RESEARCH

Differences between Altmetric Data Sources – A Case Study

Judit Bar-Ilan[1], Gali Halevi[2] and Staša Milojević[3]

[1] Department of Information Science, Bar-Ilan University, Ramat Gan, IL

[2] Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, US

[3] School of Informatics, Computing and Engineering, University of Indiana, Bloomington, IN, US

Corresponding author: Judit Bar-Ilan (Judit.Bar-Ilan@biu.ac.il)

This paper examines the data accuracy and number of altmetric counts reported by Mendeley, altmetric. com and PlumX at two points in time: June 2017 and April 2018 for the dataset of 2,728 articles and reviews published in JASIST between 2001 and 2014. The findings show growth in the number of citations and Mendeley readers over time. In addition, the results also show that there are differences in the altmetric counts reported by the two-altmetric aggregators Altmetric.com and PlumX.

Introduction

Since its inception in 2010 (Priem et al., 2010) altmetrics has been actively promoted as a new set of indicators appropriate to be used for evaluating and capturing the broader impact of scholarly output. Through the years, several studies questioning the meaning of altmetric indicators and what they actually measure were published (e.g., Rasmussen & Andersen, 2013; Haustein, Bowman, & Costas, 2016). One of the major challenges with the use and interpretation of altmetrics indicators is directly related to questionable data quality (Haustein, 2016) and high dependency on commercial providers (aggregators) of altmetric data (Costas, Zahedi, & Wouters, 2015).

Readership counts most often measured by the number of document-saving events on the reference manager Mendeley, have the greatest coverage among all the altmetric indicators (Zahedi, Costas, & Wouters, 2013). Not only is this one of the most prevalent altmetric indicators currently captured (Haustein, Bowman, & Costas, 2016), but it is also the one that most highly correlates with citation counts (around .5) (Mohammadi & Thelwall., 2014; Haustein et al., 2014).

Altmetrics aggregators including the two most often used ones; Altmetric.com and PlumX report Mendeley readership. Considering data accuracy issues and the importance of Mendeley readership data as an indicator, it is important to examine the coverage of readership counts across the data source itself (Mendeley) and two major aggregators (Altmetric and Plum X). Examples of previous reliability studies are (Zahedi, Fenner, & Costas, 2014; Ortega, 2018).

Given that indicators are only as good as the data they are based on, it is not surprising that similar questions were asked in previous studies when comparing citation counts in bibliometric databases (WoS, Scopus, and Google Scholar) (Bar-Ilan, 2008; Halevi, Moed & Bar-Ilan, 2017; Trapp 2016; Boeker, Vach, and Motschall, 2013; Harzing and Alakangas, 2016; Bramer, Giustini, and Kramer, 2016). These studies have shown that there are considerable differences between the numbers reported by the databases, primarily due to the differences in coverage, types of documents covered, and errors.

Mendeley is an online reference manager, reporting the number of users who downloaded an item to their Mendeley libraries ('readers'). Altmetric.com and PlumX are aggregators that report multiple altmetrics, including the number of tweets, blogs, Wikipedia and news mentions. They also report Mendeley reader counts. Altmetric.com and PlumX might report different altmetric scores for a number of reasons. First, they might cover different news and blog sources. In addition, they might use different ways to identify mentions (e.g., by DOI, PMID, arXiv id or title-author-source-publication year). Finally, they might have different schedules for updating from the primary data sources (Zahedi et al., 2015; Ortega, 2018).

Despite of its comprehensiveness, the Mendeley database has some inherent problems that affect the data it generates. First, the users drive the database. Each user identifies a publication of interest and adds it to their personal libraries. This, in turn, makes the data more prone to errors than third party, curated databases. Secondly, there are often

errors that can be found in the metadata fields, which prevent Mendeley to correctly aggregate reader counts, and quite often there is more than one record for a given item. Finally, when using the Mendeley API to retrieve reader counts by DOI, only a single record is retrieved (which is not necessarily the record with the highest number of readers), thus the numbers reported by the aggregators might be an underestimate. Mendeley reorganizes and cleans its database from time to time (Gunn, 2016), which might result in a decrease in the number of readers reported. The Mendeley API allows to search by several fields, such as author, title, publication source, publication year and the like, but since the metadata is entered by users who do not use a predefined format, often multiple records are created for the same item. In addition, because of possible metadata errors the search query might miss records for the given item. When querying the API by article title, Mendeley often returns multiple records for the same article. Finally, special characters in the text fields do not render well in Mendeley.

The aim of this study was, therefore, to examine the data accuracy and number of altmetric counts reported by Mendeley, Altmetric.com and PlumX at two points in time: in June 2017 and in April 2018 and to compare the reported altmetrics at each data collection point. The research questions are as follows:

1. Do all platforms report the same counts when data is downloaded on the same day?
2. What are the changes in the counts over time?

Data Collection

To answer the research questions, we used JASIST (Journal of the American Society for Information and Technology, between 2001 and 2013, and Journal of the Association for Information Science and Technology from 2014 and onwards) articles and reviews published between 2010 and mid 2017 (issues 1 to 7). The initial data collection took place on June 29, 2017, and the second round of data collection on March 29, 2018. The dataset is comprised of 2,666 articles and 62 reviews, altogether 2,728 items (referred to as 'articles' or 'documents' from this point onward). Results from the first data collection point were presented at the Altmetrics17 Workshop (Bar-Ilan & Halevi, 2017).

Data from Mendeley and Altmetric.com were collected with the help of the *Webometric Analyst* tool developed by Mike Thelwall (http://lexiurl.wlv.ac.uk/). Mendeley was searched both by DOI and by title query. The title queries underwent data cleansing and aggregation of reader counts from multiple records of the same article. The main reason that data cleansing was necessary is due to the fact that when searching Mendeley using the titles it often results in multiple records, which, at times, might not be the ones searched for. In addition, there are cases where there are multiple records of the same item that need to be aggregated. Data from PlumX were downloaded from the PlumX dashboard, which has readily available downloading functionality. Both Altmetric and PlumX use DOIs as a primary source for data collection.

Data Analysis

Mendeley Data Coverage and Reader Counts

Following a search by title, aggregating and cleansing the data from the Mendeley API, using Mike Thelwall's *Webometric Analyst* we found that Mendeley had reader counts for 2,628 publications (96.3%) in the dataset in 2017, and for 2,717 articles in 2018 (99.6%). In 2018 we conducted Mendeley reader count searches also by DOI (which is what the aggregators do), and retrieved reader counts for 2,690 documents (98.3% coverage). See **Table 1**.

As **Table 1** shows, Altmetric.com reported Mendeley reader counts for 1,124 articles (40.8%), while PlumX reported reader counts 1,721 articles (63.1%) in 2017. This difference could be due to the fact that Altmetric.com records Mendeley readership counts only if there is at least one additional altmetric indicator for the document. Therefore, if

Table 1: Readership counts reported by the sources in 2017 and 2018 M – Mendeley, MA – Mendeley reader counts from Altmetric.com, MP – Mendeley reader counts from PlumX, Mtitle – reader counts retrieved for a title query from Mendeley, Mdoi – reader counts retrieved for a DOI query from Mendeley. The columns are coloured by the year of data collection.

	Mtitle17	MA17	MP17	Mdoi18	Mtitle18	MA18	MP18
sum of readers	82,040	36,678	47,617	87,917	97,697	45,555	81,449
articles with readers	2,628	1,113	1,721	2,690	2,716	1,156	2,375
% of total	96.3%	40.8%	63.1%	98.6%	99.6%	42.4%	87.1%
average # readers	31.22	32.95	27.67	32.68	35.97	39.41	34.29
std	43.34	51.32	35.28	46.02	50.31	59.52	50.60
median	18	17	17	22	22	22	20
max	732	727	361	897	899	897	916

nothing was tracked in addition to a Mendeley 'read', Altmetric.com will not count it. This can explain the lower coverage of Altmetric.com compared to PlumX. Unfortunately, we could not find a clear statement on how PlumX collects Mendeley reader counts.

In 2018 Altmetric.com reported Mendeley reader counts for 1,157 documents (42.4%), and PlumX for 2,376 (87.1%). A possible reason for the vast increase in PlumX's Mendeley coverage is that Elsevier now owns both Mendeley and PlumX. Elsevier bought Mendeley in April 2103 (Elsevier, 2013) and PlumX was acquired in February 2017 (Michalek, 2017). The integration of PlumX with Elsevier content might have taken some time therefore not having real influence on PlumX coverage data in June 2017. Now that it is well integrated, and both Mendeley and PlumX metrics are displayed on Scopus (previously the Altmetric.com donut was displayed on Scopus) an increase in Mendeley reads in PlumX counts is noted.

As also can be seen from **Table 1**, there is an increase in the total number of readers overall. Some of the growth can probably be contributed to increased coverage, but as can be deduced from the average, median and maximum counts there is an increase in the overall number of readers over time.

While the counts reported in **Table 1** indicate that there was a growth, we also wanted to test whether this growth is statistically significant. **Figure 1** shows the percentages of articles contained in different databases in 2017 and in 2018. Each data point has an error bar showing 68% significance range based on Poisson statistics. Only the change in PlumX coverage is statistically significant ($p < 0.0001$, whereas it is 0.04 and 0.11 for Mendeley and Altmetric). This can further confirm the growth of importance of PlumX as an altmetric aggregator.

Coverage overlap between Mendeley, Altmetric.com and PlumX

Figure 2a and **2b** show the number of documents for which Mendeley readership counts were reported by the three sources in 2017 and 2018 respectively. The overlap increased considerably, with only five documents with readers from Altmetric.com and PlumX not located by Mendeley in 2018. Since Mendeley is the data source for all three this number should have been 0. However, one should note the great improvement in coverage compared to 2017, where there were 65 such documents. The intersection (items with reader counts reported by all three sources) also increased from 804 to 1,021. One should also note the increased intersection between Mendeley and PlumX.

Figure 3 displays the distribution of the differences in the Mendeley reader counts per article in 2018 and 2017, as retrieved from Mendeley. Even in the primary data, one can observe that there are cases where the number of readers decreases (but only 1.4% of articles). This is possibly due to periodical rebuilds in Mendeley that include data aggregation and cleansing (Gunn, 2014).

Figures 4 and **5** display the differences in Mendeley readership counts for articles that had at least one Mendeley reader in at least one of the two data sources compared – Altmetric and PlumX. As can be observed, in most cases the differences are small despite of some outliers. In some cases, the difference is negative, i.e. the aggregator reported higher number of readers than the source (Mendeley). For Altmetric.com 80% of articles show no difference in counts,
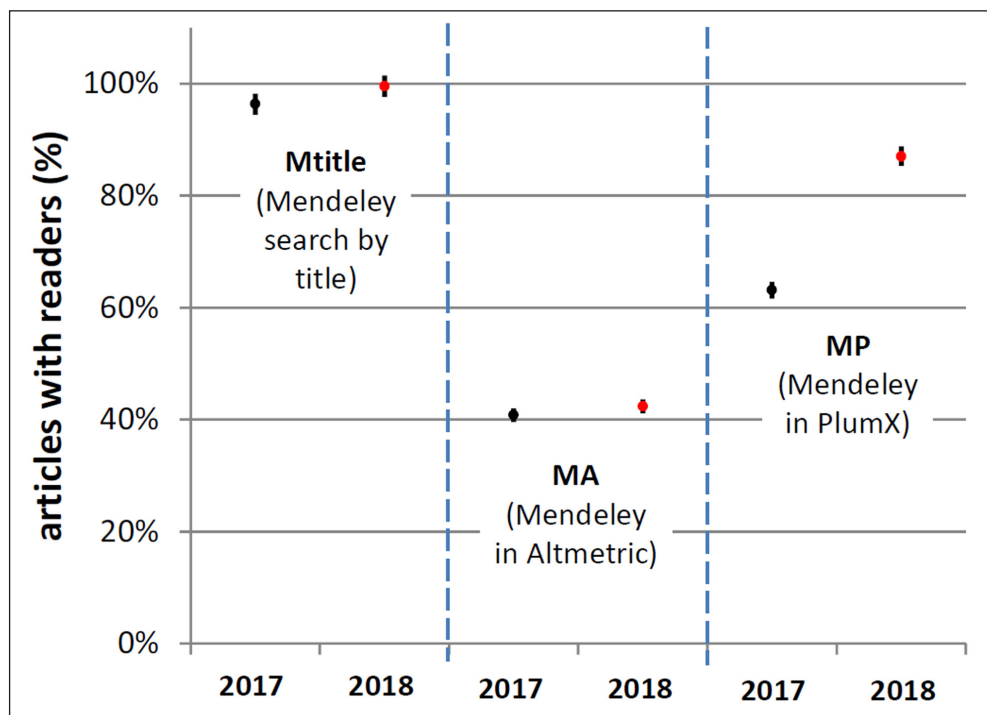
Figure 1: The percentages of articles with Mendeley readers reported by Mendeley, Altmetric.com and PlumX in 2017 and in 2018.
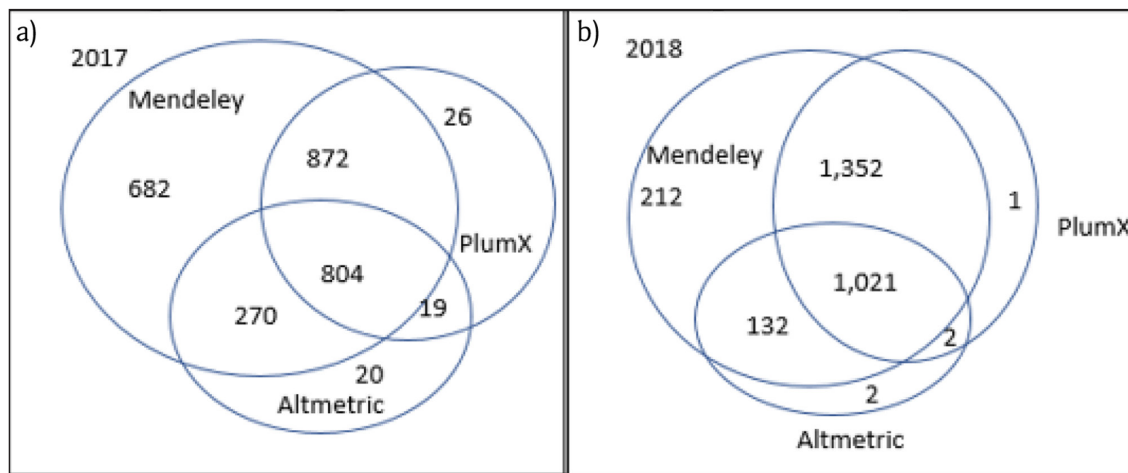
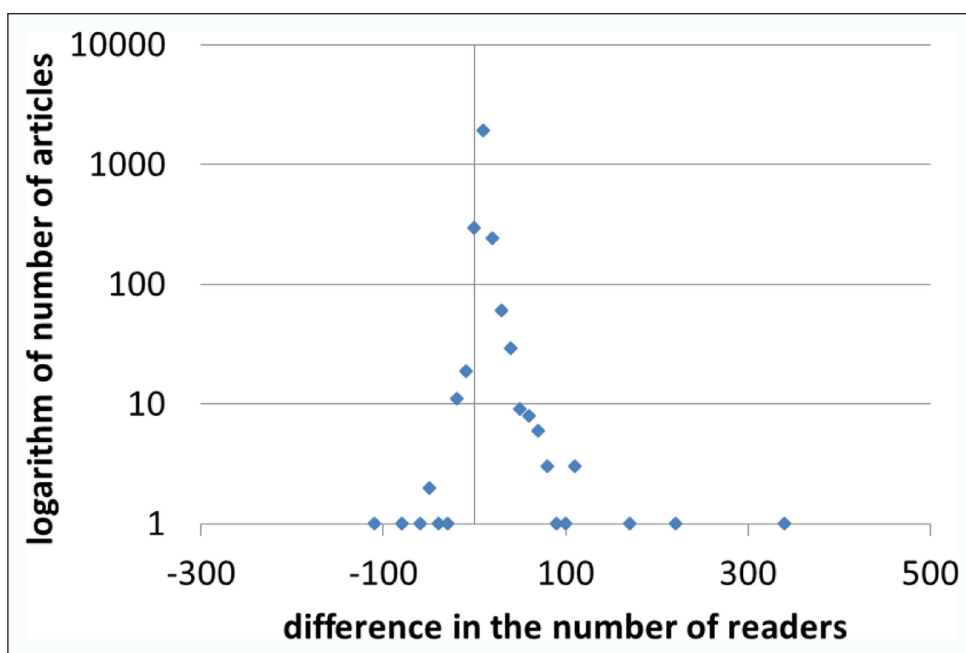**Figure 2:** Overlap in coverage between Mendeley, Altmetric.com and PlumX.

Figure 3: The distribution of the differences (bin size = 10) in the number of readers of individual articles between Mendeley (searched by title) in 2017, and Mendeley (searched by title) in 2018. The *y* axis is in logarithms because of the disproportionate number of articles with a small difference.

whereas 91% of articles have a difference of less than 10 readers. For PlumX, 73% have no difference, and for 91% it is within 10 readers.

The largest number of readers for each of the data sources both in 2017 and in 2018, except for PlumX in 2017, was the article 'The sharing economy: Why people participate in collaborative consumption' by Hamari, Sjoklint and Ukkonen published in 2016. This article had no Mendeley reader counts on PlumX in 2017. In 2017, the article 'CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature' authored by Chen and published in 2006 saw the largest reader counts on PlumX.

Twitter, Blogs, Wikipedia and Mainstream News – Almetric vs. PlumX
Our data shows reasonable Twitter activity for articles published from 2012 and onwards although Twitter was launched 6 years earlier, in 2006. This time gap could be due to the fact that it took several years for researchers to harness Twitter as a scientific communications tool. Therefore, for the purpose of Twitter mentions analysis, we considered only a subset of 1091 articles published between 2012 and mid-2017.

In the same manner we examined Mendeley readership overlaps, we also examined overlaps in Twitter coverage. **Figure 6a** and **6b** display the overlap in 2017 and 2018 respectively. Unlike Mendeley, the overlap between Altmetric and PlumX is large, and gets larger over time.
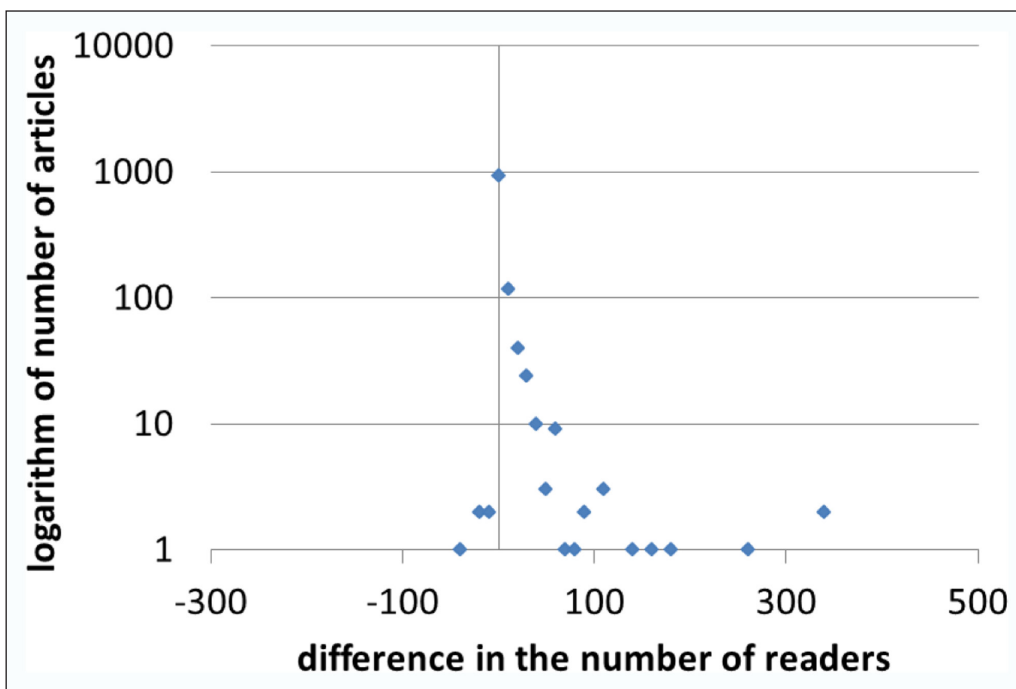
Figure 4: The distribution of the differences (bin size = 10) in the number of readers of individual articles between Mendeley (searched by title) and Mendeley in Altmetric in 2018.
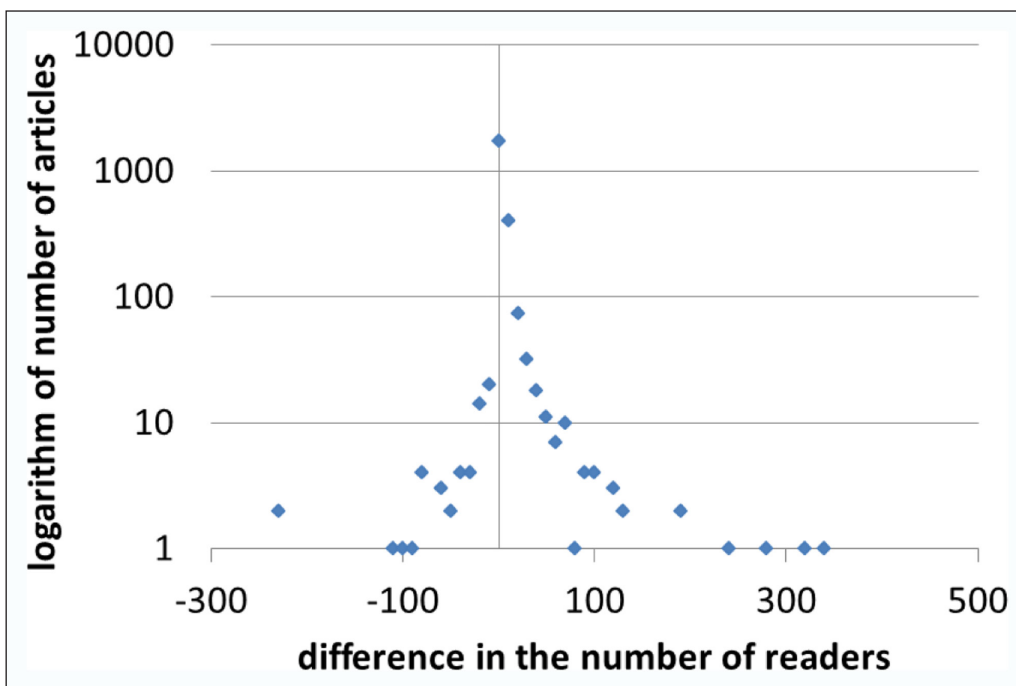

Figure 5: The distribution of the differences (bin size = 10) in the number of readers of individual articles between Mendeley (searched by title) and Mendeley in PlumX in 2018.

As can be seen in **Figure 7**, the average number of tweets per article remained more or less stable for Altmetric. com whiles the average number of tweets per article decreased for PlumX. Interesting to note that almost for all years of publication, the average number of tweets reported by PlumX.in 2018 was lower than in 2017. This can be explained by the increased coverage of twitter by PlumX in 2018 tracking fewer tweets for the newly discovered documents.

Table 2 in the appendix shows that both aggregators provided rather high Twitter coverage in the dataset, between 66% to 78%, This is contrary to previous studies which reported no more than 25% Twitter coverage of their datasets. For example, Zahedi, Costas and Wouters, (2015), studied altmetric coverage of a large dataset (more than half million articles published from 2011 and onwards) from Altmetric.com and found 13% coverage for Twitter while Thelwall,
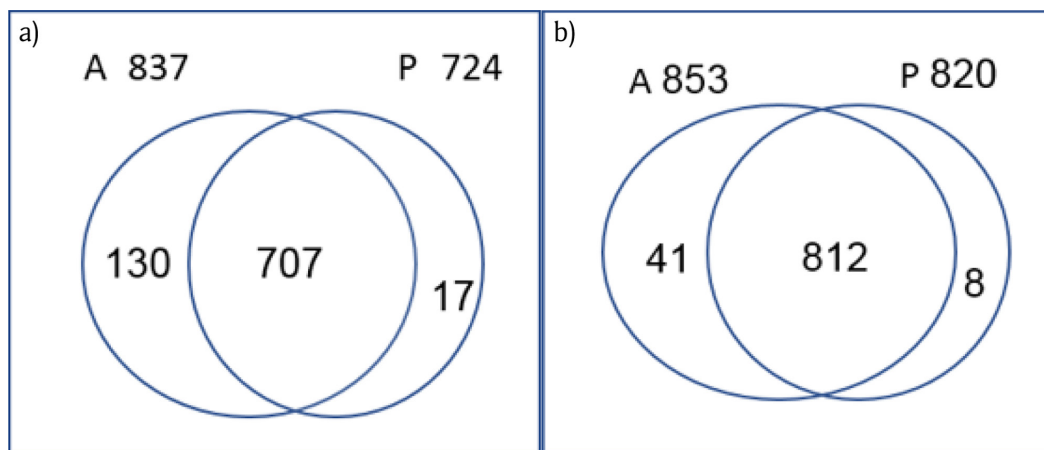
Figure 6: Overlap in coverage between Altmetric.com and PlumX for Twitter for publications between 2012 and 2017.
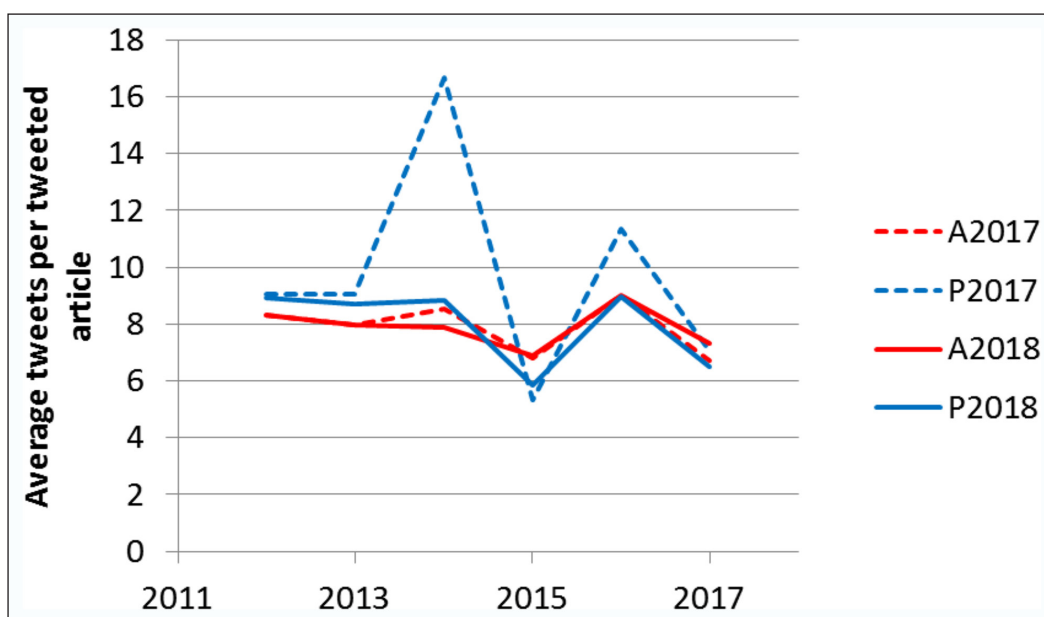
Figure 7: Average tweets per article per year of publication.

Haustein, Lariviére, and Sugimoto (2013) state that the only altmetric with reasonable coverage (besides Mendeley) is Twitter, but do not provide numbers or percentages. One of the possible reasons for the high Twitter coverage is that we studied a special dataset.

The most tweeted article based on Altmetric.com in 2017 and 2018, and PlumX in 2018 was 'The weakening relationship between the impact factor and papers' citations in the digital age' by Lozano, Lariviére and Gingras, published in 2012. It was the second most tweeted according to PlumX in 2017. The top tweeted article in PlumX in 2017 was 'Academia.edu: Social network or Academic Network?' by Thelwall and Kousha, published in 2014. This article had 723 tweets reported by PlumX in 2017 (Altmetric.com reported 16 tweets). In 2018 Altmetric reported the same number of tweets (16), and PlumX only 10. We have no explanation for this discrepancy in the number of tweets reported; it might have been an unintentional error by PlumX.

Conclusions

This paper demonstrates that overall there is a visible improvement in the coverage overlap between Altmetric.com, Mendeley and PlumX. We compared the same articles in two points in time; 2017 and 2018 and were able to see that within a relatively short amount of time, these three databases have reduced the number of coverage discrepancies. There are still evident signs of gaps in coverage but these seem to decrease over time and could be attributed to varying methodologies used by the three databases. For example, Mendeley is user-driven which makes it prone to errors. Mendeley users save records in different ways and thus there could be several instances of the same article or errors in the metadata itself that might prevent accurate account of readership. Altmetric.com only counts Mendeley readership if another altmetric indicator can be found for the article. Therefore, in cases where only Mendeley readership is found, Altmetric.com might not track the interaction until an additional altmetric indicator is found. This can also be a cause

for some of gap that we observed in the percentage of articles coverage. Although the gap seems to be reducing over time, it is recommended that altmetrics indicators and especially Mendeley readership counts will be analysed across more than one platform, even Mendeley. Because of some of its inherit metadata challenges, Mendeley data alone will, in some cases, not be accurate.

In the same manner, this article also showed that there are differences in the numbers of altmetrics indicators provided by Altmetric.com, Mendeley and PlumX. Again, this is a direct result of the manner by which each platform, tracks and reports altmetrics data as well as the sources it uses to do so. As with readership, we recommend that altmetric indicators analysis will be performed on more than one platform and compared to each other. First, one should look for articles coverage and ensure that the articles being analysed are indeed the same ones. Second, one should aggregate articles that are showing erroneous or partial metadata but are obviously the same. Lastly, one should collect the same altmetrics indicators from more than one platform and note whether there are significant differences between them. Unlike citations for example, altmetrics indicators are dynamic and are more difficult to control via standardizations. Therefore, despite of the considerable improvement in the overall overlap and coverage of articles in these databases, one should compare the results across platforms.

Limitations

This study was based on a relatively small sample in a specific field. Therefore, the results might be difficult to generalize. Further comparisons studies are needed across disciplines and years similarly to the ones that compare WoS, Scopus and Google Scholar (e.g. Halevi et al., 2017).

Additional File

The additional file for this article can be found as follows:

- **Table 2.** Twitter counts by year of publication as reported by Altmetric.com and PlumX in 2017 and 2018. DOI: https://doi.org/10.29024/joa.4.s1

Competing Interests

Judit Bar-Ilan is the editor-in-chief of this journal, but the peer review process of this article was handled by another member of the editorial board - Prof. Henk Moed.

References

Bar-Ilan, J. (2008). Which h-index? — A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, *74*(2), 257–271. DOI: https://doi.org/10.1007/s11192-008-0216-y

Bar-Ilan, J., & Halevi, G. (2017). Altmetric counts from different sources. *Presented at the Altmetrics17 Workshop*, September 26, 2017, Toronto, Canada.

Boeker, M., Vach, W., & Motschall, E. (2013). Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. *BMC Medical Research Methodology, 13*. DOI: https://doi.org/10.1186/1471-2288-13-131

Bramer, W. M., Giustini, D., & Kramer, B. M. (2016). Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: A prospective study. *Systematic reviews*, *5*(1), 39. DOI: https://doi.org/10.1186/s13643-016-0215-7

Costas, R., Zahedi, Z., & Wouters, P. (2015). Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, *66*(10), 2003–2019. DOI: https://doi.org/10.1002/asi.23309

Elsevier. (2013). Elsevier acquires Mendeley, an innovative, cloud-based research management and social collaboration platform. Retrieved from: https://www.elsevier.com/about/press-releases/corporate/elsevier-acquires-mendeley,-an-innovative,-cloud-based-research-management-and-social-collaboration-platform.

Gunn, W. (2016). Comment #00632 – clarification/correction of Mendeley saves definition – NISO RP-25-201x-3, Altmetrics Data Quality Code of Conduct – draft for public comment.pdf. Retrieved from: http://www.niso.org/apps/group_public/view_comment.php?comment_id=632.

Halevi, G., Moed, H., & Bar-Ilan, J. (2017). Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the literature. *Journal of Informetrics*, *11*(3), 823–834. DOI: https://doi.org/10.1016/j.joi.2017.06.005

Harzing, A. W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics*, *106*(2), 787–804. DOI: https://doi.org/10.1007/s11192-015-1798-9

Haustein, S. (2016). Grand challenges in altmetrics. *Scientometrics*, *108*, 413–423. DOI: https://doi.org/10.1007/s11192-016-1910-9

Haustein, S., Bowman, T. D., & Costas, R. (2016). Interpreting "altmetrics": viewing acts on social media through the lens of citation and social theories. In: Sugimoto, C. R. (ed.), *Theories of informetrics and scholarly communication. A Festschrift in Honor of Blaise Cronin*, 372–405. Berlin: De Gruyter. Retrieved from: http://arxiv.org/abs/1502.05701. DOI: https://doi.org/10.1515/9783110308464-022

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & **Terliesner, J.** (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, *101*(2), 1145–1163. DOI: https://doi.org/10.1007/s11192-013-1221-3

Michalek, A. (2017). Plum Analytics joins Elsevier. Retrieved from: https://plumanalytics.com/plum-analytics-joins-elsevier/.

Mohammadi, E., & **Thelwall, M.** (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology*, *65*(8), 1627–1638. DOI: https://doi.org/10.1002/asi.23071

Ortega, J. L. (2018). Reliability and accuracy of altmetric providers: A comparison among Altmetric.com, PlumX and Crossref Event Data. *Scientometrics*, *116*(3), 2123–2138. DOI: https://doi.org/10.1007/s11192-018-2838-z

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

Rasmussen, P. G., & **Andersen, J. P.** (2013). Altmetrics: An alternate perspective on research evaluation. *Sciecom info*, *9*(2). Retrieved from: http://journals.lub.lu.se/index.php/sciecominfo/article/view/7292/6102.

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. DOI: https://doi.org/10.1371/journal.pone.0064841

Trapp, J. (2016). Web of Science, Scopus, and Google Scholar citation rates: A case study of medical physics and biomedical engineering: What gets cited and what doesn't? *Australasian Physical & Engineering Sciences in Medicine*, *39*(4), 817–823. DOI: https://doi.org/10.1007/s13246-016-0478-2

Zahedi, Z., Costas, R., & **Wouters, P.** (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*, *101*(2), 1491–1513. DOI: https://doi.org/10.1007/s11192-014-1264-0

Zahedi, Z., Fenner, M., & **Costas, R.** (2015). Consistency among altmetrics data provider/aggregators: What are the challenges. *Presented at altmetrics15*. Retrieved from: https://altmetrics.org/wp-content/uploads/2015/09/altmetrics15_paper_14.pdf.