



## FAIR Data Guidelines for Project Outputs

|                          |  |
|--------------------------|--|
| <b>Date effective</b>    | 19 June 2020   |
| <b>Review date</b>       | 19 June 2023   |
| <b>Type</b>              | Guideline  |
| <b>Owner</b>             | Manager, engagements, ARDC   |
| <b>Approved by</b>       | EMT  |
| <b>Category</b>          | Projects   |
| <b>Version Number</b>    | [1.0]  |
| <b>Audience</b>          | External   |
| <b>Content enquiries</b> | <a href="mailto:contact@ardc.edu.au">contact@ardc.edu.au</a>   |
| <b>Related documents</b> | <a href="#">Persistent Identifiers for ARDC Investments</a>  |
| <b>Purpose</b>           | Provide guidelines to partners for ARDC Investments on how data outputs involved in the partnership will be made more FAIR in the course of the project. |

## Context

This document provides guidelines and resources for ARDC projects on how data outputs involved in the project will be made more FAIR. The ARDC is supportive and encourages the adoption of the FAIR principles as a valuable way in making research outputs more reusable, both for humans and machines. This also aligns with the requirements for research institutions as the FAIR principles are referenced in the [Guide for Management of Data and Information in Research](#) that accompanies the Code for the Responsible Conduct of Research.

ARDC is an NCRIS facility and it is the [policy](#) of the Commonwealth NCRIS program that “data generated, created, captured or stored by NCRIS funded projects will be made available to the wider research community based on the F.A.I.R. principles.”

This document describes our expectation of how FAIR will be addressed in ARDC projects.

The [FAIR data principles](#) present a number of key areas in which data can be made FAIR, without being prescriptive on how this should be achieved. The Research Data Alliance (with global input) has published a [model on](#) how to assess FAIR. The current document translates these into a practical set of actions for ARDC projects, that will make the data more FAIR. We consider these realistic aims based on the state of maturity of services and standards currently available. The links provide guidance on how the different aspects can be achieved. Please consult with your ARDC liaison for the project as they will be able to provide guidance, support and organise training on all these aspects.

If you believe there is a good reason why a certain aspect cannot be addressed, please discuss this with your project liaison as early on as possible. The reason for it not being addressed will need to be documented and provided in the progress and final report.

## Findable

|   | Action  | Required/ Recommended | Links  |
|---|---|-----------------------|--|
| 1 | Assign all data outputs appropriate PIDs (DOIs are the default identifier for this purpose)   | Required              | <a href="#">PIDs guidelines</a>  |
| 2 | All data outputs have metadata to enable discovery. The discovery metadata schema and fields to be completed for all data outputs should be discussed by partners and ARDC at project initiation. | Required              | See <a href="#">Appendix 1</a> for required fields<br><br><a href="#">List of mandatory and optional collection elements for RDA</a><br><br>Guide to <a href="#">What makes a good collection record</a> |

|   |   |             |   |
|---|---|-------------|---|
| 3 | All data outputs generated by the project must have a record in Research Data Australia                             | Required    | <a href="#">Providing metadata records to Research Data Australia</a><br><a href="#">Defining a data collection</a> |
| 4 | Records for all data outputs are registered with relevant discipline-specific discovery aggregators (if one exists) | Recommended | For a list of potential discovery mechanisms see <a href="#">Re3Data</a>  |
| 5 | The persistent identifier for the data being described must be included in the metadata                             | Required    | For example, providing the dataset DOI in the <a href="#">&lt;identifier&gt;</a> element in RIF-CS                  |

## Accessible

|    |  |                                   |   |
|----|--|-----------------------------------|---|
| 6  | All data outputs are made as openly available as possible; they are only closed where necessary  | Required                          | For guidance on what Open means see e.g. <a href="#">Open Data Handbook</a> |
| 7  | All data outputs are made available through a repository   | Required                          | See guidance on <a href="#">How to select a repository</a>                  |
| 8  | All data outputs are available as a download and/or accessible through an open, documented API   | Required where data is not closed | See <a href="#">guidance on APIs</a>  |
| 9  | If the data outputs are not openly available there is a clear description on the landing page on how to request access to the data outputs and conditions that need to be met              | Required where relevant           | See <a href="#">Guidance on Access Rights</a>                               |
| 10 | The persistent identifier for the data output will point to a landing page about the data output, even if the data output is not public.   | Required                          | See <a href="#">Best Practice for landing pages</a>                         |
| 11 | If the data output is not openly available there is an authorisation and authentication procedure to provide access to the data.   | Required where relevant           |   |
| 12 | The persistent identifier for the data output will continue to point to a landing page, even if the data output is no longer available. There is a policy to maintain these landing pages. | Required                          | See <a href="#">Best Practices for Tombstone pages</a>                      |

## Interoperable

|    |  |   |   |
|----|--|---|---|
| 13 | Data outputs use community-agreed standard data formats.   | Required where research community agreed standard formats exist   | For standards in specific disciplines see <a href="https://fairsharing.org">FAIRsharing.org</a>   |
| 14 | Metadata uses community-agreed standards   | Required where research community agreed metadata standards exist | <a href="#">ARDC Metadata guide</a>   |
| 15 | Data and metadata use community-agreed vocabularies, data models and ontologies (preferably internationally agreed ones where they exist)  | Recommended   | <a href="#">Good practice in vocabulary creation</a><br><a href="#">Research Vocabularies Australia</a><br><br><a href="#">Vocabularies and research data</a>     |
| 16 | Metadata contains persistent identifiers for research objects and entities (people, organisations) linked to the dataset(s). This includes links to ORCID, grantIDs, RAIDs, DOIs, IGSNs of related materials | Required  | Which PIDs to use: <a href="#">ARDC Persistent Identifiers (PIDs): Guidelines for ARDC Investments</a><br><br><a href="#">ARDC Citation and Identifiers guide</a> |

## Reusable

|    |   |             |  |
|----|---|-------------|--|
| 17 | All data outputs are assigned a machine readable (default choice should be a CC-BY 4.0) licence.  | Required    | <a href="#">ARDC Research Data Rights Management Guide</a>                                     |
| 18 | The licence information is available in a machine readable form on the landing page that the persistent identifier refers to.                         | Required    | <a href="#">Guidance on Machine readable reference to CC licence</a>                           |
| 19 | The landing page contains a citation statement for the data output  | Required    | <a href="#">ARDC Data Citation guide</a><br><br><a href="#">DataCite DOI Display Guideline</a> |
| 20 | Provenance information on the data output is attached alongside the data  | Recommended | <a href="#">ARDC guide on data provenance</a>  |
| 21 | Relevant discipline-specific metadata to enable reuse is captured and presented alongside the data output following research community best practice. | Recommended | See links in <a href="#">ARDC Metadata guide</a>   |

## Appendix 1 - Minimum metadata requirements: Collections

| Information type<br>(Links to RIF-CS guide) | Required/Recommended       | Meaning   |
|---|----------------------------|---|
| <a href="#">Identifier</a>                  | Required                   | A unique identifier for the resource, preferably a DOI.   |
| <a href="#">Collection Type</a>             | Required                   | The type of collection being described, i.e. collection, dataset, software, etc.  |
| <a href="#">Title</a>                       | Required                   | The name or title of the collection, should be descriptive and unique, avoid acronyms.  |
| <a href="#">Creator</a>                     | Required                   | The creator (person or organisation) of the collection (include ORCID if possible).   |
| <a href="#">Location</a>                    | Required                   | Online location of the resource   |
| <a href="#">Access Rights</a>               | Required                   | Collection access conditions. Specify one of: <b>open</b> , <b>conditional</b> or <b>restricted</b> .   |
| <a href="#">Licence</a>                     | Required                   | License conditions associated with the collection; a standard, machine understandable licence should be the default choice, e.g. Creative Commons BY-4.0. |
| <a href="#">Description</a>                 | Required                   | A summary description of the collection. Provide sufficient information to enable a user to assess suitability of the data for reuse for their purpose.   |
| <a href="#">Subject</a>                     | Recommended                | Keywords or terms to describe the topic of the resource. Include at least one <a href="#">ANZSRC-FOR</a> code.  |
| <a href="#">Spatial coverage</a>            | Recommended where relevant | The geometry for the spatial location the resource relates to.  |
| <a href="#">Temporal coverage</a>           | Recommended where relevant | The time period the resource relates to, in <a href="#">W3C Date/Time Format</a> .  |
| <a href="#">Related information</a>         | Recommended                | Links to related resources such as publications (via DOIs), websites (via URLs), funding info (via grant IDs), provenance information (via URLs), etc     |