

Insights from the network of repositories



Matt Buys

DataCite Executive Director
FAIRCORE4EOSC ...

*Insights from the network of repository
managers*

Context & Perspective

DataCite

DataCite is a global community of research organisations that shares 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 global non-profit membership organization, we work with 3500+ repositories in the world to provide DOIs for research outputs and resources.

FAIRCORE4EOSC

Developing EOSC-Core components to enable a FAIR EOSC ecosystem

The FAIRCORE4EOSC project focuses on the development and realisation of core components for the European Open Science Cloud (EOSC). Supporting a FAIR EOSC and addressing gaps identified in the Strategic Research and Innovation Agenda (SRIA). Leveraging existing technologies and services, the project will develop nine new EOSC-Core components aimed to improve the discoverability and interoperability of an increased amount of research outputs.

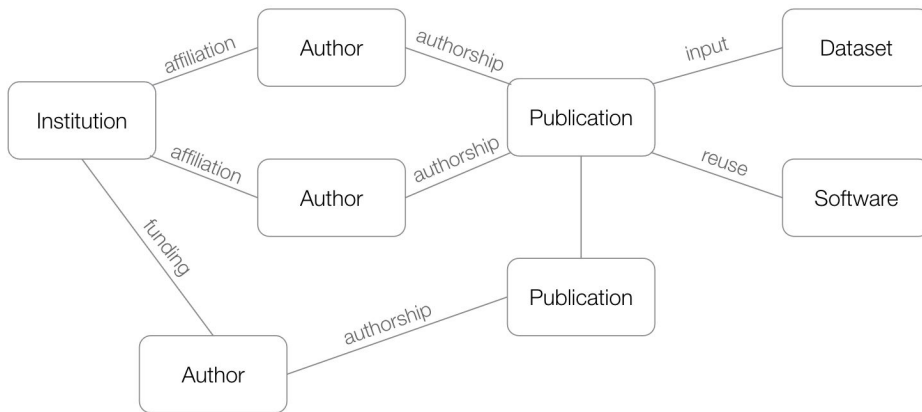
FAIRCORE4EOSC has received funding from the EU's Horizon Europe research and innovation programme under Grant Agreement no. 101057264

PID Graph Background

PID Graph – Concept. The connected searchable Graph of publications, datasets, other research outputs, people and organizations.

PID Graph – User Stories. At the beginning of the FREYA project we collected more than 40 user stories (see <https://pidforum.org/c/pid-best-practices/8>) that were difficult to address with currently available services.

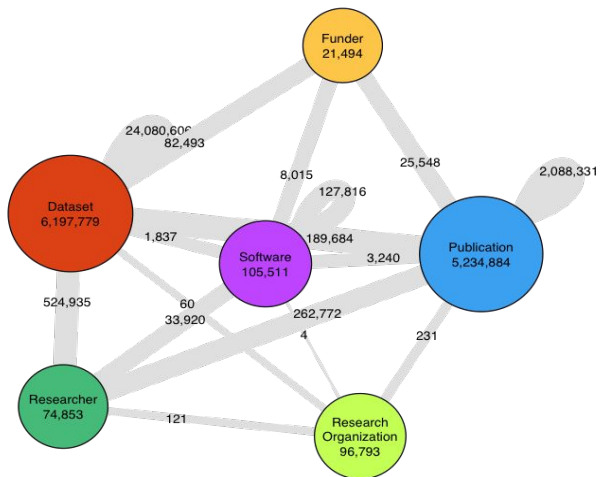
PID Graph – Architecture. We discussed and prototyped PID Graph architecture to address user stories, taking into account existing PID infrastructure, in particular from unfunded partners, and sustainability beyond FREYA project duration.



PID Graph Scaling

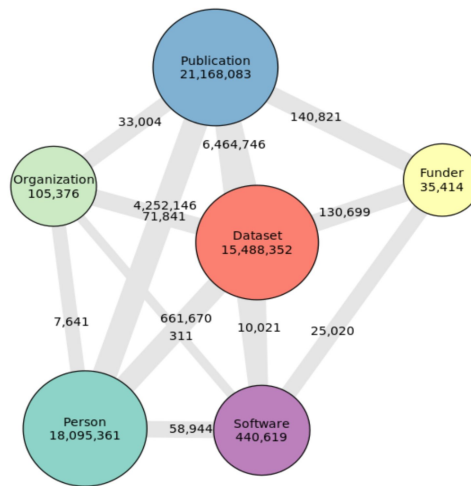
PID Graph

Nodes and connections (as of September 2019)



PID Graph

Nodes and connections (as of October 2023)



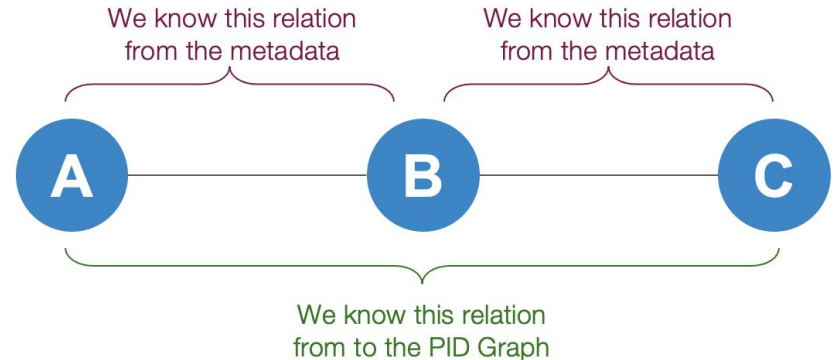
Insights: Repositories

Repositories' persistent identifiers (PIDs) and metadata feed into the PID Graph, helping to connect research outputs and resources.

FAIRCore4EOSC is making it easier to harvest metadata and understand relationships between hundreds of millions of entities.

Unique PIDs for researchers and outputs are key to linking pieces of the research landscape.

By linking PIDs through metadata relations, we can now discover connections at least two “hops” away. Bringing rigor to the scholarly records.



SKG-IF: Our Plans

The Scientific Knowledge Graphs Interoperability Framework (SKG-IF) is maturing, with a focus on adopting a standardized approach using JSON-LD for representing metadata.

We are committed to exploring how to implement and contribute to this initiative, ensuring alignment with our goals of improving research discovery and accessibility.

We advocate for interoperability that goes beyond standard frameworks, aiming for solutions that enhance metadata exchange across the full research lifecycle.

As part of our commitment to open metadata, we plan to offer an annual data dump in SKG-IF format, enhancing access to interconnected research information.

- Additionally, there is potential to extend this offering through a REST API, allowing for content negotiation and seamless data integration across platforms.

Moving forward, we will focus on planning the effective implementation of an API to support broader adoption and use of the SKG-IF format.

PID Graph: Impact for Repositories

Enhanced Reuse:

Regular data dumps related to community profiles, available through the PID Graph, support repositories in contributing to the EOSC community and beyond.

Streamlined Ingestion:

The API and data dumps, featuring DOIs and metadata from DataCite, ensure seamless integration into EOSC and related services.

Data Usage and Stats:

Integrated statistics on DataCite DOI usage help repositories track and analyze the reach and impact of their content.

Contextual Metadata:

By linking PID entities, the PID Graph enriches repositories' metadata, uncovering connections between research outputs and resources, affiliations, and contributors.

Aggregated PID Links:

Supporting PID link aggregation helps repositories highlight the interconnectedness of research, fostering deeper insights and visibility.