

Data Sharing and Long-Term Preservation

Astrid Recker

CESSDA Training at the Data Archive for the Social Sciences

GESIS - Leibniz Institute for the Social Sciences

@CESSDA_Data

The Data Spectrum

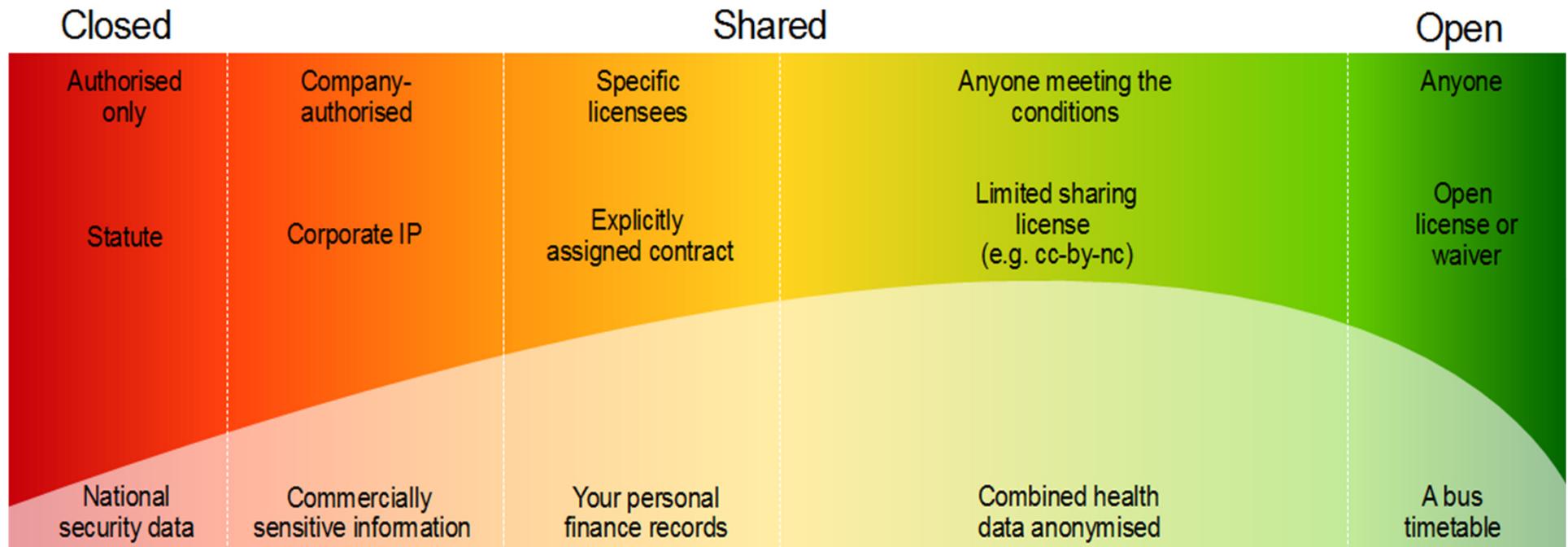


Image: Adapted from <http://theodi.org/data-spectrum> (Open Data Institute, cc-by)



Degrees of Openness and Persistence

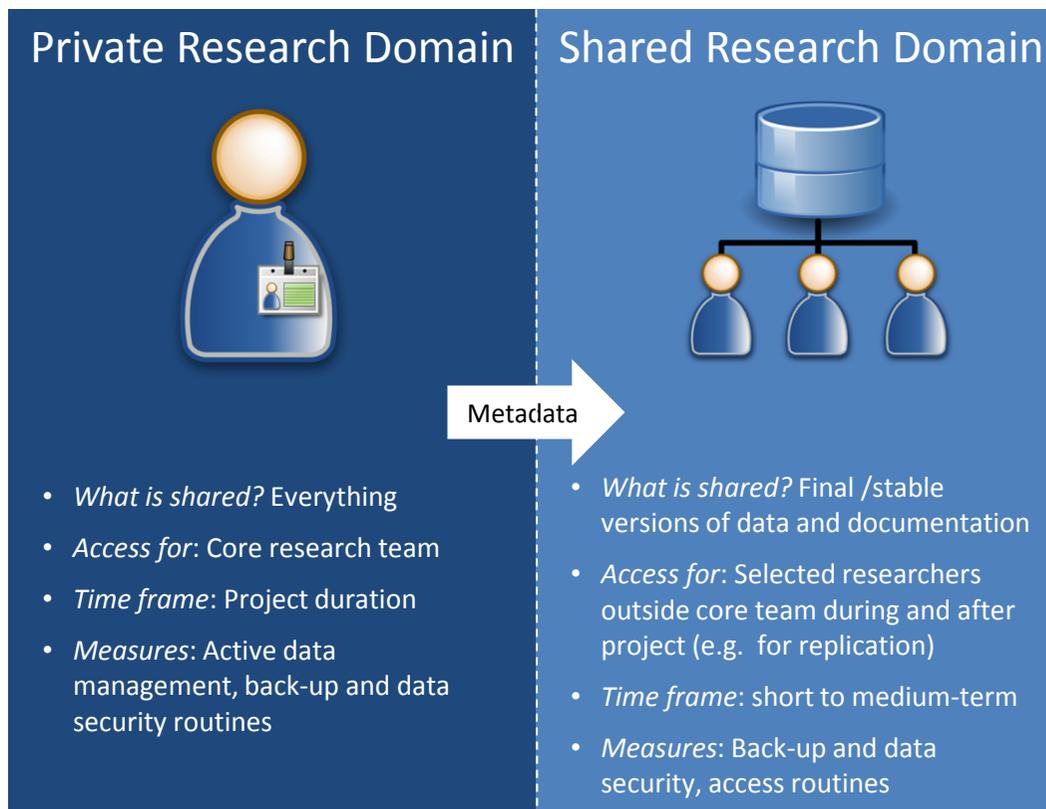
Private Research Domain



- *What is shared?* Everything
- *Access for:* Core research team
- *Time frame:* Project duration
- *Measures:* Active data management, back-up and data security routines

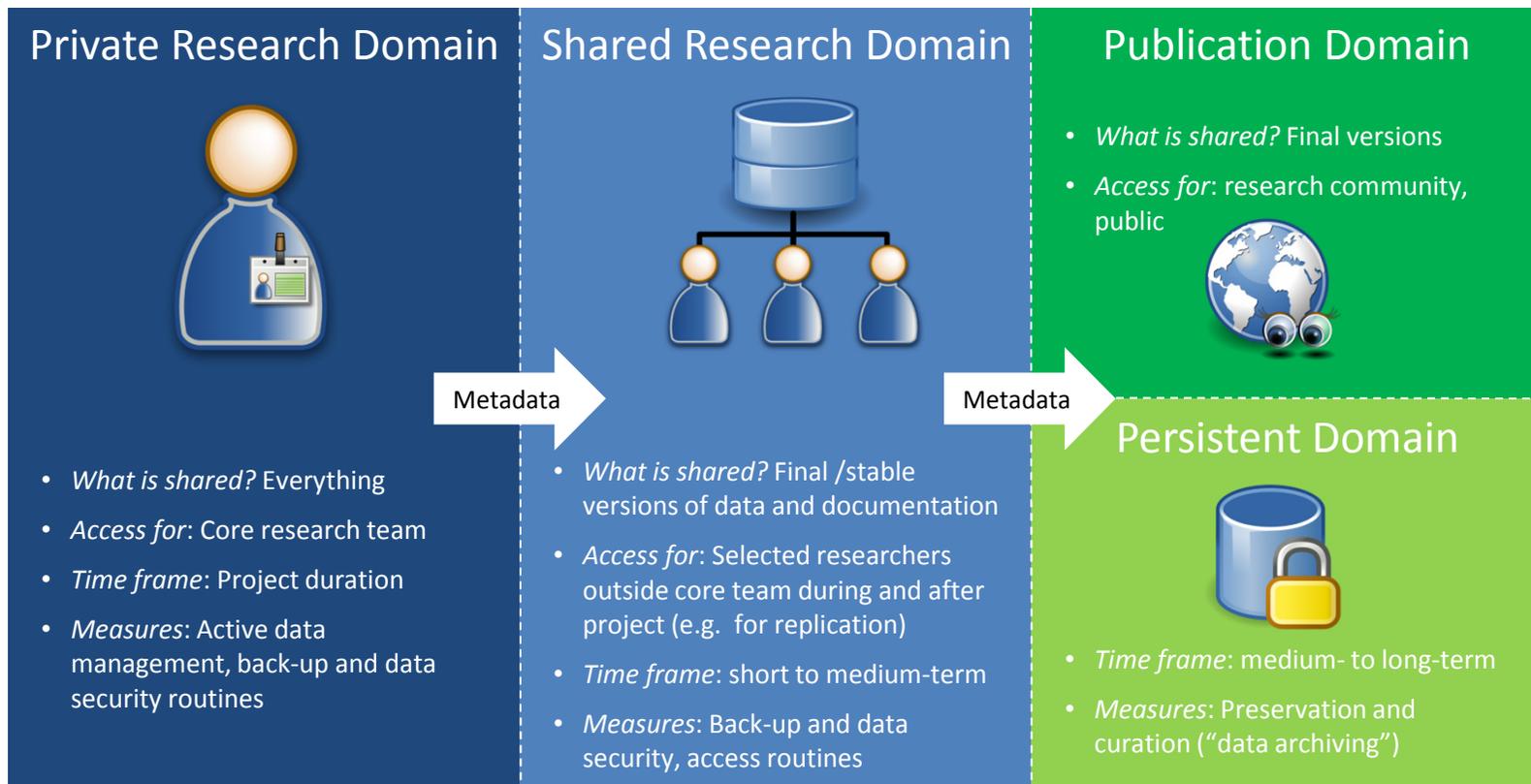
Icons by RRZE Erlangen (CC-by-sa) <https://github.com/RRZE-PP/rrze-icon-set>

Degrees of Openness and Persistence



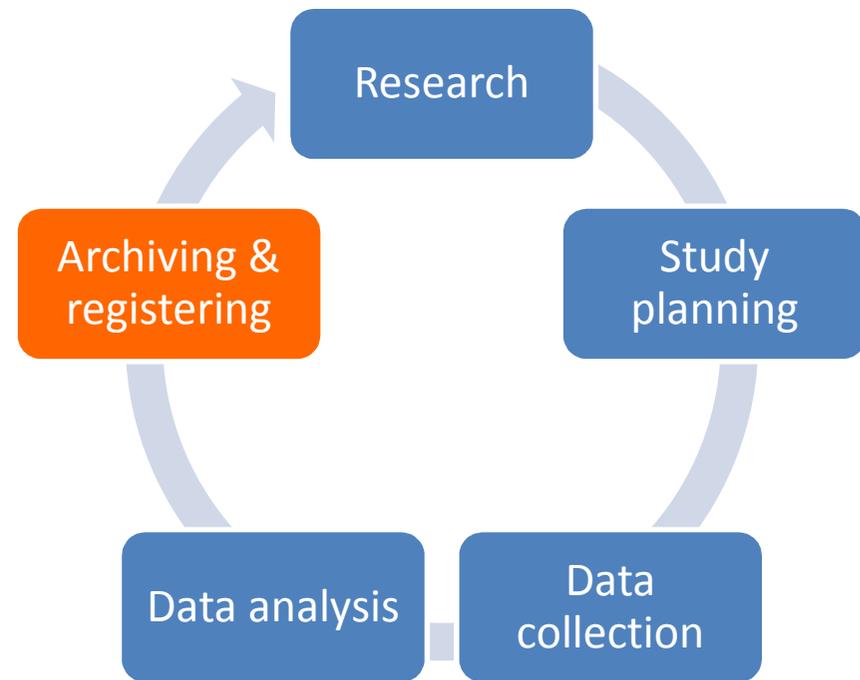
Icons by RRZE Erlangen (CC-by-sa) <https://github.com/RRZE-PP/rrze-icon-set>

Degrees of Openness and Persistence



Icons by RRZE Erlangen (CC-by-sa) <https://github.com/RRZE-PP/rrze-icon-set>

Archives in the Research Process



We Keep Data Safe!



Image: www.digitalbevaring.dk (CC-by)

We Keep Data Safe!

- Back-up and secure storage
- Long-term preservation
- Legal and ethical security
 - for data producers (IPR, licensing terms)
 - for participants in research (data protection)



Image: www.digitalbevaring.dk (cc-by)

We Make Data Accessible!

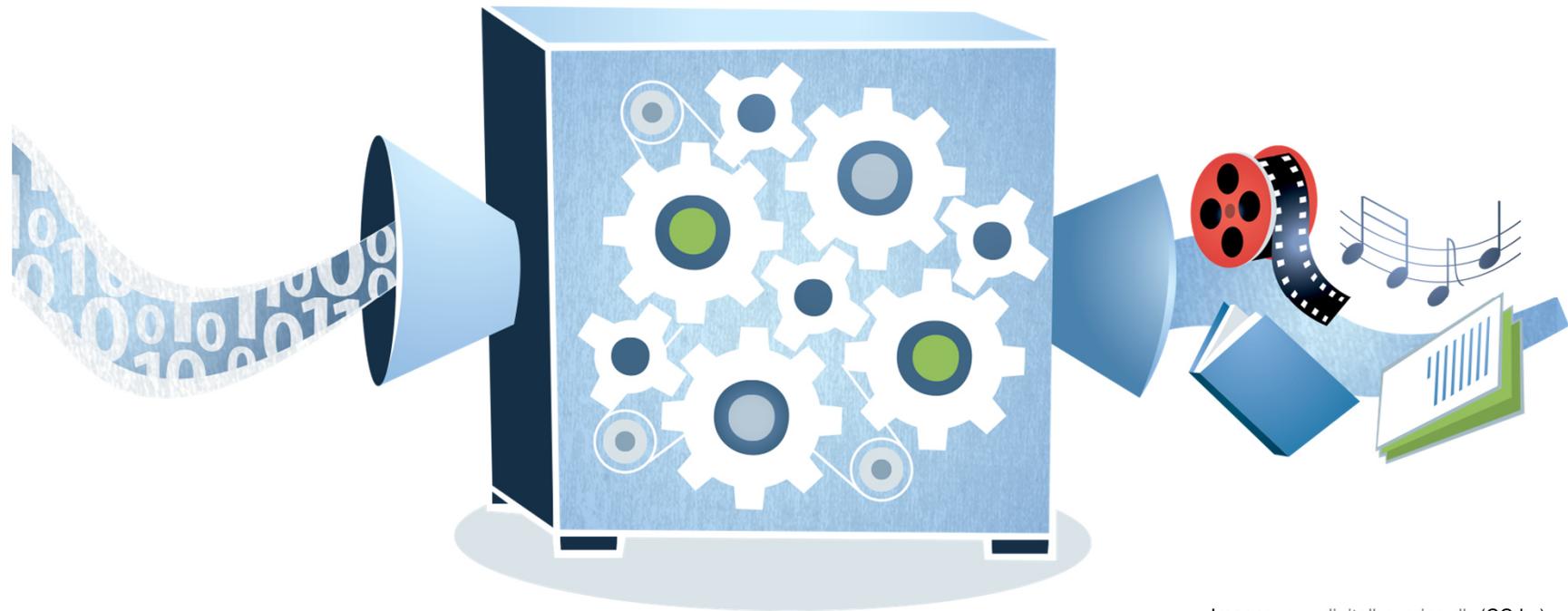


Image: www.digitalbevaring.dk (CC-by)

We Make Data Accessible!

- Accessible formats
- Increased visibility in the world wide web (e.g. through machine-readable metadata)
- Persistent identification
- High-quality documentation

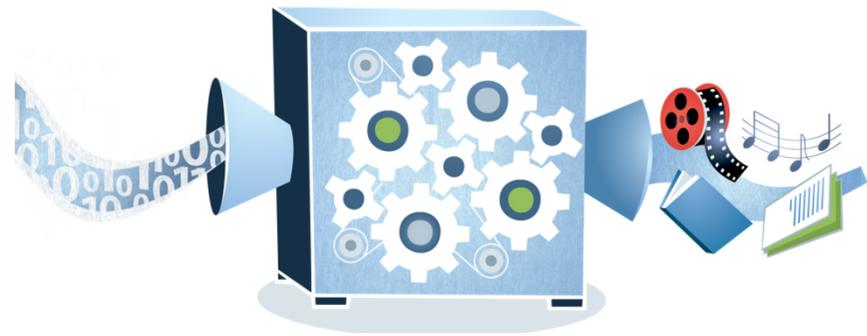


Image: www.digitalbevaring.dk (CC-by)

We Offer Support!



Image: www.digitalbevaring.dk (CC-by)

We Offer Support!

- We advise on
 - documentation
 - metadata
 - anonymization and data protection



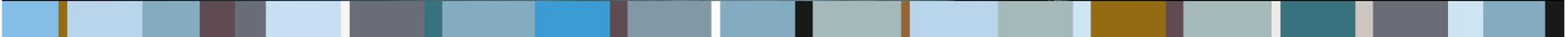
Image: www.digitalbevaring.dk (CC-by)

1. Identify a Suitable Archive or Repository

Possible selection criteria:

- data type and format
- discipline
- region or funder
- service and reputation

Search tool: re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES



2. Contact the Archive as Early as Possible

Find out

- if the archive is interested in taking your data
- which criteria for submission apply



3. Discuss When, How, and in Which Format the Data will be Submitted

A photograph of several LEGO minifigures of various colors (blue, red, green, yellow, brown) standing around a large, round, white table. The minifigures are positioned as if they are in a meeting or discussion. The table is the central focus, and the text is overlaid on it.

Clarify access conditions, legal and ethical issues.

4. Sign the Deposit Agreement

Contractual agreement clarifying rights and obligations of archive and data depositor as well as access conditions.



5. Submit the Data in the Agreed Format at the Agreed Time

Possible submission channels include

- upload,
- e-mail,
- secure file sharing service (e.g. Cryptshare).



Conclusion: Sharing and Preservation

- Start thinking about this early
- Contact potential archives at the beginning of your research
- Pay specific attention to: informed consent and legal restrictions

Wrap-up

In conclusion...

- ... provided an introduction on RDM
 - focused on the most important topics
 - gave an overview of relevant aspects

... hands-on DMP

- didn't touched some of the sections
- sometimes more a “to-do list”
than a documentation of your RDM

⇒ remember, as research is a dynamic process,
the DMP needs to be frequently up-dated



Image: CC-0



Perform an Adequate RDM

basic quality assurance		replicability		reusability	
back-up strategy within the project		back-up strategy for storing data after the project (for 10 years)		plan submission to an archive for long-term preservation	
data collection and versioning guidelines				standardization, e.g. by employing licensed scales	
minimal documentation, e.g. sampling, variable and code labels		metadata to describe the entire research process		detailed documentation for reuse	
legal / ethical issues: informed consent for use of data within the project		legal / ethical issues: data storage or making it accessible to others		legal / ethical issues: archiving and reuse (covered by informed consent?)	
file formats that fulfill the needs of the primary research group		file formats for keeping data & documentation accessible for at least 10 years		file formats that facilitate data reuse in the future	