

# Research Data Management (RDM)



[Leibniz IOER](#) - RDM Seminar 20.05.2022

*Stefano Della Chiesa*



# Outline – Today`s RDM Recipe

## 4 Modules and 1 group exercise

Short Presentation:	(9:00-9:15)
Module 1: Introduction to RDM	(9:15-10:00)
<i>Coffee break:</i>	<i>(10:00-10:15)</i>
Module 2: High-Level Policies	(10:15-11:00)
Module 3: Knowledge Bricks	(11:00-12:30)
<i>Lunch break:</i>	<i>(12:30-13:30)</i>
Module 4: Data Management Plans	(13:30-14:15)
DMP Exercise ( <i>World coffee style</i> ): in 3-4 groups Outlining a DMP	(14:15-16:00)



# Brief Presentation

## Short participant's presentation

- Briefly on your Project/PhD
- Which data you use, methods and which are the expected research outputs?



Knowledge network

# Before starting

RDM is like a multidisciplinary knowledge network

- You´ll be introduced on several concepts linked with another (be patient, feel free to ask)
- The training material contains several hyperlinks to other sources



Knowledge network



# Module 1 - Introduction

*This first module introduces RDM, giving its context in the Research Data Governance*

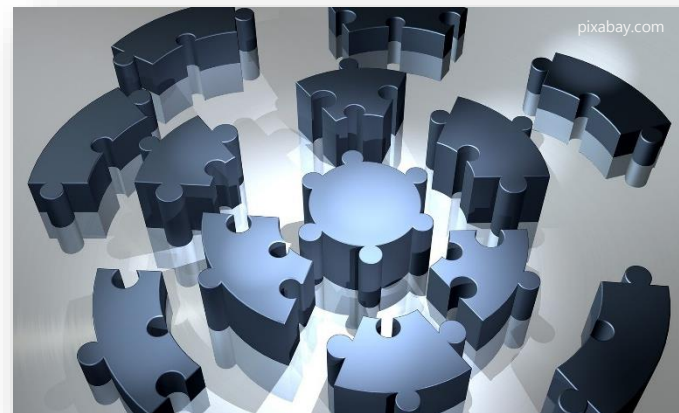
- Research Data Management (RDM)
- RDM Top Down & Bottom Up
- Linked Research Outputs
- Research Assessment



# Research Data Management - What?

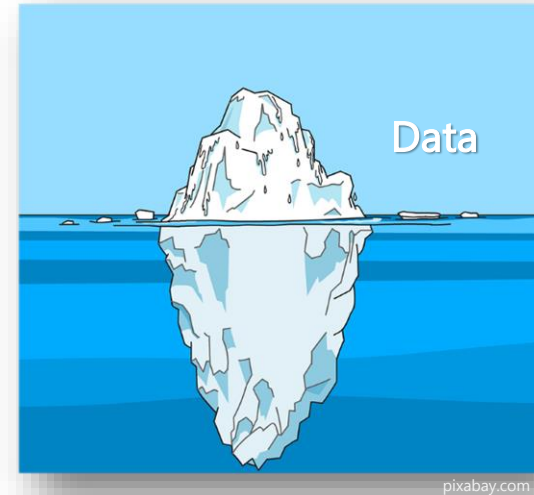
## A Puzzle of knowledge bricks

RDM is a mosaic of multidisciplinary knowledge and competencies. Hence, RDM know-how serves as a reference toolkit to address policy compliance and data management best practices throughout the project lifecycle.



# What is Research Data?

- No standard definition available.
- Research data is the fundament of any scientific work documenting its results.
- DFG specifies: "Research data includes **measurements, audio-visual information, texts, documentation, samples, questionnaires, algorithms, software, simulations**, etc..
- Research data is highly heterogeneous, with subject area having a wide variety of data type and formats.
- Hence: Research Data → **Research Input/Outputs**

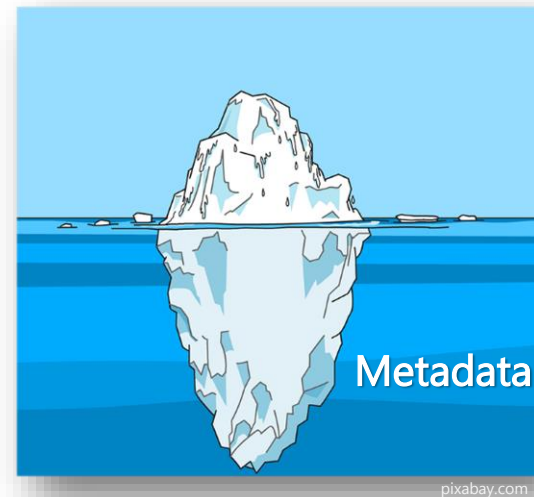


# What is Metadata (MD)?

Metadata (MD) is **data about data**.

In research Metadata represents all the **information that describe and document** in a structured way the research data.

**Structured?** It means there are several MD elements that can describe the data: Title, Creator, Abstract, Keywords, Format, Subject, PID, License, Provenance/Lineage, etc.



# Research Data Management Lifecycle

Research idea and design planning

Data retrieval (existing or new)

Data processing and analysis

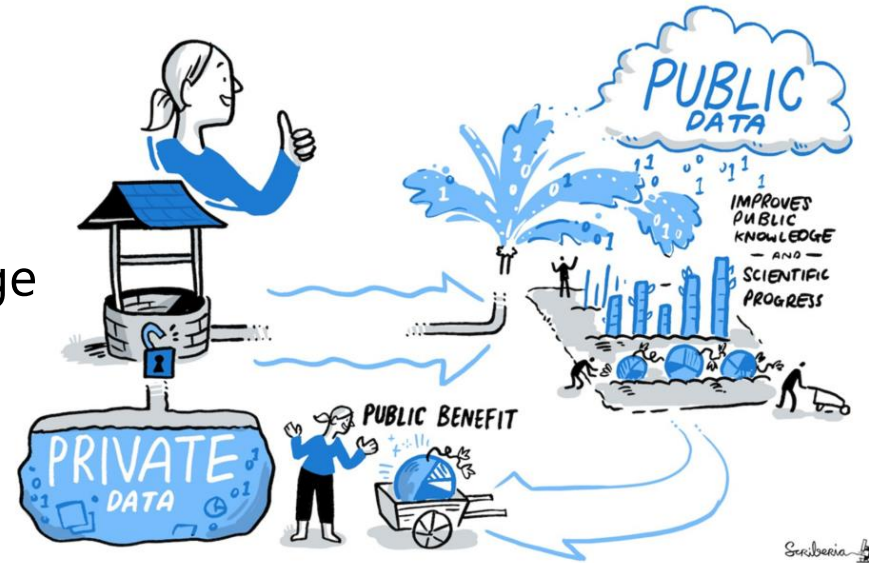
Publishing and re-use



<https://doi.org/10.5281/zenodo.6602006>

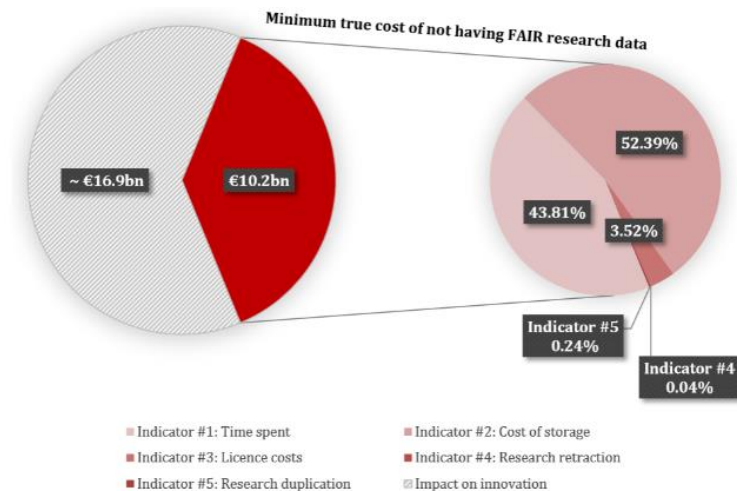
# Research Data Management - Why?

- Assure return of investment
- Efficient spending
- Increases publicly available knowledge
- Network
- Provides new unexplored benefits



# Lack of RDM best practices has a price!

- RDM practices still too “patchy and not optimal”
- Huge cost for not having “open/FAIR” data
- Minimum cost ~ €10bn<sup>-y</sup> (likely cost → almost three-fold)
- Into perspective (Horizon 2020 ~ €12bn<sup>-y</sup>)



EU Cost-benefit analysis for FAIR research data - <https://data.europa.eu/doi/10.2777/02999>

# Research Data Management - How?

Top Down Requirements



Bottom up needs!





# Top Down

## Stakeholder Requirements & High level policies

- DMP Funder requirements  
(e.g. [DFG](#), [Horizon Europe](#))
- [Open Science](#)
- [Good Scientific Practice](#)
- [Guidelines Handling Research Data](#)
- [Open Access](#)
- [FAIR Principles](#)



CC-BY 4.0 [10.5281/zenodo.5706310](https://doi.org/10.5281/zenodo.5706310)

Scriberia

# Bottom up

## Researcher and community needs!

- Save time in the long-term
- Find again your data
- Good practices & Data Culture
- Partnership & Collaboration
- Share, re-use and get credits
- Data/Method/Software Paper
- Grey Literature



# Digital & Linked Research Outputs

## Research is digital

Several research outputs associated with the formal Journal publication are stored & shared online. (e.g. datasets, software/code/algorithms, protocols, preprints, posters, documentation, etc.)



[Crossref Blog: "Linking publications to associated research outputs"](#)

# Linked Research Outputs



# Research Outputs

Peer review



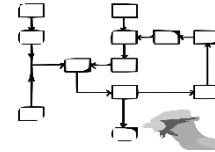
*Peer review  
publication*



*Algorithms*



*Software*



*Workflows*

**Non-Peer review  
Workflow  
(Free)**



*Peer review  
Data Paper*



*Peer review  
Software Paper*



*White papers*



*Posters*



*Documentation*



*Data\**



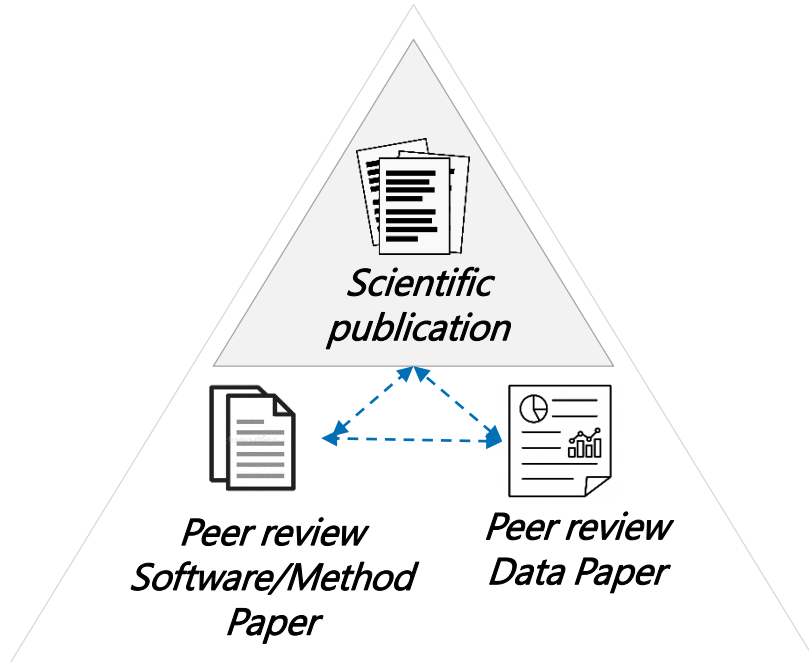
*Designs  
Artworks*

*\* General, Institutional,  
Disciplinary Repositories*

**Research Outputs**

Icons:  
pixabay.com  
Freeicons.io

# Research Outputs and Evaluation metrics



Traditional  
ISI- indexed Research Outputs



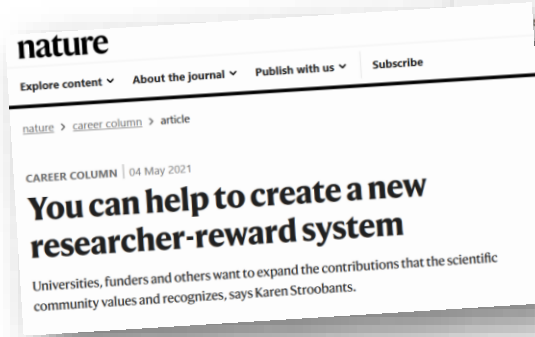
Emerging  
Non-ISI Open Research Outputs

Icons:  
pixabay.com  
Freeicons.io

# Research evaluation

...in addition to research quality and impact factor...

"...so far career-advancement criteria don't reward **open practices** but it is **changing...**"



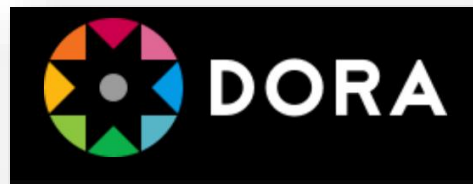
# The Declaration on Research Assessment (DORA)

Scientific Community Initiatives to Improve the ways in which researchers and the outputs of scholarly research are evaluated!

“Research assessment criteria should reward **ethics, reputation, teamwork and diversity of outputs**”

Research evaluation needs to change with the times.

Nature 601, 166 (2022). <https://doi.org/10.1038/d41586-022-00056-z>

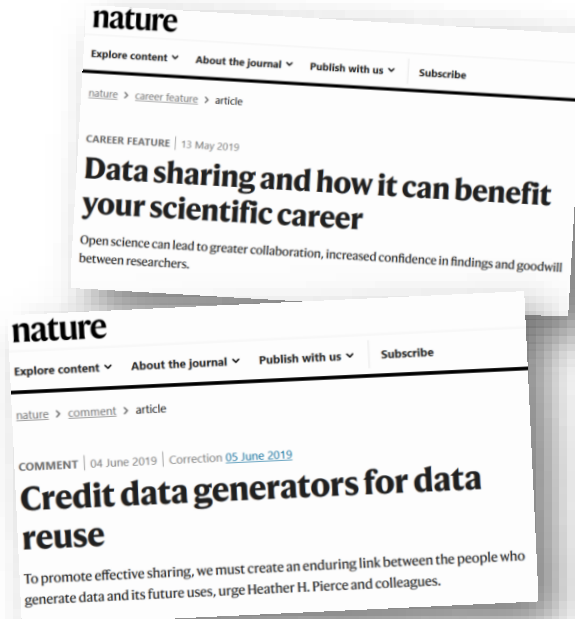


Leibniz IOER is among the 2500 organizations that signed it

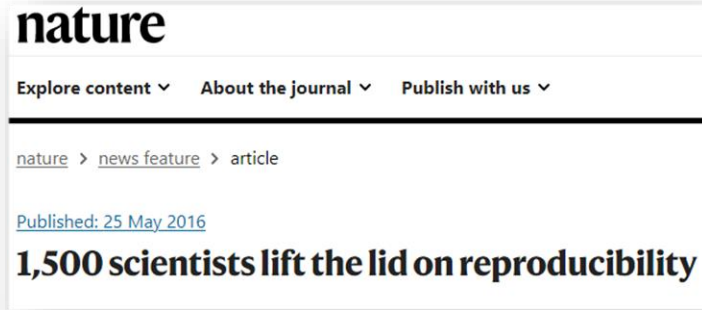




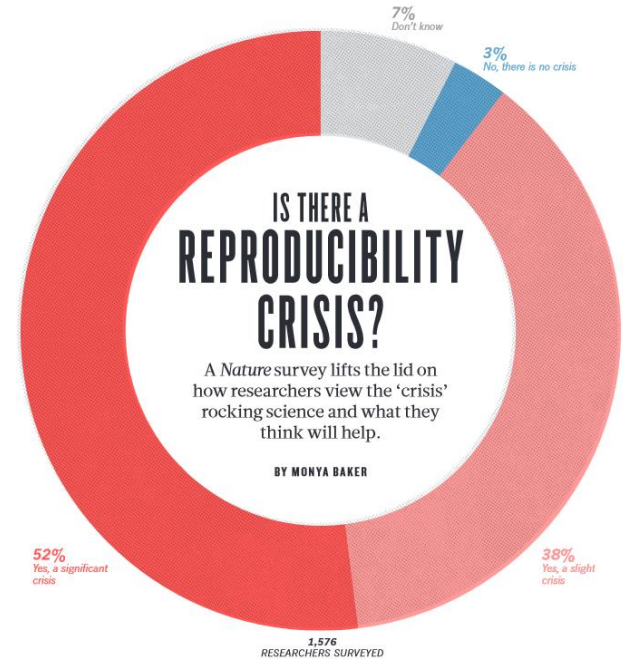
# Growing momentum toward Open Science



# Why RDM does it matter? The reproducibility challenge



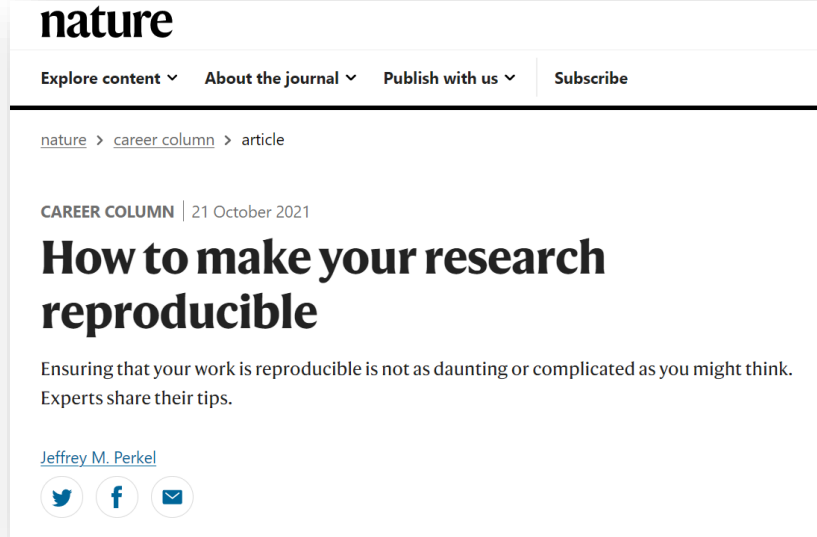
*"More than 70% of researchers have tried and **failed to reproduce** another scientist's experiments, and more than 50% have failed to reproduce their own experiments"*



Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* **533**,452–454 (2016).

# Good RDM → Reproducibility is doable

*Ensuring that your work is reproducible*  
*is not as daunting or complicated*



# Research Data – A Valuable Asset!

## Research Data = Research Asset

Beside scientist competences and skills, data is to be considered as a valuable research asset.

## Open Research outputs = Scientific Reputation

Open and Linking all the Research Outputs is key to build up your Scientific Reputation

