



Faculty Publications

2016-12-12

Conducting and Using an Academic Library Data Inventory

Holt Zaugg

Brigham Young University, holt_zaugg@byu.edu

Quincey McKeen

BYU, quinceybelle@gmail.com

Brett Hill

BYU, bretthillj@gmail.com

Ben Black

BYU, benjamin.r.black@gmail.com

Follow this and additional works at: <https://scholarsarchive.byu.edu/facpub>



Part of the [Library and Information Science Commons](#)

Original Publication Citation

Zaugg, H., McKeen, Q., Hill, B., & Black, B. (2017). Conducting and using an academic library inventory, *Technical Services Quarterly*, 34(1), 1-12.

BYU ScholarsArchive Citation

Zaugg, Holt; McKeen, Quincey; Hill, Brett; and Black, Ben, "Conducting and Using an Academic Library Data Inventory" (2016). *Faculty Publications*. 1809.
<https://scholarsarchive.byu.edu/facpub/1809>

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.

Conducting and Using an Academic Library Data Inventory

Holt Zaugg, Assessment Librarian, BYU holt_zaugg@byu.edu

Quincey McKeen, Student Researcher, BYU

Brett Hill BYU, Student Researcher, BYU

Ben Black, Student Researcher, BYU

Abstract: This article describes the need for and the processes used to create an inventory of data collected by an academic library. The study uses a survey augmented by multiple interviews to create and populate an inventory of data. The study was able to identify and sort sources of data generated by library personnel based on type of data (demographic, location, log, qualitative, and quantitative) and library division. It indicated the frequency of data collection and use, as well as where the data was stored. Finally, the inventory also identified types of data that are not currently collected but that librarians would like to collect. The article indicates how the data inventory may be used to support planning and improve library operations.

Keywords: data inventory, data audit strategic planning, library operations.

Conducting and Using an Academic Library Data Inventory

Academic libraries are known as repositories of knowledge and multiple sets of data. However, in the discharge of their duties, academic libraries also generate considerable data. This data may be used to inform library practice and determine library value (Oakleaf, 2010). Accessing this type of data helps inform librarians on how libraries operate and indicates opportunities for improvement and cooperation. However, as fiercely independent entities, libraries often collect and define their data in disparate ways that make comparison difficult.

A literature review indicated minimal discussion on conducting a data inventory (DI) or a data audit in libraries. While subtle differences occur between these two terms, the essence is that there is an accounting for data collected within a library. Creaser, Johnson, and Walton (2009) indicate that a DI helps to ensure that the data collected is used, relevant, and fits the purpose for which it was collected. The DI provides the opportunity to establish baselines of information, to allow for sharing of data, and to reduce redundancies in data collection (DiMattia & Blumenstein, 2000). In this way the DI enables strategic planning to ensure data collected is needed and used in appropriate ways. More importantly, it enables decision makers to quickly locate the type of data collected when it is needed to assist in decision making.

Several efforts directed toward the creation of a library-focused data inventory emphasize the need to collect data on international, national, regional, or local levels that define the value and service delivery of libraries. Decades ago, Herner, Vellucci, and Leyman (1972) issued a report calling for national, state, and local efforts to identify data sets, to define key terms, and to collect these data sets across multiple libraries. The intent and purpose was to be able to compare libraries on key indices of value and function. Jackson (2015) advocates for the collection and open sharing of such data to ensure open practices, to help determine the value of libraries, and to improve library practices. He advocates that such an approach may be instituted on a regional or national basis. This approach may include the perspectives of community partners who, using the provided data, could suggest ways to improve service and collaboration.

Ellis, Heaney, Meunier, and Poll (2009) worked on an international level DI with the intent of standardizing library data definitions and collection procedures. The authors felt that standardization of data definitions and data collection would make possible international comparisons that would facilitate the growth, development, and sharing of library practices to assist librarians worldwide. They identified six core indicators of libraries (accessibility, collection, usage, input, output, and potential for development and change), each with multiple types of data, designed to inform and to compare library practice. However, in pilot studies, Ellis, et al. (2009) continued to find similar problems of non-response or incomplete responses in categories where libraries do not collect the specified data for a variety of reasons.

The issue of non-response or incomplete responses is a pervasive problem with identifying data to place in a data inventory (Chiwere & Becker, 2015; de Jager & Nassimbeni, 2005; Jackson, 2015). Ellis et al. (2009) also describe a variety of reasons to account for this lack of or incomplete participation. In some cases, the desired data does not apply to the specific library situation. In others cases, there is a lack of experience and training in data collection, analysis, and use. In other situations, local librarians or government agencies funding the library do not see the value of data collection. Finally, data definitions need to be clear. For example, the concept of full-time equivalent (FTE) positions, in which two half-time positions would equal one FTE, was not clearly understood by some South American libraries. Instead, a head count of staff was used.

Henczel (2001) indicated that uniformity of definitions and data is not the only problem. She states that one needs to know where data is created and where it is stored to enable sound decision making and as an initial step in data management. She lists seven iterative steps needed to conduct an audit of data generated by the organization that provided for a full accounting of data collected, who collected it, and where it is housed. In testing Henczel's audit plan, Raliphada and Botha (2006) described two strengths of the plan: flexibility (application, scope, objectives and techniques) and cost effectiveness. They also added that the plan was effective in identifying information needs, sources, and significance.

By initiating an inventory of library data at the local library level, librarians are able to better understand what data is collected and how it is collected to inform and improve the creation of a DI. To inform current practices for conducting a data audit or inventory, a query was sent out via an ARL-ASSESS listserv. Correspondence from this query indicated a wide variety of approaches to creating an inventory or audit of library data. The approaches identify types of data, ranging from naming a few types of data to several dozen types of data. Collection methods vary from surveys to personal interviews (Barnachea, L., personal communication, September 2014; Ferguson, R. C., personal communication, September 2014; Merguerian, K., personal communication, September 2014; Neuhaus, P., personal communication, September 2014; Oakleaf, M., personal communication, September 2014; Peri, S., personal communication, September 2014; Rich, S., personal communication, September 2014; Ward, R., personal communication, September 2014).

This study describes the process and benefits of creating a data inventory for a large, private, academic library. The study includes a description of category creation and data collection, and ends with a discussion on how a local DI may be used to inform library value, planning, and practice.

Description of Library

The Harold B. Lee Library serves the students and faculty of Brigham Young University (BYU). The university has approximately 33,000 undergraduate and graduate students across 13 colleges. The library has an area of 665,000 square feet and holds over six million volumes. There are 156 full-time and part-time, non-student library employees with a yearly complement of over 400 student employees.

Method

From a background search and interviews with library leaders, potential categories for our DI were determined. Three broad categories (demographic data, specified data collected, desired data) were identified and subcategories created for each division. Because people collect and store the library-generated data, the DI focused on who collects the data within the context of their positions within the library. It is realized that, as people change positions, data collection may be expanded, maintained, or discontinued, depending on the person, his or her position, and data information needs. However, identifying what data each person collects allows researchers to connect the data to the library position via the person in that position. Table 1 lists each subcategory and content descriptions. Brief descriptions of these categories follow.

Table 1.*Category names and descriptions*

Category	Subcategory	Category Content
Demographic	Name	Employee's name (last name first)
	Job Title	Specific Title of Employee
	Division (Department)	Divisions including: Library Administration Office (LAO), Special Collection, Technical Services, Public Services, Library IT, Cataloging followed by departments specific to each division
	Email	Employee's work email
Data Specific	Data Type	Categories include demographic, location, log, qualitative, and quantitative
	Details	Brief specific details of collected data to inform a lay person.
	Frequency Collected	Categories include: as needed, hourly, daily, weekly, monthly, quarterly, semester, and yearly
	Frequency Used	Categories include: as needed, hourly, daily, weekly, monthly, quarterly, once a semester, and yearly
	Storage Location	Categories include: computer storage (i.e. m-drive, excel file, Lib-Analytics) and hard copy
Desired Data	-	Brief descriptions of desired data

Demographic Information

This part of the DI identified the person, his or her job title within the library departments and divisions, and contact information (email). This enabled the researchers to specifically connect data collected by an individual to the responsibilities of a specific library position. Associate university librarians were situated in the Library Administration Office (LAO) with a notation indicating the division for which they are responsible.

Type of Data

This section focused on the particulars of data collected. A pilot study identified five types of data (demographic, location, log, qualitative, and quantitative) with a sixth category (other) used as a placeholder for when library employees identified data but did not know how to categorize it. Data in this category was later reclassified to one of the other five types of data. For the purposes of this inventory, each type of data is briefly described below.

Demographic. This data refers to any information that identifies people, their college, their campus location, their teaching responsibilities, and/or their research interests. It may include names, discipline, contact information (e.g., address on or off campus, phone number, email), birthdate, research interests, and other background information collected by library employees. The specific categories used depended on the librarian collecting the data.

Location. This data refers to where items may be found but is not metadata. An example of this would be a general map of where items are stored in the library's on-site auxiliary storage. It may overlap with demographic data when it connects contact information with a specific

Data Inventory

location in a building or on campus. In short, it facilitates finding a person or item associated with the library.

Log. This data describes any data that is automatically collected via computers.

Qualitative. This type of data describes information that is not numeric in nature but provides comments, feedback, or descriptions of library services.

Quantitative. This information refers to numeric data that identifies how much, how often, or the amount of some item. While not all surveys are quantitative in nature, any survey data collected, unless specifically identified as otherwise, was categorized as quantitative.

In addition to identifying the type of data collected, a brief description of the data was provided to give enough descriptive information without losing context or becoming too wordy. The frequency in which the data was collected and used (i.e., as needed, automatically, hourly, daily, weekly, monthly, quarterly, once a semester, and yearly) were also identified. Finally, where each set of data was stored was also identified (e.g., on the m-drive, on paper or a hard copy) to indicate where the data could be located if needed.

Desired Data

This section identified only the type of data employees would like to collect to assist in decision-making and planning connected to their current position, but is data that is currently not collected.

All male and female, non-student, full- and part-time employees were identified as participants for the DI. No IRB was required because no identifying details within the DI would be disseminated, and the DI is used only for internal purposes.

Information for the DI was collected using three steps. First, a survey was sent to all full- and part-time librarians requesting the information, with multiple reminders to non-responders. The DI was initially populated with this information supplied by employees on the survey. Second, interviews with each employee were conducted to verify that entered data was correct and to fill in any missing gaps in the DI. Finally, the assessment librarian reviewed the DI with follow-up phone calls, as needed, to clarify entries, complete missing data, and improve descriptions.

Analysis

The purpose of creating the DI was not to conduct any indepth analysis per se. Its purpose was to identify all sources of data created by the library and to indicate where the data could be found. The intent is that any librarian could examine the DI to determine if data they needed was already being collected and, if so, where that data could be found. The librarian could then access the data to inform decisions, indicate improvement or value, and to conduct library-focused research. In this way the DI becomes a research or planning tool and not research data. As the DI is relatively new, several efforts are described in which the DI is currently being used as a research or planning tool, but these efforts have not been fully completed. It also suggests opportunities for further use. To facilitate this interaction with librarians, three tables were created—the DI sorted by type of data collected, the DI sorted by division and department, and requests for additional data collection. Data requests in the latter section were made in the form of a question. These three tables, along with current and potential uses are discussed in the Discussion section.

Findings

Types of Data

A total of 156 full- and part-time employees were included in the DI. A total of 612 sets of data were identified. As expected, over half (52%) of all data sets in the library were quantitative, with qualitative data sets accounting for about 18% of all data sets identified. The other three types of data (demographic, location, and log) equally split the remaining percent of data sets. The Public Services Division identified the most data sets, but this was somewhat expected as it has the largest number of employees. Types of data collected by each division are reported in Table 2.

Table 2.

Total data sets collected by division and type of data

Division	Number of Data Sets by Type					Total
	Demographic	Location	Log	Qualitative	Quantitative	
Cataloging	2	5	1	5	55	68
LAO	7	7	6	9	42	71
LIT	2	5	30	23	12	72
Public Services	27	25	19	49	118	238
Special Collections	11	11	7	17	47	93
Technical Services	6	5	4	8	47	70
Total	55	58	67	111	321	612

Requests for Data

Seventy-seven library employees made 153 requests to collect more data. Requests for data were organized into seven broad categories: collection, information, instruction, professional development (PD), promotion, technology, and website. Each request for additional data was sorted into its respective categories. Descriptions of each category are as follows:

Collection. These requests focus on the library collection, in whole or in part. It includes determining how the collection is made available or accessible to patrons.

Information. How the library provides service, patron satisfaction, patron usage patterns and demographics, and other operational factors are included in this section.

Instruction. These requests focus on the importance, impact, and efficacy of library instruction on student and faculty success.

Professional development. These are not requests for data, but for information on how data collection and organization tools may be used. The request is typically made by a single person, but the topics may be of interest to others.

Marketing. These requests seek to determine the efficacy of library marketing efforts to inform patrons of library services and collections.

Technology. Requests include how technology is being used in the library to create better efficiencies.

Website. These requests seek to discover patterns and trends of how patrons use the Internet or links on the library's website to meet their instruction, learning, and research needs.

Data Inventory

Three categories (collection, information, and website) were further divided into service (focusing on delivery of services) or technical (data related to technical questions) categories. Table 2 provides a summary of total requests for data by category.

No attempt was made to disaggregate requests for additional data by division, department, type of data, or why the data was needed. Each category included employee requests that data collection be done differently. While these were not requests for new data, they were included because they changed the way data was collected or used. Table 3 indicates the number of requests for additional data sets by category.

Table 3.

The number of requests for data by category

Category	Requests
Collection: Service	20
Collection: Technical	21
Information: Service	26
Information: Technical	24
Instruction	8
Professional Development	17
Promotion	11
Technology	8
Website: Service	9
Website: Technical	9
Total	153

Limitations

Several limitations affected the creation and population of the DI. First, there may be errors due to lack of understanding, transcription, or human error. As the data is viewed and used by library employees, these errors will be identified and corrected. Second, library employees will change responsibilities, arrive at or leave the library. While these changes are unavoidable and date the DI, subsequent updates will correct these changes. Third, some departments reported data collected as a group because all department employees collected the data as a group. The collected data was identical among individuals. In these cases, the same reported data was used to populate each employee's position in the DI. Finally, some library employees did not participate in the inventory for a variety of reasons. It is hoped with future inventory updates that data collected by these employees will be included.

Discussion

The DI serves several practical and potential purposes. As mentioned earlier, the DI is primarily a tool to assist librarians in the decision-making and planning efforts. The inventory also increases efficiencies (e.g., reduce double-collected data or increase data sharing), examines opportunities for collaboration, indicates how data may be used to improve service delivery, and identifies future data-collection efforts. The final DI contains spreadsheets sorted to facilitate these efforts, but sorting the DI using different metrics may affect other decisions or planning efforts.

Data Inventory

The first spreadsheet is the working document in which library employees are listed alphabetically for ease of access to add, delete, or modify entries. As librarians leave the library, change positions, or come to the library, this list is used to easily make the needed changes. The three other spreadsheets are for convenience, as they are sorted by the type of data collected, by the library division and department, and by the requests for data collection. Since the first spreadsheet is only used as a working document it will not be discussed further.

Type of Data Collected

One of the positive aspects of surveying and interviewing each employee was the opportunity for non-assessment employees to better understand the types of data collected and how the data connects to their position. There were several employees who collected data as part of their job responsibilities but did not realize that it fit into a specific category of data. This discussion helped to increase awareness of what types of data were collected and how that data may be used and shared. In other cases, employees only collected one type of data but had the potential and specific need for collecting other types of data to inform the delivery of their services.

Division and Department

When disaggregated by division, expected patterns emerge. For example, the Library Information Technology division is the largest collector of log data. This finding is not unexpected as much of their responsibilities require them to collect data and pass it on to librarians for further use. The LAO had the lowest rate of data collection because it is largely composed of library leadership who review reports and data collected by employees within their respective divisions. They do not directly collect data, but they are often the end point for data collection reports and efforts.

The Public Services Division was the largest collector of every type of data except log data. This finding, though not unexpected, illustrates how much contact employees in this division have with others both inside and outside the library. In the case of subject and liaison librarians, the differing levels of data collected, stored, and used were somewhat of a surprise. Reports of data collection and use ranged from those who kept no records of patron visits to librarians who kept detailed reports providing a host of demographic and qualitative data. Furthermore, there was a wide range of how data was stored—again ranging from no records kept to keeping only paper or electronic records. The DI highlights some of these differences and provides the opportunity to examine the type of data collected to determine a core set to be collected and how the data should be collected and stored. For example, if each subject librarian had a card reader or collected the patron ID number, this would automatically provide a host of demographic data to examine trends and patterns among patrons using this library service. If qualitative records with similar categories were kept regarding the nature of the visit, using somewhat uniform language, the data could be examined to determine the efficacy of library instruction as it highlights what instruction was helpful, what needs to be further emphasized, and which groups need targeting with initial or refresher courses. The qualitative data may also provide additional information regarding the patron experience. This information would be useful in strategic planning and service delivery.

The DI also allows employees from one division to examine data collection in related or connected divisions and departments. This comparison helps to identify work flow issues where one department or division may alter practices to facilitate better service delivery in another

department or division. It would also facilitate collaboration projects. The DI identifies who these librarians are so a cohort can be formed to examine this issue.

Another strategic planning use of the DI would be to populate a library impact map (Oakleaf, 2010). A library impact map can be used to determine the library's value to its patrons and for planning. It uses a grid system that matches library services housed within library departments and divisions on one axis with library or organization goals on the other axis. The intersection of each goal and library service indicates whether there is no data, the potential for data, data collected but not used, and data that is collected and used (Zaugg, 2015). Previous efforts to populate an impact map asked librarians if they thought such data existed, a DI would provide evidence of the data's existence, where it is stored, and who collects the data. Populating a library impact map using the data inventory would provide clear evidence of the library's value. It could also highlight areas where data requests were made to indicate data that could be potentially collected to inform library value and practice.

Finally, as new faculty and staff come to the library, the DI could be used to inform them of the types of data their predecessors collected. This would jump-start their knowledge of what types of data need to be collected and how the data is used. The new hires may choose to keep collecting the same data or they can build from the existing data information.

Each of these efforts also point to data that is being collected and may be used in library improvement processes. Library employees can reduce redundancies in data collection. Sharing data among library employees also provides the opportunity for collaborations to improve service delivery.

Desired Data

The requests for additional data present several opportunities. The requests highlight opportunities for collaboration. Employees from different departments and divisions with similar data-collection wishes could be joined together in a research group to determine best or common practices that would lead to the collection and use of this data. This collaboration would not only provide a common data collection effort with specific additions for the unique circumstances of each department, but it would also increase the understanding of what different employees do. Increasing understanding may lead to additional collaboration opportunities between library divisions and departments.

The data-collection wish list also indicates professional development opportunities that could be addressed through one-on-one instruction, study groups, webinars, or formal courses. Those seeking the additional training could be matched up with those who have had the training to form mentorships in which experienced librarians mentor other librarians. The requests for additional data could also be used to inform library professional development conferences. Finally, the list presents the opportunity to identify areas where additional data would help to inform library practice and service delivery.

Dissemination

Once the DI is complete, it needs to be disseminated throughout the library, otherwise it just becomes another source of lost or forgotten data. As mentioned earlier, all librarians may not understand the differences between the types of data and how the data may be used. This communication effort identifies who in the library is in the best position to use the DI. It also provides the opportunity for them to learn how the DI can be used and sorted to determine where information is for them to access. Presentations and workshops are methods to inform their

decisions and planning efforts. The dissemination efforts also become a source of feedback on how the DI may be improved to become more responsive to the needs of librarians (Creaser et al., 2009; Henczel, 2001). It should be noted that, while all librarians will contribute to the DI, all librarians may not use the DI.

Conclusion

The DI serves as a method for determining what data is being collected and used by librarians to inform library practice and service delivery. Patience and flexibility, by both those populating the inventory and those supplying the information, is needed to collect the data. Once collected, the result is a comprehensive look at what data is available. The DI becomes a strong tool for identifying what type of data is collected and where to find that data. Accessing the data provides librarians with the information they need to make key decisions and to plan for the future.

The data also indicates opportunities for cooperation and collaboration. This can lead to more efficient practices where data that is collected but unused or that provides no helpful information is no longer collected. It identifies duplication of data collection to reduce redundancies. It also highlights potential projects librarians may wish or need to undertake. It enables all library personnel to step out of their work silos to better understand what their colleagues do and how librarians can work together to improve the importance and effectiveness of library services. The act of creating and collecting information for the DI also provides the opportunity to instruct librarians on the types of data they collect and why it is important to collect the data. The result of a completed DI is improved knowledge and understanding of how librarians can work together and where to find the data needed to inform service and planning decisions.

References

- Chiware, E. & Becker, D. (2015), The uptake and usage of a national higher education libraries statistics database in South Africa, *South African Journal of Library and Information Science*, 81(1), 1–7.
- Creaser, C., Johnson, P. & Walton, G. (2009, August). Illumination & support, or, what do libraries do with all that data? in Ian Hall, Stephen Thornton, and Stephen Town (Eds.), *Proving value in challenging times: Proceedings of the 9th International Conference on Performance Measurement in Libraries and Information Services*, York, UK.
- de Jager, K. & Nassimbeni, M. (2005), Towards measuring the performance of public libraries in South Africa, *South African Journal of Library and Information Science*, 71(1), 39–50.
- DiMattia, S. S. & Blumenstein, L. (2000). In search of the information audit: Essential tool or cumbersome process? *Library Journal*, 125(4), 48-50.
- Ellis, S., Heaney, M., Meunier, P. & Poll, R. (2009). Global library statistics, *IFLA Journal*, 35(2), 123–130.
- Henczel, S. (2001). The information audit as a first step towards effective knowledge management, *Information Outlook*, 5(6), 48-62.
- Herner, S., Vellucci, M. J., & Leyman, E. (1972). *National inventory of library statistical practices, Volume 1: Data collection on the national, state, and local levels* (Report No. OEC-0-70-4890) U.S. Department of Health, Education, and Welfare, Office of Education, National Center for Educational Statistics (Library Surveys Branch).
- Jackson, B. (2015). The state of Canadian library data, *The Canadian Journal of Library and Information Practice and Research*, 10(1), 1-5.
- Oakleaf, M. (2010). *The value of academic libraries: A comprehensive research review and report*, Chicago: Association of College and Research Libraries.
- Raliphada, L. & Botha, D. (2006). Testing the viability of Henczel’s information audit methodology in practice, *South African Journal of Library and Information Science*, 71(1), pp. 39–50.
- Zaugg, H. (2015). Using a library impact map to assist in strategic planning in academic libraries, *Library Leadership & Management*, 20(3), 1-17.