

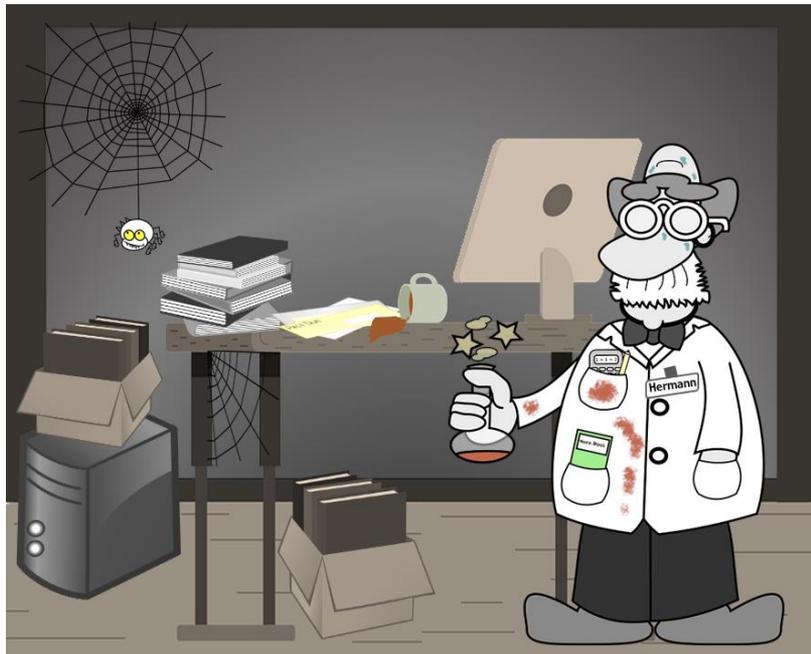
My research data and me: How it can work!

Research Data Management - Basics



Find the differences?

Former practices



FAIR practices



Find the differences, by RDM Team TU Graz, [CC by 4.0](https://creativecommons.org/licenses/by/4.0/)

Content

- Research Cycle & Good Scientific Practice
- Publications & Open Access
- FAIR Principles
- File organization basics

Research circuit

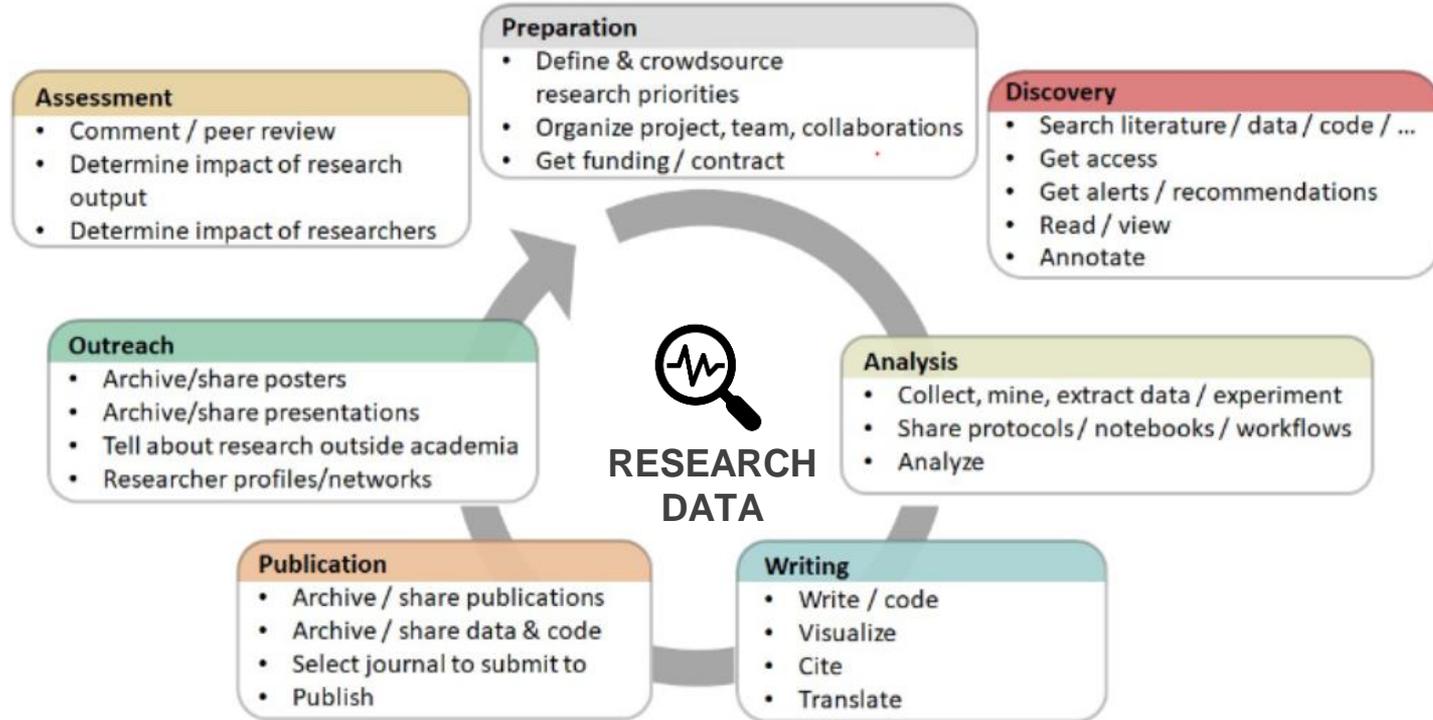
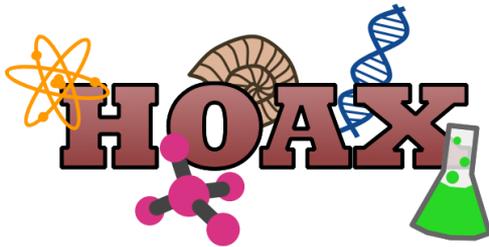


Image from <https://open-science-training-handbook.github.io/>, CC0



Good scientific practice

- Work in accordance with legal regulations & ethical standards
- Clear and traceable documentation of results
- Subject results to critical review
- Public nature of basic research



Guidelines on Safeguarding Good Scientific
Practice

RL 92000 SGWP 050E-04

CC0, by j4p4n at <https://freesvg.org/science-hoax>

Source: TU4U TU Graz 2017

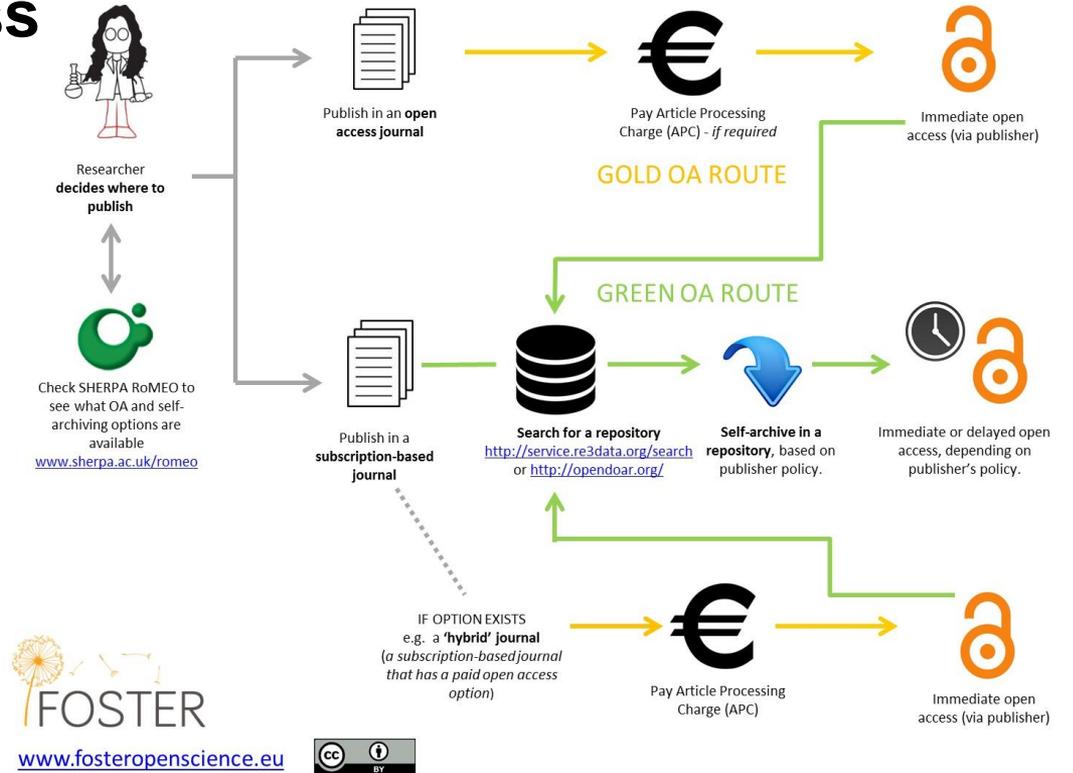
Challenges with publications



Solution: Open Access

Open Access stands for **online**, and **free access** to peer-reviewed publications with **open licences for reuse**.

[Contact: TU Graz Library](#)

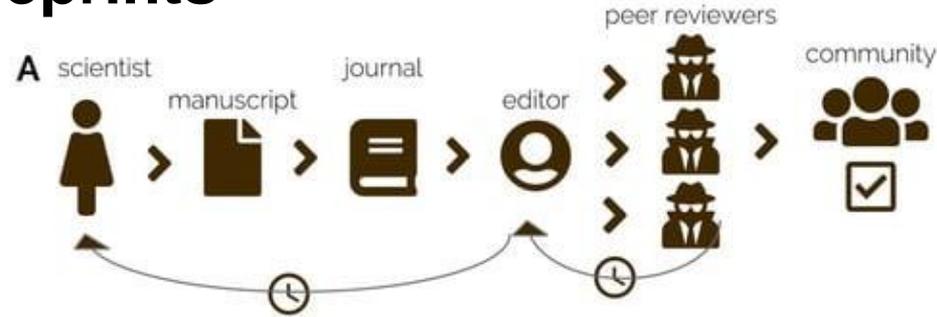


„OA-routes“, by Digital curation Center, from dcc.ac.uk/blog/fostering-open-science, [CC by 4.0](https://creativecommons.org/licenses/by/4.0/)

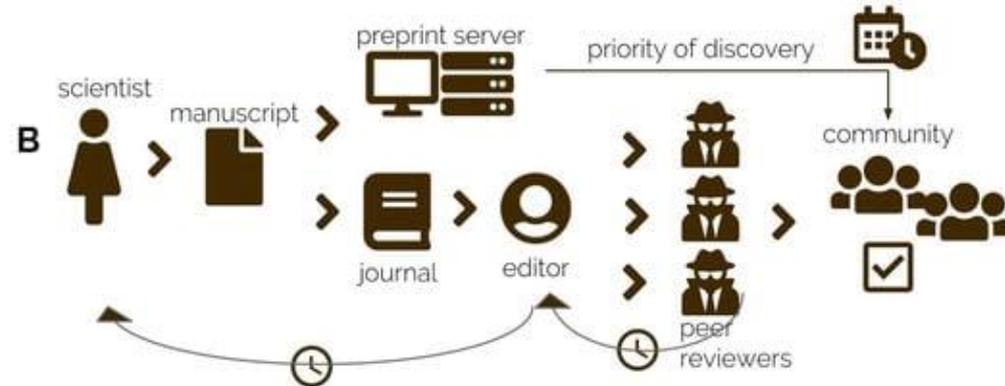


Solution: Preprints

Traditional submission



Preprint submission



Ten Hot Topics around Scholarly Publishing, Tennant et al. 2019, MDPI, [CC by 4.0](https://creativecommons.org/licenses/by/4.0/)



Solution: Registered Reports

Peer review before results are known
to align scientific values and practices



The registered report workflow, from Promoting reproducibility with registered reports. Nat Hum Behav 1, 0034 (2017). <https://doi.org/10.1038/s41562-016-0034>

Scientific publication of data

- **Scientific data**
(<https://www.nature.com/sdata/>)
- **Data in Brief**
(<https://www.journals.elsevier.com/data-in-brief>)
- **MDPI Data**
(<https://www.mdpi.com/journal/data>)
- **Data Science Journal**
(<https://datascience.codata.org>)

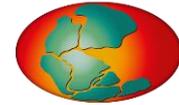


Repository, RDM Team TU Graz, [CC by 4.0](https://creativecommons.org/licenses/by/4.0/)

Data publication

As an independent information object in a research data repository

- Discipline-specific repositories, e.g.
- Cross-disciplinary repositories, e.g.
- Institutional repositories, e.g.



PANGAEA.



TU GRAZ
REPOSITORY
Bibliothek und
Archiv

Find repositories:



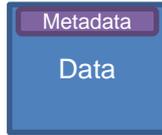
you can find repositories for your discipline.



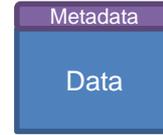
Metadata

- **Description** and **context**
- Discipline-/topic-specific schemata: <https://fairsharing.org/>
- Generic: Data cite,
<https://schema.org>,
<https://guidelines.openaire.eu/>
- Code: <https://github.com/codemeta/codemeta>,
<http://datacite.org/schema/kernel-4>

Metadata



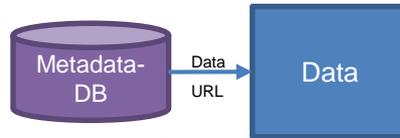
Metadata in data
(part of header of data-formats, e.g. HDF5)



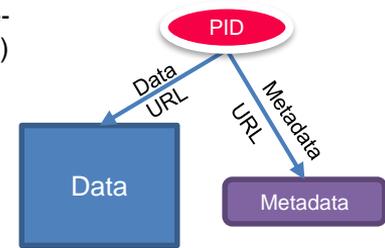
Metadata on data
(e.g. object-storage, file-and folder-designations)



Metadata with data
(e.g. readme-file, metadata)



Metadata-DB
Link to data
(e.g. search index, repository)



Persistent Identifier
linked to metadata and data
(e.g. DOI, EPIC-PID, URN)

Content from: TU9-FDM. (2019, May 3). Metadaten. Zenodo. <https://doi.org/10.5281/zenodo.2660187> - CC by 4.0



FAIR Principles

Findable - Accessible - Interoperable - Reusable

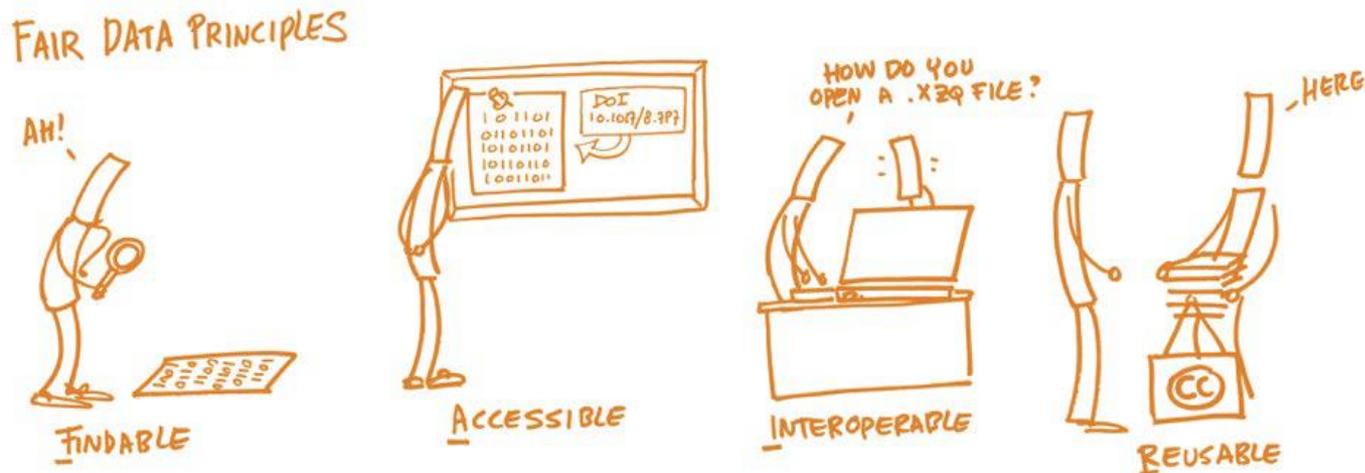


Image from <https://open-science-training-handbook.github.io/>, CC0

FAIR Principles

Findable - Accessible - Interoperable - Reusable



Image from “Resources for making your data fair”, Australian Research Data Commons, [CC By 3.0](#)

FAIR Principles

Example for FINDABLE:



Yes

Data repository



No

Personal website

| Name | Last modified | Size | Description |
|------------------|------------------|------|-------------|
| Parent Directory | - | - | - |
| foobar.zip | 2017-08-01 08:58 | 492M | |
| sec.pdf | 2017-08-01 08:59 | 38M | |

Let's make our data FAIR, by Tomasz Miksa. DOI: 10.25365/phaidra.306. [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)



FAIR Principles

Example for ACCESSIBLE:

Yes



February 26, 2021

Dataset (IV) related to publication: Decisive Role of Water and Protein Dynamics in Residence Time of p38a MAP Kinase Inhibitors

VC simulation data of **Open Data** related to the publication Percec et al. Decisive Role of Water and Protein Dynamics in Residence Time of p38a MAP Kinase Inhibitors

Other datasets related to this publication: 16 (DOI) (dataset: 4772444; 16 (DOI) (dataset: 458118; 16 (DOI) (dataset: 4581797)

- Individual ZIP files contain raw experimental trajectories (cell time files and 10 files)
- Cell files contain cell coordinates for the energy-minimized measurable data point structures

Open

Restricted Access

Normally, request access to the files in this dataset, provided by you, to the conditions below. The decision whether to grant this access is solely under the responsibility of the data provider.

Details will be made open in a later open publication.

[Request access](#)



Restricted access,
but a clear way to
request access

No



Enter password

Enter password for the encrypted file
C:\Users\... Desktop\View folder (7)\Fi... \20130607_103638.jpg
in archive Finance.zip

Enter password

Show password

Use for all archives

Organize passwords...

OK Cancel Help

Let's make our data FAIR, by Tomasz Miksa. DOI: 10.25365/phaidra.306. [CC-BY 4.0](#)

FAIR Principles

Example for INTEROPERABLE:

Yes

- XML following known XSD Schema
- MP3 for audio recordings

No

- Custom XML without any documentation
- M4P (Apple) for audio recordings

Let's make our data FAIR, by Tomasz Miksa. DOI: [10.25365/phaidra.306](https://doi.org/10.25365/phaidra.306). [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)

FAIR Principles

Example for REUSABLE:

Yes

Trusted source, permission to reuse, well defined meaning of terms used

No

Provenance and permissions not clear

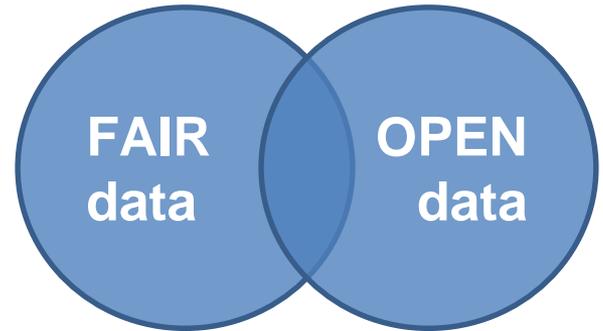
Let's make our data FAIR, by Tomasz Miksa. DOI: 10.25365/phaidra.306. [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)



FAIR Principles

Towards "as **FAIR** as possible" and "as **open** as possible"

- **FAIR** principles \neq **open** data sharing
- FAIR = **best practices** for data exchange
- Data not shared, still **metadata** defined

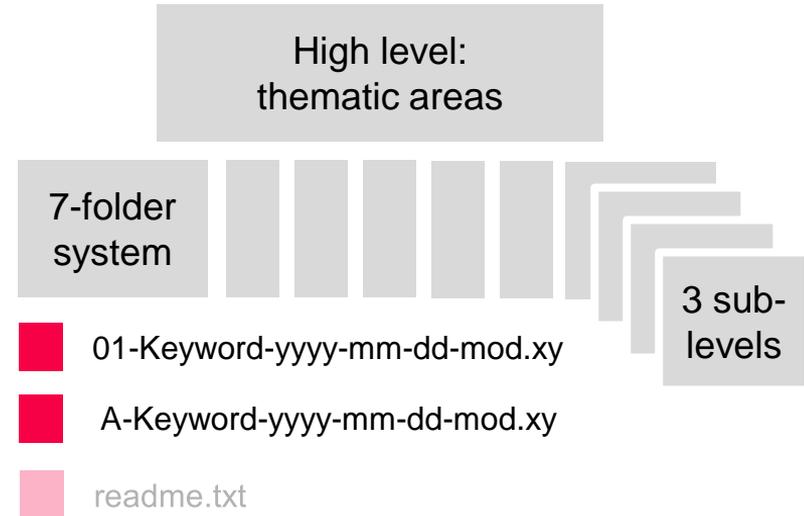


Basic Data Management – Toolless File Organization

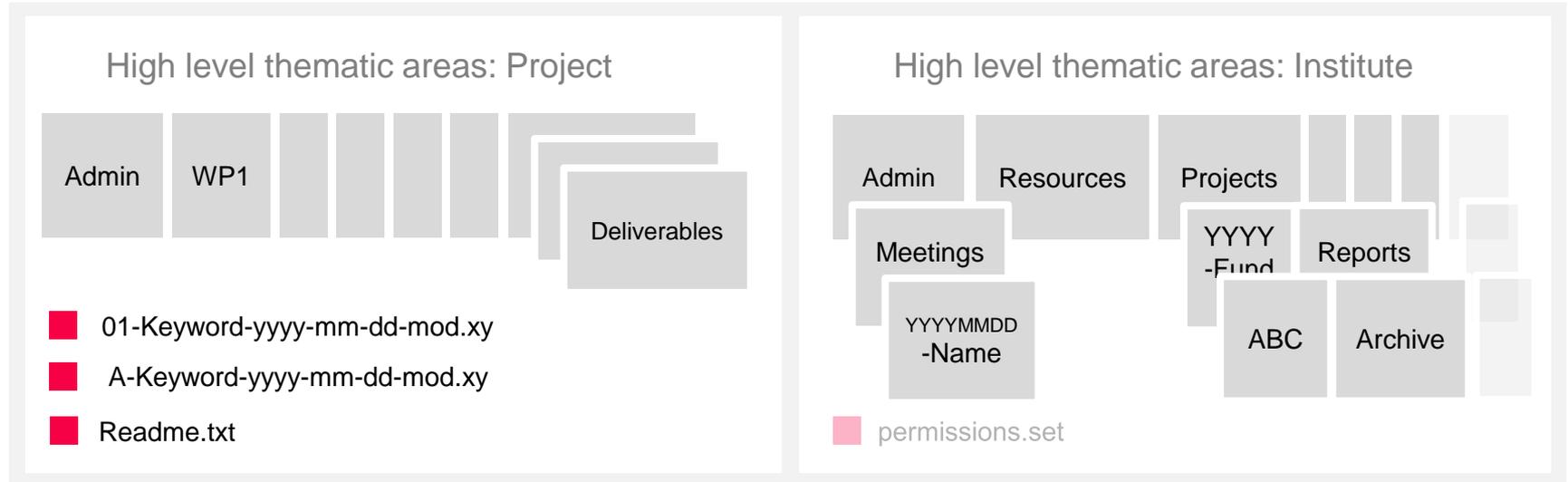
AVOID: How did I name the file, where did I put it?
What is what?

Basic Data Management – Toolless File Organization

- Clear hierarchical folder structure
 - similar subfolders
- Separate sensitive structure
- Files/records at bottom level
- Document consistency



Basic Data Management – Toolless File Organization



Toolless File Organization - Summary

(1) Define **naming rules** and options

- document rules, e.g. readme.txt in root/main folders!
- consistency, e.g. keyword, date, modification
- hierarchy structure (levels with similar subfolders, max 5)
- avoid special characters and spaces (– only hyphens _or underscores)
- metadata? Content & Provenance (author, changes, date)

(2) **Group**, e.g. projects/topics/resources - group working data together!

- set permissions

(3) Avoid duplicates/redundance! - keep necessary **versions** only!

- keep track

(4) **Numbering** & Order

- year/month/day in similar format
- number in front: importance



(5) **Revise** regularly!

Basic Data Management – Further Materials

Knowledge clip:

Keeping research data organized

CC By 3.0 Ugent Data Stewards



Disciplinary In-depth Exercise:

Working with files – FAIR in (biological) practice
(carpentries-incubator.github.io)

CC-BY 4.0 2023 by Zieliński T, Romanowski A, et al.



messy/unstructured

organised folder structure

Research Data Management @ TU Graz

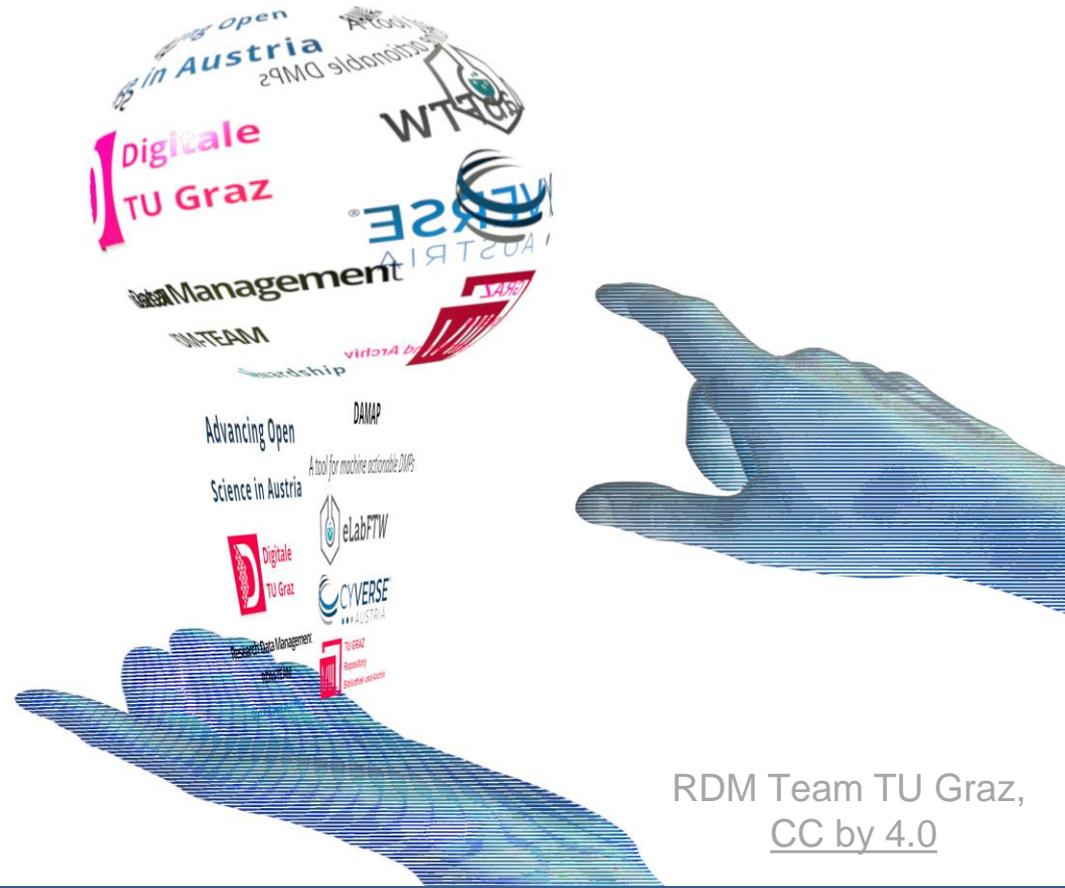
- DMP Support
- RDM Optimization Institute/Group: Workflows/Pipelines, Infrastructure
- RDM Individual support & support in Research Projects (Collaborations)
- RDM guidelines & policies
- Tool-dependent support
(workshops, introduction, training, assistance, development)

Supporting reproducible research

rdmteam@tugraz.at



<https://rdm.tugraz.at>



RDM Team TU Graz,
[CC by 4.0](https://creativecommons.org/licenses/by/4.0/)