

Bibliometrics Working Group

Co-chairs: Sarah Callaghan
(sarah.callaghan@stfc.ac.uk)
Kerstin Lehnert
(lehnert@Ideo.columbia.edu)



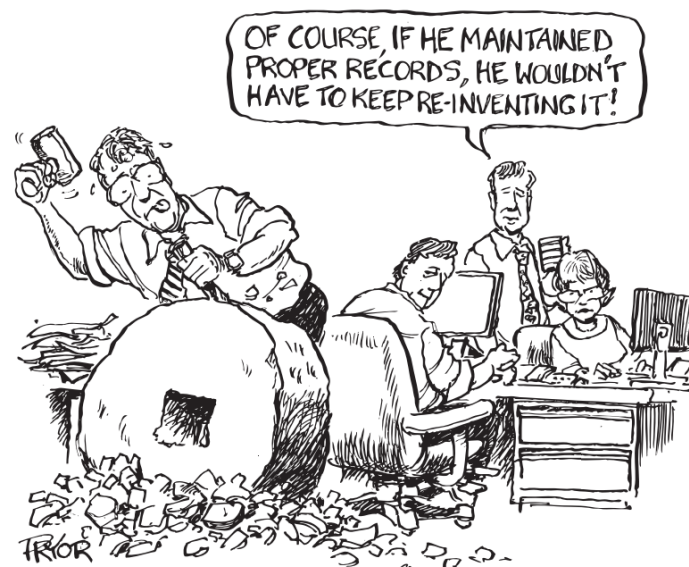
What are Bibliometrics?

- Quantitative measures to assess and measure impact and quality of research and researchers,
- by tracking and recording access and citation of scientific publication.
- Inform & influence the advancement of academic careers, investments in research and scientific journals.

The basic idea of bibliometrics is to evaluate the attention scientific publications receive within the scientific community.

Bibliometrics: Examples

- ISI Science Citation Index – *individual articles*
- h-index (= Hirsch number) – *researchers*
- Impact Factor - *journals*



<http://www.naa.gov.au/records-management/capability-development/keep-the-knowledge/index.aspx>

Bibliometrics for Data

- Quantitative measures for the use, utility, and impact of data.
- **Rationale:** Essential for a culture change toward full appreciation and recognition of data as a part of the scholarly record.
 - raise the value of data acquisition, curation, and sharing;
 - encourage more and better data citation;
 - augment the overall availability and quality of research data;



Use of existing metrics?

- Data citation is not standard practice in the research community - though it is getting more common
- Though Principles for Data Citation (Joint Declaration on Data Citation Principles, CODATA TG on DataCitation) exist, work is still being done on how to implement them (Force11)
- Data repositories are mostly not ready for the data publication concept and have not implemented formal procedures.
- There are unresolved issues of granularity, versioning, unique identification, metadata, review.

Goals

- **“conceptualize data metrics and corresponding services to overcome barriers”**
 - Summarize current and emerging *data citation practices* of all stakeholders (journals, data centers, funders, societies);
 - Understand and articulate *necessary organizational and cultural changes* in the scholarly publishing system needed to foster proper attribution of data sets;
 - Evaluate and report on *possible models* how data use and citations can be successfully tracked and measured based on existing and emerging approaches (e.g. altmetrics);
 - Identify and report on *possible barriers* for the implementation and adoption of data citation and data metrics solutions.

Members

Note that the membership of this WG is open, and members are actively seeking participation from members outside the US/EU

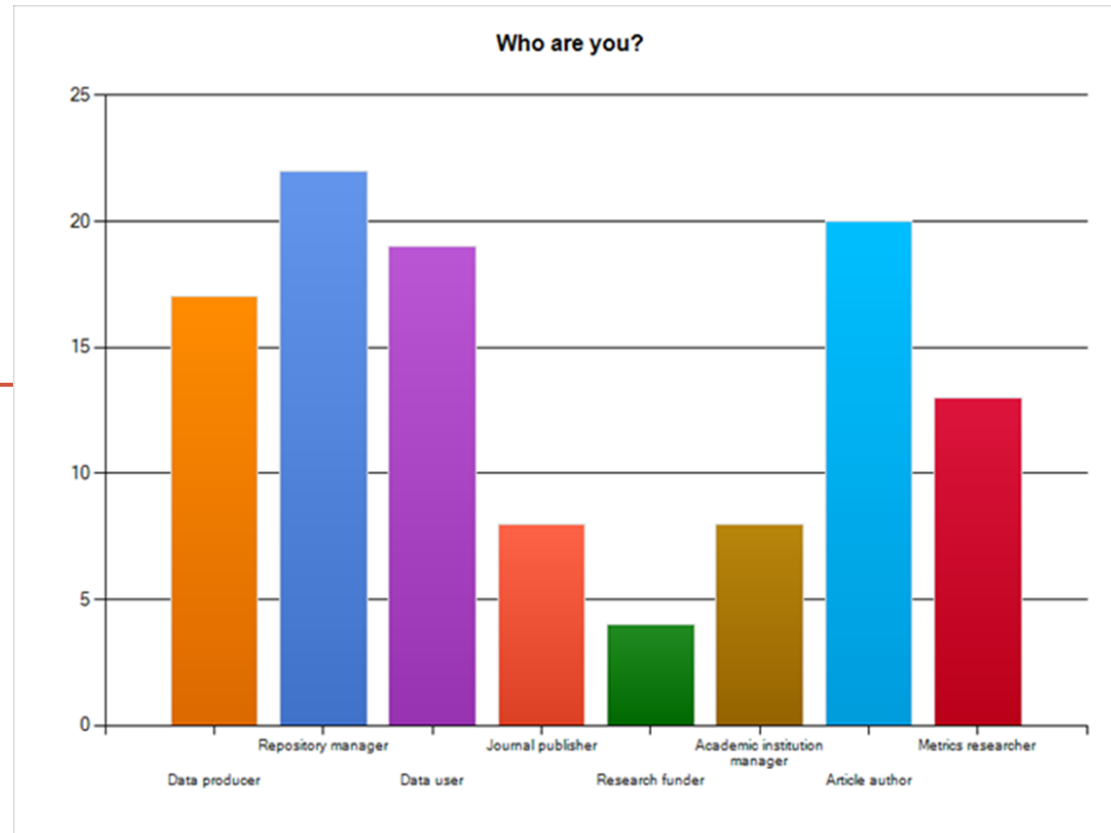
- Kerstin Lehnert (US, IEDA, WDS) [**CO-CHAIR**]
- Sarah Callaghan (UK, BADC) [**CO-CHAIR** - also co-chair of the CODATA TG on data citation]
- Jan Brase (Germany, DataCite) [also co-chair of the CODATA TG on data citation]
- Ross Cameron (The Netherlands, Scopus)
- Cyndy Chandler (US, Woods Hole Oceanographic Institution)
- Ingeborg Meijer (The Netherlands, University of Leiden)
- Fiona Murphy (UK, Wiley-Blackwell)
- Lyubomir Penev (Bulgaria, Pensoft Publishers)
- Fiona Nielsen (UK, DNAdigest.org)
- Nigel Robinson (UK, Thomson Reuters)
- Mary Vardigan (USA, ICPSR)
- Jochen Schirrwagen (Germany, Universität Bielefeld)

Preliminary survey results

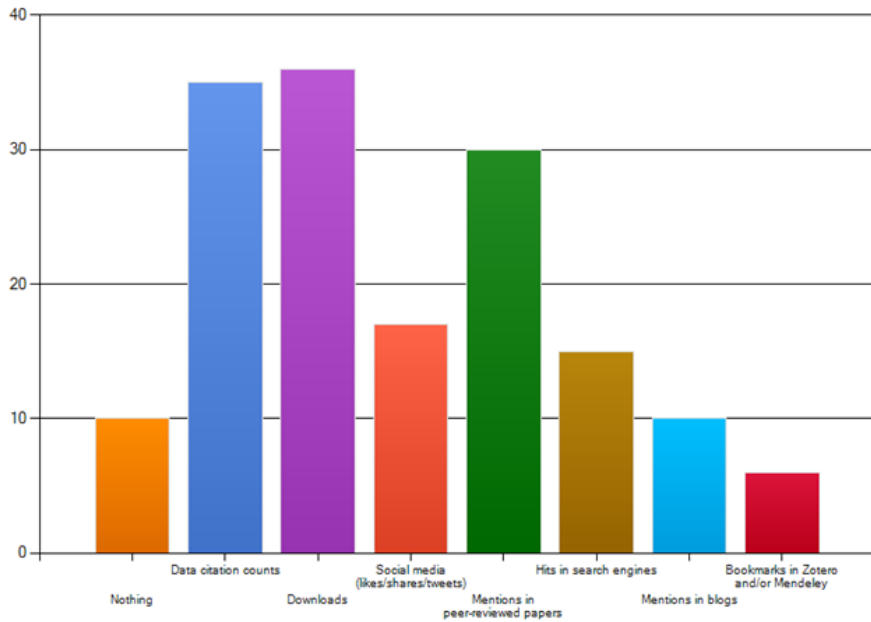
- Launched 3rd September
- As of 17th September – 63 responses
- 100% completion
- Survey link still live

https://www.surveymonkey.com/s/RDA_bibliometrics_data

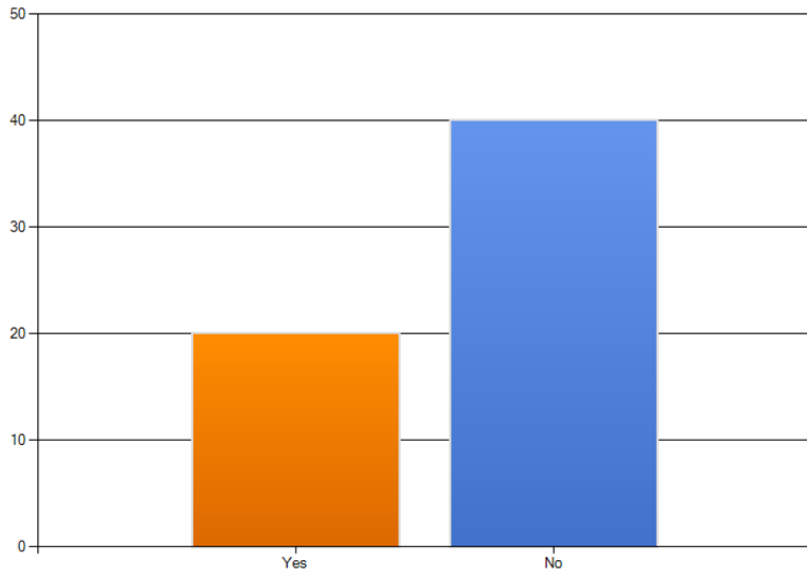
Science 3
Earth sciences 16
Physics 4
Scientometrics and bibliometrics 4
Engineering 2
Chemistry 1
Biology (inc. zoology) 2
STEM 1
Medicine & biomedical research 8
Energy 1
Admin for research 2
Computer science 4
Social science, policy and economics 4
Librarian and digital curation 11



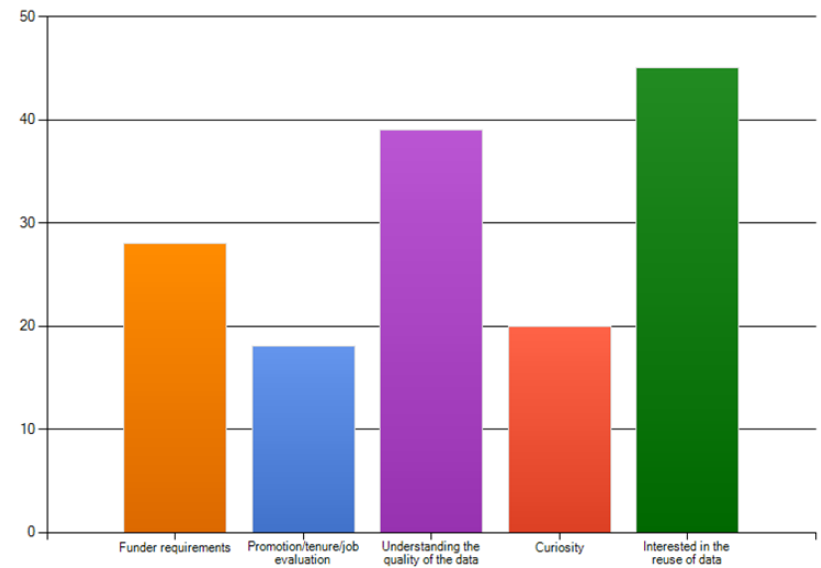
What do you currently use to evaluate the impact of data?



Are the methods you use to evaluate impact adequate for your needs?



Why do you want to evaluate the impact of data?



Future and missing

• In the future, what would you like to use to evaluate the impact of data?

• Most popular suggestions:

- Data citations
- Actual use in professional practice
- Download statistics
- Mentions in social media
- DOIs/PIDs
- Altmetrics
- Well regarded indicators

• Also pleas for:

- Easy to use and set up
- Radically different tools
- Whatever tool can provide reliable information
- Best estimate of societal benefit in \$\$ terms

What is currently missing and/or needs to be created for bibliometrics for data to become widely used?

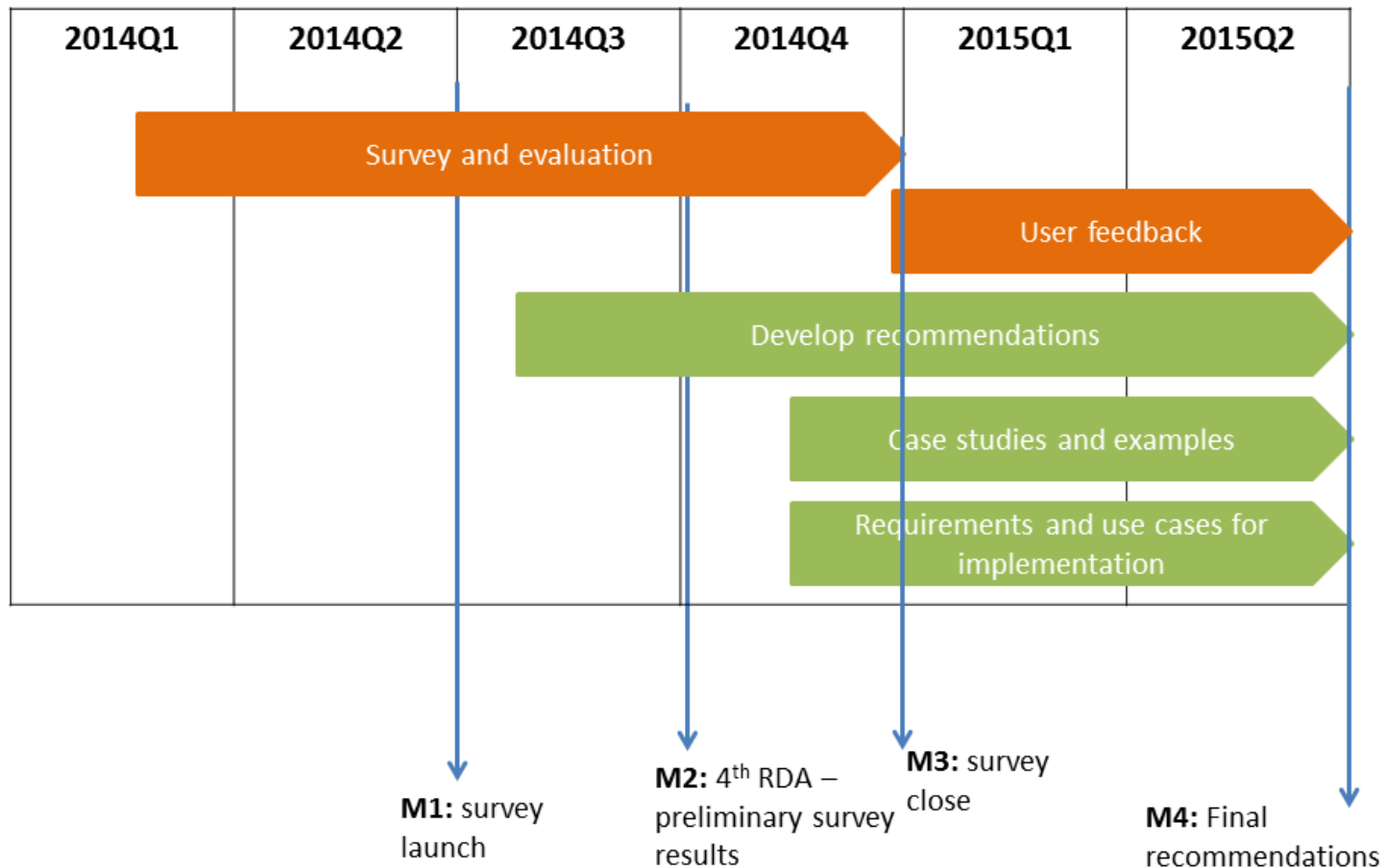
Most popular suggestions:

- Culture change!
- Principles and standards for consistent practice (and enforcement of these)
- Use of PIDs
- Mature tools for data citation, publishing, discovery and impact analysis
- Openness in papers and patents

Also:

- Research on what current metrics actually measure
- Infrastructure
- Free apps

Time plan



Feedback

- Are we missing relevant stakeholders?
- Are we missing relevant initiatives to engage with?
- Are we focusing on the essential aspects of bibliometrics and data citation?
- How can we improve our approach and outcomes?

- Please join us!

- Contact: sarah.callaghan@stfc.ac.uk,
lehnert@LDEO.COLUMBIA.EDU