



CONNECTING RESEARCH,
IDENTIFYING KNOWLEDGE

The PID Graph powered by DataCite

Matt Buys

May 10, 2023



[@datacite](https://twitter.com/datacite)

[@mjbuys](https://twitter.com/mjbuys)



Why We Exist

We are a **global community that share a common interest**: to ensure that research outputs and resources are openly available and connected so that their reuse can advance knowledge across and between disciplines, now and in the future.

As a community, we make research more effective with metadata that connects research outputs and resources—from samples and images to data and preprints. We enable the creation and management of persistent identifiers (PIDs), integrate services to improve research workflows, and facilitate the discovery and reuse of research outputs and resources.



2800+

Repositories



50

Countries



45m+

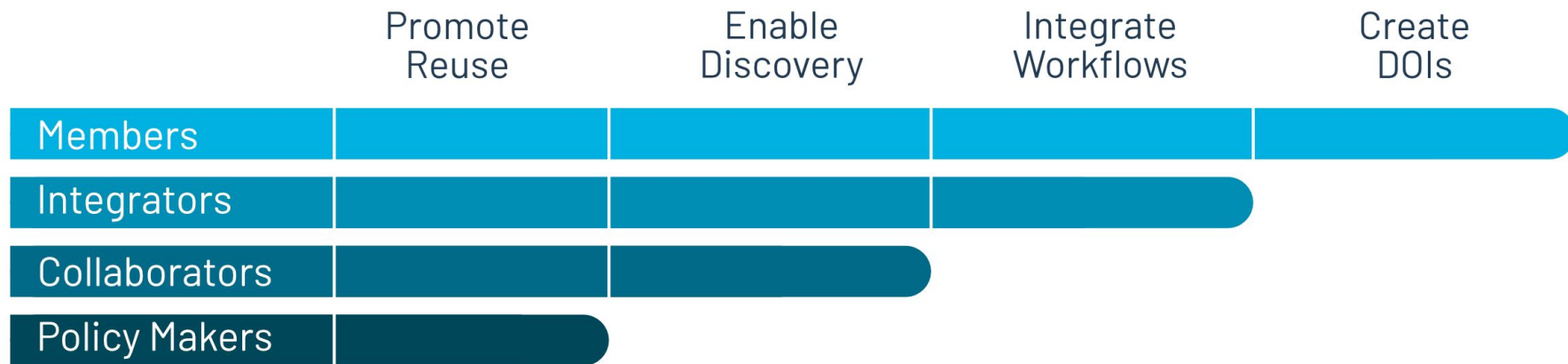
DOIs



1300+

Organizations

Community



Strategic initiatives

It takes a village

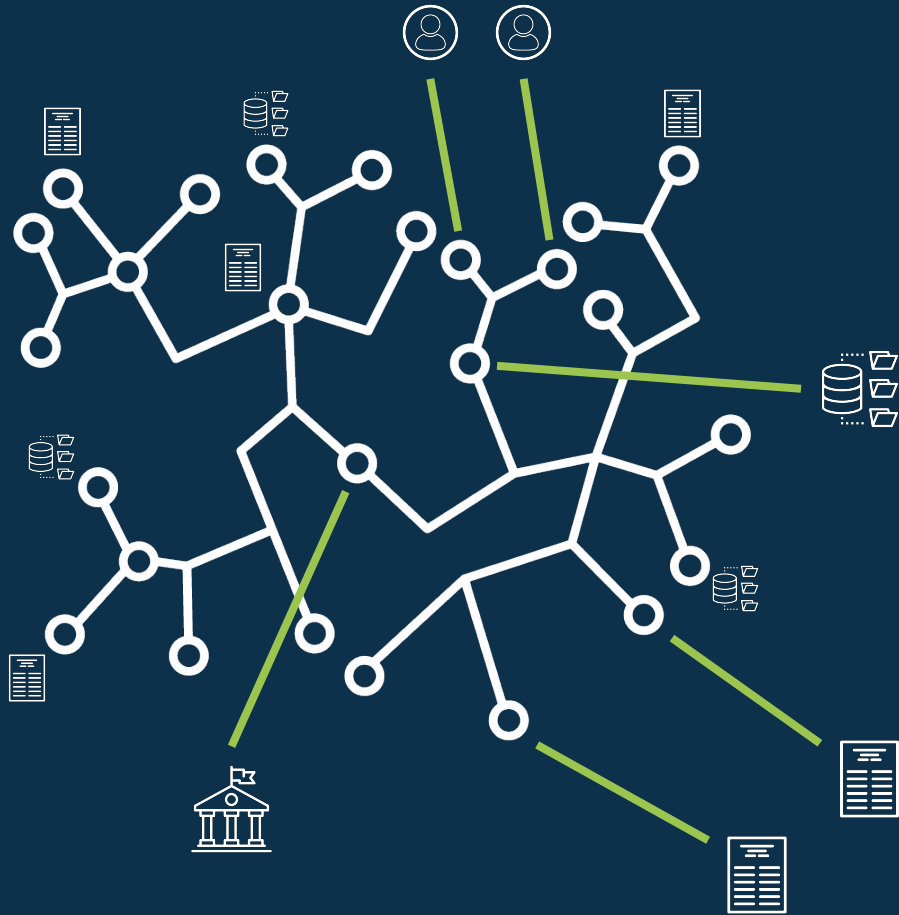
In line with our mission and vision, DataCite also actively participates and leads various initiatives through collaboration with stakeholders in the community to make open science a reality.

- **Data metrics** – We help further the adoption and implementation of responsible data metrics with, for example, the Make Data Count initiative.
- **Identifier registries** – We support community-led registries of identifiers such as the Research Organization Registry (ROR).
- **Repository discovery** – We contribute to the development of repository discovery initiatives such as re3data with collaboration and financial support.



Spotlight

PID Graph





<https://doi.org/10.5072/123456>

<https://doi.org/10.xxxx/xxxxxx>

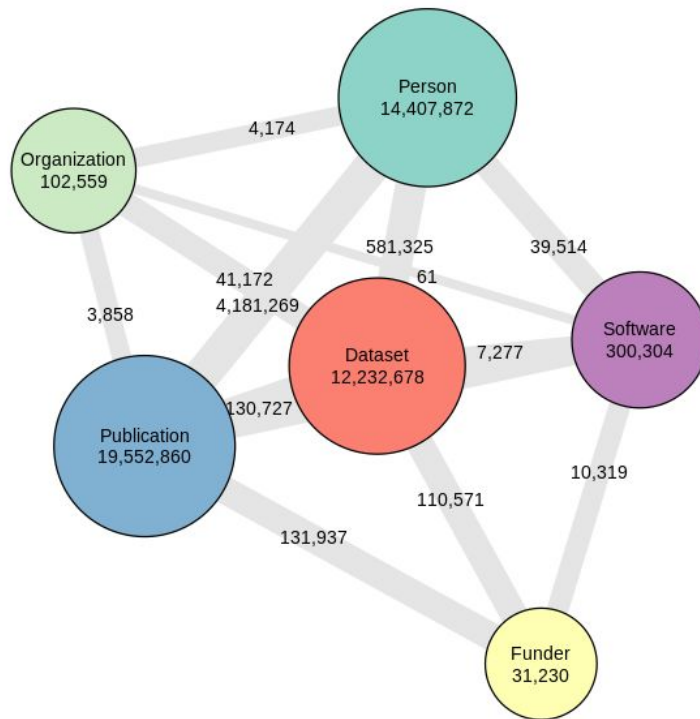
<https://doi.org/10.xxxx/xxxxxx>



Connection Metadata from Trusted sources

DataCite Metadata Schema property	Used for connections to...	Typical identifiers
relatedIdentifier	related research outputs - citations, versions...	DOIs  URLs, handles...
nameIdentifier <i>for Creators and Contributors</i>	authors and contributors	ORCID iDs (for people)  ROR IDs (for organizations) 
affiliationIdentifier <i>for Creators and Contributors</i>	affiliated organizations	ROR IDs 
funderIdentifier <i>for FundingReferences</i>	funding organizations	Crossref Funder IDs  ROR IDs 

What is in the PID Graph?

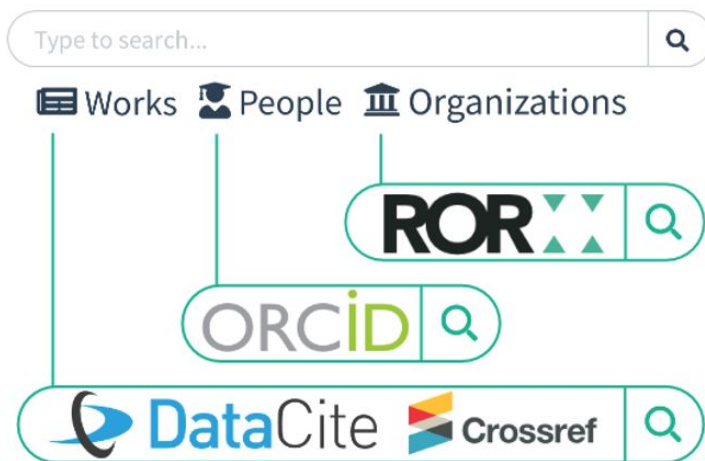


How can you access the PID graph?

DataCite Commons

Front end interface for the PID Graph

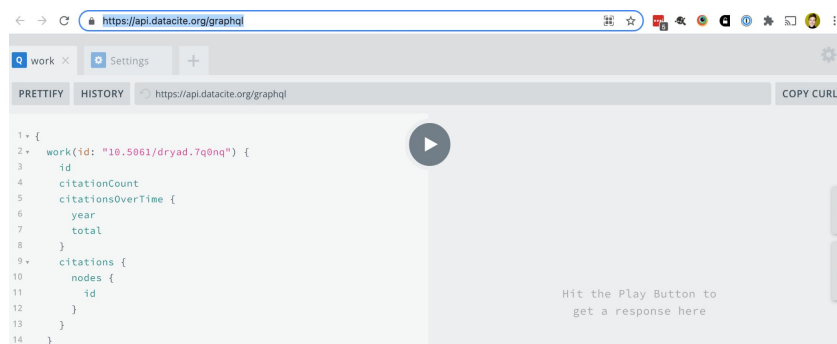
<https://commons.datacite.org/>



DataCite GraphQL API

Run your own custom queries

<https://api.datacite.org/graphql>



Characteristics of the PID Graph

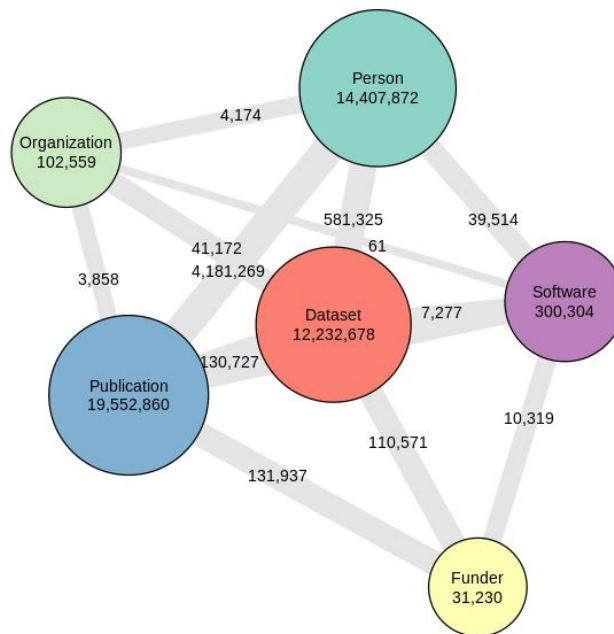
- It identifies all nodes in the graph with persistent identifiers.
- All relationships in the Graph are provided by trusted sources.
- Provides a GraphQL schema (github.com/.../schema.graphql) that describes entity types as well as minimal (and optional) set of properties for each entity.
- Relationships between the nodes are available in Scholix format in an REST API.

Syncing RDA Taskforce Efforts

With the PID Graph

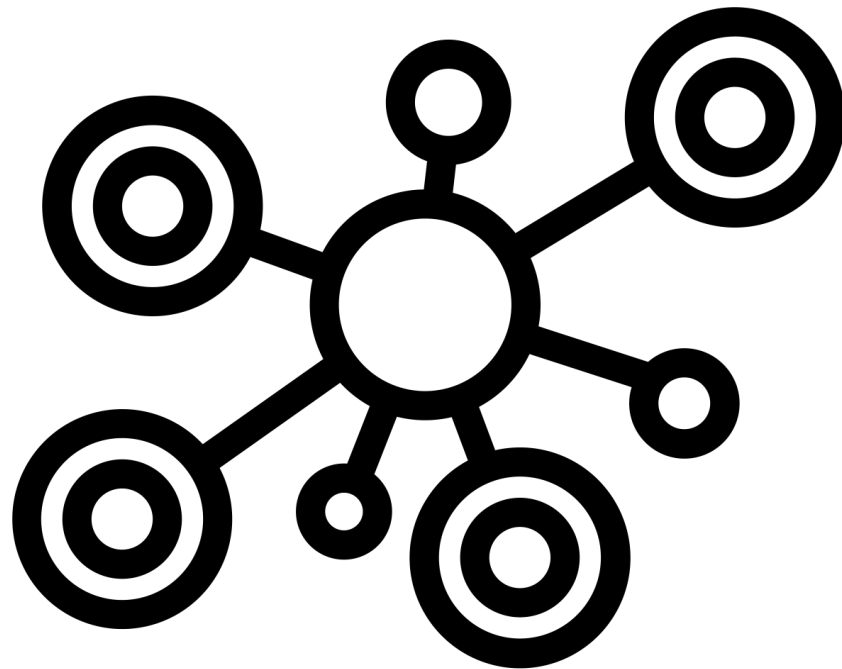
Task Force 1: Common data model

- It identifies all nodes in the graph with persistent identifiers.
- Entities: Dataset, Organisation, Funder, etc.
- On initial examination of the entity model we think that we can find compatibility for any entity that has a PID.
- Further examination would be necessary for entities that do not use widely adopted PIDs



Task Force 2: Common metadata format and crosswalk

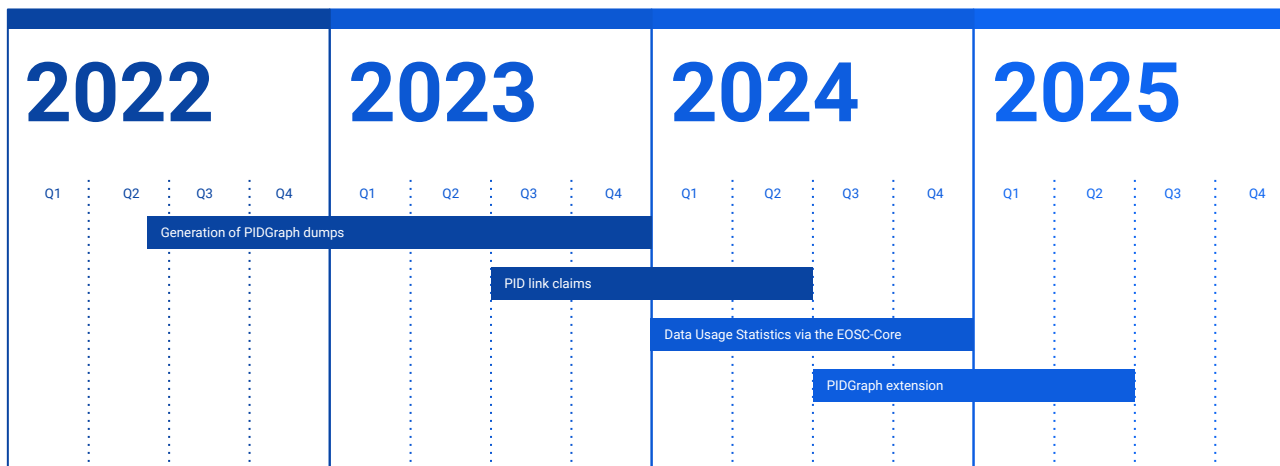
To study: Crosswalk for DataCite Schema or the PID Graph Schema.



Foreseen use cases for interoperability

- Use case #1 - Entity synchronisation
- Use case #2 - Relationship synchronisation
- Use case #3 - Bulk dump
- Use case #4 - Complex queries

Timeline to address the use cases



datacite.org/roadmap.html



CONNECTING RESEARCH,
IDENTIFYING KNOWLEDGE



info@datacite.org



pidforum.org



datacite.org
blog.datacite.org



support@datacite.org
support@datacite.org



[@datacite](https://twitter.com/datacite)



[DataCite](https://www.youtube.com/DataCite)



[@datacite](https://www.linkedin.com/company/datacite)