in a user-centric approach. The ReShare resource sharing platform relies on a centralized PostgreSQL data store hosted by Index Data. Index Data has made their underlying code base open source which increases the transparency of how resource sharing works. ReShare encourages experimentation and Index Data continues to support innovation in the library sphere based on their platforms. For more information and to explore what ReShare can do, check out their website at <https://projectreshare.org/>.

IPLC

The Ivies Plus Library Confederation (IPLC) consists of 13 academic libraries. The consortium seeks to leverage collective assets for discovery and scholarly research. They focus on collaboration in discovery, collection development, and advocacy. To achieve this, the BorrowDirect resource sharing service originated in the fall of 1999 with Columbia University, the University of Pennsylvania and Yale University, initially termed the CoPY project. Through 2016, the service was expanded to include the rest of the Ivy League Schools as well as Stanford University, University of Chicago, Duke University, Massachusetts Institute of Technology and Johns Hopkins University. The IPLC migrated to the ReShare resource sharing platform in the Winter of 2023. For more details on the confederation, see their [website](#).

MetriDoc

Around 2000, the Penn Libraries developed a data warehouse called DataFarm. DataFarm intended to gather library data from disparate sources into one platform to better understand patron behavior and to build a more comprehensive view of library services. In 2010, DataFarm was transformed into the first iteration of MetriDoc as a Groovy on Grails project supported by the Institute for Museum and Library Services ([IMLS LG-07-10-0236-10](#)). This new system was

intended to be an open-source decision support system for libraries to leverage. MetriDoc was ported to a Ruby on Rails environment in 2019 supported by the University of Pennsylvania and the IPLC. The work was preliminary to the design and construction of RSAT for the purpose of comprehensive analysis of the IPLC's resource sharing activities.

The code for MetriDoc is open source and available on [GitHub](#). The MetriDoc Data Warehouse contains a variety of data from Interlibrary Loan to EZProxy Logs to SpringShare appointments and more. The Penn Libraries are continuing to add new data sources to more fully understand the library environment.

From the first iteration of DataFarm to the current MetriDoc Rails project, we have been the reporting utility for IPLC's BorrowDirect program and previously the EZBorrow service for the Pennsylvania Academic Library Consortium Inc (PALCI). MetriDoc uses an Extract-Transform-Load (ETL) model for data ingest with an additional layer of user interaction to meet our stakeholders where they are.

Extraction – Data is extracted from sources either manually or preferably automatically. Index Data has provided read-only user credentials to their PostgreSQL database for MetriDoc to replicate data tables.

Transformation – Data can be automatically transformed and joined with other library data to form a more complete picture of activity. For ReShare data, additional tables are created to make the data more transparent for end users.

Loading – Data is loaded into a PostgreSQL database with querying capabilities.

User Interaction – The front-end webpage uses the PostgreSQL data to build indices with filtering capabilities where users can download the raw tables in CSV, XML or JSON formats. There are additional dashboards built into the website with general statistics for at-a-glance checks on services. Finally, for the more tech savvy librarians, direct connection to the PostgreSQL database is available for business intelligence tools like Tableau and PowerBI. The connection can also be used programmatically through languages such as R and Python.

Data

Typical vendors only provide access to canned reports. However, these reports are limited in scope. Without event-level data, future questions not yet asked may remain impossible to answer. Index Data provides a selection of their data tables to MetriDoc to replicate the event-level data. This is an unprecedented level of access and gives libraries the ability to react to changes in our future understanding of user behavior.

Working with Index Data, we at the Penn Libraries with other like-minded individuals across the ReShare environment developed a set of default reports for Index Data to host. This was essential for understanding the data structure. Additionally, Index Data did not restrict us to only the canned reporting but gave us access to the raw data for us to continue to explore and innovate on the types of information we can uncover.

An example of the continued collaboration between the Penn Libraries, as the user, and Index Data, as the vendor, was the exposure of the Host LMS Shelving Location Table. Originally the Patron Requests table used the Shelving Location as plain text but later moved to using a look up table, a much cleaner approach to storing this data. This table was not used in our original exploration, but later became a necessity to access the Shelving Location in a human readable format. We reached out to our connection at ReShare and the look up table was made available. This unprecedented level of collaboration continues to improve our understanding of the use of our collections and the inner workings of the ReShare platform.

Using this extraordinary access, I developed an Entity Relationship (ER) diagram of the ReShare databases (see Figure 1). At first glance, the database structure may seem complicated. As it is broken down into smaller pieces connected with specific questions, clarity begins to appear.

Different stakeholders across the libraries are interested in different types of data. Bibliographers and Subject Libraries are more interested in the bibliographic information available about the item being requested than anything else (see the items in the Bibliographic information part of the Patron Requests table). On the other hand, the consortium is interested in who is borrowing and who is lending items (Lender and Borrower Information in the Patron Requests table). A handful of lookup tables are needed to translate the hashed foreign keys into human readable objects (the right side of the ER diagram). For local interlibrary loan staff, the process each item goes through is important to understand. Each step in that process is recorded in the Patron Request Audit table which requires a lookup table for the status changes which are represented as hashed values.

One table we are just starting to explore is the Patron Request Rota. This table provides details on the order institutions will be asked for an item. We get the full list even if the first institution lends the item. As the diagram is broken down into its components, we can get a better idea of the types of questions we can ask and start to explore new areas of data analysis.

This level of data is fantastic for a data scientist who has license to explore. However, librarians should not have to be data experts to extract statistics from the data. To reduce the cognitive load on our librarians, I created a Transactions table (see Figure 2). The idea was to build a single flat record of each borrower and lender interaction with all the bibliographic, shelving, pick up and status information in a human readable format.

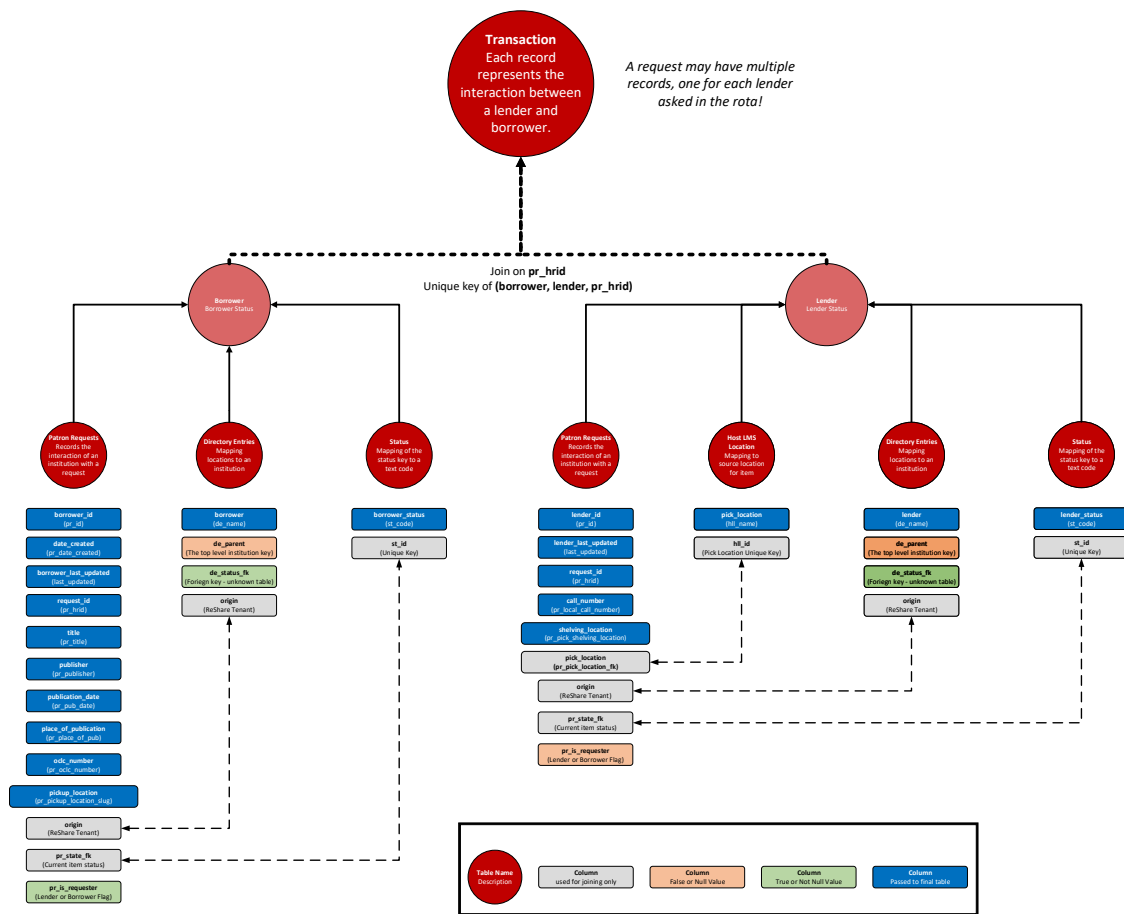


Figure 2: Construction of a Transactions Table. Image created by Microsoft's Visio.

This is where the power of MetriDoc and its capabilities come into the story. As a programmer, I need to meet my stakeholders where they are. Our users have a variety of technical abilities, we want to ensure they can access the data in the most natural way to them. On the MetriDoc website, a selection of canned statistical reports is available for quick viewing. Additionally, users can access both the original (such as the Patron Requests Table) and derived (such as the Transactions table) tables through Rails index pages with filtering, pagination and download capabilities. There is also a built-in query tool to create and run SQL queries through the website. Finally, for the more tech savvy librarians, direct connection to the PostgreSQL database is available for business intelligence tools like Tableau and PowerBI. The connection can also be used programmatically through languages such as R and Python.

Discovery

With such a robust data source, there are a variety of insights we can gain about user behavior and ReShare's process. I will only touch on a few here. The IPLC migrated to ReShare in December 2022.

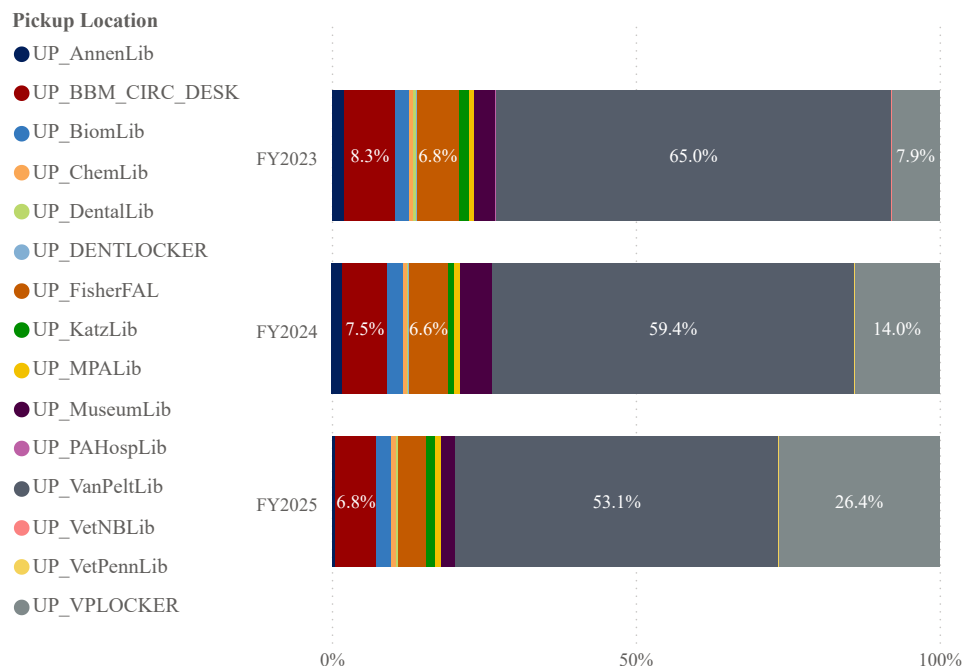


Figure 3: Borrowing by Destination. The X axis shows the percent of loans by delivery method and the Y-axis shows the Fiscal Year. Fiscal Year 2025 only contains the first quarter. The colors represent the different pickup sites and methods across the Penn Libraries. Plot created using Microsoft's PowerBI.

Operationally, we are interested in what services our patrons are using to access their materials. This data is accessible through the Transactions table. The pandemic created an opportunity for the Penn Libraries to explore alternate ways to share collections with patrons. Pick-up lockers were installed at the Van Pelt-Dietrich Library Center (our main library – VP-DLC) and at the Dental Library. The VP-DLC has seen a continued rise in usage of the Lockers (the light grey bar at the right hand of Figure 3, labeled UP_VPLOCKER). This increase may not only be due to user behavior as Access Services has changed some of their workflow related to delivery of items.

Another clear trend is a reduction in the Books By Mail service (Dark Red bars in Figure 3, labeled UP_BBM_CIRC_DESK). For current faculty, staff, and students in the United States and Canada, the Penn Libraries will ship requested items to the patron's home address. This service allows people who are not on campus to use library resources. The Books By Mail service saw a large spike in usage over the pandemic with more local people signing up for access. Around a quarter of Books By Mail activity goes to Philadelphia and its suburbs. With the return to campus, there has been a natural decrease in the program's popularity. Access Services is trying to reduce local usage due to on-going systemic issues with porch piracy in Philadelphia neighborhoods. In these cases, the patron is not held responsible for the missing item which is then lost to both the patron and the libraries.

Using the Patron Request Audit table, we can calculate detailed turnaround times. For the IPLC in Fiscal Year 2024, our response averages from half a day to three quarters of a day split by institution. Extracting the specific shipping time also lets the consortium monitor the performance of the shipping vendor. For the Penn Libraries, the shipping time depends slightly on geography, with closer institutions like Princeton University and Johns Hopkins University

averaging just over 1 day in Fiscal Year 2024. Stanford University, across the continental United States, has an average shipping time of over 7 days in Fiscal Year 2024. Being able to pull these statistics out quickly and efficiently is useful when the consortium reevaluates the shipping vendor.

Not only does the library need to be concerned about turnaround time from receipt to shipping, but it needs to track the turnaround time for notification of not having an item available. There were 11,079 requests that were unfilled by the Penn Libraries in Fiscal Year 2024. The majority of those were automatically rejected by an autoresponder based on an item's availability. This allows the request to be immediately passed to the next institution in the rota. However, 990, or 9%, of the items were manually rejected. Introducing manual rejections can greatly increase the wait time for a patron.

Understanding the different reasons for manual rejections can lead to improvements in the autoresponder. There are several reasons for needing to reject a request. If an item was requested on Friday afternoon, it may not be searched for until Monday morning. In the interim, the item could be checked out locally. Occasionally, an item will be discovered to be library-use only and not a lendable item. The Penn Libraries uses shelving locations to determine if parts of the collection are lendable or not. When this situation occurs, the shelving location gets a note and an improvement to our autoresponder is made. Another reason could be the item was placed on reserve, perhaps for a course. As Penn Libraries continues to use ReShare, the number of manually closed items should reduce due to proactive actions.

Typically, Interlibrary Loan systems have opaque rota calculations. Libraries are not sure why they are requested for items in any particular order. IndexData has developed a clear method for determining what the rota looks like for every loan. Each institution can input their preferred "Load Balancing Ratio" which is a target Lending to target Borrowing ratio. Some of our more geographically distant members choose to have a lower ratio to keep their collections closer to home. Most of the consortium targets a ratio of 1 to 1. For every one book the library borrows, they would like to lend one book.

Any time a book is requested, a rota is calculated for that instance in time using the following formula:

$$S = B_A * \frac{L_T}{B_T} - L_A$$

Where S is the score, B_A is the Actual Borrowing, L_A is the Actual Lending and L_T / B_T is the target lending to borrowing ratio (check out the [code](#)). Figure 4 shows the distribution of institutional rota positions. The geographically distant institutions tend to be later in the rota as evidenced by the pink and dark red portions of the bars.

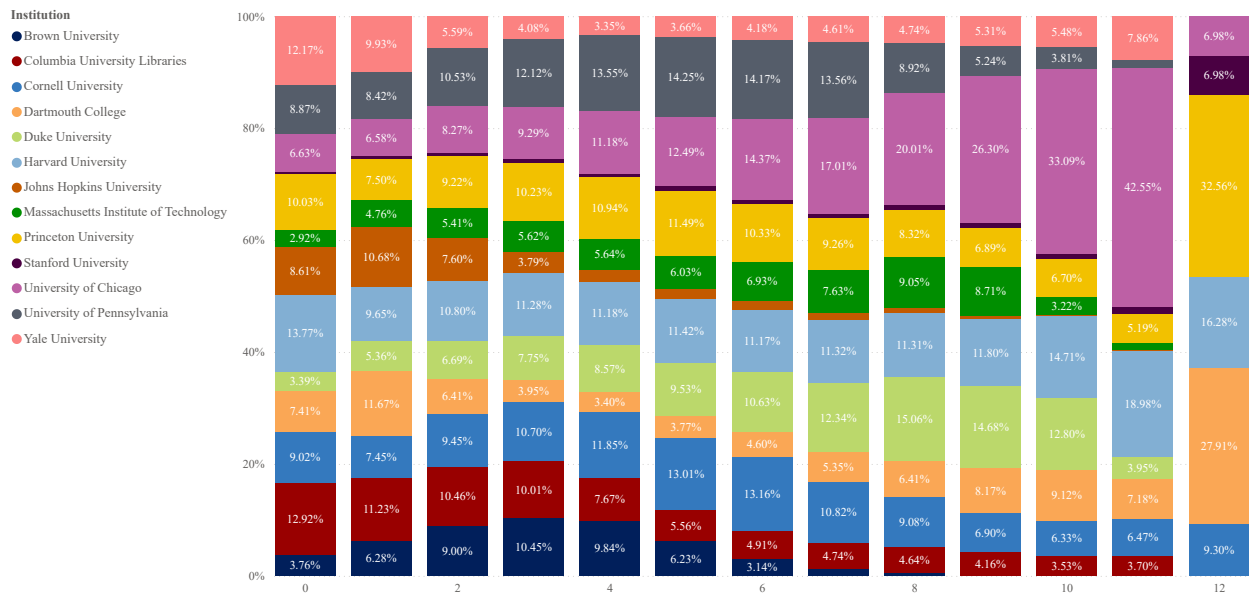


Figure 4: Consortium Rota Positions. The color bar represents the different consortial members. The X-axis shows the location in the rota. The Y-axis is the percent of the rota position. The labels on the bars show the percent of the rota position occupied by that consortial member. Plot created using Microsoft's PowerBI.

We can also look at the uniqueness of parts of our collections. Figure 5 shows the depth of the rotas created by the Library of Congress main classes of items lent by the University of Pennsylvania to our partners. The shading represents how many institutions were present in the rota for that item. When the rota depth is one institution, only the Penn Libraries had the requested item in their collection. Lighter vertical bars represent more unique items.

Philadelphia is the birthplace of the U.S. Navy, which has led to a robust collection of LC class V (Naval Science). In fact, looking across the consortium, at most 7 institutional members had items from the Penn Libraries' collection requested by the consortium. Nearly half of the items the Penn Libraries provided were unique items.

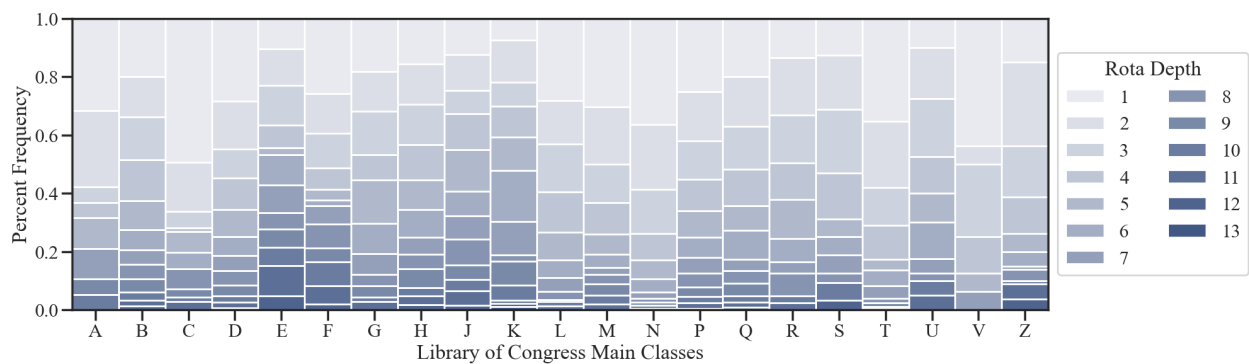


Figure 5: The rota depth created for items lent by the Penn Libraries to the IPLC separated by the Library of Congress classifications. The X-axis has the categories for the Library of Congress top-level classification and the Y-axis shows the percent of the items in a class shaded by the rota depth. Plot created using Python.

In contrast, the LC class E (History of the Americas) lent by the Penn Libraries show more commonalities amongst the consortium collections. In fact, nearly 5% of the items provided by the Penn Libraries were also available at 11 other institutions, that is, the rest of the consortium.

Conclusion

We collaboratively worked with Index Data to understand the ReShare tables and build reports and tables. We built pipelines to import data into the MetriDoc Data Warehouse. We created summary tables to reduce the cognitive load on consortium members. We built multichannel data access to meet our stakeholders at their technical abilities. Although this paper focuses on the data from BorrowDirect, because PALCI and its EZBorrow service is also a ReShare platform, all analyses mentioned can be performed with PALCI data. Using the ReShare data and statistics, we have supported institutional and consortial decision making with an eye to the future of interlibrary lending.

Bibliography

- Collins, Peter. "Fear and loathing in cooperative collection development." *Interlending & Document Supply* 40.2 (2012): 100-104. <https://doi.org/10.1108/02641611211239579>.
- Wilson, Kristen, and Jill Morris. "Project ReShare: Building a Community-Owned Resource Sharing Platform." *The Serials Librarian* 78, no. 1–4 (February 2020), 141–48. <https://doi.org/10.1080/0361526x.2020.1717323>.
- Zucca, Joe. "Business Intelligence Infrastructure for Academic Libraries." *Evidence Based Library and Information Practice* 8, no. 2 (June 2013): 172–82. <https://doi.org/10.18438/b83g75>.