Sharing data and other research outputs: FAQs

In this piece we answer the top 10 data sharing questions that researchers ask at the beginning of their data sharing journey.

1   What is a research output and why should I share my research outputs?

Research outputs include datasets, software, code, materials and intellectual property (IP). Sharing research outputs offers the potential to accelerate the pace of discoveries and therefore the benefits to society. Transforming human health will take longer if research outputs aren't managed, shared and used in ways that realise their full value. Routinely making your research outputs open is also an important part of improving the overall rigour and reproducibility of research.

2   What is an outputs management plan and why are they important?

An outputs management plan (OMP) sets out your approach for maximising the value of the research outputs that your grant will generate. An OMP should outline: the outputs that the planned research will produce, the metadata that will accompany the outputs, when and where the outputs will be available, and how these will be made discoverable to the research community. An OMP should be written at the application stage, to ensure that you know how your outputs will be managed from the beginning of your grant and to ensure that your research outputs are as open and FAIR (Finable, Accessible, Interoperable and Reusable) as possible. Writing an OMP should also help you to consider the costs required, to make your research outputs open, so you can build these into your funding proposals. OMPs should be updated throughout the duration of your grant, taking into account changes in your research. More information about OMPs can be found **here**.

3   What are Wellcome's outputs sharing requirements?

Research data, software and materials should be made available with as few restrictions as possible. As a minimum, the data underpinning your research papers should be made available to other researchers at the time of publication, as well as any original software that is required to view datasets or to replicate analyses. Where research data relates to public health emergencies, you must share quality-assured interim and final data as rapidly and widely as possible, and in advance of journal publication.

4   I am using a spreadsheet to store my data. How do I prepare spreadsheets for sharing?

By following best practices when using spreadsheets, you can ensure your data is interoperable and reusable for both humans and machines in the future. **Here** is a useful guide which outlines the best practise when sharing spreadsheets.

5   What is metadata and why is it necessary?

Metadata is the underlying information associated with research outputs, which helps others to find, understand and reuse the research output. Metadata can take many different forms, from free text to standardized, structured, machine-readable, extensible content. When possible, you should structure your metadata using an appropriate, agreed-upon metadata standard format. To find an appropriate metadata standard for your discipline, refer to the **Disciplinary Metadata guide**. When no appropriate metadata standard exists, you may consider composing a "readme" style metadata document.

Wellcome Open Research

6   What is a Digital Object Identifier (DOI) or persistent identifier (PID) and how do I assign one to my research output?

A persistent identifier acts as a permanent address and remains constant even if the location of the item changes. Digital Object Identifiers are one example of a persistent identifier and are commonly used when publishing data and journal articles. Despite the name, items do not need to be digital in order to obtain a DOI. They can be physical e.g. a book. You can obtain a DOI for your research output by publishing in a journal or depositing in a repository that issues DOIs. All DOIs start with the number 10 followed by a series of numbers, letters and symbols e.g. https://doi.org/10.12688/wellcomeopenres.15603.1. You can easily search for any research output by pasting its DOI into your search engine.

7   What is a repository and how do I know which one to use?

A repository is an online archive for the storage of digital objects. Open access repositories use open standards to ensure that the content they contain is accessible, in that it can be searched and retrieved for later use. Content that you can add to a research repository include research papers, reports, datasets, software and other objects resulting from research.

You should deposit data in recognised data repositories for particular data types where they exist, unless there is a clear reason not to do so. Wellcome Open Research maintains a **curated list of approved repositories** suitable for Wellcome-funded research. Choosing the best repository can often be daunting. Here is a **guide** to help you choose which repository to use.

8   What are licences and which one should I use for my data?

Data reusability is defined by the presence of a user licence. The licence you select will determine the freedom with which others can reuse your data. When choosing a licence, it is important that you adhere to any repository, institutional, legal or ethical obligations. **Here** is a useful guide which explains more about licences.

9   How do I share sensitive data?

How you share your data first and foremost depends on any applicable legal and ethical factors, this is especially true where the data collected is personal. You may feel that there is tension between protecting research participants and openly sharing your data, but datasets that contain personal data can often be shared following informed consent, appropriate anonymisation and/or controlled access. Read through this guide for more details on each of these aspects.

10   What is a data availability statement?

A data availability statement provides information about where the data supporting a publication is stored, and the conditions under which it can be accessed. This means that the data supporting publications is more findable and discoverable and that others can use this data to seek to replicate or build on the research.

Wellcome Open Research