NISO Recommended Practice: Outputs of the Alternative Assessment Metrics Project

Jill O’Neill
National Information Standards Organization, joneill@niso.org

Follow this and additional works at: https://digitalcommons.du.edu/collaborativelibrarianship
Part of the Library and Information Science Commons

Recommended Citation
Available at: https://digitalcommons.du.edu/collaborativelibrarianship/vol8/iss3/4

This From the Field is brought to you for free and open access by Digital Commons @ DU. It has been accepted for inclusion in Collaborative Librarianship by an authorized editor of Digital Commons @ DU. For more information, please contact jennifer.cox@du.edu,dig-commons@du.edu.
NISO Recommended Practice: Outputs of the Alternative Assessment Metrics Project

Abstract
In September of 2016, the National Information Standards Organization (NISO) published the collaboratively produced Recommended Practice, *NISO RP-25-2016, Outputs of the Alternative Assessment Metrics Project*. Funded by the Alfred P. Sloan Foundation, the project sought to establish a consensus among stakeholders whose activities require robust and precise tools for gauging the impact and reach of scholarship in a globally networked research environment—more robust than were available from impact factor and other such measures.

Contributions to this effort came from an international population via one-on-one interviews, satellite meetings at conferences, and numerous teleconference gatherings. Based on those inputs, working group efforts coalesced around the need for a consensually derived definition of terms, the development of use cases illustrating the anticipated applications for such metrics, and the crafting of a code of conduct aimed at establishing benchmarks for improving the quality of altmetric data through a focus on transparency, replicability, and accuracy.
NISO Recommended Practice: Outputs of the Alternative Assessment Metrics Project

Jill O’Neill (joneill@niso.org)
Educational Programs Manager, National Information Standards Organization

Abstract

In September of 2016, the National Information Standards Organization (NISO) published the collaboratively produced Recommended Practice, *NISO RP-25-2016, Outputs of the Alternative Assessment Metrics Project*. Funded by the Alfred P. Sloan Foundation, the project sought to establish a consensus among stakeholders whose activities require robust and precise tools for gauging the impact and reach of scholarship in a globally networked research environment—more robust than were available from impact factor and other such measures.

Contributions to this effort came from an international population via one-on-one interviews, satellite meetings at conferences, and numerous teleconference gatherings. Based on those inputs, working group efforts coalesced around the need for a consensually derived definition of terms, the development of use cases illustrating the anticipated applications for such metrics, and the crafting of a code of conduct aimed at establishing benchmarks for improving the quality of altmetric data through a focus on transparency, replicability, and accuracy.

Altmetrics and Scholarly Assessment: An Introduction

The Alternative Assessment Metrics project was launched in 2013 in response to a community need for tools of greater variety and greater precision. Institutional assessment frequently relies on quantitative data. A snapshot of the value of contribution and subsequent influence on the work of others was in the past generally drawn from such metrics as citation counts and impact factors. These metrics, while valued for decades, had certain drawbacks—the time lag involved in gathering citation data; the fact that average journal-level citation statistics cannot be applied directly to a specific article; and the gaming of the system through author self-citation, the reclassification of document types by publishers to boost a journal’s impact factor, etc. As time went by, new technologies reduced the time required for publication and dissemination of results. Digital storage systems, with their associated and increasingly granular tracking of usage, began to offer a better sense of when and how published knowledge was absorbed into the perpetual activity of knowledge creation. Accordingly, in 2010 researchers themselves called for new indicators, christened “altmetrics.”

As various stakeholders grasped the possibilities of a new system, it became evident that competitive and entrepreneurial activity might inadvertently frustrate attempts to extract meaning from the body of rapidly accumulating (and varied) forms of data. As noted in a 2013 presentation made by Nettie Lagace, NISO Associate Director of Programs, in order to be made useful:

- Altmetrics would have to coalesce around commonly held definitions, calculations, and data-sharing practices;
- There would have to be a means for auditing these altmetrics; and
- Organizations interested in applying such metrics would have to understand their meaning and be able to ensure consistent application and meaning across the industry.
There were concerns about different layers of meaning and activity that should be covered by the term itself. *Altmetrics*—did it refer to or include article level metrics, data citation, or other forms of bibliometric data? What did the number of tweets or other forms of social media reference surrounding a published paper truly accomplish? Others hesitated over what were and were not meaningful indicators of scholarly impact and contribution, while poorly defined data-gathering practices complicated any exercise in data analysis. What did a particular metric signify? To what other metric or data might it be legitimately compared?

In early 2013, NISO applied to the Alfred P. Sloan Foundation for funding to support an investigation into community needs for the establishment of foundational standards for these newly emerging alternative metrics and subsequent development of those standards. The perception of NISO in the scholarly community as a consensus-seeking organization suggested that it would be trusted to bring together multiple perspectives, nuances, and needs in developing a single set of recommendations upon which the community could rely.

**Phase I of the Alternative Assessment Metrics Project**

Once funding was secured, a NISO steering committee for the project (see [http://www.niso.org/topics/tl/altmetrics_initiative/sc_roster/](http://www.niso.org/topics/tl/altmetrics_initiative/sc_roster/)) was formed, led by Martin Fenner (at the time, he was affiliated with the Public Library of Science [PLOS] but is currently Technical Director of DataCite). The Steering Committee shaped the approach to gathering input from the wider information community with regard to current needs, practices, and priorities that would be relevant to developing standards and best practices for the implementation and use of altmetrics.

The initial phase of the project in 2013 was launched by creating opportunities for gathering community views on existing practice and conditions; it also focused on identification of specific community priorities. Scholars and researchers, university administrators, librarians, representatives of funding bodies, and publishers, as well as members of the broader public, were invited to participate in a series of related interviews and communal events. Discussions were fostered through in-person meetings held in late 2013 and early 2014 in Washington, DC; San Francisco; and Philadelphia. On-site discussions were simultaneously disseminated via streaming sessions, with additional commentary and interaction from off-site participants brought in via social media. Subsequently, NISO held a series of one-on-one interviews with participants to further elicit relevant views and appropriate context. Emerging from those discussions were more than 250 separate ideas. A considered process of filtering those ideas resulted in a more manageable 25 potential action items, which were made public in a white paper released for public comment by NISO in June of 2014.

The responses to that 2014 White Paper came from a broad array of organizations. Government funding bodies, scientists, medical researchers, international bodies advocating for individual metrics and identifiers, consultants, and a major university press offered additional suggestions for how best to advance the initiative.

**Moving into Phase II of the Alternative Assessment Metrics Project**

NISO progressed to Phase II of the Alternative Assessment Metrics Project with an August 2014 survey. The aim of this survey was to query the larger community on the priorities to be assigned to the 25 action items set forth in the Phase I White Paper. Respondents were asked to rank the items on a scale from “not important” to “very important.” Several priority action items emerged:

- Develop specific definitions for alternative assessment metrics
- Identify research types to which altmetrics can be applied
- Define appropriate metrics and calculation methodologies for specific output types, such as software, datasets, or performances
- Promote and facilitate use of persistent identifiers in scholarly communications
- Develop strategies to improve data quality through normalization of source data across providers

Also deemed to be of importance through this survey were two other items—agreement on proper usage of the term “altmetrics” or identification of an alternate preferred term and the need to develop a statement about the role of alternative assessment metrics in research evaluation. This input was reviewed by the NISO Business Information Topic Committee, one of NISO’s leadership committees, which then crafted the feedback into a consolidated new work proposal that was put before the NISO voting membership. By consolidating the effort involved in the five projects listed above into three working groups, the work was made more manageable and more easily implemented. NISO’s voting membership approved the proposed work item in the fall of 2014 and volunteers for the three working groups (A, B, and C) gathered to carry out the project and began their work in early 2014.

Working Group A included representatives from a variety of academic institutions (both U.S. and international), research institutes, libraries, and commercial entities supplying data, content, and platforms to the information community. Members of this Working Group included representatives from organizations such as the Genetics Society of America, Japan’s National Institute of Science and Technology Policy, and the University of Leiden. Working Group A took on the task of developing definitions for alternative assessment metrics as well as framing use cases that would describe how different stakeholders within the research institution might expect to apply such metrics in the course of their ordinary workflow. Eight stakeholder personas were developed—Librarian, Research Administrator, Member of a Hiring Committee, Academics/Researchers, Publishing Editor, Media or Public Information Officer, and Producer of Altmetrics Data. This last persona actually took in three separate entities—a producer of attention data, a content provider, and/or a platform provider.

The task of defining terms that was assumed by Working Group A meant that they formulated the key definition of the Project. What did this community mean by use of the term altmetrics? The final definition formulated by the group reads as follows:

Altmetrics is a broad term that encapsulates the collection of multiple digital indicators related to scholarly work. These indicators are derived from activity and engagement among diverse stakeholders and scholarly outputs in the research ecosystem, including the public sphere.

The inclusion in the definition of altmetrics of many different outputs and forms of engagement helps distinguish it from more established citation-based metrics. At the same time, it leaves open the possibility of the complementary use of these conventional metrics including for purposes of gauging scholarly impact. However, the development of altmetrics in the context of alternative assessment sets its measurements apart from conventional instances of citation-based scholarly assessment.

The definition was collaboratively generated and subsequently finalized in these broad terms in order to ensure community acceptance and understanding. It was a time-consuming process, necessitating study of the existing altmetrics literature and other communications. To adequately capture the nuances of meaning in use by the various stakeholders (and their associated use cases), the working group held numerous discussions, refining iteration upon iteration. Growing out of the group’s studies of the literature and discussions—and again written with an eye to establishing a common vocabulary and understanding in a rapidly evolving area—the final Recommended Practice included a glossary of 17 other related concepts pertinent to an altmetrics discussion, including usage, scholarly output, and research quality. Some of the terminology was less lofty in tone, as the Working Group wrestled with the practical aspects of day-to-day workflow. What types of tasks might fall under the heading of activity in discussing how a researcher might interact with scholarly output? What would be the difference between an altmetric data aggregator and an altmetric data provider? Before the Recommended Practice appeared in finalized form, the community again had an opportunity to comment on the
definitions supplied and further smooth out meaning.

Working Group B included representatives from such entities as Galter Health Sciences Library at Northwestern University, CASRAI, DataCite, and Jisc, as well as national labs and academic libraries. The group focused on work areas “related to unconventional research outputs and identifiers”, most immediately on the nascent area of data metrics. Its charge included crafting definitions for appropriate metrics and calculation methodologies for specific output types. In particular, this strategy was needed to address outputs such as software, performances, research data, and other output unique to areas of the social sciences. Blogs, CAD files, DNA sequences, diagnostic techniques and procedures may not seem particularly unconventional forms of output, but there were few metrics available to the scholarly and research community for purposes of assessing the practical impact on and value created through these contributions.

Working Group B paid particular attention to data metrics. As data sets represent a primary building block for verifying and reproducing scientific findings, the research community had been struggling to reach consensus on practices surrounding the publication and appropriate citation of data. Only if consensus were achieved would the community be able to satisfy existing governmental demands for open access to and sharing of critical scientific data, and only by achieving that consensus would scientists and other participants receive the appropriate credit for their contributions.

As the NISO group was considering the various elements associated with data publication and data citation, in 2014, the Joint Declaration of Data Citation Principles was released by the Data Citation Synthesis Group of Force11. That declaration articulated the critical importance of datasets in the context of scholarship, noting that without persistent identifiers and robust metadata, the widespread recognition and reuse of such data would be significantly hampered. Additionally, the Declaration noted that any access provided through the mechanism of citation would need to be made readable by both machines and humans. The final published report of NISO Alternative Assessment Metrics Project both acknowledges and recommends the developing standards work of the Force11 community.

The NISO working group therefore focused on how best to encourage the practice of data citation and enable appropriate citability of research datasets. Commonly agreed upon recommendations were machine-actionable persistent identifiers, required metadata, landing pages, inclusion of data citations in reference lists, and research data usage statistics. Ensuring the application and preservation of those elements falls primarily to the institution and perhaps even more specifically to the individual manager of any existing repository at that institution. Research funders bear the responsibility of identifying and providing mechanisms for support of those repositories in developing standard means of interoperability and obtaining meaningful metrics.

Standards are needed for formulating statistics regarding research data use. The COUNTER Code of Practice was recommended as a model for this ongoing process, but in the final NISO Recommended Practice, the working group noted that it is important to determine how best to differentiate between “human” downloads and downloads by research-focused non-human agents. The intent behind those downloads are not identical and the final data used for comparison and assigning value will need to reflect that.

Equally valuable output from Working Group B was the comprehensive catalog of persistent identifier players and schema, which may eventually be added to the CASRAI Data Dictionary (see http://dictionary.casrai.org/Main_Page).

The third Working Group associated with the Alternative Assessment Metrics project was tasked with the development of strategies to improve data quality through source data providers. Working Group C chose to address the issue through the creation of a Code of Conduct. As one might anticipate, the working group included such suppliers as Thomson Reuters; John Wiley and Sons, Ltd.; and Johns Hopkins Uni-
versity Press, with balance provided by representation from rOpenSci, the Scholarly Publishing and Academic Resources Coalition (SPARC), and the Université de Montréal. The resulting Code “aims to improve the quality of altmetric data by increasing the transparency of data provision and aggregation as well as ensuring replicability and accuracy of online events used to generate altmetrics.”

The Code of Conduct as it appears in the Recommended Practice consists of guidelines intended to support the focus on transparency, replicability, and accuracy. Most important is the call for an Annual Report supplied by altmetric data providers and altmetric data aggregators; the Code supplies a standard tabular format for reporting purposes and provides samples for the various suppliers that might be expected to provide such altmetric data (Twitter, Mendeley, Plum Analytics, Facebook, etc.)

All of the Working Groups offered their drafts for public comment in the first and second quarters of 2016. Responses came in from a diverse range of individuals associated with the National Library of Medicine, Research Councils UK, the University of Southampton, and Mendeley, as well as from individual scholars and scientists from a variety of disciplines. Each comment was carefully reviewed and incorporated as deemed advisable by the larger group.

In September of 2016, the final consensus Recommended Practice was released in the hope that, as Todd Carpenter, Executive Director, NISO noted in an accompanying press release, “the recommendations that our working groups so carefully crafted will guide users toward optimal uses of the newly available data that can be such a benefit to their careers and institutions.”

Next Steps

Having developed a meaningful vocabulary for discussions within the community as well as some foundational understanding of needs and requirements, the next question facing the information community at large must be “where do we go from here?” An immediate (and perhaps easy) step is to promote awareness that such a foundation for standard practices surrounding altmetrics has been laid. Already, the industry has briefs on the results of the three-year project at a variety of library conferences and publishing industry events.

Much more challenging—even with community consensus—will be active adoption of these definitions and implementation of recommended practices. Given the critical role played by persistent identifiers in digital information systems, a workflow routine of obtaining and using such identifiers must become commonplace at both the individual and institutional levels. Providers—both emergent and established—must do their part by implementing recommended practices and by ensuring the consistent quality of their data. Accommodation of business needs will be more willingly accepted if libraries and their parent institutions are persuaded that providers are committed to offering legitimate value-add through robust APIs and similar services.

In order to sustain this collaborative accomplishment and propel it forward, NISO has plans to organize and manage an ongoing Maintenance Committee (assuming such a committee will be approved through NISO governance). In addition to determining publicity and education activities, the committee’s efforts may include the support and further development of such resources as registries of compliant vendors and of identifiers and metrics information. A NISO maintenance committee would also discuss the ever-changing altmetrics landscape and be expected to undertake additional consensus work on any alternative assessment issues or challenges that might arise.

---


Outputs of the NISO Alternative Assessment Metrics Project: A Recommended Practice of the National Information Standards Organization 14 September 2016 http://www.niso.org/apps/group_public/download.php/17091/NISO%20RP-25-

Ibid.