SHARED RDM

Shared RDM Services and Infrastructure

CLUSTER FORSCHUNGSDATEN

Reproducibility in Science

Webinar, 07.02.2024

Good Scientific Practice

Scientific research is committed to GSP generating new knowledge

SHARED RDM

- **Research Integrity:** trust & confidence in the methods and findings
- **Understandability:** quality of comprehensible thought
- **Comprehensibility:** ability of stakeholders to understand relevant aspects
- **Transparency:** making the research process understandable to third parties

  → data documentation & availability for good methodological work

OeAWI Guidelines for Good Scientific Practice
Austrian Agency for Research Integrity
Vienna 2016 ; https://www.oeawi.at

Gleicher M. (2016). A Framework for Considering Comprehensibility in Modeling. Big data, 4(2), 75–88. https://doi.org/10.1089/big.2016.0007

Reproducibility

- **Reproducibility**: same data & same methods → same results
- **Replicability**: new data & same methods → same results

Reproducibility as minimum standard, particularly if replicability is not feasible

Reproducible research is fundamental for scientific integrity

Steven N. Goodman et al. ,
What does research reproducibility mean?
Sci. Transl. Med.8,341ps12-341ps12(2016).
https://doi.org/10.1126/scitranslmed.aaf5027

Reproducibility

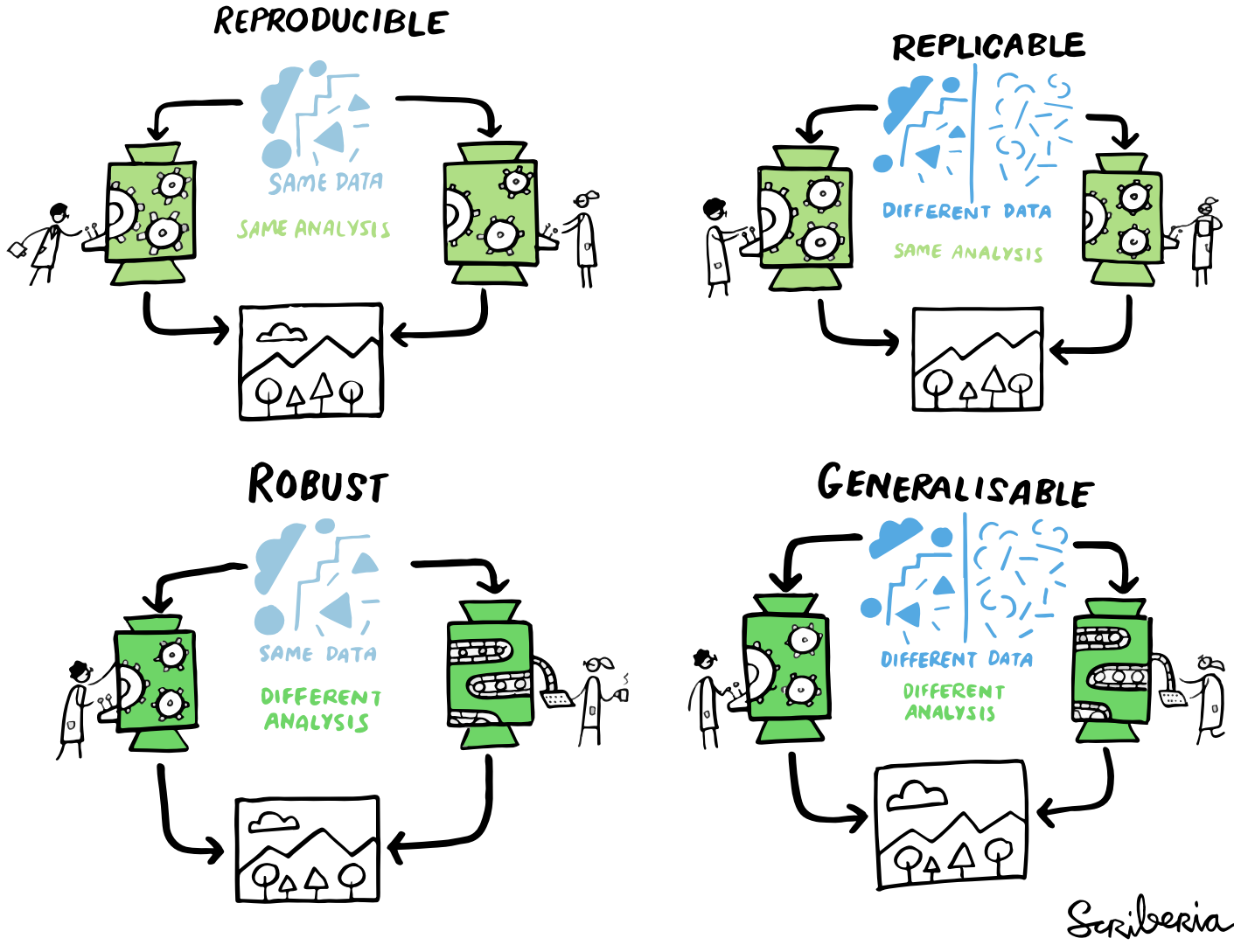
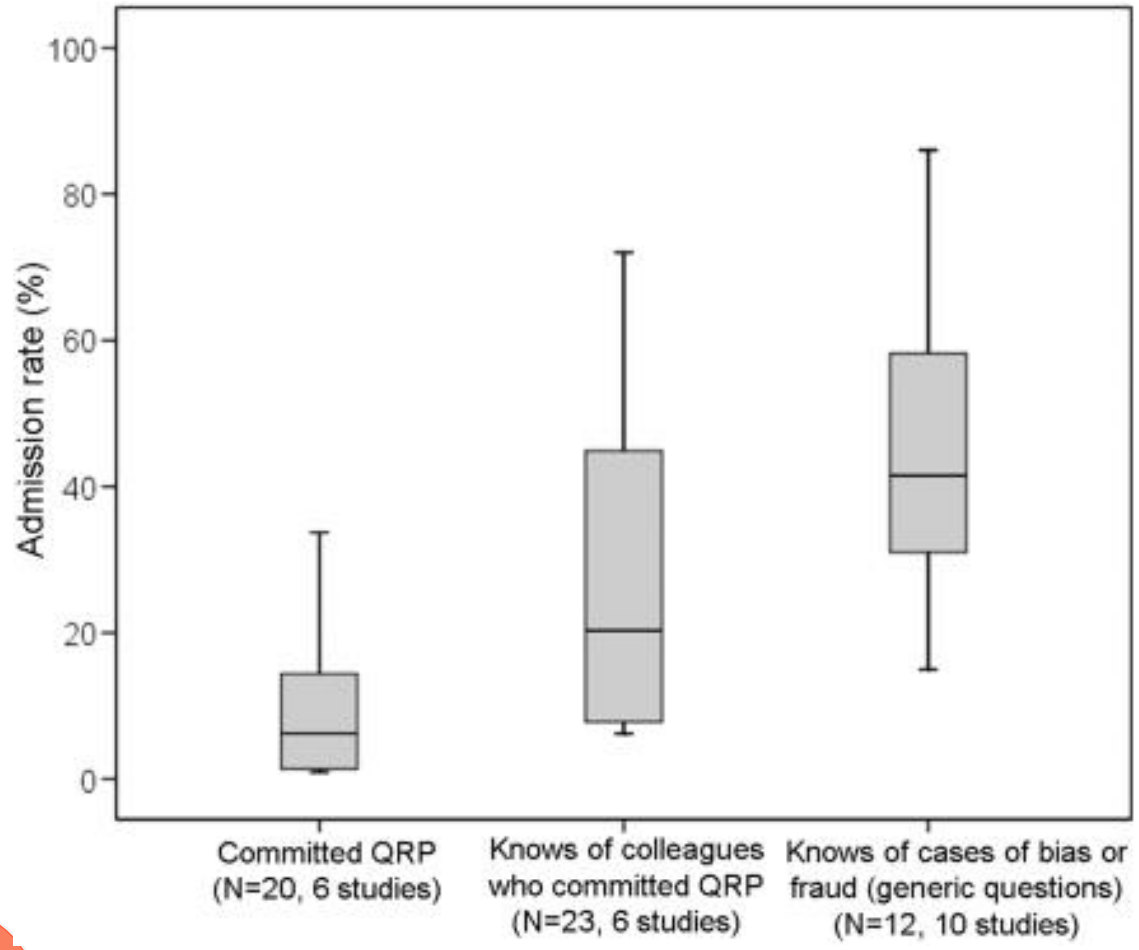


Fig. 6 The Turing Way project illustration by Scriberia.
CC-BY 4.0
https://doi.org/10.5281/zenodo.3332807

Questionable Research Practices



P-hacking

Falsification

Data trimming, Data cooking

Fabrication

Fanelli D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PloS one, 4(5), e5738. https://doi.org/10.1371/journal.pone.0005738

Reproducibility/Replication Crisis

coined over a decade ago,
recognized in Psychology & Medicine/Life Sciences

Retractionwatch.com

„Tracking retractions as window into the scientific process“

Yong, E. Replication studies: Bad copy. Nature 485, 298–300 (2012). https://doi.org/10.1038/485298a
“2011 … eventually confessed to, scientific fraud on a massive scale.”

Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452–454 (2016). https://doi.org/10.1038/533452a
„More than 70% of researchers have tried and **failed to reproduce** another scientist's experiments, and more than half have failed to reproduce their own experiments. ”

Trust but verify. *Nat. Mater.* **23**, 1 (2024). https://doi.org/10.1038/s41563-023-01790-z
„Data may not be reproducible for several reasons, ranging from **honest errors**, such as those in complicated analyses needed to extract results, **to shameful cases of data manipulation**”

Intro: Reproducible Research

… and how to implement it locally

Graz University of Technology
Postdoctoral Researcher
Graduate Research Assistant

The University of Göttingen
PhD student
Department of Diagnostic and Interventional Radiology
Max Planck Institute for Dynamics and Self-Organization
Research Assistant

Challenges & Best Practices

Irreproducible research

- Selective reporting
- Pressure to publish
- Insufficient peer review
- Insufficient mentoring & oversight
- Unavailability of data, methods & code

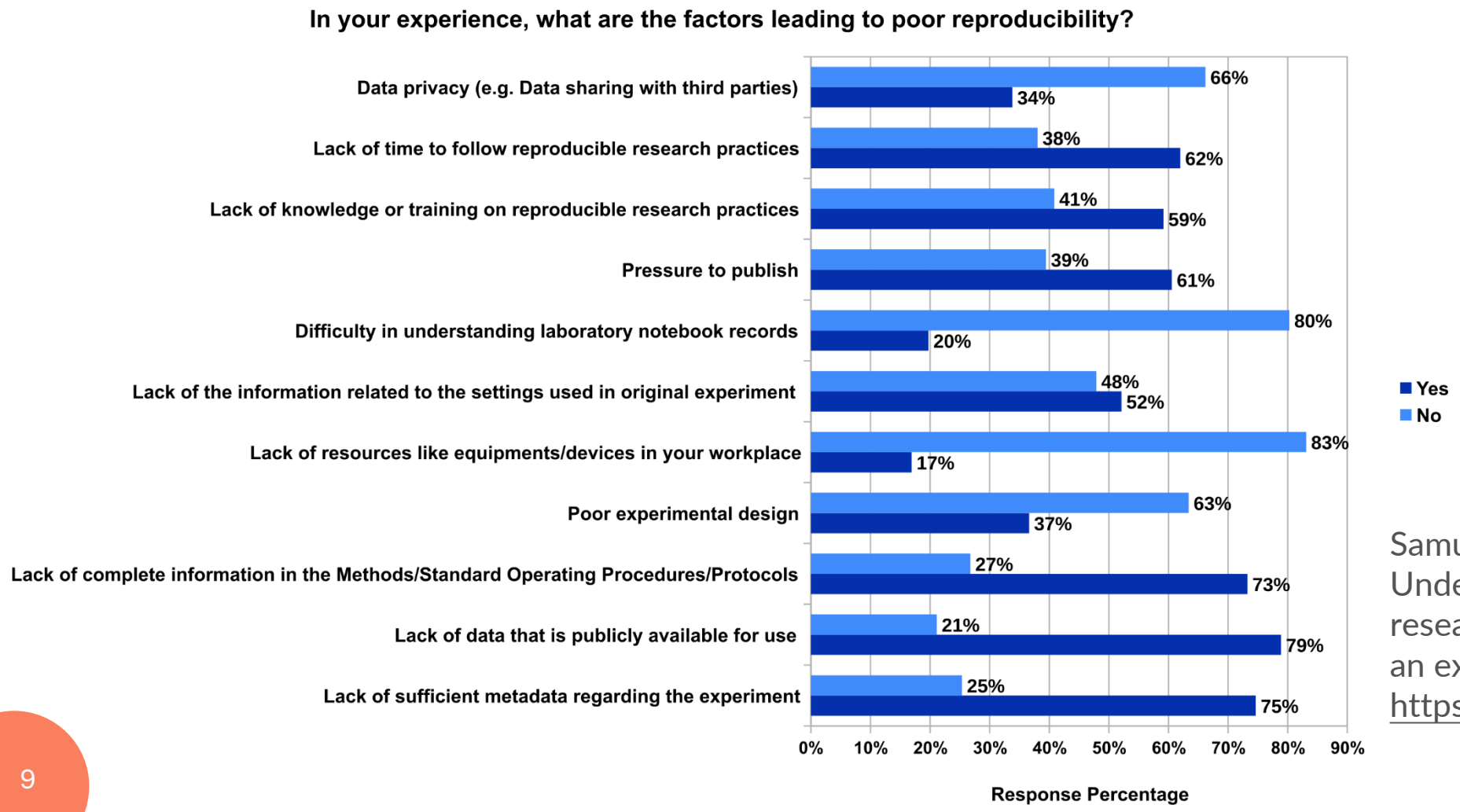
Guidance

Documentation

Research data management

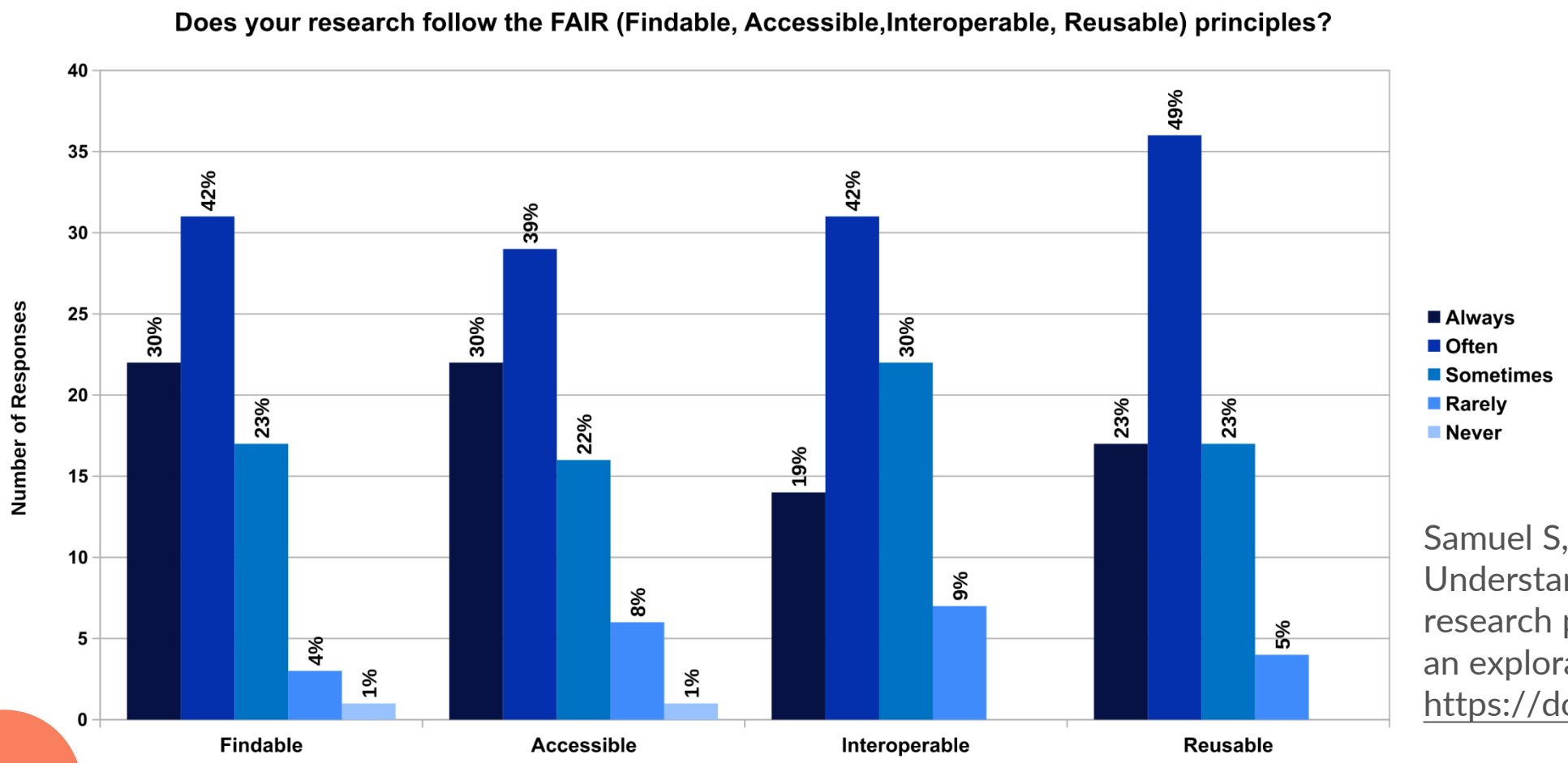
Baker, M. 1,500 scientists lift the lid on reproducibility. Nature 533, 452–454 (2016). https://doi.org/10.1038/533452a

Challenges & Best Practices



Samuel S, König-Ries B. 2021. Understanding experiments and research practices for reproducibility: an exploratory study. PeerJ 9:e11140 https://doi.org/10.7717/peerj.11140

Challenges & Best Practices



Samuel S, König-Ries B. 2021. Understanding experiments and research practices for reproducibility: an exploratory study. PeerJ 9:e11140 https://doi.org/10.7717/peerj.11140

Challenges & Best Practices

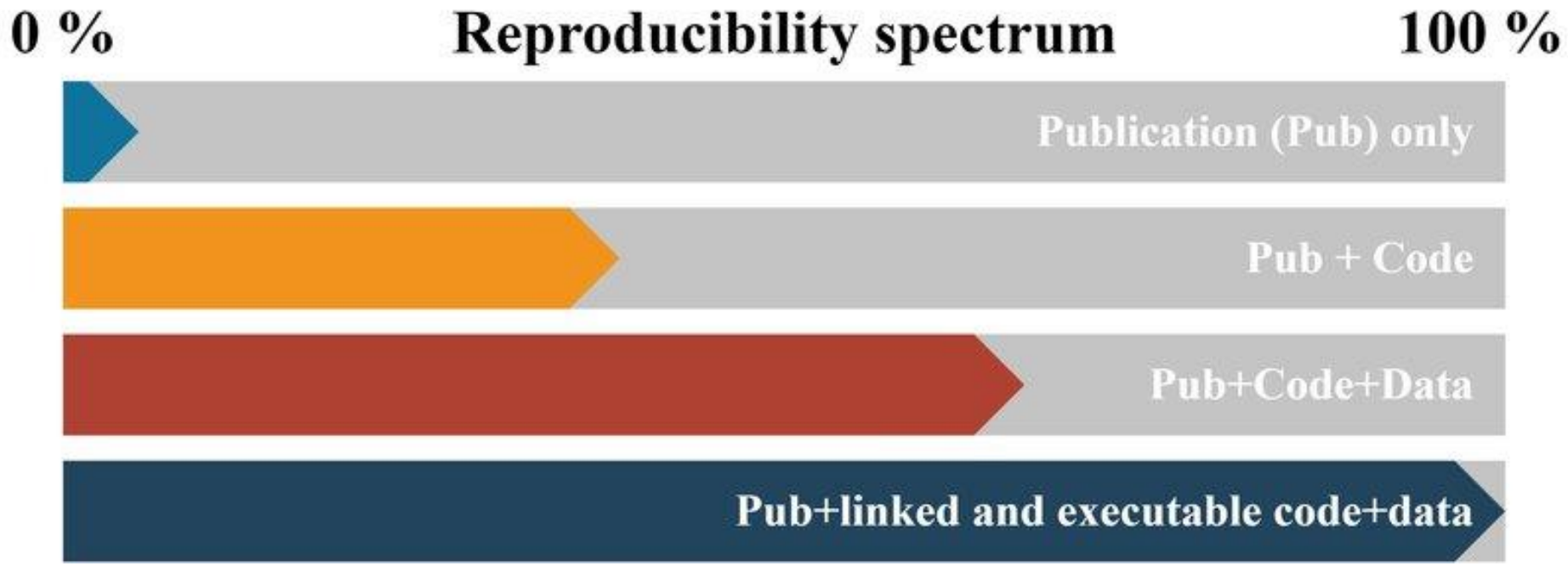
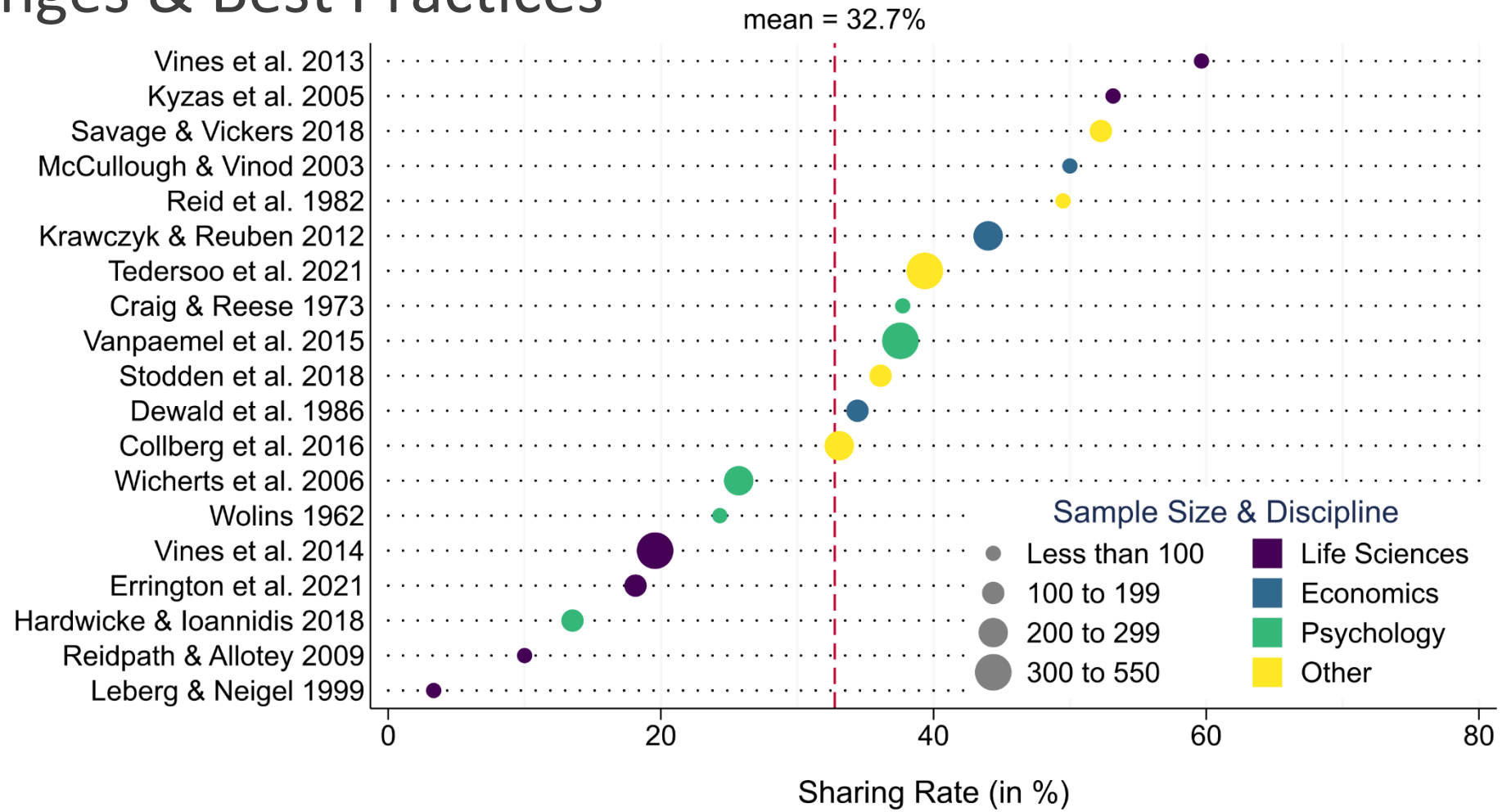


Figure 2: CC by 4.0, I. Charalampopoulos.
The R Language as a Tool for Biometeorological Research
June 2020, Atmosphere 11(7):682
http://dx.doi.org/10.3390/atmos11070682

Challenges & Best Practices



Researchers' willingness to share data/code in the literature. CC by 4.0, Krähmer D, Schächtele L, Schneck A (2023) Care to share? Experimental evidence on code sharing behavior in the social sciences. PLOS ONE 18(8): e0289380. https://doi.org/10.1371/journal.pone.0289380

Best Practices & RDM

Personal benefits?

- **Helps to avoid disaster**: Keeping a record saves you time later
- **Makes it easier to write papers**: Very transparent data and code is easier to explore
- **Helps reviewers see it your way**: Facilitates constructive reviewing processes
- **Enables continuity of your work**: Continue a project where it left off
- **Helps to build your reputation:** Honest and careful researchers are in a very good position to defend themselves

Markowetz, F. Five selfish reasons to work reproducibly. Genome Biol 16, 274 (2015). https://doi.org/10.1186/s13059-015-0850-7

Youtube: Five selfish reasons to work reproducibly

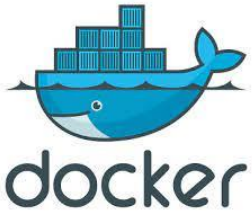
Best Practices & RDM

- **Data Organization**: file structure, version tracking, compiling information
- **Research Quality**: transparency
- **Sharing**: Preservation, Reuse, visibility, collaboration
- **Compliance**: Policies & institutional/funder/publisher requirements

M. Grenier. GitHubSlides_Research Data Management for Reproducible Research. Uvic Libraries Digital Scholarship Commons

Best Practices & RDM

Organization → Versioning
Organization → Sharing
Organization → Citation Manager

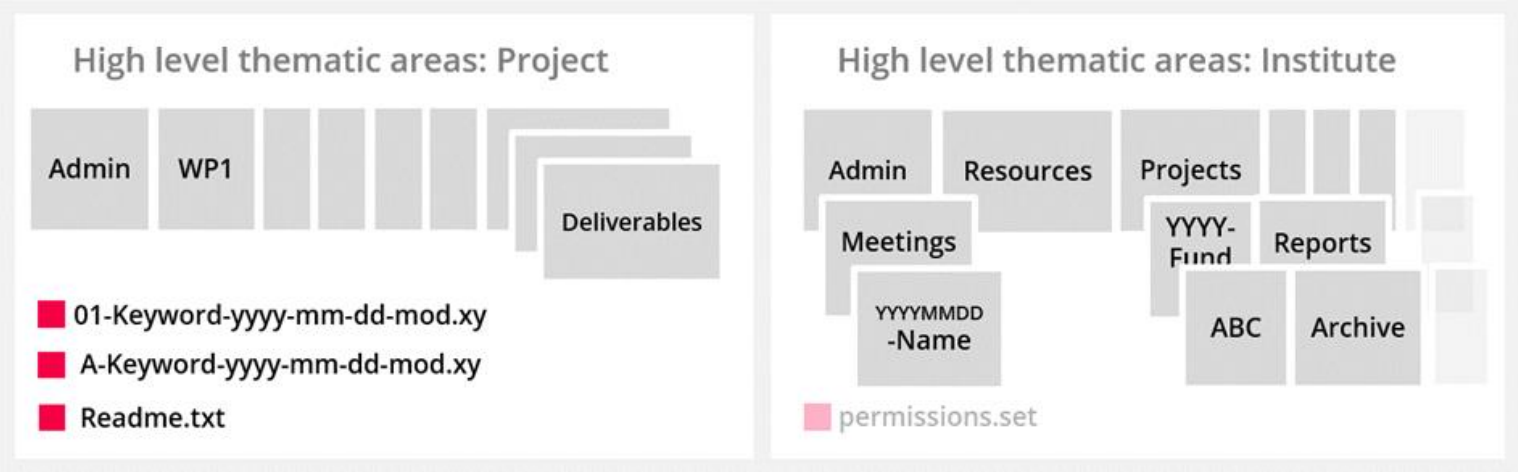


Computational environment

Literate Programming & Code Documentation

M. Grenier. GitHubSlides_Research Data Management for Reproducible Research. Uvic Libraries Digital Scholarship Commons

Best Practices & RDM for today?



Data Management Plan

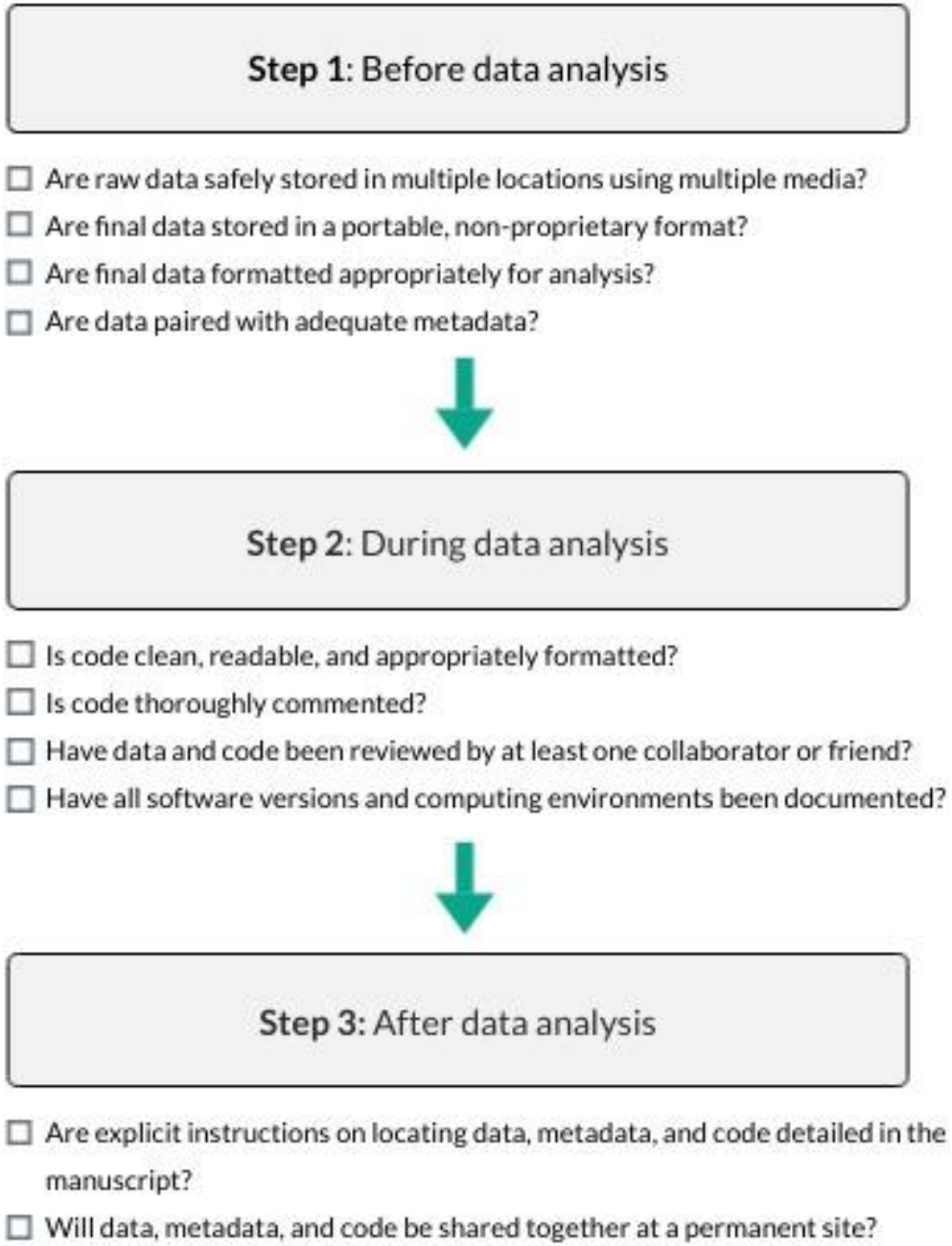
Version 1.0
responsibilites
resources

```
Import ln from library
## comment purpose

Def structure(x,y,n):
          product = 1
          for i in range(y)
                    product = x*product
          return product %  n
Print(structure(x,y,n))
## description of details
```

Contact your institutional RDM Team!

Best Practices & RDM for today?



A Beginner's Guide to Conducting Reproducible Research.
Bulletin Ecologic Soc America, 102 (2), 2021.
https://doi.org/10.1002/bes2.1801

UPCOMING EVENTS

- **Love Data Week**

  12–16.02., online

  https://forschungsdaten.info/fdm-im-deutschsprachigen-raum/love-data-week-2024/

- **National RDM Exchange**

  21.02., 09–10:00, online

- **CRIS2024 Conference**

  15-17.05., Vienna

SHARED RDM

CONTACT

contact.sharedrdm@mlist.tugraz.at
forschungsdaten.at/projekte/sharedrdm/