Does (No) Discovery Lead to (ILL) Delivery?
Comparing Transaction Logs to ILL Requests to Determine the Persistence of Library Patrons

Sherri Michaels, Indiana University Libraries
Wells Library 241G
1320 E Tenth Street
Bloomington, IN 47401
shmichae@indiana.edu

Rachael Cohen, Indiana University Libraries
Herman B Wells Library W501
1320 E Tenth Street
Bloomington, IN 47401
rachcohe@indiana.edu
Does (No) Discovery Lead to (ILL) Delivery?

ABSTRACT

This study looks at known-item searches within a university catalog that come up with a “no results” response to see whether patrons will then place an ILL request for that material and how persistent a user might be in order to obtain something that they want. Catalog search transaction logs were compared against interlibrary loan (ILL) requests placed during the same time frame. Overall, the rate of persistence, or the frequency with which ILL requests were placed when no results were found in the catalog is very low.

Key words: Interlibrary loan, Search queries, Know-item searches

INTRODUCTION/BACKGROUND

If a patron is searching for an exact title, they have generally already done some research to get to that point – previous keyword searches or they've read an article that cites the title. So, by searching the catalog for a specific title, they are presumably interested enough to want to know if the library owns the item. If the library does not own it, what are the next steps? Will they see interlibrary loan as a viable option or will they give up and move on to something else that is theoretically more readily available? How frequently does it occur that the library does not own what is being sought? This study compares transaction logs to ILL requests to answer some of these questions.

The Environment

Indiana University Bloomington (IUB) is the flagship residential campus of the eight campus IU system. The Bloomington campus consists of approximately 49,000 students (39,000 undergraduates, 6,000 graduate, and 4,000 doctoral), 3,210 faculty, and 5,600 staff (Indiana University, 2016). IUB offers more than 550 degree and programs ranging from associate to doctorate and has a Carnegie classification of Doctoral Research Extensive (Indiana University Bloomington, n.d.). The Indiana University Libraries collections include more than 9.9 million books in over 450 languages, journals, maps, films, and sound recordings. The library subscribes to approximately “1,871 databases, 60,000 electronic journal titles, and 1.9 million electronic books, as well as locally developed digital content” (Indiana University Libraries, n.d.). These collections can be accessed through the shared SirsiDynix Symphony catalog, IUCAT, which uses the open source discovery layer Blacklight for its front end. IUCAT allows users to search for known-items, single-word titles, phrase searching using various indexes (author, title, and subject) along with an advanced search, which allows users to search in nine different indexes. The search results page includes search limiters called facets that help users refine their search and quickly zero-in on the most relevant results. Each search performed by the user on the user interface is recorded in a transaction log on the back end of the server.

Despite these rich resources, no library can meet all their patron needs, so like many other universities;

1 Faculty include tenure track Administrators, Librarians, Researchers, Visiting Scientists and Academic Specialists
IUB also provides interlibrary loan services. The Document Delivery Services (DDS) department provides a suite of services including interlibrary loan, document delivery, and support for distance education programs to all IU Bloomington students, faculty, and staff. In fiscal year 2016-17, the department received over 21,000 borrowing requests, over 52,000 lending requests, almost 10,000 document delivery requests, and over 500 distance education delivery requests. IU consistently ranks in the top ten lenders on OCLC and in the top twenty borrowers in terms of interlibrary loan volume. While IUCAT is shared resource across the IU System, each campus provides their own ILL services.

To manage this volume of requests, DDS uses ILLiad, a resource sharing management software product developed by Atlas Systems and now distributed through OCLC. While ILLiad allows users to enter information into a web form to request items, one of its key features is that the software also integrates with other automated systems to allow bibliographic citation information to be imported into the system without a user needing to re-key any of the information. This facilitates the interlibrary loan process greatly, and IU has added the “request this” button to as many places as possible to allow patrons to request items with minimal interference.

LITERATURE REVIEW

A review of the literature reveals several studies that have examined a library user’s ability to search, and find, what they are looking for. Early shelf availability studies, and later, transaction log analyses have identified barriers to finding a known item within the library so that those obstacles can be reduced or eliminated. Kantor (1976) developed a mathematical model to determine the probability of user satisfaction based on the findability of library items. He identified several areas where a user was likely to encounter an issue starting with an item not being acquired and progressing through other problem areas such as circulation and shelving, cataloging errors, and user errors. Ciliberti et al. (1987) built upon Kantor’s model and conducted a similar study at the William Paterson College Library in New Jersey to determine why users could not find material on the shelves. One interesting finding was that for known-item searches, “of the 206 patrons who had correct information, 21 (19 %) were searching for titles the library had not purchased.” (Ciliberti, Cassery, Hegg, & Mitchell, 1987, p. 519). With this result, the scope of “acquisition errors” (Ciliberti et al., 1987, p. 522) becomes known and we begin to see how often users are searching for things that are not owned by the library and thus might lead to an interlibrary loan request.

More recent studies have focused on analyzing transaction logs to better understand user-searching behavior. Transaction logs capture the searches performed by users in the online catalog, and can be used to look at exactly what search terms are entered or how a query is phrased. Specht (1980), Peters (1989), Hunter (1991), Wyly (1996), and Behnert and Lewandowski (2017) all examined transaction logs in order to determine where users were encountering problems in searching an online catalog. These studies show that between 40 and 54 percent of all catalog searches result in zero hits. Searches resulting in zero hits were often described as a “failure” but Peters (1989) reminds us “a zero-hit search does not necessarily mean that the patron had a problem using the system, that the search was a failure, or that the patron went away frustrated “(Peters, 1989, p. 271). He goes on to say that “in reality, a user may conduct a good search and be perfectly content to learn that the library does not own the item sought” (Peters, 1989, p. 272). Indeed we are very interested in knowing just how often this occurs, and correspondingly, if the patron in turn submits an ILL request for the item.
Does (No) Discovery Lead to (ILL) Delivery?

At the intersection of interlibrary loan and the usability of an online catalog sits another set of studies illuminating patron-requesting behavior. Opposite the aforementioned acquisition issues, many of these studies begin by asking why patrons place interlibrary loan (ILL) requests for items that are owned and locally available, and then expand into investigating the reasons behind such behavior. Patrons frequently request items through ILL that are owned at the home institution, and these requests have made up anywhere from six and a half to twelve and a half percent of all ILL requests within a unit (Dwyer, Gossen, & Martin, 1991; Kress, Del Bosque, & Ipri, 2011; Page & Kuehn, 2009; Seaman, 1992; Yontz, Williams, & Carey, 2000). ILL librarians have been investigating for years ways to decrease this workload on understaffed departments as well as to improve patron satisfaction by getting materials to them faster. An early study by Seaman (1992) points out that using ILL requests as the basis of trying to determine catalog errors offers several advantages, not the least of which is that “each request unquestionably represents what the user believes to be a known-item search” and that it is reasonable to assume the patron has gone to some length the search the item locally” rather than wait an extensive period for delivery (Seaman, 1992, p. 113). Seaman (1992) and Dwyer, Gossen, and Martin (1991) reported several areas where a patron could go astray and not find the item such as in incorrect citation, in circulation, special location, as well as cataloging errors and incomplete cataloging records. However, for some subset of the requests, there was no discernable reason that that patron would not have found the item in question. Yontz, Williams, and Carey (2000) discussed similar findings, but also noted that catalogs had improved findability from earlier studies by noting that “we found that the catalog records largely contained no errors that would prohibit retrieval” (Yontz et al, 2000, p.125). So, if catalogs have improved findability, then why are patrons still placing requests for items that are owned? Yontz et al. found that 11 percent of requests came from patrons who “demonstrably knew” that the library owned the item (Yontz et al, 2000, p.125). These patrons did not wish to “use it in its format (microform) or did not want to wait for a recall or shelf search” (Yontz et al, 2000, p.122) as the policies in place prevented these types of requests. Page and Kuehn (2009) also examine several reasons why these requests still get placed. They found that “missing or problematic links [to full text] occurred more frequently for publications for which interlibrary loan requests were canceled due to local or electronic availability” (Page & Kuehn, 2009, p. 487). Thus, we begin to learn why users may place an ILL request, even if is for an item owned; however, little research has been done on how often an ILL request is placed when a user knows, or at least believes, an item is not available. If ten percent of requests are placed for items locally owned, then ninety percent are placed when an item is not held. But how often are users searching the catalog to find a zero search result before placing the request? How many zero results are there? What percentage of zero-result searches will turn into an ILL request? Will they simply move on and find something more readily available or persist in obtaining the item? Our study is the first step in trying to answer these questions.

METHODOLOGY

In the fall of 2015, we were able to work with our library’s developers to get a custom logging script that generated an anonymous transaction log any time a patron received the “no results page” (see image 1) in the union catalog section of Blacklight, the IU Libraries catalog interface. These transaction logs were sent to us in at tab-delimited file on a weekly schedule and showed the date and time of the search query, the index the search was performed in, any facets applied, and the exact search query performed using a 24-hour clock. These files were then imported into Excel as a workable format (see table 1).

Users are not required to log in to search IUCAT, which means search data is captured with no identifying or demographic characteristics of the user who performed it. IU has a shared catalog
Does (No) Discovery Lead to (ILL) Delivery?

amongst its eight campuses. IU Bloomington sets its default searches to the union catalog whereas the majority of the other campuses restrict their search results to a “campus view” showing only results held locally. For this study, searches performed in the campus views were not included in the collected search data.

IMAGE 1: IUCAT No search results page

![IUCAT No search results page](image)

TABLE 1: Formatted IUCAT zero search query data

<table>
<thead>
<tr>
<th>Date</th>
<th>Query Type</th>
<th>Query Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-10-05T10:29:03</td>
<td>author search with query</td>
<td>Hotle</td>
</tr>
<tr>
<td>2015-10-05T10:30:27</td>
<td>title search with query</td>
<td>drug court effectiveness: a matched cohort study in the dane county drug treatment court</td>
</tr>
<tr>
<td>2015-10-05T10:37:22</td>
<td>title search with query</td>
<td>&quot;The chemistry of cereal proteins. 2nd ed.&quot;</td>
</tr>
<tr>
<td>2015-10-05T10:43:05</td>
<td>title search with query</td>
<td>Atlanta Constitution</td>
</tr>
</tbody>
</table>
We then decided on a month in which we would compare the transaction logs to the ILL requests placed within the corresponding timeframe. October was chosen as this is usually one of the busiest months in terms of library usage and ILL requests for Indiana University Bloomington, giving us a larger data set to examine. The study was then run two years in a row in 2015 and 2016.

ILL data was captured using an SQL query of the ILLiad database and included timestamps using a 12-hour clock, request type, titles, authors, and transaction status.

The raw data for October 2015 included:

- **Number of ILL request in October:** 3,551
  - Number of Articles: 2,060
  - Number of Books: 1,491
- **Number of zero search results in IUCAT in October:** 69,670

The raw data for October 2016 included:

- **Number of ILL request in October:** 2,703
  - Number of Articles: 1,220
  - Number of Books: 1,483
- **Number of zero search results in IUCAT in October:** 79,433

Clean up of the data was the next step to be able to compare the data sets. The focus of this study was known-item searches to more easily compare those results against specific interlibrary loan requests. The transaction log data included all zero-result searches, not just those for specific items so the results needed to be filtered. The data from the IUCAT user logs, once imported into Excel, started out in an unusable form for analysis and required significant work to clean up. A student worker was hired in spring 2016 using a grant to do the initial clean-up and filtering of unneeded data for the 2015 data. Before we started the comparison, we first filtered out the following categories, as they did not fall into the "known item" scope of this project, or contained corrupt data.

1. Blank searches - a number of blank searches were recorded in the logs. These could be removed as there is no data to analyze.
2. Hyperlinked searches - hyperlinked search performed from inside detailed bib records which would not be used to find "known items" but instead used for browsing. These include Author,
Does (No) Discovery Lead to (ILL) Delivery?

Author_Title, Series, and Subject links. In 2015 and 2016, these were keyword searches instead of exact phrase searches.
3. Subject search with query - subject searches were considered “browsing” searches for this project.
4. Double dash searches – subject searches that were not performed in a subject index (i.e. “Industrial management--Employee participation--United States--Congresses”)
5. Kinsey search - an advanced subject search, searching only the Kinsey subject index.
6. Facet only - searches that contain zero search terms, and instead only search a facet value (only available for 2016 data)
7. Browse search - searches performed in the Authority Browse search. Contains Author, Series, Subject browse searches (only available for 2016 data)
8. Reserve searches - Course Number, Course Name, and Instructor searches performed through the Reserves search (only available for 2016 data)
9. Obvious gibberish/garbage - searches that were not captured correctly (i.e. æŽæ—œ—èfåÎŽé—¨ä, “ä¸šæ·æœæ·œ®æœè‰œ©æ¢Žç…” (weibo.com%2Fp%2F1005051584740914%2F)), (only available for 2016 data).

This criteria of filters removed 36,867 searches immediately from the 2015 data set and 57,111 from the 2016 (table 2). This left a total of 32,803 zero search result searches in 2015 and 22,322 in 2016 for comparison with ILL data.

<table>
<thead>
<tr>
<th>Search Type</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blank searches</td>
<td>13,555</td>
<td>10,266</td>
</tr>
<tr>
<td>Author Link</td>
<td>8,720</td>
<td>3,269</td>
</tr>
<tr>
<td>Author_Title Link</td>
<td>882</td>
<td>110</td>
</tr>
<tr>
<td>Series Link</td>
<td>1,119</td>
<td>2,654</td>
</tr>
<tr>
<td>Subject Link</td>
<td>8,288</td>
<td>3,411</td>
</tr>
<tr>
<td>Subject search with query</td>
<td>1,069</td>
<td>17,778</td>
</tr>
<tr>
<td>Double Dash subject</td>
<td>3,215</td>
<td>2,994</td>
</tr>
<tr>
<td>Kinsey search</td>
<td>19</td>
<td>80</td>
</tr>
<tr>
<td>Format only</td>
<td>NA</td>
<td>6,504</td>
</tr>
<tr>
<td>Browse search</td>
<td>NA</td>
<td>101</td>
</tr>
<tr>
<td>Reserves search</td>
<td>NA</td>
<td>16</td>
</tr>
<tr>
<td>Obvious gibberish/garbage</td>
<td>NA</td>
<td>9,928</td>
</tr>
<tr>
<td>TOTAL</td>
<td>36,867</td>
<td>57,111</td>
</tr>
</tbody>
</table>

After the initial clean-up and filtering, the data was prepared for Fuzzy Lookup in order to match searches and ILL requests for the comparison. Fuzzy Lookup is a free add-on function for Excel that is capable of searching for approximate matches between tables with specified minimum match indexes. The function automatically joins matched data and creates a new table with the associated index for each pair of matched cells.
Does (No) Discovery Lead to (ILL) Delivery?

We separated search terms by category using the index names from IUCAT, creating worksheets of just “Title search”, “Author Search”, “Numbers”, and ALL. This allowed the Fuzzy Lookup function to be more efficient, but also made it possible to catch search terms that are ambiguous and might have shown up in one of the three ILL title categories. Finally, the similarity threshold was set at greater than .85.

We then ran each of the spreadsheets against its compatible type of ILL data and then filtered to display results above .85. For example, LoanTitle vs. Title Search and, LoanTitle vs. ALL.

Each match was reviewed to verify that it was actually a correct match. We then assigned a color to the result based on a timeframe (see table 3). Initially, we only noted ILL requests that were placed within 24 hours or one week following a corresponding search. However, we noticed a pattern of ILL requests being placed just before a corresponding zero-result search, so we went back and highlighted those cases as well. Finally, each spreadsheet was checked remove any searches that may have been double-counted.

<table>
<thead>
<tr>
<th>ILL Request placed within 24 hours of zero results found query</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILL request placed within 1 week of zero results found query</td>
</tr>
<tr>
<td>ILL request placed first, zero results query less than 10 minutes later</td>
</tr>
</tbody>
</table>

**DISCUSSION**

**Zero Search Results**

Of the 32,803 catalog searches with zero results in 2015, only 512 (1.56%) corresponding ILL requests were placed within the same month. 2016 showed a slight increase of 640 ILL requests placed out of 22,322 (2.87%) zero search results. Thus, an average of 2.22% of zero search results turned into an ILL request. Most of the ILL requests were actually placed outside of the timeframes that we expected, which was one day. We anticipated that most users would search the catalog and then place an ILL request immediately. We thought that a few people would stretch out the requests over a day or two if they were doing a more thorough literature review so we also intended to track requests up to a week after the zero search result. What was found though was that a number of requests were actually placed outside of the first week but still within the month. What was most surprising was that we quickly noticed a pattern of an ILL request being placed immediately before the zero result search.
Does (No) Discovery Lead to (ILL) Delivery?

CHART 1: 2015 and 2016 Search vs. Request timeframe results

There are a few explanations as to why a search might have occurred after the ILL request was placed. One is that the patron placed the request and then immediately thought they should have searched the catalog first to verify it wasn’t owned. Another is that the search was actually placed by ILL staff upon receiving the request. A significant number of these occurred outside of the working hours of the ILL department, so that could only explain a few of these occurrences. Finally, a possibility is that the time stamps on some of the computer logs were off by enough to make it appear as though the ILL request was placed first. In many cases, it was less than a minute between the two events.

**ILL Requests**

To look at the results from the ILL request perspective, of the 3,551 ILL requests placed in 2015, 512 (14.45%) had a corresponding zero result search in IUCAT within the same month. In 2016, there were 2,703 ILL requests placed. 640 (23.67%) had a corresponding zero result search in IUCAT within the same month.
Does (No) Discovery Lead to (ILL) Delivery?

CHART 2: ILL Requests vs. Zero Search Results

This means that in 2015, 85.58% of ILL requests were not searched in IUCAT before a request was placed, and 76.32% in 2016. There are some contributing factors to this number. There are other services our patrons can use to check for availability of materials such as Worldcat.org, or our EBSCO Discovery Service. Still, given that we limited this study to focus on books, and known-items at that, we expected to see more correlation between a catalog search and an ILL request. So, where are we losing our users? One explanation is that currently in IUCAT, when a user hits the zero search results page for a query, they are presented with a page suggesting they modify their search (See image 2). There is no “did you mean” for spelling errors, no indication to try our discovery service for an article title instead, nor the suggestion of placing an interlibrary loan request instead. Users have to be conscious enough of why their search failed, and what other options the Libraries offer them in locating and retrieving the item.
Does (No) Discovery Lead to (ILL) Delivery?

IMAGE 2: IUCAT No search results page

We have made some improvements though to make ILL more visible. The fraction of ILL requests that had a corresponding zero result search in IUCAT was 64% higher in 2016 than in 2015. One possible reason is that in summer 2016, we redid the library homepage and added ILL as one of the top 4 buttons at the top (see image 3). ILL was also added to One.IU as a task. Indiana University developed the One.IU as a site to corral all of the important services available to students, staff, and faculty into one central place. Having the ILL on both the Libraries website homepage and in One.IU not only makes it easier for users to find it, but it brings attention to the service to those who were unaware of it before.
Why Searches Failed

Given past studies as discussed in the literature review, we also know that users cannot always find owned materials in the catalog even if they do look there. So, we also noted the number of ILL requests that were canceled due to ownership and availability. In 2015, 14.3% of ILL requests were canceled because we owned the item and in 2016, it was 7.3%. These results are in line, or even a bit higher than, with other studies. As we investigated these canceled requests and the zero search queries, we found that three user errors are the most likely cause for a user to request an item through ILL that the library already had in full text/print. These are: typographical errors, article searches, and wrong index use. Overall, the authors found a total of 2,489 typographical errors, 1,603 article searches, and 234 searches performed in the wrong index.

When looking for article searches the authors focused on obvious article titles, and searches that contain section (or entire) citations. For example, “Miller, K., Gault, B., &Thorman, A. (2011). Improving child care access to promote postsecondary success among low-income parents.” Indiana University does not index/load individual article titles into IUCAT. Users would need to know to use either our
Does (No) Discovery Lead to (ILL) Delivery?

Discovery Service (EBSCO Discovery Service) or another database to find these items. Out of the 32,803 searches checked, the authors found 1,603 article searches, 55 of these also contained a spelling error.

Finally, for wrong index searches, the authors found that users selected the wrong index 234 times. “All Fields” is by far the most common used index in IUCAT. It is the default for Basic Search and the first search box for Advanced Search. Users would have to deliberately select one of the other indexes to use it. Users selected a specific index (Title, Author, and ISBN) 8,862 times out the 32,803 zero search results (27.02%). Of those 8,862 searches, users searched in the wrong index 234 times (2.64%). For example, an Advanced Search in the ISBN index of “ISBN: American cameraman” will bring back zero results as it is searching marc tags which do not contain text. If the user had searched for this item in the Title or All Fields index, they would have found the video they were most likely looking for.

Limitations of this study

Search log analysis can be an economical and unobtrusive method for gathering large amounts of user interaction data to help with understanding information-seeking behaviors. However, as with any research method, using search log analysis it is not without its limitations and in the context of this study.

One major limitation is that demographic data is not collected in the IUCAT search results logs. As a result, this study excludes any demographic analysis which limits what we can conclude from the results. Without the patron data we do not know what campus they are truly searching from. The majority of the regional campuses default to their “campus views” in IUCAT, but there is nothing preventing them from searching in the union catalog. This means that some of the zero searches may not have a match to our ILL requests, because the patron performing the search belongs to another campus and would use a different ILL service. There is no way to tell what percentage of these searches could fall under this scenario. We also do not know if the same person doing the zero-result search placed the corresponding ILL request. With many of the ILL requests placed outside of the 24 hours of the zero-result search one has to question how many of these are actually coincidental and not performed by the same patron.

Another limitation is that we had to set an arbitrary time range for our study, and that may skew our results. This means that a search placed on the last day of October might not have had a corresponding ILL request because it was placed in November, or ILL requests placed early in the time frame may have had a corresponding search previous to our time frame. We also initially thought that we would look at ILL requests placed up to a week after a zero-result search, but we found that there were several examples of weeks between a search and what could be a corresponding ILL request.

Finally, a study of this type is extremely time intensive. The data required the authors to do a significant amount of cleaning and processing just to get to a point in which it was readable. After the cleanup it then further required time and effort to sort out the undesired entries, complete the comparisons, identify the searches in the correct time range, and strip out any duplication. Because of the different methods of recording time uses, each search query had to be individually checked to see if it fell in the correct time range.
CONCLUSION

While this study begins to shed light on how frequently users turn to ILL when they do not find an item in the catalog, there are many more questions yet to be explored. For example, where are these users going after they get a zero search result? Do they give up, ask a librarian, or just change their topic? That might depend on their status (faculty or grad or undergrad) but more study is needed there. Additionally, how long do patrons typically wait between searching the catalog and placing a request? We found some interesting results between the search and a possible ILL request, but do not know if they were placed by the same person who did the search. The next step will be to survey ILL users to explore some of these topics.

DATA

The data that support the findings of this study are openly available in IUScholarWorks at http://hdl.handle.net/2022/22102
REFERENCES


Does (No) Discovery Lead to (ILL) Delivery?


