File-Naming Design

# 

Consider all your current and anticipated files. The file naming conventions you create may differ based on the needs of the project or subset of files. You can use different conventions for different sets of files.

Use the table below to capture the information on all your different data file types.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Data Type** | **File Format (extension)** | **Features** | **Versions (frequency of the same file naming)** | **File name design (draft)** |
| *E.g., Microscope image* | *E.g., tiff* | *E.g., Date collected; Experiment number; Collector name ;description; Conditions* | *E.g., Y*  *More than 1 per day (~50/day)* | *E.g.,*  *YYYYMMDD\_exp001\_v01.tif* |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Considerations**

**Feature extraction:** The unique characteristics that you’ve identified in the “feature” column can be used to create a file name. You may want to shorten the information to include it in a filename. For example, person name - initials rather than full names.

*Example: Sample ID will use a code made up of: a 2-letter project abbreviation (project 1 = P1, project 2 = P2); a 3-letter species abbreviation (mouse = “MUS”, fruit fly = “DRS”); and 3-digit sample ID (assigned in my lab notebook).*

|  |
| --- |
| *My own notes* |
|  |

**Versions:** It is needed usually when you anticipate multiple versions of the same (most likely identical) file naming. If you anticipate having more than one version in a day, you can’t rely on the date alone and should add a version number to the file name. For example, YYYYMMDD\_exp001\_v01.tif. You may also want to differentiate between different states of data as it goes through your workflow.

*Example: As each image goes through my analysis workflow, I will append the version type to the end of the file name (e.g., “\_raw”, “\_processed”, and “\_composite”).*

|  |
| --- |
| *My own notes* |
|  |

If you’re using versioning software or a table to track versions of files, rather than the file name, be sure to note that in your documentation.

**Ordering:** Put the features in the order that you want your data files to sort by. For example, if you prefer to sort by date and then experiment number, the date should be first followed by the experiment number and so on, e.g., YYYYMMDD\_exp001\_. Use ISO 8601- formatted dates (YYYYMMDD or YYYY-MM-DD) at the beginning of the file names so dates sort chronologically.

Some computer systems cannot handle spaces in file names. To make file names both computer- and human-readable, use dashes (-), underscores (\_), and/or capitalize the first letter of each word (“camelCase”) in the file names.

*Example: I will use underscores to separate the characteristics used in the file name and dashes between parts of my sample ID and date. YYYY-MM-DD\_exp001\_MUS-01*

|  |
| --- |
| *My own notes* |
|  |

**Discussions:**

**Now exchange your design sheet and discuss with your teammates, explain your design concepts.**

**You can (probably should) use this design and include it your DMP (or in a README.txt and keep it with your files).**

|  |
| --- |
| *Discussion notes* |
|  |

# This worksheet is adapted from

Stobbs, R. (2022). *File-Naming and Organization Worksheet* [Teaching Resource]. AUSpace. <https://auspace.athabascau.ca/handle/2149/3684>. Licensed under a Creative Commons Attribution 4.0 International ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/legalcode)) license.