An Analysis of Bibliographic References Collected by a Social Computing Group

C. Sean Burns

School of Information Science & Learning Technologies

University of Missouri

csbc74@mail.mizzou.edu

May 15, 2010

Extended Abstract

This research presents an analysis of references collected by an established group of users who share a common interest on a social computing website. The purpose is to demonstrate how social computing functions as an instrument for measuring scholarly communication. The goal is twofold: to reveal how an analysis of the references collected by interested users compares to and complements citation studies.

Citation analysis is dependent on its data set and the traditional data set is provided by the ISI indexes. Wouters (1998) describes the Science Citation Index (SCI) as "the first citation index aimed at the scientific literature as a whole" whose aim is to create "an image of this literature like a telephone book creates one of the inhabitants of a city (p. 225). Lately, and for various reasons, there is interest in examining other indexes and databases, such as those provided by Scopus and Google Scholar (e.g., Yang & Meho, 2006; Falagas, et al., 2008; Harzing & Wal, 2008; Howland, et al., 2009).

The attempt to aggregate the literature as a whole is a shared characteristic among the above indexes and databases. Excluding for a moment the practical limitations involved with the

actual ability to capture the totality of scholarly communication, what is important is the inherent aim, or unspoken claim, that these services create an objective and quantitative database of that totality. Including the practical limitations, one can argue that the validity of some citation analysis research rests on the indexes' and databases' ability to actually capture that totality. One motivation for examining sources other than those provided by ISI is to determine what is not captured by one or the other (Meho, 2007).

If we think of citation analysis, in part at least, as the study of influence in quantitative aggregate, then if we are really interested in this notion of influence we must wonder if there are other methods for studying in quantitative aggregate what is influential. Cronin, Shaw, and La Barre (2003) pursued such research in their analysis of an article's paratext, its bylines and acknowledgements. This study, and others like it, demonstrates an interest in capturing in quantitative aggregate sources of influence, of collaboration, and of recognition not easily identified from a list of references alone.

This study examines a potentially different source of influence---the references one collects. My argument essentially states that what academics, scholars, and scientists collect is as significant as what they cite. Social computing websites such as CiteULike.org, a Web service that lets users collect, store, tag, and share references, offer us the ability to examine these collections.

This kind of research requires multiple studies, but initial results based on a collection of references created by a group of users on CiteULike.org provide enough confirmation of the arguments above to proceed. With regards to scholarly communication in general, the group's collection reveals a strong interest in e-print, open access articles. The collection also favors

articles that are neither highly cited nor originate from high impact journals, which leads to the conclusion that less visible articles have some influence. Although there are limitations with this initial study, the analysis provides strong enough evidence to conclude that these types of collections capture areas of influence that citedness is unable to address. Consequently, further research in this area will strengthen the validity of traditional citation analysis.

References

Cronin, B., Shaw, D., & La Barre. (2003). A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy. *Journal of the American Society for Information Science and Technology*, 54, 855-871. doi: 10.1002/asi.10278.

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *The FASEB Journal*, 22, 338-342. doi: 10.1096/fj.07-9492LSF.

Harzing, A. W. K., & van der Wall, R. (2008). Google Scholar as a new source for citation analysis. *Ethics in Science and Environmental Politics*, 8, 61-73. doi: 10.3354/esep00076.

Howland, J. L., Wright, T. C., Boughan, R. A., & Roberts, B. C. (2009). How scholarly is Google Scholar? A comparison to library databases. *College & Research Libraries*, 70, 227-334. doi:

Meho, L. I. (2007). The rise and rise of citation analysis. *Physics World*, 32-36.

Wouters, P. (1998). The signs of science. *Scientometrics*, 41, 225-241. doi:

10.1007/BF02457980.

Yang, K., & Meho, L. I. (2006). Citation analysis: A comparison of Google Scholar, Scopus, and Web of Science. *Proceedings of the American Society for Information Science and Technology*, 43, 1-15. doi: 10.1002/meet.14504301185.