



SLUB

Wir führen Wissen.

Lessons in Open Science

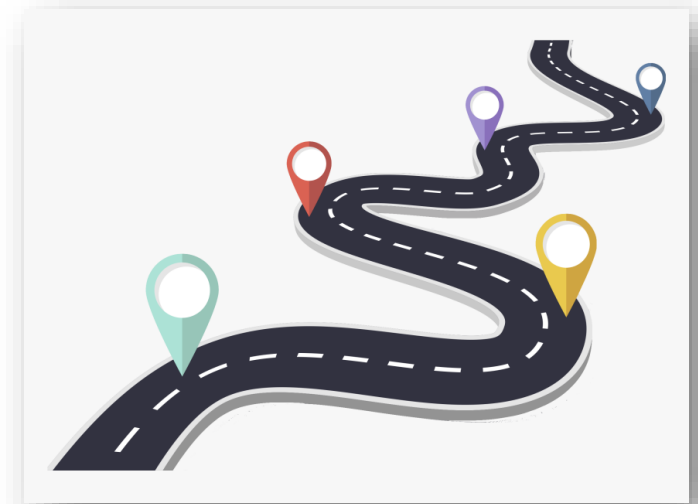
Best Practices in Personal RDM (Research Data Management)

Dr. Andreas von der Dunk
Dresden Service Center Research Data

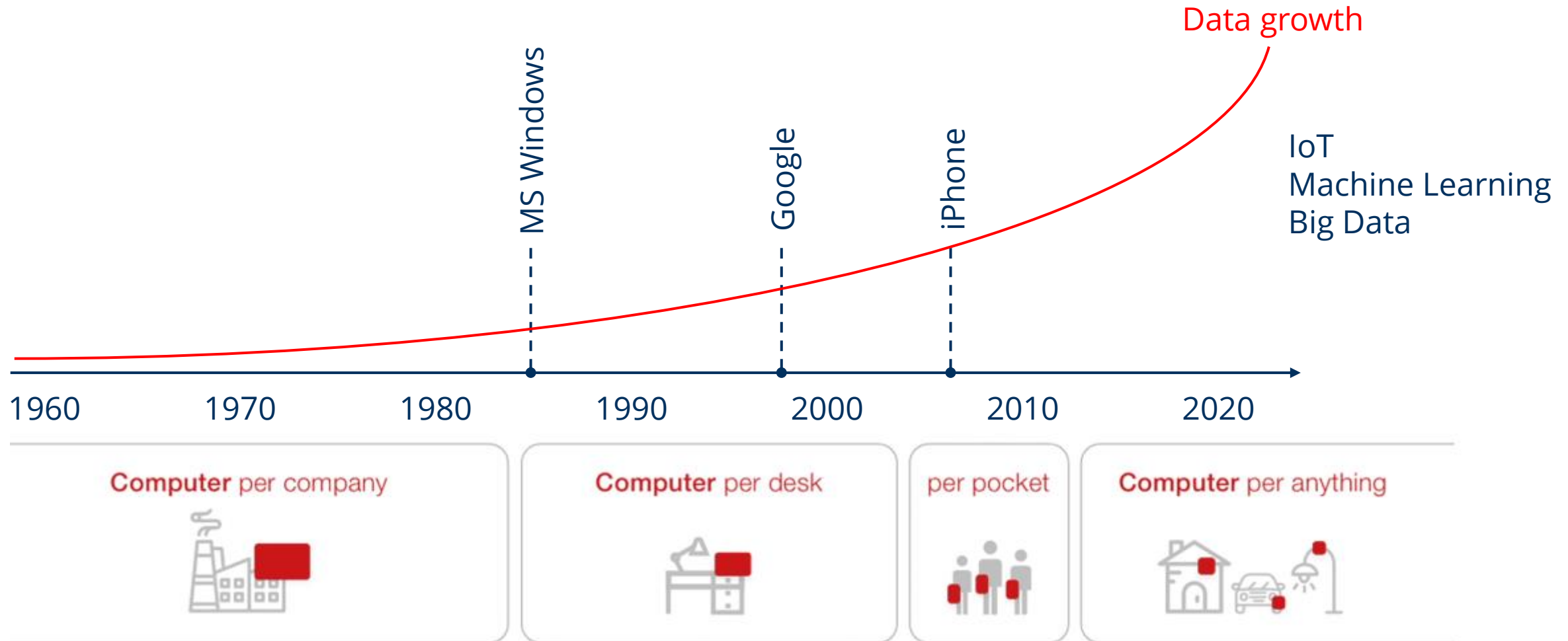
October 1, 2021

Agenda

1. **Research Data Management – why the Hype?**
2. File Organization
3. File Backup
4. Research Data Management & Open Science
5. Next Steps



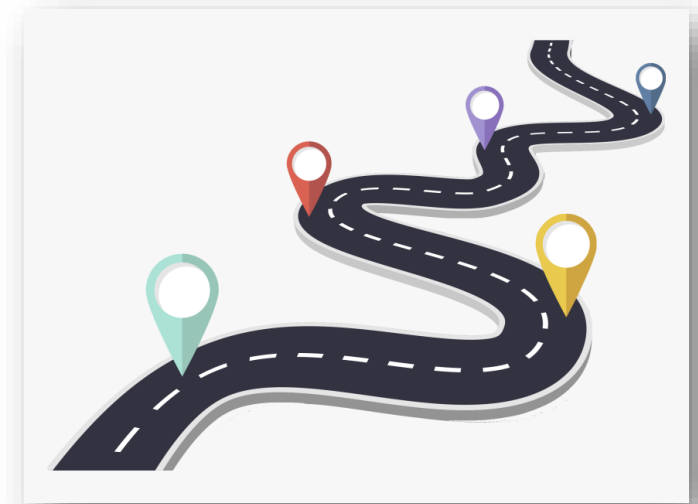
Ever increasing data volume



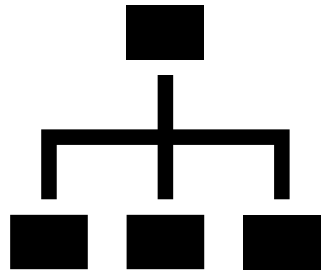
Source: <https://d-one.ai/>

Agenda

1. Research Data Management – why the Hype?
- 2. File Organization**
3. File Backup
4. Research Data Management & Open Science
5. Next Steps



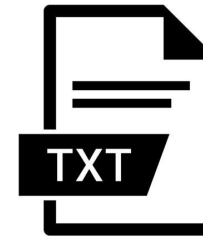
File organization



Folder structure

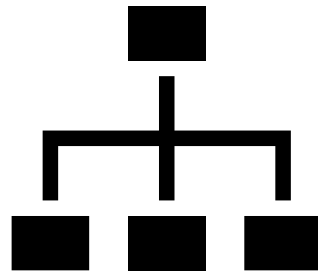


File naming



Documentation

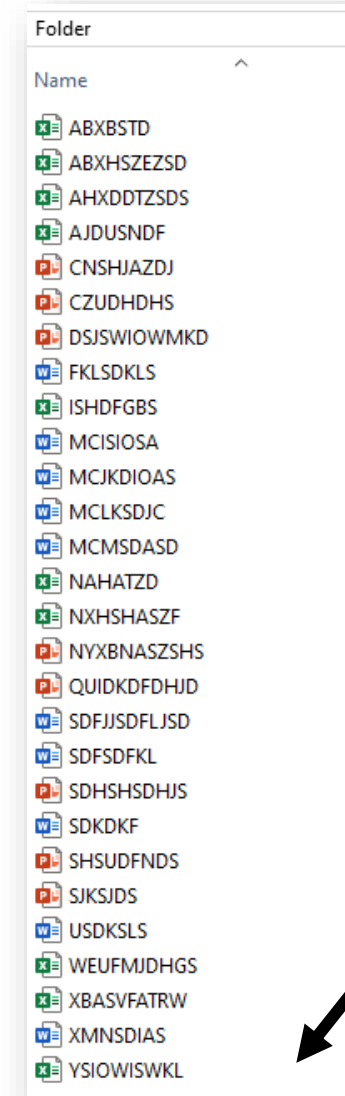
Folder Structure



Folder Structure

Do not let your folders grow **too large**

- more than 7* files in a folder?
Consider aggregating them into subfolders
- common exception: uniform datasets, e.g. time series data



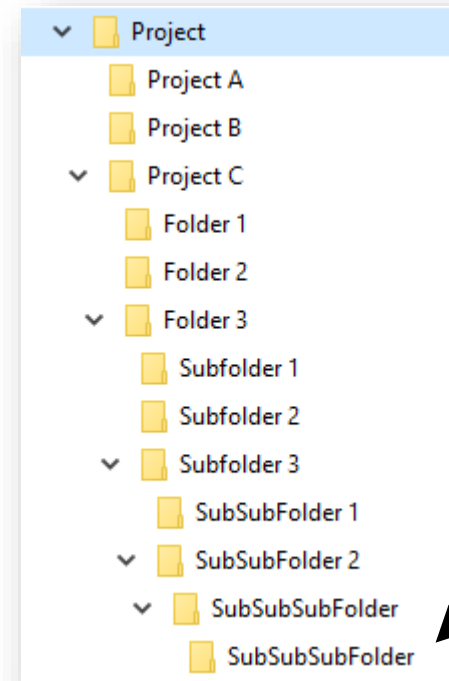
scroll, scroll, scroll ...

* for further information please see: https://en.wikipedia.org/wiki/The_Magical_Number_Seven,_Plus_or_Minus_Two

Folder Structure

Do not let your folder structure grow **too deep**

- too many hierarchies complicate navigation
- 4-5 levels of folders most efficient

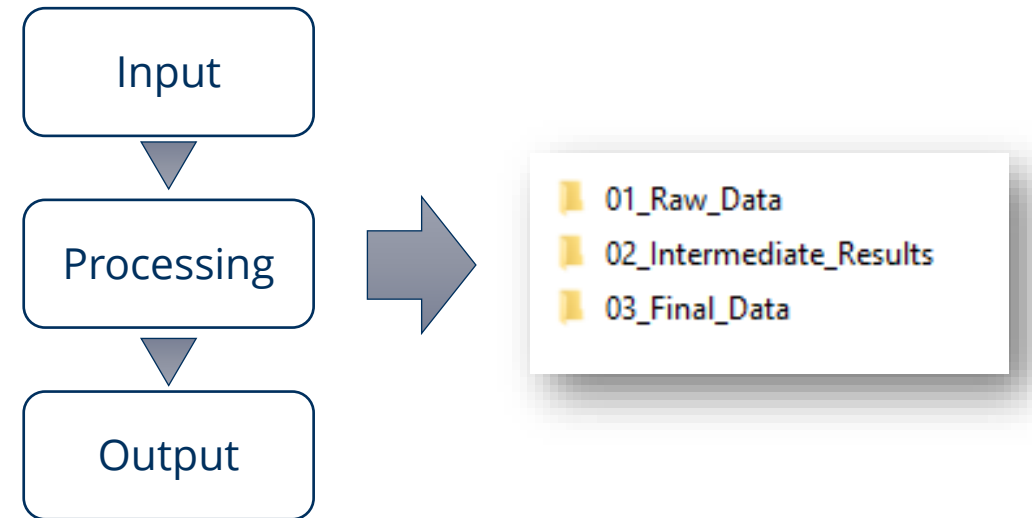


click, click, click ...

Folder Structure

Use separate folders for:

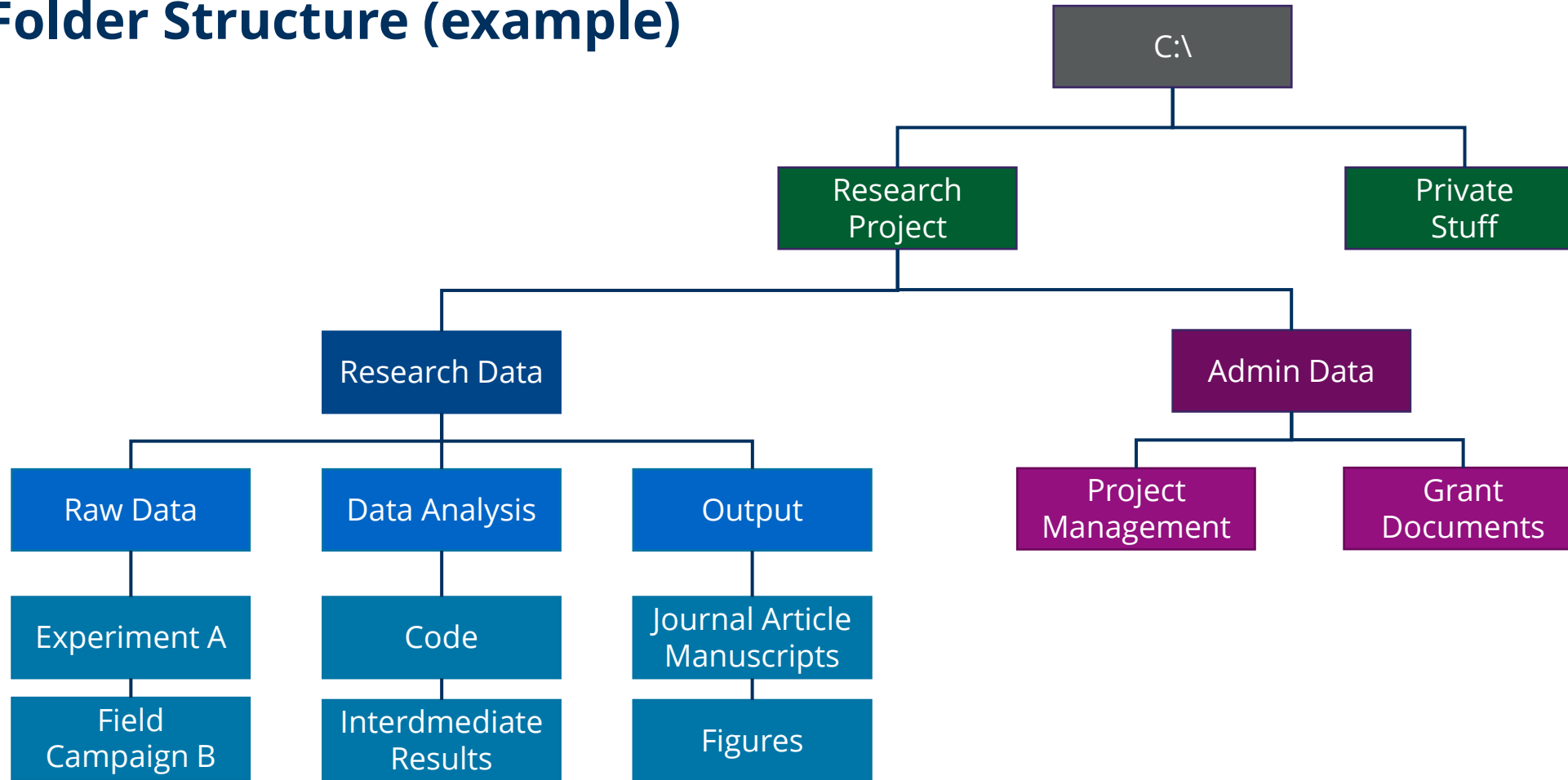
- work files and private files
- research files and administrative files
- research files: **raw** data, **processed** data and **final** data



Card sorting* is a proven technique for creating folder structures

* for further information on card sorting please see: <https://uxdesign.cc/card-sorting-what-how-the-perks-29f6cb020270>, as referred to in this article on how to create a folder structure: <https://blog.imagerelay.com/2016/folder-organization>

Folder Structure (example)

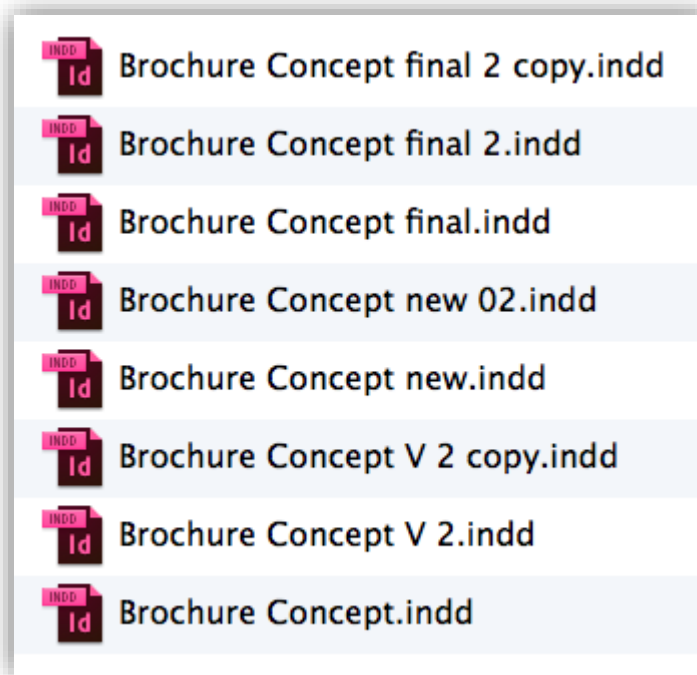


* this example inspired by: https://ethz.ch/content/dam/ethz/associates/ethlibrary-dam/documents/Aktuell/Kurse/CoffeeLectures/2021-06-23_Coffee_Lecture_File-naming_final.pdf

File naming



File naming



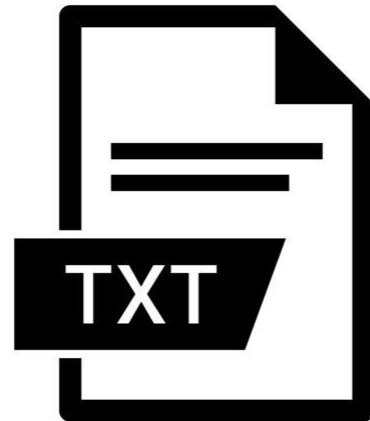
File naming

File naming convention (example): YYYY-MM-DD_Author_Topic

Tips for creating a file naming convention

- Use naming convention **consistently** throughout the project
- Use **meaningful**, self-explanatory names; avoid "work", "temp"
- **De-code** any abbreviations or IDs in a code-sheet
- Naming conventions apply to **files, folders** and (!) column **headings**
- Don't use special characters such as & , * % # ; * () ! @ \$ ^ ~ ' { } [] ? < >

Documentation



Documentation

Your data **without** documentation

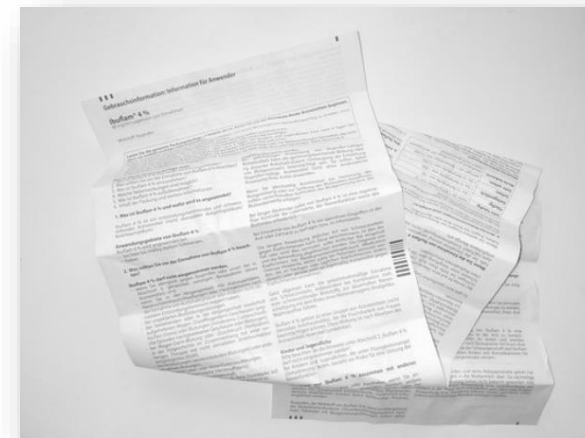


Image source: <https://pixabay.com/>

Your data **with** documentation



Image: Brett Jordan on <https://www.unsplash.com/>



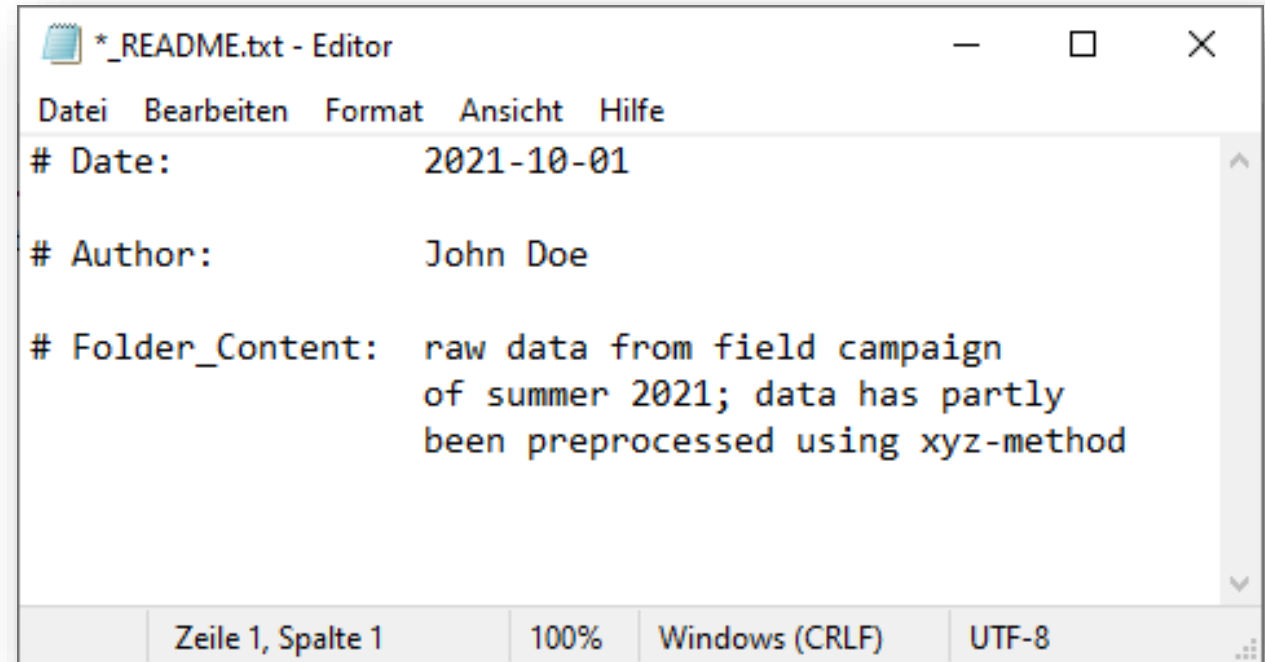
Documentation

Easiest/fastest way to document your data:

- **README.txt***
- 1 Readme.txt for each folder (ideally)

Contents of a README.txt:

- Date
- Author
- Description of folder content
- ...



The screenshot shows a text editor window titled "*_README.txt - Editor". The window contains the following text:

```
Datei  Bearbeiten  Format  Ansicht  Hilfe
# Date:                2021-10-01
# Author:              John Doe
# Folder_Content:     raw data from field campaign
                     of summer 2021; data has partly
                     been preprocessed using xyz-method
```

The status bar at the bottom of the window displays: "Zeile 1, Spalte 1", "100%", "Windows (CRLF)", and "UTF-8".

* for further information please see: <https://data.research.cornell.edu/content/readme>

Documentation

Code Sheet* (aka *Code Book*, *Data Dictionary*)

- stored in top level folder
- aim: clarifying **units** (meters? feet?) and **coded values** (male = 0 / female = 1)

Description	Range/Units/Values
Height in cm	100 - 250
Gender	Male = 0, Female = 1
Religion	Hindu = 1, Christian = 2, Muslim = 3
Family size	≤5 = 0, ≥5 = 1
...	...

*for further information please see <https://www.icpsr.umich.edu/icpsrweb/content/shared/ICPSR/faqs/what-is-a-codebook.html>

Documentation

Proper file naming / documentation releases full potential of **desktop search tools**, e.g.

- Windows Explorer Search
- Spotlight (Mac)

Also check out 3rd-party tools like

- Copernic
- Everything

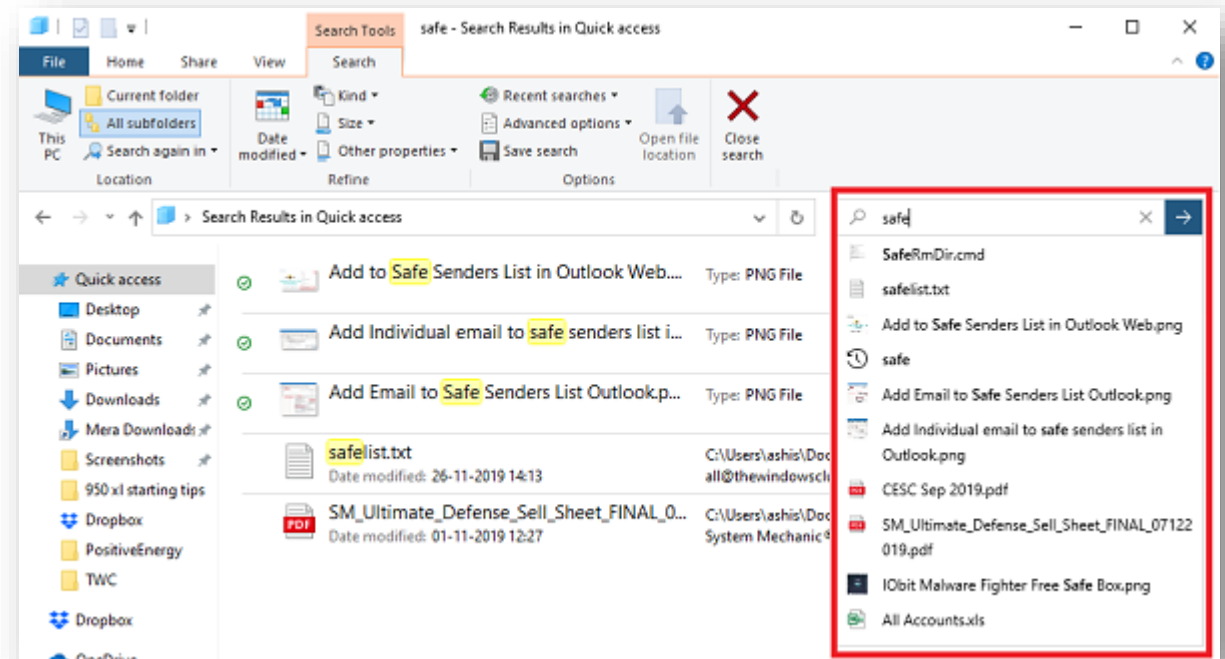


Image source: <https://www.thewindowsclub.com/windows-10-file-explorer-search-not-working-properly>

Useful tools for file organization

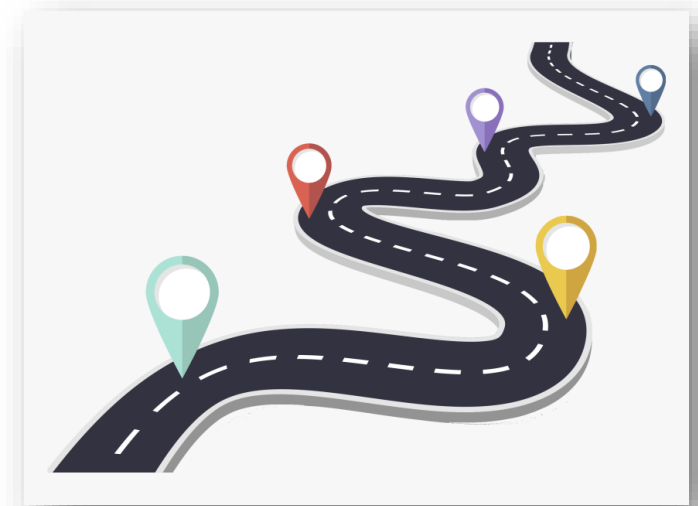
Renaming files – use batch renaming tools when renaming a large number of files, e.g. <https://www.advancedrenamer.com/>

Removing duplicates – use duplicate finder tools to clear out your storage, e.g. <http://www.alldup.de/>

Copying files – do not use Ctrl+C / Ctrl+V when copying large datasets; use e.g. Robocopy (built-in Windows utility) or rsync (Linux, Mac) <https://rsync.samba.org>

Agenda

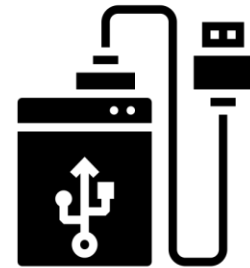
1. Research Data Management – why the Hype?
2. File Organization
- 3. File Backup**
4. Research Data Management & Open Science
5. Next Steps



Backup

3-2-1 Backup Rule:

- at least 3 copies of the data...
- ... stored on 2 different kinds of storage media...
- ... 1 copy is physically separated from the others



Backup

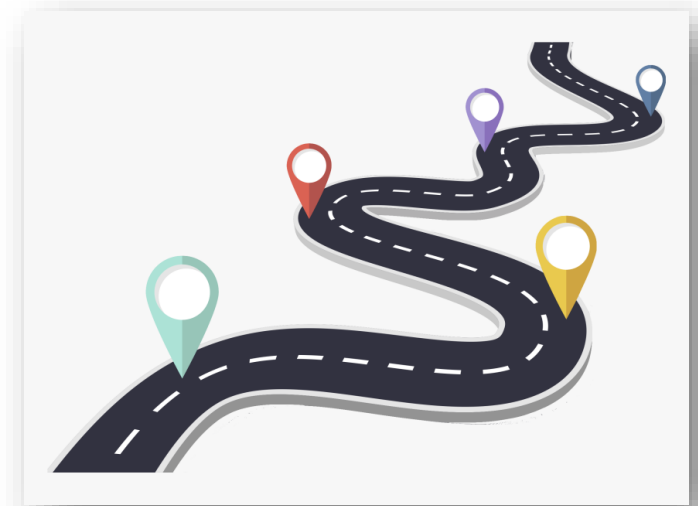
Advice: do not store your backup data on USB flash drives

- increased risk of **physical loss**
- increased risk of **accidental overwrite**



Agenda

1. Research Data Management – why the Hype?
2. File Organization
3. File Backup
- 4. Research Data Management & Open Science**
5. Next Steps



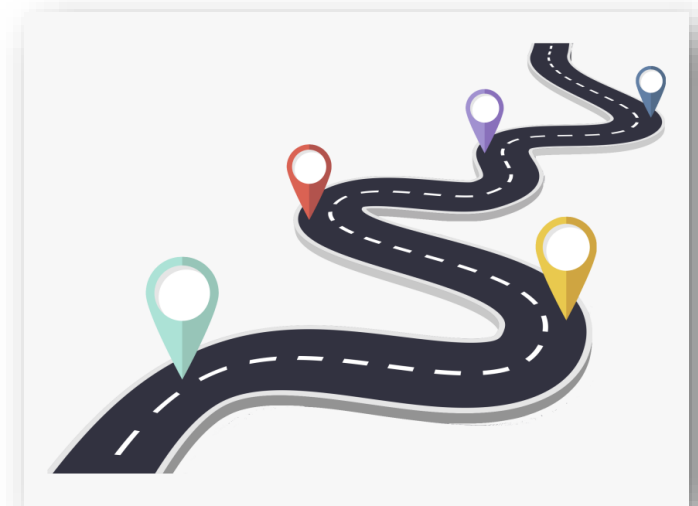
Research Data Management & Open Science

- Research Data Management \neq Open Science
- Research Data Management is a **prerequisite** for Open Science



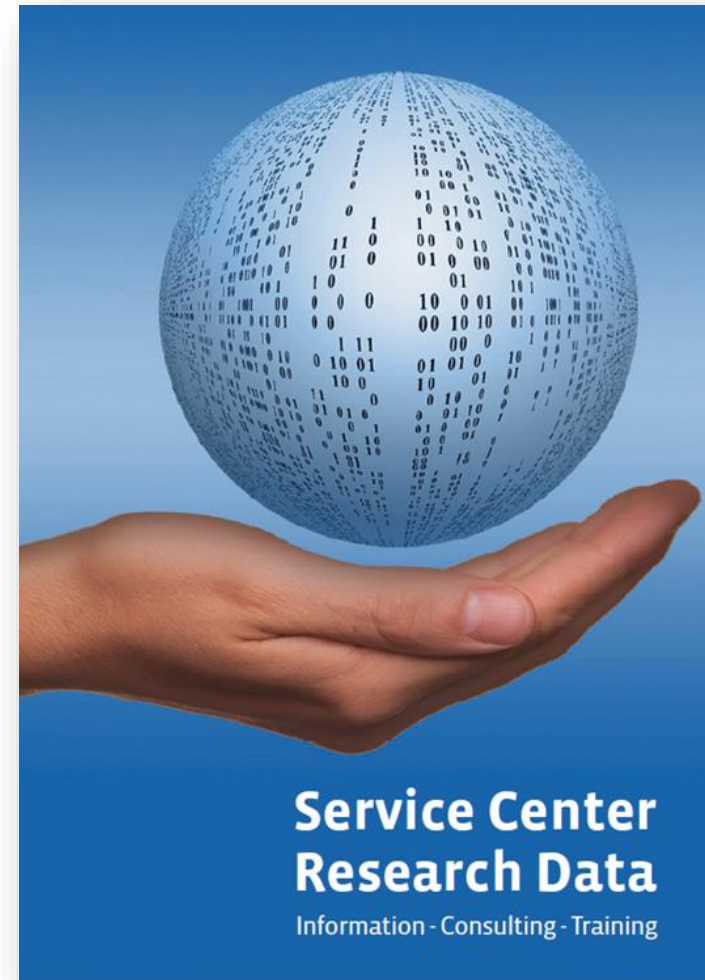
Agenda

1. Research Data Management – why the Hype?
2. File Organization
3. File Backup
4. Research Data Management & Open Science
5. **Next Steps**



Next steps

- browse this presentation
- check out the links provided
- check out the user stories on our website*
- get in touch with us, book a consultation**



*user stories: https://tu-dresden.de/forschung-transfer/services-fuer-forschende/kontaktstelle-forschungsdaten/news?set_language=en

** book a consultation: <https://tu-dresden.de/forschung-transfer/services-fuer-forschende/kontaktstelle-forschungsdaten/unser-service/unterstuetzung-bei-ihrem-fdm>