

Social Media Metrics as Indicators of Broader Impact

Cassidy R. Sugimoto*, Stefanie Haustein**, Qing Ke*, & Vincent Larivière**

* School of Informatics and Computing, Indiana University Bloomington
sugimoto@indiana.edu (corresponding author); keqing.echolife@gmail.com

** École de bibliothéconomie et des sciences de l'information, Université de Montréal
OST-CIRST, Université du Québec à Montréal
stefanie.haustein@umontreal.ca; vincent.lariviere@umontreal.ca

Over the last five years, heterogeneous social media metrics related to scholarly communication—so-called “altmetrics” (Priem et al., 2010)—have begun to be used as indicators of the impact of scientific publications and of other research products. These metrics include, for example, number of tweets, mentions in blogs, Mendeley readership, presence in policy documents, usage of code on GitHub, or views of presentations on SlideShare. Altmetrics have been promoted, in large part, by the increasing use of social media and other online tools for the dissemination of scholarly material by scholars (Van Noorden, 2014), as well as by data producers that aim at creating a market for these indicators. Simultaneously, there has been a strong demand by research funders and policy makers for scholars to demonstrate the impact of research “beyond academia” (HEFCE, 2011; Viney, 2013; Wildson et al., 2015). Altmetrics have been heralded as a potential source of this impact (Bornmann, 2014). However, serious concerns relating to the reliability and validity of these indicators have been raised (Wouters & Costas, 2012; Haustein, 2016), particularly in regards to their use to demonstrate societal impact. The extant literature has focused largely on coverage and correlative analyses—investigating the degree to which various altmetrics correlate with and predict citation counts (e.g., Coastas, Zahedi, & Wouters, 2014; Eysenbach, 2011; Haustein et al., 2014; Haustein, Costas, & Larivière, 2015; Shuai, Pepe, & Bollen, 2012). The low correlations found in most of these studies speak to something novel that is captured in these metrics. Scholars have been quick, therefore, to label social media metrics as an indicator of the “broader impact” of science, demonstrating the impact of science upon the lay public. However, despite a robust literature on the topic of scholarly use of social media and social media metrics (Sugimoto et al., 2016), there remain many misconceptions on this topic. The use of these metrics under these potentially flawed assumptions has several critical consequences for the conduct and reward of science. Therefore, for the purpose of this plenary debate, we will present key provocations on this topic and associated evidence in order to promote informed dialogue on the topic.

Altmetric indicators do not measure a broader array of genres than citations.

Evidence. PLOS article-level metrics were among the first large-scale altmetric data collection activities (Fenner, 2013). These data were, as their name implies, collected to reflect the evidence of dissemination of an individual article published on the PLOS platform. As other altmetric aggregators—e.g., Altmetric.com—began to emerge, their data also tended to use the article as the main research object. This is largely a function of the ability to identify and mine, at scale, unique identifiers (i.e., DOIs) for this research object. Unfortunately, even though the use of DOIs is increasing, it remains lower in social sciences and humanities journals, which places these journals at a disadvantage (Haustein, Costas, & Lariviere, 2015). There is also a disadvantage for these disciplines, as books form the main genre of scholarship in many. As of 2016, Altmetric.com began offering badges, that is, altmetric indicators, for books. However, these have yet to be widely adopted and tested and reflect similar initiatives in the realm of citation indicators (e.g., the creation of the Book Citation Index in WoS). In sum, while altmetric

indicators have widened the scope of sources of attention, the scholarship under examination has remained largely unchanged in comparison with traditional metrics.

Altmetrics are heterogeneous.

Evidence. To measure the impact of these journal articles, aggregators of altmetric data scour a wide range of sources to find evidence of interaction: for example, tweets, facebook posts, newspaper articles, and policy documents which link to the document. There is, however, a vast skewness in the strength of these signals: coverage of documents is highest in social reference managers (i.e., Mendeley) (Costas et al., 2014) and the highest signal comes from microblogging platform Twitter (Thelwall et al., 2013). Fewer than 1% of articles appear in mainstream media (Haustein, Costas, & Lariviere, 2015). Altmetric.com accounts for these different levels of intensity by giving each indicator a different weight—recognizing that one mention in the *New York Times* is not equal to one mention on Twitter. However, the overwhelming volume of signal from Twitter continues to weigh heavily; this is critically important when interpreting the underlying concept of this indicator. In the aggregate, the indicator measures attention rather than impact or quality (Sugimoto, 2015). However, were the data disaggregated (e.g., indicators presented per source of attention), we may be able to construct more theoretically robust frameworks for these indicators.

Altmetric indicators do not measure societal impact.

Evidence. In our work, we have sought to identify the entities producing the signals using a combination of manual, algorithmic, and network analyses. This has allowed us to identify, for example, the proportion of Twitter conversation about science originating from individuals with doctoral degrees and the disciplinary orientation of tweeters (Tsou et al., 2015; Qing, Ahn, & Sugimoto, 2016). We have found that tweets to scientific papers are likely to come from highly educated individuals, most of whom are affiliated with research institutions (Birkholz et al, 2015; Tsou et al, 2015). Present data suggests that altmetric indicators are primarily an indicator of receipt by the scientific community, not by the general public. Although tools such as Altmetric.com and Impact-Story claim to categorize the accounts generating online attention into audience types, these categories have been shown to be highly problematic (Tsou et al, 2015). Better algorithms must be constructed to more accurately represent the attention generated by those outside of the scientific community.

Altmetric indicators do not broaden dissemination across disciplines.

Evidence. There are disciplinary differences in coverage, with Twitter density being highest in the Social Sciences and lowest in Mathematics and Computer Sciences and Natural Sciences and Engineering (Haustein, Costas, & Lariviere, 2015). This corresponds to the disciplinary composition of the tweeting population: social scientists are overrepresented on Twitter, given their proportion in the scientific workforce, whereas mathematicians are underrepresented (Qing, Ahn, & Sugimoto, 2016). This alignment may be explained due to the disciplinary homophily in the scientific twitterverse: we found high assortativity in the network by discipline (Qing, Ahn, & Sugimoto, 2016), which suggests that altmetric indicators may not demonstrate a broadening of communication across disciplinary boundaries. Rather, twitter represents another space in which “invisible colleges” are replicated. One exception are highly tweeted papers which are often related to topics (and humor) that are shared across the research community.

Altmetric indicators are easily gamed.

Evidence. One concern with altmetric indicators is the presence of automated accounts on social media platforms—i.e., “bots”—and the influence of these bots upon social media metrics (Haustein et al., 2016). For example, a conservative estimate suggests that automated handles on Twitter account for nearly 10% of tweets to scientific papers (Haustein et al., 2016). These automated accounts may be run by a publisher, journal, or by an interested individual. These are not necessarily malicious, but they threaten the interpretation of altmetric indicators as a measurement of any social activity—let alone of impact—as they are more likely a marketing signal. Of course, there is gaming in citations as well as demonstrated by the establishment of citation cartels and the evidence of editors gaming the JIF. However, gaming in altmetrics is fundamentally different due to both scale and ease of entrance: there are far fewer (if any) gatekeepers in the altmetric realm and the virality of social media intensifies indicator abuses.

Altmetric indicators do not broaden the geographic dissemination of science.

Evidence. Twitter is predominately European and North American. The effect of this is a difference in the tweeting and citing profiles of the papers. For example, while there is a high local aspect in citations—with most countries citing their own work proportionally more—tweets are mostly being made by American and British users, irrespective of the country of the authors. Similarly, twitter users—irrespective of their country—are more likely to tweet papers authors by scholars from these two countries. Therefore, contrary to what has been expected, altmetrics—at least those based on Twitter—narrow the focus to specific parts of the world, both from the place of research than from the place of impact points of view.

Altmetric indicators broaden the profile of authors whose work generates attention.

Evidence. It has been suggested that altmetrics might serve to democratize impact—that is, allowing those with historically lower power positions in the scientific workforce (e.g., women, the Global South, and underrepresented minorities) to have greater voice and vote in determining impact. There is some evidence to support these assertions. Our work on gender and altmetrics found that women authors were disproportionately represented, given their proportion among publishing authors (Lariviere et al., 2013). We have also examined the gender of the authors whose work was being tweeted and found that these metrics were far more gender neutral than citation data (Paul-Hus et al., 2015). This suggests an opportunity for an indicator that can equalize the playing field for groups and subjects historically undervalued by traditional metrics.

Policy recommendations. Several initiatives have been undertaken to examine the utility of altmetrics for policy-making. The Metric Tide was a report, commissioned by the Higher Education Funding Council for England (HEFCE) to examine the role of metrics in research assessment and management (Wilsdon et al., 2015). NISO formed several working groups to examine standardization of altmetric indicators (NISO, 2016). Most recently, the European Commission has commissioned an “Expert Group on Altmetrics” which will focus on categorizing altmetrics, examining relationships with other established indicators, and work towards an agenda of responsible metrics to present to the EC as part of the Open Science Agenda (European Commission, 2016). This parallels wider efforts within the community for more responsible use of all types of indicators (e.g., Hicks et al, 2015). To these initiatives, we add our recommendations on the adoption of altmetrics, generally, and on the use of altmetrics as a measurement of broader impacts, specifically.

- **Provide disaggregated rather than composite altmetric indicators.** Altmetrics are not monolithic. Aggregating, for example, policy mentions with Twitter mentions constructs a false

indicator with no clear underlying concepts. In order to be policy-relevant, clear concepts need to be established for each altmetric indicator. This will require grouping only those sources together which share the same underlying concept.

- **Account for outliers and gaming.** Altmetric data providers should systematically identify outliers and potential gaming and construct mechanisms for accounting for these in the development of their indicators. The means by which this is done should be made transparent, to improve the confidence in these indicators.
- **Expand both sources of evidence and sources of scholarship.** Many providers have begun to diversify both evidence and sources, but the amount of data is still quite small and adoption of the new indicators is lacking. Much of this is a reflection of the systematic availability of the data online. Publishers, policy makers, and altmetric providers should work in concert to ensure that all genres are given unique identifiers and that these are routinely used in the dissemination of this scholarship.
- **Reframe the conversation around what is meant by broader impact.** Altmetrics may not provide an indication of societal relevance. It may, however, broaden the attention received by certain populations of authors, and make visible the effect that scholarship is having across the world. The nature of the impact should be well understood before adopting altmetrics as a measure of broader impact. There is, overall, a great deal of variation in the results based on idiosyncratic and highly localized approaches to studying the problem. What is needed are large-scale, domain-neutral approaches that seek to understand *who* is generating altmetric signals. If altmetrics are taken to be indicators of impact, we must ask—*impact upon whom?* Unless the producers of the impact signals are identified, no statement as to the actual impact of the work can be made. This, therefore, is the critical next step for the science policy community in regards to this novel set of indicators.

CITED REFERENCES

- Birkholz, J. M., Seeber, M., & Holmberg, K. (2015). Drivers of Higher Education Institutions' visibility: A study of UK HEIs social media use vs. organizational characteristics. In Proceedings of the 2015 International Society for Scientometrics and Informetrics (pp. 502-513). Istanbul, Turkey.
- Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, 8, 895-903.
- Costas, R., Zahedi, Z., & Wouters, P. (2014). 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*, n/a-n/a. <http://doi.org/10.1002/asi.23309>
- European Commission. (2016). Expert group on altmetrics. Retrieved from: http://ec.europa.eu/research/openscience/index.cfm?pg=altmetrics_eg
- Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research*, 13(4), e123-e123. <http://doi.org/10.2196/jmir.2012>
- Fenner, M. (2013). What can article-level metrics do for you? *PLoS Biology*, 11(10), e1001687. <http://doi.org/10.1371/journal.pbio.1001687>
- Haustein, S. (2016). Grand challenges in altmetrics: heterogeneity, data quality and dependencies, *Scientometrics*. <http://doi.org/10.1007/s11192-016-1910-9>
- Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., & Larivière, V. (2016). 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.
<http://doi.org/10.1002/asi.23456>
- Haustein, S., Costas, R., & Larivière, V. (2015). Characterizing social media metrics of scholarly papers: The effect of document properties and collaboration patterns. *PLoS ONE*, 10(3), e0120495.
<http://doi.org/10.1371/journal.pone.0120495>
- Haustein, S., Peters, I., Sugimoto, C.R., Thelwall, M., & Lariviere, V. (2014). Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *JASIST*, 65(4), 656-669.
- Hick, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520, 429-431.
- Higher Education Funding Council for England. (2011). Decisions on assessing research impact. Research Excellent Framework (REF) 2014 (No. 01.2011). Retrieved from http://www.ref.ac.uk/media/ref/content/pub/decisionsonassessingresearchimpact/01_11.pdf
- Lariviere, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C.R. (2013). Global gender disparities in science. *Nature*, 504(7479), 211-213.
- NISO. (2016). NISO Alternative Assessment Metrics (Almetrics) Initiative. Retrieved from: http://www.niso.org/topics/tl/altmetrics_initiative/
- Paul-Hus, A., Sugimoto, C.R., Haustein, S., & Lariviere, V. (2015). Is there a gender gap in social media metrics? *Proceedings of the International Society for Scientometrics and Informetrics Conference 2015*. Istanbul, Turkey.
- Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). Altmetrics: A manifesto.
- Qing, K., Ahn, Y.Y., & Sugimoto, C.R. (2016). A systematic identification and analysis of scientists on Twitter. arXiv. <https://arxiv.org/abs/1608.06229>
- Shuai, X., Pepe, A., & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: article downloads, Twitter mentions, and citations. *PloS One*, 7(11), e47523-e47523.
<http://doi.org/10.1371/journal.pone.0047523>
- Sugimoto, C.R. (2015). "Attention is not impact" and other challenges for altmetrics. Wiley Exchanges. <http://exchanges.wiley.com/blog/2015/06/24/attention-is-not-impact-and-other-challenges-for-altmetrics/#comment-2097762855>
- Sugimoto, C.R., Work, S., Lariviere, V., & Haustein, S. (2016). Scholarly use of social media and altmetrics: a review of the literature. arXiv. <https://arxiv.org/ftp/arxiv/papers/1608/1608.08112.pdf>
- Thelwall, M., Haustein, S., Lariviere, V., & Sugimoto, C.R. (2013). Do altmetrics work? Twitter and ten other social web services. *PLoS ONE*. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0064841>
- Tsou, A., Bowman, T., Ghazinejad, A., & Sugimoto, C.R. (presenter). (2015). Who tweets about science? *Proceedings of the International Society for Scientometrics and Informetrics Conference 2015*. Istanbul, Turkey.
- Van Noorden, R. (2014). Online collaboration: Scientists and the social network. *Nature*, 512(7513), 126-129. <http://doi.org/10.1038/512126a>
- Viney, I. (2013). Altmetrics: Research Council Responds. *Nature*, 494(176). <http://doi.org/10.1038/494176c>
- Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., ... Johnson, B. (2015). The metric tide: report of the independent review of the role of metrics in research assessment and management. <http://doi.org/10.13140/RG.2.1.4929.1363>
- Wouters, P., & Costas, R. (2012). Users, Narcissism and Control - Tracking the Impact of Scholarly Publications in the 21 st Century. In *Proceedings of 17th International Conference on Science and Technology Indicators* (Vol. 2, pp. 847-857).