

TRAINING: Data Science and AI for Medicine Training School 2026
Day 2: RDM

SPEAKER: Lea Kabjesz, Maryam Khodaei Dolouei, Toralf Kirsten

Using materials from Robert Haase (DSC ScaDS.AI / Leipzig University)
These slides may be reused under the terms of the CC-BY 4.0 license, unless otherwise specified.

Gefördert durch:



Bundesministerium
für Forschung, Technologie
und Raumfahrt



SACHSEN Diese Maßnahme wird gefördert durch die Bundesregierung
aufgrund eines Beschlusses des Deutschen Bundestages.
Diese Maßnahme wird mitfinanziert durch Steuermittel auf
der Grundlage des von den Abgeordneten des Sächsischen
Landtags beschlossenen Haushaltes.

AGENDA

- Theory and Terms
 - FAIR-Principles
- Medical Data
- Practice
 - Data Types & Principles in RDM
- MII Academy

Sneak Preview ;-)



Basics



What does it
actually mean?

Research Data Management (RDM)

Legal Studies

copyright licensing

Ethics

consent privacy

transparency

research integrity

Project Management

planning documentation workflows

IT & Cybersecurity

access control storage

data protection backup

Data Science

data quality curation

metadata analysis/ AI readiness



Principles in RDM

FAIR

FINDABLE
ACCESSIBLE
INTEROPERABLE
REUSABLE

data should be...

easy to discover

easy to retrieve

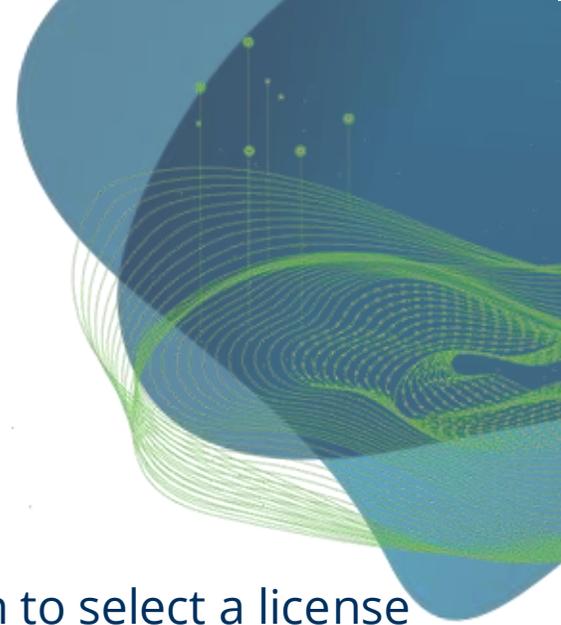
easy to combine with other data or tools

easy to use again (by people + computers)



As **open** as possible, as **closed** as necessary!

Guidelines



1) Register datasets in public repositories

- use trusted repositories with clearly defined access conditions and option to select a license
- choose a repo that assigns a **DOI** (= **permanent** string **uniquely** identifying a digital object)



<https://doi.org/10.1038/s41392-020-00243-2>



Everything is data but where can I upload it?

Data Repositories



for code



for posters, presentations



for images



Hugging Face

for DL models, datasets

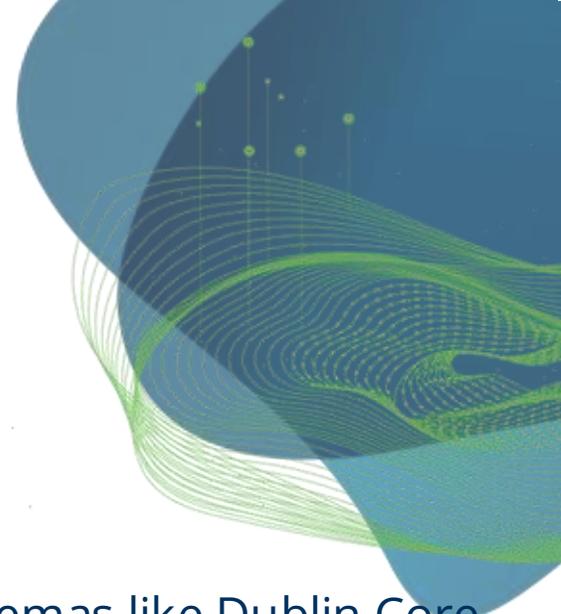


→ next talks

special considerations
for medical data



Guidelines



2) Provide documentation & metadata (= data about data)

- describe context, data structure, variables, methods using metadata schemas like Dublin Core
- use tools like ReadMe files, Jupyter Notebooks

→ watch out for open standards, file formats (e.g. CSV, TXT, JSON, XML, PDF) & schemas!

→ watch out for consistent file naming & version control!

All of this is possible...

Name	Date modified	Type	Size
↳ ú ε-τib. oo√		File folder	
↳ +1φ Γ. ó		File folder	

- 4B7C3646463D2D6B0139D159360A9D91_02_07-adverbs.jpg
- 6E839ED619188BAD4B4FF602819487B7_03_02-possessive.jpg

- 03_08-commas.mp4
- 03_09-minormarks.mp4

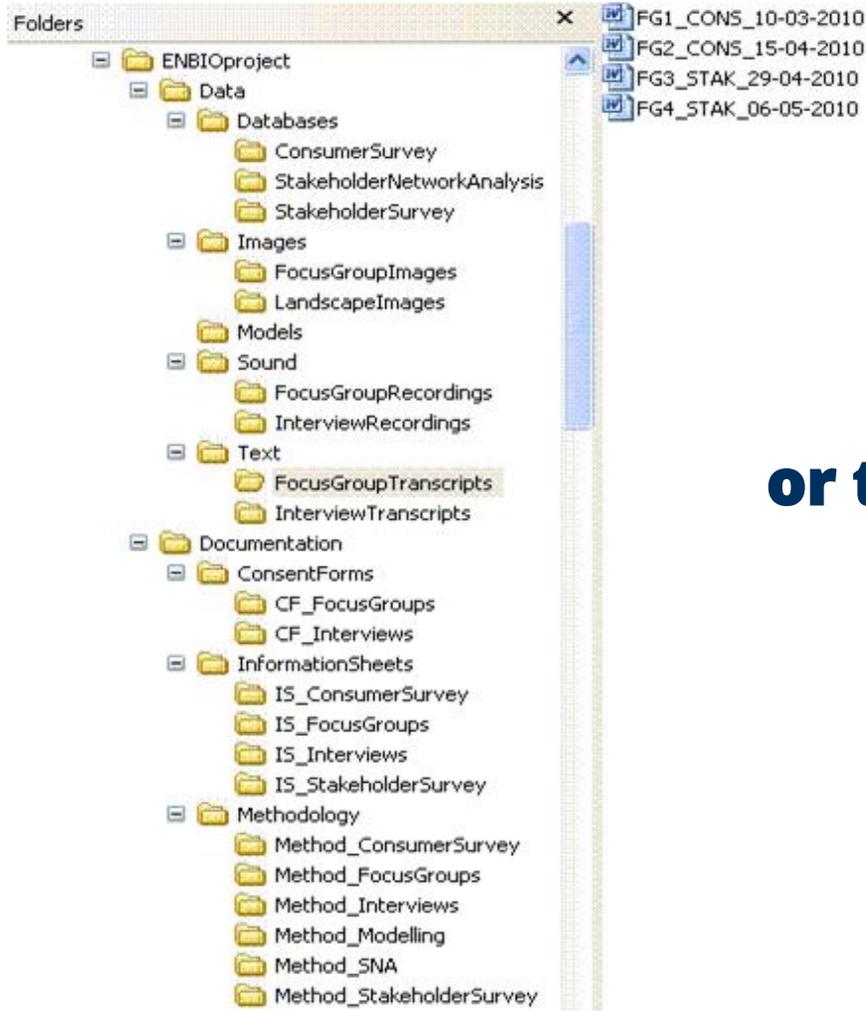
PPTX PPTX PPTX

new.pptx newfinal.pptx newfinalfinal.pptx

PPTX PPTX PPTX

newfinalest final.pptx newfinalest finalforsure.pptx newfinalscrewit thiswilldo.pptx

... but please stick to this:



or this

Name	Kind	Last Opened
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_A01.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_A02.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_A03.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...e-100-1MutantFraction_B01.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...e-100-1MutantFraction_B02.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...e-100-1MutantFraction_B03.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_C01.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_C02.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_C03.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_D01.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_D02.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_D03.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...e-100-1MutantFraction_E01.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...e-100-1MutantFraction_E02.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...e-100-1MutantFraction_E03.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...e-100-1MutantFraction_F01.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...e-100-1MutantFraction_F02.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...e-100-1MutantFraction_F03.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_G01.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_G02.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_G03.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_H01.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_H02.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...-100-1MutantFraction_H03.csv	comma-separated values	2014-05-08 8:05 PM
2013-06-26_BRAFWTNEGASSAY...0-1MutantFraction_platefile.csv	comma-separated values	2014-05-08 3:06 PM



INCENTIVES!



Increased visibility.

Higher citation of data
and publications.

More recognition in the field.

Easier collaboration.

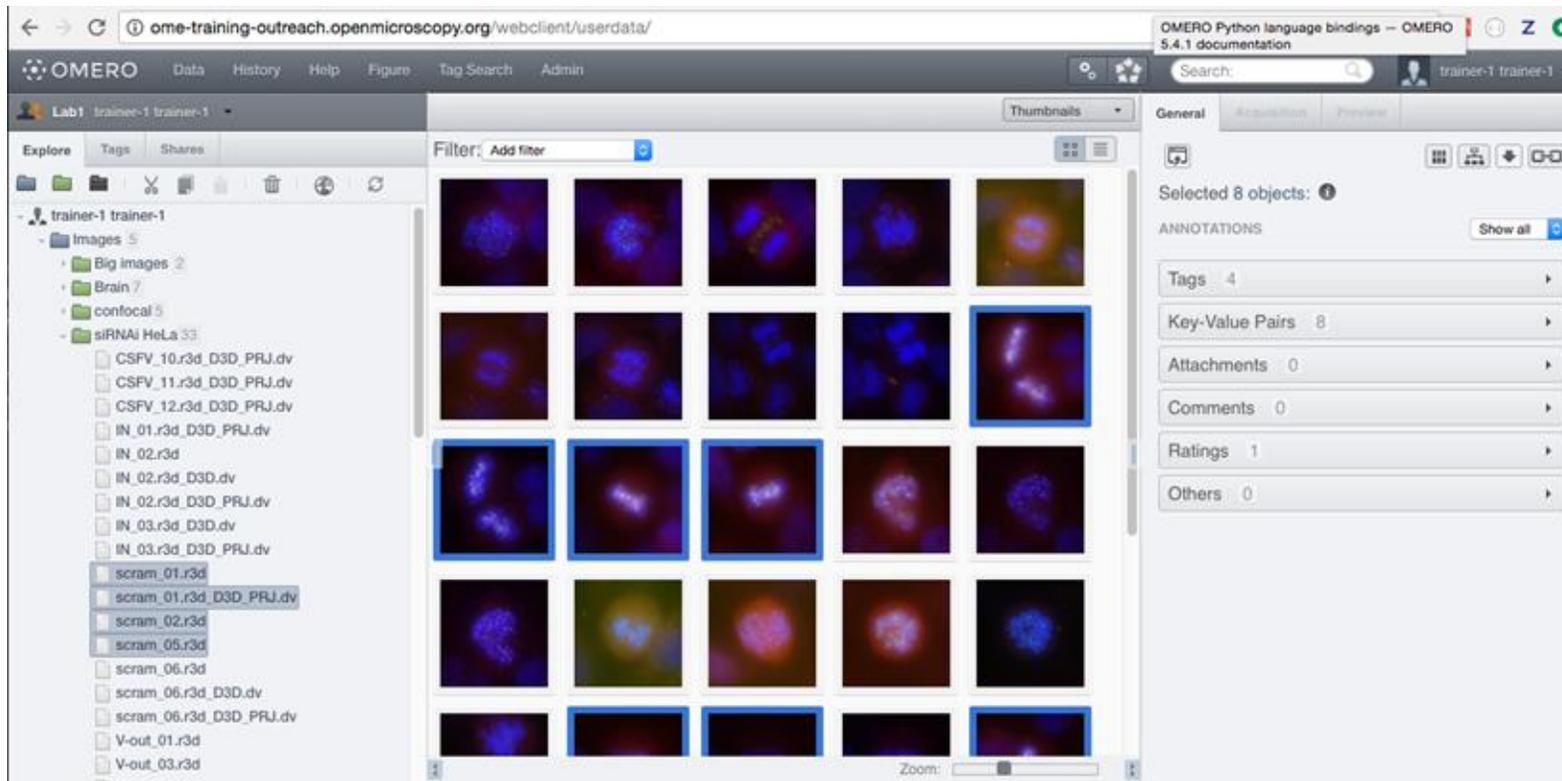
Fostering new discoveries.

Saves time.

RDM TOOLS



= An **open-source** software for **image data** management



<https://omero-guides.readthedocs.io/projects/introduction/en/latest/data-management.html>

Riccardo Massei

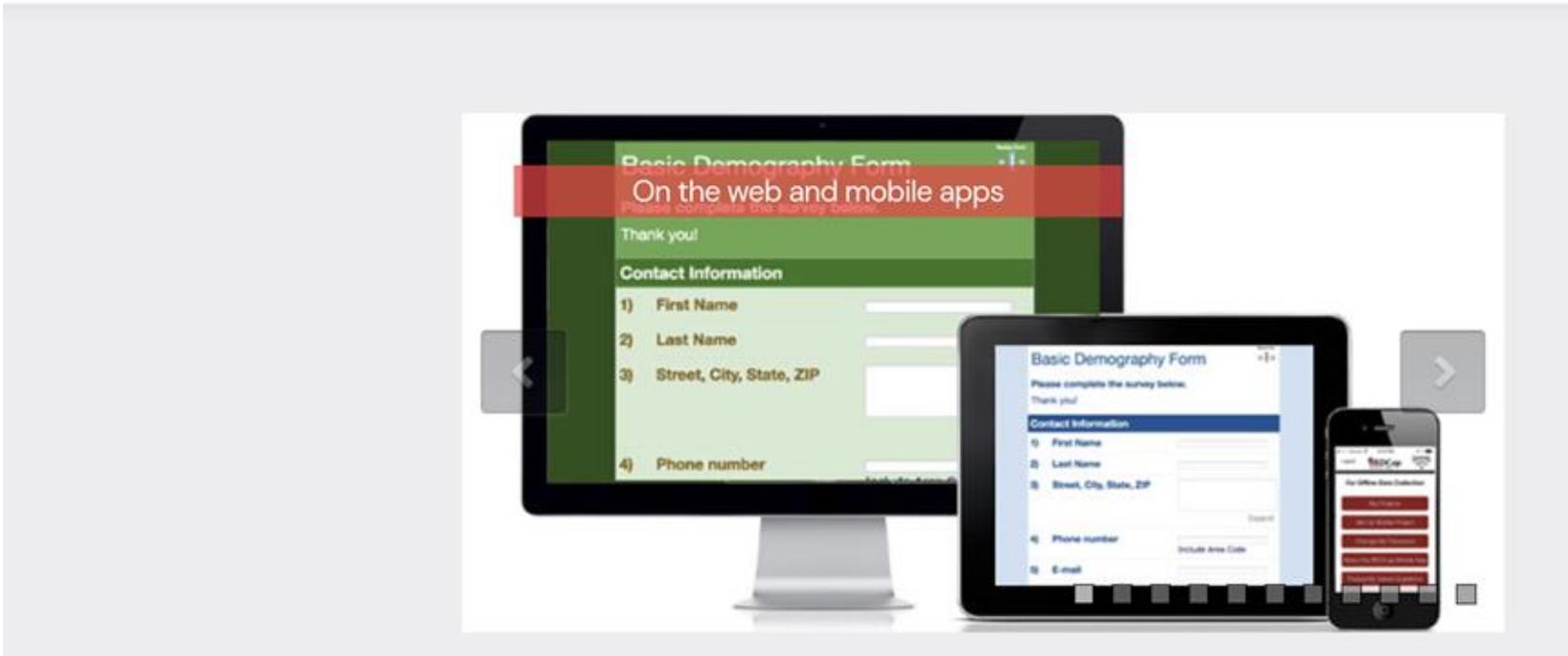


<https://zenodo.org/records/16880913>

Tutorials & Help

REDCap

= A **secure** web application for building and managing **survey data**



Christof Meigen



Contact & Help

Further Reading

Guides

- RDM-Guide (Uni Leipzig):
https://www.uni-leipzig.de/fileadmin/ul/Dokumente/190424_Grunds%C3%A4tze_Forschungsdatenmanagement.pdf
- Medical RDM-Guide (DFG):
<https://zenodo.org/records/4912388#.YPazF5dxfcu>

Contact Points

- RDM-Services (Uni Leipzig)
<https://www.uni-leipzig.de/forschung/forschungsservice/forschungsdatenmanagement>
- RDM-Services (Saxony)
<https://saxfdm.de/>
- RDM-Services (Germany)
<https://www.nfdi.de/consortia/>

Take Home Messages

If you want to make your data *findable*:

- use **unique identifiers (DOIs), rich metadata, and searchable repositories.**

If you want to make your data *accessible*:

- use **open or controlled-access repositories and clear retrieval protocols.**

If you want to make your data *interoperable*:

- use **standard formats, vocabularies, and ontologies compatible with other datasets.**

If you want to make your data *reusable*:

- use **detailed documentation, clear licensing, sufficient metadata for others to understand and use it.**

Keep FAIR in mind!

Finding Data Repositories

- re3data.org
- for storing data but also useful for finding new data

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Filter for discipline
+ DOI assignment!

Registering data in a Repo \neq Backup!



Archiving

vs.



Backup



long-term preservation for **10 years** (DFG).

data **recovery** after **loss** or **corruption**.
usually **three copies** stored in **separate** locations.

Acknowledgments



UNIVERSITÄT
LEIPZIG



Gefördert durch:



Bundesministerium
für Forschung, Technologie
und Raumfahrt

SACHSEN



Diese Maßnahme wird gefördert durch die Bundesregierung aufgrund eines Beschlusses des Deutschen Bundestages. Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.