

Reliability and accuracy of altmetric providers: a comparison among Altmetric, PlumX and Crossref Event Data

José Luis Ortega

Cybermetrics Lab, Madrid, Spain, jortega@orgc.csic.es

Abstract

The aim of this study is to measure differences between three relevant altmetric providers: Altmetric, PlumX and Crossref Event Data (CED). More than 67,000 research papers, initially extracted from PlumX, were searched in Altmetric and CED to compare their counts. Differences between services were analyzed regarding to the number of documents with an altmetric event and the counting differences in each metric. Results show that Altmetric is the provider with the best coverage of blog posts, news and tweets; while PlumX better collects Mendeley readers, and CED is the site that extracts the most Wikipedia citations. The study concludes that there are important counting differences due to technical errors and extracting criteria. The article recommends the use of specific services for particular metrics or the combination of several sources if we want to perform a global analysis.

Keywords: Altmetric, PlumX, Crossref Event Data, Altmetrics, data providers

Introduction

Since the publication of the Altmetrics Manifesto (Priem et al., 2010), a huge number of articles and books have been published around the world. The main objective of altmetric studies has been to understand the meaning of these metrics and their possible application to research evaluation. In this form, a first step has been to compare and put in context these metrics with regard to bibliometric indicators (Thelwall et al., 2013; Costas et al., 2015). The poor correlations found lead us to consider, for now, that these metrics express a new dimension of academic impact, closer to the popularization of science, societal impact and research dissemination (Ortega, 2016).

Altmetric studies have been supported by data produced by a large range of services that track the use, mention, sharing and citation of research papers in social networks, publisher platforms, repositories, etc. Faced with this great amount of metrics and sources, Altmetric providers have been a key element in the consolidation of this young research field, gathering alternative metrics under a same platform. These databases have made easier the comparison and study of different metrics, putting them in relation to each other and to the citation impact. Based on web syndication and open data, events on the Web about academic outputs are collected by these aggregators using several identifiers (DOIs, ISBN, URLs, etc.). The counting of these events makes visible a new perspective about the social and online impact of these academic results. Today, these services are inserted in the scholarly communication

system providing altmetric information to journals (PLOS, Nature), publisher platforms (ScienceDirect, SpringerLink) and digital libraries.

However, and spite of the importance of these instruments for altmetric studies, there are few studies that have treated the consistency and reliability of altmetrics across data providers. This is an important gap because many of the current results might be limited to the data supplied by these aggregators. The aim of this study is to fill this gap and to analyze the coverage and counting differences between three representative data providers: Altmetric, PlumX and CED.

Related Research

Altmetric providers constitute an indispensable part in the altmetric research because they supply data for these studies. Some works have described and explored the functionalities of these services, as a way to explain their advantages and drawbacks. On this point, it is worth mentioning the studies of Champieux (2015) and Lindsay (2016) about PlumX; the analyses performed by Adie and Roe (2013) and Trueger et al. (2015) about Altmetric and the serious criticisms of Gumpenberger et al. (2016) about the Altmetric Score. There are no studies to date about Crossref Event Data. However, other studies have analyzed the coverage of these services and have described the proportion of altmetric events in several samples. Robinson-García et al. (2014) analyzed the coverage of Altmetric and they found that 87.1% of articles had at least one tweet and 64.8% one Mendeley reader. In a similar way, Bornmann (2014) explored a set of articles from Altmetric and he observed that 71% of articles were tweeted and a high proportion of documents were mentioned in blogs (16%) and news (13%). Fraumann et al. (2015) also described the distribution of blogs and news covered by Altmetric and they found an important bias to U. S. sites. According to PlumX, there are no studies that have treated its coverage. We can only mention the recent work of Torres-Salinas et al. (2017) about the collection of books and the assessment of these materials, and the study of Ortega (2016) about the tweeting of research articles using PlumX data.

However, much fewer papers have carried out comparative studies between data providers. Jobmann et al. (2014) were the first ones to compare the altmetric coverage and counts of ImpactStory, Altmetric Explorer, Plum Analytics and Webometric Analyst by research areas. Their results show important divergences between services. Plum Analytics is the platform that better covers Mendeley and Facebook data, while Altmetric stands out gathering blogs, news and CiteULike data. Zahedi et al. (2015) explored the consistency of data across three altmetric aggregators: Altmetric, Mendeley and Lagotto. They also detected significant differences, finding that Altmetric gathers more tweets, but it is faulty collecting Mendeley readers. Baessa et al. (2015) evaluated several altmetric providers for their institutional repository and they recognized that Altmetric has a better coverage of blogs, news and government documents, while PlumX is most exhaustive regarding formats. Kraker et al. (2015) studied the collection of research data in three data sources: figshare, PlumX and ImpactStory. They observed that PlumX detects considerably more items in social media and also finds higher altmetric scores than ImpactStory. The most recent comparative approach was performed by Peters et al. (2016), who extended the before study with the inclusion of Altmetric. Their results confirmed that PlumX covers better non published materials such as research data.

Many of these studies have been focused more on coverage of altmetrics than on counting events and, in the case that events were studied, they were analyzed at research area or source level. This study thus attempts to investigate these counting differences comparing the altmetrics of each article in three representative sources. We consider that this method would bring more detailed results.

Objectives

The main purpose of this study is to compare the results of three important altmetrics providers for a same set of publications. The objective is to determine the differences and similarities between services when they extract and gather these metrics, detecting inconsistencies between different results. These outputs would allow the scholarly community to value the quality and reliability of these services as data providers. Several research questions were formulated:

- Are there differences between data providers according to the proportion of altmetric events captured?
- Are there counting inconsistencies between the metric provided by these altmetrics services? And if that is the case, is it possible to quantify these differences?
- Could these altmetric providers be used for research evaluation equally?

Methods

Three of the most important altmetric providers were selected for this study: Altmetric, PlumX and Crossref Event data (CED). Two main criteria were applied to select these platforms. First, altmetric providers should have a global coverage, not limited to specific publishers, disciplines or regions. In this way, altmetric services of publishers (i.e. PLOS, Nature Publishing Group, etc.) were rejected. Second, these services should have the ability to extract data from them, using a public API or scrapping their web pages. Due to this restriction, ImpactStory was not analyzed because it does not provide an API key (ImpactStory, 2017). Both conditions are necessary to be fulfilled in order to make a fair comparison among platforms.

Altmetric providers

PlumX: PlumX is a provider of alternative metrics created in 2012 by Andrea Michalek and Michael Buschman. This product is addressed to the institutional market, offering altmetric counts of publications for particular institutions. PlumX is the aggregator that offers more metrics, including citation and usage metrics (i.e. Views and Downloads). It covers more than 52.6 million of artifacts, being then the largest altmetric aggregator. However, only a limited set of institutions make public their profiles, therefore, not all the publications tracked by PlumX are accessible. In 2017, Plum Analytics was acquired by Elsevier and its altmetric information was added to Scopus database. This change could influence data sources, because many of their sources are EBSCO services, the former proprietary.

Altmetric: It was the first altmetric provider and it was born in 2011 by Euan Adie, with the support of Digital Science. Unlike PlumX, Altmetric.com is centered in the publishing world, signing agreements with publisher houses to monitor the altmetric impact of their

publications. This information is accessible through a public API. Today, Altmetric tracks the social impact of close to 8 million of research papers. However, this platform does not include metrics about citations and usage. In 2015, Altmetric launched Altmetric for Books, an exclusive service for books and book's chapters.

Crossref Event data (CED): CED is the youngest service, created in 2016 and officially released in 2017. Due to this, the platform claims that the service is still in beta. Unlike the previous ones, CED is not a commercial site and it provides free access to their data through a public API. Another difference is that it does not aggregate the information, but it displays the entire information about each altmetric event. For instead, it shows the information about the mention of an article on Twitter (date, user, tweet, etc.), but it not show a count of the number of tweets. For that reason, data have to be processed to be comparable with the other services.

Data extraction

An initial sample of 67,147 research articles was obtained from the public pages of PlumX. This platform was used as the starting source because it does not permit to retrieve publications searching by DOI. This drawback makes impossible the search of publications from other altmetric sources in PlumX database. Other important reason is that this platform only shows altmetric information of publications from institutions that make public their profiles. Therefore, we may suppose that those publications could be also indexed in the other services because they have no problem to display their records. Another reason is that PlumX includes publications that have or have not altmetric events, which it also allows us to detect articles that may have altmetric events in other aggregators.

The first step was to identify the organizations that make public their profiles. A search in Google (site:plu.mx) retrieved the web pages of 17 research institutions. The sites that more publications contribute to the sample are: *Mount Sinai Health System* (19,827), *Concytec* (8,141), *University of Pittsburgh* (8,103), *Seoul National University, College of Medicine* (7,871), *King Abdullah University of Science and Technology* (7,281), *University of Helsinki* (5,287), *Xi'an Jiaotong University* (2,781), *International Islamic University of Malaysia* (2,350), *Universidad del Rosario* (1,648), *Georgia Southern University* (1,083), and *Centro Hospitalar e Universitário de Coimbra* (1,009). The remaining institutions contribute with fewer papers. Only article types published after 2013 and with a DOI were selected, being in total 67,147 research papers.

This initial sample of research papers was searched again in Altmetric.com using its public API (api.altmetric.com), and in CED using its API (query.eventdata.crossref.org) as well. In this way, three samples were obtained that can be compared among them. A routine in SQL was written to crawl and extract the altmetric information from the three providers. The samples were obtained in May 23rd, 2017.

Metrics

Because each service has different numbers and types of metrics, it is necessary to precise the metrics analyzed in this study. These indicators have to be common to the three platforms and they have to be counted in a similar way. Five metrics were compared:

- **Blog posts:** number of blog posts that mention a research article. Each service has its own list of blogs where the mentions are searched. Differences between services could be caused more by the number of indexed blogs than by the way in which the mentions are extracted. The only service that provides a number of blogs is Altmetric, which monitors 11,000 blogs (Williams, 2017). The other platforms do not make public the list of blogs nor the number of blogs covered.
- **News:** number of news feeds that mention a research article. In the manner of blog posts, each platform has their own lists of news media. Therefore, differences between services could be also due to coverage criteria. Altmetric collects a list of 1,300 news outlets, which could be expanded to 80,000 thanks to a partnership with Moreover.com (A LexisNexis Company) (Williams, 2015). The number of news media covered by PlumX is unknown. CED defines this section as “web” and it includes links to other sites different from blogs, then it is possible that many of these counts do not belong to news mentions.
- **Tweets:** number of tweets that cite a research article. In this case, the source is the same for the three sites. Possible differences might be found in the way that they obtain their data (i.e. data providers) and how they search the article in the source (i.e. use of identifiers).
- **Wikipedia citations:** number of Wikipedia entries that include a citation to a research article. This is a unique source as well, and the differences might be motivated by the way in which the citations are extracted.
- **Mendeley readers:** number of users in Mendeley that include a research article in their libraries. As the two previous metrics, differences in this count would be caused by technical reasons.

Results

Distribution of altmetrics by provider

	PlumX	PlumX %	PlumX (altmetrics) %	Altmetric	Altmetric %	CED	CED %
Blogs	1,787	2.66	3.37	3,334	11.86	37	1.40
News	852	1.27	1.61	4,493	15.98	132	5.01
Tweets	11,803	17.56	22.28	25,699	91.38	896	34.00
Wikipedia	1,168	1.74	2.21	841	2.99	844	32.03
Readers	52,256	77.74	98.66	26,930	95.76	0	0
Total_records	67,147			28,123		2,635	

Table 1. Number of articles with at least one altmetric event by provider

Table 1 presents the number of documents in the three samples that have altmetric events. Overall, Altmetric includes 28,123 articles (41.84%) of the initial sample of 67,147 papers in PlumX. This proportion shows that one out of every two articles has statistics in both services, making evident a low overlapping rate. According to CED, the overlap is much higher because 63,637 articles (94.67%) of the PlumX’s sample are as well registered in CED. However, only 2,635 (3.92%) of these documents have altmetric events. This absence of data could be due to the service is still in beta and many of the events are not counted yet.

Observing the distribution of altmetric events in the three services, one can detect that there are important differences according to the percentage of articles with an altmetric event. PlumX, the original sample, shows a high proportion of articles bookmarked in Mendeley (77.74%) and mentioned in Twitter (17.56%), and in less matter, mentioned in blogs (2.66%), news (1.27%) and Wikipedia (1.74%). However, in the case of Altmetric, these percentages are much higher. Thus, the proportion of articles saved in Mendeley is 95.76% and mentioned in Twitter is 91.38%. These high percentages are due to Altmetric only records articles that at least have some altmetric event, whereas PlumX indexes articles without distinction. If the articles without altmetrics (14,251) are subtracted, PlumX would then include more articles with readers (98.66%) than Altmetric (95.76%). However, this subtraction does not prevent that the proportion of tweeted papers in PlumX (22.28%) is much lower than in Altmetric (91.38%). This significant small proportion of tweets was already noticed by Jobmann et al. (2014) and it could be due to PlumX did not use Gnip, the official Twitter data provider, until 2016.

Regarding to CED, the number of articles with an altmetric event is very low, with only 2,635 papers (3.92%) from the initial PlumX's sample. This demonstrates that CED is still in beta and the coverage of altmetric events is for now incomplete. Taking only into account papers with at least one altmetric event, 34% articles have a Twitter mention and 32.03% a citation in Wikipedia. This last is the highest percentage of all the three providers, and it suggests that CED has a special coverage of Wikipedia articles.

Counting differences between providers

The main objective of this work is to analyse the counting differences between altmetric providers. To this end, a new distribution was created from differences between two altmetric aggregators. For example, Altmetric/PlumX shows the counting difference of one metric (blog posts, tweets, citations, etc.) between Altmetric and PlumX by each article. We think that is an easier and clearer way to present and analyze these differences as well as it allows to detect systematic or random errors. In this way, the dispersion (Standard deviation) informs us about the amount of difference between two services. As more pointed and slender is a distribution, lesser different are both altmetric providers. However, the symmetry would be symptom of random or systematic errors. Thus, a skewed distribution will show that a provider always counts more or less events than other provider in a same sense. This would be caused by a technical failure or a different criterion to compute a metric.

Blogs

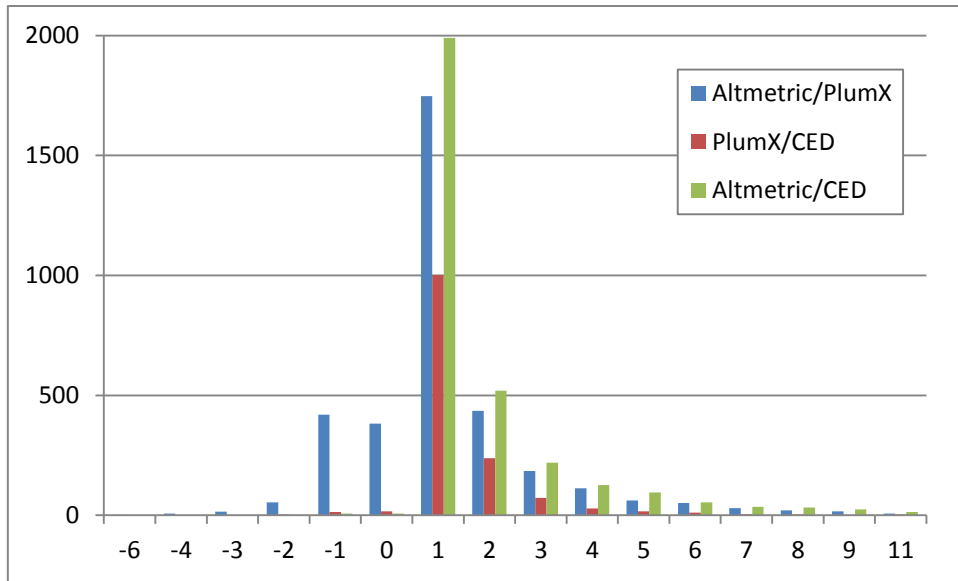


Figure 1. Histograms of counting differences between providers according to blog posts

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric/PlumX	3,585	1.52 (±.094)	2.703	13.92%	10.66%	48.73%	12.13%
PlumX/CED	1,417	1.54 (±.094)	1.792	1.2%	1.13%	70.64%	16.8%
Altmetric/CED	3,185	2.21 (±.107)	3.09	0.25%	0.22%	62.48%	16.33%

Table 2. Some parameters of the three comparative distributions according to blog posts

Figure 1 and Table 2 present the distribution of the counting differences between Altmetric, PlumX and CED regarding blog posts. Generally, Altmetric is the service that counts more blog posts. In average, 1.52 more posts than PlumX and 2.21 more than CED. These differences suggest that Altmetric has the largest list of blogs. In fact, it gathers more than 11,000 blogs. By contrast, the number of blogs that PlumX and CED track is unknown. However, the fact that 13.9% of the articles contain more blog posts in PlumX than in Altmetric and that only 10.6% of the articles show the same count in both services would mean that the PlumX's list could be quite different to the Altmetric one. It is then possible that PlumX covers an important number of blogs that Altmetric does not manage. CED shows more significant differences with respect to the other providers. In almost all the cases, CED contains less blog mentions than Altmetric (2.21) and PlumX (1.54). This result suggests that the CED's blog list is short and overlapped with the other providers.

News

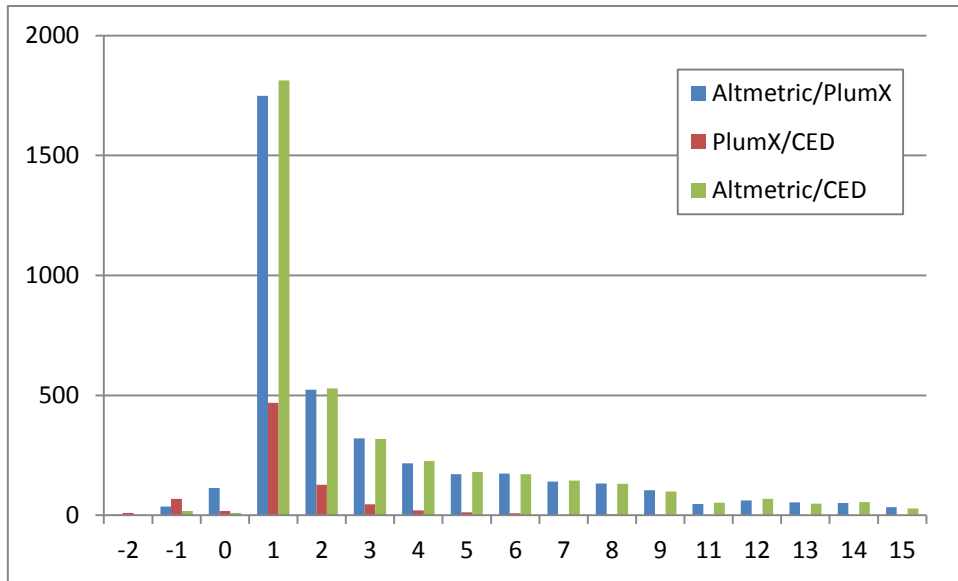


Figure 2. Histograms of counting differences between providers according to news

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric/PlumX	4,224	6.79 (±.45)	14.951	0.92%	2.63%	40.3%	12.08%
PlumX/CED	791	1.37 (±.17)	2.378	0.5%	0.23%	41.9%	12.23%
Altmetric/CED	4,314	6.82 (±.43)	14.463	12%	2.23%	57.97%	15.7%

Table 3. Some parameters of the three comparative distributions according to news

Figure 2 and Table 3 show the distribution of the counting differences between providers at the level of news mentions. Again, Altmetric is the service that gathers more comments on news media sites. In average, it collects 6.79 more than PlumX and 6.82 more than CED. These differences cause a high standard deviation of the Altmetric’s distributions. This fact makes evident that media coverage in Altmetric is much wider than in the other services. Some possible causes could be that Altmetric can track the mention of articles in the media without a specific identifier (Liu, 2013). Other possible reason would be its partnership with Moreover.com, a news data provider, which it allows to collect more news media (Williams, 2015). According to CED, it is possible that many of these counts do not exactly belong to news mentions. Therefore, the percentage of real news mentions could be even lower. Differences between PlumX and CED are smaller (1.37), which it could mean that both services have similar news sources.

Tweets

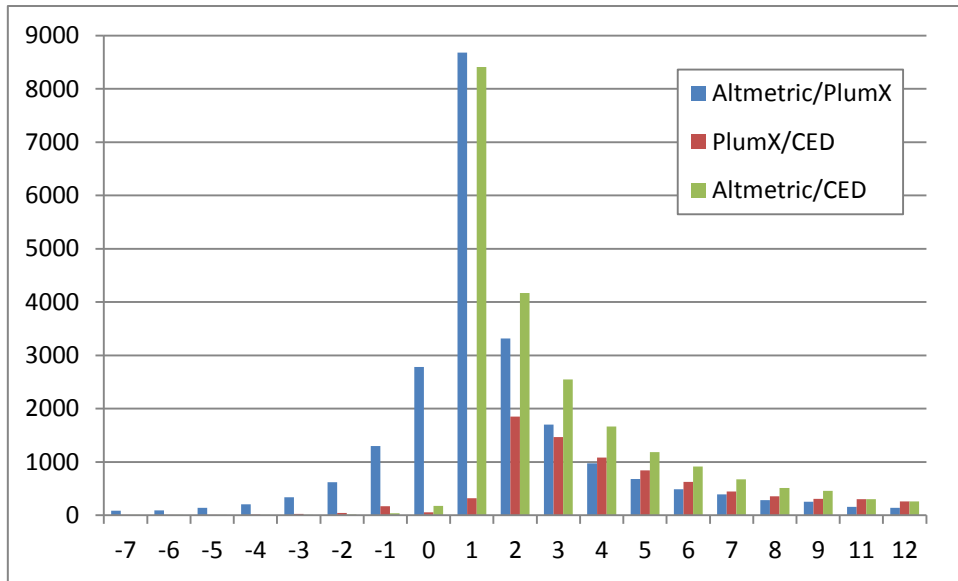


Figure 3. Histograms of counting differences between providers according to tweets

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric/PlumX	21,742	3.72 (± 0.34)	25.445	13.07%	11.34%	35.39%	13.51%
PlumX/CED	9,951	14.32 (± 1.37)	69.692	0.36%	0.72%	34.46%	17.09%
Altmetric/CED	24,227	9.21 (± 0.63)	49.676	3.38%	0.57%	3.17%	18.5%

Table 4. Some parameters of the three comparative distributions according to tweets

Figure 3 and Table 4 present the distribution of the counting differences between Altmetric, PlumX and CED regarding Twitter mentions. In average, Altmetric again counts 3.72 more tweets than PlumX and 9.21 more than CED. It is interesting to notice that there is a slight symmetry between Altmetric and PlumX. 13.07% of the articles have more tweets in PlumX than in Altmetric, 11.34% show the same number of mentions in both platforms and 35.39% have one more tweet in Altmetric than in PlumX. This symmetry could be due to Altmetric only counts tweeters, different tweets from the same account are hence not counted. On the contrary, PlumX does count those tweets. This could explain why there are 13.07% of articles that have more tweets in PlumX than in Altmetric. However, the remaining 75.6% of articles with more tweets in Altmetric than in PlumX suggest that there is a systematic error. This error could be caused by the employment of different data providers or by the use of distinct criteria in order to count tweets. For example, a possible cause could be that PlumX did not use data from Gnip, the official Twitter data provider, until 2016. This could cause that some tweets were not properly matched. A manual inspection of some articles showed that the range of URLs that PlumX uses to identify mentions is much smaller than the used by Altmetric. This absence of completeness would produce the systematic underperforming of PlumX according to Altmetric. With regard to CED, it is interesting to notice the great difference with PlumX (14.32) and Altmetric (9.21). In this case, there are more systematic variations and these could be due that CED only compiles tweets that contain the DOI prefix or the main landing page of the article, dismissing other links to repositories, publishing platforms, etc (CED, 2017).

Wikipedia

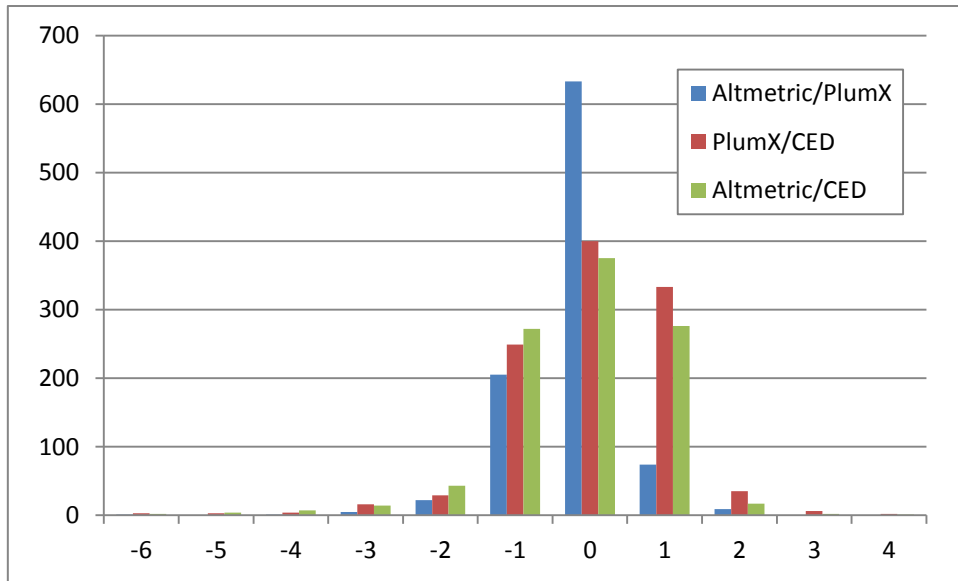


Figure 4. Histograms of counting differences between providers according to Wikipedia citations

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric/PlumX	955	-0.73 (±0.21)	1.954	24.92%	66.28%	7.75%	0.94%
PlumX/CED	1,091	-0.6 (±.33)	3.31	28.69%	36.66%	30.52%	3.21%
Altmetric/CED	1,023	-0.24 (±.25)	4.294	34.31%	36.66%	26.98%	1.66%

Table 5. Some parameters of the three comparative distributions according to Wikipedia citations

Figure 4 and Table 5 exhibit the distribution of the differences between the three data providers according to Wikipedia citations. In this case, differences are smaller than in the other metrics. This fact is demonstrated by the low averages and the high percentage of articles with the same count value. CED is the platform that gathers the most Wikipedia citations in average, with 0.6 more citations than PlumX and .24 more than Altmetric. In addition, PlumX obtains .73 more mentions than Altmetric. This metric displays the highest percentages of articles with the same count value, being 66.3% in Altmetric/PlumX, 35.6% in PlumX/CED and 35.7% in Altmetric/CED. The fact that Wikipedia citations shows such small differences across providers could be caused by the low number of references in Wikipedia, by the easy way with which the mentions are computed and by the fact that only one source is tracked. Despite of these reasons, differences between CED and the other providers are caused because Altmetric and PlumX only track mentions on the English, Finnish and Swedish-language Wikipedias (Altmetric Support, 2017), while CED explores all the Wikipedia. Another possible reason is that Altmetric does not extract mentions from incomplete references or without an identifier. However, PlumX counts mentions only matching the title of the article. This difference would explain the slightly higher number of mentions computed in PlumX according to Altmetric.

Readers

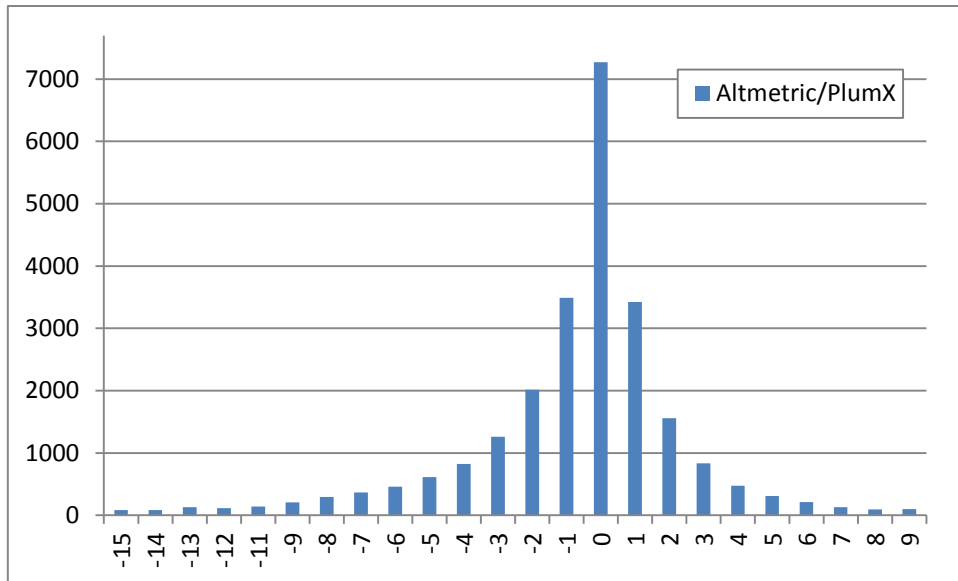


Figure 5. Histograms of counting differences between Altmetric/PlumX according to Mendeley readers

	Articles	Mean	Std. Dv.	Counting differences			
				<0	0	1	2
Altmetric/PlumX	18,436	-2.5 (±0.48)	33.421	42.35%	28.28%	13.32%	6.07%

Table 6. Some parameters of the comparative distribution of Altmetric/PlumX according to Mendeley readers

In the case of Mendeley readers, only Altmetric and PlumX include this altmetric indicator. CED does not offer information about this bibliographic service. Figure 5 depicts the distribution of the differences between Altmetric and PlumX compiling Mendeley readers, and Table 6 shows some parameters about this distribution. The histogram is nearly symmetric, with an important proportion of similar counts (28.28%). However, there is a slightly higher proportion of negative values, which means that PlumX regularly counts more readers than Altmetric. In fact, the mean of the differences (-2.5) confirms this counting error. This result is significant because we are talking about only one source and the metric is already counted by the source, that is, providers do not have to aggregate readers. This error is explained by the fact that Mendeley allows to include references that are not previously registered in its Public Catalog. The risk of this practice is that it creates duplicate records. In this case, PlumX aggregates the readers of possible duplicate records from references with similar titles, years, authors, etc. Contrarily, Altmetric only counts the readers from a unique identifier, dismissing readers from duplicate items. This could cause that in average PlumX shows more readers than Altmetric.

Discussion

The results show that there are important differences between providers when they collect and count altmetrics. In general, Altmetric captures more altmetric impact than the other services. Excepting the number of Mendeley readers and Wikipedia citations, Altmetric exhibits the highest percentage of papers with some altmetric event by each metric. Perhaps, the most significant difference is regarding the coverage of Twitter. 91.38% of articles in

Altmetric have been mentioned on Twitter, a slightly higher proportion than the found by Robinson-García et al. (2014). Meanwhile, only 34% of articles in CED and 22.28% in PlumX were tweeted. This significant difference was already noticed by Jobmann et al. (2014), mainly with regard to PlumX. This variation could be possible that this aggregator did not use Gnip, the official Twitter data provider, until 2016. Another reason could be that the number of identifiers used by PlumX and CED is lower than the ones used by Altmetric, causing a reduction of the number of captured tweets. Altmetric also has a better coverage of blogs and news, with more than the double of mentions. However, there are some metrics in which PlumX and CED perform better than Altmetric. For instance, PlumX (98.66%) covers more articles with Mendeley readers than Altmetric (95.76%); and CED captures more Wikipedia citations (32%) than Altmetric (2.99%) and PlumX (2.21%). These coverage differences allow us to suggest that, despite Altmetric is comparatively the best provider, it would be recommendable to select certain providers to analyze specific metrics. In this form, while Altmetric would be the best source for news, blog posts and tweets, PlumX would be more suitable to study Mendeley readers and CED to Wikipedia mentions.

According to counting differences, the results also show that there are important variations between providers in the five metrics analyzed. Altmetric is the service that has a better coverage of blogs, news and tweets, a fact also confirmed by Baessa et al. (2015). In these metrics, Altmetric shows more counts than PlumX and CED, being more visible in news feeds, with 6.79 more news than PlumX and 6.82 more than CED. This result makes clear that Altmetric is more concerned with having a complete list of blogs and news, which let it capture the widest media impact possible. Altmetric also gathers more tweets than their competitors (3.72 more than PlumX and 9.21 more than CED), caused perhaps by the use of more URLs identifiers such as DOIs, repository handles, landing pages, etc.

By contrast, PlumX achieves a better coverage of Mendeley readers (2.5 more) and Wikipedia citations (.73 more) in comparison with Altmetric, although these differences are smaller and they are due to technical criteria. For example, Altmetric obtains fewer Mendeley readers because it fails to aggregate duplicate records. In comparison with CED, PlumX has better figures in all metrics, excepting Wikipedia citations (.6 fewer than CED). This difference is especially significant when counting tweets, where PlumX captures 14.32 more tweets than CED.

These results have also indicated that CED is still a beta service. The number of collected events is rather low in comparison with the other providers, being this difference more notorious in Altmetric than in PlumX. This fact is evident in all the metrics, with the only exception of Wikipedia citations. In this case, CED covers all the Wikipedia, slightly collecting more citations than Altmetric (.24) and PlumX (.6).

These results have strong implications for research evaluation because they demonstrate that the counts are not uniform across providers. This means that the altmetric impact of scientific publications is different according to the provider used and it casts doubts on the reliability of these tools for measuring the altmetric impact. These differences are due to systematic errors caused by coverage problems (blog posts, news, Wikipedia citations) or technical limitations (tweets, Mendeley's readers). Such results recommend using specific providers if we want to

track particular metrics. In this sense, Altmetrics is the most suitable tool for blogs, news and tweets, PlumX for Mendeley readers and CED for Wikipedia citations. On the contrary, if we want to analyze several metrics, these results, then, suggest to use several aggregators. In this way, altmetric studies could be more consistent selecting the most appropriate sources depending on the metric to be analyzed.

Conclusions

Several conclusions were drawn from the obtained results:

There are differences in the amount of articles with altmetric impact in each provider. Altmetric is the service with the highest percentage of publications with some altmetric event, surpassing the other providers in almost all the indicators. The only exception is on Wikipedia citations where CED obtains very high ratios and on Mendeley readers where PlumX slightly exceeds it. Otherwise, PlumX yields a significantly low proportion of tweeted papers and CED only stands out extracting Wikipedia citations.

There are also important differences between providers when they come to count altmetrics. Overall, Altmetric is the platform that collects more altmetric events, obtaining the best scores on blog posts, news and tweets. PlumX is only better than Altmetric when it compiles Mendeley readers, but it is limited when it comes to extract blogs, news and tweets. CED is still a beta product with a short range of altmetrics and a deficient collection of news and tweets. However, it is the best service extracting Wikipedia citations.

Finally, these results suggest that Altmetric, PlumX and CED cannot be used separately for research evaluation because there are important differences according to coverage and counting procedures. Therefore they recommend to select particular providers for the analysis of specific metrics (i.e. CED for Wikipedia citations, PlumX for Mendeley readers) or to combine several sources if we want to perform a global analysis about the interaction of several altmetrics.

References

Adie, E. & Roe, W. (2013). Altmetric: enriching scholarly content with article-level discussion and metrics. *Learned Publishing*, 26(1): 11-17.

Altmetric Support (2017). How does Altmetric track mentions on Wikipedia?
<https://help.altmetric.com/support/solutions/articles/6000060980-how-does-altmetric-track-mentions-on-wikipedia->

Baessa, M., Lery, T., Grenz, D. & Vijayakumar, J. K. (2015). Connecting the pieces: Using ORCID's to improve research impact and repositories. *F1000Research*, 4: 195.
<https://f1000research.com/articles/4-195/v1>

Bornmann, L. (2014). Validity of altmetrics data for measuring societal impact: A study using data from Altmetric and F1000Prime. *Journal of Informetrics*, 8(4): 935-950.

Chamberlain, S. (2013). Consuming article-level metrics: Observations and lessons from comparing aggregator provider. *Information Standards Quarterly*, 25(2): 4-13.

Champieux, R. (2015). PlumX. *Journal of the Medical Library Association*, 103(1): 63.

CED (2017). Twitter – Event Data User Guide

<https://www.eventdata.crossref.org/guide/sources/twitter/>

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.

Fraumann, G., Zahedi, Z. & Costas, R. (2015). What do we know about Altmetric.com sources? A study of the top 200 blogs and news sites mentioning scholarly output. Altmetrics workshop, 9 October 2015, Amsterdam Science Park, Amsterdam.

Gumpenberger, C., Glänzel, W. & Gorraiz, J. (2016). The ecstasy and the agony of the altmetric score. *Scientometrics*, 108(2): 977-982.

ImpactStory (2017). New widget and API. ImpactStory blog. <http://blog.impactstory.org/new-widget-and-api/>

Jobmann, A., Hoffmann, C.P., Künne, S., Peters, I., Schmitz, J. & Wollnik-Korn, G. (2014). Altmetrics for large, multidisciplinary research groups: Comparison of current tools. *Bibliometrie – Praxis und Forschung*, 3, <http://www.bibliometrie-pf.de/article/viewFile/205/258>

Kraker, P., Lex, E., Gorraiz, J., Gumpenberger, C. & Peters, I. (2015). Research data explored II: The anatomy and reception of figshare. *arXiv preprint arXiv:1503.01298*.

Lindsay, J. M. (2016). PlumX from Plum Analytics: Not Just Altmetrics. *Journal of Electronic Resources in Medical Libraries*, 13(1): 8-17.

Liu, J. (2013) New Feature: News Tracker. Altmetric: <https://www.altmetric.com/blog/new-feature-news-tracker/>

Ortega, J. L. (2016). To be or not to be on Twitter, and its relationship with the tweeting and citation of research papers. *Scientometrics*, 109(2): 1353-1364.

Peters, I., Kraker, P., Lex, E., Gumpenberger, C. & Gorraiz, J. (2016). Research data explored: an extended analysis of citations and altmetrics. *Scientometrics*, 107(2): 723-744.

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

Robinson-García, N., Torres-Salinas, D., Zahedi, Z. & Costas, R. (2014). Nuevos datos, nuevas posibilidades: Revelando el interior de Altmetric.com. *El Profesional de la Información*, 23(4): 359-366.

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., Gumpenberger, C. & Gorraiz, J. (2017). PlumX as a Potential Tool to assess the Macroscopic Multidimensional impact of Books. *arXiv preprint arXiv:1707.03327*.

Trueger, N. S., Thoma, B., Hsu, C. H., Sullivan, D., Peters, L. & Lin, M. (2015). The altmetric score: a new measure for article-level dissemination and impact. *Annals of emergency medicine*, 66(5): 549-553.

Williams, C. (2015) Altmetric to provide expanded mainstream media tracking. Altmetric: <https://www.altmetric.com/blog/moreover/>

Williams, C. (2017) Help Altmetric further expand its coverage: tell us about your favourite blogs. Altmetric: <https://www.altmetric.com/blog/expandblogs/>

Zahedi, Z., Fenner, M. & Costas, R. (2015). Consistency among altmetrics data provider/aggregators: What are the challenges?. The 2015 Altmetrics Workshop, Amsterdam.