8-2019

A Brief History and Overview

Elaine M. Lasda
University at Albany, State University of New York, elasda@albany.edu

Measuring Research Impact: Introduction, A Brief History and Overview

Data-driven. Evidence-based. Outcome-oriented. Common buzzwords abound today that demonstrate our propensity as a society for quantifiable, (generally) numeric information that will enable making decisions, allocating resources, prioritizing projects and initiatives. Traditionally, the measure of scientific achievement is predicated on when and how often research output is subsequently cited in other scholarship, generally peer-reviewed journal articles (PRJAs). Citation-based metrics, known as bibliometrics, are now bolstered by other indicators such as alternative metrics, web analytics, journal usage metrics, and other measures of productivity, reach, impact, prestige, and so forth. The existence of these broader measures has been largely facilitated by electronic publishing and dissemination of scholarly output on the World Wide Web. Use of metrics such as the Journal Impact Factor (JIF), citation counts, and more recently, the h-index have primarily been utilized in academic tenure and promotion dossiers to demonstrate the success or merit of the candidate's scholarly pursuits. Evaluation of research through measures of impact extends beyond academe, and use of these indicators is manifesting in new places and in new ways. This work presents five case studies that show how a variety of research impact indicators are being used in specialized settings.

First, providing a bit about the context, history, and evolution of research impact metrics will help set the stage for each of our organizations and lend clarity to their use of metrics in organizational activities.

A matter of resource allocation

Government funding is a key support for scientific inquiry in the United States. Nonetheless, according to the Association of the Advancement of Science, the allocation of all Federal R&D funds peaked at 11.7% of the total US budget in 1965, but by 2017 all R&D funding represented a mere 2.9% of the Federal budget (American Association for the Advancement of Science, 2018). This exemplifies the ever-increasing scarcity of resources available for so-called “pure science,” i.e., phenomena studied “without regard to practical applications” (Stevenson, 2010). Resources have decreased while the number and range of disciplinary subspecialties have increased, as has overall research output. There is a need for scrutiny of research pursuits, as we have seen from well-known retracted theorems such as the vaccine-autism scare and the viability of cold fusion (Institute of Medicine, 2004; Ritter, 2003). Thus it may be only natural that funders of scientific pursuits seek additional means of distinguishing amongst project applications.

The resultant need for scientists and researchers to justify and promote one’s research agenda with funders and other constituencies has engendered a variety of metrics from which to evaluate research at all unit levels: article, author, research group, institution, discipline, country, and the like. Part of the reason for this proliferation is that we can now collect and analyze data on a scale heretofore unprecedented, and there are increasingly sophisticated means of analyzing and discerning patterns (Nowakowska, 1990; Raan, 2014).

There is a rapidly shifting landscape when it comes to measures of research impact. For decades after Eugene Garfield first conceived of his citation index schema, its strength was primarily in coverage of the hard sciences. It has long been the case that social sciences coverage in Garfield’s Social Science Citation Index was significantly less robust, and arts and humanities coverage was even further wanting. Nonetheless, the ISI indexes were the only source with citation data considered authoritative until the early 2000s, when competitors Scopus and Google Scholar began to also provide citation indexing. Clarivate Analytics, the current corporate owner of the ISI indexes appears to be both actively and proactively working assure Web of Science retains its dominant position in the research impact metrics domain by adding journal titles and new databases that cover books, datasets, emerging journals, and more (Clarivate Analytics, 2017).

For better or worse, quantitative and qualitative measures are being used to evaluate research and scholarship of all stripes, despite limitations to various indicators. Experts in bibliometrics, altmetrics and general measures of scholarly reach have long documented the pitfalls of over-reliance and irresponsible use of research impact and metrics indicators (Wilsdon et al., 2015).

Probably the main concern about indicators of scholarly impact for evaluative purposes is that it creates an incentive to play to the metric or, as Muller calls it "juke the stats" (2018, p. 2). The premise is thus: due to the research cycle reward system of increased funding and support for researchers with high research impact scores of varying ilk, Scholars will direct their research inquiries toward areas that garner attention or are hot topics, rather than towards lines of inquiry that are just as, if not more important than the high profile research but may be seen as dry or a fringe area undeserving of attention at the current juncture.

Major influencers and sources of today's research impact metrics

Additional context related to the current metrics landscape is provided by a a brief introduction of significant contributions and contributors to the scientometric landscape.

The use of the term "bibliometrics" is widely attributed to Alan Pritchard. Pritchard felt that a term was necessary to identify a term of art for this burgeoning field. He defined bibliometrics as "... the application of mathematics and statistical methods to books and other media of communication." An interesting side note: Pritchard would have preferred the term "scientology," which he felt would be a clear term implying the study of science. Unfortunately that term was by that time already in use by a well-known religious group (Pritchard, 1969).

"Documentation through the association of ideas, " and the influence of such tools as library authority tables and the legal field's *Shepard's Citations* drove Eugene Garfield's conceptualization of a citation-based scientific index (Garfield in Cronin & Sugimoto, 2015). Garfield's contributions to bibliometrics, citation indexing, and scientometrics are well-documented; for a brief but inclusive summary, see Lawler's chapter in *The Future of the History of Chemical Information* (2014). Garfield's Institute for Scientific Information (ISI), expanded the citation index repertoire to social sciences, arts and humanities. After being an independent for-profit entity, ISI has been subsumed by a firm called Clarivate, by way of an intermediate acquisition by Thomson Reuters, and the citation indexes were redubbed Web of Science along the way. The indicators contained in Web of Science are citation-based and well known. Eugene Garfield remains revered for his vision and drive in conceiving and executing the Science Citation Index and later indexes for other disciplines. Many a written work extolls his brilliance and vision, in fact Cronin & Atkins collected and edited a volume of devoted papers and essays largely singing his praises (2000). There has not been merely this reverential treatment. Cronin, fifteen years after the publication of his *Festschrift,* this time with collaborator Cassidy Sugimoto compiled an even more ponderous tome of articles and essays expounding on historical and current concerns related to the use and misuse of scholarly metrics (2015). As regards Garfield, one presupposes it is best to separate the man from the metrics. Aside from citation counts, the Journal Impact Factor (JIF), an unweighted ratio of times cited over articles published for a 2-year or 5-year time frame is the main metric associated with Web of Science. Cited half-life, and Immediacy index, also original metrics of Garfield's, are measures of the length of a reference's viability over time and the speed by which the reference gets traction and spreads, respectively. More recently, Web of Science added Eigenfactor, a weighted ratio based on the premise that some citing references have greater influence of value than others, and the Article Influence Score which corrects the Eigenfactor Score to a per article level metric. Eigenfactor and Article Influence calculations were inspired by Google's PageRank methodology (Bergstrom, 2007).

Scopus

Scopus was launched in 2004, by analytics and publishing conglomerate Elsevier with a greater set of covered publications than Web of Science, user-friendly navigability and sleek looking analytics pages. At the time of Scopus' release it was less expensive than Web of Science, easy to use, and retained quality control through panel of experts reviewing journal content. Others have documented the errors and omissions contained in the database, which were readily apparent from cursory comparisons (Franceschini, Maisano, & Mastrogiacomo, 2016). Nonetheless, the Scopus interface makes it fairly simple to notify the company of any content problems that were encountered by users through a web form easily located on most Scopus pages. The support documentation for Scopus to this day demonstrates the relative simplicity of the process (Scopus, 2019). When Scopus first came on the market, Web of Science did not have a similar prominently identifiable means of submitting corrections. In addition to error-prone data, the corporate culture of owner Elsevier is also a cause for concern among the research community (Swoger, 2013).

Scopus in the initial years opted to utilize metrics developed independently rather than internally, most notably SCImago Journal Rank (SJR) and Source Normalized Impact Per Paper (SNIP). SJR, developed by the SCImago Group at the University of Extremadura in Spain seeks to measure a journal's "average prestige per paper" using weighted rankings and network analysis (González-Pereiraa, Guerrero-Boteb, & Moya-Anegónc, 2009). The SJR is computed in a manner similar but not identical to Web of Science's Article Influence Score. Major differences include the size of the publication sets (relative to the size of the Scopus database vs. that of Web of Science) the time frame from which citations are captured (3 years for SJR, 5 years for AI), and whether or not to include self-citations (SJR caps self citation content, AI excludes it entirely) (Davis, 2015).

The premise of SNIP, developed at the Center for Science and Technology Studies at Leiden University was to create a metric that corrected for differing publication and citation rates between various disciplines. The means by which this was accomplished was to develop subject based citation networks, establish citation frequency patterns within the network, then measure the citation rate of a publication against this "citation potential" as a probability calculation (Moed, 2010).

In June, 2017 Scopus released CiteScore. At the simplest level, CiteScore is a journal's mean number of citations per publication. Dividing by number of publications corrects for the relative size of a journal; that is to say those journals which publish more articles do not automatically have a higher CiteScore. The calculation does not, however correct for the persistent issue of varying disciplinary citation patterns and practices. To address this, Scopus added percentile rankings to contextualize a journal's CiteScore (James, Colledge, Meester, Azoulay, & Plume, 2018).

Google Scholar

Probably the most controversial citation data provider is Google Scholar, which despite having no rhyme or reason to its coverage gives often significantly higher citation counts than either of the proprietary tools. Several years ago, scholars estimated the size of Google Scholar at approximately 160 to 165 million records (Orduna-Malea, Ayllón, Martín-Martín, & Delgado López-Cózar, 2015). Many researchers favor the high citation counts despite concerns that Google Scholar is inadequate for bibliometric study and research evaluation (Halevi, Moed, & Bar-Ilan, 2017). A legitimate strength of Google Scholar is that it covers more non-English language, non-First World publications than either Scopus or Web of Science, as well as a tremendous amount of "grey" literature and scholarly output other than peer-reviewed journal articles (Haddaway, Collins, Coughlin, & Kirk, 2015; Martín-Martín, Orduna-Malea, Thelwall, & Delgado López-Cózar, 2018) . In 2011, Google released an author profile tool called Google Scholar Citations, which provides author level metrics including their own *i-10 index*, simply the number of times cited in the past 10 years (Connor, 2011; Ortega & Aguillo, 2014).

h-Index

Aside from the indicators mentioned above, another well known indicator for scientific achievement is the h-index. In his seminal work proposing the h-index as a measure of scholarly activity, Jorge Hirsch appears to complain that the common suite of citation based metrics in vogue at the time (implicitly, those emanating from ISI/Web of Science), was a large amount of information for evaluators to digest and comprehend. Therefore he devised an index that would provide a simplified metric for evaluative purposes (Hirsch, 2005). The h-index is meant to be used at the author level, but other units of research production are also sometimes measured. The simplest way to explain how to compute the h-index is to take the researcher's peer reviewed publications and rank them from highest to lowest number of times cited. Plot this ranking on a graph, with times cited on the Y-axis, and label the ranked publication denoted as 1, 2, 3, etc. across the x axis. The integer where the number of times cited on the y axis equals the number of papers on the x-axis is the researcher's h-index (i.e. where x=y). Thus, h-index is a combination measure of productivity and impact, according to Hirsch. It is interesting to note that despite many concerns about using citation-based metrics for evaluation, Hirsch actually designed the h-index for the purpose of providing "a useful yardstick with which to compare, in an unbiased way, different individuals competing for the same resource when an important criterion is scientific achievement" (Hirsch, 2005, p. 16572). Gingras, on the other hand, states that the h-index is essentially a useless metric, because it is an "arbitrary" composite of research quality and quantity, and that it smacks of the precept that "any number beats no number" (Gingras, 2016, pp. 42–43). It may be the simplicity of the h-index that is so appealing to non-specialists. Reiterating the propensity to create metrics out of data that is easy to compile and analyze Gingras states: “too many bibliometricians have focused exclusively on the intricacies of counting any units they could find (citations, tweets, views, web connections, etc.) instead of asking first: what is the meaning of these measures? (Gingras, 2016, xi).

g-Index and other h-Index variations

Leo Egghe felt that Hirsch’s indicator did not properly address what is usually a skewed distribution of citations to a scholar’s oeuvre, therefore proposed a g-index where the “highest number of *g* papers that together received g^2 citations” (Egghe, 2006). The effect of squaring the citation count favors highly cited papers, and creates a more granular distinction between scholars’ scores than does h-index. Egghe posits that this is of greater merit in distinguishing between the scholarly output or scientific achievement of various entities. In addition to Egghe’s variation, many other alternatives have been made to the h-index to account for innumerable sorts of issues with one’s scholarly career (Harzing, 2010). It is not readily discernible from anecdotal evidence of the practical application of scholarly metrics that Egghe’s g-index or any of the other h-index variations appear to be widely adopted at this time. Web of Science, Scopus, and Google Scholar Citations all calculate an author’s h-index. Because these citation indexing sources have differing publication coverage, a researcher’s h-index can vary depending on which source is used.

Alternative (alt)metrics

With the advent of electronic publishing formats, recognition for good research spread across the World Wide Web on blogs, news sites, web pages, social media and other places where researchers navigate to stay on top of current issues. Researchers sometimes access information from places they don't feel are valued for scholarly rigor, such as message boards, blogs, or the various online communities where researchers gather and share information. Thus there is a tension between disciplinary standards and actual practice (Roemer & Borchardt, 2015b). Interest in a way to capture results of the sharing and dissemination of scholarly output in venues other than cited references in PRJAs started to gain momentum. For a brief overview of the essential merits and drawbacks of altmetrics, see Ann Williams' overview in *Online Information Review* (2017).

Priem & Hemminger published one of the first papers in support of using scholarly metrics based on sources other than citations to PRJAs. In their opinion, merely capturing the citing references would no longer reflect whole domains of dissemination through social bookmarking, blogs, social media and other content available on the Internet (2010). They primarily direct the utility of these metrics at promotion and tenure and evaluation of researcher/scholar productivity in terms of not only research, but teaching and service as well. First considered "webometrics," there was an early recognition that connections on the World Wide Web fostered a quick turnaround of knowledge dissemination. Priem and collaborators eventually refined this idea, dubbed these indicators "altmetrics" and generated the seminal work known as the *Altmetric manifesto*. The precise value of many altmetric indicators is not entirely recognized at this point, the manifesto explicitly states we need to "ask *how and why* as well as *how many*?" (Priem, Taraborelli, Groth, & Neylon, 2010). It is evident at the very least from the case studies in this book that altmetrics play a value in charting the path of dissemination of scholarly thought above and beyond researcher and disciplinary milieus.

Plum Analytics

A mere 3 years after Priem and his cohorts published their analysis and manifesto, Michael Buschman and Andrea Michalek published work on a similar theme. They identified five indicators of impact from non-peer reviewed journal sources: usage, captures, mentions, social media, and citations. These indicators remain the basis of the tool they created, PlumX and a visual display of their impact indicators known as the "PlumPrint." They questioned, even at this early stage whether so-called alternative metrics were even still "alternative" (Buschman & Michalek, 2013). Since that time they have added significantly more content to their altmetric mix, and can trace impact to a wide variety of scholarly outputs, not simply PRJAs ("PlumX Metrics - Plum Analytics," n.d.). In 2014, Plum Analytics was acquired by Ebsco, and then later sold to Elsevier in 2017, which owns it to this day (Michalek, 2014, 2017). One author has likened Plum's metrics as a kind of "Nielsen Ratings" (Borofsky, 2012).

Altmetric.com

Euan Adie and his company were meanwhile yet another group creating an alternative metric tool of their own: Altmetric.com (Adie & Roe, 2013). Altmetric.com should not be confused with Priem, et. al's site: *altmetric.org*, although it is an easy error to commit. Adie & Roe described their main interest with Altmetric.com to be the collection of metadata about publication mentions and attention on the web, and not developing a metric per se. Even still, the Altmetric "donut" and single-number Altmetric Attention Score appear to be an attempt at one cohesive indicator to integrate different sources such as: public policy documents, mainstream media, online reference managers, open peer review sources, Wikipedia, Open Syllabus Project, patents, blogs, citations, Faculty of 1000, social media, and multimedia (Altmetric, 2015).

Dimensions

In January of 2018, Herzog and colleagues launched Dimensions which may be the latest database or indexing innovation contributing to the study of the research process and of the evolution of scientific thought. The purpose of Dimensions is to bring together metadata about research through the entire process from grant to output. Therefore in addition to including citation, patent, and altmetric data, Dimensions also includes resources on clinical trials, findings, and data sources related to research projects in various stages of the research cycle. Using linked data, it "aims to be a system that helps the academic community to own the formulation and development of metrics that tell the best stories and give the best context to a piece of research" (Bode, Herzog, Hook, & McGrath, 2018). Dimensions looks promising and their linked data model may show a more granular transmission of scholarly ideas and thought through the research process. Further, the ready connections to open access versions of publications reder Dimensions a robust resource for research dissemination.

Becker Model

An interesting framework for measuring impact has come from the Bernard Becker Medical Library at Washington University at Saint Louis. Dubbed the Becker Model, it guides those looking to measure impact to map to real world changes that were made as a result of the research. Key areas for measurement are:

- “Advancement of Knowledge
- Clinical Implementation
- Community Benefit
- Legislation and Policy
- Economic Benefit “ (“How to Use the Model,” n.d.)

The Becker Model serves to organize and describe major points of real-world impact where the ideas and new knowledge brought forth in research can be applied. This model poses challenges for identifying when and where a specific unit of research has had an impact. Over time as ideas become more widely accepted, attribution to the original scholarly research falls away, a phenomenon known as “obliteration by incorporation” (McCain in Cassidy & Sugimoto; 2014). Naturally, when there is partial or no attribution of the original, it can be very difficult to locate the places in which the ideas put forth originally have landed in the arena of public discourse and societal improvement.

The other side of the coin: peer review

Peer review, or the judgement of experts, is critical for the contextualization and understanding of research. But peer review itself can be less than optimal. Insular communities of scholars may be resistant to new ideas; studies show peer review can be random and subjective to a certain extent; and much like the incentives provided by scholarly metrics, the popular or attention-attracting topics get reviewed favorably, whereas obscure but innovative areas of research may be ignored or rejected (Cronin & Sugimoto, 2015, pp. 621–622). Despite its own limitations, peer review is a valuable tool for those who are not experts in the field to understand the relevance and significance of a given scholarly output to the greater discipline. For some it remains preferable to any numeric metric (Cronin & Sugimoto, 2015, pp. 229–231). With open peer review and open access we may have a more public dialogue based not on blind peer review, which can be seen as removing reviewer accountability, but on all parties knowing full well their colleagues' agreement and disagreement with various theorems or research outputs. Open peer review models such as F1000, Kudos, and Publons may change the dynamic of peer review that we see above in new ways. As a result, measures of research impact and peer review remain counterpoints or checks and balances on the scholarly "rewards" system; both serve to provide differing contextual aspects about research output.

Gamesmanship and fraud

In his book *The Tyranny of Metric*s, Muller asserts many such measurements are a form of surveillance and that reliance on indicators to measure scholarly performance may create collateral behaviors which do not incentivize innovation or new lines of inquiry (2018). Evaluators' reliance on research impact metrics has lent at least some validation to concerns of gamesmanship. For example, the Journal *Journal of Criminal Justice*, indexed by Web of Science saw a dramatic leap in its JIF score when the editor undertook an extensive practice of increasing the citation count of *JCJ.* He did this by publishing a large number of articles that cited *JCJ*, the vast majority of which he authored himself (Bartlett, 2015). Through the years there has also been evidence of citation "cartels" where networks of scholars or journals essentially conspire to increase citations to each other in order for improved citation metrics across the network/cartel (Fister, Fister, & Perc, 2016). Certainly, if the JIF and other citation metrics did not hold very significant weight amongst a variety of stakeholders, the somewhat laborious undertakings such gamesmanship requires would not be worth the effort.

Nonetheless, gamesmanship or fraud while egregious, are not the most common cause of research impact metric misuse. Administrators, campus committees and those evaluators of research projects not intimately familiar with the standards and/or cultural norms of a given discipline can be inclined to view a score as a kind of summarizing shorthand that allows them to quantify the context provided by more narrative materials; for example, in the case of academia: a promotion or tenure candidate's letters of external review, his or her teaching and service oeuvre and other items listed on his or her curriculum vita. Reliance on metrics to supplant or simplify the evaluative process can be subtle or not-so-subtle, but it is a form of misuse that harms both the research entity and the evaluative entity. A glaringly common example of this type of misuse is the propensity to apply journal-level metrics to measure individuals or other units of researcher collaboration. Despite being well documented as inappropriate, scientists felt compelled to formulate a Declaration on Research Assessment (DORA) for scholars, researchers, and institutions to sign in an effort to spread knowledge and understanding about the misuse and misinterpretation of various research impact indicators, particularly the use of JIF as a measure of researcher achievement (Paulus, Cruz, & Krach, 2018; "San Francisco Declaration on Research Assessment (DORA)," 2019).

Surprisingly, fraud and gamesmanship in altmetrics do not seem to be a greater threat to metric integrity than in citation based metrics. Fraud for this type of metric usually centers around the automated creation of fake profiles or sites, known as "bots." Due to their automated nature, "bots" are thus far able to be spotted and filtered from most altmetric tools (Haustein et al., 2016; Liu & Adie, 2013; Roemer & Borchardt, 2015a). It may be worthwhile to note that similar concerns have been voiced about the creation of fake publications on Google Scholar (Delgado López-Cózar, Robinson-García, & Torres-Salinas, 2014). While it may be simple to automate fraud for these purposes, it is also possible to automate the filtering of bots as well, much like unwanted emails or "spam." Like spam filters, bot filters are effective, so long as the computer programmers remain vigilant.

The spread of scholarly metrics in specialized settings

Although they have been long-used in colleges and universities, the use of research impact metrics is increasingly persistent in more specialized settings, and for reasons other than, or perhaps in addition to, the career trajectory of researchers and scientists.

The case studies you see here represent five examples of such specialization. The domains of physical sciences, social sciences, and humanities are all touched upon. While two of the cases are directly affiliated with single institutions of higher learning, (University of Michigan Press and UC Berkeley's Institute for Transportation Studies), these cases do not represent typical academic disciplinary departments with the usual academic needs and concerns. In all cases, it is the staff of the organization's internal information centers, classified as *special libraries* that provide these services to their parent organizations (instead of "special," the term "specialized" is perhaps more descriptive and self-evident outside of the library community). Not all the staff who compile and provide this information possess library science degrees, however. Libraries and information centers are uniquely suited to providing impact metrics services as well as instructing stakeholders and constituencies on the strengths, limitations and appropriate use of such indicators. This is true for a number of reasons. For example, libraries have expertise in using bibliometrics to evaluate library materials for collection retention and acquisition policies. Libraries are cross-disciplinary and generally serve all constituencies across an organization. To take that point further, libraries have no “horse in the race,” generally they represent neutral entities in the provision of the information to the various stakeholders who need to demonstrate the impact of a body of research in order to further organizational mission and vision.

The Case Studies

Briefly, here is an overview of the cases presented in this book. Each provides insight into the breadth and depth of how research impact can be tracked, measured, and communicated to stakeholders.

National Center for Atmospheric Research: NCAR is doing very interesting and labor intensive work related to measuring the scholarly output of associated researchers, the use of a supercomputer, as well as the EarthCube infrastructure. NCAR’s library has excelled through incorporation of home-built applications and implementation of technology solutions to obtain and analyze important data about impact and reach.

University of Michigan Press: With access to an impressive suite of bibliometric, altmetric and data analytics tools, UMP leverages information about its monographs, journals and unique repository items to make informed decisions about the viability of open access and community supported publishing models. UMP also seeks to get its publications indexed in the right sources to assure they will be discoverable, and therefore citable.

Institute for Transportation Studies at UC Berkeley: Answerable ultimately to the California State Legislature, ITS at UC Berkeley has laid the foundation for tracking the dissemination and reach of multidisciplinary transportation-related projects, technical reports, and other grey literature using manual indexing and Google Scholar data, as well as other low- or no-cost sources.

United States Environmental Protection Agency: Librarians at the EPA have leveled up in their ability to create and replicate visually eye-catching reports and infographics that provide stakeholders with vital information about the reach and success of scholarly activity and its applications in enforcing environmental policy and regulations.

Natural History Museum: A proof-of-concept demonstrating the value of altmetrics tools for a humanities and social sciences museum shows that the information provided can help a museum tailor its programming for online and in-person programming, justify research expenses to donors, and complement public relations and other information in providing an understanding of the museum's overall reach and impact in a variety of sectors.

These case studies will be of primary use to a research organization's sub-unit, usually the library or information center, that seeks to provide or improve the provision of research impact services. Internal to the organization, high level administrators, researchers/scientists and associated staff may find this work helpful in understanding what is possible for their organization and its information center, if given the time, opportunity and resources. Externally all stripes of research evaluators, whether funders/donors, policy makers, or others who wish to understand the value in an organizations research output, will gain a better understanding of what information could be used in assessment. This may in turn help evaluators better communicate what impact measures and other scientometric data will effectively demonstrate success or achievement on the part of the organization. The projects described in this work will hopefully provide inspiration and food for thought at what will best work in a variety of specialized settings.

References

Adie, E., & Roe, W. (2013). Altmetric: enriching scholarly content with article-level discussion and metrics. *Learned Publishing*, *26*(1), 11–17. https://doi.org/10.1087/20130103

Altmetric. (2015, July 9). Our sources. Retrieved January 29, 2019, from https://www.altmetric.com/about-our-data/our-sources/

American Association for the Advancement of Science. (2018, April). Historical Trends in Federal R&D. Retrieved January 28, 2019, from https://www.aaas.org/programs/r-d-budget-and-policy/historical-trends-federal-rd

Bartlett, T. (2015, September 23). The Journal That Couldn't Stop Citing Itself. *The Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/article/The-Journal-That-Couldnt-Stop/233313

Bergstrom, C. (2007). Eigenfactor: Measuring the value and prestige of scholarly journals. *College & Research Libraries News*, *68*(5), 314–316.

Bode, C., Herzog, C., Hook, D., & McGrath, R. (2018). A guide to the dimensions data approach. *A Collaborative Approach to Creating a Modern Infrastructure for Data Describing Research: Where We Are and Where We Want to Take It. London: Digital Science*.

Borofsky, Y. (2012, March 14). Plum Analytics: "kind of like Nielson ratings" for academic research launches Alpha. Retrieved January 18, 2019, from https://technical.ly/philly/2012/03/14/plum-analytics-kind-of-like-nielson-ratings-for-academic-research-launches-alpha/

Buschman, M., & Michalek, A. (2013). Are alternative metrics still alternative? *Bulletin of the*

American Society for Information Science and Technology, *39*(4), 35–39. https://doi.org/10.1002/bult.2013.1720390411

Clarivate Analytics. (2017). *Web of Science Fact Book* (pp. 1–4). Retrieved from https://clarivate.com/wp-content/uploads/2017/05/d6b7faae-3cc2-4186-8985-a6ecc8cce1ee_Crv_WoS_Upsell_Factbook_A4_FA_LR_edits.pdf

Connor, J. (2011, July 20). Google Scholar Citations. Retrieved February 4, 2019, from https://scholar.googleblog.com/2011/07/google-scholar-citations.html

Cronin, B., & Sugimoto, C. R. (Eds.). (2015). *Scholarly metrics under the microscope: from citation analysis to academic auditing*. Medford, New Jersey: Published on behalf of the Association for Information Science and Technology by Information Today, Inc.

Davis, P. (2015, July 28). Network-based Citation Metrics: Eigenfactor vs. SJR. Retrieved February 3, 2019, from https://scholarlykitchen.sspnet.org/2015/07/28/network-based-citation-metrics-eigenfactor-vs-sjr/

Delgado López-Cózar, E., Robinson-García, N., & Torres-Salinas, D. (2014). The Google scholar experiment: How to index false papers and manipulate bibliometric indicators: Journal of the American Society for Information Science and Technology. *Journal of the Association for Information Science and Technology*, *65*(3), 446–454. https://doi.org/10.1002/asi.23056

Egghe, L. (2006). An improvement of the h-index: The g-index. *ISSI Newsletter*, *2*(1), 8–9.

Fister, I. J., Fister, I., & Perc, M. (2016). Toward the Discovery of Citation Cartels in Citation Networks. *Frontiers in Physics*, *4*. https://doi.org/10.3389/fphy.2016.00049

Franceschini, F., Maisano, D., & Mastrogiacomo, L. (2016). The museum of errors/horrors in Scopus. *Journal of Informetrics*, *10*(1), 174–182. https://doi.org/10.1016/j.joi.2015.11.006

González-Pereiraa, B., Guerrero-Boteb, V. P., & Moya-Anegónc, F. (2009). The SJR indicator: A new indicator of journals' scientific prestige. *ArXiv Preprint ArXiv:0912.4141*.

Harzing, A.-W. (2010). *The publish or perish book: your guide to effective and responsible citation analysis* (1st ed). Melbourne: Tarma Software Research Pty Ltd.

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. https://doi.org/10.1002/asi.23456

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(46), 16569–16572. https://doi.org/10.1073/pnas.0507655102

James, C., Colledge, L., Meester, W., Azoulay, N., & Plume, A. (2018). CiteScore metrics: Creating journal metrics from the Scopus citation index, 16.

Lawler, B. (2014). The Institute for Scientific Information: A brief history. In L. R. McEwen & R. E. Buntrock (Eds.), *The future of the history of chemical information* (pp. 109–126). Washington DC: American Chemical Society : Distributed in print by Oxford University Press.

Liu, J., & Adie, E. (2013). Five challenges in altmetrics: A toolmaker's perspective. *Bulletin of the American Society for Information Science and Technology*, *39*(4), 31–34.

https://doi.org/10.1002/bult.2013.1720390410

Michalek, A. (2014, January 15). Plum Analytics Becomes Part of EBSCO Information Services - Plum Analytics [Blog]. Retrieved January 28, 2019, from https://plumanalytics.com/plum-analytics-becomes-part-of-ebsco-information-services/

Michalek, A. (2017, February 2). Plum Analytics Joins Elsevier - Plum Analytics [Blog]. Retrieved January 28, 2019, from https://plumanalytics.com/plum-analytics-joins-elsevier/

Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, *4*(3), 265–277. https://doi.org/10.1016/j.joi.2010.01.002

Muller, J. Z. (2018). *The tyranny of metrics*. Princeton: Princeton University Press.

Nowakowska, M. (1990). Cluster analysis, graphs, and branching processes as new methodologies for intelligent systems on example of bibliometric and social network data. *International Journal of Intelligent Systems*, *5*(3), 247–263. https://doi.org/10.1002/int.4550050303

Ortega, J. L., & Aguillo, I. F. (2014). Microsoft academic search and Google scholar citations: Comparative analysis of author profiles. *Journal of the Association for Information Science and Technology*, *65*(6), 1149–1156. https://doi.org/10.1002/asi.23036

Paulus, F. M., Cruz, N., & Krach, S. (2018). The Impact Factor Fallacy. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01487

Priem, J., & Hemminger, B. H. (2010). Scientometrics 2.0: New metrics of scholarly impact on the social Web. *First Monday*, *15*(7). Retrieved from http://pear.accc.uic.edu/ojs/index.php/fm/article/viewArticle/2874

Raan, T. (2014). Advances in bibliometric analysis: Research performance assessment and science mapping (Vol. 87, pp. 17–28).

Roemer, R. C., & Borchardt, R. (2015a). Chapter 3. Issues, Controversies, and Opportunities for Altmetrics. *Library Technology Reports*, *51*(5), 20–30.

Roemer, R. C., & Borchardt, R. (Eds.). (2015b). *Meaningful metrics: a 21st century librarian's guide to bibliometrics, altmetrics, and research impact*. Chicago, Illinois: Association of College and Research Libraries, a division of the American Library Association.

San Francisco Declaration on Research Assessment (DORA). (2019, January 18). Retrieved January 18, 2019, from https://sfdora.org/

Scopus. (2019). Profile and content corrections Support Center. Retrieved February 3, 2019, from https://service.elsevier.com/app/home/supporthub/scopuscontent/

Stevenson, A. (2010). *pure science*. Retrieved from http://www.oxfordreference.com/view/10.1093/acref/9780199571123.001.0001/m_en_gb0674580

Williams, A. E. (2017). Altmetrics: an overview and evaluation. *Online Information Review*, *41*(3), 311–317. https://doi.org/10.1108/OIR-10-2016-0294

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. Unpublished. https://doi.org/10.13140/rg.2.1.4929.1363