Open Metrics as Part of Persistent Identifier Infrastructure

Martin Fenner
DataCite Technical Director
https://orcid.org/0000-0003-1419-2405

OPEN METRICS

Open Metrics are essential for Open Science.

PERSISTENT IDENTIFIER

Essential to collect and aggregate scholarly metrics.

AS PART OF

INFRASTRUCTURE

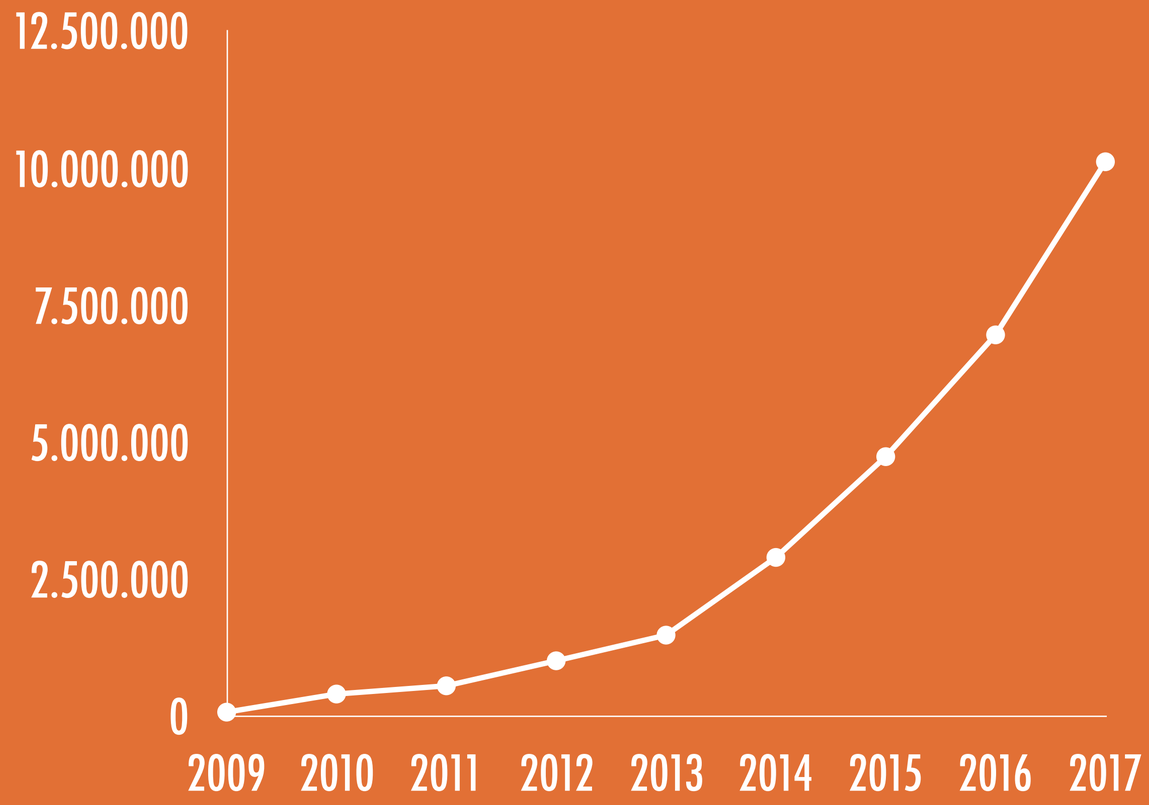
Collecting and providing metrics data should become part of scholarly infrastructure.

PERSISTENT IDENTIFIERS

DataCite is a DOI Registration Agency.

As of May 13, 2018 our users have registered 11.403,446 DOIs.

37% of these DOIs are for research data and 28% for text documents.



ROLE OF DOIS IN SCHOLARLY METRICS

CITATIONS

DOIs are essential.

Most (but not all) citations use DOIs or identifiers that can be mapped to DOIs.

USAGE

DOIs are helpful.

Globally unique identifiers can aggregate usage data from different locations.

ALTMETRICS

DOIs are sometimes not enough.

Mendeley and Wikipedia understand DOIs. Many other altmetrics sources do not.

INFRASTRUCTURE

/ˈɪnfrəstrʌktʃə/

The basic physical and organizational structures and facilities needed for the operation of a society or enterprise.

Ubiquitous
Accessible
Governance*
Sustainability*

*Principles for Open Scholarly Infrastructures https://doi.org/10.6084/m9.figshare.1314859

OPEN SCIENCE

A. Open Access
B. Open Data
C. Open Reproducible Research
D. Open Science Evaluation
    A. Open Metrics
    B. Open Peer Review

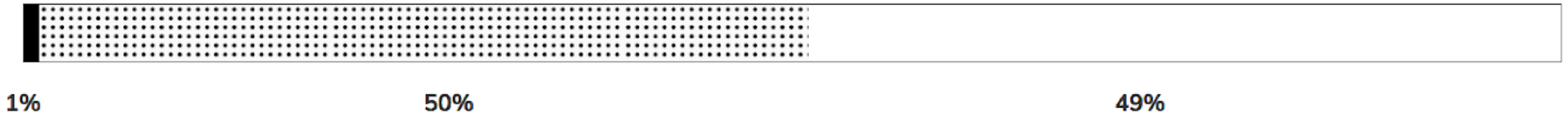
https://www.fosteropenscience.eu/taxonomy/term/4

INITIATIVE FOR OPEN CITATIONS (I4OC)

April 2018: out of 39,314,587 works with references, 51.3% are open (20,172,878). Out of 1,103,882,760 total references, 53.0% are open (584,539,432)

Mark Patterson, 25 April 2018, https://twitter.com/marknpatterson/status/989183467756904448

How many citations are open today?



RESPONSIBLE METRICS

A. **Robustness** – basing metrics on the best possible data in terms of accuracy and scope
B. **Humility** – recognising that quantitative evaluation should support, but not supplant, qualitative, expert assessment
C. **Transparency** –that those being evaluated can test and verify the results
D. **Diversity** – accounting for variation by research field, and using a range of indicators to reflect and support a plurality of research and researcher career paths across the system
E. **Reflexivity** – recognising and anticipating the systemic and potential effects of indicators, and updating them in response

http://www.universitiesuk.ac.uk/policy-and-analysis/Pages/forum-for-responsible-research-metrics.aspx

A VERY SHORT HISTORY OF ALTMETRICS

2010
Altmetrics Manifesto

2016
NISO Altmetrics Data Quality Code of Conduct

Transparency
Replicability
Accuracy

?

Evaluation

Infrastructure ≠ Evaluation

INFRASTRUCTURE

Raw data.
The data provided by infrastructure should be the same for everyone.

DataCite is involved.

EVALUATION

Adds meaning to raw data.
There is more than one way to evaluate the data and this can change over time.

DataCite is not involved.

EVENT DATA

Crossref and DataCite are jointly developing the Event Data service that provides a record of instances where research has been bookmarked, linked, liked, shared, referenced, commented on etc, beyond publisher platforms. For example, when datasets are linked to articles, articles are mentioned on social media or referenced in Wikipedia.

https://www.eventdata.crossref.org/guide/

EVENT DATA

CROSSREF EVENT DATA

Initial focus on altmetrics, in particular Twitter and Wikipedia. Available as open beta, will launch later this year.

DATACITE EVENT DATA

Initial focus is on data citations and data usage stats. Current work as part of the Make Data Count Project, funded 2017-2019 by the Alfred P. Sloan Foundation.

EVENT DATA HUB

Stores all events that involve a Crossref and/or DataCite DOI.

CROSSREF SOURCES

Crossref to DataCite references
Twitter
Wordpress.com
Reddit
StackExchange
Hypothes.is

THIRD-PARTY SOURCES

Wikipedia
F1000

DATACITE SOURCES

DataCite to Crossref references
Research data usage stats

SCHOLIX API

Events in RDA Scholix format.

CROSSREF EVENT DATA QUERY API

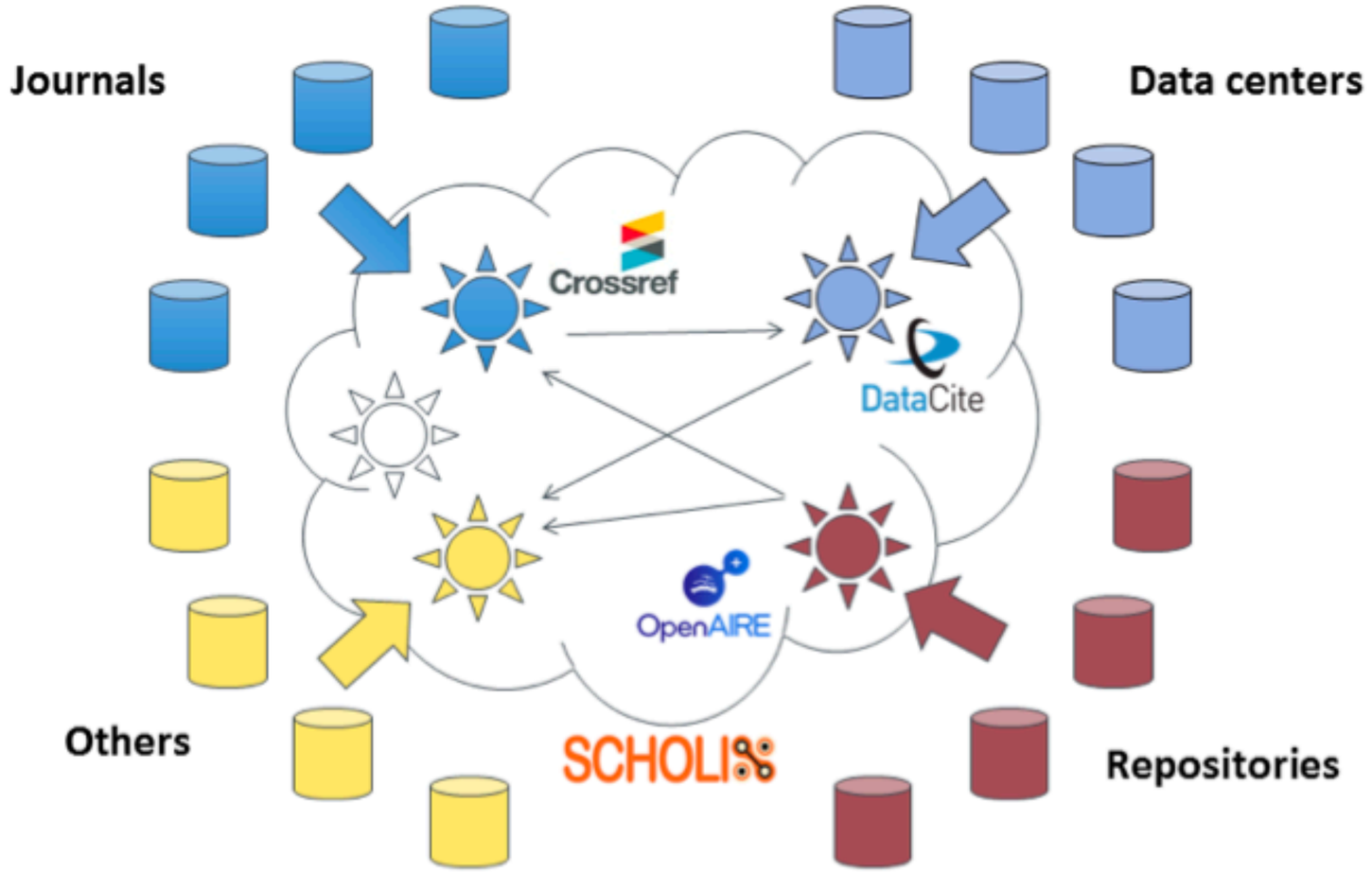
Retrieve events using an API. In open beta.

DATACITE EVENT DATA QUERY API

Work in progress, currently only for data usage events.
Will integrate with the DataCite REST API for metadata.

SCHOLIX

Research Data Alliance (RDA) working group to establish a high level interoperability framework for exchanging information about the links between scholarly literature and data.



MAKE DATA COUNT

Develop and deploy the social and technical infrastructure necessary to elevate data to a first-class research output.

A. Code of Practice for Research Data Usage (published March 2018, aligned with COUNTER release 5)
B. Develop a hub to collect and provide usage reports in standard format (May 2018, SUSHI format)
C. Hub to integrate usage stats with data citations and other metrics
D. Advocacy and training regarding the importance and use of data-level metrics

Alfred P. Sloan FOUNDATION

THANK YOU