

# A Guide to Research Data Management

Find out how to manage your research data, from organisation and storage to security and sharing.

## What is research data?

Research data is evidence used to support research conclusions. It can take many forms, such as annotations, spreadsheets and databases, lab notebooks, text corpora, measurements, statistics, survey results, images, video and audio, transcripts, measurements, models, and software.

## What is research data management (RDM)?

RDM is the process of organising research data, keeping it safe and secure, and possibly preparing it for sharing and preservation. Not all data can be shared, but all data should be managed well. It's a necessary part of any research project, and should be considered from the outset.

**Research Data in the Humanities:** Research data management isn't just important in the natural and social sciences. Humanities researchers create data too. These may include lists, tables, matrices, databases, maps, digital objects, images, and recordings and transcripts. Generally this guide will refer to digital data but some aspects of RDM also apply to non-digital data.

## Why manage your data?

- To make the most of your investment. Data is the raw material on which research is based, and it takes resources to generate, collect and collate it; good research management makes the most of that investment
- To inspire new research and collaboration. Data that is well managed can be used and reused more easily by you and others, now and in the future
- To comply with statutory requirements and funding conditions. Research organisations and research funders now issue guidelines or policies around research data, so it's worth checking what policies apply to you before starting your research.

You can improve the visibility of your research by registering data in a repository, maybe making it open in the process. When others can access your data and gain a deeper understanding of how it supports your arguments, trust in your research process is improved. Open data is a valuable resource for future research and teaching, contributing to the validation of new analysis techniques and wider studies combining multiple sources.

## Planning your data management

- As you plan your research, document how you intend to handle your data and write it up as a data management plan (DMP). Some research funders may require a formal version
- Use a DMP template, such as those provided through [DMPOnline](#), to make sure you cover all the issues across the research lifecycle
- Prepare your plan early so you consider issues like data sharing agreements and consent forms to prevent issues with sharing data later on
- If you and your colleagues find yourself writing similar DMPs for different projects repeatedly, consider a data management policy for your whole programme of research to save time.

## Organise your data

- Develop a system of organising digital and physical information so that you can easily find what you need
- Use clear folder structures and consistent file naming conventions, such as clearly labelling versions
- Write a brief 'read me' file to document your conventions, especially when working with other researchers.

## Storing your data securely

- Consider the risks of loss, corruption and unauthorised access, especially if your data is sensitive or irreplaceable
- Understand the legal frameworks governing the collection and use of information, such as data protection, copyright and intellectual property. The [UK Data Service](#) provides an overview of the key legal and ethical issues
- Plan how often to back up data, what kind of physical storage media to use, and whether to use local or cloud-based storage
- Use the 3-2-1 system for important data: keep three copies, use at least two different types of storage, keep at least one securely offsite.

## Data sharing

- Data is a valuable research output, like articles and monographs. Recognising this, many organisations and research funders have specific requirements for sharing data
- Find a robust, long term archive to store and share your data when you've finished with it. Talk to other researchers in your field about what they use, or check out the [re3data](#) database

of archives and repositories. Some institutions have their own repository, while there are many subject specific and general use repositories, such as [Zenodo](#) and [Figshare](#). These help others discover and reuse your data.

**Data repository:** Some open access publication repositories also function as repositories for datasets and other kinds of research outputs. It's possible to create a record of your data in a repository, and you may be able to share your data too. Datasets are commonly made available under Creative Commons licenses to enable reuse by other researchers in the future. See our guide on [Copyright and Creative Commons](#) to learn more.

- Many repositories work with the organisation DataCite to assign DOIs (Digital Object Identifiers) to research datasets, which means they can be more easily discovered, cited, and linked to
- Include documentation and metadata (information which describes your data) when archiving and sharing data, to ensure that you and others can accurately interpret it in future. This helps to make your data FAIR (Findable, Accessible, Interoperable and Reusable)

**Metadata** is 'data about data'. Good metadata helps others more easily find and reuse your data. Descriptive metadata should be created according to a common schema, so that the context and purpose of the research can be understood. Different research communities have created schema most appropriate to their needs. Choose one as early in the process as possible. The Digital Curation Centre (DCC) provides advice about [discipline-specific metadata](#).

- Not all data is suitable for sharing, as it may include sensitive personal data or protected intellectual property; if in doubt, aim for 'as open as possible but as closed as necessary'
- If you can't share some or all of your data, document clearly the reasons for this and consider other routes through which others may be able to access it, such as non-disclosure agreements.

## Further Resources

### [All European Academies](#)

[\(ALLEA\)](#): published a report in 2020 with recommendations for FAIR data in the humanities.

[DataCite](#): offers a search interface to find all kinds of registered research outputs, including unpublished reports and datasets.

[Digital Curation Centre](#): offers examples of DMPs, including those from researchers in the arts and humanities.

[Mantra](#): a short online course from the University of Edinburgh for all those who deal with digital data as part of a research project.

[Research Data Management and Sharing \(Coursera\)](#): an in-depth online course about research data management, run by the University of Edinburgh and The University of North Carolina at Chapel Hill.

[The Turing Way](#): provides a series of resources for ensuring data is easy to reproduce and reuse at the end of a research project.

## The British Library

Scholarly Communications Toolkit

Last updated: June 2021

DOI: 10.23636/0pbq-2r89

