

Everything you need to know
about FAIR Data

WHAT IS FAIR DATA?

The increasing production and availability of online resources requires data to be created with longevity in mind. By providing the wider research community with access to scientific data, knowledge discovery is facilitated and research transparency is improved. In light of this, in 2016, 'The FAIR Guiding Principles for scientific data management and stewardship' were published.

FAIR is a set of principles, not a standard, according to which FAIR research data are the data which can be Findable, Accessible, Interoperable, and Reusable, as given overleaf.

Besides supporting data reuse by individual scholars, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data. The Principles highlight the fact that within the current digital ecosystem, humans increasingly depend on computational support so as to be able to deal with the increasing volume, complexity, and creation speed of data. Within this framework, machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimum human intervention) that FAIR Principles bring forth, is key.



A common misconception among researchers is that FAIR data equals Open data. No, FAIR data does not have to be Open, as they can be shared under restrictions and still be FAIR.

HOW TO MAKE YOUR DATA FAIR

Turning FAIR principles into practice can vary for different disciplines, however the following guidelines can generally apply:

- Make your data findable by ensuring it has a persistent identifier, rich metadata, and is searchable and discoverable online.
- Make your data accessible by ensuring it can be retrieved online, using standardised protocols and by putting restrictions in place if necessary.
- Make your data interoperable by using common formats and standards, and by making use of controlled vocabularies.
- Make your data reusable by ensuring it is well-documented, and by having clear machine-readable licence and provenance information on how the data was formed.
- Use new innovative research approaches and tools

Findable

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

Accessible

Once the users find the required data, they need to know how these can be accessed, possibly including authentication and authorisation.

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.

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.



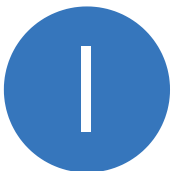
Basic steps to be followed towards Findability

- Metadata are assigned a globally unique and persistent identifier (PID)
- Data are described with rich metadata
- Metadata clearly and explicitly include the identifier of the data they describe
- Metadata are registered or indexed in a searchable resource



Basic steps to be followed towards Accessibility

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



Basic steps to be followed towards Interoperability

- Metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- Metadata use vocabularies that follow FAIR principles
- Metadata include qualified references to other metadata



Basic steps to be followed towards Reusability

- Metadata are richly described with a plurality of accurate and relevant attributes
- Metadata are released with a clear and accessible data usage license
- Metadata are associated with detailed provenance
- Metadata meet domain-relevant community standards

WHY MAKE YOUR DATA FAIR?

Making research data FAIR has manifold benefits for researchers, research communities, research infrastructure facilities and research organisations. FAIR data:

- help to gain maximum potential from data, and overall maximum impact from research increase visibility and citations
- improve the reproducibility and reliability of research
- help in staying aligned with international standards and approaches
- engage in new partnerships with researchers, business, policy and broader communities
- enable new research questions to be answered
- use new innovative research approaches and tools

EOSC AND FAIR

The European Open Science Cloud (EOSC) will strive to ensure that European scientists will be able to take advantage of the full benefits of data-driven science, by offering an environment with free, open services for data storage, management, analysis and re-use across disciplines.

It will provide a platform for European research, including a web of FAIR research data and services.

HOW FAIR ARE YOUR DATA?

Researchers in their effort to determine how FAIR their research data are and how to enhance FAIRness, they can consult already produced tools, such as:

- “How FAIR are your data” checklist (<http://doi.org/10.5281/zenodo.1065991>), or
- “A design framework and exemplar metrics for FAIRness” (<https://www.biorxiv.org/content/10.1101/225490v3>)

RESOURCES

- Wilkinson, M., et al. (2016). “The FAIR Guiding Principles for scientific data management and stewardship”, Scientific Data 3.
<https://doi.org/10.1038/sdata.2016.18>
- <https://www.openaire.eu>
- <https://www.eosc-portal.eu/>
- <https://www.fosteropenscience.eu/>
- <https://www.go-fair.org/>



www.ni4os.eu



NI4OS-Europe has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 857645