Define naming conventions for your files

Generate a name draft for each file type in your project:

- Should a versioning strategy be established?
- How should the files be sorted and the elements in a file be organised?
- What elements make each file type unique?

Include important context information in the file names: According to which criteria would you search for a file?

- Creation date?
- Persons in charge of the project or experiment? Collaboration partners?
- Research method?
- File type? (collection, spreadsheet, presentation, manuscript, …)

Which abbreviations could be used to define the key elements of the file?

How do other team members think about these files? Would your colleagues use the same keywords to search for these files?

Naming Convention Table? First aid in file naming!

The creation of a naming convenition table is very helpful in establishing naming conventions for your files.

In such a table you can list the file types that you generate in your project. For each file type, you can define elements that make this file type unique and those that are helpful when searching for it. This handout shows you what to watch out for when deciding on a naming convention for your project.

A table template is provided at the end of this handout and is also available as a separate .xlsx file for reuse.

Author: Ewa E. Bres, Research Data Service Center at the University of Bonn

Naming convention criteria

The names should be **clear**, **self-explanatory** and **consistent** (keep the same order for the same information):

- Alphanumeric sortability (for example according to the date format YYYY-MM-DD) makes searching for a file a lot easier.
- Long file names should be avoided.
  Up to 32 characters is fine: 32CharactersLooksExactlyLikeThis.txt
  (shorter names are easier to find and require a shorter file path).
- Only use a period before the file extension
  (extra periods can be confusing for both machine and human readers).
- Do not use special characters (~ ! @ # $ % ^ & * ( ) ` ; : < > ? . , [ ] { } ' " | ä) and spaces (can be equally confusing for both machine and human readers and some special characters already carry meanings).
- Do not rely solely on lowercase and uppercase letters to differentiate between files: some operating systems treat z and Z differently.
- Include clear references to the content, status, file type, person, project part, survey, source, … in the file names.
- Separate elements in file names correctly.

Naming conventions should be established, **clarified** and **accepted** by all persons involved – **and written down**!

Author: Ewa E. Bres, Research Data Service Center at the University of Bonn

DIY: File naming

How can I separate the elements in a file name correctly?

camelCase
Capitalize the first letter of each element:

2016-01-04AstroInvitroEx1SmithV01.xlsx

camelCase is more suitable for shorter names, otherwise it becomes confusing.

kebab-case
Use hyphens instead of blank spaces:

2016-01-04-astro-invitro-ex1-smith-v01.xlsx
or
2016-01-04-Astro-Invitro-Ex1-Smith-V01.xlsx

snake_case
Use underscores instead of blank spaces:

2016-01-04_astro_invitro_ex1_smith_v01.xlsx
or
2016-01-04_Astro_Invitro_Ex1_Smith_V01.xlsx

Very practical: you can combine different cases!

For example like this: 2016-01-04_Astro-Invitro_Ex1_Smith_v01.xlsx

Author: Ewa E. Bres, Research Data Service Center at the University of Bonn

File naming: some more tips

What elements describe the file you want to name? Use, for example:

- Unique identifiers like sample, chemical or experiment IDs (systemic identifiers are easier to track and are less confusing)
- Project or grant ID number
- Software name or device name

If you have a set of similar files, what makes each one of the files unique? (experiment number, date, version number, …)

Put the elements in the file name in the order you want your files to be sorted in. For example, if you would want to sort by date, the date should come first in the file name: 2021-10-08_exp001.xlsx

When using sequential numbering, use leading zeros:

For a sequence of 1-10: 01-10
For a sequence of 1-100: 001-010-100

Like this files (and folders!) can be sorted well and have uniform name lengths:

With leading zeros:
2021-11-16_MM_LrpKO_ Brain _GFAP_Cy3_x40_001
2021-11-16_MM_LrpKO_ Brain _GFAP_Cy3_x40_010
2021-11-16_MM_LrpKO_ Brain _GFAP_Cy3_x40_170

Without leading zeros:
2021-11-16_MM_LrpKO_Brain_GFAP_Cy3_x40_1
2021-11-16_MM_LrpKO_ Brain _GFAP_Cy3_x40_10
2021-11-16_MM_LrpKO_ Brain _GFAP_Cy3_x40_170

Author: Ewa E. Bres, Research Data Service Center at the University of Bonn

And finally…

The file naming should be **precise and understandable**: for you, your project partners and third parties.

Name files thoughtfully. Do not generate names with 64 characters!
Better: **short**, **simple** and **clear.**

Document your conventions. At this timepoint, you may know exactly what a particular abbreviation means, but would another researcher understand it as well? Would you still remember what this abbreviation means in a year from now?

Bulk renaming of files: tools

Windows:

- Ant Renamer
  www.antp.be/software/renamer
- RenameIT
  https://sourceforge.net/projects/renameit/
- Bulk Rename Utility
  www.bulkrenameutility.co.uk/

Mac:

- Renamer 6 (for Mac)
  https://renamer.com/
- Name Changer
  https://mrrsoftware.com/namechanger/

Linux/Unix:

- GNOME Commander
  https://gcmd.github.io/
- GPRename
  http://gprename.sourceforge.net/
- **rename** command

Author: Ewa E. Bres, Research Data Service Center at the University of Bonn

Template: Naming convention table

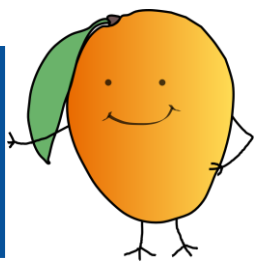
File type	File format	Unique elements & their abbreviations	Version	Naming convention with a name draft
Microscope image	.tiff	- Date (YYYY-MM-DD) - Author (Name initials, eg. Max Exampleton = ME) - Image description: animal (eg. wildtype = WT, knock-out = KO), tissue (eg. brain = Br), protein (eg. glial fibrillary acidic protein = Gfap), antibody conjugate (eg. horseradish peroxidase = Hrp) - Microscope settings (x10, x20, x40 magnification)	No	YYYY-MM-DD_Author_Animal_Tissue_Protein_Conjugate_Magnification.tiff For example: 2021-11-16_ME_KO_Br_Gfap_Hrp_x40.tiff
Protocol	.txt	- Date (YYYY-MM-DD) - Author (Name initials, eg. Max Exampleton = ME) - Method description (immunohistochemistry = IHC, Western Blot = WB) - File version	Yes, versioning done manually in the file names themselves	YYYY-MM-DD_Author_Method_Version.txt For example: 2021-11-16_ME_IHC_v01.txt

Further reading

- Batch File Renaming Tools by Christine Malinowski, MIT Libraries Data Management Services. Copyright © 2020-04-27, MIT; Licensed under CC-BY 4.0, except where otherwise noted.
- Briney, Kristin A. (2020) File Naming Convention Worksheet; Licensed under CC-BY 4.0
- Data Management: File Organization by Data Management Services. Copyright © 2022-04-12 MIT; Licensed under CC-BY 4.0, except where otherwise noted.
- Version Control Tools & Techniques by Christine Malinowski, MIT Libraries Data Management Services. Copyright © 2020-04-27 MIT; Licensed under CC-BY 4.0, except where otherwise noted.

QUESTIONS? COMMENTS?

GET IN TOUCH: researchdata@uni-bonn.de



Author: Ewa E. Bres, Research Data Service Center at the University of Bonn