Open Metrics

Silvio Peroni
silvio.peroni@unibo.it – https://orcid.org/0000-0003-0530-4305 – @essepuntato@scholar.social

Open Science (A.Y. 2024/2025)
Second Cycle Degree in Digital Humanities and Digital Knowledge
Alma Mater Studiorum - Università di Bologna

CC BY

DIPARTIMENTO DI FILOLOGIA CLASSICA E ITALIANISTICA

Discussion

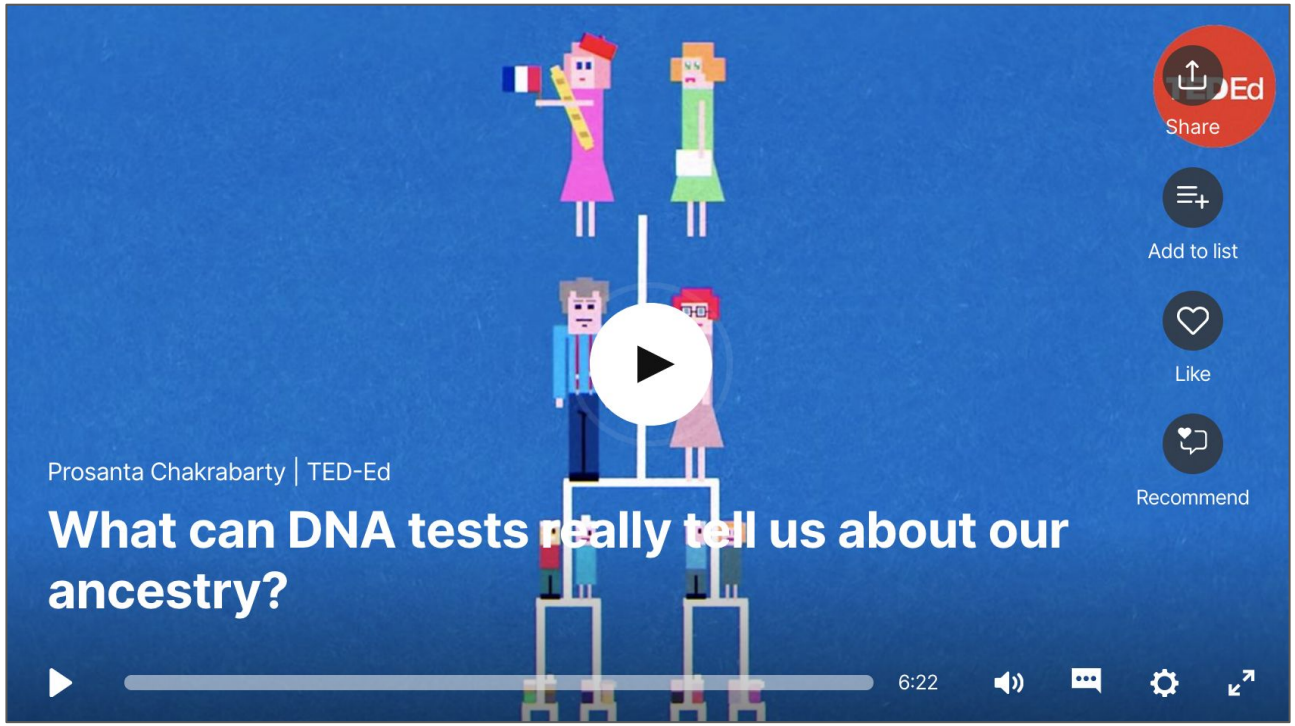
Ma, L. (2022). Metrics and epistemic injustice. Journal of Documentation, 78(7), 392–404. https://doi.org/10.1108/JD-12-2021-0240

Research workflow

We discuss about the methods that are used to assess research, with a special focus on well-know bibliometrics and the importance of enabling transparency in research assessment exercises

Kramer, B., & Bosman, J. (2015, June 18). The good, the efficient and the open—Changing research workflows and the need to move from Open Access to Open Science. CERN Workshop on Innovations in Scholarly Communication (OAI9), University of Geneva, Geneva, Switzerland. https://www.slideshare.net/bmkramer/the-good-the-efficient-and-the-open-oai9

An introductory video about DNA and ancestry



Artrake Studio. (2020, June). What can DNA test really tell us about our ancestry? https://www.ted.com/talks/prosanta_chakrabarty_what_can_dna_tests_really_tell_us_about_our_ancestry

Origin of all sins

“A citation index to science would have [...] articles that had referred to the article in question, together with an indication of whether the citing source was an original article, review, abstract, review article, patent, or translation, and so forth.

[...]

Thus, in the case of a highly significant article, the citation index has a quantitative value, for it may help the historian to measure the influence of the article – that is, its ‘impact factor’”

$$\text{Impact Factor of journal J in year Y} = \frac{\text{Citations in year Y to publications published in J between Y-1 and Y-2}}{\text{Number of publications published in J between Y-1 and Y-2}}$$

Garfield, E. (1955). Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas. Science, 122(3159), 108–111. https://doi.org/10.1126/science.122.3159.108, available in OA at http://garfield.library.upenn.edu/essays/v6p468y1983.pdf

Impact Factor (IF)

The Impact Factor of a journal measures the yearly average number of citations to recent articles (i.e. published in the past two years) in that journal

The IF was developed for helping librarians to select additional source journals to be included in a library catalogue – i.e. it was a tool to help librarians identify journals to purchase

The IF for journal is a calculated by Clarivate Analytics out of the citation data included in Web of Science (WoS), one of the most famous and proprietary citation indexes

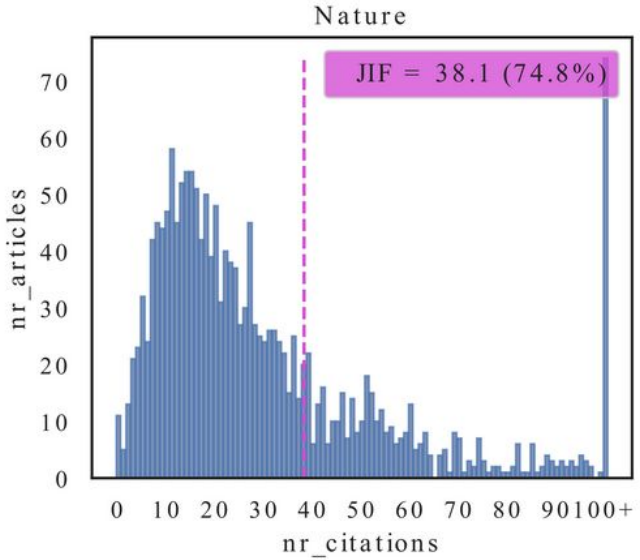
Even if the IF is a measure associated to journals, it is and has been used to measure the impact of individuals (i.e. authors) and institutions

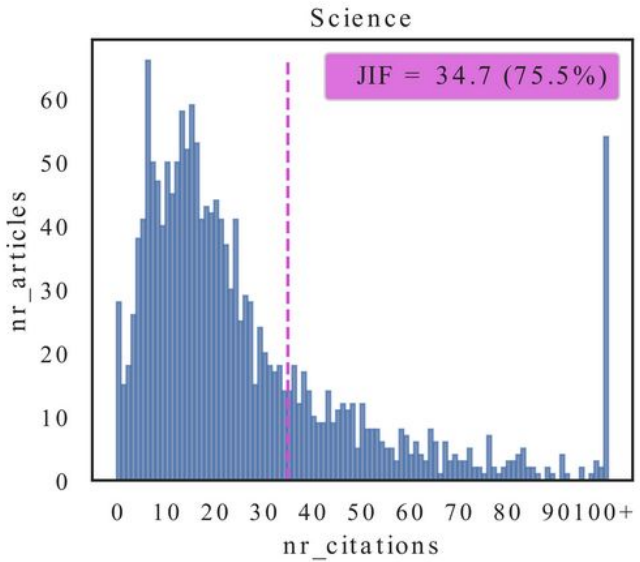
According to Eugene Garfield, its creator, the “use of journal impacts in evaluating individuals has its inherent dangers [since] in an ideal world, evaluators would read each article and make personal judgments”

Garfield, E. (2006). The History and Meaning of the Journal Impact Factor. JAMA, 295(1), 90. https://doi.org/10.1001/jama.295.1.90, available in OA at https://garfield.library.upenn.edu/papers/jamajif2006.pdf

Intrinsic issues of IF

“The distributions are clearly skewed, making the arithmetic mean an inappropriate statistic to use to say anything about individual papers”





Citation distributions follows the Pareto principle: **75% of citations** comes from **25% of articles**

Tennant, J. P., Crane, H., Crick, T., Davila, J., Enkhbayar, A., Havemann, J., Kramer, B., Martin, R., Masuzzo, P., Nobes, A., Rice, C., Rivera-López, B., Ross-Hellauer, T., Sattler, S., Thacker, P. D., & Vanholsbeeck, M. (2019). Ten Hot Topics around Scholarly Publishing. Publications, 7(2), 34. https://doi.org/10.3390/publications7020034

Reproducibility crisis in research assessment

The reproducibility crisis does not affect only science, but also the methods we currently use to assess it

Not only a critique to the ‘recipes’ (i.e. the metrics) used in research assessment, but also to the lack of transparency of the ‘ingredients’ (i.e. the data) used to compute recipes outcomes

It should always be possible to dig deeper into the data and to see what is behind a certain number; this requires scholarly metadata and citation data to be open rather than paywalled and accessible only by paying a fee – as in the case of commercial citations indexes, such as WoS and Scopus, which are the two most adopted databases in research assessment exercises

There is an urgent need of a global community effort in the scholarly domain to put such ingredients to the commons – often they are pure facts, i.e. data that cannot be copyrighted such as citations, and often they are used in metrics-based research assessment

Waltman, L. (2020, September 9). Responsible research assessment requires open scholarly metadata. Workshop on Open Citations and Open Scholarly Metadata 2020, Bologna, Italy. Zenodo. https://doi.org/10.5281/zenodo.4021492

Pushing for open data in research metrics

National Plan For Open Science, https://www.ouvrirlascience.fr/national-plan-for-open-science-4th-july-2018/

"Assessment system for researchers and research institutions must be updated to reflect the principles and practices of open science"

San Francisco Declaration on Research Assessment, https://sfdora.org

"Be open and transparent by providing data and methods used to calculate all metrics"

"Provide the data under a license that allows unrestricted reuse, and provide computational access to data"

Leiden Manifesto for Research Metrics, http://www.leidenmanifesto.org/

"Keep data collection and analytical processes open, transparent and simple"

"Allow those evaluated to verify data and analysis"

Paywalled and close citation data do not comply with the principles above – indeed such data are a threat to transparency, replicability and verifiability of research assessment exercises

A first sketch almost 30 years ago

The first embryonal introduction of open citations is in Robert Cameron's visionary article published in 1997, in which he speculates about the existence of a Universal Citation Database

- Linking every scholarly work ever written
- Freely available over the Internet
- Updated every day
- Comprehensive (no selection of particular venues, all are included)
- All types of publications (from articles to working papers and preprints)
- All publications are equally visible (even if not equally accessible)
- Decentralised

Cameron, R. D. (1997). A Universal Citation Database As a Catalyst For Reform In Scholarly Communication. First Monday, 2(4). https://doi.org/10.5210/fm.v2i4.522

Coalition for Advancing Research Assessment

The Commitments

CoARA

1. Recognise the diversity of contributions to, and careers in, research in accordance with the needs and nature of the research

2. Base research assessment primarily on qualitative evaluation for which peer review is central, supported by responsible use of quantitative indicators

3. Abandon inappropriate uses in research assessment of journal- and publication-based metrics, in particular inappropriate uses of Journal Impact Factor (JIF) and h-index

4. Avoid the use of rankings of research organisations in research assessment

5. Commit resources to reforming research assessment as is needed to achieve the organisational changes committed to

6. Review and develop research assessment criteria, tools and processes

7. Raise awareness of research assessment reform and provide transparent communication, guidance, and training on assessment criteria and processes as well as their use

8. Exchange practices and experiences to enable mutual learning within and beyond the Coalition

9. Communicate progress made on adherence to the Principles and implementation of the Commitments

10. Evaluate practices, criteria and tools based on solid evidence and the state-of-the-art in research on research, and make data openly available for evidence gathering and research

Coalition for Advancing Research Assessment. (2022). Agreement on Reforming Research Assessment [Policy]. European Science Foundation. https://coara.eu/agreement/the-agreement-full-text/

Barcelona Declaration

Research information: metadata relating to the conduct and communication of research, e.g.

- bibliographic metadata
- metadata on research software, research data, samples, and instruments
- information on funding and grants
- information on organizations and research contributors

Open research information is research information that is free to access and free of restrictions on reuse

Barcelona Declaration on Open Research Information. (2024). [Policy]. https://barcelona-declaration.org/

Mandatory text to read for the next lecture

Goudarzi, S., & Dunks, R. (2023). Defining Open Scholarly Infrastructure: A Review of Relevant Literature (Version 2). Invest in Open Infrastructure. https://doi.org/10.5281/zenodo.8064102

End

Open Metrics

Silvio Peroni
silvio.peroni@unibo.it – https://orcid.org/0000-0003-0530-4305 – @essepuntato@scholar.social

Open Science (A.Y. 2024/2025)
Second Cycle Degree in Digital Humanities and Digital Knowledge
Alma Mater Studiorum - Università di Bologna

CC BY

DIPARTIMENTO DI FILOLOGIA CLASSICA E ITALIANISTICA