

HELMHOLTZ Open Science

Introduction to Scholix – Infrastructure to Collect and Report Data Citations Martin Fenner DataCite Technical Director

> 17 June 2019 25 June 2019

The State of Data Citation



Policy	Data as first-class scholarly output	Good
Infrastructure	Support for data citation workflow	Evolving
Bibliometrics	Understand the impact of data citations	Beginning
Culture Change	Changes in researcher behaviour	Very early



Understand data citations and credit **Make Data Count**

Aggregate and show data usage and citations Make Data Count Collect and distribute data citations RDA Scholix Event Data Publish article and data DCIP







RDA/WDS Scholarly Link Exchange Working Group

research data sharing without barriers rd-alliance.org

Linking data with literature: what's the problem?

Linking Research Data with the Literature is of great value, yet current solutions are not realizing the potential

What is the problem?

- 1. Many disconnected sources (publishers, data centers, repositories, infrastructure providers, ...)
- 2. Heterogeneity of practices, for example:
 - Different PID systems (DOI, accession numbers)
 - Different ways of referencing data (formal citations, in-text references, ...)
 - Different moments of citing data (at publication, post publication, ...)





- A schema for standardizing the exchange of *scholarly link* information between scholarly infrastructure providers
 - Information Model for scholarly links representation
 - Recommendation and provision of exchange formats and protocols

Scholiss See also http://www.scholix.org/guidelines

- Scholix helps with the information exchange between infrastructure providers.
- Scholix does not change how data repositories working with DataCite or publishers working with Crossref provide data citations via DOI metadata.
- The Scholix community encourages data repositories and publishers to submit data citations to DataCite and Crossref.



Recap Scholix: connecting the dots

Past: disconnected sources using heterogeneity of practices

Future: standard set of guidelines for exposing and consuming links, supported by hubs



research data sharing without barriers rd-alliance.org

RESEARCH DATA ALLIANC

WORLD DATA SYSTEM

Scholix benefits see: www.Scholix.org







For data repositories and journal publishers

- increase their visibility and usage
- improve the user experience
- More scalable and robust

For research institutes, bibliographic service providers, and funding bodies

- Make data count
- track datasets and publications within common and comprehensive framework

For researchers:

- Easier finding and accessing
- track long-term impact of their data
- additional incentives to share data.







Information model: properties

Link Information Package

Link Publication Date (1) Link Provider (1..N) Relationship Type (1) License URL (0..1)

Source Object

Object Identifier (1) Object Type (1)

Object Title (0..1) Object Publisher (0..1) Object Creator (0..N) Object Publication Date (0..1)

Target Object

Object Identifier (1) Object Type (1) Object Title (0..1) Object Publisher (0..1) Object Creator (0..N) Object Publication Date (0..1)

Scholix in practice: the Hubs



Event Data

https://support.datacite.org/docs/eventdata-guide



Scholexplorer

https://scholexplorer.openaire.eu/



research data sharing without barriers rd-alliance.org

Numbers



The joint Crossref/DataCite Event Data service holds the following information regarding literature/data links as of 17 June 2019:





Further adoption with focus on publishers
Support and training

- Coordination between hubs
- User interfaces



User Interfaces



research data sharing without barriers rd-alliance.org

RESEARCH DATA ALLIANCI

User Interfaces

Display of citations and other connections in PID Graph using a Jupyter notebook and the DataCite GraphQL API.

https://github.com/datacite/notebooks/blob/master/pidgraph/r-grant-publications.ipynb



research data sharing without barriers rd-alliance.org