

DATA MANAGEMENT PLANS



**OFFICE OF SCHOLARLY
COMMUNICATION**

Cambridge University Libraries

What is a data management plan?



Data management plans (DMPs) are simply proposals for how a researcher intends to manage their data both over the course of their project and beyond. Different research funders and institutions will have their own specifications on what should be included but it commonly includes an outline of how the data will be collected and stored during the project and how the researcher plans to make the data available after it ends.

It is important to remember that the definition of research data is quite broad – it is any information used or produced during the research process. This can include a wide range of materials such as statistics, interview transcripts, images, physical samples and field notes.

Why are they needed?

Data management plans are becoming increasingly common in the world of research. Many funders now require a DMP prior to releasing any grant to a researcher as they want to see evidence that the proper procedures have been considered. Institutions are following suit and many have introduced mandatory data plans for research students.



Beyond this, DMPs are part of good academic practice. By planning data management at the start of a project the researcher can work more efficiently as they have already thought about many of the common problems they might encounter. The DMP is also a living document which can be revisited as needed throughout the research process. The increased emphasis on the sharing of data upon completion of a project is another important factor.

Researchers from all disciplines are now encouraged to share the data underpinning their publications and a good DMP can help make this process easier.

WHAT TO INCLUDE

The following list includes areas common to many DMPs. Remember that each funder will have its own requirements and it is always best to check these before preparing a plan.

DATA SOURCES



Researchers should consider the source of the data they will use in their project. It is likely that this data will come from various sources and may include data previously gathered by other people. Outlining how new data will be collected is also important as the person reading the DMP needs to understand how all the data used was gathered together. Researchers should be prepared to discuss the reasons for either reusing existing or creating new data.

TYPES OF DATA



A large part of the DMP will describe the type(s) of data which are being used in the research. This information will need to include details of what form the data will take, file formats, the metadata used to describe the data and why these have been chosen. At this point the researcher may have to move outside their comfort zone and think about the most appropriate format(s) for their data rather than just what they are used to using. The type of data collected will have an influence on all other sections of the plan so it is important to consider this carefully.

BACKUP STRATEGIES



Once data has been collected it needs to be kept safe, both for the current project and for future preservation. Researchers will need to outline the methods they will use to backup their data and why these are the right choice for the project in question. Physically storing data both during the project and in the long term can cost money, even if the data is digital and researchers will need to think about how this will be paid.

WHAT TO INCLUDE



DATA SHARING

The increased emphasis on sharing data is another consideration for researchers. They need to think about who needs to have access to the data during the life of the project and how this will be managed. In the case of sensitive data protection will need to be ensured. Researchers also need to consider the longer-term sharing of data and how this can be facilitated, including selecting the best repository for storage and using accessible file formats to enable people to access the data.



ETHICAL AND LEGAL ISSUES

Dealing with any type of data will require considerations around privacy but this becomes even more crucial when working with information which might be personal or sensitive. Researchers will need to think about how any ethical or legal issues should be managed including where data is stored, how it is shared both between the team and with the wider world and any copyright issues. The definition of sensitive data is wide and varies from country to country so researchers should check local practice.



RESPONSIBILITY FOR DATA

Although some researchers work alone many work in larger teams. Individual researchers may also be mobile and move on long before the completion of a project. DMPs should outline who has ultimate responsibility for the data produced by a project including who will take responsibility for backing it up, who can be contacted with questions and who will take long-term responsibility after the completion of a project.

FURTHER INFORMATION

Visit our webpages: <https://osc.cam.ac.uk/>

Read further guidance on data management plans here:
http://bit.ly/OSC_DMPs

