STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

Reproducibility or Producibility? Metrics and their masters[1]

Christian Herzog[*], Daniel Hook[**] and Euan Adie[***]

[*]christian@uberresearch.com
Digital Science, London N1 9XW, UK

[**] d.hook@digital-science.com
Digital Science, London N1 9XW, UK
Centre for Complexity Science, Imperial College London, London SW7 2AZ
Department of Physics, Washington University in St Louis, St Louis, Missouri, USA

[***] euan@altmetric.com
Altmetric, London N1 9XW, UK

In this paper, we argue that "reproducibility", though paramount to research integrity, is in some ways less important to the field of scientometrics than is "producibility". In context, we define producibility to be the ability for academics to gain access to citation data, and to use it without restriction, to develop valuable indicators that can contextualise research outputs and help to inform the important process of research evaluation.

At its core, producibility is a question of power. It is widely understood that she who controls the development and availability of scientometric indicators determines the definition of success for research and innovation. This power is apparent when one considers that Journal Impact Factors and H-indices are used when reviewing job applicants in some corners of academia; national evaluation exercises have employed metrics such as publication volume and citation counts to inform funding decisions; and that difficult to "metricize" fields such as the Arts and Humanities are increasingly at a disadvantage when asked to demonstrate their bibliometrics-based "return on investment" for publicly funded research.

The relationship between power and producibility begs a number of questions: Who should be making the metrics? Is the academy comfortable with the current gatekeepers? Should there be a "separation of powers" where data providers may not be the both the controllers of access to the data and, at the same time, the producers and controllers of indicators?

As just one group of players in the much larger academic community, we at Digital Science do not suggest that we have all the answers to these questions. Instead, we believe that an existing, widely-accepted set of best practices, codified in the Leiden Manifesto, offers an excellent framework for thinking about the relationship between metrics, their producibility, ownership and use.

First and foremost, our belief in the Leiden Manifesto is manifest in the idea that indicators developed and maintained by academia itself. That doesn't mean that evaluators, funders, commercial players and

---

[1] The authors are employed by Digital Science

governments should have no part in indicators' definition and production; rather, their development is a shared enterprise with multiple parties collaborating and bringing their unique knowledge to bear.

We also echo the Leiden Manifesto in the belief that metrics need to be open and defined collaboratively. For Digital Science, that means we have developed products like Dimensions that reflect the following principles:

- **Metrics and indicators should be a diverse collection of instruments and tools that help assess different aspects of scholarly work in different and complementary ways:** Across a number of Digital Science products, we have integrated metrics like altmetrics, Field Citation Ratios, and citation counts that map the diffusion of research across articles, data sets, patents, public policy, monographs, and other academic works.
- **Production of metrics and indicators should be close to academics and informed by those with expertise with the data and its properties:** For Dimensions, that means that we worked with researchers at the US National Institutes of Health to integrate the Relative Citation Ratio into our database and consult regularly with other members of the scientometrics community as we refine our integration of other metrics into our products.
- **Metrics and indicators should, wherever possible, be openly documented with clear methodology and should not be tied to specific datasets but rather be able to be freely applied to any dataset:** We have chosen to list a collection of metrics in Dimensions that centres around a core of openly documented, community-developed metrics such as the Relative Citation Ratio (RCR). The NIH worked with us and we have tested the RCR application across all 95 million publications that we track. (There remains the caveat that the RCR has been principally developed by the NIH to be used in biomedical science).
- **Data providers (commercial and non-commercial alike) should always make data available upon which metrics can be built, tested and reproduced:** Digital Science companies Altmetric and Dimensions both make their data available to the community via research data-access programs, which academics can use to build, test, and reproduce indicator development and research; Dimensions also offers free access to the Dimensions database for individuals.

Organisations and movements such as Crossref and I4OC are also moving towards making the Leiden Manifesto's vision possible.

Diversity of metrics is another Leiden Manifesto principle of critical importance that is affected by producibility. We cannot have diverse metrics if we lack access to adequate data to develop them. For far too long, the scientometrics community has looked at publications and citations alone, and hence has focussed on a backward-looking view of research. This is due in part to the fact that no other information was freely available to study and partly due to the siloing of data that did exist so that, for example, the connection between a publication and the grant that funded it has been challenging to find, let alone analyse. Now there are opportunities to develop more diverse metrics, this type of contextual data is becoming available to researchers and evaluators. The metrics and indicators that can be produced from these data can be much more timely than previous data allowed.

Of course, reproducibility is also a concern to the scientometrics community, one that is addressed in the Leiden Manifesto's appeal to "Allow those evaluated to verify data and analysis". If as a community we value verifiability, then we all have a role to play. Platforms should make their data as available as practically possible and research groups should make their research and intermediate datasets or supporting code available wherever possible. Many fields can claim ethical considerations that favour extreme care around the sharing, publication and handling of personal, commercial, or other sensitive data. However, much the data used scientometrics is derived from publicly available

data, much of which is the direct result of public funding. Indeed, Scientometrics as a field would seem not only to be well matched with the principles of open research and open data sharing but compelled to share data, methods and results from an ethical standpoint.

However, there are important issues in reproducibility and producibility that the Leiden Manifesto does not address. One such issue is the "separation of powers" that we suggest here, in which data providers should not be both the producers as well as controllers of indicators and the data they underlies those metrics. The current approach neither seems scientific nor open and hence is deeply at odds with the research space that it is meant to serve. It is also a single point of dependency and consequently a single point of failure in research evaluation. There is an "unvirtuous circle" in the creation of metrics that drive behaviours, but where those metrics are not transparent and where the data underlying them is not auditable.

Ultimately, (and most concerningly) this intransparency exposes our community to the risk of manipulation. As realised by the DORA community, without proper oversight, metrics can induce behaviours that are either not envisaged or not well thought through and, by extension, which are not advantageous to the research enterprise. Examples include reductionism in evaluation – it is now well understood that Journal Impact Factor is not a measure of quality (or even attention) for an individual article. However, the JIF has been a strong driver in publishing patterns, reading patterns, funding outcomes and professional development.

Dimensions is Digital Science's first step in to move towards a separation of powers where Digital Science focuses on the data and not on the development of proprietary indicators, at the same time enabling the research community to take ownership of the indicators making use of our efforts to aggregate and expand the data. At this time, we will see data becoming more commoditised and freely available. Digital Science sees this as an opportunity to drive efficiency of data collection and drive down the costs to the sector while at the same time leveraging this efficiency to be able to expand the data that is available to the sector. Thus we see a bright future for data providers who put data innovation at the heart of their relationship with the sector. Thus, our commitment is to work with the research community within the limits of our resources to enable the scientometric community to create innovative and diverse open metrics and indicators that better serve research.

Acknowledgements

Reference
Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. In *Nature*, 520, pp. 429–431