Scientometrics (2017) 112:573–581
DOI 10.1007/s11192-017-2367-1

Do Mendeley reader counts reflect the scholarly impact of conference papers? An investigation of computer science and engineering

Kuku Joseph Aduku[1] · **Mike Thelwall**[1] · **Kayvan Kousha**[1]

Received: 8 January 2017 / Published online: 13 April 2017

Abstract Counts of Mendeley readers may give useful evidence about the impact of published research. Although previous studies have found significant positive correlations between counts of Mendeley readers and citation counts for journal articles, it is not known if this is equally true for conference papers. To fill this gap, Mendeley readership data and Scopus citation counts were extracted for both journal articles and conference papers published in 2011 in four fields for which conferences are important: Computer Science Applications; Computer Software; Building and Construction Engineering; and Industrial and Manufacturing Engineering. Mendeley readership counts correlated moderately with citation counts for both journal articles and conference papers in Computer Science Applications and Computer Software. The correlations were much lower between Mendeley readers and citation counts for conference papers than for journal articles in Building & Construction Engineering and Industrial and Manufacturing Engineering. Hence, there seem to be disciplinary differences in the usefulness of Mendeley readership counts as impact indicators for conference papers, even between fields for which conferences are important.

Keywords Mendeley readers · Citation counts · Journal articles · Conference papers · Correlations

✉ Kuku Joseph Aduku
K.J.Aduku@wlv.ac.uk

Mike Thelwall
M.Thelwall@wlv.ac.uk

Kayvan Kousha
K.Kousha@wlv.ac.uk

[1] Statistical Cybermetrics Research Group, School of Mathematics and Computer Science, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK

Introduction

Mendeley readership counts are promising indicators of scholarly impact (Gunn 2013; Haustein and Siebenlist 2011; Maflahi and Thelwall 2016) and appear much earlier than citations because they are as less affected by publication delays. For example, an article may be registered in Mendeley on the day that it is published. Mendeley readership counts can also reveal readers' disciplines and nationalities, giving more specific impact evidence (Thelwall and Sud 2015). Investigations of Mendeley readership counts so far have focused on either journal articles or books, but conference papers are valuable in some engineering-related fields and so it is important to assess whether they could also be applied to conference papers.

Although there are many limitations with using citation counts in formal and informal research evaluation as scholarly impact indicators, they are more robust than indicators derived from the web because these can easily be manipulated, making them unsafe for most formal evaluations (Wouters and Costas 2012). Several years ago, Mendeley had about 2.4 million users who uploaded over 420 million documents across disciplines ranging from life science to maths to the arts and humanities (Gunn 2013). Although Mendeley can be spammed, its large user base and positive results from previous analyses with it (see below) suggest that it does not currently suffer from a substantial amount of spam.

Papers presented at conferences in many fields are seen as a stage towards the creation of journal articles (Drott 1995). Nevertheless, in some research fields conference papers are valued for being timelier, more cutting-edge and more cited than journal articles (Goodrum et al. 2001) and can be either regarded as the main outputs of research or broadly comparable to journal articles as research outputs.

The gap that this research tries to fill is to discover whether the impact of conference papers is reflected in their Mendeley readership counts in the engineering-related fields in which they are important. Although several studies have found correlations between Mendeley readership counts and citation counts (Li et al. 2011; Bar-llan 2012; Mohammadi and Thelwall 2014), the extent to which Mendeley readership counts capture the impact of conference papers is unknown. The current study compares Mendeley readership counts and citation counts for both journal articles and conference papers in four engineering-related Scopus subject categories: Computer Science Applications; Computer Software; Building and Construction Engineering; and Industrial and Manufacturing Engineering. The paper also investigates why articles can be highly cited in Scopus but have few Mendeley readers and vice versa.

Altmetrics

The term altmetrics (alternative metrics) refers to academic indicators derived from social web data. Altmetrics rely on real-time data and interactions that can be quantified and measured immediately (Galloway et al. 2013). Existing altmetrics have used a variety of data sources including article downloads (Bollen et al. 2008) views and saves, as well as tweets, blogs, bookmarking sites and wikis. These are all used by scholars to communicate different kinds of research impact (Cronin 2013).

The main, but not only, way to assess altmetrics is through correlation tests (Sud and Thelwall 2014). One altmetric, tweet counts, might not be suitable for correlation tests, based on its increasing uptake resulting in newer articles having higher tweet counts than

the older articles (Thelwall et al. 2013). However, there is no evidence that this issue also applies to Mendeley readers.

The two major shortcomings of citation counts for assessing scholarly impact are that they are slow to accumulate and only reflect scholarly impact rather than applied impact. This has led to a need for new metrics to compliment traditional citation metrics (Priem et al. 2012). However, many scholars have argued that the new metrics should not be restricted to overcoming the limitations of the previous citation indicators, but can also be expected to provide new insights into research evaluation (Priem et al. 2010; Torres-Salinas et al. 2013).

Mendeley readership

Mendeley is an academic social web site for managing references, creating online profiles and sharing research documents with peers. It has an open Applications Programming Interface (API) that can be used for compiling usage indicators with a database of 2.6 million users as of October 2014 (Mohammadi et al. 2015a). Currently, Mendeley readership statistics seem to be the most closely related to citation counts, in comparison to other altmetrics. Many studies have used correlations to assess the relationship between Mendeley readership counts and citation counts for the same articles. A study of Nature and Science articles published in 2007 shows significant and moderate correlations between Mendeley readership counts and citation counts (Li et al. 2011). A study of five social science fields with 62,647 articles and five humanities fields with 14,640 articles found low to moderate significant positive correlations for each discipline (Mohammadi and Thelwall 2014). This study provides substantial evidence that Mendeley readership could be useful for assessing scholarly impact. Generally, most studies investigating the relationships between Mendeley readers and citation counts (Li and Thelwall 2011; Haustein et al. 2014; Costas et al. 2015) have reported either weak or moderate positive correlations between Mendeley readership counts and citation counts.

Mendeley has a higher proportion of articles with non-zero metric values than most other altmetrics (Zahedi et al. 2014). Out of 19,722 publications in this study, 62.6% had at least one reader. Mendeley is particularly used by undergraduates and postgraduates, whereas only academic authors can make citations. An analysis of the 'career stages' of the different Mendeley users found that Postdocs and PhD students register more in Mendeley than any other user category (Zahedi et al. 2013). A Mendeley survey found that out of 860 Mendeley users, 55% who had bookmarked articles in Mendeley had read them or intended to read them (Mohammadi et al. 2015b). However, not all readers record their articles in Mendeley, so the data does not represent all readers, but, most importantly, the survey shows that Mendeley bookmark counts are an indicator of readership.

Articles in Mendeley may be widely read but rarely cited in Scopus-indexed publications and vice versa for a number of reasons (Thelwall 2015): Authors from countries that do not publish in Scopus journals may receive more Mendeley readers than Scopus-indexed citations; short articles may support the process of research findings but may not be cited; users of Mendeley may register for an updated version of an article but the original version may be cited by others; some communities do not use Mendeley due to the nature of their professions; multidisciplinary articles may attract many citations but few readers based on multiple categorization norms; readers may prefer to read review articles but cite the reviewed articles rather than the review.

Research questions

The primary goal of this paper is to assess the value of Mendeley readership counts in conference-based fields and, as part of this, to find why articles in Mendeley may be widely read but rarely cited and vice versa in these fields. This study focused on four Scopus subject areas: Computer science applications; Computer Software; Building and Construction engineering; and Industrial and Manufacturing engineering. These represent different fields in which conference papers are important. The following research questions drive the study.

- Do Mendeley readership counts and citation counts reflect the scholarly impact of conference papers in Computer Science Applications, Computer Software, Building and Construction Engineering and Industrial and Manufacturing Engineering?
- Does the answer to the above research question differ between engineering fields in comparison to journal articles?
- What are the causes of conference papers having many readers compared to citations or vice versa?

Methods

Correlations between Mendeley readers and citation counts were calculated to ascertain the relationship between readership counts and citation counts for both conference and journal articles. A significant positive correlation gives evidence of a common factor between readership and citation counts.

All bibliographic information and citation data for journal articles and conference papers in the four fields from 2011 was extracted from Scopus. The year 2011 was chosen to give a substantial period for citations to accrue, so that there is more chance of getting high correlations between citation and readership counts for both journal articles and conference papers. The first and last 5000 journal articles and conference papers for each subject category were downloaded from Scopus in March 2015. This has the limitation that conferences in the middle of a year for large categories may be omitted, and these may be particularly prestigious in some fields. Narrow subject categories were used to ensure more comprehensive coverage. From the Scopus computer science category, a field in which conferences are arguably more important than journals, the two categories of Computer Science Applications and Computer Software were chosen. Conferences are also known to be important in engineering in general, and so from the broad Scopus Engineering category, the two narrow categories of Building and Construction and Industrial and Manufacturing Engineering were chosen.

Spearman correlations were used to compare citations and Mendeley readers because the data are skewed. Mendeley reader data was obtained using Webometric Analyst, a free software package. The Mendeley API in Webometric Analyst was used to extract data for Mendeley readers. The Spearman rank correlation formula was used to calculate 95% confidence intervals for the correlation coefficients. The formula used was $\tanh\left(\operatorname{arctanh}(r) \pm \frac{1.96}{\sqrt{n-3}}\right)$. Here, r is the sample correlation and n is the sample size. The sampling distribution of the estimate was approximately normal on the transformed scale; hence a 95% CI was found by taking the transformed estimate and adding and subtracting 1.96 times its standard error (Dowdy et al. 2011, pp. 245–246).

To determine outliers in order to find why articles in Mendeley are widely read but rarely cited in Scopus-indexed publications and vice versa, the logarithmic transformation $\ln(1 + x)$ was used on the data set (Thelwall and Wilson 2014) for both readers and citation counts to reduce the skewness of the data, before regressing the reader counts against the citation counts. The residuals from the linear regression were used to determine the main outliers. These were then manually investigated for likely causes. The purpose of the logarithmic transformation was to avoid focusing too much on papers with high values of one or other variable but to include outliers where the anomalies were within the low or moderate end of the scale. Thus, for example, a difference between 0 and 4 citations seems more significant than a difference between 40 and 44 citations, whereas the differences would be the same for untransformed data.

Results

The Mendeley readership counts correlate strongly (0.560–0.662) with citation counts in all subject categories for journal articles (Table 1). For conference papers, readership counts correlate moderately (0.437–0.439) with citation counts in Computer Science Applications and Software. Readership counts have low correlations (0.143–0.168) with citation counts in Building and Construction and Industrial and Manufacturing Engineering. The low correlations for conference papers in Building and Construction Engineering (0.143) and Industrial and Manufacturing Engineering (0.168) might be due to the low coverage of conference proceedings in engineering subject categories, reducing their Scopus citation counts.

There are low proportions of cited papers in Scopus for Industrial and Manufacturing Engineering (17.5%) and Building and Construction Engineering (18.3%) conference papers (Table 2). This could be due to low coverage in Scopus of conference proceedings. Also, there are low Mendeley reader counts for Building and Construction Engineering

Table 1 Spearman correlations between Mendeley reader counts and citation counts for articles and conference papers in Scopus from 2011 in the four subject categories analysed

Scopus subject category	Articles	Conference papers	Spearman correlation for articles and CI 95%	Spearman correlation for conferences and CI 95%
Computer science applications	10,000	9999	0.560** (0.546, 0.573)	0.439** (0.423, 0.455)
Computer software	10,000	9974	0.572** (0.559, 0.585)	0.437** (0.421, 0.453)
Building and construction engineering	8433	4750	0.662** (0.650, 0.674)	0.143** (0.115, 0.171)
Industrial and manufacturing engineering	10,000	9999	0.660** (0.649, 0.671)	0.168** (0.149, 0.187)

95% confidence intervals are reported underneath each correlation. Categories with 9999 or 10,000 articles are incomplete (the first and last 5000 articles/papers in the year) whereas the remaining categories are complete

** Statistically significant at p=0.01

Table 2 Scopus citation counts and Mendeley readership counts, median, geometric mean and percentage coverage for both journal articles and conference papers

Scopus subject category	Journal articles		Conference papers	
	Scopus citations median, geometric mean and % with citations	Mendeley readership median, geometric mean and % with readers	Scopus citations median, geometric mean and % with citations	Mendeley readership median, geometric mean and % with readers
Computer science applications	3, 0.91 (80.9%)	3, 0.93 (64.8%)	0, 1.20 (34.4%)	0, 0.41 (47.2%)
Computer software	3, 0.91 (80.6%)	3, 0.88 (62.1%)	3, 1.43 (54.7%)	10, 1.02 68.6%)
Building and construction engineering	2, 0.71 (71.7%)	2, 0.65 (52.7%)	0, 0.08 (18.3%)	0, 0.09 (18.7%)
Industrial and manufacturing engineering	2, 0.73 (71.3%)	2, 0.64 (55.9%)	0, 0.08 (17.5%)	0, 0.28 (41.0%)

(18.7%) conference papers, which could be due to few Mendeley users or the low value of conference papers in this field.

The following were identified as likely causes of high Mendeley readership counts compared to Scopus citation counts for conference papers.

- Papers that are written based on improving the performance of an existing system. Computer Science Applications; *''Purlieus: Locality-aware resource allocation for Mapreduce in a cloud*'', has 74 Mendeley readers but no citations. *''Reducing electricity cost through Virtual Machine placement in high performance computing clouds*'', (66 readers, 0 citations) demonstrates a system that can be used to reduce electricity cost and load migration at minimum low electricity consumption rates. '' *On the duality of data-Intensive file system design: Reconciling HDFS and PVFS''* (62 readers, 0 citations).
- Papers that create public awareness, motivation and participation for new scientific discoveries. For example, three papers from Computer Software; *''Attention please! Learning analytics for Visualization and recommendation''*, (202 readers, 1 citation) *''Dusting for Science: Motivation and participation of digital citizen science volunteer*s'', (149 readers, 6 citations) and *''A survey of risks, threats and vulnerabilities in cloud computing''*, (115 readers, 0 citations).
- Papers that are relevant to daily life or religious beliefs. For example, paper in Industrial and Manufacturing Engineering, *''Halal supply chains in the food industry-A conceptual model'',* (50 readers, 6 citations).
- Practical solutions to important real-world problems. For example, Building and Construction Engineering, *''Overview of UFC 3-340-02, Structures to resist the effects of accidental explosions''*, has 29 Mendeley readers but no citation. *''Sandnet: Network traffic analysis of malicious software''* (36 readers, 10 citations) In the same Building and Construction Engineering Scopus subject category for conference paper, *''Exploiting home automation protocols for load monitoring in smart buildings''* (32 Mendeley readers 1 citation).

- Social media articles that may be of general interest to users. For example, in Computer Science Applications article papers; *''Serious social media: On the use of social media for improving student's adjustment to college''* (170 readers, 15 citations), and *''Personal Learning Environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning''* (404 readers, 74 citations).
- Practical commercial advice. In Building and Construction Engineering, an article titled; *''Characterizing entry mode for international construction markets: paving way to a selection model''* (16 readers, 0 citations).
- Articles of regional interest may attract more readers than citations. For example, Computer Software, *''A citizen-oriented approach for evaluating the performance of e-government in Sri Lanka*'' (28 readers, 0 citations).

The following were identified as likely causes of high Scopus citation counts compared to Mendeley readership counts for conference papers.

- Papers on software packages that may be cited if the software is used, without necessarily reading the paper. In Computer Software, *''MICE: Multivariate Imputation by Chained Equation in R''* (249 citations, 0 readers) and "*ContextFJ: A minimal core calculus for context-oriented programming*" (18 citations, 0 readers).
- Papers with a set model for completing a task. For example, in Computer Science Applications, *''Recommended steps for thematic synthesis in software engineering''* (20 citations, 0 readers). *"A framework for capturing distinguishing user behaviours in novel interfaces"* (13 citations, 0 readers) and Building and Construction Engineering *"A naming convention for the piano key weirs geometrical parameters"* (20 citations, 0 readers).

Discussion

This study has several limitations. The citation data is from Scopus and is dependent on the coverage of Scopus for the magnitude of the citation counts. Similarly, Mendeley is not the only reference manager and the results may have been different for another reference manager if it is more widely used by engineers. The results may also vary by year and could be different for other types of engineering that have not been investigated. The causes of higher Mendeley readership counts than citations or vice versa are based on the small sample of papers analysed in the current paper and may not apply to other years or fields. Another limitation is that in some fields a few conferences may be very important whereas others are not and so the scope of the current study (analysing all Scopus-indexed conference papers) may hide important differences between conference types.

Mendeley readership counts and Scopus citation counts have strong and significant positive correlations for journal articles in all of the engineering fields analysed and for conference papers in the two computing fields but not in the other two engineering fields, Industrial and Manufacturing Engineering and Building and Construction Engineering, which have weak but positive correlations.

For journal articles, the strong and positive correlations between Mendeley readership and citation counts for all four of the studied Engineering subject categories corroborate past studies of other areas (Li et al. 2011; Bar-Ilan 2012; Mohammadi and Thelwall 2014).

For conference papers, 68.6% of the papers in the Computer Software subject category have at least one Mendeley reader and 54.7% of the papers have at least one Scopus citation. These findings show that the impact of conferences is high in Scopus and Mendeley for computing research. In Building and Construction Engineering, conference papers have a much lower percentage coverage; 18.3 and 18.7% of the papers have at least one Scopus citation and at least one Mendeley reader, respectively. This may be due to low coverage of conference proceedings in the field of engineering for Scopus but this cannot explain the results for Mendeley. It may be that a high percentage of engineering conference papers are not of interest to publishing academics, either because of their applied focus or due to disciplinary norms in citation practices.

The list of reasons why papers may attract many readers compared to their citations, or vice versa, shows that there are legitimate causes of outliers. It is therefore important to accept that Mendeley reader counts will not always be a good approximation to Scopus citation counts for individual papers.

Conclusions

Based upon high and positive correlations in the subject categories of Computer Science Applications and Computer Software, Mendeley readership counts for conference papers in computer science should be acceptable as scholarly impact indicators. Since Mendeley readership counts are particularly useful for early impact evidence (Thelwall and Sud 2016), this may be their greatest value in computing. In contrast, the weak correlations between Mendeley readership and citation counts in the subject categories of Building and Construction Engineering and Industrial and Manufacturing Engineering, coupled with low proportions of papers with at least one reader, suggest that conference papers in these types of engineering do not support a similar claim. This may be due to the low scholarly impact of conference papers in these fields and it may be that their value is not primarily within academia. If this is the case, then new indicators, perhaps including download counts, would be needed to reflect this impact.

Acknowledgement This article is an extended version of a paper by Aduku et al. (2016) presented at the Science Technology Indicators (STI) conference, held in Valencia (Spain) on 14–16 September 2016. The authors would like to thank the conference reviewers and participants for comments on the earlier drafts.

References

Aduku, K. J., Thelwall, M., Kousha, K. (2016). Do Mendeley reader counts reflect the scholarly impact of conference papers? An investigation of computer science and engineering fields. In I. Rafols, J. Molas-Gallart, E. Castro-Martínez, R. Woolley (Eds.), *Proceedings of the 21th International conference on science and technology indicators—Peripheries, frontiers and beyond*, Valencia (Spain).

Bar-Ilan, J. (2012). JASIST 2001–2010. *Bulletin of the American Society for Information Science and Technology, 38*(6), 24–28.

Bollen, J., Van de Sompel, H., & Rodriguez, M. A. (2008). Towards usage-based impact metrics: First results from the mesur project. Paper presented at the *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries,* 231–240.

Costas, R., Zahedi, Z., & Wouters, P. (2015). Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology, 66*(10), 2003–2019.

Cronin, B. (2013). Metrics à la mode. *Journal of the American Society for Information Science and Technology, 64*(6), 1091.
Dowdy, S., Wearden, S., & Chilko, D. (2011). *Statistics for research*. New York: Wiley.
Drott, M. C. (1995). Reexamining the role of conference papers in scholarly communication. *Journal of the American Society for Information Science, 46*(4), 299.
Galloway, L. M., Pease, J. L., & Rauh, A. E. (2013). Introduction to altmetrics for science, technology, engineering, and mathematics (STEM) librarians. *Science and Technology Libraries, 32*(4), 335–345.
Goodrum, A. A., McCain, K. W., Lawrence, S., & Giles, C. L. (2001). Scholarly publishing in the internet age: A citation analysis of computer science literature. *Information Processing and Management, 37*(5), 661–675.
Gunn, W. (2013). Social signals reflect academic impact: What it means when a scholar adds a paper to mendeley. *Information Standards Quarterly, 25*(2), 33–39.
Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics, 101*(2), 1145–1163.
Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of informetrics, 5*(3), 446–457.
Li, X., & Thelwall, M. (2011). F1000, mendeley and traditional bibliometric indicators. Paper presented at the *Proceedings of the 17th International Conference on Science and Technology Indicators, 2*451–2551.
Li, X., Thelwall, M., & Giustini, D. (2011). Validating online reference managers for scholarly impact measurement. *Scientometrics, 91*(2), 461–471.
Maflahi, N., & Thelwall, M. (2016). When are readership counts as useful as citation counts? Scopus versus Mendeley for LIS journals. *Journal of the Association for information Science and Technology, 67*(1), 191–199.
Mohammadi, E., & Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology, 65*(8), 1627–1638.
Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (2015a). Who reads research articles? an altmetrics analysis of mendeley user categories. *Journal of the Association for Information Science and Technology, 66*(9), 1832–1846.
Mohammadi, E., Thelwall, M., & Kousha, K. (2015b). Can mendeley bookmarks reflect readership? A survey of user motivations. *Journal of the Association for Information Science and Technology*. doi:10.1002/asi.23477.
Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *ArXiv Preprint* arXiv:1203.4745.
Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). *Altmetrics: A manifesto*. Retrieved Jan 13, 2011.
Sud, P., & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics, 98*(2), 1131–1143. doi:10.1007/s11192-013-1117-2.
Thelwall, M. (2015). Why do papers have many mendeley readers but few Scopus-indexed citations and vice versa? *Journal of Librarianship and Information Science*. doi:10.1177/0961000615594867.
Thelwall, M., & Wilson, P. (2014). Regression for citation data: An evaluation of different methods. *Journal of Informetrics, 8*(4), 963–971.
Thelwall, M., & Sud, P. (2015). Mendeley readership counts: An investigation of temporal and disciplinary differences. *Journal of the Association for Information Science and Technology, 57*(6), 3036–3050.
Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? twitter and ten other social web services. *PLoS ONE, 8*(5), e64841.
Torres-Salinas, D., Cabezas-Clavijo, Á., & Jiménez-Contreras, E. (2013). Altmetrics: New indicators for scientific communication in web 2.0. *ArXiv Preprint* arXiv:1306.659.
Wouters, P., & Costas, R. (2012). *Users, narcissism and control: Tracking the impact of scholarly publications in the 21st century*. Utrecht: SURF foundation.
Zahedi, Z., Costas, R., & Wouters, P. (2013). *What is the impact of the publications read by the different mendeley users? Could they help to identify alternative types of impact?*. San Francisco: Plos alm workshop.
Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics, 101*(2), 1491–1513.