

Research Data Management (RDM)

February 24./25. 2026



BERGISCHE
UNIVERSITÄT
WUPPERTAL

Photo: Sebastian Jarych

Topics

Basics: What are research data, data life cycle, benefit of RDM

How can I find research data?

Persistent identifiers: DOI, URN, ORCID, ROR

Licences for research data

Requirements of our university and DFG

Data management plan, RDMO

Research data repository PublicData.fdm.uni-wuppertal.de

Technical quality assurance

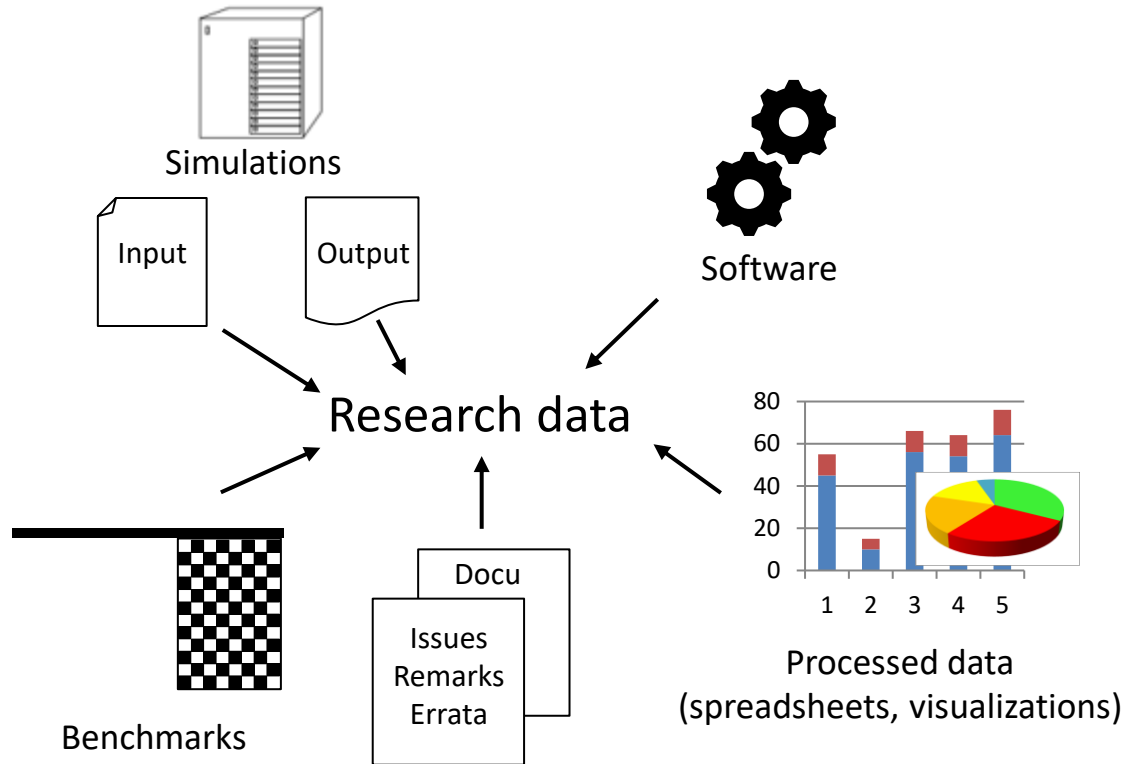
Basics

Research data management
Torsten Rathmann



BERGISCHE
UNIVERSITÄT
WUPPERTAL

What are research data? From DFG Checklist:



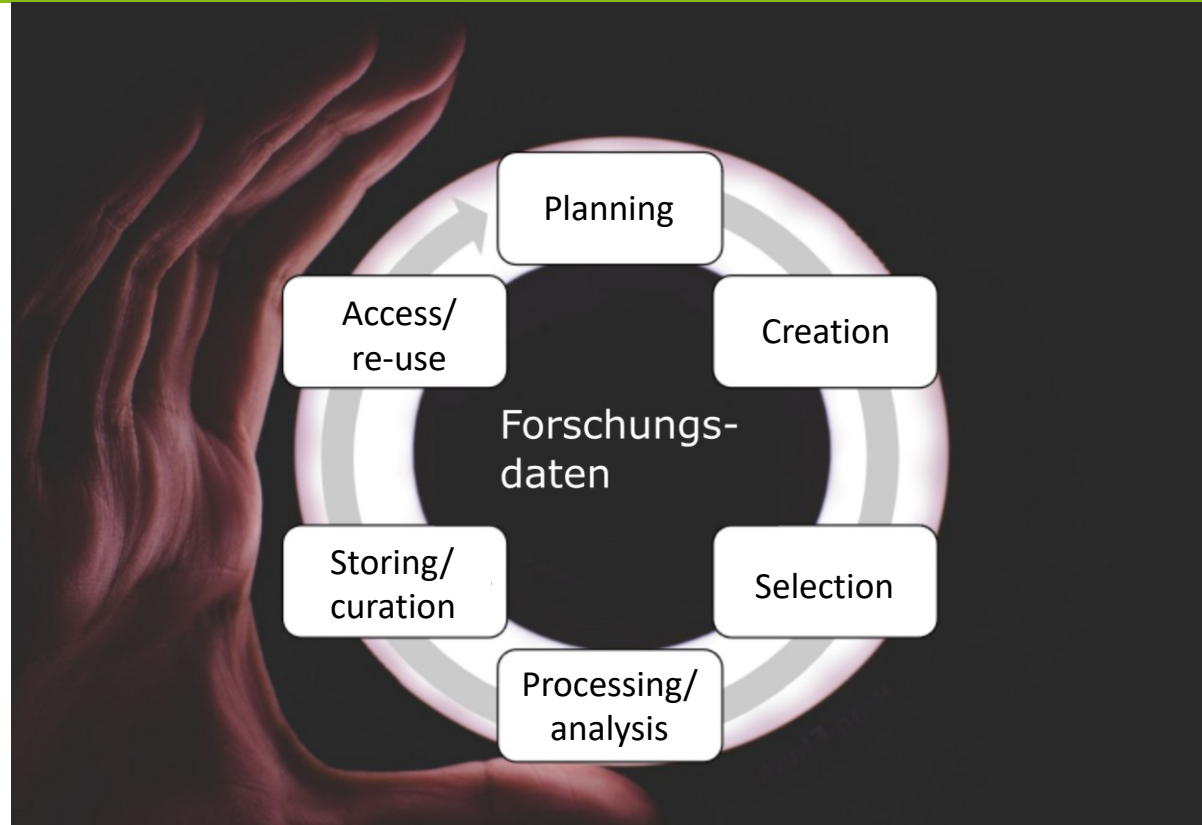
Research data includes ... texts, ..., methodological test procedures ...

Compilations, software and simulations can equally represent a central result of scientific research and are therefore also included under the term research data.

Usually not research data:

- Papers
- Application

Assemble a six-stage lifecycle for research data.

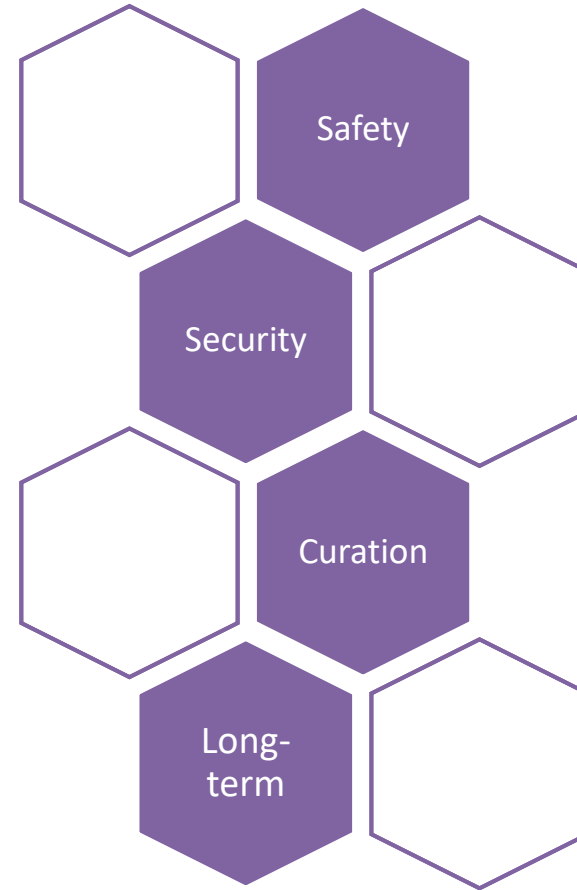


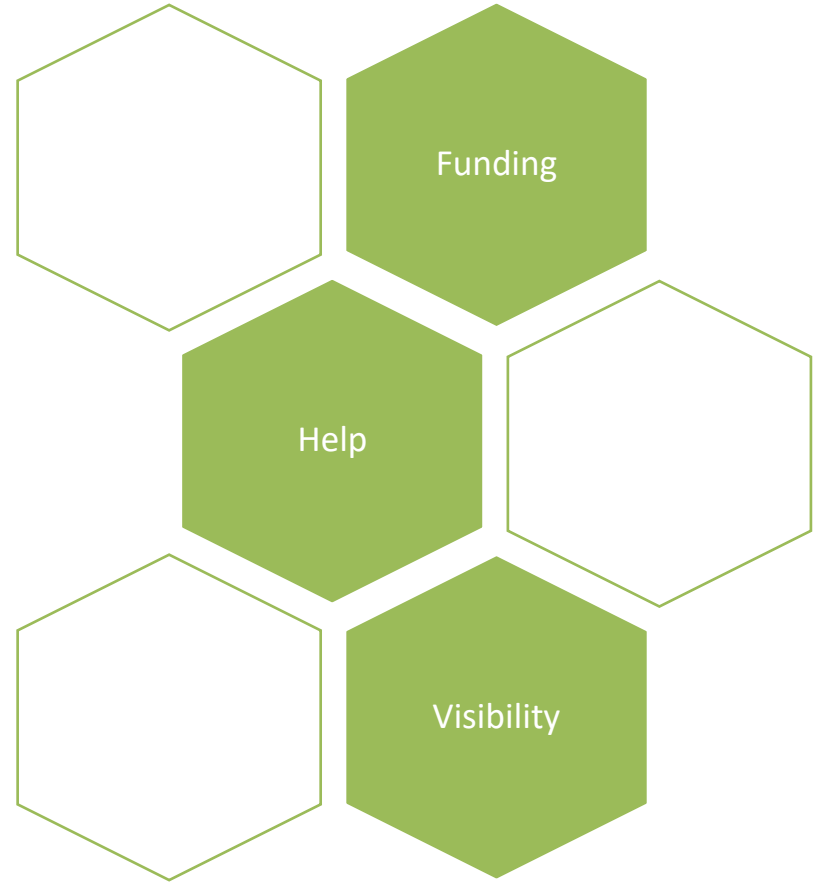
Why research data management?





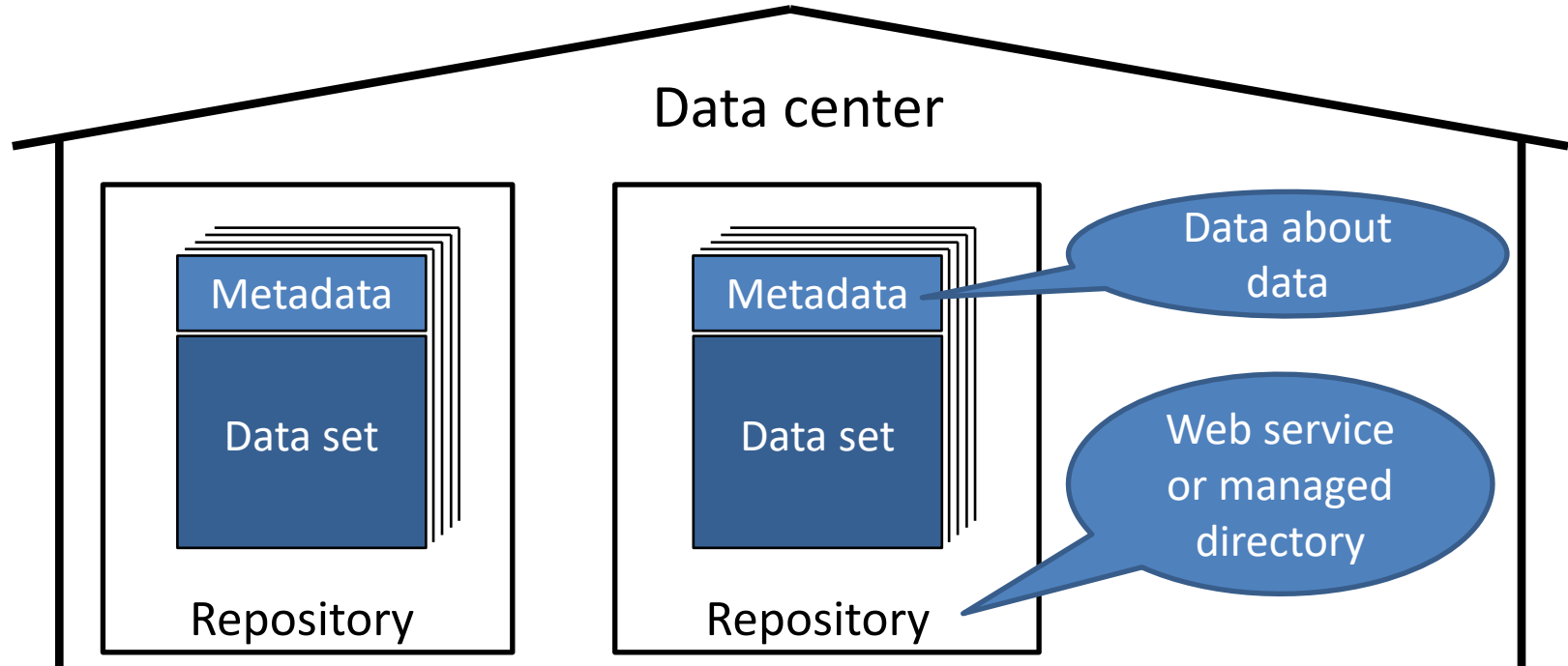
Research data management
Torsten Rathmann





Where can I find research data?

Where are publicly accessible research data stored?



How can I find research data?

Link in monograph or paper

Two-stage: a) search for repositories, b) search in repository

- re3data.org (registry of research data repositories)
- resources.dfg.de

Direct search for data sets

- Google Dataset Search, <https://datasetsearch.research.google.com/>
- Other search engines
- DataCite search, <https://search.datacite.org/>

Persistent Identifiers

Part I: Persistent identifiers for digital
objects: URN, DOI

Part II: Persistent identifier for researchers:
ORCID

Part III: Persistent identifier for institutions:
ROR

Page not found

- ▶ A URL (Uniform Resource Locator) describes the location of a resource.
Problem: The locations of digital objects can change, domains are frequently abandoned, and URLs are not redirected.
- ▶ **URLs are therefore considered unreliable.**

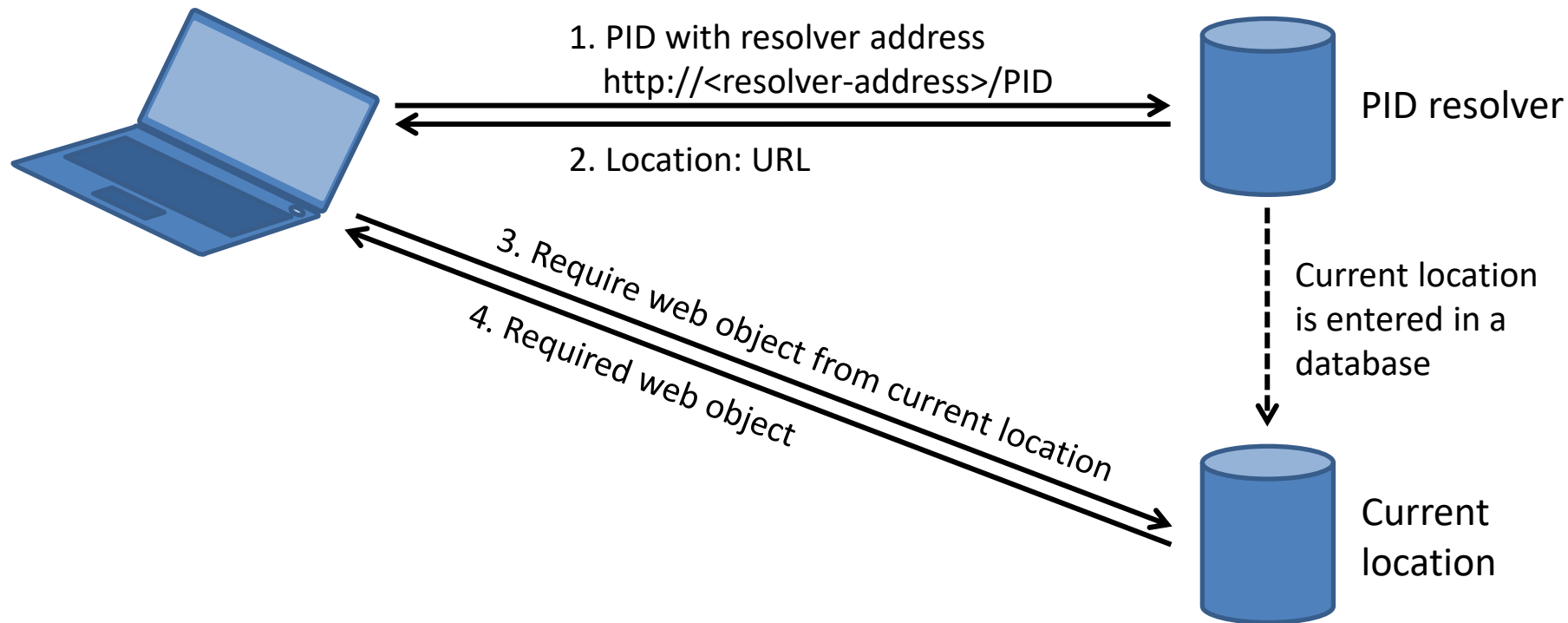
Properties of PIDs

- exist globally only one-time, i.e. unique
- permanently valid
- via software/database („resolver“) resolvable, leads to the current storage location of the object

Advantages

- **unique** and **permanent identification** of electronically published research results (e.g. journal articles, research data)
- ensuring their permanent **findability** and **citability**

Principle of PID systems



URN (Uniform Resource Name)

Example for a URN of German National Library

https://nbn-resolving.org/urn:nbn:de:hbz:468-20110509-151022-7

URL of the resolver
(for D und CH)

Name space

Sub-spaces

Object denomination
(unique number)

Check digit

- ▶ Important especially for dissertations
- ▶ Metadata of the objects not in the PID system

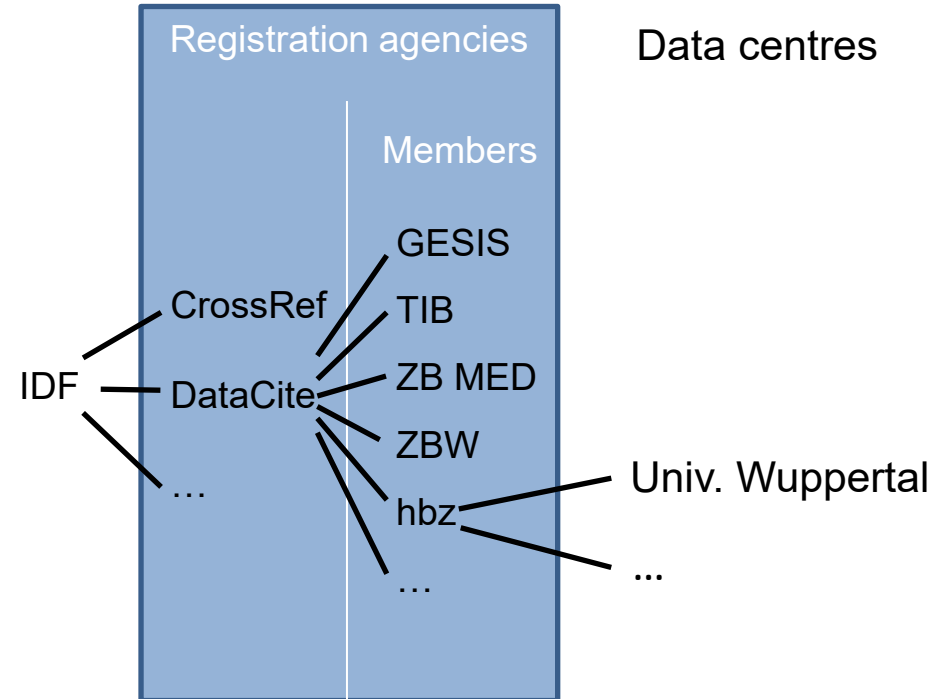
`https://doi.org/10.5281/zenodo.1230525`

The diagram illustrates the components of the DOI URL `https://doi.org/10.5281/zenodo.1230525`. It features three horizontal blue brackets below the URL segments. The first bracket, under `https://doi.org/`, is labeled 'Resolver address'. The second bracket, under `10.5281/`, is labeled 'Organization'. The third bracket, under `zenodo.1230525`, is labeled 'Object denomination'. Above the URL, two light blue boxes are positioned: 'Prefix' is centered over `10.5281/`, and 'Suffix' is centered over `zenodo.1230525`. Vertical green lines connect the labels to their respective brackets.

Resolver address Organization Object denomination

DOI (Digital Object Identifier)

- ▶ Operated by the International DOI Foundation (IDF), founded in 1997
- ▶ Currently 10 registration agencies, including
 - CrossRef (Focus on publisher publications: journal articles, E-Books)
 - DataCite (Focus on research data)
- ▶ DOI registration is only possible with the submission of metadata



Differences and similarities

URN	DOI
Provided by the DNB free of charge	The university will cover the costs.
Must not be deleted	
A valid URL points ...	
• directly to the object or • a front page (landing page)	... to a front page (landing page)
Only conditionally suitable for research data	

Each URN/DOI partner commits to the ongoing and continuous maintenance of its URNs/DOIs.

Part I: Persistent identifiers for digital
objects: URN, DOI

Part II: Persistent identifier for researchers:
ORCID

Part III: Persistent identifier for institutions:
ROR

Why are personal names not sufficient?

Several Authors can have the same name.

Names can change (e.g., upon marriage).

Various spellings of an author's name:

e.g. Karl Peter Müller

- Abbreviations of first names: K. P. Müller, Karl P. Müller
- Umlauts are resolved to “u” or “ue”

e.g. Сергей Королёв (russ.), Сергій Корольов (ukrain.),
lat. spelling depend on transliteration table



Connecting Research
and Researchers

Torsten Rathmann  <https://orcid.org/0000-0001-5880-1546>

- ▶ Non-profit organization, founded 2010, launch 2012, 11 Mill. yearly active  2025
- ▶ De facto standard, already mandatory at some publishers and specialist events
- ▶ Additional data (career, list of publications) can be entered
- ▶ Data sovereignty lies with the researcher.

*The University of Wuppertal
recommends that all authors create
an ORCID iD for the unambiguous
identification of their authorship.*

Amtliche Mitteilungen, Jg. 48, Nr. 10 vom 13.03.2019



Part I: Persistent identifiers for digital
objects: URN, DOI

Part II: Persistent identifier for researchers:
ORCID

Part III: Persistent identifier for institutions:
ROR

Open, persistent identifier for organisations: Research Organization Registry (ROR)

[ABOUT](#)[REGISTRY](#)[COMMUNITY](#)[BLOG](#)[DOCUMENTATION](#)[ROR search home page](#)

 <https://ror.org/00613ak93>

University of Wuppertal

ORGANIZATION TYPE

Education

OTHER NAMES

BUW, Bergische Universität Wuppertal

WEBSITE

<http://www.uni-wuppertal.de/en/home.html>

LOCATION

Wuppertal (GeoNames ID [2805753](#))

Germany

OTHER IDENTIFIERS

GRID [grid.7787.f](#)

ISNI [0000 0001 2364 5811](#)

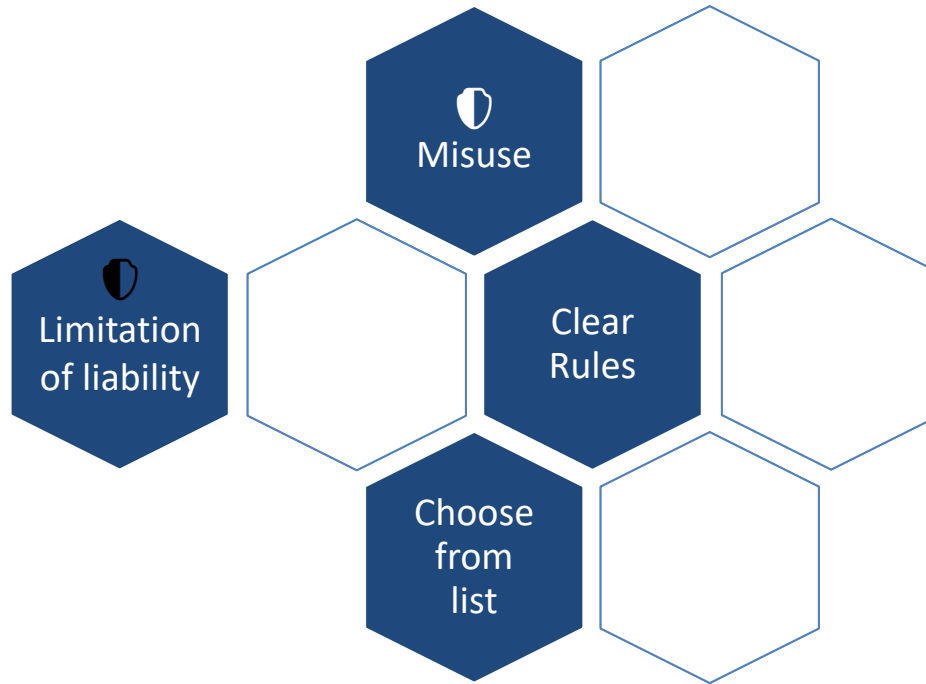
Crossref Funder ID [100012325](#)

Wikidata [Q447953](#)

Is there an issue with the data on this record? [Submit a curation request](#)

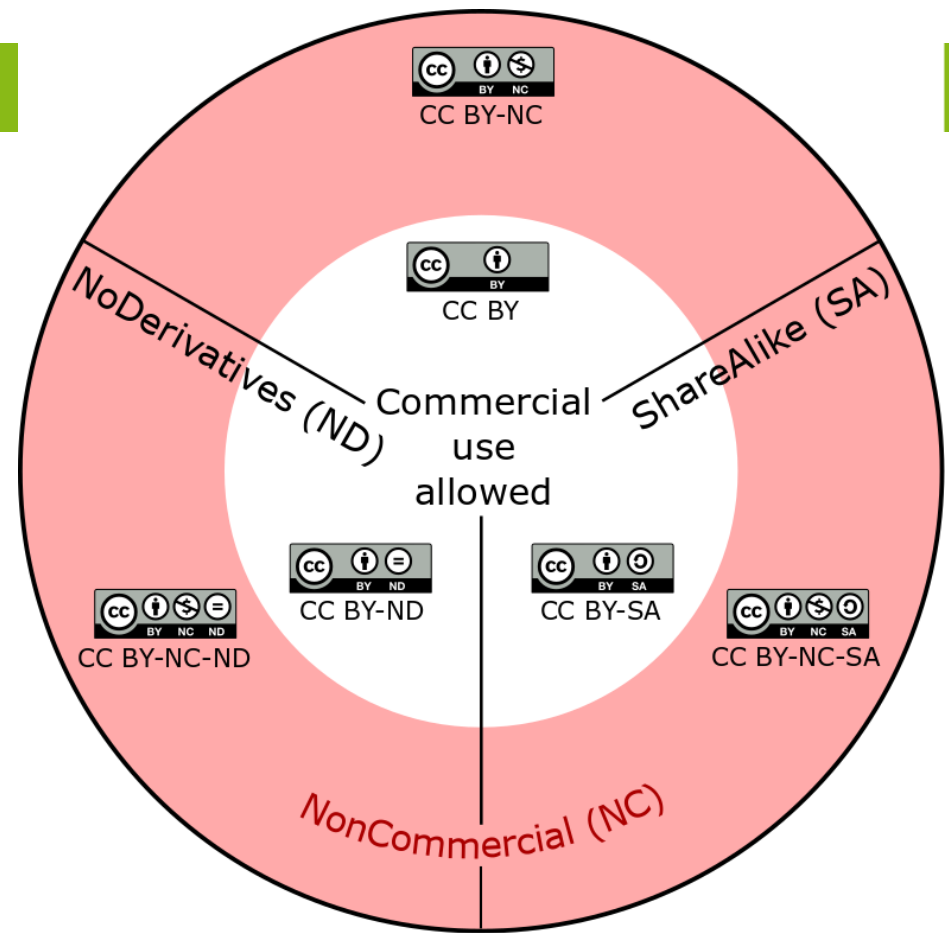
Licenses for research data

Why licenses for research data that reach the level of originality ("Schöpfungshöhe")?



Creative Commons (CC)

- Widely used
- Free of charge
- For all kinds of objects except software
- Freely useable (under CC0)
- Several different licenses
 - CC0 most extensive waiver of rights
 - CC BY: re-user of the work must provide
 - the name of the creator and attribution parties,
 - a copyright notice,
 - a license notice,
 - a disclaimer notice and
 - a link to the licensed original.

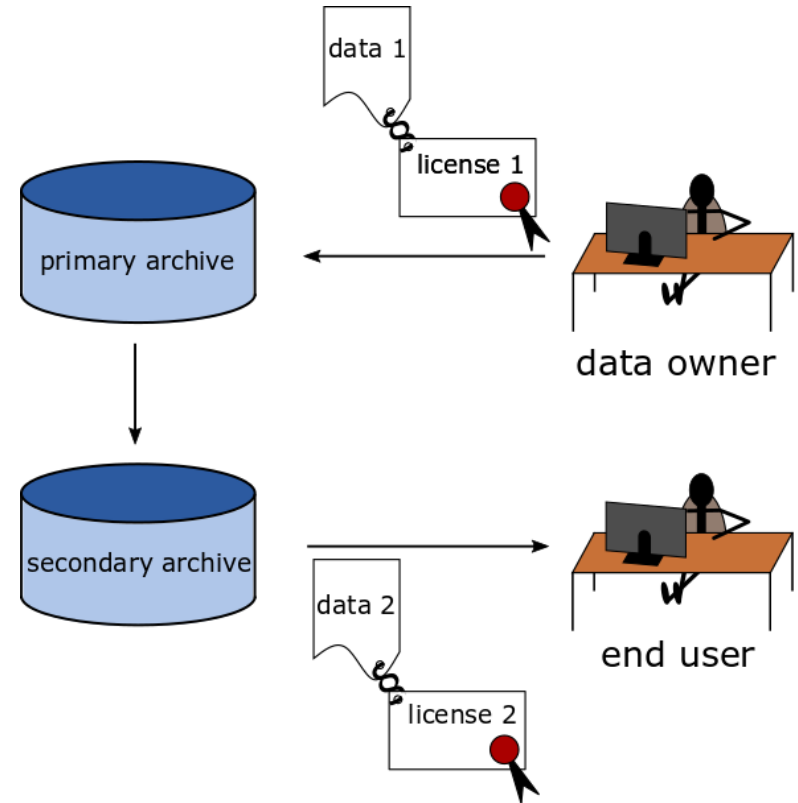


ShareAlike (SA)

Re-publication: license 1 = license 2

NoDerivatives (ND)

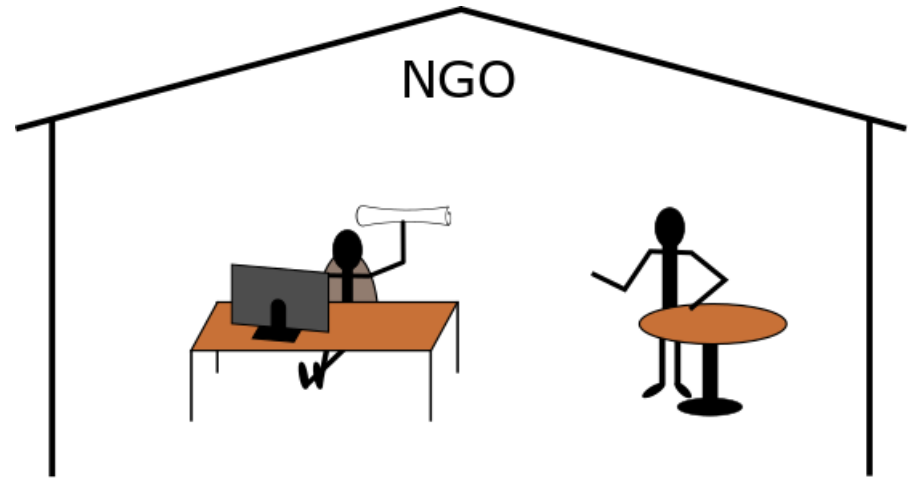
- data 1 $\approx$ data 2
- Serious restriction!
 - Derivates are important in science.
 - Does a correction already reach the level of originality?



What do NonCommercial (NC) mean?

"NonCommercial means not primarily intended for or directed towards commercial advantage or monetary compensation."

Issue:
Non Government Organizations
may be excluded too.



Licenses for software (examples)

GNU General Public License (GPL)	Source code open, Copyleft ©
GNU <i>Lesser</i> General Public License (LGPL)	as GPL, but no Copyleft, if re-use is only library call
Apache license	• Source code need not to be open. • Apache license must be included, but custom software does not need to be under the Apache license.

Overview of numerous further software licenses:

<https://www.gnu.org/licenses/license-list#SoftwareLicenses>

Open Data Commons (ODC) licenses for databases

Created to cover database rights (according to § 87b of the German Copyright Act, separate rights of reproduction and distribution for databases), because CC licenses only protect databases from version 4.0 onwards.

Public Domain Dedication and License (PDDL)

Most extensive waiver of rights similar to CC0

Open Data Commons Attribution License (ODC-By)

Similar to CC BY

Open Database License (ODC-ODbL)

- Similar to CC BY-SA
- Most important user: OpenStreetMap

What are the minimum requirements for a research data license?

Granting of a non-exclusive right of use

Disclaimer of warranty

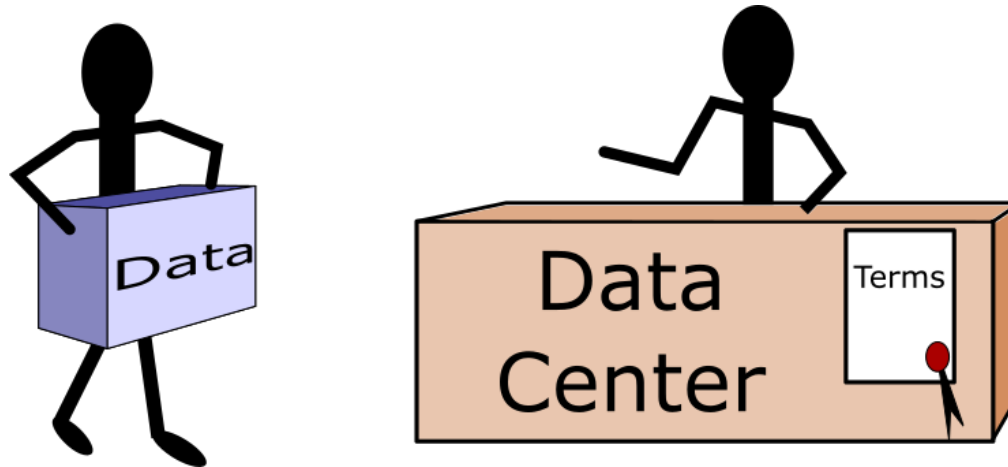
Limitation of liability

Citation obligation

Rules for databases

Attention!

Your license must be compatible with the terms of use of your repository.



Offers and requirements of our University

„Grundsätze im Umgang mit Forschungsdaten an der Bergischen Universität Wuppertal“

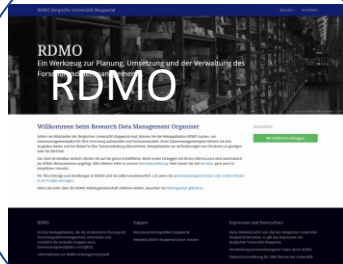
(Amtliche Mitteilungen Nr. 93 vom 27.08.2015)

- “Project leaders and independent researchers are generally responsible for the research data management of their research projects. In particular, they are obligated to ensure compliance with good scientific practice and disciplinary standards.”
- Storage and archiving in the IT infrastructure of the University of Wuppertal (BUW) or in subject-specific repositories (taking into account data protection requirements)
- Retention period: at least ten years

Offers of the Service Centre Research Data Management

Consulting
+
Training

Website



Moodle course



Repository



DFG

Requirements of DFG concerning research data management

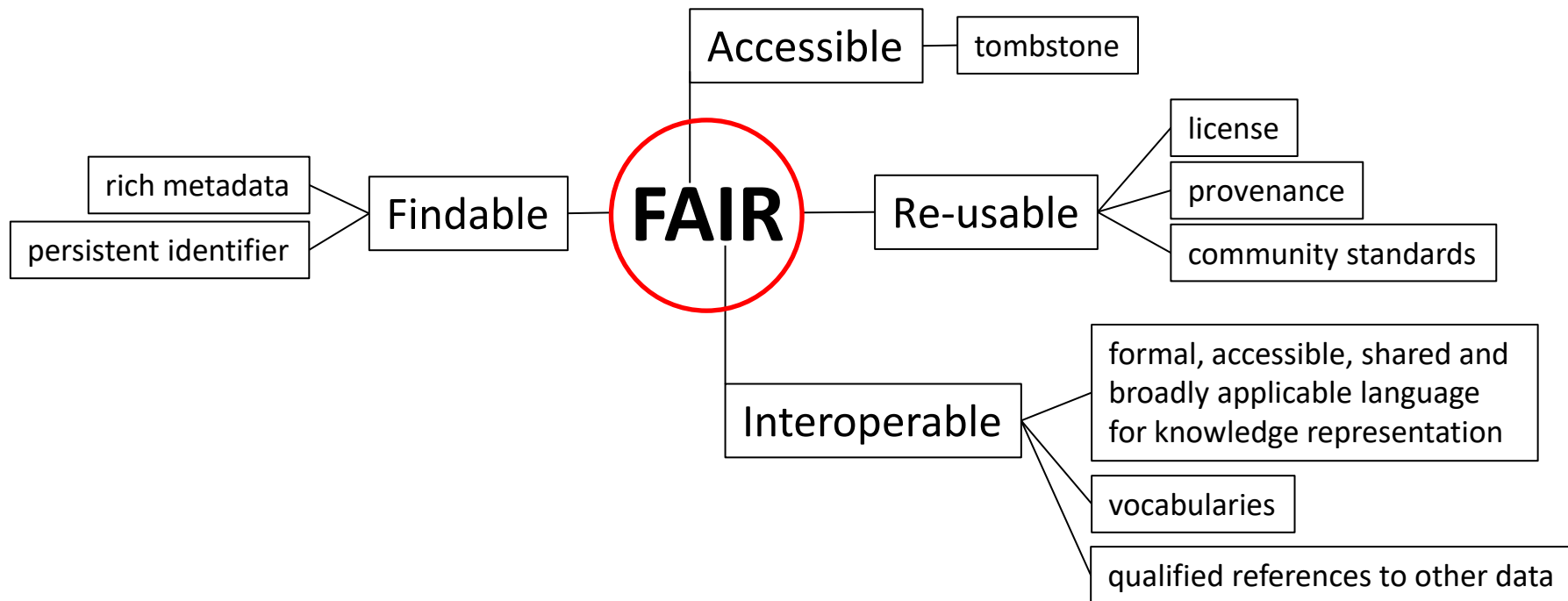


Guidelines for Safeguarding Good Research Practice

Code of Conduct



Guideline	
7	Quality assurance
12	Documentation
13	FAIR
17	Store primary data ≥ 10 years

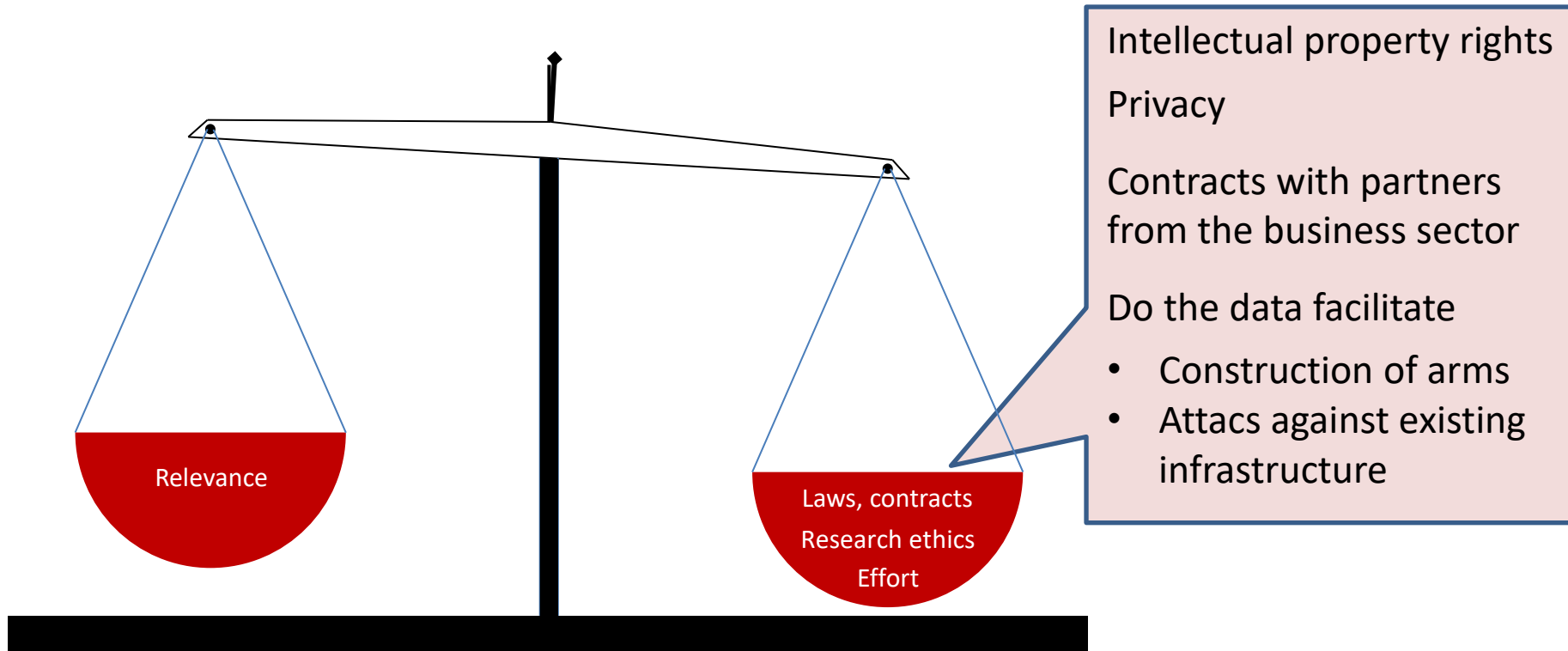


Data human- and machine-readable!

Research data should be made accessible

- as soon as possible
- the publication of research data does not conflict with the rights of third parties (in particular data protection or copyright)
- at a stage of processing that allows it to be usefully reused by third parties (raw data or structured data)

Laws, contracts, and ethics can prevent opening up. Otherwise, weigh the options.



In the application:

- Data types, origin, how processed, data volume
- Documentation, quality assurance, for re-use required software
- Storage and access control while the project is running
- Legal obligations and conditions
- Research codes or professional standards
- Which data sets are especially suitable for use in other contexts?
- Data selection
- Archiving, retention period
- Responsibilities
- Required resources

Data management plan

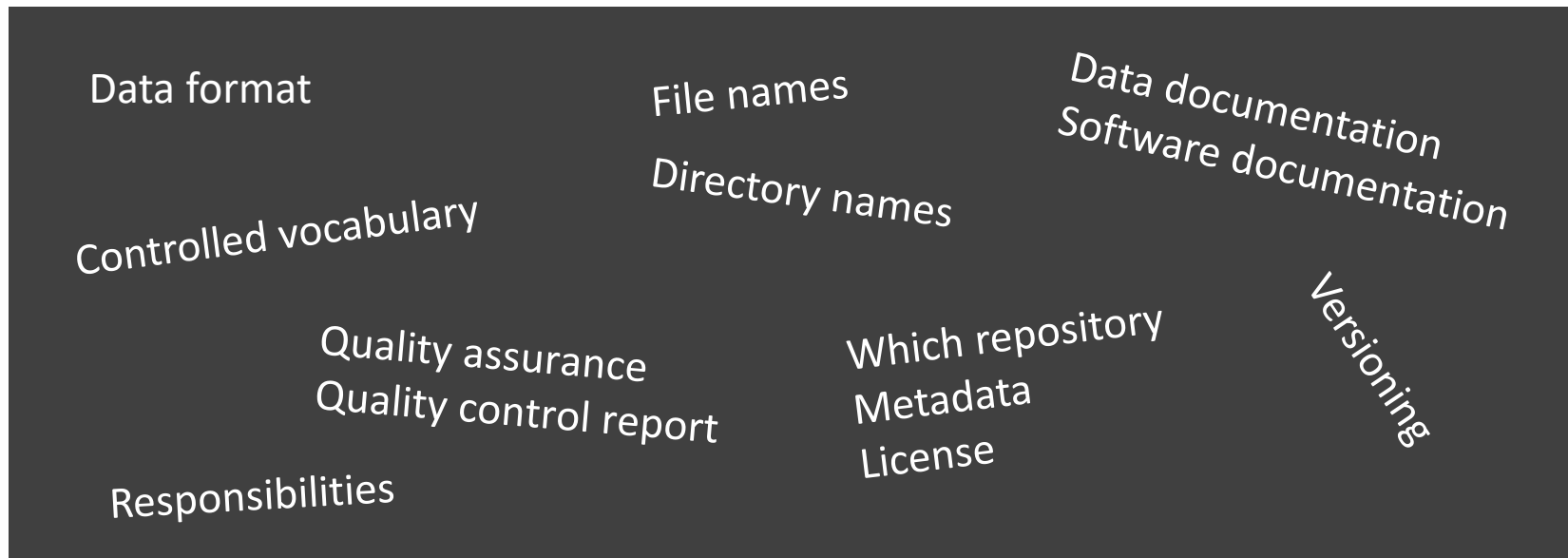
University of Virginia:

A data management plan is a formal document that outlines what you will do with your data during and after you complete your research.

- Compulsory basis for data management
- Facilitates coordination of project partners
- Helps to identify problems and develop solutions.
- Helps to avoid data loss.
- Defines responsibilities
- Regulates access rights

DMP: When should planning be started?

Early, because there are many details to plan:



Name of institution

- Bergische Universität Wuppertal
- BUW
- University of Wuppertal ← ROR-Name

Method

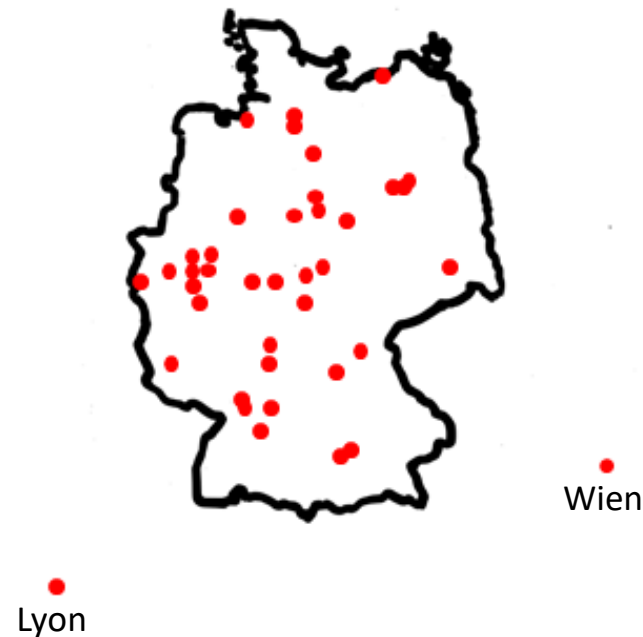
- Implicit midpoint rule
- A4

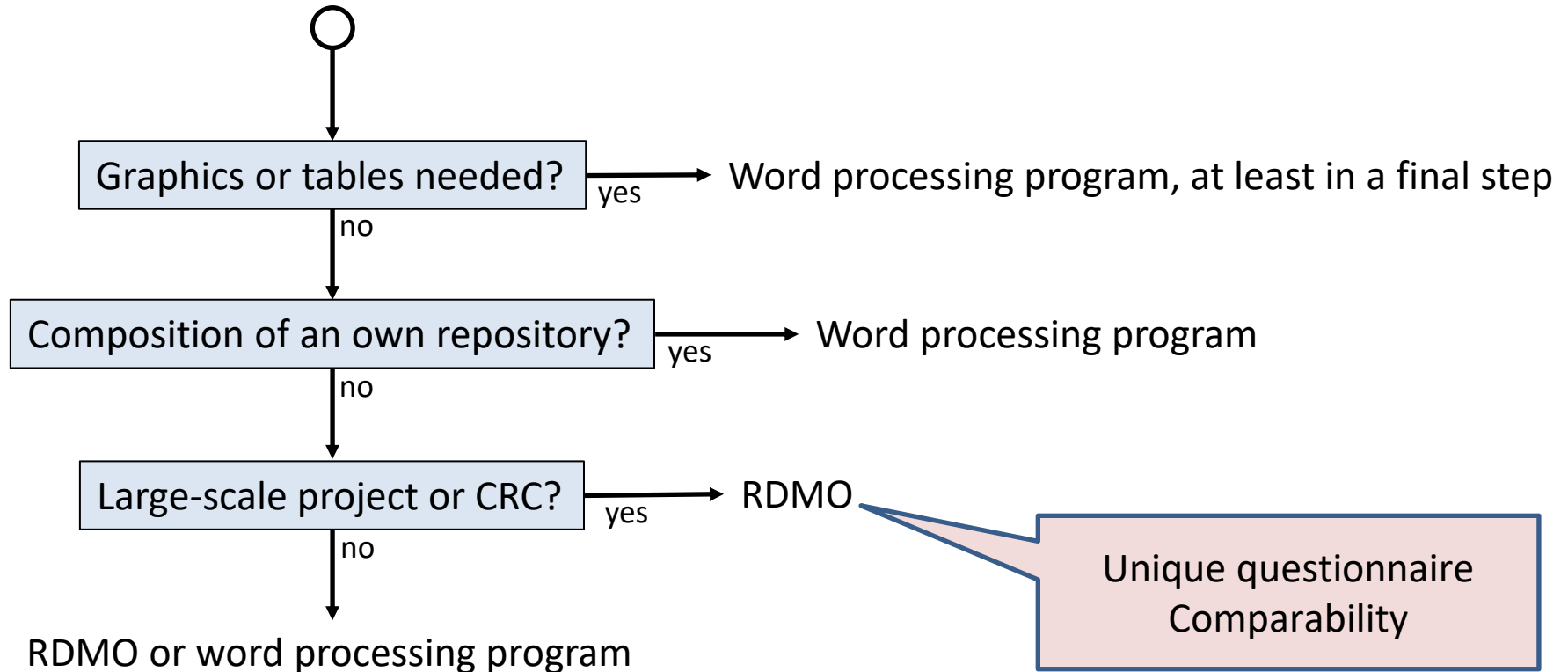
RDMO

Research Data Management Organiser

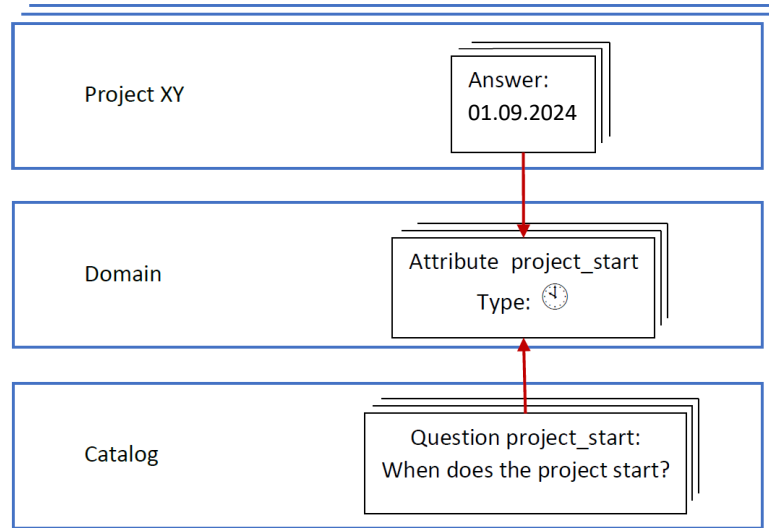
Running at our university and many other research institutions

You can export and re-import your DMP as XML, e.g. at your new work.





RDMO: question answer assignment



Questions and answers are linked via "attributes". Therefore, switching the question catalog is possible.

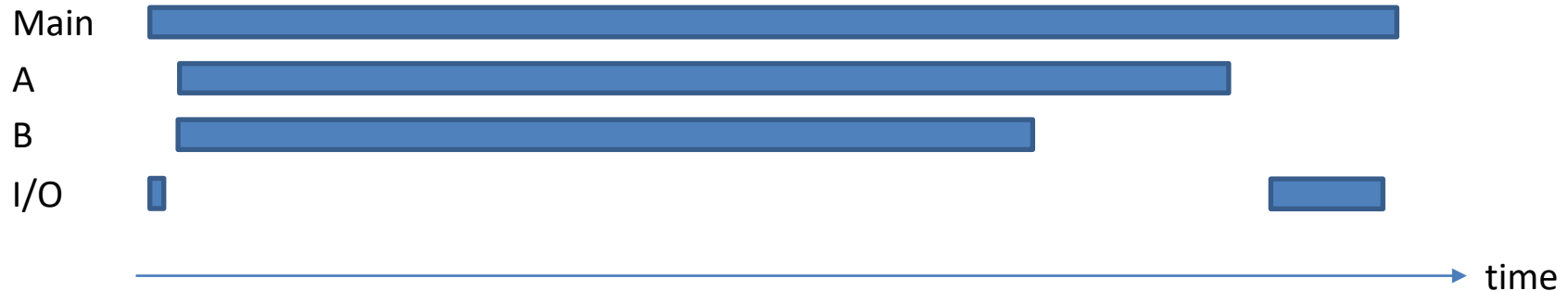
Technical quality assessment

Type	Check
Numerical	Really a number?
	Overflow?
	NaN (Not a Number)?
Integer	Really integer?
Missing value	Use everywhere the same identification marking, e.g. MV
Every type	Is the value allowed? Compare with codebook, e.g. is the value really 1, 2 or 3 for model sub-type

Check for completeness

- All files there, including model input?
- In case you use parallel computing, do processes/threads work in the right sequence?

Frequent error: empty blocks in the output if your I/O process doesn't wait



Research data repository

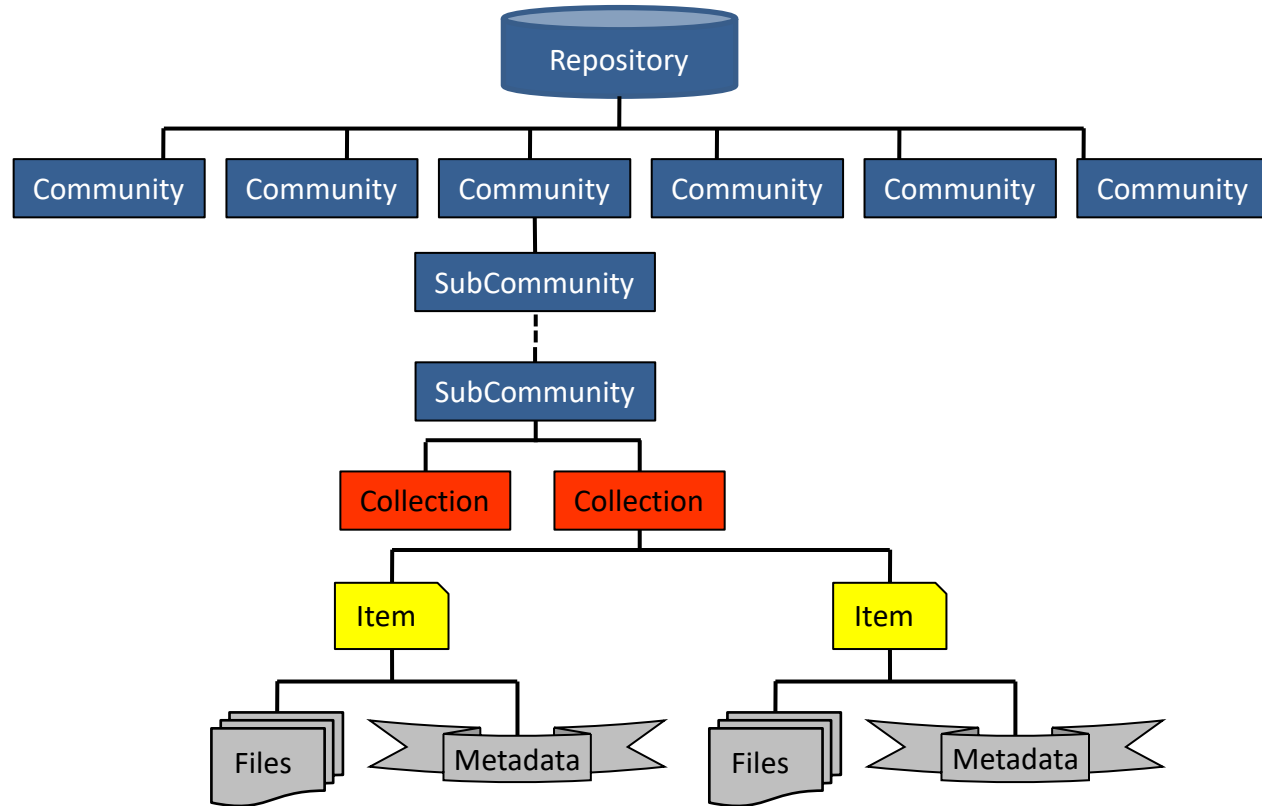
PublicData.fdm.uni-wuppertal.de

PublicData: institutional research data repository of our university

- Publicly accessible under publicdata.fdm.uni-wuppertal.de
- Storage by data creator
- DataCite DOI can be assigned.
- Versioning
- End user agreement (German) public under:

fdm.uni-wuppertal.de/de/langzeitspeichern/repo/speicherungsvereinbarung



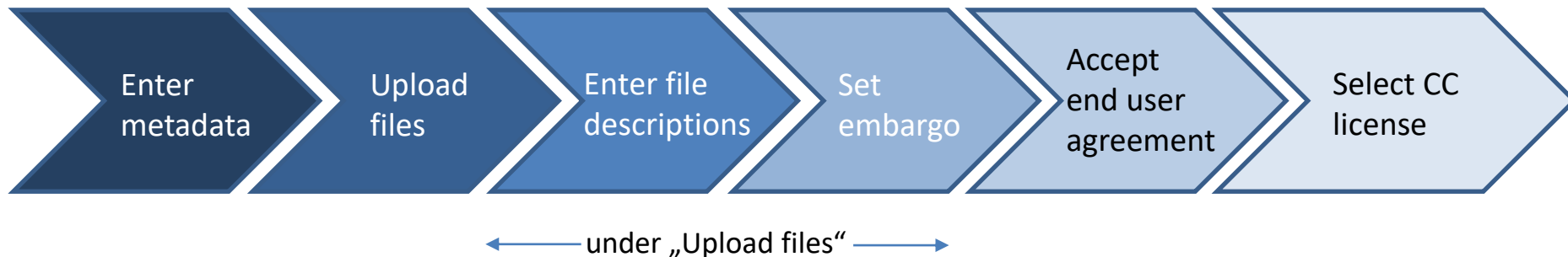


PublicData: prepare for storage



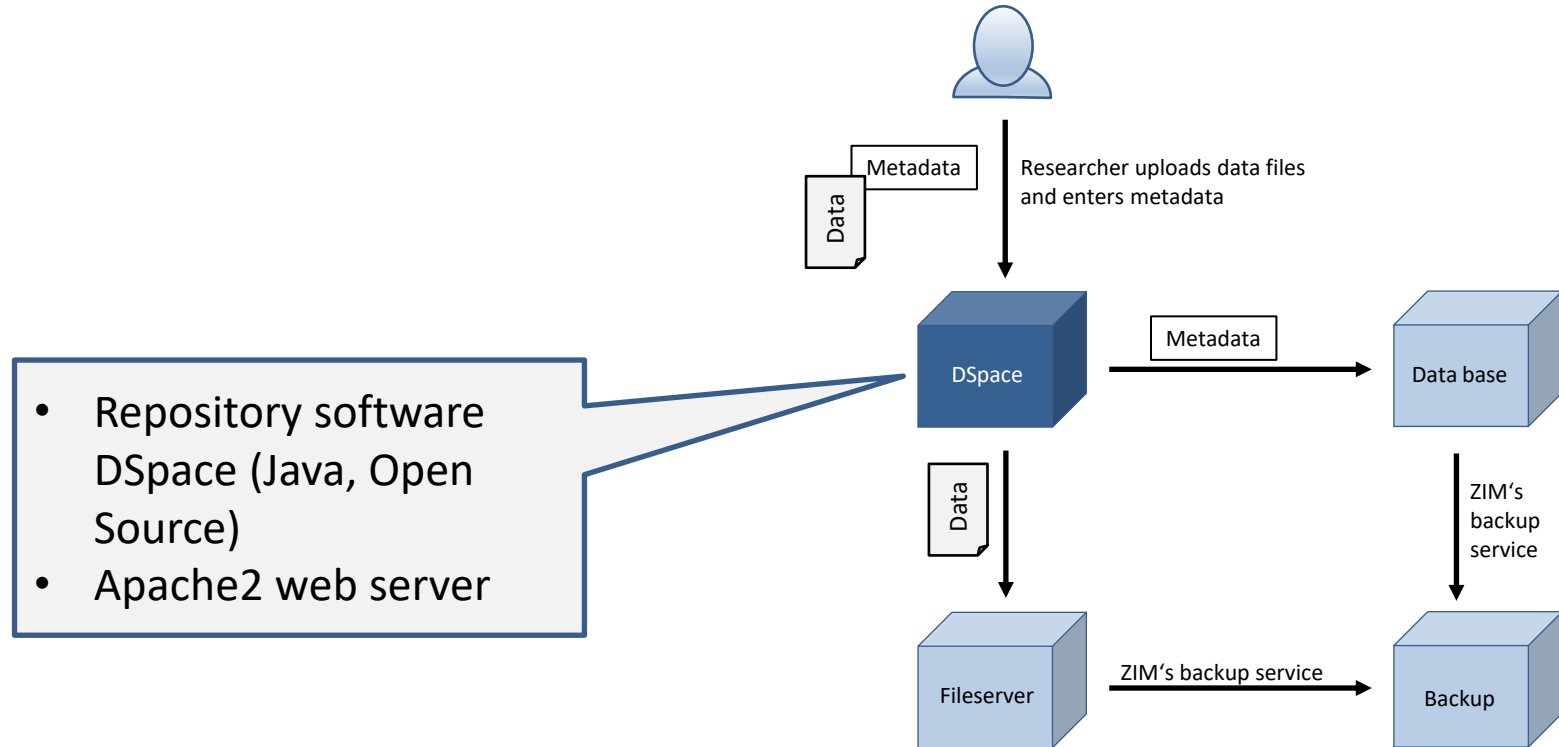
DSpace admin

Submitter



- Maximum size of a single file : 10 GB
- Maximum size of the whole Item: 512 GB
- Compressed files allowed, e.g. *.zip, *.tgz

PublicData: involved services



Metadata (metadata scheme: Dublin Core). Please enter them in English:

Compulsory metadata (metadata scheme: Dublin Core). Please enter in English:

Authors	People who have made significant contributions to this item: name, first name, ORCID iD
Title	Title: should clearly distinguish the item while still being short
Date of Issue	For secondary storage, fill in the date of the initial storage; otherwise, the current date is used.
Type	For research data usually "Dataset"
Language	The most widely used language in the data (not the metadata). If the Item doesn't contain text, e.g. for pure numerical oder image data, please use "N/A".
Subject Keywords	Can be freely assigned
Abstract	Description of the item

+ further optional metadata

Data type	Data format
Text	PDF/A; Markdown, LaTeX, (docx, odt)
Spreadsheet	CSV (Comma-Separated Value)
Raster graphics (photo, screenshot)	TIFF, TIFF-LZW, JPEG2000
Vector graphics	XML/SVG

Text: PDF (Portable Document Format)

Page description language with a focus on document output true to original

Two important specific formats	
PDF/A (PDF for Archive)	for long-term archiving
PDF/X (Xchange)	for pre-press (Druckvorstufe)

PDF/A:

- All fonts and images used in the document are embedded.
- Animations, videos, and other elements that need external software are not embedded.
- Colors as close as possible to the colors in the document
- No encryption

Microsoft Excel

	A	B	C
1	x	y	Importance
2	1	3	
3	1.5	3.5	
4	1	4	important

Export to CSV

Comma Separated Value (CSV)

```
; ; Importance  
; ;  
; ;  
; ; important
```



Your quality control report

- English
- Lists control steps, e.g. “Plausibility and completeness of the data checked through visualization with...”
- At least readability tests: “All files in xyz format can be read with software XYZ.”
- Send it to rathmann@uni-wuppertal.de

Thanks for your attention!



BERGISCHE
UNIVERSITÄT
WUPPERTAL

Photo: Sebastian Jarych