Ciencia Abierta y Reproducibilidad Científica

Susana Sánchez Expósito, Lourdes Verdes-Montenegro, Julian Garrido, Laura Darriba, Javier Moldón, Manuel Parra, Mª Angeles Mendoza

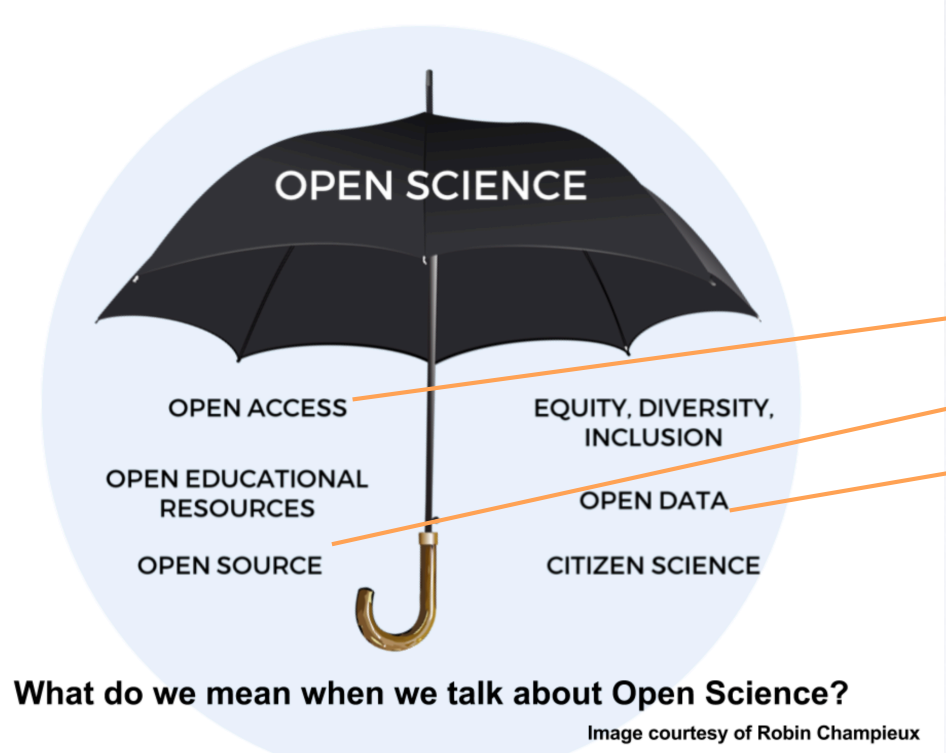
Instituto de Astrofísica de Andalucía (CSIC)

Marzo 2022

Open Science Definition

Open Science represents an approach to research that is collaborative, transparent and accessible

Open Science definition, European Commission, 2017, doi: 10.2777/75255



Improving the access to the elements involved in a scientific study:

- Article
- Method / software tools
- Data
- Etc.

Beyond the PDF initiative

http://www.clipartkid.com/rip-cliparts/

Knowledge Burying in paper publication

Research Objects: Towards Exchange and Reuse of Digital Knowledge, S. Bechhofer et al. 2011

"New mechanisms are needed that will allow us to share, exchange and reuse digital knowledge"

Need of exposing the complete scientific record, not the story and in a way the experiment can be discovered and understood

MOVING FROM NARRATIVES (LAST 300 YRS) TO THE ACTUAL OUTPUT OF RESEARCH

Open Science Definition

Open Science is the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.

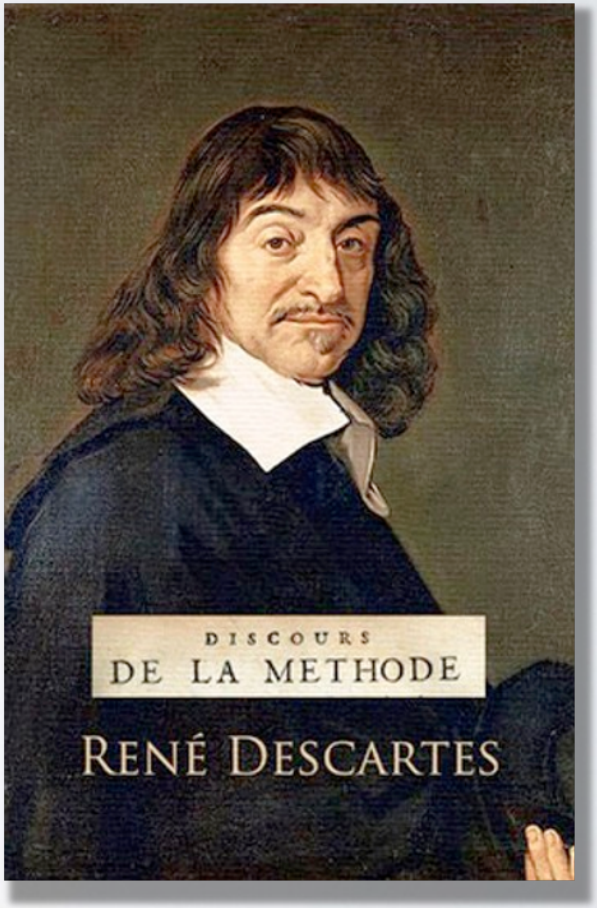
Foster project, https://www.fosteropenscience.eu/

Open Science: a new concept?

Too many **adjectives** for science:

excellent, high quality, trustable, reproducible … Open

Let's go back 383 years in time…



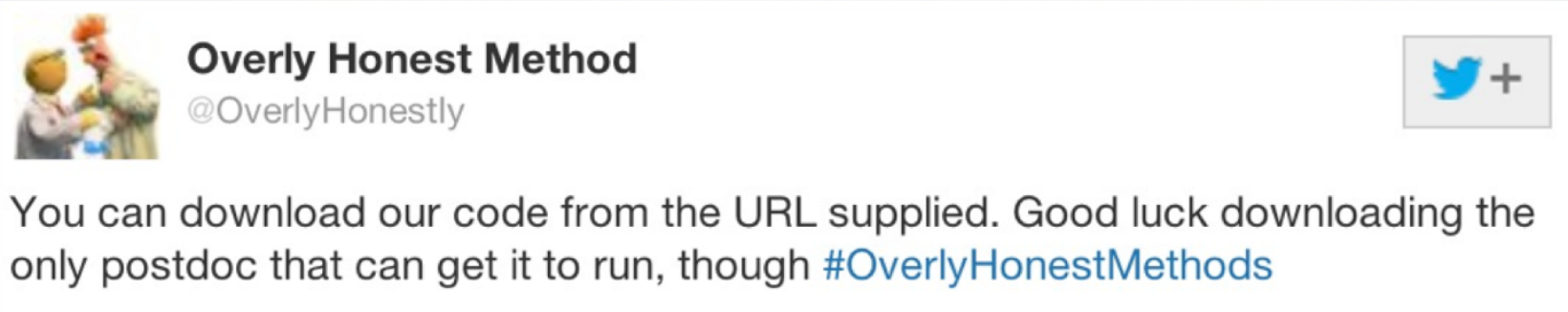
Scientific Reproducibility is a fundamental principle of the Scientific Method, a process established in the 17th century that marked the beginning of modern science and laid the foundations for the Philosophy of Science

Crisis of the scientific reproducibility

Questionnaire on reproducibility (1500 scientists)

- 70% of researchers have tried and failed to reproduce another scientist's experiments
- > 50% have failed to reproduce their own ones!
  - Chemistry: 90% (60%)
  - Biology: 80% (60%)
  - Physics and engineering: 70% (50%)
  - Medicine: 70% (60%)
  - Earth and environmental science: 60% (40%)

Baker (2016) https://doi.org/10.1038/533452a



Reproduce an experiment is not so easy

- Original data are not publicly available
- They are available but not in an automatic way
- Processed data is only available in the published PDF
- There are some scripts for processing the data on a server somewhere, but no one remembers where
- The code is in a public repository, but good luck trying to install/ execute it.

OPEN SCIENCE IS A SOLUTION FOR THOSE PROBLEMS

FAIR principles and Open Science

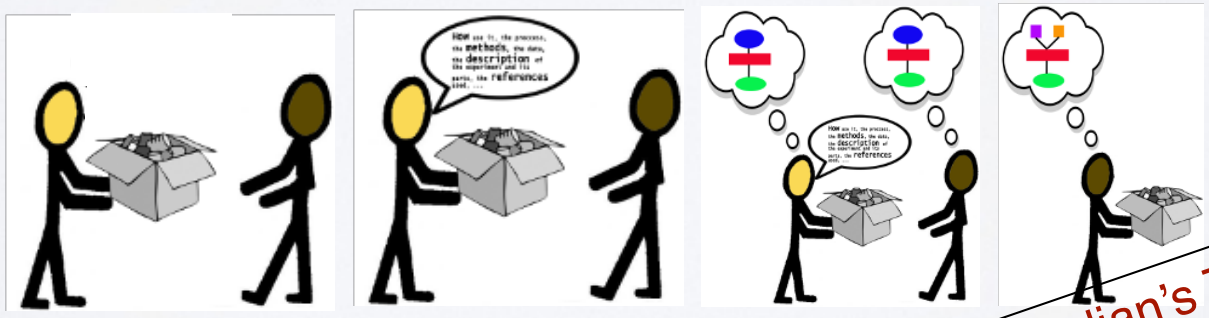
THE AVAILABILITY IS NOT ENOUGH → *FAIR - ABILITY*

Findable: data/methods can be discovered in an automatic way

Accesible: Available/stored in repositories/catalogues

Interoperable: Based on standards

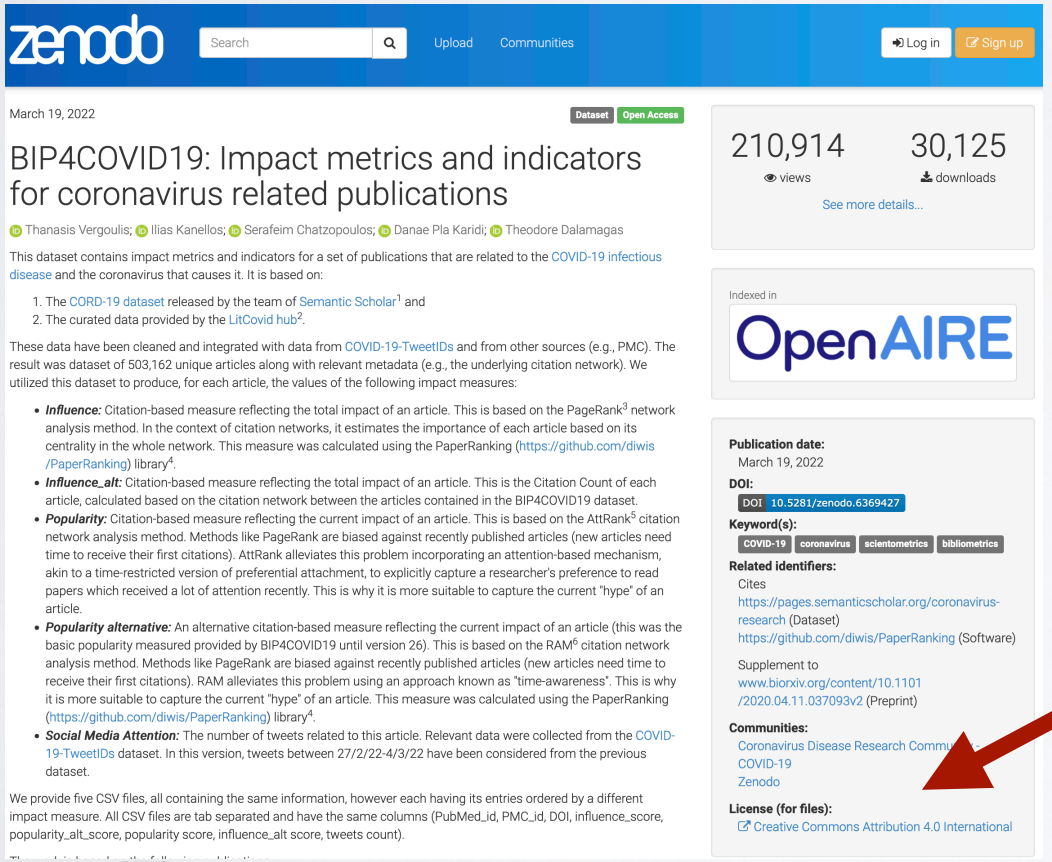
Reusable: provenance information & licenses



See Julian's Talk in next session

As open as possible

- Embargo period compliant with Open Science
- Open Source is not Free Software
- Licenses for using scientific resources: a key tool in Open Science



Barriers to Open Science

Barriers to Data and Code Sharing in Computational Science

Survey of Machine Learning Community, NIPS (Stodden, 2010):

Code		Data
77%	Time to document and clean up	54%
52%	Dealing with questions from users	34%
44%	Not receiving attribution	42%
40%	Possibility of patents	-
34%	Legal Barriers (ie. copyright)	41%
-	Time to verify release with admin	38%
30%	Potential loss of future publications	35%
30%	Competitors may get an advantage	33%
20%	Web/disk space limitations	29%

Barriers to Open Science

Barriers to Open Science

- Lack of awareness and training
- Cultural inertia and misinformation
- Challenging the establishment
- Follow the status quo to succeed
- Perceived lack of reward
- Not considered for promotion
- Requires additional skills
- Takes time
- Publication bias towards novel findings

Fig: McKiernan http://whyopenresearch.org

Whitaker (2018) https://doi.org/10.6084/m9.figshare.7140050.v2

R. Ainsworth, "Reproducibility & Open Science in the SKA Era". 2019 https://doi.org/10.5281/zenodo.2631868

Barriers to Open Science

L. Verdes-Montenegro, "Is the current measure of excellence perverting Science? A Data deluge is coming, it is time to act". ESOF 2018

Things are changing: DORA declaration. https://sfdora.org

Open Science

References

- **Love for science or "Academic Prostitution"?** L. Verdes-Montenegro. http://amiga.iaa.es/p/347-love-for-science-academic-prostitution.htm

- **Reproducibility & Open Science in the SKA Era.** R. Ainsworth et al. https://doi.org/10.5281/zenodo.2631868

- **Open Science for sustainability and inclusiveness: the SKA role model.** L. Verdes-Montenegro et al. https://www.youtube.com/watch?v=ErNT9Va9vus

Thanks!