University of Dar es Salaam Library Journal
Vol 13, No 1 (2018), pp-54-68
ISSN: 0856-1818

Analysis of top cited papers in library and information science that have stood the test of time

Emmanuel F. Elia
University of Dar es Salaam
elia@udsm.ac.tz
mklmakala@yahoo.com

Alfred S. Sife
Moshi Co-operative University
assife@mocu.ac.tz
sifesas@gmail.com

Abstract

Analysis of 10 top cited papers in the field of library and information science was carried out to understand their main characteristics and features such as bibliographic details, authorship collaboration, author affiliation, citation counts and specialty. This list of classic papers was extracted from Google Scholar Metrics in June 2018. These most cited articles were published in 2006 and the list was released in June 2017. The findings indicate that these 10 papers originated from 10 different institutions in six countries mostly in North America and Europe. The number of citations ranged from 410 to 3051, the mean number of authors per article is 2.3 and the mean number of citations per year was 98. Nine out of the 10 articles reported citation analyses. This suggests that LIS research has shifted towards exploring scientific publications and their impact through citation analysis, bibliometrics and scientometrics.

Keywords: citation analysis, classic papers, library and information science, scientometrics, bibliometrics.

Introduction

Scholarly writing involves acknowledgement of earlier publications by citing the authors and years of their publications inside newer works (in-text citation) as well as providing bibliographic details in the lists of references. Citations enable readers to identify and retrieve the referred publications, support and substantiate authors' arguments, credit other authors' ideas, acknowledge intellectual indebtedness, and signal authors' awareness of ethical conduct in writing (Ticehurst and Veal, 2000; Webster and Watson, 2002). Citations also indicate that the advancement of knowledge is incremental and new knowledge is built on existing works - "authors standing on the shoulders of giants in furthering their works" (Sellitto, 2004). Furthermore, the quality of scholarly publications, which reflects their impact, is often measured by the number of times a publication is cited. According to Bauer *et al*. (2016), research papers with higher citations indicate an authority in the discourse and progress of a scientific discipline. It is generally agreed that the higher the number of citations, the higher the impact.

The Institute for Scientific Information (ISI) has been compiling bibliometric information of scientific articles since 1945. In the 1960s, journal impact factor was established for determining the extent to which articles published in journals are cited (Blessinger and Hrycaj, 2010). Advances in information technologies have led to creation of other bibliometric tools such as Scopus and Google Scholar that collect and compile citation data. Based on these bibliometric data, citation analyses are performed to assess scientific articles, journals, individual scholars, institutions and countries in determining rankings and other decisions. Citation rank lists are often used to identify works that have greater intellectual influence. Analysis of most cited

articles within a research field can reveal topics of current interest, research trends, best journals and new techniques.

The history of classic papers articles goes back to 1969 when Garfield compiled a list of the top 50 most cited articles published in 1967 (Garfield, 1971). He produced a similar list six years later for articles published between 1961 and 1972 (Garfield, 1977). He used the term "classics" to refer to highly cited articles. In the LIS field, several studies (e.g. Åström, 2007; Levitt and Thelwall, 2009; Blessinger and Hrycaj, 2010; Bauer *et al*. 2016) have analyzed the top cited journal articles. Data for these studies was retrieved from Web of Science and Scopus databases. Åström (2007) reported a stable structure of two distinct research fields - informetrics and information seeking and retrieval. Levitt and Thelwall (2009) reported that the highest cited article received 901 citations and that two thirds of the first authors of the top cited articles had an h-index of less than eight. There was a moderate correlation between citation ranking and the number of years between peak year and year of publication.

A study by Blessinger and Hrycaj (2010) established that two journals published close to 70 percent of the highly cited articles and 31 percent of the highly cited articles were written by more than one author. Aksnes (2003) correlated the number of citations to the number of authors in a publication - the higher the number of authors the higher the citation counts. Other scholars such as Baltussen and Kindler (2004) argue that review articles often have higher number of citations that original research articles. Among other findings, Bauer *et al*. (2016) found that authors at Harvard University and Leiden University contributed the largest number of papers. The study also found that co-authorships is based on common interests of a specific topic. The

three important topics were collection and exploitation of information in clinical practices, the use of internet in public communication and commerce, and scientometrics. Furthermore, several studies (e.g. Erfanmanesh *et al*., 2010; Reddy *et al*., 2016) have repeatedly ranked USA as the most prolific country in publishing metrics in many subjects including LIS research. In Africa, Nigeria and South Africa have been reported as prolific countries in LIS publications (Ocholla and Ocholla, 2007; Sitienei and Ocholla, 2010). Other studies have analysed LIS publications in specific fields such as library operations (Blessinger and Frasier, 2007), library users and technology (Blessinger and Hrycaj (2010) as well as citation analysis (Ivanovic and Ho, 2016).

As a multidisciplinary field, Library and Information Science (LIS) encompasses diverse subdisciplines that have developed over time. Such subdisciplines include library studies, information organization, information architecture, information behavior, information retrieval, scholarly communication, information literacy, bibliometrics and scientometrics. Developments in information technology have transformed LIS in numerous ways. The technology has transformed how information is processed, packaged, preserved and disseminated. Effectiveness, efficiency and quality of library and information services now depend greatly on the use and application of information technology. The technology has also affected routines, activities, practices and behaviours of LIS professionals.

In June 2017, Google Scholar released lists of "Classic Papers" under Google Scholar Metrics. These Classic Papers are highly cited papers in various areas of research (including LIS) that have stood the test of time. For each area, the top 10 most cited articles that were published 10 years earlier (published in 2006) were listed. The listed papers were journal articles, articles

deposited in repositories or conference articles. The papers are mainly those which describe original research, written in English and which received at least 20 citations. In July 2018, Google Scholar Metrics was updated, and a list of top journals was released instead of a new version of Classic Papers. The link to the 2017 edition of Classic Papers was removed from Google Scholar Metrics but it is still accessible online[1].

In the LIS cluster, it is not known why these 10 articles were the most cited. The main research question addressed in this study was therefore "what are the characteristics and features of these classic papers?". Specifically, this analysis was done to understand the key characteristics and features of LIS Classic Papers in terms of their bibliographic details, authorship collaboration, author affiliation, citation counts and specific subject areas. Knowing these characteristics and features is useful to LIS researchers who may use these papers as models and guides for their future research. Such papers can also be used in identifying and benchmarking excellent scientific research.

Methods

A list of top 10 LIS Classic Papers was extracted from Google Scholar Metrics in June 2018 through browsing broad subject categories and subcategories. The extracted bibliographic information of the articles included:

- Title of the article
- Title of the journal, volume, issue, page numbers

[1] https://scholar.google.com/citations?view_op=list_classic_articles&hl=en&by=2006

- Author information
- Year of publication.
- Citation counts received (as of May 2017).

Additionally, further information for each article was searched to determine affiliation of first author and specific subfield of the article.

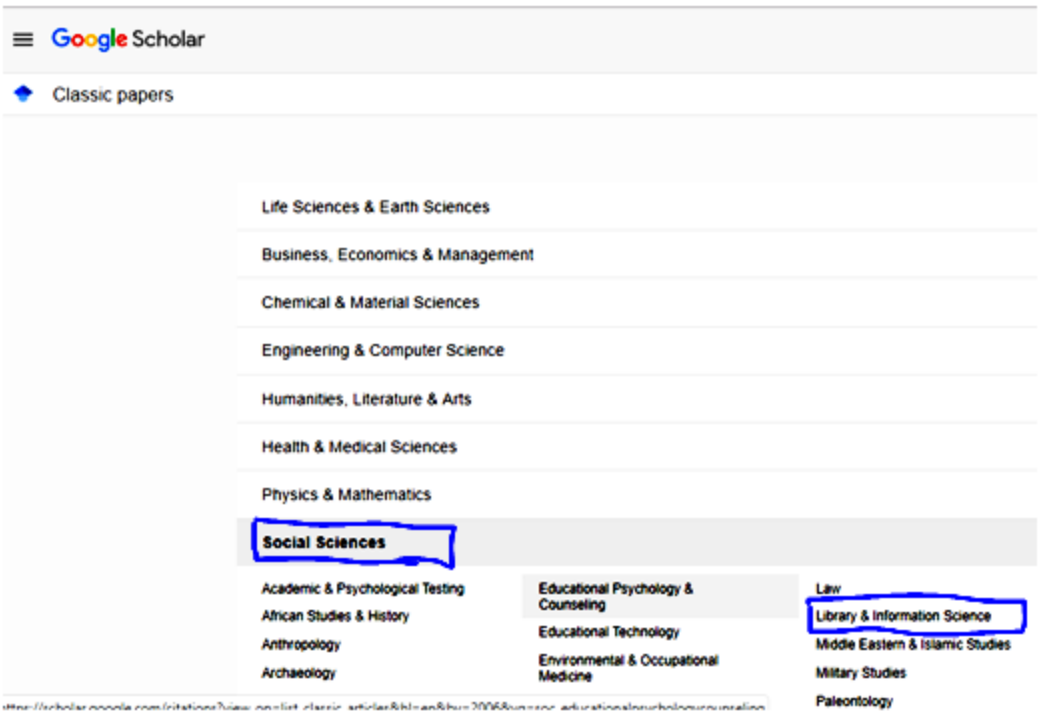

Figure 1: Google Scholar Classic Papers Subject Categories

Results and discussion

The bibliographic details of the top 10 articles in the field of library and information science are presented in Table 1 in a descending order based on their citations counts. Four papers were single-authored, six were multi-authored and the maximum number of authors in one article was 4. The mean number of authors per article was 2.3. Although Aksnes (2003) observed that publications with four authors on average tend to obtain twice as many citations compared to

those with one author, this was not the case in this study. Single-authored papers had higher citation counts compared to a number of multi-authored papers. Factors such as the quality of research, subject area and type of journal (Blessinger and Hrycaj 2010) are probable reasons for increased citations. Results also show a similarity of three prolific scholars (L. Egghe, W. Glänzel and A. F. Van Raan) who were also reported by Bauer *et al.* (2016). These are among the most proficient authors in the field of LIS as they have consistently maintained high citation counts.

The 10 papers originated from 10 different institutions from six countries. These institutions are the Cornell University, Drexel University, New York University, Los Alamos National Laboratory, University of Toronto, Université du Québec à Montréal, Universidade de São Paulo, LUC Universitaire, Hungarian Academy of Sciences and Leiden University. The countries are USA (n = 4), Canada (n = 2), Brazil (n = 1), Belgium (n = 1), Hungary (n = 1) and the Netherlands (n = 1). This means that most of these papers originated from the developed world particularly in North America and Europe. This is not surprising given the quantity and quality of research and publications published in these countries. A study conducted between 1996 and 2013 ranked USA as the first in research publishing metrics in the top 20 countries in all subject fields (Reddy et al., 2016). Similarly, Erfanmanesh *et al.* (2010) and Bauer et al. (2016) reported that USA LIS professionals contributed more articles than any other country. The current findings also show that Brazil, Belgium, Hungary and Netherlands each had one highly cited article. The results also show that although Nigeria and South Africa have been previously reported as prolific African countries in LIS publications (Ocholla and Ocholla, 2007; Sitienei and Ocholla, 2010), none of the 10 classic papers originated from these countries.

The number of citations ranged from 410 to 3051 and the mean number of citations per year was 98. These classic papers were cited 984 times on average during the period of 10 years. Three articles had over 1,000 citations which means they received over 100 citations each year. The most cited article was "Usage patterns of collaborative tagging systems" published by S. A. Golder and B. A. Huberman in the *Journal of Information Science*. This article was cited 1391 times more than the second paper on the list. The 10th article titled Journal status was published in *Scientometrics* journal.

University of Dar es Salaam Library Journal
Vol 13, No 1 (2018), pp-54-68
ISSN: 0856-1818

Table 1: Top 10 cited papers in library and information science: 2006 - 2017

Rank	Bibliographic details of articles	No of authors	Affiliation of first author	Citation	Citations per year	Specific subfields
1.	Golder, S. A., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2), 198-208.	2	Cornell University (USA)	3051	305	Information retrieval
2.	Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the Association for Information Science and Technology*, 57(3), 359-377.	1	Drexel University (USA)	1660	166	Domain visualization, Citation analysis
3.	Egghe, L. (2006). Theory and practice of the g-index. *Scientometrics*, 69(1), 131-152.	1	LUC Universitaire Campus (Belgium)	1473	147	Citation analysis: Bibliometrics/Scientometrics
4.	Eysenbach, G. (2006). Citation advantage of open access articles. *PLoS Biology*, 4(5).	1	University of Toronto (Canada)	659	66	Citation analysis
5	Van Raan, A. F. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, 67(3), 491-502.	1	Leiden University (the Netherlands)	608	61	Citation analysis: Bibliometrics/Scientometrics
6.	Braun, T., Glänzel, W., & Schubert, A. (2006). A Hirsch-type index for journals. *Scientometrics*, 69(1), 169-173.	3	Hungarian Academy of Sciences (Hungary)	564	56	Citation analysis: Bibliometrics/Scientometrics
7.	Zheng, R., Li, J., Chen, H., & Huang, Z. (2006). A framework for authorship identification of online messages: Writing-style features and classification techniques. *Journal of the Association for Information Science and Technology*, 57(3), 378-393.	4	New York University (USA)	502	50	Information retrieval, Citation analysis
8.	Batista, P. D., Campiteli, M. G., & Kinouchi, O. (2006). Is it possible to compare researchers with different scientific interests? *Scientometrics*, 68(1), 179-189.	4	Universidade de São Paulo (Brazil)	500	50	Citation analysis: Bibliometrics/Scientometrics
9.	Hajjem, C., Harnad, S., & Gingras, Y. (2006). Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. *IEEE Data Engineering Bulletin.* 28(4): 39-47	3	Université du Québec à Montréal (Canada)	411	41	Citation Analysis
10.	Bollen, J., Rodriquez, M. A., & Van de Sompel, H. (2006). Journal status. *Scientometrics*, 69(3), 669-687.	3	Los Alamos National Laboratory (USA)	410	41	Citation analysis: Bibliometrics/Scientometrics
Average				984	98	

Majority of these top cited papers were published in journals with high impact factors. Impact factor is one of the known measures of a journal's significance within the corresponding field (Garfield, 2006). Half (n = 5) of these articles were published in a journal known as *Scientometrics* with Impact Factor of 2.173. Two (n = 2) papers were published in the *Journal of the Association for Information Science and Technology* with Impact Factor of 2.835. The rest were published in the *Journal of Information Science* (n = 1) with Impact Factor of 1.372 and *PLoS Biology* with Impact factor of 9.163. The fact that half of the top cited articles came from the journal of *Scientometrics* is because of its reputation in the field of information science.

The most cited article is on information retrieval and it analyzed the structure of collaborative tagging systems and their dynamics (Golder and Huberman, 2006). The next most cited article by Chen (2006) is a citation analysis and domain visualization paper that describes development of a generic approach to detect and visualize emerging trends and transient patterns in scientific literature. Another scientometric article described g-index as an improvement of the h-index for measuring citation performance (Egghe, 2006). The fourth article by Eysenbach (2006) is a longitudinal bibliometric analysis comparing a cohort of open access and non-open access articles. Another scientometric article compared the h-index with several standard bibliometric indicators (Van Raan, 2006). Similarly, the sixth article described h-index as a useful supplement to journal impact factors (Braun, 2006). An article by Zheng et al. (2006) is a combination of information retrieval and citation analysis that developed a framework for authorship identification of online messages. The eighth article is again a scientometric paper that compared researchers with different scientific interests using h-index (Batista, 2006). Article by Hajjem et al. (2006) is a ten-year cross-disciplinary comparison of the growth of open access and how it

increases research citation impact. The last article demonstrates how PageRank algorithm obtains metrics that reflect journal prestige (Bollen et al., 2006).

Surprisingly, of the 10 articles, nine reported citation analysis mainly on bibliometrics and scientometrics. These findings suggest that LIS research has shifted to exploring scientific publications and their impact on various sectors through citation analysis, bibliometrics and scientometrics methods. A study by Bauer et al. (2016) also reported that the publication trend in LIS has changed and more literature is on performance assessment of scientific publications. The development of the field is moving from traditional library and information science to the use of technology and measuring research output in more technocratic ways (Blessinger and Hrycaj, 2010). These findings are in line with the general belief that review articles tend to have higher number of citations (Baltussen and Kindler, 2004).

These findings are contrary to several previous studies that reported that LIS research is conducted mainly on traditional library operations. Blessinger and Frasier (2007) reported that most articles published between 1994 and 2004 were published in *Library Trends* journal which focused more on practical matters of the profession including library operations including cataloging, references, information services as well as library user instructions and education. Blessinger and Hrycaj (2010) reported that the *Information Science* and *Journal of Documentation* were highly cited journals and dominant subjects were research in librarianship/users and technology. Another study by Ivanovic and Ho (2016) reported that *MIS Quarterly* and the *Journal of the American Society for Information Science and Technology* produced highly cited articles between 1956 and 2009.

Conclusion

This study presents 10 top cited papers in the field of library and information science extracted from Google Scholar Metrics, along with their bibliographic details and citation counts received between 2006 and 2017. This list of classic papers provides valuable information and it is useful in that it identifies notable papers that have contributed highly to the field. It indicates the topics that have profound influence and the new trends for future research in LIS. It suggests that LIS research has shifted to exploring scientific publications and their impact through citation analysis, bibliometrics and scientometrics. This list also provides institutions and authors that have contributed to these papers and have subsequently shaped research in LIS. The findings indicate that almost all 10 top papers were produced in the developed world. It is expected that by reading these papers, LIS scholars will learn as to what it entails to produce a classic paper.

References

Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12, 159−170.

Åström, F. (2006). Changes in the LIS research Front: time-sliced cocitation analyses of LIS journal articles, 1990–2004. *Journal of the American Society for Information Science and Technology* 58(7): 947–957.

Baltussen A., and Kindler, C.H. (2004). Citation classics in critical care medicine. *Intensive Care Med* 30: 902–910.

Batista, P. D., Campiteli, M. G., & Kinouchi, O. (2006). Is it possible to compare researchers with different scientific interests? *Scientometrics* 68(1): 179-189.

Bauer, J., Leydesdorff, L., and Bornmann, L. (2016). Highly cited papers in library and information science: authors, institutions and network structures. *Journal of the Association for Information Science and Technology* 67(2): 3095-3100.

Blessinger, K., and Frasier, M. (2007). Analysis of a decade in library literature: 1994-2004. *College and Research Libraries 68(2): 155-169.*

Blessinger, K., and Hrycaj, P. (2010). Highly cited articles in library and information science: an analysis of content and authorship trends. *Library and Information Science Research* 32:156-162.

Bollen, J., Rodriquez, M. A., and Van de Sompel, H. (2006). Journal tatus. *Scientometrics* 69(3): 669-687.

Braun, T., Glänzel, W., and Schubert, A. (2006). A Hirsch-type index for journals. *Scientometrics* 69(1): 169-173.

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the Association for Information Science and Technology* 57(3): 359-377.

Egghe, L. (2006). Theory and practice of the g-index. *Scientometrics* 69(1): 131-152.

Erfanmanesh, A., Didegah, F. and Omidvar, S. (2010). Research productivity and impact of library and information science in the web of science. *Malaysian Journal of Library and Information Science* 15(3): 85-95.

Eysenbach, G. (2006). Citation advantage of open access articles. *PLoS Biology:* 4(5).

Garfield, E. (2006). The history and meaning of the journal impact factor. Journal of the *American Medical Association* 295: 90−93.

Garfield, E. (1971). Citation indexing, historio-bibliography and the sociology of science. *Current Contents* 6: 156-157.

Garfield, E. (1974). Selecting the all-time citation classics: here are the fifty most cited papers for 1961-1972. *Current Contents* 2: 5-8.

Garfield, E. (1977). Introducing citation classics: the human side of scientific papers. *Current Contents*: 1: 5-7.

Golder, S. A., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science* 32(2): 198-208.

Hajjem, C., Harnad, S., and Gingras, Y. (2006). Ten-year cross-disciplinary comparison of the growth of open access and how it increases research citation impact. *IEEE Data Engineering Bulletin* 28(4): 39-47.

Ivanovic, D., and Ho, Y.S. (2016). Highly cited articles in the information science and library science category in social science citation index: a bibliometric analysis. *Journal of Librarianship and Information Science* 48(1): 36-46.

Koufogiannakis, D., Slater, L., and Crumley, E. (2004). A content analysis of librarianship research. *Journal of Information Science* 30(3): 227−239.

Levitt, J.M., and Thelwall, M. (2009). The most highly cited library and information science articles: interdisciplinary, first authors and citation patters. *Scientometrics* 78(1): 45-67.

Ocholla, D., and Ocholla, L. (2007). Research in Library and Information Science in South Africa: an analysis of journal research output from 1993-2006. *South African Journal of Libraries and Information Science*. 73(2): 109-118.

Reddy, K., Xie, E. and Tang, Q. (2016). Higher education, high-impact research, and world university rankings: a case of India and comparison with China. *Pacific Science Review B: Humanities and Social Sciences* 2: 1-21.

Sellitto, C. (2004). A study of missing Web-cites in scholarly articles: towards an evaluation framework. *Journal of Information Science* 30(6): 484-495.

Sitienei, G. and Ocholla, D. (2010). A comparison of the research and publication patterns and output of academic librarians in eastern and southern Africa from 1990-2006: a preliminary study. *South African Journal of Libraries and Information Science* 76(1): 36-48.

Ticehurst, G.W. and Veal, A.J. (2000). Business Research Methods, Frenchs Forest: Longman.

Van Raan, A. F. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics* 67(3): 491-502.

Webster, J. and Watson, R.T. (2002). Analyzing the past to prepare for the future: writing a literature review. *MIS Quarterly 26(2): 13-23.*

Zheng, R., Li, J., Chen, H., and Huang, Z. (2006). A framework for authorship identification of online messages: Writing- style features and classification techniques. *Journal of the Association for Information Science and Technology* 57(3): 378-393.