While special collections and archives managers have at times recognized the importance of using data to drive decision making, translating this objective into reality and integrating data analysis into day-to-day operations has proven to be a significant challenge. There have also been obstacles to formulating quantitative metrics for special collections and archives and rendering them interoperable across institutional boundaries. This article attempts to focus a conversation around two issues: 1) the importance of quantitative analysis of operational data for improving research services in special collections and archives; and 2) the need for the profession to achieve consensus on definitions for quantitative metrics to facilitate comparisons between institutions. We will begin by exploring the value of assessing special collections user services through operational data analyses and reviewing case studies of institutions that have incorporated such analyses into their decision making and management practices. We will then proceed to a discussion of the benefits that can be gained from having interoperable standards for tracking user services metrics.

Operational Data

First, it is important to clarify what we mean by operational data—or, as it called in some contexts, administrative data. The Society of American Archivists (SAA) Glossary of Archival and Records Terminology is silent on these terms. Although the Glossary defines both administrative and operational records, these definitions are geared toward records collected by an archives, not the operational records of the archives itself.

---

1. Portions of the research for this paper were made possible by grants from the National Historical Publications and Records Commission (#DS-50009) and the Andrew W. Mellon Foundation.
3. In the Glossary, “administrative records” are defined as “a document that has been preserved because it facilitates the operations and management of an agency, but which does not relate directly to programs that help the agency achieve its mission”; www2.archivists.org/glossary/terms/a/administrative-records. The Glossary defines “operational records” as “1. records that relate to the substantive activities an organization undertakes to accomplish its mission or mandate; program records. –2. Information documenting the performance and use (operation) of a piece of equipment”; www2.archivists.org/glossary/terms/o/operational-records.

*Please note: URLs for all websites referenced herein were valid at the time of article submission.*
We did find the terms discussed in several sources outside of library, archives, and museum (LAM) literature. The Organisation for Economic Co-operation and Development (OECD) Glossary of Statistical Terms defines administrative data as “the set of activities involved in the collection, processing, storage and dissemination of statistical data from one or more administrative sources.” The McKinsey Global Institute uses the term “digital exhaust data” to refer to the large quantities of data that are constantly produced as a by-product of regular business activities in the digital age. Its report “Big Data: The Next Frontier for Innovation, Competition, and Productivity” underscores the strategic advantages that can be gained by mining the data generated from day-to-day business operations for management purposes. For this paper, we will use the term operational data because it more precisely connotes the kinds of data gathered through regular business operations than administrative data. Examples of operational data in special collections and archives might include data from patron registration forms, visitor logs, circulation counts, routine statistics kept on reference sessions conducted and classes taught, and web usage statistics. By contrast, data gathered through user surveys, for example, even if collected regularly, cannot properly be called operational because such data are not compiled passively as by-products of normal business functions, but rather require active input by users that is extrinsic to the function they seek to perform (for instance, answering questions about the usability of a finding aid while attempting to identify and request relevant materials). And while survey data might be used to inform administrative decisions, they do not do so from the purely operational perspective upon which we will focus our attention here. Nevertheless, we recognize that the terms operational and administrative data may be used interchangeably in other contexts. Interchangeable in our parlance are the terms special collections and archives insofar as both refer to institutional repositories that provide access to collections of rare and unique research materials.

One of the main advantages of operational data over forms of data that are gathered through extrinsic methods such as questionnaires, interviews, focus groups, and simulations is that there is no appreciable cost or overhead associated with data collection. Stores of operational data are constantly produced by existing administrative software and systems without any additional effort on the part of staff. Other operational data, such as information transcribed manually into paper or electronic user registration forms or call slips may require additional time and effort on the part of staff to acquire but is still considered operational for our purposes.

---

since it is gathered in the course of normal business functions. However, such data can be expensive to extract, manipulate, and analyze, especially if additional labor is expended in transcribing manually created data into electronic systems (such as spreadsheets or databases) to facilitate tabulation and analysis.\(^6\) Despite such costs, ongoing or periodic analysis of operational data can help special collections make effective improvements to public-facing services and aid in decision making around programming and outreach, materials location, opening hours, staffing models, and prioritization of processing projects. Such data can also be used to support budgets and justify resource requests. Mary Jo Pugh has summarized the benefits of routine data collection to archival reference services as follows:

> Measuring the use of repository holdings is necessary to organize and manage reference services in the repository and to evaluate their effectiveness. Information about use and users is needed to allocate resources, plan staffing patterns, order equipment and supplies, plan programs to meet identified needs, and reward staff. Such information helps staff to determine whether the level of service is adequate, assess assumptions about reference services, and modify services to meet changing circumstances.\(^7\)

**Obstacles to Data-Driven Management**

Despite the considerable benefits, operational data are generally underused in many archives, since many struggle to incorporate data analysis into their management cycles. The reasons for this can be manifold, including a lack of knowledge about what operational data exist within the organization; how to extract the data from the systems where they are created and stored; a lack of time, resources, or the specialized skills needed to conduct data analysis and interpretation; an unclear understanding of what questions data analysis could help to answer; and confusion over what questions staff need to ask to support management needs.\(^8\)

Unfortunately, collecting administrative data solely for the sake of collecting data—without a clear sense of what questions the data will be used to answer—is not an uncommon practice in special collections, and libraries more generally. While institutions may collect certain metrics for mandatory reporting—for example, numbers of visitors or instructional sessions—a clear purpose for collecting data

---

\(^6\) While there are obvious benefits to automating the collection of such information and avoiding, especially, redundant manual data entry and replication, it is beyond the scope of this paper to explore solutions for doing so.

\(^7\) Mary J. Pugh, “Managing Reference Services and Evaluating the Use of Archives,” in *Providing Reference Services for Archives and Manuscripts* (Chicago: Society of American Archivists, 2005), 98.

beyond simple summation is typically never established. Yet when data are collected without a clear understanding of how they will be used or what questions staff may need to answer through data analysis, the utility of the data might be limited. Without good planning, issues such as data structure, level of granularity, or lack of controlled vocabularies often contribute to an inability to analyze or repurpose operational data for specific administrative needs.

The obstacles to data-driven management in the public service functions provided by special collections stem from both organizational culture and a scarcity of data-related skill sets in the field. Examples of obstacles stemming from organizational culture might include undocumented data production mechanisms, unclear channels of communication between management and staff with regard to the purposes for data collection, or unassigned responsibility for data analysis and reporting. While managers and staff can work together to change organizational culture and implement a more proactive approach to data collection, documentation, communication, and responsibility assignment, the issue of data-related skill sets remains a challenge. Traditional archival skill sets do not necessarily include data analysis, data modeling, database querying, or data extraction and analysis skills. At the same time, data-related skill sets are rapidly becoming a highly desired and scarce commodity on the national scene, as “big data” floods the global economy and the ability for industry to interface with and interpret such data is becoming a requirement in achieving competitive advantage.9 While the kind of data-related skills needed for successful data-driven management in archives may not be as sophisticated or technically complex as that required in the global marketplace of big data, the trending need for analytic skill sets in all sectors of library and higher education administration highlights a new inclination toward data-driven management and underscores the need to rethink traditional definitions of skill sets in our own profession. When a special collections department lacks a staff member with data extraction and analysis skills, it might mean having to rely on external technical support from IT departments to fulfill part of the data-driven management loop. Yet such external technical support is unattainable for many and limited for others.

**Literature Review**

**Evidence-Based Practice**

Evidence-based practice is gaining traction in library and information science. The open-access journal *Evidence Based Library and Information Practice* is one example of

---

this trend. Although archivists typically have not thought of the use of operational data as evidence-based practice, this is precisely what it is. Anne McKibbon defines evidence-based practice in library and information science as:

An approach to information science that promotes the collection, interpretation, and integration of valid, important and applicable user-reported, librarian-observed, and research-derived evidence. The best available evidence, moderated by user needs and preferences, is applied to improve the quality of professional judgments.

In special collections, we have not clearly identified evidence-based practice as a goal; nevertheless, some studies have highlighted the value of examining operational data as evidence. Thus far, the area that has received the most scrutiny is collections processing. A number of articles from the early 1980s through the mid-1990s measured internal processing rates and costs. Several complicating issues emerged from these studies, including variations in the ways that processing activities were measured, differences between processing manuscript and archival materials, and the many localized decisions made during processing. What we do not know is how much this information changed practice in these institutions or the profession more generally.

Evidence exists, however, that leads us to think these problems persist. In their 2005 article, "More Product, Less Process" (dubbed MPLP), Mark Greene and Dennis Meissner provided evidence that the problem of inefficient processing still remained. They suggested alternative modes of processing that would open up more collections to researchers more rapidly and eliminate backlogs or "hidden collections." As a result of their work, many individual archivists and archival repositories have begun to implement MPLP in some manner. As archivists have attempted to measure increases in processing efficiency, however, issues of metrics

10. Evidence Based Library and Information Practice is available online at http://ejournals.library.ualberta.ca/index.php/EBLIP. EBLIP is an open access, peer-reviewed journal that is published quarterly. It is supported by an international team of editorial advisors and hosted by the University of Alberta Learning Services; see www.ls.ualberta.ca/. The purpose of the journal is to provide a forum for librarians and other information professionals to discover research that may contribute to decision making in professional practice.

11. Andrew Booth, "From EBM to EBL: Two Steps Forward or One Step Back?" Medical Reference Services Quarterly 21, no. 3 (2002): 53.


(that is to say, definitions as well as how to measure) have come to the fore. In their 2010 article on the Arizona Archives Summit, Linda Whitaker and Melanie Sturgeon note the importance of definitions:

> The need for a common vocabulary was soon apparent, since the interpretation of “unprocessed” varied. After much discussion, the planners reached consensus: an unprocessed collection was one that had no inventory or finding aid.14

More recently the Processing Metrics Collaborative, discussed elsewhere in this issue of RBM, is heir to this line of inquiry.15 As in the case of the Arizona Archives Summit, the Processing Metrics Collaborative seeks to identify and agree on community metrics to enable larger-scale comparisons between repository practices.

**Transaction Data**

A decade ago, ARL SPEC Kit 268 documented how library reference and instructional services collected and used reference transaction data. Although not specific to special collections, its findings are nevertheless relevant to the framework of special collections and archival user services. SPEC Kit author Eric Novotny asked librarians to report their impressions of the quality of their library’s efforts to record, analyze, and make use of reference transaction data for assessment. He found that “as a group, the responding libraries rated their performance as below the minimum performance level in the analysis and use of reference transaction data. And, in all three areas (collection, analysis, and use of data) the library’s performance was deemed to fall far short of the desired performance levels.”16

Novotny found that the main method by which reference service was evaluated for 96 percent of responding institutions was by basic transaction count, and that 99 percent of institutions recorded data on reference transactions manually and not electronically, using tick marks on paper.17

In Novotny’s research, many institutions reported that they were interested in using reference data to contribute to national statistics and for assessment pur-

---

poses, such as monitoring trends in reference services and determining appropriate reference staffing levels. Budget justifications or resource requests were not an important consideration for data collection for the majority of respondents; only 34 percent indicated that they used reference service data for budget justifications.\(^\text{18}\)

In ARL SPEC Kit 296, *Public Services in Special Collections*, published in 2006, Florence Turcotte and John Nemmers reported that a number of special collections gathered operational data systematically to measure the volume of their services, but only a handful actively assessed either the quality or effectiveness of these services. They found that most institutions that responded did not incorporate hard data collection or analysis into their evaluation processes as a central component: 96 percent of respondents relied on informal feedback, while 66 percent relied on direct observation of service transactions as a primary means by which they collect data to be used in service quality assessment. Turcotte and Nemmers viewed this reliance on informal and anecdotal data as a concern and encouraged further quantitative analysis of special collections public services data.\(^\text{19}\)

Their survey furthermore showed that many of the responding ARL libraries considered patron registration data to be relevant only for security purposes and therefore did not view it as a dataset from which to extract information on user behavior. In 2006, 88 percent of the 74 ARL libraries that answered the question gathered patron registration data in paper form, with at least 55 percent storing the registration data only as paper—meaning that the data could not be analyzed computationally. A mere 34 percent reported storing registration data in a database, 3 percent stored it in a spreadsheet, and 11 percent stored it in other formats, some of which were electronic.\(^\text{20}\)

Turcotte and Nemmers also found that less than half of ARL respondents tracked or cross-analyzed public service transactions with patron affiliation (such as faculty, community member, or undergraduate), despite the fact that it can be easily done using basic registration and circulation data and can provide a wealth of evaluative information about who is using collections over time. More than 90 percent did track the number of reference transactions, as well as the number of instructional sessions taught in special collections spaces. Yet 77 percent of respondents did not record patron affiliation for public service transactions, about one quarter did not track the number of items retrieved from the stacks for patron use, and more than a third did not track the number of items reproduced for patrons.\(^\text{21}\)

---

In 2010, ARL SPEC Kit 317, Special Collections Engagement, found that almost a third of responding ARL libraries had no formal measures in place to evaluate outreach events, and only half attempted to evaluate the success of exhibits. Of the 61 respondents who answered the evaluation section of the survey, one third did not collect data on student use of archival materials. Meanwhile, two-thirds performed some sort of evaluation of students’ use of archives, and 45 percent of the 61 respondents said they relied on anecdotal feedback as a means of evaluating student use of archives. Over 50 percent reported that they conducted no evaluation of the effectiveness of outreach efforts directed to affiliated scholars and researchers. For those 22 institutions that reported conducting outreach evaluation, common methods included tracking usage statistics, reacting to anecdotal feedback, and formal surveys. Eight (36%) reported enacting changes in practice in reaction to evaluation findings. The authors of the SPEC Kit considered possible reasons why data collection and evaluation efforts were so limited:

Assessment is rarely easy … Institutions feel they are not able to quantify the success of their efforts, and this in turn limits the ability to compare activities within the institution or across institutions, to plan further outreach effectively, or to communicate the results of those outreach activities to the larger special collections community.

Genya O’Gara and several other authors of SPEC Kit 317 further discussed the implications of their research findings in a 2011 blog post:

Simple forms, tally marks, and baseline ARL statistics will never be able to get at the information we need to improve our practices. Special collections need to make the case for developing more appropriate evaluation methods—even though this will require a commitment of valuable staff resources—and then make the commitment to using the results of these evaluations to enhance services. Ultimately, more meaningful data will help us provide better service to the students, faculty, and researchers who rely on special collections, and it will better equip us to tell their story and our own.

While there is not much published data related to the efforts of archives to implement data-driven management and operational data analysis in user services, the

23. Berenbak et al., SPEC Kit 317, 16.
studies discussed above have shown that data collection and analysis—while perhaps understood to have value—are often prioritized behind many other essential activities. They also indicate that anecdotal and observational evidence has proven easier for librarians and archivists to understand and use than methods involving quantitative data analysis. To learn more about the obstacles to employing such methods, we conducted a study of our own.

A Survey of Current Practices

For our study, we used an exploratory methodology to get a sense of how institutions were harnessing quantitative data to drive decision making. To identify a broad range of such efforts, we sent e-mail inquiries in January 2012 to six electronic discussion lists: SAA Archives and Archivists, SAA Description Section, SAA Reference Section, ACRL Rare Books and Manuscripts Section (RBMS), Exlibris, and ARL-Assess. We received eighteen responses from individuals who reported incorporating analysis of existing quantitative data into management and decision-making processes. A follow-up phone call was conducted with one of the respondents. A majority of respondents had either begun such efforts only recently or were just beginning a project. We were encouraged by the diversity of responses and enthusiasm on the part of staff participating in such activities despite the low response rate. We have presented several of these below as case studies to encourage others to take up similar efforts and share ideas for how different data sources can be leveraged for various assessment purposes.

Overview of Findings

Among the eighteen individuals who responded, fourteen (78%) were from colleges and universities. The remaining respondents were from state archives and historical societies. Most respondents began their analysis efforts only within the past few years and were attempting to analyze operational data only from their own repositories. One respondent was working toward standardized metrics within one organization with multiple repositories and another was involved in a statewide effort to combine data from special collections.

The respondents cited a variety of operational data that they analyzed. These data included information about users (institutional affiliations, geographic origins, research interests, and total numbers), reading room visits, items paged (onsite and offsite), reference services, collections (extent, growth, processing capacity), press coverage, instructional sessions, requests for permission to publish, and event attendance. Only one respondent mentioned analyzing web analytics data. Even within this small sample, we observed distinct patterns for collecting and analyzing operational data about users, collection extent, and reading room traffic.
While it is interesting to observe what is being counted, asking why is perhaps a more important question. Why do special collections librarians and archivists collect such data? How do they want to use it? Our respondents indicated that they were interested in measuring reading room traffic to inform hours of operation and desk staffing, demographic information in relation to use data to inform outreach efforts, and processing rates and the extent of unprocessed collections to decrease backlogs. They also expressed interest in minimizing interinstitutional competition for collections through sharing acquisitions data and in increasing access and use. These objectives for data collection feed directly back into collection management and other operational concerns. One respondent (SCA03), however, expressed a deeper desire to “consistently collect numbers that are meaningful measures of archives impact/ value, productivity, outcomes.” This latter approach to metrics requires another level of consensus on what value and impact mean in a given context.

Our simple survey also revealed a variety of constituencies for whom these data were being collected. The librarians and archivists themselves were an implicit audience, yet respondents also focused on external drivers for data collection. Three respondents identified both grant funders (who demanded processing data) and institutional administrators as major audiences for data collection efforts. Additional constituencies cited included legislators, advisory boards, colleagues, and the media. From this overview of the varied interests and motivations for collecting data, we turn now to four case studies developed from our survey responses to explore these issues in greater depth.

Case Studies

In 2010, North Carolina State University (NCSU) Libraries’ Special Collections Research Center (SCRC) collaborated with staff in NCSU’s Digital Libraries Initiatives department to analyze special collections user services data. With the goals of uncovering trends in materials usage to better support data-driven management, staff harnessed pre-existing stores of quantitative data to analyze changes in the makeup of their patron base over time, how different groups within the patron base use different types of materials, which collections and collecting areas receive particularly high usage, and the points in the day, week, and semester during which usage peaks are observed. The SCRC had used a Microsoft Access database to track all patron registrations and collection usage since 2007; staff analyzed data from 2007 to 2010 using Microsoft Excel and created visualizations in Adobe Illustrator, such as figure 1.

The SCRC has been using the results to understand reading room staffing needs, determine processing priorities based on areas of researcher interest, and plan
outreach for graduate and undergraduate user populations. For example, the chart above helped to clearly articulate the amount of usage their collections were receiving on an annual basis from graduate students. The SCRC has since begun to coordinate greater graduate student outreach efforts both to NCSU’s own graduate student population as well as developing plans to extend these efforts to graduate student populations at nearby institutions. It has likewise refocused efforts to create greater awareness of NCSU collections among local peer institutions.

The University of South Florida (USF) Libraries began using Atlas Systems’ Aeon software in May 2010 to record information about reading room visits, patron demographics, patron research topics and interests, and collection usage. Since then, on a semester by semester basis, staff have used the reporting and data export capabilities of Aeon to analyze these data and drive decision making about reading room hours and staffing: for example, a cross-sectional analysis showing what times of day visitors from different university departments come to use the reading room. In 2012, USF staff will be looking into ways to use the data for decision

making around outreach, stacks reorganization, and digitization priorities. USF has been using Microsoft Access to query its Aeon database and then exporting results to Microsoft Excel for analysis. They have made use of user-defined fields in Aeon to track specific patron research topic areas and interests and analyzed these data alongside data obtained from standard elements in Aeon.27

In 2011, Duke University’s David M. Rubenstein Rare Book & Manuscript Library initiated a data dashboard, compiling statistics on a quarterly basis for management purposes. Staff cited four objectives for creating the dashboard: to track progress toward annual goals, establish benchmarks to be able to determine targets and expectations, recognize trends that may have resource or policy implications, and call attention to outcomes that they wanted to prioritize (such as press coverage). Data collected and analyzed in the dashboard include acquisitions data, processing capacity, web statistics, press coverage, reference transactions, research visits, items requested, materials retrieved from offsite storage, instructional sessions, requests for permission to publish, and event attendance. The quarterly reports are rolled

Figure 2. Data Dashboard, David M. Rubenstein Rare Book & Manuscript Library, Duke University

27. Barbara Lewis, e-mail message to Joyce Chapman, January 30, 2012.
into an annual dashboard that is attached to the department’s annual report to library administration.

The dashboard has proven to be a good way to share information between departments. Goal setting—facilitated by the dashboard’s data displays—has been helpful in setting expectations and recognizing possible gaps in capacity. The metrics related to processing capacity and research use transactions are helping management to better balance staff responsibilities and strengthen requests for expanded staffing in those areas. Administrators have also been able to use the metrics in presentations to library stakeholders and in benchmarking against peers.28

In 2011, the University of Minnesota (UMN) Libraries’ Archives and Special Collections division conducted a study of reading room user statistics and related staffing costs for a three-year period covering 2007–2010 to evaluate service models and determine when extended service hours should be offered. Staff analyzed existing data on special collections reading room usage and reference transactions by hour of the day and day of the week and then compared the results to general library reference statistics. They also analyzed the costs of current service models and alternate scenarios. Findings enabled them to make data-driven decisions about their operating hours. UMN uses Compendium’s Desk Tracker software29 to track and analyze systemwide visitor patterns and Microsoft Excel (in which reading room visit data are compiled on a monthly basis) to analyze special collections reading room visits.30

The above case studies highlight the breadth of ways in which operational data produced as a by-product of existing services can support data-driven management in special collections. It is likely that more organizations are conducting similar explorations into operational data analysis to inform decision making; nevertheless, the relative novelty of this practice (compared to other areas in which archives have long been pioneers) likely accounts for the frustrating lack of published or otherwise publicly available information for those who wish to initiate such efforts at their own institutions and draw upon experiences and best practices developed elsewhere.

The Need for Interoperable Metrics

The attempts to collect and mine operational data for organizational learning and change depicted in our case studies represent significant efforts to improve prac-

30. Lara Friedman-Shedlov, e-mail message to Joyce Chapman, February 3, 2012; Linnea Anderson, e-mail message to Joyce Chapman, April 5, 2012.
tices at individual institutions. Yet, there is another layer to consider: creating standardized metrics to look across like institutions for peer analysis and benchmarking is critical for the development of the profession as a whole.

In “The Bell Curve,” an article on cystic fibrosis reporting, Atul Gawande argues that the power of knowing what others are doing can make us all better.31 In the 1960s, the Cystic Fibrosis Foundation began collecting anonymous and eventually institutionally identifiable statistics. For the first time, cystic fibrosis centers around the United States could see how peer institutions were doing in terms of mortality rates. Large differences emerged, and it was observed that patients at several centers clearly had more positive results. Following the dissemination of these findings, the centers with the poorest outcomes improved—yet surprisingly perhaps, so did those centers with the best outcomes. While the improvements could not be solely attributed to exposing statistics on mortality, knowing how others were doing enabled everyone to see what was possible. It provided incentives both for centers with lower mortality rates, enabling them to affirm that their practices were good, and for centers with high mortality rates, providing them data to use as a benchmark toward which to aspire.

While the need for standardized metrics has long been acknowledged in the special collections and archival community, it has remained an elusive goal. This is not to say that there have been no attempts to develop consistent means for analyzing archival functions across repositories. In the 1960s, the SAA established a Committee on Uniform Archival Statistics. Its functions were:

- to collect and analyze information about existing archival statistical systems with a view towards (a) isolating and describing these aspects of archival activity which are measurable, i.e., can be expressed in numerical terms; (b) defining these characteristics with a precision that will eliminate confusion wherever a particular term is used; (c) developing standards for archival statistics that will permit meaningful comparisons and studies of archival institutions throughout the country; and (d) encouraging general adoption of these standards by archival agencies.32

---

This unheeded charge strikes us as surprisingly modern, though—half a century later—archives have not resolved the issues it presents.

In 1967, the Committee on Uniform Archival Statistics completed a survey of data collection practices in archives. It found that a majority of the data collected served to document accessions and management functions rather than use. Additionally, few respondents reported actually using the data. For those who did, the primary activities identified were better space allotment and allocation of staff time, more targeted expenditure of funds, better understanding of the “scope of the research trade,” and inclusion in annual reports. In summarizing these data, Herbert Finch, a member of the committee then at Cornell, stated:

Two general conclusions may be made about the current use of archival statistics at this repository. (1) There is very little use of numerical information and (2) the use made is almost entirely for inventory purposes and not for analytical or comparative study.33

By 1973, the Committee on Uniform Archival Statistics had become a committee on “Terminology and Statistics.” Unfortunately for the statistics, this group focused on the publication of the first SAA glossary, which did formulate common definitions useful for archival management if not operational data collection. The next attempt to standardize the data collected by archives was in 1980. In that year, SAA President Maynard Brichford created the Standard Reporting Practice Task Force “with a charge to recommend, and provide units of measure for, the small number of statistical data items that would be most valuable for archives and manuscript repositories to collect.”34 The task force recommended collecting specific information on “Reference Services” including daily visits, requests received by telephone and mail, the number of discrete individuals doing research, and materials consulted by or for researchers.35

During this same period, SAA’s Institutional Evaluation Task Force was conducting a parallel activity by examining which data to collect across repositories as a prelude to institutional accreditation. The task force also discussed capturing reference activity, including the types of users (with distinct user types identified and defini-
tions) as well as different venues for research interaction (such as phone or personal visit). This is the closest that archivists ever have come to creating metrics with data definitions.

This history provides both encouragement and cause for hesitation with regard to current efforts to develop metrics. It is encouraging that past efforts have indicated that the community has recognized the potential benefits of standardized metrics. It is discouraging, though, that previous efforts have had little impact and never progressed past the planning stages. Jackie Dooley and Katherine Luce sum up well the state of affairs, noting that “a lack of established metrics limits collecting, analyzing, and comparing statistics across the special collections community.”

The latest development in this tradition is the RBMS Metrics and Assessment Task Force, which was formed in July 2011 and charged with exploring current practices for gathering and reporting metrics that demonstrate the value and impact of special collections and archives. The group plans to provide an initial set of recommendations by June 2013, one of which may be the appointment of subsequent task forces to work on defining operational metrics for specific functional areas; the hope is that it will be in explicit and official cooperation with SAA.

Meanwhile, the library field in general has been steadily working to develop a body of standards and metrics that may not conform well to the needs of special collections libraries and archives. The agencies promulgating these standards and metrics include the American National Standards Institute / National Information Standards Organization (ANSI/NISO), the E-Metrics Instructional System, the National Center for Educational Statistics, Library Statistics Program (now only academic and school libraries), and, since 2007, the Institute for Museum and Library Services’ Office of Planning, Research and Evaluation (OPRE), which administers the

---


38. For the task force charge and related documents, see Rare Books and Manuscripts Section, Metrics and Assessment Task Force, available online at http://rbms.info/committees/task_force/metrics_assessment/index.shtml.


Data-Driven Management and Interoperable Metrics

public libraries and state library agencies surveys. Within the archives realm, the International Council on Archives (ICA) has developed the International Standard for Describing Institutions with Archival Holdings (ISDIAH), but this archival standard does not interface well with library-based standards. Finally, software applications, such as ATReference and Aeon may also be accelerating the drive toward standardized metrics. We turn now to discuss each of these factors individually and its potential relevance to analyses of operational data in special collections and archives.

The Association of Research Libraries (ARL) annual statistical survey relies on definitions from NISO and E-Metrics. ARL has long collected statistics from research libraries, including special collections and archives at those institutions, whether these are formally under the aegis of the library or not. Anecdotal evidence suggests that archives and special collections have difficulty responding to data requests from library administrators responsible for compiling the data for their institutions. There is also widely inconsistent interpretation of the data definitions; for example, the way that reading room visits and reference transactions are tabulated varies among special collections. Special collections librarians and archivists routinely express bewilderment at how to count reference interactions when an e-mail thread with a given researcher might include several interactions with different staff over several weeks.

ARL borrows many of the data definitions for special collections statistics reporting from ANSI/NISO Z39.7-2004. This standard is important for archivists because it provides guidance for basic archival metrics, even though archivists have had little or no say in its development. One interesting and problematic example is the case of cubic feet. In the original 2004 standard, cubic feet was defined as a “unit of measure for library buildings.” Thus, cubic feet measured building capacity while linear feet was reserved for the measure of collection extent. However, the definition of cubic feet was altered in 2010 to read: “Cubic feet is a measurement of volume for collections.” This change means that cubic feet now competes directly

42. Institute for Museum and Library Service, Office of Planning, Research and Evaluation (OPRE), available online at www.imls.gov/research/.
44. For information about these surveys, see: Association of Research Libraries, Statistics and Assessment at www.arl.org/stats/annualsurveys/arlstats/.
45. For example, at the University of Michigan, the independent William L. Clements and Bentley Historical Libraries are included in the ARL count.
with linear feet and there is no preferred metric for measuring the extent of archival and manuscript collections.

ARL also borrows metrics and definitions from the E-Metrics Instructional System that are incorporated into the ANSI/NISO Z39.7 standard. As creators or sources of digital content, special collections and archives are most impacted by the definitions and metrics pertaining to digital documents and computer files. For example, E-Metrics defines a "digital document" as an "information unit with a defined content that has been digitized by the library or acquired in digital form as part of the library collection. This includes eBooks, electronic patents, networked audiovisual documents and other digital documents, e.g., reports, cartographic and music documents, pre-prints, etc.; databases and electronic serials are excluded."47

In Z39.7 and the ARL statistics, this definition is used as the basis for a definition of computer files, and the corresponding metric is stated as "the number of pieces of computer-readable disks, tapes, CD-ROMs, and similar machine-readable files comprising data or programs that are locally held as part of the library’s collections available to library clients."48 While affecting many library units, this definition presents difficulties for archives, which may be accessioning gigabytes of e-mail and other administrative files in electronic formats, and for special collections, which may hold both low-resolution use copies and high-resolution digital masters of the same item. In these cases, counting items becomes difficult and the utility unclear. ARL is currently examining this definition; it may change in the future.

Some of the data definitions from the NCES/IMLS surveys are also incorporated into Z39.7 and, as a consequence, into the ARL statistics. Others remain unique to the NCES instrument. The NCES surveys of academic, public, and state libraries cater to library-related metrics, but they do have some applicability for archives and special collections (for example, data on preservation expenditures in academic libraries):

Preservation (item 306): Report expenditures associated with maintaining library and archival materials for use either in their original physical form or in some other usable way. This includes but is not limited to binding and rebinding, materials conservation, deacidification, laminating, and restoration. Also, include preservation-related contracts for services (e.g. digitization). Do not include staff salaries and wages.49

---

This is all the more interesting since ARL has dropped the preservation elements in its survey. Since there are no equivalents for special collections and archives, we have had little influence over how these measures are applied in our field.

ISDIAH purports to provide “general rules for the standardisation of descriptions of institutions with archival holdings,” yet it takes a narrative rather than quantitative approach.\(^5\) For example, it calls for narrative descriptions of buildings, holdings, and finding aids, but not counts or other metrics that could be used for benchmarking. Data definitions are broad. By way of illustration, the instructions for recording research services include a rule followed by an example:

**Rule:** Record information about the onsite services provided by the institution with archival holdings such as languages spoken by staff, research and consultation rooms, enquiry services, internal libraries, map, microfiches, audio-visual, computer rooms, etc. Record as well any relevant information about research services, such as research undertaken by the institution with archival holdings, and the fee charge if applicable.

**Example:** The search room at the Surrey History Centre contains a local studies library; seating and table space for 24 researchers; microform readers for 30 researchers and large tables for the consultation of maps. The Centre has a large conference room, which can be divided into two using an acoustic screen. The conference room is fully equipped with the latest audio-visual technology for showing videos and computer generated resources.\(^5\)

Although ISDIAH descriptions offer some basis for comparing institutions, a major opportunity was lost, in our opinion, by not also grappling with a more quantitative methodology.

Data definitions and measures are also asserting themselves into special collections through online cataloging and content management systems and other types of software. Two examples of systems that are designed to collect and manage information about users and their requests are ATReference\(^5\) and Aeon.\(^5\) ATReference records visit dates but not specific times. Aeon enables staff to record the specific times when users enter and leave the reading room, but local policy determines whether users are signed in and out during shorter or longer breaks or only once

---


\(^5\) Rockefeller Archives Center, ATReference, [https://github.com/RockefellerArchiveCenter/ATReference/wiki](https://github.com/RockefellerArchiveCenter/ATReference/wiki).

at the beginning and ending of the day. Aeon also permits institutions to define custom data fields if desired, but that means that institutions might create similar local administrative data outputs with incompatible data definitions. Neither ATReference nor Aeon enforces data definitions, though both potentially can be configured and used to collect data that conforms to community-based standards should they be developed. Archivists and special collections librarians should engage systems engineers in standards development to ensure that community standards and values are incorporated into any statistics generated by these systems.

One respondent to our e-mail query observed how multiple systems hamper the ability to develop uniform statistics at his large institution, which includes several distinct archives and special collections units:

Each repository uses the systems in different ways to record collection use and to collect statistics. [One unit] ... uses the Archivists’ Toolkit for some of our collection management functions. [Other units] ... use Aeon for patron registration and circulation requests. [Another uses] ... local patron registration/circulation systems ... [T]here’s a lack of standardization. One strong argument for using centrally supported systems is adherence to standards, and these days, with increasing attention on comparing ourselves to peer institutions, standardized statistics are seen as a good thing (SCA6).

The standards discussed above—NISO/ANSI, E-Metrics, NCES/IMLS—best support output and not outcome measures. We need both. Output measures, which joined the more traditional input measures as an important area of library reporting metrics in the 1990s, offer quantitative, descriptive perspectives on user services. If done well, they also enable benchmarking and comparison. Outcome measures focus on the effects and changes that are a result of a service; they can be measured qualitatively or quantitatively. While extensive work in outcomes-based assessment has been ongoing in higher education for decades, libraries came to the table much later, and archives came later still. Special collections librarians and archivists need to be aware of the differences between these different types of measures and make strategic choices about prioritizing output as well as outcome measures.

This article discusses output measures. However, one of the respondents to our survey explicitly mentioned collecting data to demonstrate the outcomes of

56. For further discussion of assessing impact, see Anne Bahde and Heather Smedberg, "Measuring the Magic: Assessment in the Special Collections and Archives Classroom," RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage 13, no. 2 (Fall 2012): 152–74.
impact and value, and this goal was implicit for many other respondents who cited administrators and decision makers as key constituencies. The “value proposition” of libraries has been a hot topic for several years now,\textsuperscript{57} and special collections have also begun to participate in this discussion.\textsuperscript{58} Value and impact are both difficult to define and to measure. While neither can be determined solely by quantitative means, operational metrics can be structured to better support research into value propositions and impact. For example, outputs might show the number of courses taught by a special collections librarian or archivist, the number of students in those sessions, and the number of students using the university archives for a course paper. These broad outputs might form the background or context for more directed outcome measures demonstrating how a special collections library or archives contributes to student learning.\textsuperscript{59}

Discussion
Special collections librarians and archivists have been collecting operational data for decades, but they have done little with it besides add it to annual administrative reports. In our various investigations, we found evidence of an emergence over the past few years of a more systematic approach to collecting and analyzing administrative data to provide information to improve operations—or at least a more widespread desire to do so. Yet much more can and should be done. Operational and administrative data more generally have become extremely important both locally and as a shared resource. Locally, operational data can be used to adjust services to meet the needs of patrons and staff, support collection development, processing and deaccessioning decisions, and demonstrate to higher organizational officials that goals and objectives are being met. As a shared resource, such data can be used as a benchmarking tool for evaluating operations at peer institutions.


Perhaps more important, as in the case of the cystic fibrosis data, good and truly comparable data would enable special collections librarians and archivists to see the range of practice and understand what is possible in a new way, beyond site visits and sharing at conferences.

The catalysts for this renewed look at the potential uses of operational data come both from external as well as internal sources. Externally, grant funders, institutional administrators, and software systems are driving the demand for more reporting and evidence of our efficiency, expenditures, outputs, and value as budgets tighten and ROI-driven operations become general practice on a national scale. In the process, external standards and metrics are being imposed on special collections that may not suit our operations, collections, or users. Because we have not stepped up to join the conversation, these external standards may not well represent archival practice. Internally, in times of scarcer resources, special collections librarians and archivists are attempting to use data to manage their workflows more efficiently to accomplish such goals as maximizing access, increasing the impact of outreach, managing collections processing, exploring new staffing configurations, and making intelligent use of physical spaces. Yet these objectives cannot be achieved if methods for data collection, metrics, and tools for analyzing the data are not well operationalized.

The full potential of using operational data is not being reached for several reasons. The absence of data definitions makes it difficult for repositories to model and collect data in a coherent and consistent way over time. If multiple systems with unclear definitions are problematic on the local level, the issues are multiplied across the special collections in one or more institutions and across the archival profession more generally. We hope that efforts such as the Arizona Archives Summit, the Processing Metrics Collaborative, and the newly formed RBMS Metrics and Assessment Task Force will make progress in this area, but to do so they will have to deal with the variety of standards and practices both external and internal to our field. Special collections librarians and archivists lag far behind their colleagues in academic, public, and state libraries in the collaborative establishment of interoperable metrics, and we need to catch up. Externally imposed data collection regimes may benefit those who establish those metrics more than the special collections libraries and archives charged with collecting these data.

Another reason that the full potential of operational data has not been met concerns archival education and special collections staffing models. The next generation of special collections librarians and archivists must be intelligent creators, manipulators, and consumers of quantitative data. They need skills in quantitative data analysis and database querying, as well as knowledge of data definitions and
data modeling. Moreover, they must foster critical thinking skills around quantitative data, providing tools needed to understand the proper and improper use of data analyses to support claims. Competencies such as numerical and database literacy should be added to special collections and archival education.

In terms of staffing models, if institutions have the flexibility, they should consider explicitly asking for relevant data-related skills when hiring for new positions or campaigning for a dedicated percentage of time from a data-oriented staff person in the larger organization. At a minimum, they should make training available to existing staff on how to perform data reporting. Too often, administrators assume that the compilation of requested statistics is a simple and nontechnical task, and responsibilities for data reporting are frequently assigned to staff members without thought for what skills will be required to successfully meet administrative needs. If data reporting requires knowledge of certain software or techniques (such as the ability to write functions in spreadsheet software or SQL database queries), then administrators should support their staff by providing training to develop these skills. Institutions that do not have the ability to hire staff with specialized skill sets or retrain existing staff should reach out to others in their organizations for support for data analysis and data modeling efforts. While central IT staff may be one resource, they are often overburdened and unable to provide in-depth assistance. Other staff, such as assessment librarians, data services librarians, or campuswide institutional research offices, may be better positioned to help.

In this article, we advocate for the archival and special collections community to rise to the challenge of leveraging operational data to support decision making and creating community standards to enable meaningful comparison of metrics across institutions. At present, we see promising signs that the community is interested in moving in these directions, but we have also seen past initiatives fade away. Numbers speak, perhaps more loudly than we would like at times, and they present a profile of an organization unlike any other. If special collections librarians and archivists are to manage the outline of that profile, we need to harness the power of operational data to support better insights and decisions, and we need to take control of how we are measured by identifying the metrics that are relevant to us and defining these metrics in terms that best represent our institutions, collections, and user communities.