How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications

Zohreh Zahedi · Rodrigo Costas · Paul Wouters

Abstract In this paper an analysis of the presence and possibilities of altmetrics for bibliometric and performance analysis is carried out. Using the web based tool Impact Story, we collected metrics for 20,000 random publications from the Web of Science. We studied both the presence and distribution of altmetrics in the set of publications, across fields, document types and over publication years, as well as the extent to which altmetrics correlate with citation indicators. The main result of the study is that the altmetrics source that provides the most metrics is Mendeley, with metrics on readerships for 62.6 % of all the publications studied, other sources only provide marginal information. In terms of relation with citations, a moderate spearman correlation ($r = 0.49$) has been found between Mendeley readership counts and citation indicators. Other possibilities and limitations of these indicators are discussed and future research lines are outlined.

Keywords Altmetrics · Impact Story · Citation indicators · Research evaluation

Introduction

Citation based metrics and peer review have a long tradition and are widely applied in research evaluation. Citation analysis is a popular and useful measurement approach in the context of science policy and research management. Citations are usually considered as a proxy for 'scientific impact' (Moed 2005). However, citations are not free of limitations

Z. Zahedi (✉) · R. Costas · P. Wouters
Centre For Science and Technology Studies (CWTS), Leiden University, PO Box 905,
2300 AX Leiden, The Netherlands
e-mail: z.zahedi.2@cwts.leidenuniv.nl

R. Costas
e-mail: rcostas@cwts.leidenuniv.nl

P. Wouters
e-mail: p.f.wouters@cwts.leidenuniv.nl

(MacRoberts and MacRoberts 1989; Nicolaisen 2007), they only measure a limited aspect of quality (i.e. the impact on others' scientific publication) (Martin and Irvine 1983; Bornmann and Leydesdorff 2013), their actual meaning has been broadly debated (Wouters 1999) and they also pose technical and conceptual limitations (Seglen 1997; Bordons et al. 2002). On the other hand, peer review or peer assessment is also an important instrument and is often regarded as gold standard in assessing the quality of research (Thelwall 2004; Moed 2005; Butler and McAllister 2011; Taylor 2011; Hicks and Melkers 2012), but it has its own limitations and biases as well (Moed 2007; Benos et al. 2007). Moreover, both citations and peer review are considered mostly as partial indicators of "scientific impact" (Martin and Irvine 1983) and also no single metric can sufficiently reveal the full impact of research (Bollen et al. 2009). Given these limitations, the combination of peer review with "multi-metric approach" is proposed as necessary for research evaluation (Rousseau and Ye 2013) in the line of the "informed peer review" idea suggested by Nederhof and van Raan (1987).

However, the shortcomings of these more traditional approaches in assessing research have led to the suggestion of new metrics that could inform "new, broader and faster" measures of impact aimed at complementing traditional citation metrics (Priem et al. 2012). This proposal of using and applying so-called 'alternative indicators' in assessing scientific impact has entered the scientific debate, and these new metrics are expected not only to overcome some of the limitations of the previous approaches but also to provide new insights in research evaluation (Priem and Hemminger 2010; Galligan and Dyas-Correia 2013; Bornmann 2013).

These alternative metrics refer to more "unconventional" measures for evaluation of research (Torres-Salinas et al. 2013a), including metrics such as usage data analysis (download and view counts) (Blecic 1999; Duy and Vaughan 2006; Rowlands and Nicholas 2007; Bollen et al. 2008; Shuai et al. 2012); web citation and link analyses (Smith 1999; Thelwall 2001; Vaughan and Shaw 2003; Thelwall 2008; Thelwall 2012a, b) or social web analysis (Haustein 2010). The importance of the web as a rich source for measuring impact of scientific publications and its potentials to cover the inadequacies of current metrics in research evaluation have been also acknowledged in these previous studies. For instance, the scholarly evidence of use of publications found on web are seen as complimentary to citation metrics, also as predictors of later citations (Brody et al. 2006) and being of relevance for fields with less citations (Armbruster 2007). In this sense, the more traditional metrics based on citations, although widely used and applied in research evaluation, are unable to measure the online impact of scientific literature (for example via Facebook, Twitter, reference managers, blogs or wikis) and also lack the ability of measuring the impact of scholarly outputs other than journal articles or conference proceedings, ignoring other outputs such as datasets, software, slides, blog posts, etc. Thus, researchers who publish online and in formats different than journal articles do not really benefit from citation based data metrics.

The rise of these new metrics has been framed with the proposition of the so-called "altmetrics" or social media metrics introduced in 2010 by Priem and colleagues (Priem et al. 2010) as an alternative way of measuring broader research impacts in social web via different tools (Priem et al. 2012). More specifically, altmetrics covers mentions of scientific outputs in social media, news media and reference management tools. This development of the concept of altmetrics has been accompanied by a growth in the diversity of tools that aim to track 'real-time'[1] impact of scientific outputs by exploring the

---

[1] Being immediately available compared to citations that take time to accumulate.

shares, likes, comments, reviews, discussions, bookmarks, saves, tweets and mentions of scientific publications and sources in social media (Wouters and Costas 2012). Among these tools we find F1000 (http://f1000.com), PLOS Article-Level-Metrics (ALM) (http://article-level-metrics.plos.org/), Altmetric.com (www.altmetric.com/), Plum Analytics (www.plumanalytics.com/), Impact Story[2] (www.impactstory.org/), CiteULike (www.citeulike.org/), and Mendeley (www.mendeley.com/). These web based tools capture and track a wide range of researcher's outputs by aggregating altmetrics data across a wide variety of sources. In the next section, we summarize the previous studies on altmetrics that have made use of these tools.

Background

The study of altmetrics is in its early stage but some work has already been done. The features of altmetrics tools in general (Zhang 2012) and their validation as a sources of impact assessment has been investigated in some studies. For example, Li et al. (2012) studied the strengths, weaknesses and usefulness of two reference management tools for research evaluation. Their findings showed that compared to CiteULike, Mendeley seems to be more promising for future research evaluation. Wouters and Costas (2012) compared features of 16 web based tools and investigated their potentials for impact measurement for real research evaluation purposes. They concluded that although these new tools are promising for research assessment, due to their current limitations and restrictions, they seem to be more useful for self-analysis than for systematic impact measurement at different levels of aggregation.

Shuai et al. (2012) examined the reactions of scholars to the newly submitted preprints in *arXiv.org*, showing that social media may be an important factor in determining the scientific impact of an article. The analysis of social reference management tools compared to citations has been broadly studied in the field, particularly the comparison of citations and readership counts in Mendeley, in most of the cases showing a moderate and significant correlation between the two metrics (Henning 2010; Priem et al. 2012; Li et al. 2012; Bar-Ilan 2012; Zahedi et al. 2013; Schlögl et al. 2013; Thelwall et al. 2013; Haustein et al. 2013). Also weak correlations between users' tags and bookmarks (as indicators) of journal usage, perception and citations observed for physical journals (Haustein and Siebenlist 2011) have been reported. For the case of F1000, it has been found that both Mendeley user counts and F1000 article factors (FFas) in Genomics and Genetics papers correlate with citations and they are associated with Journal Impact Factors (Li and Thelwall 2012).

Some other studies have focused on whether altmetrics can be used as predictor of citations. For example, in the case of F1000, it has been found that recommendations have a relatively lower predictive power in indicating high citedness as compared to journal citation scores (JCS) (Waltman and Costas 2013). It has been also suggested that at the paper level, tweets can predict highly cited papers within the first 3 days of publication (Eysenbach 2011) although these results have been criticized by Davis (2012) and more research should delve into this point. Moreover, most of the articles that received blog citations close to their publication time are more highly cited than articles without such blog citations (Shema et al. 2013).

Previous studies mentioned above used altmetrics as a new data source and investigated the association between altmetrics and citation impact. Most of these studies were based on journals such as *Nature* & *Science* (Li et al. 2012); *JASIST* (Bar-Ilan 2012), Information

---

[2] Previously known as Total Impact, we use IS in this study to refer to Impact Story. For a review of tools for tracking scientific impact see Wouters and Costas (2012).

System Journal (Schlögl et al. 2013); articles published by bibliometrics and scientometrics community (Bar-Ilan et al. 2012; Haustein, et al. 2013), *PLoS* and other medical and bio-medical journals in PubMed (Priem et al. 2012; Thelwall et al. 2013; Haustein et al. 2013).

However, to the best of our knowledge, little has been done to date to investigate the presence of altmetrics across various scientific fields and also for relatively ample periods of time. This study is thus one of the first in analyzing a relatively large sample of publications belonging to different fields, document types and publication years. This paper builds upon Wouters and Costas (2012) and Zahedi et al. (2013).

Our main objective in this paper is to present an exploratory analysis of altmetrics data retrieved through Impact Story (IS) focusing on the relationship of altmetrics with citations across publications from different fields of science, social sciences and humanities. For this, we examine the extent to which papers have altmetrics obtained through different data sources retrieved via IS and the relationships between altmetrics and citations for these papers. In exploring these issues, we pursue the two following research questions:

1. What is the presence and distributions of IS altmetrics across document types, subject fields and publication years for the studied sample?
2. Is there any relationship between IS-retrieved altmetrics and citation indicators for the studied sample? In other words, to what extent do the IS altmetrics correlate with citation indicators?

Research methodology

In this study, we have focused on IS. Although still at an early stage ('beta version'), IS is currently one of the most popular web based tools with some potentials for research assessment purposes (Wouters and Costas 2012). IS aggregates "impact data from many sources and displays it in a single report making it quick and easy to view the impact of a wide range of research output" (http://impactstory.org/faq). It takes as input different types of publication identifiers (e.g. DOIs, URLs, PubMed ids, etc.). These are run through different external services to collect the metrics associated with a given 'artifact' (e.g. a publication). A final web based report is created by IS which shows the impact of the 'artifacts' according to a variety of metrics such as the number of readers, bookmarks, tweets, mentions, shares, views, downloads, blog posts and citations in Mendeley, CiteULike, Twitter, Wikipedia, Figshare, Dryad, Scienceseeker, PubMed and Scopus.[3]

For this study, we collected a random sample of 20,000 publications with DOIs (published between 2005 and 2011) from all the disciplines covered by the web of science (WoS). Publications were randomly collected by using the *NEW ID ()* SQL command (Forta 2008, p. 193).The altmetrics data collection was performed during the last week of April 2013. The altmetrics data were gathered automatically via the IS REST API,[4] then the responses provided on search requests using DOI's were downloaded. Using this API we could download the altmetric data faster (one request per 18 s) compared to the manual data collection we did for the previous study.[5] The files were downloaded per API search

---

[3] For a full list see http://impactstory.org/faq.

[4] A REpresentational State Transfer (REST)(ful) API (Application Programming Interface) used to make a request using GET (DOIs) and collect the required response from impact Story.

[5] In the previous study, the data collection was performed manually directly through the web interface of IS. Manually, IS allowed collecting altmetrics for 100 DOIs per search and maximum 2000 DOIs search per day in order to avoid swamping the limits of its API, for details see Zahedi, Costas & Wouters (2013).

request separately in Java Script Object Notations (JSON) format on the basis of individual DOI's and parsed by using the additional JAVA library from within the SAS software.[6] Finally, the data was transformed into a comma separated value format and matched back with the CWTS in-house version of the Web of Science on the DOIs to be able to add other bibliometric data to them. The final list of publications resulted in 19,772 DOIs (out of 20,000) after matching.[7] Based on this table, we studied the distribution of altmetrics across subject fields, document types and publication years. Citation indicators were calculated and the final files were imported in IBM SPSS Statistics 21 for further statistical analysis.

In order to test the validity of our sample set we compared the distribution of publications across major fields of science in our sample with that of the whole Web of Science database (Fig. 1) in the same period and only those publications with a DOI. As it can be seen, the distribution of publications of our sample basically resembles the distribution of publications in the whole WOS database, so we can consider that our sample is representative of the multidisciplinarity of the database.

Results and main findings

In the first place, we present the result of our exploratory analysis of the presence of IS altmetrics over the 19,772 WOS publications published between 2005 and 2011. Then, we examine the extent to which papers are represented in the data sources both in general and also across document types, subject fields and publication years. Finally, the relationships (correlation) between IS altmetrics and citations for these papers are compared.

Presence of IS altmetrics by data sources

In our sample, the presence of IS altmetrics across publications is different from each data source. Out of 19,722 publications, 12,380 (62.6 %) papers have at least one reader[8] in Mendeley, 324 (1.6 %) papers have at least one tweet in Twitter, 289 (1.4 %) papers have at least one mention in Wikipedia, 72 (0.3 %) papers have at least one bookmark in Delicious and 7413 (37.4 %) papers have at least one citation in PubMed. Only one paper in the sample has metrics from PLoS ALM.[9] Based on this preliminary test, we decided to exclude some of the metrics from our study: PlosAlm indicators due to their low frequency as they are only available for the PLoS journals thus their presence in our sample is negligible and PubMed-based citations because they are limited only to the Health

---

[6] The additional functionality from the "proc groovy" which is a java development environment added to SAS (Statistical Analysis Systems) environment for parsing and reading the JSON format and returning the data as an object.

[7] From IS one DOI was missing. We also found that 301 DOIs were wrong in WoS (including extra characters that made them unmatchable, therefore excluded from the analysis). Also 61 original DOIs from WOS pointed to 134 different WOS publications (i.e. being duplicated DOIs). This means that 74 publications were duplicates. Given the fact that there was no systematic way to determine which one was the correct one (i.e. the one that actually received the altmetrics), we included all of them in the analysis with the same altmetrics score resulted in: $20,000 - 1 - 301 + 74 = 19,772$ final publications. All in all, this process showed that only 1.8 % of the initial DOIs randomly selected had some problems, thus indicating that a DOI is a convenient publication identifier although not free of limitations (i.e. errors in DOI data entry, technical errors when resolving DOIs via API and also the existence of multiple publication identifiers in the data sources, resulted in some errors in the full collection of altmetrics for these publications).

[8] It means that publications without any metrics were left out of the analysis.

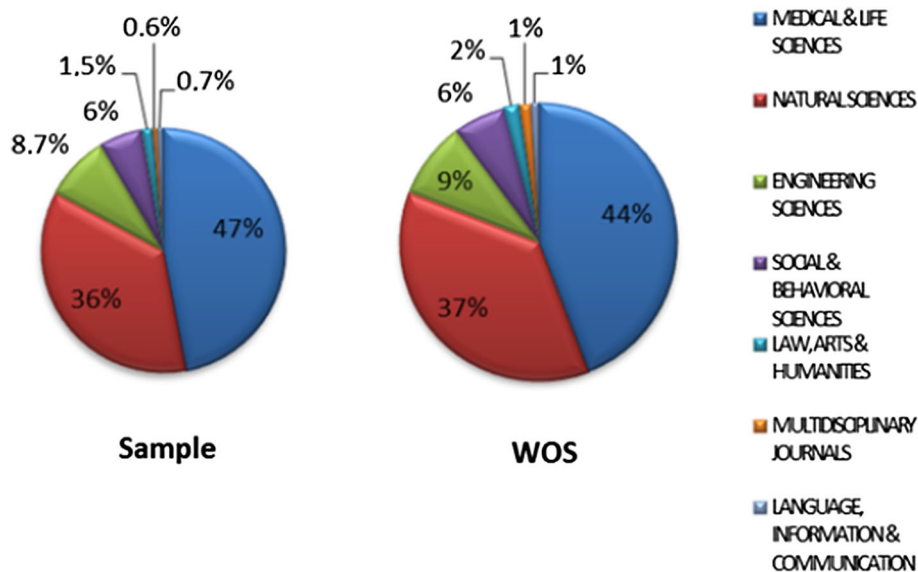[9] This was the only PLOS paper captured by our sample.

Fig. 1 Distribution of publications by major fields of science: sample vs. whole database

Sciences and they refer to citations, which we will calculate directly based on the Web of Science. We also decided to sum the metrics coming from Twitter ("Topsy tweets" and "Topsy influential tweets") given their relatively low frequency. As a result, in the current study, the data from Mendeley, Wikipedia, Twitter and Delicious were analyzed.

Table 1 shows the number and percentages of papers with and without IS altmetrics sorted by % of papers with metrics (excluding the PLOS ALM and PubMed metrics). Based on Table 1, our main finding is that, for this sample, the major source for altmetrics is Mendeley, with metrics on readerships for 62.6 % of all the publications studied. But for other data sources (Twitter, Wikipedia and Delicious), the presence of metrics across publications is very low, with more than 98 % of the papers without metrics. Thus, it is clear that their potential use for the assessment of the impact of scientific publications is still rather limited, particularly when considering a multi-year and multidisciplinary dataset as the one here studied.

Presence of IS altmetrics across document types

Regarding document type, out of 19,772 publications, there are 16,740 (84.7 %) articles, 944 (4.7 %) review papers, 487 (2.4 %) letters and 1,601(8 %) non-citable[10] items in the sample. Table 2 indicates the coverage of the sampled publications with document types across each data sources. According to Table 2, 81.1 % (766) of the review papers, 66.3 % (11,094) of articles, 25.1 % of letters and 24.9 % (398) of non-citable items in the sample have been saved (read) in the Mendeley. In Twitter, 3.4 % (32) of the review papers, 1.9 % (30) of non-citable items, 1.5 % (255) of articles and 1.4 % (7) of letters have tweets. In the case of Wikipedia, 4.6 % (43) of the review papers, 1.4 % (230) of articles and less

---

[10] Non-citable document type corresponds to all WOS document types other than article, letter and review (e.g. book reviews, editorial materials, etc.).

Table 1 Presence of IS altmetrics from data sources

Data source	Papers with metrics	%	Papers without metrics	%
Mendeley	12,380	62.6	7,392	37.3
Twitter	324	1.6	19,448	98.3
Wikipedia	289	1.4	19,483	98.6
Delicious	72	0.3	19,700	99.7

than 1 % of other document types (letters and non-citable items) are mentioned at least once in Wikipedia. Therefore, Mendeley has the highest coverage of all data sources in this sample, (81.1 % of the review papers and 66.3 % of articles in the sample are covered by Mendeley).

We also studied the total numbers of Mendeley readers, tweets, mentions and bookmarks for each document types covered in the sample (i.e. not only the number of publications with metrics, but the frequency of these metrics). Table 3 shows the result of the total sum and the average number of altmetrics scores per document types provided by the different data sources. Based on both Table 3 and Fig. 2, in general, articles have the highest values of numbers of readers, tweets and bookmarks (more than 77.5 % of all altmetrics scores are to articles), followed[11] by review papers, non-citables and letters (less than 18 % of the altmetrics scores are to the other types) in all data sources. But considering the average metrics per publications,[12] it can be seen that, Mendeley accumulate the most metrics per all document types than all other data sources. Also, in Mendeley, review papers have attracted the most readers per publications (on average there are ∼ 14 readers per review paper) than all other data sources.

Presence of IS altmetrics across NOWT Subject fields

For this analysis, we used the NOWT (High) classification which has 7 major disciplines developed by CWTS.[13] Table 4 shows the percentage of publications having at least one metrics (i.e. papers with at least one reader in Mendeley, once bookmarked in Delicious, once tweeted, or once mentioned in Wikipedia) across those major disciplines.[14] According to the results, Multidisciplinary publications ranked the highest in all data sources. The major source for altmetrics data in our sample is Mendeley with the highest proportion for Multidisciplinary fields, which include journals such as Nature, Science or PNAS. 80 % of the publications in this field, 73 % of the publications from Medical & Life Sciences[15] and 68 % of the publications from Social & Behavioural Sciences have at least

---

[11] In Delicious, articles, non-citables, letters and review papers have the highest number of metrics orderly.

[12] Average metrics per publications calculated by dividing the total numbers of metrics from each data source by total number of publications in the sample. For example, in Mendeley, average number of readers per publication equals to 99,050/19,772 = ∼5.

[13] In the previous study, we used the NOWT (Medium) with 14 subject fileds. For more details see: http://nowt.merit.unu.edu/docs/NOWT-WTI_2010.pdf.

[14] Here publications can belong to multiple subject categories.

[15] According to the Global Research Report by Mendeley (http://www.mendeley.com/global-research-report/#.UjwfTsanqgk), coverage of Mendeley in different subjects are as follows: the highest coverage are by publications from Biological Science & Medicine (31 %), followed by Physical Sciences and Maths (16 %), Engineering & Materials Science (13 %), Computer & Information Science (10 %), Psychology, Linguistics & Education(10 %), Business Administration, Economics & Operation Research (8 %), Law & Other Social Sciences (7 %) and Philosophy, Arts & Literature & other Humanities (5 %).

Table 2 Coverage of publications with different document types by different data sources

Doc type	Pub		Mendeley	%	Twitter	%	Wikipedia	%	Delicious	%
Article	16,740	84.7 %	11,094	66.3	255	1.5	230	1.4	56	0.3
Review	944	4.7 %	766	81.1	32	3.4	43	4.6	7	0.7
Letter	487	2.4 %	122	25.1	7	1.4	4	0.8	3	0.6
Non-citable	1,601	8.0 %	398	24.9	30	1.9	12	0.7	6	0.4
Total	19,772	100	12,380	62.6	324	1.6	289	1.4	72	0.3

one Mendeley reader. Among the other data sources, Multidisciplinary publications ranked the highest as well but with lower presence of publications with metrics. Regarding the top three fields with the highest percentage of altmetrics, Wikipedia has similar pattern as Mendeley: 7 % of the publications from Multidisciplinary field, 2 % of the publications from Medical & Life Sciences and 2 % of the publications from Social & Behavioural Sciences have at least one mention in Wikipedia. In Twitter, 7 % of the publications from Multidisciplinary field, 3 % of the publications from Social & Behavioural Sciences and 2 % of publications from Medical & Life Sciences are the top three fields that have at least one tweet. In Delicious, only 1 % of the publications from Multidisciplinary field, Language, Information & Communication and Social & Behavioural Sciences have at least one bookmark while other fields have less than 1 % altmetrics.

Again, the total scores of Mendeley readers, tweets, mentions and bookmarks for each discipline in the sample have been calculated. Figure 3 shows that the distributions of IS altmetrics across different subject fields is uneven. Both Medical & Life and Natural Sciences received the highest proportion of altmetrics in all data sources. In general in all data sources, more than 30 % of altmetrics accumulated by publications from Medical & Life Sciences and more than 23 % of altmetrics are to publications from the fields of Natural Sciences. Other fields, each received less than 10 % of total altmetrics. Comparing the different data sources in terms of the proportion of altmetrics across fields, different patterns arise: Medical & Life Sciences fields proportionally attracted the most attention in Wikipedia, followed by Mendeley, Twitter and Delicious while in case of Natural Sciences, Delicious, Twitter, Mendeley and Wikipedia, proportionally got the most attention orderly; moreover, for Mendeley, both Social & Behavioural and Engineering Sciences, proportionally, received the highest attention than all other fields (Table 5).

Comparison of citations per papers (CPP) and readerships per papers (RPP) across fields

Although measuring the impact of scholarly publications in social media is very important, it is not yet clear for what purposes scholarly publications are mentioned in social media and reference management tools such as Mendeley, in social bookmark manager such as Delicious, in Wikipedia and Twitter by different users/scholars, and particularly it is not clear if these mentions can be considered as measures of any type of "impact" of the publications. In case of Mendeley, it is assumed that publications are saved in users' libraries for immediate or later reading and possibly also future citation.

In any case, it is important to know how many altmetrics versus citations each publication received and what are the different pattern across different subject fields. Due to the fact that not all of scholarly publications are covered equally by citation databases and also

Table 3 Distribution of IS altmetrics per document types in different data sources

Doc type	pub	Mendeley Readers	%	Avg	Tweets	%	Avg	Wikipedia Mentions	%	Avg	Delicious Bookmarks	%	Avg
Article	16,740	82,553	83.3	4.9	3,020	94.5	0.18	292	77.5	0.02	213	87.3	0.01
Review	944	12,730	12.9	13.4	78	2.4	0.08	68	18.0	0.07	7	2.9	0.01
Non-citable	487	3,301	3.3	2.0	76	2.4	0.05	13	3.4	0.01	14	5.7	0.01
Letter	1,601	466	0.5	0.9	21	0.7	0.04	4	1.1	0.01	10	4.1	0.02
Total	19,772	99,050	100	5.0	3,195	100	0.16	377	100	0.02	244	100	0.01

Table 4 Coverage of publications with different NOWT subject fields by different data sources

NOWT high subject categories	Total number of publications		Mendeley	%	Wikipedia	%	Twitter	%	Delicious	%
Multidisciplinary journals	216	47 %	172	80	15	7	16	7	3	1
Medical & life sciences	15,637	36 %	11,353	73	284	2	301	2	67	0.4
Social & behavioral sciences	1,878	6 %	1,268	68	32	2	58	3	11	1
Natural sciences	11,935	8.7 %	6,554	55	103	1	123	1	34	0.3
Engineering sciences	2,885	0.6 %	1,558	54	7	0.2	9	0.3	2	0.1
Language, information & communication	241	0.7 %	123	51	2	1	1	0.4	3	1
Law, arts & humanities	488	1.5 %	190	39	8	2	7	1	0	0
		100								

Fig. 2 Distribution of IS altmetrics across document types
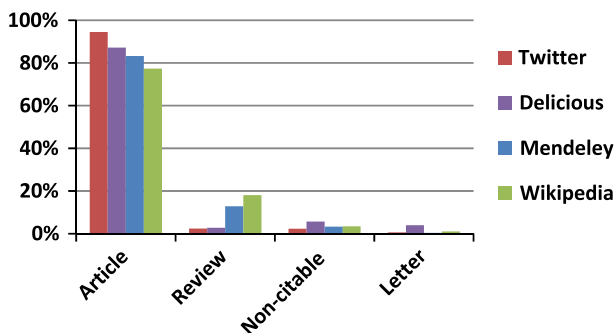
the existence of disciplinary differences in terms of citations, which vary a lot between fields, it is interesting to study both the proportion of altmetrics vs citations per publications to see which fields can benefit from having more density of altmetrics scores (i.e. altmetrics scores per paper) than citation density. Since Twitter, Wikipedia and Delicious showed an overall very low presence per paper, we focus here only on Mendeley. Both the average number of Mendeley RPP and WOS CPP across different NOWT subject fields were calculated and analyzed (Fig. 4). For calculating the citations (excluding self-citations), we used a variable citation window from the year of publication to 2012. Also a variable "readership window" was considered for Mendeley, counting readerships from the publication year of the paper until the last week of April 2013. In this analysis we have also included publications without any metrics (citations or Mendeley readers). The result (Fig. 4 sorted by RPP) shows that in general, Multidisciplinary journals have the highest values of both RPP and CPP; and Law, Arts & Humanities have the lowest values. For fields such as Multidisciplinary journals, Medical & Life Sciences, Natural and Engineering Sciences, the value of CPP is higher than RPP, while for fields such as Social & Behavioural Sciences, Language, Information & Communication and Law, Arts & Humanities, RPP outperforms CPP. The latter is an interesting result that might suggest the relevance of Mendeley for the study of Social Sciences and Humanities publications, which are often not very well represented by citations (Nederhof 2006). In order to further test the differences between RPP and CPP, we extended the same type of analysis for all 248 WOS individual subject categories, resulting that 167 out of 248 WOS subject categories have higher CPP values than RPP values. Most of the fields with higher values of CPP vs. RPP are from the Sciences (145), 18 from the Social Sciences and 4 from the Art and Humanities. On the other hand, 72 fields presented higher RPP than CPP scores (among them 31 are from Social Sciences, 27 from Science and 13 from Art and Humanities).[16] Therefore, we can conclude that citations are more dominant than readerships particularly in the fields of the Sciences (which are also the fields with the highest coverage in citation databases); while on the other hand, many sub-fields from the Social Sciences and Art and humanities received proportionally more RPP than CPP. This could be seen as a possibility for these fields with lower coverage in citation databases (such as WoS) to benefit from Mendeley in terms of having more readership impact than citation impact, although this needs further explorations.

---

[16] For 9 fields (8 fields from Art and Humanities and 1 field from Science) CPP and RPP scores were exactly the same.

Table 5 Distribution of IS altmetrics per NOWT subject fields in different data sources

NOWT subject category	Mendeley Readers		Wikipedia Mentions		Delicious Bookmarks		Tweets	
Medical & life sciences	86,347	50 %	371	64 %	92	31 %	1,958	42 %
Natural sciences	54,481	32 %	136	23 %	186	62 %	2,317	49 %
Social & behavioral sciences	14,102	8 %	35	6 %	12	4 %	112	2 %
Engineering sciences	9,800	6 %	7	1 %	2	1 %	100	2 %
Multidisciplinary journals	4,521	3 %	20	3 %	3	1 %	144	3 %
Language, information & communication	1,492	0.9 %	2	0.3 %	4	1 %	1	0 %
Law, arts & humanities	1,297	0.8 %	13	2 %	0	0 %	72	2 %
	100		100		100		100	

Trend analysis of IS altmetrics across publication years

Table 6 shows the trend analysis of number and share of publications in the sample by altmetrics sources. Regarding the publication years, the share of publications ranges from 10 % in the year 2005 to 18 % in the year 2011. The coverage of different sources is also shown in the table. In our sample, Mendeley has its peak in its proportion of publications with some readers in 2009 (66 %) and the lowest point in 2011 (57 %), although the total number of publications with some Mendeley readerships has increased during the whole period, with the exception of 2011 when there is a small drop compared to 2010. Twitter has its highest peak in 2011 (4 %) and its lowest values in the early years (around 1 % between 2005 and 2009). Wikipedia mentions are for 2 % of all the publications published between 2005 to 2008 and 1 % of all the publications published between 2010 and 2011. For Delicious, the highest peak is for the years 2007 and 2011 and the lowest one for the year 2005, also publications from 2008, 2009 and 2010 have the same presence in Delicious. All in all, it seems that Twitter and Delicious tend to cover the more recent publications better than the older ones although the values are in general very low.

The presence of overall altmetrics scores (i.e. not only publications with altmetrics, but their total counting) has been also calculated in order to know its trend over time. According to Table 7, this is quite different across different data sources. For example, for Wikipedia and Mendeley, publications from the years 2006 and 2009, accumulated most of the mentions (20 %) and readerships (17 %) respectively. In the case of Mendeley and Wikipedia we noticed a decrease in the amount of altmetrics in the last 2 years.

Both in Delicious and in Twitter, publications from the year 2008 received the highest proportion of altmetrics. In case of Delicious, 50 % of bookmarks and in case of Twitter, 34 % of tweets are to publications published in 2008. Comparing the amount of altmetrics in each year across different data sources shows that in this sample, both the oldest and the
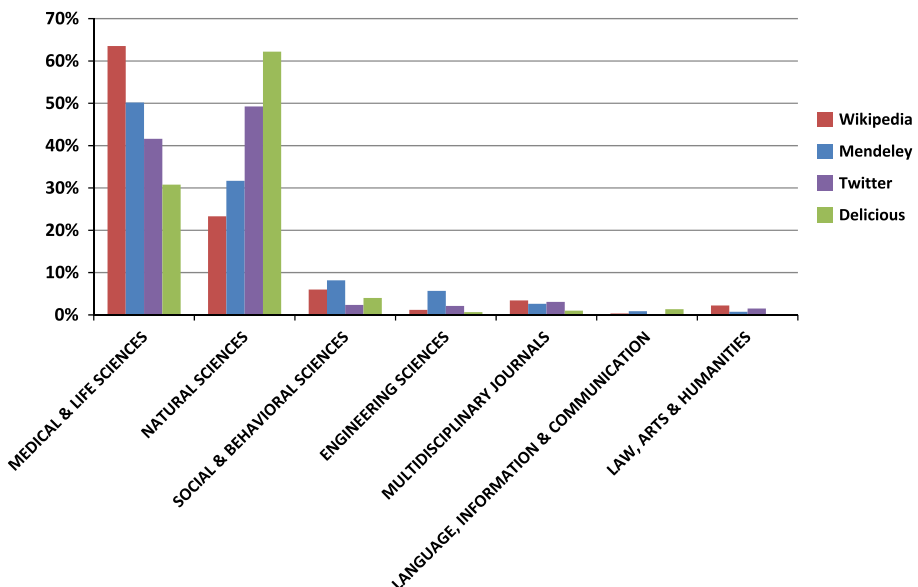

Fig. 3 Distribution of IS altmetrics across NOWT subject fields

Table 6 Coverage of publications with different publication years by different data sources

Pub year	p	%	Mendeley	%	Wikipedia	%	Delicious	%	Twitter	%
2005	2,006	10	1,263	63	39	2	3	0.1	17	1
2006	2,405	12	1,491	62	58	2	4	0.2	6	0.2
2007	2,682	14	1,702	63	41	2	13	0.5	16	1
2008	2,858	14	1,799	63	46	2	11	0.4	34	1
2009	3,039	15	2,001	66	43	1	12	0.4	31	1
2010	3,228	16	2,099	65	37	1	13	0.4	62	2
2011	3,548	18	2,020	57	25	1	16	0.5	158	4

Table 7 Distribution of IS Altmetrics across publication year

Pub year	p	%	Mendeley	%	Wikipedia	%	Delicious	%	Twitter	%
2005	2,006	10	10,814	11	48	13	51	21	835	26
2006	2,405	12	12,658	13	77	20	4	2	20	1
2007	2,682	14	13,739	14	58	15	14	6	102	3
2008	2,858	14	14,299	14	67	18	122	50	1,072	34
2009	3,039	15	16,922	17	50	13	21	9	145	5
2010	3,228	16	16,305	16	43	11	14	6	198	6
2011	3,548	18	14,239	14	34	9	18	7	823	26
		100		100		100		100		100

most recent publications in Twitter have the most altmetrics (tweets) (26 % of tweets are to publications from the year 2005[17] and 2011 respectively) and also the recent publications (2009–2010) have the most altmetrics (readerships) in Mendeley (Fig. 5).

Relationships between IS altmetrics and citation indicators

In this section we study more thoroughly the relationship between the IS altmetrics and citation indicators. Following the CWTS standard calculation of indicators (cf. Waltman et al. 2011), we calculated for all the publications the following citation indicators: citation score (CS), that is, number of citations per publications; normalized citation score (NCS), that is, number of citations per publications, with a normalization for fields differences and publication year; JCS, that is the average number of citations received by all publications in that journal of a publication; and normalized journal score (NJS), that is, the average number of citations received by all publications in that journal normalized by fields differences and publication year. For the calculation of the impact indicators, as explained before, we used a variable citation window (i.e. citations up to 2012) excluding self-citations. The result of the factor analysis, the correlation analysis and impact of publications with and without altmetrics will be presented in the next sections.

---

[17] In 2005, the two most tweeted papers are from the field of Physics, they received more than half of the total tweets in this year (472 tweets), thus showing a strong skewed distribution.
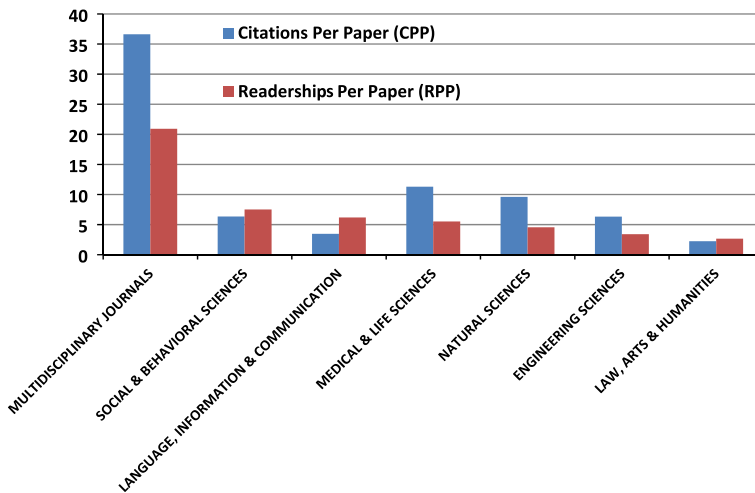
Fig. 4 Comparing CPP and RPP in Mendeley across Subject Fields

Factor analysis of IS altmetrics and bibliometrics indicators

An exploratory factor analysis has been performed using SPSS version 21 in order to know more about the underlying structure, relationship among the variables and the dimension of variables (Table 8). Principal component analysis revealed the presence of 2 main components or dimensions with eigenvalues exceeding 1, explaining 58 % of the total variance. The first dimension is dominated by bibliometric indicators. Mendeley readerships and Wikipedia mentions are also included in this dimension; although Mendeley readership counts has the highest loadings in this dimension of the two indicators. The second dimension is more related to social media metrics, showing that Twitter and Delicious are strongly correlated. These results suggest that the variables in each group may represent similar concepts.

Correlations between IS altmetrics and bibliometrics indicators

In order to overcome the technical limitation of SPSS for calculating Spearman correlation for large datasets,[18] first, rankings of variables computed using Data > rank cases and then Pearson correlation performed on the ranked variables; this method provides the spearman correlation of the original variables. Table 9 shows the result of the correlation analysis among the different altmetrics data source and citation and JCSs and their 95 % confidence intervals (calculated using the Bootstrapping technique implemented in SPSS). According to this table, citation indicators are more correlated between them than with altmetrics. In general, direct citations indicators (i.e. CS and NCS) correlate better among them than with indicators of journal impact (JS and NJS), although the correlations between the two groups are fairly high. Mendeley is correlated with Wikipedia ($r = 0.08$) and Twitter is correlated with Delicious ($r = 0.12$), this is in line with the result of the factor analysis but

---

[18] Calculating Spearman correlation analysis in SPSS for large datasets gives this error: "Too many cases for the available storage", for overcoming this limitation, we followed the process we mentioned in the text. For more details see: http://www.ibm.com/support/docview.wss?uid=swg21476714.

Table 8 Factor analysis of the variables	Rotated component matrix[a]		
		Component	
		1	2
	CS	**0.837**	*0.005*
	NCS	**0.752**	*0.009*
Extraction method principal component analysis. *Rotation Method* Varimax with Kaiser normalization. Loadings higher than 0.1 are shown in bold. 58 % of total variance explained	JS	**0.745**	*−0.011*
	NJS	**0.720**	*−0.015*
	Mendeley	**0.680**	*0.008*
	Wikipedia	**0.297**	*0.009*
[a] Rotation converged in 4 iterations	Delicious	*0.003*	**0.954**
	Twitter	*0.004*	**0.954**

the correlation values are very low. Compared to citation indicators, Mendeley has the highest correlation score with citations (moderate correlation of $r = 0.49$) among all the altmetrics sources. The other altmetric sources show very weak or negligible correlation with citation indicators.

Impact of publications with/without altmetrics

In this section, we study the differences in impact between publications with and without altmetrics. The main idea is to see whether publications with altmetrics tend to have more citation impact than those without altmetrics. Table 10 presents the bibliometric indicators and their 95 % confidence intervals (calculated using the Bootstrapping technique implemented in SPSS). For instance, according to the median values it can be observed that publications with metrics have in general higher CSs compared to those without metrics in all data sources (although, in some cases, the confidence intervals show some overlapping, thus the claim of the higher impact for these cases is less strong and probably more influenced by outliers).

Focusing on the number of Mendeley readers per publication and considering their impact as measured by the NCS and NJS, we can see how publications tend to increase in citation impact as the number of readerships increases (Fig. 6). The effect is quite strong, especially for the average number of citations per publication but this is less prominent for the NJS indicator. The same result found by Waltman and Costas (2013) for relationship between recommendations from F1000, citations and journal impact. In their study, they found that on average, publications with more recommendations also have higher citation and journal impact.

Discussion and conclusions

In this paper we have used IS[19] for gathering altmetrics for a set of randomly sampled publications. IS is an interesting open source for collecting altmetrics, however, we also see some important limitations[20] particularly regarding the speed and capacity of data

---

[19] Impact Story, was in an initial stage of development (i.e. in a 'Beta' version) at the moment of development of this study.

[20] For current limitations of IS see: http://impactstory.org/faq#toc_3_11.

Table 9 Correlation analysis of the rank values of variables

	NCS	JS	NJS	Mendeley	Wikipedia	Delicious	Twitter
CS	0.886 (0.882–0.89)	0.762 (0.756–0.769)	0.557 (0.547–0.567)	0.497 (0.485–0.508)	0.094 (0.08–0.108)	0.011 (−0.005–0.027)	0.025 (0.01–0.039)
NCS		0.528 (0.516–0.538)	0.6 (0.59–0.609)	0.467 (0.455–0.478)	0.074 (0.059–0.087)	0.019 (0.002–0.035)	0.054 (0.037–0.068)
JS			0.711 (0.702–0.718)	0.44 (0.428–0.452)	0.09 (0.075–0.105)	−0.003 (−0.018–0.012)	−0.003 (−0.018–0.011)
NJS				0.427 (0.415–0.439)	0.058 (0.044–0.072)	0.012 (−0.005–0.028)	0.039 (0.023–0.053)
Mendeley					0.083 (0.067–0.099)	0.031 (0.015–0.047)	0.07 (0.055–0.084)
Wikipedia						0.021 (−0.001–0.049)	0.056 (0.025–0.087)
Delicious							0.125 (0.073–0.185)

Table 10 Comparison of NCS and NJS of the publications with and without altmetrics

		With metrics				Without metrics			
		CS	JS	NCS	NJS	CS	JS	NCS	NJS
Mendeley	N	12,380	12,380	12,380	12,380	7,392	7,392	7,392	7,392
	Median	5	6.53	0.72	1.02	1	1.76	0.10	0.53
Confidence interval	Lower	4	6.4	0.69	1.01	0.5	1.67	0.08	0.51
	Upper	5	6.69	0.74	1.04	1	1.89	0.12	0.55
Wikipedia	N	289	289	289	289	19,483	19,483	19,483	19,483
	Median	12	13.87	1.18	1.18	2	4.43	0.47	0.86
Confidence interval	Lower	9	11.91	0.97	1.07	2	4.32	0.46	0.85
	Upper	14	15.2	1.35	1.31	3	4.57	0.49	0.87
Twitter	N	324	324	324	324	19,448	19,448	19,448	19,448
	Median	4	3.6	1	1.1	3	4.53	0.47	0.86
Confidence interval	Lower	3	3.1	0.86	0.97	2	4.39	0.46	0.85
	Upper	5	4.74	1.29	1.27	3	4.62	0.49	0.87
Delicious	N	72	72	72	72	19,700	19,700	19,700	19,700
	Median	3	3.99	0.89	1.07	3	4.52	0.48	0.86
Confidence interval	Lower	2	2.34	0.52	0.76	2	4.38	0.46	0.85
	Upper	6	5.55	1.57	1.33	3	4.62	0.49	0.87

Fig. 5 Distribution of IS altmetrics across publication years
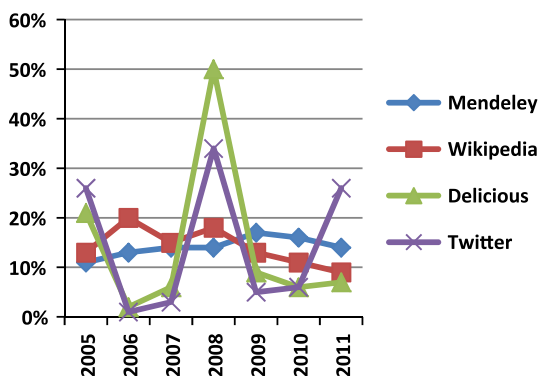
collection and formatting of the data. We detect different results comparing our current results with those presented in our previous study (Zahedi et al. 2013) mostly due to the different methodology of data collection (manually vs. automatically) and collecting the data at different points in time as it happened between our two studies, where in the first one, Mendeley was only presented in around 37 % of the publications[21] and now in more than 60 %.[22] This situation also points to the need for the tools to be transparent in how

---

[21] The time interval between the first and the second data collection was 6 months and data collection done manually versus the second one which done automatically using RESTAPI calls.

[22] Reasons for these differences can be the changes/improvements in the identification of publications by Mendeley (e.g. by merging version of the same paper, identifying more DOIs, increments in the number of users in Mendeley, etc).
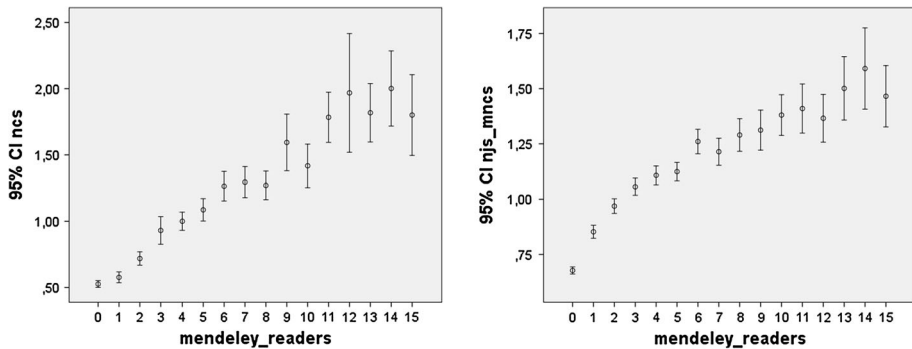
Fig. 6 Relation between number of Mendeley readerships and citation and journal impact

their data are collected and their limitations. This means that an important natural future step will be the proper assessment of the validity of the data retrieved via different altmetrics data sources (as it has been done for example for Google Scholar—cf. Delgado López-Cózar et al. 2012). This validation of the quality, reliability and robustness of the altmetrics tools is essential in order to be able to apply altmetrics for serious research assessment purposes. For these tools to be fully incorporated in regular research assessment processes, they need to meet the necessary requirements for data quality, transparency and indicator reliability and validity as emphasized by Wouters and Costas (2012) in their study of altmetric tools. Moreover, the results of this study are based on the WOS covered publications; hence, it is important to keep in mind the restrictions of this database with regards to its coverage of some fields, language and publication formats (Moed 2009; Van Raan et al. 2011; Archambault and Larivière 2006; Torres-Salinas et al. 2013b).

All in all, given the exploratory nature and the fact that basically the same results have been found with the two data collections, we can assume that our results are robust and valid for our purposes. In general, our study shows that Mendeley is the major and more useful source for altmetrics data. Mendeley has the highest coverage and proportion of altmetrics compared to Twitter, Wikipedia and Delicious for the studied publications. Out of 19,772 publications a total 12,380 cases (62.6 %) had at least one reader in Mendeley. Previous studies also showed that Mendeley is the most exhaustive altmetrics data source (Bar-Ilan et al. 2012; Priem et al. 2012) mostly for the publications from Library and Information Science field: 97.2 % coverage for JASIST articles published between 2001 and 2011 (Bar-Ilan 2012); 82 % coverage for articles published by researchers in Scientometrics (Bar-Ilan et al. 2012); and 82 % of bibliometrics literature (Haustein et al. 2013), for Multidisciplinary journals such as Nature and Science (94 and 93 % of articles published in these journals in 2007) (Li et al. 2012); and more than 80 % of PLoS ONE publications (Priem et al. 2012) covered by Mendeley. In terms of document type, review papers and articles were proportionally the most read, shared, liked or bookmarked format compared to non-citable items and letters across all data sources. Multidisciplinary fields (i.e. the field where journals such as *Nature*, *Science* or the *PNAS* are included) are the most present in all altmetrics data sources but concerning the distribution of altmetrics across different fields, more than 30 % of altmetrics accumulated by publications from Medical & Life Sciences and more than 23 % of altmetrics are to publications from the fields of Natural Sciences. Comparing both proportion and distribution of IS altmetrics across different fields among different data sources shows different patterns, particularly in

Mendeley, both Social & Behavioural and Engineering Sciences, have proportionally received the highest attention compared to all other fields. Considering citations and readerships per publication, Multidisciplinary journals have the highest and Law, Arts & Humanities have the lowest density of both citations and readerships per publications. However, according to our observation, there is a higher density of RPP than CPP in several fields of the Social Sciences and Humanities. This finding suggests that Mendeley readership counts could have some added value in supporting the evaluation and analysis of these fields, which have been traditionally worse represented by citation indicators (cf. Nederhof 2006). Another explanation for those fields with lower proportion of readers than citations could be the fact that Mendeley is relatively new and not yet widely used and adopted among all scholars from all the disciplines. Besides, differences in citation and readership behaviors and practices among fields could also explain these differences. In any case, this is an aspect that needs further analysis.

Our trend analysis shows that particularly publications with Mendeley readerships have increased over time, although there is a slight decrease in the number of readerships and proportion of publications with Mendeley readers for the last two years. The most plausible explanation for this is that the accumulation of readers takes some time. To the best of our knowledge there is no information on the 'readership history' of publications (besides the fact that readerships could conceptually decrease as the users delete or change their libraries) and so far we don't have results on the readerships pace. This means that we don't know when a paper in a given year has obtained its peak in readerships. It is highly likely, that although faster than citations, the accumulation of readerships for publications also takes some time, and this is the reason why for the most recent publications, the number of readers is slower as compared to those older publications that have had more time to accumulate readerships. Future research should also focus on disentangling this aspect.

The Spearman correlation of Mendeley readerships with citation impact indicators showed moderate correlations ($r = 0.49$) between the two variables which is also found in other previous studies (Bar-Ilan 2012; Priem et al. 2012). This indicates that reading and citing are related activities, although still different activities that would be worthwhile to explore. According to the result of comparing the impact of publications with and without altmetrics with their CSs, it can be also concluded that in general, publications with more altmetrics also tend to have both higher direct citations and are published in journals of higher impact. The issue about the potential predictability of citations through altmetric scores will be explored in follow-up research.

Finally, although citations and altmetrics (particularly Mendeley readerships) exhibit a moderate positive relationship, it is not yet clear what the quality of the altmetrics data is and neither what kind of dimension of impact they could represent. Since altmetrics is still in its infancy, at the moment, we don't yet have a clear definition of the possible meanings of altmetric scores. In other words, the key question of what altmetrics mean is still unanswered. From this perspective, it is also necessary to know the motivations behind using these data sources, for example in case of Mendeley: what does it reflect when an item is saved/added by several users to their libraries? Also, what does it mean that an item is mentioned in Wikipedia, CiteULike, Twitter and any other social media platform? Does it refer to the same or different dimension compared to citation? In the same line, besides studying to what extent different publications are presented in Mendeley and other social media tools and their relations with citation impact, we need to study for what purposes and why these platforms are exactly used by different scholars. Moreover, research about the quality and reliability of the altmetric data retrieved by the different altmetrics

providers is still necessary before any interpretation and potential real uses for these data and indicators are developed. This information in combination with the assessment of the validity and reliability of altmetrics data and tools will shed more light on the meanings of altmetrics and can help to unravel the hidden dimensions of altmetrics in future studies.

References

Archambault, É., & Larivière, V. (2006). *The limits of bibliometrics for the analysis of the social sciences and humanities literature, International Social Science Council: World social sciences report 2010: Knowledge divides* (pp. 251–254). Paris: UNESCO.

Armbruster, C. (2007). Access, usage and citation metrics: what function for digital libraries and repositories in research evaluation? Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1088453.

Bar-Ilan, J. (2012). *JASIST@ Mendeley*. Presented at ACM Web Science Conference Workshop on Altmetrics, Evanston, IL. Retrieved June 21, 2012 from http://altmetrics.org/altmetrics12/bar-ilan/.

Bar-Ilan, J., Haustein, S., Peters, I., Priem, J., Shema, H., & Terliesner, J. (2012). Beyond citations: Scholars' visibility on the social Web. In *Proceedings of the 17th International Conference on Science and Technology Indicators, Montreal, Quebec*. Retrieved from http://arxiv.org/abs/1205.5611/.

Benos, D. J., Bashari, E., Chaves, J. M., et al. (2007). The ups and downs of peer review. *Advances in Physiology Education, 31*, 145–152.

Blecic, D. (1999). Measurements of journal use: An analysis of the correlations between three methods. *Bulletin of the Medical Library Association, 87*, 20–25.

Bollen, J., Van de Sompel, H., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLoS ONE, 4*(6), e6022. doi:10.1371/journal.pone.0006022.

Bollen, J., Van de Sompel, H., & Rodriguez, M. A. (2008). Towards usage-based impact metrics. In *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital libraries (JCDL), New York, USA*. Retrieved from http://arxiv.org/pdf/0804.3791v1.pdf.

Bordons, M., Fernandez, M. T., & Gomez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance. *Scientometrics, 53*, 195–206.

Bornmann, L. (2013). Is there currently a scientific revolution in Scientometrics? *Journal of the American Society for Information Science & Technology*. Retrieved from www.lutz-bornmann.de/icons/impactrevolution.pdf.

Bornmann, L., & Leydesdorff, L. (2013). The validation of (advanced) bibliometric indicators through peer assessments: a comparative study using data from InCites and F1000. *Journal of Informetrics, 7*(2), 286–291.

Brody, T., Harnad, S., & Carr, L. (2006). Earlier Web usage statistics as predictors of later citation impact. *Journal of the American Society for Information Science, 57*, 1060–1072.

Butler, L., & McAllister, I. (2011). Evaluating university research performance using metrics. *European Political Science, 10*(1), 44–58.

Davis, P. M. (2012). Tweets, and our obsession with alt metrics. The Scholarly Kitchen. Retrieved from http://scholarlykitchen.sspnet.org/2012/01/04/tweets-and-our-obsession-with-alt-metrics/.

Duy, J., & Vaughan, L. (2006). Can electronic journal usage data replace citation data as a measure of journal use? An empirical examination. *The Journal of Academic Librarianship, 32*(5), 512–517.

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research, 13*(4), e123.

Forta, B. (2008). *Sams teach yourself SQL in 10 minutes*. USA: Sams Publishing.

Galligan, F., & Dyas-Correia, S. (2013). Altmetrics: Rethinking the way we measure. *Serials Review, 39*(1), 56–61.

Haustein, S. (2010). Multidimensional journal evaluation. In *Proceedings of the 11th International Conference on Science and Technology Indicators* (pp. 120–122), Leiden, the Netherlands.

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2013). Coverage and adoption of altmetrics sources in the bibliometric community. In *14th International Society of Scientometrics and Informetrics Conference* (pp. 1–12). Digital Libraries. Retrieved from http://arxiv.org/abs/1304.7300.

Haustein, S., & Siebenlist, T. (2011). Applying social bookmarking data to evaluate journal usage. *Journal of Informetrics, 5*, 446–457.

Henning, V. (2010). The top 10 journal articles published in 2009 by readership on Mendeley. Retrieved from http://www.mendeley.com/blog/academic-features/the-top-10-journal-articles-published-in-2009-by-readership-on-mendeley/.

Hicks, D. & Melkers, J. (2012). Bibliometrics as a Tool for Research Evaluation. In Al Link & Nick Vornatas (Eds.) *Handbook on the Theory and Practice of Program Evaluation*. Edward Elgar. Retrieved from http://works.bepress.com/diana_hicks/31.

Li, X., & Thelwall, M. (2012). F1000, Mendeley and traditional bibliometric indicators. *Proceedings of the 17th International Conference on Science and Technology Indicators* (pp. 451–551). Canada: Montréal.

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics, 91*(2), 461–471.

Lopez-Cozar, E. D., Robinson-Garcia, N., & Torres Salinas, D. (2012). Manipulating google scholar citations and google scholar metrics: simple, easy and tempting. Retrieved from http://arxiv.org/abs/1212.0638.

MacRoberts, M. H., & MacRoberts, B. R. (1989). Problems of citation analysis: A critical review. *Journal of the American Society for Information Science, 40*, 342–349.

Martin, B. R., & Irvine, J. (1983). Assessing basic research: Some partial indicators of scientific progress in radio astronomy. *Research Policy, 12*, 61–90.

Moed, H. F. (2005). *Citation analysis in research evaluation*. Berlin: Springer.

Moed, H. F. (2007). The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Science and Public Policy, 34*(8), 575–583.

Moed, H. F. (2009). New developments in the use of citation analysis in research evaluation. *Archivum immunologiae et therapiae experimentalis, 57*(1), 13–18.

Nederhof, A. J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics, 66*(1), 81–100.

Nederhof, A. J., & Van Raan, A. F. J. (1987). Peer review and bibliometric indicators of scientific performance: a comparison of cum laude doctorates with ordinary doctorates in physics. *Scientometrics, 11*(5–6), 333–350.

Nicolaisen, J. (2007). Citation analysis. *Annual Review of Information Science and Technology, 41*, 609–641.

Priem, J., Hemminger, B. H. (2010). Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. Retrieved First Monday 15, from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2874/2570.

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

Priem, J., Taraborelli, D., Groth, P., and Neylon, C. (2010). Altmetrics: a manifesto. Retrieved from http://altmetrics.org/manifesto/.

Rousseau, R., & Ye, F. (2013). A multi-metric approach for research evaluation. Chinese Science Bulletin, 10–12.

Rowlands, I., & Nicholas, D. (2007). The missing link: Journal usage metrics. *Aslib Proceedings, 59*(3), 222–228.

Schlögl, C., Gorraiz, J., Gumpenberger, C., Jack, K., & Kraker, P. (2013). Download vs. citation vs. readership data: the case of an information systems journals. In J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, & H. Moed (Eds.), *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference, Vienna, Austria* (pp. 626-634). Wien: Facultas Verlags und Buchhandels AG.

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, 314–497.

Shema, H., Bar-Ilan, J., & Thelwall, M. (2013). Do blog citations correlate with a higher number of future citations ? Research blogs as a potential source for alternative metrics. In J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, & H. Moed (Eds.), *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference, Vienna, Austria* (pp. 604–611). Wien: Facultas Verlags und Buchhandels AG.

Shuai, X., Pepe, A., & Bollen., J. (2012). How the scientific community reacts to newly submitted preprints: Article downloads Twitter mentions, and citations. Retrieved from http://arxiv.org/abs/1202.2461v1.

Smith, A. G. (1999). A tale of two web spaces; comparing sites using web impact factors. *Journal of Documentation, 55*(5), 577–592.

Taylor, J. (2011). The assessment of research quality in UK. Universities: peer review or metrics? *British Journal of Management, 22*(2), 202–217.

Thelwall, M. (2001). Extracting macroscopic information from web links. *Journal of American Society for Information Science and Technology, 52*(13), 1157–1168.

Thelwall, M. (2004). Weak benchmarking indicators for formative and semi-evaluative assessment of research. *Research Evaluation, 13*(1), 63–68.

Thelwall, M. (2008). Bibliometrics to Webometrics. *Journal of Information Science, 34*(4), 605–621.

Thelwall, M. (2012a). A history of webometrics. *Bulletin of the American Society for Information Science and Technology, 38*(6), 18–23.

Thelwall, M. (2012b). Journal impact evaluation: a webometric perspective. *Scientometrics, 92*(2), 429–441.

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, A., & Jimenez-Contreras, E. (2013a). Altmetrics: New indicators for scientific communication in web 2.0. Comunicar. Retrieved from http://arxiv.org/ftp/arxiv/papers/1306/1306.6595.pdf.

Torres-Salinas, D., Robinson-Garcia, N., Campanario, J. M., & López-Cózar, E. D. (2013b). Coverage, field specialisation and the impact of scientific publishers indexed in the book citation index. *Online Information Review, 38*(1), 24–42.

Van Raan, A. F. J., Van Leeuwen, T. N., & Visser, M. S. (2011). Severe language effect in university rankings: particularly Germany and France are wronged in citation-based rankings. *Scientometrics, 88*(2), 495–498.

Vaughan, L., & Shaw, D. (2003). Bibliographic and web citations: what is the difference? *Journal of the American Society for Information Science and Technology, 54*(14), 1313–1322.

Waltman, L., & Costas, R. (2013). F1000 Recommendations as a potential new data source for research evaluation: a comparison with citations. *Journal of the Association for Information Science and Technology*. doi: 10.1002/asi.23040.

Waltman, L., Van Eck, N. J., Van Leeuwen, T. N., Visser, M. S., & Van Raan, A. F. J. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics, 5*(1), 37–47.

Wouters, P. (1999). The Citation Culture, Ph.D. Thesis, University of Amsterdam.

Wouters, P., Costas, R. (2012). Users, narcissism and control: Tracking the impact of scholarly publications in the 21st century. Utrecht: SURF foundation. Retrieved from http://www.surffoundation.nl/nl/publicaties/Documents/Users%20narcissism%20and%20control.pdf.

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 J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger, & H. Moed (Eds.), *Proceedings of the 14th International Society of Scientometrics and Informetrics Conference, Vienna, Austria* (pp. 876–884). Wien: Facultas Verlags und Buchhandels AG.

Zhang, Y. (2012). Comparison of select reference management tools. *Medical Reference Services Quarterly, 31*(1), 45–60.