



Australian Research Data Commons

# Persistent Identifiers for Research

PRESENTED BY

Siobhann McCafferty

17/06/2021



The Australian Research Data Commons is enabled by NCRIS.



**We acknowledge and celebrate the  
First Australians on whose traditional lands  
we meet, and we pay our respect to the  
elders past and present.**

# PIDs for ARDC Projects Investment Policy

The ARDC's suite of PID services, combined with those available through the AAF led Australian ORCID Consortium, form the backbone of enabling FAIR research outputs via PIDs in Australia. Through the ARDC's [Data and Services](#) portfolio we provide to the sector:

- [DOIs](#) (for research data, software, grey literature, instruments)
- [RAIDs](#) (for research projects)
- [IGSNs](#) (for physical objects collected during the course of research)
- [PURLs](#) (for research grants)
- [Handles](#) (for data and for the foundation of RAIDs).

PIDs are captured and displayed in [Research Data Australia](#) where international partners such as OpenAire harvest the information and display it in their own discovery portals, facilitating further exposure and impact of Australian research.

---

What is a persistent identifier?

(with thanks to PIDapalooza)

# persistent identifier

an organization  
made a promise  
to keep it alive

globally unique  
string of  
characters

(known as PIDs to their friends)

# PIDs for people, places, and things in the research community

PIDs for people (researchers) include ISNIs and ORCID iDs



PIDs for places (research organizations) include GRID and ROR

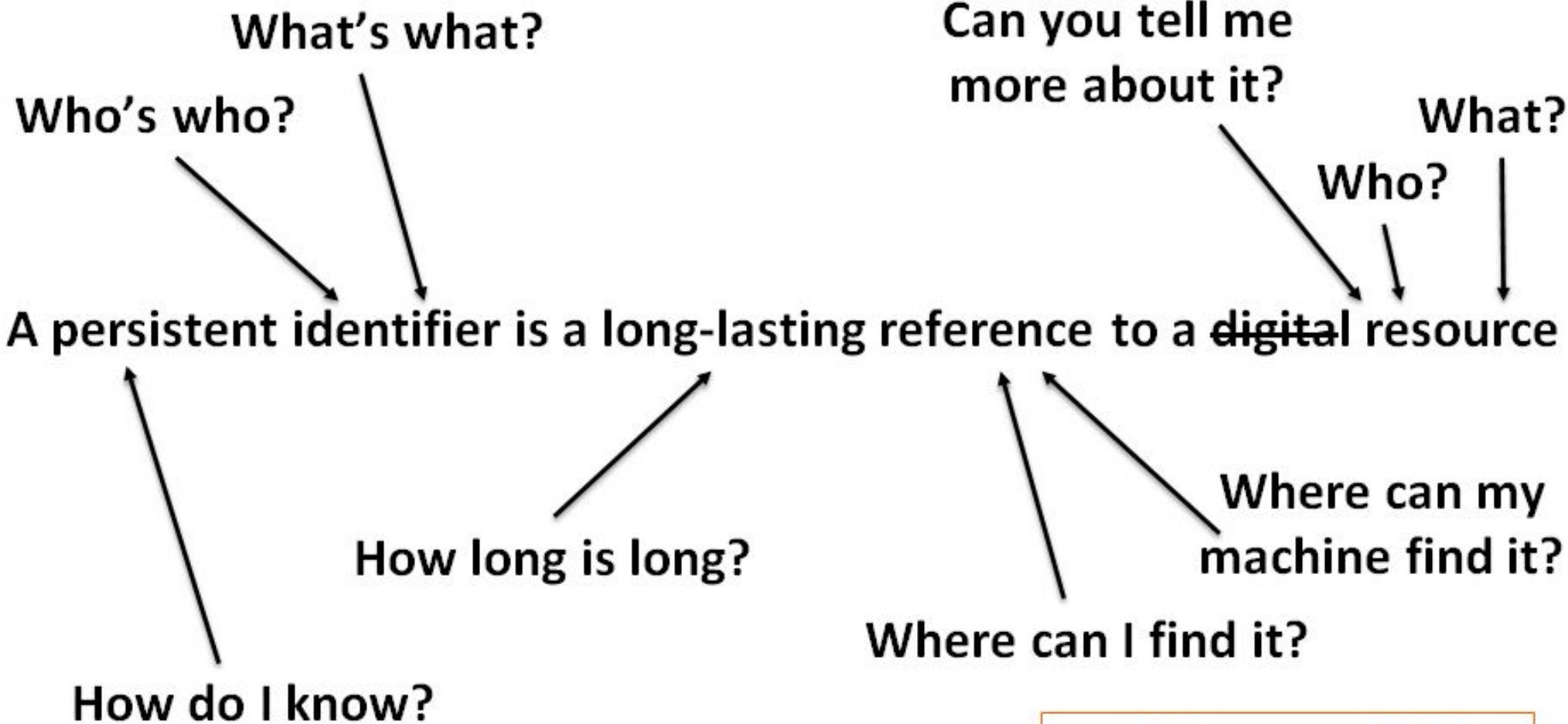


PIDs for things (research outputs/inputs like grants, reviews, preprints, projects, etc.) include Crossref and DataCite DOIs, IGSNs, RAiDs, and more



**Provenance**

**Metadata**



**Policies and Guarantees**

**Machine-readability**

---

What can PIDs \*do\* and why are they important?




---

PIDs  
disambiguate

# Robin Dasler

## ORCID iD

 <https://orcid.org/0000-0002-4695-7874>

 [Print view](#) 

## Also known as

RH Dasler, RL Dasler, RL Howard,  
Robin Howard

## Other IDs

[ResearcherID: N-9035-2013](#)

---

# PIDs support linking

## References

[Abd Ellah and Abouelmagd, 2016](#) N.H. Abd Ellah, S.A. Abouelmagd  
**Surface functionalization of polymeric nanoparticles for tumor drug delivery: approaches and challenges**  
Expert Opin. Drug Deliv., 1–14 (2016),  
[10.1080/17425247.2016.1213238](https://doi.org/10.1080/17425247.2016.1213238)

[Google Scholar](#)

[Abouelmagd et al., 2016](#) S.A. Abouelmagd, F. Meng, B.-K. Kim, H. Hyun, Y. Yeo  
**Tannic acid-mediated surface functionalization of polymeric nanoparticles**  
ACS Biomater. Sci. Eng. (2016), p. 6b00497,  
[10.1021/acsbiomaterials.6b004](https://doi.org/10.1021/acsbiomaterials.6b004)

[Google Scholar](#)

[Ahmed et al., 2016](#) S. Ahmed, S. Annu, S.S. Yudha  
**Biosynthesis of gold nanoparticles: a green approach**  
J. Photochem. Photobiol. B: Biol., 161 (2016), pp. 141-153,  
[10.1016/j.jphotobiol.2016.04.034](https://doi.org/10.1016/j.jphotobiol.2016.04.034)

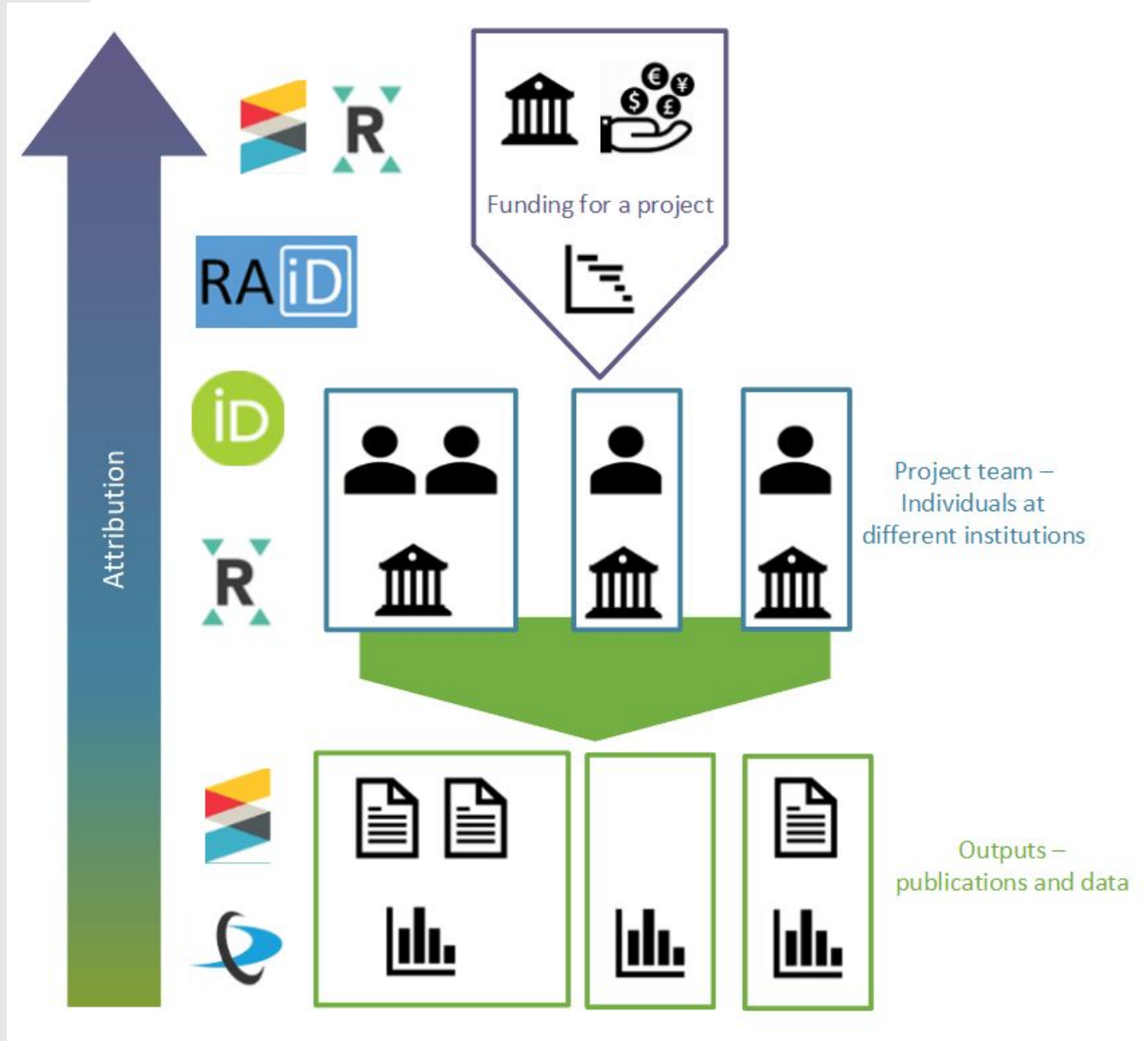
[Article](#)  [Download PDF](#) [View Record in Scopus](#)

[Google Scholar](#)

[Akhavan et al., 2011](#) O. Akhavan, R. Azimirad, S. Safa, E. Hasani

---

# PIDs enable interoperability



---

# PIDs help make research FAIR

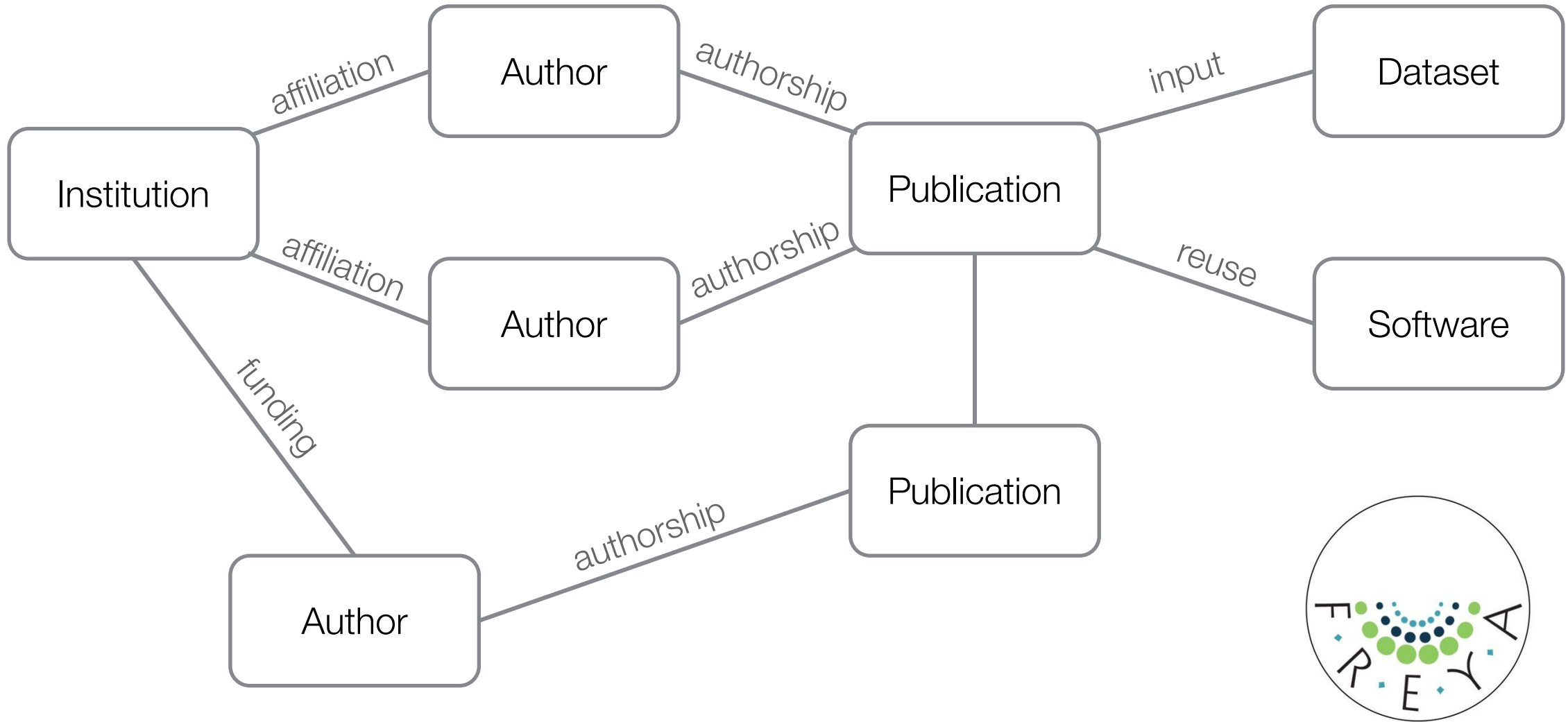
|                                     |   |
|-------------------------------------|---|
| <b>Data should be Findable</b>      | <ul style="list-style-type: none"><li>F1. (meta)data are assigned a globally unique and persistent identifier (DOI)</li><li>F2. data are described with rich metadata</li><li>F3. metadata clearly and explicitly include the identifier of the data it describes</li><li>F4. (meta)data are registered or indexed in a searchable resource</li></ul>   |
| <b>Data should be Accessible</b>    | <ul style="list-style-type: none"><li>A1. (meta)data are retrievable by their identifier using a standardized communications protocol<ul style="list-style-type: none"><li>A1.1 the protocol is open, free, and universally implementable</li><li>A1.2 the protocol allows for an authentication and authorization procedure, where necessary</li></ul></li><li>A2. metadata are accessible, even when the data are no longer available</li></ul> |
| <b>Data should be Interoperable</b> | <ul style="list-style-type: none"><li>I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.</li><li>I2. (meta)data use vocabularies that follow FAIR principles</li><li>I3. (meta)data include qualified references to other (meta)data</li></ul>  |
| <b>Data should be Reusable</b>      | <ul style="list-style-type: none"><li>R1. meta(data) are richly described with a plurality of accurate and relevant attributes<ul style="list-style-type: none"><li>R1.1. (meta)data are released with a clear and accessible data usage license</li><li>R1.2. (meta)data are associated with detailed provenance</li><li>R1.3. (meta)data meet domain-relevant community standards</li></ul></li></ul>   |

---

PIDs support a trustworthy research infrastructure



Image: University of Washington Office of Research



**Connected PIDs form a graph...**



# National PID Infrastructure

# Australian National Research Infrastructure Roadmap

“...nationally significant assets, facilities and services to support leading edge research...”

<https://docs.education.gov.au/node/43736>





# ARDC PIDs Policy - Highlights

Persistent identifiers (PIDs) are a core component of national infrastructure and key to world class, global research infrastructure. As specified in the FAIR Data Principles, the use of Persistent Identifiers (PIDs) is critical as they provide global methods to uniquely identify and connect entities in the research system such as researchers, funders, organisations, articles, datasets, software, and samples.

- By linking these entities and enabling research provenance and attribution, persistent identifiers lay the foundation for **improved tracking of research impact**.
- By precisely identifying the inputs and outputs of research, PIDs **contribute to research integrity**
- By linking scientific concepts across systems, they **enable research innovation and efficiency**

[https://ardc.edu.au/about\\_us/policies-and-guidelines/persistent-identifiers-policy/](https://ardc.edu.au/about_us/policies-and-guidelines/persistent-identifiers-policy/)

# ARDC PIDs Policy - Highlights

ARDC therefore places high value on the use of PIDs in research and offers a suite of PID services to the Australian research sector to facilitate their uptake

As part of a strategy to optimise sustainable longevity and global acceptance, ARDC PID services are offered in partnership with international PID service providers (DataCite, ORCID, CrossRef, etc).

Selection of PIDs in the ARDC services suite is based on an assessment of the attributes of good PID systems such as fitness for purpose, effective technical infrastructure, good governance, broad community adoption, cost effectiveness, and sustainability.

**The ARDC's suite of PID services, combined with those available through the AAF led Australian ORCID Consortium, form the backbone of enabling FAIR research outputs via PIDs in Australia.**

[https://ardc.edu.au/about\\_us/policies-and-guidelines/persistent-identifiers-policy/](https://ardc.edu.au/about_us/policies-and-guidelines/persistent-identifiers-policy/)

---

Which PIDs can you get from ARDC and how do you get them?

# ARDC Identifier Services



Data and associated workflows, software, models, and grey literature.



Physical samples and specimens.



General purpose identifier - datasets, collections, papers etc.



Research projects.

Each service:

- Covers different use cases in the research ecosystem and have their own governance models.
- Supports the creation and management of globally unique, persistent and resolvable identifiers.
- Is free to use by Australian researchers and research organisations/institutions.

# Web Interfaces

- Manual creation and management of identifiers.
- Primarily used to create smaller numbers of identifiers.
- Assist in creating rich metadata records.
- No technical skills/expertise required.
- Understanding of underlying metadata schema is not required.

The screenshot displays a web interface for creating identifiers, divided into two main sections: 'Primary Information' and 'Curation Details'.

**Primary Information:**

- Owner:** A dropdown menu with 'Private' selected.
- Resource Identifier:** Two input fields containing '20.500.11812/XXZT1' and 'BQXKIV'.
- Landing Page:** An input field containing 'https://test.identifiers.ardc.edu.au/igsn-portal/view/20.500.'.
- Registered Object Type:** A dropdown menu with 'Physical sample' selected.
- Metadata Visibility:** A dropdown menu with 'Publicly Visible' selected.
- Sample or Item Title:** An input field containing 'Zircons from fraser range amphibolite on epoxy SHRIMP'. Below it is a red error message: 'Provide a title for the sample or item'.

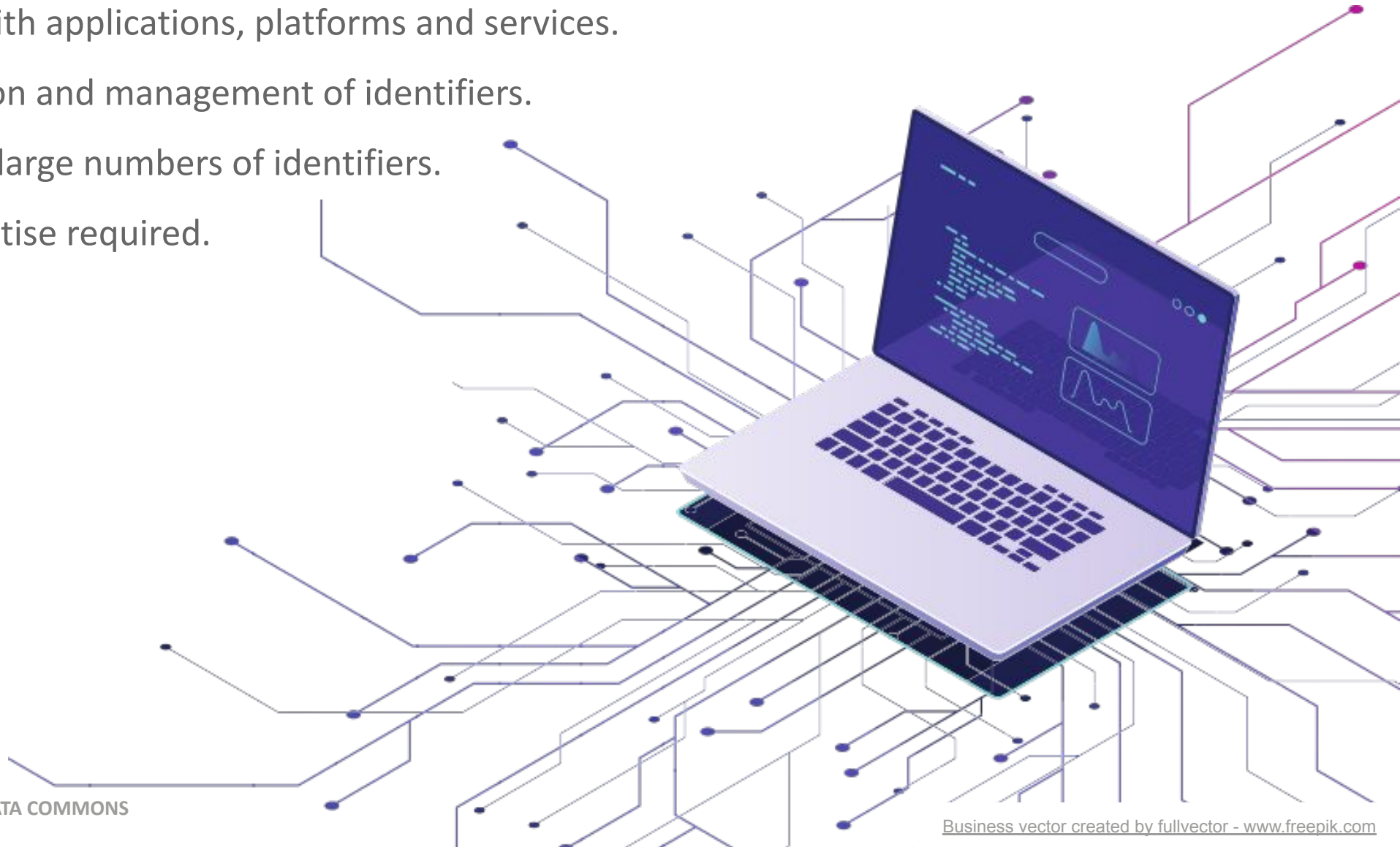
**Curation Details:**

- Curator:** An empty input field.
- ORCID Search:** A search button with a magnifying glass icon and the text 'ORCID Search'.
- Search Query:** An input field containing an empty search query. To its right is a close button '[X]'.
- Search Results:** A list of names: Ryan Sullivan, JULIA MARTIN, Tom Honeyman, Liz Stokes, Siobhann McCafferty, Rhys Williams, Matthias Liffers, Brian Ballsun-Stanton, Andrew Treloar, and Natasha Simons.
- Curator Input:** A red error message at the bottom: 'Enter the name of the resource curator'.

A yellow arrow points to the search results list in the 'Curation Details' section.

# APIs

- Enable integration with applications, platforms and services.
- Automate the creation and management of identifiers.
- Efficient for creating large numbers of identifiers.
- Technical skills/expertise required.



# Service exploration and testing

- Test credentials provided for each service.
- Web UIs and APIs are accessible.
- ARDC technical support and guidance is available.



[Box vector created by sentavio - www.freepik.com](http://www.freepik.com)

# Obtaining an ARDC Service Account

Applicant discusses their interest in using the service with the ARDC Services Team or an ARDC Engagements officer.

Applicant reads the relevant ARDC Service Policy Statement

Applicant completes and submits the relevant Service Participant agreement to [services@ardc.edu.au](mailto:services@ardc.edu.au)

The ARDC Services Team sets up a new account for the applicant and provides them with the details.



# Service Documentation

<https://documentation.ardc.edu.au>

Public Technical Documentation & Resources

Find documentation, answers, and more ...

Find information on ARDC Online Services and software release notes, and access guides for using our online services.

|   |   |   |   |  |
|---|---|---|---|--|
| <br><b>Registry Software</b>         | <br><b>Research Vocabularies</b> | <br><b>DOI Service</b>             | <br><b>IGSN Service</b>        | <br><b>Handle Service</b> |
| <br><b>Content Providers Guide</b> | <br><b>Widgets And APIs</b>    | <br><b>RIF-CS Advisory Board</b> | <br><b>Technical Updates</b> | <br><b>User Guides</b>  |

---

# PIDs for Projects: RAiD

# What is a RAiD?

RAiD is an Identifier that you will be given by ARDC when your project commences.

RAiD is a unique and persistent identifier for research projects.

It acts like an envelope for the research project activities and so it collects a list of identifiers for the people, data, publications, instruments and institutions that are involved.

RAiD is an identifier service run directly by ARDC. It is in the process of gaining ISO Certification.

## Project vs Grant

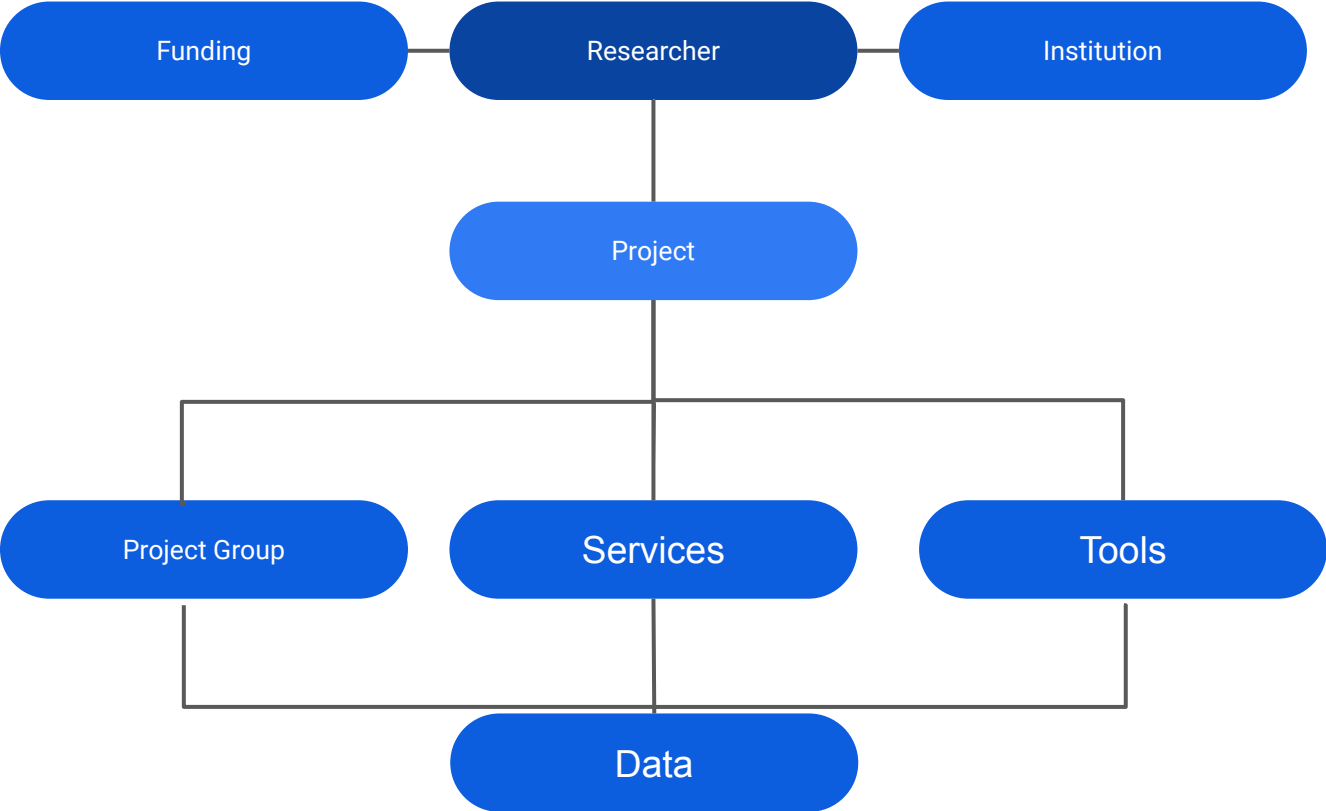
A research project is an activity.

It takes place over a period of time, has a set scope, is resourced by researchers and research support staff, and uses and produces data.

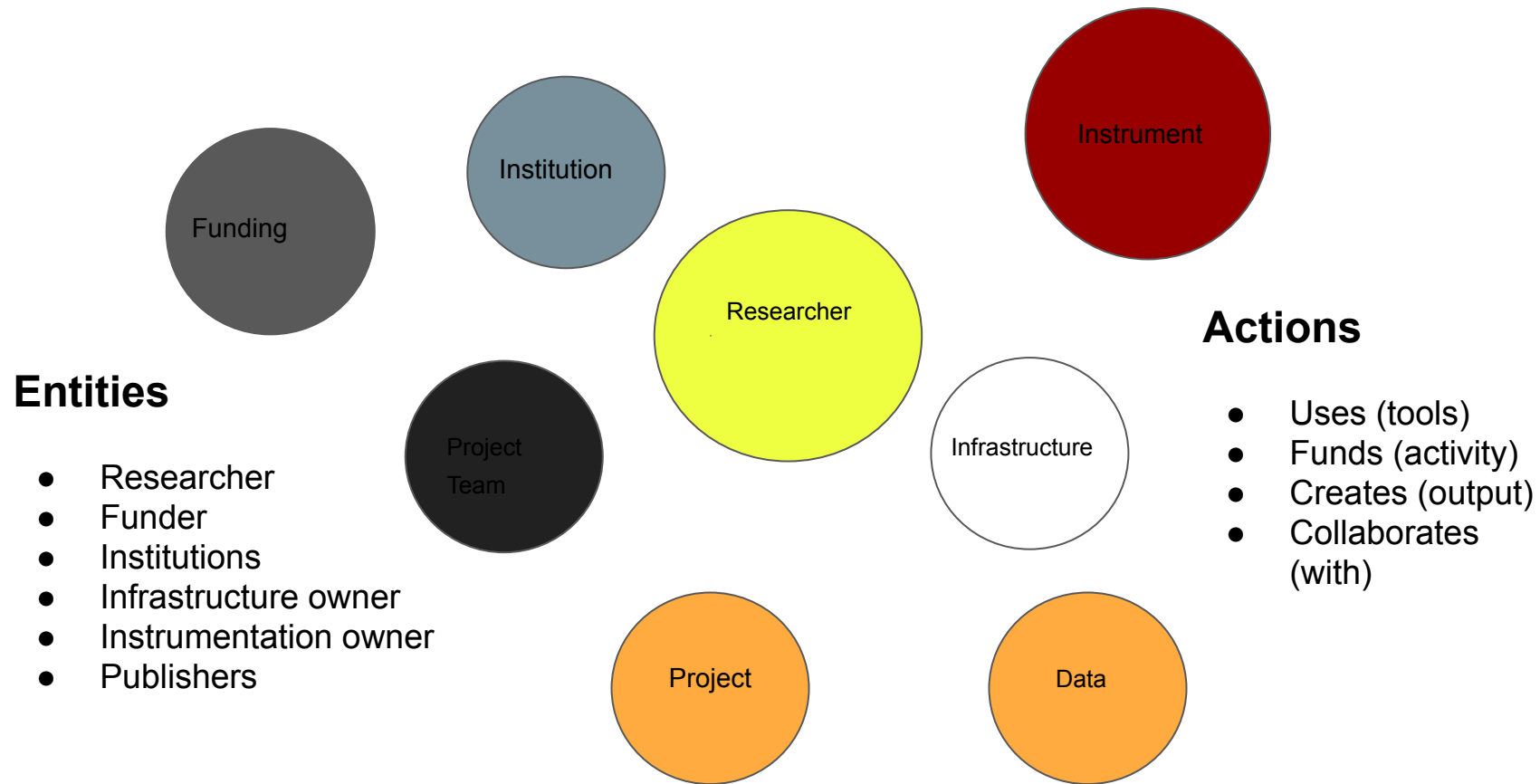
A research project is **not** a research grant.

A research project is the **activity** of doing research. A research grant is an **award** to do the research.

# Why a Project ID?

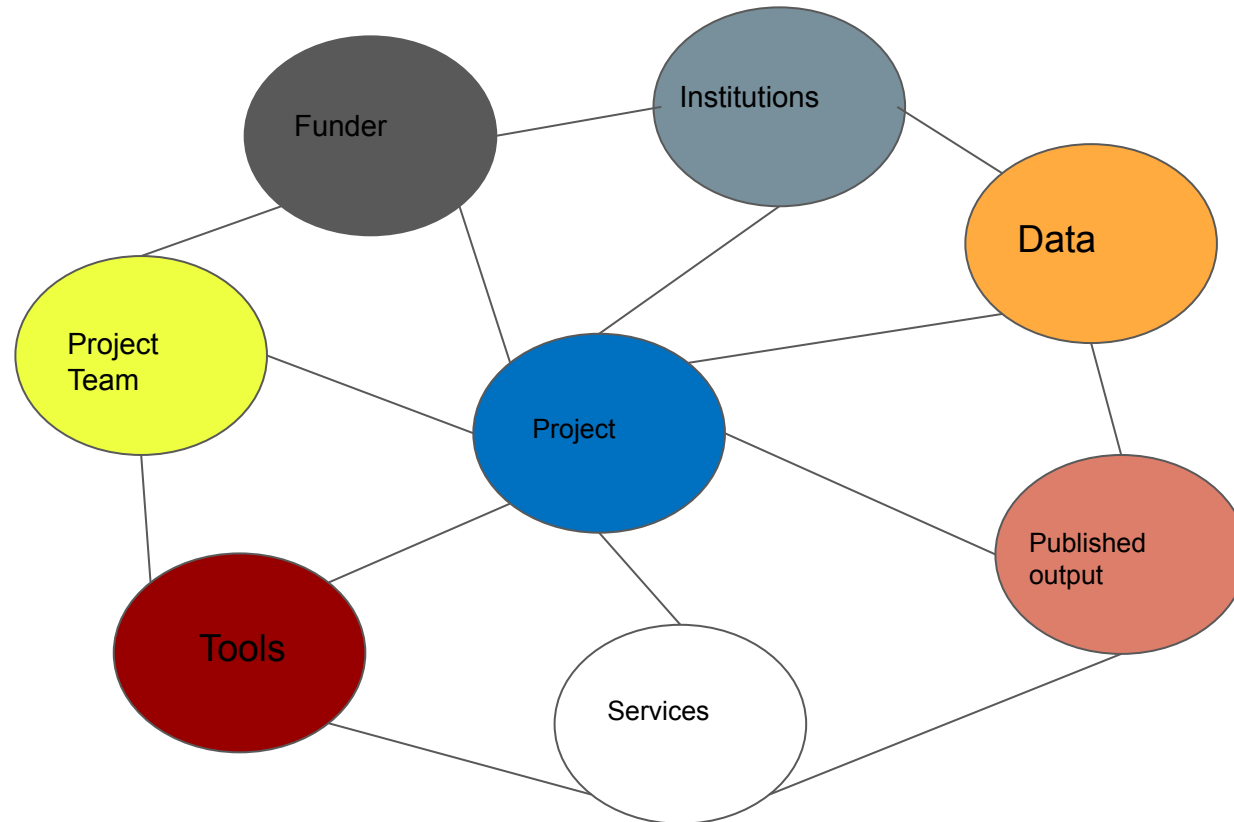


# Research Activity Components



# Project Centred Rhizome Model

- **Project has Persistent ID**
- **Entities recorded in Metadata**
- **Research actions reflected in project timeline**
- **Related PIDs recorded in metadata**



# How does RAiD work?

A RAiD has two parts: The RAiD identifier and the RAiD metadata envelope.

RAiD uses the Handle system to create its identifiers.

The RAiD Handle is a number that is like the address on an envelope. Inside the envelope are the contents of the envelope, made up of other identifiers that represent project activities and describe the relationship between them.

The metadata envelope records time/date stamped PIDs for:

- Grants and investments
  - The first identifier in your project's RAiD will be the ARDC investment identifier
- Organisations (institutions)
- Collaborators (people)
- Tools and Services (such as the Cloud)
- Data and publications

A RAiD records who and what a project interacts with during its active timeline and stores them in the metadata manifest.



# What are the benefits of using RAiD?

RAiD is used to establish data provenance, audit data access, collect metrics on facility and instrument use, grant access authorisation and reflect the roles of individuals in the research process.

## Benefits of RAiD

- Create a timeline of interactions
- Gain insights into investments and activities
- Develop better strategic intelligence on outcomes
- Generate better evidence and understanding of impact
- Save time on administration and reporting
- Provide better tools for analysis and decisions support



# Use Cases

## UQ RDM

- RAiD for storage allocation

## QLD NIF Node

- RAiD for instrument booking and data storage

The screenshot displays the Research Data Manager (RDM) interface. At the top left is the RDM logo, consisting of a grid of colored dots (R, D, M) and the text "Research Data Manager". Below the logo is the text "Designed at The University of Queensland". On the top right, the user's name "Ms Siobhann McCafferty" is displayed. Below the name are statistics: "2 records", "413 MB of storage", and "4 collaborators". At the bottom of the header area, a summary line reads: "Total UQRDM storage: 9,120TB. Total Files: 302,195,463. Total UQRDM records: 10,697. Unique active users: 12,251." Below the header are two main sections: "My records" and "My collaborators".

| My records                    |        |
|-------------------------------|--------|
| Project title                 | Status |
| test 1                        | Active |
| DLCF - Data Lifecycle Project | Active |

| My collaborators |  |
|------------------|--|
| Contact name     |  |
| Dr Andrew Janke  |  |

# RAiD Roadmap and plans

- Developing further functionality
  - Manual DOI entry of PIDs to DMR
- Improve RAiD GUI
  - Better user experience
  - Analytics
  - PID graph?
- Early Adopter program
  - Supported integration and manual use
- ISO Standard
  - Late 2021
- Supported ‘baked in’ integrations in selected platforms

---

# PIDs for Instruments: an emerging standard

# Identifiers for Instruments

## Issues:

- **No current cross-discipline standard PID or schema**
- **Lack of agreement on what an instrument is**
- **Granularity of description**
- **Community ownership of standards**

# RDA PIDs for Instruments Working Group

## Objectives

- Explore the use of a globally unique solution to persistently identify active measuring instruments
- Recommend a metadata profile to describe instruments that harmonises existing identification standards and complements existing metadata schemas
- Explore methodology/technology to register and resolve the new PID
- Operationalise the solution by engaging existing PID infrastructure providers, instrument developers and manufacturers, as well as instrument database providers

# RDA PIDInst. Working Group

## Outputs

- PIDInst Schema
  - (RDA Supporting Output)

## In progress

- Modified Datacite Schema
- White Paper Draft
- Adoption cases

<https://www.rd-alliance.org/groups/persistent-identification-instruments-wg>

# i4iOZ: A Community of Practice

## Objectives

- Support and develop best practice for instrument PIDs
- Share current identification practices and developments
- Connect activities in Australasia with international activities
- Raise awareness of technical requirements for instrument identifiers

<https://sites.google.com/ardc.edu.au/i4ioz>



# Current Practice example: National Imaging Facility

- Instrument described using agreed schema
- DOI or handle for instrument (via ARDC)
- Instrument record in Research Data Australia
  - [Bruker BioSpec MRI](#)
- Calibration and Instrument specs listed in data set record in [Research Data Australia](#)
- DOI cited in reporting and publications



The screenshot shows a record for the Bruker BioSpec 94/30 US/R MRI instrument. At the top right, there is a 'Service or Tool' icon. The main title is 'Bruker BioSpec 94/30 US/R MRI', with the subtitle 'Also known as: MRIB94T' and 'The University of Queensland' below it. To the left of the title is the University of Queensland logo. On the right side, there are social media icons for Facebook and Twitter, and a 'Viewed: 1144' counter. Below the title, there is a blue button labeled 'Access the service' with a downward arrow, and a grey button labeled 'Save to MyRDA'. Underneath these buttons, there is a 'Licence & Rights' section with a yellow 'CONDITIONAL' label and a 'View details' link. To the right of these buttons is a 'Full description' section containing the text: 'The Bruker Biospec MRI scanner consists of a 9.4 Tesla, 300 mm ultrashield refrigerated 30 cm magnet interfaced to a scalable Bruker Avance III spectrometer that allows unparalleled flexibility for MRI scanning of small and medium sized research animals and other samples.'



# Questions?



**Australian Research Data Commons**

## **CONTACT**

Level 2, AIBN (Building 75)  
Brisbane QLD 4072  
AUSTRALIA

+61 7 0405 340986  
siobhann.mccafferty@ardc.edu.au  
ardc.edu.au

[i4iOZ](#)  
[i4iOZ Google Group](#)