DataCite

CONNECTING RESEARCH,
IDENTIFYING KNOWLEDGE

Showing Data Citation & Reuse in the PARSEC Project

Rorie Edmunds, Matt Buys, Helena Cousijn, Shelley Stall

Wednesday, 19 October 2022
eResearch Australasia 2022

@datacite

CC BY

BELMONT FORUM

PARSEC

PARSEC is a project sponsored by the Belmont Forum as part of its Collaborative Research Action on Science-Driven e-Infrastructures Innovation, with funding from FAPESP, ANR, JST and NSF, with collaborators from Australia, and support from the synthesis centre CESAB of the French Foundation for Research on Biodiversity

FRB FONDATION POUR LA RECHERCHE SUR LA BIODIVERSITÉ

CESAB CENTRE FOR THE SYNTHESIS AND ANALYSIS OF BIODIVERSITY

AGENCE NATIONALE DE LA RECHERCHE ANR

FAPESP FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO

NSF

JST

Participating Countries

- BRAZIL: University of São Paulo – The São Paulo Research Foundation, FAPESP (P. Pizzigatti Corrêa) plus postdoc and technical support (FAPESP)
- FRANCE: Foundation for Research on Biodiversity (FRB), University of Toulouse III – Paul Sabatier – ANR (N. Mouquet), ERINHA (FAIR DMP advisor - R. DAVID)
- JAPAN: National Institute of Information & Communications Technology (NICT), Research Institute for Humanity and Nature (RIHN) – JST (Y. Murayama)
- USA: American Geophysical Union (AGU) – NSF (S. Stall)

PARSEC Objectives

Building New Tools for Data Sharing and Reuse through a Transnational Investigation of the Socioeconomic Impacts of Protected Areas

- Provides unique opportunity for data and synthesis scientists to collaborate and exchange in real-time towards goal of improving research outcomes, data sharing, and data reuse
- Helps pioneer new scientific and data science technologies aimed at improving management and conservation of global biodiversity

Two Interactive Teams

- Synthesis Science Team
  - Employing AI techniques to analyse satellite images and socio-economic information
  - Better predict and mitigate the effect of creation of Protected Areas that potentially threaten the livelihoods and health of local (indigenous) communities
  - Depends significantly on availability of high-quality, spatially dispersed, multidisciplinary, time-series data

- Data Science Team
  - Environmental data management professionals representative of data communities (RDA, ESIP), societies (AGU), and infrastructures for data attribution (DataCite and ORCID)
  - Developing leading practices on data citation, attribution, credit, and reuse, using Synthesis Science team as a case study of the wider scientific community
  - Providing a review of best practices for data management and stewardship to optimize data access and reuse
  - Implementing new tools to better track data usage and reuse

Data Strand

- Data science experts partner with the Synthesis Science team to build tools and processes for better data and software management that integrate into the research lifecycle and shared with the wider research community
- Recently, important papers and books have been published that support the goal of helping researchers to incorporate data and software management during the whole lifecycle
- Need to develop methods and foster cultural change to implement in practice across scientific communities. Main challenges are thus technological and cultural
- Share such considerations and work collaboratively on incremental improvements that benefit all

Outcomes

- Determine influence of natural protected areas on consumption expenditure and asset health of rural communities in a range of countries
- Recommendations for research data workflow and skills for research teams
- Tools for researchers to view how data they have deposited are used and cited, improving the number of citations to datasets and better attributing them to the data creator

Contribute to promotion of FAIR, CARE, and TRUST through authentic examples from an active research project

Increase Properly Cited Datasets, Provide Credit and Attribution, and Accurately Track Data Reuse

A. Robustly connect identifiers across papers, people, and repositories
B. Outreach and adoption campaigns on importance of PIDs and infrastructure
C. Promote and extend data usage metrics and citations generated by RDA WGs. By requiring data be cited from a trusted, community-accepted repository, the value that repositories provide to scientists and the research infrastructure can be better measured. Data sharing through citation increases the likelihood of data discovery and reuse
D. Provide guidance to Synthesis Science team and project teams from partners to optimize data reuse
E. Promote integrated guidance that will address recommendations to improve processes of data sharing, reuse, credit and reward for researchers and repositories

C. Promote and Extend Data Usage Metrics...

- Deliverables
  - Researcher profiles that make information about reuse of their data easily accessible
  - Widget displaying research data usage metrics for researchers that can be easily implemented by repositories
  - User and Usability Testing of researcher profile and widget

  Developer: DataCite, leading provider of DOI registration services for data
- Best Practices Communities and User/Usability Testing Support
  - RDA: Data Usage Metrics WG, Scholix, SHARC; ESS IG, Domain Repository IG
  - ESIP: Data Stewardship, Digital Object Citation, Usability Cluster, Repository Community
  - COPDESS: Enabling FAIR Data Commitment Statement
- Adoption Required by
  - Repositories: Provide standard usage criteria per MakeDataCount criteria
  - Publishers: Require that data be cited
  - Organizations providing researcher profiles (institution, societies): Embed widget
  - Researchers: Cite data in research; view data usage

This goal benefits all researchers in any research field internationally

About DataCite

- Our vision at DataCite is **Connecting Research, Identifying Knowledge**
- Global non-profit membership organization – community led
- Work with over 1200 research organizations (including research funders and facilities) from over 50 countries to provide the means to create, find, cite, connect, and use research
- 2700+ repositories worldwide provide DOIs for research outputs and resources

2700+
Repositories

280+
Members

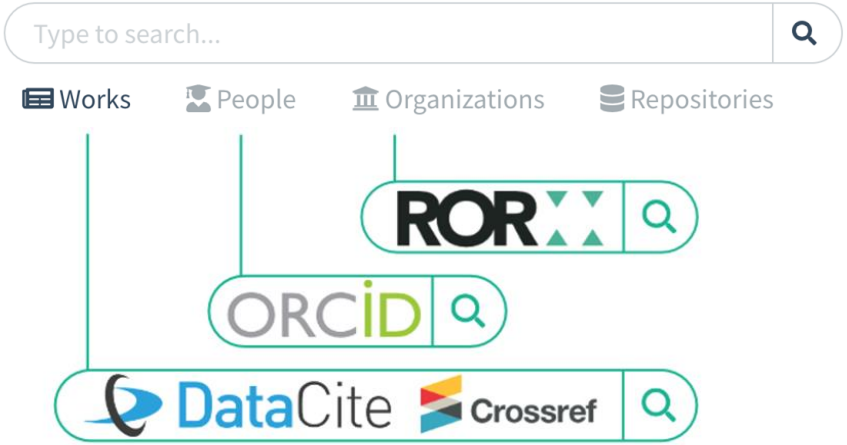
50
Countries

37m+
DOIs

1200+
Organizations

Find & Connect Research

Find Research with DataCite Commons



- Describes works, people, and organizations, and their connections, and enables users to search for them
- Replaces DataCite Search and the Repository Search, which are being retired
- Identified by persistent PIDs:
  - Works (DOI)
  - People (ORCID ID)
  - Organizations (ROR ID)
- Have standard metadata that describe them and connections to one other

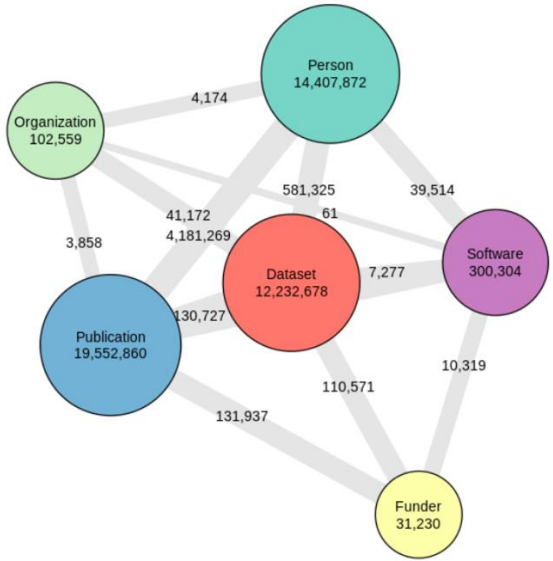
https://commons.datacite.org

Find & Connect Research

- The works, people, and organizations in DataCite Commons form the PID Graph—the resources identified by PIDs and the connections among them
- Extends the Event Data Service
- DataCite Commons thus provides a public web search interface to the PID Graph
- From the statistics page in DataCite Commons you can find a live summary of the data involved in the PID Graph: https://commons.datacite.org/statistics

PID Graph

Number of nodes and connections (June 2022)



Researcher Profiles

DataCite Researcher Profiles provide insight into data-level metrics by researchers

- How it Works
  - A researcher that has openly published research datasets can find out many times these published datasets have been viewed, downloaded, and cited by visiting their DataCite Researcher Profile page
  - Example Researcher Profile: https://commons.datacite.org/orcid.org/0000-0003-1419-2405

- How to Access
  - To access Research Profiles navigate to: https://commons.datacite.org/orcid.org/+RESEARCHER_ORCID_ID

https://support.datacite.org/docs/datacite-researcher-profiles

Profile Example

DataCite Commons

Type to search... | Pages | Support | K. J. Garza

Works | People | Organizations

https://orcid.org/0000-0003-1419-2405

Martin Fenner

Martin Hellmut Fenner

Martin Fenner is the Founder of Front Matter, an organization that provides custom software development and hosting for the scholarly community. Martin was the Technical Director of the DOI registration agency DataCite from 2015 until July 2021. Martin helped launch and maintain the infrastructure for the Research Organization Registry (ROR) from 2018 until 2021. From 2012 to 2015 he was the technical lead for the PLOS Article-Level Metrics project. He was a board member of the ORCID board from 2010 to 2012. Martin has a medical degree from the Free University of Berlin and is a Board-certified medical oncologist.

Links
Twitter

Other Identifiers
Scopus Author ID: 7006600825
ISNI: 000000035060549X
GitHub: mfenner

Other Profiles
ORCID
Impactstory
Europe PMC

Germany

https://orcid.org/0000-0003-1419-2405

Share

Email
Twitter
Facebook

Employment

Front Matter
Founder
Since August 2021

Founder
Since August 2021

DataCite
Technical Director
August 2015 to July 2021

Hannover Medical School
Clinical Fellow in Hematology and Oncology
November 2005 to May 2017

Public Library of Science
Technical lead article-level metrics project (contractor)
April 2012 to July 2015

Charité Universitätsmedizin Berlin
Resident in Internal Medicine
September 1998 to October 2005

Aggregated Citations, Views and Downloads

49 Citations | 379 Views

Accessibility Achievements

53% of the researcher's associated DOIs have metadata with rights as CC-BY, CC0 or public domain license.

251 Works

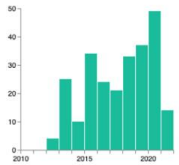
Filter Works

Type to search...

Publication Year

2021	14
2020	49
2019	37
2018	33
2017	21
2016	24
2015	34
2014	10
2013	25
2012	4

Publication Year



Work Type

License

CrowdoMeter Tweets

Martin Fenner & Euan Adie
Dataset published 2012 in figshare Academic Research System

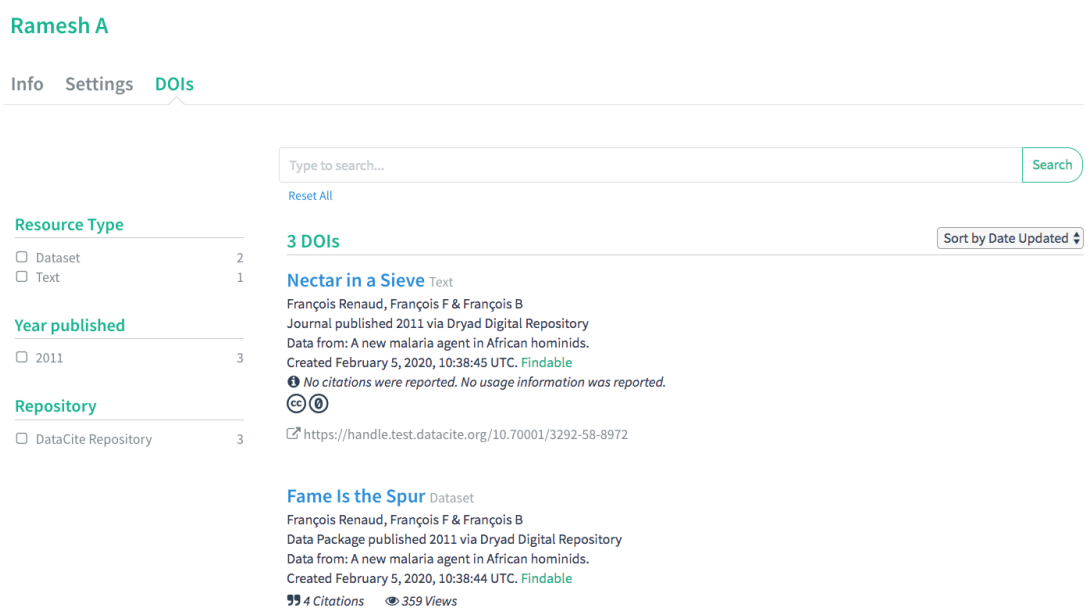
467 tweets that contain links to scholarly content, collected in October 2011. Analyzed in the CrowdoMeter project.

DOI registered June 11, 2012 via DataCite.

Dataset

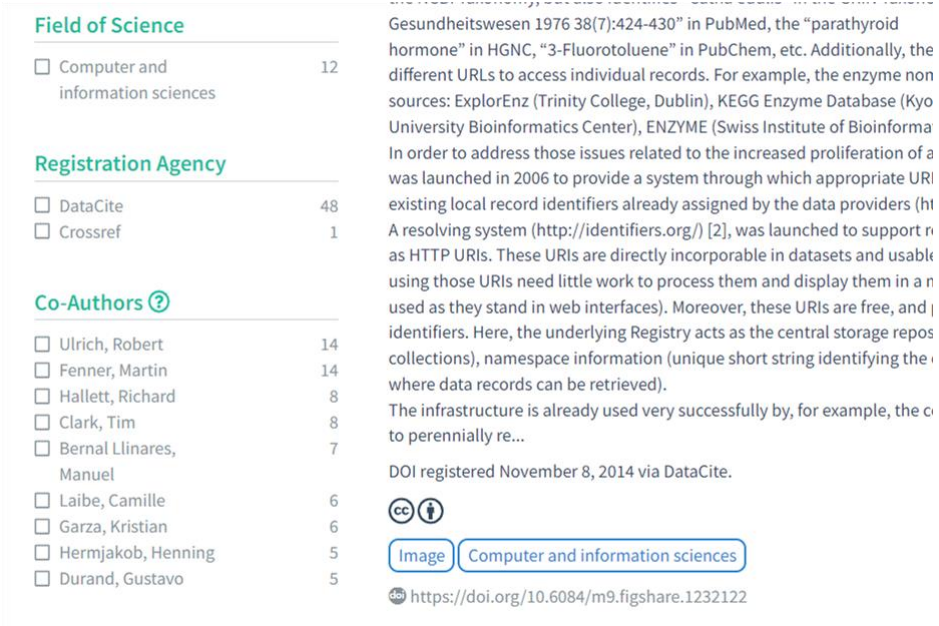
DOIs Tab

- The DOIs tab lists of all (only) the DOIs in which the researcher's ORCID iD appears in the metadata
- Can filter by Resource Type, Year Published, and Repository
- Can sort DOIs by Date Updated and Date Created
- You must be logged in through DataCite Profiles to access the DOI tab



Co-author List

- DataCite Commons also enables visualization of a researcher's co-authors
- Analyze the network of a particular researcher's collaborations and explore their profiles
- By visiting your research profile page you can see your top co-authors (N.B. the list includes only co-authors with an ORCID iD)
- Example: https://commons.datacite.org/orcid.org/0000-0002-5355-2576



Data Metrics Badge (Widget)

- Enables display in your repository of usage and citations information
- Provides open and simple way to display numbers of citations and usage statistics that research datasets and resources have received
- Easily embedded on repository landing page and will display usage and citations metrics retrieved from DataCite services
- Usage statistics processed according to Code of Practice for Research Data
- Citations stats are collected via the DataCite and Crossref EventData Service

https://support.datacite.org/docs/displaying-usage-and-citations-in-your-repository

Embed Data Metrics Badge

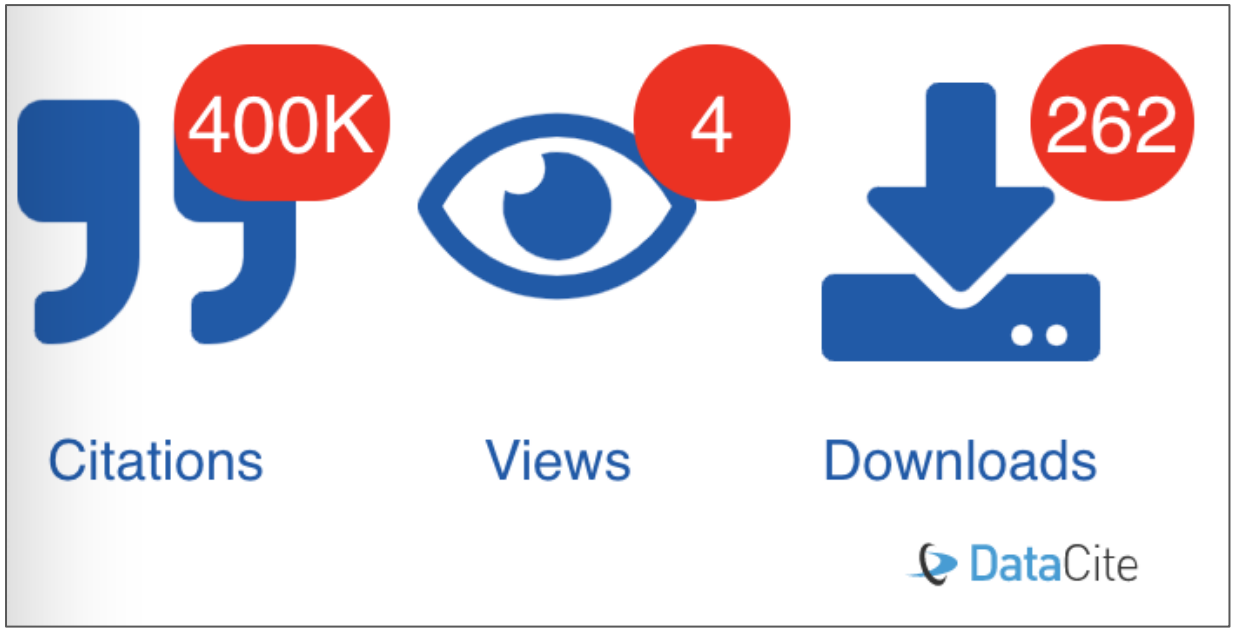
- To display, add this code in the <body> of landing pages and modify the DOI property in the <data-metrics-badge> according to your landing page's DOI

```
HTML
<body>
 <script src="https://unpkg.com/vue@^2/dist/vue.min.js"></script>
  <script
src="https://unpkg.com/@webcomponents/webcomponentsjs@2.0.0/webcomponents-loader.js"></script>
 <script src="https://unpkg.com/data-metrics-badge/dist/data-metrics-badge.min.js"></script>
 <data-metrics-badge doi="10.7272/q6g15xs4" display="regular"></data-metrics-badge>
</body>
```

- Choose between two different types of badge by changing the `display` property to either: `small` or `regular`

DataCite

Different Badges




NICT
国立研究開発法人
情報通信研究機構
National Institute of Information and Communications Technology
RIHN
Inter-University Research Institute Corporation
National Institutes for the Humanities
Research Institute for Humanity and Nature
TOKYO METROPOLITAN UNIVERSITY
UNIVERSITÉ TOULOUSE III PAUL SABATIER
UNIVERSITE MONTPELLIER
USP
SciELO
IRD
Institut de Recherche pour le Développement
cnrs
marbec
marine biodiversity exploitation & conservation
Thank You
ADVANCING EARTH AND SPACE SCIENCE
RESEARCH DATA ALLIANCE
BELMONT FORUM
The Nature Conservancy
Powell Center
NSF
WORLD DATA SYSTEM
SIP
PARSEC
FRB
FONDATION POUR LA RECHERCHE SUR LA BIODIVERSITÉ
CESAB
CENTRE FOR THE SYNTHESIS AND ANALYSIS OF BIODIVERSITY
AGENCE NATIONALE DE LA RECHERCHE
ANR
FAPESP
FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO
JST
THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

DataCite

CONNECTING RESEARCH,
IDENTIFYING KNOWLEDGE


info@datacite.org


pidforum.org

datacite.org
blog.datacite.org


support.datacite.org
support@datacite.org


@datacite


DataCite

@datacite