Submitted on: 19.10.2015

Does open source track the effects of open access in the society?

Sri Amudha S
Department of Library and Inforamtion Science, Pondicherry University, Pondicherry, India
E-mail: sriamudhas@gmail.com

Abstract:
The Open Access (OA) is a means by which all users are accessible to the peer-reviewed scholarly articles. This has been established since early 1990's and has become a major source of journals and articles in this century. As it is a successful method that increases impact, it goes without saying that wide and different target groups are accessible to it which is a very positive outcome. But, it is also a challenge to quantify such vast and different types of usage or impact. This study attempts to extract the efficiency of Open source in quantifying the level of impact created by the OA. This was accomplished by 2 objectives namely: by investigating the effects of NPG's new initiative towards Open Access to provide read only access to research articles on Open Source Altmetric providers. And to identify the non-academic intellectual impact created by OA journal articles in the society beyond academia using altmetrics, the open metrics. The study showed that the NPG's new initiative called the beggar link had no significant effect on the impact of Open Access and that the Open source does quantify the impact of OA journals. The evaluation of the data present on OA when analysed showed that the OA had higher impact on most of the altmetric platforms when compared to Non-OA articles.

Introduction

In the present world with internet being a great source for faster communication, writing software had become a community activity where they collaborate for innovation. Open source targets the needs of its users and provide social benefit for wider group. Therefore there is no doubt that it extends its scope to academia world too. Research evaluation is one domain where open source stepped in recently.

Researchers in academia have been trying to figure out methods that are more productive in creating and measuring research impact. Traditionally research evaluation is a measure ignoring a significant group of target audience – the society beyond academia. Research community is attempting to bridge this gap using altmetrics by tracking the online attention and research communication around research outcomes via online open APIs in both scholarly and non scholarly platforms. Impactstory, Altmetric.com and Plum analytics are the major altmetric service providers. Amongst these Impactstory runs on an Open source program however it serves mainly research individuals. In spite of the need to investigate the role of Open source in research evaluation, this study gathers altmetric data from altmetric.com for convenience of data and to avoid sample bias. The goal of this study is to identify and justify the role of open source in research evaluation using altmetrics. For this purpose, the paper has two objectives.

- To identify the non- academic intellectual impact created by OA journal articles thus investigating the effects of open access journals in the society beyond academia using altmetrics, the open metrics.
- To investigate if the NPG's new initiative towards Open Access to provide read only access to research articles will have affect the Open Source Altmetric providers.

This has been done by taking into consideration all other expressions of impact or usage from social web and media like Reddit threads, Bloggers, Tweeters, Google+ authors, pinterest posts, Q&A site users, Facebook walls, Weibo users, Wikipedia pages, CiteULike readers, news outlets et., along with other sites that include scholarly inputs like citations, altmetric score, F1000 reviews, Peer review sites, mendley readers etc.

Literature review

Open Access has been established since the early 1990s.The landscape of academic research has changed vastly from the year 1993. OA has thrived ever since the internet gave access to research articles for a lower cost. It is know that 8.5% of articles released in 2008 is available as OA. About 62% of the older articles are seen to be in one or the other form of OA. The major issue is that OA cannot be easily quantified or expressed within an index. In order for researchers to be able to accomplish that, there have been certain improvements made. There have been certain internet sites that have established over the years to accommodate OA. These have been organized to a certain level in order to enable us the ability to quantify OA, they are Ulrich's periodical directory, ISI's web of science, Scopus, The Directory of Open Access Journal (DOAJ) (Bjork et al, 2009).

The quantification of OA has been pondered over for years and the Altmetrics is one of the evolving metric systems that has been found to be complimentary to solve this particular mystery. Altmetrics is

considered the ‘alternate metrics’. It considers a diverse group of altmetric indicators. It uses web as the key to extract information of usage and impact of articles and that includes OA articles (Konkeil et al, 2014).

The altmetrics indicators are usually platforms from which information on different types of usage and the amount of usage can be obtained for Non-open access and open access articles. There have been several articles that find the impact by considering one or more altmetric indicators like twitter, facebook, blogs etc., which are social web from where usage data can be collected. (Thelwall et al, 2013). There are Open source codes and sites that actively collect and store the data from these sites to individual articles for further study such as ours (Altmetrics.com)

Citations can also be considered as an evaluating factor to quantify the OA data. There has been debate as to whether or not the OA has affected the citation of papers and it has been found that they do not. They have similar impact from citations like all other.

The method to actually find the efficiency of Open access when compared to Non-Open access has been debated by many research scholars. There have been comparison by means of altmetric indicators where many such web social inputs are collected for every article and compare them through t-test as it compares 2 groups with similar variables to find the significance between the groups (Thelwall et al, 2013)

The Nature Communications Journal has been one of the hybrid journals where the authors can decide whether or not their article needs be OA or Non-OA. This particular journal can be a good source for both OA and Non-OA data (Thelwall et al, 2013). This journal has recently introduced a new feature known as the beggar journal which is a read-only option of the article of choice. This is solely to obtain more accurate social usage data.

Methodology

The present study was conducted to calculate the effect of open source. In order to accomplish the objectives we based it on the articles published in Nature Communication journal during the period of 2013 – 2014. In a total of 3973 articles were taken and the number of articles that had at least one Altmetric score is 3002. For these articles, citations were collected from Scopus databases and Altmetric data were collected by submitting a request to Altmetric.com. Only the articles with atleast one altmetric score were chosen for analysis. The data obtained included the scores on altmetric indicators like citations, altmetric score, Reddit threads, bloggers, tweeters, google+ authors, F1000 reviews, pinterest posts, news outlets, Q&A site users, facebook walls, Weibo users, Peer review sites, Wikipedia pages, Policy documents, Mendeley readers, CiteULike readers. The entire dataset was formatted in two different types to find the solution for the proposed 3 objectives. The data was tabulated such that the

altmetric platforms were firstly divided into Open-Access and Non-open access. Then it was further tabulated into yearly data. The data consists of 2598 articles before 20/10/2014 and 404 articles after 20/10/2014. The numbers signify the period before and after the beggar link was introduced in Nature journal.

The number was further fed into SPSS software to accomplish t-test to find the significance between the 2 variables respectively. If the data does not show significance it implies that they are inversely proportional. The calculations gave results that included Significance, Standard error mean and t-test value.

Result

The data when tabulated appropriately as in this study gave space for various calculations and analysis. The Each record was tagged as Open and Non Open and is grouped into two different cohorts. The data was divided into the period before and after the introduction of beggar link as the 2 groups to be evaluated by t-test to provide us with the appropriate result for the first objective. The t-test value represents the impact of either OA or Non-OA in the particular altmetric indicator for the first objective.

Objective 1:

The 17 academic indicators when tabulated as given below (Table 1), was easily processed to obtain the necessary output to evaluate whether OA is quantified by the Open source and also to find if OA has higher impact when compared with Non-OA.

Table 1:

	Article Type	Mean	Std. Error Mean	t test (sig.)
Citations	Non-Open Access	2.556	0.164	2.538 (0.011*)
	Open Access	4.028	0.776	
Altmetrics Score	Non-Open Access	27.518	1.169	4.721 (0.000*)
	Open Access	38.906	2.488	
Reddit threads	Non-Open Access	0.040	0.006	3.858 (0.000*)
	Open Access	0.086	0.013	
Bloggers	Non-Open Access	0.588	0.033	3.711 (0.00*)
	Open Access	0.827	0.062	
Tweeters	Non-Open Access	6.152	0.382	8.200 (0.000*)
	Open Access	13.694	1.071	
Google+ authors	Non-Open Access	0.282	0.036	3.696 (0.000*)
	Open Access	0.573	0.086	
F1000 reviews	Non-Open Access	0.057	0.005	4.067 (0.000*)
	Open Access	0.097	0.010	
Pinterest posts	Non-Open Access	0.023	0.016	0.522 (0.601^{NS})
	Open Access	0.011	0.011	
News outlets	Non-Open Access	2.340	0.102	2.472 (0.014*)
	Open Access	2.821	0.183	
Q&A site users	Non-Open Access	0.012	0.012	0.680 (0.497^{NS})
	Open Access	0.000	0.000	
Facebook walls	Non-Open Access	0.797	0.193	2.542 (0.011*)
	Open Access	1.588	0.188	
Weibo users	Non-Open Access	0.141	0.023	0.587 (0.557^{NS})
	Open Access	0.164	0.031	
Peer review sites	Non-Open Access	0.011	0.002	0.677 (0.498^{NS})
	Open Access	0.008	0.003	
Wikipedia pages	Non-Open Access	0.079	0.010	1.780 (0.075^{NS})
	Open Access	0.114	0.019	
Policy documents	Non-Open Access	0.007	0.006	0.541 (0.588^{NS})
	Open Access	0.014	0.014	
Mendeley readers	Non-Open Access	13.354	0.426	6.439 (0.000*)
	Open Access	18.677	0.793	
CiteULike readers	Non-Open Access	0.301	0.068	1.182 (0.237^{NS})
	Open Access	0.444	0.099	

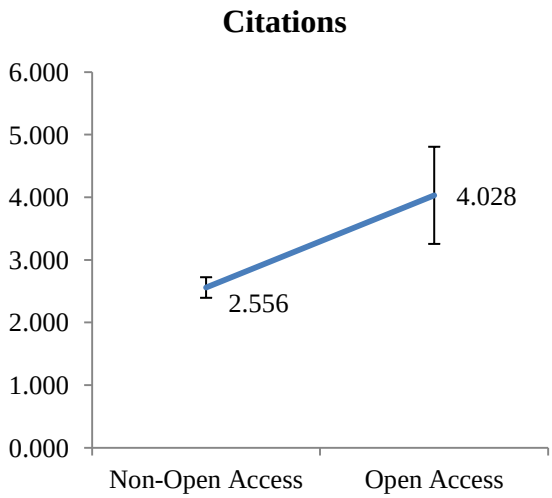
The value of OA was always seen to be high when compared to Non-OA. The t-test revealed that the tweeter had highest value of 8.200 and Pinterest had the lowest value of 0.522. The Highest mean of scores was seen in Altmetric score with an average of 38. The Standard error mean was seen high in Altmetric score with 1.071 and lowest was seen in F1000 with 0.005. All the high scores were seen in OA, and all the low scores were seen in Non-OA, except for Q&A site users where the OA value is 0 and Non-OA is 0.012 and the Peer review site. One might argue if it is essential to consider these as an indicator to impact among society beyond academia. Henceforth, this paper proposes the two terms differentiating the two impacts as below for clear understanding.

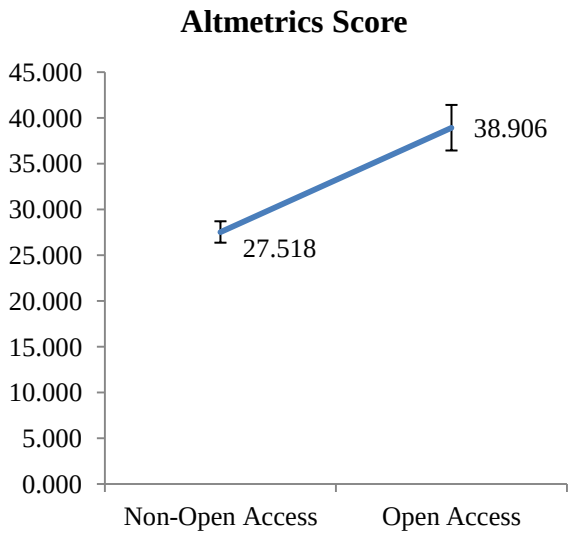
Citational Impact can be understood as the usage of knowledge inside academia resulting in creation of new knowledge. For instance, citing a research article to support a claim made in new research article, which is well captured by traditional metrics.

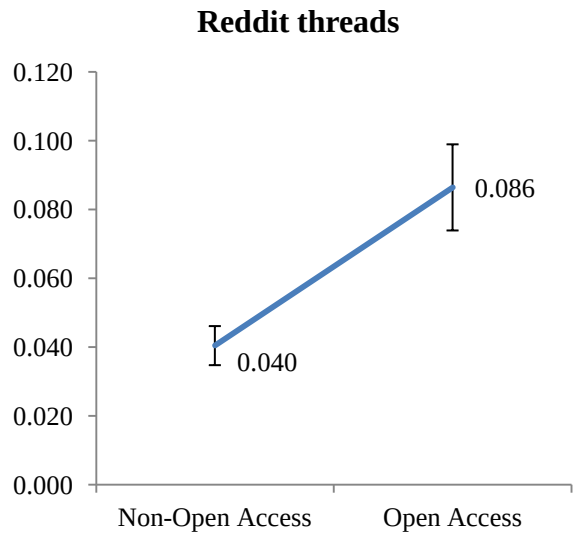
Non-citational Impact can be understood as usage of existing knowledge without resulting in creation of new knowledge. Further, Non academic impact can be of two types, Research acceptance and Research application (reuse) excluding the usage inside academia. This includes classroom teaching, debate etc. Research acceptance is included under non academic impact based on the definition of impact framed by Allen and others. (Allen et al. 2013). Research acceptance is otherwise known as Intellectual impact. These are tracked by altmetrics.

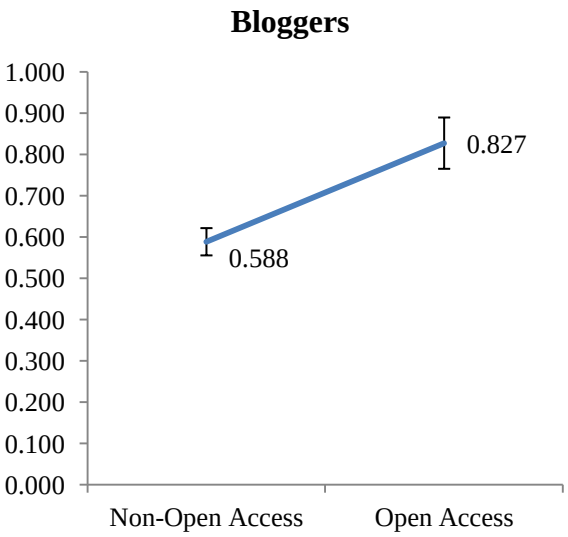
Along with this new hierarchical classification in mind, the results of the investigation clearly show that the impact of Open access is high and also that altmetrics proves to be a good indicator to measure research impact among society thus giving us an answer to our second objective. The correlation between altmetrics score and citations is observed to be -0.026 with a significance of 0.153 indicating that they are not correlated. However altmetrics score and citations are found to be inversely related to each other in this particular sample data.

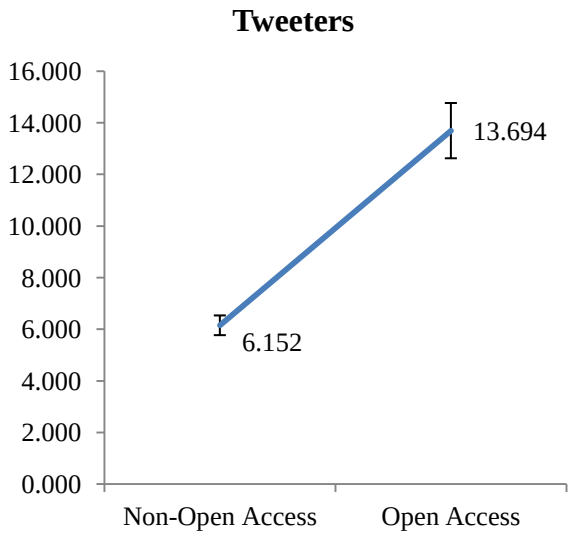
List of graphs showing the impact created by OA and Non OA articles

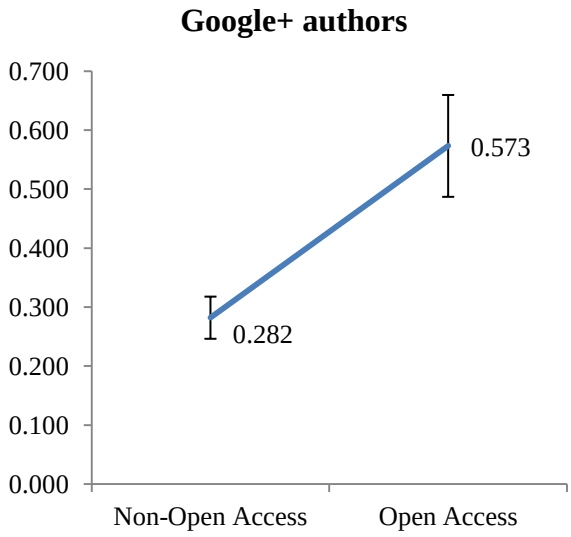


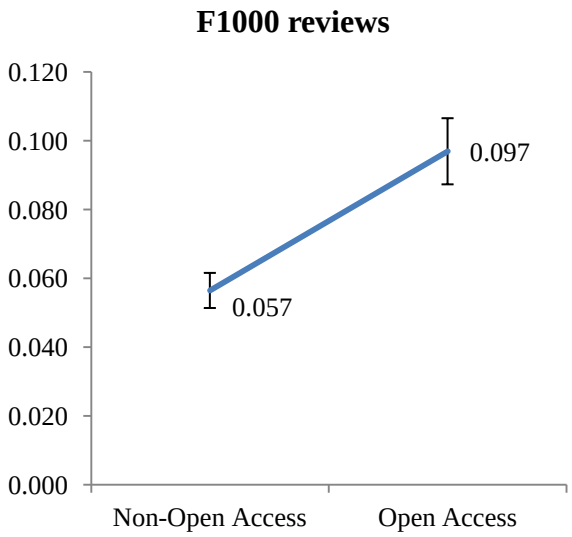


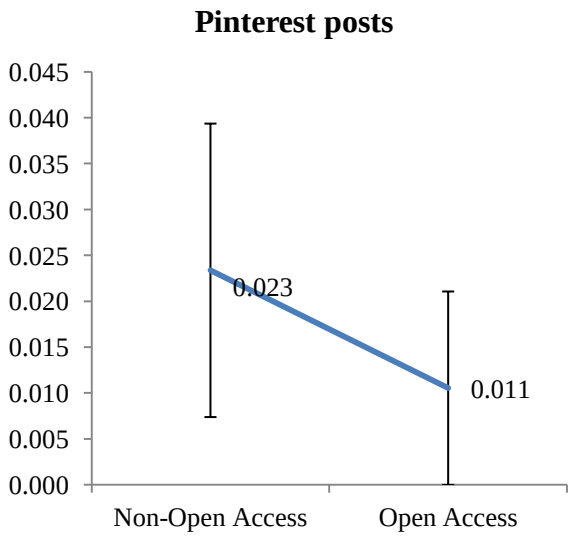


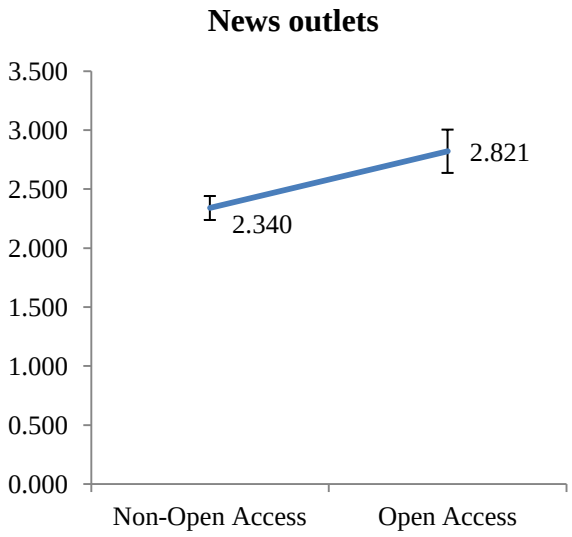


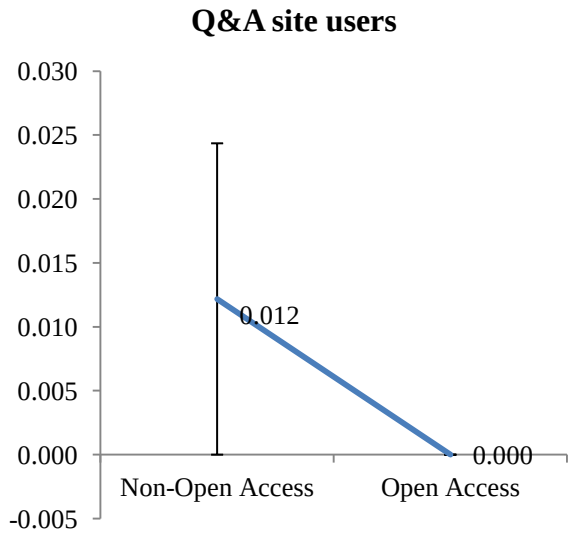

Tweeters
16.000
14.000
12.000
10.000
8.000
6.000
4.000
2.000
0.000
13.694
6.152
Non-Open Access
Open Access

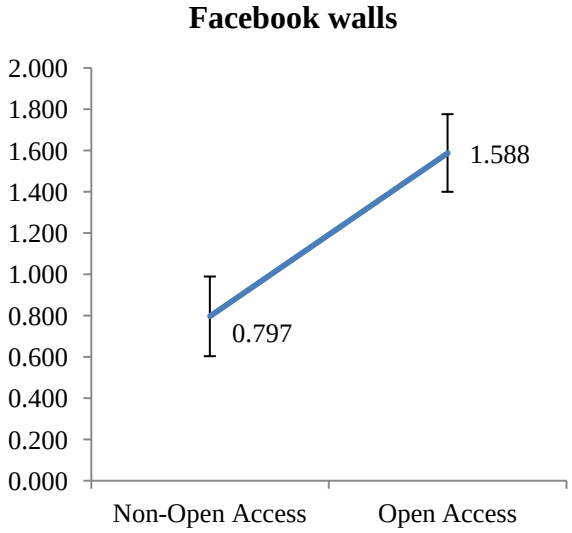

Google+ authors
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
0.573
0.282
Non-Open Access
Open Access

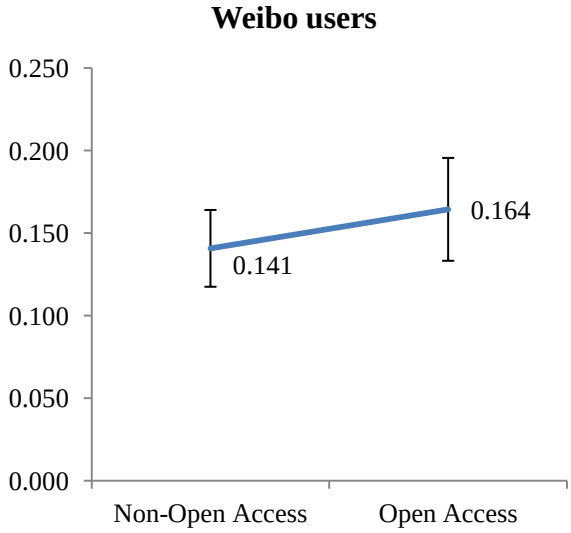

F1000 reviews
0.120
0.100
0.080
0.060
0.040
0.020
0.000
0.097
0.057
Non-Open Access
Open Access

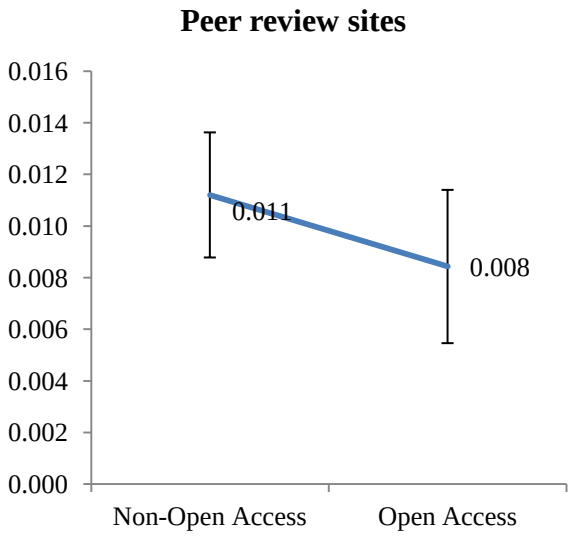

Pinterest posts
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0.023
0.011
Non-Open Access
Open Access

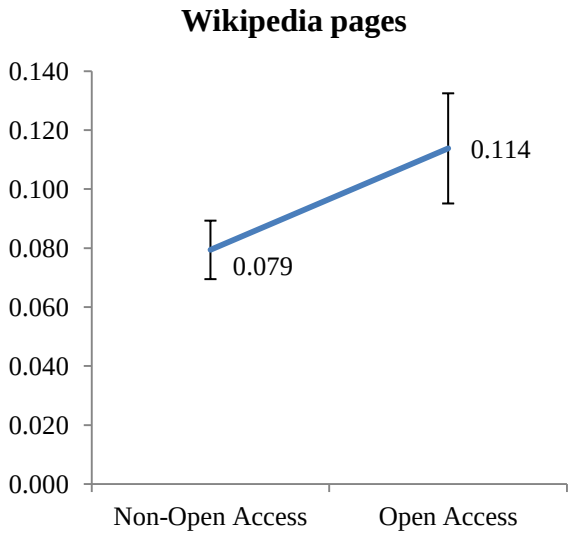

News outlets
3.500
3.000
2.500
2.000
1.500
1.000
0.500
0.000
2.821
2.340
Non-Open Access
Open Access

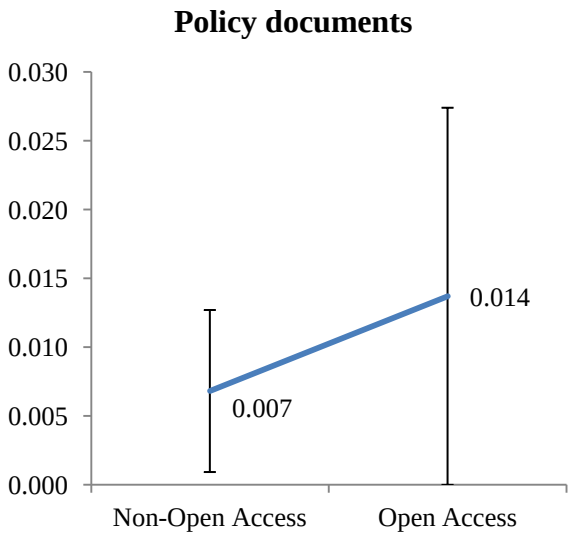

Q&A site users
0.030
0.025
0.020
0.015
0.010
0.005
0.000
-0.005
0.012
0.000
Non-Open Access
Open Access

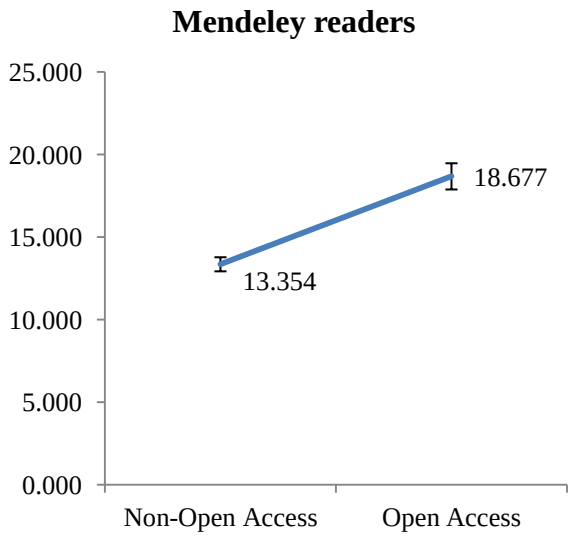

Facebook walls
2.000
1.800
1.600
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
1.588
0.797
Non-Open Access
Open Access

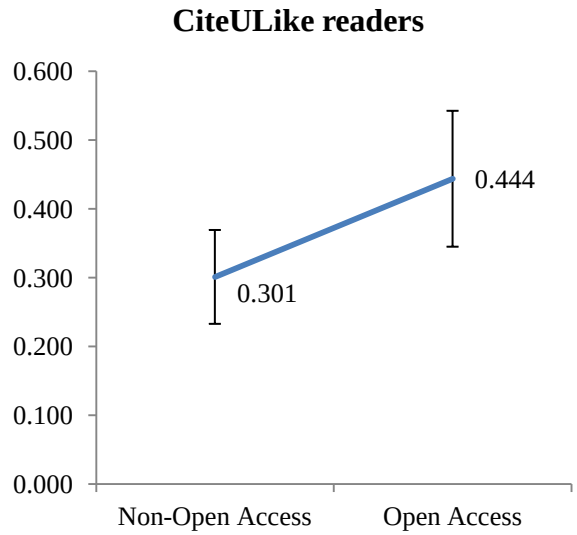

Weibo users
0.250
0.200
0.150
0.100
0.050
0.000
0.164
0.141
Non-Open Access
Open Access


Peer review sites
0.016
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
0.011
0.008
Non-Open Access
Open Access


Wikipedia pages
0.140
0.120
0.100
0.080
0.060
0.040
0.020
0.000
0.079
0.114
Non-Open Access
Open Access


Policy documents
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0.007
0.014
Non-Open Access
Open Access


Mendeley readers
25.000
20.000
15.000
10.000
5.000
0.000
13.354
18.677
Non-Open Access
Open Access



The graphs show a detailed visual idea as to how the Open Access has higher impact scores than in Non Open Access. The values are high in 14 out of 17 indicators. The tweeter and google+ authors show the highest difference where the OA is high and Non-OA is low. The 3 altmetric indicators which show reverse results are Peer riview sites, Pintrist and Q&A site users.

Objective 2

The data was modulated to find if there was any increase in Altmetric score due to the introduction of beggar links (on 20th October) the records that are published before 20th October and after 20th Oct were classified under two groups. The data consists of 2598 articles before 20/10/2014 and 404 articles after 20/10/2014

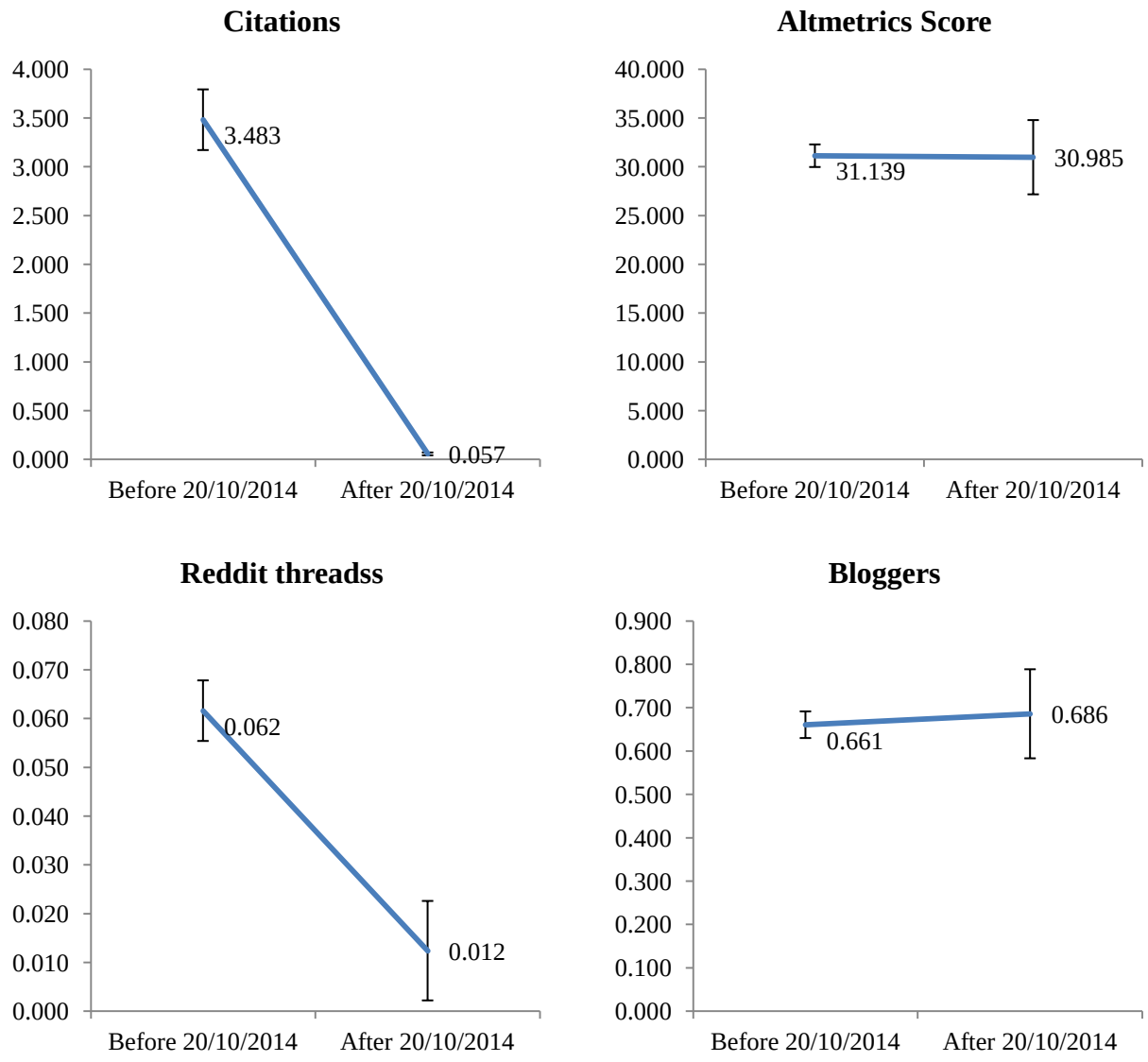
Table: 2 Results of t- test on data based on different time frame

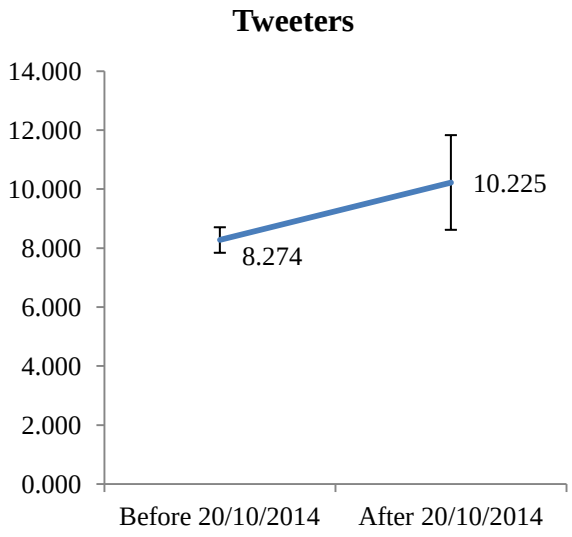
	Group based on date	Mean	Std. Error Mean	t value (sig.)
citations	Before 20/10/2014	3.483	0.311	4.343 (0.000*)
	After 20/10/2014	0.057	0.014	
Altmetrics Score	Before 20/10/2014	31.139	1.158	0.047 (0.963^{NS})
	After 20/10/2014	30.985	3.806	
Reddit threads	Before 20/10/2014	0.062	0.006	3.028 (0.002*)
	After 20/10/2014	0.012	0.010	
Bloggers	Before 20/10/2014	0.661	0.031	0.286 (0.775^{NS})
	After 20/10/2014	0.686	0.103	
Tweeters	Before 20/10/2014	8.274	0.433	1.541 (0.124^{NS})
	After 20/10/2014	10.225	1.605	
Google+ authors	Before 20/10/2014	0.380	0.039	0.429 (0.668^{NS})

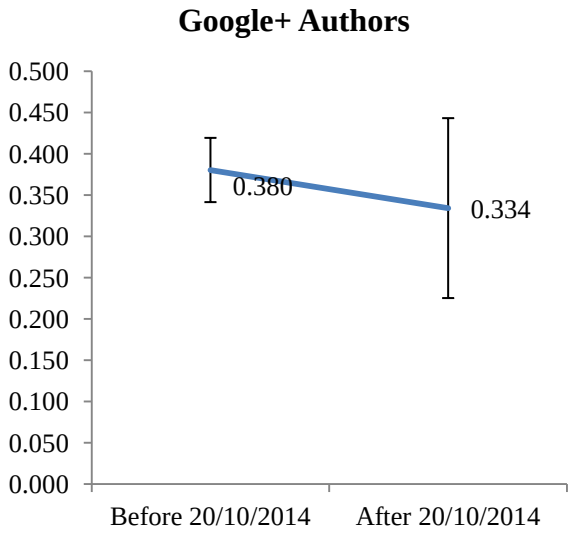
	After 20/10/2014	0.334	0.109	
F1000 reviews	Before 20/10/2014	0.073	0.005	1.894 (0.058^{NS})
	After 20/10/2014	0.047	0.011	
Pinterest posts	Before 20/10/2014	0.012	0.008	1.725 (0.085^{NS})
	After 20/10/2014	0.069	0.069	
News outlets	Before 20/10/2014	2.521	0.096	0.806 (0.420^{NS})
	After 20/10/2014	2.307	0.270	
Q&A site users	Before 20/10/2014	0.000	0.000	2.538 (0.011*)
	After 20/10/2014	0.062	0.062	
Facebook walls	Before 20/10/2014	1.042	0.161	0.095 (0.925^{NS})
	After 20/10/2014	1.082	0.297	
Weibo users	Before 20/10/2014	0.109	0.018	5.303 (0.000*)
	After 20/10/2014	0.399	0.078	
Peer review sites	Before 20/10/2014	0.011	0.002	1.113 (0.266^{NS})
	After 20/10/2014	0.005	0.005	
Wikipedia pages	Before 20/10/2014	0.098	0.010	2.115 (0.035*)
	After 20/10/2014	0.042	0.022	
Policy documents	Before 20/10/2014	0.010	0.007	0.600 (0.549^{NS})
	After 20/10/2014	0.000	0.000	
Mendeley readers	Before 20/10/2014	17.361	0.429	15.854 (0.000*)
	After 20/10/2014	0.092	0.045	
CiteULike readers	Before 20/10/2014	0.342	0.052	0.194 (0.846^{NS})
	After 20/10/2014	0.374	0.246	

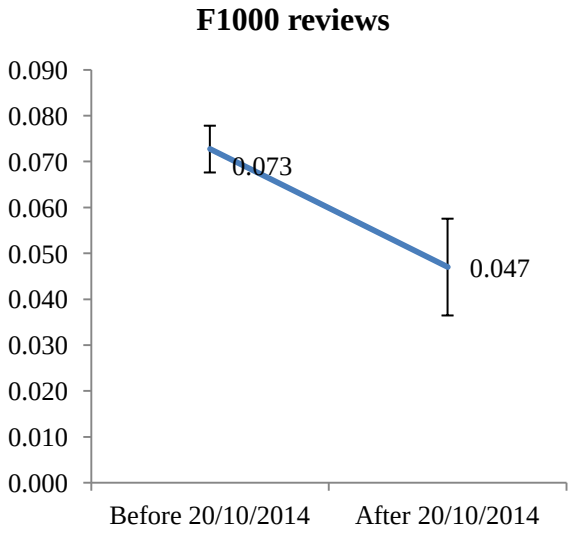
The effect of the beggar link seems to show invaluable results as each indicator has given a different reaction to the new feature. Out of 17 indicators 4 showed improvement after the introduction of beggar link, they were Q&A sites, Tweeter, Pinterest posts, blogger. The Tweeter indicator showed increase in mean after the beggar link introduction, with a value of 10.225 from 8.274 before. The Q&A improved to 0.069 from 0, the remaining 2 had similar variance. All the others showed negative impact. The highest value is seen in the Mendley reader indicator with a t-test value of 15.854. The mendley indicator had a mean of 17.367 before beggar link and dropped down to 0.092 after the same. The lowest t-test value was seen in altmetric score with 0.047, though the average score was about 30 both time showing that in some cases the beggar link does not affect it much.

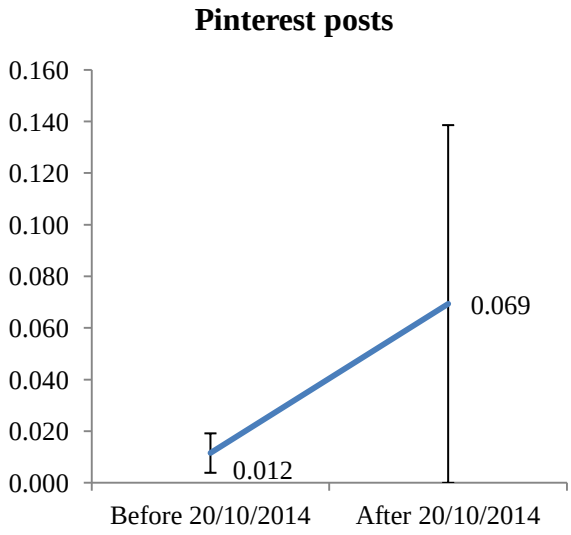
Figure 2 Graphical representation of the results of t- test on data based on different time frame

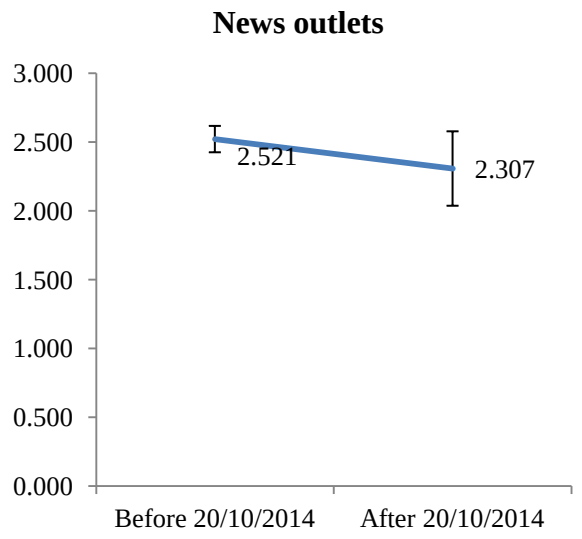


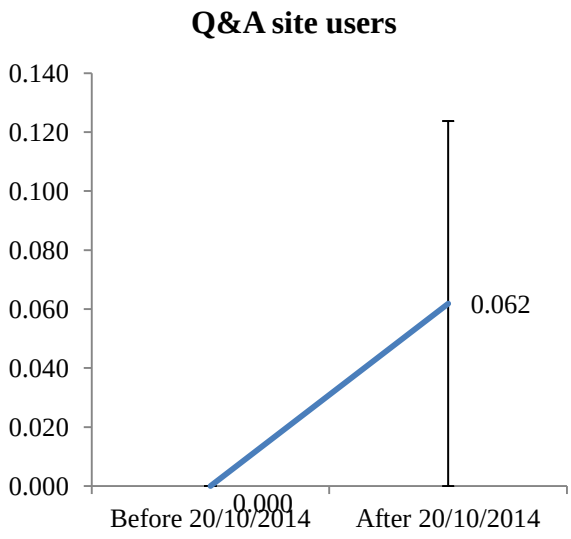

Tweeters
14.000
12.000
10.000
8.000
6.000
4.000
2.000
0.000
10.225
8.274
Before 20/10/2014
After 20/10/2014

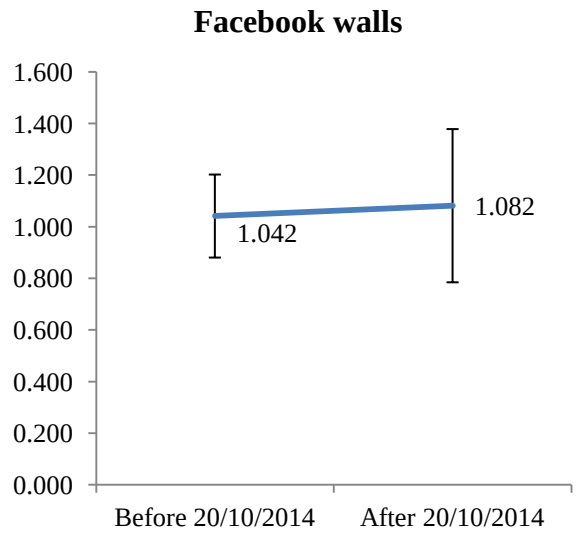

Google+ Authors
0.500
0.450
0.400
0.350
0.300
0.250
0.200
0.150
0.100
0.050
0.000
0.380
0.334
Before 20/10/2014
After 20/10/2014

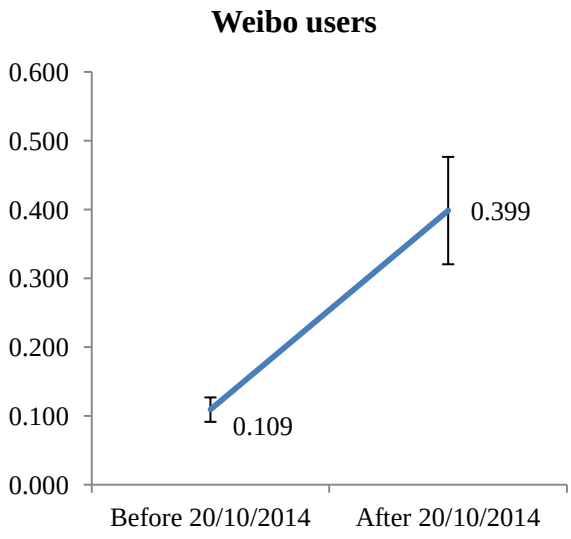

F1000 reviews
0.090
0.080
0.070
0.060
0.050
0.040
0.030
0.020
0.010
0.000
0.073
0.047
Before 20/10/2014
After 20/10/2014

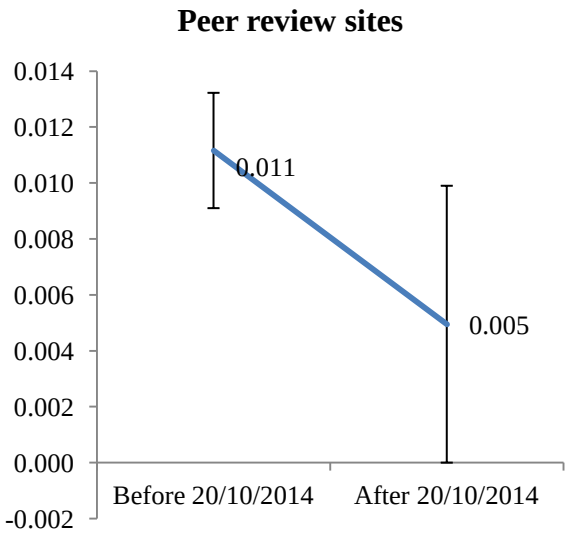

Pinterest posts
0.160
0.140
0.120
0.100
0.080
0.060
0.040
0.020
0.000
0.069
0.012
Before 20/10/2014
After 20/10/2014

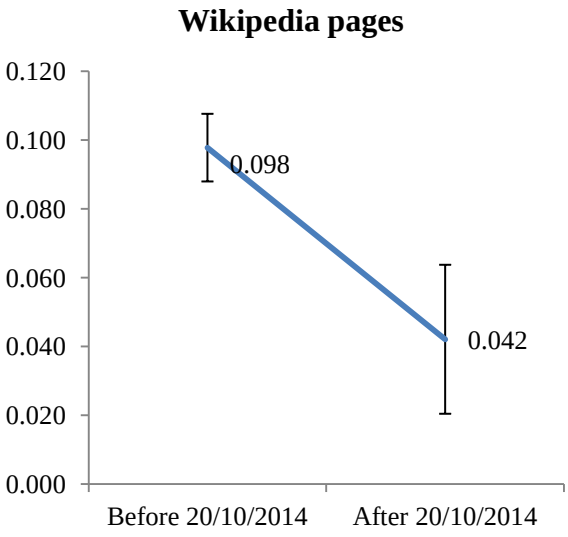

News outlets
3.000
2.500
2.000
1.500
1.000
0.500
0.000
2.521
2.307
Before 20/10/2014
After 20/10/2014

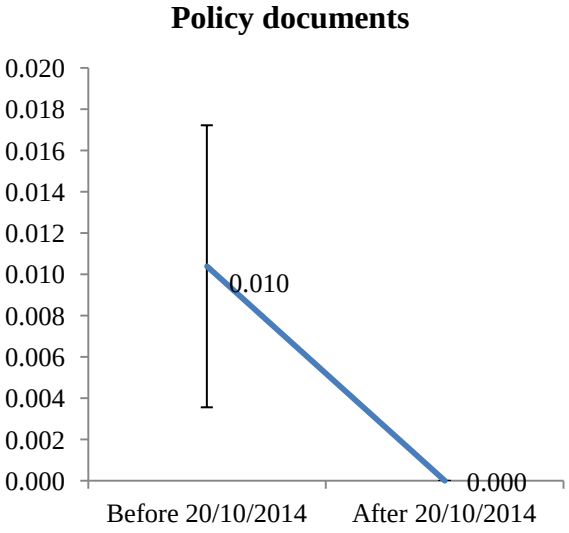

Q&A site users
0.140
0.120
0.100
0.080
0.060
0.040
0.020
0.000
0.062
0.000
Before 20/10/2014
After 20/10/2014

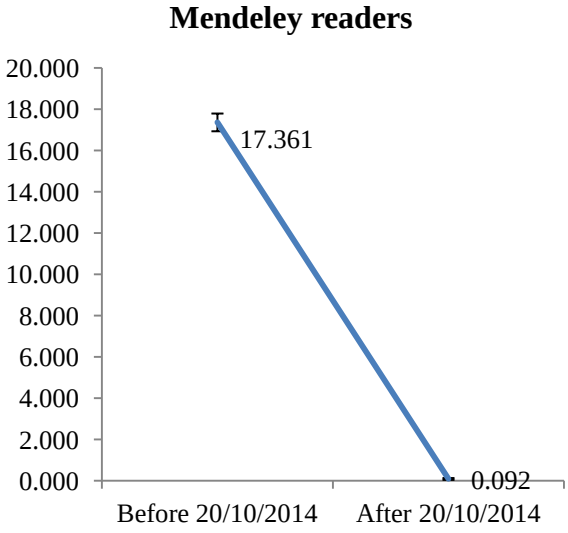

Facebook walls
1.600
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
1.042
1.082
Before 20/10/2014
After 20/10/2014

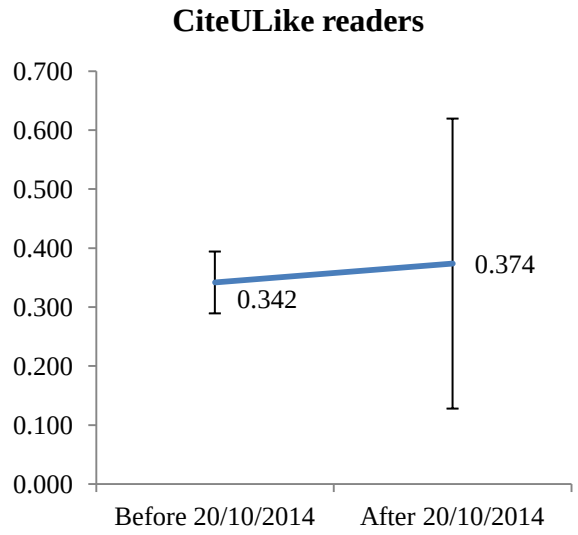

Weibo users
0.600
0.500
0.400
0.300
0.200
0.100
0.000
0.109
0.399
Before 20/10/2014
After 20/10/2014


Peer review sites
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
-0.002
0.011
0.005
Before 20/10/2014
After 20/10/2014


Wikipedia pages
0.120
0.100
0.080
0.060
0.040
0.020
0.000
0.098
0.042
Before 20/10/2014
After 20/10/2014


Policy documents
0.020
0.018
0.016
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
0.010
0.000
Before 20/10/2014
After 20/10/2014


Mendeley readers
20.000
18.000
16.000
14.000
12.000
10.000
8.000
6.000
4.000
2.000
0.000
17.361
0.092
Before 20/10/2014
After 20/10/2014



The graphs given above compliment the table 2 in their representation of the results. Instead of all the indicators showing one trend they showed four different types. The Policy documents, Mendley readers, Peer review sites, Wikipedia pages,Reddit threads and citation scores all showed a huge drop in value from before to after, whereas Weibo users, Q&A sites, Pinterest posts, Tweeters all showed a huge increase in value after the beggar link introduction. The CiteULike, facebook walls and blogger showed negligible increase, and News outlet, F1000 reviews, Google+ authors and altmetric scores showed very low decrease in values. As an overall view there are more indicators showing a decrease in value than increase.

Discussion:

The 17 different altmetric indicators or platforms that were taken into account in this study gave a very deep understanding on the dynamics of this study. Out of the 17 indicators, 14 showed higher values in case of OA and to stress the value of them all the indicators that contribute to the altmetric score showed the same. The twitter being one of the altmetric indicators showed the highest t value, standard mean error. It is also known that out of the 17 indicators only the Peer review site that is meant for scholarly users showed low OA impact, this clearly shows that OA is also more useful for the academic world and scholars. This proves that the Open Access is clearly supported by more users and has a higher impact than the Non- Open Access articles. It can be said without doubt that the Open source (Impactstory) does evaluate the effect of OA in the society which is the base of this study.

The main purpose of beggar link (read only access model) is to track the dark social communications (sharing articles via emails etc). Altmetric captures the online mentions including social media, and the

results shows no significant change of impact created in different platforms that mainly contributes to Altmetrics, therefore the introduction of read only access model introduced by Nature communication to track impact beyond academia among society cannot be considered a great success even if argued against that the major purpose of 'Read only access model' is to make scientific documents open to society. The Weibo indicator that showed increase could also be due to the fact that they have only recently been forthcoming and made it Open Access, and so may not be due to the beggar link. To add to the point all the indicators that contribute to the altmetric score showed negligible difference.

Conclusion

After this extensive work of data collection and analysis the study provided a satisfying output. The study proved that the OA (Open Access) is actively being pursued in various aspects by Open source to find and quantify the impact accomplished by it on the public. The vast data which was once considered to be impossible to collect or quantify has been methodically tackled, collected and quantified to provide us as much information as possible on the effects of OA in the society. The effect of the beggar link introduced in Nature Communications has been analysed to find that they may bring a negative impact, as the different platforms showed varied differences between the periods before and after the introduction of the beggar link. The fact they may contribute to negative impact is proven by the fact that in this study 10 out of 17 altmetrics platform showed decrease in value. But as a closure both the objectives in this study were successfully accomplished. This study could be a base for future extensive studies on evaluating the OA articles by means of altmetric indicators.

Reference

Allen, H.G. et al., 2013. Social Media Release Increases Dissemination of Original Articles in the Clinical Pain Sciences. *PloS one*, 8(7).

Antelman, Kristin.,: 2004, 'Do Open-Access Articles Have a Greater Research Impact?', College & Research Libraries, Vol 65, September

Bjork, B.,Welling, P., Laakso, M., Majlender, P., Hedlund, T., Guonason, G.,: 2010, 'Open Access to the Scientific Journal Literature: Situation 2009', PLOS

Chaire, S.H., Brody, T., 2004, 'Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals', D-Lib Magzine, June, Vol 10

Konkiel, S., Piwowar. H., Priem, J.,: 2014, 'The Imperative for Open Altmetrics', Metrics for Measuring Publishing Value: Alternative and Otherwise, Impactstory

Moed, H.,: 2012, ‘Does open access publishing increase citation or download rated’, Research Trend, May

Mounce. R.,: 2013,‘Open Access and Altmetrics: Distinct but Complemetary’. Bulletin , ASIS, April/May.

Priem, J., Hemminger, B.M.,: 2010, ‘Scientometrics 2.0: Toward new metrics of scholarly impact on the Social Web’, first Monday, July, Vol 15

Rodgers, E., Barbrow, S.,: 2013, ‘A Look at Altmetrics and Its Growing Significance to Research Libraries’, University of Michigan, September edition

Thelwall, M., Houstein, S., Lariviere. V., Sugimoto, C.R.,: 2013, ‘Do Altmetrics Work? Twitter and Ten Other Social Web Services’, PLOS, May 28th