

# The FAIR Principles

Mercè Crosas, Ph.D.


University Research Data Management Officer, HUIT

Chief Data Science and Technology Officer, IQSS

Harvard University

Open Access | Published: 15 March 2016

# The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J.G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, [Jaap Heringa](#), Peter A.C 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao & Barend Mons  -Show fewer authors

*Scientific Data* **3**, Article number: 160018 (2016) | [Cite this article](#)

**149k** Accesses | **1776** Citations | **1597** Altmetric | [Metrics](#)

## FAIR Principles Popularity

- **Published in 2016**
- **54 authors**
- **1700 citations**
- **149K accesses**
- **Adoption by funding agencies, research communities**

## European Commission Data Policy:

"The OBJECTIVES for adopting and implementing the JRC (Joint Research Centre) Data Policy include: [...]Facilitate management, broaden access and use of the JRC data, in line with FAIR Data principles (Findable, Accessible, Interoperable, Reusable);"

<https://ec.europa.eu/jrc/en/publication/jrc-data-policy>

National Health Institutes (NIH) proposed DRAFT Data Policy:

"NIH encourages data management and data sharing practices consistent with the NIH Plan for Increasing Access to Scientific Publications and Digital Scientific Data from NIH Funded Scientific Research and the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles."

[https://osp.od.nih.gov/wp-content/uploads/Draft\\_NIH\\_Policy\\_Data\\_Management\\_and\\_Sharing.pdf](https://osp.od.nih.gov/wp-content/uploads/Draft_NIH_Policy_Data_Management_and_Sharing.pdf)

## Coalition for Publishing Data in Earth and Space Science (CODPESS)

### Commitment Statement:

**"Ensuring that Earth, space, and environmental science research outputs, including data, software, and samples or standard information about them, are open, FAIR, and curated in trusted domain repositories whenever possible ..."**

<http://www.codpess.org/enabling-fair-data-project/commitment-to-enabling-fair-data-in-the-earth-space-and-environmental-sciences/>

# Emphasis on Machine-Actionability

“ The **FAIR** Principles put specific emphasis on enhancing the ability of **machines to automatically find and use the data**, in addition to supporting its **reuse by individuals**. ”

Wilkinson et al. 2016.

# Principle 1: To Be Findable

"Digital resources should be easy to find for both humans and computers. **Extensive machine-actionable metadata** are essential for automatic discovery of relevant datasets and services, and are therefore an essential component of the FAIRification process."

Jacobsen et al. 2020.

- F1: Globally unique and persistent identifier (e.g., DOI)
- F2: Rich metadata
- F3: Metadata includes ID of data
- F4: Meta(data) registered in a searchable resource (e.g., DataCite)

# Principle 2: To Be Accessible

“**Protocols** for retrieving digital resources should be made explicit, for both humans and machines, including well-defined mechanisms to obtain **authorization** for access to protected data.”

- A1: Standardized communication protocol:
  - A1.1: open, free
  - A1.2: authentication & authorization when needed
- A2: Metadata accessible even when data no longer available



# Principle 3: To Be Interoperable

“When two or more digital resources are related to the same topic or entity, it should be possible for machines to merge the information into a richer, unified view of that entity. Similarly, when a digital entity is capable of being processed by an online service, a machine should be capable of automatically detecting this compliance and facilitating the interaction between the data and that tool.”

- I1: Formal, accessible, shared and broadly applicable language
- I2: Vocabularies that follow FAIR principles
- I3: Include qualified references to other (meta)data

# Principle 4: To Be Reusable

"Digital resources are sufficiently well described for both humans and computers, such that a **machine is capable of deciding**: if a digital resource **should be reused**; if a digital resource **can be reused**, and under what conditions; and **who to credit** if it is reused."

- R1.1: Clear and accessible data usage license
- R1.2: Detailed provenance
- R1.3: Domain relevant community standards

# References and Links

Wilkinson et al. 2016. Nature-Springer Scientific Data. *The FAIR Guiding Principles for Scientific Data Management and Stewardship*. doi:  
10.1038/sdata.2016.18

A. Jacobsen, et al. 2020. FAIR principles: Interpretations and implementation considerations. *Data Intelligence* 2(2020), 10-29. doi:  
10.1162/dint\_r\_00024

Keynote European Dataverse Community Meeting: FAIR Principles and Beyond:  
Implementation in Dataverse  
<https://scholar.harvard.edu/mercecrosas/presentations/fair-principles-and-beyond-implementation-dataverse>  
<https://scholar.harvard.edu/mercecrosas> @mercecrosas