Assessment and Statistics   University Libraries

10-30-2008

White paper: Landscape of data gathering and analysis at the University Libraries, University of New Mexico

Tamara McMahon

Follow this and additional works at: https://digitalrepository.unm.edu/ulls_assessment

White paper: Landscape of data gathering and analysis at the University
Libraries, University of New Mexico

Tamara McMahon
September 20, 2005

Executive summary
Currently, the primary factor behind gathering usage data at the University Libraries (UL) at the University of New Mexico (UNM) is the necessity of its contribution to the annual Association of Research Libraries (ARL) report. Other reasons for statistical gathering, such information needed for better-informed budgetary decisions, exist but are not addressed in a consistent manner by UL. Each year, many people in the libraries spend hours compiling statistics for the ARL report. While this results in a comprehensive collection of data, it is not always accurate or efficient. Data gathering methods are not consistently recorded, resulting in varying methods for data collection each year. Additionally, data are often not reviewed for trends analysis. The current data gathering methods do not provide the UL with a good understanding of user needs and behaviors as they relate to the libraries. I propose several options for addressing these issues. I recommend reorganizing DATAC in January, 2006, as this will give me, the committee leader, additional time needed to decide on committee members and projects.

General
Current state
Much data exist within the UL defining what is being used, where, when, and how. Little data exist to aid in understanding why, but this could be obtained through focus groups and surveys. Focus groups, a form of qualitative research, compliment surveys, a quantitative research method.   These groups can address specific issues that arise in larger, traditional surveys such as LibQUAL+. Interviewing is another useful and economical method used in conjunction with surveys or statistical trend analysis (Avery 1994). Two methods for data collection exist for reporting on the UL as a whole. The primary collection is performed on an annual basis for the ARL report. Another method is LibQUAL+, a survey emailed to students, which is also performed annually. Both of these methods provide snapshots of UL, but neither is analyzed for trends in UL usage. Other statistics are created for occasional needs, such as grant proposals and management dashboard reports.   Problems resulting from the sheer complexity of the UL exist in various areas of the data collection. These problems, which will be explained in more detail in the following sections, include inaccuracies and underutilizations, among other issues.

Recommendations
- Continue LibQUAL+ and gathering ARL statistics
  These provide a good overview of the UL. They should be reviewed annually to analyze trends within the UL. In addition, problems discussed in later sections should be addressed to provide more accurate statistics in the ARL report.

- Set standards and report monthly
  I would like to see consistency and standards across all branches and departments of UL for data collection, which will eliminate many of the

inaccuracies in our current data collection methods. These standards and procedures should be reviewed annually by those involved in the creation and use of these statistics. Monthly reporting, in many areas, will provide a better understanding of how and when the library is used. In addition, monthly gathering of data will keep the methods consistent and will eliminate the struggle to gather and provide accurate annual ARL statistics. Trends analysis should become an integral part of future decision making since usage and behaviors change periodically. Therefore, I should compile these monthly statistics into a quarterly report.   I recommend that Senior Team review the quarterly reports for trends analysis.

- New employee orientation
  All new employees should be introduced to the statistical gathering methods and usage performed by his or her department during the orientation period.

- Short term: LIRIC database
  I propose moving forward with the development of the proposed LIRIC database. After it is in production, I would like to expand the database to include a section for reference statistics. I found these two areas to be most fragmented between branches.

- Long term: central repository
  I would like a comprehensive database for all data collection at UL, similar to the one developed at the University of Southern Mississippi (USM). The abstract for this poster presentation can be found at the following URL: http://www.lib.jmu.edu/org/ala/abstracts/default.asp#203

  A comprehensive database similar to that at USM will aid the UL in creation and maintenance of accurate and continuous statistics in all areas and branches of the UL. It will also facilitate analyzing trends in the UL over time. This, in turn, will aid in decision making in a variety of areas from staffing to journal cancellations.

LIBROS
Current state
LIBROS provides a wealth of statistical data, and it offers much versatility in accessing this data. LIBROS offers prepackaged reports as well as the versatility to create reports on-the-fly. Currently, circulation, acquisitions, and cataloging reports are run on a monthly basis, ARL statistics are reported annually, and other reports are run as needed.   Some methods of entering data within LIBROS have changed over the years. This has lead to inconsistencies and inaccurate data retrieval.

Recommendations
- Staff member
  The LIBROS team is a very busy and hard working team. An additional person to run the statistics would free up time for the system administrator to work on other areas of LIBROS. This would not involve hiring new staff but instead allocating a portion of a current staff member's time. I propose that I, the Data and Trends Analyst, take over this role of running the statistical reports that Eric Nudell, system administrator for LIBROS, currently handles. Dave Herzel, who currently creates and maintains the circulation statistics, should continue to be in charge of this area of LIBROS statistics. Once a central statistical repository is created, Dave and I can work together with these statistics.

- ARL Statistics for e-resources team
  A team has been formed consisting of Sever Bordeianu, Twila Firmature, Tamara McMahon, Eric Nudell, Evangela Oates, Carol Renfro, and Linda Skye. This project team will address the problems in obtaining accurate statistics for e-resources, including those found in LIBROS.

Journals – Electronic and Print
Current state
Statistics on print journal use varies among branches. For instance, CSEL keeps a tally of each time a print journal is reshelved. The numbers are entered into an Excel spreadsheet, which can then be reviewed for trends analysis.

Statistics for electronic journals are all over the board, due to external factors. Unfortunately, not all publishers provide usage statistics, and of those that do, the information is not provided in a consistent format among the publishers. Sessions, searches, and article downloads are some of the numbers provided by various publishers. It makes it difficult to obtain an accurate idea of usage when comparing inconsistent types of data. COUNTER (Counting Online Usage of Networked Electronic Resources) has established guidelines for recording and reporting usage statistics. More publishers are becoming COUNTER compliant, which is helping to alleviate these problems.

With the current budget issues, reviewing trends among both print and electronic journal usage is imperative for making decisions regarding journal cuts.

Recommendations
- Consistent print journal statistics
  Form a committee involving all branches of UL to develop standards for creating and maintaining print journal usage statistics consistently among branches. Once standards are set, this committee should review standards

and procedures on an annual basis to ensure consistency among branches and to address new statistical needs.

- Scholarly Stats
  MPS Technologies will offer Scholarly Stats (http://www.mpstechnologies.com/scholarlystats.htm) in the fall of 2005. This service compiles journal and database usage statistics from multiple publishers and vendors. These data are combined in various comprehensive reports. It provides an easy to understand overview of journal and database usage. This service is expensive. To justify the high cost, the following questions need to be reviewed:
    1. Will UL use these reports? If so, why and how often?
    2. If created in-house instead of through MPS, will the cost of employee salary be less than the cost for Scholarly Stats?
  When performing the cost analysis for salary versus product, take into consideration that this type of service will take numerous hours in-house to create. Once created, monthly maintenance will take less time.

- Trends analysis for journal cancellations
  With regard to journal cancellations, both print and electronic statistics need to be analyzed for usage trends over time. Cost analysis should be performed on the price of journals and their usage. Once a journal is cancelled, ILL usage and cost, as pertaining to a cancelled journal, should be analyzed to review the cancellation decision.

Electronic Databases
Current state
Similar to the electronic journals, inconsistencies in data provided by the vendors of databases exist. In addition to the subscription databases at UL, there are numerous open source and home-grown databases, for which little to no usage data are kept.

Recommendations
- Scholarly Stats
  See recommendation #2 under journals. This pertains to our subscription databases.

- DSpace
  The latest version of DSpace includes a statistics function. Once this version is installed and running on our servers, we should review this functionality and start analyzing these usage statistics for our DSpace databases.

- Home-grown databases
  At this time, it would be too difficult and time consuming to start a system to create and maintain usage statistics for all the home-grown databases a UL. Once other issues and recommendations discussed in this paper are

reviewed and in process, we can revisit the need to create and maintain statistics for these databases.

Electronic Books
Current state
Electronic books are relatively new and gaining in popularity.   Similar to electronic databases and journals, usage data are supplied in a variety of manners. Some publishers provide statistics by title or subject while others simply provide a number; whether it describes searches, sessions or something else entirely, we do not know.

Recommendations
- Consolidation report
  I am currently working on a method of collecting and organizing these various types of statistics for easy review. The result will be a consolidated report similar to the MPS Scholarly Stats reporting style. I am also working with a member of the LIBROS team to include usage statistics for electronic books from LIBROS.

Library instruction
Current state
Although library instruction statistics are kept for the UL, they are recorded and maintained at each branch of the library. Some are kept in an Access 97 database, while others are reported in the Public Services Statistics Count web site. While the Public Services Statistics Count web site is a great start for centralizing statistics at UL, not all branches and departments submit information through this web site. Additionally, it is confusing as to how this information is used once entered.

Recommendations
- LIRIC database
  Centralize this process through the creation of an instruction statistics database, as proposed by the LIRIC committee. This can be proposed as a project for ECE students to be completed by the end of the Fall '05 semester. This database should be created in such a way that it can be easily expanded to include other types of information, such as reference statistics, as necessary.

Reference
Current state
Reference statistics are gathered in each branch and many, but not all, departments of the UL. Gathering and analysis techniques vary between departments and branches. Reference statistics do not always include interactions resulting from email directly to subject specialists or email from the

web site. Not all departments report interactions which could be considered reference. For instance, LIT does not record or report their interactions with the public in cases where they provide technological assistance.

Recommendations
- Set Standards
  Designate a committee to review and set standards for collecting reference statistics. This includes which departments should report statistics, what needs to be reported, and how it should be reported.   Other issues that need to be addressed are how to report chat interactions, interactions with subject specialists outside of the reference desk, and email from the web site. Addressing these issues will provide more accurate statistics for the UL as a whole.

- Conduct survey and focus groups
  While it is important to know how many people we help at the UL, it is equally important to know if our interactions were useful to the patrons. This type of information will not be acquired through counting numbers. After chat reference is up and running, I suggest conducting short surveys and focus groups, similar to the process performed at the University of Illinois (Jacoby 2005). This will provide insight into the usefulness of our interactions, whether in person, via email or via chat.

- Centralize statistics
  Once standards for reference statistics are set, expand the proposed LIRIC database to include a section for reference interactions. This will centralize reference statistics among all the branches.

Web site
Current state
Statistics are being kept on the elibrary web site and the proxy server. Webalizer, an open source software, tracks site usage by month, day, and hour. It lists the URLs with the most hits, types of browsers and machines used, and other useful information. This information is used to provide statistics for the ARL report, grant reports, and to answer the occasional question. In addition, there are raw web logs that date back to 1999. The information provided by Webalizer is rarely used in web site design decisions.

No usability testing is being conducted at this time. In the past, usability testing has been conducted on an as-needed basis, such as rolling changes with the site. The testing, similar to the web pages, varied from branch to branch and designer to designer.

Recommendations
- Analyze stats and perform usability testing
  Since joining UL in April, I have heard much discussion regarding major changes with the web site. I do not believe that the UL has a good understanding of how people use and would like to use the library web sites. Before any major redesign occurs, I would like to see these questions addressed through web site statistical analysis and usability testing on the current site. The web site statistics, provided by Webalizer, will provide an insight into frequency of web page use, as well as when and how it is used. Webalizer is appropriate for our current web statistics software needs. The usability testing should consist of task analysis, using the Morae software, focus groups, and small surveys. The Webalizer analysis combined with the usability testing will provide an idea of how best to design our site to meet the needs of our users.

Space Utilization
Current State
To better understand how people use the physical spaces within UL, occupancy studies have been performed in the past. These studies provide insight into use of physical space at UL for issues such as staffing needs. Recent surveys have been conducted on areas within Zimmerman, such as the T-area.

Recommendations
- Standardization
  The data collected from the space utilization and surveys should be created and maintained in a standardized manner and stored in a central repository for UL data and statistics.

DATAC
At this time, I do not feel that I have a strong enough understanding of the UL operations and staff to adequately form the DATAC committee. I suggest postponing the reformation until January, 2006, at which time I will have a better understanding of these issues.

Acknowledgements
I would like to thank the numerous people throughout the libraries that took the time to meet with me and discuss current statistics gathering and reporting.

Bibliography
Avery, C & Zabel, D (1994). Total Quality Management in Academic Libraries: Initial Implementation Efforts. *Proceedings from the International Conference on TQM and Academic Libraries* (1st, Washington, DC, April 20-22, 1994).

Jacoby, J & O'Brien, NP. (2005). Assessing the impact of reference services provided to undergraduate students. *College & Research Libraries,* July: 324-340.