

# **FAIR Earth sciences**

#### Helen Glaves ICS-C Director











### Defining 'Open Science'

- Transparency in experimental methodology, observation, and collection of data
- Public availability and reusability of scientific data
- Public accessibility and transparency of scientific communication
- Using web-based tools to facilitate scientific collaboration

Dan Gezelter (2009) http://openscience.org/what-exactly-is-open-science/



## What is FAIR?

- set of guiding principles for data stewardship and management
- guidelines to improve the <u>F</u>indability, <u>A</u>ccessibility, <u>Interoperability, and <u>R</u>euse of digital assets
  </u>
- emphasise machine-actionability
- refer to three types of entities:
  - data (or any digital object),
  - metadata (information about that digital object)
  - infrastructure





#### Findable

- The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machinereadable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the FAIRification process.
- 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

- Once the user finds the required data, she/he/they need to know how they can be accessed, possibly including authentication and authorisation.
- 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

- The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.
- 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

- The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.
- 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

#### **Evaluating FAIRness of data**

- FAIR does not mean open and free
- FAIR principles define the conditions and provides instructions for data reuse
  - not a strict definition of FAIRness
  - open to interpretation
- FAIR assessment tools
  - domain relevant and community-defined
  - two distinct types of tool
    - manual, questionnaire based assessment e.g. ARDC FAIR assessment tool
    - automated systems to benchmark digital research assets against predefined FAIRness criteria e.g. F-UJI Automated FAIR Data Assessment Tool <u>https://www.fairsfair.eu/f-uji-automated-fair-dataassessment-tool</u>



# EPOS Research Infrastructure: promoting adoption of FAIR principles

- Research Infrastructures provide harmonised access to data e.g. EPOS – solid Earth data
- Promoting adoption of FAIR principles by data providers
- Turning principles into practical implementation
- EPOS architecture was FAIR-compliant from its inception:
  - implemented key concepts necessary for FAIR compliance e.g. metadata and ontologies, identifiers, technological interoperability
  - methodology was applied to an existing research infrastructure.



#### **Related** initiatives

- ENVRI-FAIR (<u>https://envri.eu/current-envri-projects/</u>)
  - EU-funded project promoting findability, accessibility, interoperability, and reusability of digital assets through the development of FAIR compliant common policies and technical solutions
  - fostering development of FAIR data services by participating environmental Research Infrastructures
  - integration of services to support European Open Science Cloud (EOSC)
- GO-FAIR (<u>https://www.go-fair.org/</u>)
  - contributing to and coordinating development of the Internet of FAIR Data & Services (
  - focused on priorities, policies and incentives for FAIR implementation and compliance
- Research Data Alliance (RDA) FAIR Data Maturity Model WG <u>https://www.</u> rd-alliance.org/groups/fair-data-maturity-model-wg
  - developing common set of core assessment criteria for FAIRness
  - creating a generic and expandable self-assessment model for measuring the maturity level of a dataset.
- European Open Science Cloud (EOSC) <u>https://eosc.eu/about-eosc</u>
  - aims to develop a multidisciplinary 'Web of FAIR Data and services' for science in Europe



## **Thank You**





#### References

Mark D. Wilkinson; Michel Dumontier; IJsbrand Jan Aalbersberg; et al. (15 March 2016). "The FAIR Guiding Principles for scientific data management and stewardship". Scientific Data. 3 (1): 160018. doi:10.1038/SDATA.2016.18. ISSN 2052-4463. PMC 4792175. PMID 26978244. Wikidata Q27942822.

Wilkinson, M. D. et al. A design framework and exemplar metrics for FAIRness. Sci. Data 5:180118 doi: 10.1038/sdata.2018.118 (2018).

Petzold, A., and Glaves, H. (2018). ENVRI-FAIR - the next step towards FAIRer environmental research. Am. Geophys. Union Fall Meet. 2019:15684.

Bailo D, Paciello R, Sbarra M, Rabissoni R, Vinciarelli V and Cocco M (2020) Perspectives on the Implementation of FAIR Principles in Solid Earth Research Infrastructures. Front. Earth Sci. 8:3. doi: 10.3389/feart.2020.00003

Cocco, M., Bailo, D., Jeffery, K. G., Paciello, R., Vinciarelli, V., and Freda, C.: Sustainable FAIR Data management is challenging for RIs and it is challenging to solid Earth scientists, EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-18570, https://doi.org/10.5194/egusphere-egu2020-18570, 2020

Wilkinson, M.D., Dumontier, M., Sansone, SA. et al. Evaluating FAIR maturity through a scalable, automated, community-governed framework. Sci Data 6, 174 (2019). https://doi.org/10.1038/s41597-019-0184-5



