Project FREYA

Connected Open Identifiers for Discovery, Access and Use of Research Resources

Frances Madden, British Library (orcid.org/0000-0002-5432-6116)

ORCID Tech-Workshop, 25 November 2020

www.project-freya.eu | twitter: @freya_eu | frances.madden@bl.uk

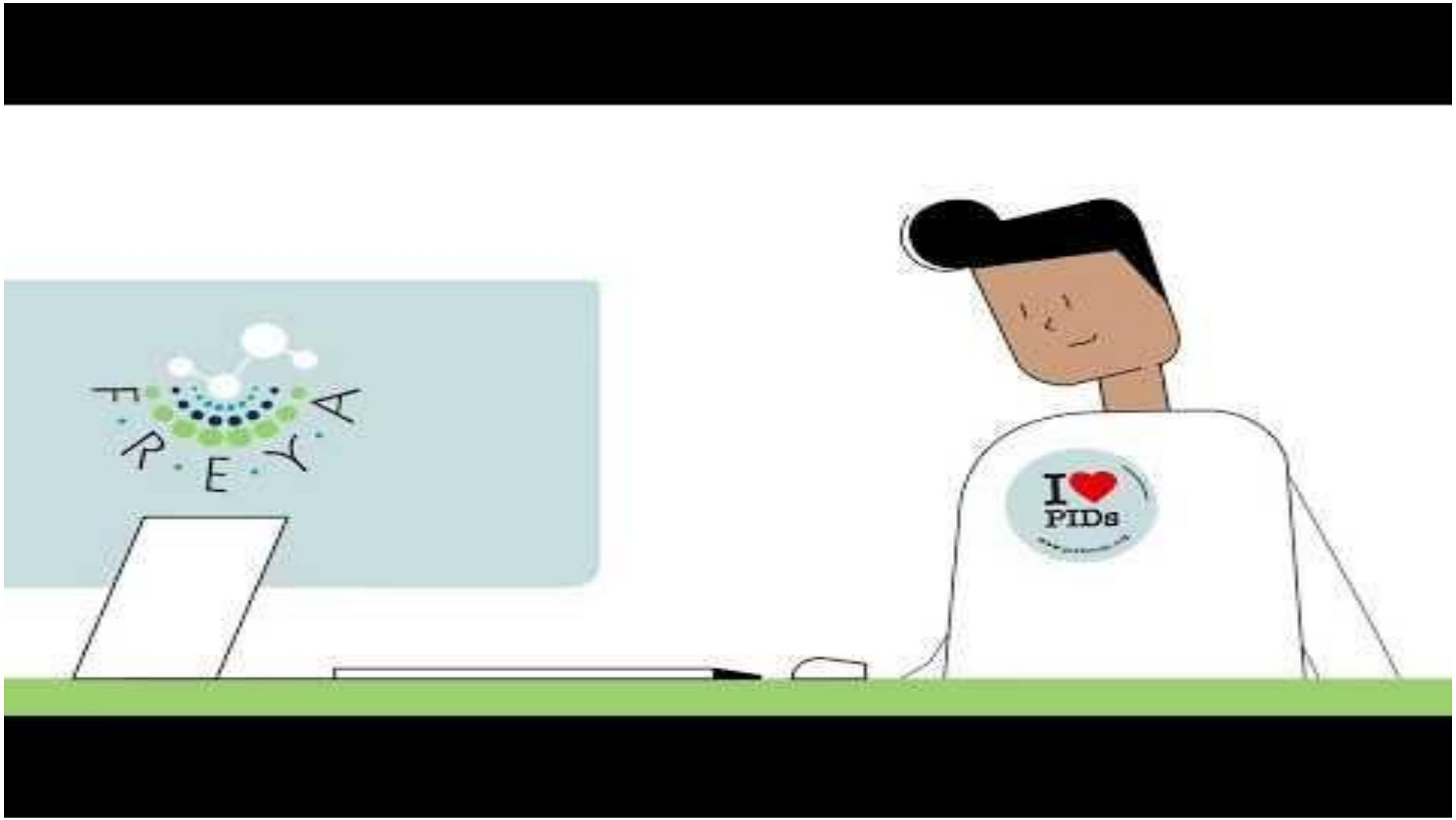
European Commission | Horizon 2020 European Union funding for Research & Innovation

The FREYA project

FREYA helps to build and extend the infrastructure for PIDs as a core component of Open Science

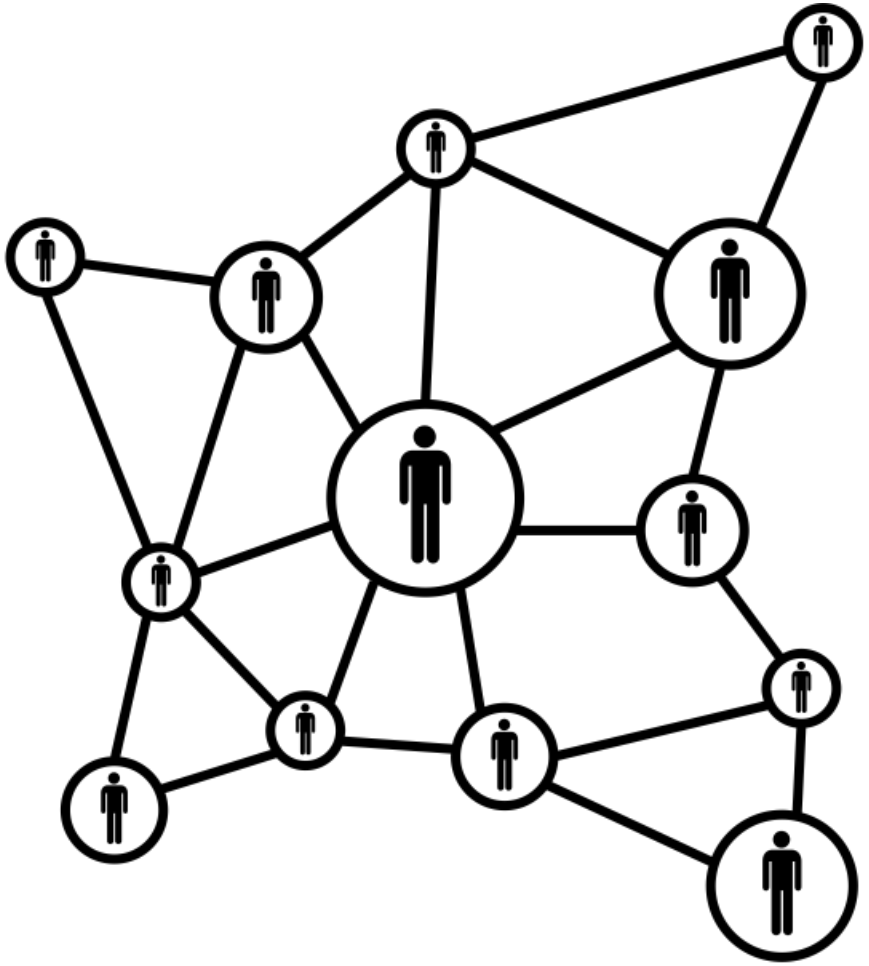
- Improved discovery and access of research data
- Develop new types of PIDs
- Demonstrate disciplinary PID systems
- Integrate services with the EOSC

The Power of PIDs

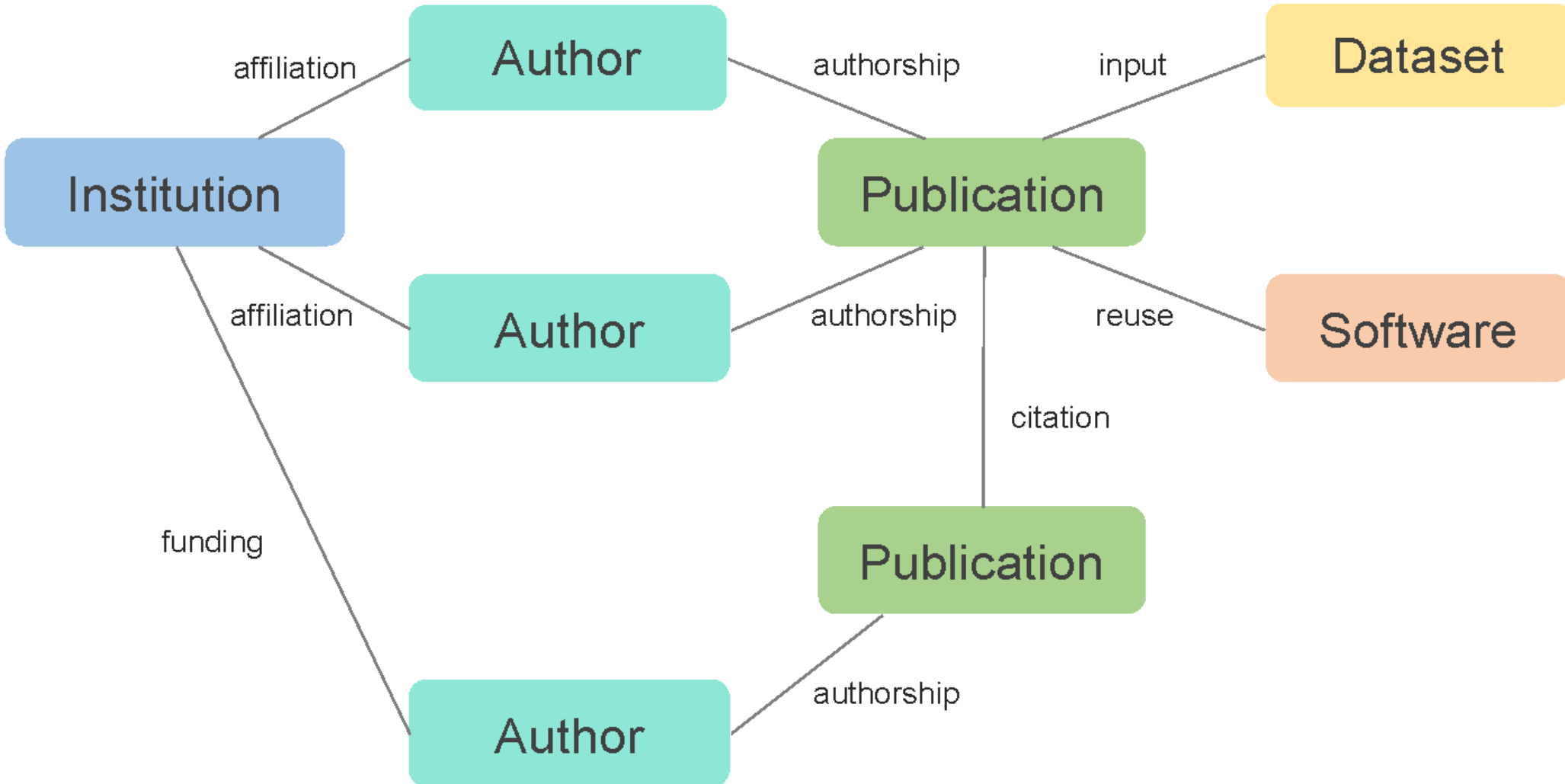


https://www.youtube.com/watch?v=9G4EMJCwCw4

“graph”

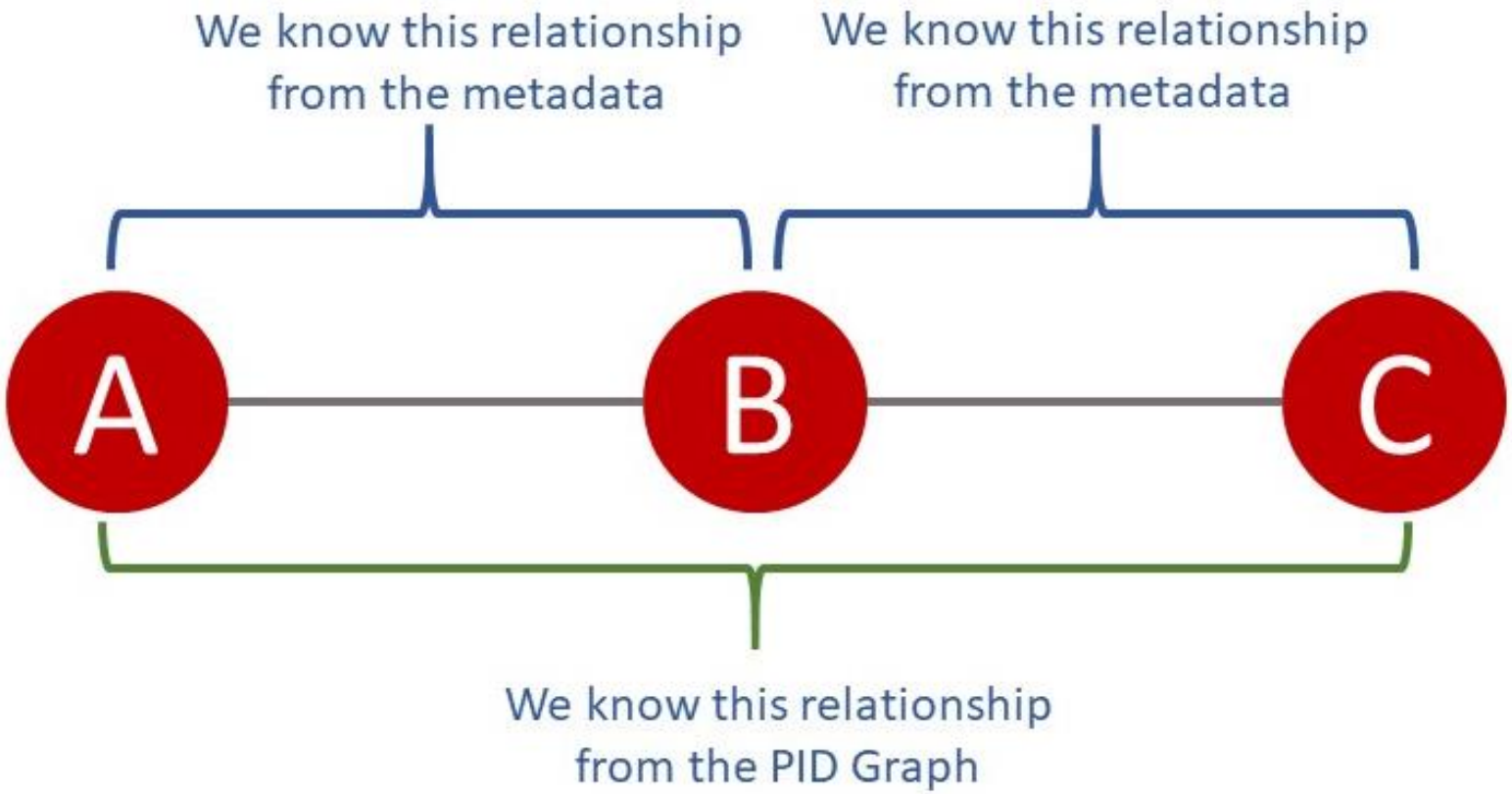


Research as a graph



The PID Graph

Connecting PIDs together so we can discover relationships two ‘hops’ away



From this…

Photo by Stephen Pedersen on Unsplash

To this…Using PIDs

Photo by Jim Witkowski on Unsplash

PIDs Connect

PIDs are the backbone of connecting research outputs because…

- PIDs are unique
- PIDs are persistent

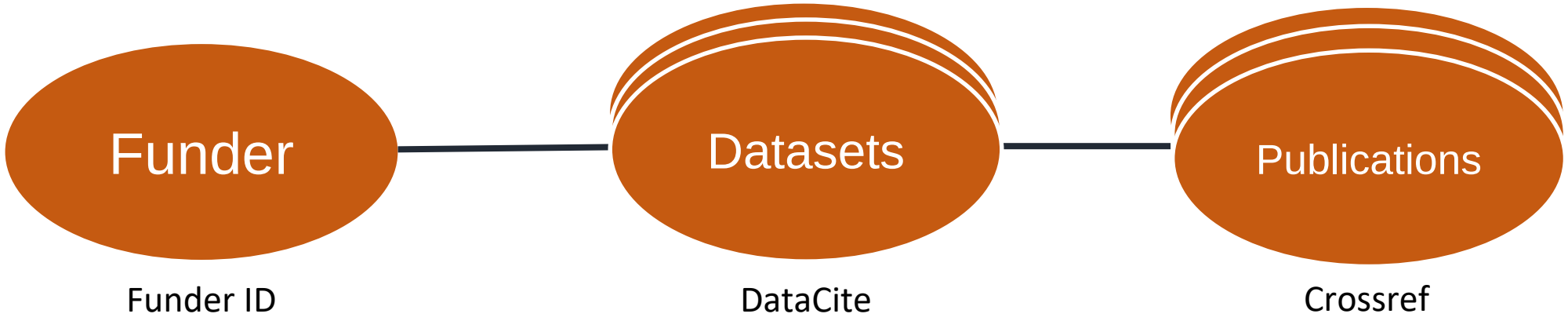
Can use these aspects to connect different research entities reliably.

Lots of this functionality is there already but we are not utilising it fully.

Photo by Stephanie Ronquillo on Unsplash

An example

Show all datasets funded by the European Commission that have been cited in journal articles.



How has the graph been applied?

DataCite GraphQL API

- Production release in May 2020
- Allows searching by several entities e.g. DOI (Crossref and DataCite), ORCID, ROR etc.
- Used as basis for several DataCite services
- PID Services Registry (pidservices.org)
- DataCite Commons (commons.datacite.org)
- Jupyter Notebooks (pidnotebooks.org)

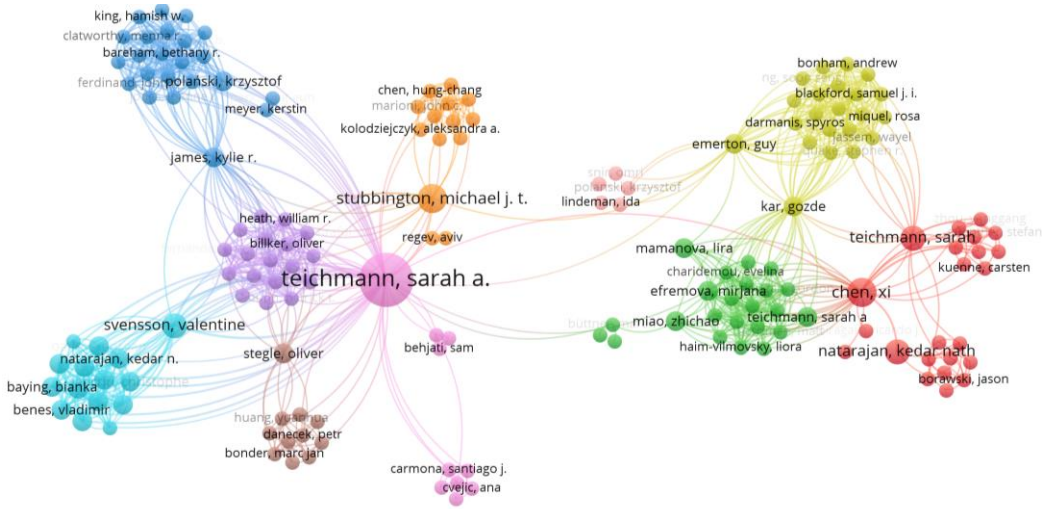
Jupyter Notebooks (pidnotebooks.org)

- 10 Notebooks addressing user stories identified by FREYA
- Demonstrates potential of PID Graph, includes visualisations etc.
- Written in Python using GraphQL API
- Available in Github (https://github.com/datacite/pidgraph-notebooks-python)
- Can be accessed in Binder
- All notebooks have a DOI

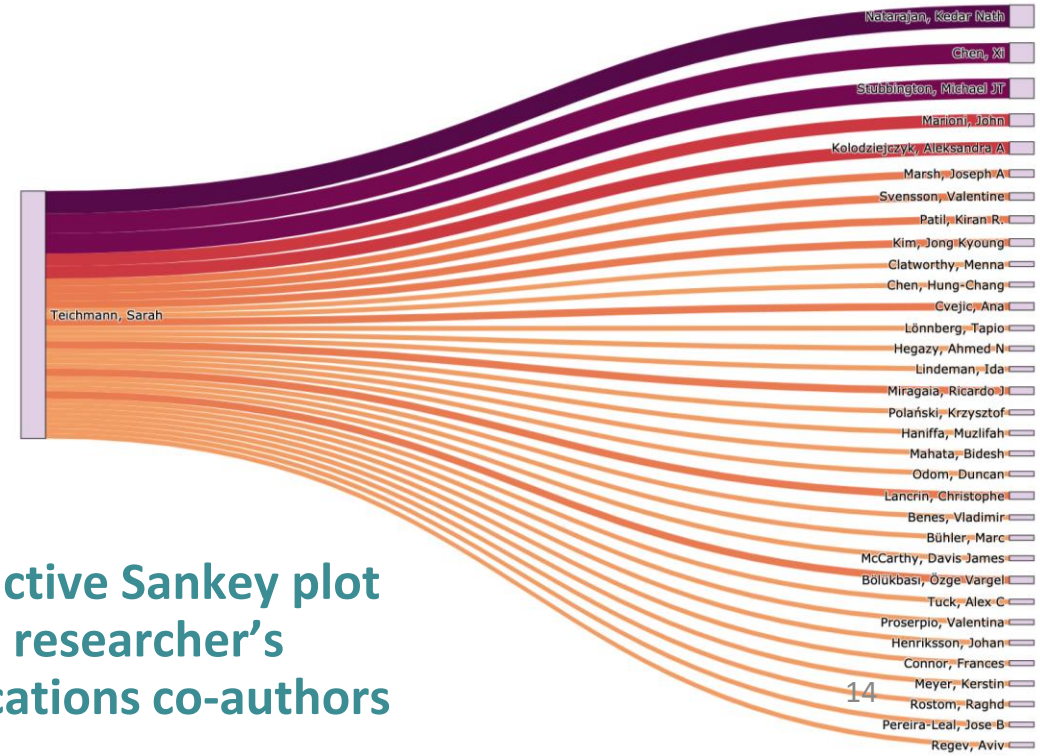
Jupyter Notebooks (pidnotebooks.org)

User Story 9: As a bibliometrician, I want to know all the co-authors of a particular researcher, so that I can do a network analysis of the researcher's collaborations.

Example: Dr Sarah Teichmann (https://orcid.org/0000-0002-6294-6366)



Visualise publications in VOS Viewer



Interactive Sankey plot of the researcher's publications co-authors

DataCite Commons (commons.datacite.org)

https://ror.org/05dhe8b71

British Library

BL

Links
Homepage
Wikipedia

Other Identifiers
GRID grid.36212.34
Wikidata Q23308

United Kingdom | Archive

https://ror.org/05dhe8b71

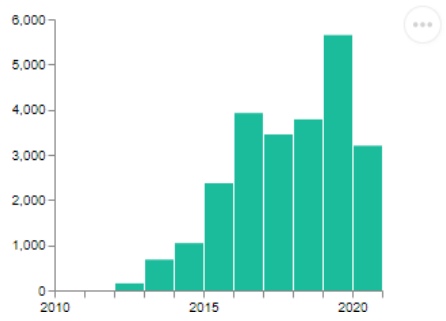
Share

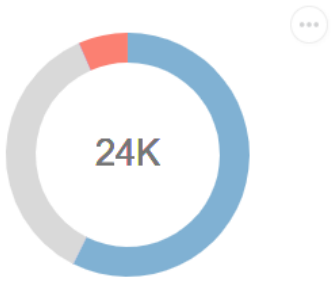
Email
Twitter
Facebook

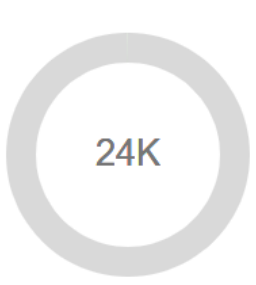
24,264 Works

Publication Year

2020	3,200
2019	5,649
2018	3,783
2017	3,449
2016	3,923
2015	2,371
2014	1,044
2013	680







DataCite Commons

The following main data sources are used in DataCite Commons for a total of currently 40,570,143 records:

DataCite
20,373,711 Works
100% of identifiers and metadata.

Crossref
9,872,811 Works
8.30% of identifiers and metadata. Import is ongoing.

ORCID
10,225,023 People
100% of identifiers. Personal and employment metadata.

ROR
98,598 Organizations
100% of identifiers and metadata.

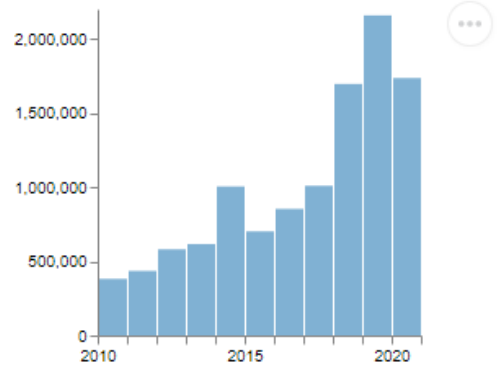
Additional information comes from these data sources:

- Wikidata: inception year, geolocation and Twitter account for organizations
- Unpaywall: download link for Open Access content via Crossref

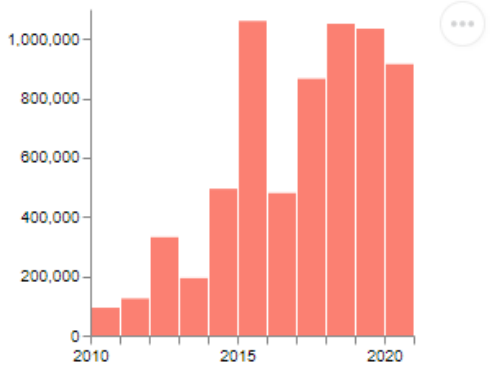
Works

DataCite Commons currently includes 30,246,522 works, with identifiers and metadata provided by DataCite and Crossref. For the three major work types publication, dataset and software, the respective numbers by publication year are shown below.

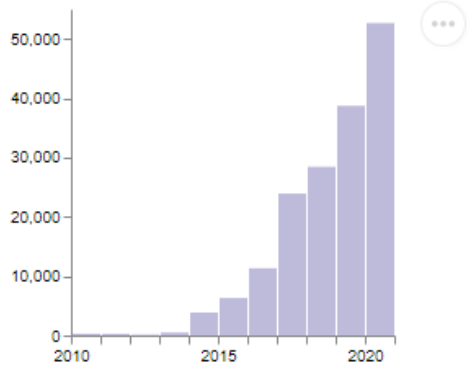
15,729,028 Publications



7,704,040 Datasets



167,477 Software



6,342,929 out of all 30,282,045 (20.95%) works have been cited at least once, including 0.93% of works registered with DataCite, and 62.33% of works registered with Crossref.

DataCite Commons

https://orcid.org/0000-0002-5432-6116

Frances Madden

Other Profiles

ORCID
Impactstory
Europe PMC

United Kingdom of Great Britain and Northern Ireland

https://orcid.org/0000-0002-5432-6116

Share

Email
Twitter
Facebook

74 Works

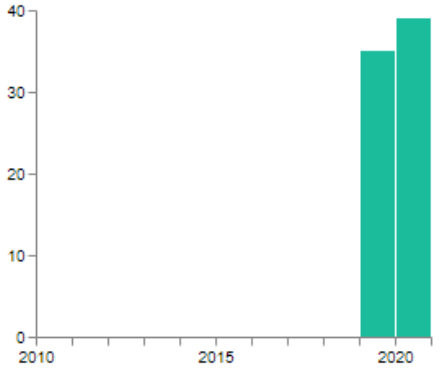
Publication Year

- 2020 39
- 2019 35

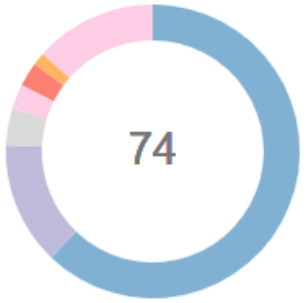
Work Type

- Text 46
- Audiovisual 10
- Software 10
- Other 3
- Dataset 2
- Interactive Resource 2
- Collection 1

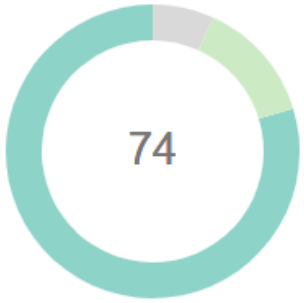
Publication Year



Work Type



License



From Standard to Community Resource: A View on ISNIs and ORG IDs

Torsten Reimer & Frances Madden
Presentation published 2019 in

FEEDBACK

DataCite Commons

DataCite Commons

fundingReferences.awardNumber:777523

About Support

This Page | Works | People | Organizations

https://ror.org/00k4n6c32

European Commission

EC

Links

Homepage
Wikipedia

Other Identifiers

GRID grid.270680.b
Crossref Funder ID 10.13039/501100000780
Crossref Funder ID 10.13039/501100000893
Crossref Funder ID 10.13039/501100000891
Crossref Funder ID 10.13039/501100000894
Crossref Funder ID 10.13039/501100000887
Wikidata Q8880
Wikidata Q20855594

Belgium | Government

https://ror.org/00k4n6c32

Share

Email
Twitter
Facebook

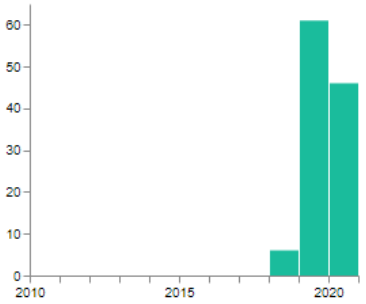
114 Works

Publication Year

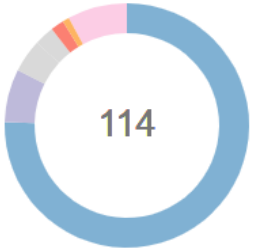
- 2020 46
- 2019 61
- 2018 6
- 2005 1

Work Type

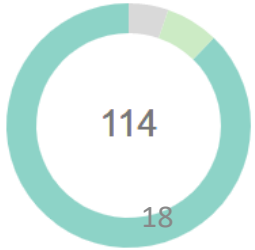
Publication Year



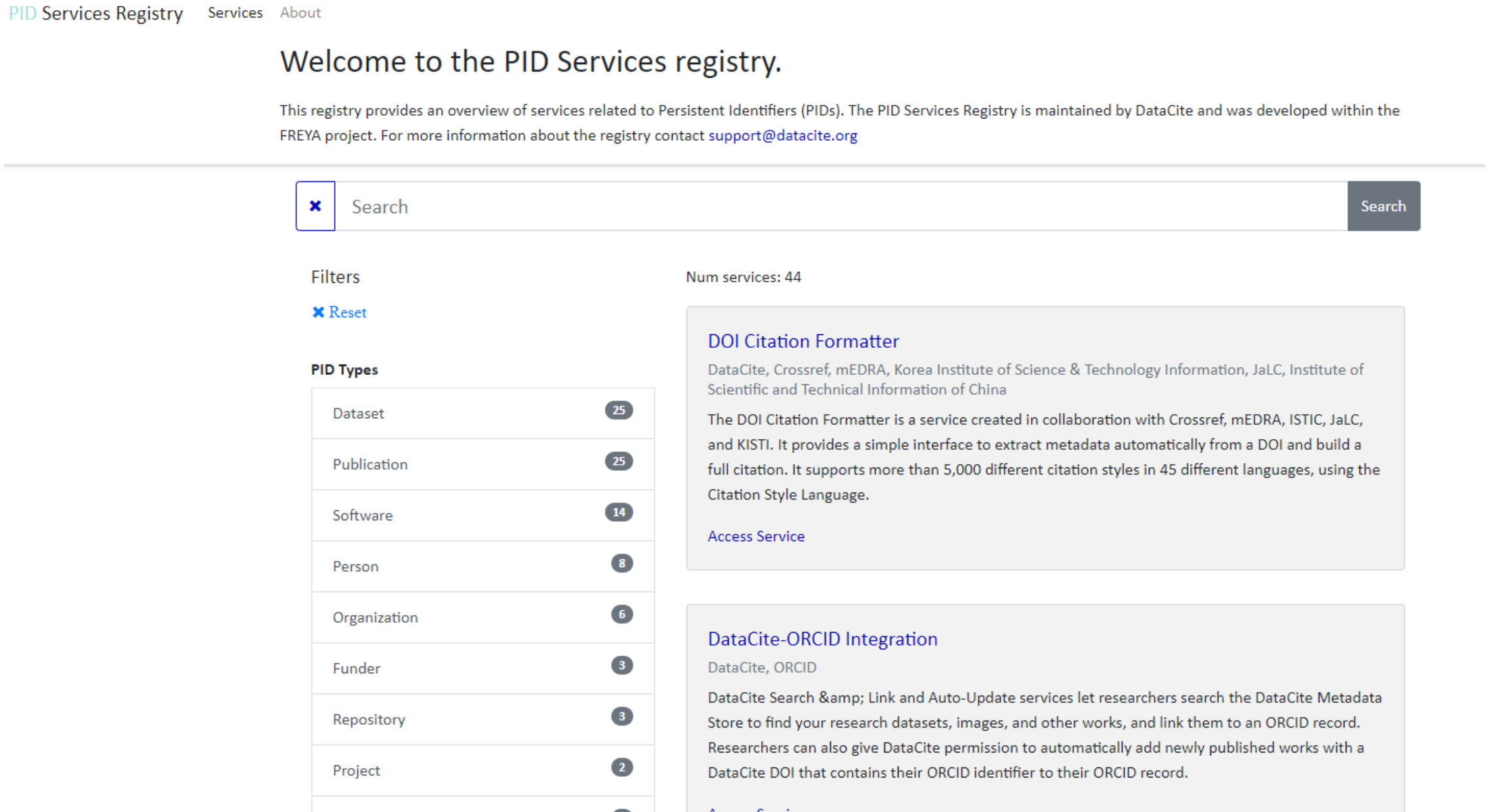
Work Type



License



PID Services Registry



Disciplinary Applications - BL

Shared research repository

Integration of new PID types (org and funder IDs) into metadata model

DATASET

UK Doctoral Thesis Metadata from EThOS

British Library ; Rosie, Heather

July 2020

ABSTRACT

The data in this collection comprises the bibliographic metadata for all UK doctoral theses listed in EThOS, the UK's national thesis service. We estimate the data covers around 98% of all PhDs ever awarded by UK Higher Education institutions, dating back to 1787. Thesis metadata from every PhD-awarding university in the UK is included. You can investigate and re-use this unique collection of UK universities' PhD thesis data to analyse trends in postgraduate research, make connections between researchers, apply large data analysis, improve citation of theses and many more applications.

Disci

Shared r

Integrati

del

MAGAZINE ARTICLE

MicroPasts: An Innovative Place for Progressing Research

Keinan-Schoonbaert, Adi; Bevan, Andrew; Pett, Daniel; Bonacchi, Chiara; Wilkin, Neil; Wexler, Jennifer; Sparks, Rachael

2014

ABSTRACT

Archaeology has always attracted enthusiastic volunteers, who have participated in excavations, surveys, site recording or artefact handling, as well as museum-related tasks such as engaging with visitors or helping with curatorial duties. However, most data have been produced by specialists. More often than not the knowledge remains in the academic or professional domains. Poorly known evidence and objects fill museum storage rooms and university archives. Traditional methods and resources do not seem sufficient to give these assets the public attention they deserve.

FILES

This is a metadata only record.

Funder Arts and Humanities Research Council (Award number: AH/M00953X/1)

Organisational unit	Digital Scholarship
Project name	MicroPasts
Funder	Arts and Humanities Research Council (Award number: AH/M00953X/1)
Journal title	British Archaeology
Volume	139

Where to find out more?

- Website – Disciplinary applications, deliverables www.project-freya.eu
- YouTube Channel – webinar recordings etc.

www.youtube.com/channel/UCQ5Jp19cvtVLPxUB2WVO5CA

- Power of PIDs Video and PID Graph Demo
- the PID Forum discussion, questions etc. www.pidforum.org/
- Zenodo Community - Slides etc. zenodo.org/communities/freyaproject/

Image Credits

Slide 5

- alex setyawan from Noun Project
- Made from Noun Project
- arjuazaka from Noun Project
- Adrien Coquet from Noun Project
- zidney from Noun Project

Thank you!

Frances Madden, British Library

orcid.org/0000-0002-5432-6116

frances.madden@bl.uk

www.project-freya.eu | twitter: @freya_eu | frances.madden@bl.uk