Data Citation Metrics: Are You Ready for Us?

CHORUS Forum: Are We Ready for Data Citation Metrics?
2021-10-28

Amy L. Nurnberger (she/hers)
MIT Libraries, Program Head, Data Management Services
Interim Head, Data & Specialized Services
Teachers College, Columbia University, Assistant Adjunct Professor in Learning Analytics

@ANurnberger | ORCID: 0000-0002-5931-072X

*The views expressed herein are those of the presenter and not necessarily her institution.*

This work is licensed under a Creative Commons Attribution 4.0 International License.
Step #1: Data Citation
We’ve been waiting...

1985, NRC: “Recommendation 12. Journals should require full credit and appropriate citations to original data collections…” [1]

2007, Altman & King: “A critical component of this community is the common language of and the universal standards for scholarly citation, credit attribution, and the location and retrieval of articles and books. We propose a similar universal standard for citing quantitative data…” [2]

2012, Mooney & Newton: “the majority of journal articles failed to include an adequate citation to data used in secondary analysis studies” [3]

Just the other day, Data professional: “…require a PID and indexing if you want to change anything.” [4]
What we have:

Including data availability statement (PLOS) [5]

*Necessary, but insufficient*

Appropriate citation of publicly available data sets (ACS) [6]

*Closer!*

What we need:

Joint Declaration of Data Citation Principles (force11) [7]

*That’s the one!*
Step #2: Metrics…we have a problem

- misuse, misrepresentation, misinterpretation
- biases
  - inclusion and representation of diverse communities around the world in these metrics?
- inequities
  - accounting for geographic and social inequities in underlying infrastructures that support deployment of metrics?
Let’s do better

Clearly lay out data citation metrics’ role and meaning for:

- Quality and impact of research products
- Incentives, rewards & recognition for researchers and institutions
- Inequities and biases of research and scholarship systems
Quality & impact

“In treating data as a first-class research product, we need clear measures that describe the quality and impact of a shared data product.”
Incentives, rewards, & recognition

Data citation metrics are necessary, but to what purpose?

How do we
• incentivize
• reward
• recognize
behaviors that contribute to healthy, productive, sustainable scholarship systems?
Inequities and biases

Let’s not replicate those that already exist in:
- evaluation and credit systems
- (and between) research cultures
- costs
- intellectual property
- ethical perspectives
- resources and infrastructure

[8]
Before we get to data citation metrics

- normative practice of data citation
- shared understandings of
  - what data citation metrics mean
  - what data citation metrics incentivize
  - what systems data citation metrics support
- infrastructure to support desired outcomes
Sources & Resources

Please respect the copyrights and licenses of the creators


Print. By Kurt Komoda https://www.flickr.com/photos/39573281@N06/7042782421 CC BY-NC-ND 2.0