This article was originally published in Issues in Science and Technology Librarianship. http://www.istl.org

Citation Analysis: A Method for Collection Development for a Rapidly Developing Field

Kristen B. LaBonte
Digital Resources and GIS Librarian
California State University, Channel Islands
Camarillo, CA
kristen.labonte@csuci.edu

Abstract

Citation analysis was used to determine if the Sciences-Engineering Library at the University of California at Santa Barbara is meeting the needs of an interdisciplinary group of 60 faculty members at the new California NanoSystems Institute. The latest three publications of each faculty member (published within the last two years) were analyzed in two ways using the Science Citation Index: 1) the journals they were published in, and 2) the journals where cited articles were published. The results indicate that the library subscribes to 98 percent of the journals in which faculty members are published or are citing frequently. This information is useful to map the citation patterns of a new interdisciplinary field and can be used for future collection management decisions.

Introduction

In December 2000, the Governor of California established four California Institutes for Science and Innovation, one of which is a multidisciplinary partnership between the University of California at Santa Barbara (UCSB), and the University of California at Los Angeles. Funding from the California State legislature and private industry established the California NanoSystems Institute (CNSI).

The mission of CNSI is "to create the collaborative, closely-integrated and strongly interactive environment that will foster innovation in nanosystems research and education." Nanosystems research concerns the manipulation of structures at the nanometer scale and how this manipulation can be used to build complex systems (California NanoSystems Institute) and it integrates the fields of electronics, microbiology, engineering and many other disciplines.

The institute at UCSB brings 60 faculty members together from nine different departments; Computer Science, Chemistry and Biochemistry, Physics, Chemical Engineering, Electrical and Computer Engineering, Music, Materials Science,

Mechanical and Environmental Engineering, and Molecular, Cellular, and Developmental Biology. Each of these departments has its own collection manager and fund codes, the CNSI does not.

The goal of this study is to use citation analysis to investigate whether the Sciences-Engineering Library at UCSB is meeting the needs of the recently established CNSI. Citation analysis measures how often items are cited in references, bibliographies, or indexing tools and compares their frequency of occurrence to collection holdings (Lockett 1989). This process assumes the items used by the authors occur as citations in their work. This assumption may fall short due to unacknowledged influences on the authors' work that are outside the scope of this project. This study is aimed at developing a core list of journals for and identifying journals that should be added to the collections in the Sciences-Engineering Library at UCSB.

Literature Review

In one of the early citation studies. Gross and Gross (1927), discovered that very few journals were cited frequently in the *Journal of the American Chemical Society*, while many journals were only cited once. Their findings concur with Zipf's Law, which states that while few items occur often, many items occur rarely (National Institute of Standards and Technology 2002). Libraries should be able to supply most of the needs of the users with a small amount of journals.

Citation analysis is now commonly used to determine what titles to purchase, to discontinue, or to weed (Smith 1981). As the costs of journal subscriptions escalated in the 1970s, the use of citation analysis was expanded to determine the ratio of serials versus monographs that should be purchased (White 1981). This use of citation analysis was employed by Kriz in 1978 and has been followed by more than 50 additional authors, covering a broad range of subjects from theology to geology (Devin & Kellogg 1990; Bowman 1990). Although this measure is easy to use, there are many factors to be considered as to why a journal is being cited, such as circulation and acceptance rate and other factors such as the group of researchers with whom the author associates (Buffardi and Nichols 1981).

Studies like Dombrowski's 1988 evaluation of journals in the fields of embryology, anatomy, and morphology show what journals rank best within these fields using ranked lists in *Journal Citation Reports*, but there is no additional information such as a citation or a circulation study to examine what is being used locally. Nanoscience is difficult to evaluate according to ranked lists as very few journals are dedicated to this interdisciplinary field, thus a local citation study is essential.

Local citation studies can also be of great value for interdisciplinary research. The study conducted by Hurd in 1992 found that in recent articles published by the chemistry department, 49% of citations were not from the field of chemistry. Findings like these can boost the argument to eliminate departmental libraries with narrow collection management scopes. The study by Delwiche (2003) mapped the literature of clinical

laboratory science by analyzing recent articles to identify the core journals of the field, the primary format of the literature, the currency of the literature, and index coverage of those titles.

Methods

The three most recent articles by CNSI faculty members published during 2002-2003 were collected and analyzed using information obtained from the *Science Citation Index* (SCI) (Thomson ISI 2003). The SCI has citations from over 5,300 science journals and adds around 17,000 new articles each week (University of California at Santa Barbara Libraries). One weakness of using SCI to obtain data for this analysis is that the SCI does not include all of the journals used by CNSI faculty. Not every faculty member had three publications in SCI, and 17 articles were not included in the study. Articles listed as review articles, editorials, or as abstracts were not included in the study, as it is difficult to determine if the work that is cited is necessary for the typical research performed at the institute. The short time span of the years studied reflects the length of time that the faculty members may have been publishing as affiliates of the CNSI, although they may have published nanosystems papers previously, and not all articles in this study are in the field of nanosystems. This study examines the most recent needs of the CNSI.

Journal titles, publication years and titles of journals that were cited were recorded. The titles and years of each journal that a faculty member published in, as well as each journal cited eleven or more times (which contain the top third of citations), were compared to library holdings in the UCSB catalog. Journals were grouped according to Bradford's 1948 Law of Scattering. By applying this law, journals are grouped into three zones. In Zone 1, a few journals produce the largest amount of citations, the second and larger group (Zone 2) has journals cited somewhat less frequently and Zone 3 contains a much larger group of journals cited relatively infrequently. Information regarding the type of citation was also collected in the groups of journal articles, conference proceedings, and other, which included personal correspondence, books, book chapters, and patents. Any citations where format could not be determined were verified using *Ulrich's International Periodicals Directory* (Bowker 2003). If the citation could not be verified, it was included in the "other" field. Articles that were written by more than one faculty member in the study were included once for analysis.

Journal titles were evaluated on the basis of the number of times they were cited. Journals that changed titles over time were included with the information for the most recent title. Journals that had split into multiple titles were analyzed using all title information. It is possible for a journal to have a lower ranking based on the original title not being included in one or both of the two new titles.

Results

Source Journals

163 source articles were published in 81 journals. 55 journals (67.9%) contained one article, and one journal (*Applied Physics Letters*) contained 19 articles. Two journals (*Quantum Information & Computation* and *Integrated Ferroelectrics*), each with one article, were not owned by UCSB. None of these titles, nor many of the cited titles, are dedicated to nanotechnology. Six titles occur in Zone 1 with six to 19 articles, 19 titles occur in Zone 2 with two to five articles, and 54 titles occur in Zone 3, all with one article. Table 1 shows all journals in Zones 1 and 2 published by faculty members in the California NanoSystems Institute at UCSB.

Table 1: Journals Most Frequently Published In	
Title	**Number of Articles**
ZONE 1	
Applied Physics Letters	19
Physical Review. B	11
Physical Review Letters	8
IEEE Photonics Technology Letters	7
Biophysical Journal	6
Journal of Chemical Physics	6
ZONE 2	
Journal of Applied Physics	5
Macromolecules	5
Journal of the American Chemical Society	4
Synthetic Metals	3
Advanced Materials	3
Chemical Communications	3
IEEE Transactions on Electron Devices	3
Langmuir	3
Physical Review E	2
Physics of Fluids	2
Proceedings of the National Academy of Sciences of the U.S.A.	2
Acta Metallurgica Sinica	2
IEEE Journal of Selected Topics in Quantum Electronics	2
IEEE Transactions on Very Large Scale Integration (VLSI)	2
Journal of Molecular Biology	2

Journal of Physical Chemistry B	2
Nano Letters	2
Nature	2
Physica E	2

Cited Journals

4,023 citations were analyzed from 643 journals, of which 318 journals (49.5%) occurred as one citation, and one journal (*Applied Physics Letters*) occurred as 267 citations (see Table 2). The mean number of citations per bibliography was 28.77 of which 90.2 percent were journals, 3.4 percent were conference proceedings, and 6.4 percent were other formats such as books, patents and personal communication. Zone 1 contains eight journal titles cited 88 to 267 times, Zone 2 contains 51 journal titles cited 11 to 83 times and Zone 3 contains 583 titles cited 10 or fewer times.

Table 2: Top 59 Cited Journals (top two-thirds of citations)	
Title	**Citations**
ZONE 1	
Applied Physics Letters	267
Physical Review Letters	261
Physical Review. B	209
Science	161
Nature	122
Journal of Chemical Physics	108
Journal of the American Chemical Society	100
Macromolecules	88
ZONE 2	
Journal of Applied Physics	83
Advanced Materials	75
Proceedings of the National Academy of Sciences of the United States of America	73
Langmuir	70
Journal of Physical Chemistry. B	57
Journal of Molecular Biology	55

Biophysical Journal	50
Biochemistry	49
Chemistry of Materials	49
Chemical Physics Letters	43
Physical Review. A	40
Angewandte Chemie - International Ed.	37
Surface Science	31
Electron Letters	27
Synthetic Metals	26
Journal of Biological Chemistry	25
Journal of Vacuum Science and Technology A	25
Journal of Physical Chemistry	24
Journal of Vacuum Science and Technology B	24
Journal of Fluid Mechanics	21
Journal of Crystal Growth	20
Limnology and Oceanography	20
Nature Structural Biology	20
Biopolymers	18
Experiments in Fluids	18
Acta Metallurgica et Materialia	16
IEEE Photonics Technology Letters	16
Inorganic Chemistry	16
Journal of Physical Chemistry A	16
Ultrasound in Medicine & Biology	16
Reviews of Modern Physics	15
Japanese Journal of Applied Physics Part 1	14
Proceedings of the Royal Society of London Series A	14
Chemical Communications	13
Journal of Colloid and Interface Science	13
Accounts of Chemical Research	12

Journal of Catalysis	12
Journal of Materials Chemistry	12
Physica Status Solidi A: Applied Research	12
Current Opinion in Colloid & Interface Science	11
IEEE Transactions on Information Theory	11
Journal of the Acoustical Society of America	11
Journal of Computational Physics	11
Journal of Physics: Condensed Matter	11
Japanese Journal of Applied Physics Part 2	11
Physics of Fluids	11
Physical Review	11
Physica B: Condensed Matter	11
Polymer	11
Proteins	11
Solid State Communications	11

All of the cited journals in Zone 1 had complete holdings at UCSB while two journals in Zone 2 had partial holdings (see Table 3).

Table 3: Zone 1 and 2 Cited Journals not owned by UCSB

Title	Citations
Current Opinion in Colloid and Interface Science^	11
Proteins#	11

^ Partial access - online only, only some represented years accessible
# Partial access - gap between when print was cancelled and when online access begins

Discussion

The faculty members of the California NanoSystems Institute rely heavily on journal articles for research, as over 90 percent of all references were from journals. This number is slightly higher than the commonly used number of 80 percent for the sciences in general (Bowman 1990; Devin & Kellogg 1990). If the library at UCSB decides to give the CNSI its own fund code, this figure can be used to determine the serials budget. Currently, 89.8 percent of the $2,111,910 Sciences-Engineering Library budget is spent on serials, so the 90 percent figure is compatible with the preexisting budget allocations.

Title dispersion is the relative "degree to which the useful literature of a given subject area is scattered through a number of different books and journals" (Stevens 1953). For example, it was found that title dispersion was low in the Gross and Gross (1927) study of chemistry citations, as 25 percent of the citations came from two journals. Stevens (1953) found that title dispersion was greater for technologies than for pure sciences, and that younger sciences had greater title dispersion than older, well established sciences because those sciences have their own discipline-specific titles. Edwards (1999) found the opposite result when analyzing title dispersion for polymer science, a young science at the time.

Nanoscience is a relatively young science and if Stevens' (1953) conclusions were applicable to nanoscience, this study should show great title dispersion. It was found that there is low title dispersion when 51 titles are required to cover 66 percent of the total citations and 318 journal titles were only cited once. This may be a result of the interdisciplinary foundation of nanosystems, which is based upon the hard sciences. Although it is an interdisciplinary field almost all of the citations come from physics, general science, chemistry and materials science. Table four shows the journal title dispersion in journal articles written by faculty members of the CNSI.

Table 4. Title dispersion in CNSI faculty articles	
Total number of journals cited	643
Number of journals needed to cover 33%	8
Number of journals needed to cover 66%	51

Although Stevens' (1953) theory of title dispersion of new sciences may not seem to hold true for nanoscience, it validates Kriz's (1977) observation that libraries can supply users with a majority of journal articles by subscribing to a small number of journals. Of the 59 cited titles in Zones 1 and 2, only two did not have complete holdings at UCSB and both of those titles are at other UC institutions and are readily available through interlibrary loan. This is also the case with the two source journals that UCSB did not have any holdings for.

Edwards (1999) and Lockett (1989) recommend that citation studies should be supplemented with shelving counts or interlibrary loan analysis. The UCSB libraries have substituted online subscriptions for many print subscriptions, so an interlibrary loan analysis would aid in the recommendation of what items to purchase if there are journals are being used for research that were not cited in the articles examined in this study.

Conclusions

Citation analysis is a practical tool to evaluate how a library is meeting the needs of local users. In this case, citation analysis has shown that the Sciences-Engineering Library at UC Santa Barbara is already adequately meeting the needs of the new California

NanoSystems Institute. The journals examined form a core of those used by faculty members. It is recommended that two titles should have expanded access. Librarians should note these titles and take advantage of any opportunities to purchase online back issues.

Citation analysis was particularly useful because of the interdisciplinary nature of the new institute and the heavy reliance on journals. The data obtained from journal articles composed the majority of the literature used for nanoscience research, and showed which journal titles are used the most. Information obtained about journals not owned can be used in collection management decisions in the future.

If the need arises to make cuts to serials budgets and collection managers are forced to cancel titles, this data can be used to find the least cited material. It should be compared with a citation analysis from the parent discipline of the material and the department that relates to it (unless it is a nanosystems journal). This method may also be used if the library needs money to purchase back issues of more heavily used journals.

The methods used and the information obtained in this study can be used as a reference to other research institutions with a newly established interdisciplinary unit composed of faculty from multiple departments. This study contributes to the extensive field of citation analysis but focuses on the most up to date information in order to evaluate an existing science collection and its relation to a new institute in a rapidly evolving field.

References

Bowker LLC. 2003. Bowker's ulrichsweb.com. [Online]. Available: http://www.ulrichsweb.com/UlrichsWeb/# [Accessed December 16, 2003].

Bowman, M. S. 1990. Format citation patterns and their implications for collection development in research libraries. *Collection Building* 11(1), 2-8.

Buffardi, L. C., & Nichols, J. A. 1981. Citation impact, acceptance rate, and APA journals. *American Psychologist* 36, 1453-1456.

Bradford, S. C. 1948. *Documentation.* London: Crosby, Lockwood.

California NanoSystems Institute. 2003. *About CNSI.* [Online]. Available: http://www.cnsi.ucsb.edu/about/about.html [Accessed: January 2, 2004].

Delwiche, F. A. 2003. Mapping the literature of clinical laboratory science. *Journal of the Medical Library Association* 91(3), 303-310.

Devin, R. B. & Kellogg, M. 1990. The serial/monograph ratio in research libraries: Budgeting in light of citation studies. *College & Research Libraries* 51, 46-54.

Dombrowski, T. 1988. Journal evaluation using Journal Citation Reports as a collection development tool. *Collection Management* 10(3/4), 175-180.

Edwards, S. 1999. Citation analysis as a collection development tool: A bibliometric study of polymer science theses and dissertations. *Serials Review* 25(1), 11-20.

Gross, P. L. K. & Gross, E. M. 1927. College libraries and chemical education. *Science* 66, 385-389.

Hurd, J. M. 1992. Interdisciplinary research in the sciences: Implications for library organization. *College & Research Libraries* 53, 283-297.

Kriz, H. M. 1977. Citation counting and the future of engineering libraries. *Engineering Education* 67, 707-710.

Lockett, B. (Ed.). 1989. *Guide to the evaluation of library collections.* Chicago: American Library Association.

National Institute of Standards and Technology. 2002. *Zipf's law*. [Online]. Available: http://www.nist.gov/dads/HTML/zipfslaw.html [Accessed: January 4, 2004].

Smith, L. 1981. Citation analysis. *Library Trends* 30, 83-106.

Stevens, R. E. 1953. Characteristics of subject literature. In *Publications Committee* (ed. by Association of College and Reference Libraries), ACRL Monograph #6, pp. 10-21. Chicago: Association of College and Research Libraries.

Thomson ISI. 2003. *ISI Web of Knowledge*. [Online]. Available: http://isi3.isiknowledge.com/portal.cgi/wos [Accessed: December 17, 2003].

University of California at Santa Barbara Libraries. 2003. *Databases*. [Online]. Available: http://www.library.ucsb.edu/databases/ [Accessed: January 2, 2004].

White, H. S. 1981. Strategies and alternatives in dealing with the serials management budget. In *Serials collection development: Choices and strategies* (ed by S. H. Lee), pp. 31-32. Ann Arbor, MI: Pierian Press.

A note about the author:

This paper was written to fulfill a requirement for a master's degree from Southern Connecticut State University. Most of it was written while the author was a library assistant at the University of California, Santa Barbara and was revised after she became a librarian at California State University, Channel Islands.