

---

# Tips for sharing data

---

## Make sure that you can actually share your data

Not all data can be openly shared, so make sure that you really can share your data. Many factors affect whether or not you can share your data. If you work in a larger research group, all **co-authors must agree** to data sharing. If you use **third-party data**, you need their **author's permission** to share. This permission can be indicated through an assigned license, or you can negotiate the terms of reuse and sharing directly with the author of the dataset. If your data include **personal information**, you need to either **remove** the information from the dataset or **anonymise** it, or you need to have a **consent to share** it. Data associated with **national security** typically **cannot be shared**. In some cases, it might not even be possible to share metadata records about their existence.

## Share documentation along with the data

You should always share documentation along with your data (e.g., in the form of a Read-me file) to help users understand your data and to reduce the risk of misinterpretation. The form of the documentation will depend on the type of data that you are sharing; it could be a list of variables and their description, aggregated demographic data of research participants, comments included in a computer script and so on. When compiling the documentation for your data, think of your future self – what information would you need to interpret your data if you were to revisit them in, say, ten years?

## License your data

When sharing research data, it is recommended to assign an appropriate [license](#) to them, so that users will know what they can and cannot do with your data. The most commonly used for these purposes are the [Creative Commons](#) licenses, which enable authors to grant some rights to potential re-users while reserving others. You can also apply a custom license to your data that better suits your needs. It is also possible to share the data under multiple licenses, such as free for non-commercial use and available for commercial use at a fee. If your dataset is not subject to copyright protection, it is recommended to label it with a [Public Domain Mark](#), so that the users know that the data can be reused without further restrictions.

## Use persistent identifiers

To make your data more findable (according to the [FAIR principles](#) principles), they should be assigned a persistent identifier like DOI or handle. Typically, authors cannot assign a persistent identifier to the data themselves, this requires a service authorised to assign identifiers. For datasets, some repositories offer this service, so when selecting a suitable repository, make sure that it provides a persistent identifier to your dataset.

## Provide a recommended citation format

To facilitate the [citation of your data](#), include a recommended citation format alongside the data. This serves two purposes: it reminds users that proper citation is needed when using your data, and it makes it easier for them to cite the data. Some repositories may include a recommended citation format on the landing page of your dataset based on the metadata that you provided. Alternatively, you can include the recommended citation format in the documentation.