arXiv:1203.4745v1 [cs.DL] 20 Mar 2012

Altmetrics in the Wild: Using Social Media to Explore Scholarly Impact

Jason Priem*, Heather A. Piwowar**, Bradley M. Hemminger*

The first two authors contributed equally to this paper.

* School of Information and Library Science , University of North Carolina at Chapel Hill

** National Evolutionary Synthesis Center (NESCent)

Abstract

In growing numbers, scholars are integrating social media tools like blogs, Twitter, and Mendeley into their professional communications. The online, public nature of these tools exposes and reifies scholarly processes once hidden and ephemeral. Metrics based on this activities could inform broader, faster measures of impact, complementing traditional citation metrics. This study explores the properties of these social media-based metrics or "altmetrics," sampling 24,331 articles published by the Public Library of Science. We find that that different indicators vary greatly in activity. Around 5% of sampled articles are cited in Wikipedia, while close to 80% have been included in at least one Mendeley library. There is, however, an encouraging diversity; a quarter of articles have nonzero data from five or more different sources. Correlation and factor analysis suggest citation and altmetrics indicators track related but distinct impacts, with neither able to describe the complete picture of scholarly use alone. There are moderate correlations between Mendeley and Web of Science citation, but many altmetric indicators seem to measure impact mostly orthogonal to citation. Articles cluster in ways that suggest five different impact “flavors,” capturing impacts of different types on different audiences; for instance, some articles may be heavily read and saved by scholars but seldom cited. Together, these findings encourage more research into altmetrics as complements to traditional citation measures.

Introduction

Social media tools are beginning to affect the research workflow; growing numbers of scholars discuss and share the research literature on Twitter, organize it in social reference managers like Mendeley and Zotero, and review it in blogs, article comments, and post-publication peer review services like Faculty of 1000. While this incorporation of social media into the research workflow may improve the overall responsiveness and timeliness of scholarly communication, it also has a powerful secondary advantage: exposing and fixing scholarly processes once hidden and ephemeral. The daily work of scholars is moving online. As it does, the background of scholarship—the dog-eared manuscripts and hallway conversations—is pushed out on to the stage.

Traditionally, this stage has been occupied almost exclusively by citations. Researchers have turned to these “pellets of peer recognition” [1] to help track the flow of scholarly ideas, and have been rewarded with an impressive and growing understanding of how scholarship is transmitted and adopted. However, citation tracking has never been able to follow the less visible—but often more important [2]—threads of invisible colleges, woven through personal connections and informal communications [3]. The increasing online reification of these ethereal threads gives us a chance to fill this gap, building a deeper, richer map of scholarly information flows.

So-called “alternative metrics” or “altmetrics” [4] build on information from social media use, and could be employed side-by-side with citations—one tracking formal, acknowledged influence, and the tracking the unintentional and informal “scientific street cred” [5]. Altmetrics could deliver information about impact on diverse audiences like clinicians, practitioners, and the general public, as well as help to track the use of diverse research products like datasets, software, and blog posts. The future, then, could see altmetrics and traditional bibliometrics presented together as complementary tools presenting a nuanced, multidimensional view of multiple research impacts at multiple time scales.

Before this can be done, however, it is important to describe and characterize social media tools as sources for metrics. Scholarly users of tools like Mendeley and Twitter remain a minority among their (often skeptical) peers; is there in fact enough data available to construct meaningful metrics? How is data distributed across tools, users, and time? How do altmetrics relate to accepted citation measures? There is a need for exploratory research to answer these and other questions, presenting early descriptive findings to guide more focused inquiry. This paper presents the results of our study investigating these questions.

Literature review

Although citation metrics have become increasingly important in the measure of research impact, they are not without flaws. Citations take time to accumulate; the “citation latency” of even high-impact articles may by “1-2 years or even longer” [6]. Citations have been used to measure only one type of research product, the peer-reviewed article; this leaves out other important products such as datasets [7–9]. Similarly, citations measure impact only on that minority of an article’s audience with the means and inclination to cite it in the literature, leaving out clinicians, the general public, and even most scientists —by some estimates, "Only about 15 to 20% of scientists in the United States have authored a refereed article" [10]. The most popular citation metric, Thomson Scientific’s Journal Impact Factor, has been a lightning rod for critics, who have argued that it fails to correlate with impact at the article level [11], is arbitrary and irreproducible [12], and is easy to game [13].

The shortcomings of citation metrics are not news. Scientometricians and others have for decades worked to create more realistically diverse measures of research impact, painstakingly gathering diverse indicators including patents [14], acknowledgements [15], doctoral committee membership [16], and many others. The Web brought with it the potential for yet more metrics, gathered electronically and at scale; swapping citation indexes for search engines, practitioners of “webometrics” [17] gathered indicators from hyperlinks and online mentions on webpages [18–20], syllabi [21], presentations [22], and other online resources. At the same time, researchers began to investigate the usage records created by online article delivery as yet another source of impact evidence [6,23].

These new approaches, however, have not completely filled the gap between citation metrics and real-world impact. Webometric approaches that rely on search engines are fundamentally limited by terms of use restrictions on automated mining of results; consequently, webometrics must be compiled manually, on a custom basis. This approach does not work at large scale. Usage metrics are automatically gathered, but also scale poorly for a similar reason: publishers are unwilling to release data for widespread use [24]. However, there is growing interest in a new type of Web-based metric that may scale more effectively and present an even broader picture of scholarly impact: metrics based on social media activities.

Social media tools are becoming increasingly important in scholars’ workflows, as several recent studies have made clear: [25] report that 13% of UK academics frequently use Web 2.0 in novel forms of scholarly communications, while [26] find that 80% of scholars have social media accounts. Carpenter et al. report that around 10% of UK doctoral students "use and value" Twitter for research, and [27] also report that social media tools are affecting the scholarly workflow. These tools include social reference managers, [28] Twitter, [29] blogs, [30] bookmarking services, [31] and more. Importantly, these tools do not *create* new types of scholarly practice so much as they *facilitate*existing practice. Social reference managers like Mendeley, for example, are an extension of paper-based bibliography collections academics have maintained for centuries, while Twitter facilitates the sort of informal conference chats that have long vivified the academy’s invisible colleges [32].

Consequently, many have suggested that these tools open a valuable window on heretofore hidden scholarly processes, finally fulfilling the early Web’s promise to “...give substance to modes of influence which have historically been backgrounded in narratives of science” [33]. Some suggest using social media to assess scholars’ authority [34], while others propose social media as a sort of “soft peer review” [35,36]. Social media tools typically offer open APIs facilitating large-scale use, unlike most search engines and most journal publishers, offering potential for a “scientometrics 2.0” [37]. More recently, “altmetrics” has been used to describe “alternative” metrics based on social media, with proponents touting their “speed, richness, and breadth” [38]. As these rallying calls have gone out, several tools have emerged to gather altmetrics, including (in approximate order of appearance) PLoS Article-Level Metrics, ReaderMeter, CitedIn, total-impact, altmetric.com, and ScienceCard.

A growing body of research has begun to test the claims of altmetrics' value, examining two main questions: first, "how much altmetrics data exists, and how is it distributed," and second, "what do altmetrics measure, and how does it correlate with citation?" Inquiries into the first question have tended to find scholarly use of social media relatively rare, but significant and growing. Examinations of journal-based commenting systems or "rapid responses," uncover a very skewed distribution: half of sampled journals in [39] have not attracted a single comment on any article, though a few active journals attracted comments on most articles. Although only an estimated 2.5% of scholars actively use Twitter, this number is growing steadily [40], and around 95% of sampled articles from the arXiv preprint repository have been tweeted [41]. Usage of scientific terms on Twitter is modest but reveals interesting usage patterns [42]. Social reference manager applications have attracted significant use. Mendeley has been particularly successful, advertising on their homepage that over 1.6 million users have collected 161 million documents; over 90% of Nature and Science articles are included in Mendeley collections [43] (although this high number no doubt in part reflecting the popularity of these publications). In contrast, find only 3.5% of physics articles bookmarked in CiteULike [24].

Researchers have also looked beyond counts of social media users and uses to assess types of use. Do social media really host scholarship, or just idle chatter? What do social media uses or citations mean? Twitter, for instance, is often perceived as a place to do little more than discuss one's lunch. This so-called "cheese sandwich problem" [44] seems to be significantly overstated, however, as several studies have presented evidence of scholars using the service to enrich academic conferences [45,46] as well as cite scholarly literature [47,48]; as many as ⅓ of tweets from scholars contain scholarly content [40]. Examinations of social reference managers have focused more on the potential uses of available data, demonstrating value in collaborative filtering approaches [49,50] and use of tag-based folksonomies [51–53]. Post-publication peer review service Faculty of 1000 (F1000) has attracted several studies of its article rankings, with inconsistent findings: [54] reports that the reviewers failed to spot many significant articles, while [55] emphasize that reviewers uncovered many "hidden gems."

Several studies have attempted to relate the unknown impacts of social media to the more accepted yardstick of citation. Citations in Wikipedia correlate with those in the Journal Citation Report, though slightly over-citing high-impact journals [56]. Similarly, while discussions in scholarly blogs pull from diverse sources, they tend to focus on articles published in highly-cited journals [57,58]. In a sample of economics blogs, [59] report correlation between pageviews and per-year citation of blogs' contributors is .56. Several studies have compared traditional citation with usage or bookmarking in social reference managers, suggesting weak to moderate correlation. Correlations of .56 and .54 are reported between Mendeley users and Web of Science citations of Science and Nature articles (.36 and .30 for CiteULike) [43], as well as a correlation of .76 between Mendeley users and citations for top biology papers [60]. Using a more heterogeneous dataset of 168,109 physics papers from 45 journals, [24] reports a correlation of .21 between saves into CiteULike and Web of Science citations, although the top paper was the same for both metrics. Data from the open-access publisher Public Library of Science's (PLoS) Article-Level Metrics (ALM) reveal a correlation of .2 between CiteULike saves and Scopus citations [61]. More recently, [62] examined articles in the Journal of Medical Internet Research that had attracted links from tweets (or "Twitter citations" [47]). Articles in the most-tweeted quartile after one week were eleven times more likely to be in the most-cited quartile after two years. For the top 100 most-tweeted articles in a sample from arXiv repository, log of tweet counts correlates weakly with log of citation counts (r=.33) [41].

Methods

Data collection

In November 2010 we collected identifiers for all 24,331 articles published to date in the seven journals of open-access publisher Public Library of Science (PLoS) and gathered a set of approximately 1.8 million altmetrics *events*, defined as a specific action on a specific article. Each event was of a certain *event type*, such as "Delicious bookmark" or "PDF download." Using common scientometrics terminology, event types could also be called *indicators*. We obtained counts for a diverse set of event types as described in Table 1 and Table S1, using a combination techniques, including:

searching the PLoS ALM API,

writing custom scripts to search APIs of external data sources and

writing a custom script to crawl and download PLoS comments

In November 2010, and again in December 2011, we gathered Thomson Reuters Web of Science citation counts manually through the Web of Science website interface.

Table 1

API	data on articles	event-level data?
Delicious	bookmarks (includes author, timestamp, tags)	yes
Wikipedia	counts of links out to articles	counts only
Mendeley	counts of users bookmarking articles (includes location, discipline of users)	counts only
Facebook	counts of clicks, likes, shares, and comments	counts only
Twitter via BackTweets *	linking tweets (includes author, timestamp, text)	yes
CiteULike via PLoS ALM	saves into reference library (includes author, timestamp, tags)	yes
Nature Blogs via PLoS ALM *	linking blog posts (incl. timestamp, author, url)	yes
Postgenomic blogs via PLoS ALM *	linking blog posts (incl. timestamp, author, url)	yes
Research Blogging blogs via PLoS ALM	linking blog posts (incl. timestamp, author, url)	yes
Scopus citations via PLoS ALM	counts of citations	counts only
CrossRef citations via PLoS ALM	articles citing (includes bibliographic information)	yes
PubMed Central citations via PLoS ALM	counts of citations (includes PubMed IDs)	counts only
Monthly pdf, html, and xml views/downloads via PLoS ALM	counts of pageviews and downloads, by month	counts only
PLoS comments count	comments on articles (incl. timestamp, author)	yes
Web of Science citations	articles citing (includes bibliographic information)	yes
Faculty of 1000 rankings	numerical ratings	counts only

Table 1: Citation and altmetrics data collected for this study. An (*) denotes metrics no longer available. For information on access methods and rate limits, see supplemental materials.

Normalization and transformation

Simply comparing raw counts is highly affected by confounding trends, such as the length of time it took for various metrics to accumulate, and changes in overall user bases of various tools over the seven-year period of our sample. Therefore, we need a way to isolate altmetrics activity on articles from other trends, in order to compare apples to apples.

To do this, we used a technique to express event counts as a percentage of what count we would *expect* for that article, compared to other articles published 1) in the same journal and 2) at around the same time. More precisely, we smooth the number of events based on the average metric value within a given time window, 180 days before and after the publication date of the current article. We apply a standard digital signal processing technique to avoid inducing noise from outliers suddenly appearing then disappearing in the averaging window: weighting the points by a Hamming window [63] entered on the current publication date. These weights are shown in Figure 1.

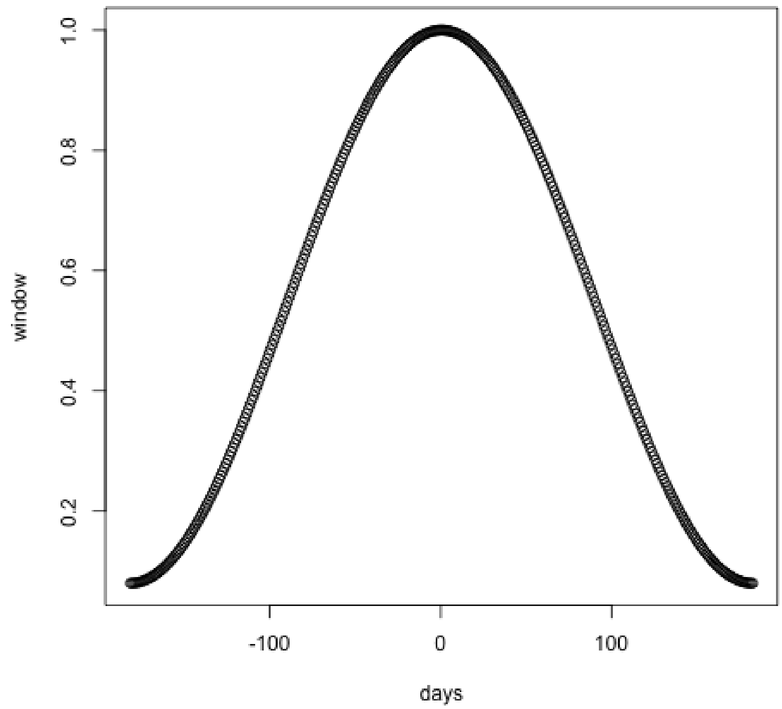


Figure 1: *Hamming window used for weighting metric values included in smoothed averages, centered on article publication date and extending 180 days before and after.*

Because different communities have different levels of social media adoption and readership levels (see http://www.plosone.org/static/journalStatistics.action), we needed to also normalize by journal; when we normalized an article, we only included other articles from the same journal in the normalization window. The normalized values were set to “Missing” for metric values without at least 25 nonzero metric points in the same journal in the surrounding year.

Finally, we know that the distribution of these counts is heavily skewed, with most articles receiving ones or zeroes (depending on the metric) and just a few with very high counts. We log-transformed the normalized counts to even out the distributions (adding 1 to all counts since there is no log of zero). Since F1000 evaluation scores are not count data, they were not log-transformed.

Statistical analysis

Data was analyzed using the R statistics environment version 2.12.0 [64] with the following libraries: clValid [65], ggdendro [66], ggplot2 [67], GPArotation [68], gplots [69], nFactors [70], plyr [71], psych [72], reshape2 [73], rms [74], RWeka [75], signal [76], sqldf [77], xtable [78], and zoo [79].

Correlations between metrics and citation count were computed on unnormalized data for articles published in 2010. We applied the Spearman nonparametric ranking correlation method to estimate correlations.

The overall correlation matrix was computed across all articles, calculating Pearson correlation coefficients on pairwise-complete normalized and transformed indicator values. First-order exploratory factor analysis was performed with the fa function in the psych library on this overall correlation matrix, using the minimum residual (minres) solution and a promax oblique rotation.

We used k-means cluster analysis to group papers based on patterns of impact represented by a subselection of indicators. K-means cluster centers were calculated for PLoS ONE articles published before 2010. Cluster membership was assigned for PLoS ONE papers published in 2010 based on Euclidian distance to these centers, and rules to assign these same cluster memberships were derived using the Weka JRip algorithm [80] through rWeka. Evaluation of the accuracy of rules to predict k-means cluster membership was done through cross-validation using the evaluate_Weka_classifier function.

Data and code availability

Datasets and statistics scripts are availableonGitHub (https://github.com/jasonpriem/plos_altmetrics_study) and from the Dryad Digital Repository ([citation to be included pending manuscript acceptance]) under a CCZero (http://creativecommons.org/publicdomain/zero/1.0/) waiver.

Results and discussion

In this section, we first present an overview for the dataset, followed by results and discussion structured around our four main research questions:

- How much and what kind of altmetrics data exist?
- How are altmetrics distributed over time?
- How do altmetrics relate to one another and to traditional citations?
- Can we cluster articles of different impact types using altmetrics?

Dataset attributes

We collected altmetrics data for 24,331 PLoS publications between 2003-08-18 and 2010-12-23. Editorials, perspectives, etc. were excluded from the analysis dataset, resulting in 21,096 research articles.

The research articles were distributed across the seven PLoS journals as shown in Table 2.

	pbio	**pcbi**	**pgen**	**pmed**	**pntd**	**pone**	**ppat**	**NA**
	PLoS Biology	PLoS Computational Biology	PLoS Genetics	PLoS Medicine	PLoS Neglected Tropical Diseases	PLoS ONE	PLoS Pathogens	
2003	33	0	0	0	0	0	0	33
2004	174	0	0	13	0	0	0	187
2005	176	55	66	86	0	0	36	419
2006	184	136	173	145	0	138	97	873
2007	203	203	210	135	26	1229	173	2179
2008	197	251	325	110	133	2716	271	4003
2009	181	343	436	76	178	4399	423	6036

2010	177	363	409	78	284	5596	459	7366
Sum	1325	1351	1619	643	621	14078	1459	21096

Table 2: Distribution of articles by journal and publication date.

As shown previously in Table 1, we collected data from a wide variety of sources. There were two main types of data: events and event counts. Events were actions applied to articles, such as bookmarks or tweets. Events included metadata such as timestamps and creator identifiers; this supported fine-grained examination of, for instance, number of events within one week of article publication or mean number of event creators per article. Event counts were simply the total number of events of that type to date.

One particularly distinct type of event was downloads, separated into PDF, HTML, and XML. The latter is mostly for machine consumption and is consequently disregarded in most of the following analysis. HTML "downloads" are likely browser views in almost all cases. Downloads are binned into monthly download "events," whose magnitude represents the number of views that month.

How much and what kind of altmetrics data exist?

Unsurprisingly, nearly every article had at least some PDF and HTML downloads. According to the PLoS article-level metrics website, these statistics do not include robot or crawler downloads. A full 80% of articles had at least one bookmark on Mendeley while 31% had been bookmarked on CiteULike; this finding is in keeping with earlier suggestions that the Mendeley community is significantly more active than the CiteULike community [43]. In December 2010, 65% of PLoS research articles had received one citation in Web of Science and 50% of articles had citations from papers in PubMed Central (a free source of citation information). By December 2011, more than 92% of research articles had received at least one citation in Web of Science (90% for PLoS ONE).

As shown in Figure 2, other types of events had less activity. Ten to twelve percent of articles had been bookmarked on Delicious, tweeted, shared on Facebook, or received a comment on the PLoS website (although note that actual tweet counts are likely higher since Twitter data was incomplete). About 7.5% of articles were the topic of a blog post or had received an F1000 rating. About 5% of articles had been cited on Wikipedia, Liked on Facebook, or Commented on in Facebook.

There are two groups evident in the activity graph, with the separation occurring between Facebook (15% of articles shared) and saves into CiteULike (30% of articles). This may be explained by the novel kinds of work required by some event types. The high-activity group tracks activities – reading, saving, citing – that are requirements of traditional scholarly communication, in contrast to the low-activity group tracking new communication modes unrewarded in the traditional publication cycle.

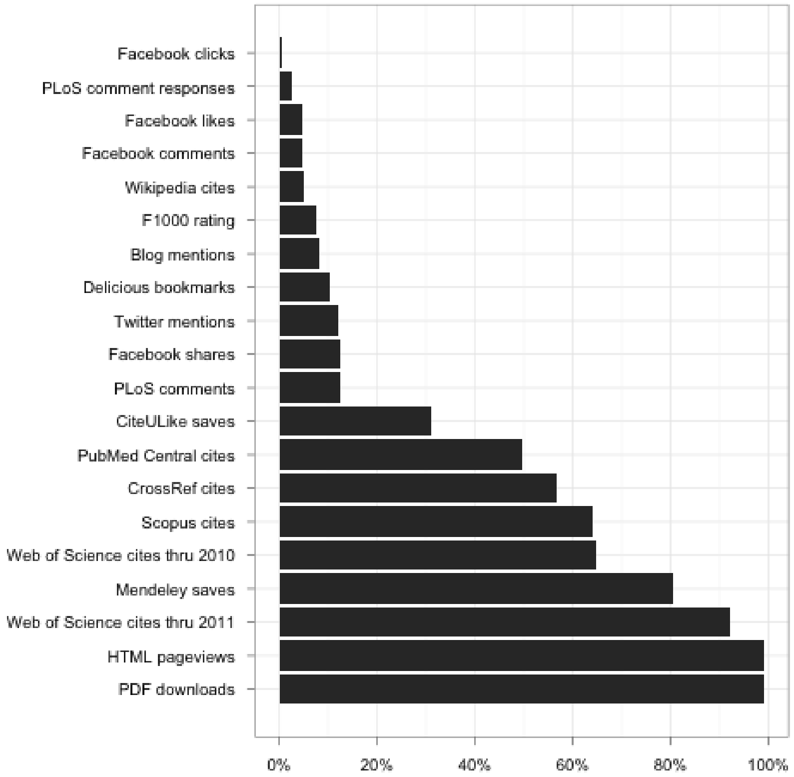


Figure 2: *Proportion of articles with at least one event, by metric.*

Attention across the metrics was not concentrated on only a few papers. We defined an "engaged indicator" for a given paper as an indicator having a value of at least one for that paper. For example, an article with no events except 100 PDF views and a single tweet would have two engaged indicators. Counting engaged indicators per article is a good way to see whether attention to articles is spread out over multiple systems.

Figure 3 shows a histogram of papers by number of engaged indicators (citations are represented here with just the 2010 Web of Science source). Half of the papers in our dataset had at least four engaged indicators – usually HTML page views, PDF downloads, maybe saves into Mendeley and a citation. A quarter of the papers had at least five engaged indicators in our dataset; 10% of papers received attention from six or more event types.

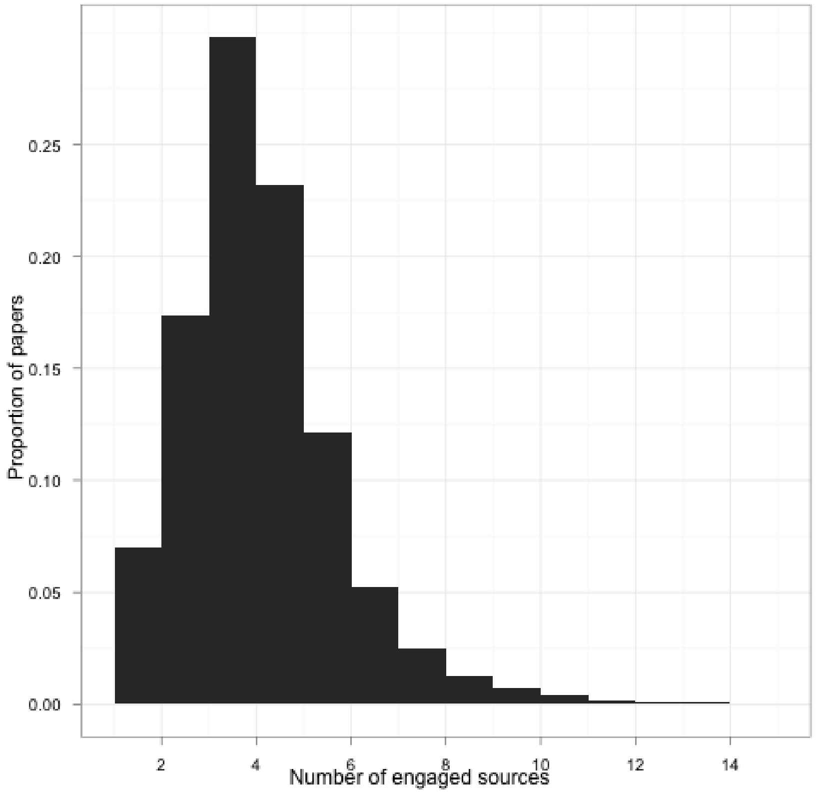


Figure 3: *Histogram of papers by the number of metrics with at least one event.*

As described in Table 1, some of the indicators had associated qualitative information while others are simply event counts.

Collected qualitative information differed between services, with some sources returning much richer information. This is in keeping with previous findings that social tags on research articles “add a third layer of perception besides the author and indexer perspectives” [51]. Figure 4 demonstrates this for a single exemplar article, comparing tagclouds of tags returned by PLoS metadata, Delicious, and CiteULike.

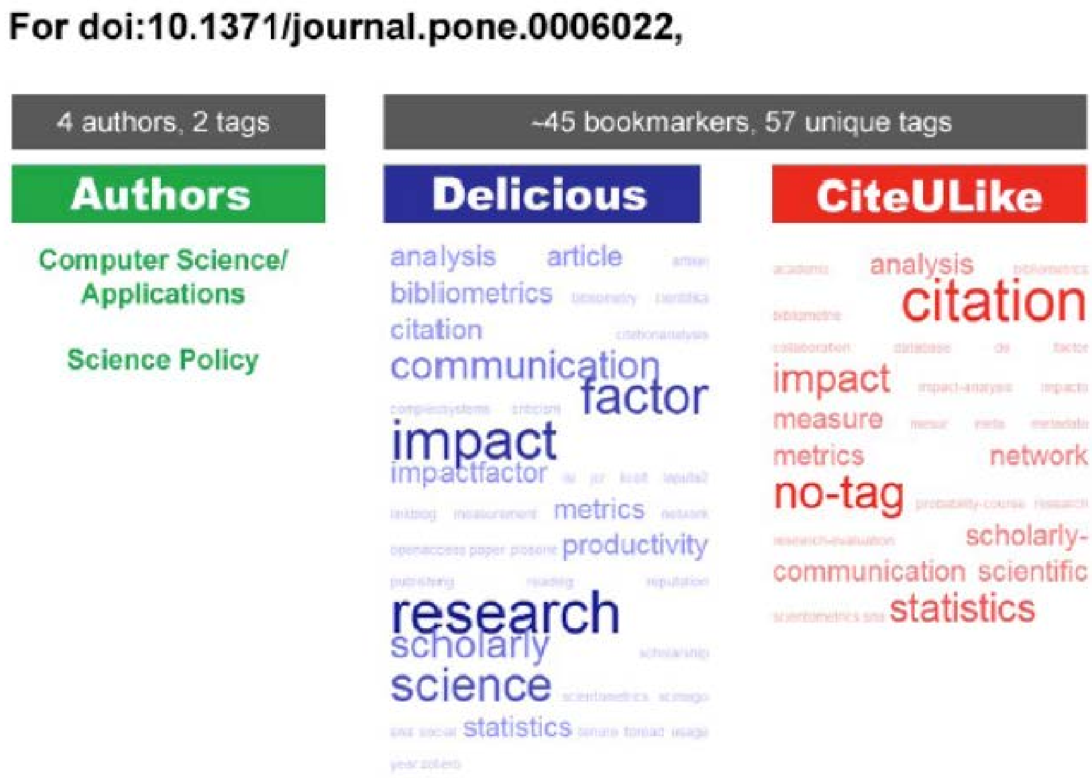


Figure 4: *Illustration of tag metadata captured by PLoS, delicious, and citeULike for one article.*

In Figure 5, the number of tweets and bookmarks by content creators roughly approximates a straight line on a log-log plot. This familiar distribution indicates that a few power users post the lion's share of events (CiteULike saves, Delicious bookmarks, or tweets) on articles in the dataset, while the majority of users reside in the "long tail," contributing just a few events.

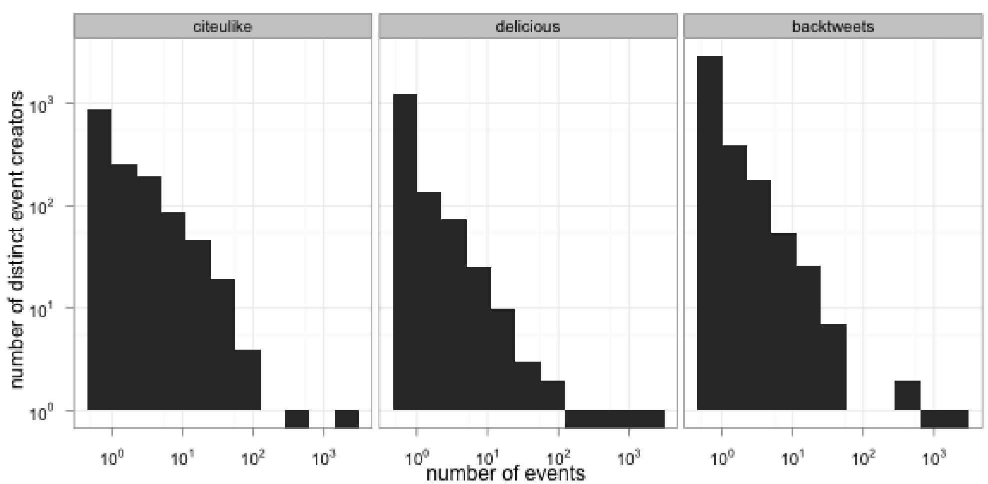


Figure 5: *Histogram of number of distinct PLoS-related events from each creator, by service (log-log scale, PLoS ONE). The bars that drop below the 100 line represent the value of zero.*

How are metrics distributed over time?

We first examined the delay between article publication and different types of events, following in the footsteps of earlier such investigations of download [6] and citation latency [81] as well as more recent examinations of latency between article publication and tweets [41,47,62]. Figure 6 shows the results. When we look at the timing of events in aggregate across the lifecycle of an article, we can see that the rates of bookmarking, commenting, and sharing decay more quickly after publication than website views and PDF downloads. CiteULike saves do not drop off quite as quickly, which may be related to its prominent position within scholars' document workflows; this bears additional investigation.

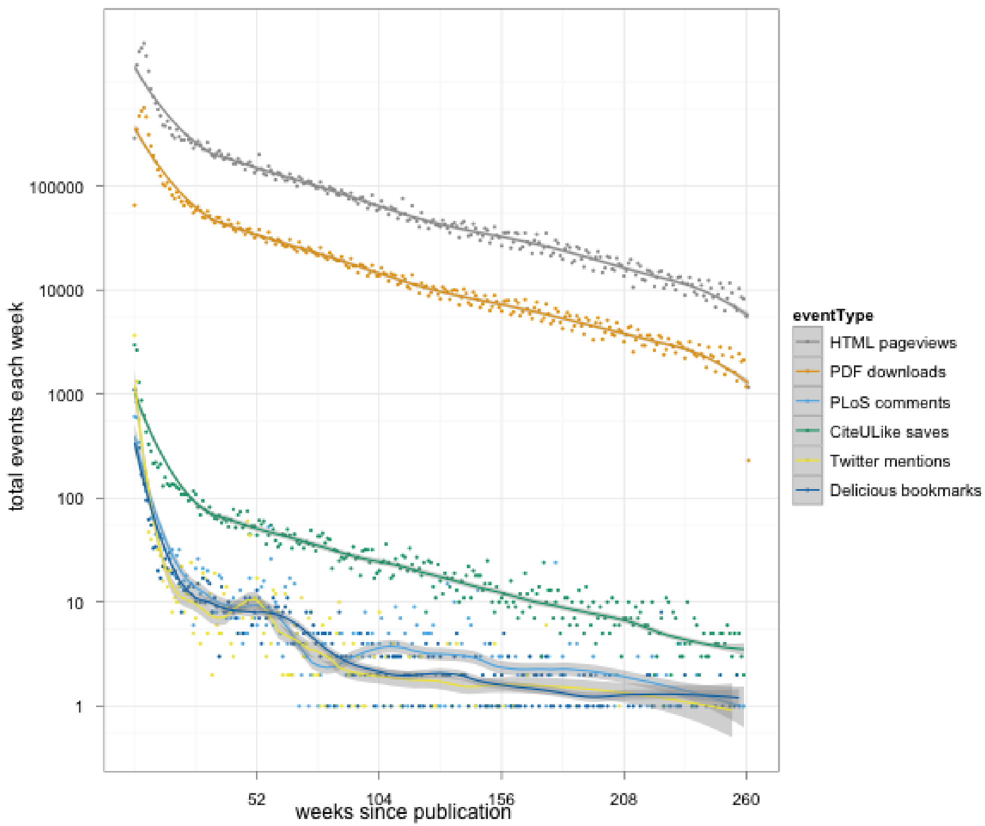
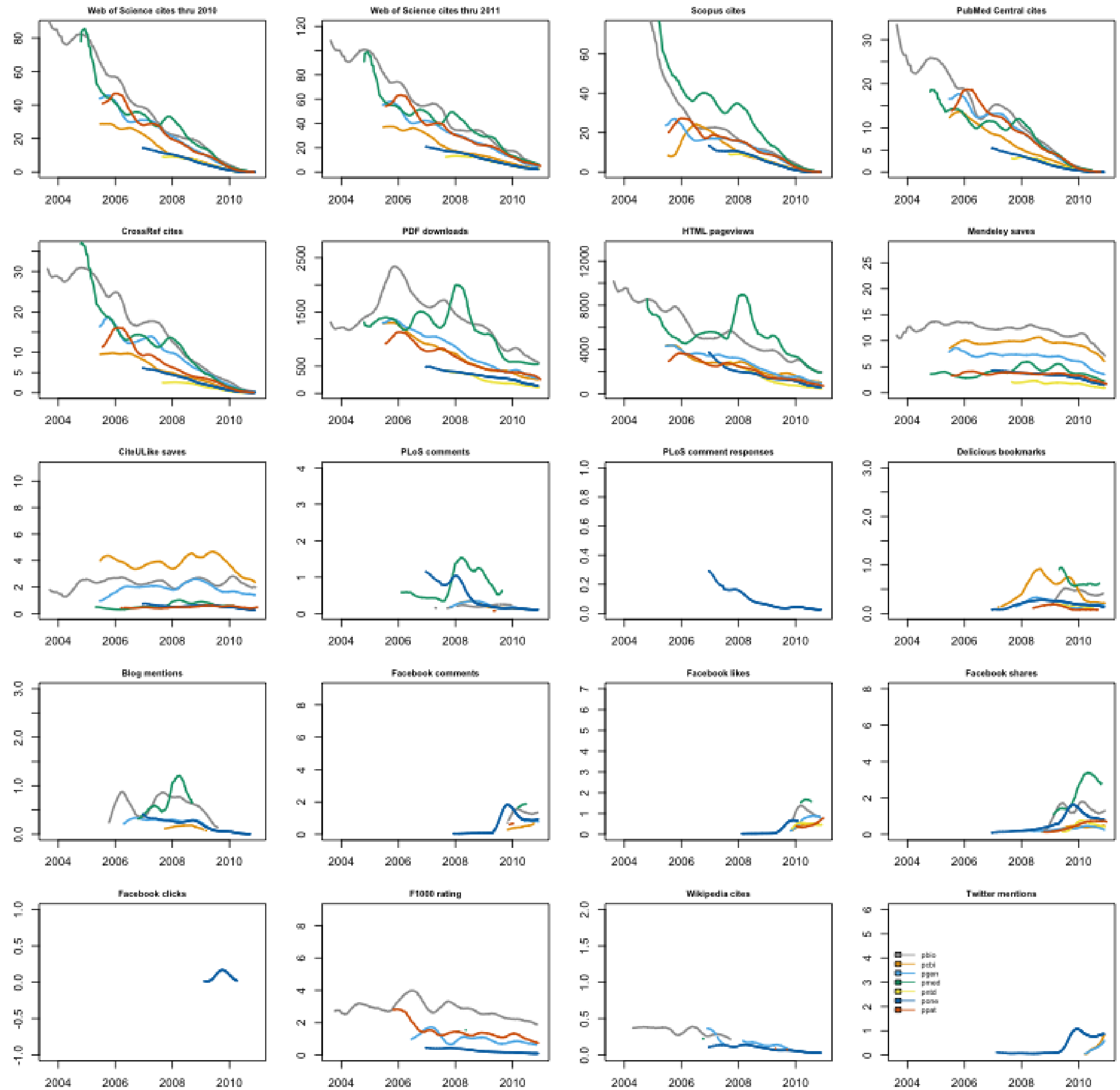


Figure 6: *Total number of events each week aggregated across all articles vs time after article publication, by metric type, for metric types with timestamped events (log scale).*

We next examined the distribution of events by article publication date and event type, as shown in Figure 7. Here three main distributions are visible. Citations, pageviews, and Wikipedia citations all show linear relationships between article age and number of events; the older an article is, the more events it has, suggesting that events are accumulating steadily (if more slowly) over time. Social reference managers CiteULike and Mendeley, as well as Delicious bookmarks and F1000 ratings, fall into a second distribution, in which the number of events is relatively unaffected by articles' ages. This is likely due to a much smaller half-life in these environments (as suggested in Figure 6); after just a year, articles have accumulated about as much activity in these environments as they ever will. Finally, a third distribution is displayed by Twitter, Facebook, PLoS comments, and blogs. In all these cases data are somewhat confused due to changes in adoption of the services and factors relating to data collection and quality (for example, PLoS discontinued posting editorial comments in PLoS ONE, blogging networks come and go, and our Twitter source only goes back one year). This demonstrates the challenge of gathering data from sources that can appear, become popular, evolve, and decline all in a few years, and underscores the importance of conditioning and normalizing raw data.



Differences in behavior across journal can be observed in Figure 7. For example, articles published in PLoS Medicine (green) were cited often, relative to articles in PLoS Computational Biology (light orange). In contrast, articles in PLoS Computational Biology were saved in Mendeley or CiteULike more often than articles in PLoS Medicine.

Figure 7: *Average metric levels over time, by journal. Averages are displayed for metric sources with events across at least 25 articles in the surrounding year in the given journal.*

Finally, Figure 8 looks more closely at a specific article, illustrating how events reveal patterns over time and in relationship to one another. Unsurprisingly, Delicious, CiteULike, and PLoS comments are active immediately after the article's publication. Interestingly, though, in this case a second wave of activity arrived around four months later, possibly initiated by readers alerted to the article by its first external citation (which may reported slightly late here due to issues of granularity in journal publication dates). Although beyond the scope of this paper, this sort of data could be aggregated from thousands of individual articles to identify and classify patterns of dissemination. These in turn could inform powerful predictive models.

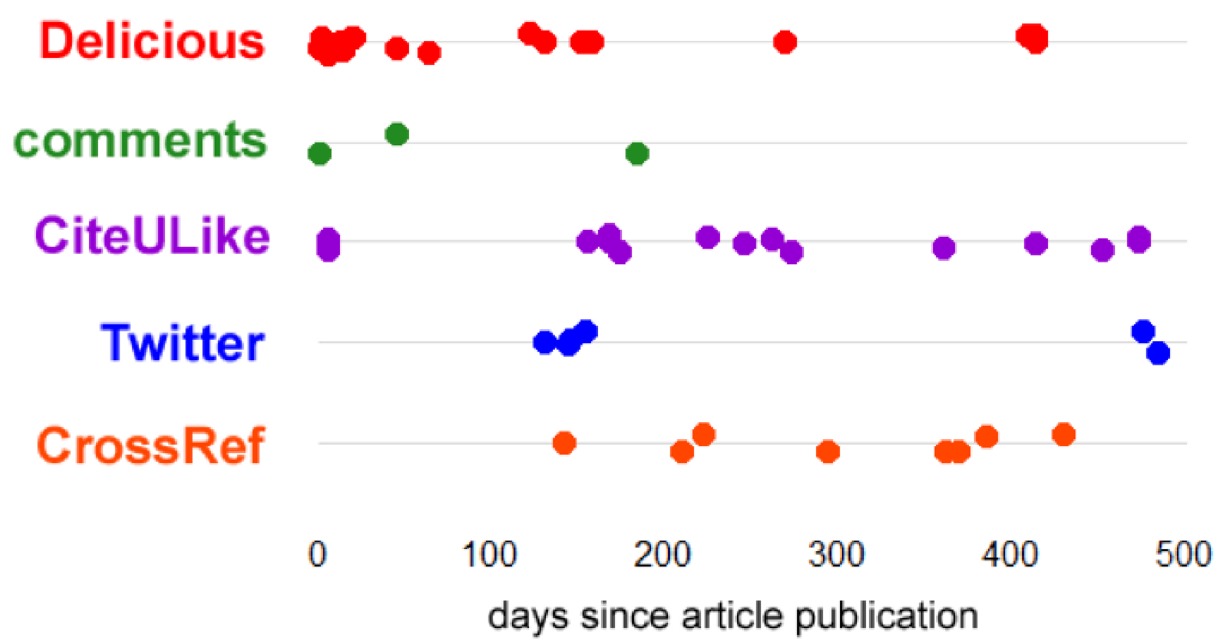


Figure 8: *An illustrative example of altmetric activity timelines for the article doi:10.1371/journal.pone.0006002. Each point represents a single event.*

Results: how do altmetrics relate to one another and to citations?

In exploring the potential of altmetrics it is of great importance to see how metrics relate to one another. In particular, it is useful to establish the relationship between altmetrics and the relatively well-researched metric of traditional citation counts. To do this, we first examined correlations between normalized event counts of different types; then, we performed factor analysis before turning to examine the altmetrics-citation relationship more closely.

Correlation

Figure 9 presents Pearson correlations between all gathered metrics. This graphic is quite similar to the visualization of PLoS Article-Level Metrics correlations in [61]; however it differs in that it contains data from sources not reported by PLoS, and that correlations are on log-transformed counts that have been normalized by metric type, date, and journal.

A few trends are immediately apparent. Citation measures cluster closely together, joined by social reference managers Mendeley and CiteULike; this is unsurprising given the overwhelmingly scholarly audience of these tools and their integration in the citing process. PDF and HTML downloads correlate at moderate to high levels with almost every other indicator; this too is to be expected, given that most types of use both require that someone has looked at the article, and tend to encourage others to do so. F1000 is notable in its relative lack of correlation with anything but citations and social reference managers. Finally, it is interesting that most indicators correlate better with 2011 Web of Science citation counts than with 2010 counts. This may be because later counts are higher, overcoming noise from low early counts.

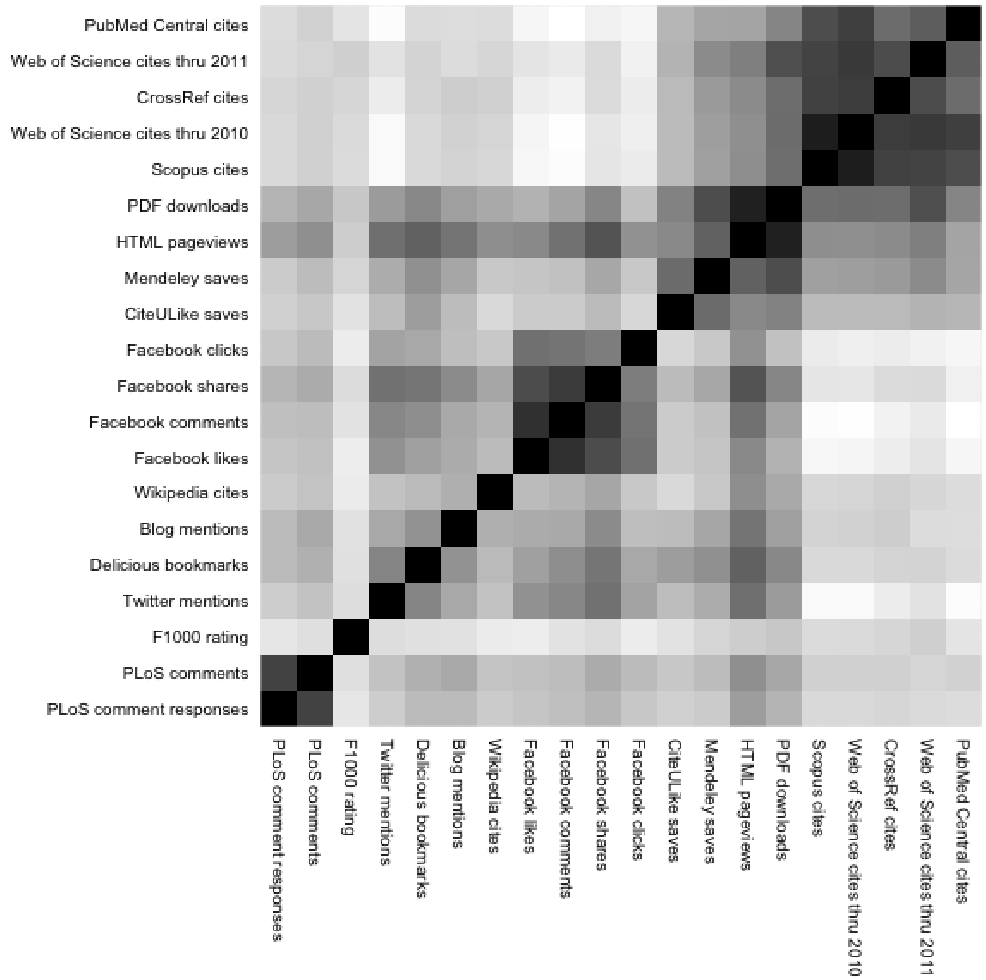


Figure 9:*Pearson correlations of all journals, all years, normalized and transformed.*

Factors' dimensions of impact

To examine the intercorrelation among indicators, we turned to factor analysis. Exploratory factor analysis is a statistical approach to distill variability shared across observed variables in terms of unobserved variables, called factors. Factors often facilitate an understanding of the underlying structure of data and can be used for dimension reduction.

We chose to explore a solution with six factors based on scree plots and the size of our dataset. The rotated first-order factors are given in Table 3 with loadings larger than 0.3 or less than −0.3. Factors in Table 3 are presented in the order of factor extraction: the factor that describes the most shared variation is the column furthest to the left. Some of the loadings are greater than one. This is not unexpected since the factors are oblique and thus the loadings in the pattern matrix represent regression coefficients rather than correlations. Correlations between attributes and the factors are given in the structure matrix in Table S2. The six factors accounted for 53% of the total common variance.

After examining the relative loadings of variables on the factors *post-hoc*, we interpreted and named the six impact factors as impact signal reflected through: 1) citations, 2) page views and shares, 3) Facebook-hosted discussion, 4) PLoS-hosted comments, 5) social reference manager saves, and 6) PDF downloads.

The factors must be interpreted in the context of the other factors, in the order of factor extraction. Common citation variability was identified first and is represented by the Citation Factor. From the remaining variability, a commonality was detected between HTML page views, Twitter mentions, Delicious bookmarks, etc: this was extracted as the second factor which we have dubbed Page Views and Shares. The remaining factors are interpreted similarly. For example, the sixth factor – PDF downloads – needs to be interpreted as the variability shared between PDF downloads, HTML page views, and 2011 Web of Science cites *that was not captured* by any of the earlier factors. Interestingly, the F1000 indicator did not have shared variability with any of the derived factors.

	citations	**pageviews and shares**	**facebook-hosted discussion**	**plos-hosted comments**	**social ref saves**	**pdf downloads**
Web of Science cites thru 2010	1.02					
Web of Science cites thru 2011	0.57					
Scopus cites	0.93					
PubMed Central cites	0.68					
CrossRef cites	0.67					
PDF downloads		0.50				0.77
HTML pageviews		0.93				0.42
Mendeley saves					0.68	
CiteULike saves					0.69	
PLoS comments				0.92		
PLoS comment responses				0.69		
Delicious bookmarks		0.51				
Blog mentions		0.51				
Facebook comments			0.86			
Facebook likes			0.93			
Facebook shares		0.40	0.50			
Facebook clicks			0.39			
F1000 rating						
Wikipedia cites		0.40				
Twitter mentions		0.50				

Table 3: Loading of variables onto exploratory rotated factors. Factors are given in the order of extraction and should be interpreted in this context.

Because the underlying dimensions of impact are likely intercorrelated, we chose a factoring algorithm that allowed the derived factors to be correlated with one another. Correlation patterns can be seen in Figure 10. We found the Citation factor was correlated with both the PDF Download factor and the factor representing Social Reference Manager Saves. Indeed, Social Reference Manager Saves were relatively well correlated with all of the derived factors. The Page Views and Shares signal was particularly correlated with Facebook-hosted Discussions and PLoS-hosted Comments.

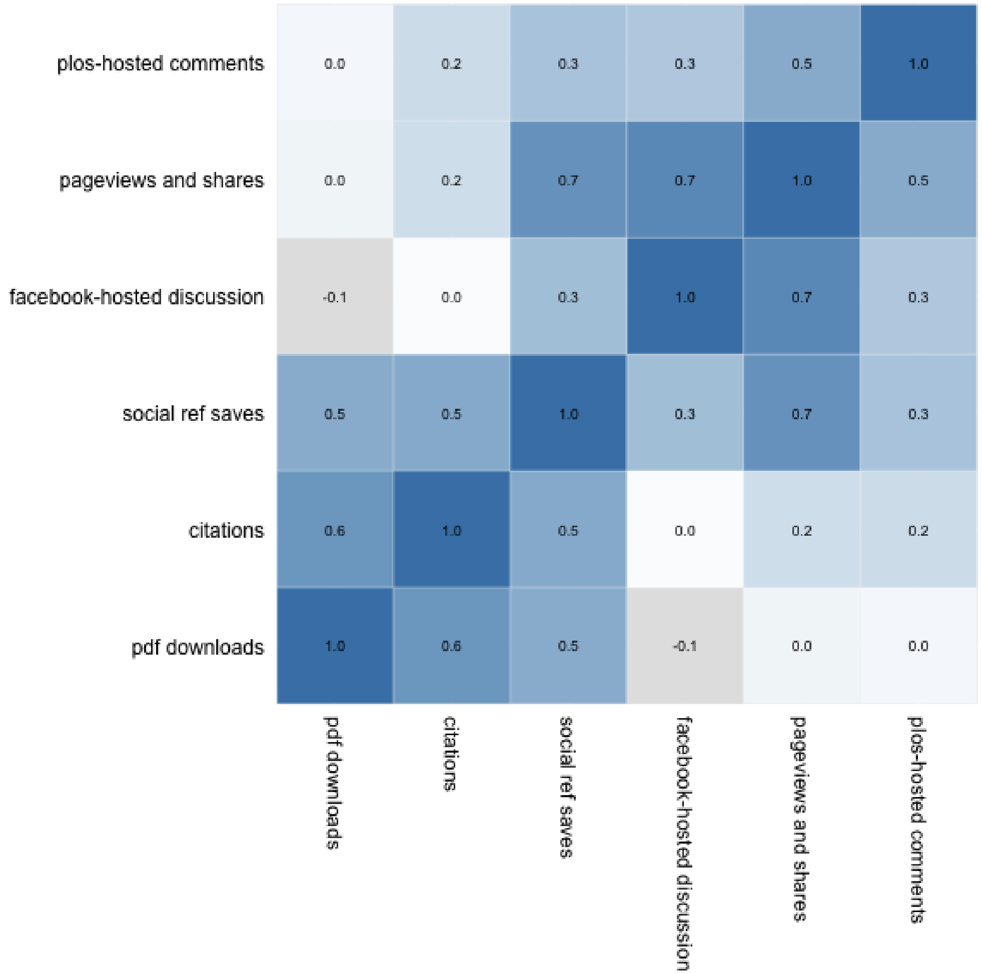


Figure 10: *Correlation between derived factors, based upon the loading matrix.*

Future work incorporating additional altmetrics streams would facilitate a more nuanced understanding of factor structure. Second-order factors might also be interesting: it appears there may be a higher-order dimension around citing and PDF access, another dimension around social engagement, and a third bridging signal related to social reference management.

Relation to citation

Although the validity of citation counts as a measure of research impact has not gone unchallenged [82,83], citation has become a gold standard for impact in both theory and practice. Consequently, it is important to examine altmetrics against this existing standard. However, high correlation with citations is not essential or even perhaps desirable. If altmetrics do in fact track important forms of impact not reflected in the citation record, we would expect low to moderate correlations.

Figures 11 and 12 present correlations between all gathered indicators and Web of Science citation counts, each using a different method to address noise from differing publication dates. Figure 11 includes correlations of all research papers, with indicator values that have been normalized by date and journal and transformed, as described above. Figure 12 reports correlations on a restricted sample: only articles published in 2010. In this case the indicator values were left unnormalized and untransformed, so Spearman correlations were used because of high skew in the untransformed data. Both 11 and 12 focus show correlations for three journals: PLoS ONE (very high-volume, interdisciplinary), PLoS Pathogens (medium volume, specialist), and PLoS Biology (low volume, high-impact).

The patterns in the two figures are quite similar, with the largest difference being much higher correlations between Web of Science citations and citation counts from other indicators. This may be because Figure 11 includes many older articles; these have had enough time to accumulate a large proportion of their lifetime citation counts, informing a more reliable signal. Since the two figures are similar, this discussion will focus on results reported in Figure 11 except when noted otherwise.

All three journals sampled display moderately strong relationships between citation count and pdf/html download count (r=0.34-0.65). Of the altmetrics indicators, Mendeley and CiteULike displayed the highest correlations to Web of Science counts for all three journals(r=0.15-0.43). Interestingly, for all three journals Mendeley bookmark counts correlate more closely to Web of Science citations counts than expert ratings of F1000. What is more, in the age-restricted sample (Figure 12), the correlations between Mendeley and Web of Science citations rivaled or surpassed those of Scopus, PubMed, and CrossRef citations for all three journals.

Correlation patterns across the three journals differ in many ways as well. PLoS ONE correlations are weaker almost across the board, with negligible or zero values for altmetrics other than Mendeley. The strongest correlations are from PLoS Biology, with PLoS Pathogens tending to fall in between. This difference may be related to the number of articles a journal publishes per year. PLoS Biology is likely read in its entirety by many knowledgeable scholars, who in turn probably decide to share, cite, or discuss only the best articles – there may be a sort of ongoing curation from the same consistent group of experts likely to cite. PLoS ONE, on the other hand, likely publishes too many articles, across too many different fields, for any expert curator to personally rank. Consequently, altmetrics data is fragmented, inconsistent, and likely reflects impact on populations less likely to cite. PLoS Pathogens lies somewhere in between, both in size and correlations between citations and altmetrics. This hypothesis is but one possible explanation and deserves further investigation; the important finding here is that the correlation between altmetrics and citations appears quite sensitive to the publishing journal.

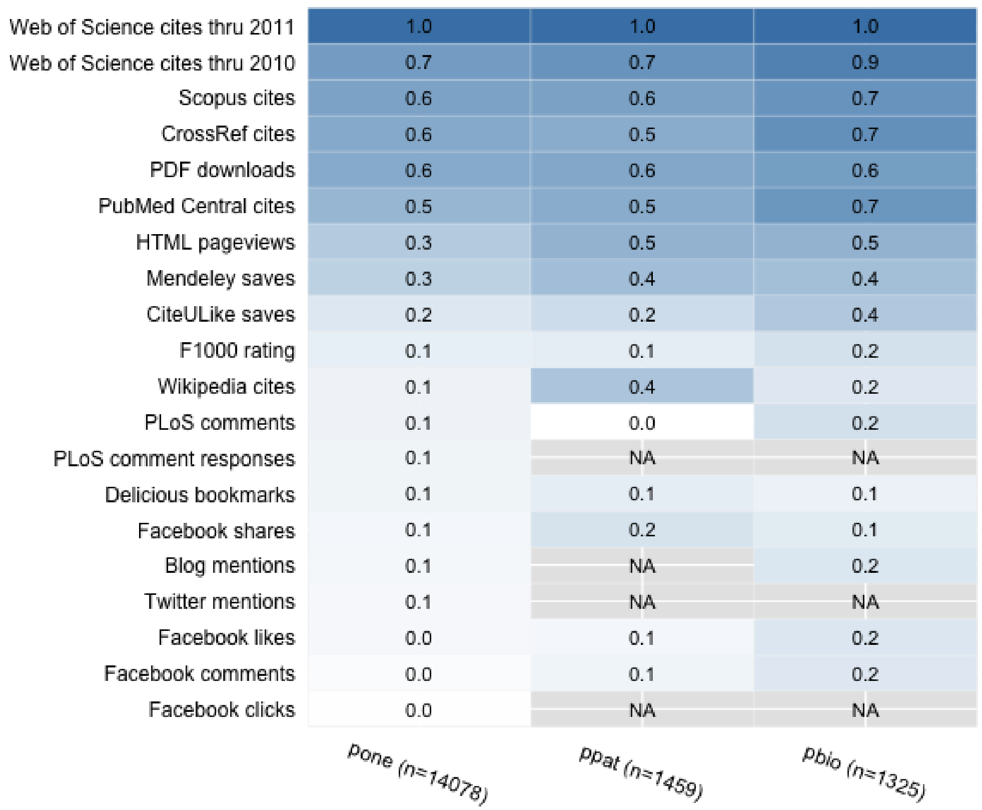


Figure 11: *Pearson correlations coefficients for normalized and transformed 2011 citation counts with other normalized, transformed indicators on articles published all years, by journal.*

	pone 2010 (n=5596)	ppat 2010 (n=459)	pbio 2010 (n=177)
Web of Science cites thru 2011	1.0	1.0	1.0
PDF downloads	0.5	0.6	0.6
HTML pageviews	0.4	0.6	0.5
Web of Science cites thru 2010	0.4	0.6	0.5
Scopus cites	0.4	0.5	0.5
CrossRef cites	0.4	0.4	0.5
PubMed Central cites	0.3	0.4	0.4
Mendeley saves	0.3	0.5	0.4
CiteULike saves	0.1	0.2	0.2
F1000 rating	0.1	0.2	0.3
Wikipedia cites	0.1	0.1	0.2
Facebook shares	0.1	0.2	0.3
Delicious bookmarks	0.1	0.1	0.1
Blog mentions	0.1	0.1	0.2
PLoS comment responses	0.1	0.0	0.1
PLoS comments	0.0	0.0	0.2
Twitter mentions	0.0	0.0	-0.1
Facebook comments	0.0	-0.0	0.2
Facebook likes	0.0	0.0	0.2
Facebook clicks	0.0	0.1	NA

Figure 12: *Spearman correlations coefficients for unnormalized 2011 citation counts with other metrics on articles published in 2010, by journal.*

Can we cluster articles of different impact types using altmetrics?

Factor analysis helps us understand which metrics are similar and how many different dimensions of impact signal are represented in our data. We would also like to know how these signals are reflected in different types of papers. For that, we clustered the papers based on the patterns of attention they've received across various metrics.

To ground these results most closely to our observed data we clustered based on selected variables rather than factor scores. Based on correlations, factor analysis results, and popularity we chose five metrics to be representative of impact metric dimensions: 2011 Web of Science citation counts, Mendeley saves, HTML page views, F1000 score, and a composite we call shareCombo. We constructed the shareCombo variable to keep collinearity and number of clustering dimensions low; it consists of the aggregation of events from four sources: Facebook shares, Delicious bookmarks, Blog mentions, and mentions on Twitter.

We used normalized metric values from papers published in PLoS ONE before 2010 to derive the clusters. Based on scree plots, stability metrics, and interpretability we directed the clustering algorithm to identify five cluster centers.

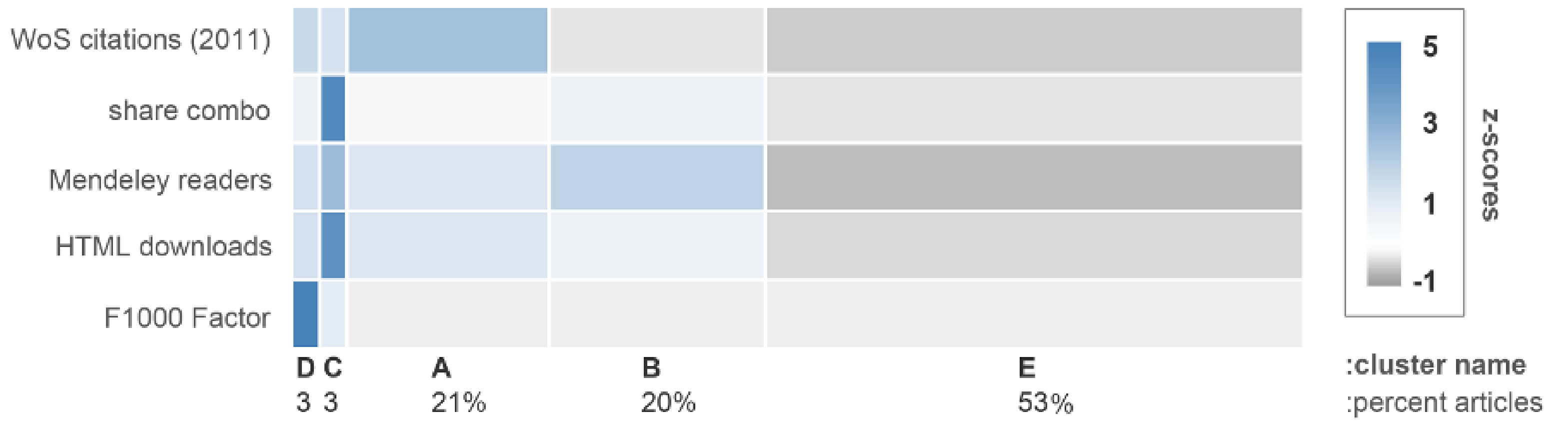


Figure 13: *Cluster centers. Columns represent the centers of clusters; rows show indicators. A dark blue cell in a given cluster indicates that the cluster centers on a relatively high standardized value (z-score) for the given indicator; grey cells represent low values.*

The derived clusters were quite stable across random starting locations: the centers of the clusters are shown in Figure 13. The center of Cluster A is a paper with a high number of citations from Web of Science and a low F1000 score, whereas the center of Cluster B is a paper with a relatively few citations that has received quite a few saves into Mendeley. Clusters A and B varied somewhat across runs; clusters C, D, and E were quite stable. The clusters were relatively distinct and non-overlapping in some projections, as demonstrated by the existence of simple rules to accurately predict cluster membership, as described at the end of this section.

The clusters of impact patterns could be considered the “impact flavor” of the research article [84]. By analyzing patterns in what people are reading, bookmarking, sharing, discussing, AND citing online we might identify what kind—what flavor—of impact a research output is making in a way that citations alone cannot rival. The goal is not to compare flavors: one flavor is not objectively better than another. However, recognizing different types of contributions might help us appreciate scholarly products for the particular needs they meet.

We might consider papers in cluster D, for example, to be “expert picks” – papers with high F1000 evaluations and relatively high citation counts. Cluster C could be dubbed “popular hits” given the high degree of social media attention they received. Papers in cluster A were viewed and cited often, whereas those in cluster B are heavily included in reference managers but rarely cited. Cluster E included papers that didn’t receive much attention in any of the variables in this analysis, though it is worth remembering they may have made impact in ways not captured in this analysis.

We do not claim this analysis reports canonical impact flavors—more research is needed to identify stable, reproducible, validated clusters of research impact. However, this does illustrate what it might look like to begin describing research impact of papers and other scholarly product types with a full flavor palette.

Cluster exemplars

Exemplar membership in each cluster is illustrated in Table 4. Exemplars were chosen as papers with metric values closest to the center of each cluster. There are few obvious emergent patterns, save that cluster C titles, unsurprisingly, suggest articles of particular appeal to a “general interest” audience. Future work might more systematically examine articles for internal features (like topic, keywords, reading level, article length, and so on) that predict cluster membership.

Cluster	**% articles in cluster**	**Examples**

A: Read and cited	20%	Phylogeny in Aid of the Present and Novel Microbial Lineages: Diversity in Bacillus Contact Networks in a Wildlife-Livestock Host Community: Identifying High-Risk Individuals in the Transmission of Bovine TB among Badgers and Cattle Refuge or Reservoir? The Potential Impacts of the Biofuel Crop Miscanthus x giganteus on a Major Pest of Maize
B: Read, saved, and shared	21%	Vision and Foraging in Cormorants: More like Herons than Hawks? Rich Pickings Near Large Communal Roosts Favor 'Gang' Foraging by Juvenile Common Ravens, Corvus corax Wind, Waves, and Wing Loading: Morphological Specialization May Limit Range Expansion of Endangered Albatrosses
C: Popular hit	3%	Cognitive Processes Associated with Sequential Tool Use in New Caledonian Crows Claims of Potential Expansion throughout the U.S. by Invasive Python Species Are Contradicted by Ecological Niche Models Why Men Matter: Mating Patterns Drive Evolution of Human Lifespan
D: Expert pick	3%	An Inhibitory Sex Pheromone Tastes Bitter for Drosophila Males Substantial Alterations of the Cutaneous Bacterial Biota in Psoriatic Lesions Community Analysis of Chronic Wound Bacteria Using 16S rRNA Gene-Based Pyrosequencing: Impact of Diabetes and Antibiotics on Chronic Wound Microbiota
E: Not much attention using these metrics	53%	Effects of Endolithic Parasitism on Invasive and Indigenous Mussels in a Variable Physical Environment Sexual Conflict and Sexually Antagonistic Coevolution in an Annual Plant Antibiotic Treatment of the Tick Vector Amblyomma americanum Reduced Reproductive Fitness

Table 4: Exemplar members of each cluster, from the field of Ecology

Identifying clusters with rules

Clusters would be most useful if they could be simply defined and operationalized through a simple set of rules. This would allow researchers or evaluators to characterize the impact flavor of an article simply by examining its pattern of activity. To illustrate this concept and explore the distinctness of the clusters, we attempted to derive simple rules for accurately predicting the k-means cluster membership.

The K-means centers described above were derived on papers published before 2010. We now used these centers to assign cluster membership to the 4,294 PLoS ONE papers in our dataset published during 2010; it was for these papers that we derived associative rules to predict the k-means assigned cluster membership. Five simple rules, listed below, were sufficient to assign the identical cluster membership to Euclidean distance algorithms with 11% misclassification error (kappa 0.81).

cluster D if ((F100- rating >= 6) and (html pageviews <= 4332))

ELSE cluster C if ((html pageviews >= 2031) and (tweets >= 4))

ELSE cluster A if ((Web of Science Citations >= 7) OR ((Web of Science Citations >= 5) AND (Mendeley saves >= 1)))

ELSE cluster B if ((Mendeley saves >= 3) OR ((Mendeley saves >= 2) AND (tweets >= 1)))

ELSE cluster E

The confusion matrix from the rule classification evaluation, given in Table 5, provides some insight into which of the clusters were easily mistaken for one another. Papers which had been classified as Cluster B (“read, saved, and shared”) were classified relatively often by the simple rules as members of Cluster A (“read and cited”) or Cluster E (“not much attention”), whereas classification into cluster D (“expert pick”) was very consistent across k-means and the simple rules.

	cluster C	**cluster D**	**cluster E**	**cluster B**	**cluster A**
cluster C	78	0	3	18	14
cluster D	0	85	0	0	0
cluster E	0	0	2088	79	86
cluster B	3	0	112	817	38
cluster A	3	0	74	39	757

Table 5: Confusion Matrix (cross classifications) of k-means. Clusters reported in rows vs. rule-based classifications reported in columns.

Conclusion

This descriptive study presents a great many findings, but three are perhaps especially salient: first, there is no shortage of data from altmetrics sources, although different indicators vary greatly in activity. Around 80% of sampled articles have been included in at least one Mendeley library, and a quarter of articles have nonzero data from five or more different sources. Second, altmetrics and citations track forms of impact that are distinct, but related; neither approach is able to describe the complete picture of scholarly use alone. There are moderate correlations between Mendeley and Web of Science citation (comparable to that between Web of Science and Scopus), but many altmetric indicators seem mostly orthogonal to citation. Third, articles cluster in ways that suggest several different impact “flavors,” that capture impacts on different audiences and of different types; for instance some articles (cluster B) may be heavily read and saved by scholars but seldom cited.

Limitations

These results are subject to several important limitations. First, data quality is a challenge, since different services come and go over time. This is a major source of noise in the impact signal, and while our normalization approach is helpful in reducing this noise, it is not perfect.

Second, our sample should be carefully considered before attempting to generalize results. In particular, PLoS publishes only open-access journals, exposing each article to a much broader audience and a much wider range of potential interactions than a closed-access publisher would. The corpus of articles is also dominated, particularly in later years, by the huge volume published by PLoS ONE. This unique publication has a volume and scope dwarfing that of traditional journals, inescapably giving rise to similarly unique patterns of readership. There are also limitations in our sampling of indicators. Although this sample is diverse, it is not complete; Google Scholar citations, for instance, are conspicuously absent, and likely would have correlated more closely with many altmetric indicators than citations from Web of Science [43,85]. Also, given the growing scholarly use social media, it is likely that many counts are low compared to what we would see if replicating the study today.

Third, counts for altmetric indicators may be relatively easy to game. This is more a limitation for future applications, since few of these counts are today important enough to fabricate. However, it is a limitation that will need to be addressed before altmetrics can be used for serious evaluation. Indeed, experience suggests that the more attention is paid to these metrics, the greater problem gaming will become [86]; this has certainly been the case with citations, which are routinely gamed by academics to varying degrees of success [13,87]. Experience at Google and Wikipedia, however, suggest that careful data mining is an effective technique for stopping gaming, while community-driven policing at social sites like Digg and Reddit have also been effective [37]. The Social Science Research Network (SSRN), a scholarly preprint repository, has also reported success with these approaches [88].

Finally, it is important to emphasize that that while new metrics are useful as alternatives to citation-based measure for some purposes, they are by no means a replacement. Instead, they should be deployed alongside one another, complementing each other’s strengths.

Future research and applications

Much work to expand this research will center around reducing noise that obscures the impact signal – or, more accurately, isolating and identifying different types of impacts on different audiences. Most obviously, this will involve investigating altmetrics in other contexts, with more sources, and (especially) sampling from different journals and publishers. This also must involve observational, interview, content-analytic, and ethnographic studies. These are crucial to understand what the events informing alternative metrics actually *mean*, following the examples set by early investigators of citation [83]. These investigations of context should also expand, refine, and validate clusters or impact flavors, perhaps by asking readers or expert judges to identify article features influencing different patterns of use. Metrics must move beyond simply reporting counts, and take network properties into effect, as Google does with PageRank; a tweet from a highly-connected, expert scholar should mean something different from one authored by a casual observer. Research should continue to examine whether early altmetrics counts can predict later citation.

Investigation should also expand to examine altmetrics for scholarly products other than articles, such as datasets, software, and blog posts. Since creators of these products may find recognition via traditional citation difficult, altmetrics could be crucial in evaluating and promoting Web-aware scholarship. However, it must be emphasized that much additional research is needed before altmetrics can be relied upon in high-stakes evaluation like tenure and promotion decisions.

Tools to gather altmetrics data will be essential to fuel this additional research, as well as to support low-stakes experimentation with altmetrics as source of formative evaluation data. It is essential that these tools supply open data and make their collection procedures completely transparent, avoiding the much-lamented opacity, closedness, and consequent irreproducibility of Thomson's Impact Factor [89–91]. Several promising, open-source altmetrics tools have begun to appear, including CitedIn, ReaderMeter, ScienceCard, and total-impact.

In the future, tools like these may allow researchers to keep a "live CV" showing up-to-date indicators of their works' impact next to each product they have produced. Funding organizations could keep timely portfolios of their grants impact across different populations and communities. Tenure, promotion, and hiring committees could assemble research teams whose members have different sets of "impact specialties," just as sport managers might assemble teams of players with different physical and mental skill sets. Altmetrics also have great potential for supporting personalized recommender systems; citation-based versions of these have been created before [92], but suffer from the long lag between an idea's insemination and citation; altmetrics could solve this problem. Sufficiently advanced recommendation tools allow the technical act of publishing to be decoupled from relevance and quality filtering [93], allowing every researchers to essentially consume their own private journal, peer reviewed organically by their scholarly network.

While many of the data sources described in this study will no doubt disappear over time, scholars' presence on the social Web will not. Invisible colleges are not a fad. The social Web is merely the latest in a series of technologies improving the reach and density of these communities, which have always lay at the heart of scholarship. In 2005, Cronin argued: "It is clear that we will soon have access to a critical mass of web-based digital objects and usage statistics with which to develop multi-dimensional models of scholars' communication behaviors" [94]. Seven years later, our data suggest that he was correct.

Acknowledgements

Thanks to Todd Vision for comments, and Kaitlin Costello for formatting and proofreading. Thanks also to the Public Library of Science for creating and maintaining Article-Level Metrics as a free and open data source, as well as all other providers of free altmetrics data.

References

1. Merton RK (1988) The Matthew Effect in Science, II. ISIS 79: 606–623.
2. Kuhn T (1977) Second thoughts on paradigms. The essential tension: Selected studies in scientific tradition and change. Chicago: University of Chicago Press.
3. Bernal JD (1944) The social function of science. George Routledge. 512 p.

4. jasonpriem (2010) I like the term #articlelevelmetrics, but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics. Available: https://twitter.com/#!/jasonpriem/status/25844968813.

5. Cronin B (2001) Bibliometrics and beyond: Some thoughts on web-based citation analysis. Journal of Information Science 27: 1. doi:10.1177/016555150102700101.

6. Brody T, Harnad S, Carr L (2006) Earlier Web usage statistics as predictors of later citation impact. Journal of the American Society for Information Science and Technology 57: 1060–1072. doi:10.1002/asi.20373.

7. Piwowar HA (2010) A method to track dataset reuse in biomedicine: Filtered GEO accession numbers in PubMed central. Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem - Volume 47. Silver Springs, MD, USA: American Society for Information Science. pp. 179:1–179:2. Available: http://dl.acm.org/citation.cfm?id=1920331.1920551. Accessed 21 February 2012.

8. Chao TC (2011) Disciplinary reach: Investigating the impact of dataset reuse in the earth sciences. Proceedings of the American Society for Information Science and Technology 48: 1–8. doi:10.1002/meet.2011.14504801125.

9. Palmer CL, Weber NM, Cragin MH (2011) The analytic potential of scientific data: Understanding re−use value. Proceedings of the American Society for Information Science and Technology 48: 1–10. doi:10.1002/meet.2011.14504801174.

10. King D, Tenopir C (2004) An evidence-based assessment of the "author pays" model. Nature web focus: Access to the literature. Available: http://www.nature.com.libproxy.lib.unc.edu/nature/focus/accessdebate/26.html. Accessed 31 January 2012.

11. Seglen PO (1994) Causal relationship between article citedness and journal impact. Journal of the American Society for Information Science 45: 1–11. doi:10.1002/(SICI)1097-4571(199401)45:1<1::AID-ASI1>3.0.CO;2-Y.

12. Rossner M, Van Epps H, Hill E (2007) Show me the data. Journal of Experimental Medicine 204: 3052–3053. doi:10.1084/jem.20072544.

13. Falagas M, Alexiou V (2008) The top-ten in journal impact factor manipulation. Archivum Immunologiae et Therapiae Experimentalis 56: 223–226. doi:10.1007/s00005-008-0024-5.

14. Pavitt K (1985) Patent statistics as indicators of innovative activities: Possibilities and problems. Scientometrics 7: 77–99. doi:10.1007/BF02020142.

15. Cronin B, Overfelt K (1994) The scholar's courtesy: A survey of acknowledgement behaviour. Journal of Documentation 50: 165–196. doi:10.1108/eb026929.

16. Sugimoto CR, Russell TG, Meho LI, Marchionini G (2008) MPACT and citation impact: Two sides of the same scholarly coin? Library & Information Science Research 30: 273–281. doi:10.1016/j.lisr.2008.04.005.

17. Almind TC, Ingwersen P (1997) Informetric Analyses on the World Wide Web: Methodological Approaches to "WEBOMETRICS." Journal of Documentation 53: 404–426.

18. Ingwersen P (1998) The calculation of Web impact factors. Journal of documentation 54: 236–243.

19. Thelwall M (2003) What is this link doing here? Beginning a fine-grained process of identifying reasons for academic hyperlink creation. Information research 8: 8–3.

20. Vaughan L, Shaw D (2008) A new look at evidence of scholarly citation in citation indexes and from web sources. Scientometrics 74: 317–330.

21. Kousha K, Thelwall M (2008) Assessing the impact of disciplinary research on teaching: An automatic analysis of online syllabuses. Journal of the American Society for Information Science and Technology 59: 2060–2069. doi:10.1002/asi.20920.

22. Thelwall M, Kousha K (2008) Online presentations as a source of scientific impact? An analysis of PowerPoint files citing academic journals. Journal of the American Society for Information Science and Technology 59: 805–815.

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

24. Haustein S, Siebenlist T (2011) Applying social bookmarking data to evaluate journal usage. Journal of Informetrics 5: 446–457. doi:10.1016/j.joi.2011.04.002.

25. Procter R, Williams R, Stewart J, Poschen M, Snee H, et al.

26. Adoption and use of Web 2.0 in scholarly communications. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 368: 4039 –4056. doi:10.1098/rsta.2010.0155.

27. Tinti-Kane H, Seaman J, Levy J (2010) Social media in higher education: The survey.

28. Rowlands I, Nicholas D, Russell B, Canty N, Watkinson A (2011) Social media use in the research workflow. Learned Publishing 24: 183–195. doi:10.1087/20110306.

29. Hull D, Pettifer SR, Kell DB (2008) Defrosting the digital library: Bibliographic tools for the next generation Web. PLoS Computational Biology 4. doi:10.1371/journal.pcbi.1000204

30. Ebner M, Schiefner M (2008) Microblogging-more than fun. In: Arnedillo Sánchez I, Isaías P, editors. Proceedings of IADIS mobile learning conference. Algarve, Portugal, Vol. 155. p. 159.

31. Efimova L, Hendrick S (2004) In search for a virtual settlement: An exploration of weblog community boundaries. Telematica Instituut. Milan, Italy. Available: https://doc.telin.nl/dsweb/Get/Document-46041/.

32. Lund B, Hammond T, Flack M, Hannay T, NeoReality I (2005) Social bookmarking tools (II). D-Lib Magazine 11: 1082–9873. Available: http://www.dlib.org/dlib/april05/lund/04lund.html.

33. de Solla Price DJ, Beaver D (1966) Collaboration in an invisible college. American Psychologist 21: 1011.

34. Cronin B, Snyder HW, Rosenbaum H, Martinson A, Callahan E

35. Invoked on the Web. Journal of the American Society for Information Science 49: 1319–1328. doi:10.1002/(SICI)1097-4571(1998)49:14<1319::AID-ASI9>3.0.CO;2-W.

36. Jensen M (2007) The New Metrics of Scholarly Authority. The Chronicle of Higher Education. Available: http://www.lexisnexis.com/us/lnacademic/frame.do?tokenKey=rsh-20.918629.3985793049&target=results_listview_resultsNav&reloadEntirePage=true&rand=1254107715289&returnToKey=20_T7441126159&parent=docview. Accessed 28 September 2009.

37. Taraborelli D (2008) Soft peer review: Social software and distributed scientific evaluation. Proceedings of the 8th International Conference on the Design of Cooperative Systems (COOP ’08). Carry-Le-Rouet.

38. Patterson M (2009) Article-level metrics at PLoS – addition of usage data. Available: http://blogs.plos.org/plos/2009/09/article-level-metrics-at-plos-addition-of-usage-data/. Accessed 1 March 2012.

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

40. Priem J, Taraborelli D, Groth P, Neylon C (2010) alt-metrics: a manifesto. Available: http://altmetrics.org/manifesto/. Accessed 15 August 2011.

41. Schriger DL, Chehrazi AC, Merchant RM, Altman DG (2011) Use of the Internet by Print Medical Journals in 2003 to 2009: A Longitudinal Observational Study. Annals of Emergency Medicine In Press, Corrected Proof. Available: http://www.sciencedirect.com.libproxy.lib.unc.edu/science/article/B6WB0-51S6FN7-1/2/2edccd79138b204deee17536d13a2eae. Accessed 23 December 2010.

42. Priem J, Costello K, Dzuba T (2011) First-year graduate students just wasting time? Prevalence and use of Twitter among scholars Metrics 2011: New Orleans, LA, USA. Available: http://jasonpriem.org/self-archived/5uni-poster.png.

43. Shuai X, Pepe A, Bollen J (2012) How the Scientific Community Reacts to Newly Submitted Preprints: Article Downloads, Twitter Mentions, and Citations. arXiv:12022461. Available: http://arxiv.org/abs/1202.2461. Accessed 18 February 2012.

44. Uren V, Dadzie A-S (2011) Relative trends in scientific terms on Twitter Koblenz, Germany. Available: http://altmetrics.org/workshop2011/uren-v0/. Accessed 15 August 2011.

45. Li X, Thelwall M, Giustini D (2011) Validating online reference managers for scholarly impact measurement. Scientometrics. Available: http://www.springerlink.com.libproxy.lib.unc.edu/content/35146th23t1j1284/. Accessed 14 January 2012.

46. Costello KL, Priem J (2011) Archiving scholars' tweets. Proceedings of the Fourth Annual Society of American Archivists Research Forum. Chicago, Il: Society of American Archivists. Available: http://www2.archivists.org/sites/all/files/KCFinal.pdf.

47. Letierce J, Passant A, Breslin J, Decker S (2010) Understanding how Twitter is used to spread scientific messages. Web Science Conference: Raleigh, NC. Available: http://journal.webscience.org/314/2/websci10_submission_79.pdf.

48. Ross C, Terras M, Warwick C, Welsh A (2011) Enabled backchannel: conference Twitter use by digital humanists. Journal of Documentation 67: 214–237.

49. Priem J, Costello KL (2010) How and why scholars cite on Twitter. Proceedings of the 73rd ASIS&T Annual Meeting. Pittsburgh PA, USA. Available: http://mail.asis.org.libproxy.lib.unc.edu/asist2010/proceedings/proceedings/ASIST_AM10/submissions/201_Final_Submission.pdf.

50. Weller K, Dröge E, Puschmann C (2011) Citation analysis in Twitter: Approaches for defining and measuring information flows within tweets during scientific conferences. Proceedings of Making Sense of Microposts Workshop (# MSM2011). Co-located with Extended Semantic Web Conference, Crete, Greece.

51. Jiang J, He D, Ni C (2011) Social reference: aggregating online usage of scientific literature in CiteULike for clustering academic resources. Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries. JCDL '11. Ottawa, Ontario, Canada: ACM. pp. 401–402. doi:10.1145/1998076.1998155.

52. Bogers T, van den Bosch A (2008) Recommending scientific articles using citeulike. Proceedings of the 2008 ACM conference on Recommender systems. ACM New York, NY, USA. pp. 287–290.

53. Peters I, Haustein S, Terliesner J (2011) Crowdsourcing in Article Evaluation. Proceedings of the ACM WebSci 11 – 3rd International Conference on Web Science. Koblenz, Germany. Available: http://www.websci11.org/fileadmin/websci/Posters/89_paper.pdf.

54. Kipp MEI (2007) Tagging Practices on Research Oriented Social Bookmarking Sites Montreal, Quebec. Available: http://hdl.handle.net/10760/10385. Accessed 28 January 2012.

55. Good B, Tennis J, Wilkinson M (2009) Social tagging in the life sciences: Characterizing a new metadata resource for bioinformatics. BMC Bioinformatics 10: 313. doi:10.1186/1471-2105-10-313.

56. Wardle DA (2010) Do “Faculty of 1000” (F1000) ratings of ecological publications serve as reasonable predictors of their future impact? Ideas Eco Evo 3. Available: http://library.queensu.ca/ojs/index.php/IEE/article/view/2379. Accessed 12 July 2010.

57. Wets K, Weedon D, Velterop J (2003) Post-publication filtering and evaluation: Faculty of 1000. Learned Publishing 16: 249–258. doi:10.1087/095315103322421982.

58. Nielsen F (2007) Scientific citations in Wikipedia. First Monday 12(8). Available: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1997/1872.

59. Groth P, Gurney T (2010) Studying scientific discourse on the Web using bibliometrics: A chemistry blogging case study. Web Science Conference, Raleigh, NC: US. Available: http://journal.webscience.org/308/. Accessed 13 September 2010.

60. Shema H, Bar-Ilan J (2011) Characteristics of Researchblogging.org science Blogs and Bloggers. Altmetrics '11, Koblenz, Germany. Available: http://altmetrics.org/workshop2011/shema-v0/. Accessed 15 August 2011.

61. Mixon FG, Upadhyaya KP (2010) Blogometrics. Eastern Economic Journal 36: 1–10.

62. Henning V, Reichelt J (2008) Mendeley - A Last.fm For Research? eScience, 2008. eScience ’08. IEEE Fourth International Conference on. pp. 327–328. doi:10.1109/eScience.2008.128.

63. Yan K-K, Gerstein M (2011) The Spread of Scientific Information: Insights from the Web Usage Statistics in PLoS Article-Level Metrics. PLoS ONE 6: e19917. doi:10.1371/journal.pone.0019917.

64. Eysenbach G (2012) Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. Journal of Medical Internet Research

65. doi:10.2196/jmir.2012.

66. Hamming RW (1989) Digital filters. Englewood Cliffs, NJ: Prentice-Hall. 306 p.

67. R Development Core Team (2010) R: A Language and Environment for Statistical Computing. Vienna, Austria. p. Available: http://www.R-project.org/.

68. Brock G, Pihur V, Datta S, Datta and S (2011) clValid: Validation of Clustering Results. Available: http://CRAN.R-project.org/package=clValid.

69. Vries A de (2011) ggdendro: Tools for extracting dendrogram and tree diagram plot data for use with ggplot. Available: http://CRAN.R-project.org/package=ggdendro.

70. Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer New York. Available: http://had.co.nz/ggplot2/book.

71. Bernaards CA, Jennrich RI (2005) Gradient Projection Algorithms and Software for Arbitrary Rotation Criteria in Factor Analysis. Educational and Psychological Measurement 65: 676–696.

72. Warnes GR (2010) gplots: Various R programming tools for plotting data. Available: http://CRAN.R-project.org/package=gplots.

73. Raiche G, Magis D (2010) nFactors: Parallel Analysis and Non Graphical Solutions to the Cattell Scree Test. Available: http://CRAN.R-project.org/package=nFactors.

74. Wickham H (2010) plyr: Tools for splitting, applying and combining data. Available: http://CRAN.R-project.org/package=plyr.

75. Revelle W (2010) psych: Procedures for Psychological, Psychometric, and Personality Research. Evanston, Illinois. p. Available: http://personality-

project.org/r/psych.manual.pdf.

76. Wickham H (2007) Reshaping data with the reshape package. Journal of Statistical Software 21. Available: http://www.jstatsoft.org/v21/i12/paper.

77. Harrell FE Jr (2010) rms: Regression Modeling Strategies. Available: http://CRAN.R-project.org/package=rms.

78. Hornik K, Buchta C, Zeileis A (2008) Open-source machine learning: R meets Weka. Computational Statistics 24: 225–232. doi:10.1007/s00180-008-0119-7.

79. Short T (2010) signal: Signal processing. Available: http://CRAN.R-project.org/package=signal.

80. Grothendieck G (2010) sqldf: Perform SQL Selects on R Data Frames. Available: http://CRAN.R-project.org/package=sqldf.

81. Dahl DB (2009) xtable: Export tables to LaTeX or HTML. Available: http://CRAN.R-project.org/package=xtable.

82. Zeileis A, Grothendieck G (2005) zoo: S3 Infrastructure for Regular and Irregular Time Series. Journal of Statistical Software 14: 1–27.

83. Cohen WW (1995) Fast effective rule induction. In: Prieditis A, Russell S, editors. Proceedings of the 12th International Conference on Machine Learning. Morgan Kaufmann. pp. 115–123. Available: http://citeseer.ist.psu.edu/cohen95fast.html.

84. Burton RE, Kebler RW (1960) The “half-life” of some scientific and technical literatures. American documentation 11: 18–22.

85. MacRoberts MH, MacRoberts BR (2010) Problems of citation analysis: A study of uncited and seldom-cited influences. Journal of the American Society for Information Science and Technology 61: 1–12. doi:10.1002/asi.21228.

86. Bornmann L, Daniel HD (2008) What do citation counts measure? A review of studies on citing behavior. Journal of Documentation 64: 45.

87. Piwowar H (2012) 31 Flavors of Research Impact through #altmetrics. Research Remix. Available: http://researchremix.wordpress.com/2012/01/31/31-flavours/. Accessed 2 March 2012.

88. Kousha K, Thelwall M (2007) Google Scholar citations and Google Web/URL citations: a multi-discipline exploratory analysis. Journal of the American Society for Information Science and Technology 58: 1055–1065.

89. Espeland WN, Sauder M (2007) Rankings and Reactivity: How Public Measures Recreate Social Worlds. American Journal of Sociology 113: 1–40.

90. Wilhite AW, Fong EA (2012) Coercive Citation in Academic Publishing. Science 335: 542–543. doi:10.1126/science.1212540.

91. Edelman BG, Larkin I (2009) Demographics, Career Concerns or Social Comparison: Who Games SSRN Download Counts? SSRN eLibrary. Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1346397. Accessed 12 February 2012.

92. Rossner M, Van Epps H, Hill E (2008) Irreproducible results: a response to Thomson Scientific. Journal of Experimental Medicine 205: 260.

93. Brumback RA (2008) Worshiping false idols: The impact factor dilemma. Journal of child neurology 23: 365.

94. The PLoS Medicine Editors (2006) The Impact Factor Game. PLoS Med 3: e291. doi:10.1371/journal.pmed.0030291.

95. McNee SM (2006) Meeting user information needs in recommender systems University of Minnesota. PhD dissertation. Available: http://www-users.cs.umn.edu/~mcnee/mcnee-thesis-preprint.pdf.

96. Priem J, Hemminger BH Decoupling the scholarly journal. Frontiers in Computational Neuroscience. Available: http://www.frontiersin.org/computational_neuroscience/abstract/14455. Accessed 2 March 2012.

97. Cronin B (2005) A hundred million acts of whimsy. Current Science 89: 1505–1509.