

The Accuracy of Newspaper Citation Count Reported and Actual Citation Found in Web of Science Citation Database

Ravikumar* and Sweety Warjri

Department of Library and Information Science, North Eastern Hill University, Umshing Mawkynroh, Shillong - 793022, Meghalaya, India; ravikumarnehu@gmail.com, sweetiewarjri03@gmail.com

Abstract

Bibliometric indicators are widely used in evaluating research output. Citation information about a publication can be obtained from various databases like Web of Science (WoS), SCOPUS and Google Scholar. Bibliometricians solely depend on the above databases to get citation data. WoS is considered a credible source for citation data. In this paper, we have tried to study the accuracy of the citation count reported and actually found with special reference to three newspapers namely Wall Street Journal, Washington Post, and Guardian. It was found that only 69% of citation data was accurate for above three newspapers.

Keywords: Bibliographic Database, Citation, Information Accuracy, Web of Science (WoS), SCOPUS

1. Introduction

Citation is the process of giving references to the source that is being cited. (Naan, 1976) defined as "... a citation is an acknowledgment that one document received from another". Citation data are available in various databases such as WoS, Scopus, Google Scholar, etc. These Citation databases have been made developed as a means of evaluating publications. WoS is the research database for citation tracking which indexes peer-reviewed journal literature. It is the oldest database that has a strong coverage with citation data and bibliographic data that goes back to 1900 (Boyle & Sherman, 2006). Scopus is also one of the databases that were introduced by Elsevier in 2004 providing coverage of articles from journals in science and social sciences, along with citation tracking. Google Scholar was introduced in November 2004 which is a tool that indexes the content of scholarly materials and allowing the researchers to locate a wide array of scholarly literature on the web including scholarly journals, abstract, peer reviewed articles, theses, books, dissertations, etc. Various citation studies have been done on different sources such as books, journals, thesis, dissertation etc. Newspapers are a primary source that

play an important role in the communication process publishing scientific research findings from day to day across the world and it has also become one of the sources that is found to be cited in the scholarly journals. In this study, only newspapers are taken to identify how newspaper citations are being reported in the WoS database and if there are any errors relating to these citations and also to find out the accuracy of citation data for the indexed newspaper in WoS.

2. Literature Review

Many bibliographic studies have been carried out using the three sources mentioned earlier. WoS has strong coverage compared to Scopus because its data goes back to 1900 (Joshi, 2016). A comparative study between the above three databases revealed that Google Scholar has higher coverage than the proprietary databases (Harzing & Alakangas, 2016).

A country-specific study Portuguese Universities found better coverage by SCOPUS than WoS to (Vieira & Gomes, 2009). In another comparative study of the three databases using the journal as a source, it was found that Google Scholar provided higher citation than WoS and

*Author for correspondence

Table 3. Variation in Number of Citations : Washington Post.

Washington Post			
Years	Web of Science Reported Citation	Web of Science Actual Found Citation	Percentage (%)
2001	1156	808	69.89%
2002	1453	980	67.44%
2003	1127	696	61.75%
2004	1834	1137	61.99%
2005	2239	1412	63.06%
2006	2054	1324	64.45%
2007	2010	1315	65.42%
2008	1862	1254	67.34%
2009	1860	1232	66.23%
2010	2190	1545	70.54%
Total	17785	11703	65.80%

*source WOS

Table 4. Variation in Number of Citations: Guardian.

Guardian			
Years	Web of Science Reported Citation	Web of Science Actual Found Citation	Percentage (%)
2001	846	584	69.03%
2002	885	602	68.02%
2003	1124	766	68.14%
2004	1194	798	66.83%
2005	1478	948	64.14%
2006	1684	1069	63.47%
2007	1759	1160	65.94%
2008	2289	1494	65.26%
2009	3362	2329	69.27%
2010	4362	3193	73.20%
Total	18983	12943	98.18%

*source WOS

WoS The Table 1–4 indicate that error occurred for all the years for all the three newspapers.

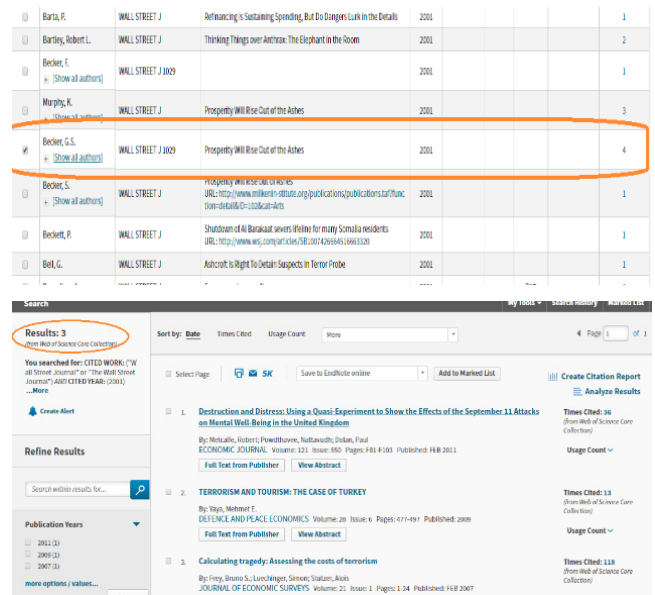


Figure 1. Screenshot of Wall Street Journal.

*source WOS

In the Figure 1 Wall Street Journal is shown to have received 4 citations from the WoS database but while downloading it reports only 3 citations.

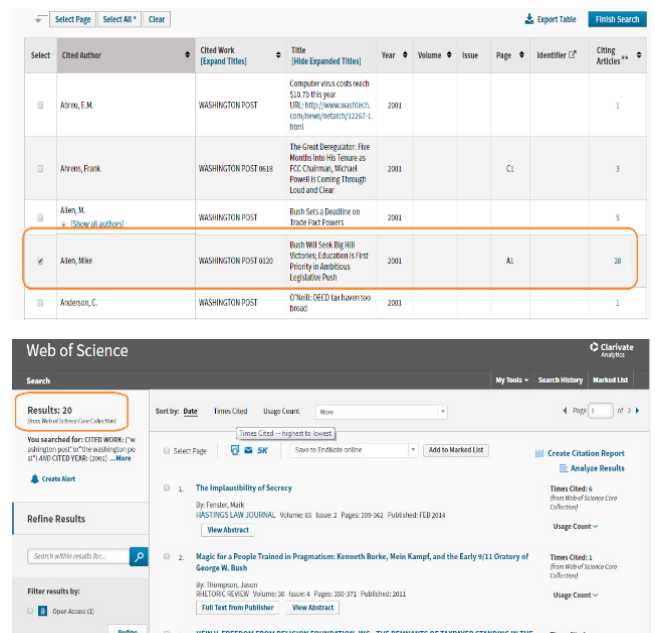


Figure 2. Screenshot of Washington Post.

*source WOS

Figure 2 indicates that Washington Post got 28 citations from the WoS database but while downloading it reports only 20 citations.

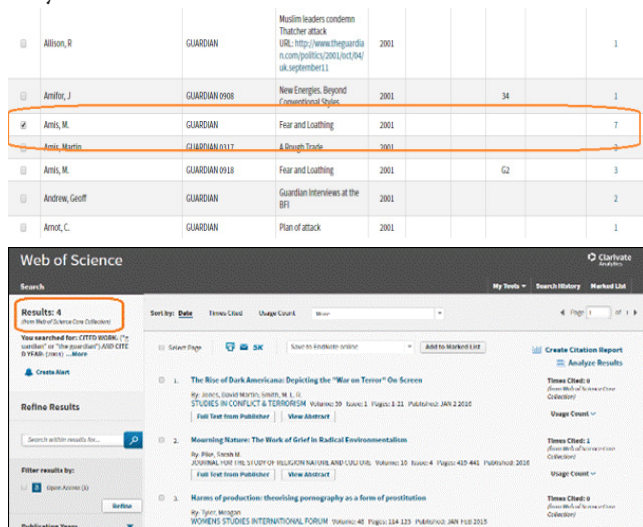


Figure 3. Screenshot of Guardian.

*source WOS

In the above Figure 3 Guardian article got 7 citations from the WoS database but while downloading it reports only 4 citations. In all the above figures, after downloading the actual data was found to be lower than the data reported by the WoS.

6. Analysis by Citation Counts

For Wall Street Journal, WoS shows 12837 over all citations but after downloading it was found to be only 9238 which show an accuracy rate of about 71.96%. Similarly for Washington Post, WoS shows reports 17785 citations but the actual result after downloading was found to be only 11703 which show about 6082 errors in citation count (error rate of 34.19%). Most of the errors related to reporting a citation frequency of 2, while actually there was only 1 citation. The same is true of Guardian (error rate of 31.81%).

7. Finding and Analysis

This study reveals that the data reported in the WoS database is not accurate. Of the three newspapers studied, the error rate was highest for *Washington Post*.

The examination also found that some of the author's name and article title was not given and was defined as 'Anonymous' in WoS.

8. Conclusion

From the study, it is found that the WoS database has many discrepancies in providing citation information. The citation count reported by the system and the actual number found did not match in many cases. Measures need to be initiated to increase the level of accuracy of data reported.

9. Reference

Adriaanse, L.S. and Rensleigh, C. (2013). WoS, Scopus and Google Scholar: A content comprehensiveness comparison. *The Electronic Library*, 31(6), 727–44. <https://doi.org/10.1108/EL-12-2011-0174>

Bakkalbasi, N. and Bauer, K. et.al. (2006). Three option for citation tracking: Google scholar, scopus and WoS. *Biomedical Medical Digital Library*, 3(7), 1–8.

Bauer, K. and Bakkalbasi, N. (2005). An examination of citation count in a new scholarly communication environment. *D-Lib Magazine*, 11(9), Retrieved from <http://www.dlib.org/dlib>, <https://doi.org/10.1045/september2005-bauer>.

Blanchard, M. (1998). *History of the Mass Media in the United States: An Encyclopedia*. London and Chicago: Fitzroy Dearborn Publishers.

Borstein, B.A. and Morillo, F. (2017). Funding acknowledgements in the WoS completeness and accuracy of collected data. *Scientometrics*, 1(12), 1793–812.

Boyle, F. and Sherman, D. (2006). Scopus: The product and its development. *The Serials Librarian*, 49(3), 147–53. https://doi.org/10.1300/J123v49n03_12.

Chandegani, A. and Salehi, H. (2013). A comparison between two main academic literature collections WoS and Scopus databases. *Asian Social Science*, 9(5), 18–26.

Drewry, J. M. (2007). *Google Scholar, Windows Live Academic Search and beyond: A study of new tools and changing habits in ARL libraries*, University of North Carolina at Chapel Hill; 2007. p. 63. Retrieved from <http://hdl.handle.net/1901/429>, <http://rightsstatements.org/vocab/InC/1.0/>.

Early history of the English newspaper. British Library Newspaper. Retrieved from http://find.galegroup.com/bncn/topicguide/bbcn_03.htm.

- Erfanmanesh, M.A and Didegah, F. (2013). A comparison of WoS and scopus for Iranian publications and citation impact. *International Journal of Information Science and Management*, 3(1), 11–27.
- Franceschini, F. and Maisano, D. (2015). Error in DOI indexing by bibliometric databases. *Scientometrics*, 10(2), 2181–6. <https://doi.org/10.1007/s11192-014-1503-4>.
- Franceschini, F., Maisano, D. and Mastrogiacomo, L. (2014). Scientific journal publishers and omitted citations in bibliometric databases: Any relationship? *Journal of Informetrics*, 8(3), 751–65. <https://doi.org/10.1016/j.joi.2014.07.003>.
- Harzing, A. and Alakangas, S. (2016). Google scholar, Scopus and WoS: A longitudinal and cross disciplinary comparison. *Scientometrics*, 10(6), 787–804. <https://doi.org/10.1007/s11192-015-1798-9>.
- History of Newspaper. Retrieved from <http://open.lib.umn.edu/medianandculture/chapter/4-2-history-of-newspaper/>
- History of printed newspaper. Retrieved from <http://www.peprint.com/resources/history-of-the-printed-newspaper/>.
- Joshi, A. (2016). Comparison between Scopus and ISI WoS. *Journal Global value*, 7(1), 1–11.
- Lambert, T.A. Brief history of newspaper. Retrieved from www.localhistories.org/media.html.
- Lewison, G. (2001). Evaluation of books as research outputs in history of medicine. *Research Evaluation*, 10(2), 89–95. <https://doi.org/10.3152/147154401781777051>.
- Lopresti, R. (2010). Citation accuracy in Environmental Science Journals. *Western Libraries and the Learning Commons at Webster CEDAR*, 85(2). <https://doi.org/10.1007/s11192-010-0293-6>.
- Meho, L. and Roger, Y. (2008). Citation counting, citation ranking, and h-index of human computer interaction researchers: A comparison between scopus and WoS. *Journal of the American Society for Information Science and Technology*, 59(11), 1711–26. <https://doi.org/10.1002/asi.20874>.
- Meho, L. and Kiduk, Y. (2007). Impact of data sources on citation counts and ranking of LIS faculty: WoS versus scopus and Google scholar. *Journal of the American Society for Information Science and Technology*, 58(13), 2105–25. <https://doi.org/10.1002/asi.20677>.
- Mikki, S. (2009). Google scholar compare to WoS: A literature review. *Nordic Journal of Information Literacy in Higher Education*, 1(1), 41–51. <https://doi.org/10.15845/noril.v1i1.10>.
- Ndonye, M. and Telewa, V. (2013). The future of print journalism: Internet technology and the 21th century newspaper in Kenya. *International Journal of Research in Social Sciences*, 3(3), 96–105. http://shodhganga.inflibnet.ac.in/bitstream/10603/18612/8/08_chapter%203.pdf.
- Stephens, M. (2008) History of Newspapers. For Collier's Encyclopedia. Retrieved from <http://www.nyu.edu/classes/stephens/Collier's%20page.html>
- Stohr, R. (2011). WoS citation data. *Communication Studies Theses, Dissertations, and Student Research*, 1(1). Data accessed: 28 May, 2018. Retrieved from <http://digitalcommons.unl.edu/commstuddiss/15>.
- Thomas, P.R. and Watkins, D.S. (1998) Institutional research rankings via bibliometric analysis and direct peer review: A comparative case study with policy implications. *Scientometrics*, 41(3), 335–55. <https://doi.org/10.1007/BF02459050>.
- Trapp, J. (2016). WoS, Scopus and Google Scholar citation rates: A case study of medical physics and biomedical engineering: What gets cited and what doesn't? *Australasian College of Physical Science and Engineers in Medicine*, 3(9), 817–23. <https://doi.org/10.1007/s13246-016-0478-2>. PMID: 27578318.
- Vieira, E.S. and Gomes, J. (2009). A comparison of Scopus and WoS for a typical university. *Scientometrics*, 81(2), 587–600. <https://doi.org/10.1007/s11192-009-2178-0>.