

Handout “Data organisation”

What is it about?

Data organisation encompasses all strategies for structuring data in research data management and is therefore part of structured working. This includes, above all, the creation of meaningful and expandable directory structures, meaningful naming at the file level and the development of and adherence to a versioning concept.



Why is this important?

The more data accumulates over a certain project period, the more important uniform concepts for data organisation and its documentation become in order to ensure structured work.



Good data organisation leads to better traceability and makes it possible to say with certainty, even years later, what was done, how and why. In addition, in collaborative projects with appropriate concepts, joint work on the data is facilitated because everyone follows the same naming and filing strategy. This means that data on the data carriers can be searched for and found more quickly. The use of versioning concepts also allows the current status of an investigation or a particular experiment to be viewed directly and old measurements to remain available as comparative values.

How do I implement this?



For **file naming**, you should follow either the Camel Case (= ExperimentNumberOne) or the Pothole or Snake Case (= experiment_number_one). Date formats should be written in the formats YYYY-MM-DD or YYYYMM-DD to ensure clear sorting from old to new. You also need to determine what information belongs in the file name and how it is abbreviated. Finally, you should create a ReadMe file in the appropriate folder for these conventions.

For **versioning**, you should follow the Major.Minor.Revision versioning concept: experiment_v1_0_0. The first value indicates a major change, the second a minor one and the third value, for example, a correction of a spelling mistake. However, if several identical measurements do not occur on one day, it is often sufficient to use the date as a version indicator. With **directory structures**, you should make sure that the concept is neither too detailed, nor that it is not expandable at all. On the right you can see an example of a directory structure that can be used as a basis.

