The Life Cycle of Altmetric Impact: A Longitudinal Study of Six Metrics from PlumX

José Luis Ortega
Cybermetrics Lab, Madrid, 28006
Email: jortega@orgc.csic.es

Abstract

The main objective of this study is to describe the life cycle of altmetric and bibliometric indicators in a sample of publications. Altmetrics (Downloads, Views, Readers, Tweets, and Blog mentions) and bibliometric counts (Citations) (in this study, the indicators will be capitalized to differentiate them from the general language) of 5,185 publications (19,186 observations) were extracted from PlumX to observe their distribution according to the publication age. Correlations between these metrics were calculated from month to month to observe the evolution of these relationships. The results showed that mention metrics (Tweets and Blog mentions) are the earliest metrics that become available most quickly and have the shortest life cycle. Next, Readers are the metrics with the highest prevalence and with the second fastest growth. Views and Downloads show a continuous growth, being the indicators with the longest life cycles. Finally, Citations are the slowest indicators and have a low prevalence. Correlations show a strong relationship between mention metrics and Readers and Downloads, and between Readers and Citations. These results enable us to create a schematic diagram of the relationships between these metrics from a longitudinal view.

Keywords: altmetrics, PlumX, Citations, Readers, Tweets, longitudinal study

1. Introduction

Altmetrics studies started as a result of the rise of new metrics that count the number of events related to research outputs on the Web (Tattersall, 2016). Academic social networks, publishing houses and repositories, content providers, and academic search engines are computing events or actions for any academic document hosted in their databases. Starting from the conceptual framework of bibliometrics, altmetrics has established a parallelism between citations and social mentions, publication venues and social networks, etc., but without first exploring the real origin and meaning of these measures. Because of this dependence on bibliometric principles, altmetrics lack their own conceptual foundation (Priem and Hemminger, 2010; Haustein et al., 2015). In contrast to bibliometric indicators, which count only mentions (citations) and production (publications) of scholarly outputs in the academic publishing world, altmetric indicators consider a wider and more complex range of actions related to the usage, mentioning, sharing, and bookmarking of research publications. This wider perspective on the impact introduces more theoretical questions, as each indicator expresses different actions that occur in different varying times and contexts (Haustein, 2016). Consequently, the interpretation of each metric is very different from other altmetric indices and some of these metrics cannot be associated with concepts such as research impact or quality. For that reason, altmetric indicators have been the target of criticisms on the real value of these metrics for the assessment of research papers and therefore on the utility of altmetrics for policy-making (Bornmann, 2014a; Sugimoto et al., 2016). These criticisms are based on the absence of a definite and consistent theoretical background that explains the academic nature of these metrics (Nature Materials, 2012; Wouters and Costas, 2012; Sugimoto, 2014; Haustein, 2016). This makes it clear that more fundamental analyses are needed to elucidate the meaning, origin, and effect of these alternative metrics, which would simplify the employment of altmetrics in research evaluation processes.

This article attempts to explore five altmetric indicators and one bibliometric indicator from a dynamic perspective: Views, Downloads, Readers, Tweets, Citations, and Blog mentions. The purpose is to describe the evolution and frequency of these metrics throughout the life cycle of research papers and to understand how these metrics are related among each other and how they could influence citation impact.

2. Related research

Before the “altmetrics” concept was coined, many works studied the relationship of usage metrics (views, downloads, hits, etc.) and citations (Perneger, 2004; Brody et al., 2006; Bollen et al., 2009). The results showed that online dissemination favored the number of citations of research papers. Moed (2005) reported that both citations and downloads followed different time trends, but that the increase in downloads could be influenced by citations during the first months after an article is cited. Similar results were obtained by Schlögl et al. (2014), who found that most downloads were made in the publication year, while citations might take several years to accrue. Watson (2009) concluded that downloads provide a useful indicator of eventual citations after observing high correlations between both metrics.

However, it was not until the advent of academic social networks and altmetric aggregators that researchers began examining the interaction of different alternative metrics, including tweets, readers, and mentions (Priem et al., 2012; Liu et al., 2013). Thelwall et al. (2013) correlated 11 altmetric indicators and citations and found that only tweets, blog mentions, and research highlights showed a slight relationship with citations. Costas et al. (2015) reached similar results, finding weak correlations between altmetrics and citations. They also detected that only blog mentions could estimate future highly-cited publications. Ortega (2015) analyzed altmetric indicators from the main academic social networks at the author level. His results showed that bibliometric indicators correlated across platforms, while networking and usage metrics were highly dependent on their own sites.

Other studies have reflected the relationship between specific altmetric indicators and citations (Thelwall, 2016). Thus, for example, the number of tweets and retweets has been compared to citations but with different results. In some cases, tweets have been considered possible estimates of citations (Eysenbach, 2011; Shuai et al., 2012); in other cases, no significant relationship was found between both metrics (De Winter, 2015; Ortega, 2016). However, the bookmarking of articles, specifically Mendeley reader counts, has shown better results. Mohammadi and Thelwall (2014) and Mohammadi et al. (2015) found positive correlations between citations and Mendeley reader counts. Haunschild and Bornmann (2016) also observed positive correlations between normalized citations and Mendeley reader counts at the institutional level. At present, Mendeley readers are considered the best proxy for predicting highly cited articles in some disciplines (Thelwall and Sud, 2016; Maflahi and Thelwall, 2016). Other studies have also explored the ties between citations and other metrics, such as blog mentions (Shema et al., 2014), recommendations (Bornmann, 2014b; Zuccala et al., 2015), bookmarks (Lin and Fenner, 2013), and Wikipedia citations (Shuai et al., 2013). In all the cases, the results exhibited moderate correlations.

Nevertheless, few studies have assessed the time evolution of altmetric indicators. Many of these works have focused only on the increase in Mendeley reader counts in comparison to the number of citations. Maflahi and Thelwall (2016) correlated the number of readers in Mendeley and the number of citations in Scopus as a function of the age of papers. Their results showed that correlations increase as papers age. Pooladian and Borrego (2016) also analyzed the relationship between readers and citations and found that the overlap between the most frequently bookmarked and the most cited papers increased over time. Fewer publications have dealt with the longitudinal dynamics of other altmetric indicators. Eysenbach (2011) tracked articles from the *Journal of Medical Internet Research* and showed that papers that were frequently tweeted during the first three days were more likely to be highly cited. Xia et al. (2016) discovered that correlations between tweets and citations changed based on the publication year. However, no article has studied the joint evolution of altmetric indicators, which could show mutual interactions among metrics from a dynamic perspective.

3. Objectives

The main objective of this study was to analyze the life cycle of five altmetric indicators and one bibliometric indicator: Views, Downloads, Readers, Tweets, Citations, and Blog mentions. This study attempted to observe when and in which form these metrics appear over the life of academic documents. To this end, research papers were analyzed following a synchronous approach. Articles were tracked according to their publication age (in months) in order to observe the evolution of the different altmetric indicators. Data provided by PlumX were used.

In addition, correlations among the metrics were calculated as a function of the documents' age. The purpose was to observe the relationships between altmetrics over time and provide a dynamic perspective on these relationships.

The following research questions were addressed:

- When do altmetrics appear in the life cycle of a publication?
- How do altmetrics evolve over time?
- How do the life cycles of different metrics correlate among each other?

4. Methods

This study followed a synchronous approach. This means that the evolution of altmetrics was measured according to the age of the articles and not tracking altmetric events from the same set of papers throughout the duration. The advantage of this procedure is that the observation window can be reduced. For example, in a synchronous study, we need to count only the altmetrics of an article at a particular time, avoiding tracking the performance of this publication over many years. Another advantage is that we can reuse several observations of the same object at different times. For example, a document observed at different times has distinct ages. Therefore, each age can be analyzed as an independent observation. One possible drawback is that this information can be seen only from a cumulative view because it is impossible to know to what extent an observation has changed with respect to a previous time.

4.1. Data sources

PlumX: PlumX is a provider of alternative metrics. This means that this platform obtains metrics from secondary sources (for example, social networks, repositories, publishing platforms, etc.) in order to describe the performance of scholarly documents in different online environments. Created in 2012 by Andrea Michalek and Michael Buschman, this platform enables the aggregation of altmetric counts by author and organization. This allows the presentation of graphics and statistics on the online impact of researchers, departments, and universities. PlumX was selected because it offers an easy way to extract information and provides a wide and complete range of metrics about the usage, mention, and impact of documents. Specifically, the advantages of PlumX with regard to other providers are:

- It is the only platform that contains usage statistics (i.e., Views and Downloads).
- It is also the only platform that includes citations from Scopus and other services (Crossref, PubMed, etc.).
- It has a search interface that allows the filtering of results by document type and date.

However, PlumX has some disadvantages as a data provider:

- It covers only publications from institutions that contract the service. These organizations decide to make their altmetrics results available on a web page (for example, the University of Helsinki in plu.mx/helsinki/g/). However, since 2017, when PlumX was acquired by Elsevier, the coverage has been extended to every document indexed in Scopus.
- Until 2016, PlumX obtained Twitter data from a non-official provider, which could cause undercoverage of Twitter events (Jobmann et al., 2014). At present, PlumX obtains tweets via Gnip, the official provider of Twitter.

Crossref: Crossref is a consortium of academic publishers created in 2000 for the free exchange of bibliographic references and the improvement of citations across scholarly journals (Crossref, 2017). Its main product is the DOI (digital object identifier), a unique alphanumeric string assigned by a registration agency (the International DOI Foundation) to identify content and provide a persistent link to its location on the Internet (APA, 2018). Crossref contains more than 80 million records and offers a public application programming interface (API) to search and extract records. Crossref can also be used to obtain the date of the online publication of each document and check the reliability of the data provided by PlumX.

4.2. Data extraction

PlumX enables its customers to make public reports on the altmetric impact of their publications. In this way, the web pages of universities and research organizations can be freely visited. For this study, the institutional web pages of Concytec (plu.mx/concytec/g/), Coimbra Hospital and University Center (CHUC) (plu.mx/chuc/g), the International Islamic University of Malaysia (plu.mx/iium/g), Georgia Southern University (plu.mx/georgiasouthern/g/), the University of Helsinki (plu.mx/helsinki/g/), and the University of Pittsburgh (plu.mx/pitt/g/) were found and used to extract publications. In addition, publications by researchers from the home page of PlumX (plu.mx/plum/g/) were also collected. These organizations were considered because they make available their altmetrics on a plu.mx page, all are multidisciplinary (except for Coimbra Hospital), and they represent organizations from around the world.

At present, the life cycle of a research paper does not start when it is published, but earlier, when it is uploaded to the Web. Many publishing houses make available the accepted papers on their websites only when they are formatted and a DOI is assigned. Thus, the real "birth date" is not when an article is published, but when this handle is allocated. As a result, publication dates available in PlumX are not suitable to track the whole life of a research article. To solve this problem, Crossref API (http://api.crossref.org/) extract the date when the DOI identifier was assigned to each publication. For that reason, only papers with DOIs were included in this study.

From September 2016 to January 2017, publications indexed in PlumX's institutional web pages were extracted each month. This allowed us to obtain the age of a document at five different times and, therefore, to increase the size of each sample. For example, a paper published in May 2016 and extracted in the sample of September 2016 is four months old. This paper is also eight months old in the next sample of January 2017. In this way, a document has different ages according to the sample's date. The use of repeated measures is a common practice in medicine and psychology because it allows to easily broaden a sample with minimal effort. This procedure helps perform longitudinal studies before the sample reaches its maturity in long-lasting events (Baltes and Nesselroade, 1979; Hand and Crowder, 1996). Therefore, the data set in this study is composed of 5,185 publications and 19,186 observations. These observations correspond to the different ages of the 5,185 publications between September 2016 and January 2017. The age was calculated in months, subtracting the date of DOI registration from the date when the sample was taken. From these observations, only 13,636 articles contained any metric (71%).

A web crawler was designed to extract articles from PlumX's institutional web pages. Then a harvester was written to capture the altmetric events of each article in PlumX and Crossref. Both bots were designed ad hoc using an adaptation of the SQL language for extracting data from the Web. This language was used to automatically query PlumX's pages and retrieve, in a structured form, the data included in plain text html code (web scraping).

4.3. Metrics

PlumX captures a wide range of metrics from different sources. Many measure the same action, but from different websites (e.g., EPrints Downloads, SSRN Downloads, etc.). In these cases, the metrics were aggregated into a generic measure (i.e., Downloads or Views). Other metrics showed a very low incidence (e.g., Scores, Facebook comments, etc.) and were thus dismissed. These metrics are capitalized herein to be differentiated from generic meanings. Then five altmetric and one bibliometric indicators were analyzed.

Indicator	Metrics	Sources	Definition
Downloads	Downloads	Airiti Library, bepress, Dryad, DSpace, EPrints, Figshare, GitHub, institutional repositories, Pure, RePEc, SlideShare, SSRN	The number of times a publication is downloaded from different platforms
	Clicks	bit.ly	The number of times an article's URL is clicked through bit.ly
	Link outs	EBSCO databases	The number of times an article's URL is clicked

Views	HTML views Abstract views PDF views	Airiti Library, bepress, CABI, DSpace, EBSCO, EPrints, PLOS, RePEc, SSRN	The number of times an article has been viewed
Readers		Mendeley	The number of saves of a document into a user's library
Citations		Scopus	The number of times that a paper is cited by other publications
Tweets		Gnip (Twitter)	The number of tweets and retweets that mention a research paper
Blog mentions		PlumX	The number of blog posts written about one article

Table 1. Metrics, sources, and definitions of the six indicators used in this study

5. Results

5.1. Metric events

Metric events refer to when a paper is mentioned, cited, marked, read, etc., and are counted by an altmetric service. This section analyses when an altmetric or bibliometric event occurs according to the age (in months) of a document. The purpose is to observe at what age these events occur and how they evolve.

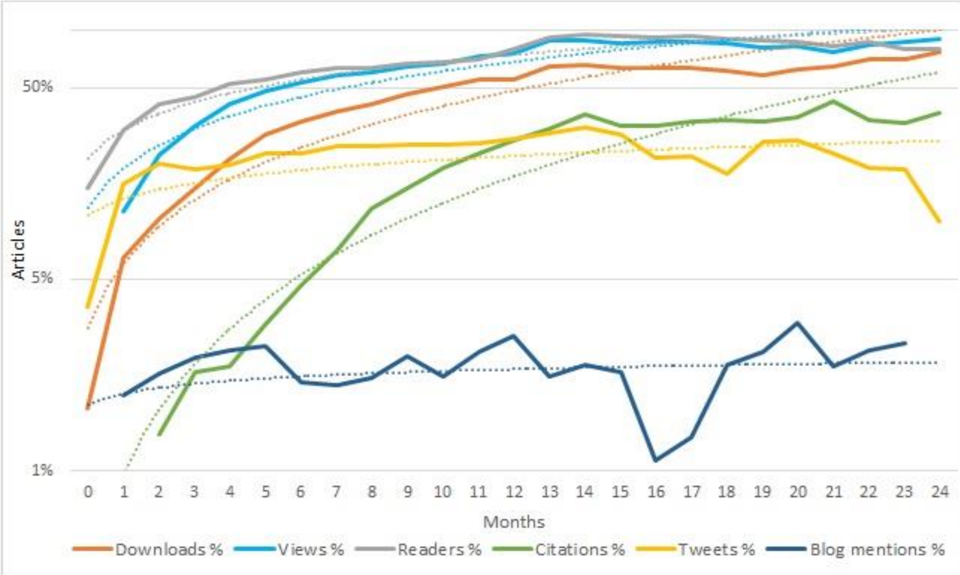

Figure 1. Distribution of the percentage of research papers that have an altmetric or bibliometric event by age (log-normal)

Metrics	*1 month*	*3 months*	*6 months*	*12 months*	*24 months*	*k*	R^2
Downloads	6.5%	14.8%	33.3%	55.3%	76.7%	1.11	0.906
Views	11.4%	31.3%	53.3%	75.1%	90.0%	0.686	0.878
Readers	30.0%	44.9%	59.9%	79.9%	80.0%	0.485	0.902
Citations	0.0%	1.6%	4.7%	26.5%	36.7%	1.916	0.924
Tweets	15.7%	18.7%	22.6%	27.1%	10.0%	0.278	0.284
Blog mentions	1.2%	1.9%	1.5%	2.5%	-	0.001	0.07

Table 2. Percentage of research papers that have an altmetric or bibliometric event by age

Figure 1 shows the percentage of documents that have been read, cited, tweeted, viewed, and mentioned by their ages in months. Table 2 displays similar information but selects the percentage of documents that have an altmetric or bibliometric event at five specific times. These distributions are expressed in percentages because the number of papers in each month is very different. All the metrics, with the exception of Blog mentions, follow a positive power law distribution ($y=ax^k$) because the number of events is aggregated as time progresses. High values of the scaling exponent (k) express slow increases, while low values mean that the distributions grow at a fast rate. Although these percentages increase over time, the figure shows some declining peaks. This is because the percentages are calculated over the number of papers in each month; therefore, these numbers can vary based on the size of the sample at each time. This effect is more appreciable in small samples (i.e., the oldest observations) where we can find greater randomness.

A brief assessment distinguished different trends concerning the appearance of metrics over time. Readers are the most frequent metric because 30% of the papers are already included in a Mendeley library during the first month, and 80% of the articles published 24 months prior have at least one reader. These percentages illustrate an elevated use of Mendeley as a social bookmarking service and the substantial coverage of its Web catalog (Maflahi and Thelwall, 2016; Ortega, 2016). Tweets are the second earliest metric, with 15.7% of papers mentioned on Twitter during the first month of life. This percentage slowly increases to 27.1% after one year of publication. Surprisingly, after two years, the proportion of tweeted papers decreases to 10%. This is the only indicator that diminishes and could be because the activity on Twitter is higher than two years ago, which could show a lower percentage of tweeted articles (Thelwall et al., 2013). In comparison with Readers, the proportion of tweeted papers is quite low but the evolution of this metric is faster during the first months (k=.278), although it suffers a rapid deceleration after six months. Views describe a continuous growth process, with 11.4% of documents viewed during the first month and 90% after two years. Its growth is slower (k=.686) than the previous but constant over time. Downloads follow a similar pattern, but with a much slower increase rate (k=1.11). Only 6.5% of the articles are downloaded at least once during the first month, reaching 76.7% after two years. However, Citations describe a much slower rhythm (k=1.92), with no articles cited during the first months. In fact, only 1.6% of the articles are cited after three months and 36.7% two years after their publication. This late appearance of citations suggests that this metric occurs in the latter part of the life cycle of articles. In some disciplines, this delay could last one or two years after the formal publishing of the paper, the time at which the citing articles are published. Finally, Blog mentions are the only indicators that do not fit any trend. This is caused by the low number of mentions, which introduces more randomness and uncertainty. Nevertheless, this metric shows a short life cycle because after one month, the percentage of mentioned papers (1.2%) increases very slowly. This last percentage suggests that, after that time, the number of new mentions drops considerably.

5.2. Average of metric events

This section analyses the distribution of the number of events by document age. This allows us to describe the evolution of these metrics throughout the months, with their differences and magnitudes. Counts were log-transformed (*ln(1+c)*) to employ averages.

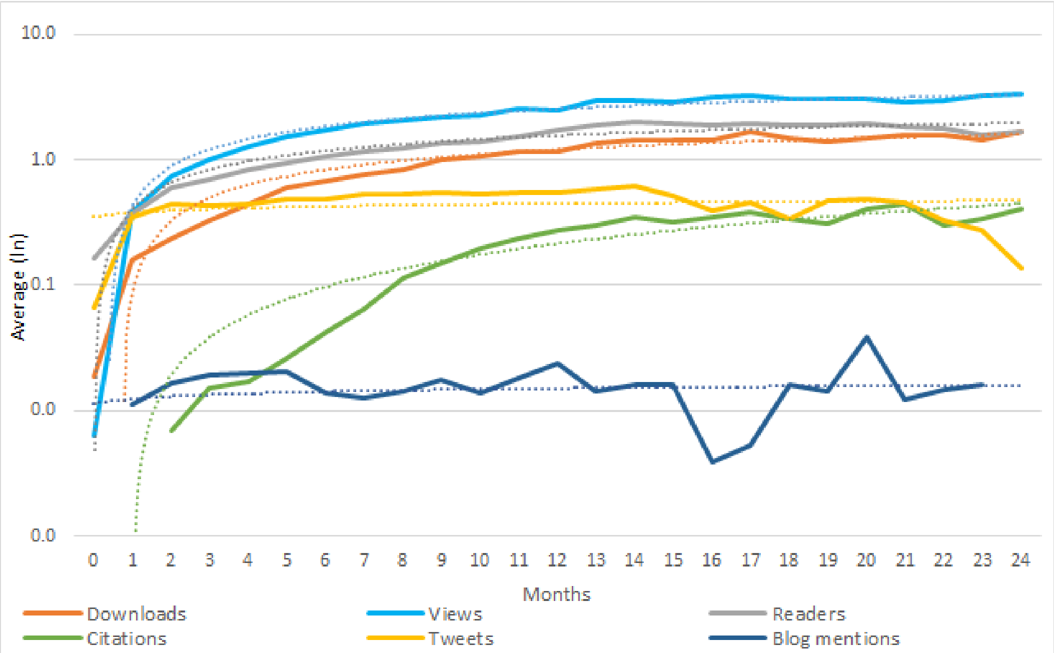

Figure 2. Distribution of the average number of altmetric and bibliometric events by document age (log-normal)

Metrics	*1 month*	*3 months*	*6 months*	*12 months*	*24 months*	*Coefficient*	*R2*
Downloads	0.16 (448%)	0.33 (304%)	0.67 (178%)	1.18 (121%)	1.68 (106%)	0.600	0.929
Views	0.38 (312%)	1.01 (180%)	1.72 (124%)	2.48 (84%)	3.32 (58%)	1.137	0.964
Readers	0.37 (184%)	0.71 (137%)	1.07 (103%)	1.70 (69%)	1.68 (65%)	0.618	0.896
Citations	0.00 (0%)	0.02 (877%)	0.04 (495%)	0.28 (197%)	0.40 (154%)	0.020	0.877
Tweets	0.35 (261%)	0.42 (237%)	0.48 (211%)	0.56 (187%)	0.14 (304%)	0.038	0.057
Blog mentions	0.01 (960%)	0.02 (807%)	0.01 (869%)	0.02 (679%)	0.00 (0%)	0.001	0.024

Table 3. Average of altmetric and bibliometric events by document age (log scale),coefficients of variation are in parentheses

Figure 2 and Table 3 show the average number of altmetric and bibliometric events by month in a logarithmic scale. In this case, the distributions follow a logarithmic trend (*y=kln(x)-a*), in which the increases are lower than the previous distributions (Section 5.1). The only exception is Citations, which follow a linear path (*y=k(x)-a*). Thus, while most of the altmetrics describe strong initial increases, the bibliometric indicator presents a slower and continued growth. Coefficients of variations decrease in all of the cases, demonstrating that the homogeneity of the samples increases as time passes.

Views are the indicator with the greatest incidence and highest values because they are the most common action and require less effort. The increase in Views is strong (k=1.137) and constant during the observation time period, doubling their averages in the first year. Downloads follow a similar pattern (k=.6) but with less strength and amount of events, suggesting that the downloading of articles is a more deliberate and responsible action. Readers describe a growth closer to Downloads (k=.618), quadrupling the average of Readers in just one year. Citations are the only metric that follow a linear trend (R^2=.877). Averages of Citations are almost nonexistent during the six first months, describing a paused increase from that time onward. In fact, the curve estimates more Citations during the first six months than in the observed ones. This later and paused growth demonstrates that the citation of papers requires much more effort and time than the altmetrics. Tweets experience a particular behavior. They show a strong initial increase in the first month (0.35) and then a slower increase during the first year (k=.038). This pattern describes an ephemeral indicator that suffers a rapid and massive appearance of events during the first months, followed by a very low incidence over subsequent months. Tweets have a very poor fit (R^2=.057), possibly because the average number of tweets decreases during the second year. This could due to increasing activity on Twitter, which means that old papers are mentioned less often than new papers (Thelwall et al., 2013). If we consider only articles with less than one year, the fit improves considerably (R^2=.864). Finally, Blog mentions describe a very irregular pattern that is not fit by any model (R^2=.024). The number of mentions remains constant over two years, suggesting that the mention of articles in blogs is not related to any time factor. However, and similar to Tweets, if we consider only six-month-old publications, the logarithmic fit improves (R^2=.71). This suggests that the mention of articles in blogs can be observable during the first six months, which is an even more ephemeral indicator than Tweets.

5.3. Correlations

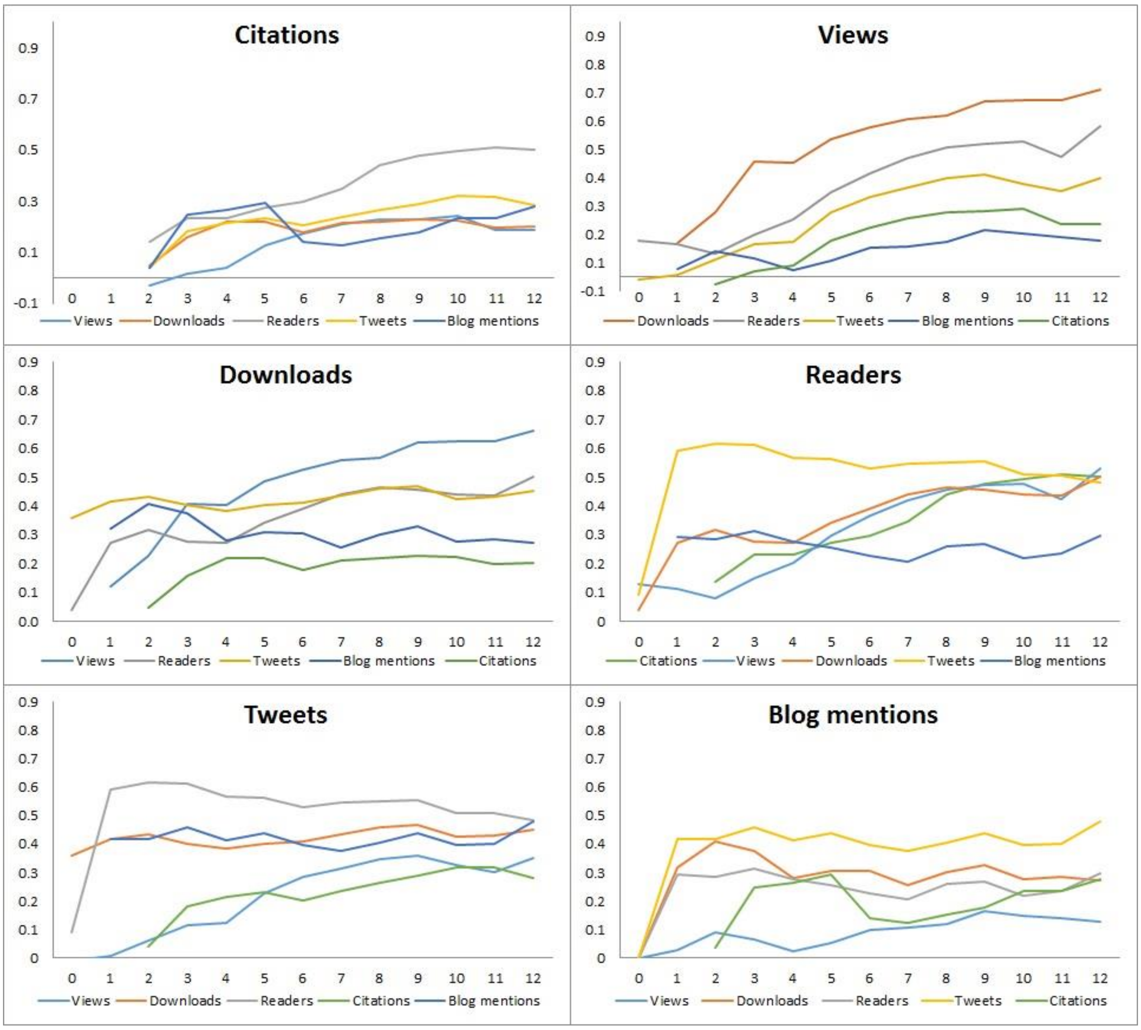

Figure 3. Pearson's correlation coefficients of altmetric and bibliometric indicators by document age

Figure 3 presents the correlations between the five altmetric indicators and one bibliometric indicator by document age. Variables were transformed to logarithms before correlations. The purpose is to see how these correlations change over time. In general, all correlations increase as months pass because there is a cumulative factor that reinforces the previous correlations. In the case of Citations, correlations increase over time, indicating that the influence of altmetrics is usually noticeable after several years. Readers are the metric that provide the best correlations with Citations, confirming previous studies (Mohammadi et al., 2015; Maflahi and Thelwall, 2016).

Views also show increasing correlations, similar to Citations. This increase could be because Views are very sensitive to other interactions, because any save, tweet, or mention may produce a view. Obviously, the best correlations are found with Downloads, the other usage metric. However, Downloads describe a different pattern.

Initially, they have high correlations with Tweets and Blog mentions, which suggest that mention indicators could significantly influence the download of papers.

Readers also show strong correlations with mention indicators (Tweets and Blog mentions) in the first months of publication. In the manner of Downloads, the number of readers of a publication could be influenced by the mention of this document in social networks. However, this statement has to be cautiously understood because both metric types occur almost at the same time and, therefore, the influence would be mutual. This relationship remains stable over time, while the correlation with Citations gradually increases to reach the best correlations after 12 months.

Tweets describe strong early correlations with Readers, Blog mentions, and Downloads that remain constant over time. These relationships confirm the immediate appearance of Tweets and their early connection with other metrics. Correlations with Views and Citations occur later and with less intensity. In the manner of Tweets, Blog mentions also correlate early with Tweets, Readers, and Downloads. This demonstrates the strong relationship between mention metrics and suggests that these indicators are very early and ephemeral.

6. Discussion

This study enables us to assess the life cycle of research papers from an altmetric and bibliometric view. The results indicate that mention indicators are the earliest metrics in a research paper, specifically Tweets and Blog mentions. Tweets increase during the first three months (18.7%) and then remain constant (22.6%), describing an abrupt increase and a short life (k=.278). The life cycle of Blog mentions is even shorter, reaching the top percentage after three months (1.9%). Both metrics present a descending trend in the average number of events, which means that mention metrics have a rapid and short life cycle. This was previously observed by Shema et al. (2014) and Groth and Gurney (2010), who found that most of the papers commented on a weblog were published in the current year. According to Twitter mentions, Eysenbach (2011) reported that more than 40% of the papers are mentioned days after the first publication. Next, Readers is the most growing metric after the mention metrics and the indicator with the highest incidence, because 30% of the papers are read after the first month of publication and 80% after two years. Similar percentages were observed by other studies (Haustein et al., 2014; Zahedi et al., 2014), confirming the high prevalence of this metric over other altmetric and bibliometric counts. Usage metrics, Views (k=.686), and Downloads (k=1.11) are the next metrics to appear. Overall, 11.4% of papers are viewed after one month and 90% after two years. These percentages are higher than Downloads (first month=6.5%; 24th month=76.7%) because the effort and interest in viewing an article are much less than downloading it. That is, viewing an article is a way to quickly browse the article's content, while a download could express a more detailed reading. This is better seen with the number of events. Articles are more viewed (.38) on average than downloaded (.16) after the first month, and this rate remains constant over the ensuing months. Finally, Citation is the final metric to appear and has the slowest growth rate (k=1.92). Thus, 1.6% of articles are cited three months after publication and 36.7% of publications are cited after two years. The low prevalence and late appearance of citations suggest that the bibliometric indicator is at the end of the life cycle of research papers and requires a substantial effort to appear (Bollen et al., 2009; Thelwall et al., 2013).

In summary, these different stages in the appearance of altmetric and bibliometric indicators provide a view into the life cycle of research papers. This life cycle starts with the mention of publications in social networks (Twitter, Facebook, blogs, etc.). Then the articles are bookmarked/saved in reference-management tools (Mendeley, CiteUlike, etc.), followed by their usage through views and downloads in publisher websites or repositories, and finishing with bibliographic citations. However, the reality may be much more complex and it is possible that these stages could overlap. The results suggest that mention metrics are situated at the starting point in the publication life cycle and they could influence the coming of the rest of the metrics. Thus, the high correlations between mention metrics and Downloads and Readers could suggest that the first ones influence the second ones. Allen et al. (2013) demonstrated that the mention of research papers in blog posts increased the number of views and downloads. Shuai et al. (2012) confirmed that early Twitter mentions favored the download of arXiv papers. Hawkins et al. (2014) observed the same influences when they analyzed tweets and downloads of articles from the *Journal of the American College of Radiology*. According to the specific relationship between mention metrics and Readers, Haustein et al. (2014) found moderate correlations between tweets and Mendeley reader counts, but did not note a direct influence. Finally, the strong and increasing correlation between Readers and Citations, in addition to the fact that the bibliometric indicator is located at the end of the document life cycle, suggests that Mendeley readers could influence the number of citations. This close relationship between readers and citations has been widely confirmed by several studies that stated that Mendeley reader counts are the best alternative metric associated with the research impact (Mohammadi et al., 2015; Maflahi and Thelwall, 2016).

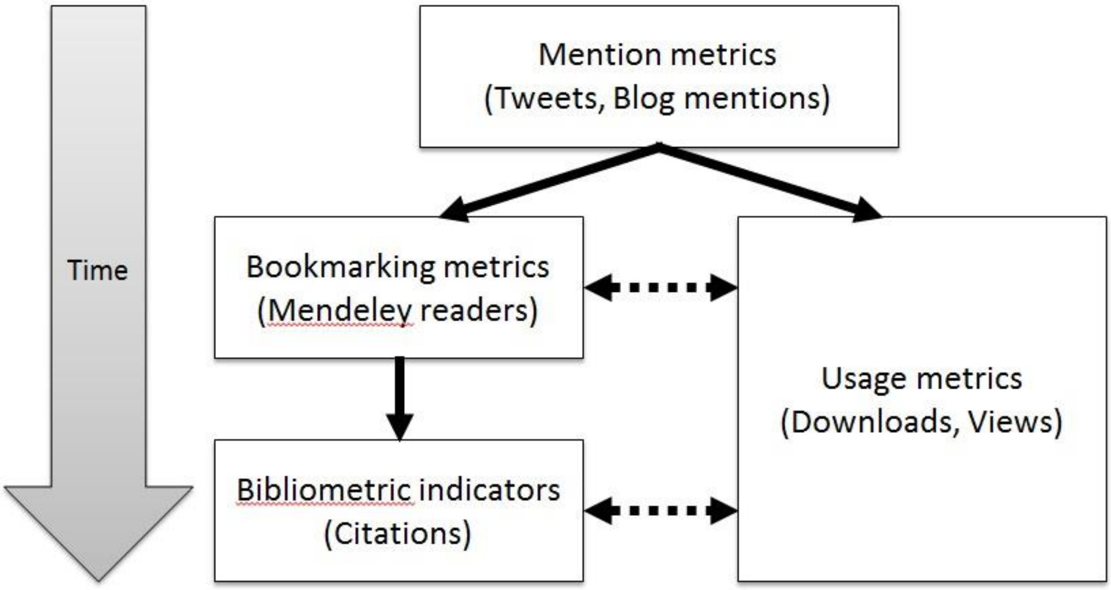

Figure 4. The life cycle of a publication from an altmetric and bibliometric view

This relationships' network and its discussion in the previous studies allow us to hypothesize a relational framework that could illustrate the evolution of the altmetric and bibliometric indicators throughout the life cycle of publications. Figure 4 is a schematic approximation derived from the empirical results of the relationships between these metrics. Thus, the first event in the life cycle of a publication is its mention in websites and social networks. Then this activates the bookmarking of the paper in reference-management sites and its usage in publisher platforms and repositories. Finally, the saving of the document into bookmarking sites might affect its future citation count. In this model, usage metrics are viewed as long-range indicators that appear in any life stage of the publication. Whenever a paper is mentioned, saved, or cited, its usage increases. Obviously, this is a hypothetical proposal originating from the results and the literature discussion, so new studies in this line would be welcome to confirm or refute this model.

7. Limitations

The principal limitation of this study results from PlumX as a data source. This altmetric provider presented some problems when it came to counting tweets in the previous years (Jobmann et al., 2014). Since April 2016, PlumX has had the support of Gnip, Twitter's official data provider. This has reinforced the quality and reliability of tweets gathered by PlumX by broadening the covered links. However, PlumX is the only source that offers usage data. This type of data has an important limitation: publications may be available through multiple sources (publisher platforms, repositories, personal home pages, social networks, etc.) that produce their own usage statistics. This may result in Downloads and Views being incomplete and unrepresentative of the global usage of these publications. This would explain why these metrics do not reach higher values. Another problem could be the delay between the time at which the event occurs and this is reported by the providers. Although this delay may be very brief, it should be considered in longitudinal studies.

According to the current study, the synchronic method allows us to observe only the evolution of metrics using articles published at different times, not measuring the changes in the same group of articles over time. This prevents us from knowing the status of articles before the observation, and therefore from measuring the changes caused. This indicates that the results are presented as a cumulative distribution. Another possible limitation could be that the current study measures the time in months, not days. Daily counts would be more precise for measuring the trend of ephemeral metrics such as tweets and blog mentions. However, this detail is impossible with our current technical means.

8. Conclusions

Several conclusions can be drawn from the results. First, the synchronic approach has made it possible to analyze the life cycle of publications from an altmetric view. The results have shown that mention metrics (Tweets and Blog mentions) are the first to appear, followed by usage (Downloads and Views) and bookmarking (Readers) metrics, and finally, bibliometric indicators (Citations).

Because of the synchronic nature of this study, the distribution of the number of documents with an altmetric and bibliometric event follows a positive power law, while the average number of altmetric and bibliometric events follows a logarithmic path. Mention metrics are the measures that increase more quickly despite having an ephemeral life cycle. Readers are the metrics with the highest prevalence and the second fastest growth. Views and Downloads show a continuous increase over time, demonstrating that these indicators have the longest life cycles. Citations are the slowest indicator and have a low prevalence, which is a result of the effort required to include a citation in a publication and demonstrates the current importance of bibliometric indicators as research impact metrics.

Correlations have shown that mention metrics are highly related to Readers and Downloads. Starting from the assumption that mention metrics are the earliest, Tweets and Blog mentions could influence the counts of Readers and Downloads. Next, the strong correlation between Readers and Citations allows us to suggest that Readers can affect the number of Citations. These interactions have led us to propose a hypothetical scheme that illustrates the appearance and influence of altmetric and bibliometric indicators, with the aim of speculating on the evolution and interaction of these metrics.

9. Acknowledgments

We would like to thank the anonymous referees for their important suggestions and to John McBurney for his helpful proofreading comments.

10. References

Allen, H. G., Stanton, T. R., Di Pietro, F., & Moseley, G.L. (2013). Social media release increases dissemination of original articles in the clinical pain sciences. PLOS ONE, 8(7): e68914.

APA (2018). APA Style: What is a digital object identifier, or DOI? http://www.apastyle.org/learn/faqs/what-is-doi.aspx

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In: J. R. Nesselroade y P. B. Baltes (Eds.), Longitudinal research in the study of behaviour and development. New York: Academic Press.

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.

Bornmann, L. (2014a). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. Journal of informetrics, 8(4): 895-903.

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

Brody, T., Harnad, S., & Carr, L. (2006). Earlier web usage statistics as predictors of later citation impact. Journal of the Association for Information Science and Technology, 57(8): 1060-1072.

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.

Crossref (2017). Crossref.org http://www.crossref.org/

De Winter, J. C. F. (2015). The relationship between tweets, citations, and article views for PLOS ONE articles. Scientometrics, 102(2): 1773-1779.

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.

Groth, P., & Gurney, T. (2010). Studying scientific discourse on the Web using bibliometrics: A chemistry blogging case study. In: WebSci10: Extending the Frontiers of Society On-Line, Raleigh, NC. http://wiki.few.vu.nl/sms/images/9/9c/Websci10-FINAL-29-4-2010f.pdf

Haunschild, R., & Bornmann, L. (2016). Normalization of Mendeley reader counts for impact assessment. Journal of Informetrics, 10(1): 62-73.

Hand, D. & Crowder, M. (1996). Practical longitudinal data analysis. London: Chapman & Hall.

Haustein, S. (2016). Grand challenges in altmetrics: heterogeneity, data quality and dependencies. *Scientometrics*, *108*(1), 413-423.

Haustein, S., Bowman, T. D., & Costas, R. (2015). Interpreting" altmetrics": viewing acts on social media through the lens of citation and social theories. *arXiv preprint arXiv:1502.05701*.

Haustein, S., Larivière, V., Thelwall, M., Amyot, D., & Peters, I. (2014). Tweets vs. Mendeley readers: How do these two social media metrics differ? IT-Information Technology, 56(5): 207-215.

Hawkins, C. M., Hillman, B. J., Carlos, R. C., Rawson, J. V., Haines, R., & Duszak, R. (2014). The impact of social media on readership of a peer-reviewed medical journal. Journal of the American College of Radiology, 11(11): 1038-1043.

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

Lin, J., & Fenner, M. (2013). The many faces of article-level metrics. Bulletin of the American Society for Information Science and Technology, 39(4): 27-30.

Liu, C. L., Xu, Y. Q., Wu, H., Chen, S. S., & Guo, J. J. (2013). Correlation and interaction visualization of altmetric indicators extracted from scholarly social network activities: dimensions and structure. Journal of medical Internet research, 15(11): e259.

Maflahi, N., & Thelwall, M. (2016). When are readership counts as useful as citation counts? Scopus versus Mendeley for LIS journals. Journal of the Association for Information Science and Technology, 67(1): 191-199.

Moed, H. F. (2005). Statistical relationships between downloads and citations at the level of individual documents within a single journal. Journal of the Association for Information Science and Technology, 56(10): 1088-1097.

Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (2015). Who reads research articles? An altmetrics analysis of Mendeley user categories. Journal of the Association for Information Science and Technology, 66(9): 1832-1846.

Nature Materials (2012). Alternative metrics. Nature Materials, 11(907): 1476-1122

Ortega, J. L. (2015). Relationship between altmetric and bibliometric indicators across academic social sites: The case of CSIC's members. Journal of Informetrics, 9(1): 39-49.

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.

Ortega, J. L. (2016). Social Network Sites for Scientists: A Quantitative Survey. Cambridge: Chandos Publishing

Perneger, T. V. (2004). Relation between online “hit counts” and subsequent citations: prospective study of research papers in the BMJ. British Medical Journal, 329(7465): 546-547.

Priem, J., & Hemminger, B. H. (2010). Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*, *15*(7).

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

Pooladian, A., & Borrego, Á. (2016). A longitudinal study of the bookmarking of library and information science literature in Mendeley. Journal of Informetrics, 10(4): 1135-1142.

Schlögl, C., Gorraiz, J., Gumpenberger, C., Jack, K., & Kraker, P. (2014). Comparison of downloads, citations and readership data for two information systems journals. Scientometrics, 101(2): 1113-1128.

Shema, H., Bar-Ilan, J., & Thelwall, M. (2014). Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics. Journal of the Association for Information Science and Technology, 65(5): 1018-1027.

Shuai, X., Pepe, A., & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: Article downloads, twitter mentions, and citations. PloS one, 7(11): e47523.

Shuai, X., Jiang, Z., Liu, X., & Bollen, J. (2013). A comparative study of academic and Wikipedia ranking. In Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries (pp. 25-28). ACM.

Sugimoto, C. (2014). The Failure of Altmetrics. https://scholarworks.iu.edu/dspace/bitstream/handle/2022/19208/dlbb-sugimoto.pdf

Sugimoto, C. R., Haustein, S., Ke, Q., & Larivière, V. (2016). Social Media Metrics as Indicators of Broader Impact. OECD Blue Sky III Forum on Science and Innovation Indicators. Gent (Belgium), September.

Tattersall, A. (2016). Road map: from Web 2.0 to altmetrics. In: Tattersall, A. Altmetrics: A practical guide for librarians, researchers and academics. London: Facet Publishing, pp. 224

Thelwall, M. (2016). Interpreting correlations between citation counts and other indicators. Scientometrics, 108(1): 337-347.

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.

Thelwall, M., & Sud, P. (2016). Mendeley readership counts: An investigation of temporal and disciplinary differences. Journal of the Association for Information Science and Technology, 67(12): 3036–3050.

Watson, A. B. (2009). Comparing citations and downloads for individual articles at the Journal of Vision. Journal of Vision, 9(4): i-i.

Wouters, P., & Costas, R. (2012). Users, narcissism and control – tracking the impact of scholarly publications in the 21st century. Rochester, NY: SURF Foundation. www.surf.nl/binaries/content/assets/surf/en/knowledgebase/2011/Users+narcissism+and+control.pdf

Xia, F., Su, X., Wang, W., Zhang, C., Ning, Z., & Lee, I. (2016). Bibliographic Analysis of Nature Based on Twitter and Facebook Altmetrics Data. PloS one, 11(12): e0165997.

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.

Zuccala, A. A., Verleysen, F. T., Cornacchia, R., & Engels, T. C. (2015). Altmetrics for the humanities: Comparing Goodreads reader ratings with citations to history books. Aslib Journal of Information Management, 67(3): 320-336.