Evidence Based Library and Information Practice

Research Article

Analyzing and Assessing a Library Collection Using Faculty Citations Via OpenAlex and R

Sylvia Orner
Collections and Resource Management Librarian
Weinberg Memorial Library
The University of Scranton
Scranton, Pennsylvania, United States of America
Email: sylvia.orner@scranton.edu

Received: 18 Dec. 2023 **Accepted**: 8 Aug. 2024

© 2024 Orner. This is an Open Access article distributed under the terms of the Creative Commons-Attribution-Noncommercial-Share Alike License 4.0 International (http://creativecommons.org/licenses/by-nc-sa/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly attributed, not used for commercial purposes, and, if transformed, the resulting work is redistributed under the same or similar license to this one.

DOI: 10.18438/eblip30493

Abstract

Objective – Citation analysis is becoming a popular means of analyzing and assessing library collections due to its relatively unobtrusive nature and the growing accessibility of citation data. The primary goal of this study was to assess whether the library at the University of Scranton is successfully meeting the research needs of faculty based on analysis of faculty publication and citation data from OpenAlex's application programming interface. Secondarily, this study analyzed faculty publication and citation patterns to help identify opportunities for the library to better support faculty in their research and publishing.

Methods – This case study focused on a citation analysis of the University of Scranton's faculty publications from 2013 to the present. Using OpenAlex and R computing language as non-proprietary sources of data and data analysis, faculty publications and citations were examined and compared to current library holdings.

Results – Overall, 16,786 unique citations from 1,045 unique faculty publications were examined and compared to a list of current library holdings. Findings concluded that approximately 65% of citations were available through library holdings. Further analysis of faculty publication practices

suggested that there are a growing number of faculty publishing open access which indicates that there may be additional opportunities to support faculty in this area.

Conclusion – While this case study represented specific needs and use cases at the University of Scranton, the ultimate importance of this study is the process itself. The use of non-proprietary tools and data sources like OpenAlex and R create exciting new opportunities for others who wish to conduct similar studies at their own institutions without relying on proprietary tools and data sources or resorting to more labor-intensive methods.

Introduction

Citation analysis has long been used as an unobtrusive method for collecting data on library materials and collections that are being used in both student and faculty research. Depending on the tools being used, it can also be a labor-intensive process. Sources of scholarly data like Scopus and Web of Science can be limiting unless an institution has a subscription, and cross-referencing citations against library holdings can be time consuming if done manually.

With the 2022 launch of OpenAlex, a free and open catalogue of scholarly data, new possibilities emerge for librarians who wish to conduct citation analyses without relying on Scopus or Web of Science. Since OpenAlex relies on application programming interface (API) for data retrieval, it is easy enough to create an automated or semi-automated workflow using programming languages like R or Python.

The need for this kind of case study at the University of Scranton's Weinberg Memorial Library arose after the completion of a full library collection analysis. After examining the collection for general age, subject representation, and types of materials collected, a way to better understand if and how this collection was meeting the needs of users was required. An analysis of faculty publications and citations was chosen because faculty data were deemed to be most readily available and, with the use of OpenAlex, most easily accessible. As noted by Watson (2010), there are unlikely to be any privacy concerns with this type of data since faculty are generally publishing with the intention to share and disseminate their work. Primarily, the goal of the study was to assess whether the library is successfully meeting the research needs of faculty based on an analysis of faculty publication and citation data via OpenAlex. Secondarily, the study analyzed faculty publications and citation patterns in an effort to identify opportunities for the library to better support faculty in their research and publishing.

To achieve these goals, the following questions were considered:

- Approximately how many publications have faculty produced in the past 10 years?
- What types of publications are faculty producing, and what types of publications are they citing?
- What specific publications are being cited by faculty, and how often is the library able to provide access to those publications?
- Are there any significant gaps in the library's current resource collection as evidenced by publications cited that the library is unable to provide access to?

However, before embarking on such an analysis, careful consideration of sources of data and methodology was needed. The University of Scranton does not maintain an institutional repository or

other definitive list of faculty publications. For the purpose of this study, an efficient way to generate a suitable list from scratch was required.

Literature Review

In general, libraries collect vast quantities of data. Kelly and O'Gara (2018) acknowledged that the combination of collections data, usage data, and citation date represent a massive quantity of information and advised selecting appropriate data points to answer specific questions or address specific assessment goals. While specific outcomes for the studies consulted varied from institution to institution, the most common goals associated with citation analyses included creating a core journal list (Kayongo & Helm, 2009; Martindale, 2020; Vaaler, 2018), assessing the usefulness of a collection (Fernández-Ramos et al., 2023; Matos, 2016; Peñaflor & Aliwalas, 2022; Smith, 2003), and informing collection development decisions (Feyereisen & Spoiden, 2009; Gao, 2016; Wilson & Tenopir, 2008; Zhang, 2007). Tucker (2013) also used citation analysis as a more general tool to glean insights on faculty research and publishing practices.

Some studies used Trueswell's 80/20 rule as a method of assessment. Applying this rule to a library journal collection, the assumption is that 80% of all usage will come from 20% of subscribed journals.(Trueswell, 1969) Of the studies that sought to apply this rule, Pastva et al. (2018) and Vaaler (2018) found that only certain disciplines adhered to the 80/20 rule. Kohn and Gordon (2014) and Martindale's (2020) findings did not agree with the 80/20 rule at all indicating that this rule may not be a useful indicator for assessment, particularly when considering a large, interdisciplinary collection.

The use of citation analyses and methodologies are well documented in the library and information science literature, though the scope and tools used tend to vary greatly. Hoffman and Doucette (2012) provided a comprehensive review of methodologies but note that specificity and reproducibility vary from publication to publication. Earlier researchers (Currie & Monroe-Gulick, 2013; Feyereisen & Spoiden, 2009; Sylvia, 1998) tended to rely on processes that required manual extraction of publication and citation data and additional labor to gather, clean, and manipulate the data in a meaningful way. They primarily relied on storing and analyzing data in Excel spreadsheets. Because of the time-consuming nature of these types of analyses, limitations on scope were often imposed.

More recently, researchers like White (2019), Kumpulainen and Seppänen (2022), and Pastva et al. (2018) focused on automated or semi-automated approaches using either Python or R scripts in conjunction with citation databases like Scopus or Web of Science to create a more streamlined approach to harvesting data. These researchers were able to generate more citation data than in many earlier citation analysis studies but also relied on knowledge of basic coding and API usage in order to achieve the required results. In the study conducted by Kumpulainen and Seppänen (2022), data from both Scopus and Web of Science were combined in order to create a more complete dataset. These researchers noted the complexity and added labor of bringing together two or more datasets.

In general, most previous researchers using citation analysis utilized data from 5-10 years' worth of publications. In some cases, analysis was limited to certain schools or departments (Currie & Monroe-Gulick, 2013; Gao, 2016; Kayongo & Helm, 2009; Ke & Bronicki, 2015; Martindale, 2020; Peñaflor & Aliwalas, 2022). Others utilized undergraduate research (Kohn & Gordon, 2014; Sylvia, 1998) or master's and honors theses (Feyereisen & Spoiden, 2009; Smith, 2003). However, Feyereisen and Spoiden (2009) did note that relying on student data may not necessarily indicate appropriateness of a collection because

students may not be as confident in identifying sources and may tend to rely more on what is readily available.

Methodology

During the review of relevant citation analysis literature, it was found that most studies utilized faculty citation data from a period of time between 5 and 10 years. Since the University of Scranton is not primarily a research institution and faculty publication output may be somewhat smaller when compared to larger universities or research institutions, the decision was made to focus on 10 years of data with the thinking that it would generate a larger dataset that was more representative of the organization as a whole.

For this study, the primary source of data was faculty publication data extracted from OpenAlex. OpenAlex was chosen as a data source because it is fully open source and easily accessible via API. It was launched in 2022 to replace Microsoft Academic Graph, and, at time of launch, it contained metadata for 209 million works and 2,013 million disambiguated authors (Priem et al., 2022). Since its launch, those numbers have grown significantly with the total number of works being 243 million at the time of this writing. It draws data primarily from Microsoft Academic Graph and Crossref, however, it also relies on data from other sources like ORCID, Research Organization Registry (ROR), and Unpaywall to create a robust and comprehensive data source searchable through a single interface (OpenAlex, n.d.). By comparison, Scopus and Web of Science, two of the most cited data sources in previous studies, contain approximately 87 million works each (Open Alex, n.d.).

The inclusion of additional data from organizations like ORCID and ROR is key here. OpenAlex is not merely a catalogue of publications. While publisher metadata does not often include information like ROR numbers or other affiliation data, OpenAlex employs the use of algorithms to disambiguate and create connections between the many data sources it harvests (OpenAlex, n.d.). This allows for the creation of a more robust dataset than might have been obtained if looking at publisher metadata alone.

So rather than trying to create a comprehensive dataset using multiple sources like Kumpulainen and Seppänen (2022) did, OpenAlex could be used as a single source of citation data. Additionally, given that no comprehensive list for the University of Scranton faculty publications currently exists, and the practice of extracting the data from faculty curriculum vitae is both labor intensive and dependent on faculty making their curriculum vitae easily available and up to date, OpenAlex was deemed the best choice for identifying the University of Scranton's faculty publication data. While it will certainly be not a complete list, it was the best way to generate a good sample of data.

The initial dataset was retrieved via the OpenAlex API using R and the openalexR package (Aria & Le, 2023) by looking at all works published on or after January 1, 2013, that had at least one author whose affiliation matched the University of Scranton's ROR identification number. This brought back a list of 1,192 works, each identifiable by a unique OpenAlex ID number, as well as the title of the work, the authors, author affiliation, type of publication, date of publication, publisher, series or journal title (where applicable), and list of unique OpenAlex IDs for all citations as well as the open access (OA) status of each work.

The dataset was then examined for accuracy as far as OpenAlex's ability to correctly associate authors with the correct institution. Because OpenAlex included information for all co-authors of a work, code was run to immediately remove any authors not associated with the University of Scranton's ROR ID.

Once this list was established, it was run against a list of current University of Scranton faculty. The resulting comparison turned up 147 works out of 1,192 where the author did not match the current faculty list. On further examination, the authors of these works were found to be associated with a nearby institution, Penn State Scranton, and were incorrectly associated with the University of Scranton's ROR ID. Once identified, the publications of these non-University of Scranton authors were removed from the dataset.

Additionally, the decision was made to consider the faculty publication dataset as a whole rather than focus on publications of faculty from particular academic departments as has been the practice in some of the previous analyses cited. This decision was made because the primary focus of this project is to determine whether or not the library is successfully meeting the research needs of faculty, and since the library supports faculty across all academic departments and disciplines, it made the most sense to use the entire faculty publication dataset as generated by OpenAlex.

The academic department of each faculty author was considered to ensure that there was representation from each department and that no academic department or individual faculty member accounted for a disproportionately large percentage of dataset. Finding that all academic departments were represented with no single department or faculty member accounting for a significant portion of the data, the OpenAlex dataset was deemed reasonable to accomplish the primary goal. It should be noted that the OpenAlex dataset should not be considered a complete dataset of every single faculty publication but rather a sample of that data. Works that do not have digital object identifiers may be excluded from the OpenAlex list, but given the examination of the dataset, it was decided that it accounted for a reasonable sample of faculty publications and citations.

With the decision to use the OpenAlex dataset of the University of Scranton's faculty publications, the next step was to examine the faculty publication dataset for any trends that might be relevant to the secondary goal of identifying opportunities for the library to better support faculty in their research and publishing by exploring faculty publishing practices. To that end, information about publication type and OA status was gathered and analyzed.

Once the faculty publication dataset had been generated, R code was used to gather and deduplicate the list of unique OpenAlex IDs for each list of citations. This list was then used to retrieve another dataset via the OpenAlex API. Similar to the faculty publication dataset, this citation dataset included information on the title of the work, the authors, type of publication, date of publication, publisher, series or journal title (where applicable), and a list of unique OpenAlex IDs for all citations as well as OA status of each work. Here, again, the works cited were analyzed for trends as far as types of publications and OA status.

Additionally, in order to determine the general age of the publication when it was cited by the University of Scranton author, the citation dataset was joined to the faculty publication dataset using the referenced works column. This allowed for easier comparison of the date the cited work was published with the date the citing work was published.

Since the data suggested that faculty were largely citing journal articles (93% of the time), the decision was made to focus specifically on the library's ability to meet faculty research needs with regards to journal subscriptions.

With the focus being journal subscriptions, the citation dataset had additional cleanup code added to remove citations that were not journal articles (e.g., books, book chapters, reference entries). Any gold OA titles were also removed. The citation dataset was run against a list of current print and electronic journal holdings extracted from the library's primary knowledge base, EBSCO Holdings Management. The code attempted to match each citation to current holdings based on journal title and date of publication. Various data points were considered in order to find a reasonable match point between the journal titles in the citation dataset and the journal titles in the holdings dataset. Unfortunately, the holdings dataset from Holdings Management was missing ISSN information in more than 30% of its records so this was deemed to be unsuitable as a match point. Ultimately, journal title was selected as the match point. In order to maximize the success rate for the code, both sets of journal titles underwent a normalization process where all capitalization and punctuation was removed. Also removed were leading "the"s, and ampersands were replaced with the word "and". Coverage dates in the holdings dataset were also modified so that they were similar in syntax to the date of publication in the citation dataset and coverage end dates for ongoing resources were set to the current date. It should be noted that various iterations of data normalization were conducted with both the citation dataset and the holdings dataset. This is the one that produced the highest match rate.

The final dataset was generated by joining the citation dataset to the holdings dataset in R with title as the match point and additional filtering to ensure that the date of publication in the citation dataset was within the coverage period of the holdings dataset. This dataset represented all journals cited that are currently held in the library's collections based on both date of publication and journal title. A secondary dataset was created to account for all journals cited that either had no holdings in the library's collection or where the article cited fell outside of current coverage dates. This was achieved in R by creating a new dataset of all titles appearing in the citation dataset but not appearing in the overlap dataset. This no-access dataset was further examined to determine if certain journals were being cited with a frequency that might warrant a subscription or if there were identifiable gaps in our journal collections.

Results and Discussion

With the removal of the publications where Penn State Scranton's authors were mistakenly associated with the University of Scranton, a total of 1,045 University of Scranton faculty publications were examined. As far as the types of materials faculty members publish, out of 1,045 publications, 932 were journal articles. There were also 81 book chapters, 12 books, 4 reference entries, 2 editorials, and 14 others. So, in the past 10 years, approximately 89% of faculty publications were journal articles.

Figure 1 offers more detailed information about the number of publications year over year. In general, University of Scranton faculty averaged approximately 97 publications per year with an average year to year change of about +/- 15 publications.

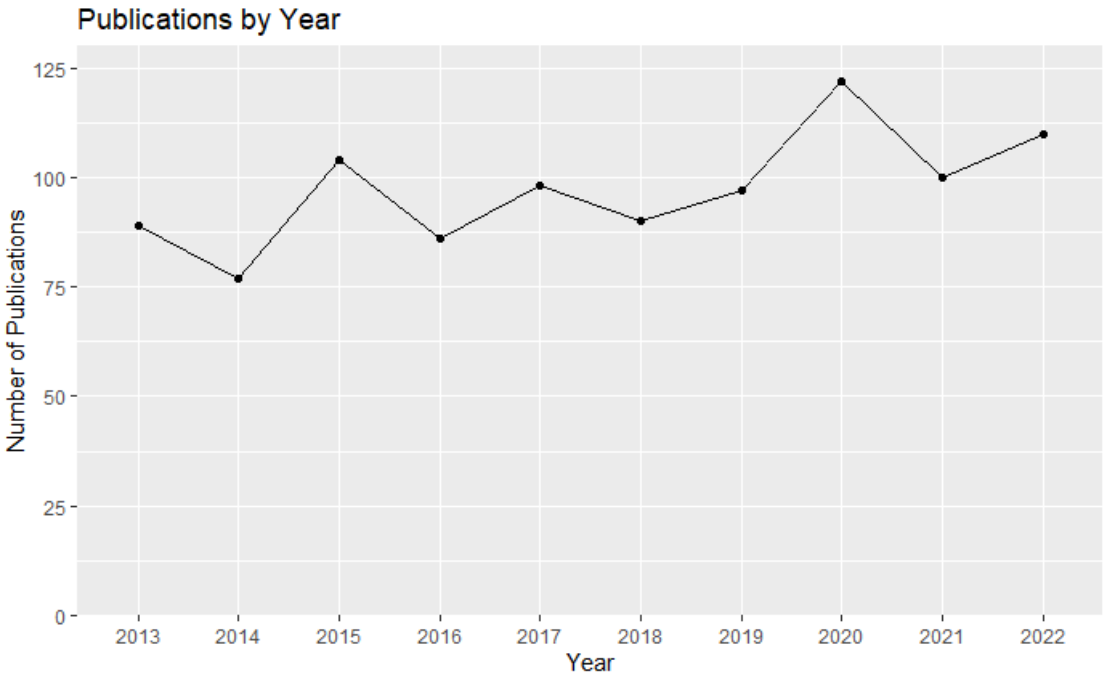

Figure 1
Publications by year.

Since OpenAlex also provides information on OA materials and since the library actively promotes the use of OA and open education resources, publications were further examined to determine if faculty are publishing OA or using OA materials in their research. Findings here were quite encouraging. Of 1,045 publications, 248 were OA. Overall, that accounts for approximately 24% of faculty publications. Furthermore, when looked at on a year-to-year basis, OA publishing among the University of Scranton's faculty appears to be growing (as seen in Figure 2). From 2017 to 2022, there has been an approximately 66% increase in the number of OA publications by University of Scranton faculty.

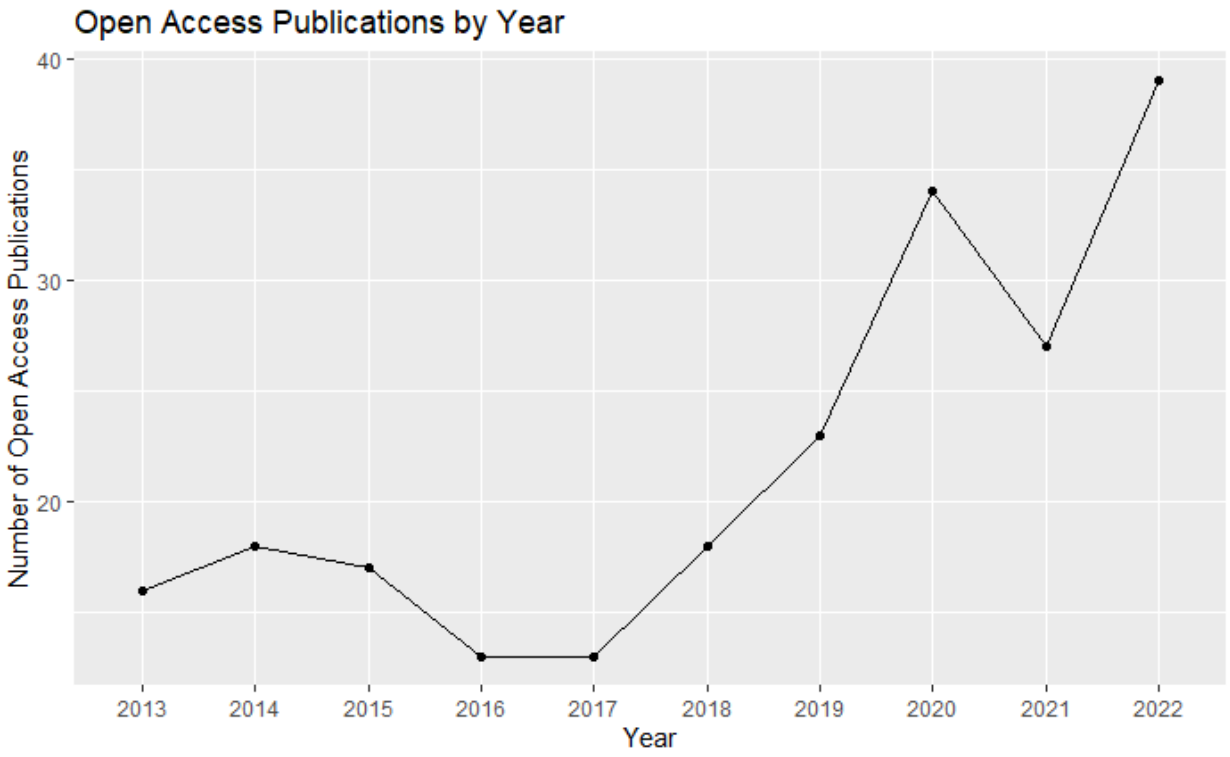

Figure 2
OA publications by year.

When considering the citation data, 16,786 unique citations were evaluated. Age of publication when cited was examined. By comparing the date of publication of the cited article with the date of publication of the citing article, the general age was determined. I found that the majority of faculty citations were from publications that were less than 20 years old at the time of citation (see Figure 3). Overall, it is not surprising that more recent publications were utilized most often, but this information may help inform future decisions about material retention especially when considering how long to keep print materials and when to purchase electronic backfiles.

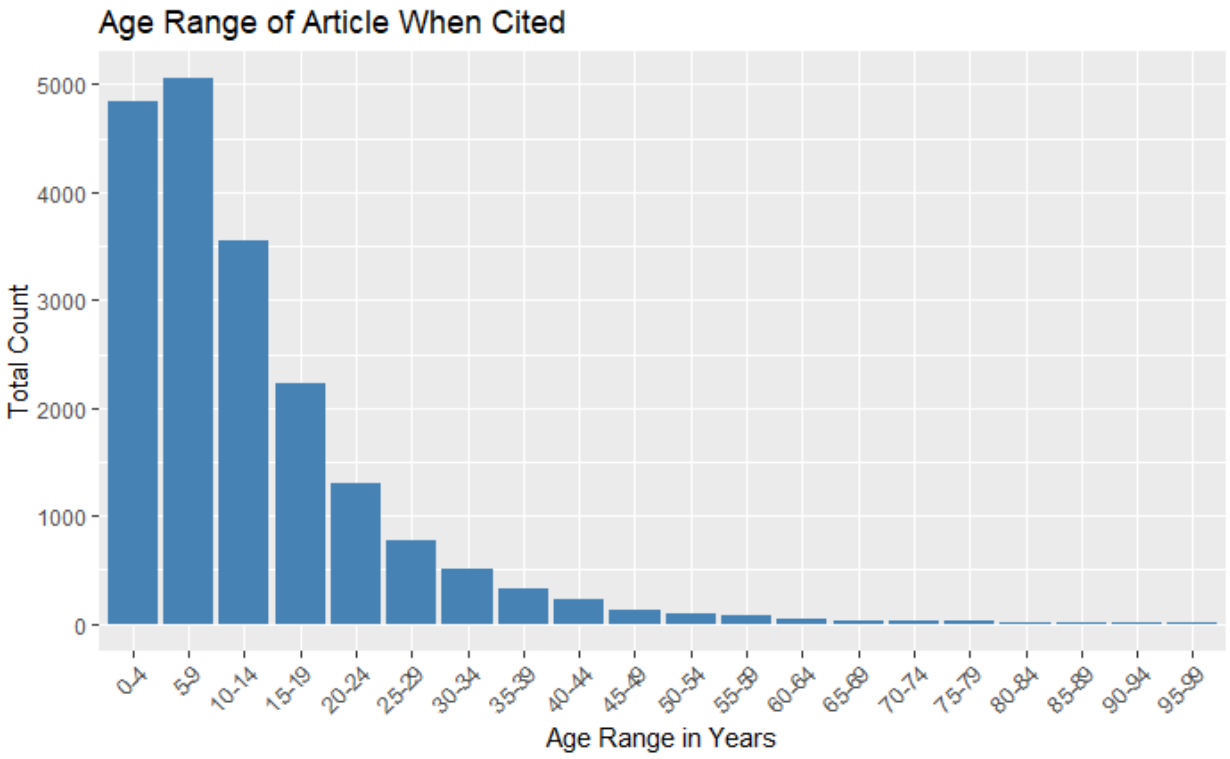

Figure 3
Age range of publications when cited.

Citations skewed largely toward articles (approximately 93%) which led to the decision to focus primarily on the library's ability to meet research needs with journal subscriptions specifically.

Citation of OA materials was considered, and it was found that 24% of articles cited were OA. Additionally, when considering journal titles cited, approximately 12% of journals cited were gold OA.

After the removal of the gold OA journal titles, additional R code was run to determine unique journal volumes and titles. There were 7,598 unique journal volumes from 3,803 unique journal titles found in the citation dataset. With 2,490 unique journal titles appearing in the overlap list created by running the citation dataset against the holdings dataset, it was found that the library was able to provide access to approximately 65% of articles cited within the examined timeframe.

Further consideration was given to the remaining 1,313 journal titles that appeared in the no-access dataset. Overall, the median number of times a no-access title was cited was one. Only 1% of journal titles (13 titles) in this dataset were cited more than 20 times. On further examination of the higher usage no-access titles, it was found that 1 title was incorrectly appearing on the list with 24 articles cited where we had no access. In reality, all of those articles were covered by backfiles access, but due to a gap in our coverage for this particular title, the code was not successfully able to match those articles to our backfiles coverage. Fortunately, such gaps in coverage are not common in the University of Scranton's holdings, and this was the only title found where a coverage gap led to an incorrect match.

Additional consideration was given to the remaining 12 unsubscribed titles with higher citation counts to determine whether a subscription would be beneficial and cost effective. While use and usefulness is clearly suggested by the number of times these titles have been cited, cost must also be considered. The combined yearly subscription cost for all 12 titles would have been approximately $20,000 with the least expensive title being approximately $1,200. Considering the average cost of interlibrary loan at the University of Scranton (approximately $40 per article), it was ultimately deemed not cost effective to pursue subscriptions to the 12 highly cited titles.

Analysis and Assessment

As previously stated, the primary goal of this study was to assess whether the library is successfully meeting the research needs of faculty based on data gathered from faculty citations via OpenAlex compared to current holdings data. Overall, the Weinberg Memorial Library's journal subscriptions are meeting the research needs of our faculty approximately 65% of the time in the given timeframe. Considering budgetary limitations and the wide variety of disciplines and academic departments the library supports; this is a better number than expected. With regard to no-access journals being used, overall, data suggest that these titles are not being used with enough frequency to warrant adding them to current subscriptions especially when subscription costs were compared to interlibrary loan costs.

Additionally, with the general age of cited materials falling mostly within the past 20-30 years, this may be an indication that older materials are only used situationally and that the purchase of electronic journal backfiles can be more selective.

As a secondary objective of this project, faculty citation and publication data were examined in an effort to identify areas where the library might be able to offer additional support for faculty. Given the increase in OA materials being published by University of Scranton faculty, it may be worthwhile for the library to find ways to support OA publishing endeavors. Article processing charges are often a barrier for faculty seeking to publish OA. While the Library can currently help offset those costs through a few transformative and read-and-publish agreements with vendors, it may be worthwhile to consider expanding those offerings, especially now that we have a better idea what publishers faculty are most likely to use.

Limitations of This Analysis

There are some limitations to conducting a citation analysis of this scope. The first being the limitation of the subscription data. The file used provided a snapshot of the moment in time when the data were extracted. Since this project considered citations from the past ten years, there is no way of knowing if the library was able to provide access at the moment in time when it was needed. It can only be said whether access can be provided now. It must be assumed that was also the case when the article was originally used. The same can be said of embargoed content. While the code could account for current embargoes, it would have proven too difficult to ascertain whether a particular article was under embargo at the time of citation. Additionally, while the library has recently begun to acquire journal backfiles to address holes in coverage, it is reasonable to assume that, in the case of older articles, there may have been a hole in coverage that is now resolved. In this case, given that most articles cited were published in the 1990s or later, this discrepancy is not too impactful for the University of Scranton. However, it may be a consideration at other institutions.

With regard to coverage, there was also some difficulty accounting for split or noncontinuous coverage dates. As noted in the Results and Discussion section, it did come up for one title in the no-access list where there is a small gap between current subscription and backfile access. For this reason, the journal incorrectly appeared in the no-access list as having 24 cited articles outside of the library's subscriptions when, in actuality, they were covered by the backfiles. Since this is a rare occurrence in our subscriptions, overall results were not adversely affected. Still, it may be a consideration for other collections with more coverage gaps.

Overall, this study can only consider works that were cited in faculty publications. It must be acknowledged that over the course of the research process, many more articles and materials may be considered and ultimately not used. However, current faculty citation patterns will serve as a good general representation of research patterns and works typically consulted.

Next Steps

Many of the citation analyses consulted also incorporate usage data. Using an R script, it would be easy to combine usage statistics with the citation list, but the Weinberg Memorial Library's current process for retrieving usage statistics is very manual. Since set up is underway for usage consolidation within the library's EBSCO Holdings Management, a more streamlined way for obtaining usage data may soon be available. Since Holdings Management also provides information on titles in aggregate databases, it may also be interesting to explore how often this content is used as opposed to content provided through vendor subscriptions.

This study also focuses primarily on statistics and numbers as related to faculty publications. A good next step in the assessment process may be to conduct focus groups with faculty to gain more qualitative data.

Finally, many studies consulted chose to focus on specific disciplines or departments. It may be interesting to conduct additional analyses by department or school to obtain some more granular insights.

Conclusion

Over the course of this project, 10 years of University of Scranton faculty publications were examined. Those publications were further examined to compose a list works that were cited by faculty in their publications, and that list was compared to current library subscriptions and holdings. Overall, the library was able to provide access to approximately 65% of the cited titles during the time period examined.

The use of faculty citation analysis in library collection development and assessment is well documented in the literature of library and information science. While this case study represents specific needs and use cases at the University of Scranton's Weinberg Memorial Library, the ultimate importance of this study is the process itself. The use of non-proprietary tools and data sources like OpenAlex and R create exciting new opportunities for others who wish to conduct similar studies at their own institutions without relying on proprietary tools and data sources or resorting to more labor-intensive methods. This increased accessibility may afford opportunities for new research and allow others to assess and analyze their collections in ways that may not have previously been possible.

Final Notes

This project made extensive use of the R programming language and RStudio both for querying the OpenAlex API and for conducting statistical analysis. The complete code and additional project notes can be found on GitHub: https://github.com/sylviaorner/citationanalysis.

References

Aria, M., & Le, T. (2023). openalexR: Getting bibliographic records from 'OpenAlex' database using 'DSL' API (R package version 1.2.1) [Computer software]. https://CRAN.R-project.org/package=openalexR

Currie, L., & Monroe-Gulick, A. (2013). What do our faculty use?: An interdisciplinary citation analysis study. *The Journal of Academic Librarianship*, *39(6),* 471–480. https://doi.org/10.1016/j.acalib.2013.08.016

Fernández-Ramos, A., Rodríguez-Bravo, B., & Diez-Diez, Á. (2023). Use of scientific journals in Spanish universities: Analysis of the relationship between citations and downloads in two university library consortia. *Scientometrics*, *128(4),* 2489–2505. https://doi.org/10.1007/s11192-023-04670-0

Feyereisen, P., & Spoiden, A. (2009). Can local citation analysis of master's and doctoral theses help decision-making about the management of the collection of periodicals?: A case study in psychology and education. *The Journal of Academic Librarianship*, *35(6),* 514–522. https://doi.org/10.1016/j.acalib.2009.08.018

Gao, W. (2016). Beyond journal impact and usage statistics: Using citation analysis for collection development. *The Serials Librarian*, *70(1-4),* 121–127. https://doi.org/10.1080/0361526X.2016.1144161

Hoffman, K., & Doucette, L. (2012). A review of citation analysis methodologies for collection management. *College & Research Libraries*, *73(4),* 321–335. https://doi.org/10.5860/crl-254

Kayongo, J., & Helm, C. (2009). Citation patterns of the faculty of the anthropology department at the University of Notre Dame. *Behavioral & Social Science Librarian*, *28(3),* 87–99. https://doi.org/10.1080/01639260903089040

Ke, I., & Bronicki, J. (2015). Using Scopus to study researchers' citing behavior for local collection decisions: A focus on psychology. *Journal of Library Administration*, *55(3),* 165–178. https://doi.org/10.1080/01930826.2015.1034035

Kelly, M., & O'Gara, G. (2018). Collections assessment: Developing sustainable programs and projects. *The Serials Librarian*, *74(1-4),* 19–29. https://doi.org/10.1080/0361526X.2018.1428453

Kohn, K. C., & Gordon, L. (2014). Citation analysis as a tool for collection development and instruction. *Collection Management, 39(4),* 275–296. https://doi.org/10.1080/01462679.2014.935904

Kumpulainen, M., & Seppänen, M. (2022). Combining Web of Science and Scopus datasets in citation-based literature study. *Scientometrics*, *127(10)*, 5613–5631. https://doi.org/10.1007/s11192-022-04475-7

Martindale, T. (2020). More than collection development: Using local citation analysis to begin a career in business librarianship. *Collection Management*, *45(4)*, 321–334. https://doi.org/10.1080/01462679.2020.1715315

Matos, M. A. (2016). Do we have it? A comparative analysis of library journal holdings and works referenced in faculty publications. *The Serials Librarian, 70(1-4)*, 260–265. https://doi.org/10.1080/0361526X.2016.1157738

OpenAlex. (n.d.). About. https://openalex.org/about

Pastva, J., Shank, J., Gutzman, K. E., Kaul, M., & Kubilius, R. K. (2018). Capturing and analyzing publication, citation, and usage data for contextual collection development. *The Serials Librarian*, *74(1-4)*, 102–110. https://doi.org/10.1080/0361526X.2018.1427996

Peñaflor, J., & Aliwalas, A. (2022). Research output and information use: A citation analysis of faculty publications in engineering. *Collection Management*, *47(4)*, 300–315. https://doi.org/10.1080/01462679.2022.2081830

Priem, J., Piwowar, H., & Orr, R. (2022). *OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts*. ArXiv. https://arxiv.org/abs/2205.01833

Smith, E. T. (2003). Assessing collection usefulness: An investigation of library ownership of the resources graduate students use. *College & Research Libraries*, *64(5)*, 344–355. https://doi.org/10.5860/crl.64.5.344

Sylvia, M. J. (1998). Citation analysis as an unobtrusive method for journal collection evaluation using psychology student research bibliographies. *Collection Building, 17(1)*, 20–28. https://doi.org/10.1108/01604959810368965

Trueswell, R.L. (1969). Some behavioral patterns of library users: The 80/20 rule. *Wilson Library Bulletin 43(5):* 458–461.

Tucker, C. (2013). Analyzing faculty citations for effective collection management decisions. *Library Collections, Acquisitions, and Technical Services, 37(1-2*), 19–33. https://doi.org/10.1016/j.lcats.2013.06.001

Vaaler, A. (2018). Sources of resources: A business school citation analysis study. *Journal of Business & Finance Librarianship, 23(2)*, 154–166. https://doi.org/10.1080/08963568.2018.1510252

Watson, A. P. (2010). Tips for conducting citation analysis in an academic setting. *Mississippi Libraries, 7 (1)*, 14–16.

White, P. B. (2019). Using data mining for citation analysis. *College & Research Libraries, 80(1),* 76–93. https://doi.org/10.5860/crl.80.1.76

Wilson, C. S., & Tenopir, C. (2008). Local citation analysis, publishing and reading patterns: Using multiple methods to evaluate faculty use of an academic library's research collection. *Journal of the American Society for Information Science and Technology, 59(9),* 1393–1408. https://doi.org/10.1002/asi.20812

Zhang, L. (2007). Citation analysis for collection development: A study of international relations journal literature. *Library Collections, Acquisitions, and Technical Services, 31(3-4),* 195–207. https://doi.org/10.1016/j.lcats.2007.11.001