



Australian Research Data Commons

FAIR for the Future: embracing all things data

PRESENTED BY

Keith Russell, Natasha Simons & Liz Stokes

14 February 2019



A R D C

Presentation Overview

Data sharing

What is it?

Who shares and why? Why not?

Why is the culture of data sharing growing?

What does data mean in STEM and HASS?

Why data should be FAIR

What are the FAIR principles?

Why are they of growing importance?

ARDC work in FAIR and HASS

Policy developments

Coming soon: update to the Code of Conduct

Data Policy Standardisation for Journals

What is data sharing?

- **Data sharing** is the practice of making data used for scholarly research available.
- **Types** of shared data might include:
 - Observational
 - Experimental
 - Simulation
 - Derived/compiled
 - Reference or canonical
- **Formats** of shared data might include:
 - Text, Image, Spreadsheets
 - Audio/video files
 - Drawings, Lab notebooks
 - Numeric data etc



Photo by [chuttersnap](#) on [Unsplash](#)

Studies on data sharing

Figshare open data survey 2017:

- 82% aware of open data sets
- 80% willing to reuse open data sets in own research
- 60% routinely share their data (frequently or sometimes)
- 21% have never made a data set openly available
- 74% are now curating their data for sharing
- 77% value a data citation the same as an article

Science, Digital (2017): The State of Open Data 2017 Report - Infographic.
figshare. <https://doi.org/10.6084/m9.figshare.5519155.v1> pp. 7-11

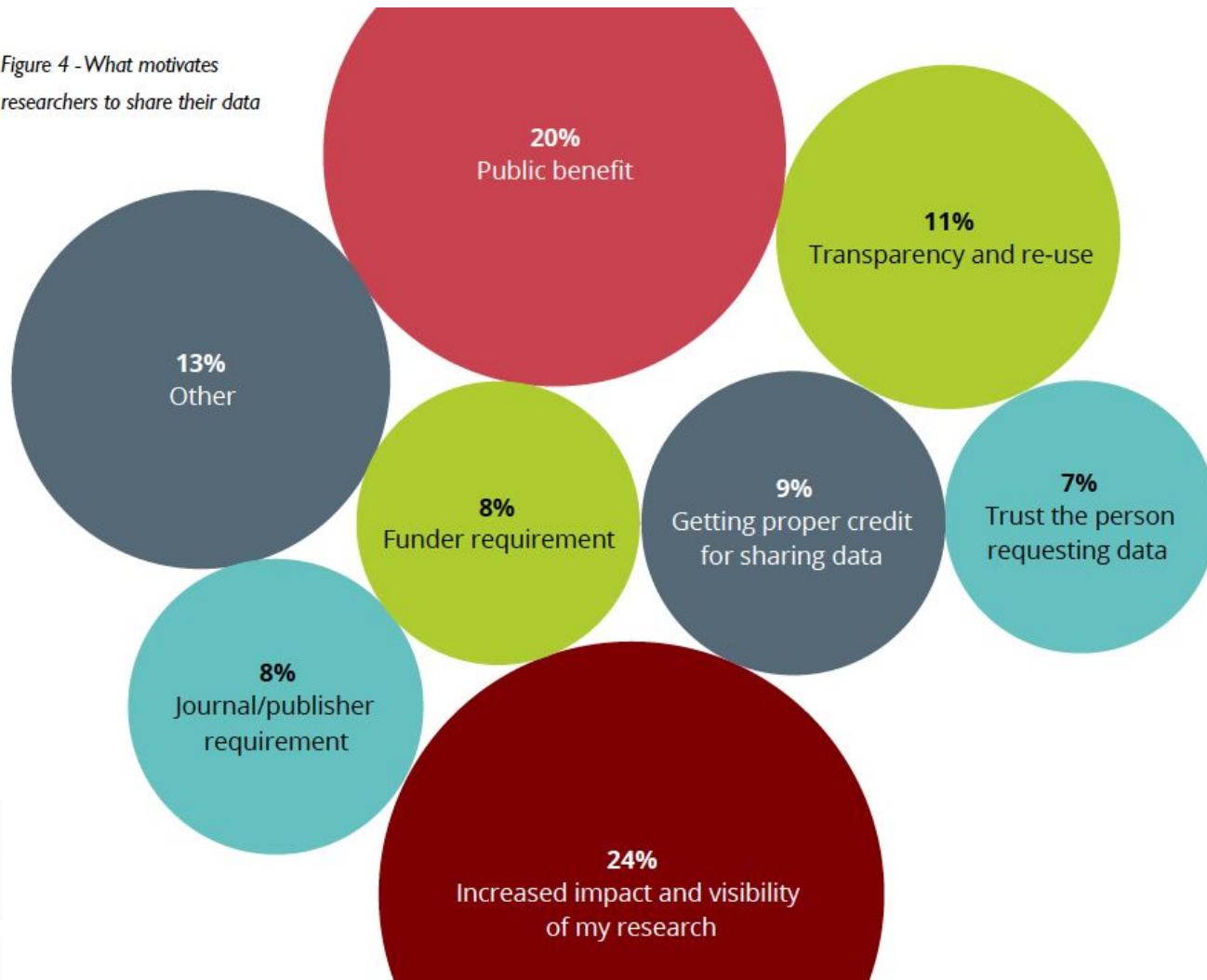
Studies on data sharing

Figshare open data survey 2017

- Where respondents publish their data:
 - +30% article appendix, -30% data repository, 20% data journal
- Why use open data?
 - +50% to validate own results
 - 50% to avoid duplication or complement own data
- 36% have lost data they were working on!

Motivations for sharing data

Figure 4 - What motivates researchers to share their data



We can see strong signals that open data is becoming more embedded [but] there is still a lack of confidence around open data.

Figshare open data survey 2017

Studies on data sharing

A 2011 study of 500 papers that were published in 2009 from 50 top-ranked research journals showed that only 47 papers (9%) of those reviewed had deposited full primary raw data online.

As another study notes, the number of datasets being shared annually has increased by more than 400% from 2011 to 2015, and this pace will likely continue.

What Constitutes Peer Review of Data? A Survey of Peer Review Guidelines by Todd A. Carpenter. Scholarly Kitchen blog post 11 April 2017.

<https://scholarlykitchen.sspnet.org/2017/04/11/what-constitutes-peer-review-research-data/>

Studies on data sharing

Global Data Sharing Trends*

Over 4,600 Wiley authors from 112 countries completed our 2016 Wiley Open Science Researcher Insights Survey

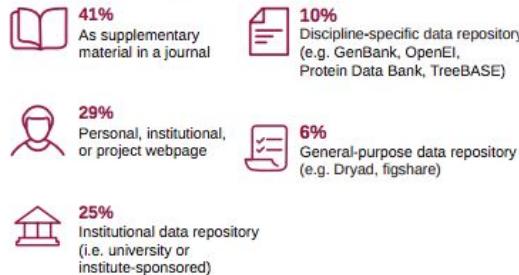
By collating results of our Wiley authors from surveys on Open Science topics in 2013, 2014, and 2016, we have started to build a valuable dataset for analysis and trend identification. Despite geographical and subject-level differences among authors, there are underlying commonalities in Open Science practices. The insights reported by our respondents show a willingness to move forward with open initiatives, but confusion around the best ways to do so.

Data sharing in 2016



More than two thirds of Wiley researchers reported they are now sharing their data. Though this varies geographically and across research disciplines we are seeing that more researchers are sharing their data and taking efforts to make it reproducible. Archiving in institutional repositories, public repositories, and personal web pages has almost doubled since 2014.

Ways data is shared

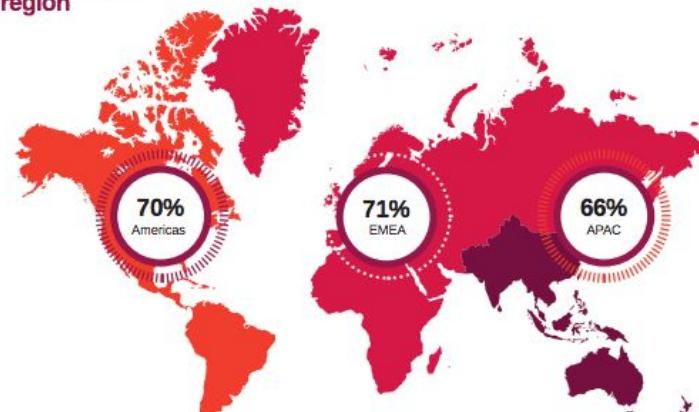


Researchers also report sharing their data in other ways including: 49% are sharing their data at conferences while 34% of researchers share their data upon informal request (email, direct contact, etc).

Top 4 researcher motivations for sharing data



Researchers sharing data by region



Data accessibility trends



Top 4 reasons why researchers are hesitant to share their data

- 1 50% - Intellectual property or confidentiality issues
- 2 31% - Ethical concerns
- 3 23% - I am concerned about misinterpretation or misuse of my research
- 4 22% - I am concerned that my research will be scooped

More than two thirds of Wiley researchers reported they are now sharing their data. Though this varies geographically and across research disciplines we are seeing that more researchers are sharing their data and taking efforts to make it reproducible.

Wiley Global Data Sharing Infographic June 2017.

<https://authorservices.wiley.com/author-resources/Journal-Authors/licensing-open-access/open-access/data-sharing.html>

Why data sharing is growing

- Funder data sharing policies e.g. [Wellcome Trust](#).
- Publisher/Journal data sharing policies e.g. [PLoS](#), [Springer Nature](#), [Elsevier](#), [Wiley](#)
- Government open data initiatives e.g. [USA](#), [Australia](#), [Europe](#)
- Publishers are at the forefront of data sharing policy initiatives e.g. [COPDESS](#), [TOP Guidelines](#), [JDAP](#)

What does “data” mean in STEM and HASS?

Data sharing in STEM

- Machines to humans and back again
- Transferring big data over long distances
- Hardware & software development
- Collaboration across disciplines
 - IGSN for physical samples
 - Electronic lab notebooks
 - Licensing data for reuse

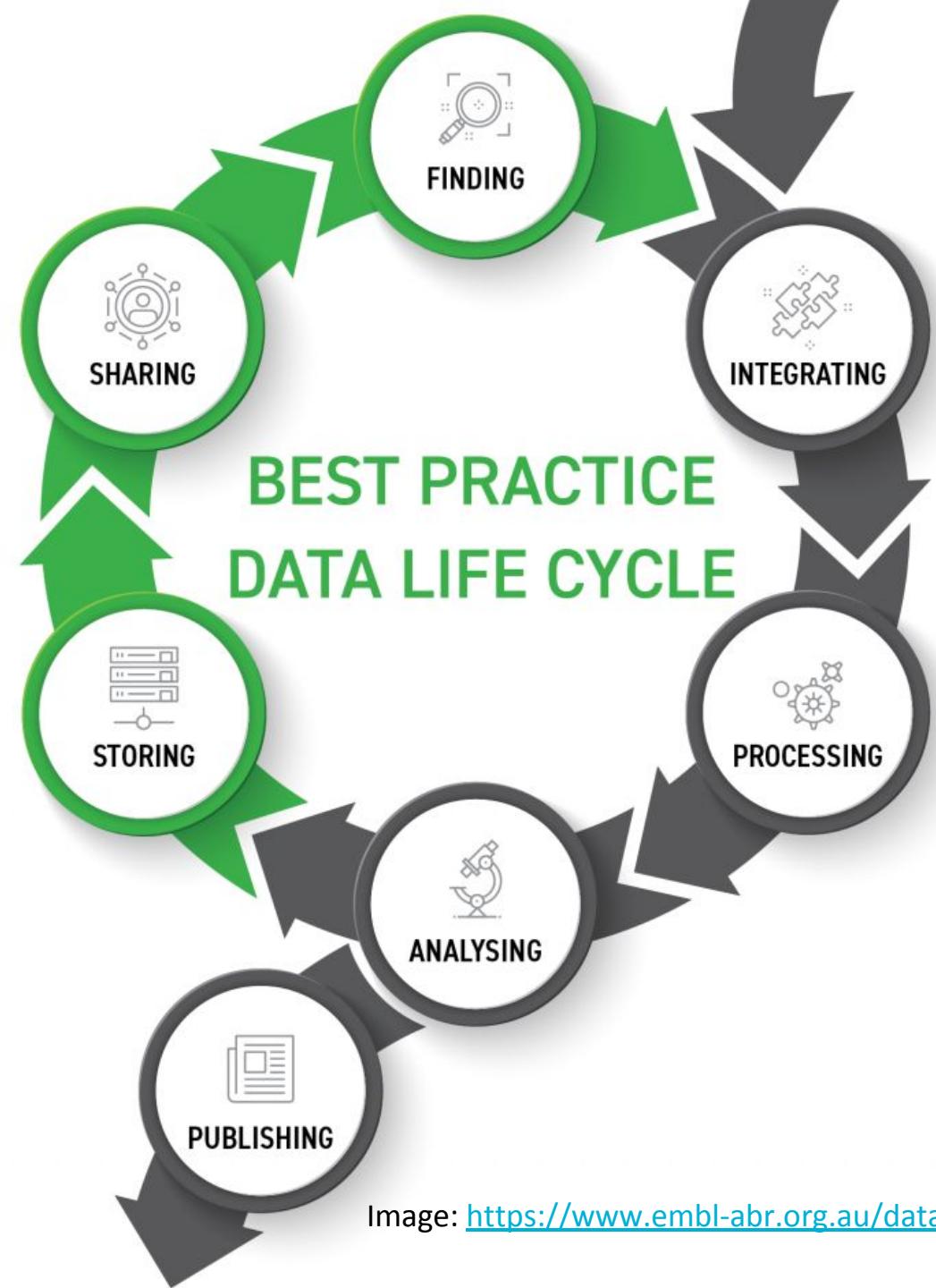


Image: <https://www.embl-abr.org.au/data/>

Data sharing in HASS

- You say “data” but what does that really mean?
- Scientific workflows do not translate directly to humanities scholarship
- Many digital humanities projects = open data principles, data viz
- Humanities data is not meant to spark joy.



History & information about Australian orphanages, children's Homes & other institutions



The Prosecution Project



Unlocking and uniting Australian
cultural data: combine, collect, connect
and collaborate

Data sharing in HASS

Map of Indigenous massacres grows to include more sites of violence across Australia

By national Indigenous affairs correspondent [Bridget Brennan](#)

Updated 27 Jul 2018, 3:17pm



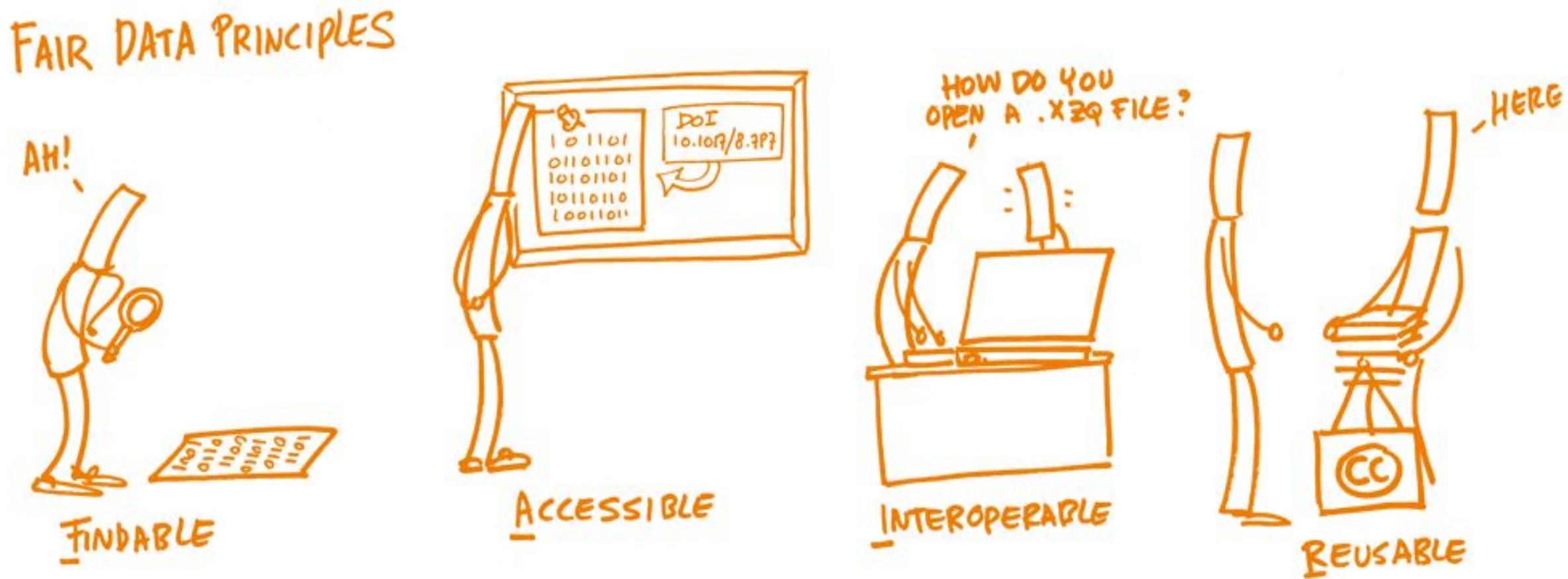
ARDC work in the humanities

Humanities, Arts & Social Sciences Data Enhanced Virtual Lab (Tinker)

- ARDC supported project, \$1 million including co-investment
- Twelve project partners



The FAIR data principles



Findable

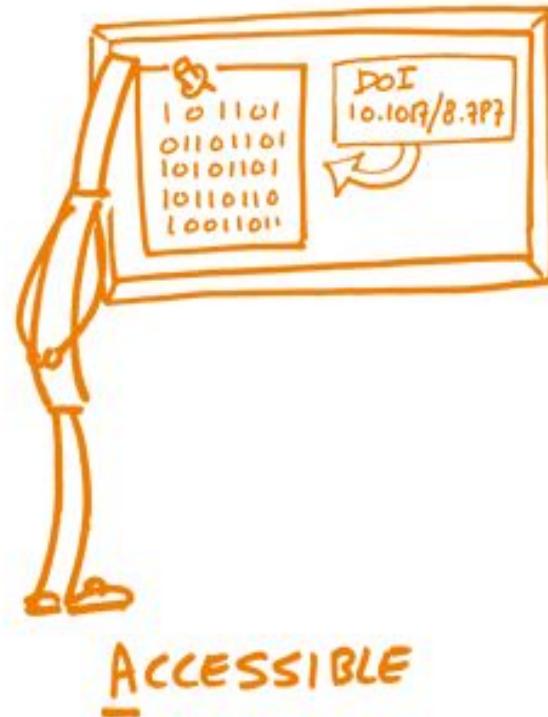
FAIR DATA PRINCIPLES



FINDABLE

Accessible

FAIR DATA PRINCIPLES



Interoperable

FAIR DATA PRINCIPLES

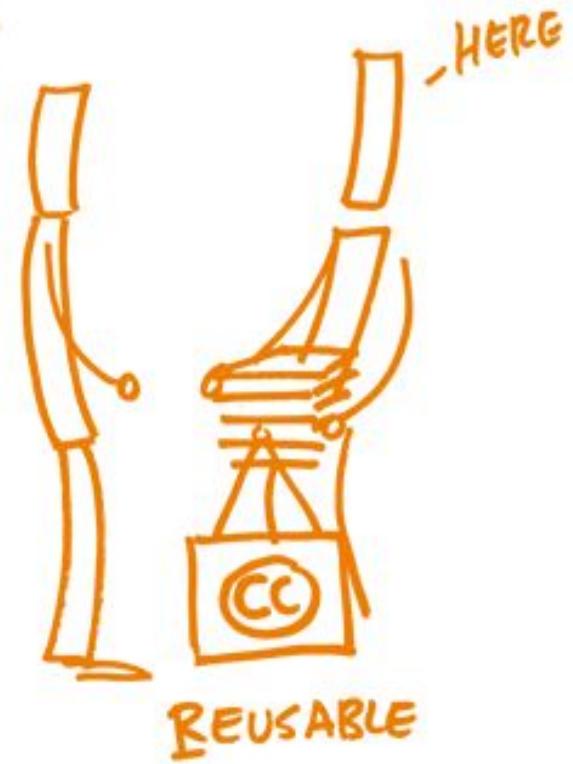
HOW DO YOU
OPEN A .XZQ FILE?



INTEROPERABLE

Reusable

FAIR DATA PRINCIPLES



FAIR in action - advocacy

COPDESS

Coalition for Publishing Data in the Earth and Space Sciences

The Coalition for Publishing Data in the Earth and Space Sciences ▾

Enabling FAIR Data Project ▾



COMMITMENT STATEMENT IN THE EARTH, SPACE, AND ENVIRONMENTAL SCIENCES

[HOME](#) / [ENABLING FAIR DATA PROJECT](#) / COMMITMENT STATEMENT IN THE EARTH, SPACE, AND ENVIRONMENTAL SCIENCES

SIGN ON

SIGNATORIES

FAQS

Imagine a world where the preponderance of Earth, space, and environmental science data, software, and models are routinely shared in ways that allow easy discovery, recombination, reuse, and to test reliability, and where information about samples, methods, and tools are standardized, available, and linked across publications.

FAIR in action - repositories



Accessing Data

Depositing Data

Australian Data Archive certified with the Core Trust Seal

Feb 15, 2018 | News



FAIR in action - research support

Findable	
Does the dataset have any identifiers assigned?	No identifier
Is the dataset identifier included in all metadata records/files describing the data?	No
How is the data described with metadata?	The data is not described
What type of repository or registry is the metadata record in?	The data is not described in any repository
Accessible	
Interoperable	
Reusable	
Total across F.A.I.R	

FAIR in action - skills



Top 10 FAIR Data & Software Things

February 1, 2019

Christopher Erdmann, Natasha Simons, Reid Otsuji, Stephanie Labou, Ryan Johnson, Guilherme Castelao, ... Tim Dennis. (2019, February). Top 10 FAIR Data & Software Things. Zenodo. <http://doi.org/10.5281/zenodo.2555498>

FAIR data principles in action

Researchers

ORCID

Know your institutional
repository

ASK for DOIs

Think about the long term

Editors

Promote trusted
repositories
Standardise journal data
policy
Make good data practice
easy

Research Support

Facilitate good data
practices

Connect researchers
with the right tools and
communities

2018 Revised Australian Code for the Responsible Conduct of Research

Timeline



For Institutions

R8 Provide access to facilities for the safe and secure storage and management of research data, records and primary materials and, where possible and appropriate, allow access and reference

[2018 Australian Code for the Responsible Conduct of Research](#)

For Researchers

R22 Retain clear, accurate secure and complete records of all research including research data and primary materials. Where possible and appropriate, allow access and reference to these by interested parties.

R27 Cite and acknowledge other relevant work appropriately and accurately

[2018 Australian Code for the Responsible Conduct of Research](#)

Data Policy Standardisation for Journals

 **RDA**
RESEARCH DATA ALLIANCE

O&A Members 55
Active Organisational & Affiliate members

MEMBERSHIP Members: 7852
Becoming a member of RDA is simple and open to both individuals and organizations
[Register now](#)

WORKING GROUPS Groups: 101
Discover what RDA Working and Interest Groups and all other Groups are up to and find out how to join them. [Explore Groups](#)

ABOUT RDA ▾ GET INVOLVED ▾ GROUPS ▾ RECOMMENDATIONS & OUTPUTS ▾ RDA FOR DISCIPLINES ▾ PLENARIES & EVENTS ▾ NEWS & MEDIA ▾ Q

Data policy standardisation and implementation IG

[Home](#) » [Working And Interest Groups](#) » [Interest Group](#) » [Data Policy Standardisation And Implementation IG](#)

IG **Group details**

Status: Recognised & Endorsed
Chair (s): Iain Hrynaszkiewicz, Natasha Simons, Simon Goudie, Azhar Hussain
Secretariat Liaison: enquiries@rd-alliance.org
TAB Liaison: Devika Madalli

 IG Established

[History](#)

Introduction

Increasing the availability of research data for reuse is in part being driven by research data policies and the number of funders and journals and institutions with some form of research data policy is growing. The research data policy landscape of funders, institutions and publishers is however too complex (Ref: <http://insights.uksg.org/articles/10.1629/uksg.284/>) and the implementation and implications of policies for researchers can be unclear. While around half of researchers share data, their primary motivations are often to carry out and publish good research, and to receive renewed funding, rather than making data available. Data policies that support publication of research need to be practical and seen in this context to be effective beyond specialist data communities and publications.

Data policy standardisation and implementation IG

Status: Recognised & Endorsed
Secretariat Liaison: enquiries@rd-alliance.org
TAB Liaison: Devika Madalli

Public - accessible to all site users

[Join Group](#)

Index	Add new content	
-----------------------	---------------------------------	--

 Click here to create a wiki index for this group.
 Group Mailing list Archive

Co-chairs from Springer Nature, Wiley, Jisc UK and ARDC



Policy stakeholders

Stakeholder	Contributes	Challenges	Who can help?
Researchers	Data	Clear requirements? Help to make data available?	Journal Editors Data Librarians Publishers
Journal Editors	Policy	Where to start? What should be included?	Publishers Others e.g. ANDS
Data repositories	Technical infrastructure	How to support?	Publishers Journal Editors Researchers Data Librarians
Peer reviewers	Reviews	Clear guidelines? Workflow support?	Journal Editors Publishers Repositories

It's in your hands

“Technologies can be consolidated, and processes automated, but collection, creation, curation and research use/reuse of data is ultimately dependent on the domain and disciplinary expertise of the humans who know them best.”

Hervé L'Hours, UK Data Archive

<https://www.linkedin.com/pulse/formal-certification-data-repositories-key-fair-eosc-herv%C3%A9-l-hours/>



Australian Research Data Commons

CONTACT

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

+61 7 3365 1120
info@ardc.edu.au
ardc.edu.au

keith.russell@ardc.edu.au
natasha.simons@ardc.edu.au
liz.stokes@ardc.edu.au

