



# PID Graph and ROR

RDC Webinar

27 February 2020

Robin Dasler, DataCite

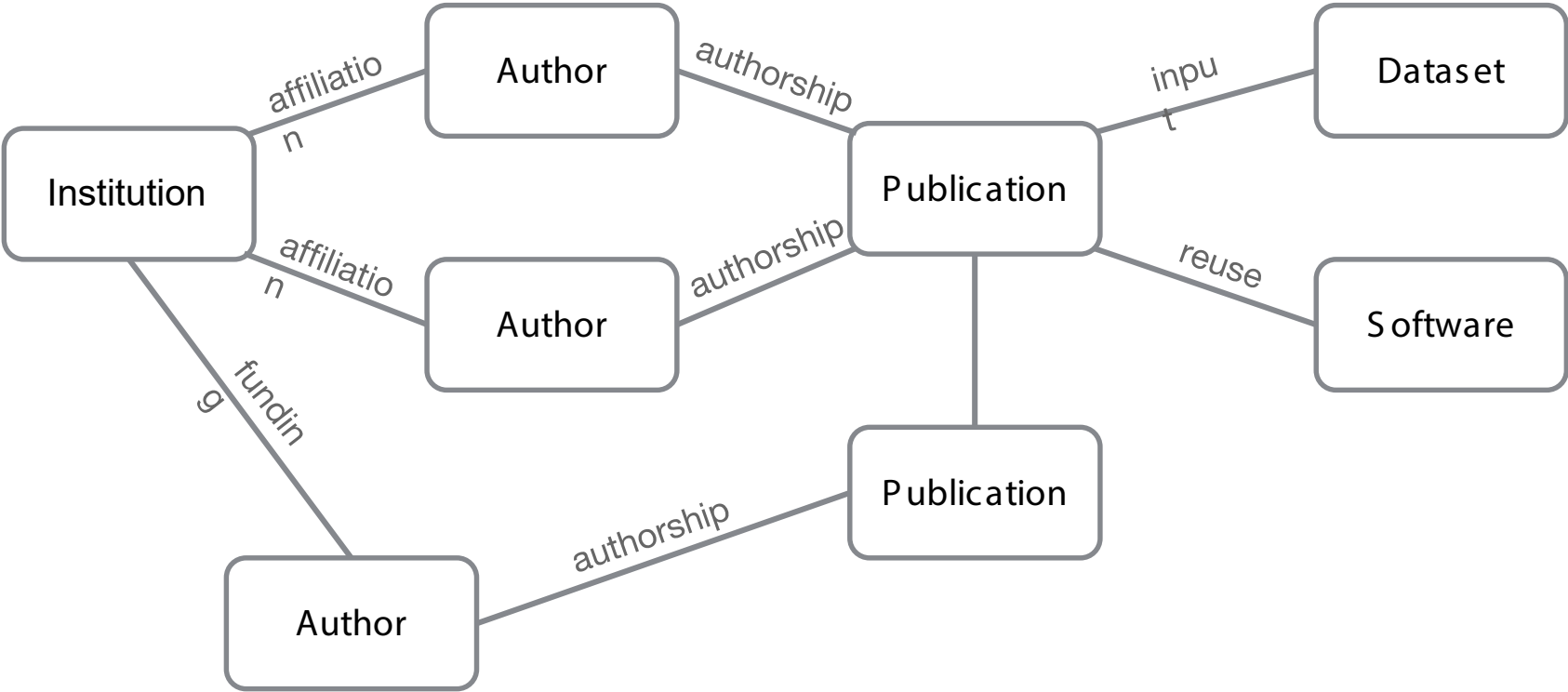
# What is the PID Graph?

# Research is already a graph

Researchers, institutions, publications, datasets, and more are interconnected.

Entities and the relationships between them form a conceptual graph of the connected research landscape.

# It could look like this



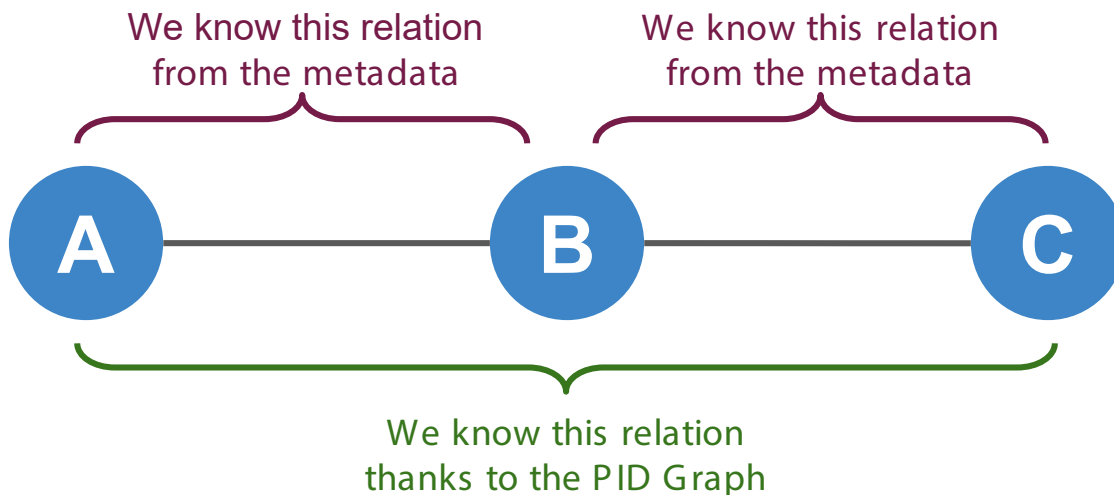
## **PIDs are the backbone of connected research**

Having unique persistent identifiers for researchers and their outputs is crucial to connecting pieces of the research landscape together.

PIDs already have the potential to enable the connected research graph, but we're not yet taking full advantage of their connecting powers.

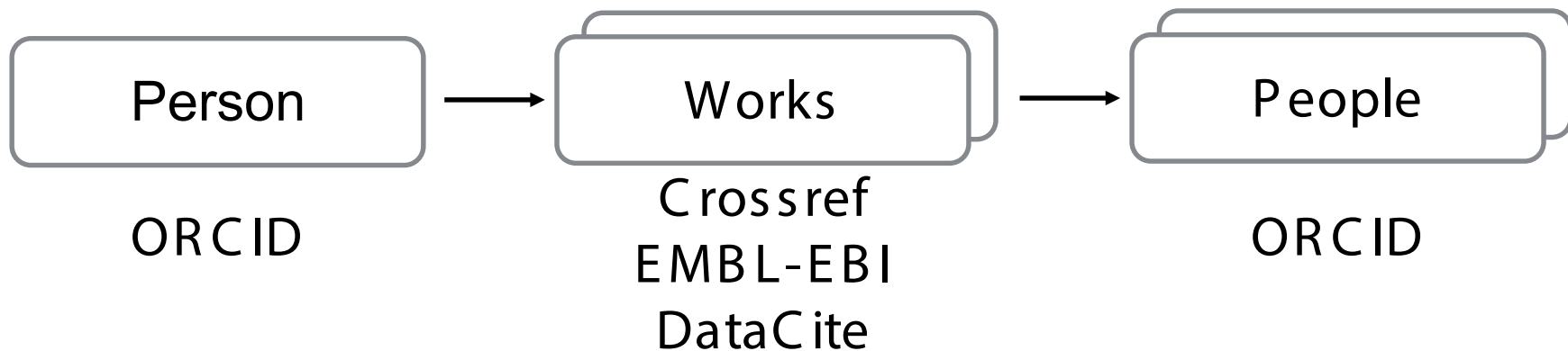
## Enter the PID Graph

We can link PIDs together via relations in their metadata to enable the discovery of connections at least two “hops” away.



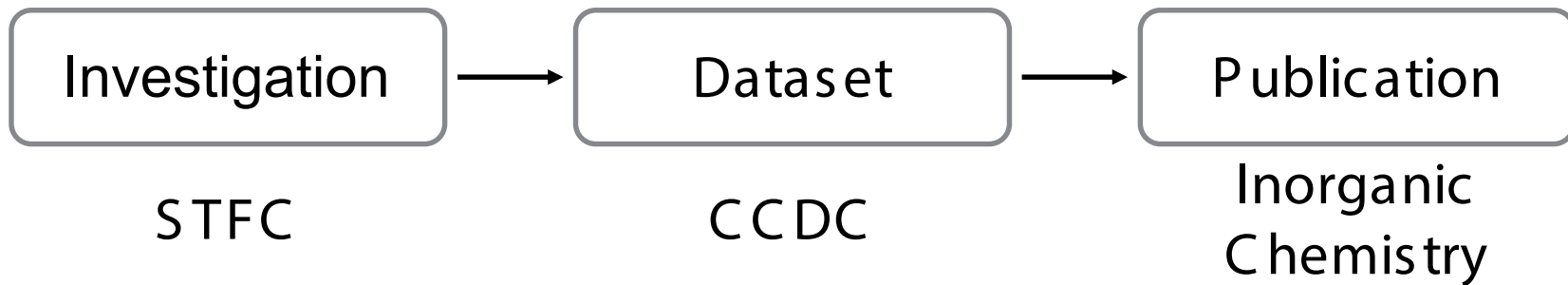
**What questions could the PID Graph  
answer?**

# Who are all the co-authors of a particular researcher?

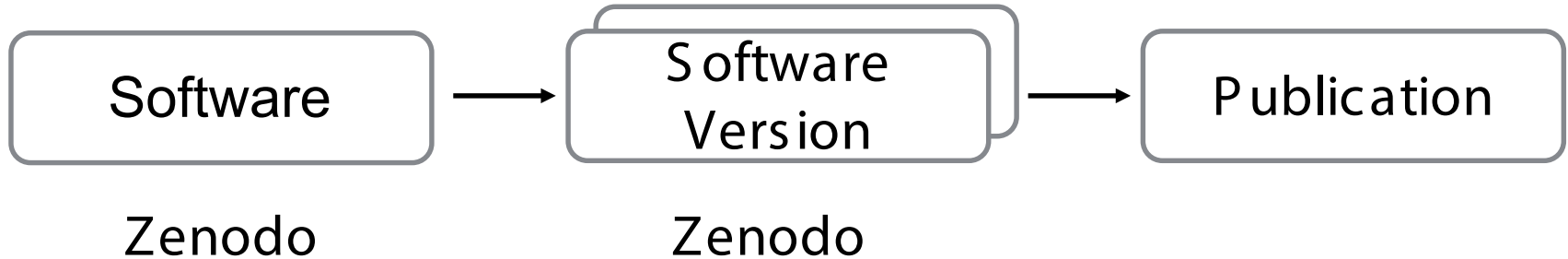




## Which publications used STFC investigations for the underlying data?



# Which publications cite any version of a piece of software?



# Enabling the PID Graph at DataCite

## DataCite GraphQL API

DataCite has developed a GraphQL API. This is the most convenient way to consume the PID Graph with DataCite metadata as a starting point.

GraphQL is a query language that's specially built for graphs. It lets you specify exactly which information you want and in what structure you'd like to receive it.

The DataCite GraphQL API is currently in pre-release, with a stable release planned for around RDA in March.

GraphQL Endpoint

Method

POST

[Edit HTTP Headers](#)

GraphiQL



Prettify

History

[< Docs](#)

```
1 {
2   publications(query: "creators.name:dasler") {
3     id
4     titles {
5       title
6     }
7     descriptions {
8       description
9     }
10    creators {
11      id
12      name
13      familyName
14    }
15    fundingReferences {
16      funderIdentifier
17      funderName
18      awardTitle
19      awardNumber
20    }
21  }
22 }
23
```

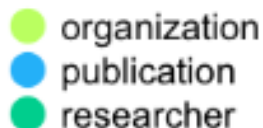
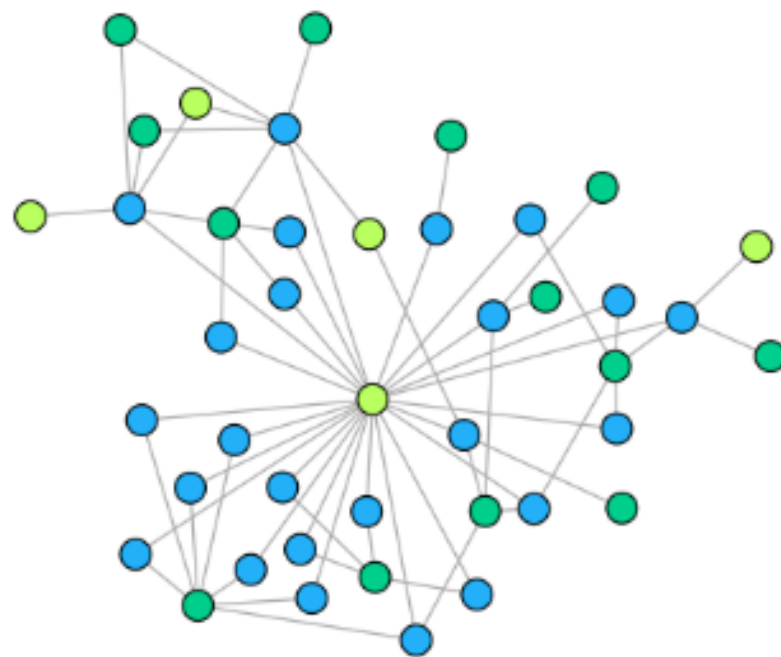
```
{
  "data": {
    "publications": [
      {
        "id": "https://doi.org/10.5281/zenodo.1064000",
        "titles": [
          {
            "title": "Pid Service Adoption"
          }
        ],
        "descriptions": [
          {
            "description": "This presentation describes how the uptake of persistent identifiers can be measured and gives an overview of the main results of the ORCID adoption study."
          }
        ],
        "creators": [
          {
            "id": null,
            "name": "Dasler, Robin",
            "familyName": "Dasler"
          }
        ]
      }
    ]
  }
}
```

## Surfacing connections

Graph of all the publications associated with DataCite, plus all the researchers and organizations associated with those publications.

Answers the question :

**With whom has DataCite collaborated on their publications?**

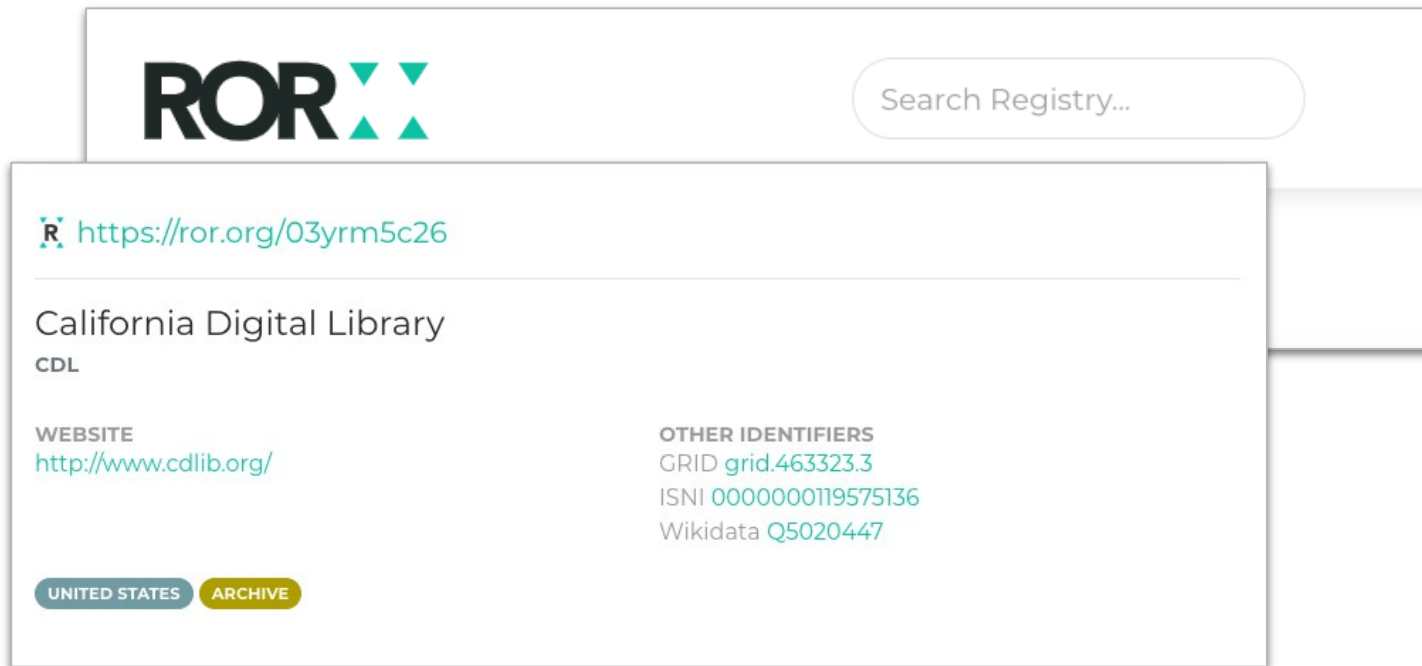


# Expanding the PID Graph with ROR


The Research Organization Registry  
is a **community-led project**  
to develop an **open**, sustainable,  
usable, and unique identifier  
for every **research organization**  
in the world.




# ROR registry: <https://ror.org/search>



The image shows a screenshot of the ROR registry search results. At the top left is the ROR logo, and at the top right is a search bar labeled "Search Registry...". The main content area displays the ROR ID <https://ror.org/03yrm5c26> for the California Digital Library (CDL). Below the name, there are two columns of information: "WEBSITE" with the URL <http://www.cdlib.org/>, and "OTHER IDENTIFIERS" listing GRID [grid.463323.3](https://www.grid.ac/details/grid.463323.3), ISNI [0000000119575136](https://www.isni.org/0000000119575136), and Wikidata [Q5020447](https://www.wikidata.org/wiki/Q5020447). At the bottom left, there are two tags: "UNITED STATES" and "ARCHIVE".

**ROR** 

Search Registry...

 <https://ror.org/03yrm5c26>

---

California Digital Library  
CDL

**WEBSITE**  
<http://www.cdlib.org/>

**OTHER IDENTIFIERS**  
GRID [grid.463323.3](https://www.grid.ac/details/grid.463323.3)  
ISNI [0000000119575136](https://www.isni.org/0000000119575136)  
Wikidata [Q5020447](https://www.wikidata.org/wiki/Q5020447)

**UNITED STATES** **ARCHIVE**

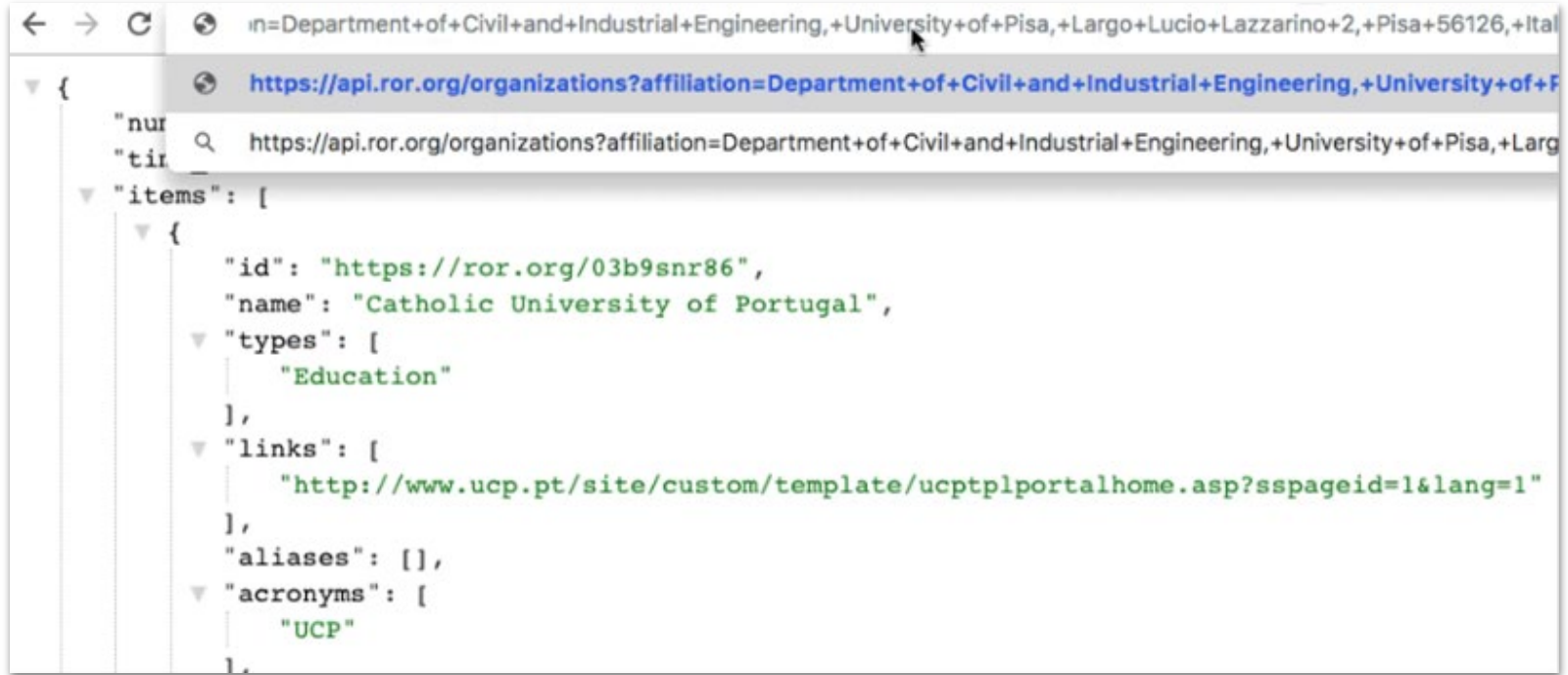
# ROR API



```
{
  "number_of_results": 91625,
  "time_taken": 1,
  "items": [
    {
      "id": "https://ror.org/01kj2bm70",
      "name": "Newcastle University",
      "types": [
        "Education"
      ],
      "links": [
        "http://www.ncl.ac.uk/"
      ]
    }
  ]
}
```

<https://api.ror.org/organizations>

# Affiliation matching (via the API)



The screenshot shows a web browser with a search bar containing the affiliation string: "Department+of+Civil+and+Industrial+Engineering,+University+of+Pisa,+Largo+Lucio+Lazzarino+2,+Pisa+56126,+Italy". Below the search bar, the browser's address bar shows the URL: <https://api.ror.org/organizations?affiliation=Department+of+Civil+and+Industrial+Engineering,+University+of+Pisa,+Largo+Lucio+Lazzarino+2,+Pisa+56126,+Italy>. The browser's developer tools are open, displaying the JSON response from the API. The response is a list of organizations, with the first one being the Catholic University of Portugal.

```
{
  "items": [
    {
      "id": "https://ror.org/03b9snr86",
      "name": "Catholic University of Portugal",
      "types": [
        "Education"
      ],
      "links": [
        "http://www.ucp.pt/site/custom/template/ucptplportalhome.asp?sspageid=1&lang=1"
      ],
      "aliases": [],
      "acronyms": [
        "UCP"
      ]
    }
  ]
}
```

# Submitting to DataCite

```
xmlns="http://datacite.org/schema/kernel-4" xsi:schemaLocation="http://datacite.org/schema/kernel-4
4 http://schema.datacite.org/meta/kernel-4.3/metadata.xsd">
  <identifier identifierType="DOI">10.5072/example-full</identifier>
  ▼<creators>
    ▼<creator>
      <creatorName nameType="Personal">Miller, Elizabeth</creatorName>
      <givenName>Elizabeth</givenName>
      <familyName>Miller</familyName>
      <nameIdentifier schemeURI="http://orcid.org/" nameIdentifierScheme="ORCID">0000-0001-5000-
0007</nameIdentifier>
      <affiliation affiliationIdentifier="https://ror.org/04wxnsj81"
affiliationIdentifierScheme="ROR">DataCite</affiliation>
    </creator>
    ▼<creator>
      <creatorName nameType="Personal">Carberry, Josiah</creatorName>
      <givenName>Josiah</givenName>
      <familyName>Carberry</familyName>
      <nameIdentifier schemeURI="http://orcid.org/" nameIdentifierScheme="ORCID">0000-0002-1825-
0097</nameIdentifier>
      <affiliation affiliationIdentifier="https://ror.org/05gq02987"
affiliationIdentifierScheme="ROR">Brown University</affiliation>
      <affiliation affiliationIdentifier="grid.268117.b" affiliationIdentifierScheme="GRID"
schemeURI="https://grid.ac/institutes/">Wesleyan University</affiliation>
    </creator>
```

# So what?

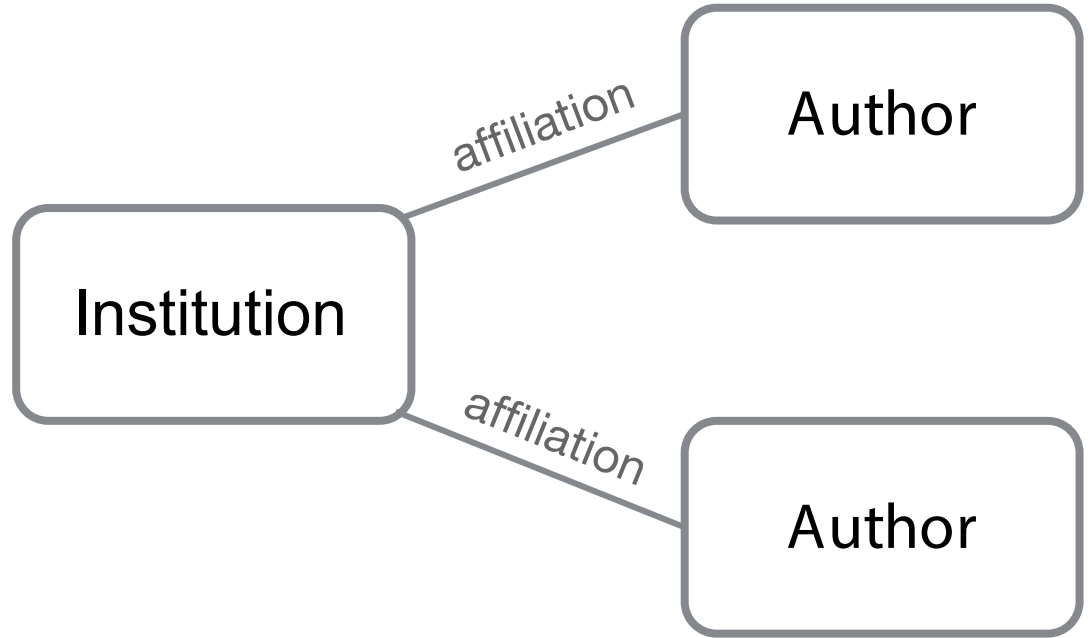
*As a university administrator, I want to get a list of all datasets and software published by our researchers, so that I can get a comprehensive view of our research outputs.*

This has been unnecessarily hard for too long.

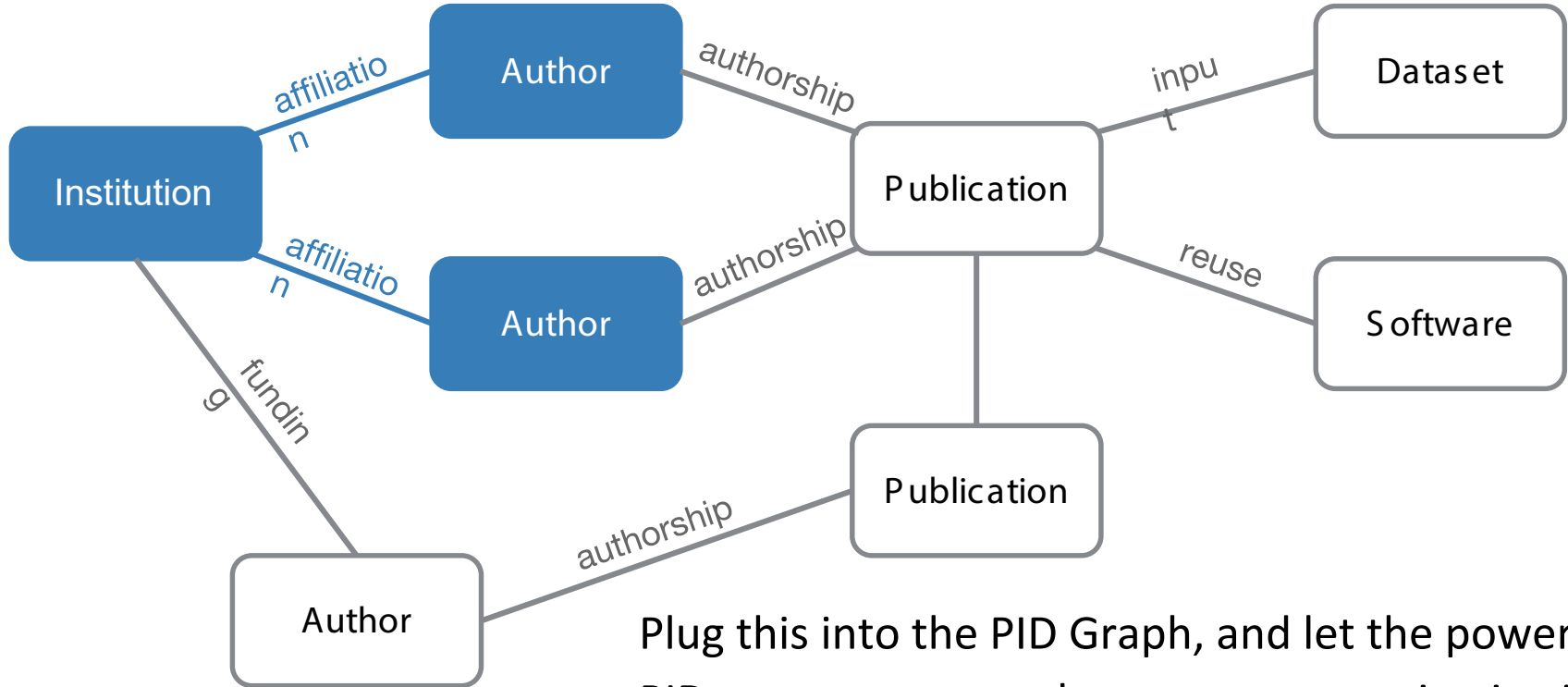
All because we couldn't definitively say who **“our researchers”** are.

# ROR and the PID Graph

**ROR** definitively identifies your institution, no matter how many names it has, so you can make sure all your authors are affiliated with the right place.



# ROR and the PID Graph



Plug this into the PID Graph, and let the power of PIDs connect research outputs to your institution.

PRETTIFY

HISTORY

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

COPY CURL

1



Hit the Play Button to  
get a response here

DOCS

SCHEMA

All your datasets in 40 seconds