



What is it?

After assuring the quality of your data, a detailed description of the data set should be provided, so other users can easily find your data, understand the context and content, reuse and cite your data. Your published article may not be sufficient to gain this information. The description of data generates structured information, so-called **metadata**, which answer the following questions:

WHY were the data generated? **WHO** created the data? **WHERE** and **WHEN** were the data collected? **WHAT** is the content of the data? **HOW** were the data assessed?

How to do it?

1. To save time, start early with the description of your data – as long as the information is present (Description-Tool soon implemented into [GFBio-Portal](#)).
2. Define how you want your data to be cited.
3. Use appropriate metadata standards where possible to avoid errors and to **make your metadata compatible**. (GFBio-Standardization-Tool soon implemented; Conversion-Tool for transforming your metadata afterwards, e.g. to EML.)
4. The description will include the **technical context** (names of datasets and data files in the datasets, versioning, file format, (processing of data), hardware/software used for data collection. (How?) (Methods, tools, instruments)
5. Describe who was and is involved in the study (collectors, stakeholders, funders, contact person for questions).
6. Describe the **scientific context**: Why were the data collected (hypothesis)? What kind of data were collected? Where were the data collected and when? Which standards or calibrations were used?
7. Describe the parameters: How (precisely) was each parameter generated/measured? What are the units and formats? Which codes and abbreviations were used?
8. Use the GFBio-Terminology-Service (soon implemented) to avoid the using of different names for e.g. the same species. By accounting for synonyms and acronyms, terms are kept consistent and reliable and your data can be annotated.

Who does it?

Ideally everybody who is producing data (e.g. a geneticist).

Key elements:

- Create metadata (Description-Tool soon provided at the [GFBio-Portal](#)).
- Alternatively, convert your metadata afterwards to a compatible standard (EML, ABCD) (Conversion-Tool soon available at the GFBio-Portal).
- Make sure the 6 questions can be answered (Who, How, Where, When, What and Why)

Useful links

<https://www.dataone.org/best-practices> (Data One Best practice)

<http://www.dcc.ac.uk/sites/default/files/documents/metadata%5B1%5D.pdf> (DCC)

http://www.youtube.com/watch?v=7IN_SD5B43U (MANTRA Video with Lynn Jamieson)

<http://www.dcc.ac.uk/resources/metadata-standards> (overview of metadata standards)

<https://www.youtube.com/watch?v=-MIH8PkuUo4&feature=relmfu> (a data file called SAM)