



Coffee Lectures

# Research data documentation

Best practices in file naming and folder structure

Dr. Fabian Schmid

23 June 2021, Webinar via Zoom

# DISCLAIMER

If not otherwise indicated, proprietary images in this slide deck are licensed under CC BY 4.0 by Andres Bucher. Please also visit <http://andresbucher.ch/cc/library.html> in order to ascertain the images this license is applicable for.

Icons of third-party services are the property of the respective companies.

# Learning Goals

- You learn about best practices and conventions in file naming
- You become familiar with different file categories and you get to know appropriate file folder structures for organizing your research data

# What is consistent file naming and folder organization good for?

## Benefits

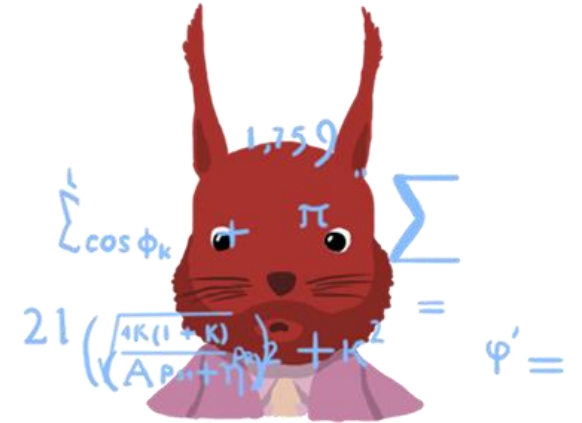
- Identify content **without opening** the file
- Easily **locate and filter** files
- Quickly identify **missing files**

## Implications

- Makes your research data **easier to find** (e.g., chronological order, file version numbering)
- Structures your research data and **avoids duplication** of files and folders
- Simplifies **collaborative work** (e.g., collaborations with external partners)
- Execution of your research project will be **more efficient**

## Advice

- Check **established naming conventions and folder structure** in your scientific discipline
- **Create folders before** you have even started with your research project



# General advice: File naming



- Use **unique** file names that reflect the **content** (not only for yourself!)
- Limit the file name to **30 characters** (preferably less)
- Use **ASCII** characters, **no spaces** and **points only before file extension** (special characters such as ~ ! @ # \$ % ^ & \* ( ) ` ; < > ? , [ ] { } ' " and | should be avoided)
- Unique file names should be **independent of upper- and lower-case** letters (not all OS are case-sensitive)
- Write dates like **YYYY-MM-DD** or **YYYYMMDD** (you will get chronological order of your files, if at the beginning of the file name)
- Use **leading zeros** in a **sequential numbering system** (e.g. file\_v01 to file\_v10 for a 1 to 10 numbering sequence)

# Examples: File naming



Good practice	Example	Comment
<b>NO</b>	2021.05.14_myFile.txt	first two points and upper case

# Examples: File naming



Good practice	Example	Comment
<b>NO</b>	2021.05.14_myFile.txt	first two points and upper case
<b>NO</b>	These-are-already_too-many_characters.docx	42 characters

# Examples: File naming



Good practice	Example	Comment
<b>NO</b>	2021.05.14_myFile.txt	first two points and upper case
<b>NO</b>	These-are-already_too-many_characters.docx	42 characters
<b>NO</b>	my file.txt	no spaces



# Examples: File naming



Good practice	Example	Comment
<b>NO</b>	2021.05.14_myFile.txt	first two points and upper case
<b>NO</b>	These-are-already_too-many_characters.docx	42 characters
<b>NO</b>	my file.txt	no spaces
<b>NO</b>	myfile_comments&suggestions.csv	wrong character [&]

# Examples: File naming



Good practice	Example	Comment
<b>NO</b>	2021.05.14_myFile.txt	first two points and upper case
<b>NO</b>	These-are-already_too-many_characters.docx	42 characters
<b>NO</b>	my file.txt	no spaces
<b>NO</b>	myfile_comments&suggestions.csv	wrong character [&]
<b>YES</b>	v12_myfile_FMS_20210516	in numerical order

# Examples: File naming



Good practice	Example	Comment
<b>NO</b>	2021.05.14_myFile.txt	first two points and upper case
<b>NO</b>	These-are-already_too-many_characters.docx	42 characters
<b>NO</b>	my file.txt	no spaces
<b>NO</b>	myfile_comments&suggestions.csv	wrong character [&]
<b>YES</b>	v12_myfile_FMS_20210516	in numerical order
<b>YES</b>	2021-05-16_FMS_myfile_v12	in chronological order

# Renaming of many files

You have already collected **plenty of files** and you do want to **reorganize and rename** them?  
The following applications can help you for **automatic renaming**:

- Bulk Rename Utility (Windows, free)  
[http://www.bulkrenameutility.co.uk/Main\\_Intro.php](http://www.bulkrenameutility.co.uk/Main_Intro.php)



- Renamer 4 (Mac)  
<http://renamer4mac.com/>



- PS FileRenamer (Windows, free for 30 days)  
<http://www.polarsoft.de/ps-filerenamer.html>

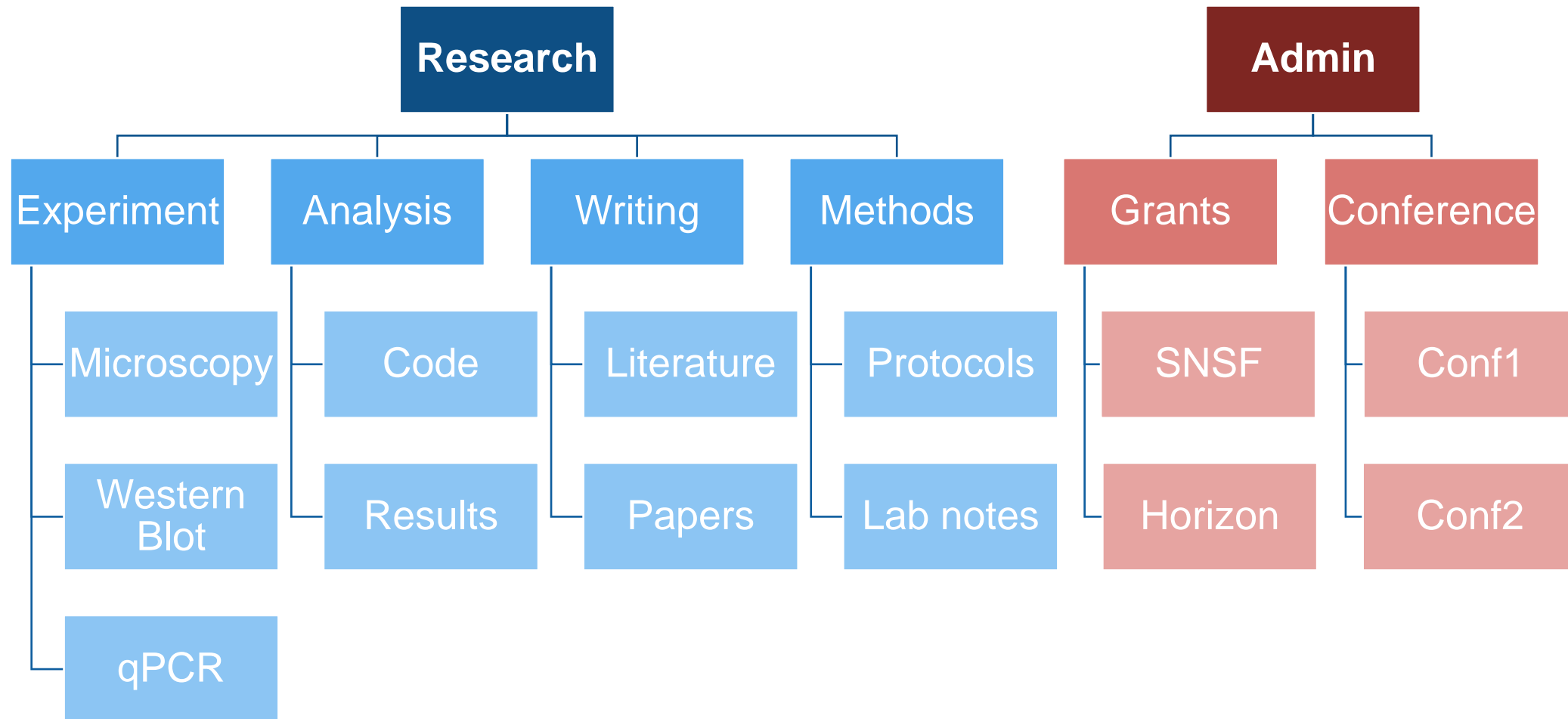


# General advice: File categories and file folder structure

- DISTINGUISH BETWEEN.....
  - **work** and **private** material
  - **own** and **others'** work (papers vs. literature)
  - **research** vs. **administrative** content
  - **processed** vs. **raw** data
  - **experiment** vs. **analysis**
  - **experimental runs/replicates** (if necessary)
- Do not use **more than 5 levels** of folders
- Try to avoid **overlapping names** in categories (e.g., Data vs. Presentation-Data)
- Relocate files you are **no longer working** with to a steady storage location (backed up)
- Keep **path names short** (< 255 characters)



# Possible file folder structure for a PhD project



# README file for documentation of systematic file naming and folder organization method

## File and folder schema (example)

File type	Filename schema	Schema key	Example filename
Microscopic image	[Date]_[microscope]_[signal]_[image number]	Date: Date of image capture in YYYYMMDD format, microscope: name of microscope, signal: source of light or dye used image number: sequential order of taken images from 1:100	20210623_Isconf28_GFP_001.tiff
...			

### Modified from

README: File & Folder Schema Example by MIT Libraries Data Management Services. Copyright © 2018 MASSACHUSETTS INSTITUTE OF TECHNOLOGY, licensed under a [Creative Commons Attribution 4.0 International License](#) except where otherwise noted.

Access at [https://www.dropbox.com/s/ritd1mwzyaz2dh6/Sample\\_README\\_fileOrg.docx?dl=0](https://www.dropbox.com/s/ritd1mwzyaz2dh6/Sample_README_fileOrg.docx?dl=0)

# README file for documentation of systematic file naming and folder organization method

## File and folder schema (example)

File type	Filename schema	Schema key	Example filename
Microscopic image	[Date]_[microscope]_[signal]_[image number]	Date: Date of image capture in YYYYMMDD format, microscope: name of microscope, signal: source of light or dye used image number: sequential order of taken images from 1:100	20210623_Isconf28_GFP_001.tiff
...			

## File name abbreviations (example)

File name descriptor	Abbreviation's key
Name of microscope	Isconf28: laser scanning confocal microscope located in room 28 of the microscopy facility at...
...	

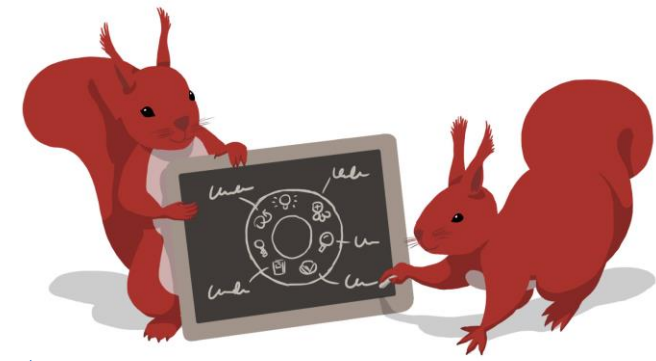
### Modified from

README: File & Folder Schema Example by MIT Libraries Data Management Services. Copyright © 2018 MASSACHUSETTS INSTITUTE OF TECHNOLOGY, licensed under a [Creative Commons Attribution 4.0 International License](#) except where otherwise noted.

Access at [https://www.dropbox.com/s/ritd1mwzyaz2dh6/Sample\\_README\\_fileOrg.docx?dl=0](https://www.dropbox.com/s/ritd1mwzyaz2dh6/Sample_README_fileOrg.docx?dl=0)



# Further reading and resources to credit



- Data management guide, University of Cambridge  
<http://www.data.cam.ac.uk/data-management-guide/organising-your-data>
- Advice on organising files and folders  
<http://www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub/Browse-by-Subject/Organising-files-and-folders.htm>
- Data management services, Stanford Libraries  
<https://library.stanford.edu/research/data-management-services>
- Data management support, MIT Libraries  
<https://libraries.mit.edu/data-management/>
- Guide from Cornell University explaining content and structure of README files  
<https://data.research.cornell.edu/content/readme#recommendedcontent>

Dr. Fabian Schmid  
Research Data Management and Digital Curation  
fabian.schmid@library.ethz.ch  
ETH Library  
Research Support Services

Rämistrasse 101  
8092 Zurich

[www.library.ethz.ch/RDM](http://www.library.ethz.ch/RDM)  
[data-management@library.ethz.ch](mailto:data-management@library.ethz.ch)



# Survey





# Thanks for your attention!

Next Coffee Lecture:

What do I have to consider when entering  
research data into the Research Collection

30 June 2021, Webinar via Zoom  
[www.library.ethz.ch/coffee-lectures](http://www.library.ethz.ch/coffee-lectures)