

Making public data FAIR

Last Updated November 26, 2018

FAIR data principles have evolved from a collective effort of stakeholders seeking to make data Findable, Accessible, Interoperable, and Reusable. The Internet of Water supports the adoption of these principles.

The Future of Research Communication and e-Scholarship (FORCE11) are working towards facilitating knowledge creation and sharing. They convened a 2014 workshop in the Netherlands that found that all research objects should be **Findable**, **Accessible**, **Interoperable**, and **Reusable** (FAIR). These [principles](#) were further expounded upon.



Data and metadata are easily discoverable (even if not accessible) by both humans and computers. Machine readable metadata are essential for automatic discovery of relevant datasets and services. Metadata are data about the data. Characteristics of findable data include:

- ◆ Data (and metadata) have a *globally* unique and *eternally* persistent identifier.
- ◆ Data are described *with* rich metadata.
- ◆ Data (and metadata) are *registered* or *indexed* in a hub (searchable resource).
- ◆ Metadata specify the data identifier.



Data are accessible to those who are given permission to use the data. Limitations on the use of data, and protocols for querying or copying data, are made explicit to both humans and machines. Characteristics of accessible data include:

- ◆ Data (and metadata) are *retrievable by their identifier* and use a *standardized communications protocol*. The protocol must be open, free, and universally implementable. It must also allow for authentication and authorization when required.
- ◆ Metadata are accessible, *even when the data are not available*.



I nteroperable

Enables systems and services to create, exchange, and consume data with clear, shared expectations for the contents, context, and meaning of that data. This makes it clear how data relate to one another. This is perhaps the most challenging component of FAIR.

- ◆ Data (and metadata) use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- ◆ Data (and metadata) use *vocabularies* that are documented and standardized.
- ◆ Data (and metadata) include *qualified references* to other data (and metadata).



R eusable

Data and metadata are sufficiently well described for both humans and computers that the data (and the process of converting data to information) can be replicated or combined with future research.

- ◆ Data (and metadata) are richly described with a plurality of accurate and relevant attributes.
- ◆ Data (and metadata) are released with a clear and accessible data usage license.
- ◆ Data (and metadata) are associated with detailed provenance (meaning there is a clear origin story / history of the data).
- ◆ Data (and metadata) meet domain-relevant community standards.

For more information:

[FAIR Principles Explained](#)

Wilkinson, et al. 2016. [The FAIR Guiding Principles for scientific data management and stewardship](#).