

# Data Management in EU-funded Projects

Sebastian Netscher

GESIS - Leibniz Institute for the Social Sciences

Laserlab Workshop on Data Handling and Open Data

7 March 2017, Berlin (Germany)

# Data Management in EU-funded Projects

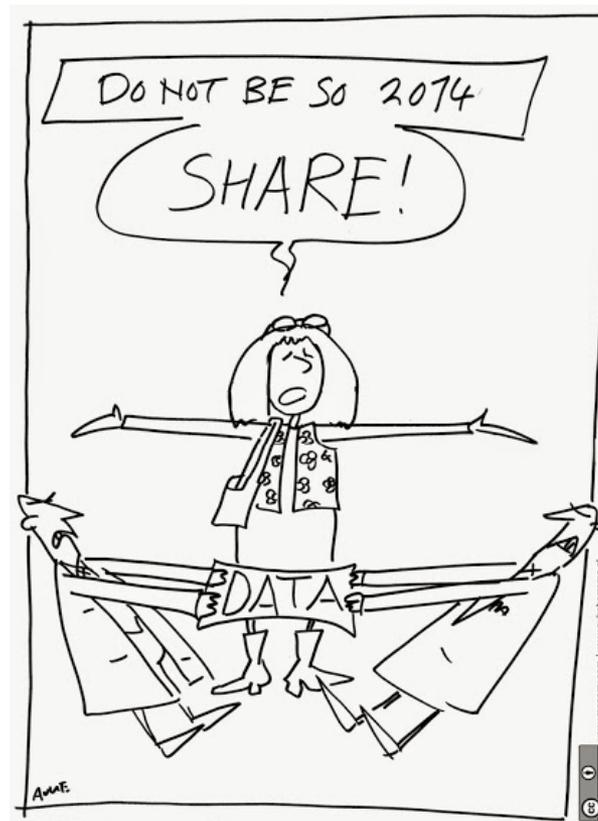
## *“Open Access to Research Data (...)*

*Regarding the **digital research data** generated in the action ('data'), the beneficiaries must (...) **deposit in a research data repository**”*

*“Actions (...) must draw up a '**data management plan (DMP)**' (...). The DMP must support the **management life-cycle** for all **data that will be collected, processed or generated** (...).”*

*(EU: Horizon 2020: Grant Agreement Art. 29.3)*

# Open Access



# Open Access & ...

*Possibility “for third parties to **access, mine, exploit, reproduce and disseminate** (...) the data”*

- using public money efficiently
- fostering research
- enabling reproducibility
- boosting reputation



# ... the FAIR Data Principles

Findable:

controlled vocabulary, standardized metadata, persistent identifiers etc.

Accessible:

access conditions and conditions of reuse

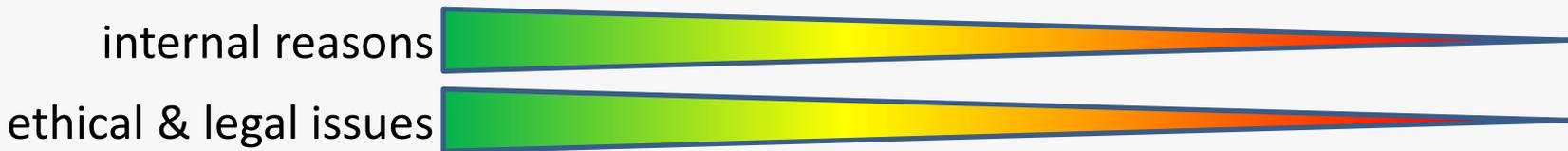
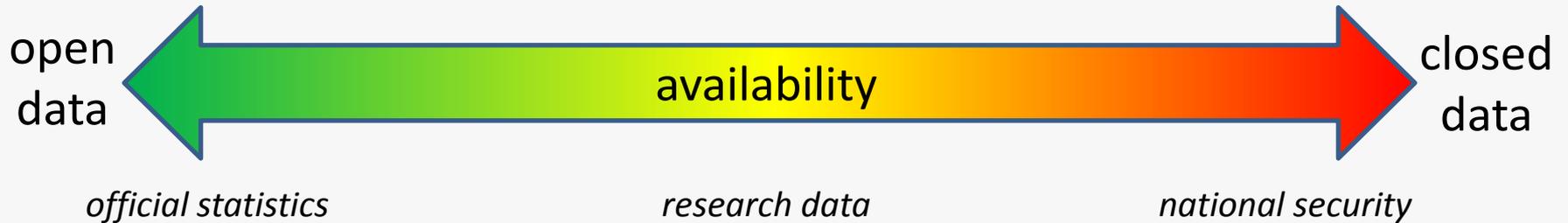
Interoperable:

technical (re-)usability, (standardized) data, file formats etc.

Re-useable:

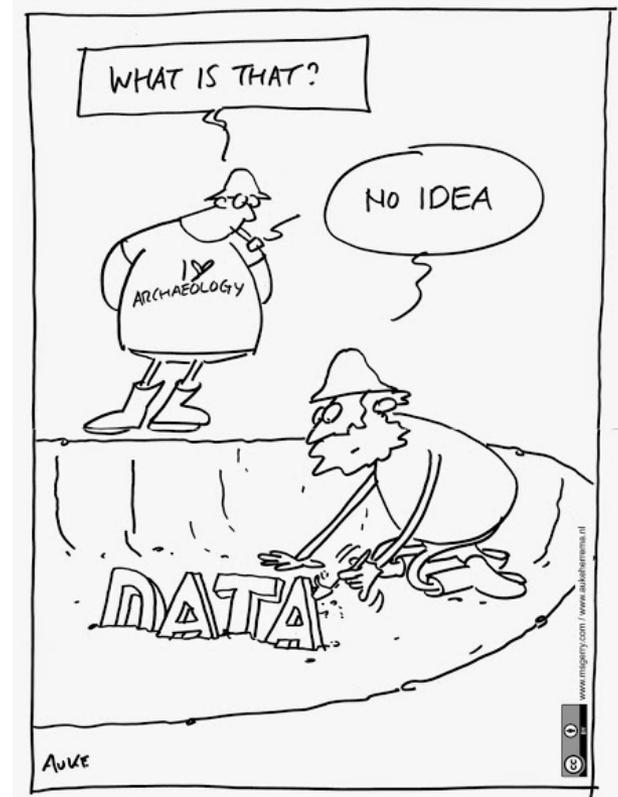
analytical (re-)usability, understandable and interpretable data

# Open Is Not Necessarily Free

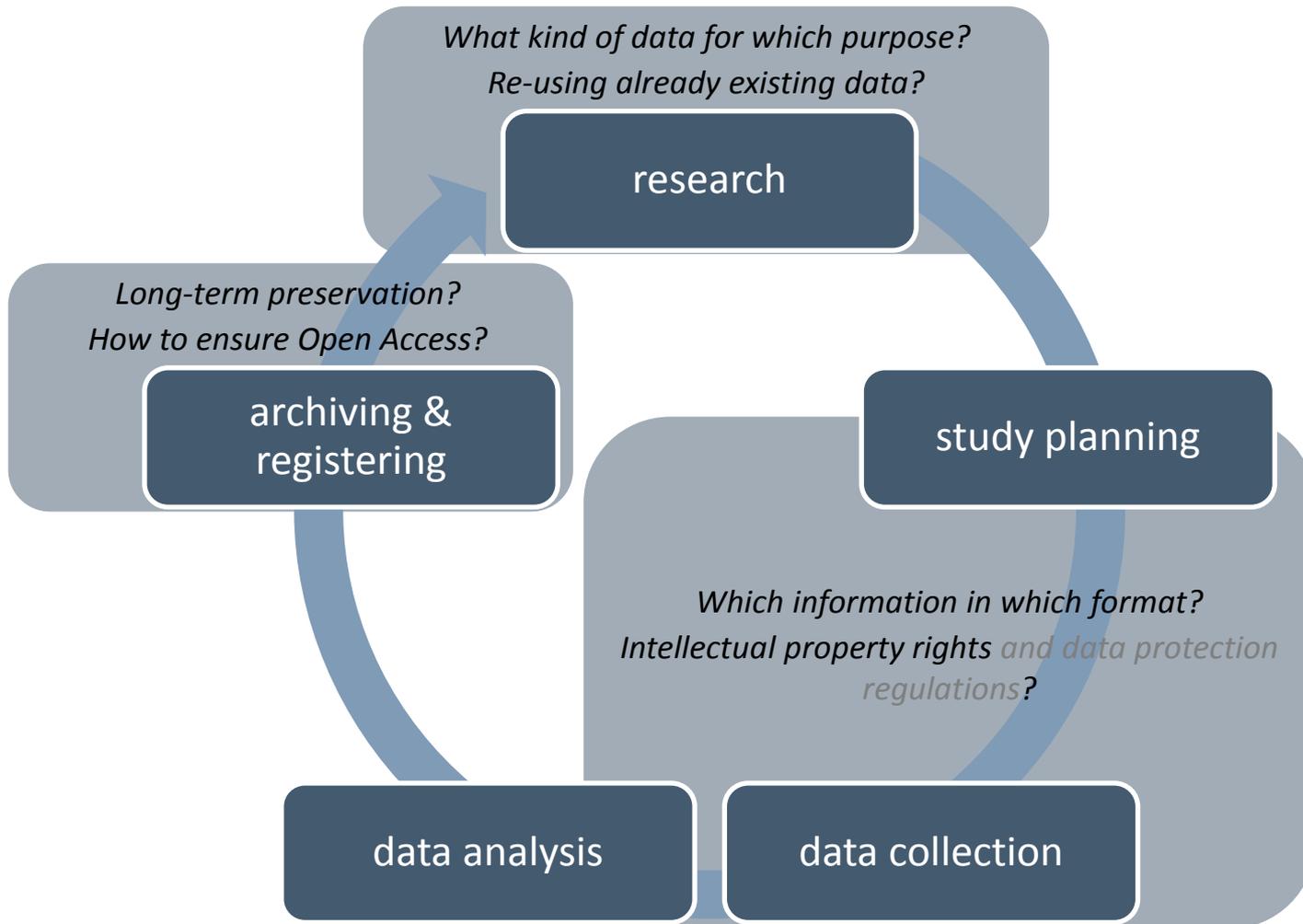


⇒ „as open as possible, as closed as necessary”

# Research Data Management

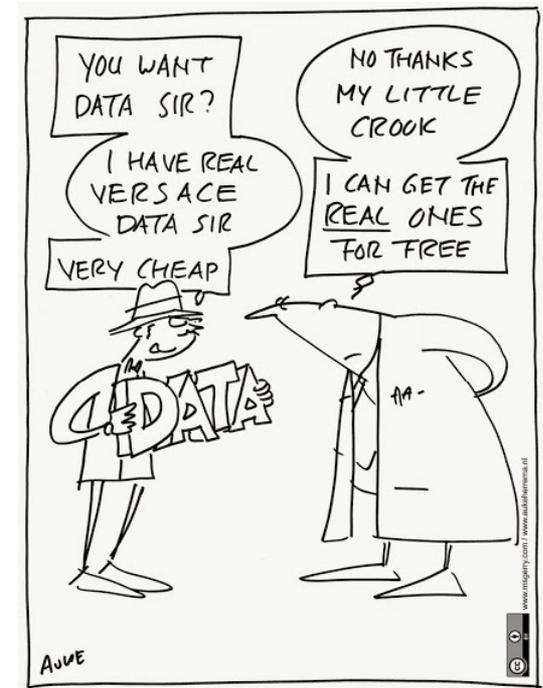


# The Lifecycle of (Research) Data



# Research Data Management

- All strategies, processes and measures to maintain
  - data quality
  - interpretability of research results
  - (re-)usability of research data
- Issues of RDM
  1. handling (research) data
  2. ethical and legal issues
  3. (long-term) preservation



# 1. Handling (Research) Data

- Collecting (raw) data
  - ⇒ collecting information to setup data base
- Processing (research) data
  - ⇒ cleaning, manipulating, calibrating etc. (raw) data to process research data
- Documentation of generating data
  - ⇒ understanding (research) data and interpreting research results
- Data administration
  - ⇒ organizing and storing data

accessible

interoperable

re-useable

findable

## 2. Ethical and Legal Issues

- Data protection regulations  
(personal information)
  - ⇒ informed consent, anonymization  
and secure protection
- Intellectual property rights
  - ⇒ on (research) data
  - ⇒ of others

accessible

### 3. (Long-term) Preservation

- replication
  - ⇒ reproducing research findings  
and replicating (research) data
- sharing with others
  - ⇒ for re-use in new research contexts

findable

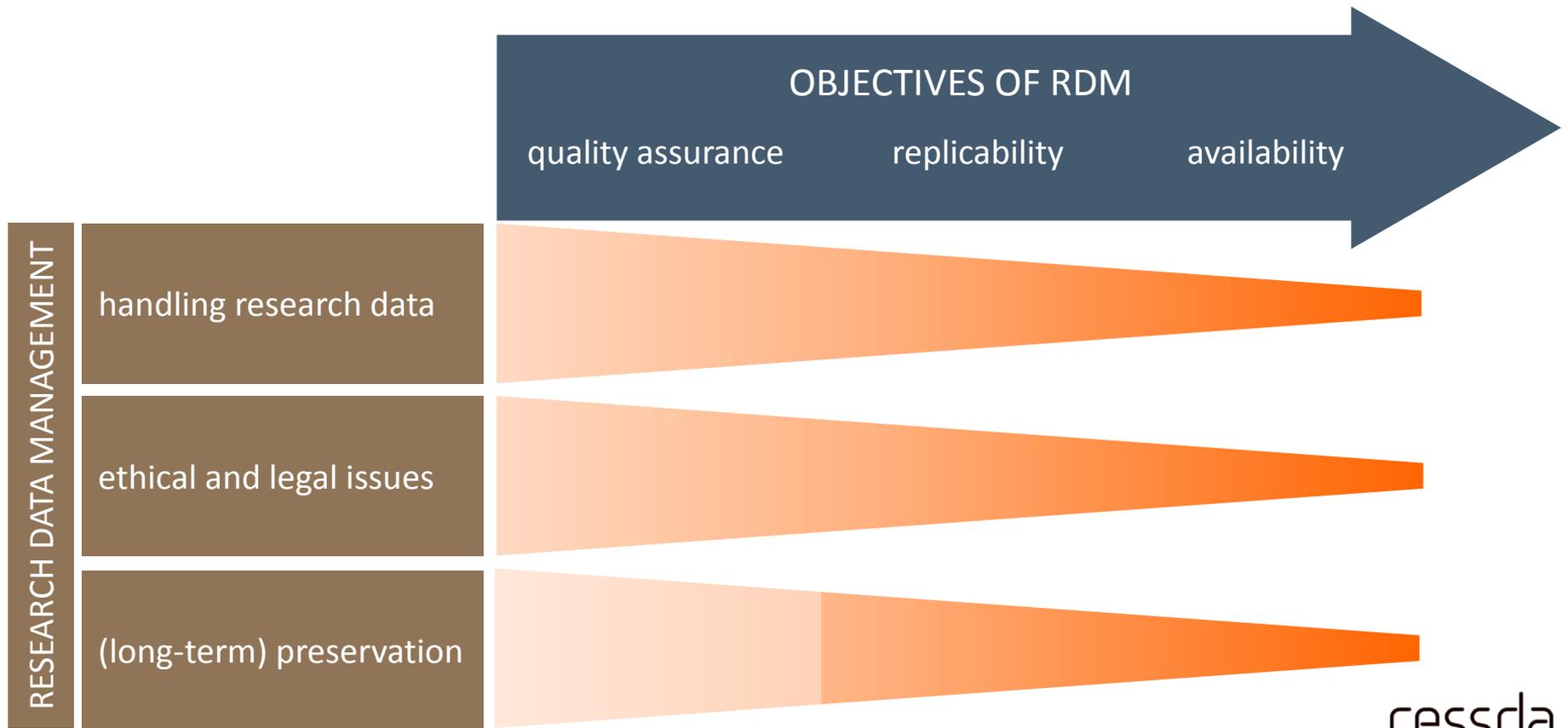
accessible

interoperable

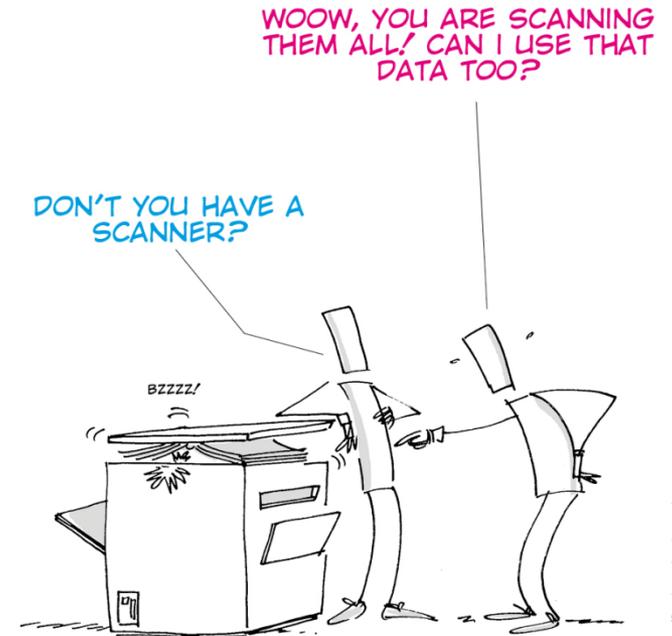
re-useable

# Data Management Plan (DMP)

⇒ systematical documentation of RDM

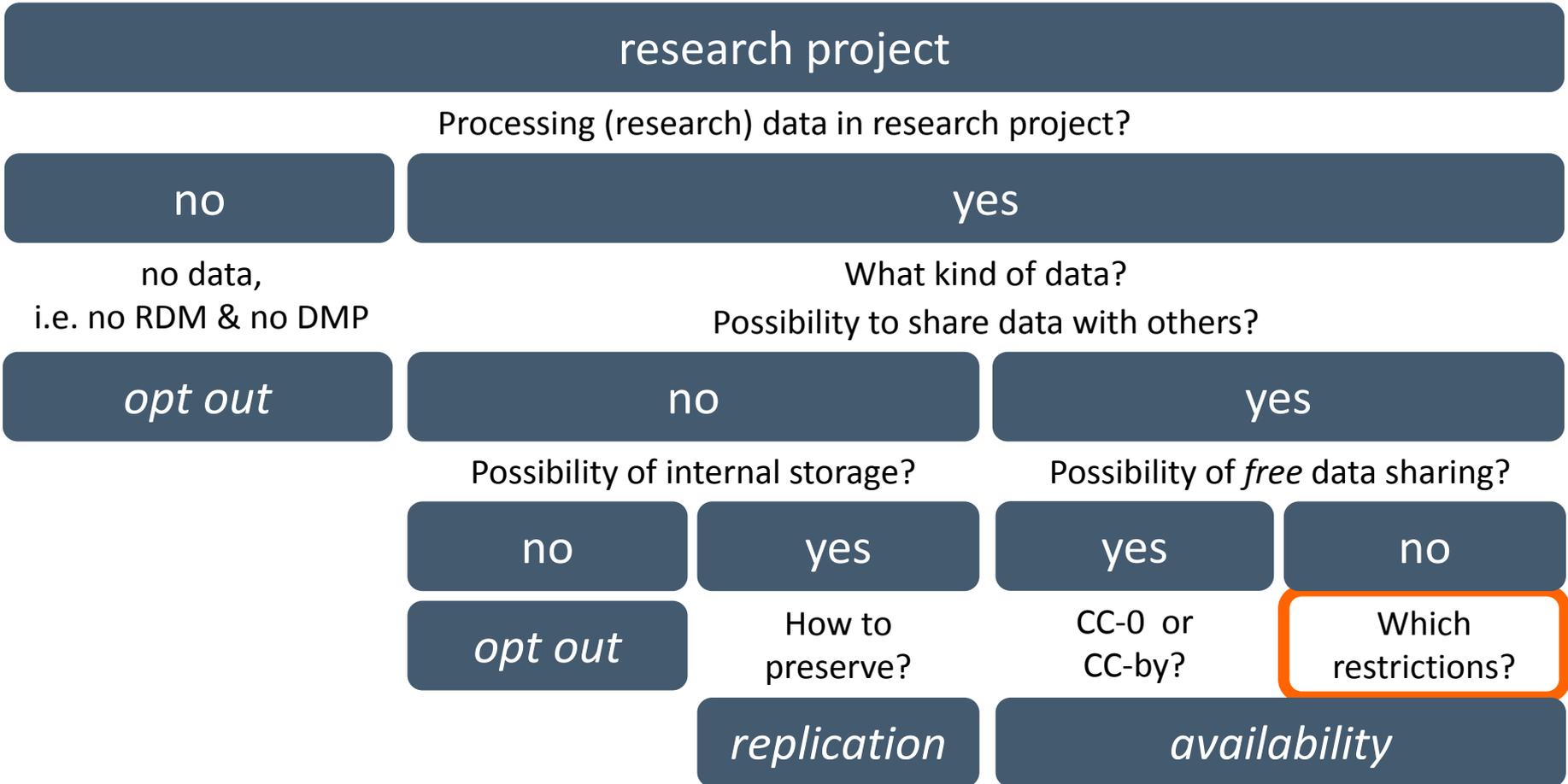


# Open Access & Data Management Plans



cc-by <http://hochstetbach.wordpress.com>

# Research Data and Availability



# Restrictions on (Open) Access

- Internal reasons (project)
  - reaching the aim of the research project
  - reaching additional aims, e.g. individual qualification
- Ethical and legal issues
  - data protection regulations
  - intellectual property rights
    - on generated (research) data
    - of others, e.g. on used methods, standards etc.
  - further requirements, e.g. of an ethics committee

⇒ Which restrictions, why and how?

# Availability and Archiving of Research Data

long-term preservation and availability of research data

Which data shall be preserved?

How to preserve research data and how to ensure availability?

by the research project

How and how long shall data be preserved?

How to find an access these data?

by a repository

findability and accessibility

repository

discipline, region, type of data etc.



# Data Archiving and Data Repositories

- *“Open access (...) via a research data repository (...) an online archive for research data”*
- Repositories
  - care about
    - findable data, e.g. by persistent identifiers, structured metadata etc.
    - accessible data, e.g. by licensing
  - support generating interoperable and re-useable research data



# Documenting RDM to ensure Open Access

- DMP is a living document, i.e.  
*„needs to be updated over the course of the project“*
  - first proposal within first six month of funding
    - ⇒ strategy to ensure Open Access
  - updated DMP for periodical reporting
    - ⇒ ensuring FAIR-data that can be shared
  - DMP in final project report
    - ⇒ data are (will be made)
      - findable and accessible at ...
      - interoperable and re-useable according to ...

# Guidelines on How to Write a DMP

- EU provides guidelines on
    - how to process FAIR data
    - how to document data processing in a DMP
- ⇒ *Horizon 2020 Programme:  
Guidelines on FAIR Data  
Management in Horizon 2020.  
Version 3.0. 26 July 2016*



# Literature

- Force 11 (2016): The FAIR Data Principles.
- Horizon 2020 Programme (2016): Annotated Model Grant Agreement. Version 2.1.1. 1. July 2016.
- Horizon 2020 Programme (2016): Guidelines on FAIR Data Management in Horizon 2020. Version 3.0. 26 Juli 2016.
- Horizon 2020 Programme (2016): Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020. Version 3.1. 25 August 2016.
- Horizon 2020 Programmme (2016): Guidelines on the Implementation of Open Access to Scientific Publications and Research Data in Projects supported by the European Research Council under Horizon 2020. August 2016.