

FAIR principles

and how to?

Ngala kwop biddi.
Building a brighter
future, together.

GRDC Data Management Workshop - 26th September 2023

Presenter: Dr Viet Dang



Short summary

Q: What is FAIR?

A: FAIR stand for **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable

Q: If I following FAIR principle, will my data be accessible to anyone?

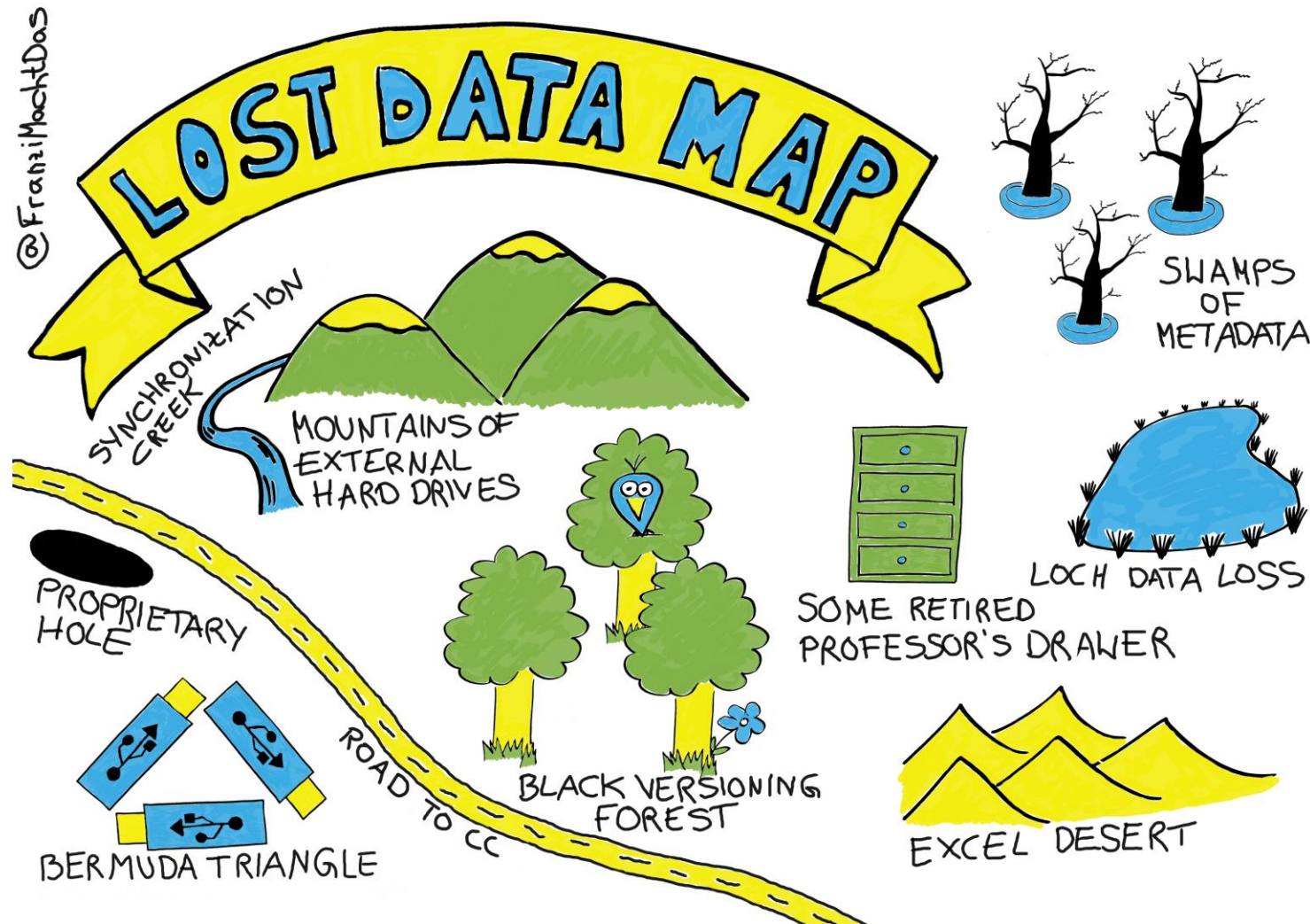
A: **NO**

Long Answer:

- The data can only be accessible as described in the access right
- Even with access, data can only be reuse with specific license

Why FAIR?

The Lost Data Map of research data management challenges and pitfalls



Maximum
potential of
data

Boost visibility
and citations

Attracting
partnerships

Aligned with
international
standards

Enabling new
research

Improve reproducibility and
reliability

F

Findable

The important of Metadata

To be Findable

F1. (meta)data are assigned a globally unique and persistent identifier

F2. data are described with rich metadata

F3. metadata clearly and explicitly include the identifier of the data it describes

F4. (meta)data are registered or indexed in a searchable resource

What is Metadata?



Pam Brophy - <https://www.geograph.org.uk>

Trial data

- Grain yield
- Grain size
- Plant height
- ...

Trial information

- Variety information
- Sowing time
- Harvest time
- Plant density
- Management practices
- Irrigation/Fertiliser
- ...

Metadata

- Identifier (i.e. DOI)
- Dates
- Title of dataset
- Version
- Creators/ Contributors
- Location
- Keywords
- Methodology
- Data processing
- Technical detail
- Access
- Rights

Metadata and data shall be findable by both humans and machines

The important of Unique identifier

Objects

Researchers/Contributors

Organisations

About this article



Cite this article

[Download citation](#)

Received	Accepted	Published
23 March 2023	20 June 2023	28 June 2023

DOI
<https://doi.org/10.1038/s41586-0>

Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Provided by the Springer Nature Sharedit content-sharing initiative

Subjects

[Biochemistry](#) [Structural biology](#)



ORCID
Connecting research and researchers

<https://ror.org/00r4sry34>

Murdoch University

ORGANIZATION TYPE

Education

WEBSITE

<http://www.murdoch.edu.au/>

<https://ror.org/02xwr1996>

Grains Research and Development Corporation

ORGANIZATION TYPE

Government

LOCATION

Canberra (GeoNames ID 2172517)
Australia

Rich metadata

https://sandbox.zenodo.org/

April 1, 2023

Barley phenotype data

John Doe

Phenotype data collected in 2023

Dataset

Closed Access

Edit

0 views

0 downloads

See more details...

Indexed in

OpenAIRE

Publication date:

April 1, 2023

DOI: [10.1234/test.dataset](https://doi.org/10.1234/test.dataset)

Keyword(s): Barley

Share

Cite as

John Doe. (2023). Barley phenotype data (Version 1) [Data set]. <https://doi.org/10.1234/test.dataset>

Start typing a citation style...

Data

Files (5.2 kB)

Closed Access

Files are not publicly accessible.

Name	Size	
barleydata.xlsx	6.2 kB	Download
md5:0b2eabf1debdaac409d9e163f30b702		

Citations

Show only: ☐ Literature (0) ☐ Dataset (0) ☐ Software (0) ☐ Unknown (0)

☐ Citations to this version

No citations.

VS

DATASET

Determining the genetic control of grain size and heat stress tolerance during flowering in barley (Phenotype data)

Calum J Watt, David Moody, Yong Han, Chengdao Li, Xiao-Qi Zhang, Camilla B Hill, Penghao Wang and Western Crop Genetics Alliance [Show details for 8 authors](#)

Murdoch University

01/03/2017 - 30/10/2020

Share

Export

Files and links

ZIP

UMU1903-003RSX Phenotype | 2.01 MB

Phenotype data collected from 2017 to 2019 | [CC BY V4.0](#), Embargoed Access, Embargo ends: 31/12/2023

Abstract

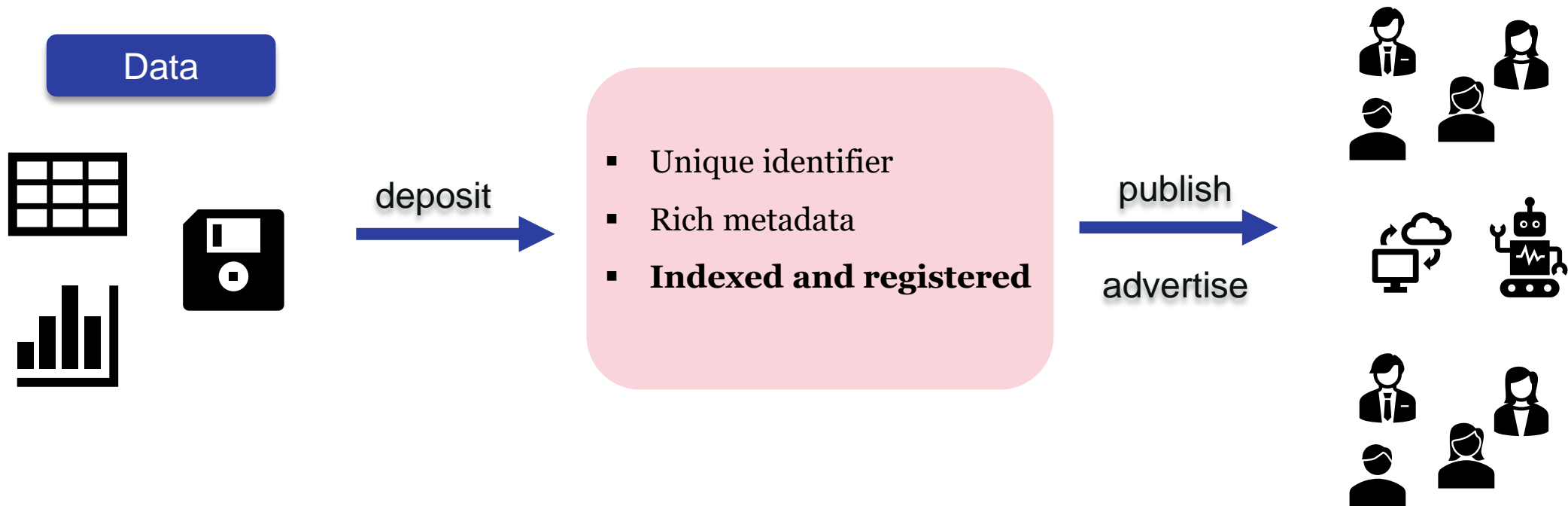
Cereal (Grains) | Barley | Abiotic stress | Grain shape | Heat stress | QTL mapping | Crop and pasture improvement (incl. selection and breeding) | Non-genetically modified uses of biotechnology

This collection comprises phenotype data collected from field trials in Wongan Hills, Merredin, Williams, and Perth from 2017 to 2019. For each trial, the phenotypic data including yield, grain width, grain thickness, grain length, thousand-grain weight, and flowering date were collected. There is also multi-environment trial analysis data which is spatially adjusted to determine genotype, year, and location effects. These results were outputs of the GRDC-funded project GRS (Calum Watt) - Determining the genetic control of grain size and heat stress tolerance during flowering in barley (UMU1903-003RSX).

Collapse

Registered or indexed metadata

Identifiers and rich metadata doesn't guarantee the findability of data



A

Accessible

To be Accessible

A1. (meta)data are retrievable by their identifier using a standardized communications protocol

A1.1 the protocol is open, free, and universally implementable

A1.2 the protocol allows for an authentication and authorization procedure, where necessary

A2. metadata are accessible, even when the data are no longer available

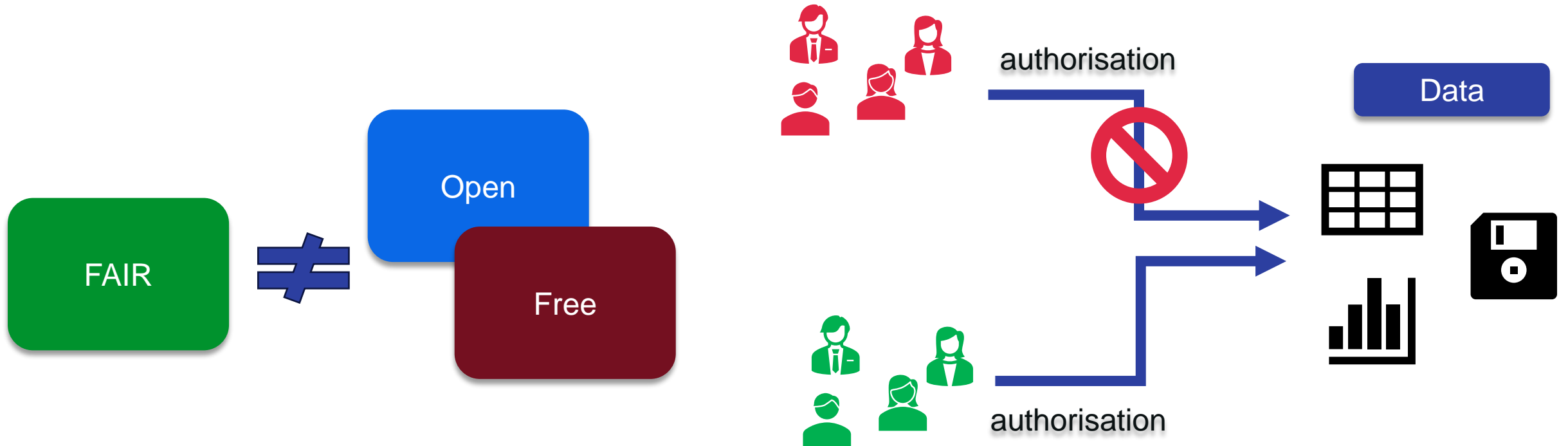
Data/Metadata can be retrieved by identifier

Data/metadata should be findable, accessible, interoperable and reusable by both **human** and **machine**

Metadata: how data can be accessed, authentication and authorisation

In case the deposited data is lost, the data owners, institutions and related publications still can be retrieved by metadata

Accessible is not Open



Metadata include: The conditions under which the data are accessible

Even protected data can be FAIR



Interoperable

To be Interoperable:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (meta)data use vocabularies that follow FAIR principles

I3. (meta)data include qualified references to other (meta)data


Metadata standard

General	Science
<ul style="list-style-type: none">• <u>Dublin Core</u>• <u>Metadata Encoding and Transmission Standard (METS)</u>• <u>Metadata Object Description Schema (MODS)</u>	<ul style="list-style-type: none">• <u>Astronomy Visualization Metadata (AVM)</u>• <u>CSMD-CCLRC Core Scientific Metadata Model</u>• <u>Darwin Core</u>• <u>Ecological Metadata Language (EML)</u>

<https://libguides.murdoch.edu.au/RDM/documentation>

Data standard / Vocabularies

<https://fairsharing.org>

 search through all content

Q SEARCH

LOGIN

STANDARDS

DATABASES

POLICIES

COLLECTIONS

ORGANISATIONS


ADD CONTENT

STATS

GENERAL INFORMATION



ACTIONS

This record replaces or incorporates the following deprecated resources:
[Cereal Plant Development](#)



R

Plant Ontology (PO)

 [10.25504/FAIRsharing.3ngg40](https://doi.org/10.25504/FAIRsharing.3ngg40) 


Type Terminology artefact

Registry Standard

Description The Plant Ontology is a structured vocabulary and database resource that links plant anatomy, morphology and growth and development to plant genomics data. The PO is under active development to expand to encompass terms and annotations from all plants.

Homepage <http://plantontology.org/>

Year of Creation 2002

Maintainers [cooperl09](#) 


Countries developing this resource [United States](#)







Subjects [Botany](#) [Life Science](#) [Plant Anatomy](#)

Domains [Expression Data](#) [Plant Development Stage](#) [Life Cycle](#)

Taxonomic Range [Viridiplantae](#)

User Defined Tags None

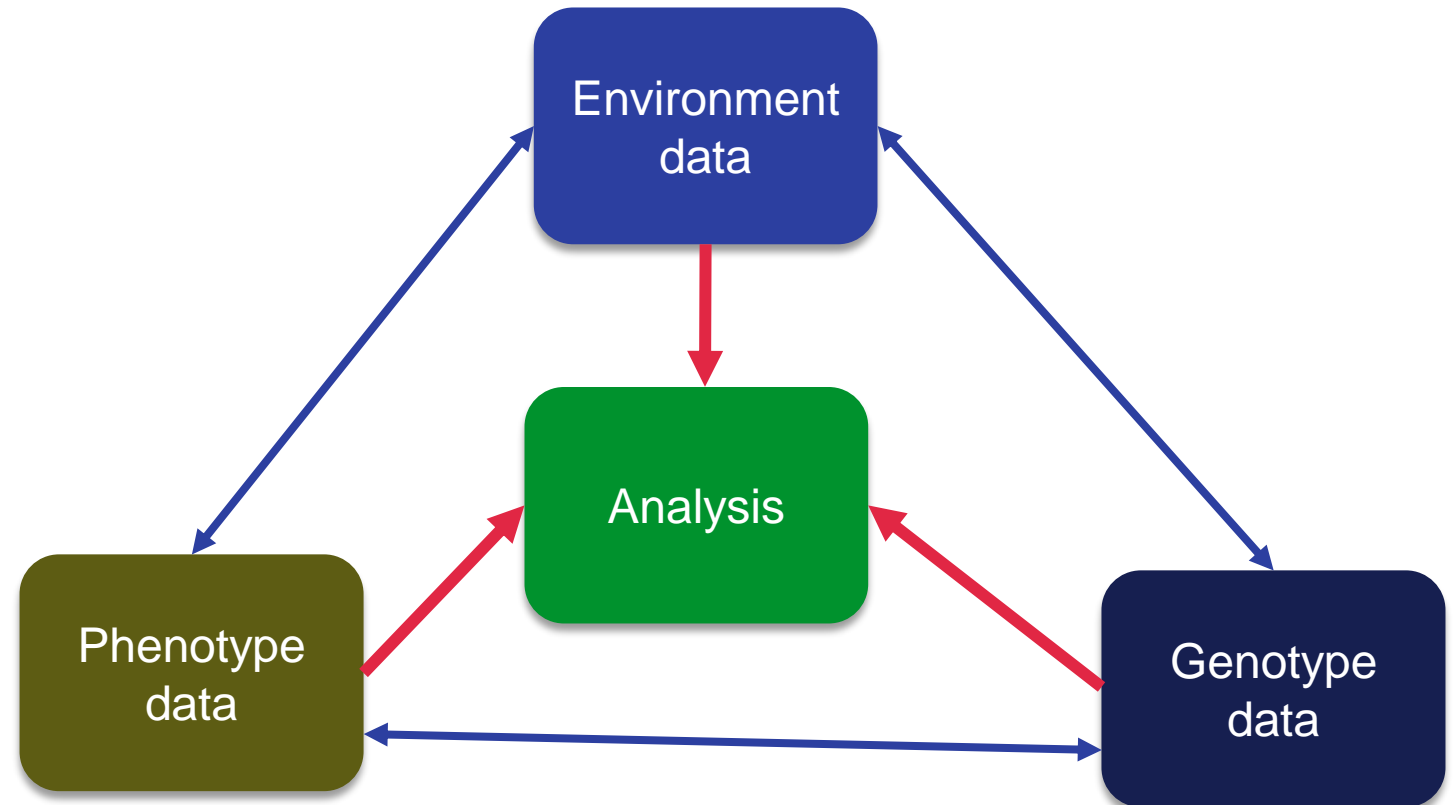
 VIEW RELATION GRAPH



Reference to other data/metadata

Specify if:

- The dataset is builds on another dataset(s)
- Additional dataset(s) is need to complete the data
- Complementary information in another dataset



R

Reusable

To be Reusable

R1. meta(data) are richly described with a plurality of accurate and relevant attributes

R1.1. (meta)data are released with a clear and accessible data usage license

R1.2. (meta)data are associated with detailed provenance

R1.3. (meta)data meet domain-relevant community standards

Data usage license

IDENTIFIERS 991005566667707891

COPYRIGHT Attribution 4.0 International (CC BY 4.0)

MURDOCH AFFILIATION Centre for Crop and Food Innovation

LANGUAGE English

RESOURCE TYPE Dataset

LOCATIONS Latitude: -31.653 Longitude: 116.666



**Attribution-NonCommercial 4.0
International (CC BY-NC 4.0)**



Attribution 4.0 International (CC BY 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#), [Disclaimer](#).

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.



Under the following terms:



Attribution — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.


Associated with detailed provenance

What is the origin of the data (who created, affiliation, where, ...)

Who to cite when the data is reused?

Domain-relevant community standards

MIABi



Minimum Information About a Bioinformatics investigation

The Minimum Information about a Bioinformatics investigation (MIABi) initiative specifies, through a series of documentation modules, the minimum reproducibility amongst the bioinformatics and computational biology community.

[Bioinforma...](#) [Not applic...](#)

» Related Standards

» Implementing Databases


» Endorsing Policies

1

1

0

MIENS



Minimum Information about an ENvironmental Sequence

An early checklist from the Genomic Standards Consortium that has been replaced by the Minimum Information about any (x) Sequence (MInxS) family data for marker genes retrieved from the environment.

[Environme...](#) [Environme...](#) [Environme...](#)

» Related Standards

» Implementing Databases


» Endorsing Policies

2

0

0

MITAP



Minimum Information Model for toIAPC

The minimum information model for toIAPC (MITAP) guideline is an initiative of members of the toIAPC field to provide a reporting framework that will of autoimmune diseases and for the prevention of destructive immune responses after transplantation. The methodologies for generating toIAPC vi

[Immunology](#) [Biomedical...](#) [Translation...](#) [Homo sapi...](#) [one more tag](#)

» Related Standards

» Implementing Databases


» Endorsing Policies

0

0

0

MIARE



Minimum Information About a RNAi Experiment

Minimum Information About an RNAi Experiment (MIARE) is a set of reporting guidelines that describes the minimum information that should be rep include a data model, data exchange format, controlled vocabulary and supporting software tools.

[Life Science](#) [Annotation](#) [Gene Expr...](#) [Regulation...](#) [All](#) [+8 more tags](#)

» Related Standards

» Implementing Databases

» Endorsing Policies

1

1

0

Use appropriate:

- File format
- Common template/
documentations
- Contain sufficient information

S

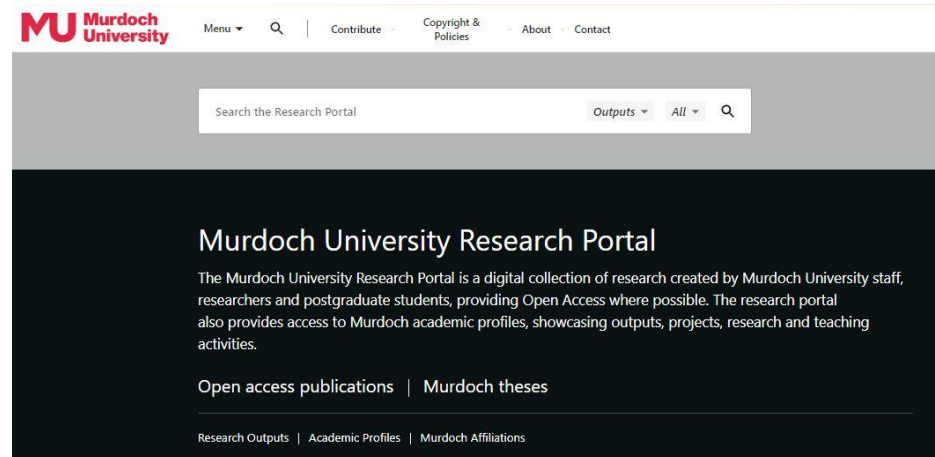
Summary

Data

- Collect as much information as possible
- Using standard vocabularies

Metadata

- Collect not only required metadata but as much as possible
- Clearly specify **Access rights** and **License**



Publish

Advertise

Collaboration

Assign Unique Identifier

Indexed and Searchable

Thank you

Ngala kwop biddi.
Building a brighter
future, together.



Useful Links

- <https://ror.org>
- <https://orcid.org>
- <https://fairsharing.org>
- <https://creativecommons.org/licenses>

Resources and references

- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>
- Australian Research Data Commons - <https://ardc.edu.au>
- Australian Research Data Commons - FAIR training resources
- University of Mannheim FAIR-Data-Week resources - <https://github.com/UB-Mannheim/FAIR-Data-Week>