Fall 9-14-2019

Studying Relationship between Citation and Altmetrics of Top Chemistry Researches' Articles

Mozhdeh Salajeghe
Shahid Bahonar University of Kerman, Kerman, Iran, msala@uk.ac.ir

Javad Zaraat Pisheh
Shahid Bahonar University of Kerman, Kerman, Iran

Adel Soleimani Nejad
Shahid Bahonar University of Kerman, Kerman, Iran

Studying Relationship between Citation and Altmetrics of Top Chemistry Researches' Articles

Abstract:

The main objective of the present research is to examine the relationship between the number of citations and the level of altmetrics for testing the validity of these new metrics, at least in terms of being alignment with the test established index. The present research population consist of articles from the top chemistry writers that were profiled at the Scopus Citation Database in 2010. Sample research is the articles by 20 top author. The present research is applied in terms of purpose, and is descriptive and correlative in terms of data collection. Data extraction was performed using Webometric analyst software and citation data was collected from Scopus. SPSS software was used to analyze the data.

The research findings show that the articles in question have little presence on social networks. In terms of the amount of attendance and distribution Mendeley, CiteUlike, Twitter, Facebook, Blogs, Google Plus and News, had the largest number of articles and altmetrics respectively. Also, the results show that Mendeley and Twitter have the most relationship with citations. Also, articles have at least one higher citation average altmetric (25.14%) than those with no altmetric (7.58%). In terms of citations' relationship, the Spearman correlation test showed a strong correlation between the number of Mendeley readers, news, and citations. Also, there was a weak correlation between Twitter, CiteUlike, and citations. Finally, there was not a meaningful relationship between Facebook posts, blog posts, Google plus, and citations.

Keyword

Mendeley, Papers Level Metrics, Scientometrics, Citation, Chemistry Articles, Top Authors, Altmetrics Tools, Altmetrics Source

Introduction

The massive amount of information production after World War has made it essential to study and evaluate the texts contains information, in order to distinguish high-quality information among them (Rismanbaf and Osareh, 1386). Although the best approach in assessing is studying all publications, and presenting experts' opinion (peer review), But time and budget constraints have always been a barrier.

As a result, over time, attempts have been made to provide quantitative measures for evaluations. In 1955, Eugene Garfield, the founder of the Scientific Information Institute, introduced the Impact Index 3 to the scientific community (West, 2008). With the advent of the social web and its impact on the communication of individuals in different societies and the creation of new Research communication among researchers, the inability of traditional metrics to measure these new communications became more visible. Therefore, the altmetrics or supplementary metrics were proposed as a level of papers' metrics (Salajegheh and Mohammadian, 1394.) Altmetrics has been defined to mentioning the scientific works in Social media such as Facebook.

Wikipedia.com, Twitter.com, Weblogs, Reference management Tools like Mendeley.com, Marking sites as CiteUlike.org, Delicious.com, News Media, and etc. These metrics include the number of clicks, Notes, stores, Tweets, likes, Shares, Recommend, Tags, posts, Discussions, and Comments. Number of downloads and views as well as the number of articles cited at citation databases such as Scopus.com Web, Web of Science and publishers like Public Library of Science (PLOS) are provided articles' level metrics by some tools that are not altmetrics.

In this research, the Altmetric.com service is used to extract the data from the altmetrics. To get the number of citations, the Scopus citation profile database is also used. The present study is going to examine the correlation between the number of citations and the number of altmetrics of articles in this field. In terms of discussion on scientific blogs, Facebook, Twitter, Reddit, Pinterest, newspapers and news sites, Comments, Google Plus, Questions and Answers, their storage rates in Mendeley, CiteUlike and Connotea, as well as the number of recommendations received from the F1000 site.

The current research is important because the subject of the research is new in scientific research, in addition, it has used altmetrics which is a new method for measuring scientific production, and study them from other dimensions such as the amount of discussion, storage, visit and received recommendations on social networks and bookmarking tools and citation management. There has not been conducted many research on this level of articles and with this range of metrics in the country. In terms of application, can use the results of this research at the research departments of universities also by chemistry researchers. The main goal of this study is studying the relationship between the number of citations and altmetrics of the top 2010 chemistry writers' articles.

Research hypothesis

1. In Mendeley service, there is a relationship between the number of citations and the amount of saving articles written by top chemistry writers in 2010.

2. In the CiteUlike, there is a relationship between the number of citations and the tags of articles written by top chemistry writers in 2010.

3. In the scientific blog, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010.

4. In news site, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010

5. In Twitter, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010

6. In Facebook, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010

7. In Google plus, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010

Previous Research

The study of the history of research carried out both inside and outside the country shows that studies which are done in the country is limited due to the new issue of altmetrics. Dayari, in her thesis, found that there is a relationship between citation indices other than IF with mean altmetrics score (Dayari, 1394). Zahedi et al research (2014) showed a positive but weak correlation (r = 0.179) between the number of citations and the storing of papers in Mendeley. Sotudeh et al (2015) found that there is a statistically significant, positive and weak relation between the number of citations and the tags of the articles. Li et al (2012) found that there is a moderate correlation between readers of Mendeley with correlation (r = 0.55) and CiteUlike (r = 0.34), and citations from articles received from Web of sciences. Bar-Ilan et al (2012) discovered that there is a significant correlation (r = 0.45) between the number of citation of Mendeley and readers and the

weak correlation (r = 0.23) with the number of CiteUlike's tags. The results obtained by Priem, Piwowar, and Hemminger (2012) have shown a moderate correlation (r =0.34-0.65) between the number of PDF and Html downloads; also it has been observed a moderate correlation between readers of Mendeley and tags of CiteUlike and between recommendations received from F1000 with the number of web of sciences citations. Other research results showed that there was a significant correlation between the number of altmetrics and citations for articles with positive altmetric score. The results provided strong evidence that only six metrics (Twitter, Facebook, F1000 blogs, News media, and associations) from eleven alternative metrics have a significant correlation with the number of citations. Due to the lack of sufficient data, there was no correlation between citations and altmetrics of LinkedIn, Pinterest, and Question and Answer websites. However, the coverage of all alternative metrics, other than Twitter, was reported below 20%. They considered the use of social web, especially Twitter, as a way of evaluating effective scientific resources (Thelwall, et al., 2013). Haustein et al., (2013) found that there was a weak correlation between the number of citations and tweets of articles. Costas et al (2014) found that there is a relatively weak correlation between the number of citations and recommendations received. Costas et al. (2014) discovered that there is positive but weak correlation between the numbers of citations with altmetrics. They reported a stronger relationship between the number of blog posts and the mention of works on news websites than other types of altmetrics. In another study, the findings showed that the very weak correlation (r = 0.003) between the citations and the volume of discussion on Twitter and blog posts (r = 0.12) and citeUlike tags (r = 0.23), but The moderate correlation between readers of Mendeley (r = 0.51) and the number of citations reported Bornmann (2015). Alhoori and Furuta (2015), reported weakly correlation between altmetrics of articles level, except numbers, such as Mendeley, with the number of citations. Shrivastava and Mahajan (2016) in their research found that there was a positive correlation between the number of readers and citations, and their correlation in 2010 was reported higher than in 2005.

Research Methodology

Because this research is for better understanding the existing conditions, it is among descriptive research with a kind of correlation and a Scientometric approach. The tops chemistry writers are the present research population that was profiled in Scopus database in 2010. Chemistry papers were selected due to they have been cited most and they reflected more rapidly in social media. Since the citation is a time-consuming process, based on the field, it may take 2-5 years to cite a paper (Thelwall, et al., 2013 quoted in Mansourian and Navidi ,1393). Chemistry published articles were selected in 2010 to provide enough time to attract citations by the articles and the chance to achieve a high correlation between the number of citations and altmetrics. The purposeful research sample has been chosen from the top 20 top writers in the field of Chemistry, which was profiled at the Scopus Citation Database in 2010, and has an article's digital indicator. First, to search articles in the field of chemistry at the Scopus database, the letter "a" was searched for all Fields field of documents and the entire production of the Scopus database (50817119) was retrieved. In the next step, these products are restricted through the field of the year to 2010, through the field of the subject area to the chemistry, and through the field of document type to article. In total (165,905) documents were obtained that is account all articles in the field of chemistry in the year of 2010 on the Scopus database. Then from the field of author name, the first 20 authors were selected as the top writers. The search on the Scopus Citation Database showed that the number of articles of top chemistry writers is 1842. In order to extract articles with a digital indicator, Scopus output based on this indicator, the number of citations received, and in the TEXT format were extracted; in the end, 1782 articles have the digital indicator that they were selected as the research sample.

Data collection was carried out in June 2016, as follows: To categorize articles into three groups of papers with a weak, medium, and high citations, primarily the papers extracted from Scopus database, were sorted

based on the number of citations from high to low in excel software, then by using the formula PERCENTILE.INC (A1:A1 782; 33%) the articles were divided into three groups based on citations (each group33%). As a result of this division, the first group of papers with weak citations included the articles with zero and one citation, the second group of articles with a moderate citation included articles with two to five citations and the third group of articles with more citations included articles with an above six citations.

After grouping the articles, the indicators for each group were arranged in a notepad file and the altmetrics' data were extracted using the webometric analyst software. An online webometric analyzer needs a keyboard account to run, and this software is only launched through Microsoft's system, so first by visiting Microsoft site, an account was created in Outlook.com. Then the key type was received from Windows Azure Marketplace Account page. Finally, by entering the key in the box "the Windows Azure Marketplace Account Key" the webometric analyst software has been activated and altmetrics' data, related to each of the groups, have been extracted separately by using the altmetric section of the software. In this study, for descriptive analyzing of the data (frequency, mean, standard deviation) the Excel software has been used. Inferential analysis of data was done using SPSS software version 19. Then, to investigate the hypothesis of the research, Spearman's correlation coefficient test has been used to determine the relationship between the number of citations and the altmetrics of the reviewed articles based on the purpose of the research.

The first research hypothesis:

In Mendeley service, there is a relationship between the number of citations and the amount of saving articles written by top chemistry writers in 2010.

Since the number of articles in three groups stored in Mendeley was not so much, the correlation test for articles was calculated in three levels of low, medium and high citation. As shown in Table 1 of the correlation test, the decision criterion values or P-VALUE obtained in the two groups of low and medium citation is greater than 0.05, therefore the zero assumption is confirmed and it can be concluded that there is no relationship between the number of citations and the number of articles stored in Mendeley in these two groups. However, the results obtained from the group of high citation indicate that the criterion of the decision is less than 0.05; therefore, the zero assumption is rejected, and it can be concluded that there is a significant relationship between the number of citations and the number of articles stored in Mendeley in the group of papers with high citation. (Correlation coefficient= 0.7). The results of the study show that there is a relationship between total acetone and the number of articles stored in Mendeley (Correlation coefficient=0.74)

Citation	Mendeley		
	Number	Correlation	Decision criteria
Weak	15	-0.069	
Medium	23	0.147	
Great	131	0/702**	
Total	169	0.738**	

Table 1: Correlation between Citation number and Article storing in Mendeley

The second research hypothesis:

In the CiteUlike, there is a relationship between the number of citations and the tags of articles written by top chemistry writers in 2010. As the correlation test in Table 2 shown, the decision criterion value is less than 0.05. Therefore, the zero hypothesis has been rejected and can be deduced that between the number of citations and the number of the Articles' tags in CiteUlike have a weak relationship.

CiteUlike			
Citations	Number	Correlation	Decision criteria
	72	0.304**	0.009

Table 2: correlation between Citations number and article comments in CiteUlike

The third research hypothesis:

In the scientific blog, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010. As shown in Table 3 of the correlation test, the decision criterion value is greater than 0.05; Therefore, the zero hypothesis is verified and it can be concluded that there is no relationship between the number of citations and the amount of discussing of articles in blogs.

Weblogs			
Citations	Number	Correlation	Decision criteria
	36	0.166	0.334

Table 3: correlation between Citations number and articles Discussion in Weblogs

The forth research hypothesis:

According to Table 4, in news site, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010.

News			
Citations	Number	Correlation	Decision criteria
	14	0.63*	0.016

Table 4: correlation between Citations number and article Discussion in News

The fifth research hypothesis:

In Twitter, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010. As shown the correlation test in Table 5 the decision criterion value is less than 0.05; so the zero hypothesis is rejected and can be deduced that there is a moderate relationship between the number of citations and the amount of discussion of articles in Twitter.

Twitter			
Citations	Number	Correlation	Decision criteria
	68	0.309*	0.010

Table 5: correlation between Citations number and article Discussion in Twitter

The sixth research hypothesis:

In Facebook, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010

As shown the correlation test in Table 6, the decision criterion value is more than 0.05; so the zero hypothesis is verified and can be deduced that there is no relationship between the number of citations and the amount of discussion of articles in Facebook.

Facebook			
Citations	Number	Correlation	Decision criteria
	38	-0.077	0.645

Table 6: correlation between Citations number and article Discussion in Facebook

The seventh research hypothesis:

In Google Plus, there is a relationship between the number of citations and the amount of discussing of articles written by top chemistry writers in 2010. As shown the correlation test in Table 7, the decision criterion value is more than 0.05; so the zero hypothesis is verified and can be deduced that there is no relationship between the number of citations and the amount of discussion of articles in Google Plus.

Google Plus			
Citations	Number	Correlation	Decision criteria
	25	-0.348	0.088

Table 7: correlation between Citations number and article Discussion in Google Plus

Discussion and conclusion

As shown in table 1, the correlation between the number of citations and the amount of articles stored in Mendeley was calculated in three citation levels due to the sufficient number of articles. Results from the citation group showed a weak and negative correlation (-0.069). The weak citation group included articles with zero and one citation. Therefore, the receipt of altmetrics by articles with zero citation is inevitable, which can affect the amount of correlation obtained. On the other hand unlike Scopus, Mendeley does not provide statistics on unsaved articles. Hence, the negative correlation obtained cannot be confirmed as no relationship. But a possible reason can be storing articles with zero citation in Mendeley. Therefore, it can be deduced that the amount of reserve in Mendeley is not affected by time, and this will be overcome over time. Accordibg to results in table 1, there is no relationship between the number of citations and the amount of storage. However, the obtained results of the two high citation groups and all the papers examined in Mendeley, confirmed high relationship. The amount of correlation coefficient is conformity of positive, significant, and still strong relationship between the number of citations and the amount of saving articles in Mendeley. The findings of this research are the same directions of the Lee and others,2012; Bar-Ilan and others, 2012; Zahedi, 1393 results and some moderate relations of Prima et al., 2012; Thelwall and Wilson, 2014; Bornmann, 2014; Srivastava and Furuta, 2015; Mohammadi et al., 2015. The possible reason for this can be the dramatic increase in Mendeley users over time. Another possible reason could be the more tendency of users to come up with newer articles. As in the study of Srivastava and Furuta (2015), the articles of the two years 2005 and 2010 were reviewed and the results of the research showed that the

relationship in articles of 2010 is stronger than in 2005, and it was due to that users have paid more attention to new articles.

As shown in table 2, the correlation coefficient obtained confirms the existence of a positive, significant and, at the same time, weakly relationship between the number of citations and the number of tags among the papers examined in CiteUlike. This conclusion confirms the findings of previous research (Bar-Ilan et al. 2012; Li et al. 2012; Sotudeh et al., 1394; Bornmann, 2015; Alhoori and Furuta, 2015). The probable cause of the low amount of correlation CiteUlike rather than Mendeley can be found in the rate of Users use. Another reason for the weakness of the relationship can be found in the level of these two tools.

The results from table 3 show that there is no relation between the number of citations and the amount of discussion of articles in scientific blogs. Study of the results of the research has reported poor correlation (Thelwall, et al., 2013; Costa and others, 2014; Bornmann, 2015; Alhoori and Furuta, 2015). Some research also reported no relationship, for example, Dayari(1393) reported the absence of the relationship between the number of citations and blog posts.

The results obtained from the correlation test in table 4 showed, there is no relationship between the number of citations and the amount of articles' discussion in the news media. Results obtained from this test are in the same way of the previous research findings (Thelwall, et al., 2013; Dayari 2014; Costas et al. 2014 and Alhoori and Furuta, 2015).

The results of table 5 show that there is a weak relationship between the number of citations and the amount of articles' discussion on Twitter. Matching the findings of this study is in line with the results of previous studies (Costas et al., 2014; Haustein et al. 2014; Bornmann, 2015; Alhoori & Furuta, 2015). This is while Dayary (1394), in her research reported no relationship, and Thelwall et al. (2013) reported correlation (-0.19) in their research. The study year in this study could be the result of the weak correlation obtained. It's because that the older articles, due to less use of the social web at the time of publication, receive fewer altmetrics score and newer articles gaining more altmetrics scores with the impact of the same articles, because in practice, researchers are looking for the latest articles.

The results of table 6 show no relationship between the number of citations and the amount of articles' discussion on Facebook. The study of the previous research findings (Thelwall, et al. 2013; Dayary 1394; Alhoori and Furuta, 2015) indicate a lack of alignment with the results of this study. On the other hand, a weak relationship in research can be attributed to the policy of collecting Facebook data by altmetric tools. Unlike other data collection tools, altmetrics collected only the statistics of the article's discussion in the public section of Facebook and ignores the personal section.

The results of table 7 show the absence of a relationship between the number of citations and the amount of the article's discussion in the Google Plus. The result can be seen in line with the findings of previous research (Dayary, 1994; Thelwall et al., 2013; Costas et al., 2014; Alhoori and Furuta, 2015). According to the results of these studies, the most altmetrics of the papers take place in the early days of their publication. As stated in the theoretical framework section, the Google Plus network was launched on June 28, 2011, while the year under review was in the 2010 study, which this matter can be effective on the amount of altmetrics data obtained by the present study.

conclusion

The results of this research indicate that the citation has a significant and strong relationship with the number of Mendeley readers and the amount of discussion in the news media, as well as a weak relationship with the number of tags in CiteUlike and Twitter posts. Meanwhile, there was no meaningful relationship

between the number of citations and posts received from Facebook, blogs, and Google Plus. On the other hand, the results show that with increasing research years, the correlation coefficient has also increased. As new research has a higher correlation coefficient. The negative correlation coefficient in the weak citation group of Mendeley indicates that articles with a low number of citations in Mendeley are bold. Overall, Mendeley as a citation management tool and Twitter as a social network have supremacy both in coverage of articles and obtaining altmetrics, and in terms of the existence of a relationship with other sources. Due to the low presence of articles in the sources of altmetrics, the existence of a weak relationship and non-relation to citations, as well as the reflection of the users' different effects and the evaluation of different scientific outputs, these metrics can be used as a complementary in the citation indices in the research evaluation, hence the use of the term Complementary metrics rather than alternative metrics seems to be more appropriate.

References

Alhoori, H., and Furuta, R. (2014). F1000 Recommendations as a Potential New Data Source for Research Evaluation: A Comparison with Citations, J. Assoc. Inf. Sci. Technol. vol. 65, no. 3, pp. 433-445.

Bar-Ilan, J., Haustein, S., Peters, I., Priem, J., Shema, H., and Terliesner, J. (2012). Beyond citations: Scholars' visibility on the social Web.in Proceedings of the 17th International Conference on Science and Technology Indicators, vol. 52900, pp. 98–109.

Bornmann, L. (2015). Alternative metrics in scientometrics: a meta-analysis of research into three altmetrics. Scientometrics, 103:1123–1144.

Costas, Rodrigo, Zohreh Zahedi, and Paul Wouters(2014). Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. ArXiv preprint arXiv: 1401.4321.

Costas, R., Zahedi, Z. and Wouters, P. (2014). Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. arXiv preprint arXiv:1401.4321.

Dayari, Sareh (1394). Comparison of Citation and Altmetric of the Top Medical Journals. Master thesis, Shahid Bahonar University, Kerman

Haustein, S. Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2013). Coverage and adoption of altmetrics sources in the bibliometric community, 12. Digital Libraries. Retrieved from http://arxiv.org/abs/1304.7300

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. Scientometrics, 91(2), 461–471. Doi: 10.1007/s11192-011-0580-x.

Mansourian, Yazdan, and Navidi, Fatemeh (1393) Introduction to Altometrics: Alternative metrics to Study the Impact of Research with Emphasis on Social Web. Journal of scientometrics. 1(2), 15-34.

Priem, J., Piwowar, H.A., & Hemminger, B.M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. Available at: arXiv: 1203.4745v1.

Rismanbaf, A, Osareh, F. (1386). Informetrics from Genesis to now. Faslnameh Ketab, 71(3). 50-70.

Salajegheh, M. and Mohammadian, S. (1394) Altmetrics: A new way in Scientometrics. Quarterly Journal of National studies on librarianship and information organization(101), 71-84.

Shrivastava, R., Mahajan, P. (2016). Relationship between citation counts and Mendeley readership metrics: A case of top 100 cited papers in Physics. New Library World. Vol. 117 Iss: 3/4, pp.229 – 238

Sotudeh, H. Mazarei, Z. and Mirzabeigi, M. (1394) Investigation of The relationship between Citation Indicators and CiteUlike's tags: A Case Study the articles in the area of Information Science and Librarianship Sciences in 2012-2004,Journal of Information Processing & Management, 936-939

Thelwall, M., Haustein, S., Larivière, V., and Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. PLoS One, vol. 8, no. 5, p. e64841.

Thelwall, M., & Wilson, P. (2015). Mendeley readership altmetrics for medical articles: An analysis of 45 fields. *Journal of the Association for Information Science and Technology*, n/a-n/a. Retrieved from http://doi.wiley.com/10.1002/asi.23501

West, J. (2008). Eigenfactor: Ranking and mapping scientific knowledge. .The 6th SPARK Japan Seminar 2008, National Institute of Information. M. Richardson (Ed.), 1-15.

Zahedi, Z., Costas, R., Wouters, P. (2013). How well developed are Altmetrics? Cross disciplinary analysis of the presence of 'alternative metrics' in scientific publications (RIP). In Proceedings of the 14th 24 International Society of Scientometrics and Informetrics Conference, Vienna, Austria (pp.876-884). Wien: Facultas Verlags Buchhandels AG.

Zahedi, Z., Fenner, M., & Costas, R. (2014). How consistent are altmetrics providers? Study of 1000 PLOS ONE publications using the PLOS ALM, Mendeley and Altmetric.com APIS. Altmetrics14: Expanding Impacts and etrics, 5–6. Retrieved from http://dx.doi.org/10.6084/m9.figshare.1041821.

-