



Depositing and repositories

Research data repositories are based on sharing and openness. Rigorous data management is required to make sure data are accessible and reusable over time.

How to choose a data repository

Follow the recommendations

In disciplines that are structured for data sharing like astronomy or genomics, data producers can access **repositories specific to their own discipline**.

The best guide is the practices in your own scientific community but **directories of repositories** do exist. The [re3data](#) directory lists repositories at the international level and enables you to **filter results according to a number of criteria** (discipline, allocation of a persistent identifier, types of data or formats accepted, long-term data preservation, choice of a licence, certification, etc.).

Alternatively, data producers can use the **repository of the institution** which they are affiliated to.

Examples

The genomics community uses [UniProt](#) and [GenBank](#) as reference databases.

Humanities and social sciences researchers use [Nakala](#) to share, publish and enhance all types of documented digital data (text files, sounds, images, videos, 3D objects, etc.). It is a secure repository that complies with the FAIR principles.

If you find no disciplinary or institutional repository that responds to your requirements, **you are recommended to deposit your data in the national multi-disciplinary [Recherche Data Gouv](#) repository**. It is important to choose this repository for evidence purposes if your data supports a scientific publication. It offers permanent and reliable access to the version of the dataset used to obtain the results.

Did you know?

The **Recherche Data Gouv** repository is organised into institutional areas for publishing and reporting data from institutions wishing to participate. It was developed on the basis of the Dataverse open source web application and also offers a generic space where researchers who do not have a dedicated institutional space can deposit their datasets.

How to submit data

Preparing and checking data

For data sharing to be effective, it needs to be prepared in accordance with the FAIR principles if it is to be totally or partially opened, secure, etc.



Filling in the form

When everything is ready, log on to the repository website, fill in the deposit form and upload your files.

Tip: Don't just fill in the mandatory fields! Filling in additional fields will make your data more findable, interoperable and reusable.

Examples

- Personal data should be totally or partially anonymised,
- An appropriate distribution licence should be selected for your data,
- The metadata and source codes to be associated with your data should be prepared, etc.

Focus on where and how to deposit in a repository

Selecting a repository from a directory by using criteria to filter

What type of repository do you prefer? Disciplinary?
Multi-disciplinary?

Is the repository certified and internationally recognised?

What types of data does the repository accept?

Which formats are accepted?

Does the repository impose a metadata standard?

Is a licence required? Is it the right one for you?

Does it assign a persistent identifier to the data?

Is the platform easy to use and is site help available?

Will you be able to modify the data you deposited and access older versions?

Is the repository located in France?

Can the repository restrict data access or apply an embargo?

Is the repository committed to preserving long-term data access?

Do you have to pay to deposit data?

Does the repository offer a link between the data and its associated publications?

Are usage, consultation and download statistics available?

You can find answers to these questions on repository policy pages or by running a search with filters on a repository directory.

✓ Checking data before submitting

- ✓ The datasets to be shared have been correctly selected
- ✓ Ethical principles are respected
- ✓ Dissemination rights have been checked
- ✓ The terms of access have been defined
- ✓ Files are organised and have been explicitly named
- ✓ Files are in permanent, open formats
- ✓ The size of your files does not exceed the maximum
- ✓ The data are described and documented using standards where possible
- ✓ A unique and persistent identifier is assigned to the data and source codes
- ✓ The data are assigned a licence