



**HAL**  
open science

# Opening science through preregistration, standardization and documentation

Gaëlle Leroux

## ► To cite this version:

Gaëlle Leroux. Opening science through preregistration, standardization and documentation. Master. NeuroConferences “Open science, Integrity and equity in research”, France. 2025, pp.65. <hal-05317045>

**HAL Id: hal-05317045**

**<https://hal.science/hal-05317045v1>**

Submitted on 16 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License



# Opening science through preregistration, standardization and documentation

Gaëlle Leroux, PhD

CNRS research engineer

<https://orcid.org/0000-0003-2729-4945>



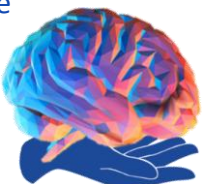
15<sup>th</sup> Oct. 2025



[San Francisco Declaration on Research Assessment](#)



Centre de Recherche en Neurosciences de Lyon



Methodological support for imaging projects



**RERSOL**  
Réseau Référents Science Ouverte Lyon 1 Univ.



# **1. Introduction**

## 2. Preregistration

## 3. Data Management Plan

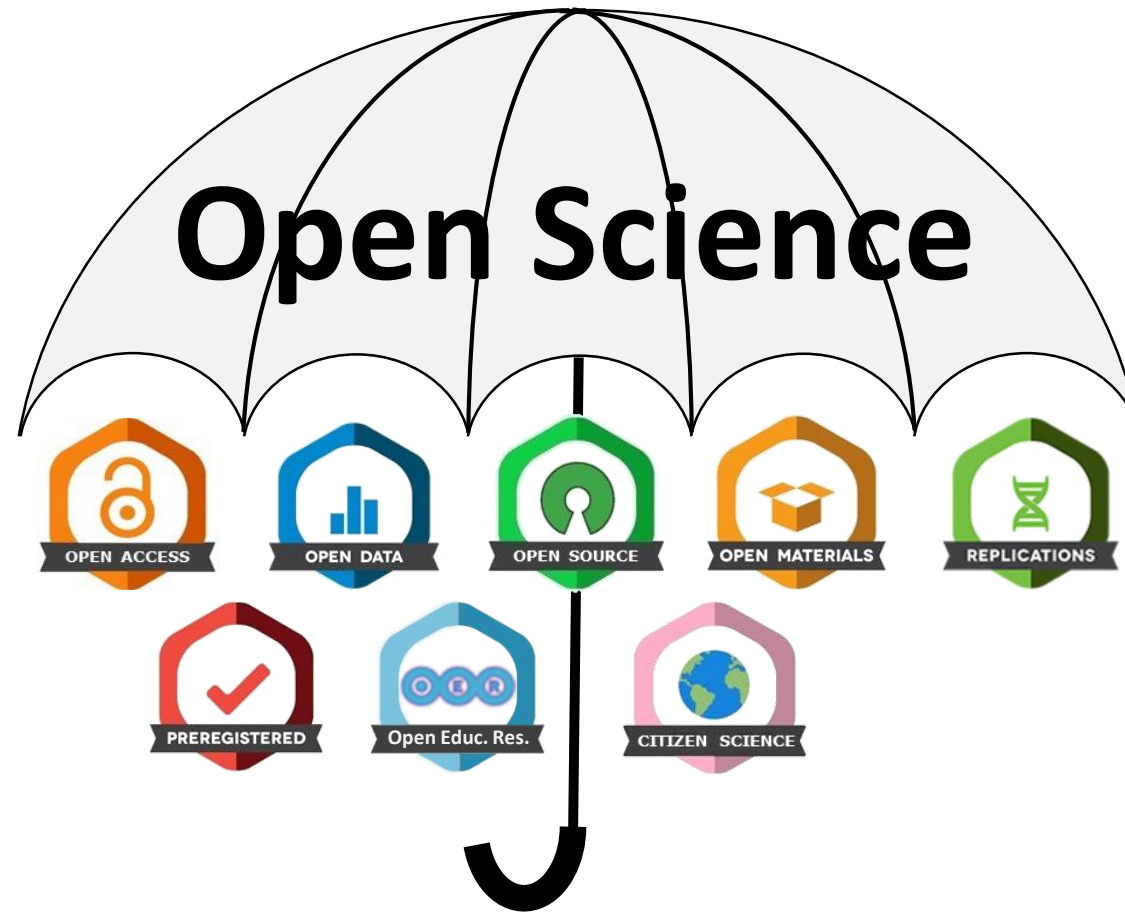
## 4. Standardisation

## 5. Documentation

## 6. Resources

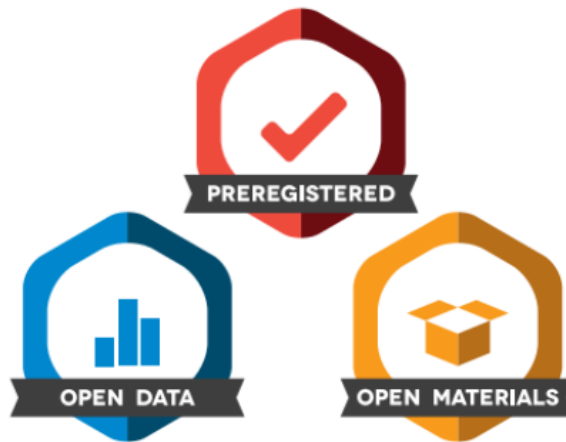


# Context: the Open Science “umbrella term” illustrated with the open badges



Adapted from an [academic blog of Vu Univ., NL](#)


# Open Science Badges enhance openness, a core value of scientific practice.



What are Open Science Badges?

- Badges to acknowledge open science practices are incentives for researchers to share data, materials, or to [preregister](#)
- Badges signal to the reader that the content has been made available and certify its accessibility in a persistent location.
- Currently, over **100 journals** offer [Open Science Badges](#) to signal and reward when underlying data, materials, or preregistrations are available, see below.

Badges seem silly. Do they work?

- Yes. Implementing badges is associated with increasing rate of data sharing ([Kidwell et al, 2016](#) ) , as seeing colleagues practice open science signals that new community norms have arrived.
- For this to occur, badges must be visible on the table of contents and authors must be consistently asked to participate. See the current issue of [Psychological Science](#) for an example of how these new norms look in practice.

→ List of journals that issue open science badges: [TOPFactor.org](https://www.topfactor.org)

<https://www.cos.io/initiatives/badges>

# Academic impact of open science: increase of citations

An article published as a **prepublication** is associated with **19% additional citations**.

An article published in **open access** is associated with **8.6% more citations** than an article not in open access.

An article **sharing source code** is associated with **13.5% additional citations**.

An article **sharing data** is associated with **14.5% additional citations**.

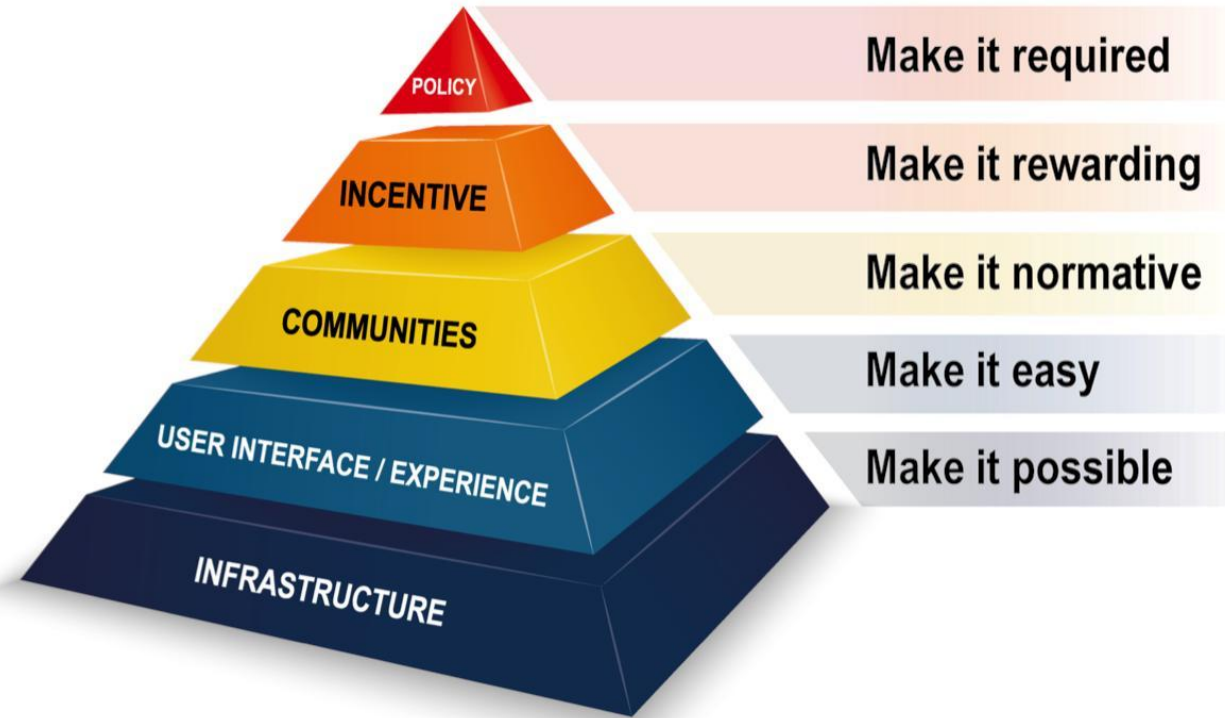
Notes:

These results vary considerably across disciplines.

The study covers a massive corpus of more than 500,000 scientific papers.

Colavizza et al. *An analysis of the effects of open science indicators on citations in the French Open Science Monitor*  
Aug. 2025. Preprint. <https://doi.org/10.48550/arXiv.2508.20747>

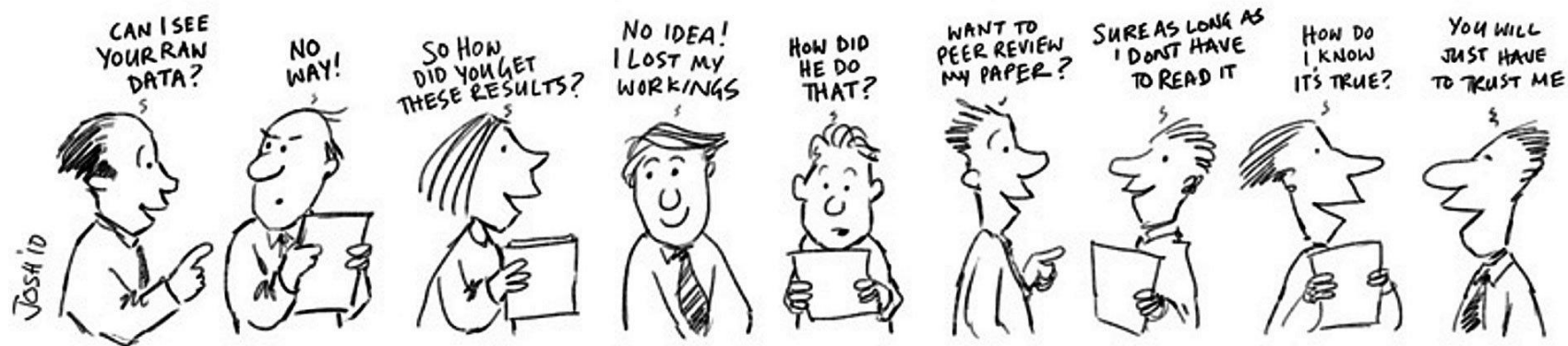
# Open science is a cultural change



I might add another layer at the bottom of the pyramid:  
*“Raise awareness of the problem.”* Burk, 2023. [Source](#)

Image credit: ["EPFL Open science" twitter account](#)  
[Original by Nosek](#)

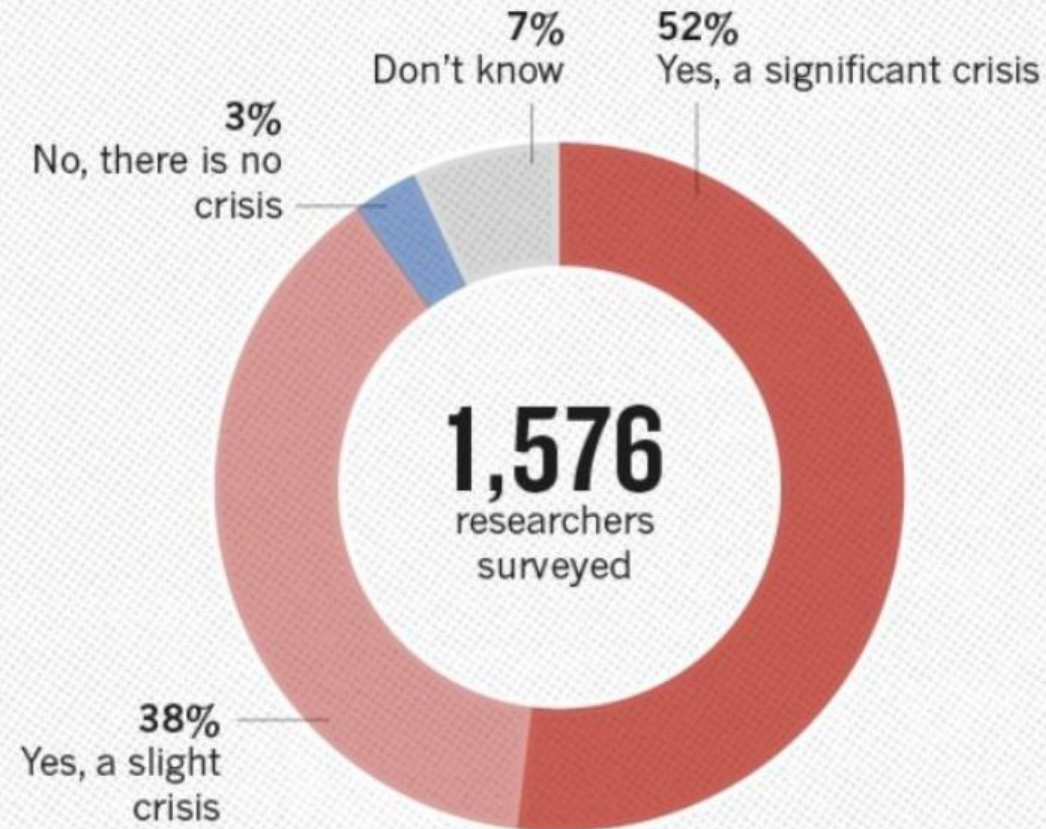
# NO TRANSPARENCY NO CONSENSUS



[https://pbs.twimg.com/media/Cjc\\_UC4XAAAKaJl.jpg:large](https://pbs.twimg.com/media/Cjc_UC4XAAAKaJl.jpg:large)

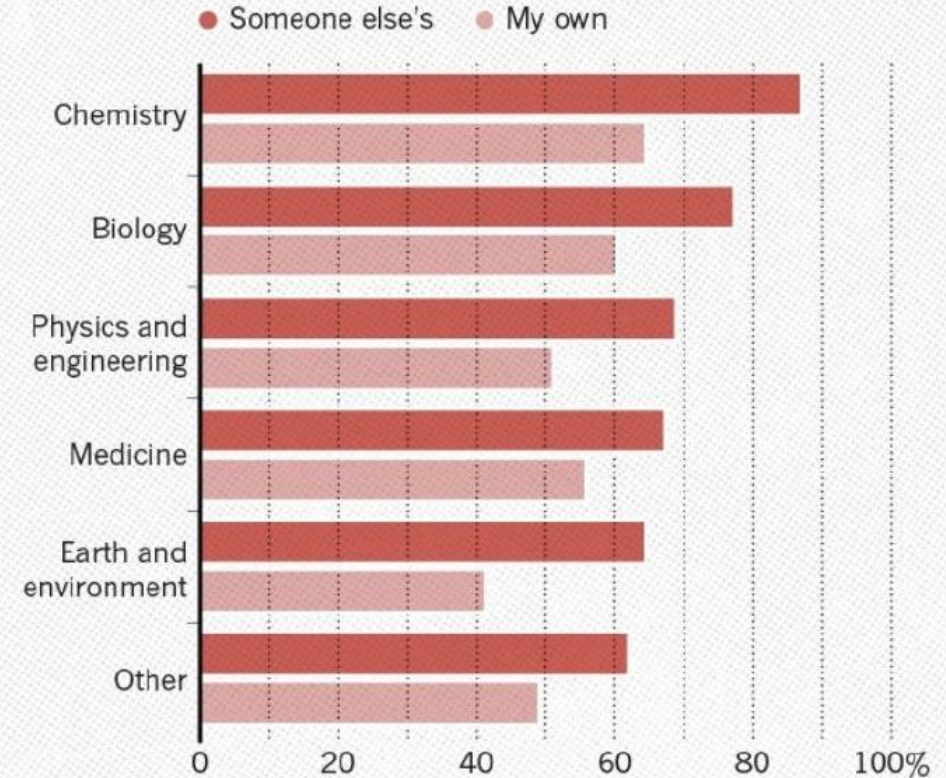
# An overview of the “reproducibility crisis”: awareness still low in 2016

## IS THERE A REPRODUCIBILITY CRISIS?



## HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



Baker. 1,500 scientists lift the lid on reproducibility (2016). <https://doi.org/10.1038/533452a>

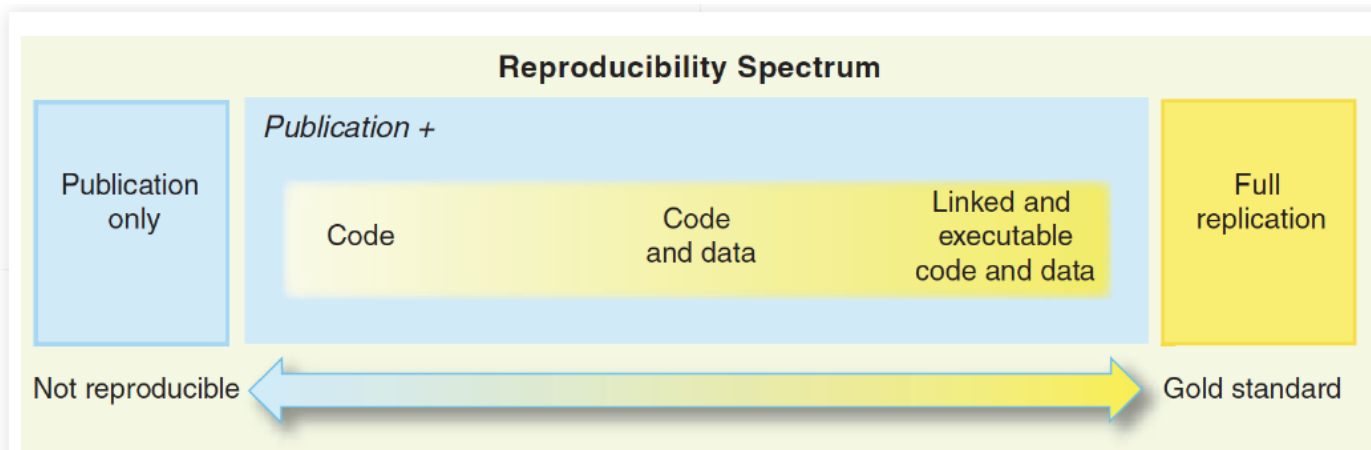
# Reproducibility is not binary



The Turing Way

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

The Turing Way Community. (2022). The Turing Way: A handbook for reproducible, ethical and collaborative research. Zenodo. doi: 10.5281/zenodo.3233853 ([link to the schema](#))

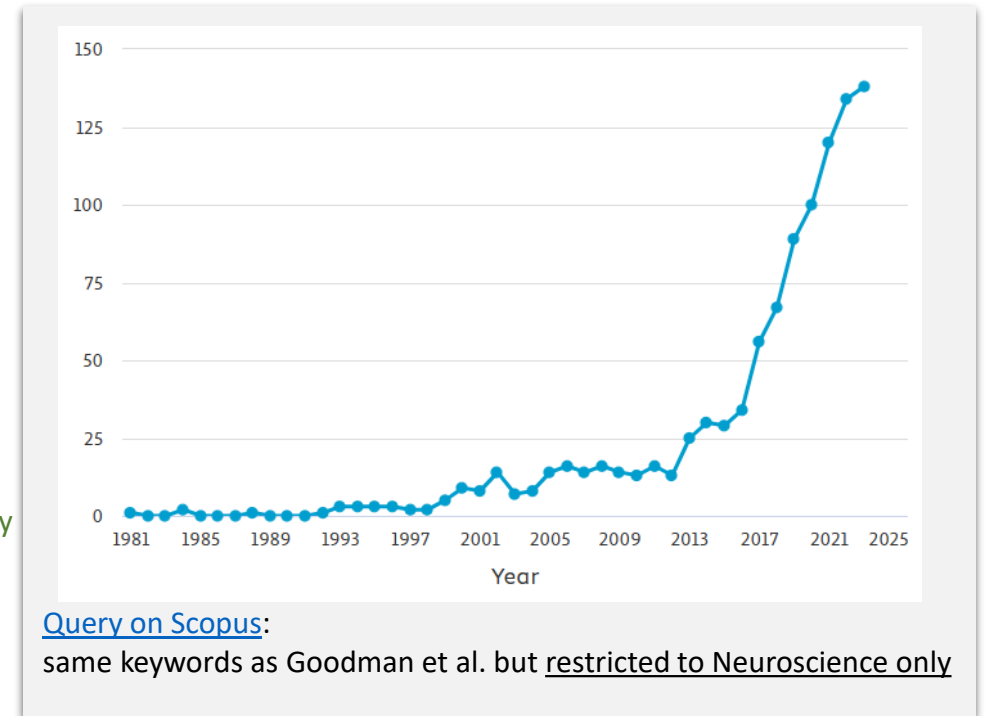
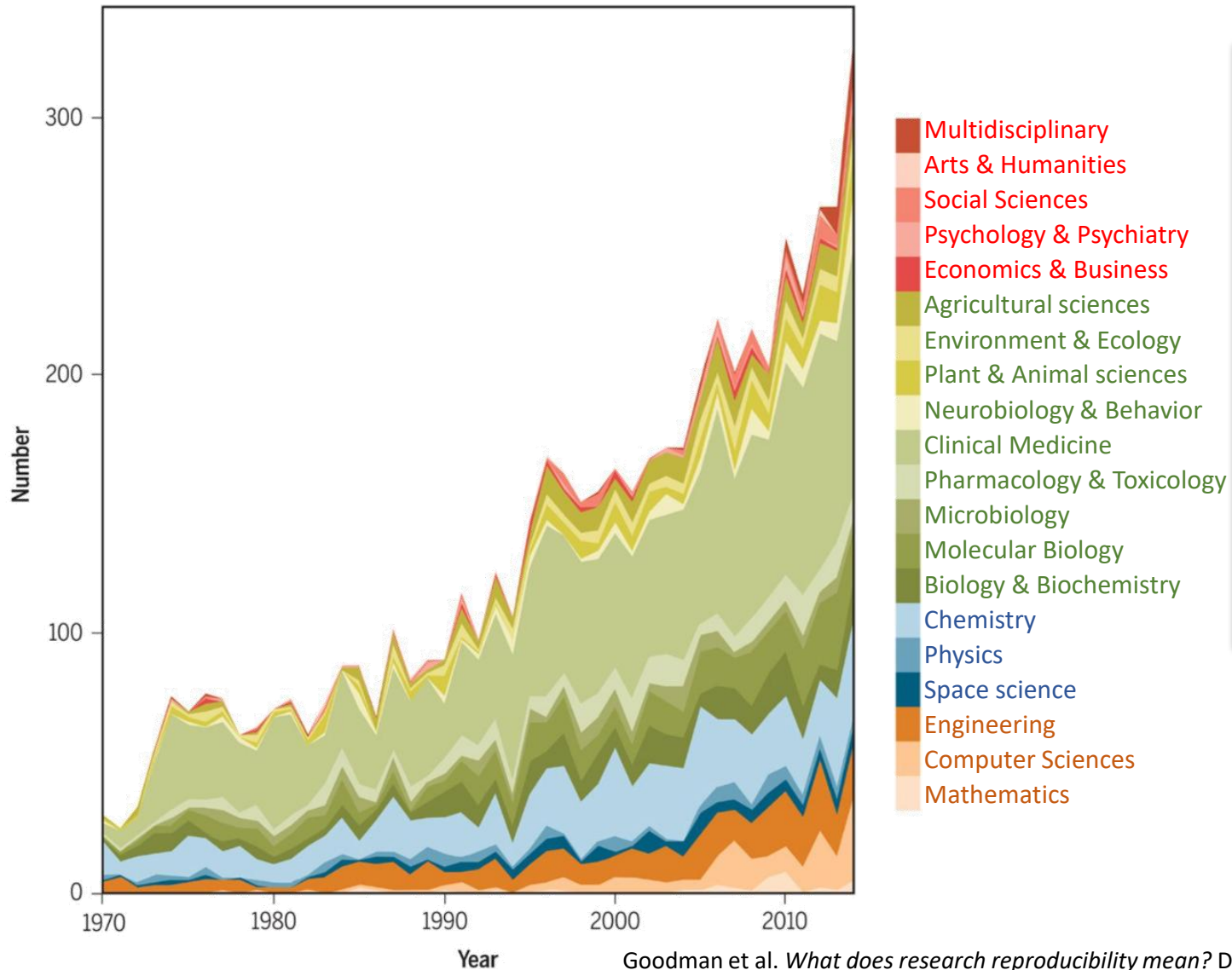


**Key message: reproducibility is not black and white, but rather a continuum**

Peng (2011) *Reproducible Research in Computational Science* <https://doi.org/10.1126/science.1213847>  
 Image credit: icon made by Freepik on [www.flaticon.com](http://www.flaticon.com)

# Research reproducibility

**Keywords** (at least 1 of the following expressions): research reproducibility, reproducibility of research, reproducibility of results, results reproducibility, reproducibility of study, study reproducibility, reproducible research, reproducible finding, or reproducible result



**Conclusion:** “any scientific field, require a clear specification of the kind of reproducibility being discussed (methods, results, or inferences)”

# Why Most Published Research Findings Are False

John P. A. Ioannidis

74,946 Save	7,388 Citation
3,001,940 View	9,576 Share

## Inadequate methodologies

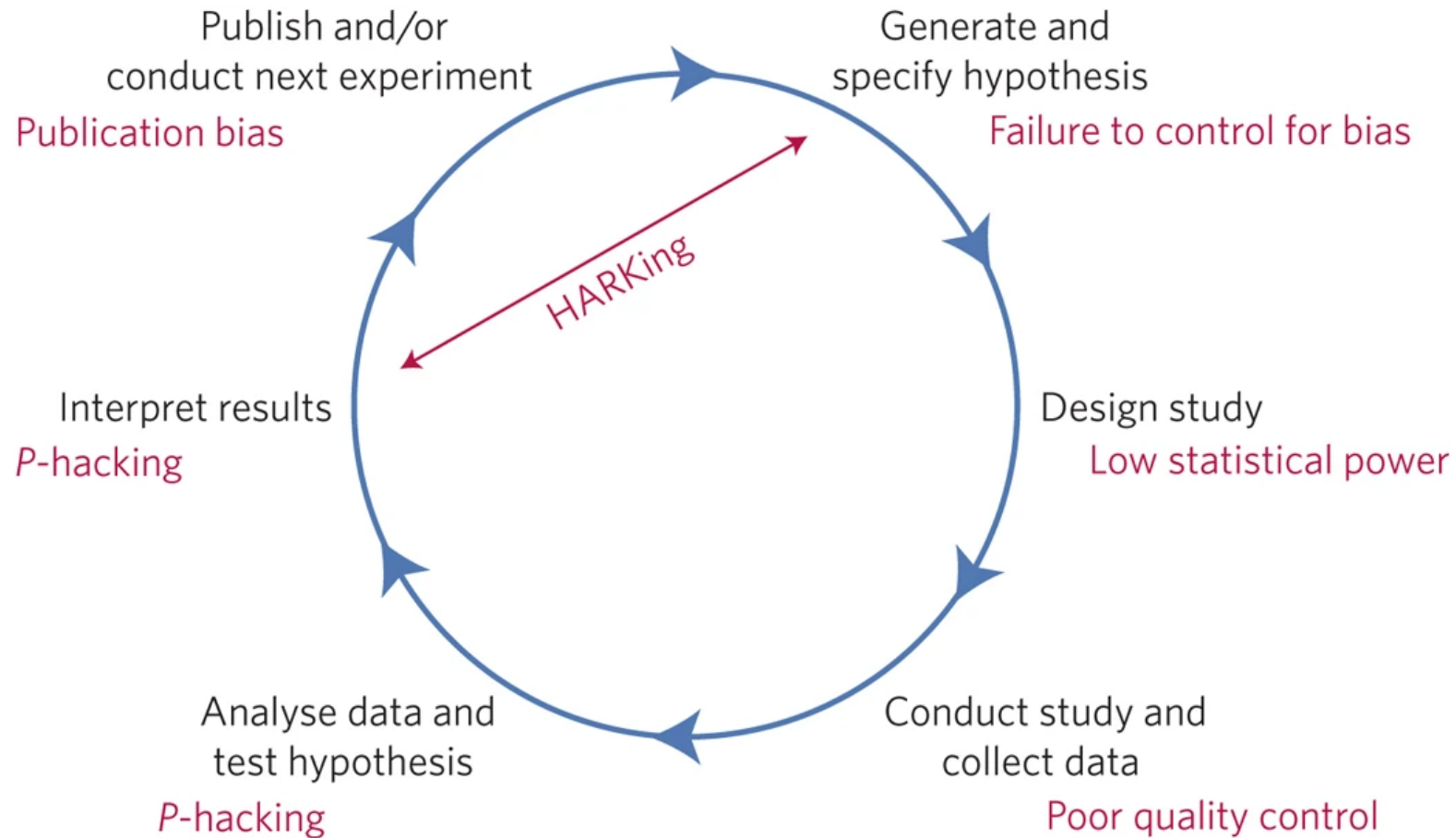
A result is all the more likely to be false when:

- **effect sizes** are small
- there is a **greater number and lesser preselection** of tested relationships
- there is **greater flexibility** in designs, definitions, outcomes, and analytical modes
- there is greater **financial and other interest** and prejudice

+ the propensity of **journals to publish only positive results**

+ the "**publish or perish**" culture for researchers

# Lack of reproducibility: can occur at all steps...



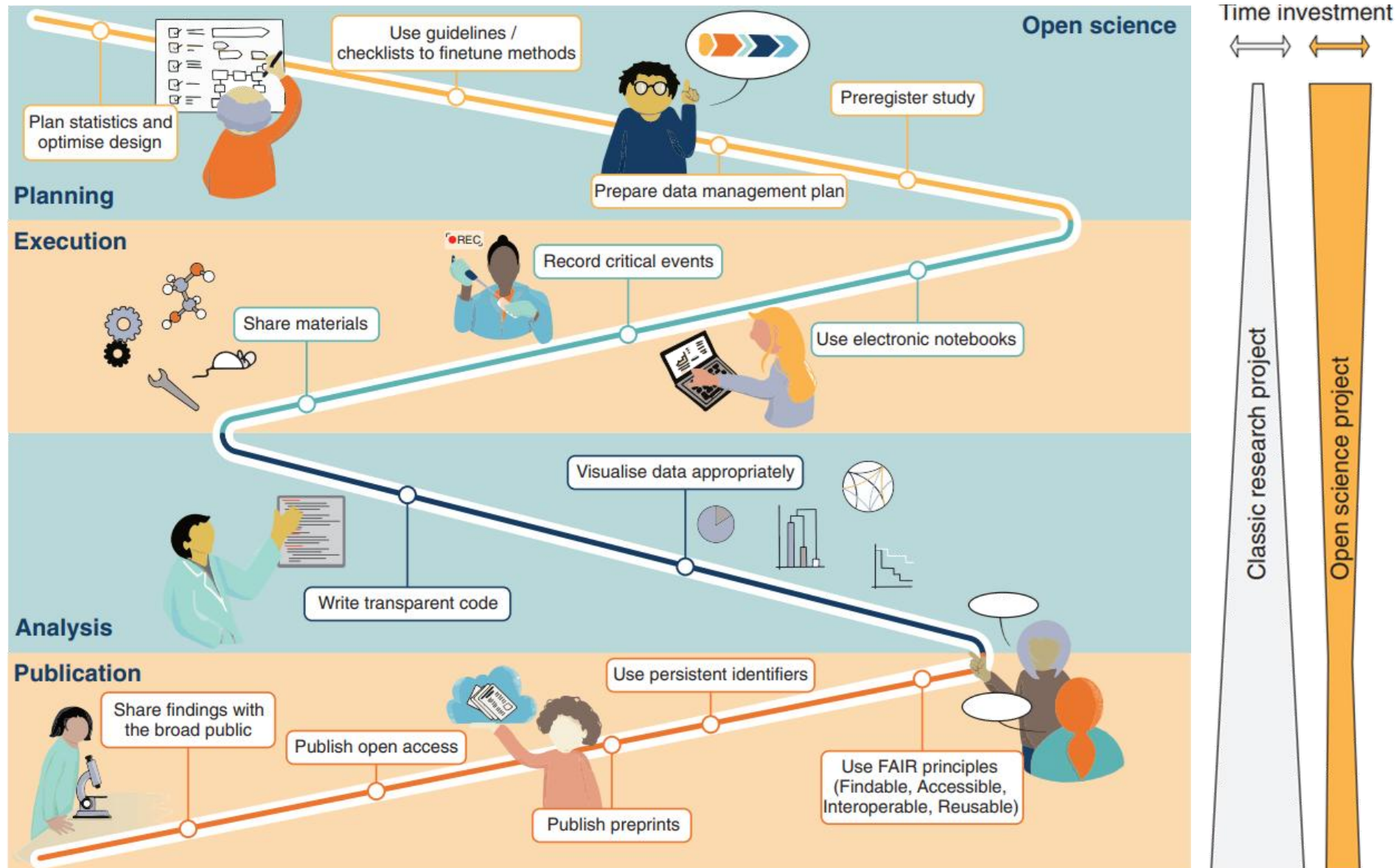
## Legend

Blue: An idealized version of the hypothetico-deductive model of the scientific method

Pink: threats to this model

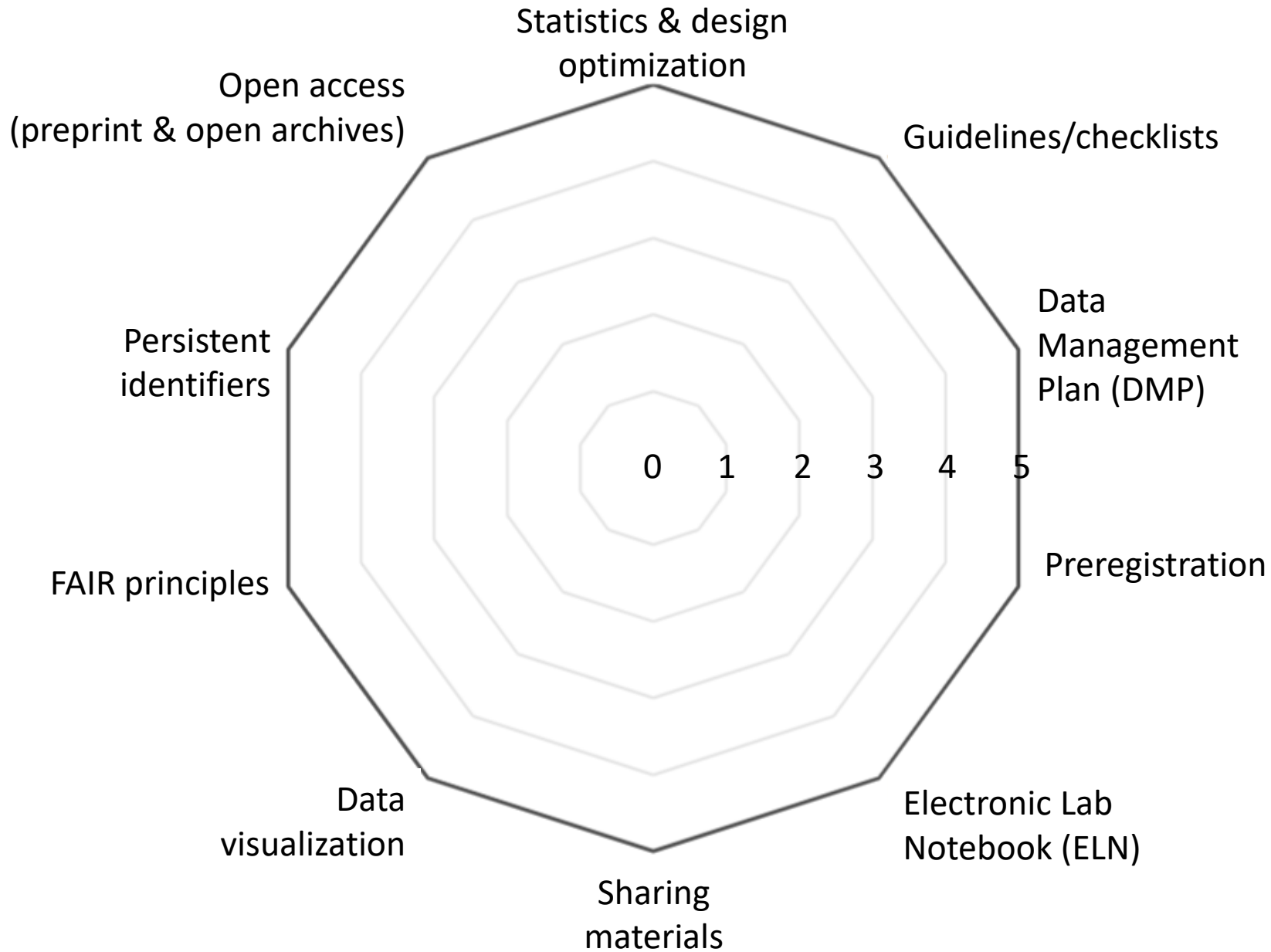
Munafò *et al.* *A manifesto for reproducible science.* (2017). doi: 10.1038/s41562-016-0021

# Overview of *(the main)* open science practices throughout the research lifecycle



Bertram et al. 2023. Open science. <https://doi.org/10.1016/j.cub.2023.05.036>

# Open science practices: self-evaluation through the radar technique



## ~~Science~~ Broken Science

- Paywall
- Non-shared data
- Closed tools & Software
- Non-reproducible
- Anonymous peer-review
- Publish or Perish



**ReScience**  
Reproducible science is good. Replicated science is better.

## ~~Open science~~ Science

- Open access
- Open data
- Open source
- Open methodology
- Open peer-review
- Much more fun...



Roesch et al. (2019) *ReScience X: Redoing experiments*. [osf.io/ede88](https://osf.io/ede88) | slides: [osf.io/uybc6/](https://osf.io/uybc6/)

1. Introduction

**2. Preregistration**

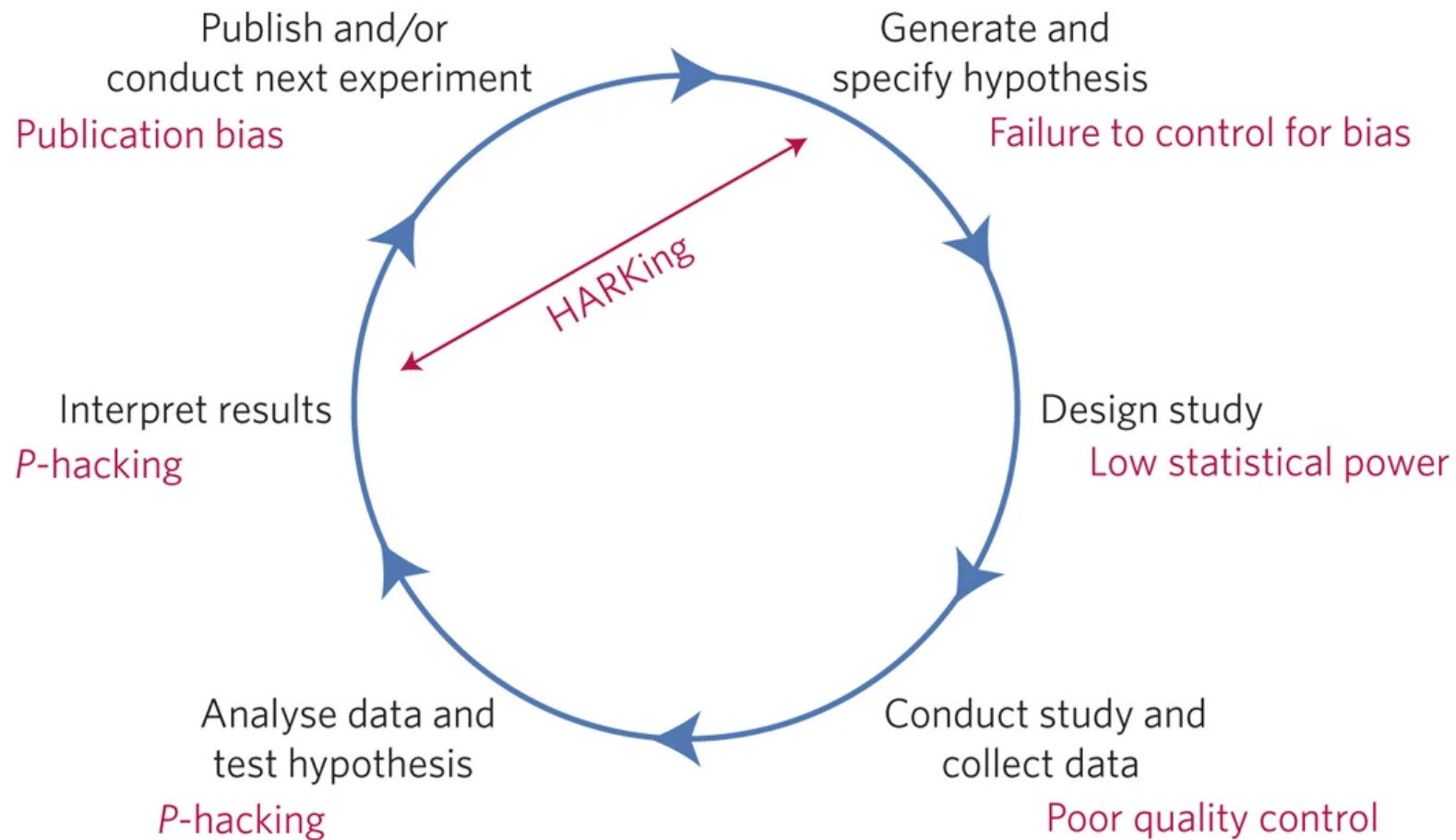
3. Data Management Plan

4. Standardisation

5. Documentation

6. Resources

# What happens when researchers are pressured to get « good results »?

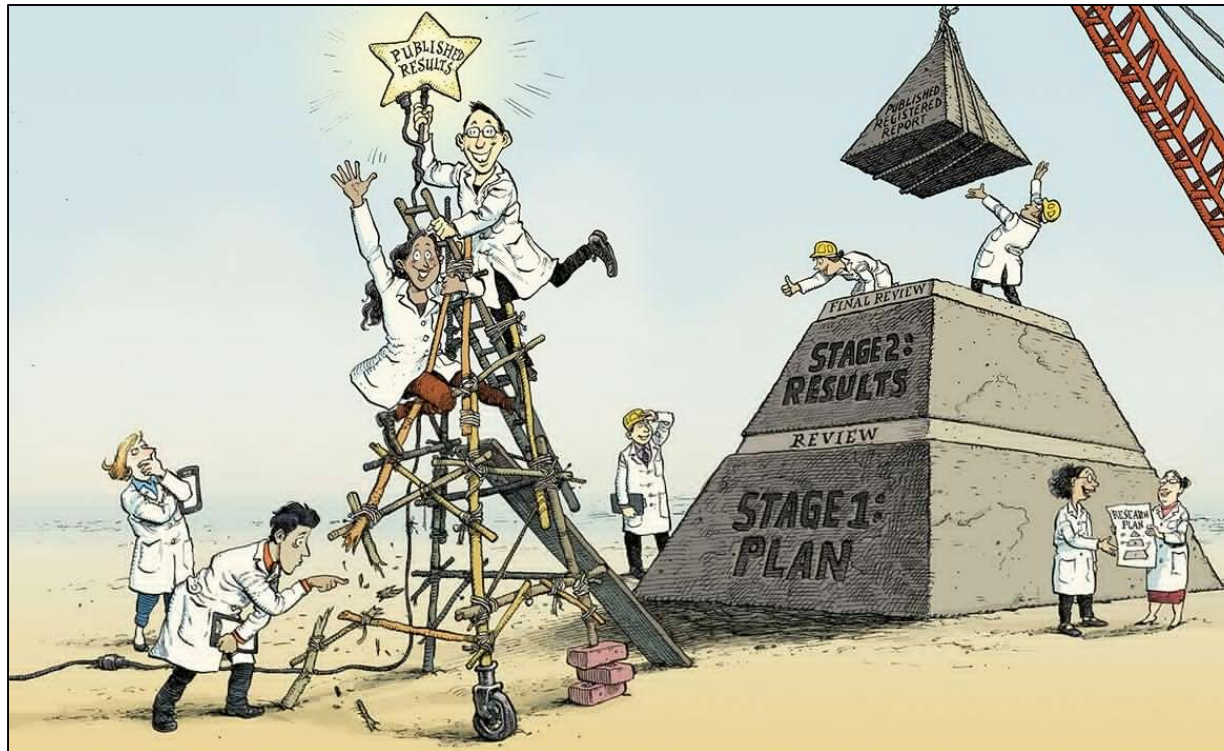


## Legend

Blue: An idealized version of the hypothetico-deductive model of the scientific method

Pink: threats to this model

Munafò et al. *A manifesto for reproducible science*. (2017). doi: 10.1038/s41562-016-0021



→ 2 types of preregistration:

1/ Reviewed PreRegistration (RRs aka “Registered Report”)

Registration is submitted to a journal and peer-reviewed by experts

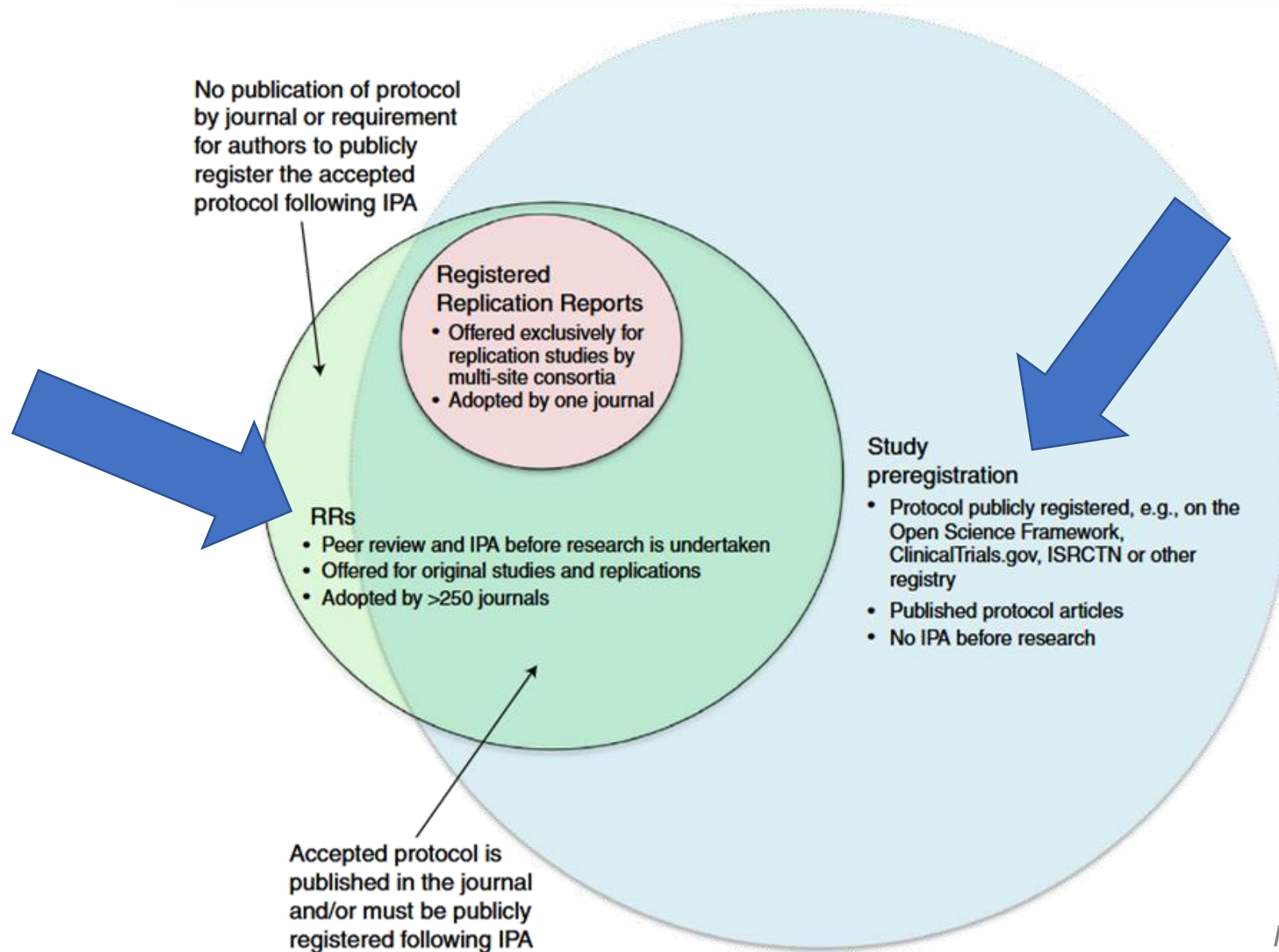
2/ Unreviewed PreRegistration (UPR)

Preregistration is deposited in a public or private repository

[Adapted from slides by Justine Épinat-Duclos](#)

CRNL NeuroImaging Club meeting, 2021

# Registered reports vs preregistration: how to spot the difference



*IPA: In Principle Acceptance*

Chambers & Tzavella. *The past, present and future of Registered Reports*. 2021. <https://doi.org/10.1038/s41562-021-01193-7>



## → For Reviewed & Unreviewed Preregistration:

Open Science Framework (OSF):



<https://osf.io/>

Preregistration in OSF

<https://www.cos.io/initiatives/prereg>



Registered Report in OSF

<https://www.cos.io/initiatives/registered-reports>

Registration forms and templates



## → For Unreviewed Preregistration:



<https://aspredicted.org/>



HOME

Nosek et al. (2018) *The preregistration revolution*. <https://doi.org/10.1073/pnas.1708274114>

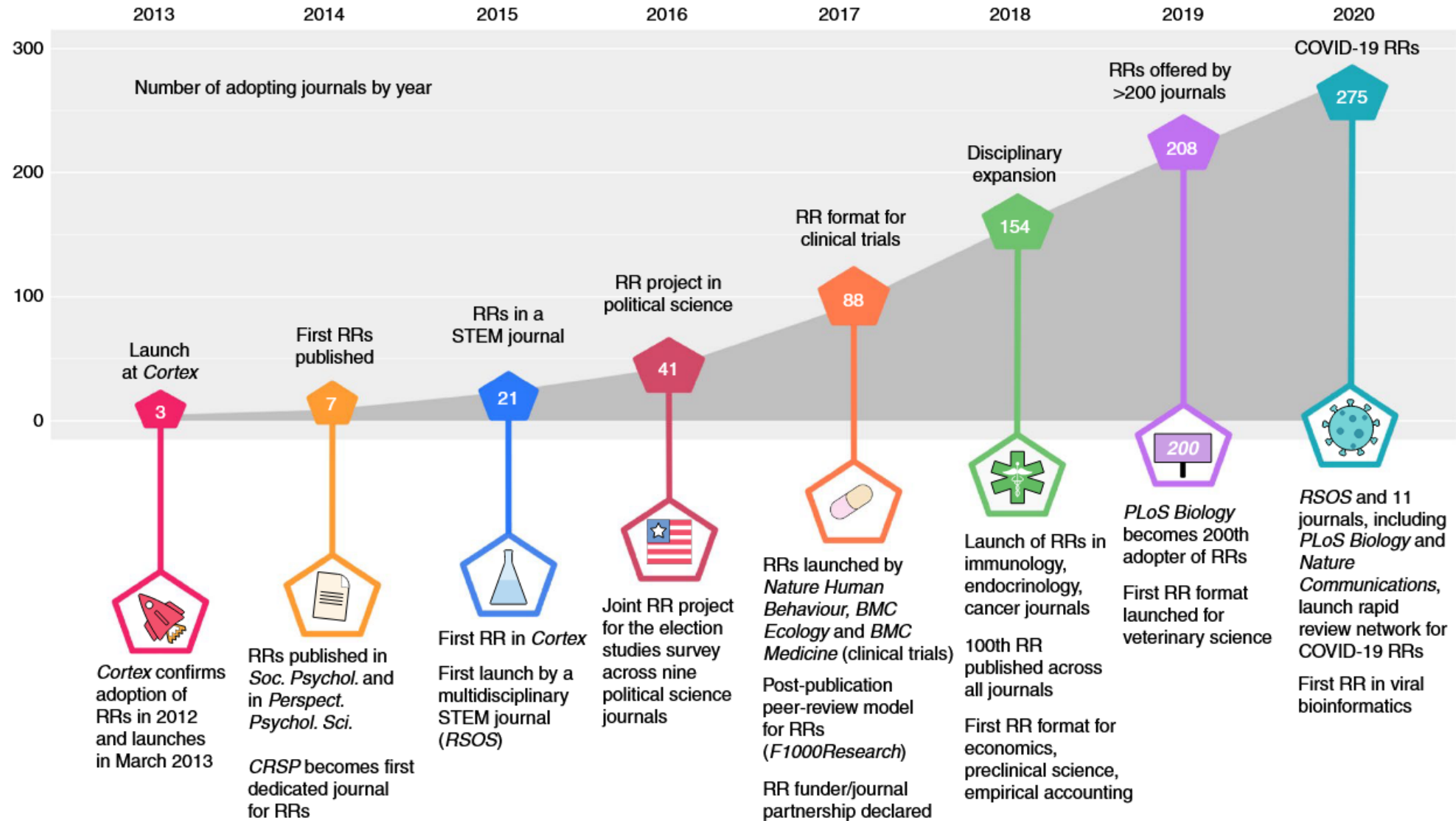
Nosek & Lakens (2014) *Registered reports: a method to increase the credibility of published results*

<https://doi.org/10.1027/1864-9335/a000192>

Adapted from slides by Justine Épinat-Duclos

CRNL NeuroImaging Club meeting, 2021

# Key milestones in the evolution of “registered reports” (RRs)



Chambers & Tzavella. *The past, present and future of Registered Reports*. 2021. <https://doi.org/10.1038/s41562-021-01193-7>



# Why?



## Benefits for science & scientific community

- Rigorous review of theory and methods
- Elimination of publications with a strong bias and/or reporting bias
- Clear distinction between exploratory and confirmatory research

## Benefits for the researcher

- Peer review when it is most helpful
- Pre-analysis review feedback (RR only)
- Publication guaranteed regardless of the results (RR only)

## Benefits for students

- Pedagogical: you really learn how to do science properly
- Producing outcomes sooner
- Add something on your CV that has already undergone peer review and is almost complete.

(Simmons et al. 2021)

More information: *Registered Reports and PhD's – What? Why? How?* An Interview with Chris Chambers  
<https://www.cos.io/blog/rrs-phds-what-why-how-interview-chambers>

Adapted [slides by Justine Épinat-Duclos](#)  
CRNL NeuroImaging Club meeting, 2021

# Pre-registration of a project with peer-reviews: “registered reports”

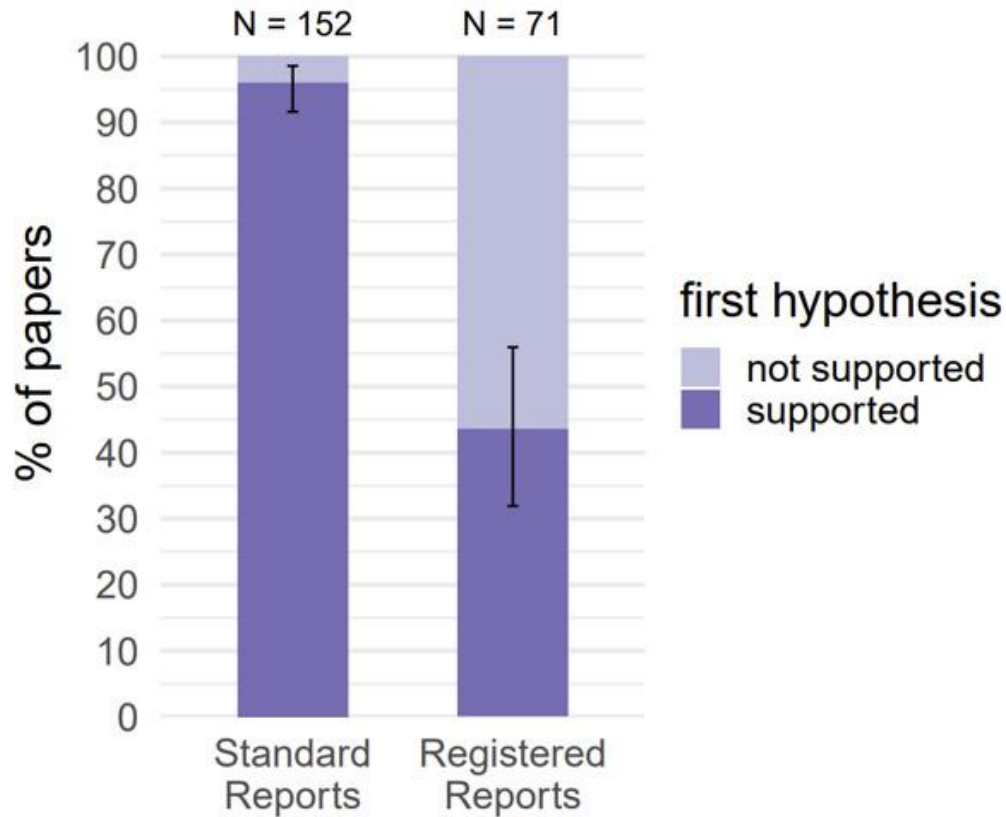
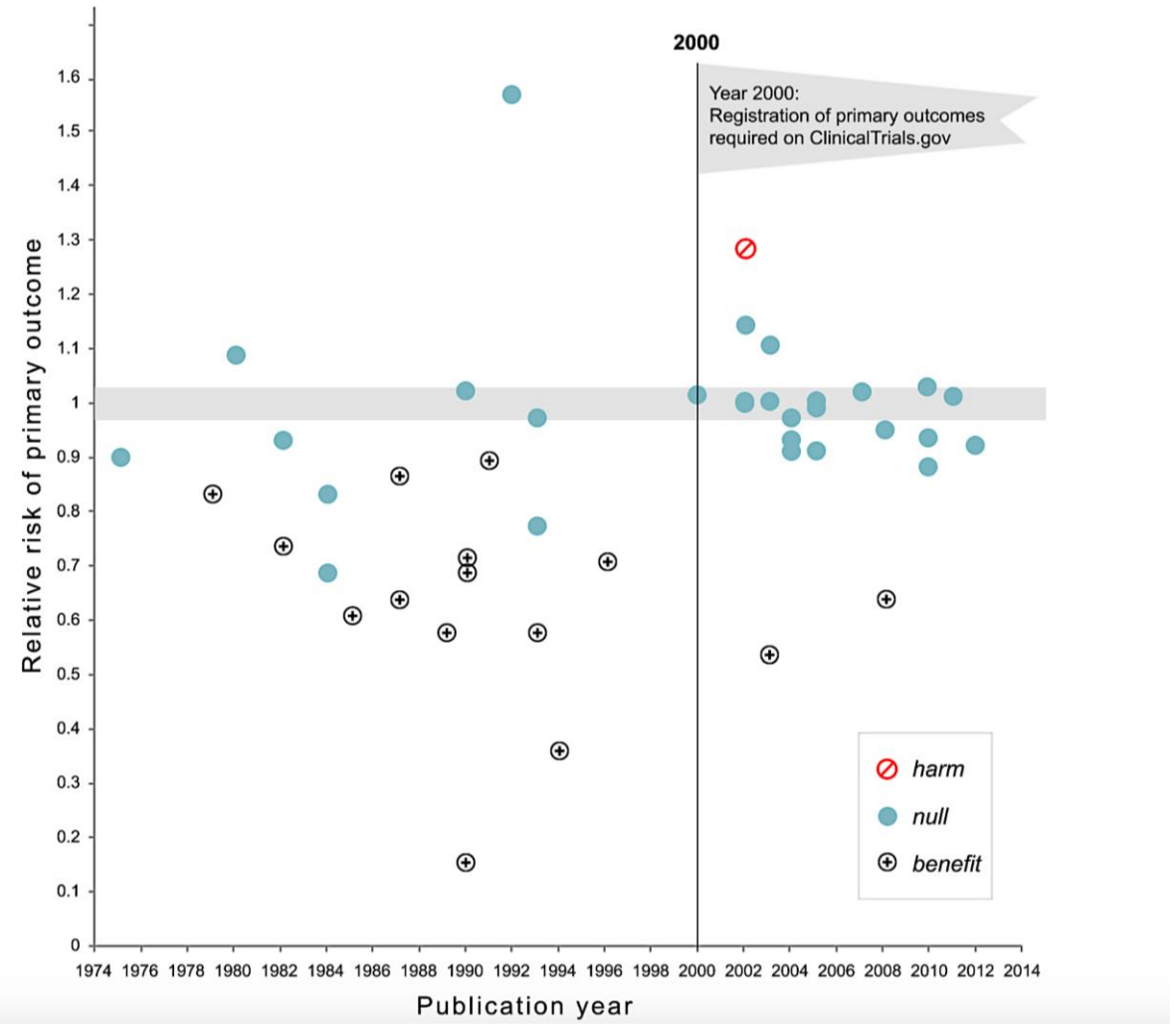


Figure 2. Positive result rates for standard reports and Registered Reports. Error bars indicate 95% confidence intervals around the observed positive result rate.

Scheel et al. (2021) *An excess of positive results: Comparing the standard Psychology literature with Registered Reports*. <https://doi.org/10.1177/25152459211007467>

## Effects of Preregistration



### Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time

Robert M. Kaplan, Veronica L. Irvin

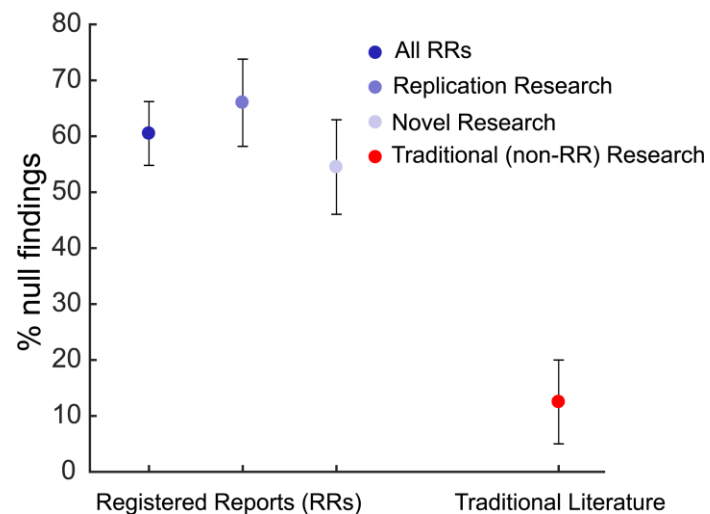
Published: August 5, 2015 • <https://doi.org/10.1371/journal.pone.0132382>

Adapted slides by Justine Épinat-Duclos  
CRNL NeuroImaging Club meeting, 2021

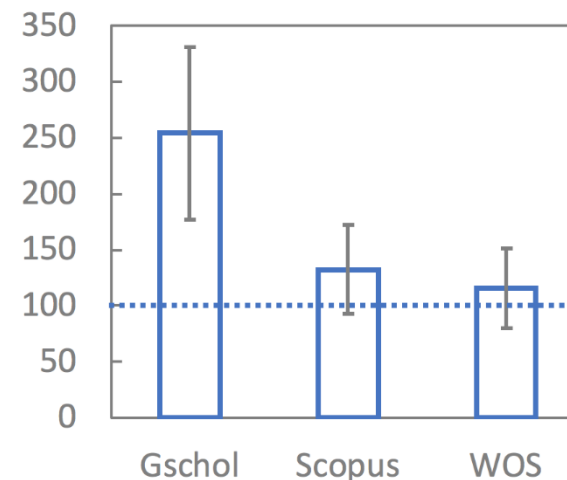
# Positive effects

Registered Reports appear to be working as intended

Percentage of null findings



% citations relative to JIF



Well cited -- at or above respective journal impact factor

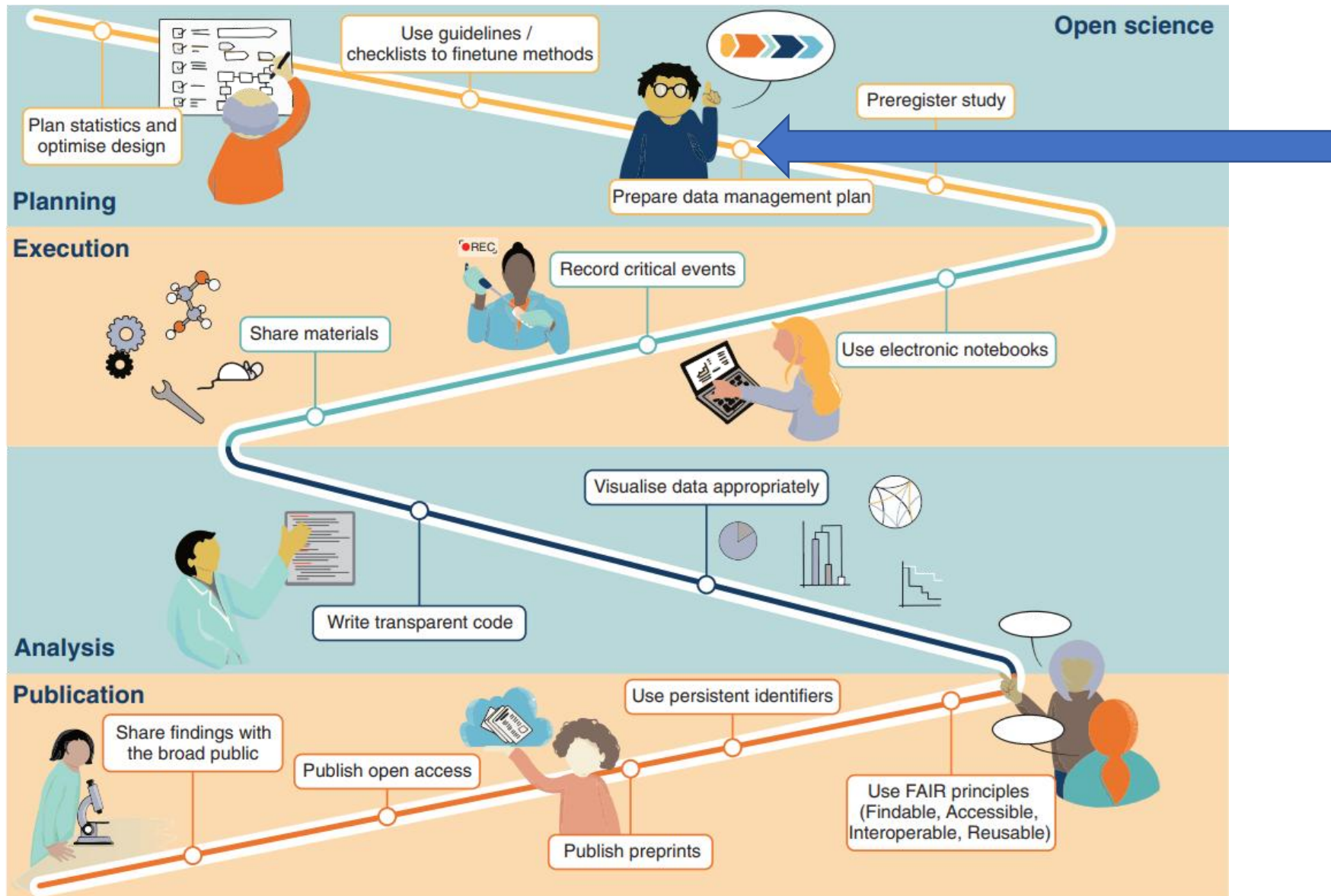
<https://tinyurl.com/RR-citations>

Hypotheses are ~5 times more likely to be **unsupported** in Registered Reports compared with regular articles

(Allen & Mehler; 2019)

Adapted [slides by Justine Épinat-Duclos](#)  
CRNL NeuroImaging Club meeting, 2021

1. Introduction
2. Preregistration
- 3. Data Management Plan**
4. Standardisation
5. Documentation
6. Resources

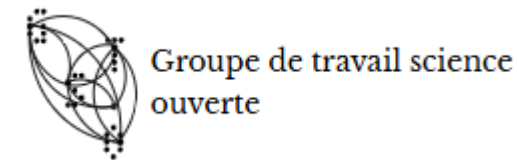


Bertram et al. 2023. Open science. <https://doi.org/10.1016/j.cub.2023.05.036>

# Best advice #2: write a Data Management Plan (DMP) to anticipate the future!



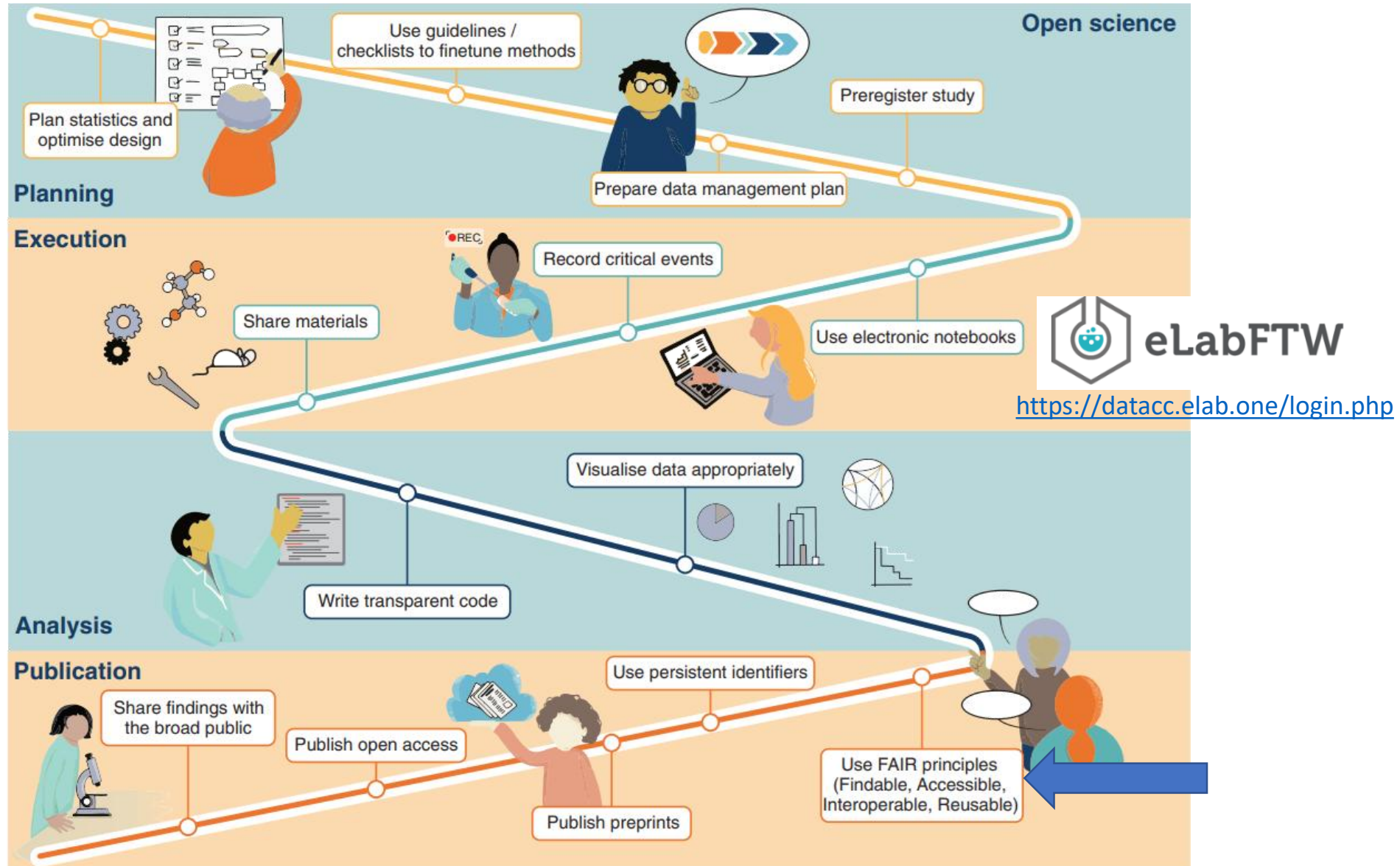
[Source](#)



<https://gtso.couperin.org/gtdonnees/sos-pgd/>

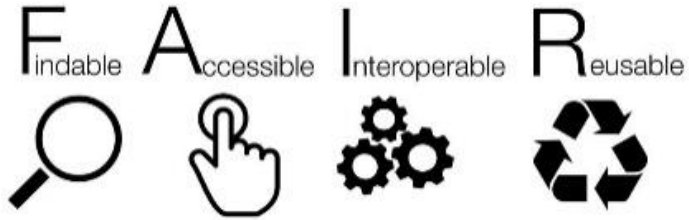
1. Introduction
2. Preregistration
3. Data Management Plan
- 4. Standardisation**
5. Documentation
6. Resources

# Standardization can be applied almost at all steps!



Bertram et al. 2023. Open science. <https://doi.org/10.1016/j.cub.2023.05.036>

# FAIR principles



Comment | [Open access](#) | Published: 15 March 2016

## The FAIR Guiding Principles for scientific data management and stewardship

[Mark D. Wilkinson](#), [Michel Dumontier](#), [IJsbrand Jan Aalbersberg](#), [Gabrielle Appleton](#), [Myles Axton](#), [Arie Baak](#), [Niklas Blomberg](#), [Jan-Willem Boiten](#), [Luiz Bonino da Silva Santos](#), [Philip E. Bourne](#), [Jildau Bouwman](#), [Anthony J. Brookes](#), [Tim Clark](#), [Mercè Crosas](#), [Ingrid Dillo](#), [Olivier Dumon](#), [Scott Edmunds](#), [Chris T. Evelo](#), [Richard Finkers](#), [Alejandra Gonzalez-Beltran](#), [Alasdair J.G. Gray](#), [Paul Groth](#), [Carole Goble](#), [Jeffrey S. Grethe](#), ... [Barend Mons](#)  [+ Show authors](#)

<https://doi.org/10.1038/sdata.2016.18>



The Turing Way project illustration by Scriberia. Used under a CC-BY 4.0 licence.  
DOI: [10.5281/zenodo.3332807](https://doi.org/10.5281/zenodo.3332807)



- ❖ The data can not always be opened, but it benefits a minimum to be FAIR.
- ❖ FAIR data may not be shared and open... but open data must be FAIR.
- ❖ Metadata should always be FAIR however.

Image credit: icons made by Freepik, VectorPortal, SBTS2018 on [www.flaticon.com](http://www.flaticon.com)

# File naming is a serious issue... for scientists

Ziemann et al. *Genome Biology* (2016) 17:177  
DOI 10.1186/s13059-016-1044-7

Genome Biology

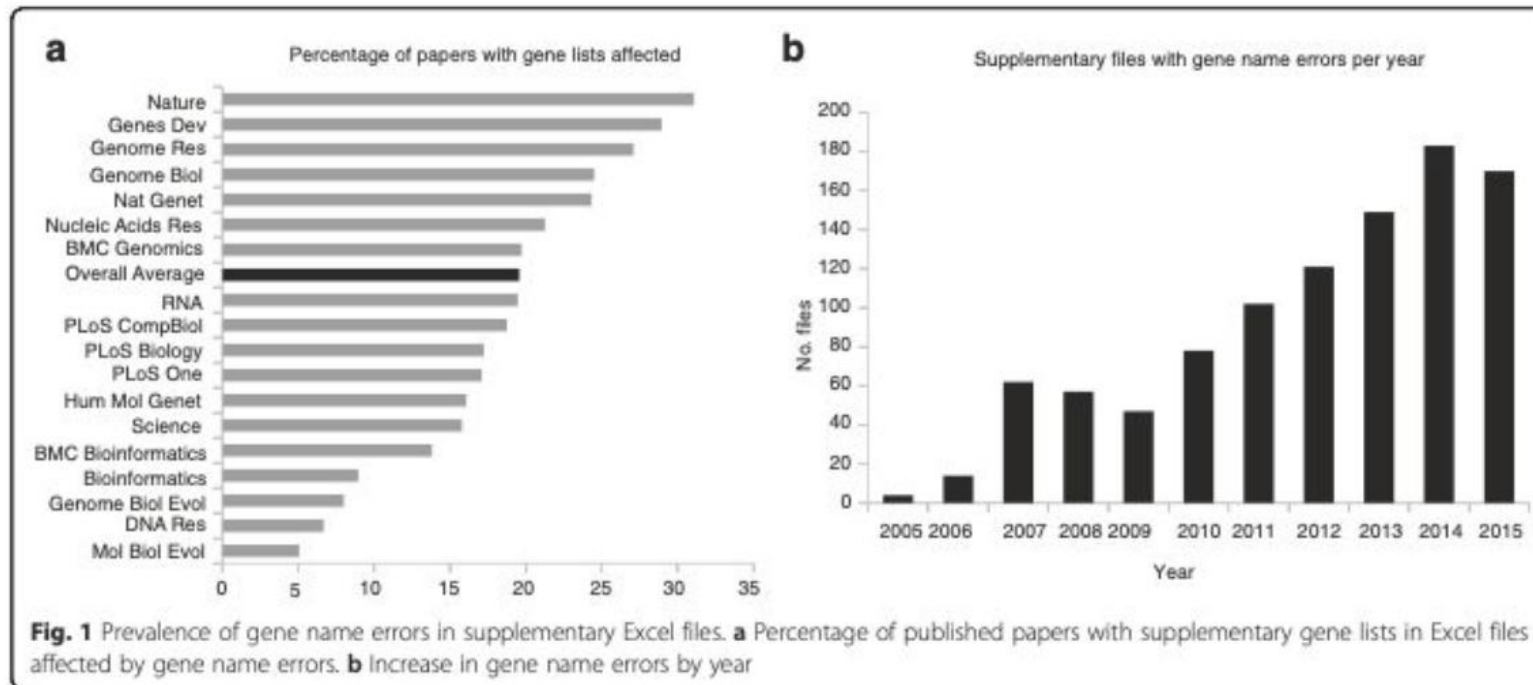
COMMENT

Open Access



## Gene name errors are widespread in the scientific literature

Mark Ziemann<sup>1</sup>, Yotam Eren<sup>1,2</sup> and Assam El-Osta<sup>1,3\*</sup>



[Inserm webinar](#)

09/2025. Data management & DMP

## Research data is a serious issue... for citizens



[https://www.youtube.com/watch?v=yeX\\_Zs7zztY](https://www.youtube.com/watch?v=yeX_Zs7zztY)

<https://scienceetonnante.com/2020/04/17/austerite-excel/>

<https://github.com/scienceetonnante/Reinhart-Rogoff>

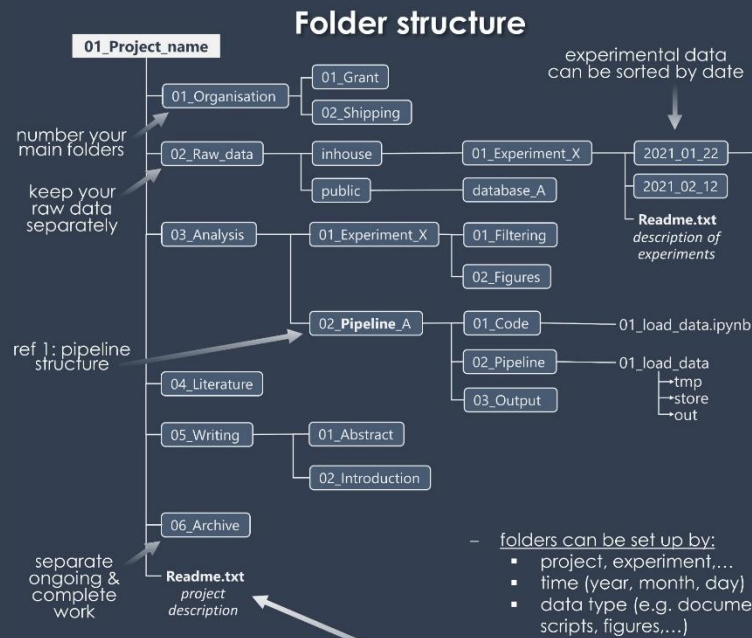


GOAL of good data management  
→ optimise the discovery & reuse of data

## Questions to ask yourself



Are my files organised in a way that I can easily find what I am searching for?  
What information would I need to understand and use my data in 20 years?  
Could others understand and use my data?



## File naming



### General naming tips for folders & files

- use unique, meaningful names
- not too long (not >30-40 characters)
- no spaces, dots, or special characters (\$!"&\*"()+=[:;~@)
- hyphens (-) & underscores (\_) to separate elements

- find the balance between a shallow & deep folder hierarchy
  - too deep → too many clicks might be needed to get to the right file
  - too shallow → too many files might end up in one folder (organise them in subfolders)



### Friendly Reminder

Comment your code! 😊

## Metadata

Which information is necessary to interpret, understand, and use a given dataset?

readme.txt files can be used to describe projects, folders, and files

## References

- [1 https://towardsdatascience.com/how-to-keep-your-research-projects-organized-part-1-folder-structure-10bd56034d3a](https://towardsdatascience.com/how-to-keep-your-research-projects-organized-part-1-folder-structure-10bd56034d3a)
- <https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Doing/Organising-files-and-folders.htm>
- <https://www.massey.ac.nz/massey/research/library/library-services/research-services/manage-data/organise.cfm>
- <https://library.bath.ac.uk/research-data/working-with-data/organising-data>
- <https://www.helsinki.fi/en/research/organizing-data-folders-with-5sdata-method>
- <https://mantra.edina.ac.uk>
- <https://old.dataone.org/education-modules>

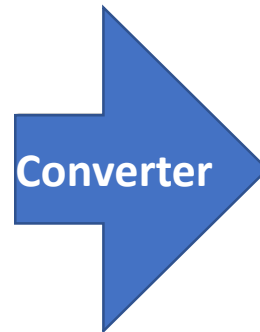
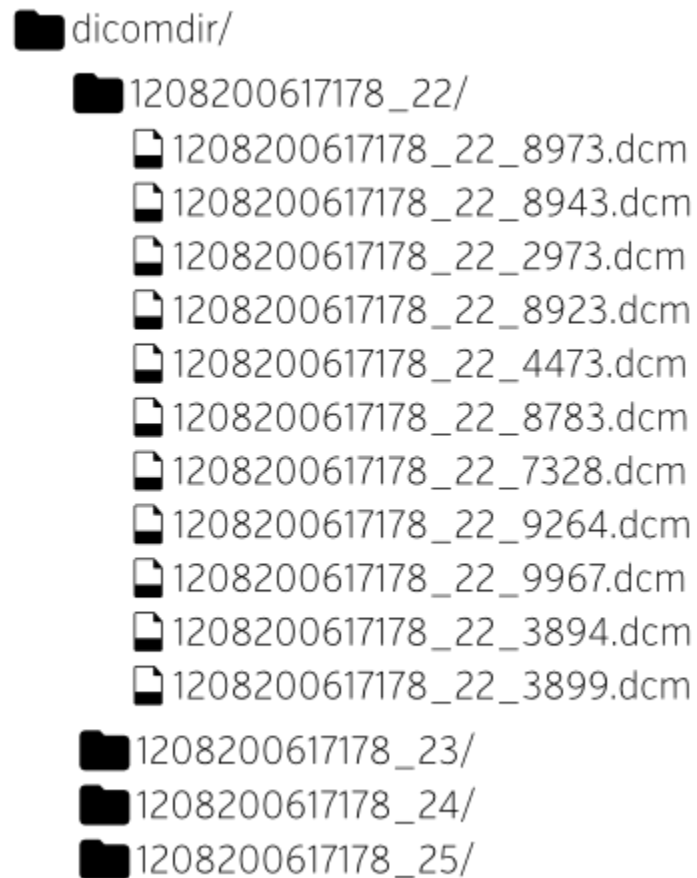


## Important information

- Who** has created the data?
- What** is the content of the data?
- Which** questions have been answered?
- When** were the data created?
- How** were the data developed (methods)?
- Why** were the data developed?
- With whom** can the data be shared?

Source: [Kira Höffler](#)

# Description of the BIDS standard



[List of converters](#)

# Description of the standard: file naming

key1 - value1 \_ key2 - value2 \_ suffix .extension

[source](#)


- **Suffixes** are preceded by an **underscore**
- Entities are composed of **key-value** pairs separated by **underscores**
- There is a limited set of **suffixes** for each data type (anat, func, eeg, ...)
- For a given **suffix**, some entities are **required** and some others are **[optional]**.
- **Keys**, **value** and **suffixes** can only contain letters and/or numbers.
- Entity **key-value** pairs have a specific order in which they must appear in filename.
- Some entities **key-value** can only be used for derivative data.

## [anat filename template](#)

sub-<label>[\_ses-<label>][\_task-<label>][\_acq-<label>][\_ce-<label>][\_rec-<label>][\_run-<index>][\_part-<mag|phase|real|imag>]\_<suffix>.json  
sub-<label>[\_ses-<label>][\_task-<label>][\_acq-<label>][\_ce-<label>][\_rec-<label>][\_run-<index>][\_part-<mag|phase|real|imag>]\_<suffix>.nii[.gz]

## [func filename template](#)

sub-<label>[\_ses-<label>]\_task-<label>[\_acq-<label>][\_run-<index>]\_bold.nii.gz  
sub-<label>[\_ses-<label>]\_task-<label>[\_acq-<label>][\_run-<index>]\_bold.json  
sub-<label>[\_ses-<label>]\_task-<label>[\_acq-<label>][\_run-<index>]\_events.json  
sub-<label>[\_ses-<label>]\_task-<label>[\_acq-<label>][\_run-<index>]\_events.tsv



Legend:  
< to\_be\_specified >  
[ optional ]

Adapted from slides from Remi Gau. 22/11/2023. BIDS. <https://allmylinks.com/remi-gau>

# Description of the standard: open formats

Prefer formats that can be opened by most software, including simple text editors.

Stay away from **.docx, .xlsx, .mat, .pickle...**

⇒ **TSV files** : Tabulation Separated Values

Ex: **participants.tsv**

participant_id	sex	age
sub-01	M	28
sub-02	F	21
sub-03	F	27
sub-04	M	25
sub-05	F	20
sub-06	M	20

⇒ **JSON files**: JavaScript Object Notation

- for attribute-values pairs
- as “side-car” to a data file

Ex: **\*\_events.json**

```
{
  "trial_type": {
    "LongName": "Emotion image type",
    "Description": "Type of emotional face from Karolinska database that is displayed",
    "Levels": {
      "afraid": "A face showing fear is displayed",
      "angry": "A face showing anger is displayed",
      "sad": "A face showing sadness is displayed"
    }
  },
  "identifier": {
    "LongName": "Karolinska (KDEF) database identifier",
    "Description": "ID from KDEF database used to identify the displayed image"
  }
}
```

# BIDS validator

<https://bids-standard.github.io/bids-validator/>

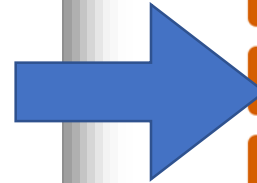
## BIDS Validator

Select a BIDS dataset to validate.

Select Dataset Files

*BIDS Validator version: 2.2.0*

Note: Selecting a dataset only performs validation. Files are never uploaded.  
The legacy BIDS Validator is available [here](#).



## BIDS Validator

Reselect Files

▶ ERROR: TSV\_VALUE\_INCORRECT\_TYPE (EDINBURGH\_HANDEDNESS)

▶ ERROR: TSV\_VALUE\_INCORRECT\_TYPE (CONTRACEPTION)

▶ ERROR: TSV\_VALUE\_INCORRECT\_TYPE (ONSET)

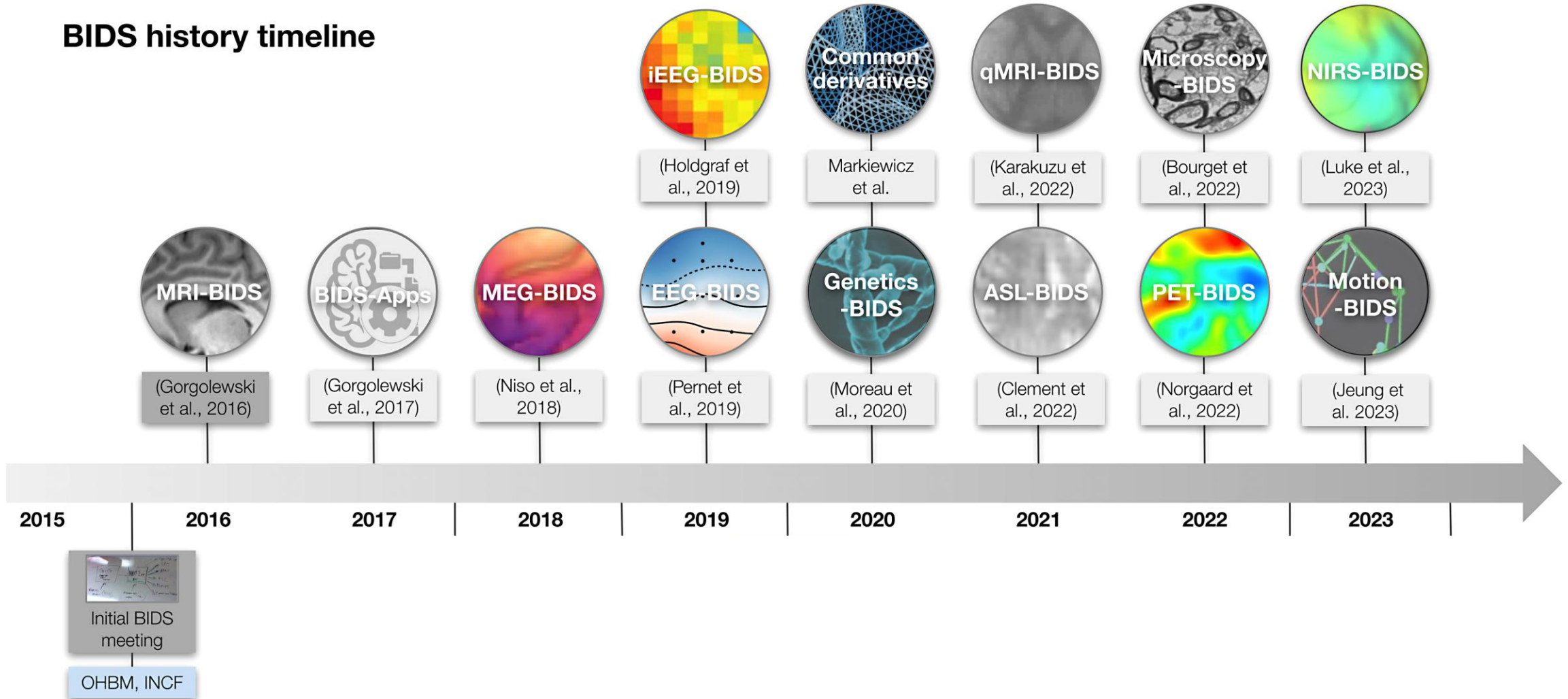
▶ ERROR: TSV\_VALUE\_INCORRECT\_TYPE (TRIAL\_TYPE)

▶ WARNING: JSON\_KEY\_RECOMMENDED (GENERATEDBY)

▶ WARNING: JSON\_KEY\_RECOMMENDED (SOURCEDATASETS)

# Description of the BIDS standard

## BIDS history timeline



Poldrack et al. (2024) *The past, present, and future of the brain imaging data structure (BIDS)* [https://doi.org/10.1162/imag\\_a\\_00103](https://doi.org/10.1162/imag_a_00103)

# Best advice #3: use standards when available

<https://www.incf.org/resources/sbps>

## BIDS-App

<https://bids.neuroimaging.io/tools/bids-apps.html>

Among the list:

MRIQC (Esteban et al., 2017)

fMRIPrep (Esteban, Markiewicz, et al., 2019)

QSIPrep (Cieslak et al., 2021)

PyMVPA BIDS-App (Torabian et al., 2023)



ABOUT PRODUCTS SECRETARIAT SERVICES ACTIVITIES NETWORK JOIN/SUPPORT NEWS Q

### INCF Standards and Best Practices portfolio



The Standards and Best Practices portfolio contains all community standards endorsed by the INCF using the criteria developed by the INCF Standards and Best Practices Committee, as well as community standards in the process of being endorsed.

The purpose of this portfolio is to provide the community with an index of robust, well-validated standards and best practices that adhere and support the FAIR principles. The portfolio provides the community with descriptions of appropriate use cases, links to tools/infrastructures, and tutorials for each standard and best practice indexed. Users can search the portfolio by data type, format, or subdomain of neuroscience. By offering this service, we hope to increase the discoverability and use of standards and best practices that support open and FAIR neuroscience.

- Submit a Standard
- How standards are endorsed
- Join the Standards reviewer pool
- Recommend a new Standards area
- INCF Infrastructure Portfolio

19 SBPs Submitted	3 SBPs Under review	13 SBPs Endorsed
----------------------	------------------------	---------------------

- All
- 3D reconstruction
- Anatomy
- Best practices
- Clinical observational studies
- Computational neuroscience
- Data format
- Data management
- Data model
- EEG
- Electrophysiology
- Experimental design
- fMRI
- HDF5
- Mass-univariate studies
- MATLAB
- MEG
- Metadata
- Metadata format
- Microscopy
- Modeling
- MRI
- NEST
- Neuroimaging
- Neuromorphology
- NEURON
- Neuronal modeling
- PET
- Physiology
- Provenance
- Python
- Quality control
- Research resource identifiers
- Simulation
- Standardized vocabularies
- Standards
- XML

1. Introduction
2. Preregistration
3. Data Management Plan
4. Standardisation
- 5. Documentation**
6. Resources



# Metadata

Metadata can be used to **describe data** more accurately.  
These are “**data on data**”.



2 types of metadata:

- **embedded metadata**: these are generated **automatically** by cameras, sound and measuring devices. This is typically the case for the photo or video of the smartphone.  
Examples of metadata generated: GPS data, device type, date, technical calibration, etc.
- **enriched metadata**: these are added **by the author(s)**.  
Examples: keywords, subject, author, laboratory or organization, project name, license, provenance, etc.

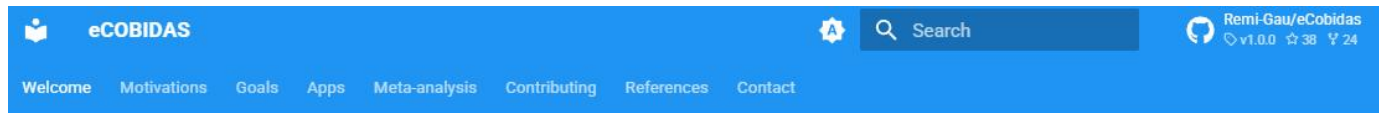
Source : [Doranum](#)

# Checklist and guidelines

## eCOBIDAS checklist

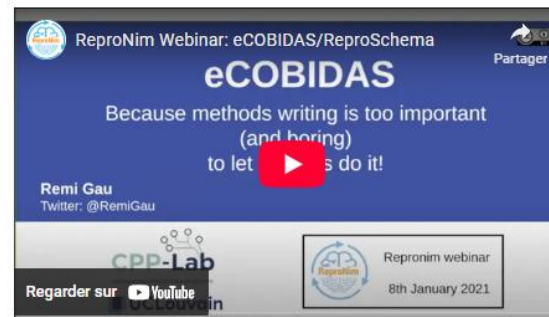
### Committee On Best practices In Data Analysis and Sharing

1. experimental design reporting,
2. image acquisition reporting,
3. preprocessing reporting,
4. statistical modeling,
5. results reporting,
6. data sharing,
7. reproducibility



DOI:10.17605/OSF.IO/ANWQY

Welcome to the eCOBIDAS documentation



#### Motivations

Poor methods and results description hinders the reproducibility and the replicability of research.

This project is built on the hope that improving methods and results reporting could improve our research.

[→ Motivations](#)

<https://ecobidas.readthedocs.io/en/latest/>

#### Goals

We want to create a web-based checklists for best practices for methods reporting that help automatically generates the method section of neuroimaging papers.

[→ Goals](#)

#### Demos

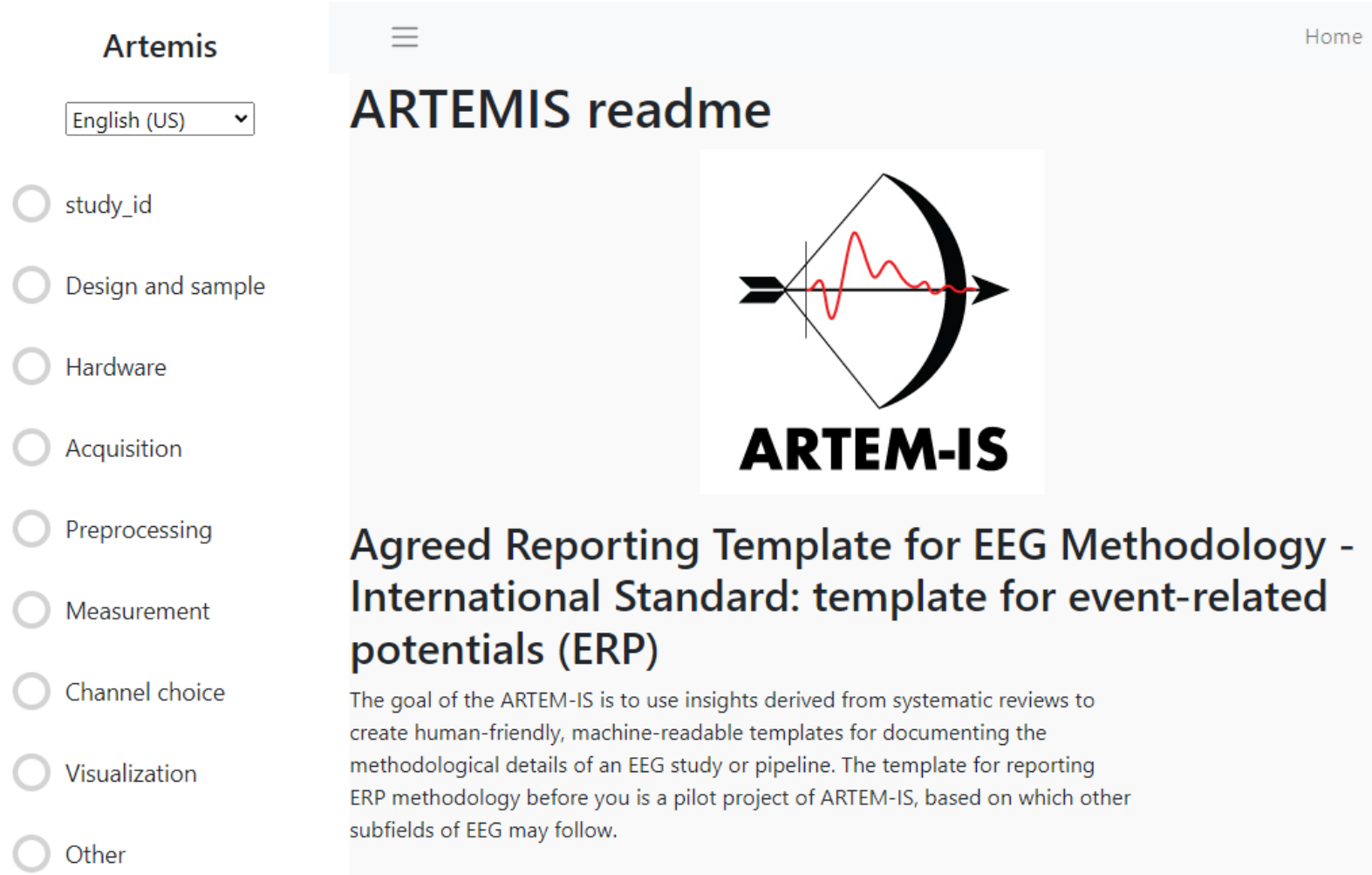
You want to try the checklists that already exist?

[→ Available apps](#)

<https://github.com/ohbm/eCOBIDAS>

# Best advice #4: use checklist and guidelines when reporting

Committee On Best practices In Data Analysis and Sharing – [COBIDAS](https://github.com/ohbm/eCOBIDAS) - <https://github.com/ohbm/eCOBIDAS>

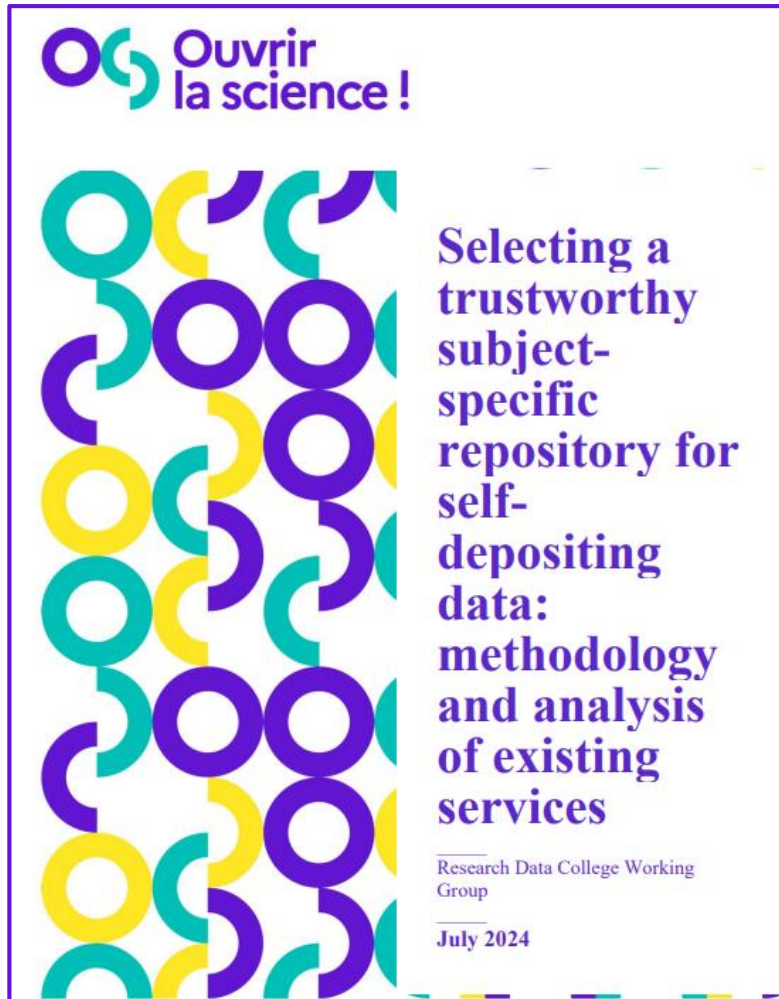


The screenshot shows the Artemis website interface. On the left, there is a navigation menu with the title "Artemis" and a language dropdown set to "English (US)". Below this is a vertical list of checklist items, each with a radio button: "study\_id", "Design and sample", "Hardware", "Acquisition", "Preprocessing", "Measurement", "Channel choice", "Visualization", and "Other". The main content area displays the "ARTEMIS readme" with a logo featuring a bow and arrow with a red waveform, and the text "ARTEM-IS". Below the logo is the title "Agreed Reporting Template for EEG Methodology - International Standard: template for event-related potentials (ERP)" and a paragraph explaining the goal of the ARTEM-IS project.

[https://remi-gau.github.io/artemis\\_checklist/#/](https://remi-gau.github.io/artemis_checklist/#/)

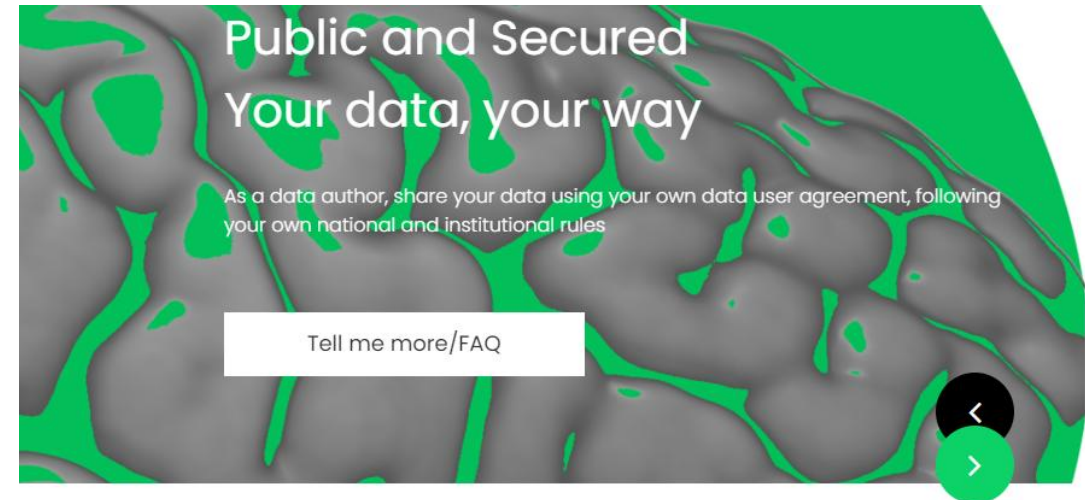
Pernet et al. (2018) *Best Practices in Data Analysis and Sharing in Neuroimaging using MEEG*. <https://doi.org/10.31219/osf.io/a8dhx>  
Pernet et al. (2020) *Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research*.

# Best advice #5: share your data and codes



List updated in April 2025

<https://www.ouvriirlascience.fr/selecting-a-trustworthy-subject-specific-repository-for-self-depositing-data-methodology-and-analysis-of-existing-services/>



About

## GDPR compliant data sharing

Thanks to the [Open Neuro PET project](#), we have built a GDPR compliant repository to share brain imaging data as FAIRly as possible. The principle is similar to [OpenNeuro](#) where each dataset is shared independently. As EU data have to be protected, we provide the necessary privacy protection.



<https://publicneuro.eu/>

## Best advice #5: share your data and codes




# Software Heritage

<https://www.softwareheritage.org/>

Home / Mission

Our ambition is to **collect**, **preserve**, and **share** all software that is publicly available in source code form. On this foundation, a wealth of applications can be built, ranging from cultural heritage to industry and research.

## Best advice #3bis: document your code

 **Sphinx** generates API documentation from Python packages ([link](#))

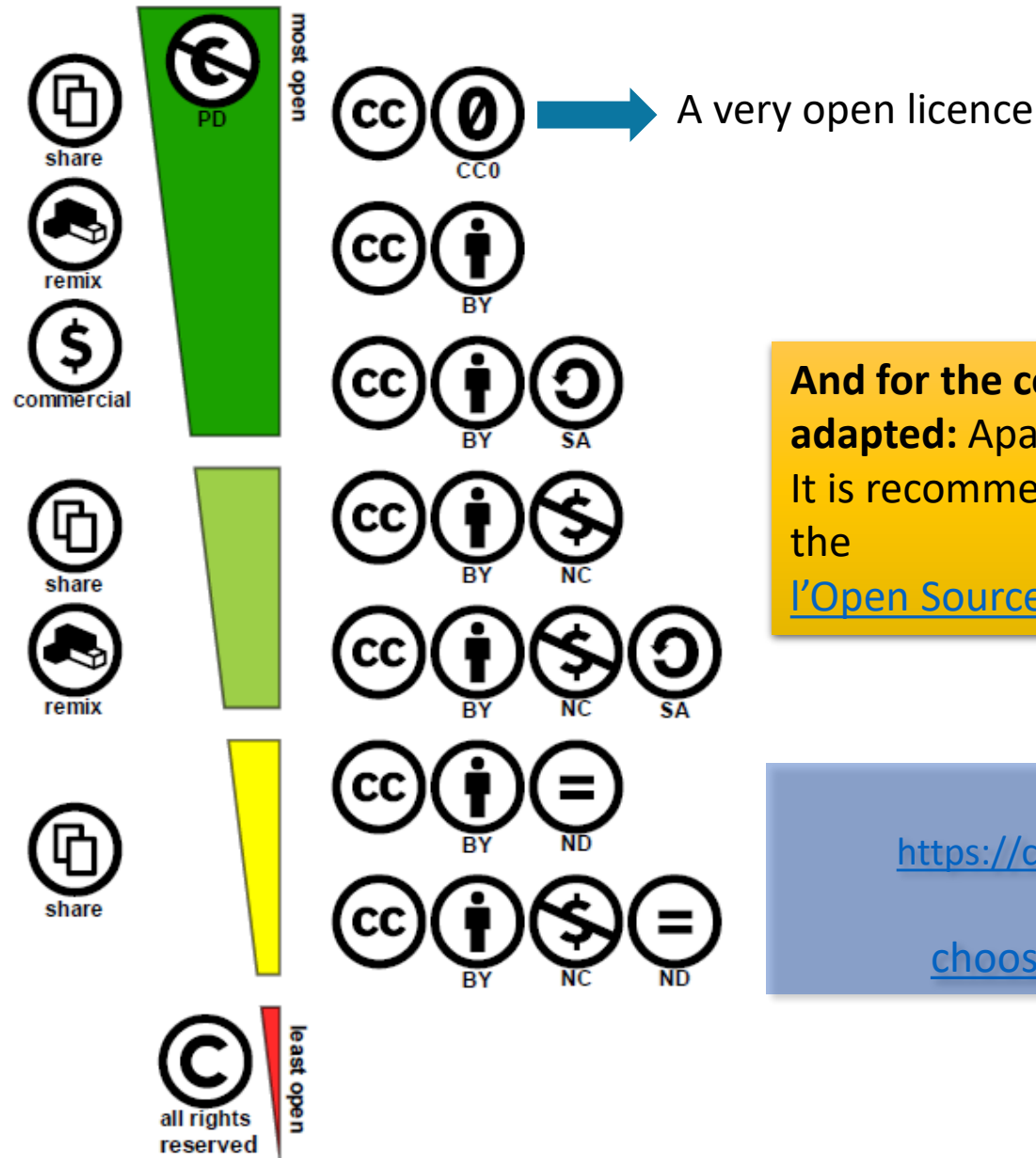
[Javadoc](#): generates HTML pages of API documentation from Java source files

[Swagger](#): to document APIs

[Doxygen](#): generates documentation for C, Java & Python languages

API: Application Programming Interface

# Best advice #6: share but protect your work



**And for the code, other licenses can be more adapted: Apache, Cecill, etc.**

It is recommended to use a licence approved by the

[l'Open Source Initiative.](https://opensource.org/)

***You hesitate?***

<https://chooser-beta.creativecommons.org/>

or

[choosealicense.com](https://choosealicense.com) (more general)

# Best advice #7: use the open ecosystem of open tools

## STUDY INCEPTION AND PLANNING



### Study Preparation

Ensure quality of data and design

- Release Standard Operating Procedures
- Pilot your experiment



<https://compose.neurosynth.org/>

protocolexchange

 protocols.io



<https://dmp.opidor.fr/>



### Pre-Registration

Have a thorough research plan: study design, hypotheses and analysis

- Follow best practices
- Use pre-registration / Registered Reports, report deviations



*ClinicalTrials.gov*



### Ethics and Sharing plan

Meet ethical and regulatory requirements

- Informed consent
- Internationally oriented data governance



Open Brain  
Consent



### Monitoring Quality

Ensure quality of acquired data

- Establish quality control checkpoints
- Quality assurance
- Include positive control analyses



# Best advice #7: use the open ecosystem of open tools

## DATA ACQUISITION



### Brain Data

Transparent and reproducible data acquisition

- Use open hardware
- Use open software
- Report full acquisition details



### Stimulus

Reproducible stimulus presentations

- Report detailed stimuli, timings, responses, etc.
- Use open frameworks and share stimulus materials



## RESEARCH DATA MANAGEMENT



### Data organization and standards

Ensure data are interoperable and reusable

- Organize your neuroimaging data following the Brain Imaging Data Structure (BIDS)
- Utilize available tools, converters and apps that work with BIDS



### Metadata and data annotation

Ensure data are understandable, interoperable and reusable

- Annotate data using established structures
- Track provenance with Neuroimaging Data Model (NIDM)



### Data management and tracking

Track and version control data for update, retrieval, and reuse

- Use open data hosting services
- Use tools for decentralized data management



# Best advice #7: use the open ecosystem of open tools

## DATA PROCESSING AND ANALYSIS



### Software

Choose from the many open tools available

- Implement best practices when writing code, and share openly
- Test for analytical robustness



### Standardizing workflows

Enhance reproducibility

- Use standardized workflows for automation, reproducibility and transparency
- Use containers to ensure computational reproducibility



### Statistical modelling

Focus on reproducibility and transparency

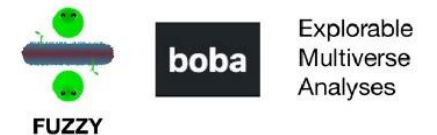
- Use analysis pipelines with automatic (machine readable) documentation
- Use standardized structures



### Multiverse analysis

Explore effects of variations in analysis pipeline on study results

- Vary critical analysis parameters
- Re-run analysis with different tools and test for robustness



### Key message:

Open data and analysis software are essential for reproducibility in data analysis

Niso et al. (2022) *Open and reproducible neuroimaging: From study inception to publication*. <https://doi.org/10.1016/j.neuroimage.2022.119623> - <https://oreoni.github.io/index.html>

# Best advice #7: use the open ecosystem of open tools



## DISSEMINATION

### **Data sharing**

**FAIR (Findable, Accessible, Interoperable, Reusable)**

- Standardize data with rich metadata, and share it under persistent identifiers on public repositories and a license
- Publish Data Descriptors

OpenNEURO GIN DANDI NDA NITRC brainlife.io

Child Mind Institute INDI BIDS OSF zenodo

### **Code sharing**

**FAIR for reproducible analysis**

- Document code and share it on public repositories + license
- Use version control
- Use containers

git jupyter docker A

GitHub GitLab binder CODE OCEAN OSF zenodo GIN

### **Derived data sharing**

**FAIR derived data and results**

- Share results of analyses in an interoperable and understandable way, so they are reusable and reproducible

NEUROVAULT brainlife OSF

TemplateFlow EBRAINS zenodo

Niso et al. (2022) *Open and reproducible neuroimaging: From study inception to publication*. <https://doi.org/10.1016/j.neuroimage.2022.119623> - <https://oreoni.github.io/index.html>

## DISSEMINATION

### Data sharing

#### FAIR (Findable, Accessible, Interoperable, Reusable)

- Standardize data with rich metadata, and share it under persistent identifiers on public repositories and a license
- Publish Data Descriptors



### Code sharing

#### FAIR for reproducible analysis

- Document code and share it on public repositories + license
- Use version control
- Use containers



### Derived data sharing

#### FAIR derived data and results

- Share results of analyses in an interoperable and understandable way, so they are reusable and reproducible




### Paper publication

#### Publish with openness and reproducibility in mind

- Follow best practices
- Publish on preprint servers and/or open access journals



 Software Heritage  
<https://www.softwareheritage.org/>

  
the first certification  
agency for scientific  
code & data  
<https://www.cascad.tech/>

  
The multidisciplinary  
archive for research  
<https://hal.science/>

**DISSEMINATION**

**Data sharing**

**FAIR (Findable, Accessible, Interoperable, Reusable)**

- Standardize data with rich metadata, and share it under persistent identifiers on public repositories and a license
- Publish Data Descriptors



**Code sharing**

FAIR for reproducible analysis

- Document code and share it on public repositories + license
- Use version control
- Use containers



**Derived data sharing**

FAIR derived data and results

- Share results of analyses in an interoperable and understandable way, so they are reusable and reproducible



**Paper publication**

Publish with openness and reproducibility in mind

- Follow best practices
- Publish on preprint servers and/or open access journals



**Beyond publication**

Disseminate scientific results to the broader scientific community and to the society in general

- Conferences, workshops, mass media, social networks



**Inclusive, diverse community research**

Open resources and welcoming spaces, diversity of ideas and approaches

- Equity, diversity and inclusivity
- Impact on careers, work-life balance, and mental health



The multidisciplinary archive for French research



**Where ever we are on the Open Science journey: there are tools to make it easier and more reproducible.**

**You're not alone!  
You're surrounded by  
people happy to help  
in your communities**



Humans make mistakes: self-correct the bugs and check the impact of them.

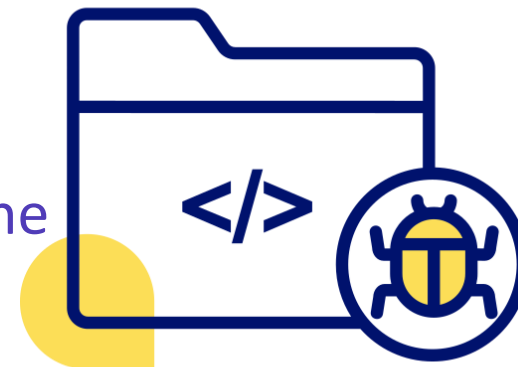
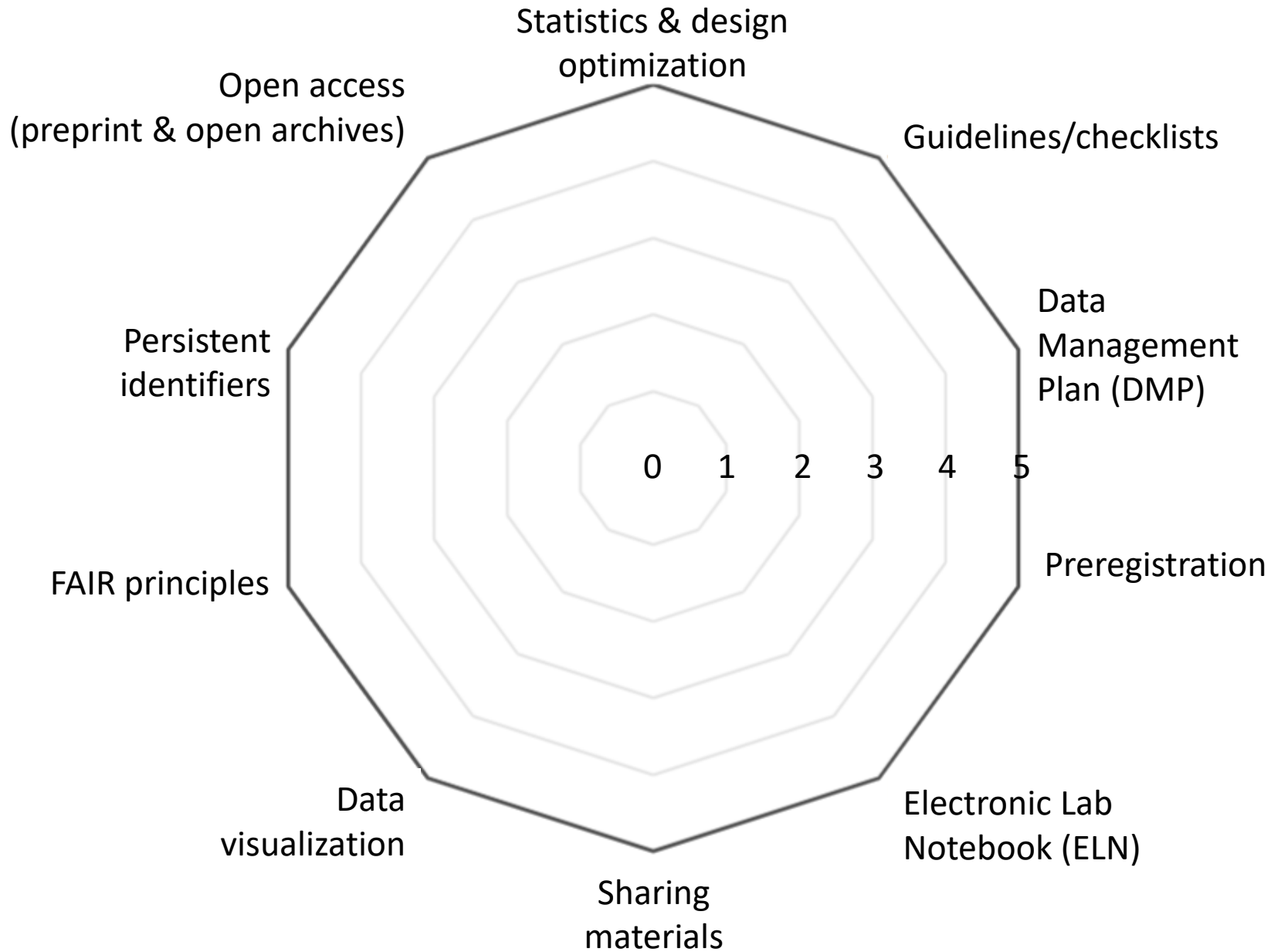


Image credit: icons made by Freepik, VectorPortal, SBTS2018 on [www.flaticon.com](http://www.flaticon.com)

# Open science practices: self-evaluation through the radar technique



1. Introduction
2. Preregistration
3. Data Management Plan
4. Standardisation
5. Documentation
- 6. Resources**

# Adopt Responsible Digital Practices

Labos 1point5

Collectif Publications Médiathèque Outils Ressources Actualités

## Reducing the environmental footprint of our research activities

Labos 1point5 is an international, cross-disciplinary collective of academic researchers who share a common goal: to better understand and reduce the environmental impact of research, especially on the Earth's climate.

### APPS 1POINTS

The GDR launches a scientific study on the carbon footprint of French public research and is making several tools available on a **dedicated platform**, including :

- GES 1point5 : analyse
- Scénario 1point5 : get into action
- Transition 1point5 : get connected

#CHG INVENTORIES **3608** #LABORATORIES **1595**

### LABOS 1POINTS

Labos 1point5 is a **Research Group (GDR)**, a **Seminar**, a **Reflection** team and a **Arts&Sciences** team.

For more information and to take part in the activities, visit the pages for the various teams.

### NEWS

**Webinar - GES 1point5 campus**  
Tuesday, May 6, 2025 at 10 a.m.  
Replay

**Doctoral training - 2nd edition**  
Thursday, May 22, 2025  
Replay

**Conference - Teaching ecological and social transitions in higher education (ETES)**  
7, 8 and 9 July 2025  
Lyon  
Schedule and registration

<https://labos1point5.org/>

## [GDS EcoInfo](#)

[Guide de bonnes pratiques numérique responsable pour les organisations](#) (in French)



Logo of the « Science and Climate » group at the CRNL

Details: <https://www.crnل.fr/fr/page-base/nos-valeurs>



# Tools for pre-registration



→ **Journals** that have adopted Registered Reports ([see a list](#))

→ [Table of Journal Features](#) for Registered Reports

→ [Examples](#) of existing preregistrations in OSF

→ **How to register: [a practical Guide](#)**

## [Create a Preregistration: Step by step](#)

1. Start a preregistration
2. Supply basic metadata about your research project
3. Enter your research plan
  - Add attachments to the form
  - Remove attachments from the form
4. Preview your preregistration
5. Register
6. Approve or cancel your preregistration

→ [COS Ambassadors](#) who can help and answer questions about preregistration

→ Check out the [Zotero database](#) for **completed examples of Stage 2** Registered Reports, and the the **OSF archive of registered Stage 1** protocols

[Adapted from slides by Justine Épinat-Duclos](#)

CRNL NeuroImaging Club meeting, 2021

# Resources

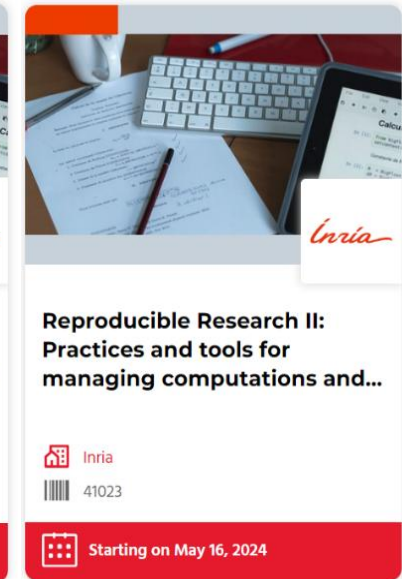
Open science in the doctoral school <https://nsco.universite-lyon.fr/science-ouverte/>



[Data management clusters](#)



<https://www.ouvrirlascience.fr/category/ressources/>



[fun-mooc](#)

## French network for Reproducible Research



[The Turing way project](#)



<https://www.recherche-reproductible.fr/>

Mailing-liste, actualités, GT, ressources

## Small Encyclopedia of Open Science

<https://encyclo.ouvrirlascience.fr/>



[Open and reproducible neuroimaging: from study inception to publication](#)

# Conference: data protection and regulation compliance

**Gaëlle Bujan, CNRS Data Protection Officer (DPO)**  
**Kodjo Kokou, lawyer, CNRS Data Protection Service**

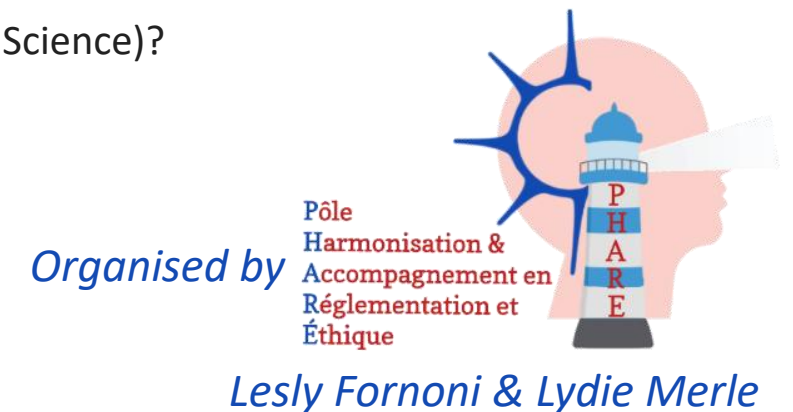
**Date:** Monday, November 17, 2:00-5:00 PM

**Location:** [Amphitheater of the Neurocampus](#), here!

## **Agenda:**

1. How to protect research data?
2. How to share experimental data (EEG, MEG, MRI...) while respecting regulations?
3. Are electrophysiological data (EEG, MEG...) or imaging data (MRI...) inevitably identifying data?  
What are the conditions for sharing them (Open Science)?
4. What means and rules should be implemented to build a data sharing server (Open Science)?
5. What are the rules regarding the use of AI in research?

**[Respond to the survey to book a seat](#) before **November 6, 2025****



# Conference: OHBM



<https://www.humanbrainmapping.org/>



# Acknowledgements



This work has been only possible with the valuable help of:

**Olivier Bertrand, previous Director of the CRNL**

**Laurent Bezin, current Director of the CRNL**



**The Open Science group at the CRNL:** Pierre-Emmanuel Aguera, Olivier Bertrand, Laurent Bezin, Anne Cheylus, Justine Épinat-Duclos, Alexandra Fargeot, Guillaume Sescousse\*



**CRNL think tanks on scientific integrity** Fabien Chauveau, Nicolas Fourcaud-Trocmé, Nicolas Grimault, Nicola Kuczewski, Marion Walter; **on Methodology** Alexandra Corneyllie & Françoise Lecaigard and **on NeuroImaging** with Justine Épinat-Duclos, Isabelle Failletot & Élise Météreau\*

**CRNL working groups** on Open Access/HAL/ORCID, Data Management Plans, Electronic Lab Notebooks



**Inspiring colleagues in networks & communities**



\* Alphabetical order