Document Delivery and Resource Sharing: Transforming Resource Sharing in a Networked Global Environment
10-11 August 2016
Library of Congress, Washington, DC

A Pinhole Approach to Understanding ILL Costs and Trends, or, What a Dutch Master Can Teach Us About Analyzing Resource Sharing Data

Dennis Massie
OCLC Research
massied@oclc.org

Abstract:

In an environment where library administrators emphasize the virtues of evidence-based decision making, the resource sharing community has access to only a small proportion of the transactional data it needs. Most have access to data from only those groups in which they are active, generated by the resource sharing systems they use. No complete global picture is available. Worse, the most often quoted cost data associated with interlibrary loan activity is more than a decade old and does not cover newer models of sharing library materials.

This paper reports on recent efforts by OCLC Research, in consultation with library practitioners, to draw fresh conclusions about current resource sharing costs and trends by studying carefully-selected snapshots of data. Those snapshots include: 1) five years' worth of collection-sharing data contributed by libraries from two large resource sharing consortia; 2) the results of structured interviews conducted with interlending staff from those consortia; and 3) lessons learned from building an ILL Cost Calculator designed to function as a real-time virtual ILL cost study. By extrapolating from these small samples of data and experience, OCLC Research demonstrates how we might come to a better understanding of our "macro" collection-sharing world by going selectively "micro" – much as Vermeer is rumored to have created some of his most famous masterpieces using the camera obscura *technique.*

Keywords: Resource sharing, cost study, data mining, consortia, interlibrary loan.

Asking Big Questions is the Easy Part

It's a big universe out there. Vast. Incomprehensible. We as human beings sometimes like to contemplate that vastness, and to seek some understanding of our place in it. Often such pondering gives rise to profound questions, such as "How big *is* the universe?" Or: "Will it just keep expanding, for all eternity? And does it support alien life forms?" Sometimes the questions are focused a bit closer to home: "Is library collection-sharing activity going up or down across the ecosystem? And, in today's

environment, which costs more for a lender to process – a request for an article from a printed volume, or a request for that same article from an e-journal?"

Unfortunately, we don't know the answer to any of these questions.

A "Water, Water Everywhere" Sort of Problem – and Some of the Water is not Fresh

I get paid to spend time thinking about that second set of questions. I work for OCLC, where we have a long-standing interest in supporting collection-sharing activity. Naturally, it behooves us to try to understand some basic things about the universe in which we operate. But we have ready access to only a tiny portion of the data that describes that universe. Our own data, drawn from our own resource sharing infrastructure, indicates that interlibrary lending traffic is trending slightly downward—from about 10.3 million items requested in 2009 to less than 8 million by 2015, a drop of twenty-five percent over six years **(Birch 2014).**

Figure 1:
Declining OCLC ILL statistics

	FY15	FY14	FY13	FY12	FY11	FY10	FY09
ILL requests	7,714,710	8,235,069	8,858,368	9,192,189	9,587,429	10,248,942	10,279,215
Annual +/-	-6%	-7%	-4%	-4%	-6%	-0.3%	

Between FY09 and FY15, OCLC ILL has seen a 25% reduction in total number of ILL requests.

Is this happening everywhere, across all ILL systems worldwide? We have no way of knowing. While experience tells us that sharing collections generates a vast amount of statistical information, no one collects such data from all parts of the globe and puts it in one place where we can all get at it. Even within the confines of a single continent – say, North America -- there is no central source of ILL data. The Association of Research Libraries, which includes 125 academic, state, national and public libraries in the United States and Canada, collects data from its members each year, including the numbers of items borrowed and loaned via ILL. But the numbers are expressed as a total for each institution, with one big number representing items loaned, and another number for items borrowed. There is no indication of how the requests were sent, whether the items loaned were print or electronic, or what proportion of the materials were provided in returnable formats. And there is certainly no indication of the "why" behind any changes that occur from year to year. Another source of uncertainty is the fact that the numbers are self-reported, not gathered directly from an automated resource sharing system. If one spends some time exploring the footnotes of ARL's annual statistical reports, one finds things such as "Sorry, we've been double-counting this one category of ILL requests for the past few years, but we promise not to do that any longer."

We also have no idea how much it costs to share collections. The figures quoted today when folks talk about the costs of resource sharing are culled from studies conducted in North American and Australia more than a decade ago. The findings referenced most often are from a 1992 study conducted jointly by the previously-mentioned ARL, in association with the Research Libraries Group, which no longer exists; this study of 76 large North American academic libraries found that processing an average borrowing request cost $18.62 USD, while processing the average lending request cost $10.93 USD

(Roche 1993). Lending fees were not included as part of the costs – these numbers are strictly overhead. The report itself does go into some detail about the difference costs between copies and loans, but these gross numbers are typically what still get quoted.

In a world where comprehensive library data is difficult if not impossible to get hold of, what's a library administrator (or OCLC Research program officer) to do?

For me the answer was simple: this called for the old pinhole approach. That's right. *La camera obscura*. According to at least one art historian, many great Renaissance paintings were accomplished using this technique, including several by Vermeer *(Steadman 2001)*. You've seen diagrams of how this works. The artist (or, in our case, analyst) allows a small amount of light (or, in our case, data) to flow into his possession through a pinhole, after which it is projected upside down on an opposite wall or screen. Once there the image (data) can be observed, traced over, captured, preserved, interpreted, and ultimately shared with others. Why shouldn't this same idea be applied to processing and understanding library collection sharing statistics? "If it's good enough for a Dutch Master, it's good enough for OCLC Research," is one way to sum up my line of reasoning.

Enter the Borrow Direct Consortium

Off I went in search of a collection-sharing pinhole. I knew that a long-time colleague, about to retire, was closely connected with the Ivies Plus consortium. The eight Ivy League institutions – Brown, Columbia, Cornell, Dartmouth, Harvard, Pennsylvania, Princeton, and Yale -- had joined together with a few like-minded institutions – the *Plus* part of Ivies Plus – such as MIT, University of Chicago, and Johns Hopkins University, to create a resource sharing consortium called Borrow Direct *(Nitecki 2004)*.

Borrow Direct is based on a model of resource sharing known, for lack of a better phrase, as consortial borrowing, or, perhaps more usefully, expanded circulation, or circ-to-circ. The way this model works is that the institutions involved create a big virtual group catalog that shows all the monographs available for borrowing via the scheme at all of the participating libraries, as well as real-time availability status. Library users at any of the institutions can order a book directly from any of the others, and it shows up at the user's home institution in three-to-five days. So it's fast. And because the availability of each item is known, the fill rate is about 95%. *And* because many institutions allow their users to order books that are owned but checked out at the home library, rather than instituting a recall, this model of sharing is enormously popular. I chose the Borrow Direct group to study because they are active, well-funded, well-run, and innovative. But mostly I chose them because my retiring colleague could make the introductions and smooth the way for me to get all the collection sharing data from the eleven (at the time) Ivies Plus institutions. Not just the Borrow Direct transactions, but all resource sharing data across all the libraries across all the ILL systems, both within the group and also with partners outside the group. Staff at each institution graciously agreed to collect and hand over to me five years' worth of their collection sharing data, which they did, quickly, and cheerfully, on very short notice, during an exceedingly busy part of the year. *Voila!* I had my pinhole.

The Borrow Direct Policy Committee and I agreed on a set of main questions to be explored:

- What numbers of borrows and loans has each institution executed in each of the various resource sharing systems and venues in the past 5 years?
- What factors determine the requesting method, system, or venue used for each request?
- How/why is all this changing over time?
- How will it most likely change in the future?

Upon receiving data from the eleven Borrow Direct members, the first thing I did was add it all up. *(See Figure 2 below.)* The trend in activity clearly seemed to be going up, year after year.
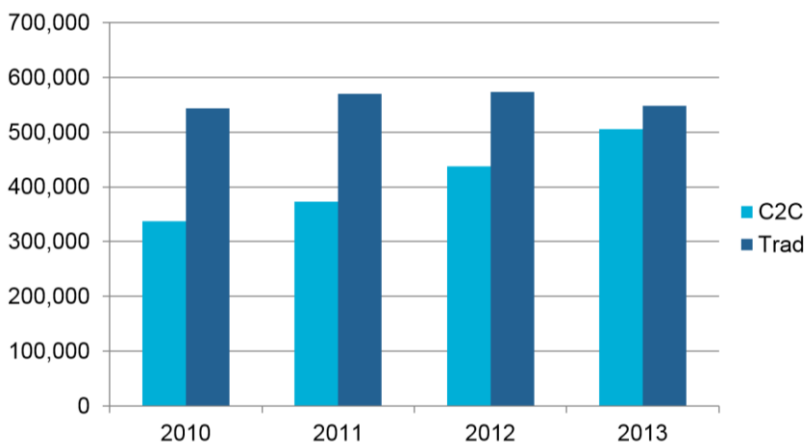
Figure 2:
Total overall sharing activity for all BD members
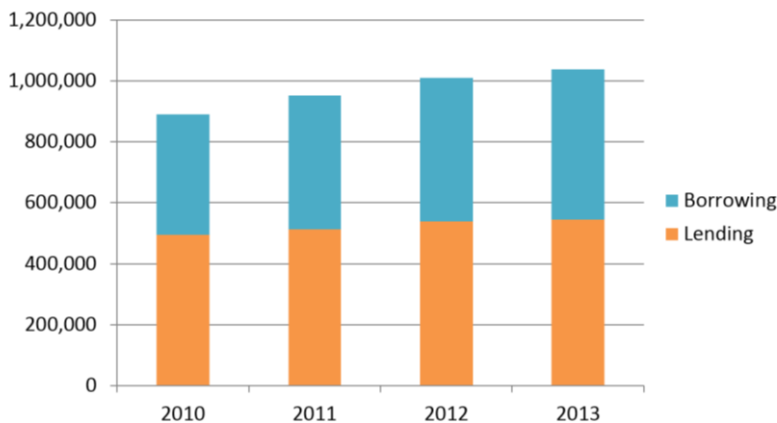
And there were two big fat juicy collection-sharing trends just waiting to be discovered.

First, all that collection-sharing growth noted in **Figure 2** turned out to be exclusively because of a sharp and steady increase in consortial borrowing activity on the part of the eleven Borrow Direct institutions. Remember, "consortial borrowing" refers to a model of collection sharing where consortium members merge the portions of their catalogs containing monographs into a virtual shared catalog from which eligible users from any of the institutions can order available materials directly from any of the other partners. As noted earlier, Borrow Direct itself is a consortial borrowing arrangement. Four of the eleven Ivies Plus libraries are also involved with an additional consortial borrowing scheme outside of Borrow Direct. This additional activity really boosts the proportion of consortial borrowing transactions that make up the total collection-sharing picture for these eleven institutions, and more so each year, when compared to the proportion of requests sent and received by what one might call traditional means – that is, via a model constructed around a physical union catalog into which all the partners load bibliographic and holdings data, and ILL requests are routed among the participating institutions by a resource sharing system built on top of the union list. *(See Figure 3 below.)* **Figure 3** illustrates this trend over four years, with the lighter columns representing the combined consortial borrowing traffic, going steadily up, while the darker columns represent the more traditional union-list-based ILL activity, including OCLC's WorldShare ILL, RapidILL, and Docline, with the numbers peaking in 2012 and then showing a slight decline.

Figure 3:
Growth for BD members is in circ-to-circ.

The second major trend I found in the Ivies Plus data was that all the growth in collection-sharing activity across the group was confined to the new members. *(See Figure 4 below.)* **Figure 4** shows all collection-sharing activity of the Ivies-Plus libraries, broken into three slices. The top light colored line, represents all the sharing activity of the three founding members of Borrow Direct. The second line represents all the sharing activity of the four institutions that joined Borrow Direct in 2002. Notice that they mirror each other, with almost exactly the same shape, and with the original group doing just a tick more sharing than the 2002 group. The line on the bottom shows the activity of those four institutions that joined Borrow Direct after 2011. Again, this is all sharing activity, with all partners, across all systems, borrowing and lending. Look how the volume takes off for this last group starting in 2011.

Figure 4:
Growth within BD confined to new members.
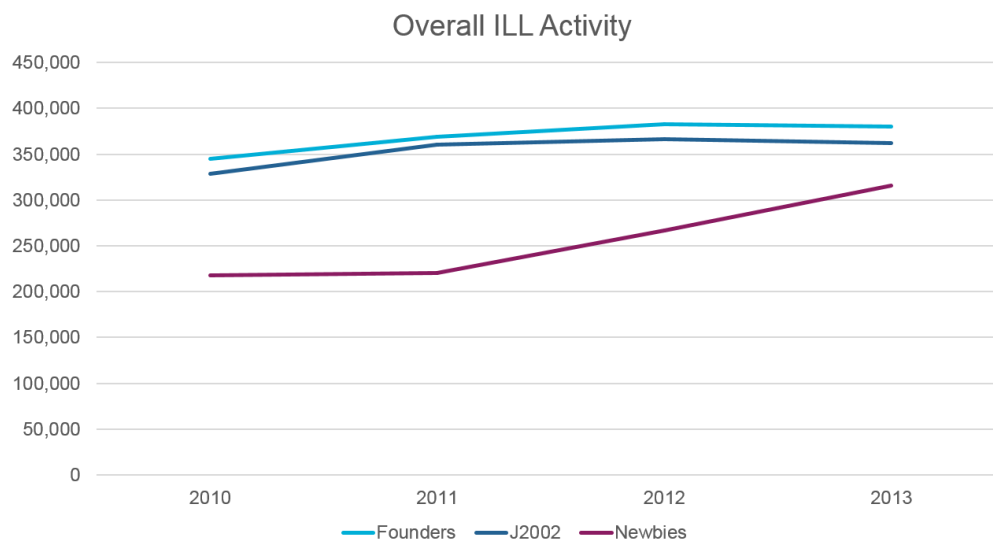
This surprised me. It's probably the one thing that surprised most of the Ivies Plus ILL staff. I know this because it was at this point in the study when I went back and conducted interviews with the ILL folks at each of the eleven Borrow Direct institutions.

If the Numbers Won't Tell You Why, Ask the People Behind the Numbers

From the beginning I suspected that staff interviews would be the most important and interesting aspect of this study. In the end, I was not disappointed.

I prepared individual snapshots of the data for each institution, showing their activity versus the group averages. Mostly this helped us find several mistakes – some mine, some theirs – that explained some of the differences in the data reported to ARL and to me. Then we spent an hour together via WebEx, talking about the answers to these questions:

- What is the strategic thinking behind the consortial groups you join? Technologies you adopt? Workflows you establish?
- What forces are at work driving your choices?
- Who are your users, and what needs do they hope will be met by your collection sharing services?
- What user experience do you offer? What do you aspire to offer?

- What would you like to learn from this data, or any data?
- How will collection sharing evolve over the next 5 years?

For the most part, their answers were exactly what you might expect them to be.

Why do they join certain groups, adopt certain workflows, embrace certain technologies? Their goal has long been to automate as much routine ILL traffic as possible. Arrangements such as consortial borrowing schemes and a journal-article-delivery system from the US called RapidILL *(Delaney 2012)* allow them to do just that. In both Borrow Direct and RapidILL, institutions commit to delivering (with expectations of a super-fast turnaround) certain portions of their collections. The routine stuff. Borrow Direct's Relais software shows real-time availability of the items, so requests are sent only to libraries where the material is available. RapidILL takes the holdings data for each library that joins and normalizes it so that a machine can route article requests to places where the volume and issue are actually owned. Both have fill rates approaching 95%. Both systems take staff largely out of the equation, and what tasks remain can be done by student workers. Both in consortial borrowing and in RapidILL, libraries arrange their routing tables so they trade with their best friends first, and then in ever-widening circles until they get to the occasional distant acquaintance.

What are some of the forces at work that drive the decisions that administrators at these libraries are making about collection-sharing? Their users are finding things. Lots of things. They have high expectations. Digital surrogates do not always suffice. No library can buy as much as they once did, even if they are spending more. Libraries need partners if they are to fulfill the needs of their users. ILL has earned a reputation for taking too long, costing too much, and yielding mixed results, all while being bounded by restrictive policies for the user. Monographs -- which are requested via ILL by a 2 to 1 margin over articles in the US – are still king, and they are available mostly in print. So consortial borrowing is now the absolute monarch of collection sharing models.

Who are the users, and what needs do they expect to be met by your collection sharing services? Nearly half the users are graduate students, with faculty making up the next largest group. They want what they want when they want it – preferably free of charge. They have high expectations. They think everything is fair use.

What user experience do you offer? What do you aspire to offer? Some of the Borrow Direct libraries have programmers on hand who have developed hidden algorithms that route user requests into the fulfillment method that makes the most sense, and which give users one place to check for all their use history. Others have started experimenting with off-the-shelf products that allow them to link up different systems they use, so that updating one updates the other. Some are still at a place where they stick a couple of buttons up on a Web site, explain what each of the buttons means – this is ILL, it can take weeks; this is Borrow Direct, it can take 2-3 days – and let the patron pick which one they want to try. Staff at all eleven institutions say that they share the aspiration of having one place where a user puts in a request, and the system routes it where it needs to go. They say they want this even as they lament that it will eliminate the branding of their service – the very effective and popular brand of Borrow Direct, which many patrons know by name and in fact happily use whenever talking about any resource sharing service the library offers, whether it is really a Borrow Direct partner that is providing the material they requested or not.

There were a number of things that the ILL folks wish they could learn, either from the data I've collected, or from what they've collected in the Borrow Direct data farm, or from any data anywhere. What percent of what's borrowed do they already own? Is there any correlation between materials expenditures or number of students and faculty and the collection sharing data? Who among their prospective users are they not serving, and why?

Staff at the eleven Borrow Direct institutions had a few predictions about what will come to pass in the next five years:

- Libraries will increasingly share more than just ILL requests. There will be shared collections. Shared facilities. Shared staff expertise.
- Some predicted that the decline in circulation will stop and possibly even reverse itself. But others felt that, going forward, the library would participate in a smaller proportion of the information-gathering transactions initiated by scholars.
- Print will stay popular.
- Special trusted partnerships will be needed if you're going to get the hard stuff.

Finally, the ILL folks at the eleven Borrow Direct institutions explained the causes behind the two big trends I had noticed in their data:

1. Consortial borrowing traffic is going up because the institutions are thinking of it as a way to provide access to supplemental copies of recent publications that are likely to be needed by their users. Many of these are things they already own. But the consortial borrowing workflows are so efficient that it is faster, easier, and cheaper to let a user borrow a copy of a book that is owned but checked out than it is to recall that same book from the current user. So many of these consortial borrowing requests were *not* things we would have seen ordered through the more labor-intensive, more restrictive ILL.

2. All the collection sharing growth is happening among the new members because Borrow Direct is a mature service at those institutions that are long-time members. The Borrow Direct market is saturated there. But the new members come on board and promote the new service aggressively, so of course they are the ones seeing all the growth.

Mysteries solved. Pinhole data explained.

There are still a few things left for me to work on with Borrow Direct, and most of it involves enhancing the picture of the group's collection-sharing activity with data from other sources. For instance, I plan to track interactions among the members where the requests go through OCLC ILL, because that's data to which I have ready access. It will be interesting to see how that traffic evolved as new members joined the Borrow Direct cohort. The Borrow Direct folks also want me to isolate returnable requests across all the systems, and nonreturnable requests across all the systems, rather than comparing OCLC ILL with RapidILL, and consortial borrowing with OCLC. Finally, they'd like for me to overlay data about materials budgets and relative size of the user bases to see if that correlates at all with the sharing activity.

Figure 5:
New BD member's OCLC ILL Interactions with BD partners

2011-2015

— Borr Orig — Borr Copy — Lend Orig — Lend Copy
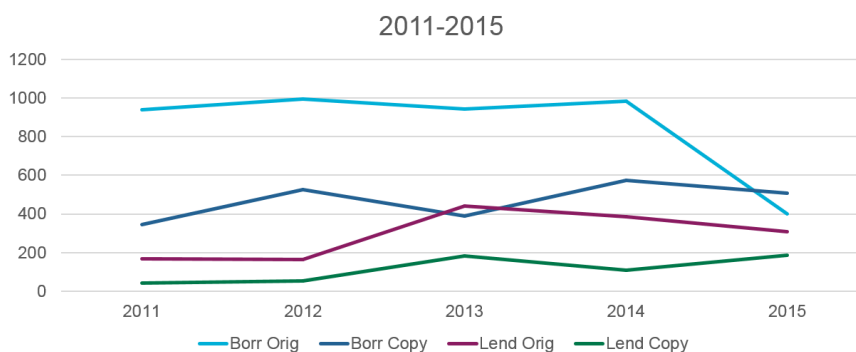
I've started on some of this "next step" work. For example, **Figure 5** above shows how one of the newer Borrow Direct members interacted with the other partners via OCLC ILL over the years. You can see a steep and sudden drop in the borrowing of originals, which is no surprise. But look how healthy the borrowing and lending of copies has remained, possibly because now the new member is tied more closely than ever with the Borrow Direct libraries, and interacts with them more frequently than before in all their resource sharing venues, not just Borrow Direct.

Having the Borrow Direct data has allowed me to take a long look at a pinhole reflection of one busy consortium's collection-sharing activity over a five-year period. Even though the numbers were self-reported and, in some instances, demonstrably inconsistent, I believe that they revealed the outlines of some significant trends that may shed light on what is happening across the entire spectrum of libraries – at least across the spectrum of large, well-funded North American research libraries. I believe that the data, rough as it is, gives us some license to start painting a portrait of what we think may be going on with Collection Sharing writ large. True, I'm working from a small and rather unscientific sample. There is really only one thing I can ever do about that: conduct more research.

If Having One Pinhole View is Good, isn't Having Two, by Definition, Twice as Good?

I'm in the midst of a second 5-year collection-sharing study with another US-based consortium called the Committee on Institutional Cooperation, or the CIC. Last summer I received five years' worth of data from 14 CIC institutions, who have a consortial borrowing set-up that they call UBorrow. Again, this is a group of large US academic institutions, so one could suppose that the results I would see from examining their data wouldn't differ that much from the Borrow Direct data. Right?

One difference is that doing the study was the CIC's idea, so they were more deeply involved in the planning and wrangling. Interestingly, they share a member with the Borrow Direct group. And they opted to run their consortial borrowing requests through the ILL infrastructure rather than mimicking circulation, which is what Borrow Direct does. So that's an interesting difference.

I mentioned earlier that I intend to overlay data on expenditures and demographics on top of the collection-sharing data, to see if there is any correlation. The graphs in **Figure 6** below show that there might possibly be some interesting differences between the two groups simply in terms of how they are funded. The Ivies-Plus institutions are mostly private institutions with large endowments, while the CIC is mostly made up of state schools that are publicly funded. You can see here the differences in expenditures per student for the two groups as of 2012.

Figure 6:
Per-Student Expenditures, 2012 NCES* Survey
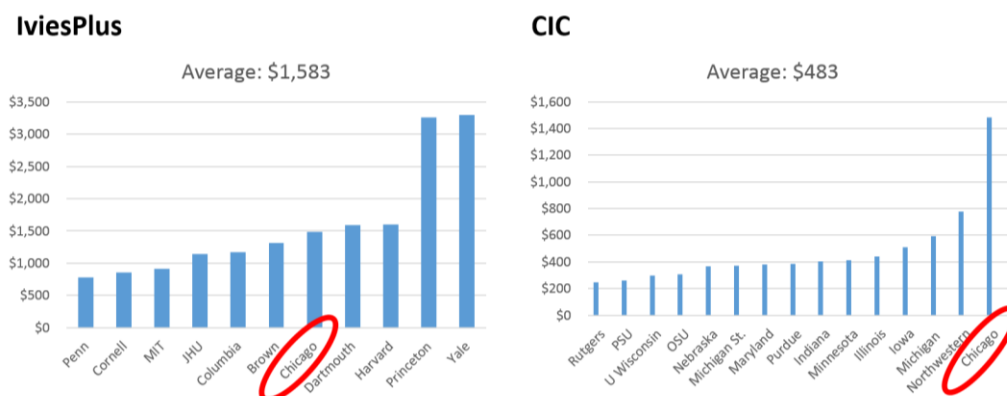
*National Center for Education Statistics

I'm at the beginning stages of investigating this sort of thing. And I've had my first pass through the CIC ILL data. I totalled it all up and threw it onto a chart, like I did initially with the Borrow Direct data. And guess what?

It's nothing like the Borrow Direct data. Total activity is trending *down*.

True, the consortial borrowing among the CIC group is skyrocketing. So in this one way the CIC experience resembles what we saw with the Borrow Direct group. But remember, this is a fresh service to the CIC, being heavily promoted by most of the partners. If you look at the consortial borrowing activity of the CIC libraries in schemes other than UBorrow, we see, at least in the lending, a distinct downward trend. *(See Figure 7 below.)* Excuse me? What happened to consortial borrowing being the king of collection sharing models? What's going on here?

As of this writing, I haven't spoken to the CIC ILL folks yet. But it seems that many of them are involved in state consortial borrowing networks where the level of activity is dropping. This line of inquiry is worthy of some further attention.

Two pinhole views really *are* better than one.

Figure 7:
Other CIC Consortial borrowing trending down.

CIC Other Consortial Borrowing Activity
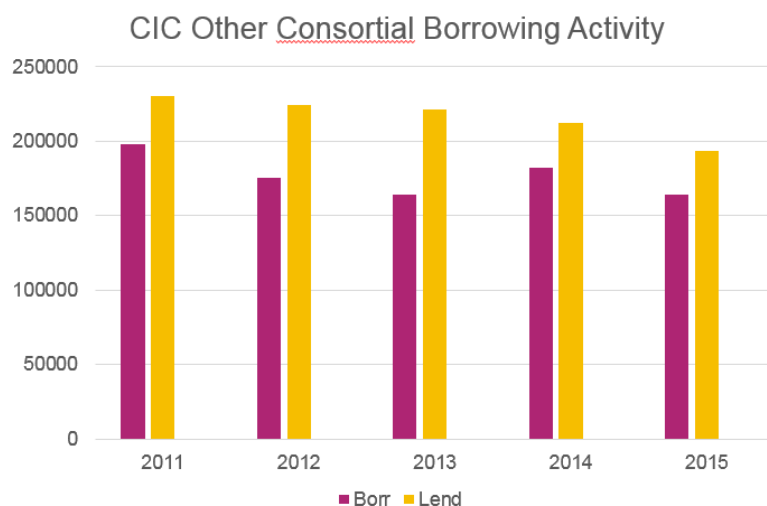
In the meantime, I've been hearing a different sort of question about the various models of resource sharing, particularly regarding consortial borrowing. Sure, folks are still wondering if collection-sharing traffic is going up or going down, or if it is switching from one model to the next. But increasingly people have been asking me what the various models and methods cost.

At this moment, I don't know the answer to that question. But I know how we can find out.

Enter the OCLC ILL Cost Calculator – a Different Type of Pinhole

This project has been about five years in the making. It started with the SHARES resource sharing consortium's pricing committee debating whether or not it is crazy to value the provision of a book the same as the provision of an article. It was Margarita Moreno of the National Library of Australia's idea to take that question a step farther. Why not design a tool that would answer many such questions, for many years to come?

Wish granted.

A SHARES working group was formed, and after several years of thrashing around we are finally closing in on having a functioning tool. We basically want this thing to function as a real-time, virtual ILL cost study. It will cover current models that weren't around when the old studies were done. It will help us identify the characteristics of high-performing operations, and to share those as suggested best practices. And it will allow real-time comparisons of your own unit costs with those of anonymized peers.

We started out by imagining various use cases for our tool. We wanted it to be something that folks could use to help them make practical decisions about how to conduct their operations. The success of the tool will depend largely upon the willingness of library administrators to allow their data to be included. We realize that all of the cost data, but especially the salary information, is quite sensitive. Only those submitting data will be able to see their own sensitive stuff. Detailed salary information will not be sent to the calculator. No report except each institution's private reports will identify any data as belonging to that institution. The privacy of individuals, institutions, and service providers has been uppermost in our minds all the time we've been working on this.

The beta testers are putting the registration and data submission processes through their paces right now. Within a few months we hope they'll be testing out the reporting functions. And then we'll be asking for a second wave of early goers to try out the whole process after we've added improvements suggested by the betas.

Stay tuned for more on this. You can Google "OCLC ILL Calculator" to keep current on developments.

Wouldn't the "Drinking from a Firehose" Metaphor be Faster?

It is frustrating to live in a world where so much library data exists and to be able to get my hands on only a tiny portion of it. But this type of research is a little bit like that popular definition of politics that is often heard: it is the art of the possible.

The pinhole approach to studying the collection-sharing universe has not been perfect or comprehensive, but it has generated some interesting food for thought. I am very grateful to the administrators and staff at the CIC and Borrow Direct institutions for so generously providing me with their data and taking the time to answer my questions and to think big thoughts with me. My hope is to continue to gain access to transactional resource sharing data from more groups of libraries. Moving forward it would be interesting to concentrate on different types of libraries than those in the CIC and Borrow Direct – libraries in other countries, public libraries, libraries that are not so well funded. Attracting many types of libraries from all over the world to enter their data into the ILL Cost Calculator will also be an important focus for me in the coming months.

With each added piece of data that comes in through the pinhole, hopefully the image of the collection-sharing universe that is being projected upside down upon the opposite wall will come into an ever sharper focus. And if we're lucky we can trace it quickly enough to capture a good likeness before that universe evolves into something not yet imagined.

References

Batt, Steve. (2012) Public Tableau profile. Academic Libraries Survey 2012, National Center for Education Statistics, U.S. Department of Education, interactive Web site. https://public.tableau.com/profile/ct.census.data#!/vizhome/AcademicLibrariesSurvey2012/Story1

Birch, Katie, and Dennis Massie. (2014) *The Borrow Direct Cohort: a proposed survey of the group's overall resource sharing activity.* Unpublished presentation to the Borrow Direct Policies Committee, June 24, 2014. Data updated via email from OCLC product manager Tony Melvyn, February 9, 2016.

Delaney, Thomas G., and Michael Richins. (2012), "RapidILL: an enhanced, low cost and low impact solution to interlending," *Interlending & Document Supply*, Vol. 40 No. 1, pp. 12-18.

Nitecki, Danuta, A. and Jones, Carol, L. (2004), "Borrow Direct: its impact on service quality at Yale University library," *Interlending & Document Supply*, Vol. 32 No. 3, pp. 146-151.

Phan, T., Hardesty, L., and Hug, J. (2014). *Academic Libraries: 2012* (NCES 2014-038). U.S. Department of Education, Washington, DC: National Center for Education Statistics. Retrieved July 3, 2016 from http://nces.ed.gov/pubsearch.

Roche, Marilyn M. (1993) *ARL/RLG Interlibrary Loan Cost Study: a joint effort by the Association of Research Libraries and the Research Libraries Group.* Washington, D.C.: The Association, 1993.

Steadman, Phillip. (2001) *Vermeer's Camera: uncovering the truth behind the masterpieces.* New York: Oxford University Press, 2001.