

NATIONAL HIGH

# MAGNETIC FIELD LABORATORY

## Building a More FAIR MagLab

David S. Butcher, FAIR Data Specialist

October 12, 2021

**FAIR DATA WORKSHOP**



**Los Alamos**  
NATIONAL LABORATORY



# FAIR Principles

To be **FAIR**, data must be **Findable, Accessible, Interoperable, and Reusable\***

## Findable

- F1. (Meta)data are assigned a globally unique and persistent identifier
- F2. Data are described with rich metadata (defined by R1 below)
- F3. Metadata clearly and explicitly include the identifier of the data they describe
- F4. (Meta)data are registered or indexed in a searchable resource

## Accessible

- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol
  - A1.1 The protocol is open, free, and universally implementable
  - A1.2 The protocol allows for an authentication and authorisation procedure, where necessary
- A2. Metadata are accessible, even when the data are no longer available

## 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

## 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

# Why go FAIR?



## NSF Data Sharing Policy

*“Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing.”*

- Dissemination and Sharing of Research Results, NSF Data Management Plan Requirements  
(<https://www.nsf.gov/bfa/dias/policy/dmp.jsp>)



*“The NIH expects and supports the timely release and sharing of final research data from NIH-supported studies for use by other researchers. Starting with the October 1, 2003 receipt date, investigators submitting an NIH application seeking \$500,000 or more in direct costs in any single year are expected to include a plan for data sharing or state why data sharing is not possible.”*

- Final NIH Statement on Sharing Research Data  
(<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>)

# Why go FAIR?

## Science

*“...the Science Journals generally require all data underlying the results in published papers to be publicly and immediately available. Post-publication embargoes are not permitted, nor are stipulations for readers to contact the authors...”*

- <https://www.science.org/content/page/science-journals-editorial-policies>

## nature portfolio

*“A condition of publication in a Nature Portfolio journal is that authors are required to make materials, data, code, and associated protocols promptly available to readers without undue qualifications.”*

- <https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards>

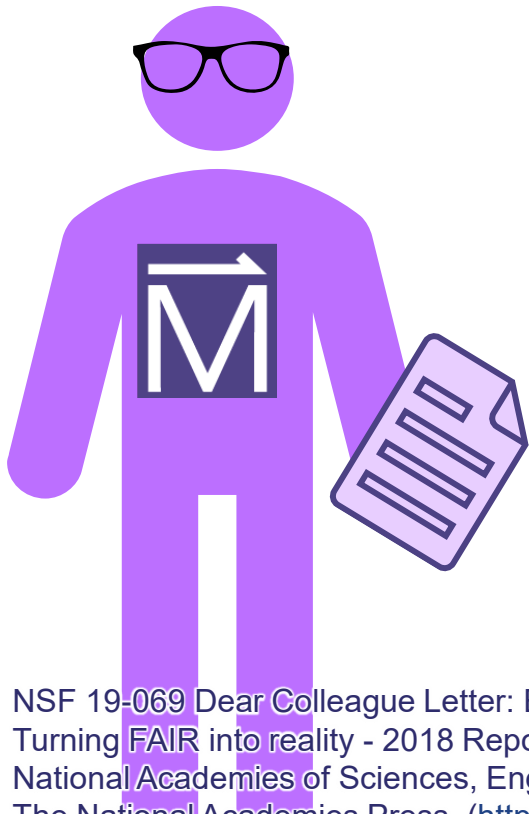


*“PLOS journals require authors to make all data necessary to replicate their study’s findings publicly available without restriction at the time of publication. When specific legal or ethical restrictions prohibit public sharing of a data set, authors must indicate how others may obtain access to the data.”*

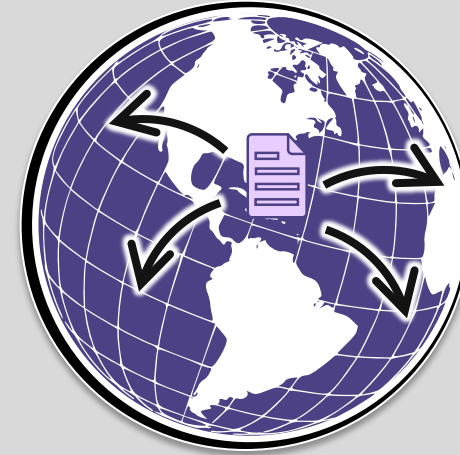
- <https://journals.plos.org/plosone/s/data-availability>

# Why go FAIR?

Making data FAIR  
(and/or open)  
provides a variety  
of benefits ...



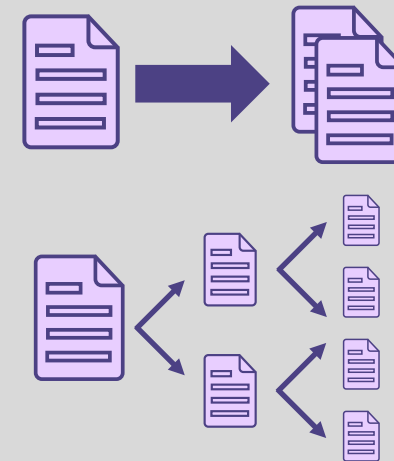
Enables  
reuse



Allows  
broader  
access



Provides  
greater  
ROI



Facilitates  
replication  
and  
expansion

1. NSF 19-069 Dear Colleague Letter: Effective Practices for Data (<https://www.nsf.gov/pubs/2019/nsf19069/nsf19069.pdf>)
2. Turning FAIR into reality - 2018 Report and Action Plan from the European Commission Expert Group on FAIR Data (<https://doi.org/10.2777/54599>)
3. National Academies of Sciences, Engineering, and Medicine 2018. Open Science by Design: Realizing a Vision for 21st Century Research. Washington, DC: The National Academies Press. (<https://doi.org/10.17226/25116>)

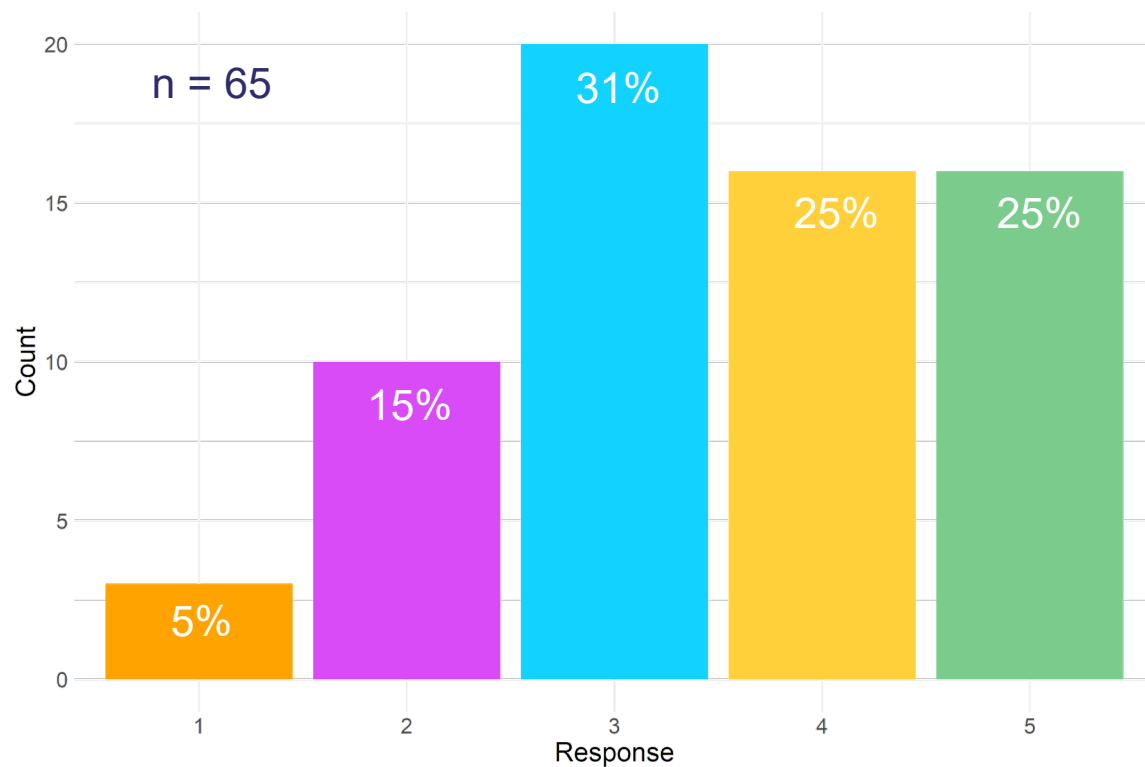
# Supporting users

## How does the MagLab plan to support users in making data FAIR?

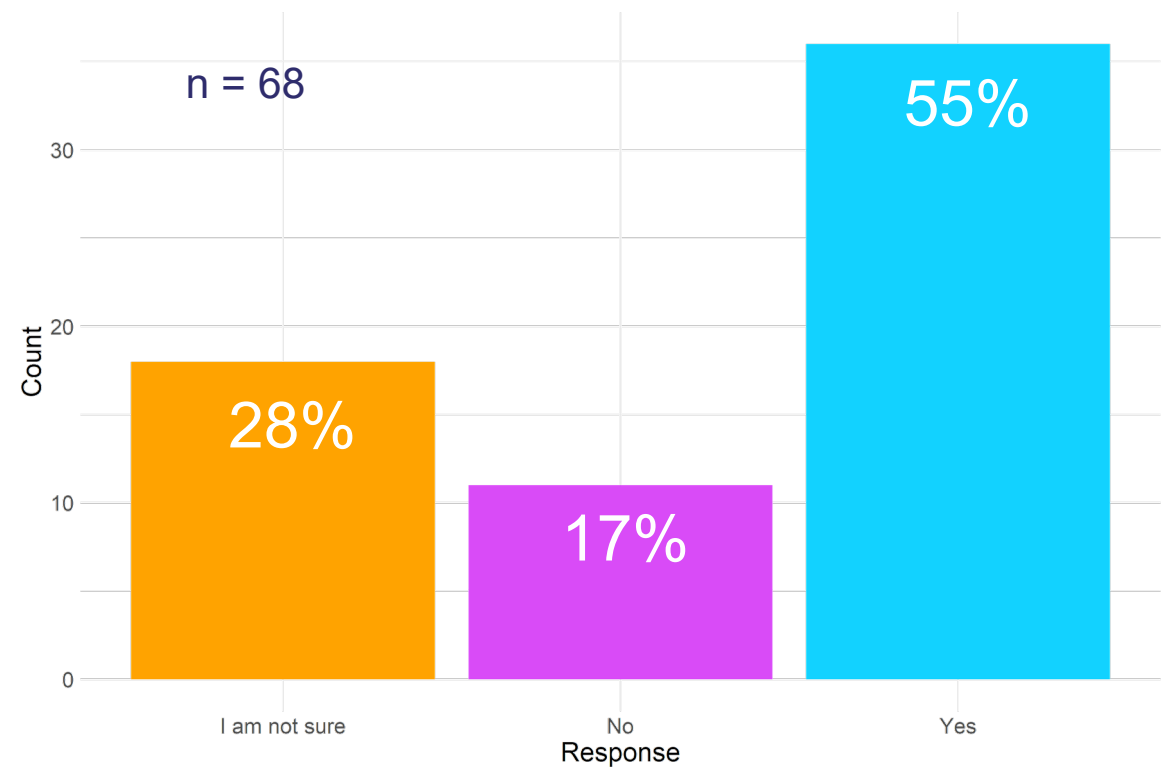
- Promote awareness of FAIR principles by developing FAIR-aligned data management plans and teaching users and staff best practices for FAIR data management and access
- Provide support in the capture, structure, and formatting of data/metadata to adhere to open, machine-actionable community standards
- Provide support in the use of FAIR-aligned data repositories
- Interface with international organizations to survey existing and emerging community standards for making data FAIR
- Communicate MagLab FAIR data progress with other user facilities, especially those under NSF-DMR

# Survey results

Please rate your familiarity with the FAIR principles on a scale of 1 (no familiarity) to 5.

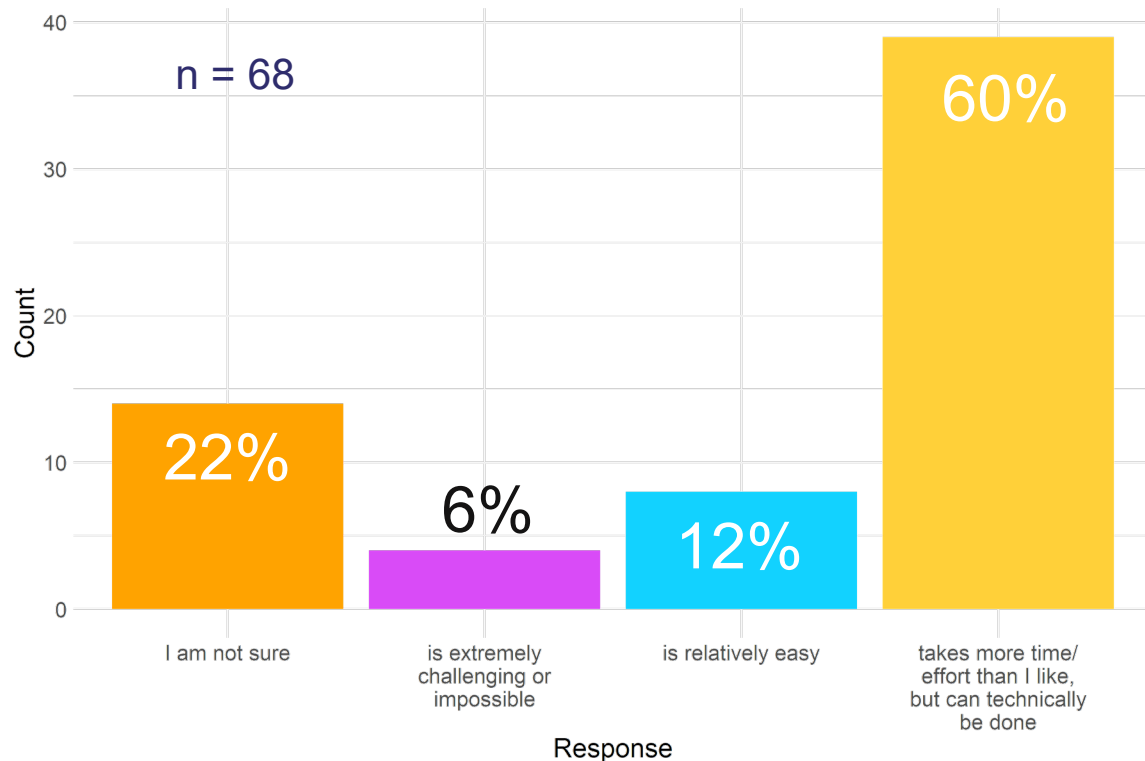


Do you (or members of your group) disseminate data via publicly accessible repositories (e.g. Materials Cloud, Protein Data Bank, Marine Geosciences Data System, figshare, etc.)?

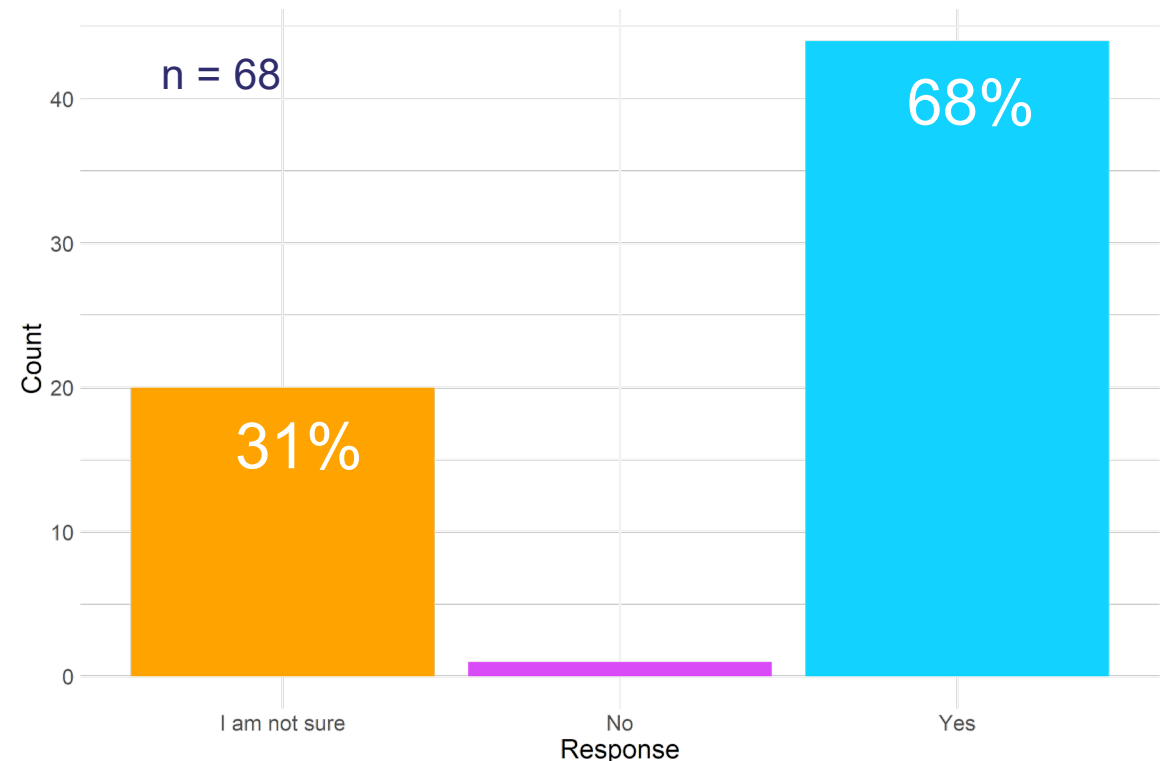


# Survey results

Please complete the following sentence:  
"I think that ensuring that the data I  
produce are FAIR \_\_\_\_\_."



Would you be willing to work with us to ensure that  
all products of research related to your user  
proposal that are generated at the MagLab adhere  
to the FAIR principles?





We look forward to hearing your feedback about how we at the MagLab can work together with users to make our data more **FAIR!**



David S. Butcher  
([dbutcher@magnet.fsu.edu](mailto:dbutcher@magnet.fsu.edu))



# FAIR and Open Data

*Open science aims to ensure the free availability and usability of scholarly publications, the data that result from scholarly research, and the methodologies, including code or algorithms, that were used to generate those data.*

- Open Science by Design: Realizing a Vision for 21st Century Research\*

***FAIR is not equal to Open: The ‘A’ in FAIR stands for ‘Accessible under well defined conditions’. There may be legitimate reasons to shield data and services generated with public funding from public access. These include personal privacy, national security, and competitiveness. The FAIR principles, although inspired by Open Science, explicitly and deliberately do not address moral and ethical issues pertaining to the openness of data.***

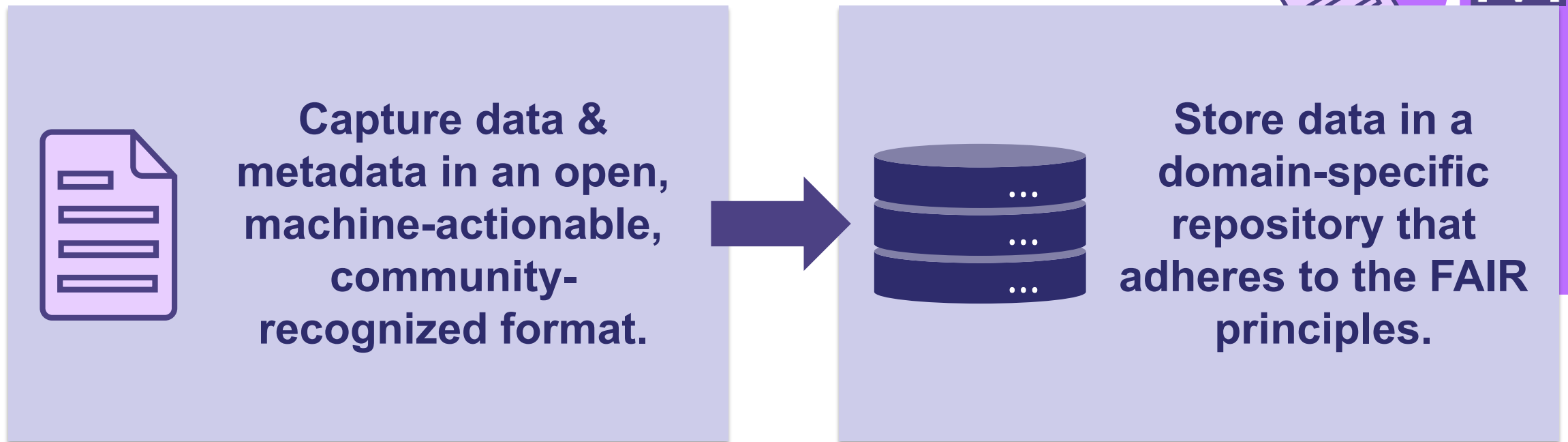
- Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud†

\*: National Academies of Sciences, Engineering, and Medicine 2018. *Open Science by Design: Realizing a Vision for 21st Century Research*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25116>.

†: 1.Mons, B. et al. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. *Information Services & Use* 37, 49–56 (2017). <https://doi.org/10.3233/ISU-170824>

# How do we go FAIR?

The procedure for making data FAIR will vary depending on the nature of the data and the research domain, but generally we should...



# Why go FAIR?

*“Scientific data are burgeoning — thousands of petabytes were collected in 2018 alone. But these data are not being used widely enough to realize their potential. Most researchers come up against obstacles when they try to get their hands on data sets. Only one-fifth of published papers typically post the supporting data in scientific repositories — as has been shown by PLoS ONE<sup>1</sup>. Too much valuable, hard-won information is gathering dust on computers, disks and tapes.”*

**- Make scientific data FAIR\***

\*: Stall, S. *et al.* Make scientific data FAIR. *Nature* **570**, 27–29 (2019).

1: Federer, L. M. *et al.* Data sharing in PLOS ONE: An analysis of Data Availability Statements. *PLOS ONE* **13**, e0194768 (2018).

# Challenges of FAIR Data

## Making data FAIR comes with a variety of challenges:

- Simultaneously ensuring accessibility and protection of personal data<sup>1</sup>
- Providing access to metadata and workflows linked to FAIR data<sup>2</sup>
- Automation of processes dealing with data, metadata, and workflows<sup>3</sup>
- Providing the necessary attributes to make data FAIR without posing an unreasonable burden on the data producer<sup>4</sup>
- Developing tools and services which are useful across a wide range of research disciplines
- Lack of mechanisms to recognize and reward FAIR data and data stewardship<sup>5</sup>

1. Landi, A. et al. *Data Intelligence* 2, 47–55 (2020).

2. Goble, C. et al. *Data Intelligence* 2, 108–121 (2020).

3. Weigel, T., Schwarzmann, U., Klump, J., Bendoukha, S. & Quick, R. *Data Intelligence* 2, 40–46 (2020).

4. Thompson, M., Burger, K., Kaliyaperumal, R., Roos, M. & da Silva Santos, L. O. B. *Data Intelligence* 2, 87–95 (2020).

5. Turning FAIR into reality - 2018 Report and Action Plan from the European Commission Expert Group on FAIR Data (<https://doi.org/10.27777/54599>)

# Challenges of FAIR Data

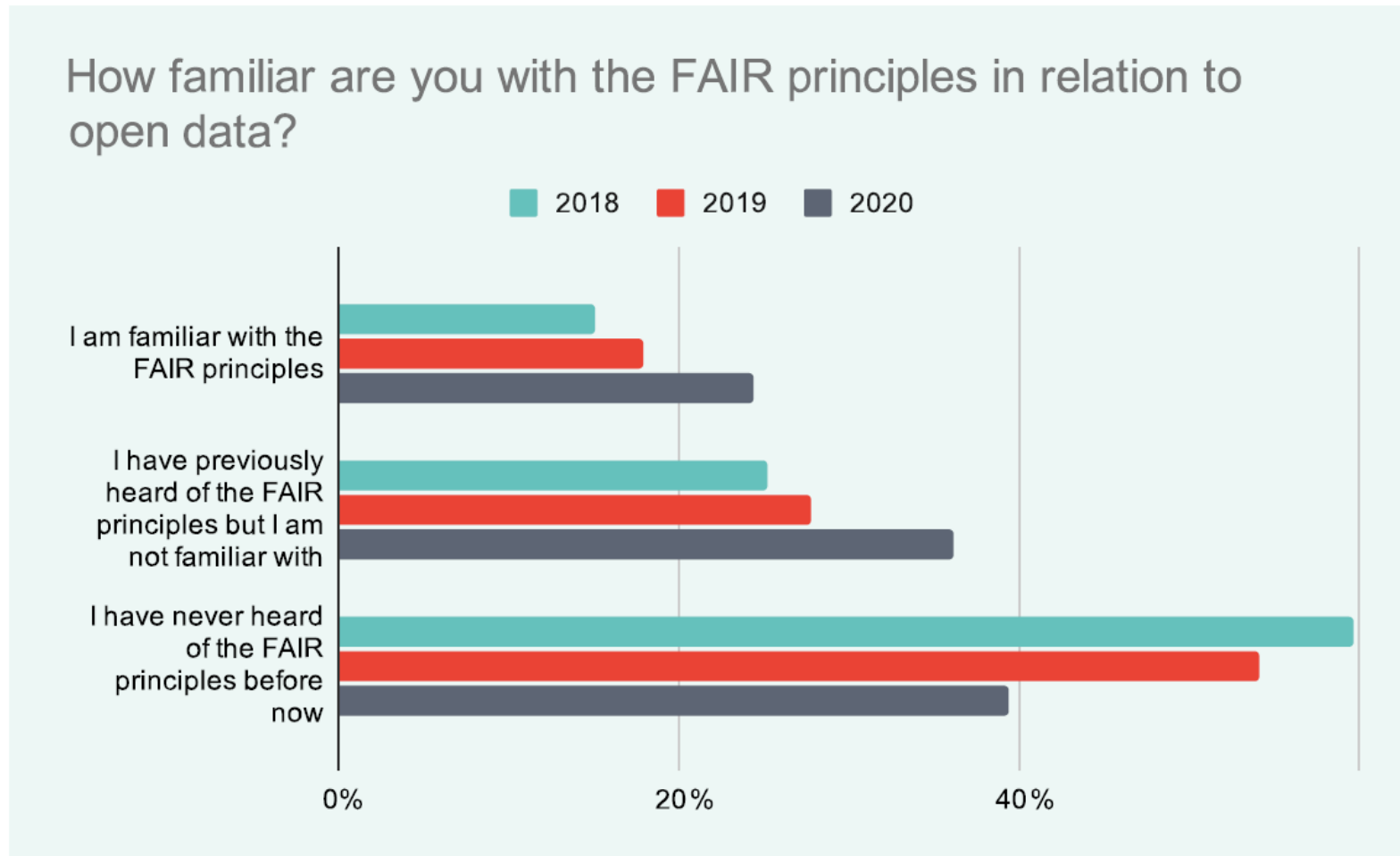


Figure from The State of Open Data 2020 report (<https://doi.org/10.6084/m9.figshare.13274744>). Available under a [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.