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 **Findable**, **Accessible**, **Interoperable**, and **Reusable**

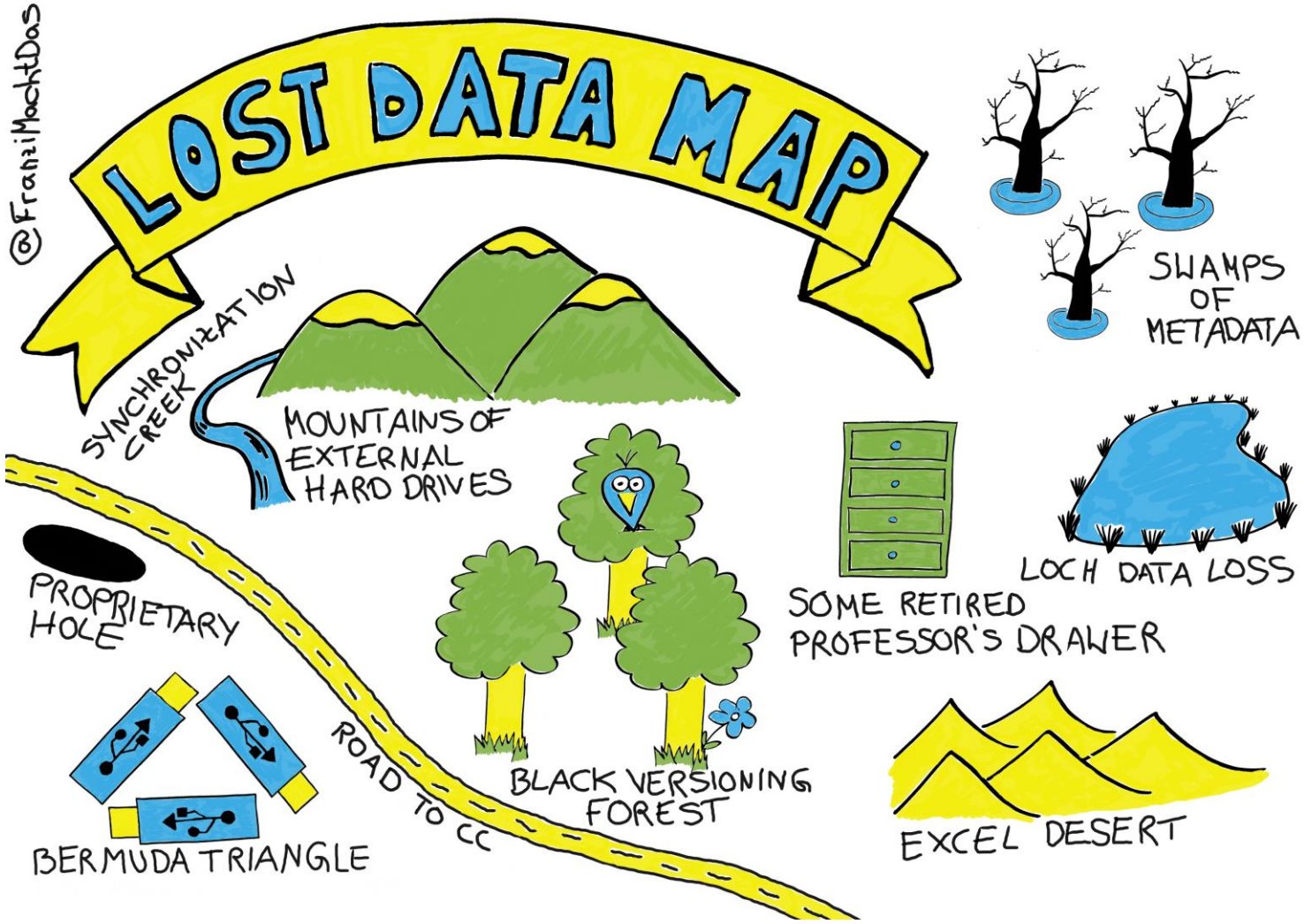
Q: If I following FAIR principle, will my data be accessible to anyone?

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**

Mau, Franziska. (2019). Sketchnote: Lost Data Map. Zenodo. https://doi.org/10.5281/zenodo.4388672

F

Findable

The important of Metadata

To be Findable

Murdoch University

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

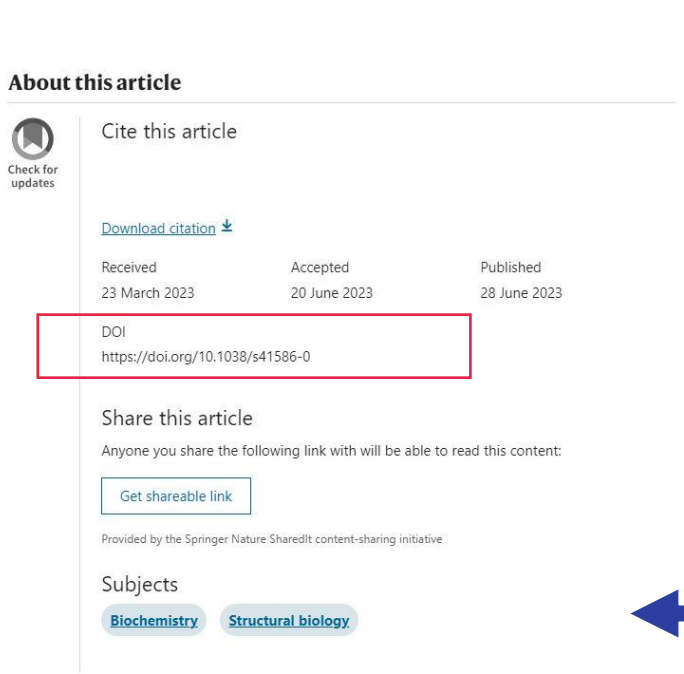
The important of Unique identifier

Murdoch University

Objects

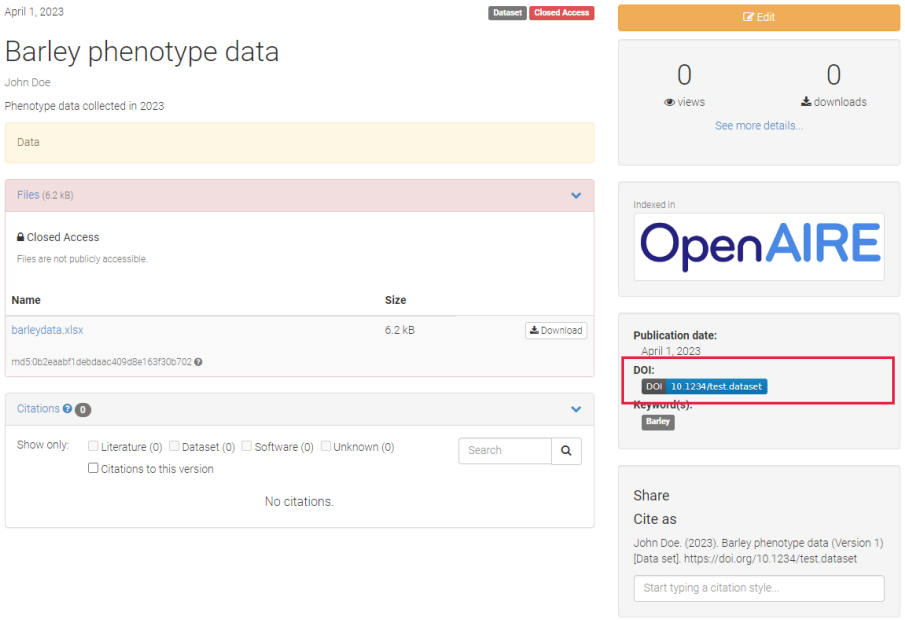
Researchers/Contributors

Organisations

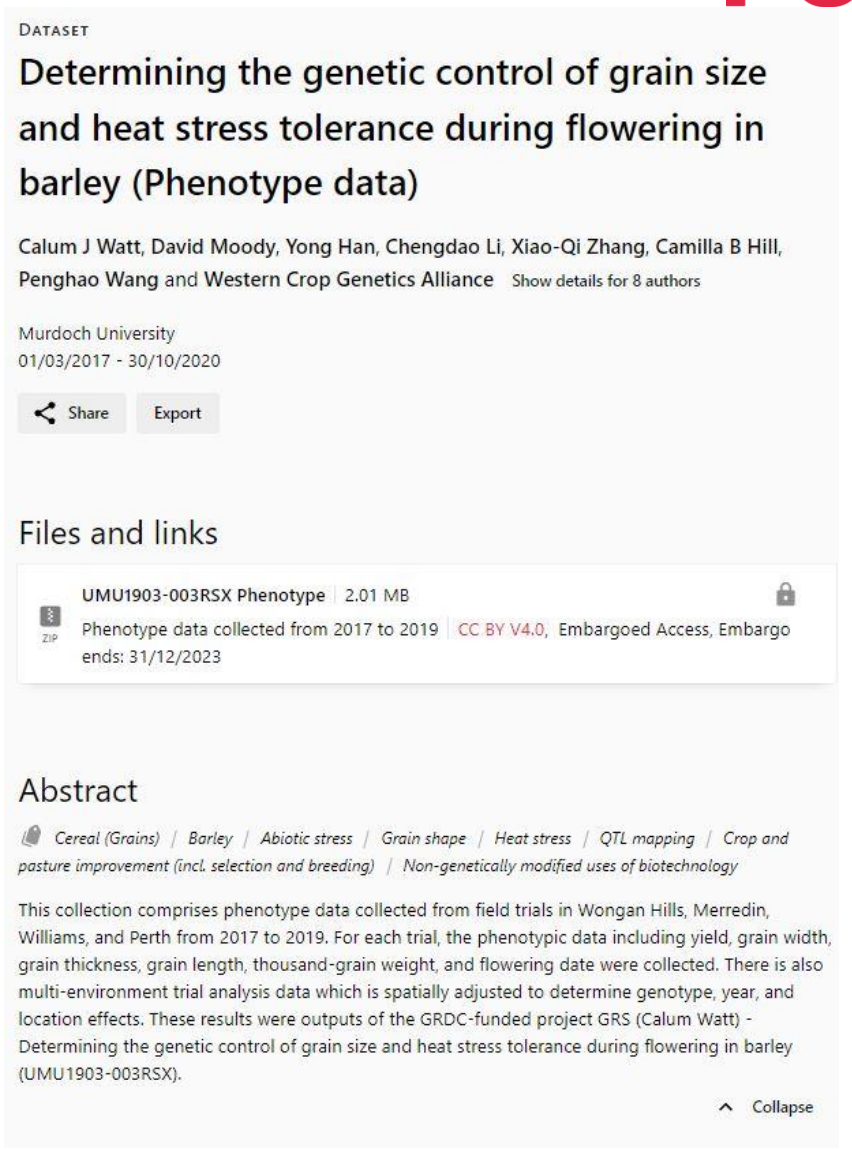


Rich metadata

https://sandbox.zenodo.org/



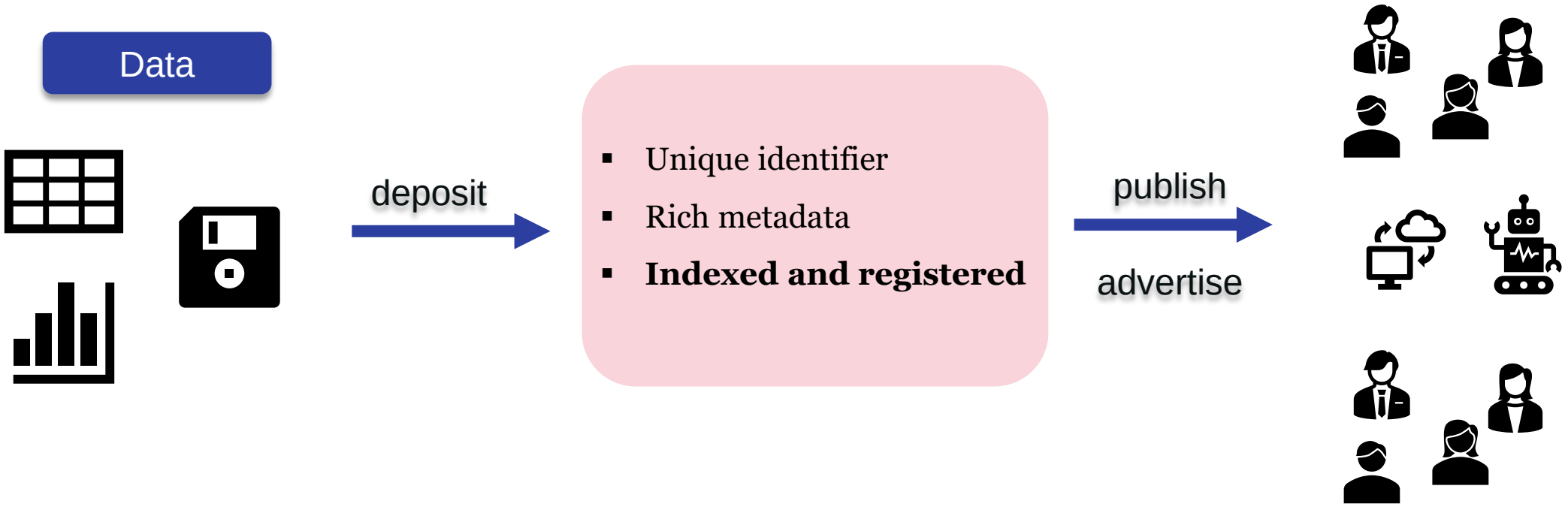
VS



Registered or indexed metadata

Murdoch University

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



A

Accessible

To be Accessible

Murdoch University

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

Murdoch University

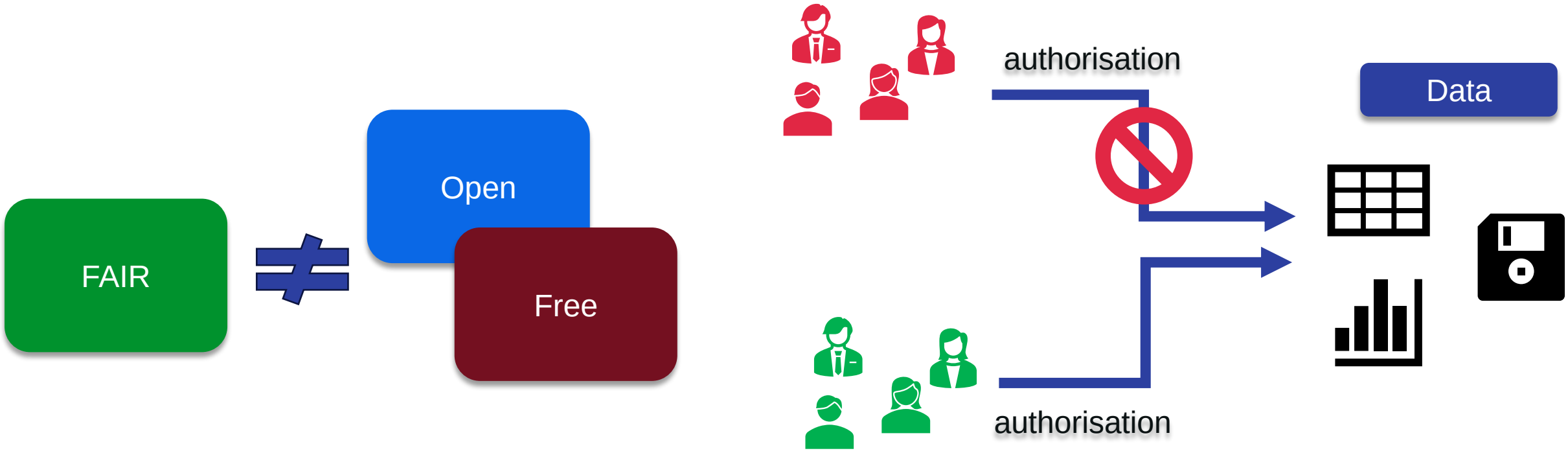
Data/metadata should be finable, accessible, interoperable and reusable by both **human** and **machine**

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

Murdoch University



Metadata include: The conditions under which the data are accessible

Even protected data can be FAIR

Interoperable

To be Interoperable:

Murdoch University

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

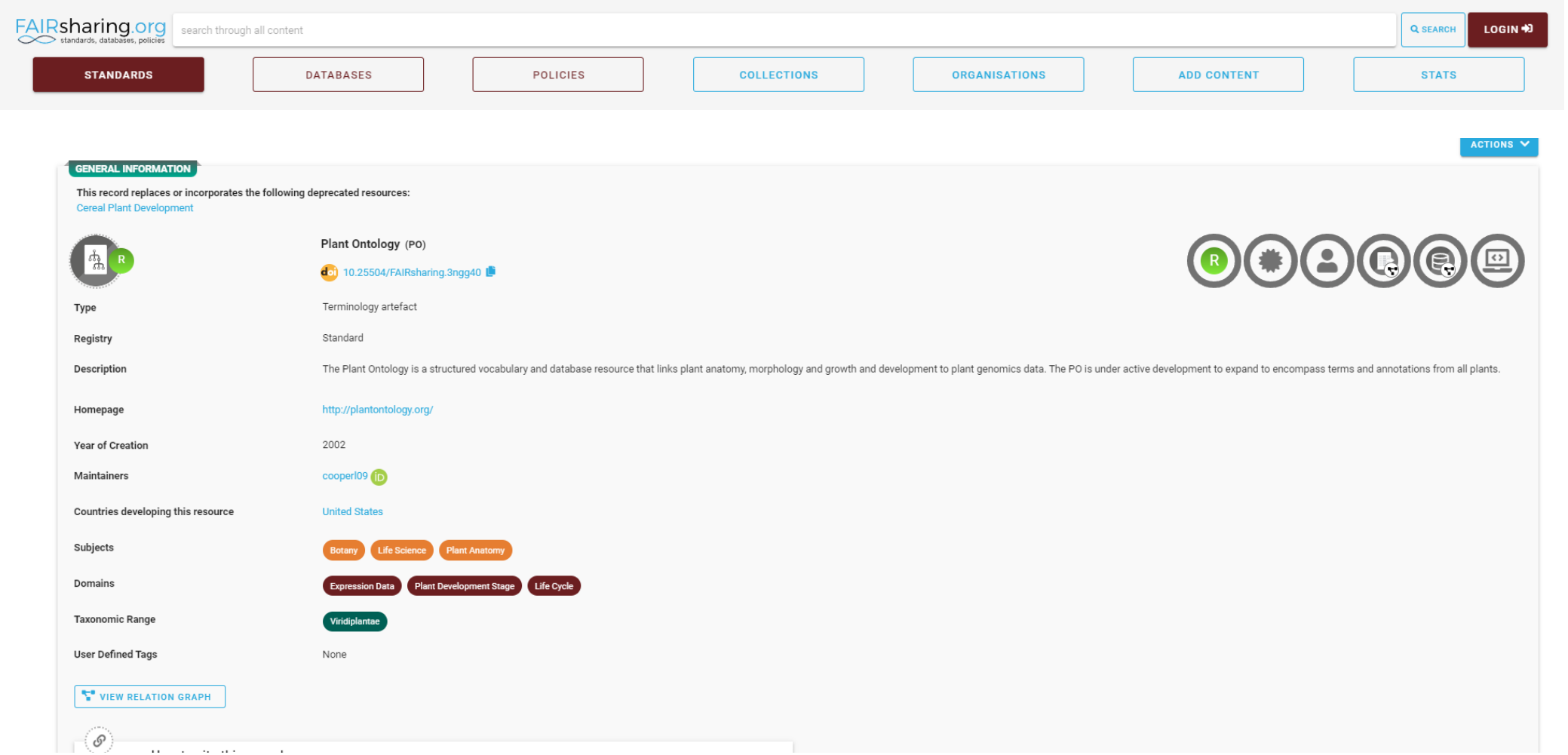
General	Science
• Dublin Core • Metadata Encoding and Transmission Standard (METS) • Metadata Object Description Schema (MODS)	• Astronomy Visualization Metadata (AVM) • CSMD-CCLRC Core Scientific Metadata Model • Darwin Core • Ecological Metadata Language (EML)

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

Data standard / Vocabularies

Murdoch University

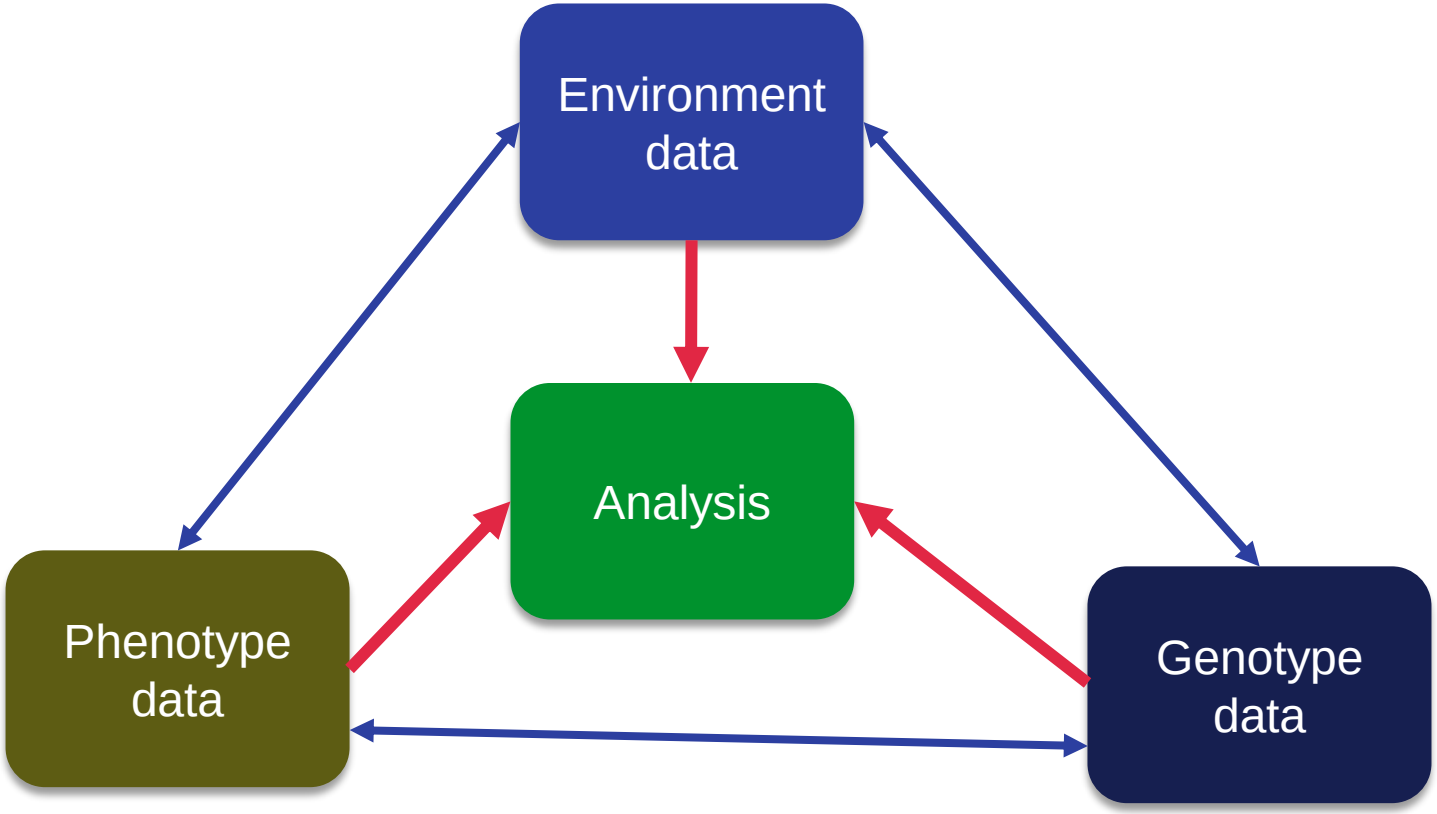
https://fairsharing.org



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

Murdoch University

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

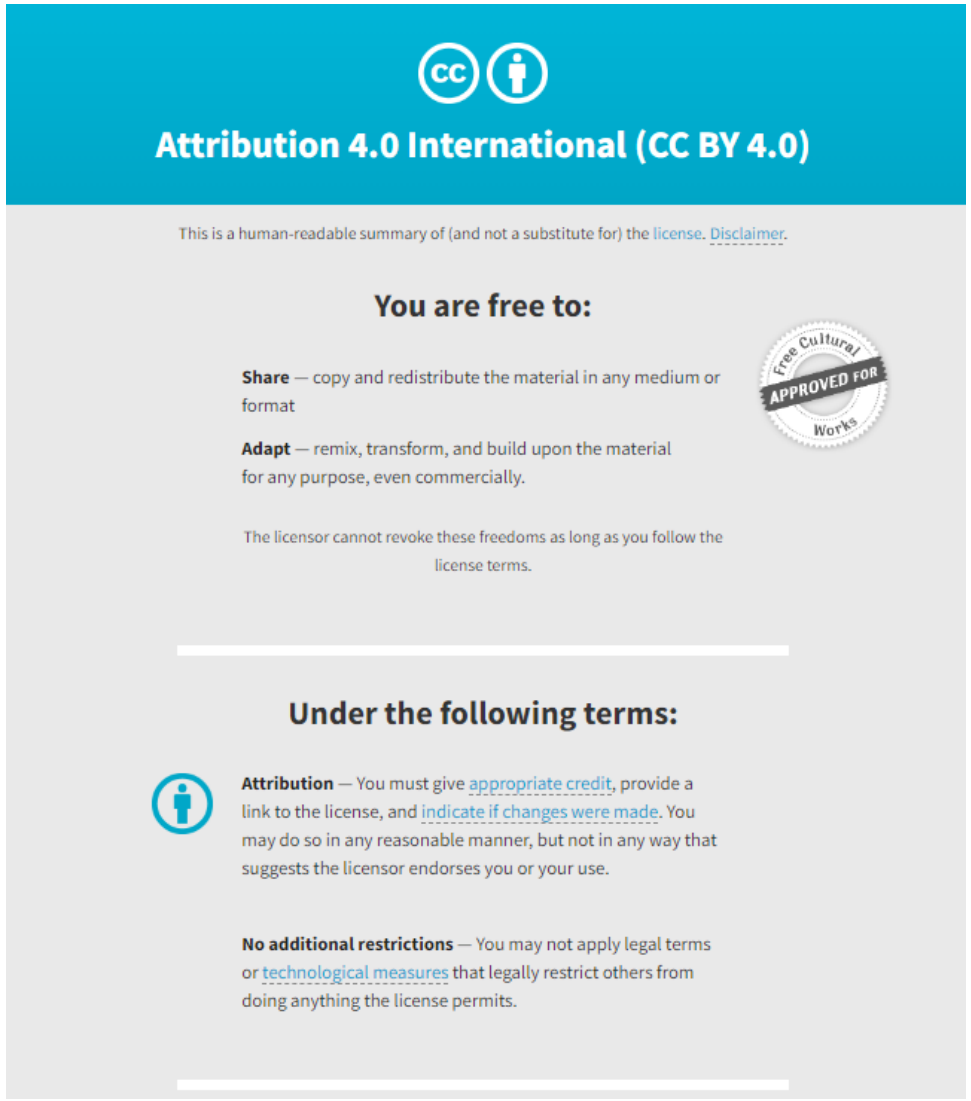
R1.2. (meta)data are associated with detailed provenance

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

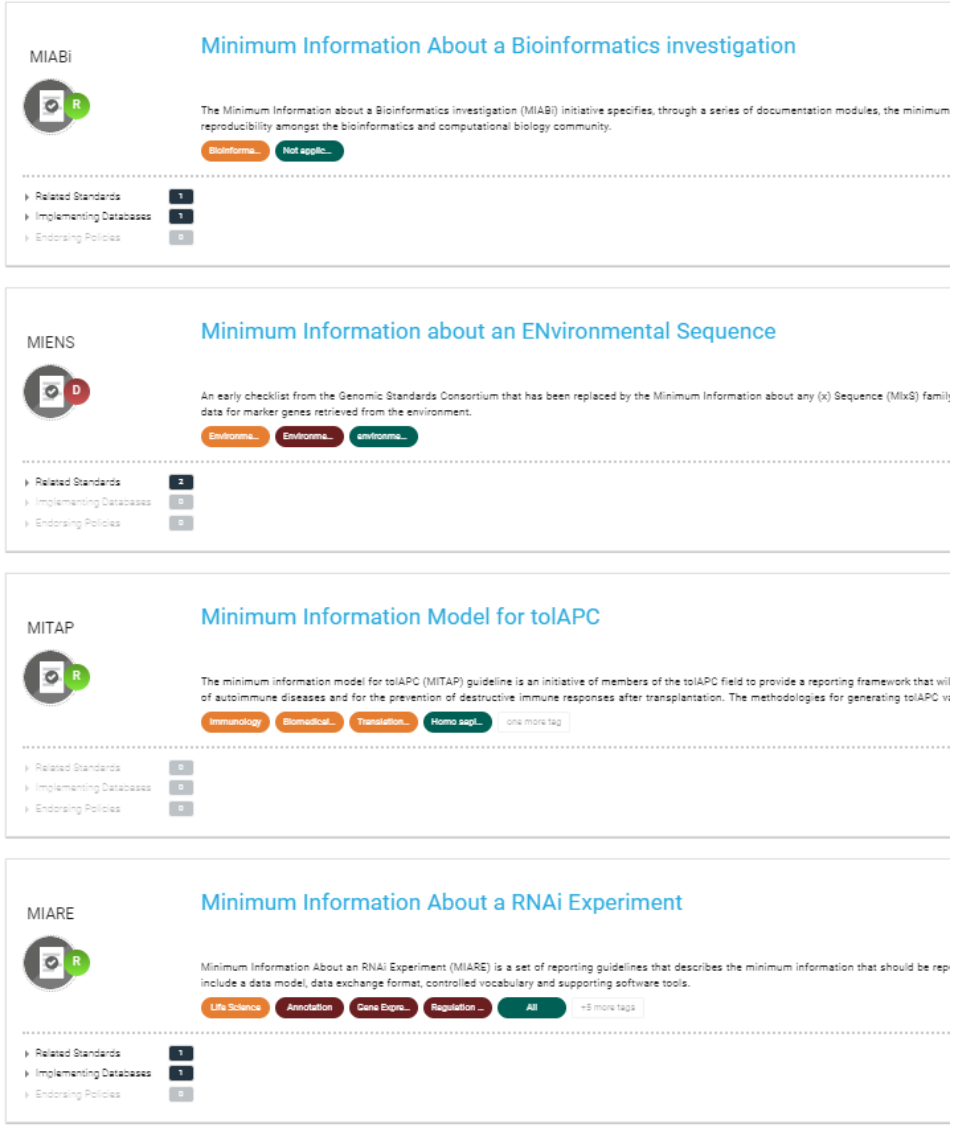
Murdoch University

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

Who to cite when the data is reused?

Domain-relevant community standards

Murdoch University



Use appropriate:

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

S

Summary

National Center for Biotechnology Information (NCBI)

Murdoch University

ENA European Nucleotide Archive

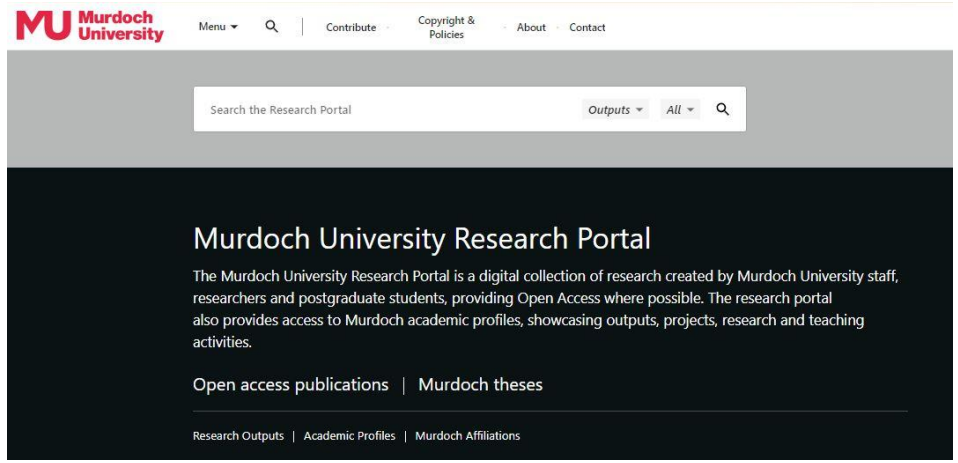
ARDC Research Data Australia

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**



Assign Unique Identifier

Indexed and Searchable

Publish

Advertise

Collaboration

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