

File-Naming and Organization Worksheet

This worksheet will help you identify and prioritize your file organization needs and develop a system for naming and organizing your files and folders.

Once you have completed the worksheet, create a readme file that can be included in your top-level folder for your team's reference. If you make changes to your naming and organization practices, be sure to update the readme.

- A template for creating a readme is available: [File-naming Convention and Folder Organization Readme Template](#).

Contact library@athabascau.ca with any questions.

Part A. Defining your organizational needs

1. How do you currently organize your files? Consider the file hierarchy on your computer and/or storage device(s). Take a screenshot of your folders window or sketch that out here.
2. Reflect on this current system and how you have recently used it. What works and what doesn't work? What do you like / don't like about your system?
3. Who are you designing your organization system for? You? Collaborators? Your PI? List all the people that will need to access the files in your system. If you know of any organization preferences or needs of this group, note those here too.

Part B. Creating a file-naming schema

1. Create your file inventory

Consider all your current and anticipated files. If it's difficult to think of ALL your files, start with a specific project or subproject. The file naming conventions you create may differ based on the

Last Updated: 2022-09-17

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.

Column A: What different **types of data** are you using/creating? (e.g., microscope image, field observations in tabular form, interview transcripts, figures, protocols, etc.)

Column B: What are the **file formats**? (e.g., .csv, .jpeg)

Column C: What are the **unique characteristics** of this data file (e.g., date created, project name, experimental conditions)? Are there standard abbreviations for any of these characteristics? You may have more than one unique characteristic.

Column D: Will you have **multiple versions** of the same file? If so, will there be multiple versions in a given day? How many?

Column E: Keep blank; we'll tackle that in the next section.

A. Data Type	B. File Format	C. Unique characteristics	D. Multiple versions?	E. Draft file name
<i>E.g., Microscope image</i>	<i>E.g., tiff</i>	<i>E.g., Date collected Experiment number Collector name Image description Microscope setting</i>	<i>E.g., Y More than 1 per day (~50/day)</i>	<i>E.g., YYYYMMDD_exp001_v01.tif</i>

2. Draft your file names.

The following provides guidance on doing so based on the inventory table you create in question 1. For more specifics on file naming best practices, see the [Name Your Files](#) page of the Library Guide. Put your draft file names in Column E of the table.

Unique Characteristics: The unique characteristics that you've identified in Column C can be used to create a file name. Pick the characteristics that you think you will use to find these files. You will need to decide how to shorten the information to include it in a filename; for example, if you are including the name of the person who collected the data, you may want to use initials

rather than full names. Creating two or three letter codes can be useful, especially if you have numerous characteristics to include.

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).

Versioning: If you will have multiple versions of a file (indicated in Column D), make sure to add version information into your file name. 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").

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 unique characteristics 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

Part C. Creating a systematic file folder hierarchy

1. Review your different data types in the table in Part B. Think about how you would group these files together into categories. By file type? By time period? By project? By instrument?

MIT Libraries Data Management Services. (2020). *Naming and Organizing your Files and Folders*. [Teaching Resource].

https://www.dropbox.com/s/xx26a1onsu1qdpc/Worksheet_fileOrg.docx?dl=0 Licensed under a Creative Commons Attribution 4.0 International ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)) Licence.