Research Data

What are research data

Research data can be characterised as any information that has been collected, observed, generated, or created to validate or reproduce your research findings. Research data can take various forms and may be digital as well as non-digital. Some examples of research data may include:

- Spreadsheets, documents
- Audio and video recordings
- Images, photographs
- Questionnaires, test responses, interview transcripts
- Code, software
- Laboratory notebooks, field notebooks, diaries
- Samples, specimens, artefacts

Research Data Management (RDM)

Research data management refers to the activity of organising, storing and preserving the data generated during the research project. Even though managing data effectively may be challenging, there are many benefits in it not only for you but for the wider community, as well. Here is a list of some examples:

- Demonstration of research integrity, enhancing your reputation of an honest and careful researcher, subsequently leading to greater impact
- Helps your research to be robust and replicable
- Helps you anticipate potential issues that may occur during the research process
- Makes writing and revising papers easier
- Helps you (and others) to find your data
- Reduces the risk of article retraction due to mixing up or mislabelling the data
- Reduces the risk of data loss
- If there is a problem with your paper, you will be in a good position to defend yourself, or at least to prove that you reported your results in good faith
- Ensuring continuity in long-term projects (where you have multiple postdocs coming and going) and consistency in projects with multiple researchers involved
- Ensures your research meets the requirements set out by research funders and publishers
- Advancing research worldwide through reuse of data

Data management in your research project
1. Before your project
You should address the issue of data already before your project begins. While **planning your research project,** you should consider what kind of data you will need, how you plan to acquire the data (will you create your own, or can you use existing data?), where you will store them, who will be responsible for them and so on. You can include the information you collect in the first version of your Data Management Plan (DMP).

2. During your project
During your research project, it is important to ensure that your data are well-organised and stored securely. You should describe your data accurately and carefully (e.g., how the data were generated, what the values mean, mechanisms for version control), and mind where you store the data and the backup and whether the storage is secure, especially if you work with sensitive data.

3. At the end of your project
At the end of your project, you should consider what happens to the data when the research project is over. Think about which data can be safely deleted and which data need to be preserved (you can use this guide to help you decide), and consider sharing your data. If you decide to share your data, you should make them FAIR, and you should also keep personal data protection in mind and anonymise the data, if necessary. To anonymise your data, you can use the tool Amnesia , for example, which is available on the OpenAIRE website. If you publish your data openly, it is recommended that you license your work, so that others know what they can and cannot do with the data. Whether or not you decide to publish your data, consider depositing them in a repository to ensure long-term preservation.

Useful Resources

CESSDA: Data Management Expert Guide (aimed at social science researches)

Markowetz, F. 2015. Five selfish reasons to work reproducibly . Genome Biol 16(274). https://doi.org/10.1186/s13059-015-0850-7

Wilkinson et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship . Scientific Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

Open AIRE: A research data management handbook

MANTRA: Research Data Management training . Edinburgh: The University of Edinburgh