
# Reproducibility and Open Science

Dr. Rachael Ainsworth
Research Software Community Manager
Software Sustainability Institute, University of Manchester

@rachaelevelyn
@rainsworth
https://doi.org/10.6084/m9.figshare.14710110

CC BY

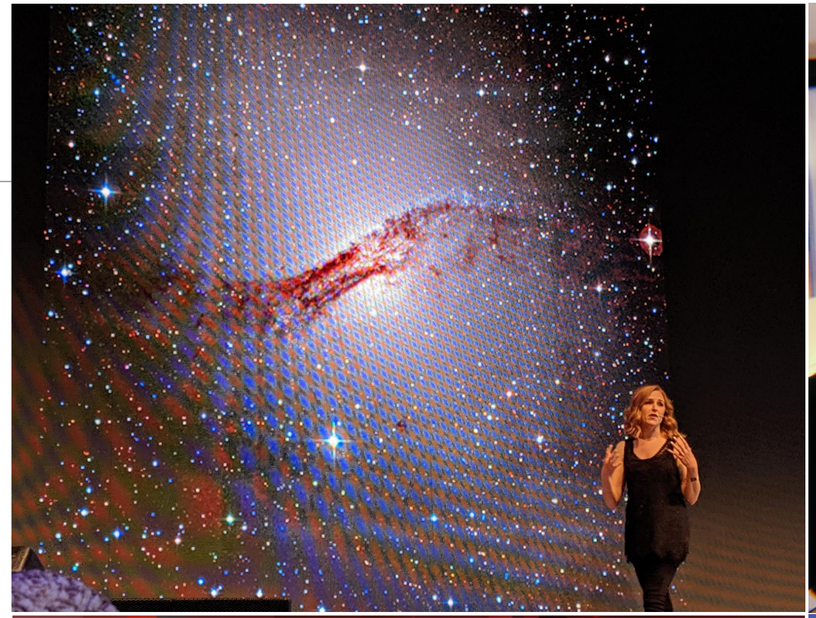
# Outline

- About me
- Reproducibility and research culture
- Open Science
- Barriers to Open Science
- Benefits of Open Science
- How to open up your research workflow
- Open Science examples in Astronomy
- More information, resources and takeaways

# About me

# About me

- Research Software Community Manager for the UK's Software Sustainability Institute
- SKA Regional Centre Steering Committee - Science Archive Working Group member
- Research background in radio Astrophysics (youtu.be/914KncrM6PM)
- Passionate about openness, transparency, reproducibility, wellbeing and inclusion in STEM/research
- TEDx speaker (youtu.be/c-bemNZ-IqA)





# About the Software Sustainability Institute

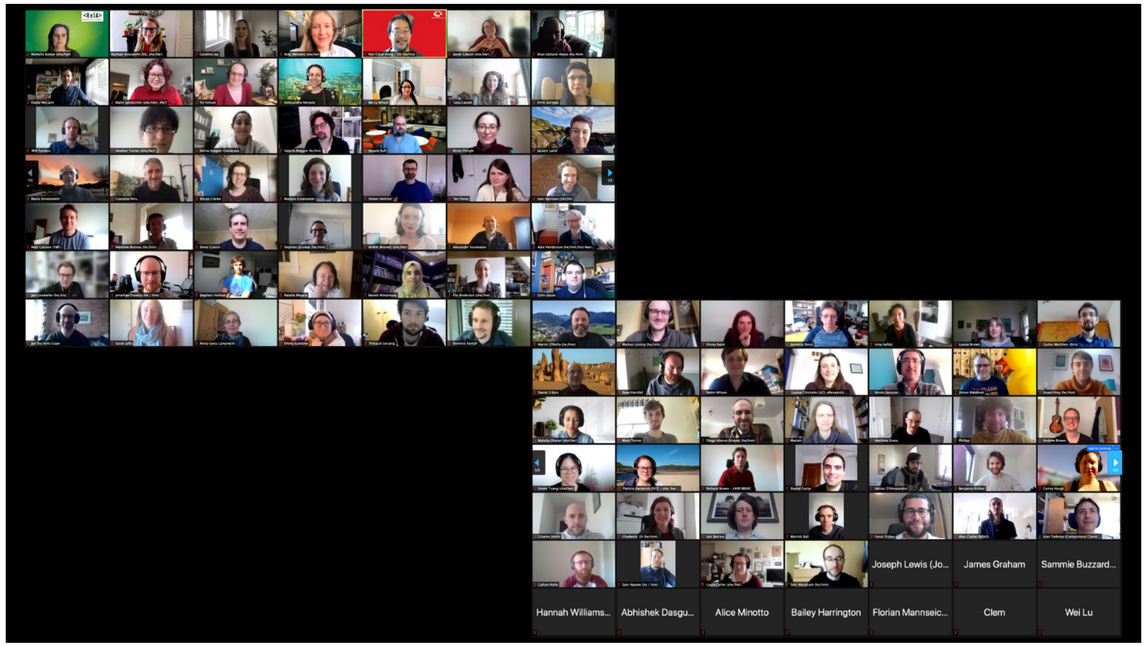
Software Sustainability Institute

- The UK's national facility for promoting the advancement of software in research by cultivating better, more sustainable, research software to enable world-class research: ***"Better software, better research"***

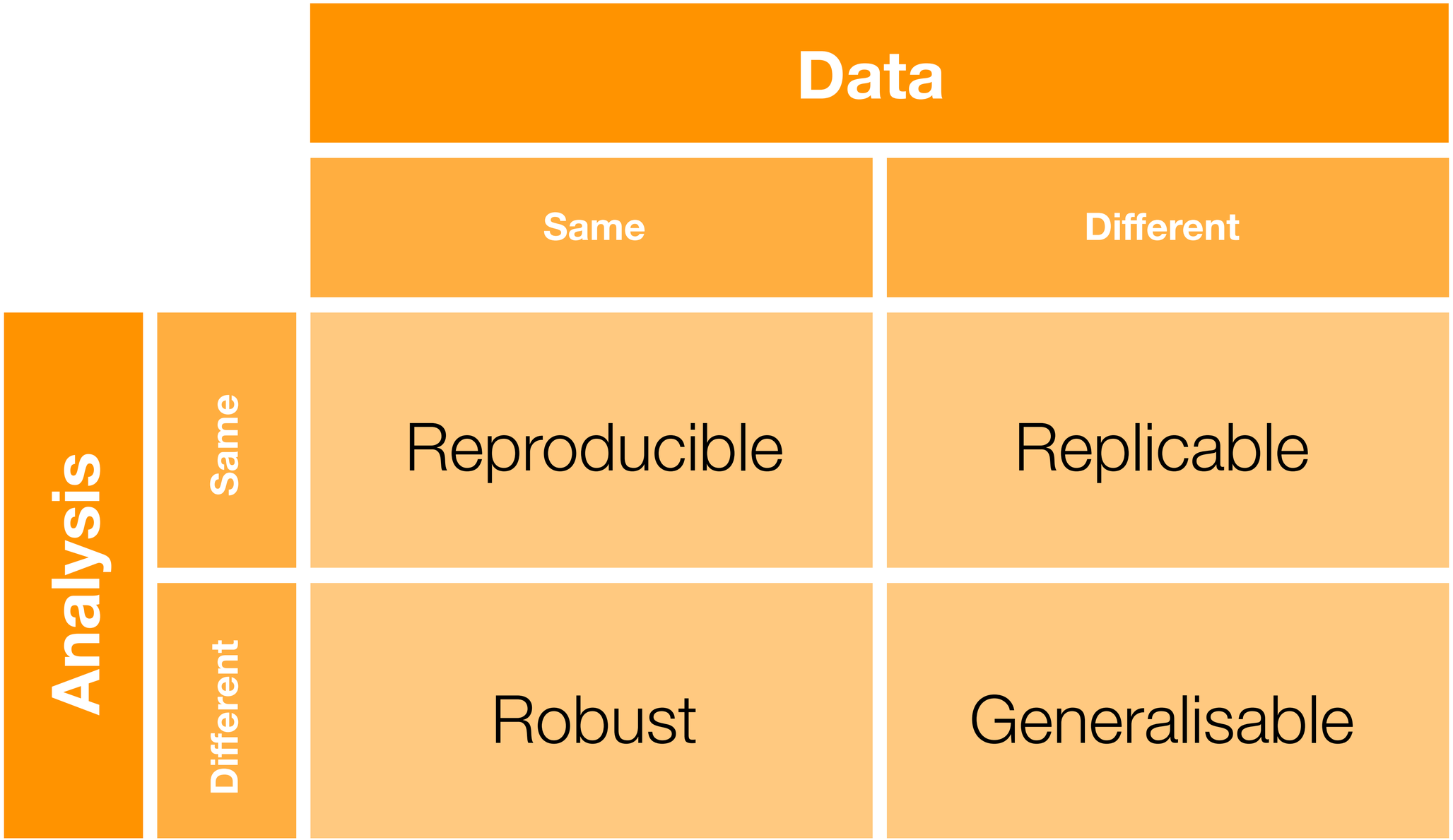
- Fellowship Programme to engage with and support natural ambassadors of better software practice in their research domains

- Collaborations Workshop is our annual unconference which brings together researchers, developers, innovators, managers, funders, publishers, leaders and educators to explore best practices and the future of research software

- https://www.software.ac.uk



# Reproducibility and research culture




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

nature International weekly journal of science



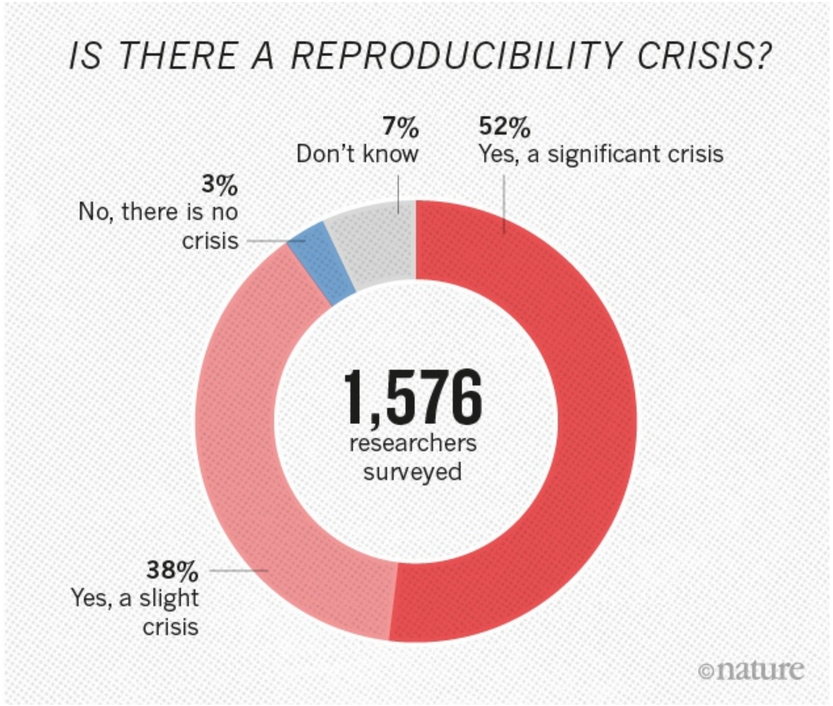
*NATURE* | NEWS FEATURE

# 1,500 scientists lift the lid on reproducibility

**Survey sheds light on the 'crisis' rocking research.**

**Monya Baker**

25 May 2016 | Corrected: 28 July 2016




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

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video

Archive > Volume 533 > Issue 7604 > News Feature > Article

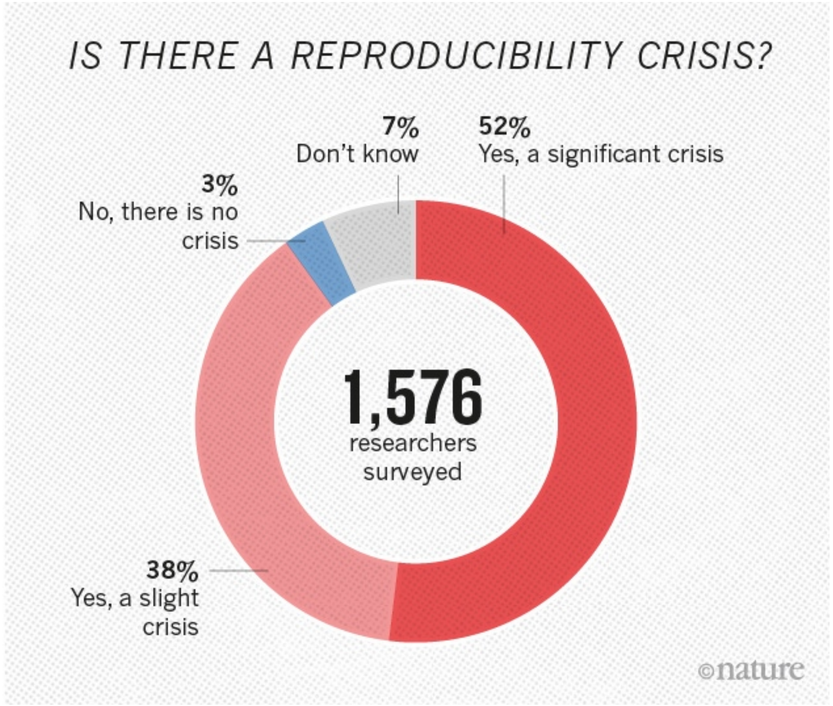
*NATURE* | NEWS FEATURE

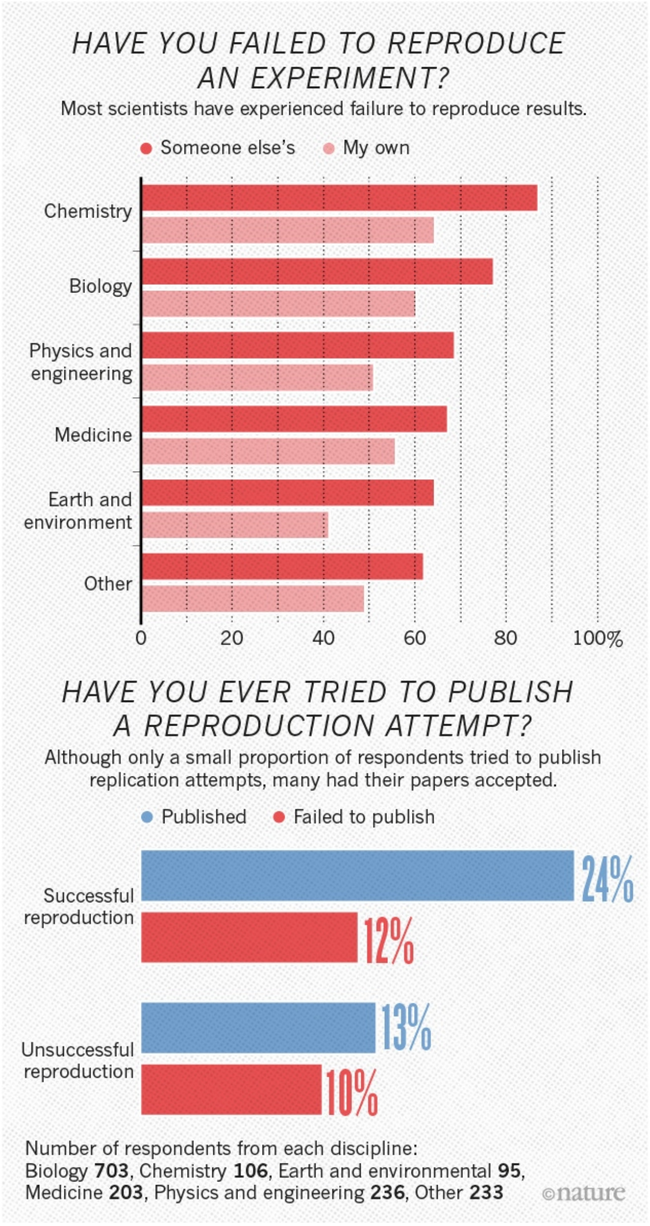
# 1,500 scientists lift the lid on reproducibility

**Survey sheds light on the 'crisis' rocking research.**

**Monya Baker**

25 May 2016 | Corrected: 28 July 2016




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

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video

Archive > Volume 533 > Issue 7604 > News Feature > Article

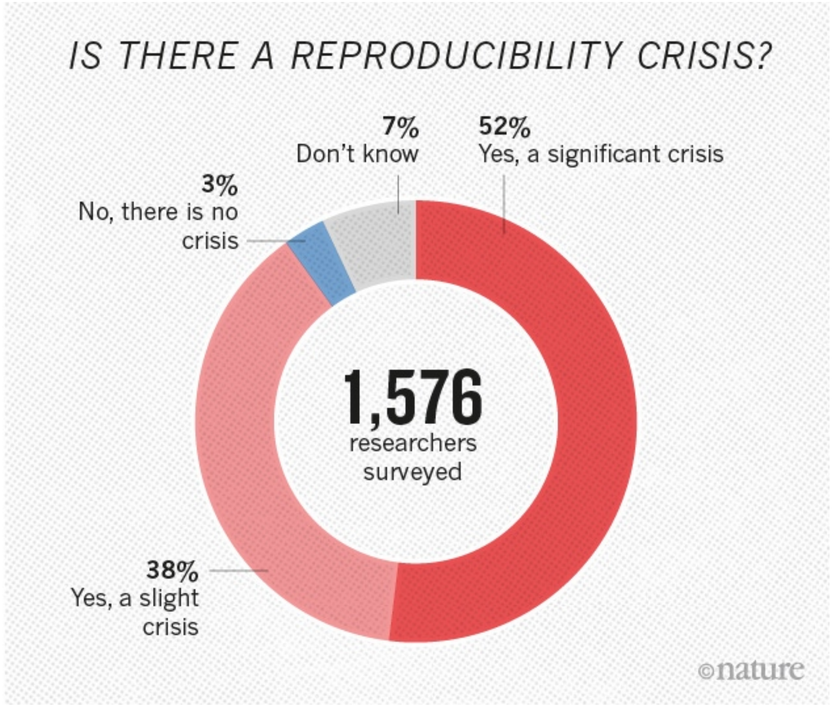
NATURE | NEWS FEATURE

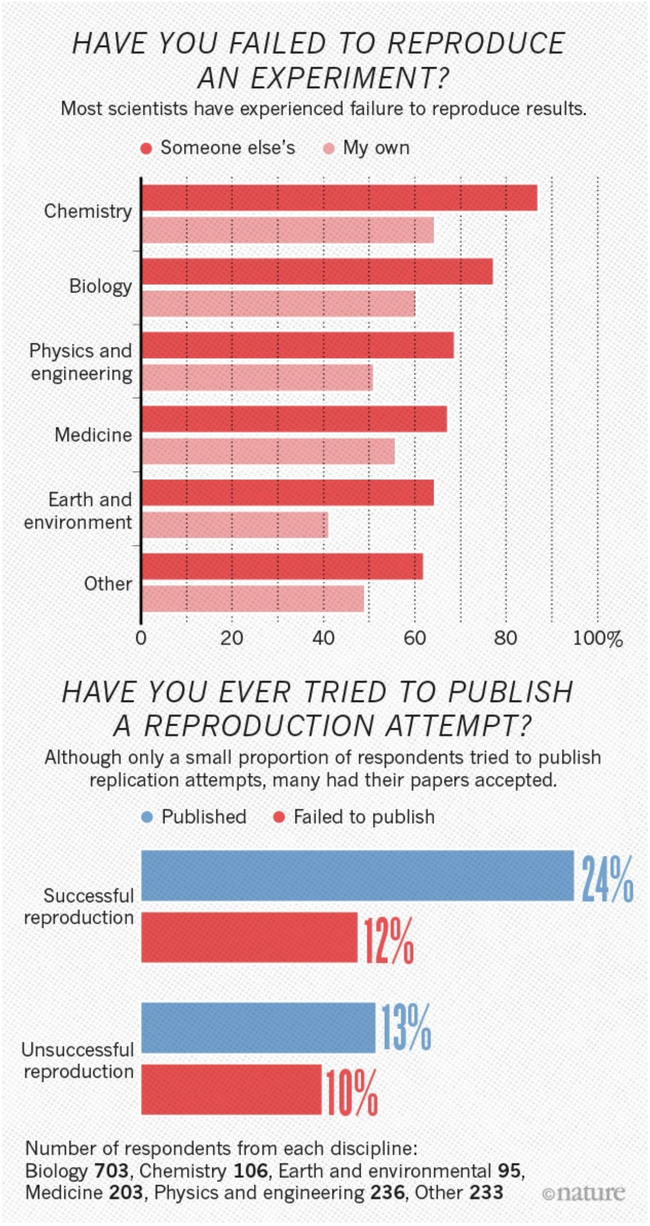
# 1,500 scientists lift the lid on reproducibility

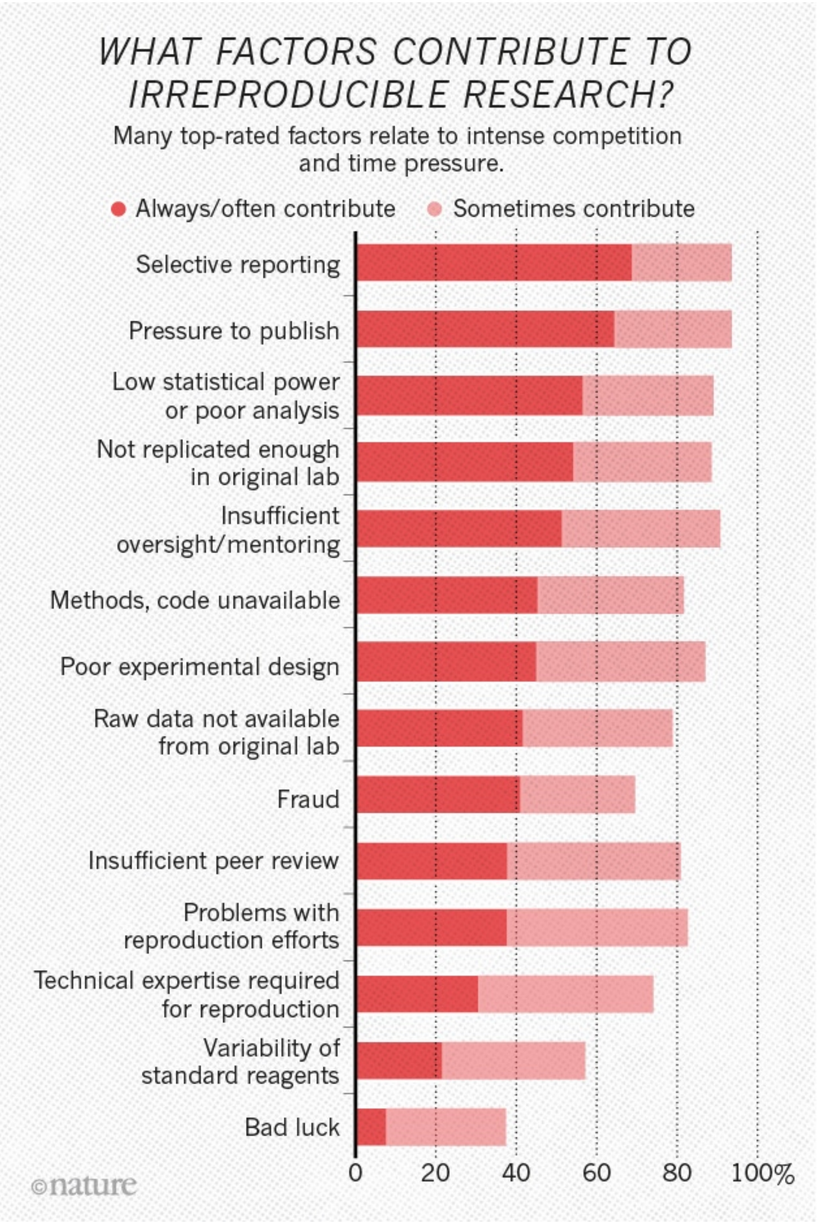
**Survey sheds light on the 'crisis' rocking research.**

**Monya Baker**

25 May 2016 | Corrected: 28 July 2016




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

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video

Archive > Volume 533 > Issue 7604 > News Feature > Article

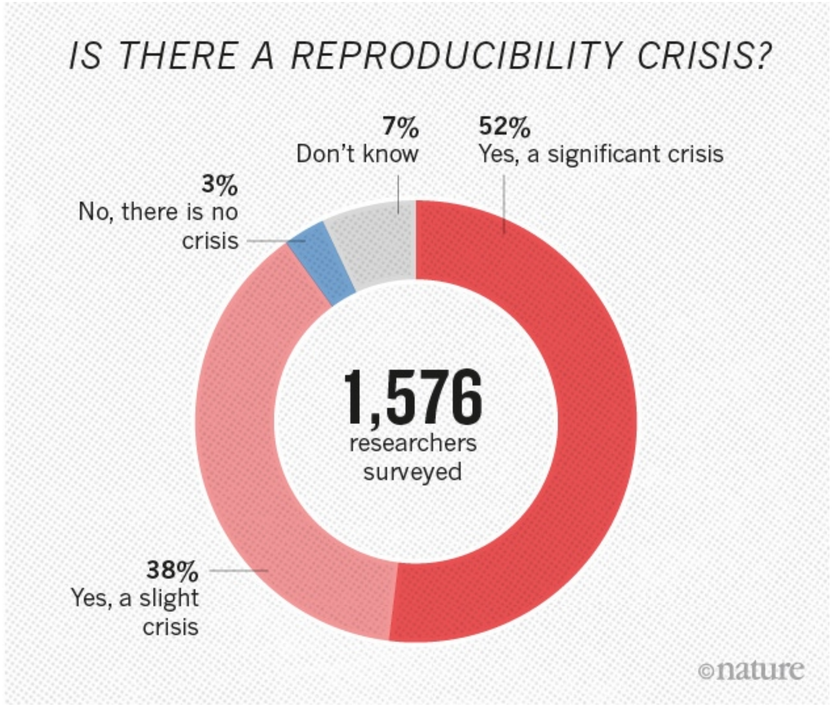
NATURE | NEWS FEATURE

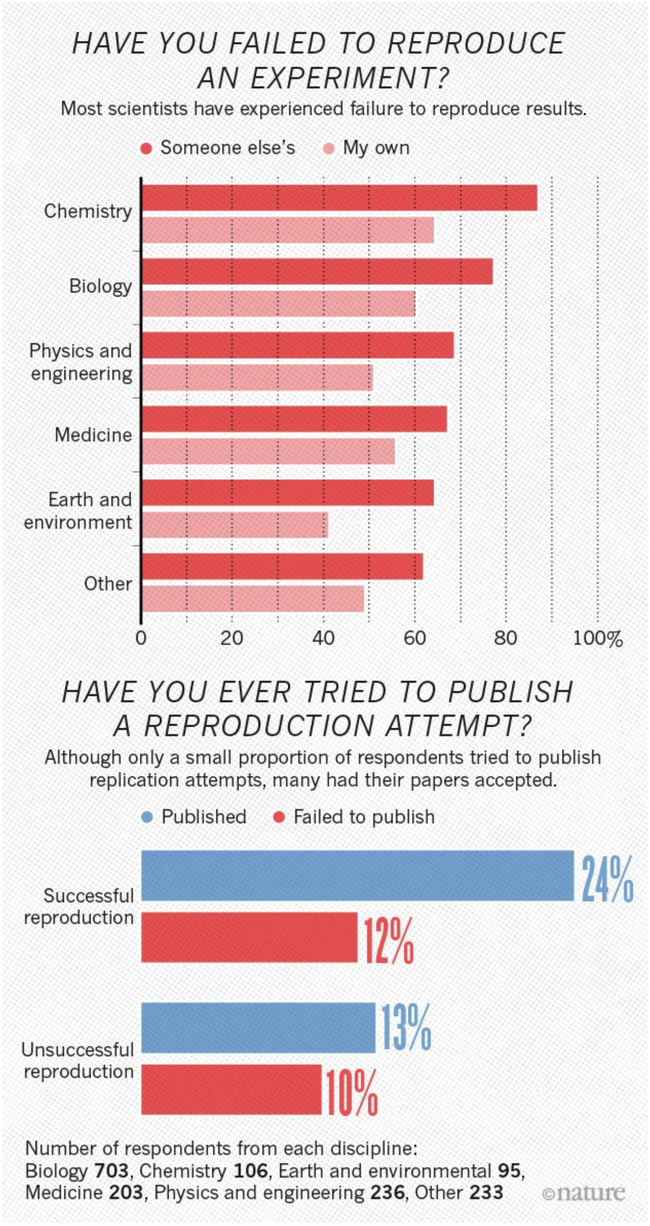
# 1,500 scientists lift the lid on reproducibility

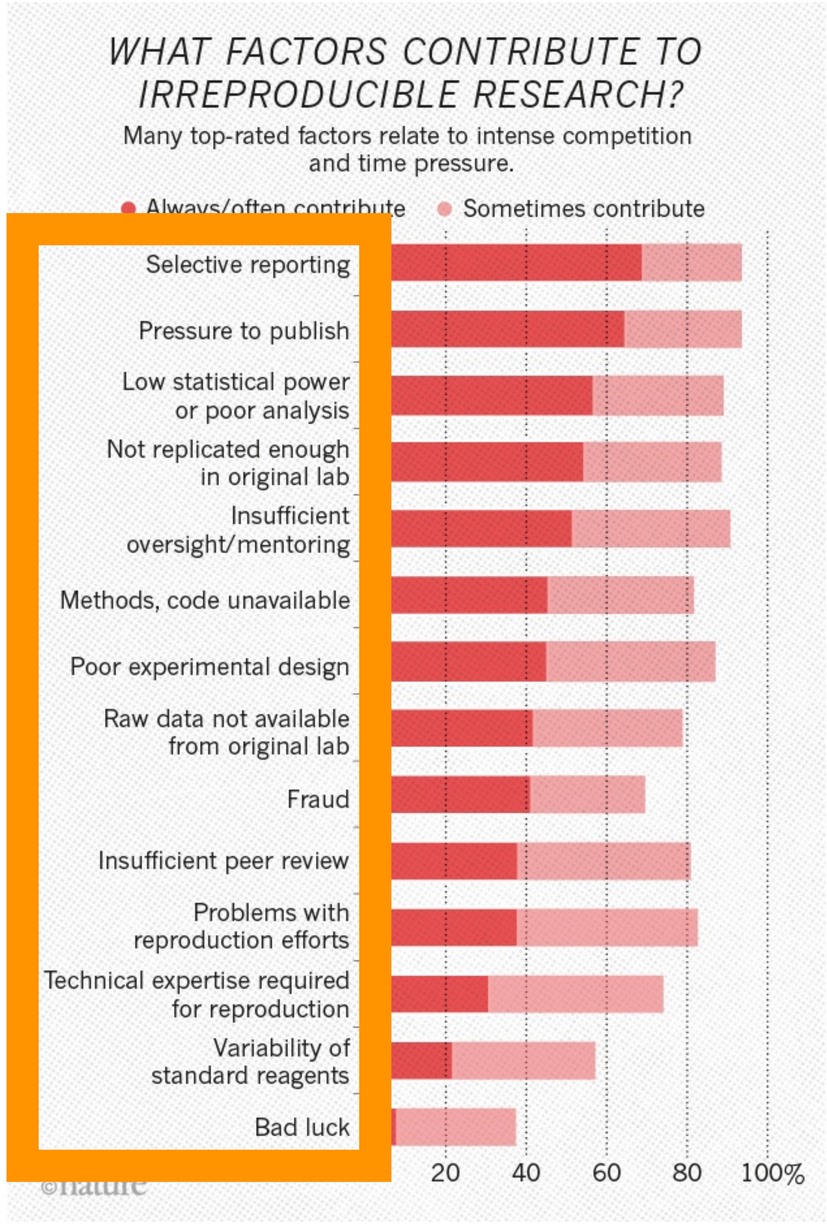
**Survey sheds light on the 'crisis' rocking research.**

**Monya Baker**

25 May 2016 | Corrected: 28 July 2016




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

*“Failing to reproduce results is a rite of passage. When I was a student, I tried to replicate what looked simple from the literature, and wasn't able to. Then I had a crisis of confidence, and then I learned that my experience wasn't uncommon.”*

**– Marcus Munafo, University of Bristol, UK**
**(Baker 2016 https://doi.org/10.1038/533452a)**

NATURE | NEWS

# Sluggish data sharing hampers reproducibility effort

**Initiative trying to validate 50 cancer papers finds difficulty in accessing original study data.**

Richard Van Noorden

03 June 2015

**RIO DE JANEIRO, BRAZIL**

Rights & Permissions

An initiative that aims to validate the findings of key cancer papers is being slowed by an unexpected hurdle — problems accessing data from the original studies.

The Reproducibility Initiative: Cancer Biology consortium aims to repeat experiments from 50 highly-cited studies published in 2010–12 in journals such as *Nature*, *Cell* and *Science*, to see how easy it is to reproduce their findings. Although these journals require authors to share their data on request, it has taken two months on average to get the data for each paper, said William Gunn, a co-leader of the project, at the 4th World Conference on Research Integrity in Rio de Janeiro, Brazil, on 3 June.

For one paper, securing the necessary data took a year. And the authors of four other papers have stopped communicating with the project altogether. In those instances, the journals that published the studies are stepping in to remind researchers of their responsibilities.

E-alert RSS Facebook Twitter


What matters in science — and why — free in your inbox every weekday.

Sign up

Listen

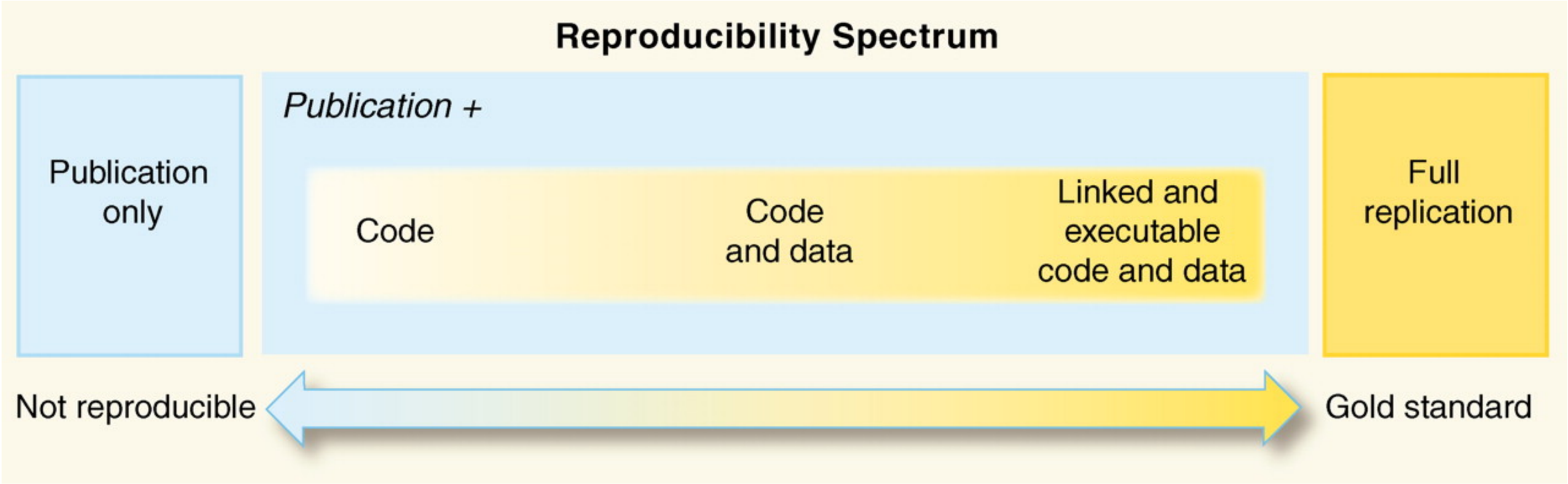
Nature Podcast

Our award-winning show features highlights from the week's edition of *Nature*, interviews with the people behind the science, and in-depth commentary and analysis from journalists around the world.

Science jobs from naturejobs

South China Normal University sincerely invite

Van Noorden (2015) https://doi.org/10.1038/nature.2015.17694

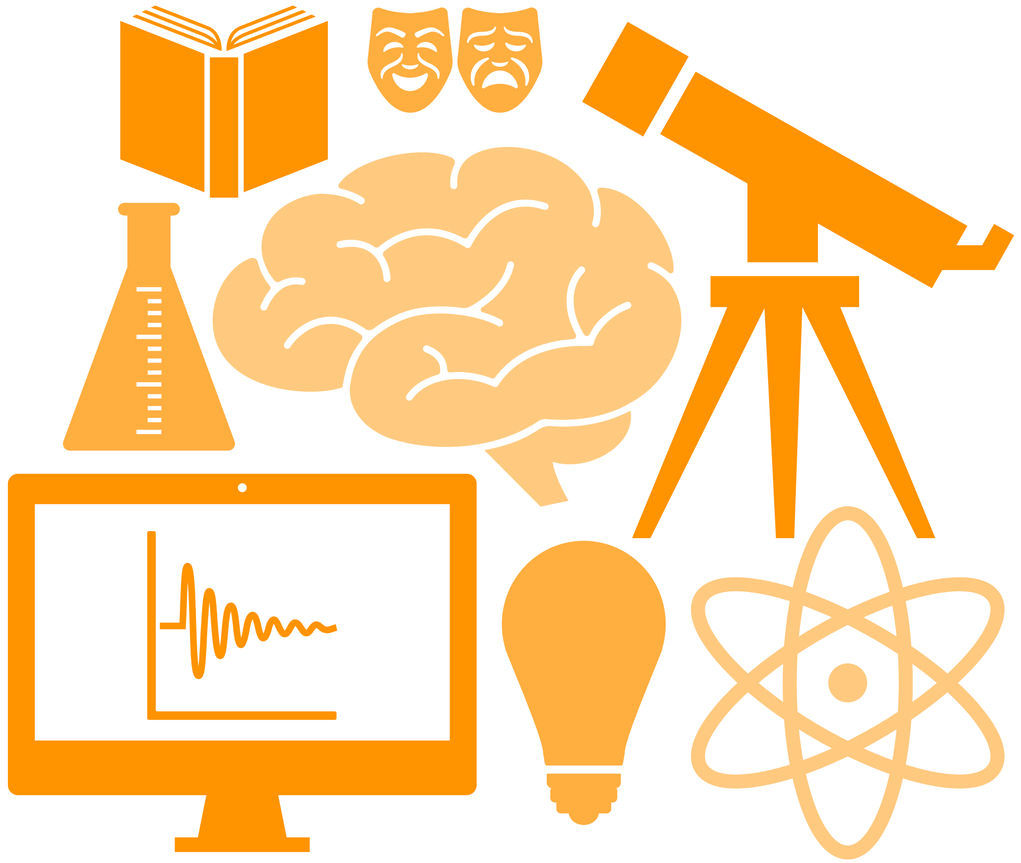



"Computational science has led to exciting new developments, but the nature of the work has exposed limitations in our ability to evaluate published findings. Reproducibility has the potential to serve as a minimum standard for judging scientific claims when full independent replication of a study is not possible."

Roger D. Peng (2011) https://doi.org/10.1126/science.1213847

# Research Culture

- Royal Society policy project on research culture: https://royalsociety.org/topics-policy/projects/research-culture/
  - Encompasses the behaviours, values, expectations, attitudes, and norms of research communities.
  - It affects who does research, what research is done, how it is done and how it is disseminated.
  - There are ongoing concerns around issues such as: research integrity, career paths, permeability between sectors, recognition and reward, diversity, and support for collaboration and interdisciplinarity.



# Research Culture

- Royal Society policy project on research culture: https://royalsociety.org/topics-policy/projects/research-culture/
  - Encompasses the behaviours, values, expectations, attitudes, and norms of research communities.
  - It affects who does research, what research is done, how it is done and how it is disseminated.
  - There are ongoing concerns around issues such as: research integrity, career paths, permeability between sectors, recognition and reward, diversity, and support for collaboration and interdisciplinarity.

***All of the issues have the same underlying causes:***

**Highly competitive environment**

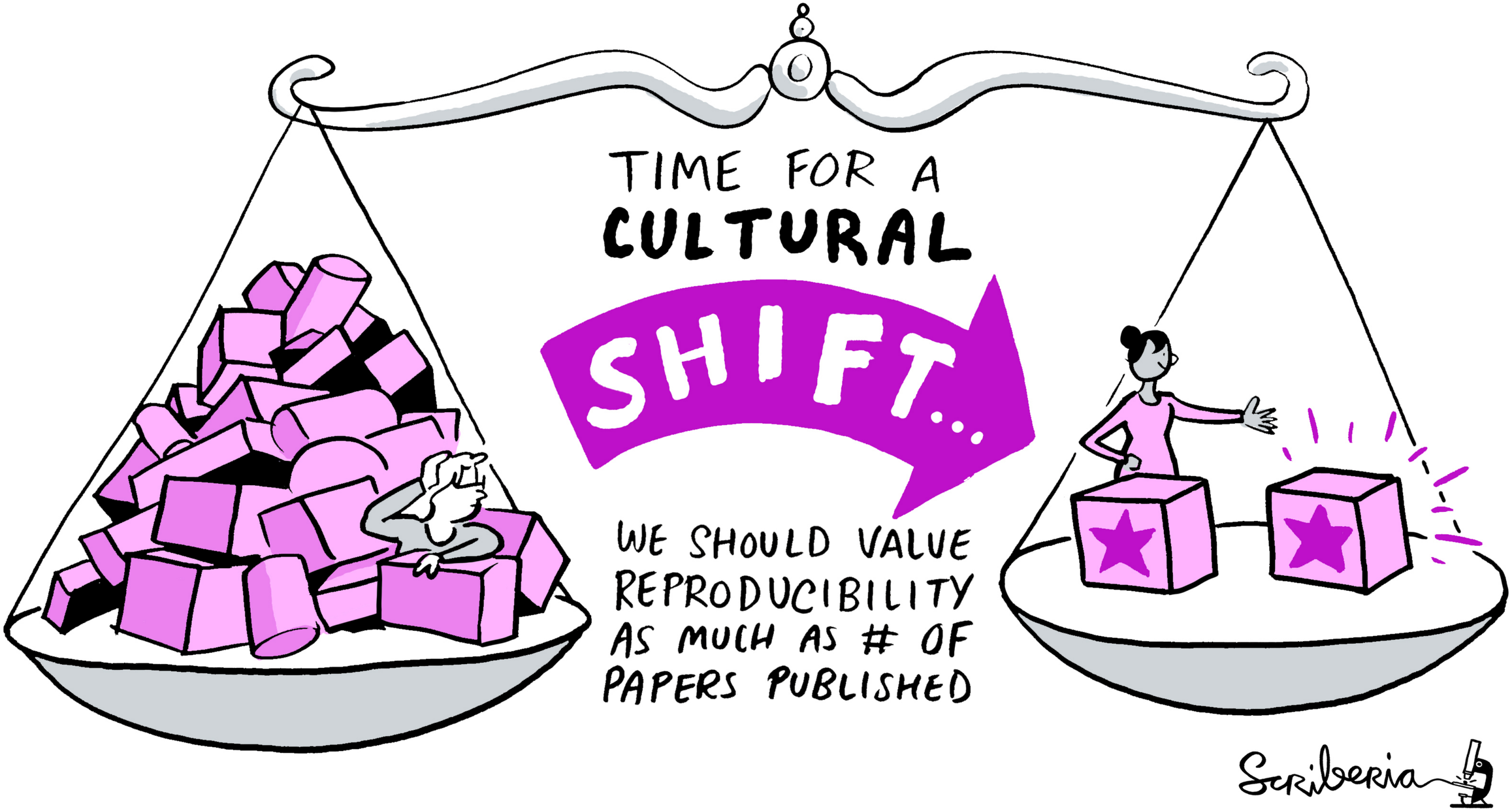
**+**

**Narrow definitions for success**

# Research Culture

- Wellcome key issue and report on research culture: https://wellcome.ac.uk/what-we-do/our-work/research-culture
  - Poor research culture is leading to unhealthy competition, bullying and harassment, and mental health issues
    - 78% of researchers think that high levels of competition have created unkind and aggressive conditions.
    - Nearly two-thirds of researchers (61%) have witnessed bullying or harassment, and 43% have experienced it themselves. Just one in three (37%) feel comfortable speaking up, with many doubting appropriate action will be taken.
    - Just over half of researchers (53%) have sought, or have wanted to seek, professional help for depression or anxiety.


Words that researchers would use to describe research culture. (Wellcome, https://wellcome.ac.uk/reports/what-researchers-think-about-research-culture)




This image was created by Scriberia for *The Turing Way* community and is used under a CC-BY licence. https://doi.org/10.5281/zenodo.3332807

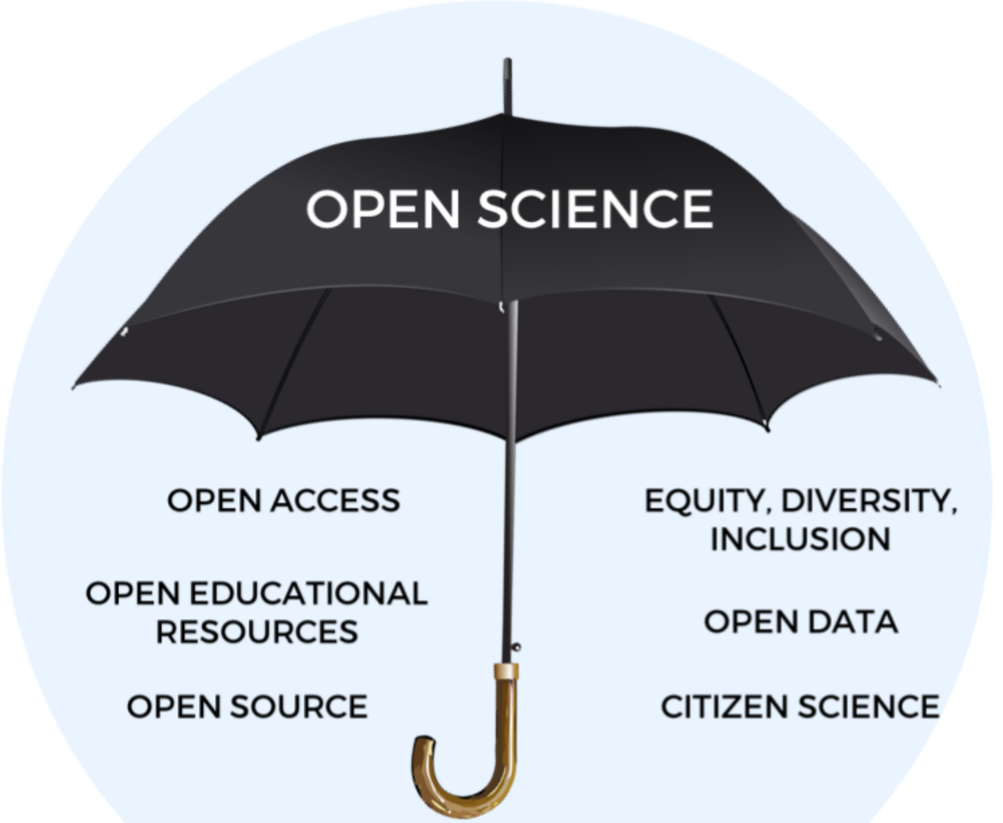
# Open Science

# What is Open Science?

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.

Open Science is about increased rigour, accountability, and reproducibility for research. It is based on the principles of inclusion, fairness, equity, and sharing, and ultimately seeks to change the way research is done, who is involved and how it is valued. It aims to make research more open to participation, review/refutation, improvement and (re)use for the world to benefit.

(FOSTER Open Science Training Handbook https://book.fosteropenscience.eu)




**What do we mean when we talk about Open Science?**

**Image courtesy of Robin Champieux**

# Barriers to Open Science

# Barriers to Open Science

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


Fig: McKiernan http://whyopenresearch.org

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

# Barriers to Open Science

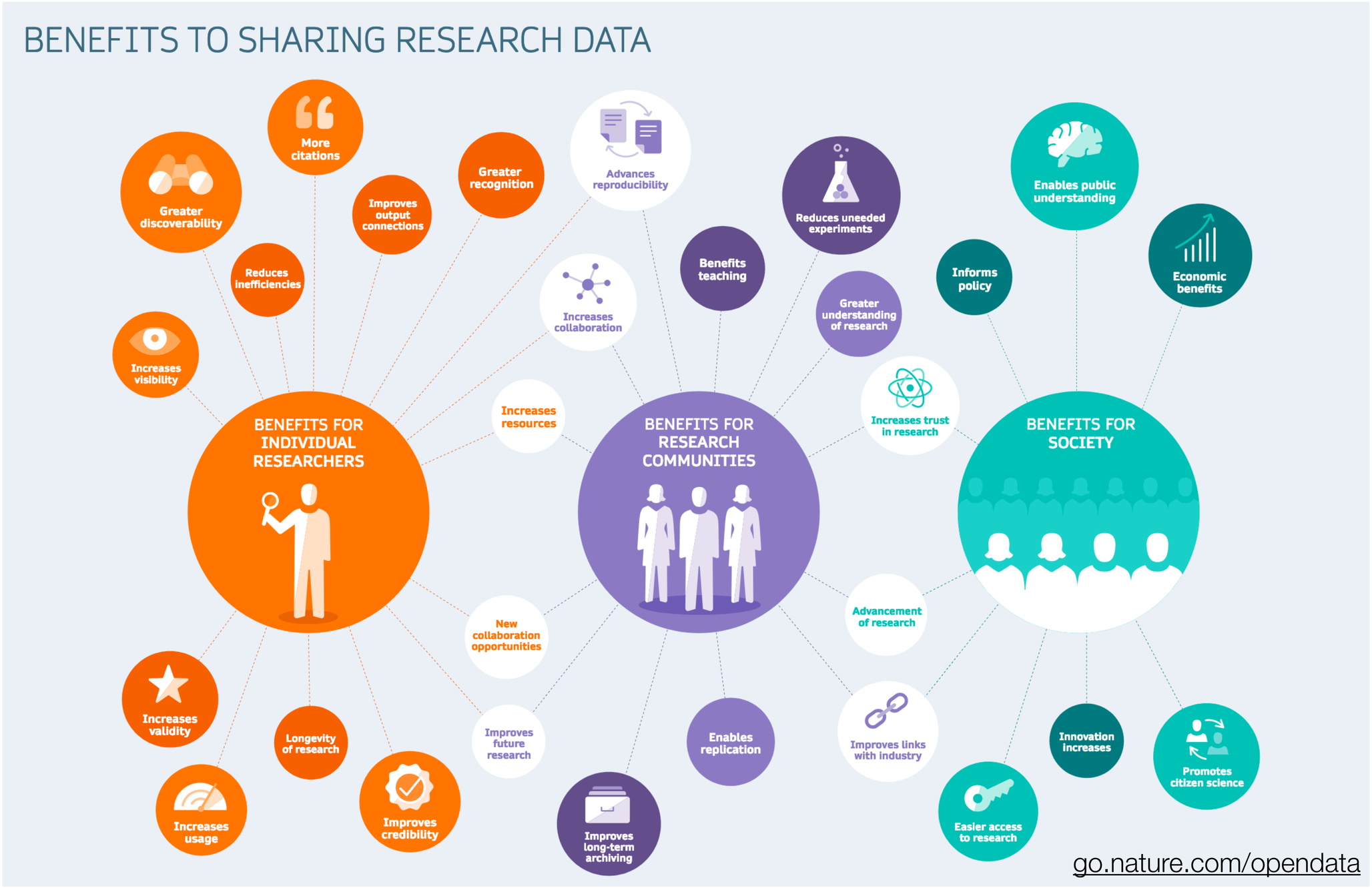
**Fear** of

- Scooping or ideas being stolen
- Not being credited for ideas
- Errors and public humiliation
- Risk to reputation
- Reduced scientific quality
- Information overload


Tennant (2017) https://doi.org/10.6084/m9.figshare.5383711.v1

https://doi.org/10.6084/m9.figshare.5558653

# Benefits of Open Science


BENEFITS TO SHARING RESEARCH DATA
More citations
Greater discoverability
Improves output connections
Greater recognition
Advances reproducibility
Reduces uneeded experiments
Enables public understanding
Reduces inefficiencies
Benefits teaching
Informs policy
Economic benefits
Increases collaboration
Greater understanding of research
Increases visibility
Increases resources
BENEFITS FOR INDIVIDUAL RESEARCHERS
BENEFITS FOR RESEARCH COMMUNITIES
Increases trust in research
BENEFITS FOR SOCIETY
New collaboration opportunities
Advancement of research
Increases validity
Longevity of research
Improves future research
Enables replication
Improves links with industry
Innovation increases
Promotes citizen science
Increases usage
Improves credibility
Improves long-term archiving
Easier access to research
go.nature.com/opendata

*“It is about mind-sets and culture: An unsung part of open software are its communities that promote and enable a more inclusive, kinder culture.”*

**– Julia Stewart Lowndes, Open Software Means Kinder Science**
**https://blogs.scientificamerican.com/observations/open-software-means-kinder-science**

BIOCHEMISTRY AND CHEMICAL BIOLOGY

# Point of View: How open science helps researchers succeed

Erin C McKiernan, Philip E Bourne, C Titus Brown, Stuart Buck, Amye Kenall, Jennifer Lin, Damon McDougall, Brian A Nosek, Karthik Ram see all »
National Autonomous University of Mexico, Mexico; National Institutes of Health, United States; University of California, Davis, United States; Laura and John Arnold Foundation, United States; BioMed Central, United Kingdom; CrossRef, United Kingdom; University of Texas at Austin, United States; Center for Open Science, United States; University of California, Berkeley, United States see all »

FEATURE ARTICLE Jul 7, 2016

CITED 66 VIEWS 18,445 ANNOTATIONS 3 CITE AS: eLife 2016;5:e16800 DOI: 10.7554/eLife.16800

Article
Figures and data
Side by side
Jump to

## Abstract

Open access, open data, open source and other open scholarship practices are growing in popularity and necessity. However, widespread adoption of these practices has not yet been achieved. One reason is that researchers are uncertain about how sharing their work will affect their careers. We review literature demonstrating that open research is associated with increases in citations, media attention, potential collaborators, job opportunities and funding opportunities. These findings are evidence that open research practices bring significant benefits to researchers relative to more traditional closed practices.

https://doi.org/10.7554/eLife.16800.001

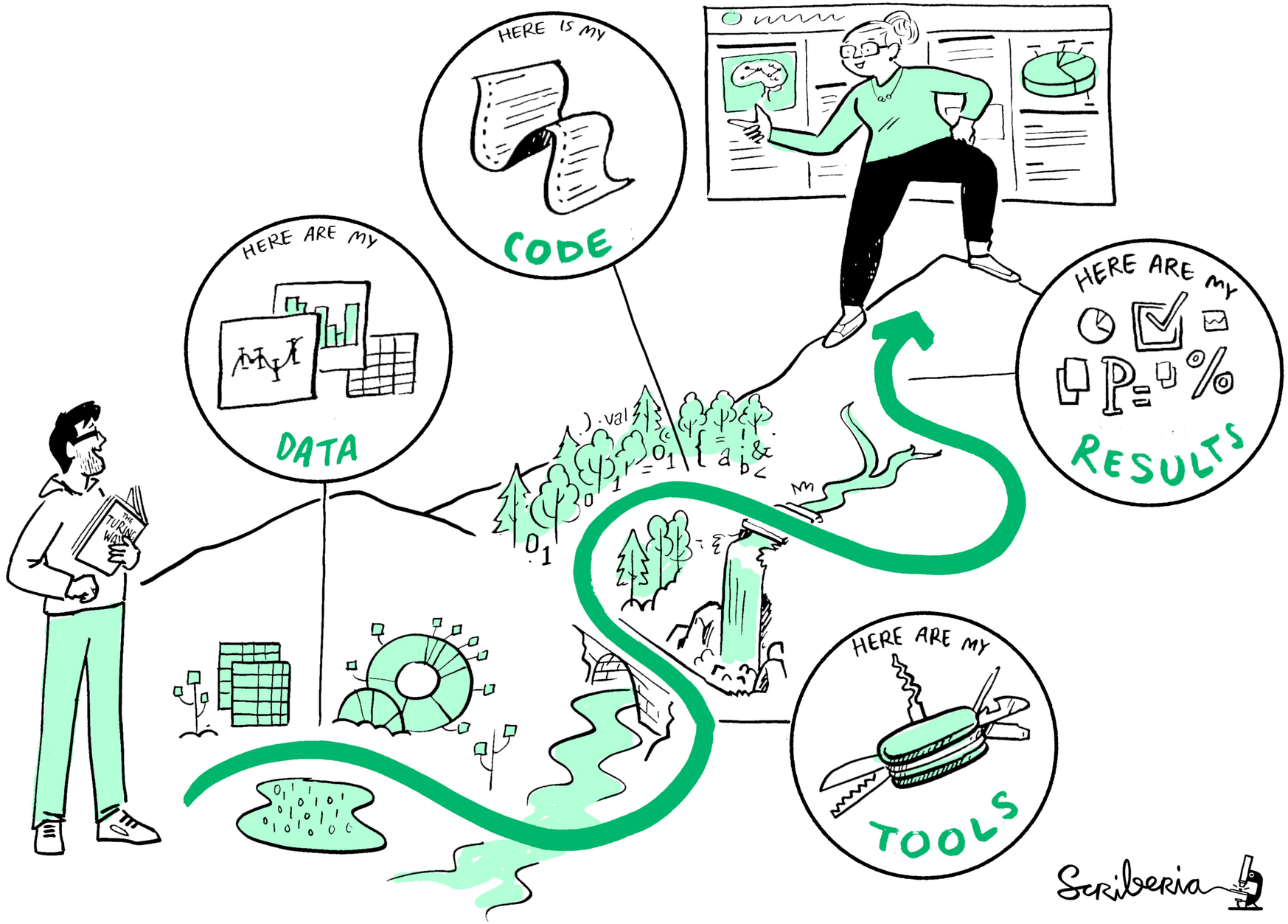
OF INTEREST
In the open
PODCAST
Further reading »

Your primary collaborator is yourself 6 months from now, and your past self doesn't answer emails.

**– Software Carpentry**

https://dynamicecology.wordpress.com/2015/02/18/the-biggest-benefit-of-my-shift-to-r-reproducibility/

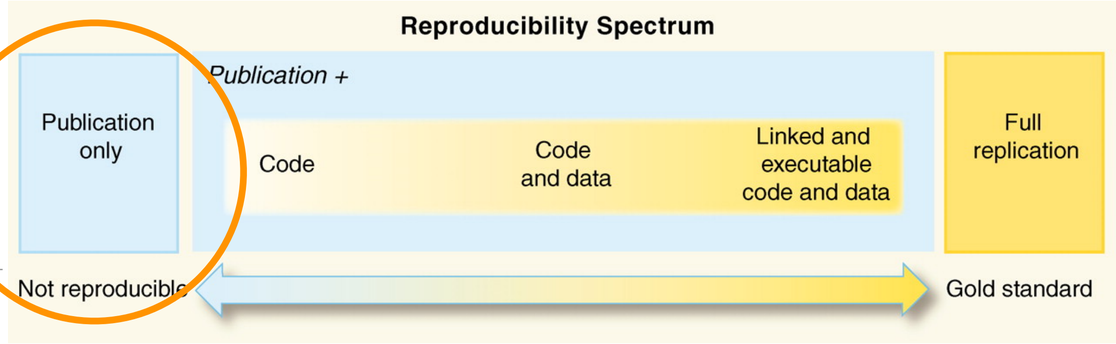
# How to open up your research workflow

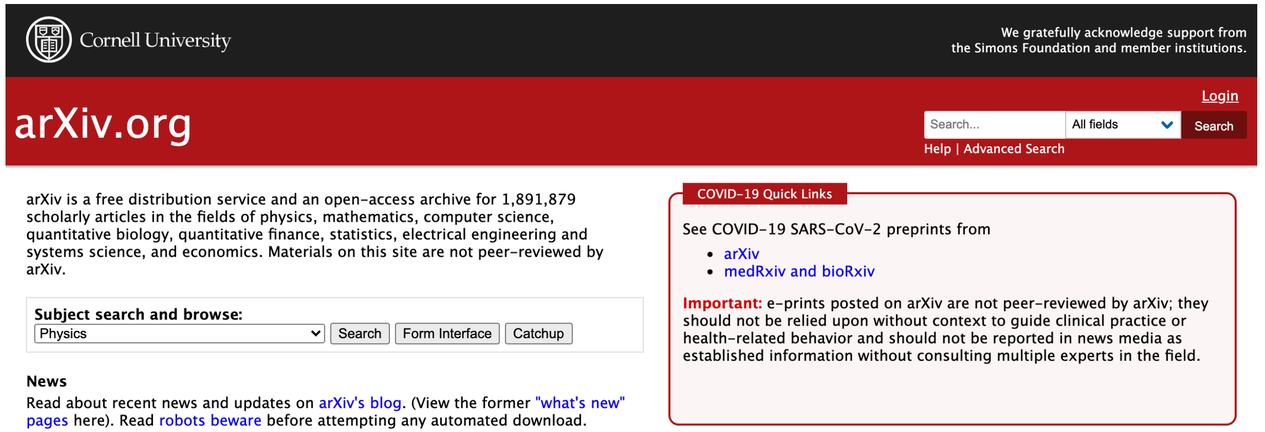



This image was created by Scriberia for *The Turing Way* community and is used under a CC-BY licence. https://doi.org/10.5281/zenodo.3332807

# Publish your research articles Open Access

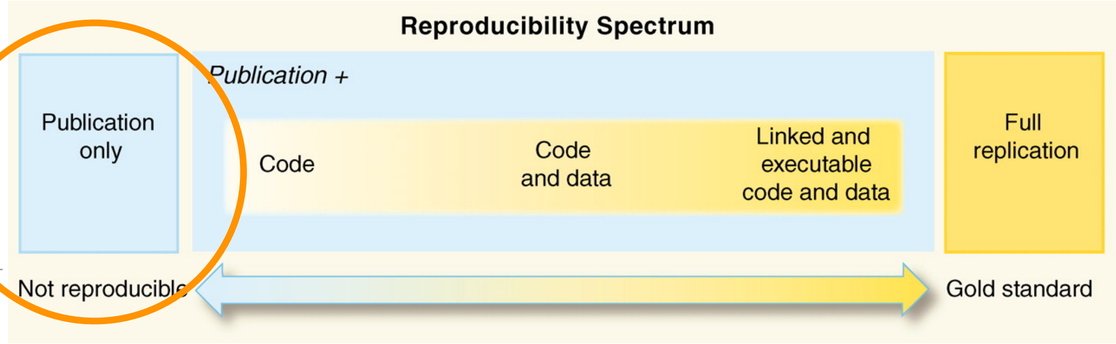
- Gold Open Access
  - Publish in an Open Access journal
  - Can incur high article processing charges (APCs)
- Green Open Access
  - Self-archive a version of the manuscript in a repository (such as arXiv)
  - Share a “pre-print” of your manuscript before peer review
  - Share a “post-print” of your manuscript after peer review

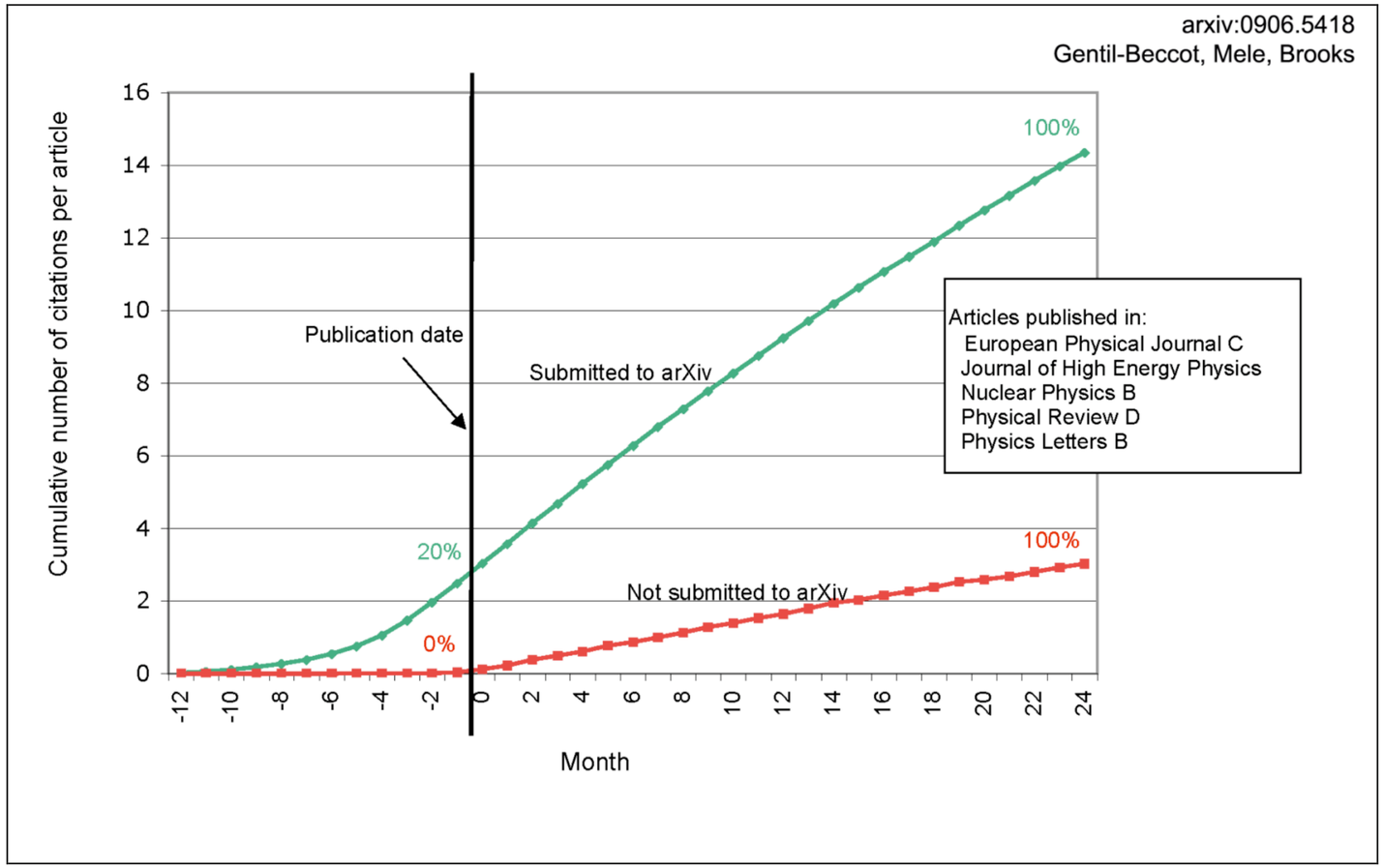


# Publish your research articles Open Access

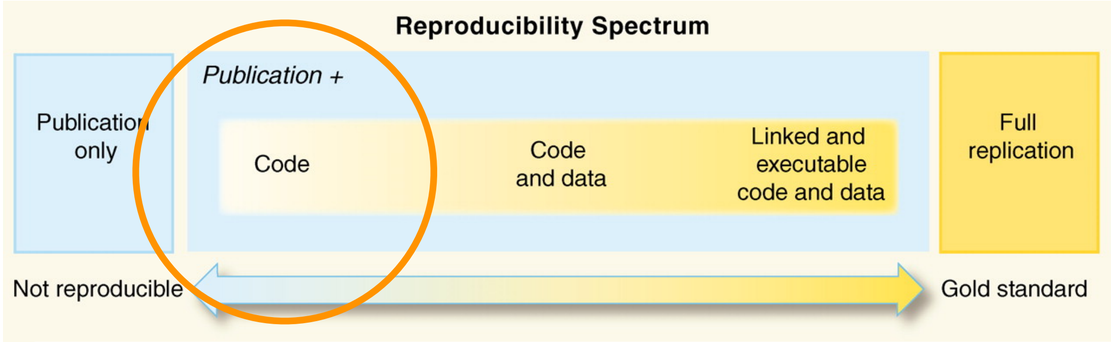
- Gold Open Access
  - Publish in an Open Access journal
  - Can incur high article processing charges (APCs)
- Green Open Access
  - Self-archive a version of the manuscript in a repository (such as arXiv)
  - Share a “pre-print” of your manuscript before peer review
  - Share a “post-print” of your manuscript after peer review



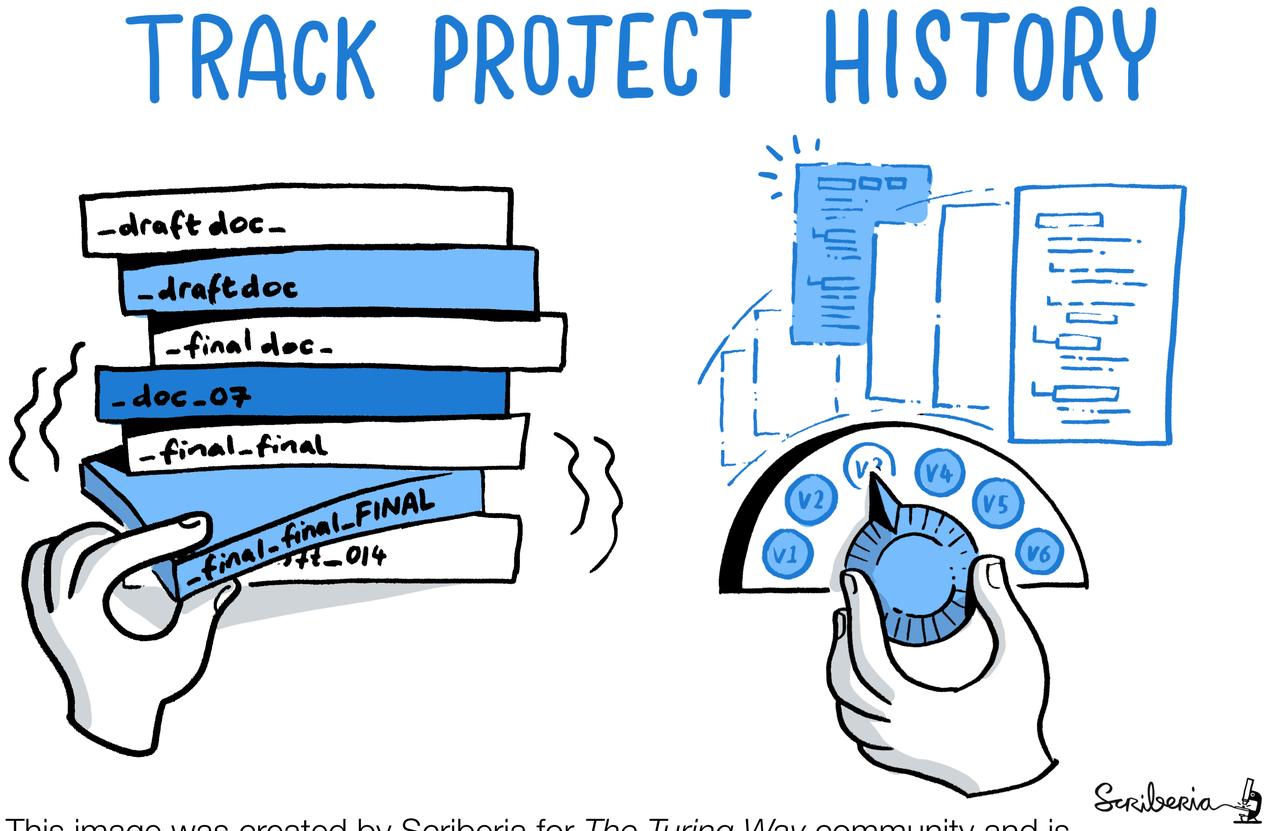




Gentil-Beccot, Mele, Brooks (2009), https://arxiv.org/abs/0906.5418

# Open source and version control your methods



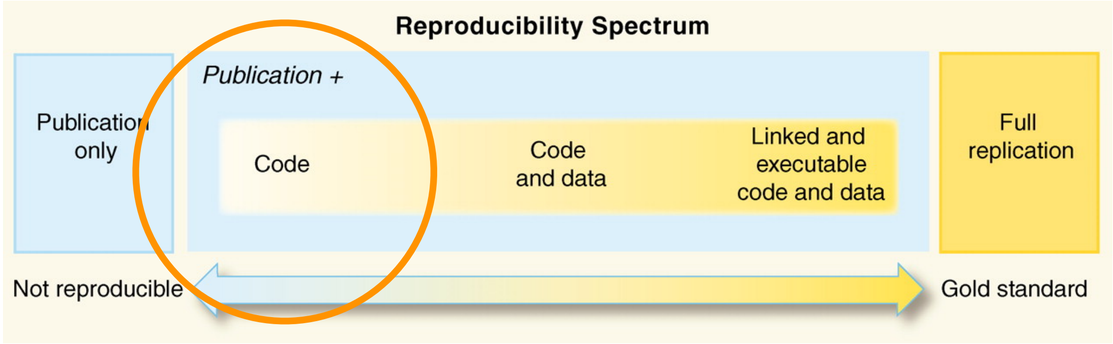

- Version control is the practice of managing and tracking changes in source code and text files (such as through using Git - see talks by Maximilian Nöthe)
  - It stores a history of changes and who made them, allowing you to revert or go back to earlier versions of those files, and understand how contributions by different contributors have changed the project over time (like having a time machine for your project's history!)
- Services such as GitHub, GitLab and Bitbucket add a web-based social and user interface to version control, which facilitates open and collaborative research by enabling you and others to work together on projects from anywhere
  - Added benefit: Can also be an online portfolio and webpage for your work

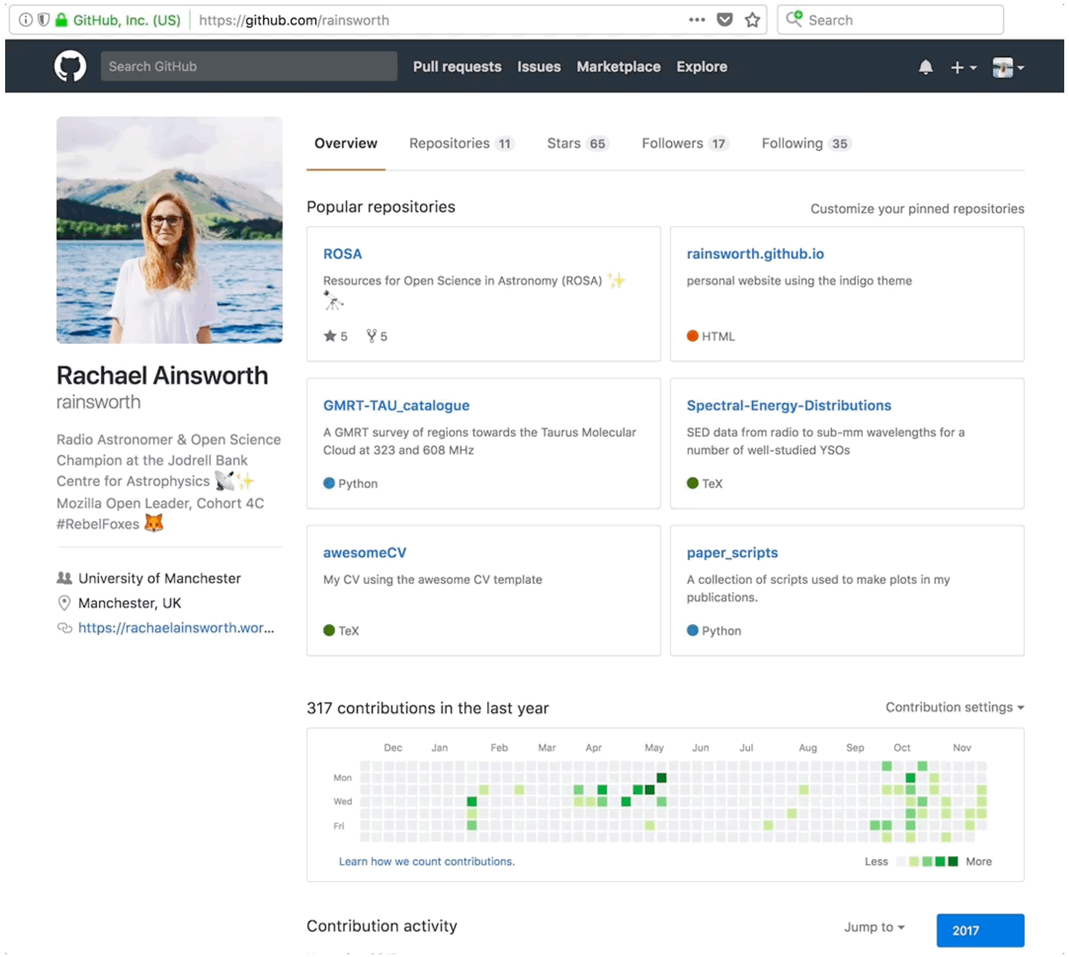



This image was created by Scriberia for *The Turing Way* community and is used under a CC-BY licence. https://doi.org/10.5281/zenodo.3332807

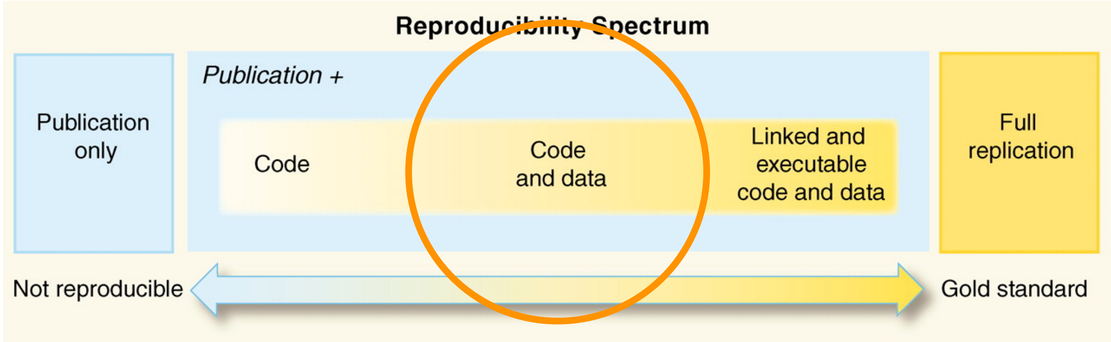
# Open source and version control your methods

- Version control is the practice of managing and tracking changes in source code and text files (such as through using Git - see talks by Maximilian Nöthe)
  - It stores a history of changes and who made them, allowing you to revert or go back to earlier versions of those files, and understand how contributions by different contributors have changed the project over time (like having a time machine for your project's history!)
- Services such as GitHub, GitLab and Bitbucket add a web-based social and user interface to version control, which facilitates open and collaborative research by enabling you and others to work together on projects from anywhere
  - Added benefit: Can also be an online portfolio and webpage for your work

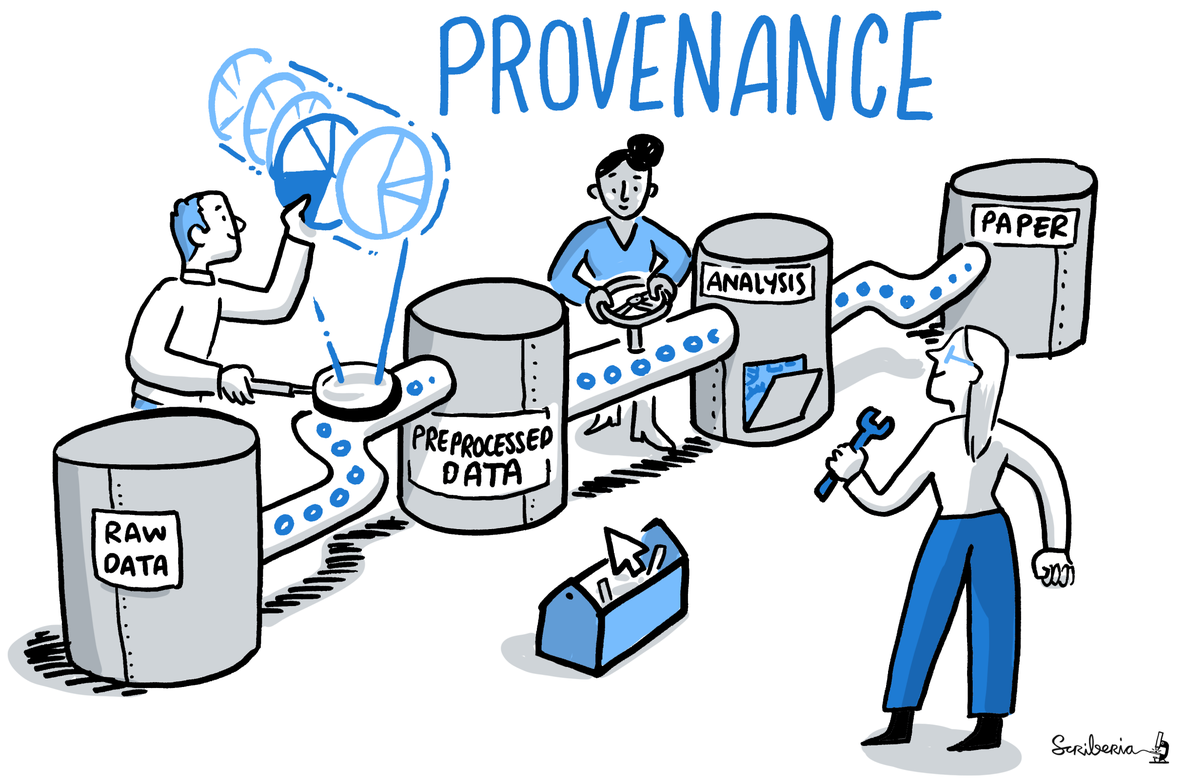


# Make your data open and FAIR




- **Findable**: Data should be described with rich metadata, assigned a globally unique and persistent identifier (such as a DOI), and registered in a searchable resource
- **Accessible**: Provide information on how to access the data (such as via a DOI link)
- **Interoperable**: The data usually need to be integrated with other data and/or interoperate with applications or workflows for analysis, storage, and processing
- **Reusable**: Share data with a clear and accessible data usage license and detailed provenance so that results can be reproduced and built upon

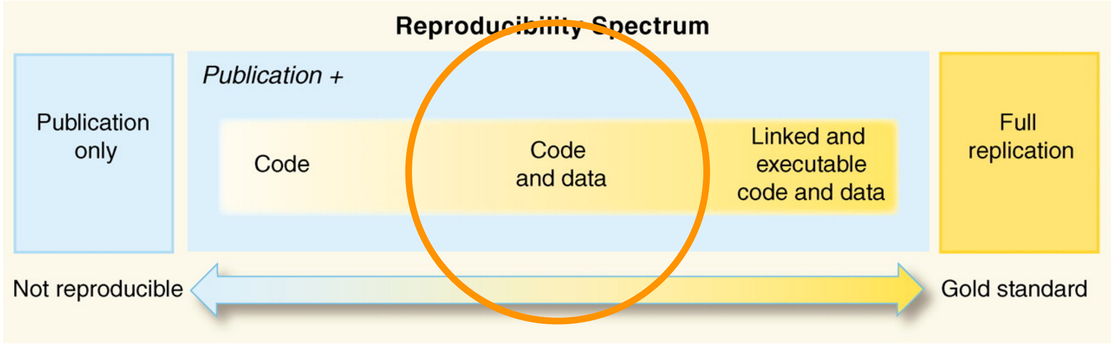


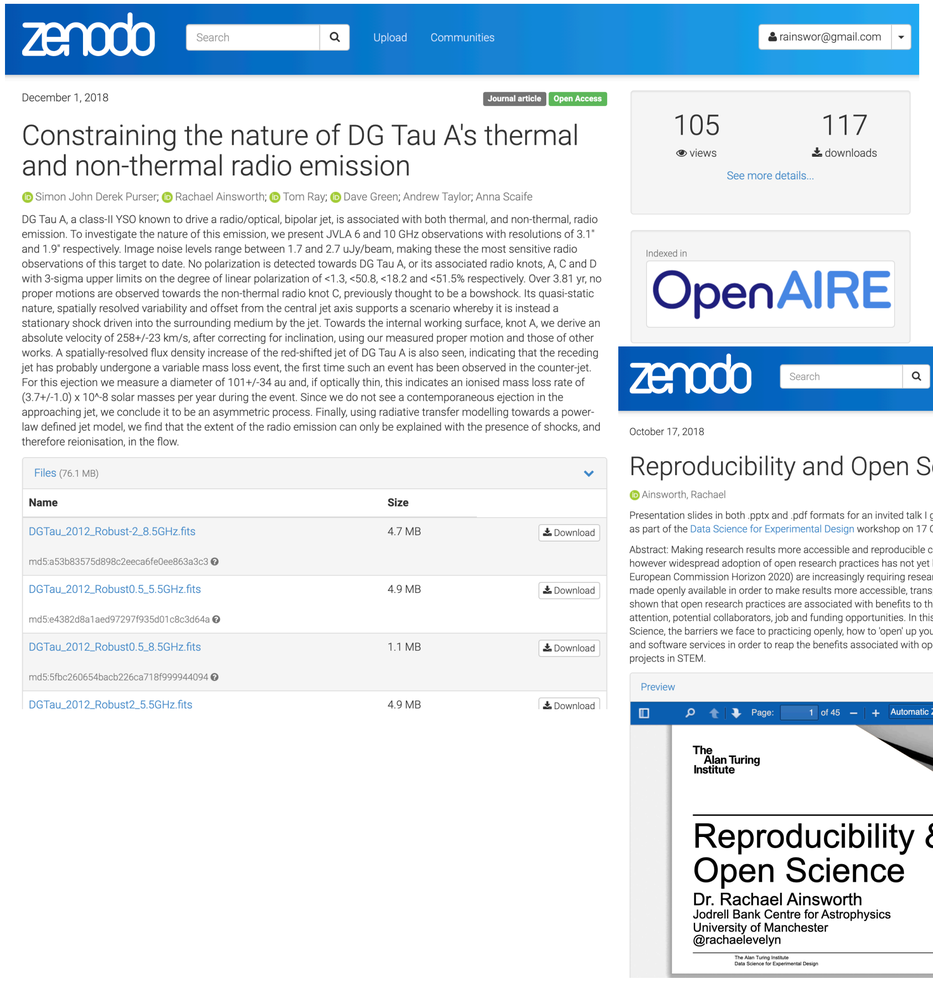

This image was created by Scriberia for *The Turing Way* community and is used under a CC-BY licence. https://doi.org/10.5281/zenodo.3332807

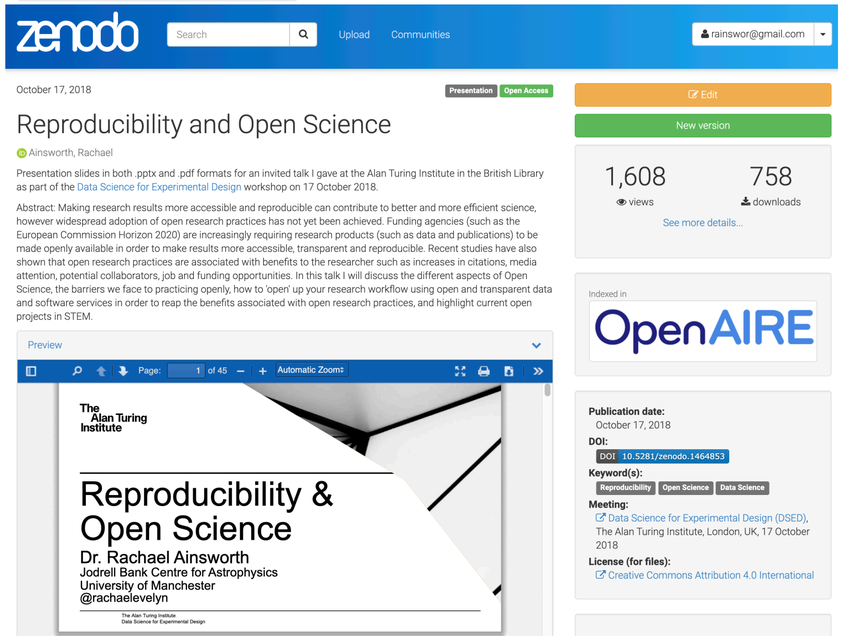
# Make your data open and FAIR

Catch-all repositories such as Zenodo and Figshare enable you to:

- Share research outputs in a wide variety of formats including text, datasets, audio, video & images across all fields of science
- Display your research outputs & get credited by making the research results citable and integrating them into existing reporting lines to funding agencies like the EU
- Easily access and reuse shared research outputs
- Archive your GitHub repository and make citable with Zenodo!

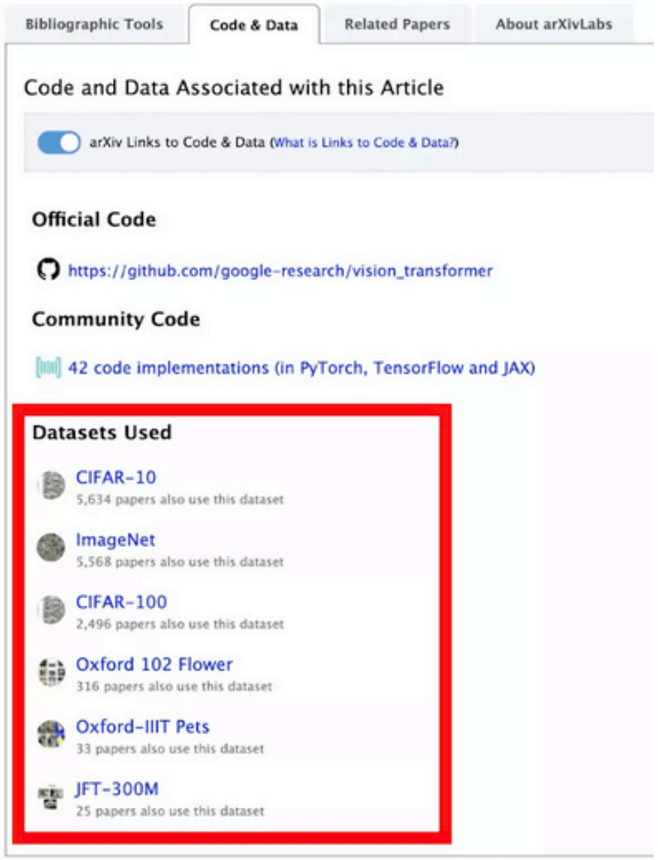


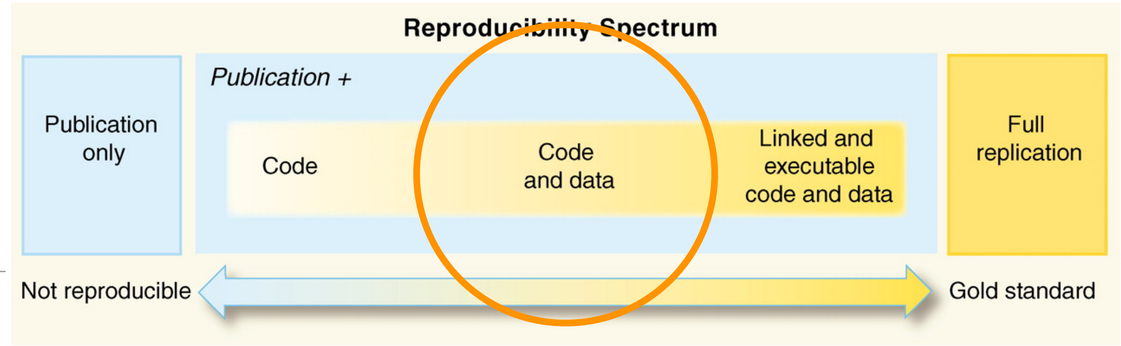



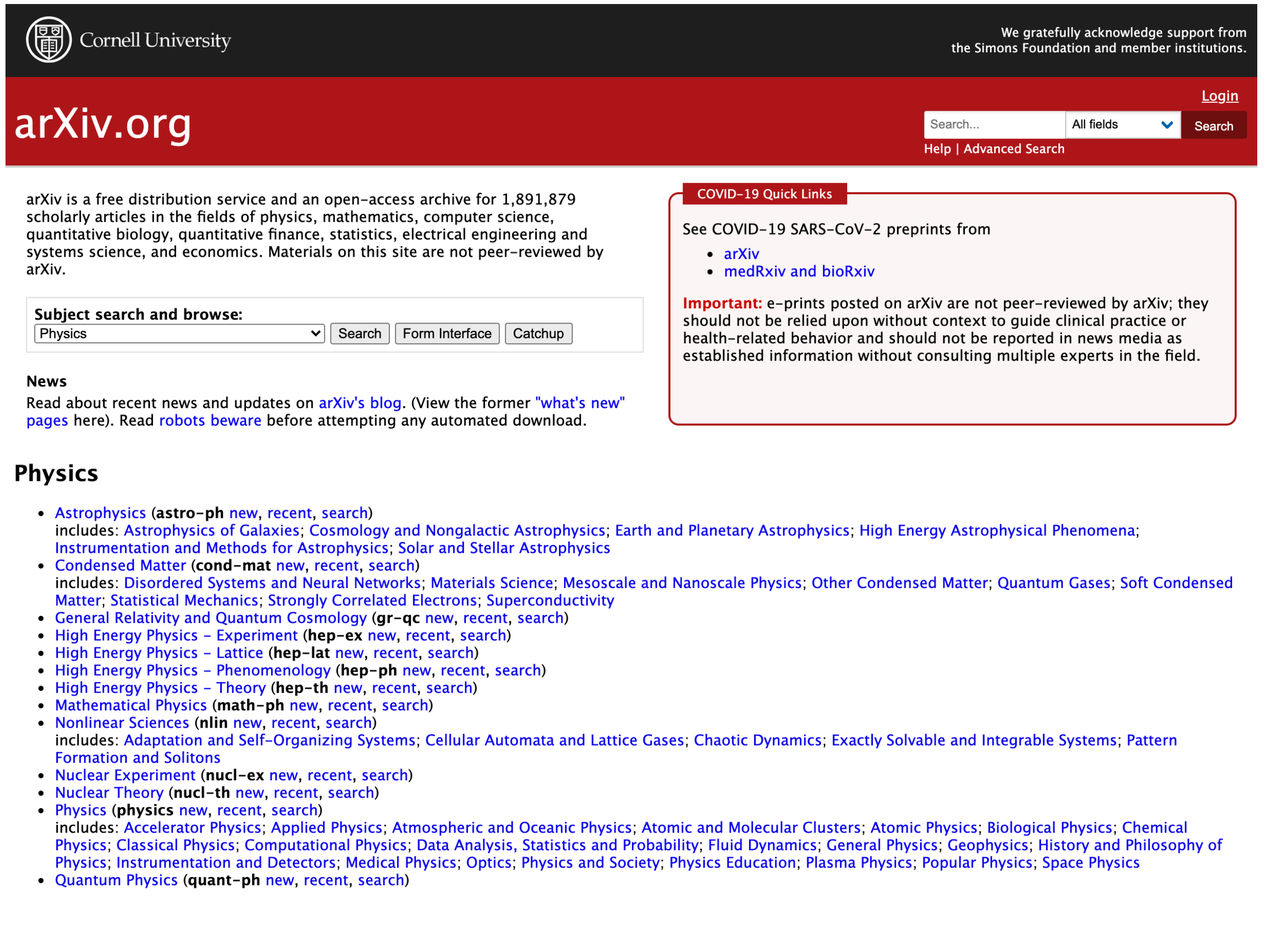



# Make your data open and FAIR

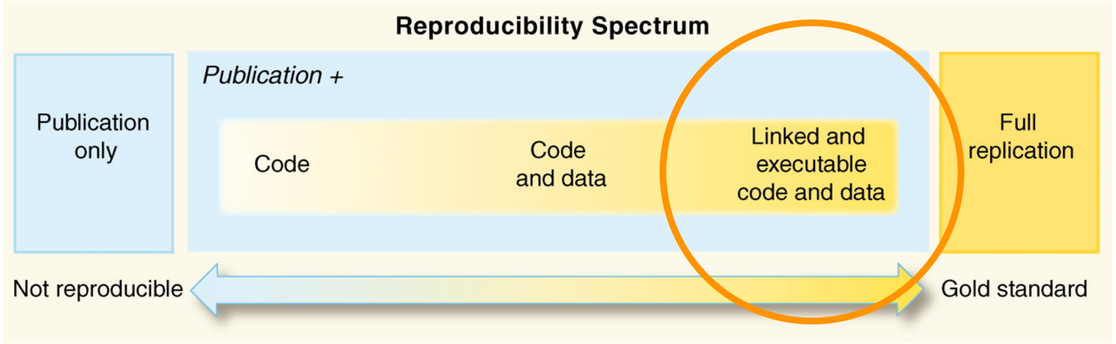
- arXiv now supports sharing code and data associated with your articles



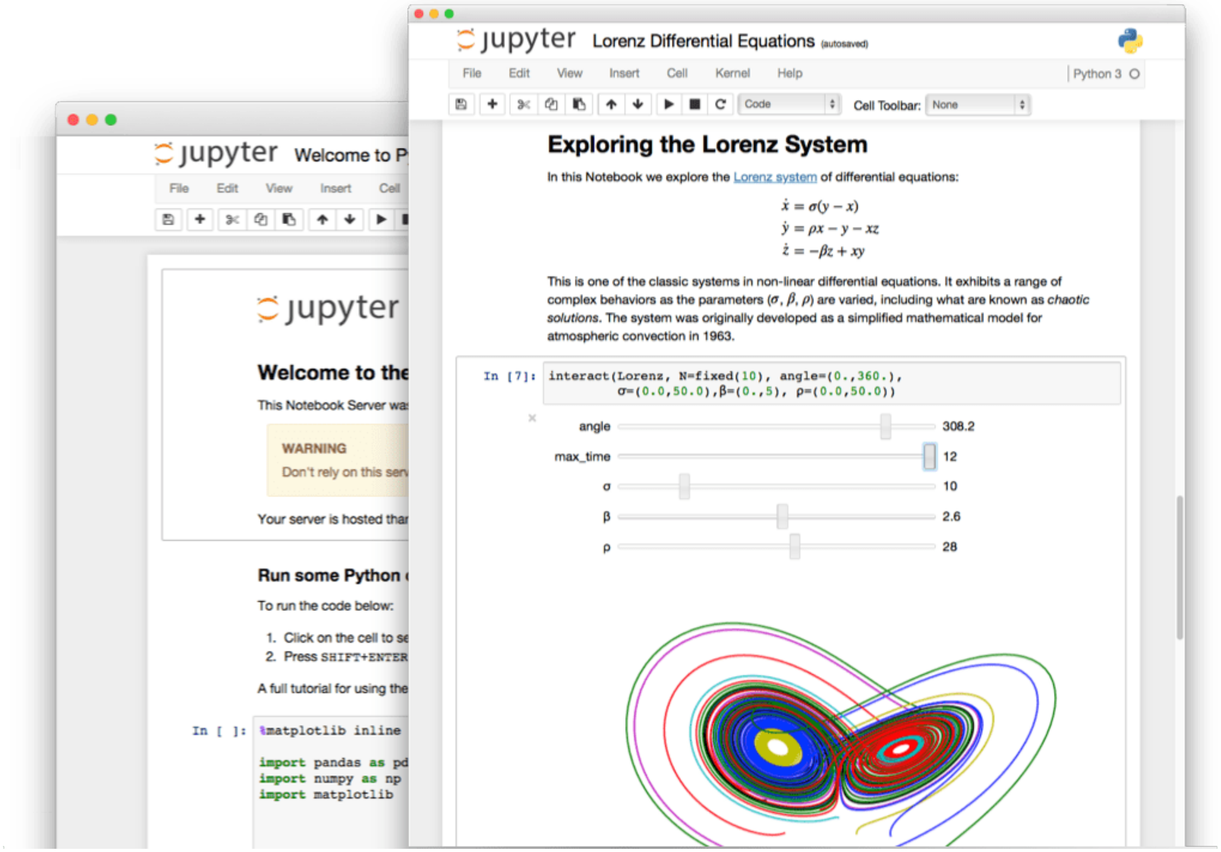
# Document methods and share analyses using Open Notebooks




- Open Notebooks (such as Jupyter) are documents that contain equations, visualisations, narrative text and live code that can be executed independently and interactively, with output visible immediately beneath the input (see talk by Enrique Garcia)

- Notebooks bring together analysis descriptions and results, which can be executed to perform the data analysis in real time

- Added value:
  - Transparency in the analysis of the data
  - Reproducibility
  - Documentation of the entire workflow



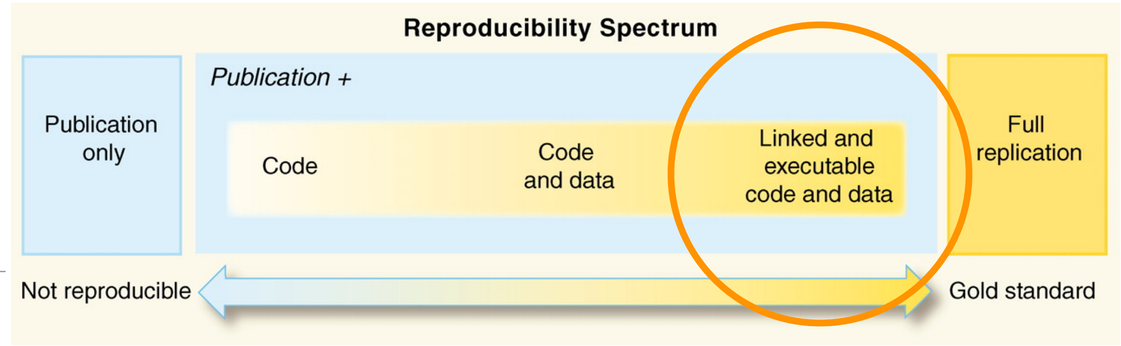

https://jupyter.org/

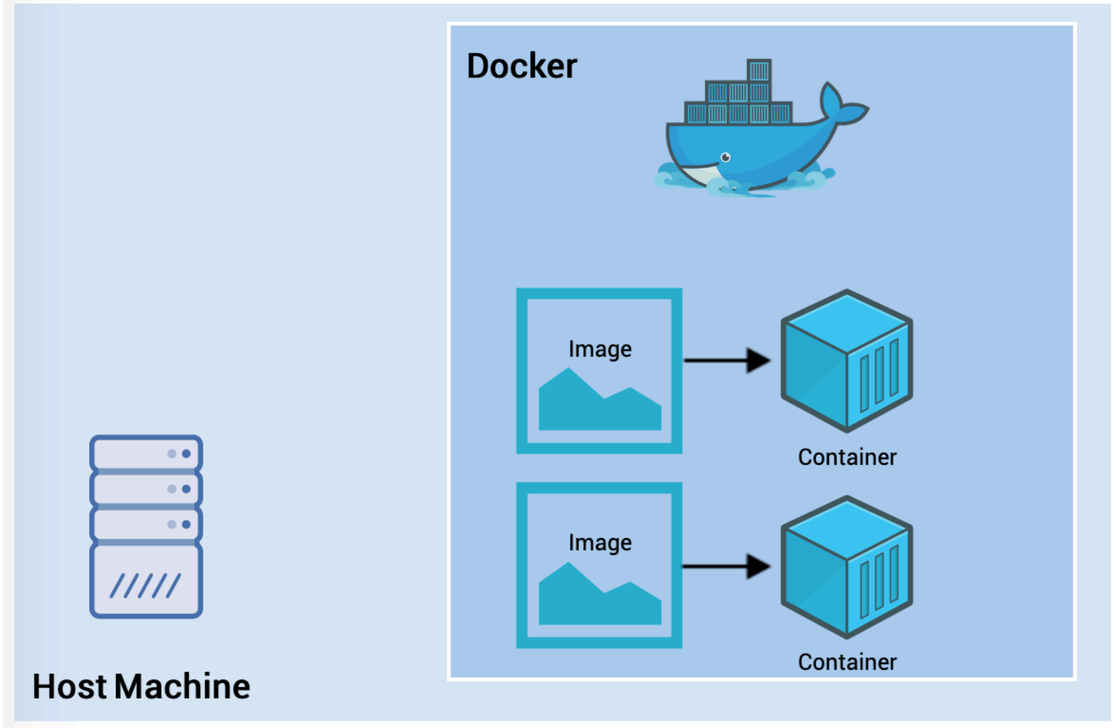

https://doi.org/10.6084/m9.figshare.14710110

“But it worked on my computer...”

# Capture and share your computational environment

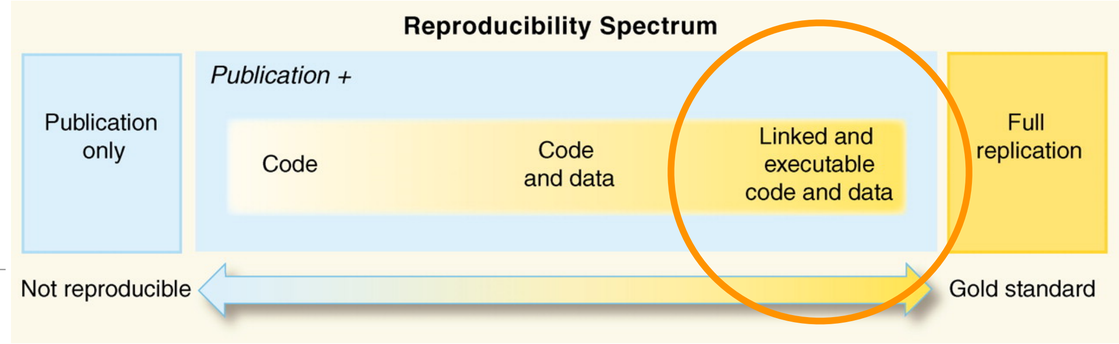
- Containerisation (such as through Docker or Singularity) allows you to package up your entire research workflow - including data, software, analyses and computational environment - so that it can run uniformly and consistently on any infrastructure
- Containers are much more lightweight than virtual machines
- Advantages: bundled dependencies, easy to distribute/share, and stackable

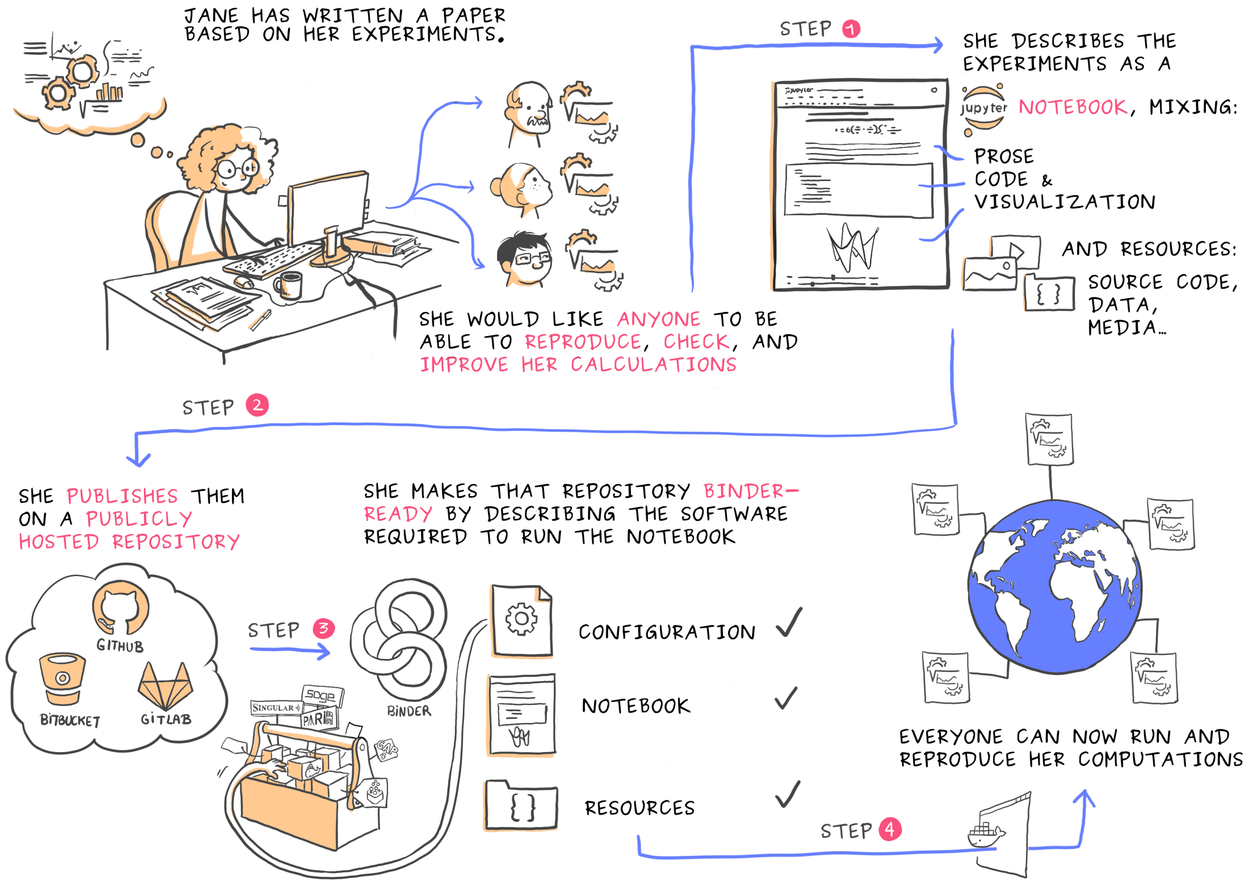


# Make it as easy as possible for others to reproduce your work

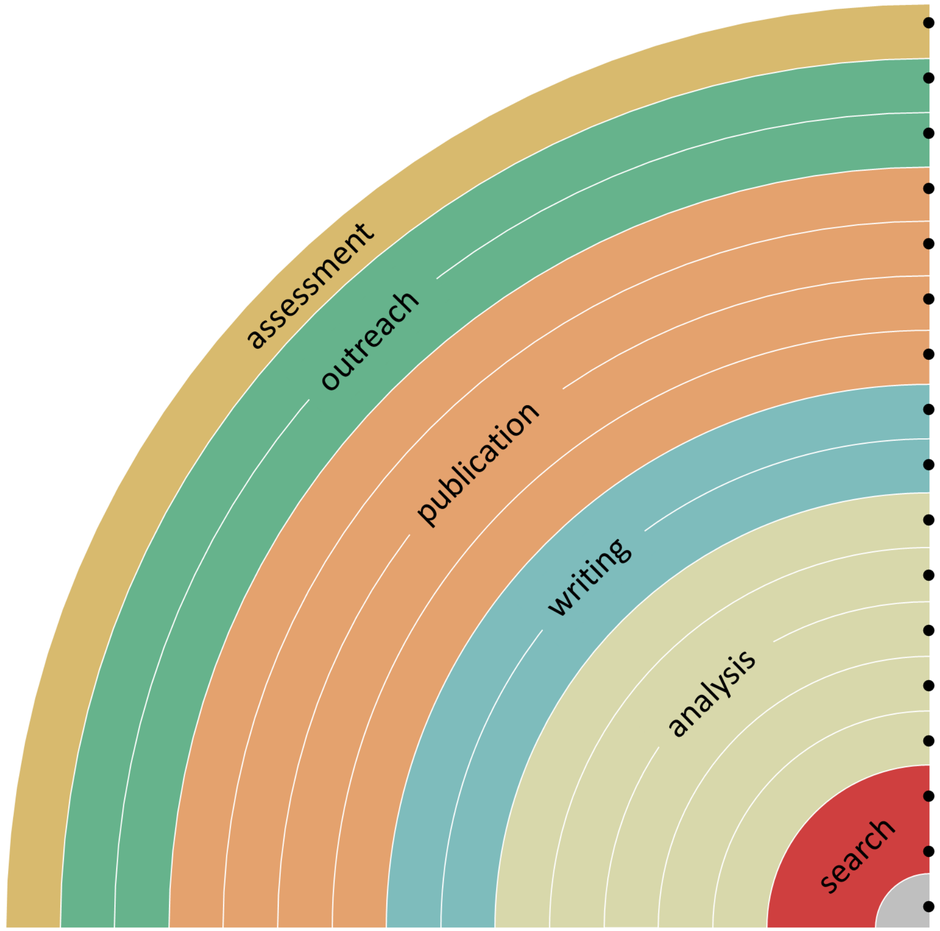
- Binder (mybinder.org) makes it simple to generate reproducible computing environments from a Git repository
  - Generates a Docker image from this repository which will have all the components that you specify along with the Jupyter Notebooks inside
  - You will be able to share a URL with users who can immediately begin interacting with this environment via the cloud




Juliette Taka, Logilab and the OpenDreamKit project

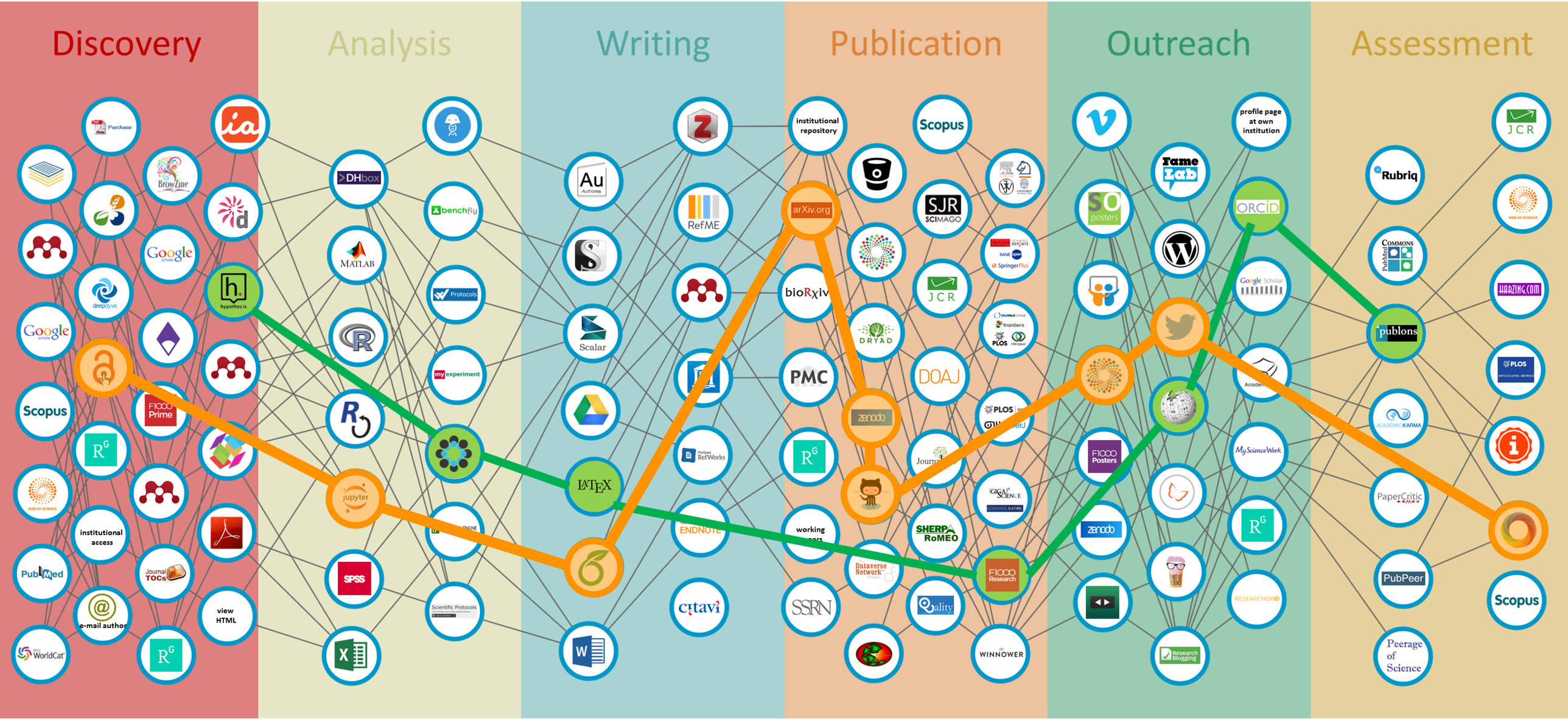
# You can make your workflow more open by…




- adding alternative evaluation, e.g. with altmetrics
- communicating through social media, e.g. Twitter
- sharing posters & presentations, e.g. at FigShare
- using open licenses, e.g. CC0 or CC-BY
- publishing open access, 'green' or 'gold'
- using open peer review, e.g. at journals or PubPeer
- sharing preprints, e.g. at OSF, arXiv or bioRxiv
- using actionable formats, e.g. with Jupyter or CoCalc
- open XML-drafting, e.g. at Overleaf or Authorea
- sharing protocols & workfl., e.g. at Protocols.io
- sharing notebooks, e.g. at OpenNotebookScience
- sharing code, e.g. at GitHub with GNU/MIT license
- sharing data, e.g. at Dryad, Zenodo or Dataverse
- pre-registering, e.g. at OSF or AsPredicted
- commenting openly, e.g. with Hypothes.is
- using shared reference libraries, e.g. with Zotero
- sharing (grant) proposals, e.g. at RIO

*Bianca Kramer & Jeroen Bosman https://101innovations.wordpress.com*

DOI: 10.5281/zenodo.1147025




Jeroen Bosman and Bianca Kramer - https://101innovations.wordpress.com/workflows/

# Open Science examples in Astronomy

# Open Science in Astronomy Examples

**Open Access**:

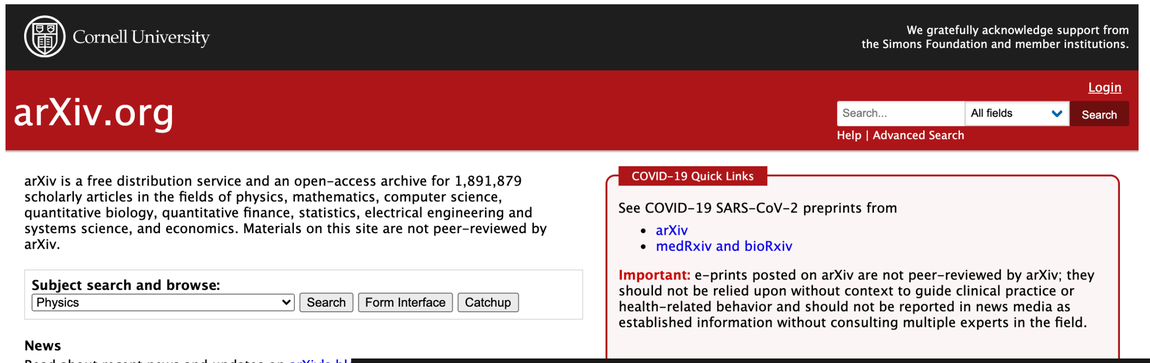
- arXiv! Started in August 1991 and provides open access to 1,891,879+ e-prints in (Astro)Physics and many other fields

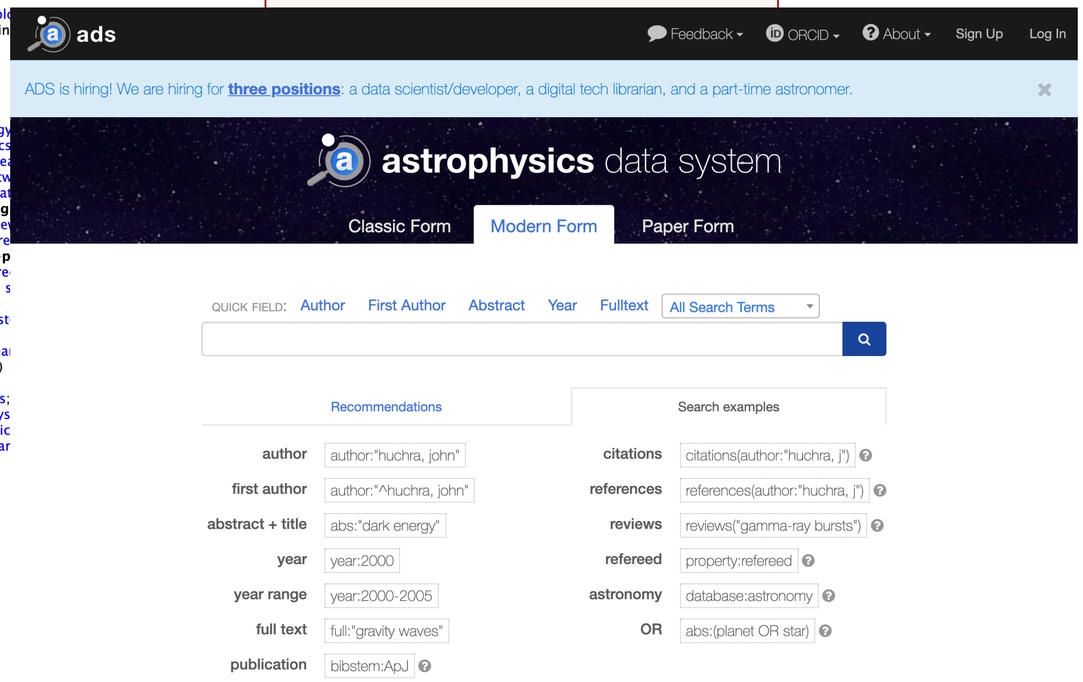
**Open Data**:

- SAO/NASA Astrophysics Data System (ADS)
- Raw data via instrument archives
- Surveys through VizieR
- Meta-data through Simbad

**Open Source**:

- Projects and tools such as Astropy
- The CASA pipeline for e-MERLIN data




http://ileo.de/2017/11/13/astronomy-as-an-example-for-an-open-science/

# Open Science in Astronomy Examples

**Citizen Science**:

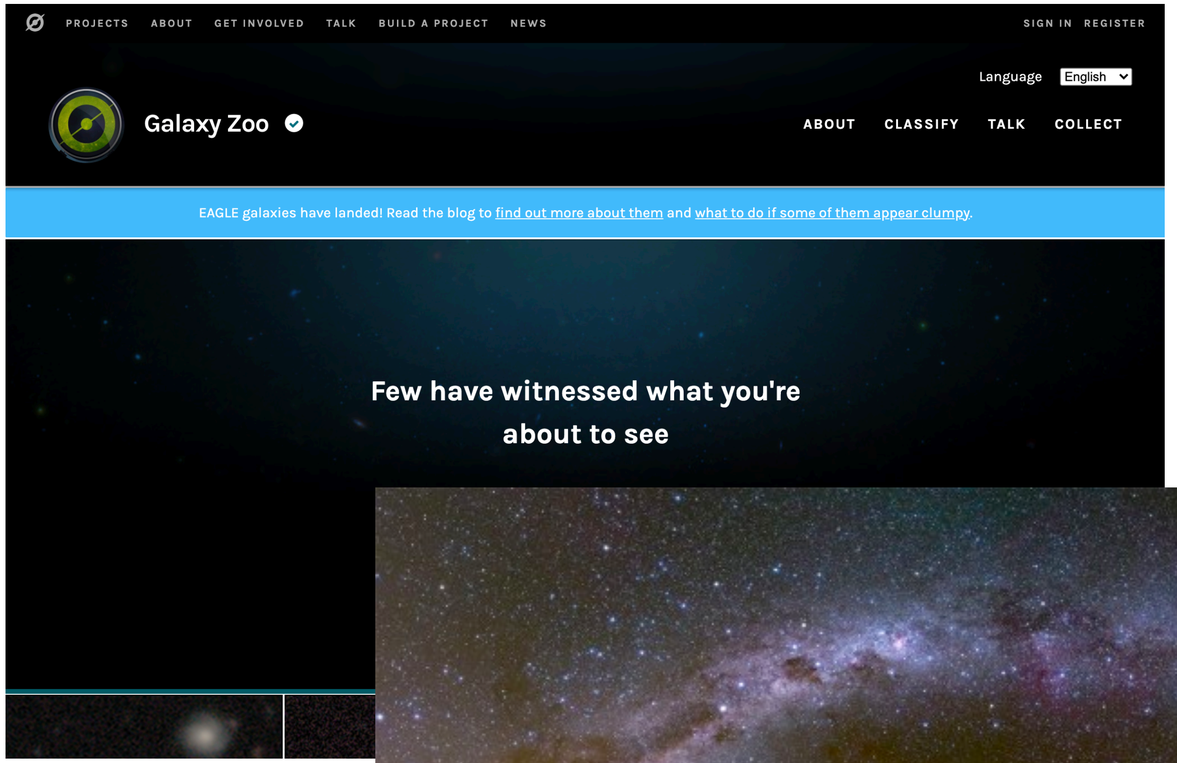
- Pulsar Hunters
- Galaxy Zoo

**Open Educational Resources**:

- ESCAPE Summer School
- astroEDU

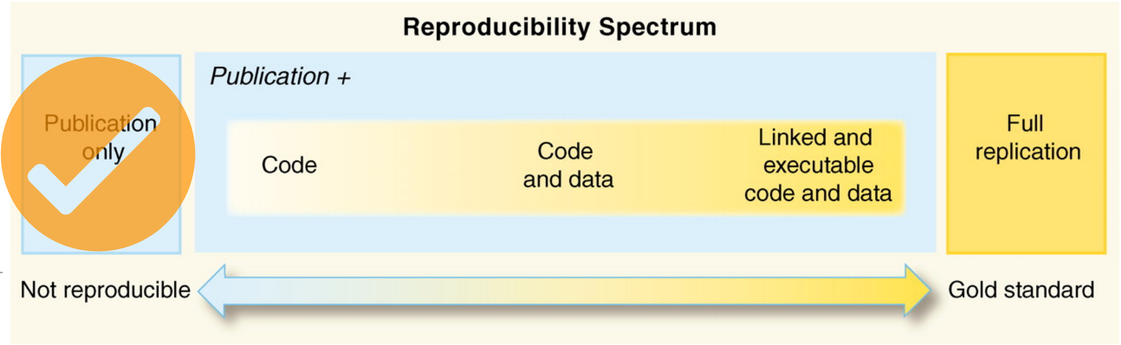
**Equity, Diversity and Inclusion**:

- AstroBetter Diversity wiki
- European Astronomical Society (EAS) Annual Meeting (formerly known as EWASS) sessions dedicate to EDI
- Development in Africa with Radio Astronomy (DARA)




Nature Astronomy focus on Diversity, Equity and Inclusion best practices and solutions. 2019. https://www.nature.com/collections/cagjjdfjaa/

# Case Study 1:
# Open Access publication




arXiv.org > astro-ph > arXiv:1804.01548

Search... All fields Search

Help | Advanced Search

Astrophysics > High Energy Astrophysical Phenomena

## Constraining Redshifts of Unlocalised Fast Radio Bursts

C. R. H. Walker, Y.-Z. Ma, R. P. Breton

*(Submitted on 4 Apr 2018)*

The population of fast radio bursts (FRBs) will continue to diverge into two groups depending on their method of discovery: those which can be localised, and those which cannot. Events potentially less useful for astronomical and cosmological purposes due to limited localisation will accumulate with the advent of new facilities and continued efforts by, e.g., the SUPERB collaboration, which may require afterglows or multi-wavelength counterparts for sub-arcsecond localisation. It is important to exploit these sources to their maximum scientific potential. We perform analysis of FRB dispersion measures (DMs), considering different theoretical FRB progenitors with view to place more rigorous constraints on FRB redshifts, in particular for large statistical samples, via their DMs. We review FRB DM components, and build redshift-scalable probability distributions corresponding to different progenitor scenarios. We combine these components into a framework for obtaining FRB DM probabilities given their redshifts. Taking into account different possibilities for the evolution of progenitors across cosmic time we invert this model, thus deriving redshift constraints. Effects of varying FRB progenitor models are illustrated. While, as expected, host galaxy DM contributions become decreasingly important with increasing redshift, for AGN-like progenitor scenarios they could remain significant out to redshift 3. Constraints are placed on redshifts of catalogued FRBs with various models and increasingly realistic models may be employed as general understanding of FRBs improves. For localised FRBs, we highlight future prospects for disentangling host and intergalactic medium DM components using their respective redshift scaling. We identify a use for large samples of unlocalised FRBs resulting from upcoming flux-limited surveys, such as with CHIME, in mapping out the Milky Way contribution to the DM.

Comments: 13 pages, 8 figures, submitted for publication in Astronomy & Astrophysics on 04/04/2018
Subjects: **High Energy Astrophysical Phenomena (astro-ph.HE)**
Cite as: **arXiv:1804.01548 [astro-ph.HE]**
(or **arXiv:1804.01548v1 [astro-ph.HE]** for this version)

### Submission history

From: Charles Walker [view email]
**[v1]** Wed, 4 Apr 2018 18:03:06 UTC (897 KB)

*Which authors of this paper are endorsers? | Disable MathJax (What is MathJax?)*

**Download:**
- PDF
- Other formats
(license)

Current browse context:
**astro-ph.HE**
< prev | next >
new | recent | 1804

Change to browse by:
astro-ph

References & Citations
- INSPIRE HEP
(refers to | cited by )
- NASA ADS

Google Scholar

Bookmark

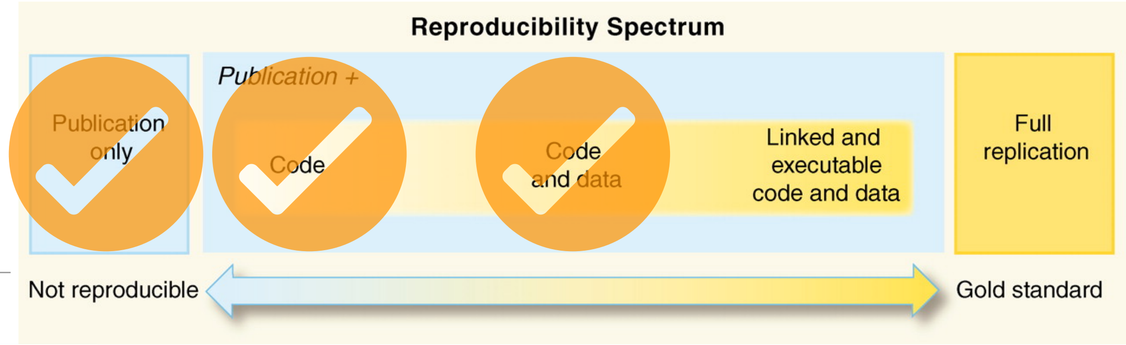
### 4.4. Concluding remarks

We present a framework for exploration of the statistical relationship between FRB redshifts and dispersion measures, which provides the basis for:

1. Qualitative assessment of host galaxy contributions to FRB DMs using realistic models. We find that all our host models may contribute large amounts of DM ($> 400$ pc cm$^{-3}$) in the rest frame, and as expected, that $\mathrm{DM_{host}}$ is most significant for FRBs of lower source redshifts, becoming negligible as redshift increases. For the most extreme scenarios where FRBs originate close to galactic centers, this component still contributes significantly to overall $P(\mathrm{DM}|z_s)$ profiles out to $z_s = 3$.
2. More rigorous uncertainties to be placed on FRB redshifts than are currently standard practice. By consulting $P(z_s|\mathrm{DM})$ probability distributions created from our (or similar) models, this may additionally provide an innovative way to narrow down the potential host galaxies for unlocalised FRBs, and allow insight into FRB progenitors to be drawn from large source populations. A repository containing our Python code and examples may be found online at `https://doi.org/10.5281/zenodo.1209920`.
3. The disentanglement of individual FRB dispersion measure components. For example, the MW components for given sightlines could be extracted from $\mathrm{DM_{obs}}$ by comparing DM probability distributions from a flux-limited survey (e.g. CHIME) at different sky locations and looking for systematic offsets in their profiles. This technique would not require redshift measurements, thus further increasing the usefulness of unlocalised FRBs. It also could be possible to separate $\mathrm{DM_{IGM}}$ and $\mathrm{DM_{host}}$ using their respective redshift dependences.

Walker, Ma & Breton, https://arxiv.org/abs/1804.01548

# Case Study 1:
# Open Source code & Open Data




zenodo | Search | Upload | Communities | rainswor@gmail.com

April 2, 2018 — Software — Open Access

## mbcxqcw2/EEModel: Master DOI release

mbcxqcw2

Updated zenodo DOI link to always link to the latest github version

Preview

EEModel-v1.03.zip

- mbcxqcw2-EEModel-268b3da
  - ExcessElectronLib.py 18.1 kB
  - ExcessElectronModel.ipynb 505.5 kB
  - FRBcat_FRB_DMs.csv 2.2 kB
  - README.md 760 Bytes
  - cosmo_consts.py 907 Bytes
  - host_galaxies
    - OB_FRBs_list.txt 17.8 MB
    - elliptical_FRBs_list.txt 17.8 MB
    - msp_FRBs_list.txt 17.8 MB
    - old_FRBs_list.txt 17.8 MB
    - young_FRBs_list.txt 17.8 MB
  - lin_mp_files
    - planck_lin_mp0.0.dat 21.5 kB
    - planck_lin_mp0.1.dat 21.5 kB
    - planck_lin_mp0.2.dat 21.5 kB
    - planck_lin_mp0.3.dat 21.5 kB
    - planck_lin_mp0.4.dat 21.5 kB
    - planck_lin_mp0.5.dat 21.5 kB

Files (39.6 MB)

| Name | Size | |
|---|---|---|
| mbcxqcw2/EEModel-v1.03.zip | 39.6 MB | Preview Download |

md5:e8e62d8974085cea865f585488722a0b

315 views — 191 downloads

See more details...

Available in GitHub

Indexed in OpenAIRE

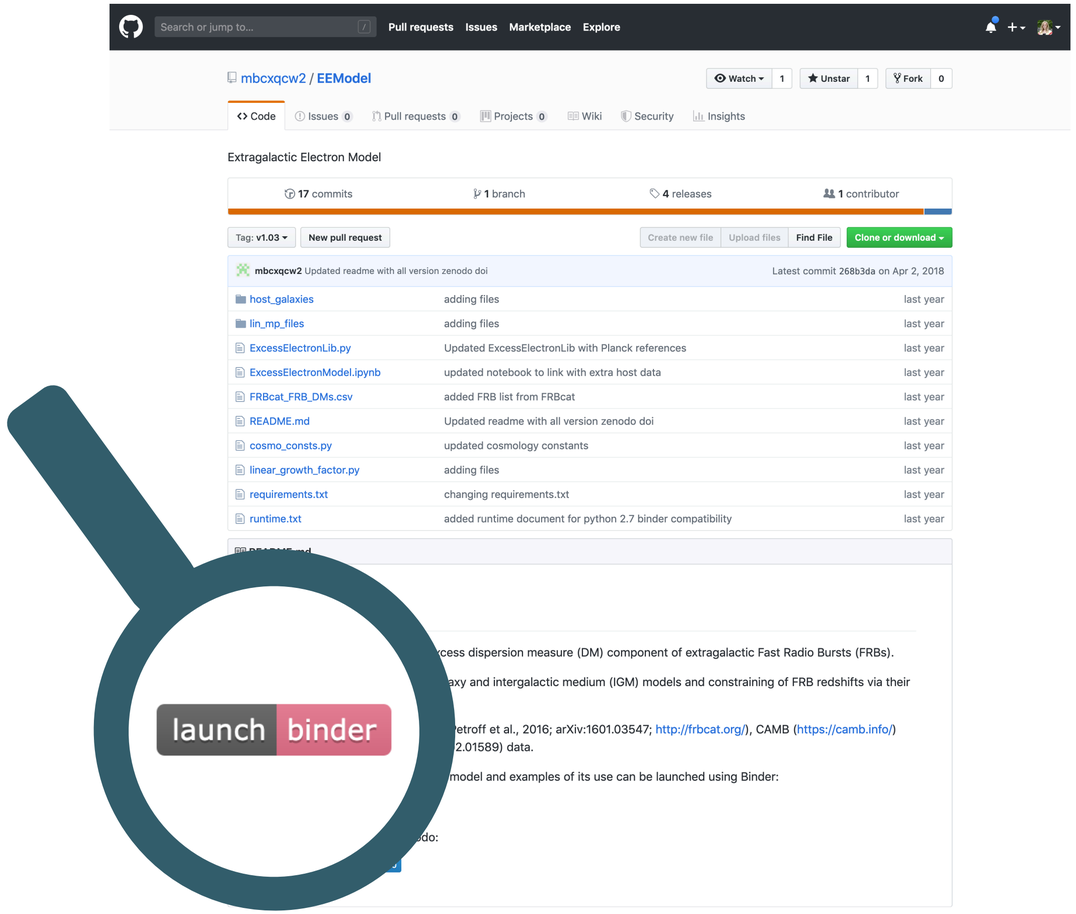
**Publication date:** April 2, 2018

**DOI:** DOI 10.5281/zenodo.1211089

**Related identifiers:** Supplement to https://github.com/mbcxqcw2/EEModel/tree/v1.03

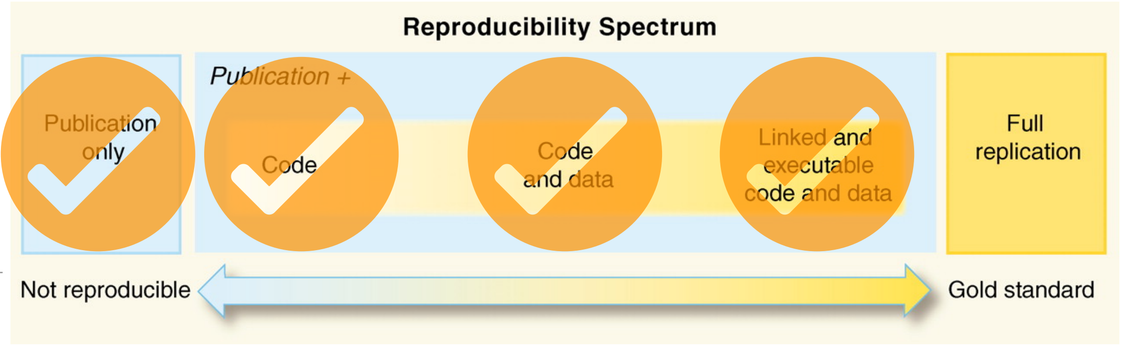
**License (for files):** Other (Open)

https://doi.org/10.5281/zenodo.1209920




https://github.com/mbcxqcw2/EEModel/tree/v1.03

# Case Study 1: Reproducible Computational Environment




binder

Starting repository: mbcxqcw2/EEModel/master

New to Binder? Check out the Binder Documentation for more information.

Build logs show

Here's a non-interactive preview on nbviewer while we start a server for you. Your binder will open automatically when it is ready.

jupyter nbviewer JUPYTER FAQ

EEModel master

| Name |
| --- |
| mbcxqcw2's repositories |
| host_galaxies |
| lin_mp_files |
| ExcessElectronModel.ipynb |
| ExcessElectronLib.py |
| FRBcat_FRB_DMs.csv |
| README.md |
| cosmo_consts.py |
| linear_growth_factor.py |
| requirements.txt |
| runtime.txt |

jupyter ExcessElectronModel (unsaved changes)

File Edit View Insert Cell Kernel Widgets Help Not Trusted Python 2

Run Code

In [1]:
```
%reset
%matplotlib inline
print "Done."
```

```
Once deleted, variables cannot be recovered. Proceed (y/[n])? y
/home/charlie/anaconda2/lib/python2.7/site-packages/matplotlib/font_manager.py:273: UserWarning: Matplotlib is buildi
ng the font cache using fc-list. This may take a moment.
  warnings.warn('Matplotlib is building the font cache using fc-list. This may take a moment.')
Done.
```

## Imports

In [2]:
```
#standard imports
import numpy as np
from matplotlib import pyplot as plt
from scipy.interpolate import interp1d as i1d

#other imports
from linear_growth_factor import E
#from WMB_scratch import Plognormal
from ExcessElectronLib import Prob_IGM # IGM distribution
from ExcessElectronLib import Convolve # Convolution function
from ExcessElectronLib import NormConv as Normalise # Normalisation function for P(DM|Z)
from ExcessElectronLib import NormTranspose # Normalisation function for P(z|DM)
from ExcessElectronLib import find_nearest # function to find nearest value in discrete array to specified value
from ExcessElectronLib import FindErrorRange # function to find min/max bounds for a PDF

print 'Imports done.'
```

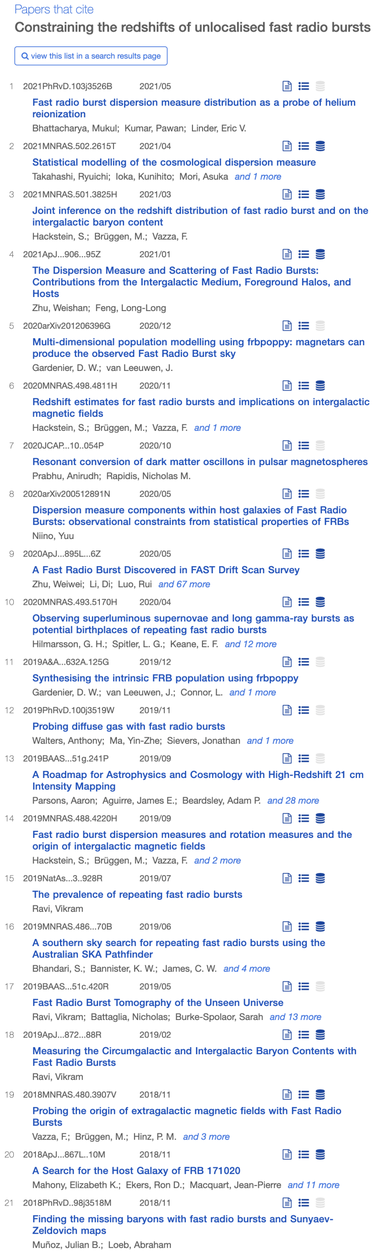
```
Imports done.
```

## Import Host Galaxy Distributions

In [3]:
```
########################################
###Stellar distributed FRBs in spirals###
########################################

##OB STARS##
print "OB..."
OB_data=np.loadtxt('./host_galaxies/OB_FRBs_list.txt')
OB_DMs = zip(*np.array(OB_data))[0][:]

##YOUNG PULSARS##
print "YPSR..."
YPSR_data=np.loadtxt('./host_galaxies/young_FRBs_list.txt')
YPSR_DMs = zip(*np.array(YPSR_data))[0][:]

##OLD PULSARS##
print "OPSR..."
OPSR_data=np.loadtxt('./host_galaxies/old_FRBs_list.txt')
OPSR_DMs = zip(*np.array(OPSR_data))[0][:]

##MSPS##
print "MSP..."
MSP_data=np.loadtxt('./host_galaxies/msp_FRBs_list.txt')
MSP_DMs = zip(*np.array(MSP_data))[0][:]

"""
#Note: these are commented out to prevent importing huge numbers of files.

########################################
#Homogenously distributed FRBs in spirals#
########################################
```

# Case Study 1: Impact

- 4 April 2018:
  - Submitted manuscript to journal
  - Deposited pre-print to arXiv
- 9 April 2018: Received referee report!
- 9 June 2020: Officially published by the journal https://doi.org/10.1051/0004-6361/201833157
  - 14 citations before officially published by the journal
  - 21 citations to date



Papers that cite

Constraining the redshifts of unlocalised fast radio bursts

view this list in a search results page

1 2021PhRvD.103j3526B 2021/05
**Fast radio burst dispersion measure distribution as a probe of helium reionization**
Bhattacharya, Mukul; Kumar, Pawan; Linder, Eric V.

2 2021MNRAS.502.2615T 2021/04
**Statistical modelling of the cosmological dispersion measure**
Takahashi, Ryuichi; Ioka, Kunihito; Mori, Asuka *and 1 more*

3 2021MNRAS.501.3825H 2021/03
**Joint inference on the redshift distribution of fast radio burst and on the intergalactic baryon content**
Hackstein, S.; Brüggen, M.; Vazza, F.

4 2021ApJ...906...95Z 2021/01
**The Dispersion Measure and Scattering of Fast Radio Bursts: Contributions from the Intergalactic Medium, Foreground Halos, and Hosts**
Zhu, Weishan; Feng, Long-Long

5 2020arXiv201206396G 2020/12
**Multi-dimensional population modelling using frbpoppy: magnetars can produce the observed Fast Radio Burst sky**
Gardenier, D. W.; van Leeuwen, J.

6 2020MNRAS.498.4811H 2020/11
**Redshift estimates for fast radio bursts and implications on intergalactic magnetic fields**
Hackstein, S.; Brüggen, M.; Vazza, F. *and 1 more*

7 2020JCAP...10..054P 2020/10
**Resonant conversion of dark matter oscillons in pulsar magnetospheres**
Prabhu, Anirudh; Rapidis, Nicholas M.

8 2020arXiv200512891N 2020/05
**Dispersion measure components within host galaxies of Fast Radio Bursts: observational constraints from statistical properties of FRBs**
Niino, Yuu

9 2020ApJ...895L...6Z 2020/05
**A Fast Radio Burst Discovered in FAST Drift Scan Survey**
Zhu, Weiwei; Li, Di; Luo, Rui *and 67 more*

10 2020MNRAS.493.5170H 2020/04
**Observing superluminous supernovae and long gamma-ray bursts as potential birthplaces of repeating fast radio bursts**
Hilmarsson, G. H.; Spitler, L. G.; Keane, E. F. *and 12 more*

11 2019A&A...632A.125G 2019/12
**Synthesising the intrinsic FRB population using frbpoppy**
Gardenier, D. W.; van Leeuwen, J.; Connor, L. *and 1 more*

12 2019PhRvD.100j3519W 2019/11
**Probing diffuse gas with fast radio bursts**
Walters, Anthony; Ma, Yin-Zhe; Sievers, Jonathan *and 1 more*

13 2019BAAS...51g.241P 2019/09
**A Roadmap for Astrophysics and Cosmology with High-Redshift 21 cm Intensity Mapping**
Parsons, Aaron; Aguirre, James E.; Beardsley, Adam P. *and 28 more*

14 2019MNRAS.488.4220H 2019/09
**Fast radio burst dispersion measures and rotation measures and the origin of intergalactic magnetic fields**
Hackstein, S.; Brüggen, M.; Vazza, F. *and 2 more*

15 2019NatAs...3..928R 2019/07
**The prevalence of repeating fast radio bursts**
Ravi, Vikram

16 2019MNRAS.486...70B 2019/06
**A southern sky search for repeating fast radio bursts using the Australian SKA Pathfinder**
Bhandari, S.; Bannister, K. W.; James, C. W. *and 4 more*

17 2019BAAS...51c.420R 2019/05
**Fast Radio Burst Tomography of the Unseen Universe**
Ravi, Vikram; Battaglia, Nicholas; Burke-Spolaor, Sarah *and 13 more*

18 2019ApJ...872...88R 2019/02
**Measuring the Circumgalactic and Intergalactic Baryon Contents with Fast Radio Bursts**
Ravi, Vikram

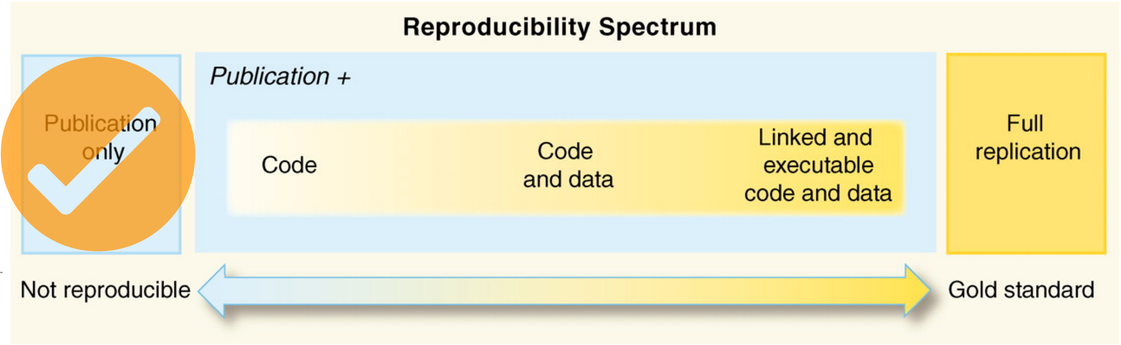
19 2018MNRAS.480.3907V 2018/11
**Probing the origin of extragalactic magnetic fields with Fast Radio Bursts**
Vazza, F.; Brüggen, M.; Hinz, P. M. *and 3 more*

20 2018ApJ...867L..10M 2018/11
**A Search for the Host Galaxy of FRB 171020**
Mahony, Elizabeth K.; Ekers, Ron D.; Macquart, Jean-Pierre *and 11 more*

21 2018PhRvD..98j3518M 2018/11
**Finding the missing baryons with fast radio bursts and Sunyaev-Zeldovich maps**
Muñoz, Julian B.; Loeb, Abraham

https://ui.adsabs.harvard.edu/abs/2020A%26A...638A..37W/citations

# Case Study 2:
# Open Access publication




THE ASTROPHYSICAL JOURNAL LETTERS

AAS | IOP

## Focus on the First Event Horizon Telescope Results

Shep Doeleman (Founding Director) on behalf of the EHT Collaboration

April 2019

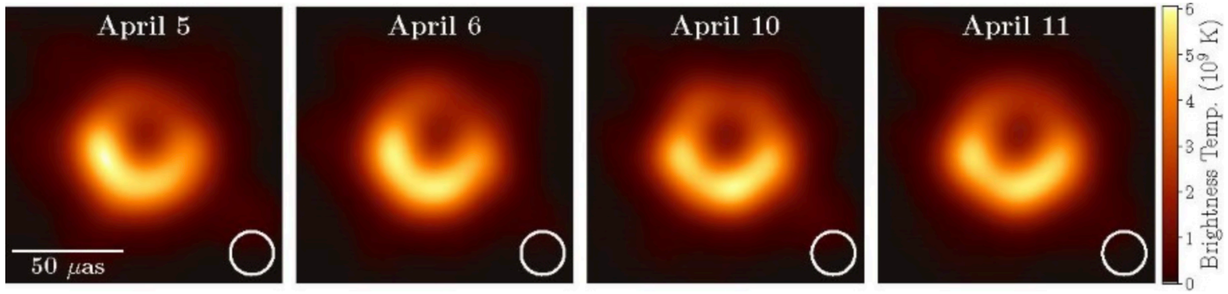



Figure 1. EHT images of M87 on four different observing nights. In each panel, the white circle shows the resolution of the EHT. All four images are dominated by a bright ring with enhanced emission in the south. From Paper IV (Figure 15).

We report the first image of a black hole.

First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole
The Event Horizon Telescope Collaboration *et al.* 2019 *ApJL* **875** L1

First M87 Event Horizon Telescope Results. II. Array and Instrumentation
The Event Horizon Telescope Collaboration *et al.* 2019 *ApJL* **875** L2

First M87 Event Horizon Telescope Results. III. Data Processing and Calibration
The Event Horizon Telescope Collaboration *et al.* 2019 *ApJL* **875** L3

First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole
The Event Horizon Telescope Collaboration *et al.* 2019 *ApJL* **875** L4

First M87 Event Horizon Telescope Results. V. Physical Origin of the Asymmetric Ring
The Event Horizon Telescope Collaboration *et al.* 2019 *ApJL* **875** L5

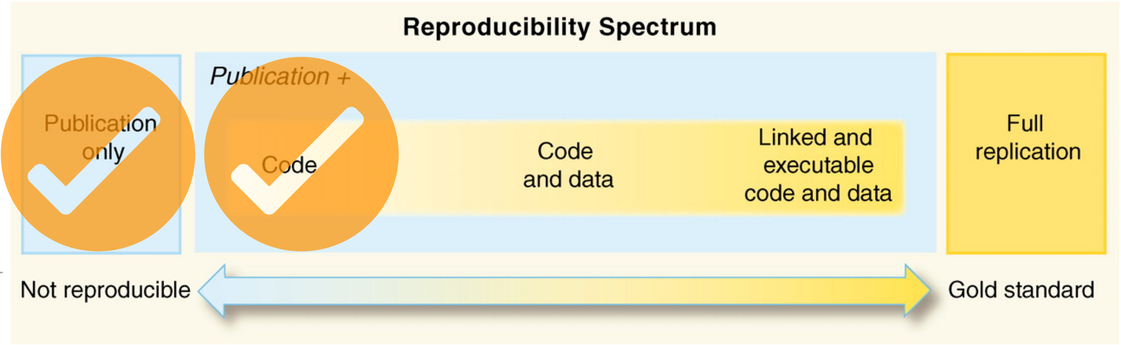
First M87 Event Horizon Telescope Results. VI. The Shadow and Mass of the Central Black Hole
The Event Horizon Telescope Collaboration *et al.* 2019 *ApJL* **875** L6

First M87 Event Horizon Telescope Results. VII. Polarization of the Ring
The Event Horizon Telescope Collaboration *et al.* 2021 *ApJL* **910** L12

First M87 Event Horizon Telescope Results. VIII. Magnetic Field Structure near The Event Horizon
The Event Horizon Telescope Collaboration *et al.* 2021 *ApJL* **910** L13

Event Horizon Telescope (EHT) Collaboration, https://iopscience.iop.org/journal/2041-8205/page/Focus_on_EHT

# Case Study 2: Open Source code




Event Horizon Telescope

Home | About EHT | Science | Technology | Press & Media Resources | For Astronomers

## EHT Data Products

| EHT Data Product Code | Title | Author(s) | Last Updated | DOI | Reference |
|---|---|---|---|---|---|
| 2021-D02-01 | Multi-wavelength Observations of M87 During the 2017 EHT Campaign | The EHT MWL Science Working Group et al. | 14 Apr 2021 | http://doi.org/10.25739/mhh2-cw46 | The EHT MWL Science Working Group et al. 2021, ApJL, 911, L11 |
| 2020-D03-01 | EHT Visibility Amplitude Data of M87* in 2011-2013 | The EHT Collaboration et al. | 23 Sep 2020 | https://doi.org/10.25739/daft-jz11 | Wielgus et al. 2020, ApJ, 901, 67 |
| 2020-D01-01 | First 3C279 EHT Results: Calibrated Data | The EHT Collaboration et al. | 12 Aug 2020 | https://doi.org/10.25739/vty0-ve39 | Kim et al. 2020, A&A, 640, A69 |
| 2019-D01-02 | First M87 EHT Results: Imaging Pipelines | The EHT Collaboration et al. | 24 Apr 2019 | https://github.com/eventhorizontelescope/2019-D01-02 | EHT Collaboration et al. (2019d), M87 Paper IV |
| 2019-D01-01 | First M87 EHT Results: Calibrated Data | The EHT Collaboration et al. | 10 Apr 2019 | https://doi.org/10.25739/g85n-f134 | EHT Collaboration et al. (2019c), M87 Paper III |

Search or jump to... / Pull requests Issues Marketplace Explore

eventhorizontelescope / 2019-D01-02

Watch 4 Star 23 Fork 28

Code Issues 2 Pull requests Actions Projects Wiki Security Insights

master 1 branch 0 tags Go to file Add file Code

rndsrc Improvement: use 'Agg' backend so the script works without X11 80d230e on 20 Feb 2020 3 commits

| | | |
|---|---|---|
| difmap | 2019-D01-02: First M87 EHT Results: Imaging Pipelines | 2 years ago |
| eht-imaging | Improvement: use 'Agg' backend so the script works without X11 | 16 months ago |
| smili | 2019-D01-02: First M87 EHT Results: Imaging Pipelines | 2 years ago |
| README.md | 2019-D01-02: First M87 EHT Results: Imaging Pipelines | 2 years ago |
| run.sh | 2019-D01-02: First M87 EHT Results: Imaging Pipelines | 2 years ago |

README.md

### First M87 EHT Results: Imaging Pipelines

**Authors:** The Event Horizon Telescope Collaboration et al.

**Date:** April 10, 2019

**Primary Reference:** The Event Horizon Telescope Collaboration, et al. 2019d, ApJL, 875, L4 (M87 Paper IV)

**Data Product Code:** 2019-D01-02

**Brief Description:**

We release three imaging pipelines (DIFMAP, eht-imaging and SMILI) used in the parameter survey of M87 Paper IV Section 6 and later. All imaging pipelines create images from calibrated uvfits files (see M87 Paper III) simultaneously released (data product code: 2019-D01-01). For more detailed instructions, please see the

About

First M87 EHT Results: Imaging Pipelines

eventhorizontelescope.org/for-astr...

Readme

Releases

No releases published

Packages

No packages published

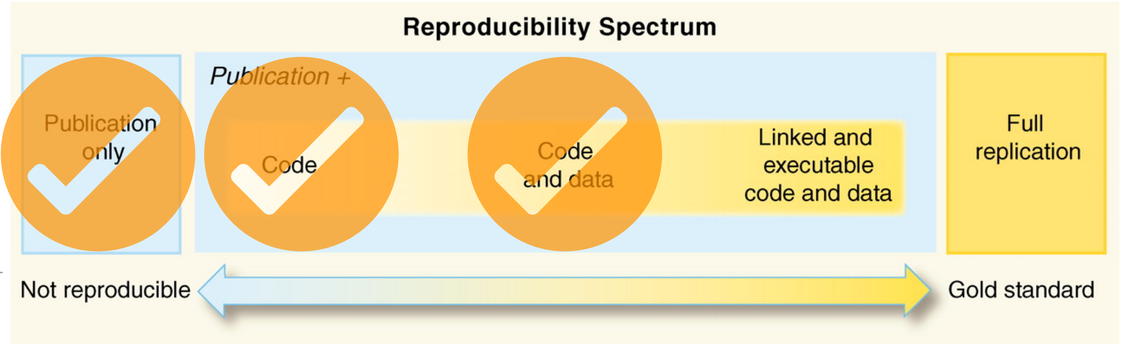
Languages

Python 89.6% Shell 10.4%

https://eventhorizontelescope.org/for-astronomers/data

https://github.com/eventhorizontelescope/2019-D01-02

# Case Study 2: Open Data




Event Horizon Telescope

Home | About EHT | Science | Technology | Press & Media Resources | For Astronomers

## EHT Data Products

| EHT Data Product Code | Title | Author(s) | Last Updated | DOI | Reference |
|---|---|---|---|---|---|
| 2021-D02-01 | Multi-wavelength Observations of M87 During the 2017 EHT Campaign | The EHT MWL Science Working Group et al. | 14 Apr 2021 | http://doi.org/10.25739/mhh2-cw46 | The EHT MWL Science Working Group et al. 2021, ApJL, 911, L11 |
| 2020-D03-01 | EHT Visibility Amplitude Data of M87* in 2011-2013 | The EHT Collaboration et al. | 23 Sep 2020 | https://doi.org/10.25739/daft-jz11 | Wielgus et al. 2020, ApJ, 901, 67 |
| 2020-D01-01 | First 3C279 EHT Results: Calibrated Data | The EHT Collaboration et al. | 12 Aug 2020 | https://doi.org/10.25739/vty0-ve39 | Kim et al. 2020, A&A, 640, A69 |
| 2019-D01-02 | First M87 EHT Results: Imaging Pipelines | The EHT Collaboration et al. | 24 Apr 2019 | https://github.com/eventhorizontelescope/2019-D01-02 | EHT Collaboration et al. (2019d), M87 Paper IV |
| 2019-D01-01 | First M87 EHT Results: Calibrated Data | The EHT Collaboration et al. | 10 Apr 2019 | https://doi.org/10.25739/g85n-f134 | EHT Collaboration et al. (2019c), M87 Paper III |

Data Commons | About | Data Store | Discovery Environment | Atmosphere | Search | Submit

shared / commons_repo / curated / EHTC_FirstM87Results_Apr2019

### First M87 EHT Results: Calibrated Data

**DOI:** 10.25739/g85n-f134

**Creator:** The Event Horizon Telescope Collaboration

**Description:** We release a data set to accompany the First M87 Event Horizon Telescope Results paper series (EHT Collaboration et al. 2019a,b,c,d,e,f, see the README.md file for full references). The data set is derived from the Science Release 1 (SR1) of the Event Horizon Telescope (EHT)'s April 2017 observation campaign (EHT Collaboration et al. 2019c). This data set contains M87 data for both low and high bands for all observed days (April 5th, 6th, 10th, 11th, 2017). Data from the 2017 observations were processed through three independent reduction pipelines (Blackburn et al. 2019, Janssen et al. 2019, Paper III). This release includes the fringe fitted, a-priori calibrated, and network calibrated data from the EHT-HOPS pipeline, which is the primary data set for the First M87 EHT results. Independent flux calibration is performed based on estimated station sensitivities during the campaign (Issaoun et al. 2017, Janssen et al. 2019). A description of the data properties, their validation, and estimated systematic errors is given in Paper III with additional details in Wielgus et al. (2019). The data are time averaged to 10 seconds and frequency averaged over all 32 intermediate frequencies (IFs). All polarization information is explicitly removed. To make the resulting uvfits files compatible with popular very-long-baseline interferometry (VLBI) software packages, the circularly polarized cross-hand visibilities RL and LR are set to zero along with their errors, while parallel-hands RR and LL are both set to an estimated Stokes I value. Measurement errors for RR and LL are each set to sqrt(2) times the statistical errors for Stokes I.

**Publisher:** CyVerse Data Commons

**Publication Year:** 2019

**Rights:** data: ODC PDDL, code: GPLv3

**Citation:** The Event Horizon Telescope Collaboration (2019). First M87 EHT Results: Calibrated Data. CyVerse Data Commons. DOI 10.25739/g85n-f134

Export Formats: BibTeX Endnote

show more | Download Metadata

| Name | Size | Created | Last Modified |
|---|---|---|---|
| csv | -- | Apr 9, 2019 12:08:12 AM | Feb 20, 2020 10:05:33 PM |
| txt | -- | Apr 9, 2019 12:07:56 AM | Feb 20, 2020 10:05:33 PM |
| uvfits | -- | Apr 9, 2019 12:08:22 AM | Feb 20, 2020 10:05:33 PM |
| EHTC_FirstM87Results_Apr2019_csv.tgz | 1.4 MB | Apr 9, 2019 12:08:30 AM | Apr 12, 2019 2:07:30 AM |
| EHTC_FirstM87Results_Apr2019_txt.tgz | 1.5 MB | Apr 9, 2019 12:08:09 AM | Apr 12, 2019 2:07:30 AM |
| EHTC_FirstM87Results_Apr2019_uvfits.tgz | 1.5 MB | Apr 9, 2019 12:08:22 AM | Apr 12, 2019 2:07:32 AM |
| INVENTORY.txt | 1.4 kB | Apr 24, 2019 1:38:35 PM | Apr 24, 2019 1:38:35 PM |
| LICENSE.txt | 34.5 kB | Apr 9, 2019 12:08:28 AM | Apr 12, 2019 2:07:18 AM |
| README.md | 4.8 kB | Apr 9, 2019 12:08:12 AM | Apr 12, 2019 2:07:19 AM |
| run.sh | 959.0 bytes | Apr 9, 2019 12:08:11 AM | Apr 12, 2019 2:07:18 AM |

Displaying 1-10 of 10 items

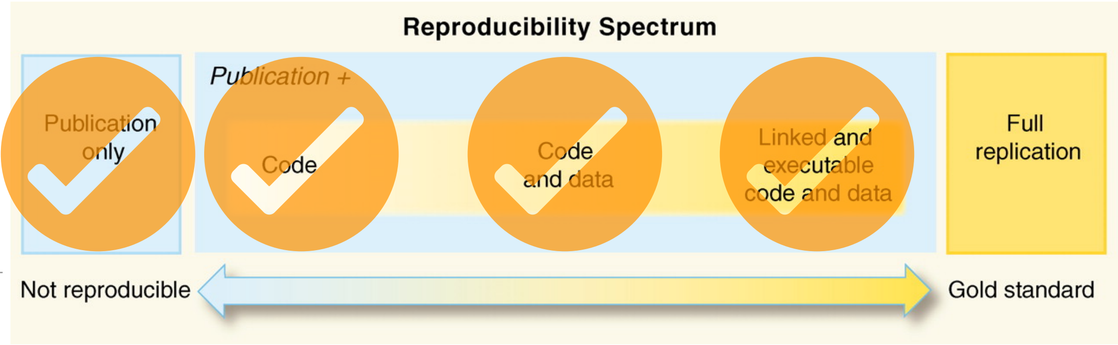
Data Commons Mirrors v2.7.0 © 2021 CyVerse

CyVerse Home | Discovery Environment Application | Atmosphere Application

https://eventhorizontelescope.org/for-astronomers/data

https://doi.org/10.25739/g85n-f134

# Case Study 2: Reproducible Computational Environment




Search or jump to... / Pull requests Issues Marketplace Explore

**The Event Horizon Telescope**

The Event Horizon Telescope is an international collaboration aiming to capture the first image of a black hole by creating a virtual Earth-sized telescope.

International Collaboration http://eventhorizontelescope.org

Repositories 9 Packages People 1 Projects

Pinned repositories

**2019-D01-01**
First M87 EHT Results: Calibrated Data
Python 11 4

**2019-D01-02**
First M87 EHT Results: Imaging Pipelines
Python 23 28

**2020-D01-01**
First 3C279 EHT Results: Calibrated Data
Python 1

**2020-D03-01**
EHT Visibility Amplitude Data of M87* in 2011-2013

**docker-recipes**
A Dockerfile library for creating Docker images for the EHT
Shell 1

**ehtplot**
Forked from liamedeiros/ehtplot
Plotting functions for the Event Horizon Telescope
Python

Find a repository... Type Language Sort

**kubernetes-examples**
A kubernetes configuration file library for setting up cloud computing tasks for the EHT
0 0 0 0 Updated 6 days ago

**eht-mayan-edms**

Top languages
Python Shell

People 1

https://github.com/eventhorizontelescope

Search or jump to... / Pull requests Issues Marketplace Explore

eventhorizontelescope / **docker-recipes**

Watch 8 Star 0 Fork 1

Code Issues 1 Pull requests Actions Projects Wiki Security Insights

master 2 branches 4 tags Go to file Add file Code

rndsrc Updated "theory-env/Dockerfile" for `theory-env:latest` ... ce5d6b3 on 19 Feb 260 commits

| | | |
|---|---|---|
| .github/workflows | Include a GitHub Action for selftag | 14 months ago |
| aips-stack | Merge commit '498a7fcd': pull the latest Dockerfiles | 2 years ago |
| bin | Pull latest scripts from "l6a/docker-recipes" | 14 months ago |
| eat | Correct README.md files | 3 months ago |
| ehtpython | Correct README.md files | 3 months ago |
| img-env | Use the AstroContainers version of AIPS container | 2 years ago |
| theory-env | Updated "theory-env/Dockerfile" for `theory-env:latest` | 3 months ago |
| .gitignore | Add a ".gitignore" file | 14 months ago |
| LICENSE | Use the Apache License | 2 years ago |
| README.md | Initial commit | 2 years ago |

README.md

## Docker Image Recipes for the Event Horizon Telescope

This repository contains recipes for building Docker images for the Event Horizon Telescope (EHT).

**About**
A Dockerfile library for creating Docker images for the EHT
Readme
Apache-2.0 License

**Releases**
4 tags

**Packages**
No packages published

**Languages**
Shell 57.9% Dockerfile 42.1%

https://github.com/eventhorizontelescope/docker-recipes

# Case Study 2: Impact

## Paper I

- 1 March 2019:
  - Submitted manuscript to journal
- 12 March 2019: Accepted by journal
- 10 April 2019: Officially published by the journal
  - 1001 citations to date
- 26 June 2019:
  - Deposited post-print to arXiv


https://ui.adsabs.harvard.edu/abs/2019ApJ...875L...1E/citations

# Case Study 2: Impact

Paper I

- 1 March 2019:
  - Submitted manuscript to journal
- 12 March 2019: Accepted by journal
- 10 April 2019: Officially published by the journal
  - 1001 citations to date
- 26 June 2019:
  - Deposited post-print to arXiv


https://eventhorizontelescope.org/

# More information, resources and takeaways

# *The Turing Way*

- Project led by Kirstie Whitaker at The Alan Turing Institute to make reproducible research "too easy not to do"
- In short: *The Turing Way* encompasses a handbook, community, collaboration, workshops and training
- Team of researchers, research software engineers, librarians and YOU!
- Demonstrates open and transparent project management and communication with future users, as it is openly developed at our GitHub repository: https://github.com/alan-turing-institute/the-turing-way


This image was created by Scriberia for *The Turing Way* community and is used under a CC-BY licence. https://doi.org/10.5281/zenodo.3332807

# More information and resources

- See talks on Reproducible Science in practice by Arturo Sánchez Pineda
- Introduction to using GitHub: https://github.com/rainsworth/intro-to-github
- Reproducible Research through Containerisation: Docker and Singularity: https://github.com/rainsworth/osip2019-containerisation-workshop
- Software and Data Carpentry workshops: https://carpentries.org/

# Takeaways

- “Reproducibility is like brushing your teeth. It is good for you, but it takes time and effort. Once you learn it, it becomes a habit.”
  - Irakli Loladze (Baker 2016 https://doi.org/10.1038/533452a)

- Start small! Test out one platform or open up one stage of your research workflow, such as sharing data via Zenodo and linking to the DOI in your publications.

- Help us build a kinder and more inclusive research culture!