

# How to create a data management plan (DMP)

A well-structured Data Management Plan (DMP) clarifies **how** and **what** data will be created, processed, and published. A DMP helps you to ask all the right questions concerning data management before starting the actual work. Also, Data Management Plans are increasingly required as a mandatory proposal part by research institutions and funding agencies. Having a high-quality DMP saves you a lot of time and frustration in the long run, and increases the transparency and the integrity of your work.

In this How-to, you learn about preparing a DMP with the [GFBio Data Management Plan Tool -DMPT](#) and about the DMP Support Service GFBio offers to help you create a state-of-the-art DMP.

## Starting up with your DMP

You have just heard about DMPs and now want to write your own? That's great! If you have never considered data management planning before, it might be good to get some general information:

Data management planning is the starting point in the [data life cycle](#), and DMPs are ideally established at this point. But a DMP is not only created at the beginning of a project, it is also a living document, and you should return to it at each step of the cycle. Ideally, the DMP should be revisited often throughout the lifetime of a project to ensure proper data documentation and management. You can think of a DMP as an addition to the data policy of a larger umbrella project.

In the following steps, you will learn about the most important aspects of data management planning and learn to explore the DMPT (you can continue reading the How-To without using the tool in parallel).

The GFBio DMPT contains the most important questions that you should ask when starting with data management planning of a biological or biodiversity-related research project and is based on the [DFG Guidelines on the Handling of Research Data in Biodiversity Research](#). Be aware that it does not provide a ready-to-use DMP that you can just attach to your project proposal. Please use our personal DMP support to contact our experts (you will find a button at the end of the tool).

### Step 1 - Provide general information about your project

It is all about the [planning!](#) You can start to consider general information about your project right at the beginning. Besides the **project name** and the **people involved**, you already give a first impression of the **research data** you will produce.

What is your **main subject**? - In the GFBio DMPT, you can choose between several categories from *Algae & Protists* to *Zoology*. What **type of project** is it? - Perhaps you are doing *Laboratory and/or Field Work*.

The **funding** you are applying for may also influence the content or structure of your DMP. Some funders have special requirements for a DMP as a mandatory part of the proposal (e.g., DFG, Horizon 2020). Many funders, as well as institutions or publishers, also have their own **policies and guidelines for data management**. You do not need to know all the details of these at the beginning of your data management planning, but it is important to know about these requirements and to document them. If you are unsure which policies or guidelines you have to stick to, just let us know, and we can support you in identifying them.

[blocked URL](#)

### Step 2 - Outline the data you will collect and your methods for data collection

Depending on your research questions and methods, you create different **types of data, such as text files, code, GIS, numeric, molecular data, or multimedia files**. In environmental and ecological research, it is common to have multiple data types. Studies and projects are a complex composition of various methods. If you are an experienced researcher in your field, you might already have an idea of the **data formats** (e.g., .xlsx, .csv, .docx, .txt, .shp) you will create. In this case, you can specify your data types, even to the level of data formats. Please document the data types and formats as detailed as possible when submitting your DMP support request to us. The more information we get, the better we can support you in your data management planning.

**Data volume** and the **number of files** you will create are important pieces of information regarding data management. They affect questions concerning backup strategies and techniques, as well as the publication of your research data. **Standards, methodologies, and tools for data collection, data assurance**, and data management may also provide important information for planning your data management. Are you part of a larger umbrella project and have to stick to conventions on standards, methodologies, and software tools? That's fine, this information is very important and should be documented in your DMP.

Working in environmental sciences often implies the collection of **physical objects**, which in some cases have to be archived as well. With GFBio, you can (if you want) submit and archive several kinds of objects related to your research data. By means of some basic questions on them (integrated in the DMPT), we can estimate if your physical objects can be archived in one of the GFBio data centres if you would like it to do so in the future. Your DMP should also contain a notice if you are working with **sequence data**, as it requires special expertise and has to follow specific standards.



### Step 3 - Describe your metadata

Have you already thought about **metadata**?

Metadata is more than the description of your research data. They contain important information on the data content, collection, and processing (who, when, what, where, how, how often, etc.). Only by means of metadata, data portals like the [GFBio Data Search](#) can browse, filter, and navigate through huge research data collections. They also make your data findable. There are certain metadata standards and defined terminologies that enable the integration of research data from different partners. You can, for example, use our [VAT tool](#) and intersect your own data sets with layers from GFBio data centers.

Are you already familiar with metadata standards and know how to handle them? - That's great, list all the standards you will comply with in your DMP.

You have no idea about what metadata is, or you don't know yet what your metadata will look like? - Don't worry, just let us know, and we will find out together with you.

Also, have a look at our data life cycle fact sheet [Describe](#) to learn about metadata and how to handle it.



### Step 4 - Consider ethical and legal aspects

In some cases, data cannot or should not be published immediately when submitting to a data center. One possible reason might be that you collect research data within the scope of your PhD thesis, and you have to publish your studies before making data available for others. In this case, you can define **access restrictions** for your research data. This means you can already submit your data to a public archive and get an identifier for them (e.g., [DOI](#)) - which makes them citable and reusable, even though you don't want your data to be published immediately.

Unlike your data, your metadata does not have any access restrictions. Thus, people will find the information about your research data in a search engine, but they cannot access the data themselves. Since metadata should always contain information about contact persons, people will be able to contact you if they are interested in your research. When the access restriction expires, the data will also become available.

You might also collect **sensitive data**, which is subject to ethical issues or legal restrictions and requirements. Personally identifiable information, Red List Species, and genetic resources must be handled according to specific legal requirements and may not easily be published. If you feel uncertain about legal requirements, we support you in identifying them.

Conscientious data management implies **licensing your data**. GFBio supports the idea of open access to research data. But open access does not mean everyone can use your data at will. Data can (or must) be cited in the same manner as publications. Licenses, such as the [Creative Commons licenses](#), define citation demands as well as further terms of use, for example, whether the data may be transformed.



## Step 5 - Outline the preservation and sharing of your data and metadata

Design a plan for **data backup**. Which technologies in which locations will be used to store your data? Are there regular backup routines, or will you back up manually? Who is responsible for data backup, and which service providers (e.g., local IT support) are involved?

Find a suitable **long-term archive** for your data. An advantage of submitting your research data to GFBio is that we can spread your data over our ten [data centers](#). Each of our data centers has specific fields of specialization on which you can build.

Think of a **submission** plan. When will your data be submitted to (a) long-term data center(s)? Remember, at this point, you already have the final and quality-assured version of your data sets. You can submit data linked to an article at the same time as the article's publication. Alternatively, you can continuously submit data during a project runtime (e.g., in defined time intervals) or submit all datasets at the end of the project. And you can also define any other reasonable submission plan for your project data. There is no one-size-fits-all plan for data submission; it always depends on your research conditions.

Find more information on data preservation in our data life cycle fact sheets [Preserve](#) and [Submit](#).



## Step 6 - Elevating your DMP with support from GFBio

Once you have completed the questionnaire in the DMPT, you can always download your entries as a structured PDF file and request free personal [DMP support](#) from our experts!

You can also register for free, and by logging in, you can save your entries to your personal account and revise them whenever needed.

You can also request free personal DMP support from our experts. If that is the case, we recommend filling in the DMPT as much as possible and then sending a support request (you can do this at the end by completing the questionnaire). We will then contact you to talk about open questions. In your support request, you can also inform us of any questions you have or the specific aspect of the DMP for which you require further information and support.

For general questions, you can [contact us](#) anytime. In the meantime, have a look at the [training](#) we offer and more!



## Verwandte Artikel

- [How to create a data management plan \(DMP\)](#)
- [How long does it take to have a GFBio approved Data Management Plan \(DMP\) for my research proposal?](#)
- [Recommended biological and environmental training datasets](#)
- [NFDI4Biodiversity meets NFDI4Earth 2025 - Call for contributions](#)
- [What data types are suitable for submission?](#)



**Do you have questions, feedback or need help?**

[Contact our Helpdesk](#) for direct support.

