# Research Trends

## Special issue on alternative metrics

ELSEVIER

# Research Trends Issue 37: Special issue on alternative metrics

This special issue of Research Trends is dedicated to altmetrics, or, as some may prefer, alternative metrics. The growing interest in the development of alternative measurements of scientific productivity resulted in the 2010 Altmetrics manifesto in which the term "altmetrics" was introduced. The manifesto notes that "in growing numbers, scholars are moving their everyday work to the web", and that new online tools "reflect and transmit scholarly impact". This "forms a composite trace of impact far richer than any available before; we call the elements of this trace altmetrics".

In his historical account, **Mike Thelwall** covers the use of social web services. He dedicates attention to Mendeley and Twitter. He also underlines the need to further validate altmetrics, by investigating the degree to which they correlate with – or predict – citation counts and other traditional measures. **Ron Daniel** addresses the issue of how past citation prediction studies showed so little consistency in their results, and indicates the potential of altmetrics. **Hadas Shema** introduces another promising altmetric data source: scholarly blogs.

The rapid ICT development is a first principal driver of new metrics in general. It is not only reflected in the emergence of social web services, but also, for instance, in the further digitalization of scientific information. Electronic user log files of publication archives provide traces of another aspect of scholars' everyday work, namely their literature browsing and perhaps their reading behavior. In this sense, indicators of full text downloads of scholarly publications can be conceived as altmetrics as well.

A team of 5 researchers headed by **Christian Schlögl** interprets correlations between citation, full text download and readership data in terms of the degree in overlap between the user communities of the three systems from which data was extracted. **Vicente P. Guerrero-Bote** and **Félix Moya-Anegón** also examine statistical correlations between downloads and citations. They focus on the role of publication language, and analyze less visible journals publishing in languages other than English and clearly show how the two types of measures are complementary – one type may reveal patterns that are invisible in the other. As **Euan Adie** outlines in his contribution, another aspect of digitalization is that more and more grey literature including pre-prints and policy documents become available for research and as a source for new metrics. However, in a discussion with **Mike Taylor**, **Juan Pablo Alperin** warns that differences in access to new technology exist between scientifically developing and developed countries that can have negative implications for the former group of countries even in altmetrics research.

A second principal driver of the development of new metrics is the Open Science movement, directed towards making scientific research, data, and dissemination accessible to all levels of an inquiring society, whether amateur or professional. Perhaps the base notions of this movement can be characterized as expressions of the fundamental ethos of science in a digital or computerized age. As **William Gunn** explains, research funding agencies seek to maximize the potential of their funded outputs including papers, methods, and data. This urges us to develop new metrics of reuse that go beyond classical citations.

This brings us to a third driver: the research policy domain. This is also clearly reflected in the contribution by **Judit Bar-Ilan** who introduces the portfolio concept developed in the ACUMEN project (Academic Careers Understood through Measurement and Norms) funded by the European Commission, aimed at "studying and proposing alternative and broader ways of measuring the productivity and performance of individual researchers". The author shows how online and social media presence and altmetrics are well represented in the expertise, output and influence sub-portfolios.

We believe the contributions in this Special Issue cover the major trends in the development of new metrics, and are written by leading researchers in the field. We hope you will enjoy reading them.

Please share your thoughts and feedback with us. You can do this in the comments section following each article on our website or by sending us an email (researchtrends@elsevier.com). We look forward to hearing from you!

Kind regards,

**Mike Taylor, Gali Halevi and Henk F. Moed**
Guest editors & Editor-in-Chief







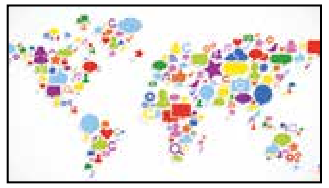








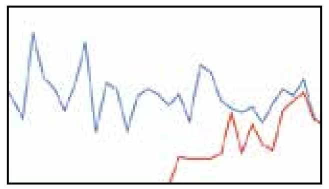












Section 1:
Research overview

# A brief history of altmetrics

**Professor Mike Thelwall, PhD.**
Statistical Cybermetrics Research Group,
University of Wolverhampton, UK

"No one can read everything. We rely on filters to make sense of the scholarly literature, but the narrow, traditional filters are being swamped. However, the growth of new, online scholarly tools allows us to make new filters; these altmetrics reflect the broad, rapid impact of scholarship in this burgeoning ecosystem. We call for more tools and research based on altmetrics (1)".

The above manifesto signaled the birth of altmetrics. It grew from the recognition that the social web provided opportunities to create new metrics for the impact or use of scholarly publications. These metrics could help scholars find important articles and perhaps also evaluate the impact of their articles. At the time there was already a field with similar goals, webometrics, which had created a number of indicators from the web for scholars (e.g. 2) and scholarly publications (e.g. 3), including genre-specific indicators, such as syllabus mentions (4). Moreover, article download indicators (e.g. 5) had also been previously investigated. Nevertheless, altmetrics have been radically more successful because of the wide range of social web services that could be harnessed, from Twitter to Mendeley, and because of the ease with which large scale data could be automatically harnessed from the social web through Applications Programming Interfaces (APIs). Academic research with multiple different approaches is needed to evaluate their value, however (6).

### 1. Scholarly use of the social web

Some research has investigated how scholars use social web services, giving insights into the kinds of activities that altmetrics might reflect. In some cases the answers seem straightforward; for example Mendeley is presumably used to store the academic references that users are interested in – perhaps articles that they have previously read or articles that they plan to read. Counts of article "Readers" in Mendeley might therefore be similar to citation counts in the sense that they could reflect the impact of an article. Mendeley has the advantage that its metrics could be available sooner than traditional citations, since there is no publication delay, and its user base is presumably wider than just publishing scientists. Nevertheless, there are biases, such as towards more junior researchers (7).

In comparison to Mendeley, Twitter has a wider user base and a wider range of potential uses. Nevertheless, it seems that only a minority of articles get tweeted – for example, perhaps as few as 10% of PubMed articles in the Web of Science 2010-2012 have been tweeted (8). Scholars seem to use Twitter to cite articles, but sometimes indirectly (9), which may cause problems for automatically harvesting these citations. Moreover, most tweet (link) citations seem to be relatively trivial in the sense of echoing an article title or a brief summary rather than critically engaging with it (10). There are also disciplinary differences in the extent to which Twitter is used and what it is used for (11) and so, as with citations, Twitter altmetrics should not be used to compare between fields. Another problem is that users may also indicate awareness of others' work by tweeting to them or tweeting about their ideas without citing specific publications (12).



### 2. Evidence for the value of altmetrics

If article level altmetrics are to be useful to help direct potential readers to the more important articles in their field then evidence would be needed to show that articles with higher altmetric scores tended to be, in general, more useful to read. It would be difficult to get direct empirical verification, however, since data from readers about many articles would be needed to cross-reference with altmetric scores. Perhaps the most practical way to demonstrate the value of an altmetric is to show that it can be used to predict the number of future citations to articles, however, since citations are an established indicator of article impact, at least at the statistical level (more cited articles within a field tend to be more highly regarded by scholars, e.g. 13), even though there are many individual examples of articles for which citations are not a good guide to their value. This has been done for tweets to one online medical journal (14) and for citations in research blogs (15). This approach has double value because it shows that altmetric scores are not random but associate with an established (albeit controversial) impact measure and also shows that altmetrics can give earlier evidence of impact than can citation counts.

A second way of getting evidence of the value of altmetrics is to show that their values correlate with citation counts, without demonstrating that the former preceded the latter (of course, correlation does not imply causation and a lack of correlation does not imply worthlessness, but a correlation does imply a relationship with citation impact or at least some of the factors that cause citation impact). This gives some evidence of the validity of altmetrics as an impact indicator but not of their value as an early impact indicator. For example, a study showed that the number of Mendeley readers of articles in the Science and Nature magazines correlated with their citations, but did not prove that Mendeley reader data was available before citation counts (16).

Although the above studies provide good evidence that some altmetrics could have value as impact indicators for a small number of journals, larger scale studies are needed to check additional indicators and a wider range of journals in order to get more general evidence. In response, a large-scale study investigated 11 different altmetrics and up to 208,739 PubMed articles for evidence of a relationship between citations and altmetric scores gathered for 18 months from July 2011. The study found most altmetrics to have a statistically significant positive (Spearman) correlation with citations but one that was too small to be of practical significance (below 0.1). The exceptions were blogs (0.201), research highlights (0.373) and Twitter (-0.190). The reason for the negative correlation for Twitter, and perhaps also for the low correlations in many other cases, could be the rapid increase in citing academic articles in social media, leading to more recent articles being more mentioned even though they were less cited. This suggests that, in most cases, altmetrics have little value for comparing articles published at different points in time, even within the same year. To assess the ability of altmetrics to differentiate between articles published at the same time and in the same journal, the study ran a probabilistic test for up to 1,891 journals per metric to see whether more cited articles tended to have higher altmetric scores, benchmarking against approximately contemporary articles from the same journal. The results gave statistical evidence of an association between higher altmetric scores and citations for most of them for which sufficient data was available (Twitter, Facebook, research highlights, blogs, mainstream media, forums) (17). In summary, it seems that although many altmetrics may have value as indicators of impact, differences over time are critical and so altmetrics need to be normalized in some way in order to allow valid comparisons over time, or they should only be used to compare articles published at the same time (exception: blogs and research highlights).

### 3. Other uses for altmetrics

Altmetrics also have the potential to be used for impact indicators for individual researchers based upon their web presences, although this information should not be used as a primary source of impact information since the extent to which academics possess or exploit social web profiles is variable (e.g. 18; 19; 20). More widely, however, altmetrics should not be used to help evaluate academics for anything important, unless perhaps as complementary measures, because of the ease with which they can be manipulated. In particular, since social websites tend to have no quality control and no formal process to link users to offline identities it would be easy to systematically generate high altmetric scores for any given researcher or set of articles.

A promising future direction for research is to harness altmetrics in new ways in order to gain insights into aspects of research that were previously difficult to get data about, such as the extent to which articles from a field attract readerships from other fields (21) or the value of social media publicity for articles (22). Future research also needs to investigate disciplinary differences in the validity and value of different types of altmetrics. Currently it seems that most articles don't get mentioned in the social web in a way that can be easily identified for use in altmetrics (e.g. 23), but this may change in the future.

**References:**


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

2. Cronin, B., Snyder, H.W., Rosenbaum, H., Martinson, A. & Callahan, E. (1998) "Invoked on the Web", Journal of the American Society for Information Science, Vol. 49, No. 14, pp. 1319-1328.

3. Vaughan, L. & Shaw, D. (2003) "Bibliographic and web citations: what is the difference?", Journal of the American Society for Information Science and Technology, Vol.54, No. 14, pp. 1313-1322.

4. 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, Vol. 59, No. 13, pp. 2060-2069.

5. Shuai, X., Pepe, A., & Bollen, J. (2012) "How the scientific community reacts to newly submitted preprints: Article downloads, Twitter mentions, and citations", PLOS ONE, Vol. 7 No. 11, e47523.

6. Sud, P. & Thelwall, M. (2014) "Evaluating altmetrics", Scientometrics, Vol. 98, No. 2, pp. 1131-1143.

7. Mohammadi, E., Thelwall, M., Haustein, S. & Larivière, V. (in press) "Who reads research articles? An altmetrics analysis of Mendeley user categories", Journal of the Association for Information Science and Technology.

8. Haustein, S., Peters, I., Sugimoto, C.R., Thelwall, M. & Larivière, V. (in press) "Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature", Journal of the Association for Information Science and Technology.

9. Priem, J., & Costello, K.L. (2010) "How and why scholars cite on Twitter", Proceedings of the American Society for Information Science and Technology, Vol. 47, pp. 1-4.

10. Thelwall, M., Tsou, A., Weingart, S., Holmberg, K. & Haustein, S. (2013) "Tweeting links to academic articles", Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics, Vol. 17, No. 1, paper 1.

11. Holmberg, K. & Thelwall, M. (in press) "Disciplinary differences in Twitter scholarly communication", Scientometrics.

12. Weller, K., Dröge, E. & Puschmann, C. (2011) "Citation analysis in Twitter: Approaches for defining and measuring information flows within tweets during scientific conferences", In Proceedings of Making Sense of Microposts Workshop (# MSM2011).

13. Franceschet, M. & Costantini, A. (2011) "The first Italian research assessment exercise: A bibliometric perspective", Journal of Informetrics, Vol. 5, No. 2, pp. 275-291.

14. 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, Vol.13, No. 4, e123.

15. 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, Vol. 65, No. 5, pp. 1018–1027.

16. Li, X., Thelwall, M. & Giustini, D. (2012) "Validating online reference managers for scholarly impact measurement", Scientometrics, Vol. 91, No. 2, pp. 461-471.

17. Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C. (2013) "Do altmetrics work? Twitter and ten other candidates", PLOS ONE, Vol. 8, No. 5, e64841. doi:10.1371/journal.pone.0064841.

18. Bar-Ilan, J., Haustein, S., Peters, I., Priem, J., Shema, H. & Terliesner, J. (2012) "Beyond citations: Scholars' visibility on the social Web", Proceedings of 17th International Conference on Science and Technology Indicators (pp. 98-109), Montréal: Science-Metrix and OST.

19. Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H. & Terliesner, J. (in press) "Coverage and adoption of altmetrics sources in the bibliometric community", Scientometrics.

20. Mas Bleda, A., Thelwall, M., Kousha, K. & Aguillo, I. (2013) "European highly cited scientists' presence in the social web", In 14th International Society of Scientometrics and Informetrics Conference (ISSI 2013) (pp. 98-109).

21. Mohammadi, E. & Thelwall, M. (in press) "Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows", Journal of the Association for Information Science and Technology.

22. 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, Vol. 8, No. 7, e68914.

23. Zahedi, Z., Costas, R. & Wouters, P. (in press) "How well developed are Altmetrics? Cross-disciplinary analysis of the presence of "alternative metrics" in scientific publications", Scientometrics.




Section 2:
Behind the data

# Predicting citation counts

**Ron Daniel Jr., Ph.D.**
Director, Elsevier Labs

**Abstract**

Many articles have been written about efforts to predict how many citations a research article will receive, based on indicators available before or shortly after publication. These efforts have widely varying results, with one effort predicting 14% of the variance in citation while another study ten years later reached over 92%. What was learned in that decade? What can this tell us about potentially valuable altmetrics, and are there areas in which new altmetrics might be discovered?

## 1. Introduction

This special issue of Research Trends is about altmetrics – alternatives to the use of citation counts as the metric for assessing the impact of an article, a researcher or a journal. Citation counts do not tell the whole story (e.g. they don't value useful research software tools, useful advisory papers to young researchers, or research that can't be published for commercial or government security reasons). Having additional metrics to provide a more complete picture is a very welcome development. However, even in a future in which additional metrics are available to assess impact, citation counts will remain first among equals because of their intimate connection with the text of the article and the article's basis on prior work.

The current and continued importance of citation counts has led to the desire to predict how often an article will be cited in order to predict its future importance. Such predictions could be used to decide if an article should be published in one journal vs. another, to flag new research for scrutiny before citation counts have had time to accrue, to assess the development of a young researcher before many counts could have accrued, etc. Many articles have been written on this topic, but there has been very little consistency in their results. Four studies between 2002 and 2012 found that they could predict 14% (1), 20% (2), 60% (3) and 90% (4) of the variance in citation counts a few years after publication based on features available before or shortly after publication. A discrepancy this great requires some explanation!

This article has three goals. The first is to explain the discrepancy in the previous research results. The second is to evaluate the various indicators, a.k.a. features, which were used in the four articles. Features that are predictive of eventual citation counts might be particularly valuable altmetrics that serve as leading indicators of an article's merit. We need to be cautious when comparing results across the studies as they use different scientific domains, make predictions over different time periods, use different statistical methods, obtain results through different procedures, etc. For example, one study measured "newsworthiness" by having readers estimate it; another did so by searching news archives. Both found it to be a notable factor but not necessarily statistically significant. All of this means the results are only loosely comparable. However, we can look within each study to see which features did have significant effects and the relative magnitude of those effects. If the same feature is found to be significant, or not, across all studies then we are fairly safe in drawing conclusions about its utility. The third goal is to see if we can draw conclusions about potentially valuable altmetrics and areas where new altmetrics might be discovered.

## 2. Prediction Features for Pre-Publication Articles

From the time an article is first conceived, features begin to accrue that we can use to predict its future citation counts. This section looks at features available from the inception of the article up to the time where it has been accepted for publication in a journal. The features can be further subdivided into those that apply to the article itself, to the authors of the article, and to the journal which has accepted the article for publication. Those three categories are named as Content, Author, and Venue (4). What we will see is that even before an article is published, we have enough information to make fairly good predictions about its future rate of citations.

### 2.1 Content

**Study Design Factors:**

The earliest article we review (1), published in 2002, made the assumption that high quality research would be more heavily cited. They thought about what made high quality research and looked for corresponding features such as sample size, controls, blinding, etc. Sample size and the presence of a control group were found to have some effect, but not to the level of statistical significance. The other factors (blinded, randomized, prospective v. retrospective) were even weaker. The second article (2), published in 2007, also looked at study design factors and found them to have little effect. What they did find, however, was that large studies funded by industry, and with industry-favoring results, were statistically significant predictors of higher rates of citation in the future. These features are understandably important in the medical therapeutic space. Such studies are likely to show drugs and other therapies soon to be available. These factors don't seem likely to generalize to other domains.

**Topic:**

Unlike the first study, which was confined to emergency medicine, the second study (2) considered the effect of the topic of the article. They found that cardiovascular and oncology articles were more likely to be cited than those on other topics such as anesthesiology, dermatology, endocrinology, gastroenterology, etc. Given the relative death rates of heart disease and cancer to the implications of the other specialties, this seems reasonable. Similarly, the third article (3), published in 2008, found that articles which provided therapeutic information were more cited, as were those which provided original research as opposed to review articles. That study also found that longer articles were cited fewer times, in a weak but statistically significant way. It also found that the more references an article contained, the more likely it was to be cited, although this effect was weak and not significant. The fourth article (4), published in 2012, found a weak effect that the more topics an article covered the higher the number of citations it received.

Table 1 lists the content-based features available before publication which were used in the four studies. Statistically significant values are highlighted. The key things to notice in this table are how few content-based features are significant, and how few of the features are used in multiple studies.

| | Callaham 2002 (1) | Kulkarni 2007 (2) | Lokker 2008 (3) | Yan 2012 (4) |
|---|---|---|---|---|
| # study participants | 26.5% | 3.1, p=.04 | < .001, p=.295 | |
| Newsworthiness score | 26% | 13.5, p<.001 | .133, p=.161 | |
| Control group | 24.3% | | | |
| Quality score | 15.8% | | | |
| Explicit hypothesis | 4.7% | | | |
| Prospective v. retrospective study | 2.7% | 3.6, p=.01 | .477, p=.009 | |
| Type of study participants | 2.1% | | | |
| Blinded | .07% | | | |
| Randomized | 0 | 13.4, p=.01 | | |
| Positive results | 0 | | | |
| Industry funding | | 19.9, p<.001 | | |
| Industry favoring result | | 19.4, p<.001 | | |
| Location of study | | 11.9, p=.001 | | |
| Topic | | 17.8, p=.001 | | |
| Original v. review article | | | .477, p=.009 | |
| # pages | | | -.011, p< .001 | |
| Structured abstract | | | -.8, p=.002 | |
| # cited references | | | .004, p=.008 | |
| Multicenter study | | | .367, p=.014 | |
| Therapy v. other article | | | .339, p=.023 | |
| Word count of abstract | | | -.0003, p=.658 | |
| Semi-structured abstract | | | .071, p=.746 | |
| Nation of first author | | | -.037, p=.762 | |
| Novelty | | | | .059 |
| Topic rank | | | | .079 |
| Diversity of topics in article | | | | .157 |

**Table 1:** Content-based features available pre-publication.



### 2.2 Author

The effects of the author were not considered in (1). The second study (2) only looked at whether the author byline indicated group authorship. This was found to be the most significant prediction feature in their study! This was a very important result. It indicated that article importance or quality was not easily measured by the presence or absence of some features we might call "good research practice". That realization led to significantly improved prediction accuracy in later work.

The last two papers (3, 4) looked at author-related features in more detail.

Both Lokker (3) and Yan (4) looked at the count of the number of co-authors. Lokker (3) found that count to be a significant factor, but Yan (4) did not. Yan looked at several other author-related features. The Maximum Past Influence of the Author (MPIA) is the citation count for the author's most-cited paper. The Total Past Influence of the Author (TPIA) is the sum of the citation count across the author's body of work. The MPIA was found to be predictive but the TPIA was essentially useless.

A strong result in (4) was the author's rank in citation counts. The citation counts for all the author's works were averaged, and the average counts were sorted to rank the authors. Figure 1, reproduced from (4), shows that being a very highly cited author is predictive of future citation counts. The rich get richer in other words. As can be seen however, this effect is limited and is only strong for authors in the top ranks of citation frequency.

Considerable attention has been paid to author-related factors in articles beyond the four we review here. (3) provides citations of articles that look at other effects such as nationality, gender, and alphabetic order of the author names.

Table 2 summarizes the effect of the author-based features available before publication. The key thing to notice is that the earliest study made no use of author information, while the latest and most accurate article tried many author-based features.

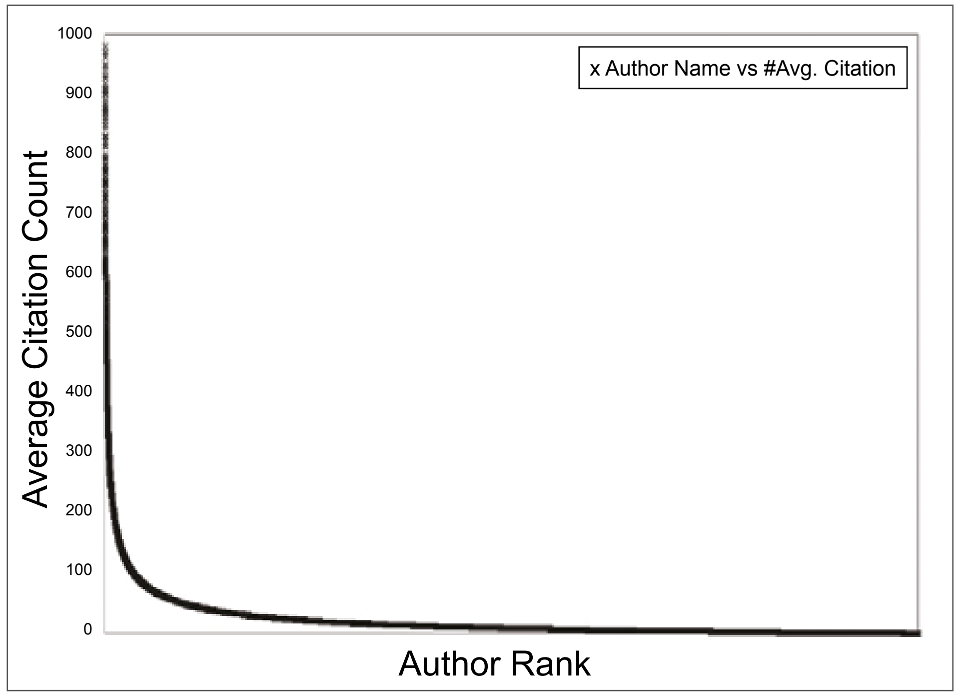



**Figure 1:** Citation Counts vs. Rank of Author's Average Citation Count. Figure reproduced from Yan et al (2012) (4). We have sought permission for re-use of this figure.

| | Callaham 2002 (1) | Kulkarni 2007 (2) | Lokker 2008 (3) | Yan 2012 (4) |
|---|---|---|---|---|
| # authors | | 20.3, p<.001 | .087, p<.001 | .056 |
| Nation of first author | | | .037, p=.762 | |
| Author rank (by citations) | | | | .593 |
| h-index | | | | .244 |
| MPIA (Max Past Influence) | | | | .585 |
| TPIA (Total Past Influence) | | | | .048 |
| Productivity | | | | .198 |
| Sociality | | | | .249 |
| Authority | | | | .155 |
| Versatility | | | | .160 |
| Recency | | | | .101 |

**Table 2:** Author-based features available pre-publication.

**2.3 Venue:**

The only statistically significant variable found in the first study (1) was the impact factor of the journal in which the article was published. This was an early indication of the power of the venue in determining future citation counts. If we know the journal the article will be published in, we can make more confident predictions about its eventual citation count.

The second study only considered three of the top-line medical journals – JAMA, NEJM, and The Lancet. Nevertheless, they found a significant difference in citation rates between articles in those publications.

The third study did not use the impact factor, as it did not apply to all their sources for content. They discovered other measures that also reflected the article's venue. The strongest are the number of databases that index the journal, and the proportion of articles from the journal which are abstracted within two months by Abstracting & Indexing services and synoptic journals.

Table 3 summarizes the effect of the venue-based features available before publication. Note that no feature is used in more than one study. Curiously, impact factor was the only significant feature found in (1), but it is not used in the later studies. Perhaps the most surprising outcome summarized in this table is the strong effect due to the venues chosen by secondary publication sources like databases, A&I services, and synoptic journals. Given the concerns we all have about infoglut, it is both interesting to see the strength of this effect, and concerning that these effects do not seem to have been featured in any previous altmetric studies. More research in this direction seems justified.

| | Callaham 2002 (1) | Kulkarni 2007 (2) | Lokker 2008 (3) | Yan 2012 (4) |
|---|---|---|---|---|
| Impact factor of publishing journal | Strongest factor, relative contribution = 100% | | | |
| Accepted for presentation at meeting | 5.5% | | | |
| Journal | | 16.3, p<.001 | | |
| Month of publication | | 0.7, p=.5 | | |
| Proportion of articles abstracted | | | 8.18, p<.001 | |
| # databases indexing | | | .039, p<.001 | |
| Venue rank | | | | .337 |
| Venue centrality | | | | .049 |
| Max past influence of venue | | | | .329 |
| Total past influence of venue | | | | .023 |

**Table 3:** Venue-based features available pre-publication.

The previous section showed that venues whose articles were frequently selected for abstraction tended to have more highly cited articles. For a single article, the number of times it is abstracted is also a statistically significant predictor (3) which is not available until shortly after publication. That study also showed that articles which were judged "clinically relevant" by the staff of a recommendation service were significantly more likely to have more citations in the future. These results are notable for the same reason as the venue results in the previous section – secondary publication sources have a predictive effect which is not being captured in current altmetrics.

**3. Prediction Features for Newly Published Articles**

By publication time, we know many facts about the Content, Author, and Venue. In the newly published phase of the article's lifecycle we shift our attention to early perceptions of the quality of the article, and to early indications of the use of the article.

There are many features that could give us early indications of how often articles are being used, or the perceptions that the early users have of them. Those include:

- Preprint access counts from arXive, etc.
- General Social Media mentions (Twitter, Facebook, …)
- Scientific Social Media mentions (Mendeley, del.icio.us, CiteULike, …)
- Sentiments expressed in early mentions
- Early download counts from services like ScienceDirect
- Early citations of the article shown in services like Scopus



| | Callaham 2002 (1) | Kulkarni 2007 (2) | Lokker 2008 (3) | Yan 2012 (4) |
|---|---|---|---|---|
| Newsworthiness score | 26% | 13.5, p<.001 | .133, p=.161 | |
| Abstracted in evidence based medicine journals | | | .839, p<.001 | |
| Clinical relevance score | | | .418, p<.001 | |
| # disciplines rating the article | | | .038, p=.371 | |
| Time to article being rated | | | -.009, p=.513 | |
| # views or alerts sent | | | -.069, p=.938 | |

**Table 4:** Features available in first months of publication.

These features were not used in the four studies, but there is good reason to believe that these features will be useful in predicting future citation counts. As mentioned in (3):

"Thirty three percent of the variance in citation counts of BMJ articles were found to be based on counts of online hits and number of pages (5)".

Table 4 shows the effect of features available shortly after the article is published. The most noticeable aspect of this table is that very few post-publication features were used in the studies other than (3).

**4. Prediction Features for Mature Articles**

The fourth article (4) looked at temporal factors such as age of the article, as well as regression constants to control the growth and decay of citation rates over time. These results were not strong and other studies did not look at features for mature articles so a summary table is not provided. While none of the studies made significant use of features that become available later in the publication lifecycle, there is no shortage of possibilities. For example, we might look at a Page-Rank like scoring of the influence of the papers citing the particular paper of interest.

Nevertheless, the short story is clear. By the time an article is a few months old, we can make good predictions of its likelihood of future citations - especially for those articles which end up being highly cited. Lokker noted that for the papers with the highest citation counts at two years after publication, "Cited articles in the top half and top third were predicted with 83% and 61% sensitivity and 72% and 82% specificity" (3). In other words, only about 20% of the papers which ended up being highly cited were not predicted to be that way.

**5. Conclusion**

Despite low performance in early studies (14% in 2002), it has become clear over time that it is possible to make good predictions (92% in 2012) of the frequency of future citations. How was this advance achieved? Quite simply, the features being used in the later studies are very different from those used in the earliest ones. The early studies tried to use features around the content, but later work found those to be the weakest while features around the Author and Venue were the most predictive. If we set the power of the Author features to 1.0, the relative power of the Venue and Content features would be about .63 and .25, respectively. We cannot directly compare results across columns, and it is not safe to predict the accuracy any new study might achieve. All of the studies used different domains of literature, predictions over different time periods, different statistical measures, etc. Nevertheless, the pattern seems clear.

It is also interesting, and mildly reassuring, to see that the strongest of these measures operate, to some degree, in a manner independent of each other. Author and Venue are the two most predictive features. However, selecting an article for a journal is usually done in a peer review process that is blind to the identity of the author. Note that this also means these measures are not well-suited for an editorial board to choose articles, since the Venue would be constant and they could not look at the author's publication rank.

In a perfect world, the content of an article would determine its future citation count. We do not, however, have any easily-computed metric for the intrinsic quality and merit of an article. This is where Lokker's results about the importance of secondary sources such as the databases and synoptic journals are most interesting. We see that in the absence of reliable, easily-computed metrics, the subjective human-in-the-loop procedures of peer review, editorial boards, selection for secondary publications, and scientific reputation provide existing mechanisms which fill that void. This provides a potential area of altmetric research to obtain such measures in various fields and compare them with current altmetrics for a variety of purposes.

**References:**

1. Callaham, M., Wears, R.L., Weber, E. (2002) "Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals", JAMA, Vol. 287, pp.2847-50, Available at http://www.ncbi.nlm.nih.gov/pubmed/12038930/

2. Kulkarni, A.V., Busse, J.W., Shams, I. (2007) "Characteristics associated with citation rate of the medical literature", PLOS ONE; 2:e403, Available at http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0000403

3. Lokker, C., McKibbon, K.A., McKinlay, R.J., Wilczynski, N.L. and Haynes, R.B. (2008) "Prediction of citation counts for clinical articles at two years using data available within three weeks of publication: retrospective cohort study", BMJ, Mar 22, 2008; Vol. 336, No. 7645, pp. 655–657. Available online: http://dx.doi.org/10.1136/bmj.39482.526713.BE

4. Yan, R., Huang, C., Tang, J., Zhang, Y., and Li, X. (2012) "To Better Stand on the Shoulder of Giants", JCDL, Available at: http://keg.cs.tsinghua.edu.cn/jietang/publications/JCDL12-Yan-et-al-To-Better-Stand-on-the-Shoulder-of-Giants.pdf

5. Perneger, T.V. (2004) "Relation between online "hit counts" and subsequent citations: prospective study of research papers in the BMJ", BMJ, Vol. 329, No. 7465, pp. 546-547, Available at: http://www.bmj.com/content/329/7465/546

Section 3:
Value of bibliometrics

# Scholarly blogs are a promising altmetric source

**Hadas Shema, MA[1&2], Professor Judit Bar-Ilan[1] and Professor Mike Thelwall, PhD[3]**

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

[2] Israeli Inter-University Center for e-Learning (MEITAL)

[3] Statistical Cybermetrics Research Group, University of Wolverhampton, UK

"…Russel Lyons who posits that Christakis' and Fowler's work is a great example of statistical illiteracy, and that the conclusion drawn from their data, that obesity is socially contagious, is severely flawed and can't be made".

Blogger Yoni Freedhoff, MD, in his blog "Weighty Matters" (1).

Scholarly blogs are one of the most prominent information sources for altmetrics and are reported in the main altmetric services (e.g. ImpactStory.org, altmetric.com). National Geographic, the Nature Group, Scientific American, and the PLOS (Public Library of Science) journals all have science blogging networks. Scholarly blogs have been defined as "blogs written by academic experts that are dedicated in large part to scientific content" (p. 171) (2). This definition is rather vague, because of the difficulty defining an "expert" and "scholarly content". A scholarly blog can also be defined by its platform (e.g. SciLogs, Scienceblogs.com), by the media outlet hosting it (e.g. Scientific American, The Guardian), by the affiliation or education of its blogger(s), by its contents, discipline, links to other blogs (blog roll) or any combination of the above. A blog by a single graduate student posting about subjects related to her research can be as 'scholarly' as a blog by several experienced researchers posting across disciplines.

A number of studies have implicated blog coverage as an indicator of scholarly impact. The commercial firm altmetric.com's data from July 2011 up to January 1st, 2013 has been used to study the association of potential altmetric sources (Twitter, Facebook wall posts, research highlights, blogs, mainstream media, forums, Google+, LinkedIn, Pinterest, question and answer sites, and Reddit) and Web of Science (WoS) citations (3). The blogs in the sample came from the Nature.com blogging network and the blogging aggregators ResearchBlogging (RB) and ScienceSeeker. The study compared the number of times an article was covered in blogs (they calculated each altmetric source separately), with two articles that have also received a mention in an altmetric source (not necessarily a blog), one published shortly before the article in question and the second published shortly after. The authors concluded that "In summary, there is clear evidence that three altmetrics (tweets, FbWalls, blogs) tend to associate with citations at the level of individual journals" (p. 4).

Another study of altmetric.com data (4) looked at altmetric mentions of articles (with DOIs - Digital Object Identifiers) in various metric sources (reference managers excluded) from July 2011 to mid-October 2013, and correlated them with articles indexed in WoS and the citations they accumulated in 2012 (in part of the analysis the corpus was used in full, in another part only the July-December 2011 data, to allow a full year of citations). They found a relatively strong correlation between blog and news outlet mentions and citations. A factor analysis found that blog and news outlet mentions belonged to one dimension, while other altmetrics (Twitter, Google+ and Facebook walls) belonged to another. This aligns with Taylor and Plume (5), who studied altmetric.com data from the last four months of 2013. Taylor and Plume classified the altmetric data sources into four categories: social activity (e.g. Facebook, Twitter), scholarly activity (e.g. bookmarking on Mendeley), scholarly commentary (e.g. blogs, F1000Prime) and mass media coverage. They found that between the top 0.5% of articles in each category, the highest chance of overlap was between mass media coverage and scholarly commentary.




**Figure 1:** A typical RB post snippet.

A small-scale study (6) looked at the effect of blog post coverage on 16 clinical pain PLOS ONE articles. The blog posts were published in the blog BodyinMind.org, which had at that time over 2,500 unique views per week, and were disseminated by social media (RB, Twitter, Facebook, LinkedIn). In the week after the blog post coverage of each article, there were on average about 3 additional downloads of the article per day and 12 additional HTML views. The authors did not find a correlation between Scopus citations a year after the blog post publications and social media metrics or HTML views, but did find a moderate correlation between PDF downloads and citations.

### The structured blog citation

Scholarly bloggers often comment on material from peer-reviewed journals, but unlike authors of peer-reviewed articles, they are not obligated to reference their sources in a formal way. Despite this, scientific bloggers have mentioned in interviews that they would have liked to use references in a similar way to the way that they cite in scholarly articles (7).

The aggregator ResearchBlogging.org (RB) was built to answer this need. Launched in late 2008, it aggregates blog posts referring specifically to peer-reviewed research. It is a self-selecting aggregator that allows bloggers to refer to peer-reviewed research in an academic citation format. Bloggers discussing peer-reviewed research can register with the aggregator and after they mark relevant posts in their blog, these posts appear on the aggregator site, giving one-stop access to a variety of research reviews from different authors. The site's human editors ensure that blogs submitted to the aggregator follow its guidelines and are of appropriate quality. RB already has an altmetric role; it currently serves as one of the article level metrics (ALM) displayed for each article in the journal PLOS ONE (8). By the end of 2011, RB had more than 1,230 active blogs and about 27,000 posts (9). These posts seem to be a transitional phase between traditional scholarly discourse and rapid, informal blog writing - a scientometric Archaeopteryx.

The first study of RB, which looked at its Chemistry category, found that most blog posts were about current research and came from high-impact niche journals as well as prestigious multidisciplinary journals (10). Similar results were also found in subsequent studies (9), (11), (12) for other RB categories. Bloggers prefer to cover articles from top multidisciplinary journals, the most popular being (in alphabetical order) Nature, PLOS ONE, Proceedings of the National Academy of Sciences of the United States of America (PNAS) and Science. Most of the posts aggregated in RB are written in English. The bloggers classify their posts into pre-defined categories, the most popular categories being Biology, Health Sciences, Neuroscience and Psychology (9), (11).

RB (see Figure 1) was the data source for our blog study of the association between blog coverage and traditional citations (13). We took a different approach than (3) and (4), not taking into account the number of times an article was covered in blogs, but only whether it was covered or not. We compared journal articles from 2009 and 2010 which were covered in blog posts from the same year (i.e. a 2009 article covered in a 2009 post, a 2010 article covered in a 2010 post) with the general population of articles from the same journal in the same year, to see if these articles received a higher number of citations in the years after their publication in comparison to articles from the same journal and year not covered in blogs. In 2009 a total of 58% (7 out of 12 journals), and in 2010 a total of 68% (13 out of 19 journals) of journals published articles covered by blogs that attracted more citations than articles from the same journal and year that were not covered by blogs. The most striking difference in medians was between articles covered by the New England Journal of Medicine (NEJM) in 2009 (172) and NEJM articles from 2009 which were not covered in blogs (76). We also found an association between coverage of the NEJM articles in blogs and their coverage in the Reuters and New York Times websites. Twenty-one out of the 26 NEJM articles in our 2009 sample (81%) and 20 out of 38 (53%) NEJM sample articles in 2010 were covered by Reuters, the New York Times, or both. This aligns with the findings of (5) as well as those presented in (4). News coverage has been known to correlate with a higher level of citations (14), and it is a possibility that the higher level of citations that many articles covered in blogs enjoy reflects the bloggers' tendency to choose articles covered by mass media. We cannot tell if this tendency comes from the direct influence of mass media coverage on its scholarly blogger consumers, or if the bloggers' tastes simply align with those of the mass media.

### In conclusion

There is evidence that blog coverage of scholarly articles associates with increased visibility and impact. Unfortunately, there are a number of obstacles that might limit the use of blog posts as an altmetric source. First, only a small percentage of articles is covered in blogs (e.g. 1.9% of the articles studied in (4)).

Second, the definition of "scholarly blogs" and the decision about which blog data to use is problematic. When relying on certain aggregators or networks for blog data we miss the impact of articles covered by blogs outside the data collection range. The coverage problem is not specific to blogs, or even to altmetrics, but extends to bibliometric databases, which also have to choose which sources to index.

Third, there is a lack of sustainability. While most peer-reviewed journals enjoy professional archiving and printed copies, blogs can close down or move without leaving a trace (except perhaps in archive.org and similar sites). For blog-derived data to be reliable, they have to be better indexed and archived.

> If peer-reviewed journals citations are "frozen footprints," (15, abstract) then citations in blogs, and altmetrics in general, are footprints in quicksand. In spite of these limitations, we consider blogs to be an especially promising altmetric source.

The effort required to write a blog post (assuming it isn't spam or computer-generated) is much greater than the effort needed to tweet, "like" or bookmark an article. Scholarly blogging at its best can be a type of post-publication peer-review, scholarly commentary or citizen journalism and its presence can be used as an impact indicator.

### References:


1. Freedhoff, Y. (2011, June 20) "Obesity is contagious, or is it? A sober second look at obesity and social networks". [BlogPost]. Retrieved from http://www.weightymatters.ca/2011/06/obesitys-contagious-or-is-it-sober.html

2. Puschmann, C., Mahrt, M. (2012) "Scholarly blogging: A new form of publishing or science journalism 2.0?" In: A. Tokar, M. Beurskens, S. Keuneke, M. Mahrt, I. Peters, C. Puschmann, K. Weller, & T. van Treeck (Eds.), Science and the Internet (pp. 171-182). Düsseldorf: Düsseldorf University Press.

3. Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C. (2013) "Do altmetrics work? Twitter and ten other social Web services", PLOS ONE, Vol. 8, No. 5, e64841.

4. Costas, R., Zahedi, Z. & Wouters, P. (2014) "Do altmetrics correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective". ArXiv preprint arXiv:1401.4321.

5. Taylor, M. & Plume, A. (2014) "Party papers or policy discussions: an examination of highly shared papers using altmetric data", Research Trends, Issue 36, 17-20.

6. 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, Vol. 8, No. 7, e68914.

7. Kjellberg, S. (2010) "I am a blogging researcher: Motivations for blogging in a scholarly context" First Monday, Vol. 15, No. 8.

8. PLOS ONE. (n.d.) "Article-level metrics information", Retrieved from http://www.plosone.org/static/almInfo

9. Fausto, S., Machado, F.A., Bento, L.F., Iamarino, A., Nahas, T.R. & Munger, D.S. (2012) "Research blogging: Indexing and registering the change in science 2.0", PLOS ONE, Vol. 7, No. 12, e50109.

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

11. Shema, H., Bar-Ilan, J. & Thelwall, M. (2012) "Research blogs and the discussion of scholarly information", PLOS ONE, 7(5), e35869.

12. Shema, Bar-Ilan & Thelwall (in press) "How is research blogged? A content analysis approach", Journal of the Association for Information Science and Technology.

13. Shema, H., Bar-Ilan, J. and 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. Advanced online publication doi: 10.1002/asi.23037

14. Phillips, D.P., Kanter, E.J., Bednarczyk, B. & Tastad, P.L. (1991) "Importance of the lay press in the transmission of medical knowledge to the scientific community", New England Journal of Medicine, Vol. 325, No. 16, pp. 1180-1183.

15. Cronin, B. (1981) "The need for a theory of citing", Journal of Documentation, Vol. 37, No. 1, pp. 16-24.




## Section 4:
Behind the data

# A comparison of citations, downloads and readership data for an information systems journal

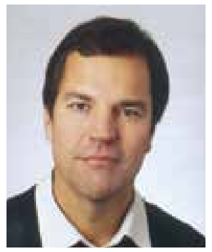

**Associate Professor Dr. Christian Schlögl[1], Dr. Juan Gorraiz[2], Dr. Christian Gumpenberger[2], Kris Jack, PhD[3], & Dr. Peter Kraker[4]**

[1] University of Graz, Institute of Information Science and Information Systems,

[2] University of Vienna, Vienna University Library, Dept. of Bibliometrics

[3] Mendeley, London

[4] Know-Center, Graz

### Introduction

In the past, citations were the prime source for measuring scholarly impact. With the advent of altmetrics, it is possible to detect the use and consumption of scholarly publishing on a much broader basis (1). According to Plum Analytics, besides citations, metrics can be provided on the basis of usage, captures, mentions, and social media (2). In this contribution we will elaborate on the similarities and differences between one example from each of the first three metrics types mentioned above: citations from Scopus; downloads from ScienceDirect; and readership counts from Mendeley. As a use case, we chose the Information Systems journal Information and Management, including all issues from 2002 to 2011.

Information and Management is one of the leading Information Systems journals. It usually publishes eight issues per year and has a geographical focus on Anglo-American and South East Asian countries with regard to authorship and associate editors. From the nearly 600 research articles in the period of analysis, half were published by authors from the U.S. and approximately one third by authors from Taiwan, China, South Korea and Singapore.

Citations and downloads were provided by Elsevier in the framework of the Elsevier Bibliometric Research Program (EBRP) (3). For the publications of the analyzed Information Systems journal all monthly downloads were made available from ScienceDirect (4) and all monthly citations from Scopus (5). Furthermore, we received the readership counts from Mendeley (6). Mendeley is a social reference management system which helps users with the organization of their personal research libraries. The articles, provided by users around the world, are crowd-sourced into a single collection called the Mendeley research catalogue. This makes it possible to calculate the readership frequencies of an article which indicates how many Mendeley users have added it to their personal research library. At the time of writing, this catalogue contains more than 110 million unique articles, crowd-sourced from over 2.5 million users, making it an interesting source of data for large scale network analysis.

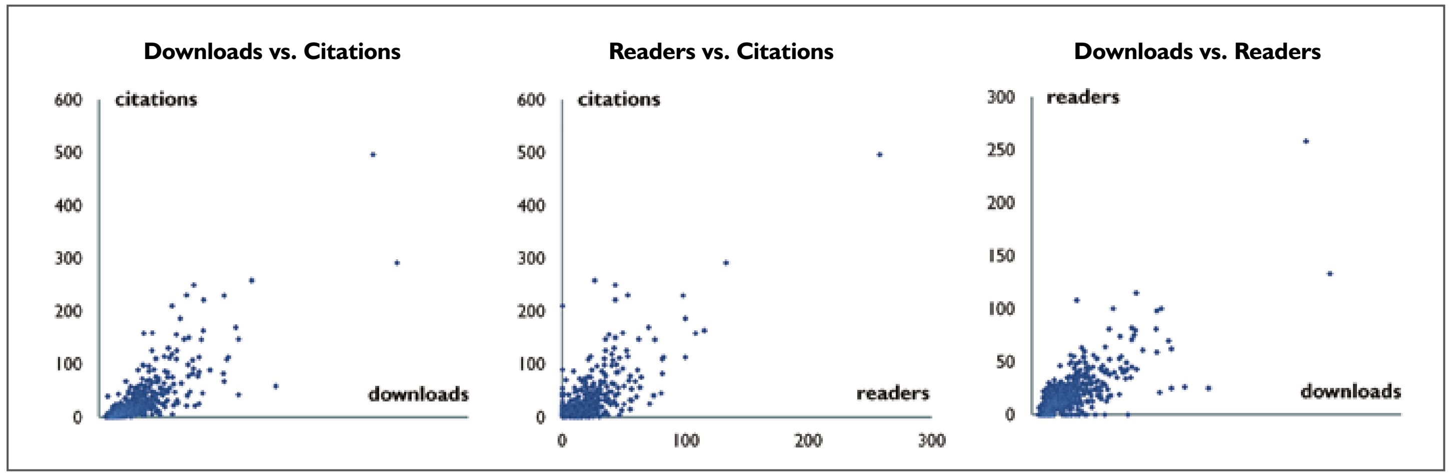



**Figure 1:** Downloads vs. cites vs. readers (publication year: 2002-2011, doc type: full-length article).

**Relation between citations, downloads and readership counts**

Figure 1 shows the relationship between downloads, citations and readership frequencies for all full-length articles (7) published between 2002 and 2011. Data were provided mid 2012 for citations and downloads and in October 2012 for readership data. As can be seen, articles that are downloaded more often are in general cited more frequently. Furthermore, the more frequently an article can be found in Mendeley user libraries (number of readers), the more often it is usually downloaded and cited.

This is also reflected through the rank correlations (Spearman) among these three indicators, which are 0.76 between citations and downloads, 0.66 between downloads and readership counts, and 0.59 between citations and readership counts. Similar correlations were computed for another Information Systems journal (Journal of Strategic Information Systems) (8). The fact that there is a strong but not a perfect correlation between these three indicators gives a first indication that they measure partly different aspects of scholarly communication. Therefore, we will look deeper into each measure. In a first step, we will investigate possible differences in obsolescence characteristics. Since Mendeley started only in 2009 and had a high growth in its user base since then, we will perform the obsolescence analysis only for citations and downloads.

**Obsolescence characteristics of citations and downloads**

Figure 2 shows the year-wise citations and the year-wise downloads (for privacy reasons, the download numbers are not specified) for an article (9) published in Information and Management in 2004. Since the article was put online in ScienceDirect on October 14th, 2003, it was already downloaded before the print publication year. Typically, the download numbers peak in the (print) publication year. In the following years, the download volume normally decreases slowly. However, a new increase is possible, for instance, due to the citation impact of an article. To some degree, also the general rise of downloads (users) in ScienceDirect might have some effect. In contrast, citations are low in the year of publication and reach their maximum several years later.

To give a more general picture, we show the year-wise downloads for all full-length articles published in Information and Management from 2002 and 2011 in Table 1. For privacy reasons, we only give relational numbers. As a matter of fact, the download numbers are one “magnitude” higher than the citation counts. As can be seen, the download maximum (formatted in bold) always (besides 2002) occurs in the (print) publication year. However, for older volumes (publication years: 2002 - 2005) a re-increase in the downloads can be observed in the years 2008 and 2009 after a decline in the previous years. Table 2 displays the year-wise citations for the corresponding document types in Scopus (article, proceedings paper, and review) and confirms what was already mentioned above.

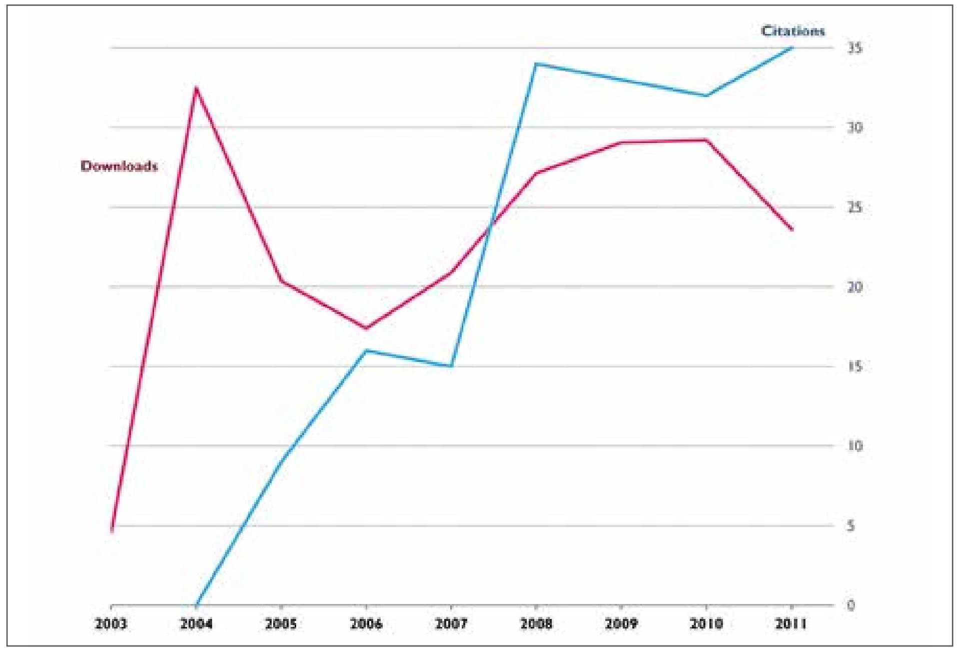



**Figure 2:** Year-wise downloads and citations for the article by Amoako-Gyampah and Salam (2004) (9).



| Pub year | n | Download year | | | | | | | | | | | Downloads |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | All | per FLA – relations[1] |
| 2002 | 46 | 1.7 | 1.6 | 1.2 | 1.1 | 1.0 | 1.2 | 1.7 | 2.1 | 2.2 | 1.9 | 15.8 | 7.6*x |
| 2003 | 73 | 0.5 | 3.1 | 2.2 | 1.5 | 1.3 | 1.4 | 1.9 | 2.2 | 2.2 | 1.9 | 18.3 | 5.6*x |
| 2004 | 71 | | 0.4 | 4.2 | 2.7 | 2.0 | 2.1 | 2.5 | 2.9 | 3.0 | 2.6 | 22.3 | 7.0*x |
| 2005 | 61 | | | 0.6 | 3.6 | 2.1 | 1.7 | 2.0 | 2.5 | 2.4 | 2.1 | 17.1 | 6.2*x |
| 2006 | 78 | | | | 0.4 | 3.5 | 3.0 | 2.6 | 3.0 | 3.0 | 2.6 | 18.1 | 5.1*x |
| 2007 | 48 | | | | | 0.0 | 2.6 | 2.2 | 2.0 | 2.1 | 1.7 | 10.7 | 4.9*x |
| 2008 | 62 | | | | | | 0.0 | 4.0 | 3.9 | 3.2 | 2.7 | 13.8 | 4.9*x |
| 2009 | 56 | | | | | | | 0.0 | 3.8 | 3.1 | 2.4 | 9.3 | 3.7*x |
| 2010 | 42 | | | | | | | | 0.2 | 2.9 | 2.1 | 5.2 | 2.8*x |
| 2011 | 44 | | | | | | | | | 0.0 | 2.0 | 2.0 | 1.0*x |
| All | 581 | 2.2 | 5.1 | 8.3 | 9.2 | 9.9 | 12.0 | 16.9 | 22.7 | 24.1 | 22.0 | 132.5 | |

[1] Since the download numbers are very sensitive, we did not provide the absolute figures but only the relations among them.

**Table 1:** Year-wise relation of downloads per print publication year (2002-2011), document type: full-length article - FLA (n=581).

| Pub year | n | Citation year | | | | | | | | | | | Cites per cited doc |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | All | |
| 2002 | 46 | 5 | 64 | 150 | 205 | 304 | 315 | 366 | 431 | 439 | 338 | 2617 | 56.9 |
| 2003 | 72 | | 15 | 91 | 215 | 287 | 341 | 438 | 479 | 433 | 527 | 2826 | 39.3 |
| 2004 | 72 | | | 6 | 143 | 325 | 453 | 582 | 698 | 681 | 715 | 3603 | 50.0 |
| 2005 | 62 | | | | 17 | 138 | 232 | 375 | 471 | 521 | 524 | 2278 | 36.7 |
| 2006 | 77 | | | | | 26 | 151 | 302 | 461 | 522 | 523 | 1985 | 25.8 |
| 2007 | 54 | | | | | | 16 | 134 | 271 | 367 | 439 | 1227 | 22.7 |
| 2008 | 61 | | | | | | | 28 | 158 | 346 | 456 | 988 | 16.2 |
| 2009 | 50 | | | | | | | | 21 | 151 | 263 | 435 | 8.7 |
| 2010 | 33 | | | | | | | | | 18 | 107 | 125 | 3.8 |
| 2011 | 6 | | | | | | | | | | 8 | 8 | 1.3 |
| All | 533 | 3 | 44 | 84 | 165 | 223 | 287 | 355 | 405 | 472 | 496 | 16092 | |

**Table 2:** Year-wise citations per publication year (2002-2011), document types: article, review, conference paper (n=533, only cited documents).

### User analysis of Mendeley readers

Mendeley enables their users to create and maintain user profiles that include, among other information, their professional status. This makes it possible to conduct an analysis of the user structure of Mendeley "readers". As can be seen in Figure 3, more than two thirds of the readers of the analyzed journal are students (most of them PhD and master students). Professors, associate professors and assistant professors, who might have a considerably higher proportion in the Scopus publications, account for only 15% of Mendeley users. These results are in line with those found when investigating another Information Systems journal (10).

### Conclusions

In our analysis we identified a high (though not a perfect) correlation between citations and downloads which was slightly lower between downloads and readership frequencies and again between citations and readership counts. This is mainly due to the fact that the used data (sources) are related either to research or at least to teaching in higher education institutions. In the research process, papers are downloaded (for instance, from ScienceDirect) and, more or less frequently, their bibliographic data are entered into a reference management system (for instance, Mendeley). Later on, the very same papers may be cited by an article which, when accepted in a journal covered by a citation index such as Scopus, will increase their citation impact. Though being used in a similar "context", the three data sources have several differences. They concern, among others, the contents and the user population.

The Mendeley catalogue with its 110 million unique documents is the largest data source among the three. It includes articles not only from journals (also from journals not included in Scopus) but also grey literature, proceedings articles and monographs. Since an article must be entered by at least one user in Mendeley, not all of the journal articles from Scopus are necessarily covered by Mendeley. In particular, coverage varies between disciplines (11). ScienceDirect is a full-text service, providing a subset of Scopus articles (see Figure 4). All three are owned by Reed Elsevier.

Since the analyzed journal was almost fully covered by the three data sources (more than 95% of ScienceDirect's full-length articles published between 2002 and 2011 were covered by Mendeley in October 2012), one of the strongest remaining influencing factors onto the relation between citations,

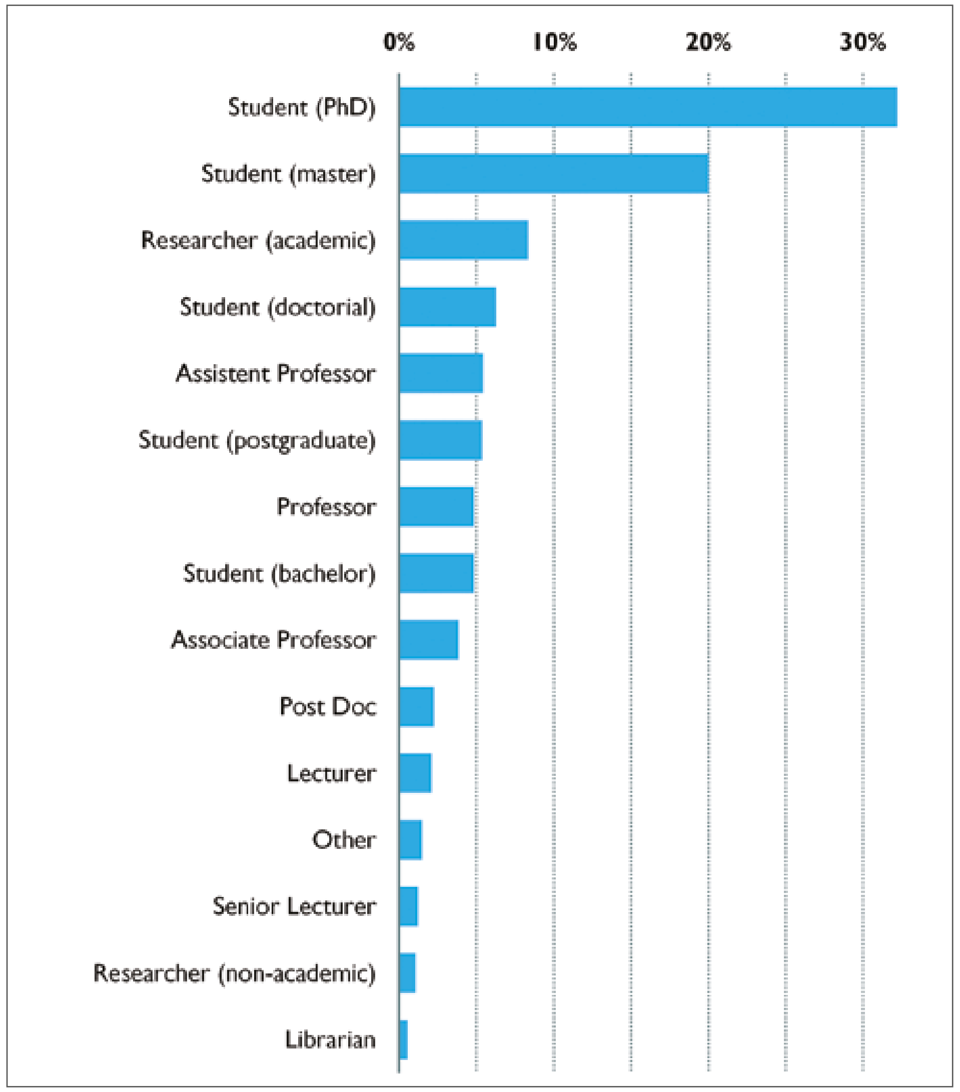



Figure 3: Readership structure of the articles in Mendeley (2002-2011) (data extraction: October 2012).

downloads and readership frequencies might be their user structure (see Figure 5).

As was reported before, two thirds of the Mendeley users are students. Contrary to bachelor and master students (approximately 25% of all Mendeley users), PhD and doctoral students are often also engaged in publication activities in particular in the Natural Sciences. Nevertheless, senior researchers might have the highest publication output in Scopus. ScienceDirect might have the broadest user base covering also users who are not actively involved in scholarly publishing (for instance, university teachers). Due to the different user structure the motives for downloading, reading and citing articles will be different too. Therefore, a perfect relation between the three indicators cannot be expected.

### Acknowledgement

This report is based in part on the analysis of anonymous ScienceDirect usage data and Scopus citation data provided by Elsevier within the framework of the Elsevier Bibliometric Research Program (EBRP). Readership data were provided by Mendeley. The authors would like to thank both Elsevier and Mendeley for their great support and the reviewers from Research Trends for their useful comments.



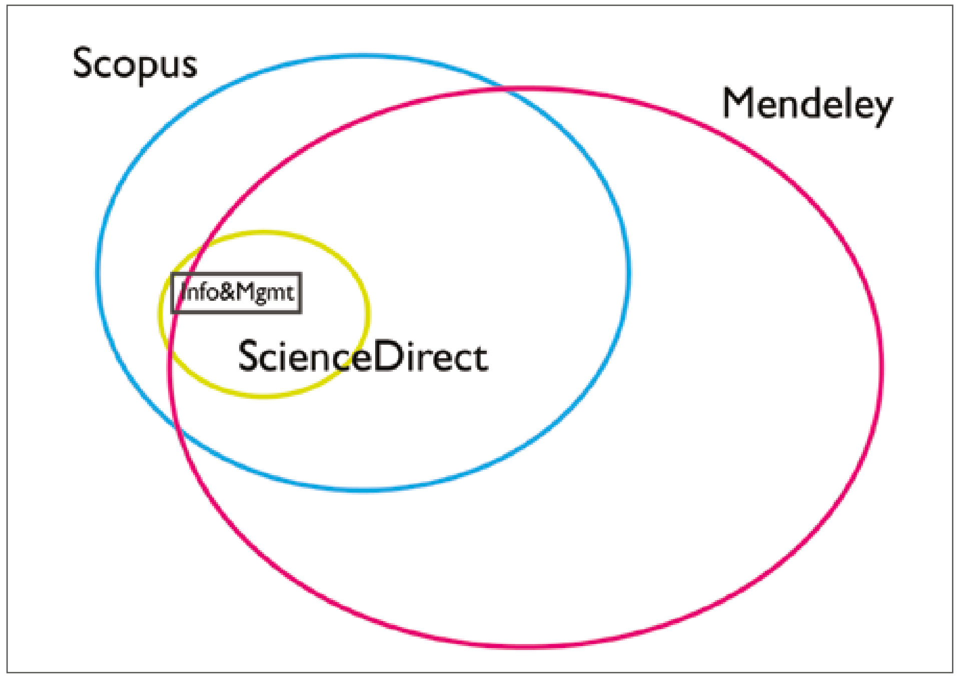



Figure 4: Coverage of ScienceDirect, Scopus and Mendeley (size of the ovals does not represent the real relations in size among the data sources; the rectangle represents the articles from the analyzed journal Information and Management).

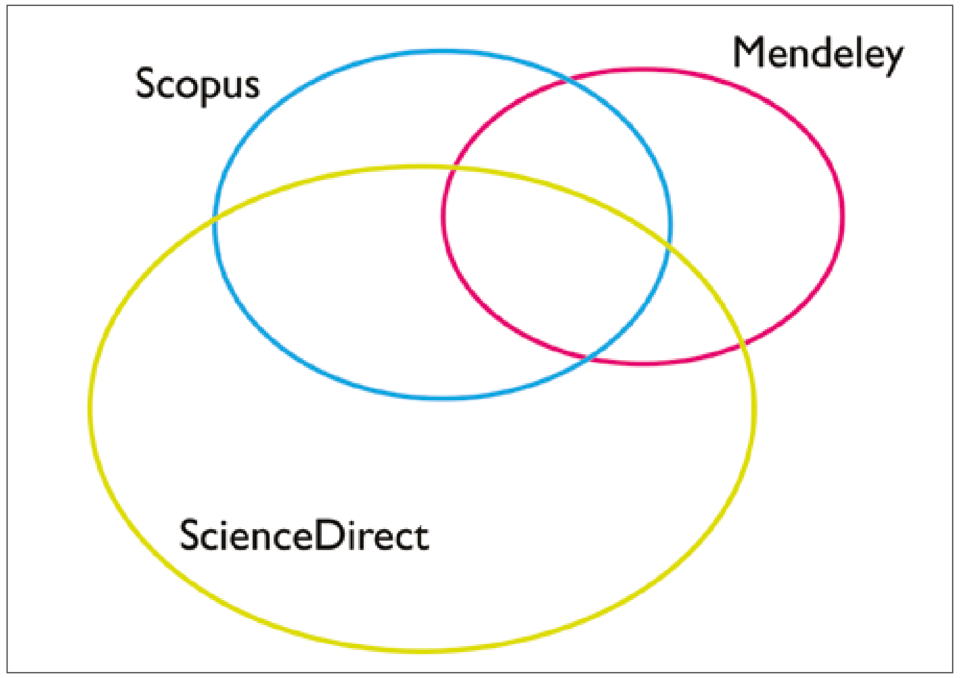



Figure 5: Size of user communities of ScienceDirect (downloading users), Scopus (publishing and citing authors) and Mendeley (readers) (size of the ovals does not represent the real relations in size among the user numbers).

### References:

1. Taylor, M. (2013) "The Challenges of Measuring Social Impact Using Altmetrics", Research Trends, Issue 33, June 2013, 11-15.

2. www.plumanalytics.com/metrics.html, accessed March 4th, 2014.

3. ebrp.elsevier.com, accessed March 4th, 2014.

4. www.sciencedirect.com, accessed March 4th, 2014.

5. www.scopus.com, accessed March 4th, 2014.

6. www.mendeley.com, accessed March 4th, 2014.

7. Scopus and ScienceDirect do not only use different record identifiers but also different document types. The Scopus document types "article", "proceedings paper" and "review" correspond mainly to the ScienceDirect document type "full length article".

8. Schlögl, C., Gorraiz, J., Gumpenberger, C., Jack, K., Kraker, P. (2013) "Download vs. citation vs. readership data: The case of an information systems journal", "Proceedings of the 14th International Society of Scientometrics and Informetrics Conference", AIT Austrian Institute of Technology GmbH, Vienna, pp. 626-634.

9. Amoako-Gyampah, K., Salam, A.F. (2004) "An extension of the technology acceptance model in an ERP implementation environment", Information and Management, Vol. 41, No. 6, pp. 731-745. The year-wise citation data were retrieved from Citation Overview Function of Scopus database on March 4th, 2014.

10. Gorraiz, J., Gumpenberger, C., Jack, K., Kraker, P, Schlögl, C. (2013) "What do citations, downloads and readership data of an information systems journal have in common and where do they differ?". In: Hinze, S., Lottmann, A. (Eds.) "Translational twists and turns: Science as a socio-economic endeavor", Proceedings of the 18th International Conference of Science and Technology Indicators (STI 2013), ENID and iFQ, Berlin, pp. 140-145.

11. Kraker, P., Körner, C., Jack, K. & Granitzer, M. (2012) "Harnessing User Library Statistics for Research Evaluation and Knowledge Domain Visualization", Proceedings of the 21st International Conference Companion on World Wide Web, ACM, Lyon, pp. 1017–1024.

Section 5:
Behind the data

# Downloads versus citations and the role of publication language

**Professor Vicente P. Guerrero-Bote[1] and Professor Félix Moya-Anegón[2]**

[1] Scimago Group, University of Extremadura, Department of Information and Communication

[2] Scimago Group, CSIC, CCHS, IPP

Since scientific literature is now published and distributed mainly online, a number of initiatives have been developed to attempt to measure scientific impact from download data. Such data would allow scientific activity to be observed immediately after publication, rather than having to wait for the citations. Shepherd (1) and Bollen et al. (2) propose a Download Impact Factor as a journal metric. It consists of the average download rates of articles published in a journal, similar to the citation-based Journal Impact Factor (JIF). COUNTER (3) define as standard a Journal Usage Factor using the median rather than the mean. Bollen et al. (2, 4) have demonstrated the feasibility of a variety of social network metrics calculated from the download networks extracted from the information contained in the clicks recorded in download logs.

Bollen et al. (5) conducted a principal component analysis of the rankings of journals produced by 39 measures of academic impact calculated from both citation and download log data. Their results indicate that the notion of scientific impact is multi-dimensional, and cannot be adequately measured by a single indicator, although some might be more suitable than others. In particular, they observed greater significance with indicators based on downloads, possibly because of the great amount of download data that can be collected.

Although Kurtz et al. (6) show how the citation obsolescence function (7) and readership follow similar trajectories over time, Schloegl & Gorraiz (8, 9) find that downloads and citations have different patterns of obsolescence. While Darmoni et al. (10) and Bollen et al. (5) report that a journal's download frequency does not to any great degree correspond with the impact factor, Schloegl & Gorraiz (9) calculate a strong correlation at the journal level between citation and download frequency when absolute values are used, and a moderate to strong correlation between the number of downloads and the journal impact factor. In this sense too, Wan et al. (11) define a download immediacy index.

### Download as predictor of citation

In recent papers (12, 13) we have used data from Scopus (citations) and ScienceDirect (downloads) to study the relationship between downloads and citations and the influence of publication language. Therefore we studied these parameters for the journals in non-English languages in ScienceDirect, specifically, for those with more than 95% of their articles in French, German, or Spanish in the period 2003-2011. We also defined a control group of journals in English in order to establish the differences with the non-English language journals. For each non-English journal, we selected as control at least one English-language journal that was present in both databases, that belonged to the same specific subject area, and had a similar number of published articles. To look deeper into the phenomenon, we compared the geographical origins of the downloads and of the citations of the two groups. It must be noted that the set of German- and Spanish-language journals is too small to draw any significant separate conclusions.



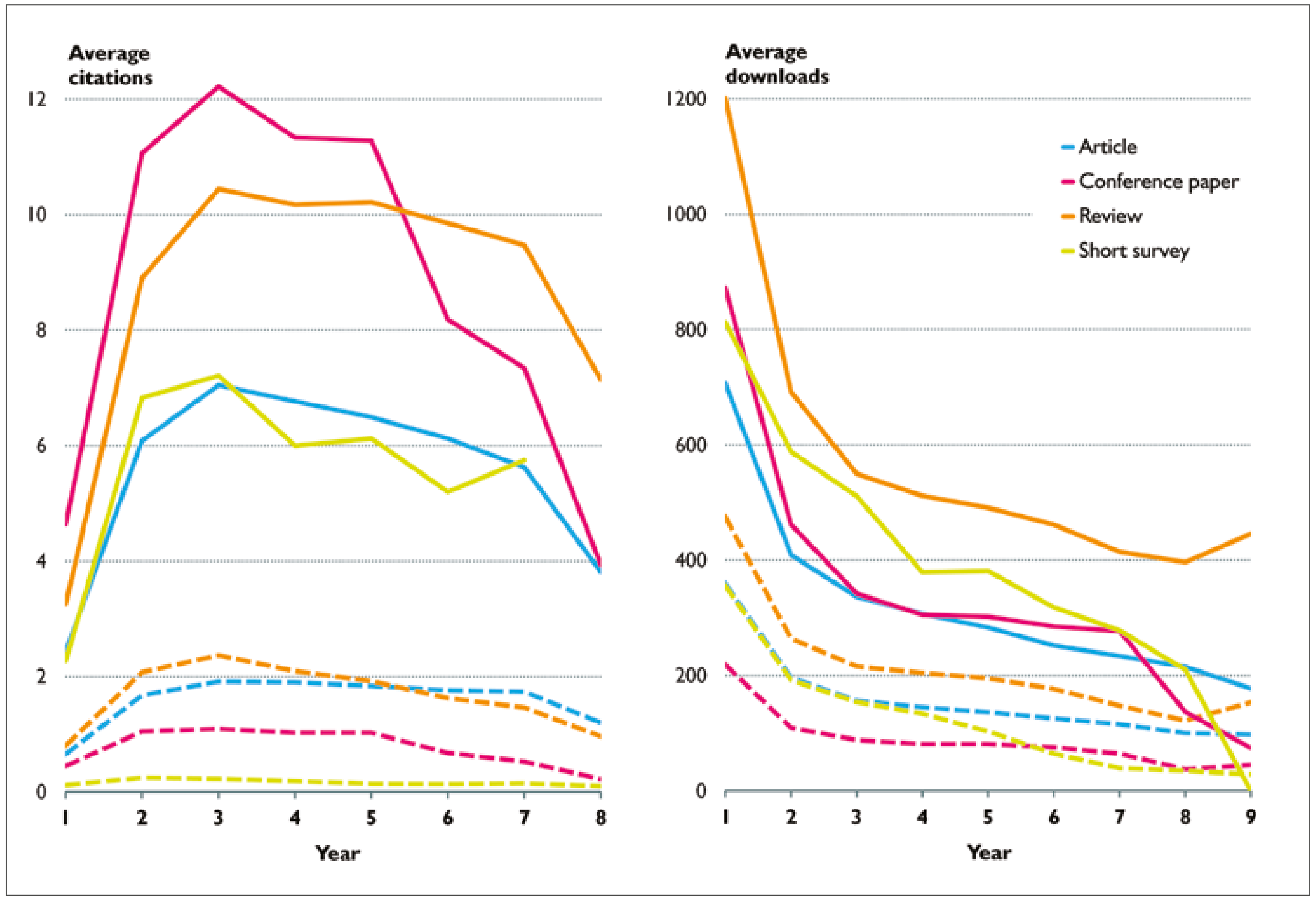



Figure 1: Left panel: Mean primary citations for Scopus document types by age of the document in years. Right panel: Mean downloads of the main document types corresponding to Scopus by age in years after the online publication date. Comparing the data for "excellent" papers (solid lines) with those for other papers (dashed lines).

Scopus and ScienceDirect cover different numbers of papers. This is because the latter includes all papers, while the former does not include conference/meeting abstracts or book reviews. The divergence between them is mainly due to the conference/meeting abstracts. The time-obsolescence curves of citations and downloads differ (see Figure 1). One appreciates the effect in the former of the time it takes for a paper to be cited, and in the latter of novelty in the downloads. The proportional difference between the downloads received by reviews and other document types increases relative to the citations.

The "excellent" papers (those belonging to the top 10% cited in the corresponding specific subject area, document type, and year) (14) showed a great difference in mean downloads with respect to the non-excellent papers throughout the period. The percentage difference was greater both at the end of the period and for the document types of medium or low download levels. The order of the subject areas in mean citation does not coincide with that in mean downloads: while Psychology was always behind Medicine in citations, it was always ahead of Medicine in downloads. This may reflect different habits in different areas, with some areas seeming to read proportionally more than they cite.

There were positive and statistically significant correlations between downloads and citations by journal and by age in years for the entire set of journals, both English and non-English (0.77 on average), but these were greatly reduced both in value and in statistical significance in the case of the non-English language journals.

In the control journals, it seems that there is a novelty effect at the beginning, with there being many downloads that do not result in citations. This may be the reason that the correlations are weakest in the first year after publication. Interestingly, the strongest correlations are found in the seventh year after publication. This may correspond to when researchers are looking for a specific paper, probably redirected by some citation.

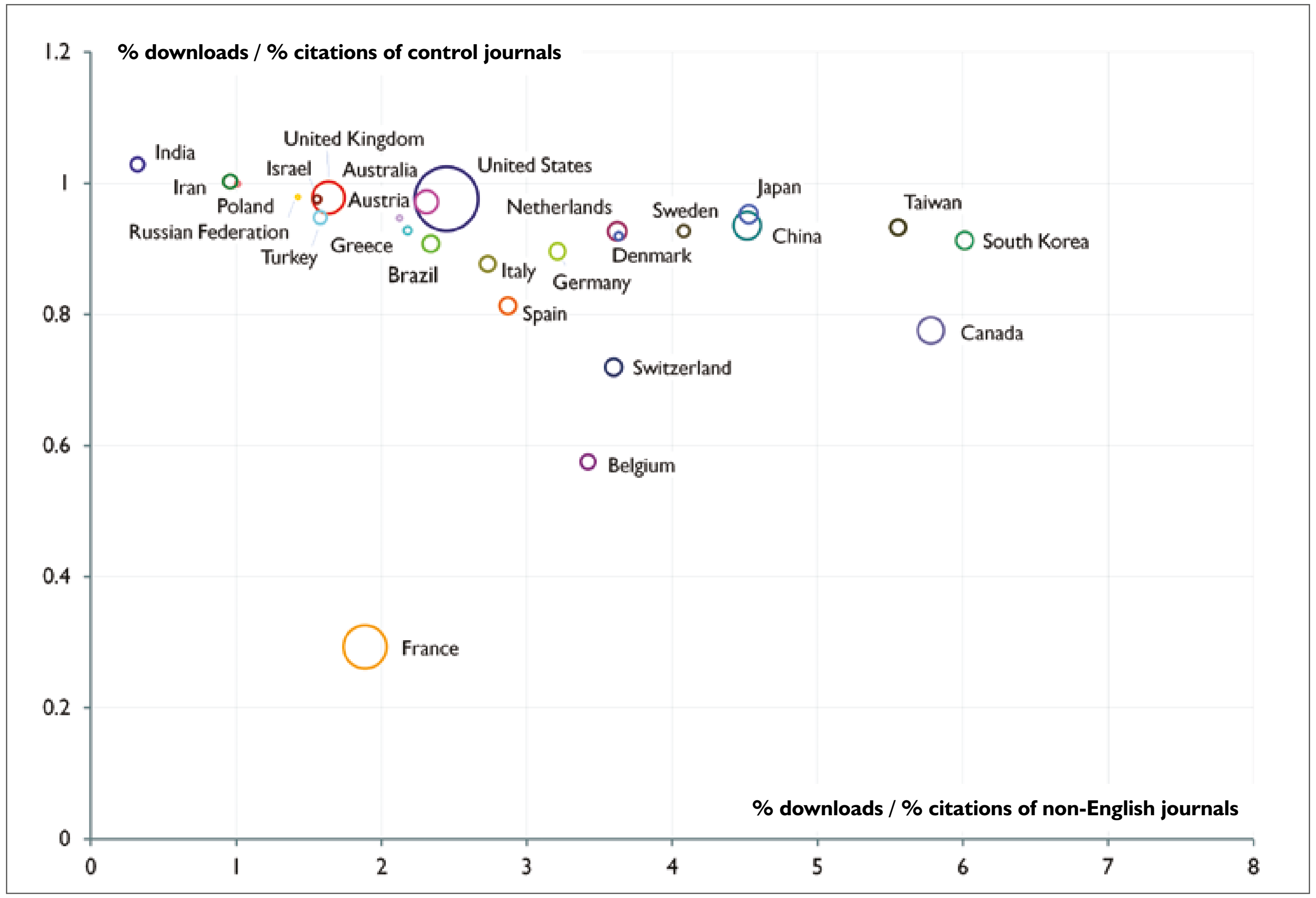



Figure 2: Plot, for the 27 greatest scientific production countries, on the vertical axis the ratio of the percentage of downloads from the control journals and the percentage of citations to these journals, against on the horizontal axis the ratio of the percentage of downloads from the French, German, and Spanish journals and the percentage of citations to these non-English journals. The area of each circle is proportional to that country's total number of downloads.

The correlations at the level of individual papers are considerably weaker (0.42 on average) than those at journal level, but markedly more significant statistically because of the far greater sample size. Nonetheless, the relative weakness of the correlation (around 55% of the correlations of the journals) may be indicative that the number of downloads, besides being a function of the quality of the paper (reflected in its citations), largely depends on the diffusion of the journal and on the effect of novelty itself. Thus, articles published in journals of wide circulation and diffusion, with high mean impact, have many downloads, even though for some papers this does not lead to many citations. Also, works published in journals of lower mean impact have fewer downloads, regardless of whether or not some of those papers later receive many citations.

All this means that the potential usefulness of download data as a predictor of citation is limited, especially so given that it is in the early years when the significance is the lowest. This circumstance was even more marked in the case of non-English language journals.

### Origin of Download/Citation and language

Figure 2 reveals that the control journals are downloaded proportionally slightly less than they are cited by the most productive countries. Instead, the non-English journals studied are downloaded proportionally more than twice as much as they are cited. This may reflect that a part of the citation impact of these non-English language journals is invisible to Scopus, because the authors who download those papers cite them in articles published in journals that are not indexed in Scopus. For example, Belgium has a percentage of downloads of control journals that is 42% less than the percentage of citation to the same journals, while having a percentage of downloads from the non-English journals which is 242% higher than the percentage citation to these journals.

In the 50 most productive countries, there is an association between the control journals' citations or downloads and a proportional increase in their downloads relative to their citations. This is to say that users who frequently download or cite the control journals download them proportionally more than they cite them. This effect is not observed for the non-English language journals studied.



In francophone regions, there is a proportionally greater decrease of downloads from control journals than of citations to those journals. In the German and Spanish language cases, the equivalent results have little significance because of the very few journals involved, some of which have been loaded into ScienceDirect retrospectively.

In sum, there seems to be a part of the citation impact of non-English language journals that is invisible to Scopus, which makes the number of downloads proportionately greater than the citations. This also has its effect on the lack of correlation between downloads and citations in these non-English journals, which means that if one wants to predict the citation rate for these titles, it will be difficult to use download data to do so.

### Acknowledgments

This work was supported by Elsevier as part of the Elsevier Bibliometric Research Program (EBRP), and financed by the Junta de Extremadura, Consejería de Empleo, Empresa e Innovación and by the Fondo Social Europeo as part of the research group grant GR10019.

### References:

1. Shepherd, P.T. (2007) "The feasibility of developing and implementing journal usage factors: a research project sponsored by UKSG", Serials: The Journal for the Serials Community, Vol. 20, No. 2, pp. 117-123.

2. Bollen, J., Van de Sompel, H. and Rodriguez, M.A. (2008) "Towards usage-based impact metrics: First results from the MESUR project". In Joint Conference on Digital Libraries (JCDL2006), Pittsburgh, PA, June 2008.

3. COUNTER (2014). "Usage Factor a COUNTER standard". Available at: http://www.projectcounter.org/documents/Draft_UF_R1.pdf.

4. Bollen, J., Van de Sompel, H., Smith, J. and Luce, R. (2005) "Toward alternative metrics of journal impact: a comparison of download and citation data", Information Processing and Management, Vol. 41, No. 6, pp. 1419-1440.

5. Bollen, J., Van de Sompel, H., Hagberg, A, and Chute, R. (2009) "A principal component analysis of 39 scientific impact measures", PLOS ONE, Vol. 4, No. 6: e6022. doi:10.1371/journal.pone.0006022.

6. Kurtz, M.J., Eichhorn, G., Accomazzi, A., Grant, C.S., Demleitner, M. & Murray, S.S. (2005) "The bibliometric properties of article readership information", Journal of the American Society for Information Science and Technology, Vol. 56, pp. 111-128.

7. Egghe, L. & Rousseau, R. (2000) "Aging, obsolescence, impact, growth, and utilization: Definitions and relations", Journal of the American Society for Information Science, Vol. 51, No. 11, pp. 1004–1017.

8. Schloegl, C. & Gorraiz, J. (2010) "Comparison of citation and usage indicators: The case of oncology journals", Scientometrics, Vol. 82, No. 3, pp. 567–580.

9. Schloegl, C. & Gorraiz, J. (2011) "Global Usage Versus Global Citation Metrics: The Case of Pharmacology Journals", Journal of the American Society for Information Science and Technology, Vol. 62, No. 1, pp. 161–170.

10. Darmoni, S.J., Roussel, F., Benichou, J., Faure, G.C., Thirion, B. & Pinhas, N. (2000) "Reading factor as a credible alternative to impact factor: a preliminary study", Technol. Health Care, Vol. 8, No. 3-4, pp. 174–175.

11. Wan, J.-K., Hua, P.-H., Rousseau, R. & Sun, X.-K. (2010) "The journal download immediacy index (DII): Experiences using a Chinese full-text database", Scientometrics, Vol. 82, No. 3, pp. 555–566.

12. Guerrero-Bote, V.P. & Moya-Anegón, F. (2013) "Relationship between Downloads and Citation and the influence of language". In: J. Gorraiz, E. Schiebel, C. Gumpenberger, M. Hörlesberger & H. Moed (Eds.), Proceedings of the 14th International Conference on Scientometrics and Informetrics – ISSI 2013 (pp. 1469–1484). Vienna: Austrian Institute of Technology.

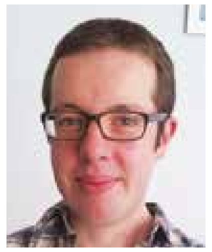
13. Guerrero-Bote, V.P. & Moya-Anegón, F. (2014) "Relationship between Downloads and Citations at Journal and Paper Levels, and the Influence of Language", Scientometrics (in press).

14. Bornmann, L., Moya-Anegón, F. & Leydesdorff, L. (2012) "The new excellence indicator in the world report of the SCImago institutions rankings 2011", Journal of Informetrics, Vol. 6, No. 2, pp. 333-335.

Section 6:
Value of bibliometrics

# The grey literature from an altmetrics perspective - opportunity and challenges

**Euan Adie**
Altmetric LLP



The field of altmetrics encompasses both alternative metrics (data beyond citation counts or impact factors) and alternative research outputs (like datasets and software).

But some material falls into both camps.

Grey literature – theses, posters, preprints, patents and policy documents and similar – are created by researchers and informed by research but aren't usually viewed as first class citizens of the scholarly literature. They are not all tracked in citation indexes like Web of Science or Scopus and can be difficult to cite in academic journals, with some editors discouraging any formal citation of preprints and similar types of document. For example, the Oxford Journals author guidelines (1) states that the reference section must not include manuscripts not formally accepted for publication, e.g. preprints. There can be good reasons for this, which we'll explore further later in the article.

The term 'grey literature' comes from their position in the fuzzy grey area between academic and popular literature (2). Importantly they are resources that aren't typically controlled by academic publishers, traditionally the gatekeepers of the scholarly record. Publishers generally take this role seriously, and there is an established technical infrastructure as well as standard processes to support them doing so. It is reasonable nowadays to expect the majority of publishers to belong to an ethics program like COPE, to assign unique and persistent identifiers like DOIs and to participate in long term archiving projects like CLOCKSS (Controlled LOCKSS). This is a not-for-profit joint venture between the academic publishers and research libraries with the ambition of developing a sustainable, geographically distributed dark archive to ensure the long-term survival of Web-based scholarly publications (www.clockss.org).

No such infrastructure or processes exist for grey literature. That is part of their appeal: you can upload a preprint or present a poster without having to go through a lengthy peer review, typesetting and publishing process, or publish a report without having to go through an intermediary. It is unfortunately also a hindrance to anybody trying to mine or analyze them. Analyzing them is exactly what altmetrics initiatives should be trying to do, because policy documents and patents are potentially very interesting indicators of impact beyond scholarly impact.

### The opportunity

It's not hard to imagine some use cases illustrating why altmetrics groups might want to get a handle on the subject:

- If my research is on the economic impact of river flooding then citations in other journals aren't the only thing that's important to me. I want to be kept aware of government policy that cites my work, too.
- If my work is referenced by a patent in a completely different field, I'd like to know about it.
- When looking at research outputs of my department, I don't just care about peer-reviewed research in journals, but patents, reports and policy documents too.

Being cited as evidence in a government policy report isn't impact in and of itself - perhaps the report will be locked in a filing cabinet and never acted on. It is still a valuable indicator, though, that's not easily obtainable anywhere else. It's not unusual for even the authors of a paper to not know about everywhere that their work has turned up.



### The challenges

Discovering what research the grey literature material cites is just one potential opportunity to enrich impact analysis, but the challenges are fairly formidable. We've spent a lot of time and effort on building up systems to track, parse and analyze policy documents and patents and some of the more interesting challenges we've faced are:

- Identifying relevant documents
- Extracting metadata & references
- Permanence

Let's look at each one in turn.

### Identifying relevant documents

The first challenge to mining grey literature is simply to find it.

It is a publisher's job (at least traditionally) to disseminate research, and there is a well-established ecosystem of discovery tools and indexing services to help individuals find and access scholarly literature that is relevant to them.

There is no such ecosystem for the grey literature, though valuable initiatives like greylit.org can give researchers a head start. Without knowing even how much grey literature material is created each year, let alone by whom, it is difficult to make assumptions about how complete any index you may build is.

### Extracting metadata & references

Once relevant documents are found, you ideally want to associate basic bibliographic metadata with them – a title, some authors, a publication date.

Central databases like CrossRef or PubMed can help do this for traditional literature, returning bibliographic records originally supplied by the publisher when queried by a unique identifier.

Policy documents, to take one example, have no such canonical metadata available and they have often been published online in ways that make automatic metadata extraction impractical. A government report may be provided only as a typeset PDF, with the title and authors (if mentioned at all) in a graphic on the first page.

For the purposes of altmetrics we are interested in the research that documents cite, and common practice in scholarly articles is to keep these to a single references section. There is often no such common practice for grey literature, where references can be in figure captions, in footnotes, tables, or separate appendices to name but a few common scenarios. Furthermore, without manual curation it is hard to figure out what's a citation at all in the traditional sense of the word: we have come across medical guidelines that explicitly list out papers that may have seemed relevant but were not used in any way.

### Permanence

A core principle shared by most altmetrics groups is that the raw data that any numbers or assertions are based on should be available to the end user.

So if we are to report that a particular policy document links to a paper then we need to make sure that users can get to that policy document.

This leads to a couple of classic online publishing problems: firstly will we always be able to find the document again in the place we found it and secondly will the document always be available online.

There is nothing to stop an NGO or government agency from redesigning its website, shifting its online publications to a different part of the site and breaking all our links. There is also no 'dark archive' of documents to ensure that we will always have a copy even if the group that originally created it ceases to exist.

### How does the grey literature fit in with other altmetrics?

One oft-mentioned advantage of altmetrics indicators is that they are usually high volume and quick to accrue, with the first data being collected within hours of publication instead of months as is usually the case with citations.

Citations to papers from policy documents buck this trend, where, anecdotally, we have seen that most of the biomedical papers referenced are five or more years older than the policy document itself – this is even slower than you might expect from traditional literature.

It is possible to imagine the attention paid to at least biomedical research on a continuum (see Table 1).

| **Within:** | Hours | Days | Months | Years |
|---|---|---|---|---|
| **Activity seen:** | Altmetrics: First mention on social media | Altmetrics: First pickups on blogs & in news outlets | Bibliometrics: First citations in the rest of the literature | Altmetrics: First appearances in policy documents |

Table 1: The attention potentially paid to research on a continuum.

Why might citations from policy documents only appear years after a paper is published? We don't know, though it would be interesting to find out. One possibility is that it takes a long time for some types of policy document or report to actually get published, so the citations are to papers that may have actually been relatively new when the authors were still discussing whatever issue the document is addressing.

### How could we improve things?

The flexibility of grey literature is a strength but also a weakness. The grey literature lacks many of the important pieces of infrastructure and best practices used by academic publishers.

Might it be possible to pull over some of the good things from academic publishing workflows, without losing too many of the benefits of occasionally being able to opt out of scholarly publishing processes?

A few key changes to the way grey literature is produced would make life much easier for anybody interested in the altmetrics that they might provide, though these must be balanced with the needs of creators who may have little interest in metrics of any kind and so lack the motivation to support change.

### Use of persistent identifiers

Use of something like the Handle System (in which resources are assigned a unique identifier that can be resolved to a URL by the creator) would help ensure that groups can track documents even if they move around the internet.

### Minimum standards of metadata

The best way to add basic metadata to scholarly PDFs and web pages is a problem that publishers solved long ago. PRISM (Publisher Requirements for Industry Standard Metadata) is a publisher driven initiative to agree on a standard set of metadata for academic publications (see idea alliance for more detail). Dublin Core is a broad set of standard metadata terms that can be applied to documents, videos, images and other resources. They provide standard ontologies; in PDFs these can be inserted using authoring tools or, after creation, using XMP which is a standard way of adding metadata to images and PDFs. On web pages the publishing industry has settled on <meta> tags, not least because for many journals this is a prerequisite for indexing by Google Scholar.

### An index of the grey literature

An open, central index of scholarly grey literature that enforced a minimum level of metadata for each item could make searching and linking documents much easier for tool makers and help the groups authoring them with discoverability (as users would have one place to look for relevant documents) and attribution.

An alternative would be to maintain a central index of grey literature repositories – the websites of each group authoring them, perhaps – and to allow harvesting from each through a standard like OAI-PMH (Open Archives Initiatives – Protocol for Metadata Harvesting), already well adopted by institutional repositories and open access publishers.

This would allow third parties to independently provide centralized tools to search or preserve content held on each group's website, making it easier to track and discover documents.

### Conclusion

The grey literature presents great opportunities for alternative metrics, providing data and indicators that cannot be found anywhere else.

Those opportunities come with great challenges, both social and technical. To work with grey literature, tools need some basic infrastructure to be put in place, but is this something that authors really want or will it compromise the advantages of publishing grey literature in the first place?

References:

1. http://www.oxfordjournals.org/our_journals/molbev/general_author_guidelines.html#References

2. Auger, C.P., Ed. (1989) Information Sources in Grey Literature (2nd ed.). London: Bowker-Saur. ISBN 0862918715.

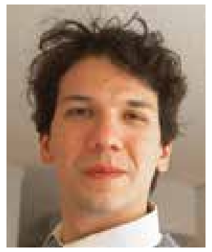


Section 7:
Dialogue

# Science without borders: are technology and policy limiting internation-alization? A conversation between Juan Pablo Alperin and Mike Taylor

**Juan Pablo Alperin**, PhD Candidate, Public Knowledge Project, Stanford University.

**Mike Taylor** is a researcher with Elsevier Labs.



**Juan Pablo Alperin** (@juancommander) is a PhD candidate in the Stanford Graduate School of Education and a researcher and systems developer with the Public Knowledge Project (PKP). Juan leads several research and development projects on improving the quality, impact, and reach of Latin American research, and is currently studying the alternative and public impact of open access (http://flacso.org.br/oa/category/proyectos/?lang=en).

**MT:** Juan, I heard you speak last year at the PLOS Article Level Metrics workshop in San Francisco. You gave a very powerful presentation on some of the problems facing researchers and journals based in the developing world. In particular, I was struck by your observation that when the developing world decides to innovate the use of things that we take for granted - for example the Impact Factor or DOIs (Digital Object Identifiers) - we effectively exclude many researchers who don't have access. In your recent blog posting (1), you state that only 4% of Latin American journals are indexed by Web of Science (WoS), and that it's argued that the excluded journals don't fall into the "mainstream" of science. To what extent do you feel that the category of mainstream is defined by access to technology?

**JPA:** I do not think that "mainstream science" is itself defined by access to technology. Scholarship is a networked process, which naturally lends itself well to a core-periphery framing. It is not my preferred characterization, but one that is arguably a reality. That is, if we were to network all the literature or form a network of all those contributing to scholarship, we may be able to identify that there is, in fact, a core which could be said to be the "mainstream".

What has been achieved through technology is to demarcate what should be considered for inclusion in that overall network; for example, if your articles are contained in an abstract and index database such as Scopus or WoS, then your work can be entered into citation analysis and therefore be considered part of the mainstream. To make matters worse for those that lack access, technologies provide a way of essentially excluding in a way that appears to be democratic and objective, but is actually far from being either.

This is not to say that technology cannot also be used for eliminating boundaries. Google Scholar is an example that offers results from small, independent journals next to those from large commercial publishers in a way that blurs the distinction between the two. It is not uncommon to find technology optimists who think that all technologies are equally unifying. The reality, however, is that access to technology can just as easily foment a false dichotomy, creating two classes of scholars (those that have access and those that do not), with the consequence that the scholarship of those in the latter group is perceived to be inherently less valuable.

**MT:** At a conference in Mexico recently, I heard a speech from Abel Packer, SciELO Brazil (2) on the threat that emerging mega-journals may have for local research journals. In short, the argument was that while these new platforms are more attractive to researchers (they provide international visibility with and access to DOIs, JIF, etc., whilst frequently being able to waive fees), the inevitable migration will lead to a decrease in the use of local journals. And that as these become less popular and less attractive to authors (particularly those writing in English), the potential loss of local journals will result in a loss of a valuable part of the academic infrastructure - for example, editorial boards, peer-review, conferences and workshops. Do you share this concern, or is the gradual death of local publishing inevitable? What do local journals have to offer that mega-journals do not?

**JPA:** Local, institutional, and student journals serve as an important learning ground for novice scholars to learn the ropes about communicating scholarship and, as you mention, they play a critical role in the research infrastructure. Their demise would be tragic: it would weaken research culture, yield more of the research agenda to those running mega-journals, and eliminate the necessary stepping stones for scholars to improve their research communication to the standards of their international peers. Given their critical importance, yes, I do worry about their decline.

However, I do not think it is inevitable or even imminent, at least not in Latin America, although there is definitely a risk. The funding model in Latin America has been very different than in the North. Currently, APCs (article processing charges) are virtually non-existent and most journals are funded through public funds (primarily funds channeled through public universities). So far, government agencies have been reluctant to shift financing from local journals to APCs, and I hope it remains this way. Unlike subscriptions or APCs, the current financial model in Latin America excludes neither reader nor writer. That said, if the APC model becomes the only model for Open Access elsewhere, it may begin to take hold within Latin America. If that happens, then international mega-journals will likely end up killing the local journals.

**MT:** I'm curious on the independence of this form of funding in Latin America - the extent to which it's subject to governmental policy or not. Generally speaking, do the funds that support journals come directly from Government, or are there intermediate bodies - research councils, or organizations similar to the UK's JISC (an independent body that is neither for-profit nor purely governmental, but which exists to support an independent academic infrastructure)?

**JPA:** We did a survey of journals some years back (3), and I know there have been other studies that corroborate the general finding, that the majority of journals in Latin America receive support from their university, most of which are themselves publicly funded. I believe a lot of it comes as in-kind support from the university (server space, technical staff, etc.). Science councils also play a big role, as they set incentive structures for researchers, write guidelines or define lists of "approved" journals, special support programs, and sometimes provide financial or technical support directly to individual journals.

**MT:** When it comes to building infrastructure, or developing a higher international profile, is there a potential advantage in more regionalism? For example, I know that there are attempts to share platforms between countries that have similar cultures, for example, Scandinavian and Baltic countries. Does a shared regional infrastructure make collaboration within the region more likely? Obviously an Ecuadorean researcher is going to be more interested in child obesity in Mexico than in (for example) Lithuania or the US, but do you feel that there is a beneficial regional level of collaboration that has yet to be explored - or should we just push for complete internationalization?

**JPA:** A shared research interest is only one reason for regionalism. Regional collaboration and a shared regional infrastructure also take advantage of similar economic models, incentive structures, levels of technical expertise, and a shared research culture. The potential is not just increased collaboration in the form of co-authorships, but also in avoiding duplicate efforts and benefiting from economies of scale.

Some great examples of this can be seen in Latin America, including the two major initiatives, SciELO and RedALyC.org. But even there, a lot more could be done. These platforms are taking advantage of economies of scale to increase visibility and are centralizing some of the technical aspects of publishing, but as of yet they still have done little to increase collaboration between scholars, build a network of copy and layout editors, share personnel, or otherwise bring together those working in the publishing process.

**MT:** Do you think it would be sensible to work towards having a regional impact factor (Latin American Impact Factor, African IF, etc.) using journal level analysis (even if not the traditional JIF formulae), or would that risk the ghettoization of developing world publishing?

**JPA:** I don't think it makes sense to create regional versions of a journal-level citation metric. I think the critiques of the IF, including that of those that have backed DORA (Declaration on Research Assessment), would still equally apply to each of these instances. Moreover, they would create the same problems I have been describing, but in the reverse: they would exclude research published outside of the region and therefore penalize researchers who are publishing locally, but are being read and cited from outside the region.

SciELO provides citation counts and an IF for journals contained within SciELO (4), but I do not think the metric has been widely used, and it certainly has not supplanted the view that Thomson-Reuters' IF is the one that "matters".

The purpose of regional portals has to be to improve quality, gain efficiencies, and increase visibility, not to isolate the regions into systems that are completely decoupled from the rest of the world.



**MT:** Much of the work in altmetrics falls into two categories at the moment: finding patterns between different social networks (for example Twitter and Mendeley), and looking for the relationship between altmetrics and citation. Needless to say, this work focusses on looking for DOIs and the resolving URLs, and this will obviously exclude any article without a DOI. Furthermore, Impactstory.org has recently adopted Altmetric.com's Twitter feed, and this has had the effect of removing the ability to look for tweets linking to a non-DOIed article's URL. What can we - as researchers interested in altmetrics - do to extend the focus of our research to the developing world? To what extent do we need to look at regional variations in platforms (for example, we know that some cultures use Facebook in a more scholarly way than Twitter, and that some countries - most notably China - have a strong cultural or politically mandated preference for their own platforms, e.g. Weibo)? Would the development of local language versions of research tools or a movement towards a community-driven identification of local language blogging and review sites be positive in extending the focus of altmetrics to the developing world?

**JPA:** As you mention, the dependence on DOIs is by far the most limiting aspect for studying altmetrics in developing regions. Despite CrossRef's efforts (and to be fair, I do believe they are making a concerted effort), DOIs are still not commonplace everywhere. For many journals, even in medium income countries, the US$1.00 per article fee remains prohibitive. As long as this is the case, and as long as altmetric tools rely on DOIs, it will be impossible to evaluate altmetrics on a large scale for journals running on low budgets.

As I mentioned in my talk at ALM 13 (5), there is a strong parallel between the use of WoS for evaluation and the use of altmetrics dependent on DOIs. If only tweets to articles with DOIs can be studied, then scholars publishing in venues without DOIs will be once again discounted. An altmetric provider that works for arbitrary URLs is therefore absolutely necessary (funding agencies, tool builders, and altmetric providers: take note!).

Second, we need studies that look at altmetrics, even in the two ways you describe above, for a set of journals from developing regions, even if we start with those that do have DOIs. The existing studies have almost exclusively focused on well-resourced journals from the global North. It is possible, and even likely, that the patterns are different a) for journals with lower visibility; and b) where the use of social Web tools is different (as you allude to above). The focus on journals from publishers like Nature and PLOS sets expectations and guides the research agenda on altmetrics.

With such studies, we would at least know the levels of penetration in the currently studied platforms, and to what extent they differ between journals. I think you are right that consultations with scholars from other parts of the world may turn up other sources that are useful for other communities.

I should mention that these issues are important enough to me that they are the focus of my dissertation work. With the help of SciELO, RedALyC, and Altmetric.com, I am studying download, citation, and altmetrics data for Latin American journals. Euan Adie from Altmetric.com has been kind enough to provide special handling for a set of URLs, so that it is possible to have altmetrics on those, even if they do not have DOIs. I will be releasing some preliminary results soon (stay tuned to my Twitter feed, @juancommander). I hope to reveal some of the ways in which altmetrics vary between contexts, and open new lines of research into these alternative metrics.

**MT:** How can international organizations – whether not-for-profits, like CrossRef, Orcid, PLOS, or commercial companies such as Thomson-Reuters, Elsevier or Altmetric.com – work with the developing world so they can increase their visibility and access to global infrastructure, while permitting their regional and national characteristics to thrive?

**JPA:** Those aiming to improve scholarly communications, including those international organizations you mention, must remember that access to the scholarly communication infrastructure is often not a technological limitation. Much of the time, it is other factors, such as an editorial decision on part of Thomson-Reuters and Elsevier that prevents a journal from being indexed, or a lack of finances that limits the use of DOIs. Giving access to the existing infrastructure is a first step, but it is not enough. The next step, if we take our global/international commitment seriously, is to be willing to make changes to that infrastructure: a) by being as acutely aware as possible of the ways in which scholars from developing regions are disadvantaged by the existing models and tools; and b) by consulting and actively engaging with scholar communities in developing regions.

**MT:** Juan, thanks for taking the time to answer my questions. Perhaps you'd be kind enough to write a piece on some of your findings for a later issue of RT!

**JPA:** Thank you for your interest, and thank you for posing questions that gave me the opportunity to talk about issues that are important to me.

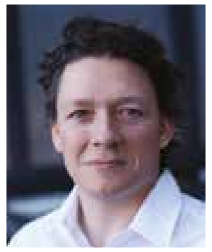
**References:**

1. http://blogs.lse.ac.uk/impactofsocialsciences/2014/03/10/altmetrics-for-developing-regions
2. http://blog.scielo.org/en/2013/07/15/38/#.UzXnUV4tXVI
3. http://www.itidjournal.org/index.php/itid/article/download/639/274
4. http://www.scielo.org/applications/scielo-org/php/citations.php
5. https://speakerdeck.com/jalperin/altmetrics-propagating-global-inequality

Section 8:
Scholarly communication

# Gauging openness, measuring impact

**William Gunn**
Head of Academic Outreach,
Mendeley Ltd



This article examines the linked concepts of openness and usability as applied to scholarly works. Openness is used to mean many different things, from transparency about influence when used in a political context, to the lack of restrictions on use when used in a software context. In the scholarly domain, openness generally refers to unrestricted, free availability of a research product over the internet. A work is considered open if there are no permission or price barriers between the work and an individual seeking to make use of the work. However, there are different levels of openness, which are defined by the types of reuse permitted.

Sir Tim Berners Lee introduced the concept of 5 star open data back in 2006 to describe the continuum from a table rendered as a PDF through to data marked up as RDF and connected to the web of Linked Open Data (1). This system clearly explained the benefits of open data by demonstrating how more value was added at each successive step of openness.

A similar scenario is presented with scholarly works. The more open it is, the more useful it is to the author and the audience (2). The first level is simply availability online, as opposed to only as a printed copy. The next level is free to read - you can read the paper without any subscription barriers. A work which is explicitly openly licensed is even more open, but the variety of open licenses leaves many works encumbered with provisions that make it impractical to reuse other than on an item-by-item level (3). Using a license without those provisions would be a further level up. This is the level of the accepted standard license for open access works, CC-BY (4). With CC-BY, there are no explicit barriers to reuse, up to the point that simply tracking and attributing all the providers themselves becomes an unmanageable task. The final level would be fully open with no restrictions of any kind, as with CC0. Each of these levels raises the ceiling value for the amount of reuse possible, while making no statement about the desirability of the work, or the sustainability of the access. Simply put, a work that's more open has, in theory, higher usability than one that is less open. If an open work is also useful to a sufficient number of people, sustainability of access is generally easier to maintain than for closed works through the LOCKSS principle (5), because at least one copy will exist for each researcher who finds it useful. In contrast, closed works can fall into "orphan" status, where reproduction is desired but not permitted, because the rights holder can no longer be identified. Openness is particularly important for works where a long incubation time may be required before the work finds its full potential. Indeed, many great historical works would have been lost were it not for the diligent copying and recopying by centuries of scribes.

★ make your stuff available on the web (whatever format)

★★ make it available as structured data (e.g. excel instead of image scan of a table)

★★★ non-proprietary format (e.g. csv instead of excel)

★★★★ use URLs to identify things, so that people can point at your stuff

★★★★★ link your data to other people's data to provide context

**Figure 1:** Tim Berners Lee's 5 star open data scale.

### What kinds of reuse exist?

The ways in which research can be reused can be divided into five general categories based on application: inspiring new research, mining existing data for novel associations, application or implementation, contribution to the popular understanding, and meta-analysis. The various types of reuse and how these can be tracked for discovery and assessment, briefly discussed below, will be the subject of a forthcoming NISO whitepaper.

The first kind of reuse, inspiring new research, is well covered by the traditional databases which track citations, but is limited



in that a subsequent piece of research points to a prior piece, but the prior piece does not reciprocally point back to the subsequent research it inspired. This type of reuse is inhibited through lack of access to the research. Additionally, the pointer is at the document level, which gives poor resolution of the details of the reuse. Another needed improvement for understanding citation behavior is to enrich a citation by adding distinguishing characteristics that would allow the different types of citations to be distinguished from one another. See the Citation Typing Ontology (CiTO) for the current work in this area (6).

Tracking mining of datasets, the second category of reuse, is often done via tracking the papers which describe them (7). However, more datasets are appearing on sites such as Figshare and Dryad, which assign DOIs (Digital Object Identifiers) to the data directly (8), instead of just a paper describing the data. Creating URIs (Uniform Research Identifiers) which point to the data directly promotes the data to equal standing with a research paper, because the data can now be referenced directly and can accrue reuse separately from the paper. As with citation of papers, access to data is a barrier to reuse, but technical skills and equipment to handle the data are also needed.

When you move out of the scholarly realm and into applications, there are less explicit mentions of the original works themselves. Detection of a reuse event in a commercial application can be done via looking for references in patent applications or publications arising from academic/industry collaborations, but this only shows first-order impact at best. As you move further away from the publication into the inventions or policies that it may have enabled or informed, the trail gets very difficult to follow, even as the raw number of possible reuse events grows. This is where individual efforts such as the implementation of a Becker Model analysis (9) become necessary, though this is prohibitive to do at scale.

Looking at the reuse of a scholarly work by the public is done much as with an application or implementation. The main source of reuse events in this category are mentions in popular media, although there is a significant "long tail" of lay communities online which discuss research: patient communities, space aficionados, citizen scientists, and teachers in non-professorial roles. Interestingly, PubMed Central reports that the majority of the page views to research papers hosted there come from non-institutional domains (10). Another notable feature of reuse within the public domain is that the direction of flow is reversed: external events such as natural disasters, celebrity endorsements, or other news events often drive increased public reuse events (11, 12), whereas availability of a technology facilitates the application.

Meta-analysis is its own category of reuse. There is a growing movement to conduct and publish replication studies of existing work, such as the Reproducibility Initiative and the Reproducibility Project: Cancer Biology, a partnership between the Reproducibility Initiative and the Center for Open Science. The aims of these projects are to understand and promote replication of research as a type of reuse. The replication studies contain pointers to the original research and explicitly identify which experiments were carried out and what the results were. This enables the creation of a separate discovery layer, to highlight and identify the more reproducible or the most reusable work, facilitating downstream commercial application or reduction to practice.

### Bootstrapping discovery of reuse

Open Access and Open Data have now become funder priorities across the world. Because funding agencies such as the NIH and Wellcome are now paying for openness in order to maximize the reuse potential of their funded outputs, it has become important be able to aggregate reuse events and to understand their relative impacts. Detecting a reuse event is challenging with current technology, primarily because reuse events don't always point back to the original item. To serve these needs, the Association for Research Libraries, with funding from Sloan and the Institute for Museum and Library Services, is building the Shared Access Research Ecosystem, an event aggregator, which will consume data sources which report on research events. Additionally, the scholarly metadata organization CrossRef is working on a service called Prospect, which aims to facilitate text and data mining of proprietary content (i.e. the data is open at the one star level, but efforts are made to make it as usable as possible). Together with technologies such as Mendeley and Impact Story, we are developing an ever clearer understanding of the importance and value of openness to the research world and society at large.

**References:**

1. 5 star Open Data. Available at: http://5stardata.info/

2. Piwowar, H.A., Day, R.B.S. & Fridsma, D.S.B. (2007) "Sharing detailed research data is associated with increased citation rate", PLOS ONE, Vol. 2, No. 3, e308.

3. Dryad (2011) "Why does Dryad use CC0?", Dryad news and views on WordPress.com. Available at: http://blog.datadryad.org/2011/10/05/why-does-dryad-use-cc0/

4. Open Access Scholarly Publishers Association (2012) "Why CC-BY?". Available at: http://oaspa.org/why-cc-by

5. LOCKSS, "Preservation Principles". Available at: http://www.lockss.org/about/principles

6. S. (2013), "CiTO, the Citation Typing Ontology". Available at: http://www.essepuntato.it/lode/http://purl.org/spar/cito

7. Piwowar, H.A. (2010) PhD Thesis: "Foundational studies for measuring the impact, prevalence, and patterns of publicly sharing biomedical research data", D-Scholarship@Pitt (Database). Available at: http://etd.library.pitt.edu/ETD/available/etd-04152010-115953/

8. Piwowar, H.A. & Vision, T.J. (2013) Data from: "Data reuse and the open data citation advantage". Available at: http://doi.org/10.5061/dryad.781pv

9. Holmes, K.L. & Sarli, C.C. (2013) "The Becker Medical Library Model for assessment of research impact – an Interview with Cathy C. Sarli and Kristi L. Holmes", Research Trends, Issue 34. Available at: http://www.researchtrends.com/issue-34-september-2013/the-becker-medical-library-model/

10. Plutchak, T.S. (2005) "The impact of open access", J. Med. Libr. Assoc., Vol. 93, No. 4, pp. 419-421.

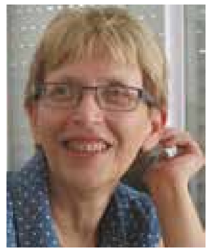
11. Bauer, M.W., Allum, N. & Miller, S. (2007) "What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda", Public Underst. Sci., Vol. 16, No. 1, pp. 79–95.

12. Monzen, S. et al. (2011) "Individual radiation exposure dose due to support activities at safe shelters in Fukushima Prefecture", PLOS ONE, Vol. 6, e27761.

Section 9:
Research assessment

# Evaluating the individual researcher – adding an altmetric perspective

**Professor Judit Bar-Ilan**
Department of Information Science, Bar-Ilan University, Israel



ACUMEN was an EU funded research project aimed at "understanding the ways in which researchers are evaluated by their peers and by institutions, and at assessing how the science system can be improved and enhanced" (1). This project was formed to answer an FP7 call that requested "studying and proposing alternative and broader ways of measuring the productivity and performance of individual researchers including new and improved bibliometric indicators and evaluation criteria for research careers, project evaluations, and scientific publications" (2). FP7 was the Seventh Framework Program of the European Union for the funding of research and technological development in Europe. The ACUMEN Consortium was comprised of nine institutions. The main outputs of the project are the ACUMEN portfolio and the Guidelines for Good Evaluation Practices (both available from http://research-acumen.eu/portfolio).

In the following article we will provide a brief introduction to the portfolio concept and then concentrate on how altmetrics are utilized in the portfolio.

### The ACUMEN Portfolio

The ACUMEN portfolio allows the researcher to present herself through a brief narrative in which she highlights her past achievements and future goals. This narrative is backed up by structured information available in the sub-portfolios: the expertise, the output and the influence sub-portfolios. For each factor in the sub-portfolios evidence is provided to support the claims. For example, if the researcher claims to have specific methodological expertise, he backs up this claim by providing references to works where this method was applied.

A more detailed description of the three sub-portfolios:

- In the expertise sub-portfolio there are factors for scientific/scholarly expertise, technological expertise, teaching expertise, knowledge transfer, communication skills and organizational expertise.
- The output sub-portfolio is comprised of factors for scholarly outputs, teaching outputs, outputs communicated to the general public including online presence and online contributions, datasets, software and tools created by the researcher, patents and grants received.
- The influence sub-portfolio provides information on citations and various citation-based indicators, scholarly prizes, prizes for teaching, membership in program committees and editorial boards, invited talks, advice given based on subject expertise, economic influence in terms of income, spin-offs, consultancies and patents, textbook sales, download counts of publications and datasets, followers on various social media platforms, Mendeley readership counts, tweets and blog posts about the researcher's work, views of online presentations, online syllabi mentions and popular articles written about the portfolio owner.

Thus the portfolio provides a holistic view of the researcher's achievements, expertise and influence. Most of the factors have detailed sub-factors, and information that the portfolio owner is interested in conveying that does not match any of the above-mentioned factors can be provided in the "other" factor of each of the sub-portfolios. Since time spent in academia is crucial for fair evaluations, the ACUMEN project introduced the "academic age", which is the time from the date the PhD was awarded with allowances for having children, illness and part-time work.

As said above, for each factor/sub-factor evidence is provided to back up the claims. The evidence is not everything that can possibly be listed, but only the "best" evidence for each factor and not more than three items. "Best" is subjectively decided by the researcher creating the portfolio. "Best" is for the specific factor; for example, in the output sub-portfolio the portfolio owner is requested to list his top three journal papers and in the influence sub-portfolio his top three most cited papers. It is possible that a different set of papers is provided for the two factors, in case he considers one of his recent works which has not accrued citations yet to be among his best works, or if he considers one of his less cited works to be among his best contributions.

### Altmetrics in the Portfolio

As can be seen from the description of the sub-portfolios, online and social media presence and altmetrics are well represented. In the portfolio, online presence is viewed as an output; the researcher is asked to list accounts in social media used for academic purposes, academic network accounts, digital repository accounts and websites that were created or used for dissemination. These include academic social media sites such as ResearchGate and Academia.edu, sites where research outputs can be published such as SlideShare, figshare, YouTube or Vimeo, and also blogs and Twitter accounts. She is also asked to indicate her activity level (e.g. average number of posts per year or month) on these sites.



Altmetrics are even more emphasized in the influence sub-portfolio. The researcher is asked for the number of followers on social media sites, where scholarly information is published or discussed. Examples of such sites are academia.edu, ResearchGate, Twitter and blog(s) maintained by the portfolio owner. The guidelines for filling in the portfolio explain that these numbers should only be provided if viewed substantial.

The researcher is asked to provide details of a maximum of three articles that were tweeted or reviewed in blogs. It was shown recently (3) that articles that are reviewed in science blog posts close to their publication date have a good chance of being cited within three years, and receive more citations than the median number of citations for articles published in the given journal and the given year that were not reviewed in science blogs. Significant associations were also found between higher number of tweets and blog mentions and higher number of citations (4).

For the portfolio the researcher is requested to list download counts for a maximum of three publications. Some publishers provide this information, and download counts are also available for example from academia.edu and ResearchGate. The ACUMEN team is aware that influence cannot be measured through publications only; therefore download counts of the top three most downloaded datasets and software are also requested.

Mendeley readership counts are currently viewed as the most promising altmetric indicator (5). Mendeley has impressive coverage, for example 93% and 94% of the articles published in 2007 in Science and Nature respectively are on Mendeley (6). Similarly, extremely high coverage (97%) was found for articles published in JASIST (Journal of the American Society for Information Science and Technology) between 2001 and 2011 (7). In (5) the coverage of Mendeley for 20,000 random publications was only 63%, but still Mendeley had by far the greatest coverage of all currently studied altmetric sources. In the ACUMEN portfolio, the user is requested to report the number of readers of up to three publications. Mendeley readership counts can possibly be useful in the Social Sciences and the Humanities, where the coverage of the citation databases (WOS and Scopus, but also Google Scholar to a smaller extent) is far from perfect. Mendeley readership counts may also reflect influence in other areas, especially for newly published items that have not received a large number of citations yet, because it takes much longer to cite an item than to be a "reader" of the item. On the other hand, it should be taken into account that it may take some time for a research result to prove its significance; receiving attention in a very early stage does not necessarily mean that the impact is stable over longer time periods. In addition, populations that do not publish in the scholarly system (e.g. students) may also be interested and influenced by scholarly work without being authors (and citers). Mendeley readership counts capture the influence of scholarly work on non-publishing, interested individuals as well. This is supported by correlations of around 0.5 in several works between readership counts and citations – indicating that Mendeley readership counts reflect impact that is different from the impact reflected by citation counts (8). It was shown (9) that PhD students, postgraduates and postdocs are the main readers of articles in Mendeley.

Educational impact can also be measured by altmetrics. Many universities have YouTube channels where they upload videos of lectures (e.g. the YaleCourses YouTube channel). Conferences also often upload videos of talks to the Web, and presentations can uploaded to Slideshare. Interest in the materials available on these sites can be measured by the number of downloads and/or the number of views. Finally, if works of the portfolio owner are referenced in online syllabi this indicates educational impact of her work. Download counts and views of the "top" items in these categories are reported in the portfolio. In addition, the researcher is encouraged to provide details of three interesting web mentions of her, or of her work not mentioned elsewhere. Thus the altmetric data appearing in the portfolio supplement information on the scientific impact and also reflect on the societal impact of the researcher and his work.

### Discussion and conclusion

Altmetrics is an emerging subfield of informetrics. Currently there are no clear guidelines on how to interpret the altmetric data in the portfolio. This is problematic both for the person filling in the portfolio and for the evaluator receiving portfolios. The best advice ACUMEN can provide at the moment is to compare with other researchers in the same field and at the same career stage. Traditional bibliometrics rely mainly on citations, whereas there are a multitude of altmetric sources. This further complicates interpretation, since we do not know how to (and probably cannot and should not) compare between tweets, downloads, blog mentions and readership counts. We are also aware that some of the altmetric indicators can be manipulated quite easily.

The aim of the ACUMEN Portfolio is to provide a holistic picture of the researcher's achievements and capabilities. To achieve this aim it is necessary to include as many facets of the achievements as possible. The ACUMEN team believes that altmetric data complement traditional bibliometric data; they indicate influence not necessarily captured by citations, and thus provide additional value.

The ACUMEN portfolio can also be used for self-assessment. The portfolio template is available here (10), and the readers are most welcome to create their own portfolio. But beware: preparing the portfolio is quite time consuming. Have fun!

### References:

1. ACUMEN (2014). Retrieved from http://research-acumen.eu/

2. European Commission (2009). Work Programme 2010. Capacities. Science in Society.

3. 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, Vol. 65, No. 5, pp.1018-1027.

4. Thelwall, M., Haustein, S., Larivière, V. & Sugimoto, C. (2013) "Do altmetrics work? Twitter and ten other candidates", PLOS ONE, Vol. 8, No. (5), e64841. doi:10.1371/journal.pone.0064841.

5. Zahedi, Z., Costas, R. & Wouters, P. (in press) "How well developed are altmetrics? Cross disciplinary analysis of the presence of 'alternative metrics' in scientific publications", Scientometrics, DOI:10.1007/s11192-014-1264-0.

6. Li, X., Thelwall, M. & Giustini, D. (2012) "Validating online reference managers for scholarly impact measurement", Scientometrics, Vol. 91, No. 2, pp. 461-471.

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

8. Bar-Ilan, J., Shema, H. & Thelwall, M. (2014) "Bibliographic References in Web 2.0". In: Blaise Cronin and Cassidy R. Sugimoto (Eds.) Beyond Bibliometrics – Harnessing Multidimensional Indicators of Scholarly Impact (pp. 307-325). Cambridge, MA: MIT Press.

9. Mohammadi, E., Thelwall, M., Haustein, S. & Larivière, V. (in press) "Who reads research articles? An altmetrics analysis of Mendeley user categories", Journal of the Association for Information Science and Technology.

10. http://research-acumen.eu/wp-content/uploads/Blank_AcumenPortfolio.v13x.pdf

Section 10::

# Did you know… that interest in altmetrics is growing fast?

**Dr. Daphne van Weijen**

Although altmetrics is still a new discipline, it is attracting more and more interest, both online and in the broader scholarly community. For example, a comparison of the results of a Google search for ‘altmetrics’ versus ‘bibliometrics’ revealed that the number of searches for both terms, relative to the total number of Google searches over time, is now more or less the same, while two years ago, the interest in ‘altmetrics’ in Google search terms was virtually non-existent. See Figure 1 to see the Google Trends graph containing the results, with altmetrics in red and bibliometrics in blue.

In terms of scholarly interest, a Scopus.com search (on altmetric* or alt-metric*) revealed that the number of scholarly papers on altmetrics is also increasing at an impressive rate. In 2011 there were just 2 papers in Scopus, followed by 12 in 2012, 32 in 2013 and 8 to date in 2014. Surprisingly those papers weren’t only published in traditional journals related to bibliometrics, but also in fields such as Plant Sciences (2), Chemistry (3), and Neuroscience (4).

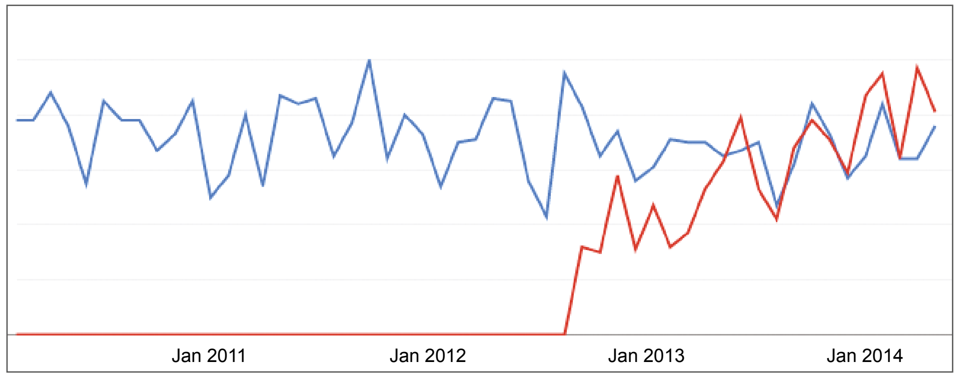



**Figure 1:** Figure showing Google trends for ‘bibliometrics’ (blue) versus ‘altmetrics’ (red).

**References:**

1. Huggett, S. & Taylor, M. (2014) Elsevier expands metrics perspectives with launch of new altmetrics pilots, Editors’ Update, Issue 42 - March 2014.

2. Osterrieder, A. (2013), The value and use of social media as communication tool in the plant sciences, Plant Methods, 9:26 doi:10.1186/1746-4811-9-26.

3. Oisaki, K. (2014), How “blog” is changing researches in synthetic chemistry? Yuki Gosei Kagaku Kyokaishi/Journal of Synthetic Organic Chemistry, Volume 72, Issue 1, 2014, Pages 74-78.

4. Garnett, A.a, Lee, G.b, Illes, J.B. (2013), Publication trends in neuroimaging of minimally conscious states, PeerJ, Volume 2013, Issue 1, 2013, Article number e155.



# Research Trends: Editorial Board



You can find more information on www.researchtrends.com or contact us at researchtrends@elsevier.com