
Where to find existing data

The easiest way to find already existing data relevant to your research is by looking at **underlying data published alongside research articles**. As an increasing number of academic journals require that underlying data are shared, you can find links to relevant datasets as part of published articles (for example, as link shared in a Data availability statement or as Supplementary material). While citing datasets in publications is not yet very common, it is becoming increasingly frequent and the number of articles citing underlying data is expected to rise.

A reliable source of data are **research data repositories**. These can be either general (such as [Zenodo](#) , [Figshare](#) , [Dryad](#) , or platforms like [OSF](#)) or subject specific (you can use the [Registry of Research Data Repositories](#) to find subject specific repositories).

Another reliable source of existing research data are **data journals**, which publish peer-reviewed papers that describe published datasets and so ensure that the datasets are well described and of high quality.

In addition to these sources, you may also use dataset search engines such as [Mendeley Data](#) or [Google Dataset Search](#) .

In your research, you can also use **public administration data** published as open data which you can find in the [National Open Data Catalogue](#) (NKOD).

When you find any data related to the issue that you are researching, **make sure that they are really suitable for your research** and that you can use them.

Make sure that:

- the data come from a **trusted source** (e.g., a certified repository, a well-known author, peer-reviewed data journal)
- the data are sufficiently **described** and include **context** in which they were collected or generated (e.g., who were the participants, what were the conditions etc.)
- the data are in a **format** that you can use (is there any specialised software that you need to work with the data?)
- you are **allowed to reuse** the data and under what conditions, e.g., does the license specify the reuse conditions? Do you have the author's permission?
- the data will be **preserved** (is there a risk the data will be deleted or lost, e.g., if they are published on the author's website? Can you make a copy?)