This article is based on a presentation by Aaron Shrimplin, Andy Revelle, & Susan Hurst, "Seeing the Forest by Counting the Trees: Using a Variety of Data Sources to See the Big Picture" at LITA Columbus, OH Oct. 6, 2012.

"Seeing the Forest by Counting the Trees: Using a Variety of Data Sources to See the Big Picture"

Susan Hurst, Business Librarian, Miami University, hurstsj@muohio.edu 219A Laws Hall, Miami University Oxford, OH 45056

Andy Revelle, Social Sciences Librarian, Miami University, revellaa@muohio.edu 206 King Library, Miami University Oxford, OH 45056

Aaron Shrimplin, Assistant Dean for Reference and Collection Services, Miami University, shrimpak@muohio.edu, 221 King Library, Miami University Oxford, OH 45056

ABSTRACT

We are swimming in a sea of data. Librarians often contribute to this by counting every possible patron interaction in an attempt to both define their current situation and to predict future staffing, budgetary, and collection needs. This investigation assessed the effectiveness of using various data sources in predicting future library activity and needs. The authors collected data on in-person and chat reference transactions, electronic journal downloads, database queries, and catalog searches from 2009-2012. By analyzing these data points, the authors hypothesized they would find correlations that might be predictive of changes in related library services. Results indicated that the strongest correlations track activity over the course of the academic calendar. While none of the data points examined had strongly predictive properties, the strong correlations between the data points themselves over the period of time studied indicated that any one of them might serve as a stand-alone indicator of usage.

Keywords – data, library services, leading indicators, lagging indicators, normalization of data, reference, chat reference, e-journals, database usage.

INTRODUCTION

It is estimated 2.5 quintillion bytes of data are created every day, 90% of the data in the world today was created in the last two years (IBM 2012). By any measure, this is a vast trove of data waiting to be harvested to find new insights and to answer important questions about the world in which we live. Much of this data is a result of the Internet and advancements in information technology. Many fields, from health care to economic forecasting, are using data analytics to extract value from new streams of data. As the ability to track and store data increases, the natural inclination is to collect as much information as possible, in hopes that it will prove useful in both day-to-day management as well as in the bigger picture. As librarians and researchers, this compunction to collect and analyze various data streams seems stronger than ever. The authors analyzed various data sets and hypothesized that correlations could be identified which might prove to be predictive or leading indicators. Ideally identifying factors that could be used predict later events would have practical value by enabling improvements in concrete areas such as staffing of service points, allowing continuing improvements in customer service while keeping staffing levels lean.

Miami University of Ohio is primarily a residential undergraduate teaching institution, with approximately 16,000 students on the main campus in Oxford, OH. There are four libraries on the Oxford campus, with King Library, housing the humanities and social sciences, being the largest. Based on experiences as reference librarians working in King Library, the authors formed two hypotheses.

One hypothesis was that Web site hits or chat reference transaction data might serve as leading indicators for traffic at service points. Second, by comparing various types of data across both Fall and Spring semesters, the researchers' goal was to determine if preconceived opinions about levels of activity and how they changed over time were accurate. In examining these statistics the authors also hoped to reduce the amount of data collected and stored locally. By eliminating the collection and management of superfluous data, staff time currently used for these tasks could be reallocated.

LITERATURE REVIEW

Library data analysis is not a new topic. A search in the Library, Information Science & Technology Abstracts database for "data' and "libraries" as subject terms returned over 2,900 hits, dating back to 1965. However, the field of data analytics in libraries is much more recent; with entries for "data analytics" going back only to 2001. About the same time, phrases such as "data driven decision making" and "evidence-based management" also started to be seen in the library literature. Following trends in business and technology, "big data" is now rearing its many heads in academic librarianship and elsewhere. With increases in computing power (and more importantly, data storage), along with the fact that so much work is now done online, the amount of data available for investigation or supervision has increased enormously. The trick, as always, is determining what should be done with it.

Looking at previous studies, data have been used to evaluate areas of library services that had not been explored fully. For instance, studies focusing on reference desk statistics often investigated the reference transaction themselves, rather than the bigger picture of overall library services (McLaughlin 2010; Garrison 2010; Mosley 2007; Todorinova et al. 2011). Jean McLaughlin found that 55-56% of New York State libraries using reference transaction data

included it in staffing and service hour decisions, meaning 44% did not incorporate this data. Only 21% were currently using it for collection development decisions (McLaughlin 2010, 12).

Circulation statistics are one of the original sources of quantitative data used for library decision making and are still used today. Librarians at Duke University Medical School used circulation data to analyze their monograph collection to determine how to improve selection decisions (Grigg et al. 2010). Several types of data, including circulation statistics and study room usage data, were examined by researchers in Japan and Korea for purposes of library marketing (Minami and Kim 2010). With increasingly large sections of collections available online as e-books and e-journals, Locke Morrisey (2010) examined usage of online resources compared to their costs, which is one of the most common uses of data in decision making.

Database usage is another source for data that drives research. Despite efforts by the COUNTER group to make databases COUNTER-compliant and standardize statistical reporting through their "Code of Practice for e-Resources," there are still many problems with database usage statistics (COUNTER 2012). Research by Gayle Baker and Eleanor Read (2008, 52) found that "more time was spent on the non-intellectual processes of gathering and preparing vendor-supplied usage data than on the analytical processes that lead to an understanding of an institution's use of its electronic resources."

Some newer sources for quantitative data include library website analytics and statistics generated by students' "swipe-card" usage. Tabatha Farney (2011) discusses various click analytic tools and how they can be used to analyze web site design, by showing which links are most or least used. Jennifer Jones (2011) at University of Georgia describes a system which requires students to swipe ID cards to enter the library. Besides enhancing security, this provides a rich source of quantitative data that has been used for staffing, operational, and marketing

decisions. These types of data are increasingly generated by various administrative processes and are typically generated automatically, so there is no additional effort or cost involved in its collection (Chapman and Yakel 2012). However, using it effectively can still be challenging.

Another source of administrative data comes from chat reference. Most systems automatically collect data about each transaction, so as chat services have taken off, they have also provided a fire hose of data. Librarians are now starting to investigate this data, both for research purposes and to use it in decision making. Mississippi State University libraries included chat transcript data in their analysis of how best to overhaul their library web site, garnering usability information based on the questions asked through chat (Powers et al. 2011). Mississippi State University libraries also gathered data from surveys of email and chat reference users to examine who was using these services in an attempt to better publicize them (Nolan et al. 2012). University of Minnesota was a pioneer both in using chat reference and then using the data from the transcripts to determine who was using the system and what types of questions were being asked in order to better manage their libraries' staffing (Houlson et al. 2006).

Researchers at Eastern Illinois University investigated a variety of data sources, including library satisfaction surveys, shelf inventory data, and circulation statistics, but used each of these sources individually to focus on discreet problems rather than tying the data together to look at overall library trends (Sung et al. 2006). University of Las Vegas used LibQUAL+ data as well as a library survey of faculty, faculty focus groups, and usage statistics for monographs and databases, in an effort to determine how library liaisons should organize their time and talents most effectively (Brown and Tucker 2010). Overall, this broad range of research indicates that while data are being used in a variety of creative ways, researchers have still typically focused on a single source of usage statistics over time, rather than comparing multiple data sources over

time. The authors hypothesize that integrating a variety of data sources over time will lead to a better understanding of library usage at Miami University.

METHODOLOGY

To determine if correlations exist or if specific types of data might serve as leading indicators to predict library usage or traffic, the researchers first set out to collect as many types of objective data as possible. The final data sets chosen to research included: chat and in-person reference statistics, database and e-journal downloads, and catalog searches. Finding comparable data sets was problematic, since each data source had its own unique restrictions and limitations including: lack of granularity within the data (i.e. for many of the databases, only monthly usage data were available), missing data points, out of date data, only the most recent year of data was sometimes retained, or technical problems with data collection devices, etc.

For some types of data that were originally planned to be included, such as door counts, it was determined that no accurate source of that data existed, due to mechanical and technical issues. Twitter data turned out to only be available for the most recent six months, a time period for which most of the other data sources did not have available data. Pay to print data were only collected by fiscal year and only reported at the end of each fiscal year. These difficulties in lining up data sets for similar time periods point out one reason that multiple data points are not often compared. Consistent time-series data is often difficult on its own to obtain, and obtaining comparable data from multiple sources for the same time periods can be one of the most challenging parts of the analysis.

The authors initially focused on the time period from August 2009 to December 2011 which was a period when the majority of the data points were available. This time period provided data sources covering five academic semesters, providing opportunities to compare

differences over time and also variations specific to the nuances of the academic calendar. For some additional areas of analysis, the focus was solely on 2010 or the spring semester of 2011, based on the areas of interest and the data available.

Another reason for the selection of the 2009-2011 time period was that in fall of 2009 Miami University Libraries began recording in-person reference statistics using LibStats software (FOSS4LIB 2012). Previously, manual tallies were used to record desk transactions, making compiling and comparing statistics a difficult and time-consuming chore. With online data collection, it became much easier to track in-person reference desk traffic on a daily or even hourly basis, as well as gathering data on many more categories of reference interactions. At the same time, Miami University Libraries also switched to Libraryh3lp software for chat reference (Nub Games 2011). Besides greatly improving the chat reference service interface, Libraryh3lp also improved data collection, because it automatically recorded each chat reference transaction. These more accurate and easily accessible sources of data were major components of this research project. Database and e-journal usage data were provided by the vendors, with some providing more detailed and timely data than others. Since the catalog search screen is the default homepage on public computers in the library, search data were determined based on the number of hits on the results pages received.

In order to ascertain the correlation between two disparate sets of data points, the authors employed Microsoft Excel's Correlation function which returns the Pearson's-product-moment correlation coefficient (more commonly known as Pearson's – *r*) values for any two data arrays. These values run from +1 to -1, with any value above .5 showing a high degree of positive correlation (Sheskin).

RESULTS

The authors began by examining in-person reference statistics. Anecdotal evidence and intuition indicated that in-person reference interactions were on a steady decline. For the time period examined, however, this was not found to be the case. While the in-person reference statistics varied widely across the time period, the data tracked closely with the academic calendar, peaking in the early Fall when first-year students arrive on campus and falling off most dramatically in the Summer, when both faculty and students tended to be on vacation (Figure 1). **[PLACE FIGURE 1 HERE]** Looking at the 29-month period for which there was data, in-person fall semester reference numbers actually increased each year; the total number of transactions was 9,890 in 2009, 10,548 in 2010 and 11,033 in 2011. The same upward trend can be seen during the spring 2010 and 2011; the total number of in-person transactions at the main reference desks at Miami University Libraries for January-April 2010 was 6,709, while for the same period in 2011 the sum was 8,371.[1] Thus, one commonly held belief did not hold true at Miami.

Week-by week comparisons of in-person reference statistics for three consecutive spring semesters (2010-12) also demonstrated consistent patterns of use during the course of the semester. The most recent semester (2012) had the highest number of transactions (Figure 2). The dramatic downward spike in transactions at week nine was due to Spring Break and Figure 3 shows that same data but with the week of Spring Break removed. By removing spring break week from the analysis, one can see the variability in traffic both across and within the semesters. By using weekly data, versus the monthly data seen in Figure 1, the timing of

[1] Spring 2009 is not included because LibStats data collection was not implemented until July 2009.

assignments in each semester can also be observed. The key aspects of the data are the high degree of correlation between the spring semesters across different years. The Pearson-*r* values when comparing the semesters week-by-week range from .59 to .66, indicating a high degree of correlation in the patterns of in-person transactions.

[PLACE FIGURE 2 & 3 HERE]

The authors next compared reference statistics on a daily and hourly basis. This more detailed analysis was also a useful tool for scheduling reference staff. Week five, an average week, was chosen to highlight difference between days of the week of the spring 2011 semester. Week 5's activity was relatively high on Monday, Tuesday and Wednesday, with some drop-off on Thursday and a significant drop-off for Friday through Sunday (Figure 4). Sundays were thought of as a "busier" day by those staffing the reference desk, and desk coverage by two librarians was typically provided for a full nine-hour period (from 1-10 PM), as compared to Saturday's four-hour desk coverage (1 PM -5 PM) provided by a single librarian. While traffic was marginally higher on Sunday during this particular week, it does not justify the extra staff time. In order to determine which hours to cut, Sunday hourly data was examined (Figure 5.) It is important to note that Figure 5 reports all desk transactions, while Figure 4 isolates those questions classified as "reference." Additionally, Figure 5 shows the total number of questions on Sundays for the entire semester, not just a specific week. Looking at the number of transactions throughout the day, the overall pattern is relatively flat. Also note that the last two hours of the nine hour period staffed by librarians (1 PM -10 PM), had the lowest numbers of total transactions. As librarians have often lobbied for a shorter shift on Sundays, this data would seem to support their argument.

[PLACE FIGURE 4 & 5 HERE]

Next, chat reference transactions from fall 2009-Fall 2011 were examined (Figure 6). Though data trends are similar for in-person transactions (Figure 1) and chat transactions (Figure 6), by overlaying the data (Figure 7) one can examine differences in timing and usage of these varying services. It is a common perception that in-person reference transactions are declining and chat reference transactions are increasing. However, as seen in Figure 7, this was not the case at Miami University. By viewing both sets of data points on one graph, one can observe the sheer numbers of in-person reference traffic as compared to chat, with the peaks of chat reference transactions often being similar to the troughs of the in-person data.

[PLACE FIGURE 6 & 7 HERE]

The large discrepancies between in-person and chat transactions created difficulties when comparing the two data sets on the same graph. In-person reference activity ranged from between 500 to 3,500 interactions per month, while the chat reference transactions ranged from 77 to 691 per month. Statistically, the correlation between these two datasets was strong, with a Pearson's *r* value of .82; however, the relatively large discrepancy between these two sets of numbers made the similarities (and differences) harder to visualize.

To ameliorate this problem of visualizing data on different scales, the data points needed to be transformed so that they were comparable, either through percentages or by normalizing the data. The researchers normalized the data by dividing each point by the maximum value in that set, so that all data points are transformed to a figure between zero and one. This technique allowed data to be compared as trend lines, providing a way to visualize the peaks and troughs of multiple data sets regardless of scale, a common practice used by Google Trends analysis

(Google). The authors used this technique throughout the data analysis to facilitate comparisons among the disparate sets of data. For example, both Figure 7 and Figure 8 show the same two data sets, comparing in-person and chat reference transactions, but the normalization process in Figure 8 makes the similarities across time much more striking. **[PLACE FIGURE 8 HERE]**

To bring in additional information beyond reference transactions, the authors next considered library research resources, such as database and catalog searches. However, sources such as EBSCO searches and e-journal downloads did not have complete data for the entire 29-month period used above. For this reason, the calendar year of 2010 was chosen because it had the most complete data available. By normalizing the data, a strikingly similar pattern appeared between total reference transactions (in-person and chat) compared to the number of searches in EBSCO's Academic Search Complete database over the calendar year 2010 (Figure 9). Also, it was apparent that the EBSCO Academic Search Complete searches were a lagging indicator in the fall semester as compared to the spring, in that peaks in EBSCO searching occurred after the peaks in reference transactions. One explanation could be related to library instruction, both in the classroom and at the reference desk. During fall semesters, librarians typically conducted more library instruction sessions and encountered more new students at the reference desk, many of whom were not familiar with the library resources. By the spring semesters, it is hypothesized that the fewer in-person reference transactions may have been due to the increasing familiarity of the students with navigating the library and its resources. Many second semester students might have already become familiar with how to use a basic resource such as Academic Search Complete during the fall term and used it as soon as they began projects and assignments in the spring. The graphs provide calendar year data rather than academic years; therefore the second

part of the calendar year (August to December) represents fall semester data, whereas January-May is the spring semester of the previous academic year.

[PLACE FIGURE 9 HERE]

The next data comparison (Figure 10) examined e-journal usage, based on download data for calendar year 2010 from EBSCO's Academic Search Complete, JSTOR, and Elsevier's Science Direct Journal platform. Due to the different scales of the data sets, the data were normalized to facilitate visualization for comparison using the same method described above. Despite their different interfaces and disciplinary coverage, JSTOR and Elsevier had nearly identical usage patterns. EBSCO downloads were also similar. In this case, the differences between the spring and fall semesters were less pronounced. One explanation may be that students took similar classes each semester and these classes had similar requirements for papers and assignments that required journal research. In each semester, e-journal usage usually peaked near the end of the semester, as major papers were due, and then usage fell sharply as classes ended. These trends in e-journal usage would have been difficult to visualize without normalizing the data, due to great differences in the monthly downloads from each provider. For example, in November 2010, 48,333 EBSCO journal articles were downloaded by Miami patrons, while J-STOR and Elsevier titles had 16,249 and 5,674 downloads respectively. If represented in a line chart using the raw numbers, the change in Elsevier downloads would barely appear.

[PLACE FIGURE 10 HERE]

The final data comparison for this project was between database downloads from EBSCO's Academic Search Complete and library catalog searches (Figure 11). Of all of the data

sets examined in this project, these two were the most highly correlated, with a Pearson's *r* value of .91. Again, by normalizing the data, the usage trends were mirrored, with both resources rising and falling throughout the academic year. In the spring semester, the use of the catalog began high and rose to its peak in April, as opposed to downloads from the EBSCO databases which (at least in 2010) peaked in the fall semester. Also, there was less variation in the number of catalog searches throughout the year, so that even at the lowest point in July, the number of catalog searches was approximately one-third of the peak number of searches in April. With EBSCO, the lowest number of downloads was less than one-fifth of its peak. It appeared that the catalog saw steady use by students, faculty, and staff throughout the year, as compared to more periodic usage of the research databases, based on assignments.

[PLACE FIGURE 11 HERE]

DISCUSSION

The key finding of this study was that library activities displayed a remarkable degree of correlation amongst themselves. It was hypothesized by the researchers that there would be leading and lagging indicators in comparing the various data points, but for the one year period (2010) studied most closely in this research, it was found that resource usage of all types tended to track very closely with each other throughout the academic semester. While chat reference transactions were lower than in-person reference transactions, the overall patterns were very similar. For the year examined in this study it was interesting to note that catalog searches and e-journal usage occurred in a nearly identical pattern, indicating that when one was being used, so was the other. Through examination of specific e-journal providers for both humanities (JSTOR), and the sciences (Elsevier), similar patterns emerged; apparently procrastination was a factor

across all disciplines in that usage of all resources peaked at the end of the semesters immediately before final assignments are typically due.

One notable exception was that the peak in EBSCO searches lagged behind in-person reference transactions during the fall semester. This was explained by the high number of reference transactions during the initial weeks of the fall semester, which in turn was largely driven by new, first-year students being unfamiliar with library services and functions. The correlation in the peaks of the two data points during the spring semester was more indicative of the true research patterns of the Miami University community, after students have baseline knowledge of the university and the libraries.

Another interesting finding, which became apparent through the use of the normalized data comparisons, was that the chat transactions peaked later in the semester than the highest number of in-person reference interactions. This could indicate that patrons were following up on information they received earlier, or that as they get more involved in their research projects they connected to librarians through the chat widgets embedded directly in the library and database homepages, rather than leaving their work to ask a question in person.

The high degree of correlation found across all of these data sets over time points to an important opportunity. By being aware of the rhythms of the semester and academic calendar to usage of all types of library resources (reference assistance, library catalog, research databases, e-journal usage), variable service point staffing could be achieved to match usage patterns. Instead of having a consistent level of staffing at the reference desk throughout the entire semester, additional librarians could be assigned to the desk during peak periods, while freeing them to do other tasks during slower parts of the semester. One likely change to occur at Miami

University Libraries is that staffing changes on specific days of the week and hours of the day (particularly during weekends) is being considered based on these findings. Since the data indicates that the third to fifth weeks of each semester are a slower period for the reference desk, and knowing that these are peak weeks for information literacy instruction, the reference desk staffing has been adjusted during those time periods. The libraries moved to staffing the reference desk with only a single librarian during those weeks, rather than double-staffing the desk.

CONCLUSIONS

While more and more sources of data appear to be available, getting the data into a useable format that allows for easy comparison among data sets during the same time period is often a more difficult task than one imagines. However, with increasingly automatic mechanisms for recording data and access to larger storage devices, data analysis will only continue to become more feasible. Access to improved sources for reference statistics make much more detailed analysis possible and can assist greatly with staffing decisions as well as the compilation of administrative reports and documentation. The benefits of increasingly granular data are also seen here, as daily and hourly statistics provide solid evidence in addition to the anecdotal reports of particularly busy shifts at the reference desk or dead weekends. As it becomes more evident that the overarching factor behind resource usage is the timing of the academic semester, having weekly or daily data for services would allow improved research into these trends and provide objective information to make more informed decisions relating to staffing.

Improved collection methods also increase the accuracy of the data, particularly with the chat reference software, which records each interaction automatically. While the in-person

transactions still require a human action to record them, the online LibStats collection data are much more likely to be accurate than the hash marks and tabulations that were used previously. Improved timeliness is also an advantage of these systems, as well as the ability to easily export the data into readily usable formats for analysis, including graphs and spreadsheets.

Overall, this research indicates that the demise of in-person reference is somewhat overstated, as chat reference provides a valuable addition but does not appear to be replacing it, based on the number of transactions for each service. Also, while no strong leading indicators were identified, the fact that most of the data points closely mirrored each other is another finding worth noting. By isolating differences in usage of resources across the fall and spring semesters, patterns indicate that new students coming into an institution in the fall are indeed transformed by library interactions and use resources earlier and may require less in-person reference services in the spring.

Future research areas may examine LibGuides usage data and more granular database usage data. Miami University Libraries is now collecting tweets made within a five-mile radius of the library in order to assess social media discussion of research topics and how they relate to reference activity and resources usage. Based on this research, librarians may also use their knowledge of the patterns of the academic semester to begin to proactively reach out to students via social media during specific research-intensive weeks. Because Miami University is undergoing a change to the academic calendar starting in the fall 2013, with each semester being one week shorter and the addition of a winter term, it may also provide a natural experiment to see how these changes affect library usage. Lastly, it is a goal to build a central online repository of data collected by the Miami University libraries, facilitating this type of research and data-driven decisions.

As increased and improved sources of data are available and longer time series of data are collected, more trends and associations may become apparent. The goal at Miami will be to capture reliable and timely data for analysis and decision making. Ultimately, it is the hope that data-driven decisions will improve service quality and reduce costs.

REFERENCES

Baker, Gayle, and Eleanor J. Read. 2008. "Vendor-supplied usage data for electronic resources: a survey of academic libraries." *Learned Publishing* 21 (1): 48-57. doi:10.1087/095315108X247276.

Brown, Jeanne M., and Cory Tucker. 2010. "Using Multiple Data Sources to Develop a Liaison Agenda." *College & Undergraduate Libraries* 17 (4): 365-385. doi:10.1080/10691316.2010.525424.

COUNTER. 2012. "The COUNTER Code of Practice for e-Resources: Release 4." Last modified April. http://www.projectcounter.org/r4/COPR4.pdf.

Farney, Tabatha A. 2011. "Click Analytics: Visualizing Website Use Data." *Information Technology & Libraries* 30 (3): 141-148. doi: 10.6017/ital.v30i3.1771.

FOSS4LIB. 2012. LIBSTATS. Last modified January 9. http://foss4lib.org/package/libstats.

Garrison, Judith S. 2010. "Making Reference Service Count: Collecting and Using Reference Service Statistics to Make a Difference." *Reference Librarian* 51 (3): 202-211. doi:10.1080/02763871003733372.

Google. 2012. "Is the data normalized?" Last modified October 16. https://support.google.com/trends/answer/87284?hl=en

Grigg, Karen S., Bethany A. Koestner, Richard A. Peterson, and Patricia L. Thibodeau. 2010. "Data-Driven Collection Management: Through Crisis Emerge Opportunities." *Journal of Electronic Resources In Medical Libraries* 7 (1): 1-12. doi:10.1080/15424060903585685.

Houlson, Van, Kate McCready, and Carla Steinberg Pfahl. 2006. "A Window into Our Patron's Needs: Analyzing Data from Chat Transcripts." *Internet Reference Services Quarterly* 11, (4): 19-39. doi:10.1300/J136v11n04_02

Jones, Jennifer Link. 2011. "Using library swipe-card data to inform decision making." *Georgia Library Quarterly* 48 (2): 11-13. http://digitalcommons.kennesaw.edu/glq/vol48/iss2/4

McLaughlin, Jean. 2010. "Reference Transaction Assessment: A Survey of New York State Academic and Public Libraries." *Journal of the Library Administration & Management Section* 6 (2): 5-20. doi: 10.1108/00907321111186631

Minami, Toshiro, and Kim, Eunja. 2010. "Data Analysis Methods for Library Marketing in Order to Provide Advanced Patron Services." *International Journal of Database Theory & Application* 3 (2): 11-20. http://www.sersc.org/journals/IJDTA/vol3_no2/2.pdf

Morrisey, Locke. 2010. "Data-Driven Decision Making in Electronic Collection Development." *Journal of Library Administration*, 50 (3), 283-290. doi:10.1080/01930821003635010

Mosley, Pixey Anne. 2007. "Assessing User Interactions at the Desk Nearest the Front Door." *Reference & User Services Quarterly* 47 (2): 159-167. http://rusa.metapress.com/content/l6w7126u075010m5/fulltext.pdf

Nolen, David S., Amanda Clay Powers, Li Zhang, Xu Yue, Rachel E. Cannady, and Judy Li. 2012. "Moving Beyond Assumptions: The Use of Virtual Reference Data in an Academic Library." *Portal: Libraries & The Academy* 12 (1): 23-40. http://muse.jhu.edu/journals/portal_libraries_and_the_academy/v012/12.1.nolen.pdf

Nub Games. 2011. LibraryH3lp. Accessed December 20, 2012. http://libraryh3lp.com

Powers, Amanda Clay, Julie Shedd, and Clay Hill. 2011. "The Role of Virtual Reference in Library Web Site Design: A Qualitative Source for Usage Data." *Journal Of Web Librarianship* 5 (2): 96-113. doi:10.1080/19322909.2011.573279

Sheskin, David J. 2004. *Handbook of Parametric and Nonparametric Statistical Procedures,* 3rd. ed. Boca Raton, FL: Chapman & Hall/CRC, Sung, Jan, Lanham Allen, and Sung Nackil. 2006. "Transforming Data to Build Knowledge for Healthy Libraries." *IFLA Conference Proceedings* 1-20. http://archive.ifla.org/IV/ifla72/papers/141-Sung_%20Lanham_Sung-en.pdf

Todorinova, Lily, Andy Huse, Barbara Lewis, and Matt Torrence. 2011. "Making Decisions: Using Electronic Data Collection to Re-Envision Reference Services at the USF Tampa

Libraries." *Public Services Quarterly* 7, (1/2): 34-48.

doi:10.1080/15228959.2011.572780