Aims and Approaches in Special Collections Assessment: A Case Study from the Osler Library

Introduction
The Osler Library of the History of Medicine, a branch of the McGill University Library in Montreal, Quebec, is a research center for the history of medicine and science with significant rare book and archival holdings. As part of an institutional review launched in 2013, the Osler Library decided to look into methods of collections analysis to compile data about its collections and uncover subject strengths among items, enabling the library to better promote and communicate information about holdings to users, plan for growth, and target collections development.

Collections assessment initiatives in rare books and special collections represent the coalescing of two major professional concerns: rendering discoverable uncatalogued collections, and developing assessment tools to inform the entire range of activities and services of these institutions. A movement since the 1990s to clear backlogs of unprocessed materials has done much to make previously “hidden” collections accessible. More recently, an increased interest in creating a culture of assessment in special collections has emerged as administrative models have shifted. Joyce Chapman and Elizabeth Yakel point to a new attention toward data-driven management in both library and university administration, arguing that archives and special collections must better adapt to this environment. The drive to promote hidden collections, combined with this interest in data-driven management, has launched many initiatives designed to systematically evaluate collections to determine their level of processing, preservation needs, and research value.

Literature Review
For over a decade, librarians and archivists have heeded the call launched by the Association of Research Libraries in 2001 to uncover special collections that were previously hidden due to insufficient cataloguing and processing backlogs. ARL’s decision in the late 1990s to begin collecting data on special collections in university libraries brought to light a crisis of processing backlogs that had left many materi-

This spurred many institutions to conduct surveys of the state of their materials and levels of processing, resulting in a number of different approaches to the data collection. For example, Teper and Erekson describe a collectionwide, item-level survey designed to assess not only the state of processing and description of rare-books material but also conservation information. Although efforts to develop a standardized process for carrying out large-scale surveys largely failed in the 2000s, recent moves to apply library assessment methods to rare books and special collections have reignited interest.

The literature on special collections assessment has been increasingly abundant, including two major reports from OCLC and ARL. Both of these found a general frustration with the lack of standardized metrics to demonstrate success and to aid in data-driven planning. Since then, many authors have written on various aspects of assessment in rare books, including in a 2012 special issue of RBM. In general, this scholarship tends to recognize two types of data that can be collected and applied to improving library services and administration. The first type of information dealt with in the literature is operating data. Chapman and Yakel describe the need to collect data about internal workflows, for example, in archival arrangement and description, digitization procedures, paging, and reference services, to improve processes and increase efficiency, in what the authors term a form of evidence-based archival practice. Novak Gustainis provides a case study at the Center for the History of Medicine at the Countway Library, where a system of collecting data about archival processing has been implemented with the goal of rendering resources usable and readily accessible more quickly. This first type of operational assessment extends as well to transactions between users and the library, including reference services provided to users either individually or as a part of group teaching, and usage statistics for tools offered, such as online catalogues, digital exhibitions, websites, and finding aids.

The second type of assessment discussed in the literature deals with collections data:
accumulating comprehensive information about collection size, subject areas represented, and circulation of materials. Having detailed data about collections is beneficial in various areas, including acquisition, outreach, institutional planning, and the formulation of future projects. Lisa R. Carter refers equally to both rare books collections and programming when she writes, "A longer-term, sustainable future for collections and programs depends on more objective evidence of relevance and efficient, strategic, and judicious prioritization of effort." While there is an existing body of literature that discusses collections evaluation, it is directed largely to librarians in charge of secondary resource collections. Other authors have focused specifically on the problem of assessing archival collections. One of the few studies that are particularly focused on rare books includes Ascher and Ferris' study of number of collection-level survey projects to compare the types of metadata structures used to carry out these projects in an attempt to discern data-collection best practices.

Many institutions have made progress toward capturing new types of operational data or finding new ways to harness existing data, but there are fewer large-scale projects designed to amass collections data. The Consortial Survey Initiative undertaken by the Philadelphia Area Consortium of Special Collections Libraries (PACSCL) is an example of a collections-level survey. The aim was to assess the libraries’ backlog of unprocessed collections and establish priorities for the treatment of collections by identifying those collections that had a particularly high research value. A survey by the Black Metropolis Research Consortium (BMRC) was modeled after the PACSCL survey, with a more thorough evaluation of research value. The Special Collections Library Unprocessed Collections Survey project at the University of Michigan was designed to collect data on access, condition, and material "themes."

Collections Assessment at the Osler Library

The Osler Library’s institutional review was designed to include a self-study of resources and activities, similar to what is done in academic departments, as well as an external evaluation. The stated aims included the following: study the library’s success in supporting the mission and strategic goals of the larger McGill University Library, assess its current performance and long-term goals, and evaluate the stature of the library’s holdings against national and international benchmarks. To address this final point, the Osler Library decided to investigate methods of collections analysis. The assumption was that having more detailed data about our collections would assist in promoting their usage, both in person and digitally, and in continuing to develop holdings. In particular, the library identified three areas in which detailed collections data would be valuable: promotion and communication of holdings, institutional planning, and targeted collections development.

Promotion and Communication of Holdings

A user survey conducted by the Francis A. Countway Library of Medicine found that chief among readers’ criticism of the library was a lack of information about its holdings. Similarly, the Osler Library identified increased knowledge of our holdings as a useful outcome of evaluation, to allow us to better promote our collections among communities of existing and potential researchers. Being able more accurately to assess the cohesiveness of research materials in thematic areas in which a library has major holdings can aid in the development of collections descriptions, which could then be publicized through the website and promotional materials to inform potential patrons. Equipped with a more complete picture of strengths that could be exploited within the collections, librarians would also be able to make overtures to a potentially more diverse group of classes and researchers.

Institutional Planning

The second goal of collections evaluation was preparedness: to have data available to apply toward the planning of upcoming projects. Both analog and digital exhibitions are easier to plan and design with quantifiable data about collection strengths and uniqueness. In addition, institutions must be able to respond quickly and effectively with hard data to take advantage of opportunities for grants and special projects funding for digitization that may arise. In the case of the Osler Library, a branch library within a larger university system, another benefit is being able to quantifiably demonstrate the library’s value to its home institution. As Carter argues, “Reduced resources combined with challenges to traditional academic models and the rapid transformations occurring in research, teaching, and learning require special collections and archives to demonstrate greater impact with more

efficient methods than ever before.” Through our evaluation efforts, the library hoped to highlight the richness and uniqueness of its collections through comparison with similar libraries with significant holdings in the history of medicine to communicate tangible information about the relative value of our collections.

**Targeted Collections Development**

The third objective was to take the information unearthed about subject strengths and weaknesses and apply it to collections development in a meaningful way. A major draw for researchers to a certain institution is the network of related materials on a particular subject, by an author, or from a particular time period that allows a scholar to do sustained research in one location. For example, the Osler Library allows for in-depth research in the history of neurology in North America because of the extensiveness and evidentiary value of its neurological materials: the archival collections of Wilder Penfield and William Vernon Cone, two notable early twentieth-century neurologists, are supplemented by our substantial collection of twentieth-century works on neurology and historical medical journals. These comprehensive networks of thematically related collections provide the primary research materials to carry out major research on a particular subject and a particular time and place. The library hopes to better ground collections development in the data generated by these evaluation efforts to expand on thematic strengths and develop these important networks of research materials.

The challenge, however, in uncovering hidden strengths and describing current ones is clear: how does one define and evaluate qualities such as completeness, evidentiary value, and strength in a meaningful way? Efforts to evaluate research value are discussed somewhat in the literature. This aspect is recognized to be among the most difficult data to assess. In an appendix to her 2012 article, Conway recommends uncovering the research value of a collection through surveying items and using a rating scale from 1 to 5 to indicate the documentation interest of an item (Does it contribute to the understanding of a subject and reinforce other collections?) and its documentation quality (How well does it shed light onto a topic? How complete is the source?). In their report for OCLC on archival assessment, Conway and Proffitt provide some recommendations for evaluating research value of archival collections, acknowledging the trouble and subjectivity attendant to this as well as the value for planning and prioritization purposes. The Columbia University Libraries Special Collections Materials Survey gives another example of a rating system devised to determine intellectual value of a collection (ranging

---

The BMRC’s survey, as well, gives deep thought into how to rate research value. Initiatives to uncover subject strengths are somewhat less developed, although some surveys have attempted to gather information on topical areas of collections. The model used by the University of Michigan in their Unprocessed Collections Survey assigned each item to a particular theme, allowing for both a primary categorization and secondary categorization based on a predetermined list of subject areas. This approach, however, is based on the assumption that one can identify a comprehensive list of strong topical areas—something that the Osler Library sought to define rather than presume.

The library sought out technological methods for evaluation rather than the more resource-intensive models. The first approach consisted of engaging a company called Sustainable Collections Services (SCS) that uses catalogue data in comparison with WorldCat records to generate reports about the amount of unique holdings in a collection. The second approach is one has been used at McGill’s Rare Books and Special Collections by Jillian Tomm in her analysis of the composition of McGill’s eighteenth-century collections. It consists of extracting catalogue information and processing it using digital humanities (DH) content analysis and data visualization tools.

**Gauging “Uniqueness”**
The McGill Library worked with SCS to generate a report on the uniqueness of Osler’s holdings as compared to both peer libraries and global holdings listed in WorldCat. The libraries used for comparison were those in the American Association of Universities (AAU), the Canadian Association of Research Libraries (CARL), and the Bureau de coopération interuniversitaire (BCI), the organization that encompasses Quebec’s universities. For comparison, rare books were divided by time period of publication to have counts of the percentage of unique items per time period. Of the two reports produced, the first contained uniqueness counts for titles that have no specific edition match in WorldCat. The second consisted of uniqueness counts for titles that have a match in WorldCat, based on any of the book’s editions. A title-by-title comparison list was also provided that gives the number of books found overlapping for each individual title.

---


24. University of Michigan, “Special Collections Library Unprocessed Collections.” Other collections surveys that have included information gathering about scope, content, and subject areas include the Bancroft Library at the University of California, Berkeley, “Manuscript Survey Project.”

Both reports show a high number of uniquely held items. Overall, twenty-four percent of titles in Osler subcollections were reported as unique. The “all editions” match showed lower rates of item uniqueness because the search is more generous and retrieves more titles. The edition-specific matching in WorldCat is more conservative and likely underreports the actual amount of same items held in other institutions. This undercounting is due to editions that may sometimes be “lost” due to misclassification. WorldCat creates a single record for each book title entered, and then different editions of that title are added to it. When a book is not catalogued uniformly, WorldCat may not be able to add it to the correct existing record. Despite the inexactness arising from such cataloguing issues, the numbers are impressive, particularly in comparison to other libraries in CARL and BCI and in respect to pre-1800 materials, confirming the library’s role as a unique resource in Canada and Quebec.  

Nonetheless, it is problematic to make claims of absolute uniqueness of items based on these data. Special collections are still faced with significant backlogs of uncatalogued material or items for which online catalogue records are unavailable. The library could state that, according to the report, a certain book is unique based on comparison with others that have included their books in WorldCat; but when we consider the breadth of materials that may not be locatable through WorldCat, this caveat becomes quite significant.

Another issue that interferes with the usefulness of this data is inconsistent cataloguing. The problem of consistent rare books cataloguing standards and best practices is long-standing. For many years, rare books had their own “cataloguing” standard of descriptive bibliography, entailing complete title-page transcription and close physical description. In recent decades, special collections have frequently become centralized into a wider institutional context, including in the sharing of resources and staffing, and rare books cataloguing has adapted to changing cataloguing codes. As technologies and standards have changed, rare books cataloguers have sought how to best serve the needs of rare books, which, with their level of detail and unique physical characteristics, sometimes resist uniform description. “Rare books are also prone to inconsistent cataloguing because they require specialized knowledge above and beyond general cataloguing.” In our case, this

26. Of 5,228 pre-1800 titles counted, 5,132 were unique as compared to BCI libraries (98%) and 4,265 as compared to CARL libraries (81.5%).
30. For example, “Special collections cataloguing may require use of additional descriptive standards beyond Anglo-American Cataloging Rules and familiarity with other formats, as well as knowledge of the history of printing and publication.” Russell, "Special Collections Cataloguing," 299.
problem of cataloguing nonconformity has likely led the SCS report to undercount the amount of edition-specific matches.

Analysis
The library analyzed these uniqueness reports using a method developed by Jillian Tomm to verify a number of items identified as unique in the edition-specific matching. The Osler Library’s Head Librarian, Christopher Lyons, provided a list of libraries with medical collections thought to be comparable to those of the library.31 Twenty titles from various time periods were chosen for verification against the holdings of these libraries through searching their catalogues, as well as double-checking in WorldCat. The records were selected at random to avoid any proximity or clustering in the list of titles that would affect the results.

Although the selected group is only a small percentage of the total number of “unique” books that the report found, the verification process did bring up issues that question the results of the uniqueness reports globally. One book identified as unique illustrates the problem of the large extent of material that cannot be found through online library catalogues or WorldCat. The Anatomia of Mondino dei Luizzi (Pavia, 1492) appears to be “unique” according to the criteria laid out for the verification process (that is, no records of other copies in WorldCat and no copies held at the comparison libraries), but it is not unique in absolute terms: there are at least two other known copies of the same edition in Italian libraries.32 Other books illustrate the great variety in the way in which early printed books are catalogued such that even basic descriptive elements, like title and author, can differ widely. Judicium conjunctionis duorum superiorum of Paulus de Polonia (Rome, 1484) has the alternative spelling Judicium coniunctionis in WorldCat, which brings up a record for another copy of the same incunable at the British Library. A later book identified as unique, Contemplatio oculi quam inclytae Facultatis Philosophicae (Leipzig, 1639), exists in multiple copies locatable through WorldCat, but all with highly variant information in both the title and author fields. This finding illustrates a previous caveat about a lack of uniform cataloguing affecting the meaningfulness of such uniqueness analysis. Considering that standards for describing rare books in print handlists and catalogues have changed over the years, as have electronic cataloguing standards, it is no surprise that many rare book records found in online catalogues do not follow current practice. At the Osler Library, for example, when the OPAC was introduced, bibliographical information was taken directly from the early twenti-

31. These were: the Wellcome Library, the National Library of Medicine, the Countway Library, Yale University Libraries (including the Cushing Whitney Library), the New York Academy of Medicine, the College of Physicians of Philadelphia, Johns Hopkins University Libraries, the Bibliothèque interuniversitaire de santé, and the Bibliothèque nationale de France.
32. The libraries of the University of Cagliari and the University of Bologna at Forlì; located through the British Library Incunabula Short Title Catalogue, available online at www.bl.uk/catalogues/istc/index.html [accessed 2 September 2013].
Aims and Approaches in Special Collections Assessment: A Case Study from the Osler Library

eth-century catalogue, the Bibliotheca Osleriana, and was input without revision to up-to-date cataloguing standards. The McGill Library record for Contemplatio oculi is sparse and the title improperly transcribed. A preliminary search for the correct title and another variation reveals at least three other copies in WorldCat.

Text Analysis
The second approach to collections evaluation relied on the use of a digital humanities tool designed to analyze large corpora of text. Many tools designed for content analysis work on the basis of drawing out common words and phrases and plotting their appearance over time. Such text analysis tools can be used with cataloguing data to uncover information about the breadth, scale, and subject strengths of a collection.

FIGURE 1. Word Cloud Created with Cirrus from Osler Title Data, 1450–1913

The first step in performing content analysis was acquiring cataloguing data from Aleph, McGill’s integrated library system. A large search was run to isolate all of the records of books in Osler published between 1450 and 1913. The data were

34. Two copies in the British Library and one in the state library in Berlin.
extracted with the help of McGill Library’s Digital Initiatives and exported to an Excel spreadsheet using the program MarcEdit.\textsuperscript{35} Duplicates, symbols, and other superfluous information were removed from the data. After this processing was finished, the resulting lists of text were run through a DH tool called Voyant,\textsuperscript{36} which provides an environment in which to experiment with different methodologies of text analysis and ways to visualize textual data.

Text analysis was run on the list of Osler rare book titles from 1450 to 1913 using a stopword list in English, French, German, and Italian, including articles (the, a, le, un, and the like) as well as genre terms unrelated to a book’s contents (études, observations, report, and such). The Cirrus visualization tool was one tool selected to review the results.\textsuperscript{37}

The first thing that jumps out of the text cloud is the predominance of French: fièvre is the most frequent word. This makes sense when we consider historical perceptions of illness—fever, from medieval medicine until the advent of twentieth-century biomedicine, was not considered a symptom but a category of illness. The predominance of this word does not lead us to a subject area, but it confirms something historians know about the history of medicine. Because of the frequent appearance of French words, it is likely that our collection of over 30,000 Paris medical theses is skewing the results. A subsequent test used the subject headings rather than title data to eliminate this issue, since the French theses have no subject headings in the catalogue.

**Challenges**

There was one problem with the cataloguing data observed from the outset. The original search to locate items printed between 1450 and 1913 only retrieved items with dates for which all four digits are present. This means that any uncertain dates in which symbols are used to replace numbers, such as 15– or 154?, are not included in the search. There are also a couple of issues to take into consideration when analyzing the usefulness of the results. The title word clouds miss items that are not in the two major languages represented. There may be holdings in languages other than French or English that add to the strength of a particular subject, but the keywords do not appear frequently enough comparatively to appear in the word cloud. Another obvious caveat is that Voyant can compute only Roman character languages; the substantial number of Arabic-language items in the collections is discarded in this analysis. Many items are, as well, better catalogued than others and have more or fewer subject headings assigned depending on the quality of the record. There is also more that can be done to improve the accuracy of results.

\textsuperscript{35} Terry Reese, “MarcEdit,” available online at http://marcedit.reeset.net/ [accessed 2 October 2013].


\textsuperscript{37} Sinclair and Rockwell, “Voyant Tools.”
Implications for Practice and Conclusions

The first benefit of text analysis is identifying a number of possible paths for deeper collections research to take place; for example, having uncovered the predominance of the subject headings obstetrics or public health, a next step would be further to investigate the completeness of the group of related holdings. This would ideally be carried out by a curator or subject specialist librarian, drawing on existing work in the field and aided by traditional tools such as subject bibliographies and checklists, to further refine the comprehensiveness and evidentiary value of the coverage. Previously underacknowledged subject strengths could open the possibility to reach more varied groups of researchers, students, and instructors.

A future direction for this approach to collections analysis would be to design a search that would balance the discrepancy in quantity between earlier and later material. This analysis could be done by separating out titles or subject headings into groupings by century and then running these groupings through Voyant. After that, the frequency counts generated in Voyant could be calculated with the totals of items per century as reported by the SCS reports to determine the actual percentage of the most frequently appearing words. These percentages could then be compared across periods to arrive at a truer glimpse of subject strengths. The technique can also be used for any metadata fields, including publication locations, authors, and so on.

Another tangible end product that libraries may consider pursuing is a set of detailed collections descriptions to promote collections to research communities, particularly on the web. Much has been written on the new audiences to which special collections are appealing, and on the importance of their being able to identify libraries and research collections of interest through the Internet. Content analysis tools can be used to sound the depths of a particular collection, compiling information on date concentrations, countries and cities of publication, subtopics, percentage of illustrated materials, and the like, to promote collections online and aid in discovery. Comparative data about relative scarcity could also be brought to bear in collections descriptions to appeal to researchers.

Improved targeting of collections development is also a natural outcome of subject strength analysis. Both in selecting items for purchase based on their relationship to significant groups of thematically related material and in justifying important
purchases or attracting interested donors, the addition of quantifiable data about the depth, coverage, and relative scarcity of subject subcollections is useful. Such new data could also be applied in reviewing collections development policies.

The conclusion drawn from analysis of the Sustainable Collections Service uniqueness report is that, while limitations exist on the meaningfulness of the data due to variant cataloguing and other issues, the information can be brought to bear in institutional decision-making. As well, digital humanities tools for analysis of cataloguing data offer a flexible, scalable, and highly customizable way to conduct collections assessment. Both the techniques outlined appear to be promising approaches to large-scale collections assessment that require fewer resources and staff time than more traditional surveys, provided that collections have some level of cataloguing already.

In the realm of institutional planning, the title-by-title list from SCS has already demonstrated its usefulness locally. In 2014, the McGill Library identified the a better way to direct planning and decision making for digital projects as an institutional priority. A group of librarians from digitization, rare books, archives, and collections was assembled to “create [a] matrix of highest digitization priorities for Archives and Rare Books & Special Collections, to drive future digitization activities and guide responses to digital project and digitization requests from users, partners, and library staff.” The group met seven times between the summer of 2014 and the summer of 2015 to design a matrix that would determine the relative priority of digital projects. The completed matrix takes into account factors such as rights and privacy status, research value, demand, condition, and format. The availability of uniqueness information about specific titles has informed a number of criteria on the new digitization matrix. Uniqueness and relative scarcity, relationship to subject and collecting strengths, and relationships to other collections and partners beyond McGill are now criteria factored into decision making—all pieces of information that are on hand through the uniqueness report from SCS.

Finally, although none of the items from the verification group drawn from the SCS report were found to be “unique” per se, the results nonetheless demonstrate that the library has a quantifiably excellent collection of relative scarcity. In an age of underfunding, any information that can be gathered about a library’s stature must be seized upon for communications with administrators, donors, and stakeholders. While imperfect, the SCS report nonetheless provides a way of assessing collections against local, national, and international benchmarks. Information such as this can define quantitatively what special collections already know to be their distinctive expression of a library, university, or community’s identity.