Grand challenges in altmetrics: heterogeneity, data quality and dependencies

Stefanie Haustein[1]

Abstract   With increasing uptake among researchers, social media are finding their way into scholarly communication and, under the umbrella term altmetrics, are starting to be utilized in research evaluation. Fueled by technological possibilities and an increasing demand to demonstrate impact beyond the scientific community, altmetrics have received great attention as potential democratizers of the scientific reward system and indicators of societal impact. This paper focuses on the current challenges for altmetrics. Heterogeneity, data quality and particular dependencies are identified as the three major issues and discussed in detail with an emphasis on past developments in bibliometrics. The heterogeneity of altmetrics reflects the diversity of the acts and online events, most of which take place on social media platforms. This heterogeneity has made it difficult to establish a common definition or conceptual framework. Data quality issues become apparent in the lack of accuracy, consistency and replicability of various altmetrics, which is largely affected by the dynamic nature of social media events. Furthermore altmetrics are shaped by technical possibilities and are particularly dependent on the availability of APIs and DOIs, strongly dependent on data providers and aggregators, and potentially influenced by the technical affordances of underlying platforms.

Introduction

Social media have profoundly changed how people communicate. They are now finding their way into scholarly communication, as researchers increasingly use them to raise their visibility, connect with others and diffuse their work (Rowlands et al. 2011; Van Noorden

✉ Stefanie Haustein
   stefanie.haustein@umontreal.ca

[1]   EBSI, Université de Montréal, CP 6128, Station Centre-Ville, Montreal, QC H3C 3J7, Canada

2014). Scholarly communication itself has remained relatively stable; in the course of its 350-year history the scientific journal has not altered much. Even in the digital age, which has facilitated collaboration and increased the speed of publishing, the electronic journal article remains essentially identical to its print counterpart. Today, the peer-reviewed scientific journal is still the most important channel to diffuse scientific knowledge.

In the context of the diversification of the scholarly communication process brought about by the digital era, social media is believed to increase transparency: ideas and results can be openly discussed and scrutinized in blog posts, some journals and designated platforms are making the peer-review process visible, data and software code are increasingly published online and reused, and manuscripts and presentations are being shared on social media. This diversification of the scholarly communication process presents both an opportunity and a challenge to the scholarly community. On the one hand, researchers are able to distribute various types of scholarly work and reach larger audiences; on the other, this can lead to further information overload. At first, altmetrics were seen as an improved filter to overcome the information overload stemming from the diversification and increase in scholarly outputs (Priem et al. 2010). In that sense, quite a few parallels exist between the development of bibliometrics and altmetrics:

> It is too much to expect a research worker to spend an inordinate amount of time searching for the bibliographic descendants of antecedent papers. It would not be excessive to demand that the thorough scholar check all papers that have cited or criticized such papers, if they could be located quickly. The citation index makes this check practicable (Garfield 1955, p. 108).

> 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 (Priem et al. 2010, para. 1).

While altmetrics rely on the users of various social media platforms to identify relevant publications, datasets and findings, Garfield (1955), and before him Gross and Gross (1927), believed that citing authors would outperform professional indexers in identifying the most relevant journals, papers and ideas. Both altmetrics and citation indexing thus rely on collective intelligence, or the wisdom of the crowds, to identify the most relevant scholarly works. Citation indexing can, in fact, be described as an early, pre-social web version of *crowdsourcing*. Quite similarly to the reoccurring question about the meaning of altmetrics, early citation analysts admitted that they did "not yet have any clear idea about what exactly [they were] measuring" (Gilbert 1977, p. 114) and the interpretation of citations as indicators of impact remains disputed from a social constructivist perspective on the act of citing. However, the fundamental difference between indicators based on citations and those based on social media activity is that the act of citing has been an essential part of scholarly communication in modern science, whereas researchers are still exploring how to use social media.

Both bibliometrics and altmetrics share the same fate of being (too) quickly identified and used as indicators of impact and scientific performance. Although early applications of bibliometric indicators in research management emphasized their complementary nature and the need for improving, verifying and triangulating available data through experts (Moed et al. 1985), citations soon became a synonym for scientific impact and quality. Consequently, bibliometric indicators were misused in university and journal rankings as

well as in individual hiring and promotion decisions, which has in turn led to adverse effects such as salami publishing and self-plagiarism, honorary authorship, authorship for sale, as well as strategic citing through self-citation or citation cartels (for a recent over-view refer to Haustein and Larivière 2015). Even though altmetrics are presented as a way to counter-balance the obsession with and influence of indicators such as the impact factor or h-index, and make research evaluation fairer by considering more diverse types of scholarly works and impact (Piwowar 2013; Priem and Hemminger 2010), they run the risk of causing similar effects, particularly as they have been created in the midst of a tech-nological push and a policy pull. This paper focuses on the current challenges of altmetrics; a particular emphasis is placed on how the current development of this new family of indicators compares with the past development of bibliometrics.

Challenges of altmetrics

Altmetrics face as many challenges as they offer opportunities. In the following section, three major—what the author considers the most profound—challenges are identified and discussed. These include the heterogeneity, data quality issues and specific dependencies of altmetrics.

Heterogeneity

The main opportunity provided by altmetrics—their variety or heterogeneity—represents also one of their major challenges. Altmetrics comprise many different types of metrics, which has made it difficult to establish a clear-cut definition of what they represent. The fact that they have been considered as a unified, monolithic alternative to citations has hindered discussions, definitions and interpretations of what they actually measure: why would things as diverse as a mention on Twitter, an expert recommendation on F1000, a reader count on Mendeley, a like on Facebook, a citation in a blog post and the reuse of a dataset share a common meaning? And why are they supposed to be inherently different from a citation, which can itself occur in various forms: a perfunctory mention in the introduction, a direct quote to highlight a specific argument, or a reference to acknowledge the reuse of a method? In the following, the challenges associated with their heterogeneity and lack of meaning are discussed by addressing the absence of a common definition, the variety of social media acts, users and their motivation, as well as the lack of a conceptual framework or theory.

Lack of a common definition

Although altmetrics are commonly understood as online metrics that measure scholarly impact alternatively to traditional citations, a clear definition of altmetrics is lacking. Priem (2014, p. 266) rather broadly defined altmetrics as the "study and use of scholarly impact measures based on activity in online tools and environments", while the altmetrics man-ifesto refers to them as elements of online traces of scholarly impact (Priem et al. 2010), a definition that is similar to *webometrics* (Björneborn and Ingwersen 2004) and congruent with the *polymorphous mentions* described by Cronin et al. (1998). Moed (2016, p. 362) conceptualizes altmetrics as "traces of the computerization of the research process". A less abstract definition of what constitutes an altmetric is, however, absent and varies among

authors, publishers, and altmetric aggregators. As it was demonstrated that these new metrics are complementary rather than an alternative to citation-based indicators, the term has been criticized and suggested to be replaced by *influmetrics* (Rousseau and Ye 2013) or *social media metrics* (Haustein et al. 2014). There is also much confusion between alt-metrics and article level metrics (ALM), which as a level of aggregation can refer to any type of metric aggregated for articles. While metrics based on social media represent the core of altmetrics, some also consider news media, policy documents, library holdings and download statistics as relevant sources, although derived indicators for those have been available long before the introduction of altmetrics (Glänzel and Gorraiz 2015). Despite often being presented as antagonistic, some of these metrics are actually similar to journal citations (e.g., mention in a blog post), while others are quite different (e.g., tweets). As a consequence, various altmetrics can be located on either side of citations on the spectrum from low to high levels of engagement with scholarly content (Haustein et al. 2016a). Moreover, the landscape of altmetrics is constantly changing. The challenge of the lack of a common definition can thus be only overcome if altmetrics are integrated into one metrics toolbox:

> It may be time to stop labeling these terms as parallel and oppositional (i.e., alt-metrics vs bibliometrics) and instead think of all of them as available scholarly metrics—with varying validity depending on context and function (Haustein et al. 2015c, p. 3).

Following Haustein et al. (2015c) and employing the terminology and framework used by Haustein et al. (2016a), scholarly metrics are thus defined as indicators based on recorded events of acts (e.g., viewing, reading, saving, diffusing, mentioning, citing, reusing, modifying) related to scholarly documents (e.g., papers, books, blog posts, datasets, code) or scholarly agents (e.g., researchers, universities, funders, journals). Hence, altmetrics refer to a heterogeneous subset of scholarly metrics and are a proper subset of informetrics, scientometrics and webometrics (Fig. 1).
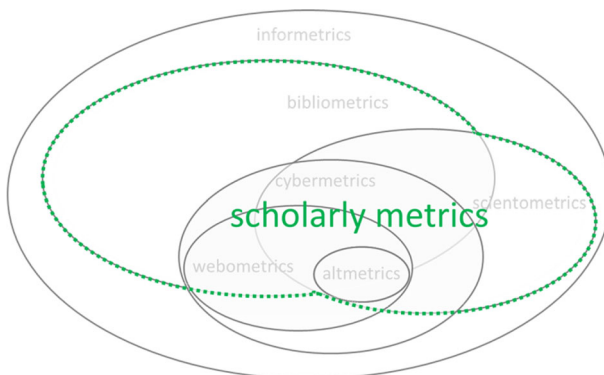

Fig. 1 The definition of scholarly metrics and the position of altmetrics in informetrics, adapted from Björneborn and Ingwersen (2004, p. 1217). Sizes of the ellipses are not representative of field size but made for the sake of clarity only

Heterogeneity of social media acts, users and motivations

As shown above, "altmetrics are indeed representing very different things" (Lin and Fenner 2013, p. 20). Even if one considers only those based on social media activity, altmetrics comprise anything from quick—sometimes even automated—mentions in microposts to elaborate discussions in expert recommendations. The diversity of indicators is caused by the differing nature of the platforms, which entail diverse user populations and motivations. This, in turn, affects the meaning of the derived indicator. It is thus futile to speak about the *one* meaning of altmetrics but rather the meaning of specific types or groups of metrics. Although many platforms now incorporate several functions, which aggravates classification, the following seven groups of social media platforms used for altmetrics are identified:

(a) social networking (e.g., Facebook, ResearchGate)
(b) social bookmarking and reference management (e.g., Mendeley, Zotero)
(c) social data sharing including sharing of datasets, software code, presentations, figures and videos, etc. (e.g., Figshare, Github)
(d) blogging (e.g., ResearchBlogging, Wordpress)
(e) microblogging (e.g., Twitter, Weibo)
(f) wikis (e.g., Wikipedia)
(g) social recommending, rating and reviewing (e.g., Reddit, F1000Prime)

The different purposes and functionalities of these platforms attract different audiences to perform various kinds of acts. For example, recommending a paper is inherently different from saving it to a reference manager and blogging about a dataset differs from tweeting it (Taylor 2013). These differences are reflected in the metrics derived from these platforms. The selectivity and high level of engagement associated with blogging and recommending is mirrored in the low percentage of papers linked to these events, for instance. Less than 2 % of recent publications get mentioned in blog posts and the percentage of papers being recommended on F1000Prime is equally sparse. On the other hand, Twitter and Mendeley coverage is much higher at around 10–20 % and 60–80 %, respectively (Haustein et al. 2015b; Priem et al. 2012; Thelwall and Wilson 2015; Waltman and Costas 2014). This reflects both the lower level of engagement of tweeting and saving to Mendeley compared to writing a blog post or an F1000Prime review, as well as the size of the Twitter and Mendeley user populations in comparison to the number of bloggers and F1000Prime experts. Although the signal for Mendeley is quite high at the paper level, the uptake of the platform among researchers is rather low (6–8 %). Most surveys report that around 10–15 % of researchers use Twitter for professional purposes, while ResearchGate, LinkedIn and Facebook were more popular, although passive use prevailed (Mas-Bleda et al. 2014; Rowlands et al. 2011; Van Noorden 2014). The different types of uses and compositions of users echo in correlations with citations. Moderate to high positive correlations were found between citations and Mendeley reader counts and F1000Prime recommendations (Bornmann and Leydesdorff 2013; Li et al. 2012; Thelwall and Wilson 2015), which alludes to the academic use and users of these platforms. F1000Prime 'faculty members' are a selected group of experts in their fields, while the majority of Mendeley users have been shown to consist of students and early-career researchers (Zahedi et al. 2014b). On the other hand, correlations between citations and tweets are rather weak, which suggests non-academic use (Costas et al. 2015; Haustein et al. 2015b). More detailed information about user demographics and particularly their motivation to interact with scholarly contents on social media is, however, still mostly lacking.

The heterogeneity of altmetrics is not only apparent between the seven types of sources but also within. For example, Facebook differs from ResearchGate and Reddit from F1000Prime. Even the exact same event can be diverse, for example, tweets can comprise the promotion of one's own work (self-tweets), the diffusion of a relevant paper, the appraisal of a method or the criticism of results. A Mendeley readership count can imply a quick look at or intense reading of a publication. Moreover, not all recorded events—e.g., a tweet linking to a paper—are based on direct acts, but could be automatically created (Haustein et al. 2016b), and not all acts result in a recorded event, if, for example, the absence of identifiers to sources prevents capturing them.

Lack of conceptual frameworks and theories

The lack of a theoretical foundation coupled with its pure data-drivenness is a central and reoccurring criticism of altmetrics. What constitutes an altmetric today is almost entirely determined by the availability of data and the ease with which it can be collected. To be fair, after decades of debates the field of scientometrics has not been successful in implementing an overarching theory either (Sugimoto 2016). And with the creation of the Science Citation Index, the field of bibliometrics and the use of citation analysis in research evaluation have been driven by the availability of data to the point that it has come to shape the definition of scientific impact. Despite being "inherently subjective, motivationally messy and susceptible to abuse" (Cronin 2016, p. 13), the act of citing has been a central part of the scholarly communication process since the early days of modern science, while social media is still searching for its place in academia. In the context of altmetrics, the lack of a theoretical scaffold is caused by the sudden ease of data collection and the high demand from funding bodies to prove societal impact of research. It is further impeded by the previously described heterogeneity of metrics. Based on a framework of various acts, which *access*, *appraise* or *apply* scholarly documents or agents, Haustein et al. (2016a) discussed different altmetrics in the light of normative and social constructivist theories, the concept symbols citation theory as well as social capital, attention economics and impression management. These theories were selected due to the affinity of altmetrics with citations, as well as the imminent social aspect of social media. The fact that the different theories apply better to some of these acts, and are less suitable for others, further emphasizes the heterogeneity of altmetrics. Haustein et al. (2016a) demonstrated that the discussed theories help to interpret the meaning of altmetrics but stress that they are not able to fully explain acts on social media. Other theories as well as more qualitative and quantitative research are needed to understand and interpret altmetrics.

Data quality

The central importance of data quality cannot be emphasized enough, particularly in the context of research evaluation. Data quality, including metadata quality and errors, and coverage, have thus been a reoccurring topic in bibliometrics, with a focus on comparing the three major citation indexes Web of Science, Scopus and Google Scholar. In altmetrics, data quality is a major challenge and transcends the known errors and biases for citation data. In the context of citations, errors mostly represent discrepancies between the act and the recorded event. These can be discovered and measured either through triangulation of different data aggregators (i.e., citation indexes) or by referring to the original source (i.e., the reference list in a publication). While bibliometrics sources are static documents, most data sources in the context of altmetrics are dynamic, which can be altered or deleted

entirely. Accuracy, consistency and replicability can be identified as the main issues of altmetrics data quality. Potential data quality issues can occur at the level of data providers (e.g., Mendeley, Twitter), data aggregators (e.g., Altmetric, ImpactStory, Plum Analytics) and users. It should be noted that many providers (e.g., Twitter, Facebook, Reddit) are not targeted at academia and altmetrics data quality is thus not their priority. In the context of the NISO Altmetrics Initiative,[1] a working group on altmetrics data quality[2] is currently drafting a code of conduct to identify and improve data quality issues.

Research into altmetrics data quality is still preliminary and has so far mainly focused on the accuracy of Mendeley data (Zahedi et al. 2014a) and the consistency between altmetric aggregators (Jobmann et al. 2014; Zahedi et al. 2014c, 2015). Inconsistencies between data aggregators can be partly explained by different retrieval strategies. For example, Lagotto aggregates Facebook shares, likes and comments, while Altmetric records only public posts. Recording tweets continuously and in real time, Altmetric shows the highest coverage of papers on Twitter, while Lagotto captures only a fraction (Zahedi et al. 2015). The extent to which Altmetric's record of tweets to scientific papers is accurate or complete is unknown. Replication would only be possible through a direct verification against Twitter's data, which is precluded by the costliness of access. The replicability of altmetrics is further impeded by the dynamic nature of events. While citations cannot decrease because they cannot be deleted, Mendeley readership counts can change over time (Bar-Ilan 2014). The provision of timestamps for events and longitudinal statistics might be able to mitigate replicability issues at least to some extent.

The quality of altmetrics data is also affected by the metadata of scholarly works. One central aspect is the availability of certain identifiers (particularly the DOI), which is discussed below in the context of dependencies. Moreover, the metadata of traditional publications might not be sufficient to construct and (potentially) normalize altmetric indicators in a meaningful manner. The publication year, which is the time-stamp used for citation indicators, is not appropriate for social media events, which happen within hours of publication and exhibit half-lives that can be counted in days rather than years (Haustein et al. 2015a). Whether or not to aggregate altmetrics for various versions of documents (e.g., preprint, version of record) is also a central issue of debate. While the location—that is, the journal website, repository or author's homepage—of a document makes hardly any difference for citations, it is essential for tracking related events on social media.

Dependencies

Altmetrics have been and continue to develop under the pressure of various stakeholders. They are explicitly driven by technology and represent a *computerization movement* (Moed 2016). The availability of big data and the ease with which they can be assessed was met by a growing demand, particularly by research funders and managers, to make the societal impact of science measurable (Dinsmore et al. 2014; Higher Education Funding Council for England 2011; Wilsdon et al. 2015)—despite the current lack of evidence that social media events can serve as appropriate indicators of societal impact. Along these lines, the role of publishers in the development of altmetrics needs to be emphasized. Owned by large for-profit publishers, Altmetric, Plum Analytics and Mendeley operate under a certain pressure to highlight the value of altmetrics and to make them profitable. Similarly, many journals have started to implement altmetrics, not least as a marketing instrument.

---

[1] http://www.niso.org/topics/tl/altmetrics_initiative/.

[2] http://www.niso.org/publications/newsline/2015/working_group_connection_apr2015.html#bi0.

The majority of collected altmetrics are those that are comparably easily captured, often with the help of APIs and document identifiers. Activity on ResearchGate or Zotero is, for example, not used because these platforms do not (yet) offer APIs. The strong reliance on identifiers such as DOIs creates particular biases against fields and countries where they are not commonly used, such as the social sciences and humanities (Haustein et al. 2015b) and the Global South (Alperin 2015). The focus on DOIs also represents a *de facto* reduction of altmetrics to journal articles, largely ignoring more diverse types of scholarly outputs, which, ironically, contradicts the diversification and democratization of scholarly reward that fuels the altmetrics movement. Above all, this repeats the frequently criticized biases of Web of Science and Scopus that it seeks to overcome.

Another major challenge of altmetrics presents itself in the strong dependency on data providers and aggregators. Similarly to the field of bibliometrics, which would not exist without the Institute for Scientific Information's Science Citation Index, the development of altmetrics is strongly shaped by data aggregators, particularly Altmetric. The loss of Altmetric would mean losing a unique data source.[3] Their collection of tweets, for instance, has become so valuable that another data aggregator (ImpactStory) has decided to obtain data from them. This alludes to a monopolistic position that, in the case of the Institute for Scientific Information, has created the ubiquitous impact factor.

Even more challenging than the dependency on aggregators is the dependency on social media platforms as data providers. While a clear distinction can be made between papers' reference lists and the recorded citations in Scopus, Web of Science and Google Scholar, acts and recorded events are virtually identical on social media. If a citation database ceases to exist, it can still be reconstructed using the publications it was based on. If Twitter or Mendeley were discontinued, an entire data source would be lost and the acts of tweeting and saving to Mendeley would no longer exist,[4] as they are not independent from the platforms themselves.

The strong dependency on social media platforms culminates in how their very nature directly affects user behavior and thus how technological affordances shape the actual acts. While this aspect has not been systematically analyzed, one could hypothesize that the use of Mendeley increases the number of cited references and the citation density in a document, that the tweet button increases the likelihood of tweeting, or that automated alerts, such as that of Twitter about trending topics, triggers further (re)tweets. The latter would represent a technology-induced Matthew effect for highly tweeted papers.

Conclusions and outlook

For any metric to become a valid indicator of a social act, the act itself needs to be conceptualized (Lazarsfeld 1993). The conceptualization can hence be used to construct meaningful indicators. In the case of altmetrics, recorded online events are used without having a proper understanding of underlying acts and in how far they are representative of various engagements with scholarly work. In fact, many of these acts are actually still forming and being shaped by technological affordances. Given the heterogeneity of the acts on which altmetrics are based, one would expect that a variety of concepts are needed to provide a clearer understanding of what is measured. Some recorded events might prove

---

[3] The loss might be avoided or at least mitigated by maintaining a dark archive, which was mentioned by Altmetric founder Euan Adie in a tweet: https://twitter.com/stew/status/595527260817469440.

[4] In the short history of altmetrics, such loss can already be observed for Connotea.

to be useful as filters, others might turn out to be valid measures of impact, but many might reflect nothing but buzz. In this context, it cannot be emphasized enough that social media activity does not equal social impact. Establishing a conceptual framework for scholarly metrics is particularly challenging when one considers that citations—which represent a single act—still lack a consensual theoretical framework after decades of being applied in research evaluation. While citations, despite being motivated by a number of reasons, have been and continue to be an essential part of the scholarly communication process, social media are just starting to enter academia and researchers are just beginning to incorporate them into their work routines.

The importance of data quality cannot be stressed enough particularly in today's *evaluation society* (Dahler-Larsen 2012) and in a context where any number beats no number. The dynamic nature of most of the events that altmetrics are based on provides a particular challenge with regards to their accuracy, consistency and reproducibility. Ensuring high data quality and sustainability is further impeded by the strong dependency on single data providers and aggregators. Above all, the majority of data is in the hands of for-profit companies, which contradicts the openness and transparency that has motivated the idea of altmetrics. As indicators will inevitably affect the processes which they intend to measure, adverse effects need to be mitigated by preventing misuse and avoiding to place too much emphasis on one indicator. While citation indicators have served as a great antagonist to help altmetrics gain momentum, it would not be in the interest of science to replace the impact factor by the Altmetric donut. It needs to be emphasized that any metric—be it citation or social media based—has to be chosen carefully with a view to the particular aim of the assessment exercise. The selection of indicators thus needs to be guided by the particular objectives of an evaluation, as well as by an understanding of their capabilities and constraints.

References

Alperin, J. P. (2015). Geographic variation in social media metrics: An analysis of Latin American journal articles. *Aslib Journal of Information Management, 67*(3), 289–304. doi:10.1108/AJIM-12-2014-0176.

Bar-Ilan, J. (2014). JASIST@ Mendeley revisited. In *altmetrics14: expanding impacts and metrics, workshop at web science conference 2014*. Retrieved from http://files.figshare.com/1504021/JASIST_new_revised.pdf

Björneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. *Journal of the American Society for Information Science and Technology, 55*(14), 1216–1227. doi:10.1002/asi.20077.

Bornmann, L., & Leydesdorff, L. (2013). The validation of (advanced) bibliometric indicators through peer assessments: A comparative study using data from InCites and F1000. *Journal of Informetrics, 7*(2), 286–291. doi:10.1016/j.joi.2012.12.003.

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

Cronin, B. (2016). The Incessant chattering of texts. In C. R. Sugimoto (Ed.), *Theories of informetrics and scholarly communication. A Festschrift in Honor of Blaise Cronin* (pp. 13–19). Berlin: De Gruyter.

Cronin, B., Snyder, H. W., Rosenbaum, H., Martinson, A., & Callahan, E. (1998). Invoked on the web. *Journal of the American Society for Information Science, 49*(14), 1319–1328. doi:10.1002/(SICI)1097-4571(1998)49:14<1319:AID-ASI9>3.0.CO;2-W.

Dahler-Larsen, P. (2012). *The evaluation society*. Stanford, CA: Stanford Business Books, an imprint of Stanford University Press.

Dinsmore, A., Allen, L., & Dolby, K. (2014). Alternative perspectives on impact: The potential of ALMs and altmetrics to inform funders about research impact. *PLoS Biology, 12*(11), e1002003.

Garfield, E. (1955). Citation indexes for science. A new dimension in documentation through association of ideas. *Science, 122*, 108–111.

Gilbert, G. N. (1977). Referencing as persuasion. *Social Studies of Science*, *7*(1), 113–122.

Glänzel, W., & Gorraiz, J. (2015). Usage metrics versus altmetrics: Confusing terminology? *Scientometrics, 102*(3), 2161–2164. doi:10.1007/s11192-014-1472-7.

Gross, P. L. K., & Gross, E. M. (1927). College libraries and chemical education. *Science, 66*(1713), 385–389.

Haustein, S., Bowman, T. D., & Costas, R. (2015a). When is an article actually published? An analysis of online availability, publication, and indexation dates. In *Proceedings of the 15th international society of scientometrics and informetrics conference* (pp. 1170–1179). Istanbul, Turkey.

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

Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., & Larivière, V. (2016b). Tweets as impact indicators: Examining the implications of automated "bot" accounts on Twitter. *Journal of the Association for Information Science and Technology, 67*(1), 232–238. doi:10.1002/asi.23456.

Haustein, S., Costas, R., & Larivière, V. (2015b). Characterizing social media metrics of scholarly papers: The effect of document properties and collaboration patterns. *PLoS ONE, 10*(3), e0120495. doi:10.1371/journal.pone.0120495.

Haustein, S., & Larivière, V. (2015). The use of bibliometrics for assessing research: possibilities, limitations and adverse effects. In I. M. Welpe, J. Wollersheim, S. Ringelhan, & M. Osterloh (Eds.), *Incentives and performance: Governance of knowledge-intensive organizations* (pp. 121–139). Springer International Publishing. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-09785-5_8

Haustein, S., Larivière, V., Thelwall, M., Amyot, D., & Peters, I. (2014). Tweets vs. Mendeley readers: How do these two social media metrics differ? *It—Information Technology, 56*(5), 207–215. doi:10.1515/itit-2014-1048.

Haustein, S., Sugimoto, C., & Larivière, V. (2015c). Guest editorial: Social media in scholarly communication. *Aslib Journal of Information Management,*. doi:10.1108/AJIM-03-2015-0047.

Higher Education Funding Council for England. (2011). *Decisions on assessing research impact. Research Excellent Framework (REF) 2014 (Research Excellence Framework No. 01.2011)*. Retrieved from http://www.ref.ac.uk/media/ref/content/pub/decisionsonassessingresearchimpact/01_11.pdf

Jobmann, A., Hoffmann, C. P., Künne, S., Peters, I., Schmitz, J., & Wollnik-Korn, G. (2014). Altmetrics for large, multidisciplinary research groups: Comparison of current tools. *Bibliometrie-Praxis Und Forschung, 3*. Retrieved from http://www.bibliometrie-pf.de/index.php/bibliometrie/article/view/205

Lazarsfeld, P. F. (1993). *On social research and its language*. Chicago: University of Chicago Press.

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics, 91*(2), 461–471. doi:10.1007/s11192-011-0580-x.

Lin, J., & Fenner, M. (2013). Altmetrics in evolution: Defining and redefining the ontology of article-level metrics. *Information Standards Quarterly, 25*(2), 20–26.

Mas-Bleda, A., Thelwall, M., Kousha, K., & Aguillo, I. F. (2014). Do highly cited researchers successfully use the social web? *Scientometrics, 101*(1), 337–356. doi:10.1007/s11192-014-1345-0.

Moed, H. F. (2016). Altmetrics as traces of the computerization of the research process. In C. R. Sugimoto (Ed.), *Theories of informetrics and scholarly communication. A Festschrift in Honor of Blaise Cronin* (pp. 360–371). Berlin: De Gruyter.

Moed, H. F., Burger, W. J. M., Frankfort, J. G., & Van Raan, A. F. J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy, 14*(3), 131–149. doi:10.1016/0048-7333(85)90012-5.

Piwowar, H. (2013). Value all research products. *Nature, 493*, 159. doi:10.1038/493159a.

Priem, J. (2014). Altmetrics. In B. Cronin & C. R. Sugimoto (Eds.), *Beyond bibliometrics: harnessing multidimensional indicators of performance* (pp. 263–287). Cambridge, MA: MIT Press.

Priem, J., & Hemminger, B. M. (2010). Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. *First Monday*, *15*(7). Retrieved from http://pear.accc.uic.edu/ojs/index.php/fm/rt/printerFriendly/2874/2570

Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *arXiv Print*, 1–17.

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010, October 26). *Altmetrics: A manifesto*. Retrieved from http://altmetrics.org/manifesto/

Rousseau, R., & Ye, F. Y. (2013). A multi-metric approach for research evaluation. *Chinese Science Bulletin, 58*(26), 3288–3290. doi:10.1007/s11434-013-5939-3.

Rowlands, I., Nicholas, D., Russell, B., Canty, N., & Watkinson, A. (2011). Social media use in the research workflow. *Learned Publishing, 24*(3), 183–195. doi:10.1087/20110306.

Sugimoto, C. R. (Ed.). (2016). *Theories of informetrics and scholarly communication*. Berlin: De Gruyter.

Taylor, M. (2013). Towards a common model of citation: Some thoughts on merging altmetrics and bibliometrics. *Research Trends, 35*, 1–6.

Thelwall, M., & Wilson, P. (2015). Mendeley readership altmetrics for medical articles: An analysis of 45 fields. *Journal of the Association for Information Science and Technology,*. doi:10.1002/asi.23501.

Van Noorden, R. (2014). Online collaboration: Scientists and the social network. *Nature, 512*(7513), 126–129. doi:10.1038/512126a.

Waltman, L., & Costas, R. (2014). F1000 recommendations as a potential new data source for research evaluation: A comparison with citations. *Journal of the Association for Information Science and Technology, 65*(3), 433–445. doi:10.1002/asi.23040.

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., et al. (2015). *The metric tide: report of the independent review of the role of metrics in research assessment and management.,*. doi:10.13140/RG.2.1.4929.1363.

Zahedi, Z., Bowman, T. D., & Haustein, S. (2014a). Exploring data quality and retrieval strategies for Mendeley reader counts. *Presented at the SIG/MET workshop, ASIS&T 2014 annual meeting*, Seattle. Retrieved from http://www.asis.org/SIG/SIGMET/data/uploads/sigmet2014/zahedi.pdf

Zahedi, Z., Costas, R., & Wouters, P. (2014b). Assessing the impact of publications saved by Mendeley users: Is there any different pattern among users? In *IATUL conference*, Espoo, Finland, June 2–5 2014. Retrieved from http://docs.lib.purdue.edu/iatul/2014/altmetrics/4

Zahedi, Z., Fenner, M., & Costas, R. (2014c). How consistent are altmetrics providers? Study of 1000 PLOS ONE publications using the PLOS ALM, Mendeley and Altmetric. com APIs. In *altmetrics 14. Workshop at the web science conference*, Bloomington, USA. Retrieved from http://files.figshare.com/1945874/How_consistent_are_altmetrics_providers__5_.pdf

Zahedi, Z., Fenner, M., & Costas, R. (2015). Consistency among altmetrics data provider/aggregators: What are the challenges? In *altmetrics15: 5 years in, what do we know?* Amsterdam, The Netherlands. Retrieved from http://altmetrics.org/wp-content/uploads/2015/09/altmetrics15_paper_14.pdf