FAIR Bioinfo 2022

Best practice in your bioinformatic projects

Pierre Marin
pierre.marin@uca.fr

Université Clermont Auvergne, AuBi, Mésocentre

18 octobre 2022

Sommaire

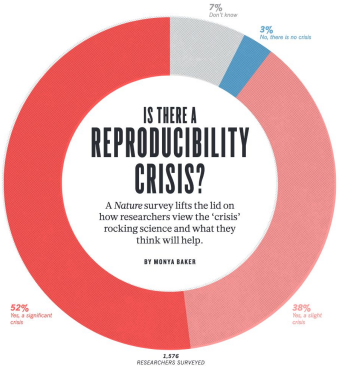
1 A reproducibility crisis

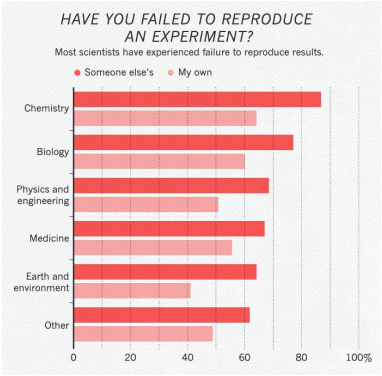
2 Open science and FAIR

3 Training content

Reproducibility crisis

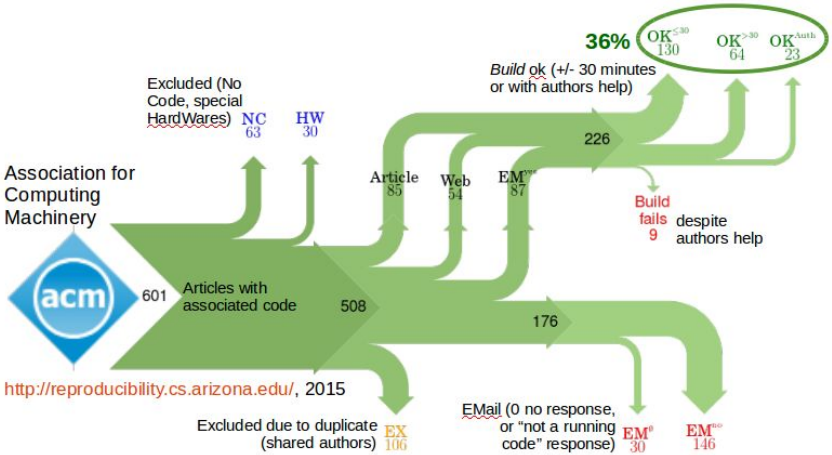
2016





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

Also in computer sciences



http://reproducibility.cs.arizona.edu/v2/data/Total.png

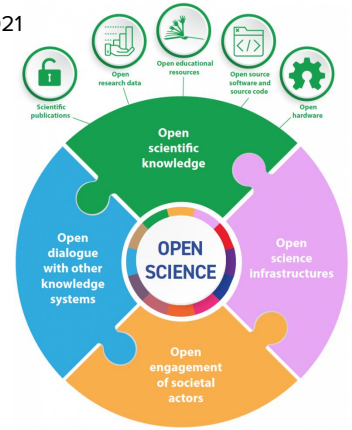
Long term negative impact of retracted papers

Article	Year of retraction	Citing Articles before retraction	Citing Articles after retraction	Total cites (journals indexed by Web of Science)
1. Primary Prevention of Cardiovascular Disease with a Mediterranean Diet. N ENGL J MED; APR **2013**. Estruch R, et al.	2018	1919	816	2735
2. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. LANCET; FEB 28 **1998**. Wakefield AJ, et al.	2010	642	867	1509
3. Visfatin: A protein secreted by visceral fat that mimics the effects of insulin. SCIENCE; JAN **2005**. Fukuhara A, et al.	2007	232	1192	1424
4. An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. PLANT J; MAR **2003**. Voinnet O, et al.	2015	896	375	1271
5. Lysyl oxidase is essential for hypoxia-induced metastasis. NATURE; APR **2006**. Erler JT, et al.	2020	977	81	1058

Retraction Watch : Top 10 most highly cited retracted papers
https://retractionwatch.com/the-retraction-watch-leaderboard/top-10-most-highly-cited-retracted-papers/ 6

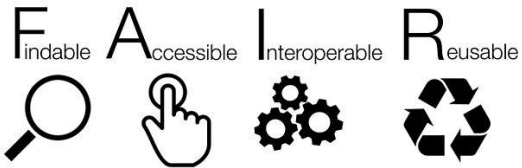
A way out: Open science and FAIR principles

2021



Graphic on page 11. UNESCO Recommendation on Open Science. CC BY IGO 3.0 C. Green

2016



Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3,** 160018 (2016). https://doi.org/10.1038/sdata.2016.18

By SangyaPundir - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=53414062

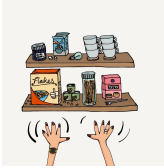
FAIR principles

Findable

By 維基小霸王 - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=88894774

PID
Repository

Accessible



https://nilsfirstworldproblems.tumblr.com/post/147555650875/i-cant-reach-the-top-shelves-of-the-kitchen

Protocols
(free, open, auth.)

Interoperable

By Unknown author - Popular Science Monthly Volume 88, Public Domain, https://commons.wikimedia.org/w/index.php?curid=22614407

Standards
(format, vocabulary)

Reusable

By Sun Ladder - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=5746428

Metadata
License
Origin

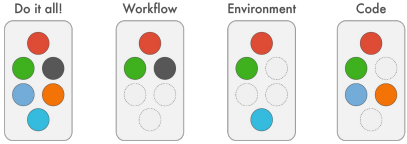
11
https://doi.org/10.1038/sdata.2016.18

Tools & use cases

Several tools but which ones to use and how? do some of them interact with each other?

3 use cases based on the previous sessions:

- E-labbook
- Reproducibility of running code
- Reproducibility in HPC



https://nbis-reproducible-research.readthedocs.io/en/course_2104/introduction/

13

FAIR session with AuBi

Objectives

- Discover FAIR practices
- Discover tools for best practices
- Use tool and best practices in practice sessions
- 5 sessions for courses and practices

FAIR session with AuBi

Objectives

- Discover FAIR practices
- Discover tools for best practices
- Use tool and best practices in practice sessions
- 5 sessions for courses and practices

The activity

- 5 sessions for courses and practice
- First session in November 2022

Contents

- Introduction to FAIR practices

Contents

- Introduction to FAIR practices
- Code control using Git
  - Git environment
  - Gitlab and Github

Contents

- Introduction to FAIR practices
- Code control using Git
  - Git environment
  - Gitlab and Github
- Encapsulation process
  - Conda environment and packages use CONDA
  - Containers as docker & singularity
  - Reproducible workflow using snakemake

Contents

- Introduction to FAIR practices
- Code control using Git
  - Git environment
  - Gitlab and Github
- Encapsulation process
  - Conda environment and packages use CONDA
  - Containers as docker & singularity
  - Reproducible workflow using snakemake
- Literate programming and documentation
  - Markdown syntax
  - Rmarkdown for R
  - Jupyterlab for Python