Making Research FAIR
With PID-centric workflows

Xiaoli Chen
May 16 2023
Edinburgh Open Research Conference

@datacite
@datacite@openbiblio.social
About DataCite

Global non-profit membership organization working with 2700+ repositories in the world to provide DOIs for research outputs and resources.
Our community

2700+ Repositories
280+ Members
50 Countries
39m+ DOIs
1200+ Organizations

(March 2022)
Implementing FAIR Workflows
A proof of concept study in the field of consciousness

To implement an exemplar FAIR and Open workflow based on the reality of an entire research lifecycle.
What makes a workflow FAIR

**FAIR Entities**
- Uniquely identified resources associated to a project
  - Researcher (ORCID ID)
  - Research organization (ROR ID)
  - Funding agency (ROR ID) and grant (Grant-ID)

**FAIR Practices**
- Sharing various types of interim outputs
  - Data Management Plan, Pre-registration, Protocol, Preprint, Code, Dataset, etc.

**FAIR Supporting Structures**
- Tools and platforms that integrate PIDs and metadata workflow
  - DMP authoring tools, Metadata templates, Data repositories, Notebooks, Collaborative research platforms, etc.

**FAIR Outputs**
- Assigning PIDs to outputs with rich metadata annotation
  - Essential descriptive and connection metadata
    - Connection between inputs and outputs
    - Relations between outputs
  - Domain specific metadata
    - Disciplinary ontological information
    - Experimental setup
What does it look like
A project centric view

1. Grant Application
   Researcher to use relevant PIDs (team member, affiliation, funder) in the grant application process.

2. Register Grant
   Funder to register grant and assign Grant ID, provide instruction for usage.

3. Submit Data Management Plan
   Researcher to develop and submit DMP, use DMP-ID in subsequent reporting.

4. Research design
   Researcher to develop research design, metadata model and experimental protocol.

5. Experiment conducted
   Researcher to conduct experiments proposed in pre-registration according to protocol.

6. Research code, data, other outputs deposited
   Researcher to deposit protocols, data, conference posters, etc.

7. Preprint submitted
   Researcher to share manuscript on preprint server.

8. Research article published
   Researcher to select appropriate publishing channel and submit manuscript to journal, journal to register articles and submit metadata to Crossref.
# Research Output and Workflow – Study 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TWCF</td>
<td>OSF</td>
<td>DMPTool</td>
<td>Github</td>
<td>Protocol.io</td>
<td>Dryad</td>
<td>PsyArxiv</td>
<td>Journal ChronosHub</td>
<td>Zenodo</td>
</tr>
<tr>
<td>PID Grant ID (DOI)</td>
<td>PID</td>
<td>PID DMP-ID (DOI)</td>
<td>PID DOI</td>
<td>PID DOI</td>
<td>PID DOI</td>
<td>PID DOI</td>
<td>PID DOI</td>
<td>PID DOI</td>
</tr>
</tbody>
</table>

**Investigators**

- Zefan Zheng (PhD student)
- Lucia Melloni (PI, advisor)
- Melissa Vo (PI, co-advisor)

**Institutions**

- Max Planck Institute of Empirical Aesthetics
- Goethe University Frankfurt

**Investigators**

- Zefan Zheng
- Lucia Melloni
- Melissa Vo

**Institutions**

- Max Planck Institute of Empirical Aesthetics
- Goethe University Frankfurt
What does it look like
A PID centric view

The PID-optimised research cycle

Actors and entities
- Funding organisations
- Research performing organisations
- Research contributors
- Research output platforms/publishers

Priority PIDs and registries
- CrossRef
- DataCite
- ORCID
- Digital Object Identifier (DOI)
- Research Activity Identifier (RAID)
- Research Organisation Registry (ROR)
What does it look like
An output centric view
Common themes of discussion around implementing FAIR practices in research.

- What make research FAIR
- Community and support
- Determining what to share
- Choosing the right tools
- Time management
- Domain-specific metadata
- Navigating data sharing policies
- Common pitfalls
- Licensing and ethics
Researcher guide

- Providing Step-by-step instructions for the key activities
- Giving local contexts to the FAIR principles
- Crash-course of PIDs and metadata
- Pooling community resources and support channels
EOSC & RDA Grant
In partnership with Research Space

“Data and metadata interoperability through the incorporation of PIDs in a research infrastructure”

· European Open Science Cloud Future & RDA grant

· RDA Working with PIDs in Tools Interest Group

· Collaboration between DataCite & Research Space

· RSpace = electronic lab notebook + sample management system, deeply integrated, and connected to an ecosystem of RDM tools
EOSC & RDA Grant
Grant outputs

1. **User research & use case definition** to understand PID workflows within research tools & their variations, iterative process

2. **High-level design** for incorporation of DataCite DOIs (incl. IGSNs) into the RSpace digital research platform

3. **Implementing a proof-of-concept** for IGSN support in RSpace: registering, populating metadata, sending to DataCite & publishing landing page

4. **Interoperability guideline document** for integrating PIDs in research infrastructures based on high-level and dev-level findings
Mentimeter