

---

# Open Research Data, FAIR Data

---

Besides open access to scientific publications, another widely discussed topic is **open access to research data** (open research data, ORD). Many research funders, such as the European Commission, acknowledge the importance of opening research data and mandate sharing data as part of the grant agreements (for example, Horizon 2020 Open Research Data Pilot)

Open research data are such data that are **freely available online to anyone and can be used, modified, and shared for any purpose**.

Contrary to popular belief, opening research data does not necessarily mean all your data should be freely accessible. In some cases, sharing data is not possible at all (e.g., [personal data protection](#), copyright infringement). The general idea is that research data should be **as open as possible but as closed as necessary**.

## Levels of openness

**Open data:** Open data are those that anyone can access, use, modify and share for any purpose

**Shared data:** Shared data may be widely accessible like Open data but there may be some conditions, such as non-commercial reuse. Sometimes, Shared data are only accessible to specific groups of users, e.g., within an institution.

**Closed data:** Some types of data, such as personal information or commercially sensitive data, may not be shared at all. Even in such cases, however, a metadata description of the data should be made available.

## Which data to share?

Generally, everything that is needed to replicate your study should be made available. Apart from the data themselves and the associated metadata, it may include software, tools, or documentation such as lab journals or code books. On top of that, you may share any other data that you think others would find useful.

## Who decides whether you can share your research data

While researchers certainly play an important part in deciding whether the research data will be shared or not, it is crucial to realise that they are not the only stakeholders in the decision-making process. The following may also be involved:

**Research collaborators:** If you work with other researchers, you need to ensure that they agree to share the research data. This should be clarified early on in the research project to avoid future conflicts and should be included in your Data Management Plan.

**Research participants:** If your research involves human participants, you need to ensure that you obtain their informed consent. The consent should include information about any plans you have for using, storing and sharing their data

**Research funders:** Some research funders may mandate that you open your research data or provide an explanation why data cannot be shared.

## Why should you share your data?

There are many reasons why research data should be as open as possible - it benefits you, as well as the wider research community.

- Opening your research data boosts the robustness of your research as it allows others to replicate the results
- Enhancing your reputation as an honest and careful researcher and thus increasing your impact
- Increased citation rates - your data can be cited as well! Moreover, [research suggests](#) that sharing your research data may increase the citation rate of the associated publication
- Compliance with the funder's or the publisher's policy
- Effective use of resources - the same data do not have to be recreated
- Sharing research data may speed up the research process
- Combining research data from multiple sources may lead to new findings
- Encouraging citizen science
- Reducing academic fraud

## How to share data?

Just like publications, data sets can be shared via **digital repositories** or **data journals**.

To find a suitable data repository for sharing your own data or searching for data you could reuse, you can use an international register [re3data.org](https://re3data.org) (Registry of Research Data Repositories). If you cannot find a suitable subject specific data repository, you can use a general-purpose repository [Zenodo](https://zenodo.org), which is developed by the European project OpenAIRE and operated by CERN. Before sharing, you should license your data so that others know what they can and cannot do with your work.

You can find more about repositories and data journals in the section [Repositories and data journals](#).

## FAIR data



In order to enhance the reusability of your research data, you should aim to make your data **FAIR**. In 2016, [The FAIR Guiding Principles for scientific data management and stewardship](#) were published in Scientific Data with the aim of optimising data sharing and reuse by humans and machines. The FAIR principles describe how data should be organised so they can be more **Findable**, **Accessible**, **Interoperable** and **Reusable**, and they are promoted by some major funding bodies such as the [European Commission](#).

The authors formulated [15 principles of FAIR data](#) :

### 1. To be Findable:

If you want to make your data reusable, the first step is ensuring that both humans and machines can find them - making the **metadata machine-readable** is key.

F1.	(meta)data are assigned a globally unique and eternally persistent identifier.
F2.	data are described with rich metadata.

F3.	(meta)data are registered or indexed in a searchable resource.
F4.	metadata specify the data identifier.

## 2. To be Accessible:

Your data should be **freely available**, ideally via a repository. Even if the access to the data is restricted, metadata should be open.

A1.	(meta)data are retrievable by their identifier using a standardized communications protocol.
A1.1.	the protocol is open, free, and universally implementable.
A1.2.	the protocol allows for an authentication and authorization procedure, where necessary.
A2.	metadata are accessible, even when the data are no longer available.

## 3. To be Interoperable:

In order to allow for your data to be integrated with other data, you should use **standardised vocabulary** to describe the data.

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.

## 4. To be Reusable:

The ultimate goal of the FAIR principles is to enhance reusability of research data. In order to achieve this, it is important that the data are sufficiently described and shared under the least restrictive license, so that users know how the data were generated, what they describe and how they can be used.

R1.	meta(data) have 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 their provenance.
R1.3.	(meta)data meet domain-relevant community standards.

## FAIR self-assessment

To assess the 'FAIRness' of your data, you can use the [FAIR self-assessment tool](#) developed by the Australian ANDS-Nectar-RDS initiative, or you can use this [checklist](#) which was created for the EUDAT summer school by Sarah Jones and Marjan Grootveld.

## Useful Resources

[GO FAIR](#) initiative website

OpenAIRE Guide for Researchers: [How to make your data FAIR](#)

OpenAIRE [Personal data and the Open Research Data Pilot](#) factsheet

ANDS, Nectar, RDS: [FAIR self-assessment tool](#)

Jones, Sarah, & Grootveld, Marjan. (2017, November). How FAIR are your data?. Zenodo. <http://doi.org/10.5281/zenodo.3405141>

Open Data Institute: [Open Data Essentials](#) - e-learning course

Drachen, T.M., Ellegaard, O., Larsen, A.V. and Dorch, S.B.F., 2016. Sharing data increases citations. LIBER Quarterly, 26(2), pp.67–82. <http://doi.org/10.18352/lq.10149>

Piwovar HA, Day RS, Fridsma DB. 2007. Sharing Detailed Research Data Is Associated with Increased Citation Rate. PLoS ONE 2(3): e308. <https://doi.org/10.1371/journal.pone.0000308>