

Planning for Data

A plan detailing how you'll manage your data, code, and other research materials (including documentation, code, and physical samples) over the course of a project will help your research proceed efficiently. Creating a comprehensive, specific, and instructive plan for your data is an important step in developing a new research project, but the best plans also evolve as a project proceeds.

What does it mean to plan for data?

Planning for data means thinking through and documenting how data and other materials will be organized, saved, prepared, analyzed, and shared over the course of your research project.

Requirements and how to meet them

Many funding agencies and institutions now require that researchers compose a short document called a Data Management Plan (DMP). A DMP provides details about the type of data to be collected and managed within a research project. It also documents the individuals responsible for managing the data, how and where data will be archived and shared, and how the financial cost of managing data will be met.

The DMPTool (<https://dmptool.org/>) is a free tool that provides guidance for creating a data management plan.

Things to think about

- Planning for data is not a one-time activity. You should create a plan as you develop a project, but you should also revisit and revise your plan as your project proceeds.
- Plans should identify the data you intend to collect, as well as how you plan to transform, analyze, and share it. Be as specific as possible.
- A plan is really only useful if people know about it and can follow it. Be sure your plans are communicated to your collaborators.
- Even if you do not have a Data Management Plan (DMP), you may have a document that describes how you plan to handle your data. For example, this information could be included in a study protocol or an IRB proposal.