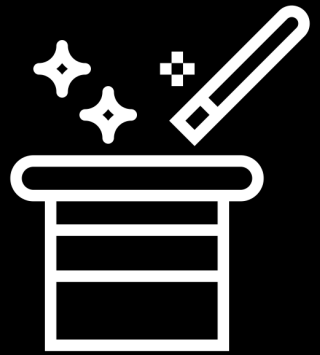
Hocus pocus: Mixing open identifiers into metadata makes connections between research work

Making the magic happen:
Helena Cousijn, DataCite: https://orcid.org/0000-0001-6660-6214
Maria Gould, ROR: https://orcid.org/0000-0002-2916-3423
Gabriela Mejias, ORCID: https://orcid.org/0000-0002-1598-7181
Rachael Lammey, Crossref: https://orcid.org/0000-0001-5800-1434



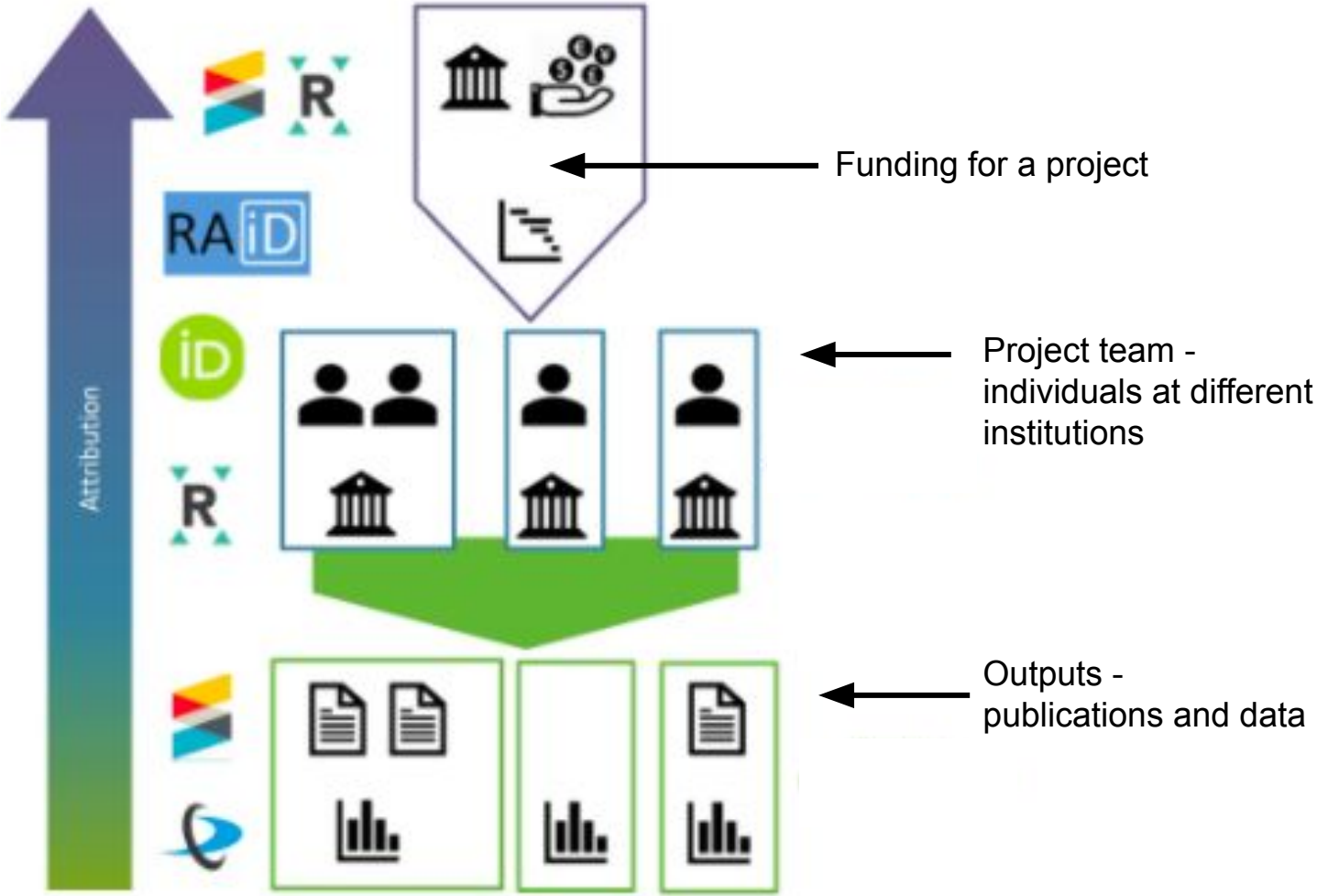
Created by Smalllike
from Noun Project

PIDs for people, places, and things in the research community

PIDs for people (researchers) include ISNIs and ORCID iDs

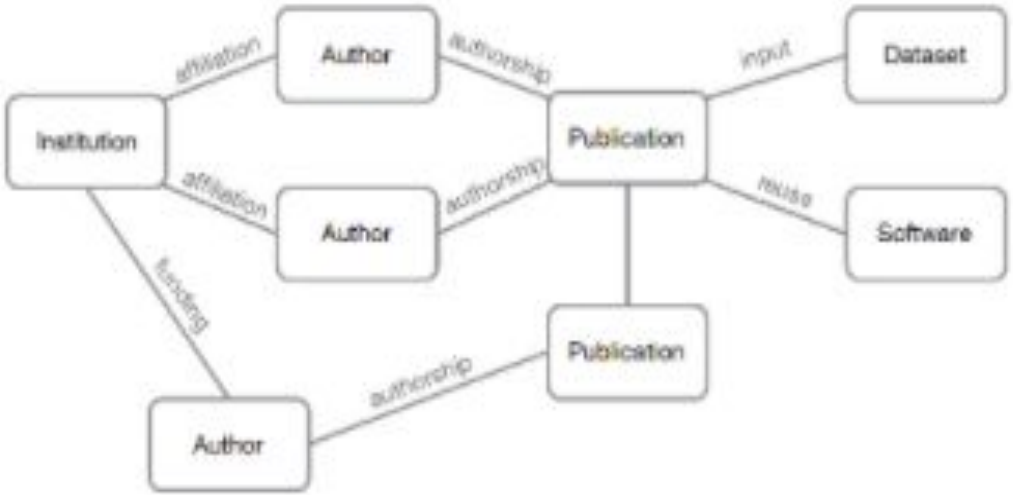
PIDs for places (research organizations) include GRID and ROR

PIDs for things (research outputs/inputs like grants, reviews, preprints, projects, etc.) include Crossref and DataCite DOIs, IGSNs, RAiDs, and more


Attribution
RAiD
iD
Funding for a project
Project team - individuals at different institutions
Outputs - publications and data

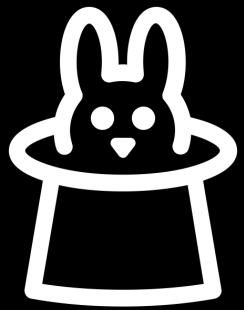
Good start, but we want more

Connecting everything reveals the true power of PIDs



https://doi.org/10.1016/j.patter.2020.100180

Connecting stuff up relies on community-driven open identifiers and metadata!

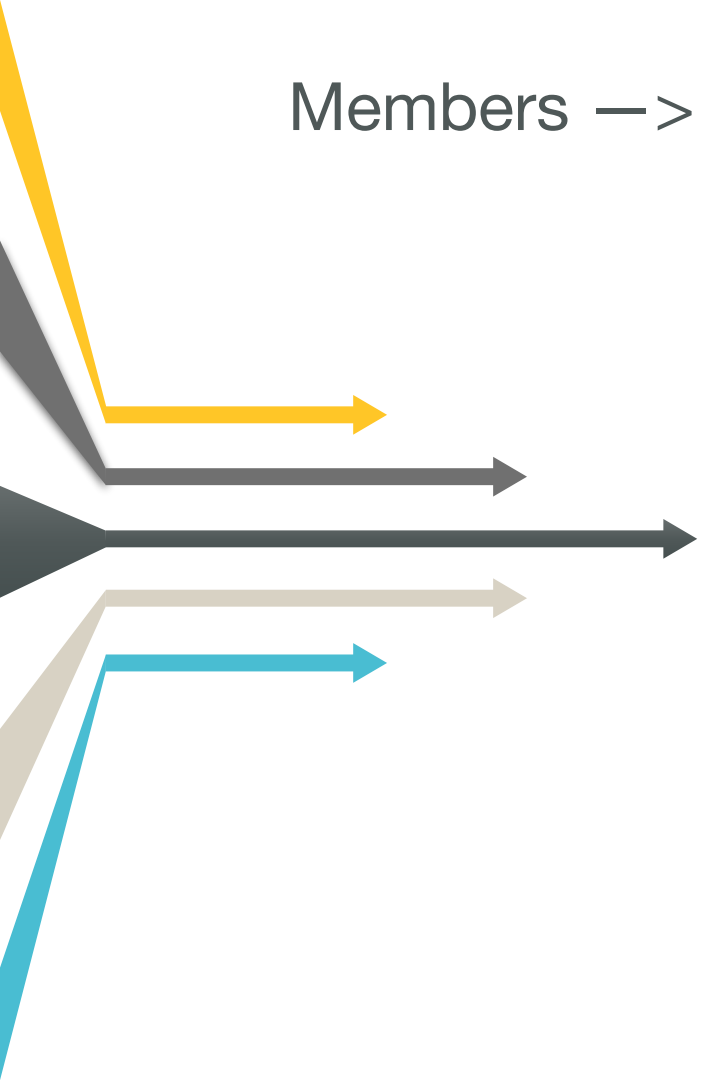


Created by Adrien Coquet
from Noun Project

Crossref

- Over 12,000 member organizations
- 16 member board, cross section of our international members
- Metadata store of over 121 million scholarly content items
- We offer a wide array of services to ensure that scholarly research metadata is registered, linked, and distributed.
- We preserve the metadata we receive and make it available via our open APIs and Search
- Committed to the Principles of Open Scholarly Infrastructure

Members —> Crossref



Basic metadata: titles; author names & **ORCID iDs**; ISSNs/ISBNs, abstracts, references

Funding Information, **Funder identifiers**, award numbers

License Information, License URIs

Full-text URIs (e.g. for text mining and Similarity Check)

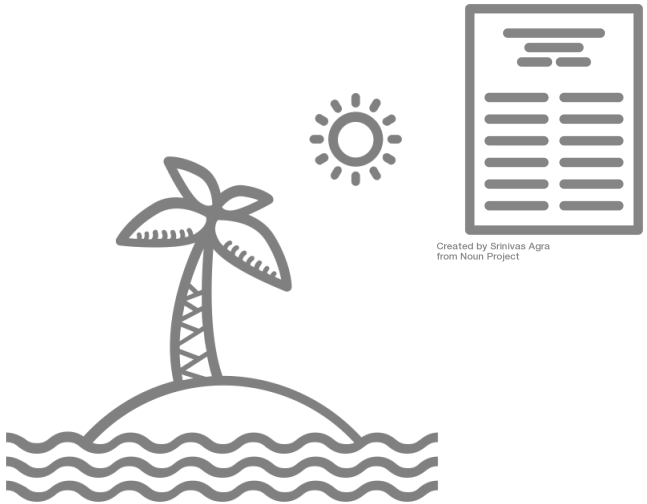
Crossmark: updates, retractions, corrections

ORCID iDs

Recently: Peer Review reports, relations. **links to data**, Grant IDs

Soon… **ROR IDs**, CRediT

Identifier-free Crossref metadata =

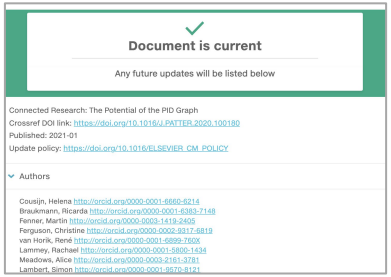


Created by Srinivas Agra
from Noun Project

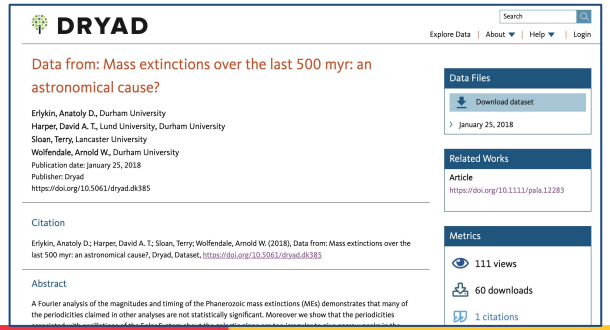
Created by AB
from Noun Project

Crossref open metadata + identifiers =

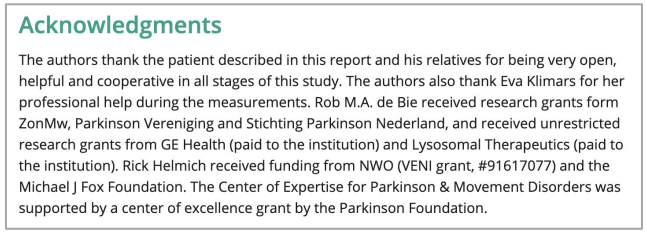
Updates and ORCID records



Links to data (& code)



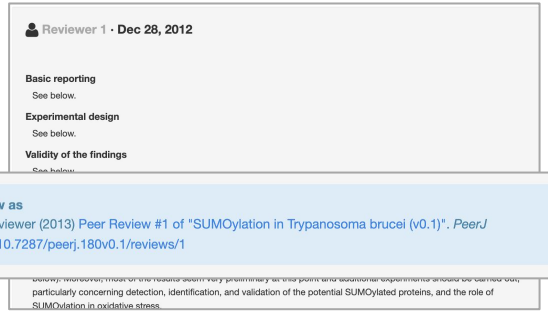
Funding

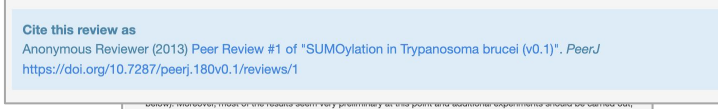


Created by Srinivas Agra
from Noun Project

Preprints & other related work

Peer reviews



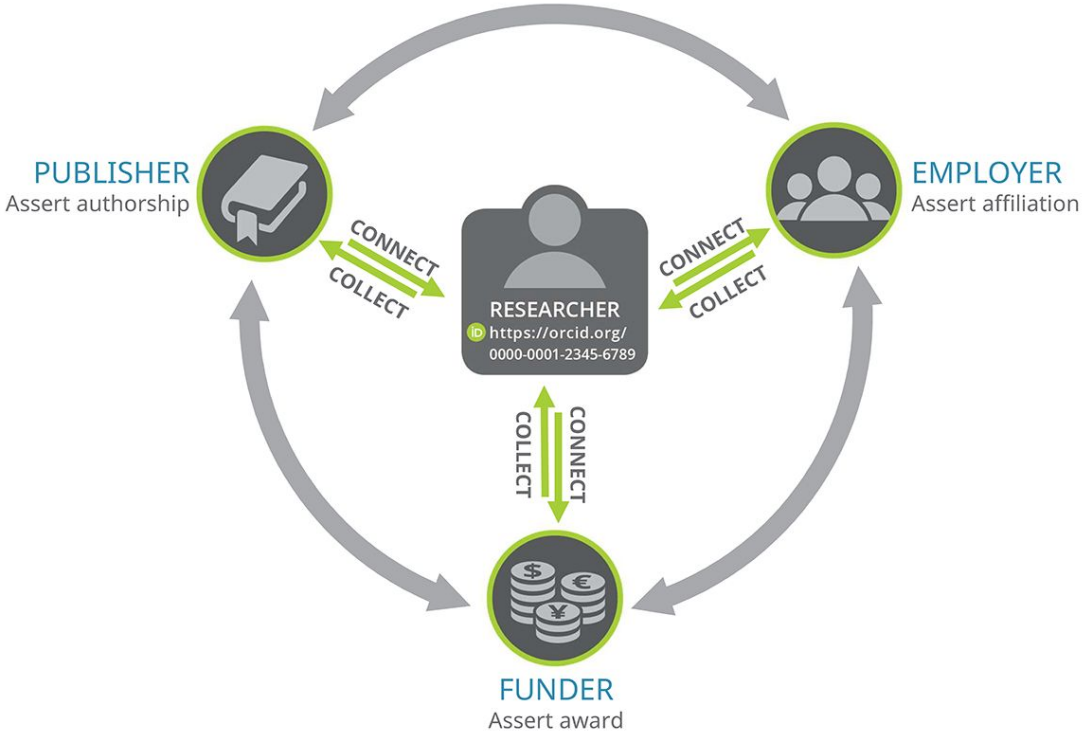


ORCID

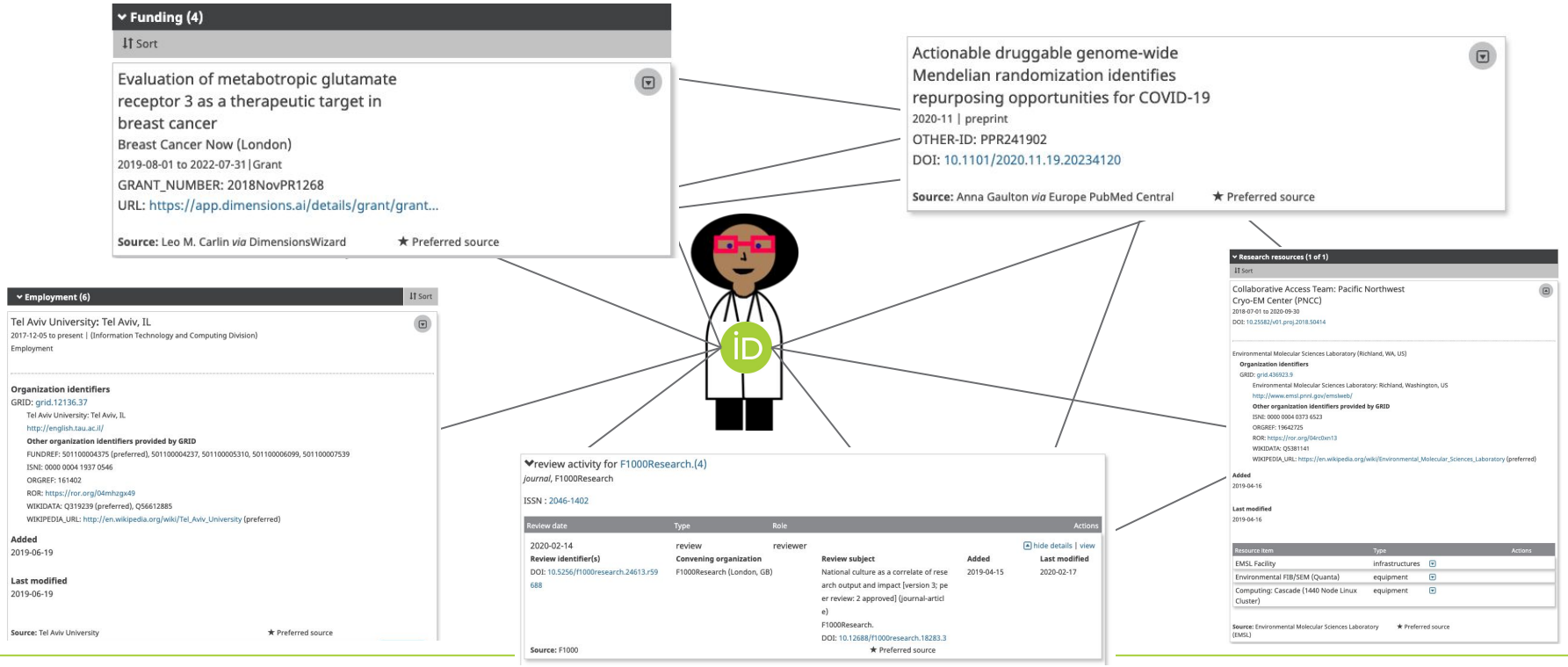
- An identifier for individuals (10M+ registered) connected to a **record** (affiliations, works, funding metadata)
- A Registry / Hub to synchronize data
- APIs for connecting researchers with other identifiers for their contributions, affiliations, activities, and more!
- A community of practice building connections and tools (1100+ members, 24 regional consortia, 3000+ Public API integrators)
- An international scale open research effort

INTEROPERABILITY

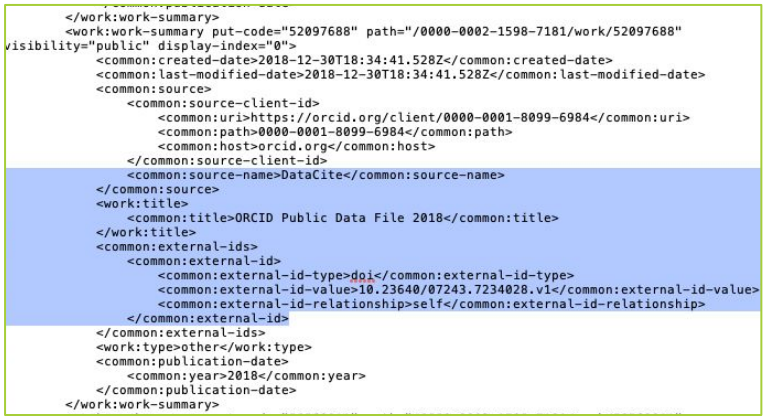
ENTER ONCE
REUSE OFTEN



Open metadata in action



The O in ORCID



```
        </work:work-summary>
        <work:work-summary put-code="52097688" path="/0000-0002-1598-7181/work/52097688"
visibility="public" display-index="0">
            <common:created-date>2018-12-30T18:34:41.528Z</common:created-date>
            <common:last-modified-date>2018-12-30T18:34:41.528Z</common:last-modified-date>
            <common:source>
                <common:source-client-id>
                    <common:uri>https://orcid.org/client/0000-0001-8099-6984</common:uri>
                    <common:path>0000-0001-8099-6984</common:path>
                    <common:host>orcid.org</common:host>
                </common:source-client-id>
                <common:source-name>DataCite</common:source-name>
            </common:source>
            <work:title>
                <common:title>ORCID Public Data File 2018</common:title>
            </work:title>
            <common:external-ids>
                <common:external-id>
                    <common:external-id-type>doi</common:external-id-type>
                    <common:external-id-value>10.23640/07243.7234028.v1</common:external-id-value>
                    <common:external-id-relationship>self</common:external-id-relationship>
                </common:external-id>
            </common:external-ids>
            <work:type>other</work:type>
            <common:publication-date>
                <common:year>2018</common:year>
            </common:publication-date>
        </work:work-summary>
```

- Public Data File
  - released annually during OA Week
  - public record data in XML
  - CC0
  - 2020 dump: 3000+ downloads

- Public API
  - + 2000 registered users
  - Read public data

https://orcid.org/content/orcid-public-data-file-use-policy
https://orcid.org/content/orcid-public-client-terms-service

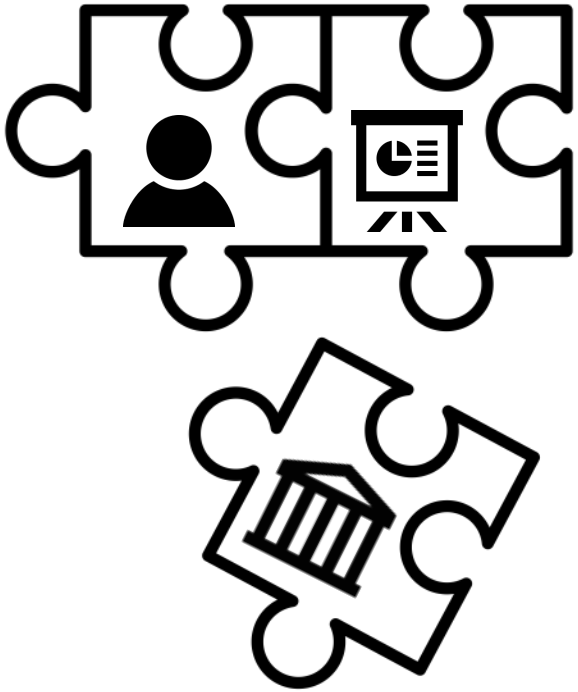
The road to open

- Add support for CRedIT
- Funded by relationship
- Add support for ROR

is never complete...

https://trello.com/b/CTB1InDi/orcid-product-roadmap

Research Organization Registry (ROR)



ROR

@ResearchOrgs

Articles by UCSF researchers ....

Robin Researcher,
University of California San Francisco

Steph Scientist,
University of California-San Francisco

Articles by UCSF researchers ....

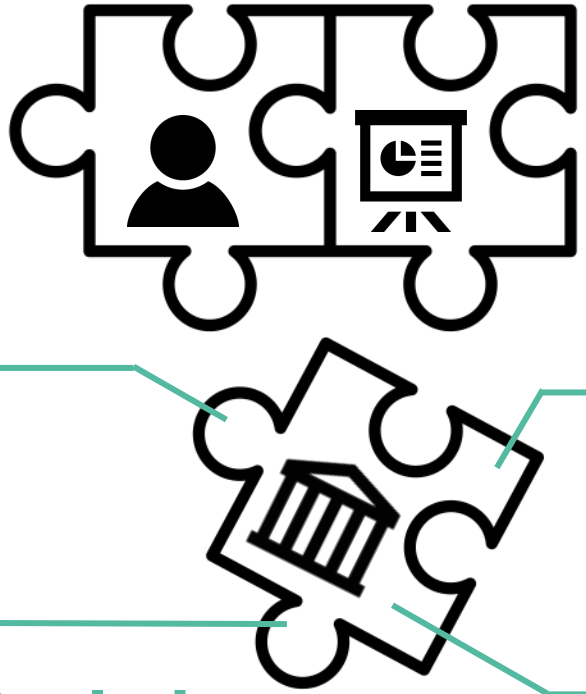
Barbara Biologist,
UCSF

Val Virologist,
UC San Francisco

Lou Labtech,
Department of Laboratory Medicine, UCSF, San Francisco CA

ROR

Research Organization Registry (ROR)



Noncommercial
Open
CC0

Designed to be integrated into open research infrastructure

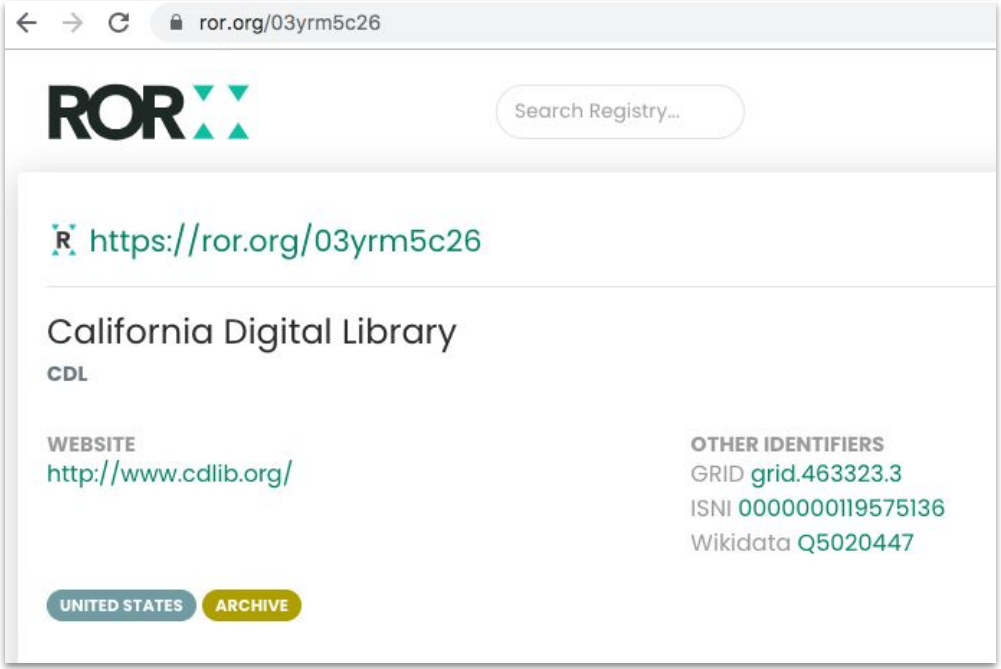
Community-led

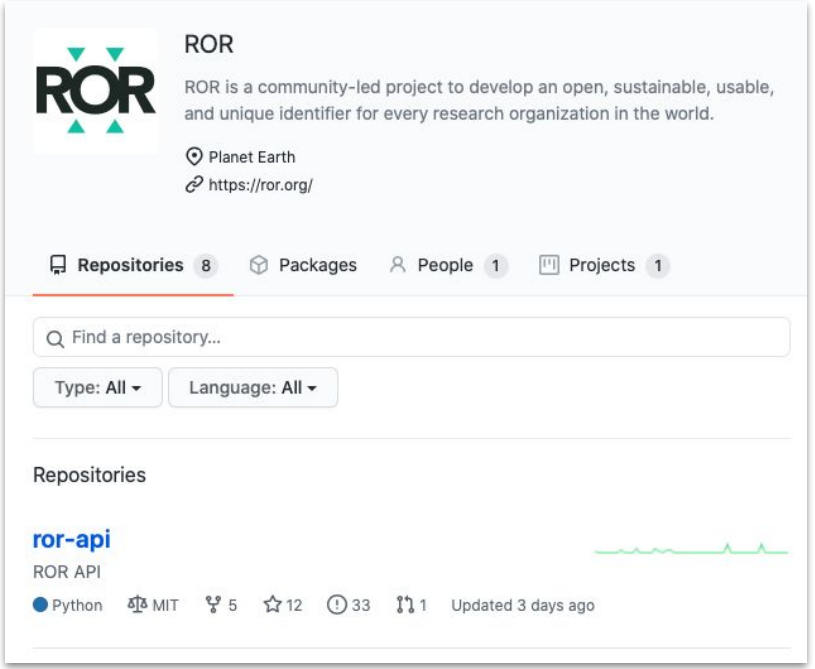
Focused on affiliations

ROR

@ResearchOrgs

Research Organization Registry (ROR)





Articles where affiliation=https://ror.org/043mz5j54

doi:10.1234/metadatawithror01

doi:10.5678/metadatawithror02

doi:10.9109/metadatawithror03

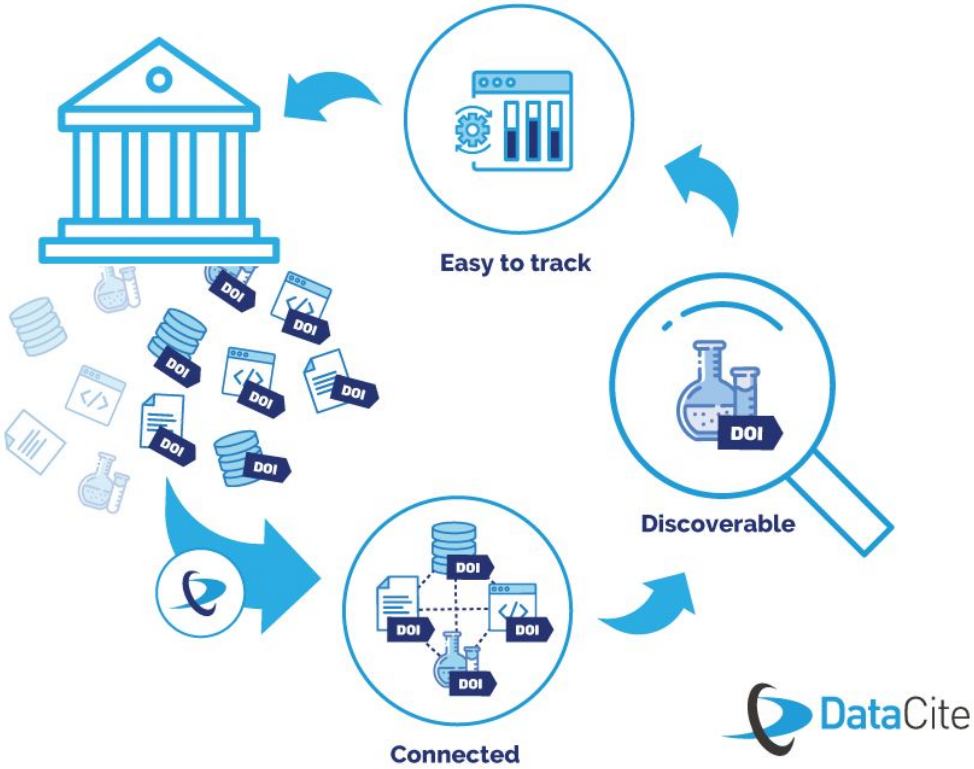
doi:10.8765/metadatawithror04

doi:10.4321/metadatawithror05

```
  <identifier identifierType="DOI">10.7272/Q6HD7SV7</identifier>
  <creators>
    <creator>
      <creatorName>Macquarie, Charles</creatorName>
      <nameIdentifier nameIdentifierScheme="ORCID" schemeURI="http://orcid.org/">0000-0003-0229-0761</
nameIdentifier>
      <affiliation affiliationIdentifier="https://ror.org/043mz5j54"
affiliationIdentifierScheme="ROR">University of California, San Francisco</affiliation>
    </creator>
    <creator>
      <creatorName>Tang, Rebecca</creatorName>
      <affiliation affiliationIdentifier="https://ror.org/043mz5j54"
affiliationIdentifierScheme="ROR">University of California, San Francisco</affiliation>
    </creator>
    <creator>
      <creatorName>Krah, David</creatorName>
      <affiliation affiliationIdentifier="https://ror.org/043mz5j54"
affiliationIdentifierScheme="ROR">University of California, San Francisco</affiliation>
    </creator>
    <creator>
      <creatorName>Ilieva, Polina</creatorName>
      <nameIdentifier nameIdentifierScheme="ORCID" schemeURI="http://orcid.org/">0000-0002-6121-100X</
nameIdentifier>
      <affiliation affiliationIdentifier="https://ror.org/043mz5j54"
affiliationIdentifierScheme="ROR">University of California, San Francisco</affiliation>
    </creator>
```

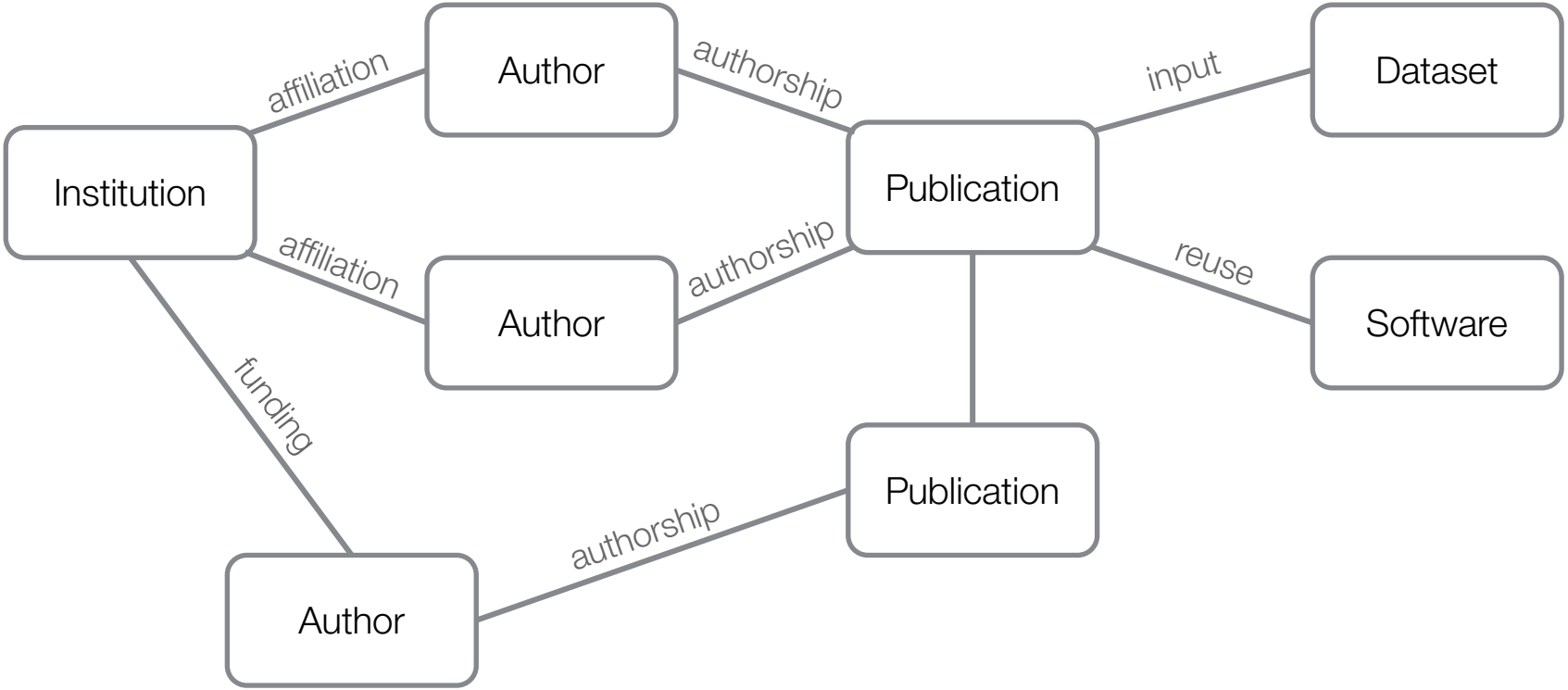
DataCite
Connecting research,
identifying knowledge

How we do that



Research is already connected

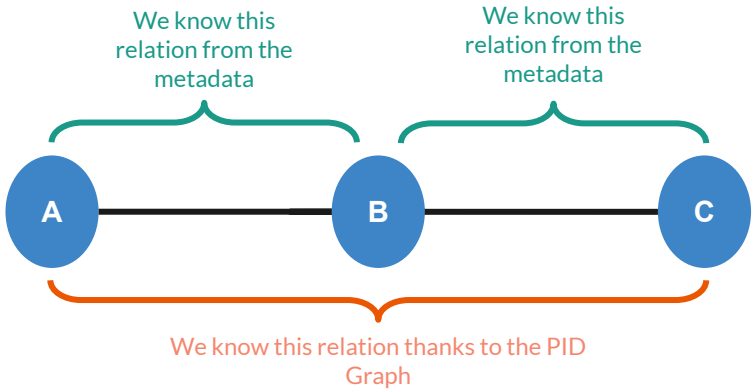
DataCite

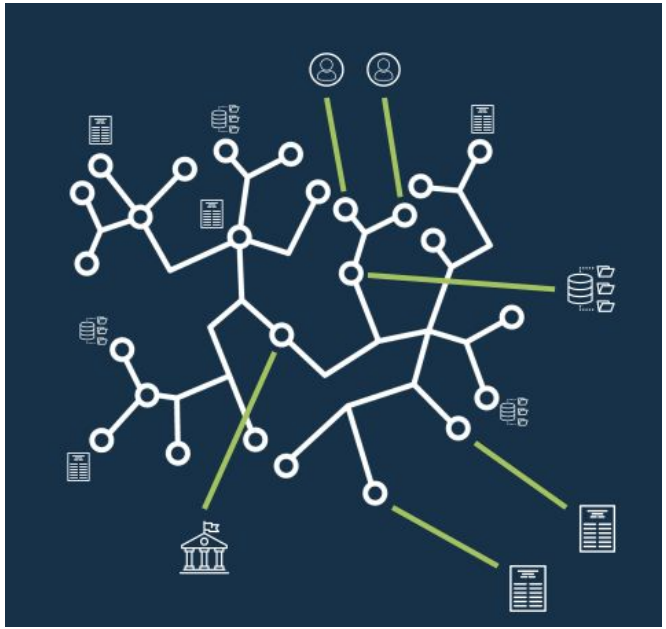


PID Graph

The PID Graph was developed within EC project FREYA (https://www.project-freya.eu).

The basic idea is to link PIDs for different entities together via relations in their metadata to enable the discovery of connections at least two steps away.

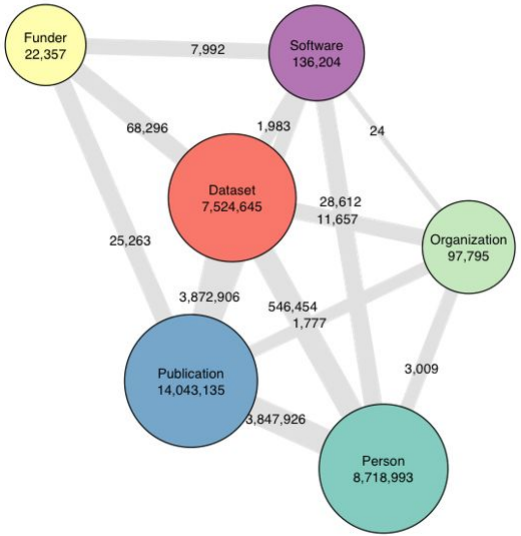




Find and Connect Research
with existing infrastructure

Find Research
with DataCite Commons

PID Graph
Number of nodes and connections (May 2020)



Cited dataset

DataCite

Data from: Impact of negative frequency-dependent selection on mating pattern and genetic structure: a comparative analysis of the S-locus and nuclear SSR loci in Prunus lannesiana var. speciosa

Kato Shuri, Teruyoshi Nagamitsu, Hiroyoshi Iwata, Yoshihiko Tsumura, Yuzuru Mukai, K Michiharu, K Saika & K Junko

Version 1 of Dataset published 2012 in DRYAD

Mating processes of local demes and spatial genetic structure of island populations at the self-incompatibility (S-) locus under negative frequency-dependent selection (NFDS) were evaluated in Prunus lannesiana var. speciosa in comparison with nuclear simple sequence repeat (SSR) loci that seemed to be evolutionarily neutral. Our observations of local mating patterns indicated that male-female pair fecundity was influenced by not only self-incompatibility, but also various factors such as kinship, pollen production and flowering synchrony. In spite of the mating bias caused by these factors, the NFDS effect on changes in allele frequencies from potential mates to mating pollen was detected at the S-locus but not at the SSR loci although the changes from adult to juvenile cohorts were not apparent at any loci. Genetic differentiation and isolation-by-distance over various spatial scales were smaller at the S-locus than at the SSR loci, as expected under the NFDS. All ele sharing distributions among the populations also had a unimodal pattern at the S-locus, indicating the NFDS effect except for alleles unique to individual populations probably due to isolation among islands, although this pattern was not exhibited by the SSR loci. Our results suggest that the NFDS at the S-locus has an impact on both the mating patterns and the genetic structure in the P. lannesiana populations studied.

DOI registered April 17, 2012 via DataCite.

CC 0

1 Citation 99 Views 16 Downloads

Dataset English

https://doi.org/10.5061/dryad.7c425

Supporting recognition

DataCite

Creators

Kato Shuri
Forestry and Forest Products Research Institute

Teruyoshi Nagamitsu
Forestry and Forest Products Research Institute

Hiroyoshi Iwata
University of Tokyo

Yoshihiko Tsumura
Forestry and Forest Products Research Institute

Yuzuru Mukai
Gifu University

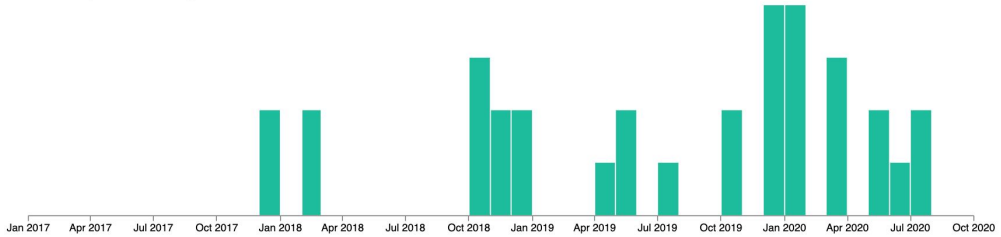
K Michiharu
Kyoto University

K Saika
Tokyo Institute of Technology

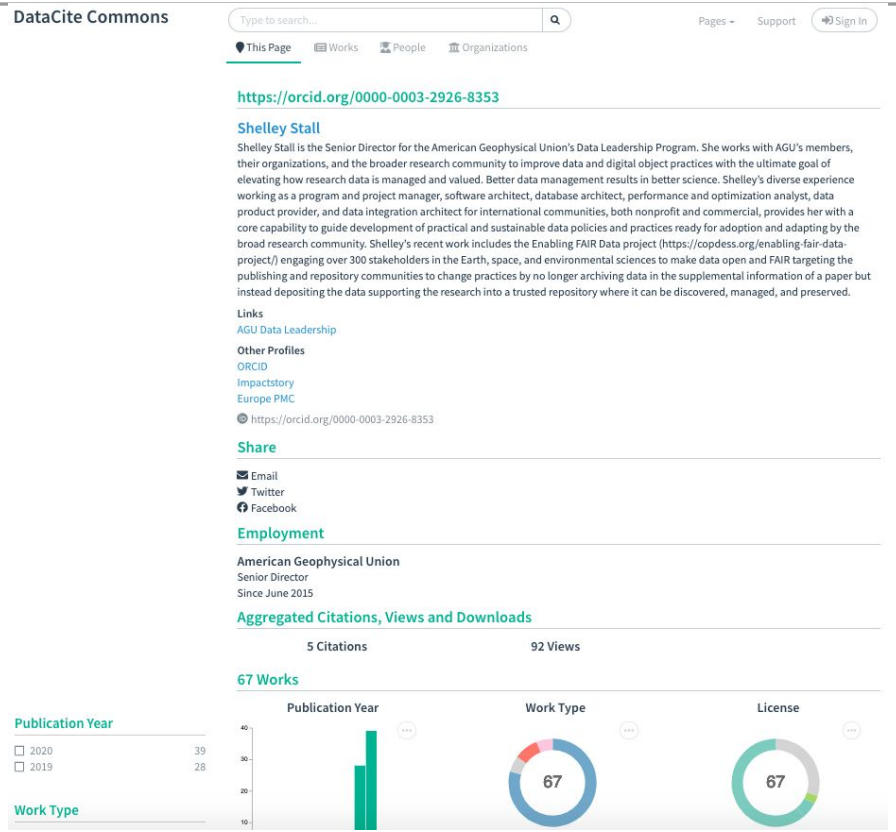
K Junko
Gunma University

99 Views 16 Views

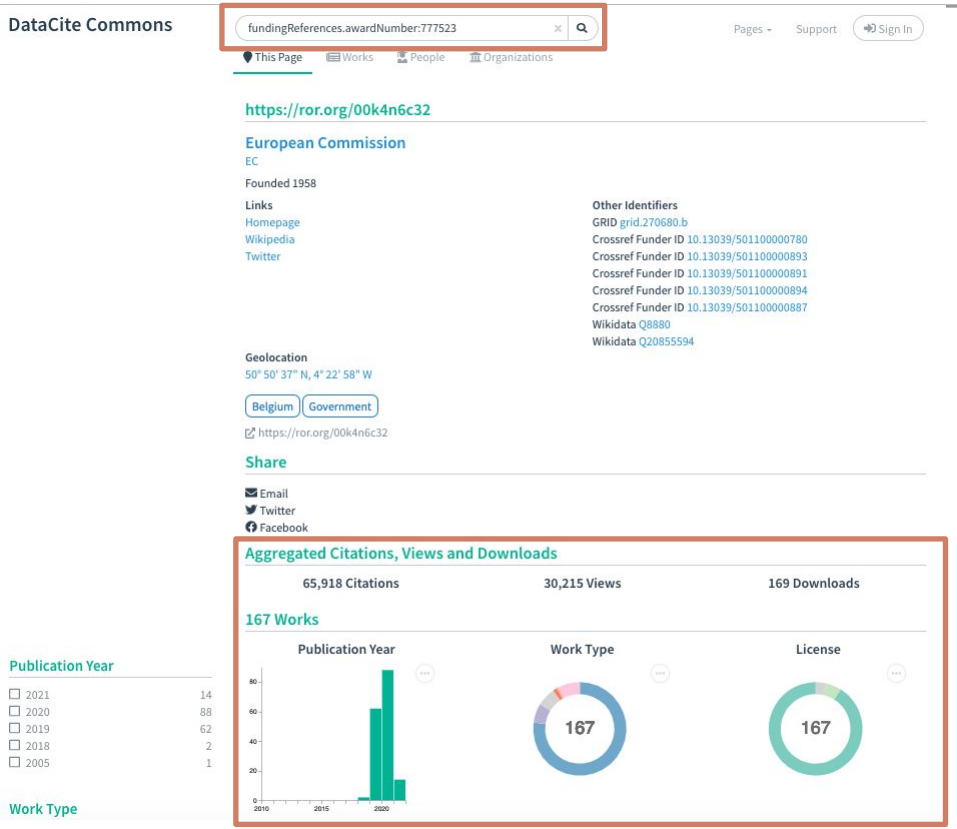
99 views reported since publication in 2012.



Leveraging ORCID iDs



Leveraging citations & usage



Benefits

Connecting open metadata helps you with:

Discovery

Improved workflows

Global collaborations

Analytics

Knowledge

What can you do?

DataCite
FIND, ACCESS, AND REUSE DATA

1. **Use PIDs for All Entities**

- Researchers should use a researcher ID in their workflows and deposit all research outputs
- Institutions should verify and use their institution ID and assist researchers with data deposition
- Funders should use a funder IDs and register DOIs and metadata for grants.
- Repositories and publishers should assign PIDs to data and other outputs.

2. **Track and Record Connections between PIDs**

- Infrastructure providers should provide a relevant metadata schema.
- Repositories and publishers should ask for information about connections between entities, such as data citations.
- Researchers and institutions should include and update the information about relations wherever possible.

3. **Make Connections Openly Available**

- Infrastructure providers should aggregate the information and make it openly available.
- Publishers and repositories should ensure the information is included in the metadata they share.

Helena Cousijn, Ricarda Braukmann, Martin Fenner, Christine Ferguson, René van Horik, Rachael Lammey, Alice Meadows, Simon Lambert (2021). Connected PIDs: The Potential of the PID Graph. Patterns 2,1. https://doi.org/10.1016/j.patter.2020.100180

Thank you!

Created by priyanka
from Noun Project