

Reproducibility and Open Science

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

 @rachaelevelyn

 @rainsworth

 [10.6084/m9.figshare.9255638](https://doi.org/10.6084/m9.figshare.9255638)

OCTOBER
21-27

🔓 OPEN ACCESS WEEK 2019

¿ABIERTO PARA QUIÉN?

EQUIDAD EN EL CONOCIMIENTO ABIERTO

मुक्त प्रवेश किसके लिए खुला?

खुले ज्ञान में समानता

OPEN FOR WHOM?

EQUITY IN OPEN KNOWLEDGE

حر لعن؟ الإنصاف في المعرفة المفتوحة

OUVERT POUR QUI?

ÉQUITÉ DANS LE SAVOIR OUVERT

<http://www.openaccessweek.org>



Outline

- About me and my Open Science journey
- Reproducibility and research culture
- Open Science/Research/Scholarship
- Barriers to open research
- Why research openly?
- How to open up your research workflow
- Open Science in Astronomy & a case study
- Takeaways



About me and my Open Science journey

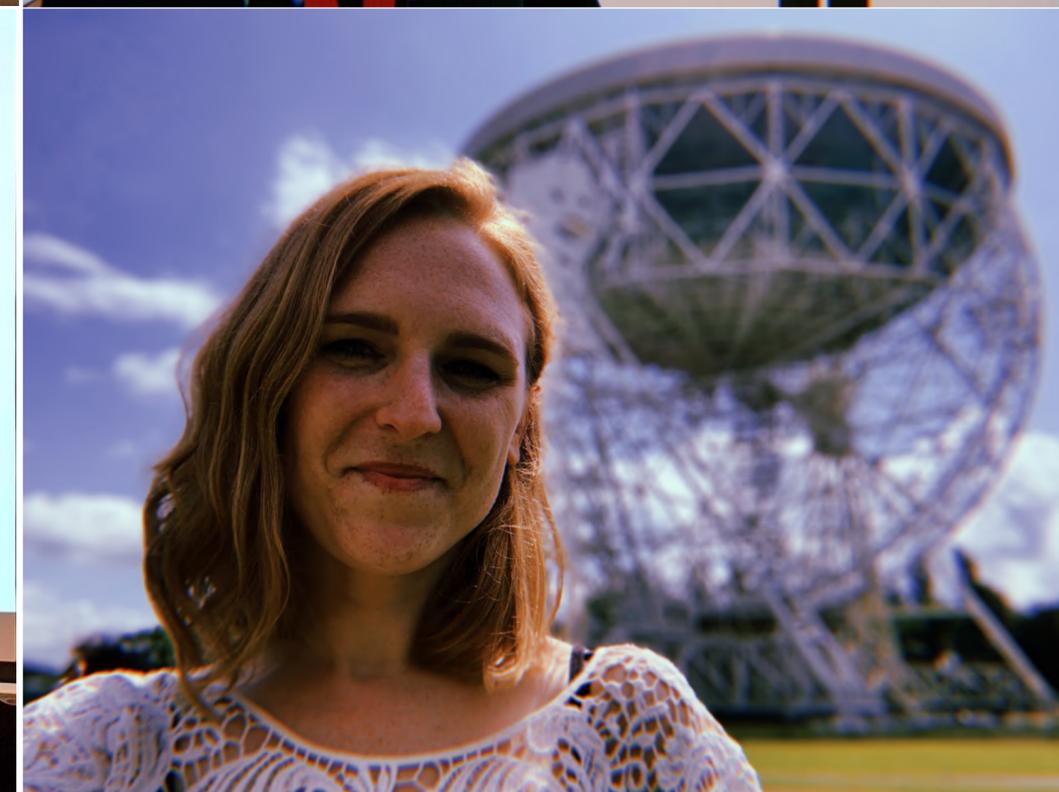
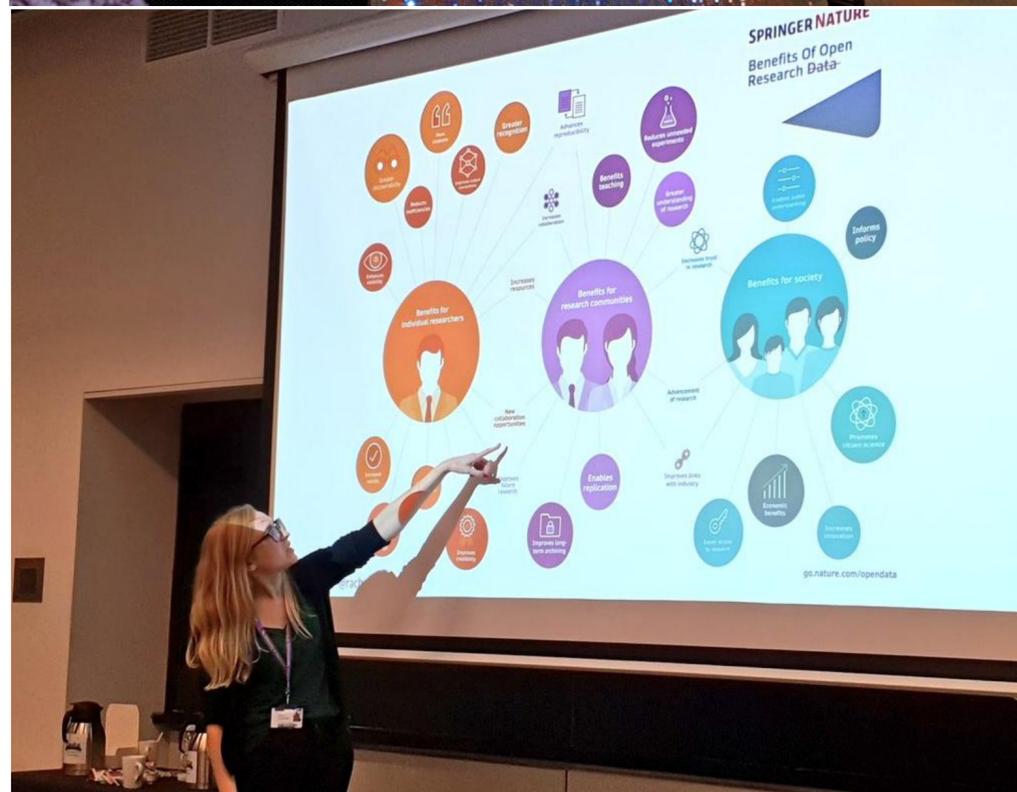


About me

- Community Manager for the Software Sustainability Institute at the University of Manchester
- Research background in Astrophysics
- Passionate about openness, transparency, reproducibility, wellbeing and inclusion in STEM
- Currently a cartoon in the UK's National Science and Media Museum Hello Universe exhibition
- Organise the Manchester women in data meetup group

HER+**Data** MCR

meetup.com/HER-Data-MCR



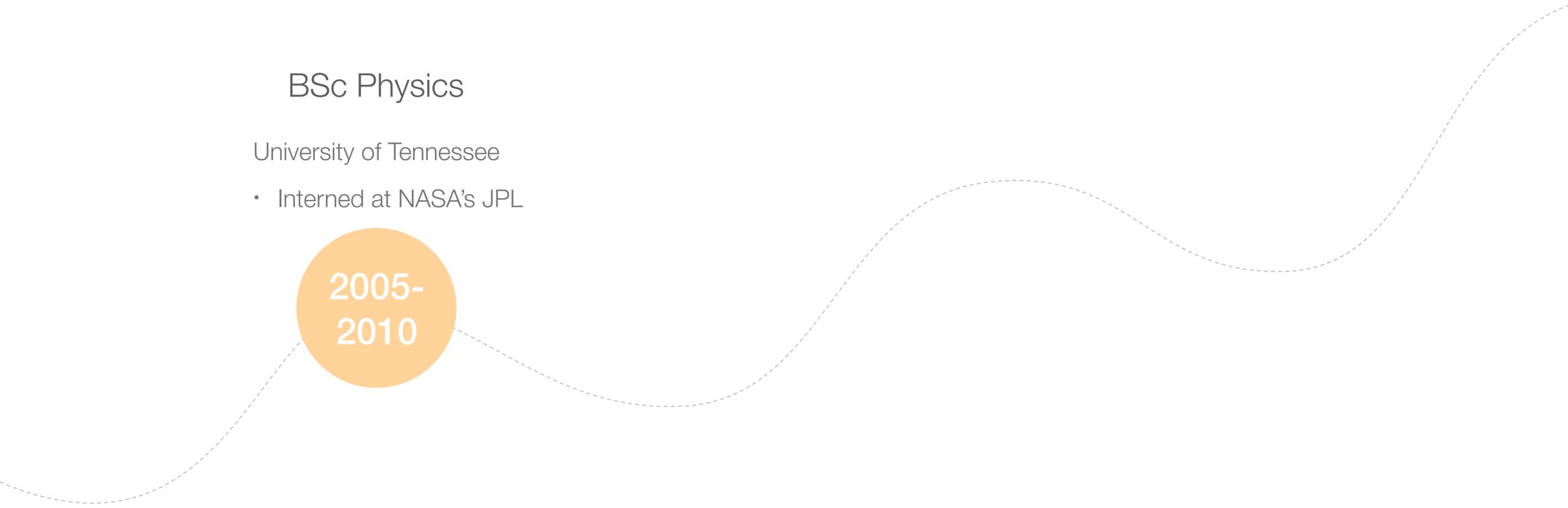
My career path and Open Science journey

BSc Physics

University of Tennessee

- Interned at NASA's JPL

2005-
2010



My career path and Open Science journey

BSc Physics

University of Tennessee

- Interned at NASA's JPL

2005-
2010

2010-
2014

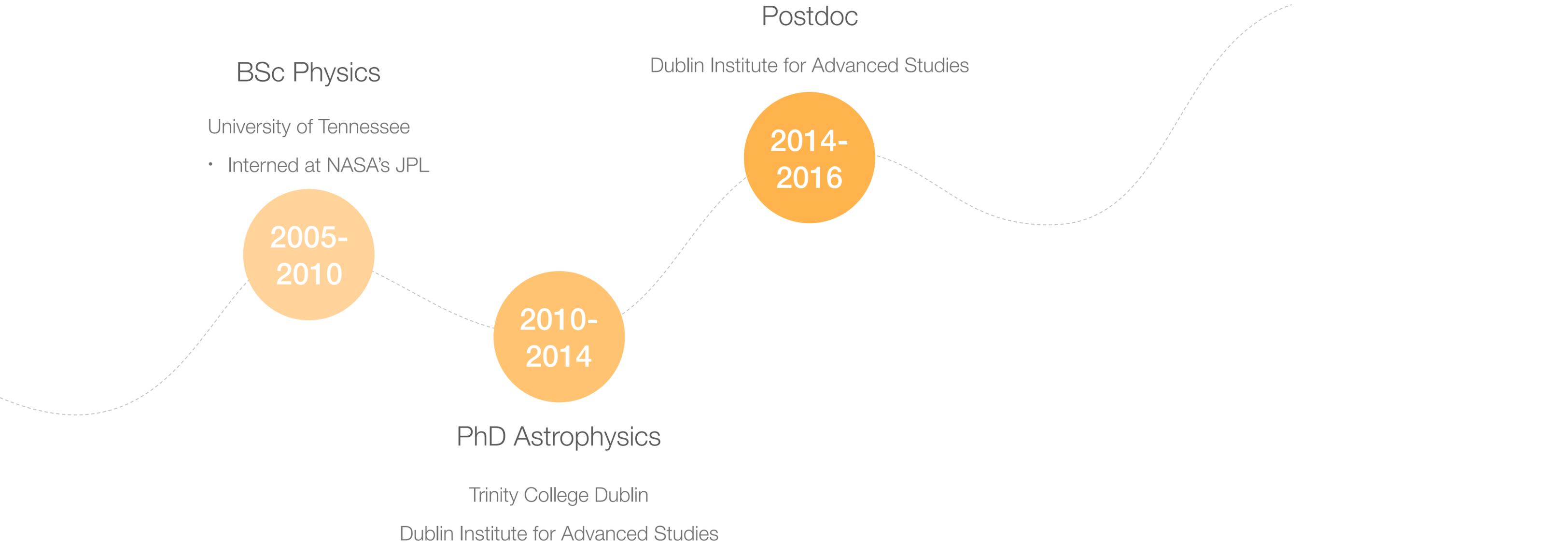
PhD Astrophysics

Trinity College Dublin

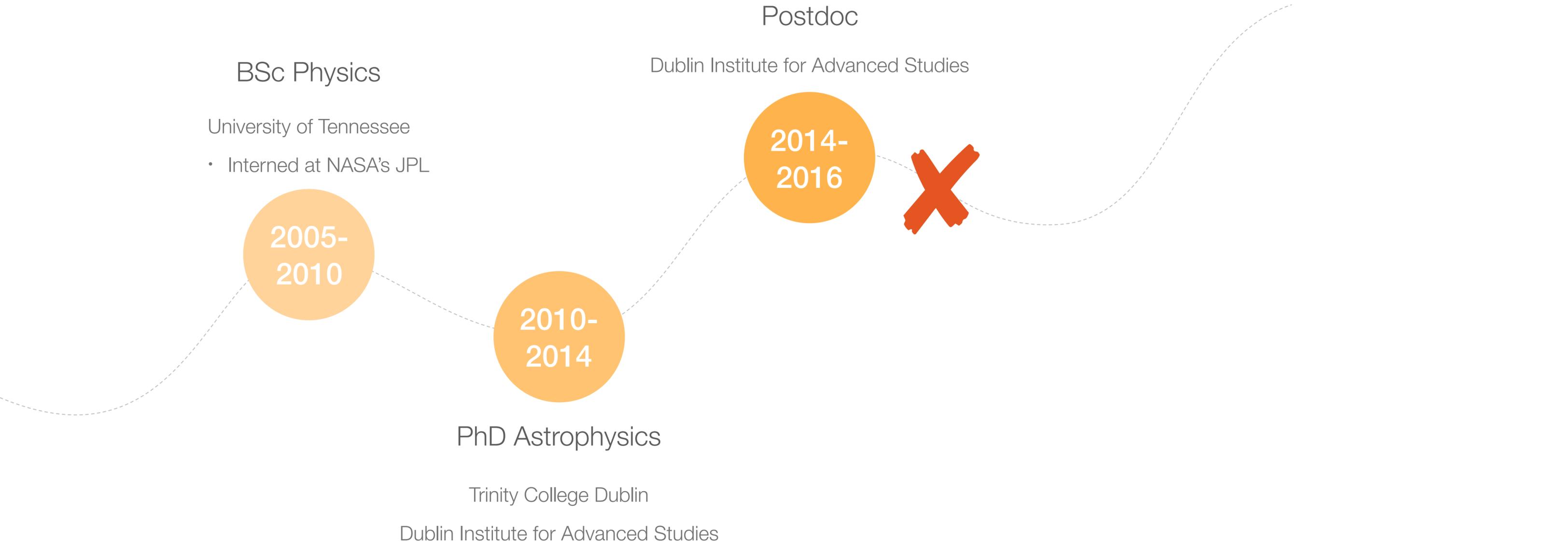
Dublin Institute for Advanced Studies



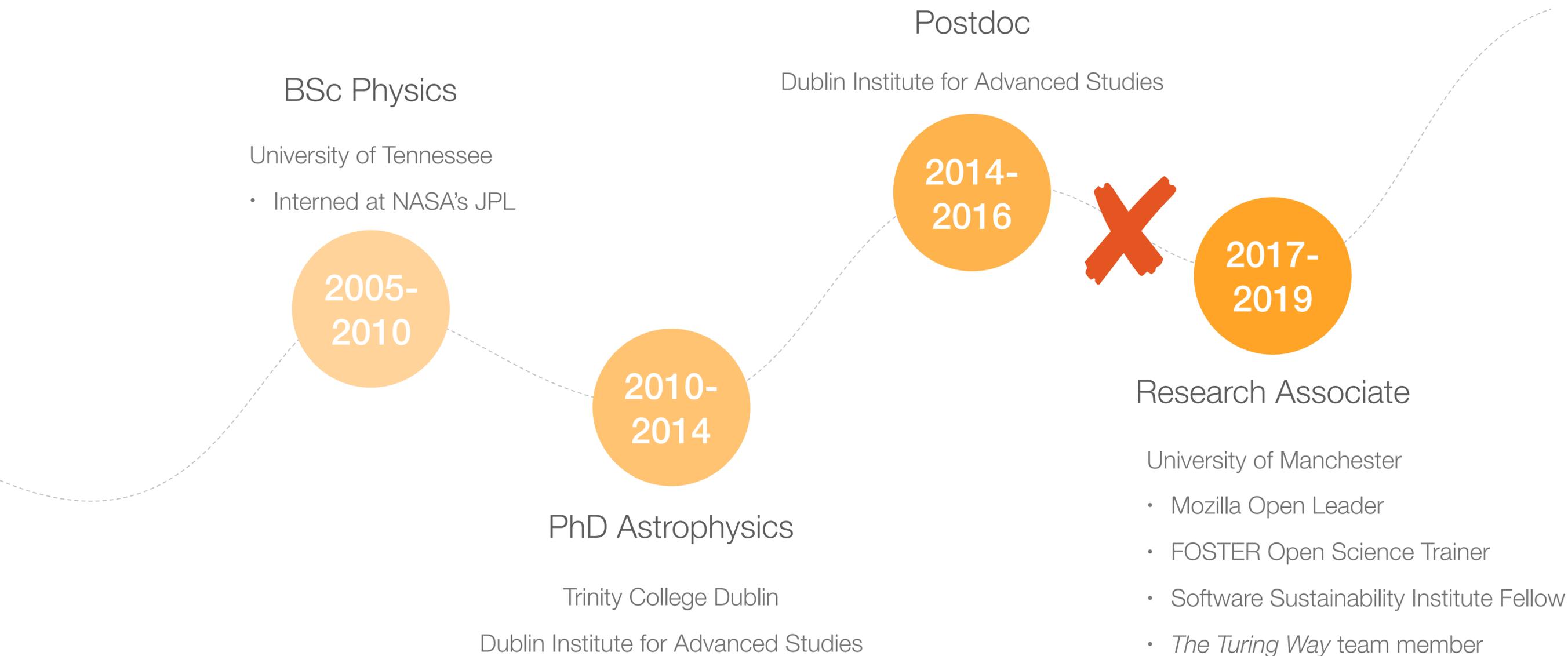
My career path and Open Science journey



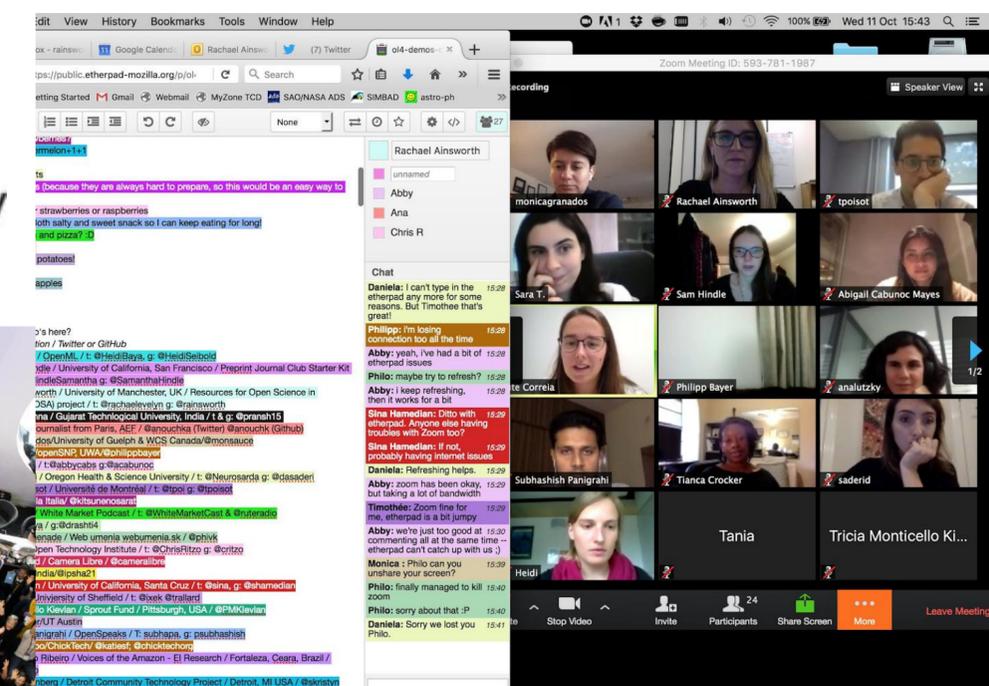
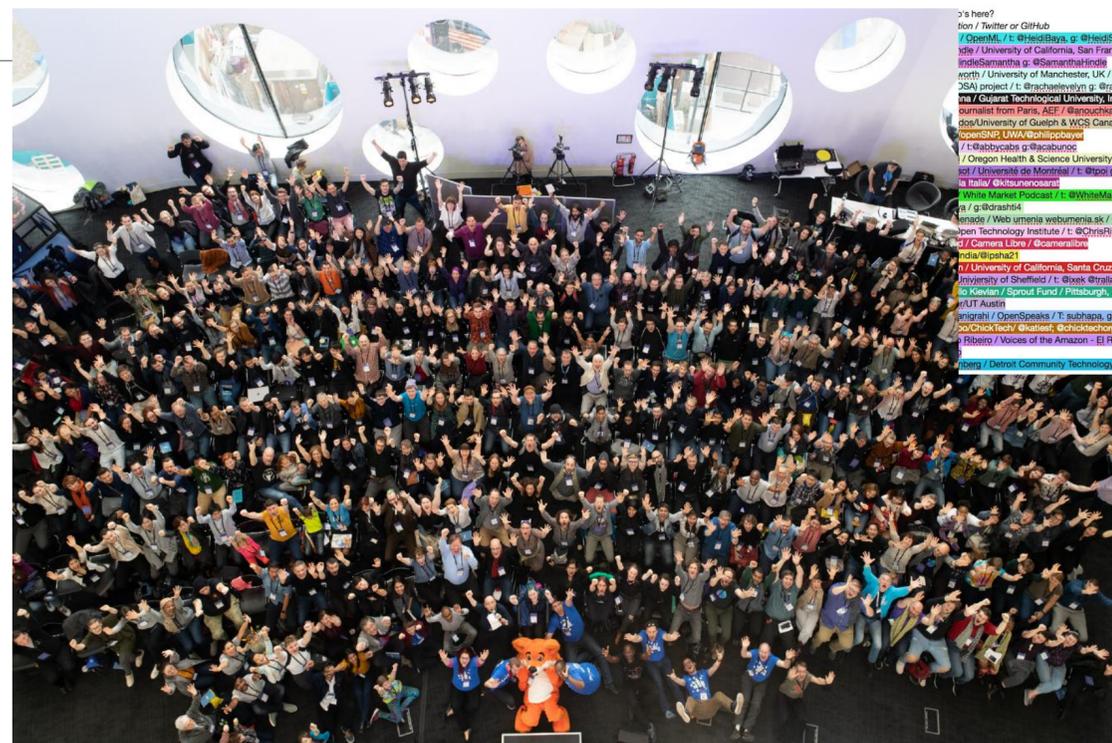
My career path and Open Science journey



My career path and Open Science journey



Open Science Advocacy



- Mozilla Open Leaders
 - Round 4 Project Lead
 - Round 5 Mentor, Cohort Host
- Mozilla Festival 2017, 2018
- OpenCon 2017
- FOSTER Open Science Trainer Bootcamp
- Mozilla Global Sprint 2018
- Speaker at Open Science events
- Software Sustainability Institute Fellow 2019
- *The Turing Way* team member
- Open Science Fair 2017, 2019





Research Culture is Broken; Open Science can [help] Fix It

<https://youtu.be/c-bemNZ-lqA>



My career path and Open Science journey



Reproducibility and research culture



| | | Data | |
|----------|-----------|--------------|---------------|
| | | Same | Different |
| Analysis | Same | Reproducible | Replicable |
| | Different | Robust | Generalisable |

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



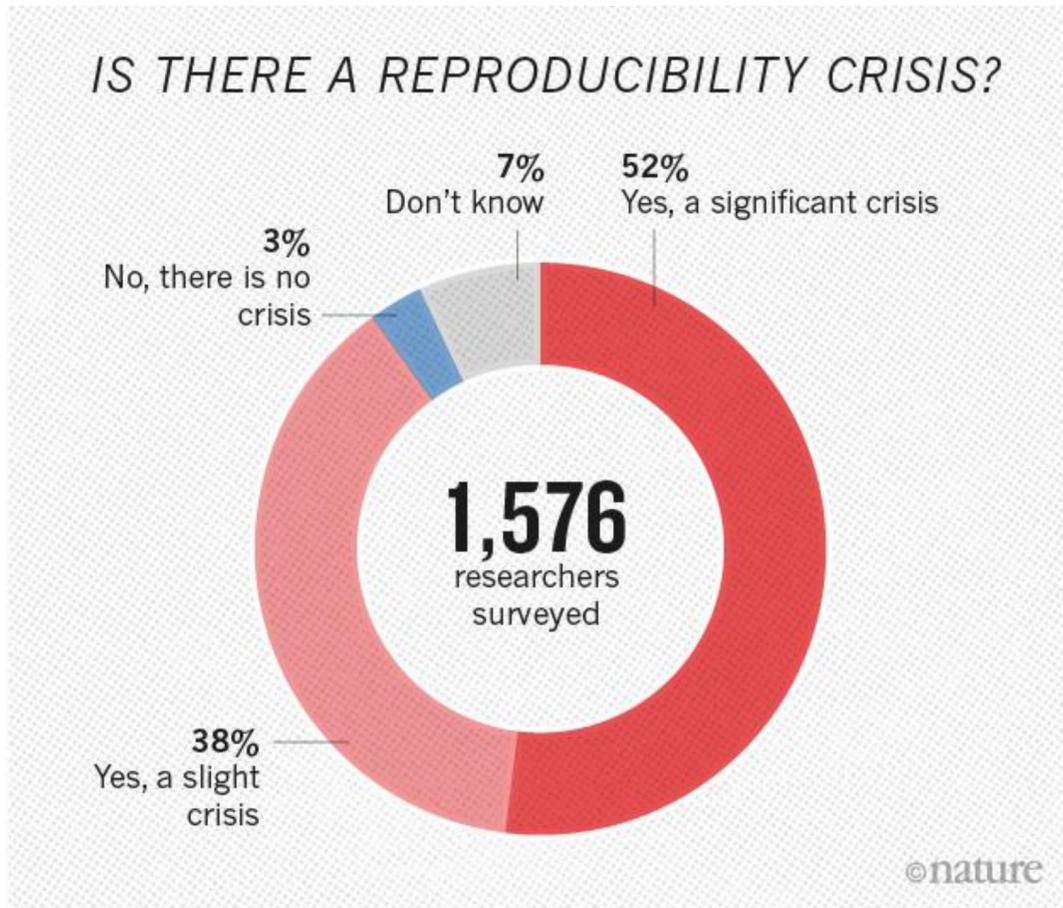


1,500 scientists lift the lid on reproducibility

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>

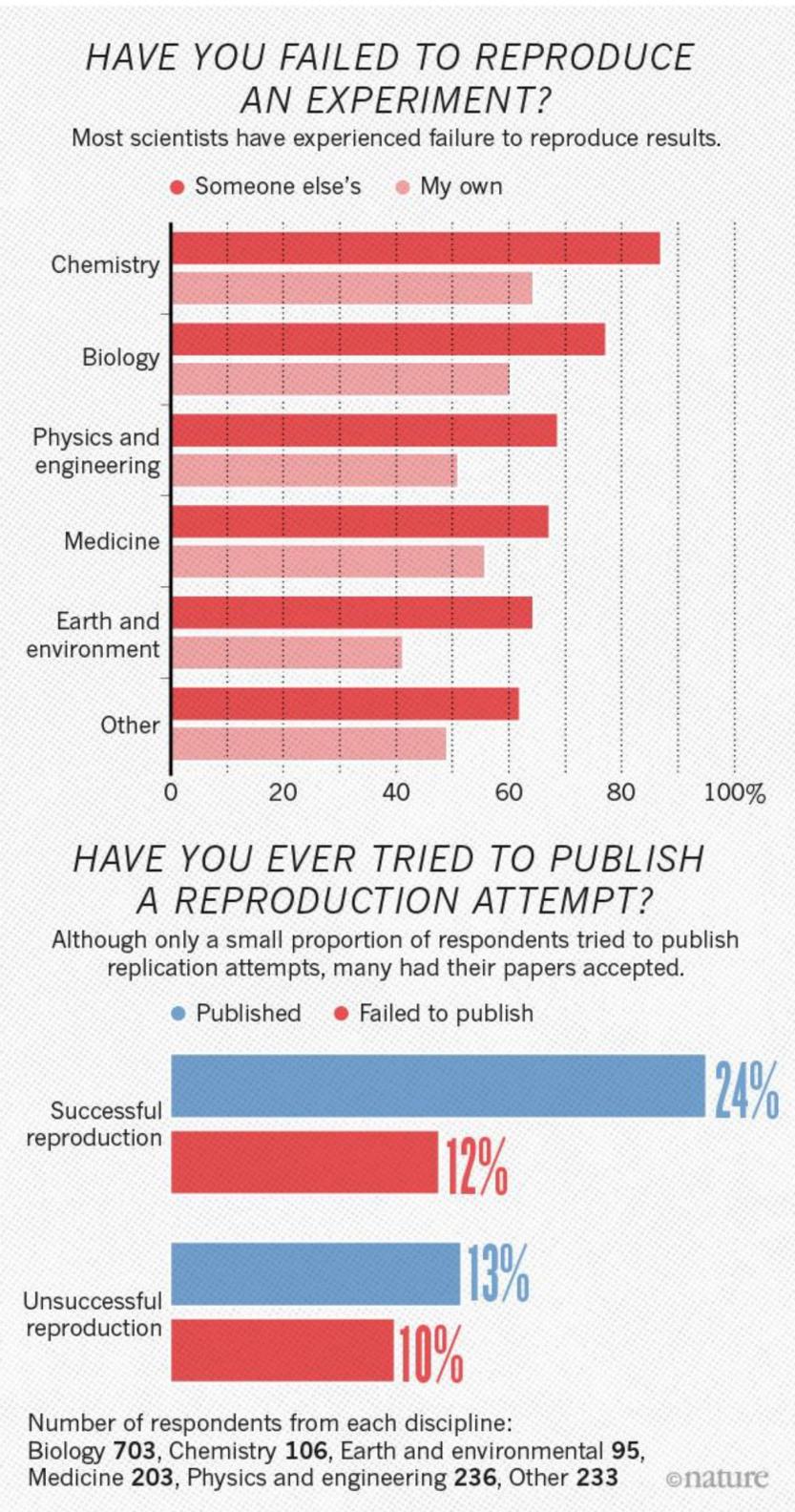
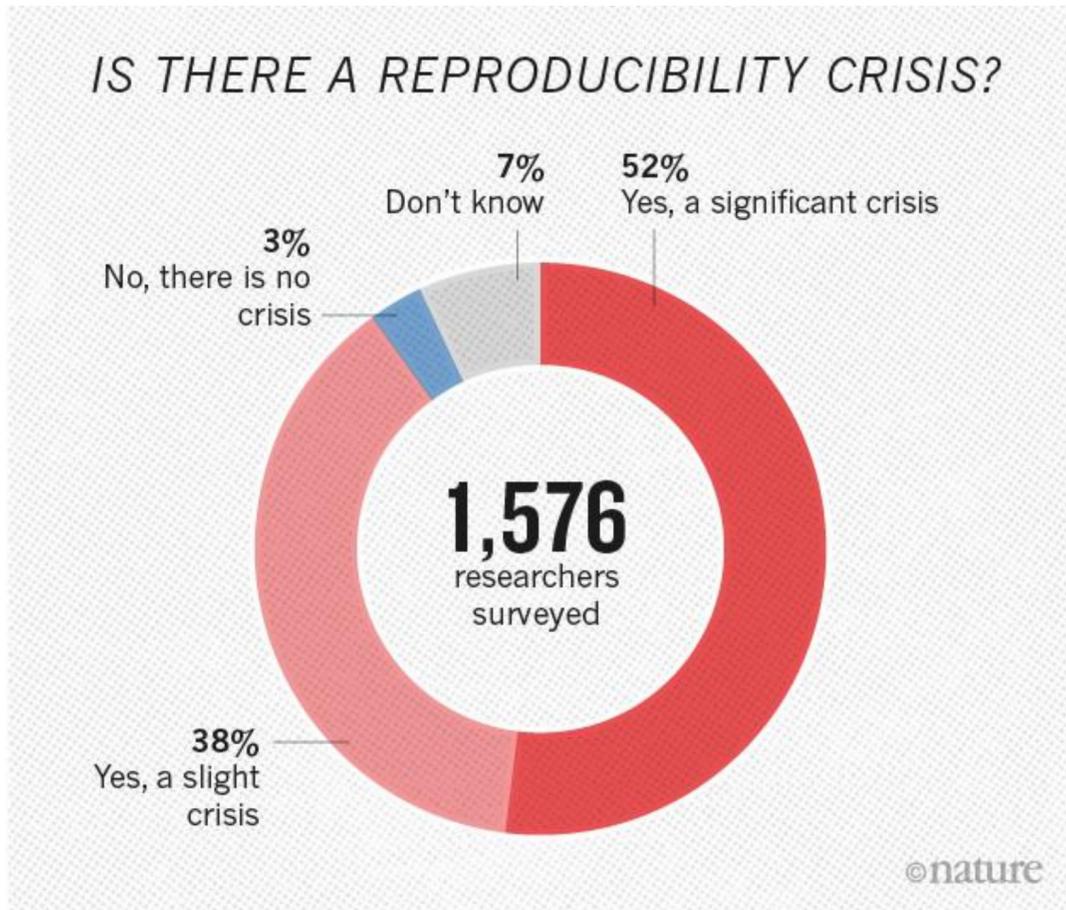


1,500 scientists lift the lid on reproducibility

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>



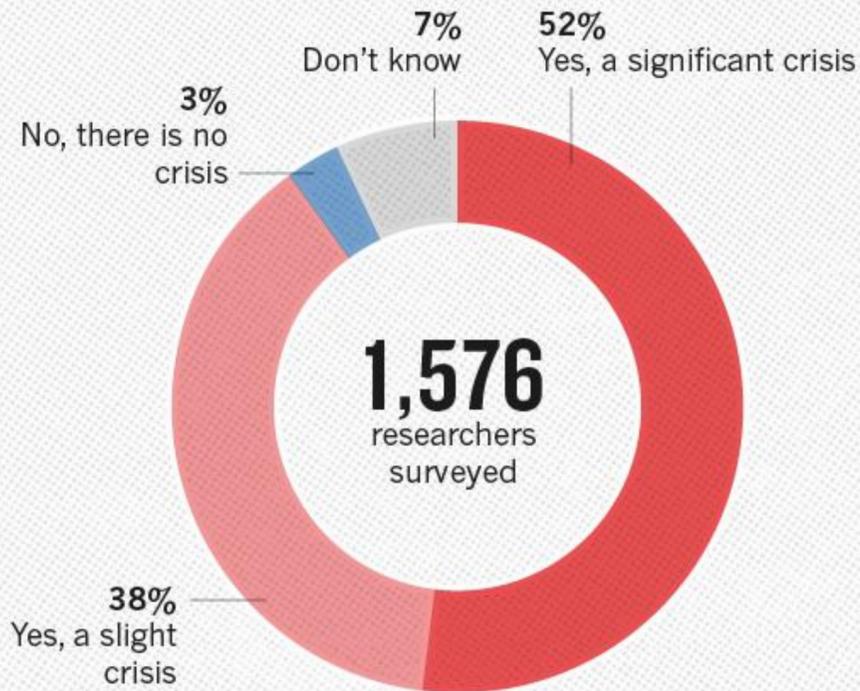
1,500 scientists lift the lid on reproducibility

Survey sheds light on the 'crisis' rocking research.

Monya Baker

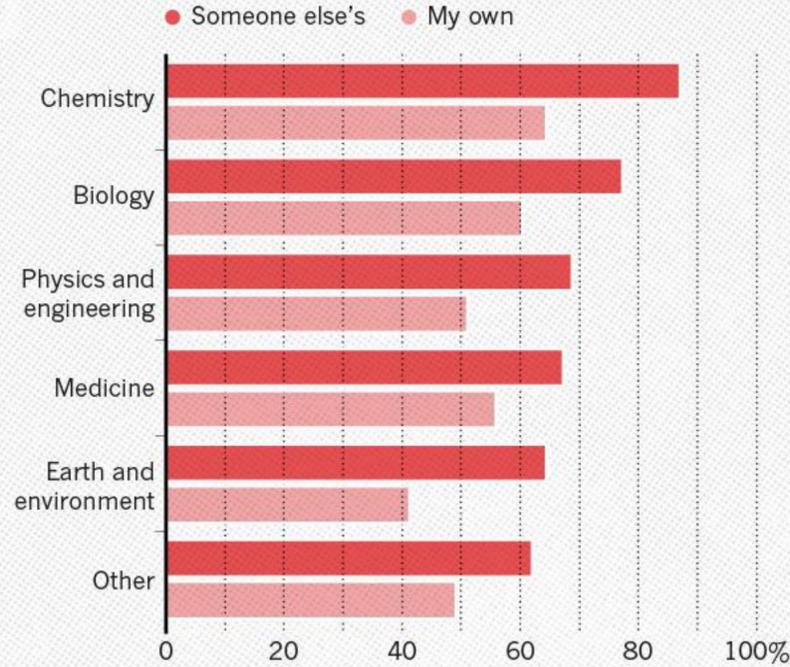
25 May 2016 | Corrected: 28 July 2016

IS THERE A REPRODUCIBILITY CRISIS?



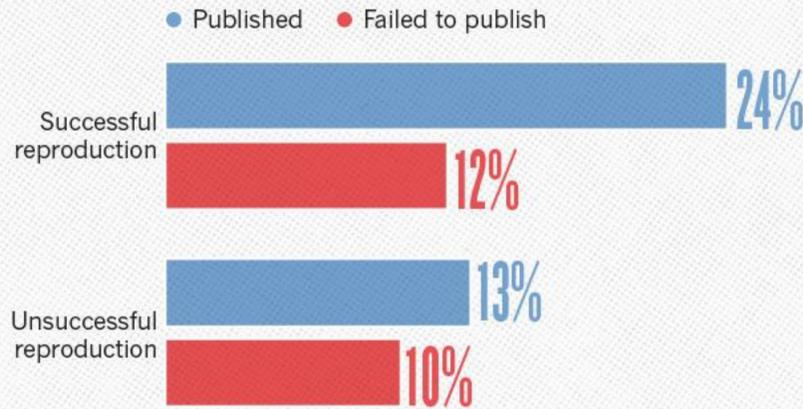
HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



HAVE YOU EVER TRIED TO PUBLISH A REPRODUCTION ATTEMPT?

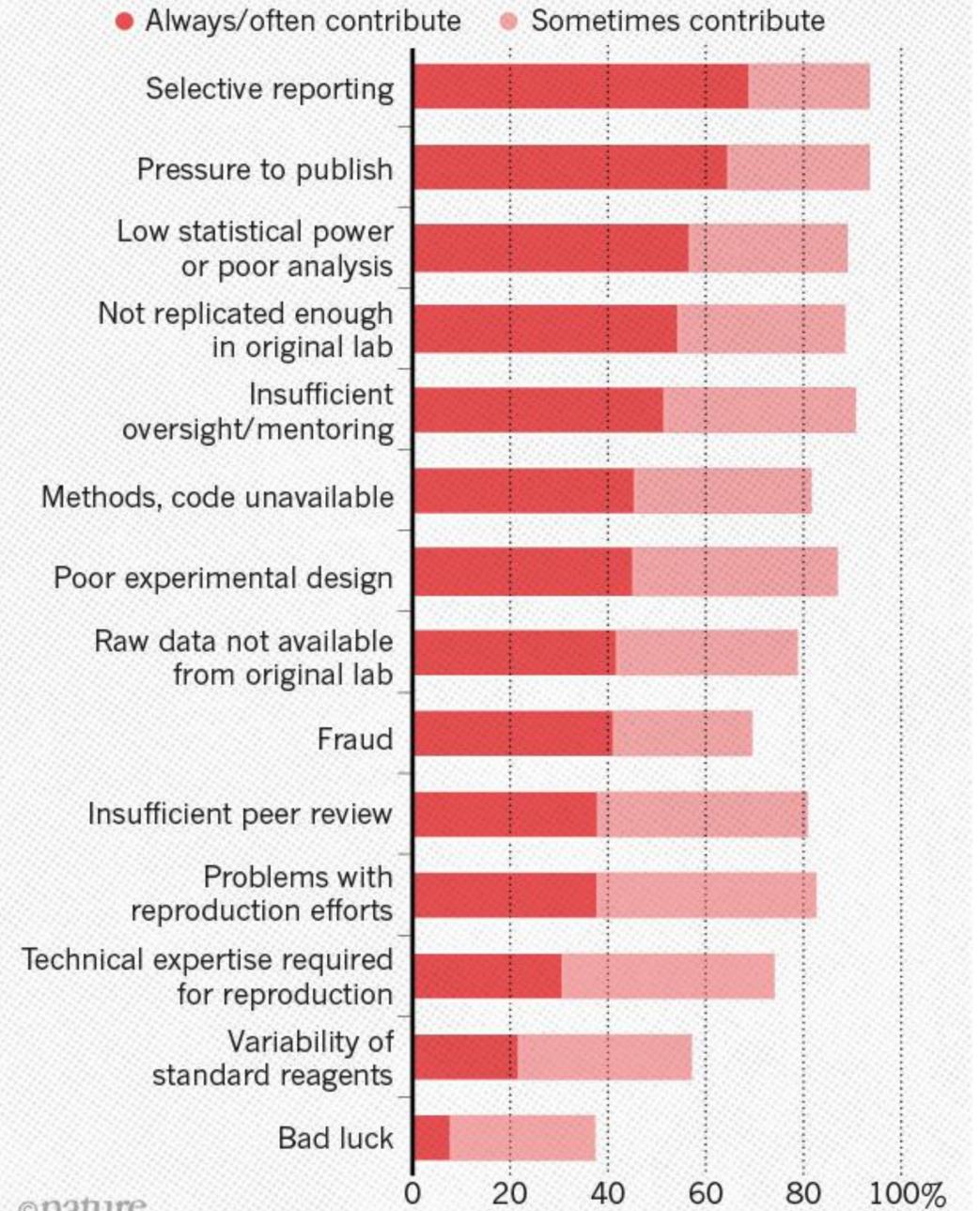
Although only a small proportion of respondents tried to publish replication attempts, many had their papers accepted.



Number of respondents from each discipline:
 Biology 703, Chemistry 106, Earth and environmental 95,
 Medicine 203, Physics and engineering 236, Other 233

WHAT FACTORS CONTRIBUTE TO IRREPRODUCIBLE RESEARCH?

Many top-rated factors relate to intense competition and time pressure.



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



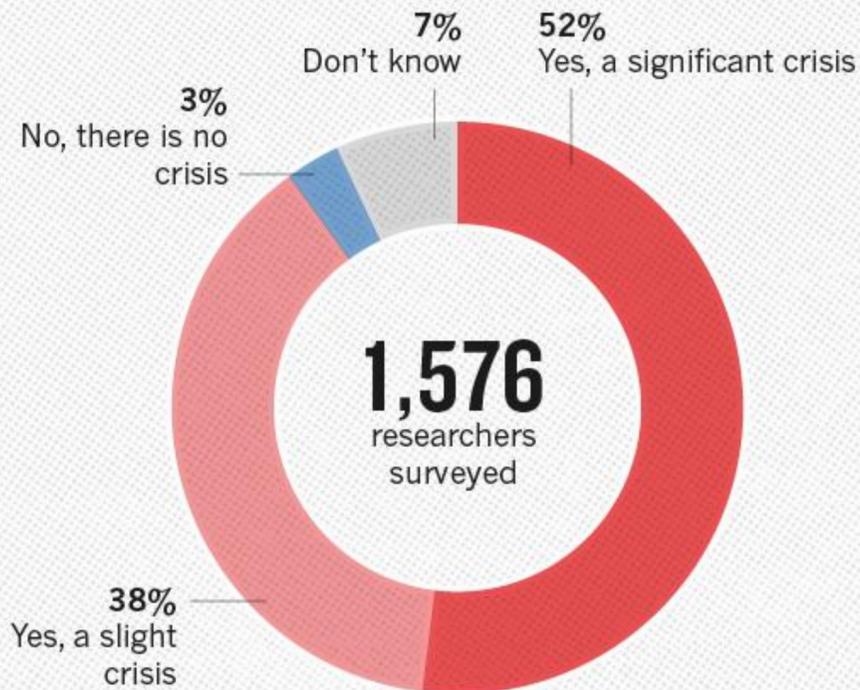
1,500 scientists lift the lid on reproducibility

Survey sheds light on the 'crisis' rocking research.

Monya Baker

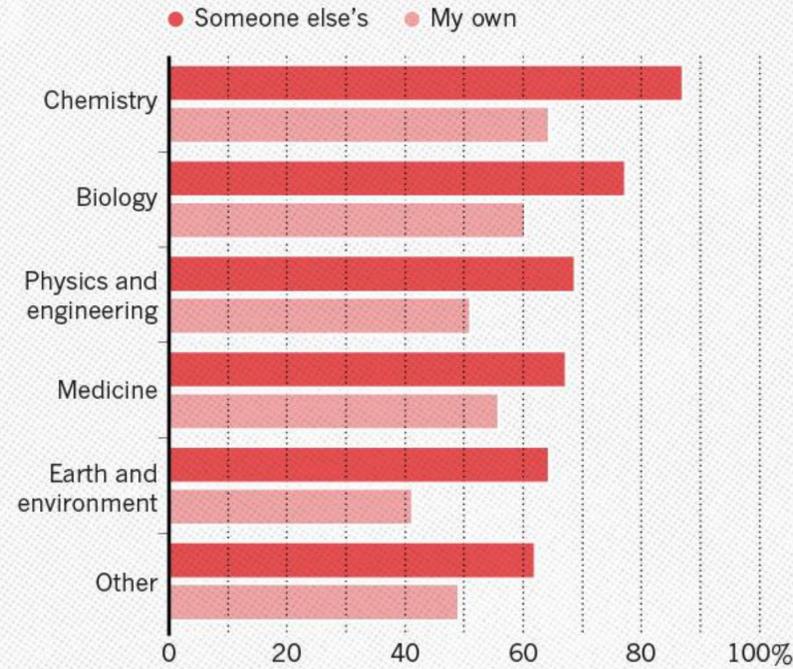
25 May 2016 | Corrected: 28 July 2016

IS THERE A REPRODUCIBILITY CRISIS?



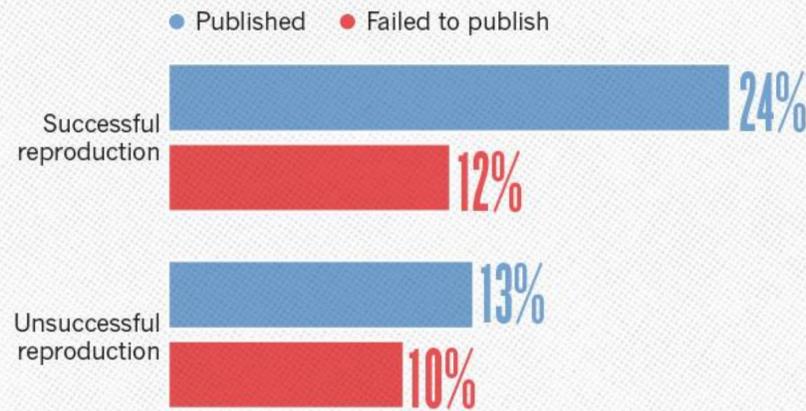
HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



HAVE YOU EVER TRIED TO PUBLISH A REPRODUCTION ATTEMPT?

Although only a small proportion of respondents tried to publish replication attempts, many had their papers accepted.

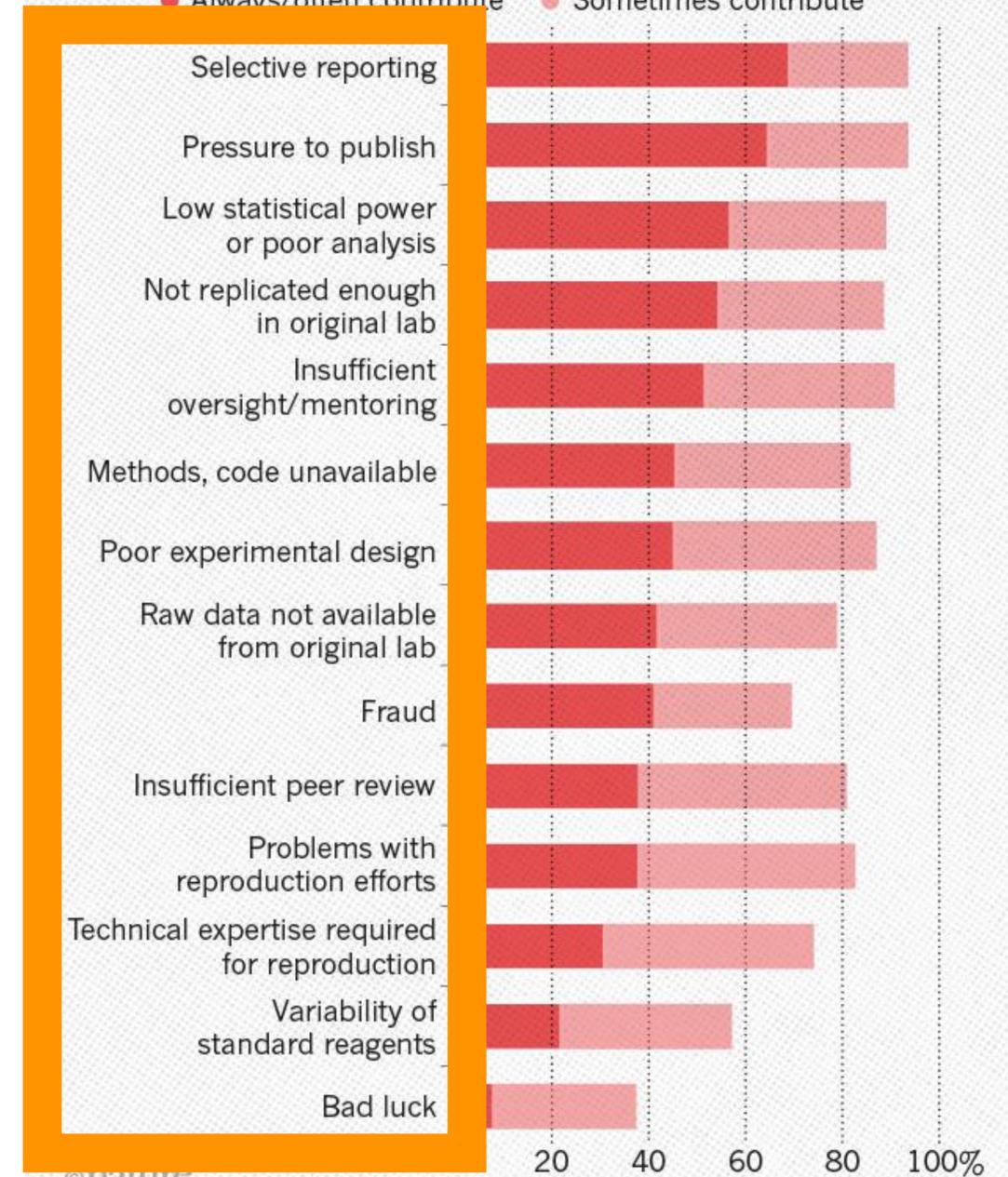


Number of respondents from each discipline: Biology 703, Chemistry 106, Earth and environmental 95, Medicine 203, Physics and engineering 236, Other 233

WHAT FACTORS CONTRIBUTE TO IRREPRODUCIBLE RESEARCH?

Many top-rated factors relate to intense competition and time pressure.

● Always/often contribute ● Sometimes contribute



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



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.



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.

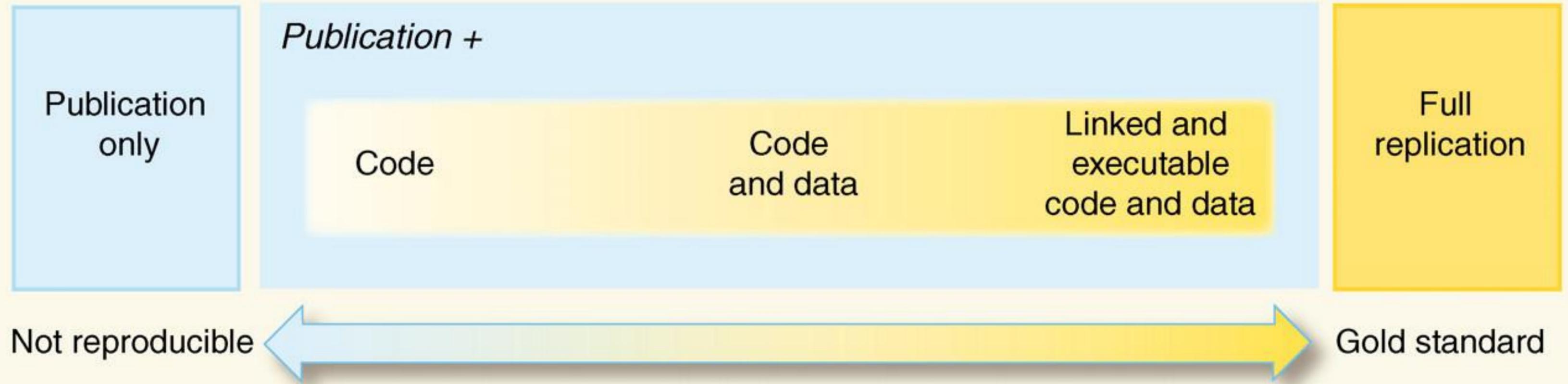
Science jobs from naturejobs

[South China Normal University sincerely invite](#)

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



Reproducibility Spectrum

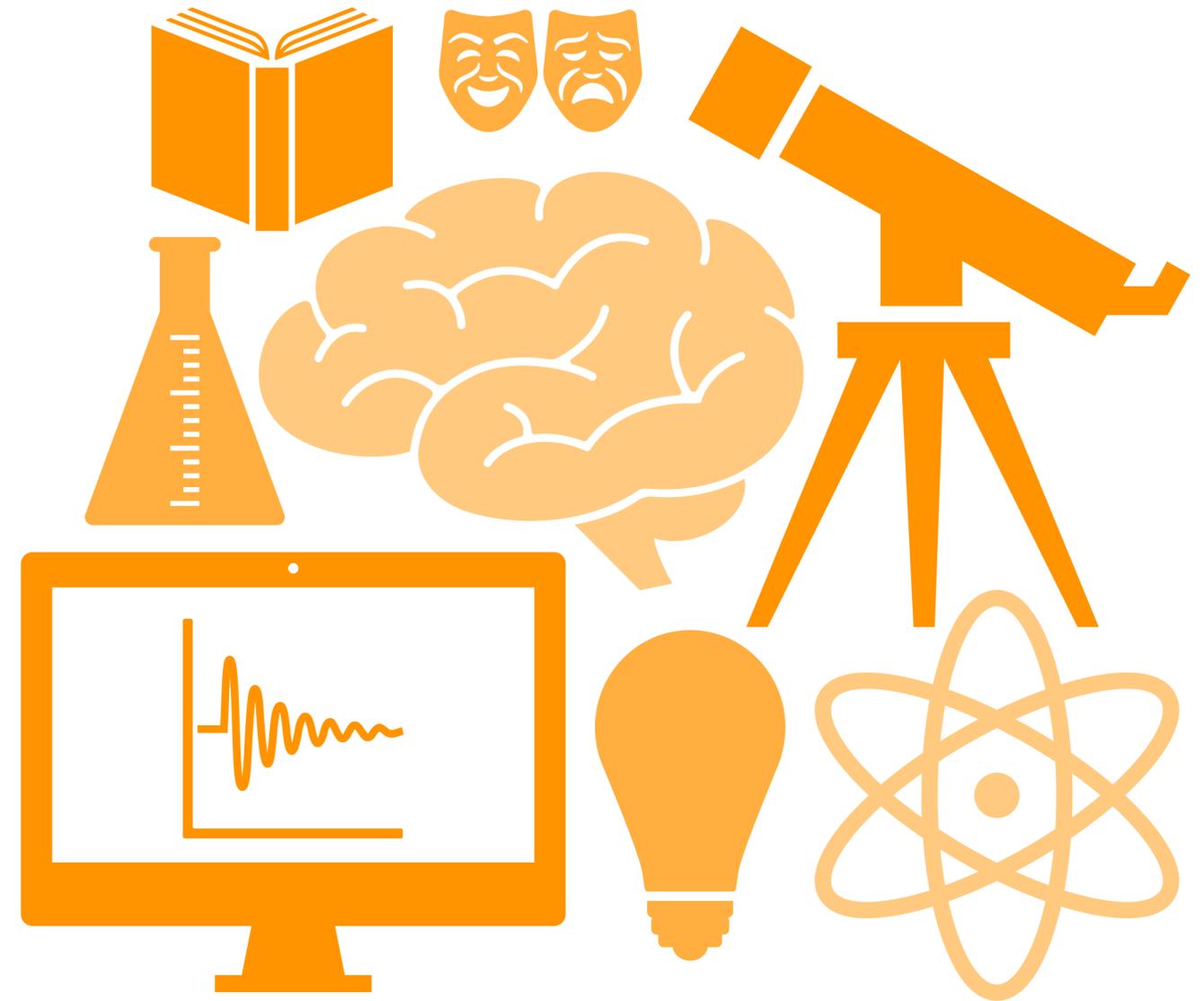


“Computational science has led to exciting new developments, but the nature of the work has exposed limitations in our ability to evaluate published findings. Reproducibility has the potential to serve as a minimum standard for judging scientific claims when full independent replication of a study is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

Peng (2011) <https://doi.org/10.1126/science.1213847>

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.



<https://royalsociety.org/topics-policy/projects/research-culture/>

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

<https://royalsociety.org/topics-policy/projects/research-culture/>



Open Science / Research / Scholarship



What is Open Science?

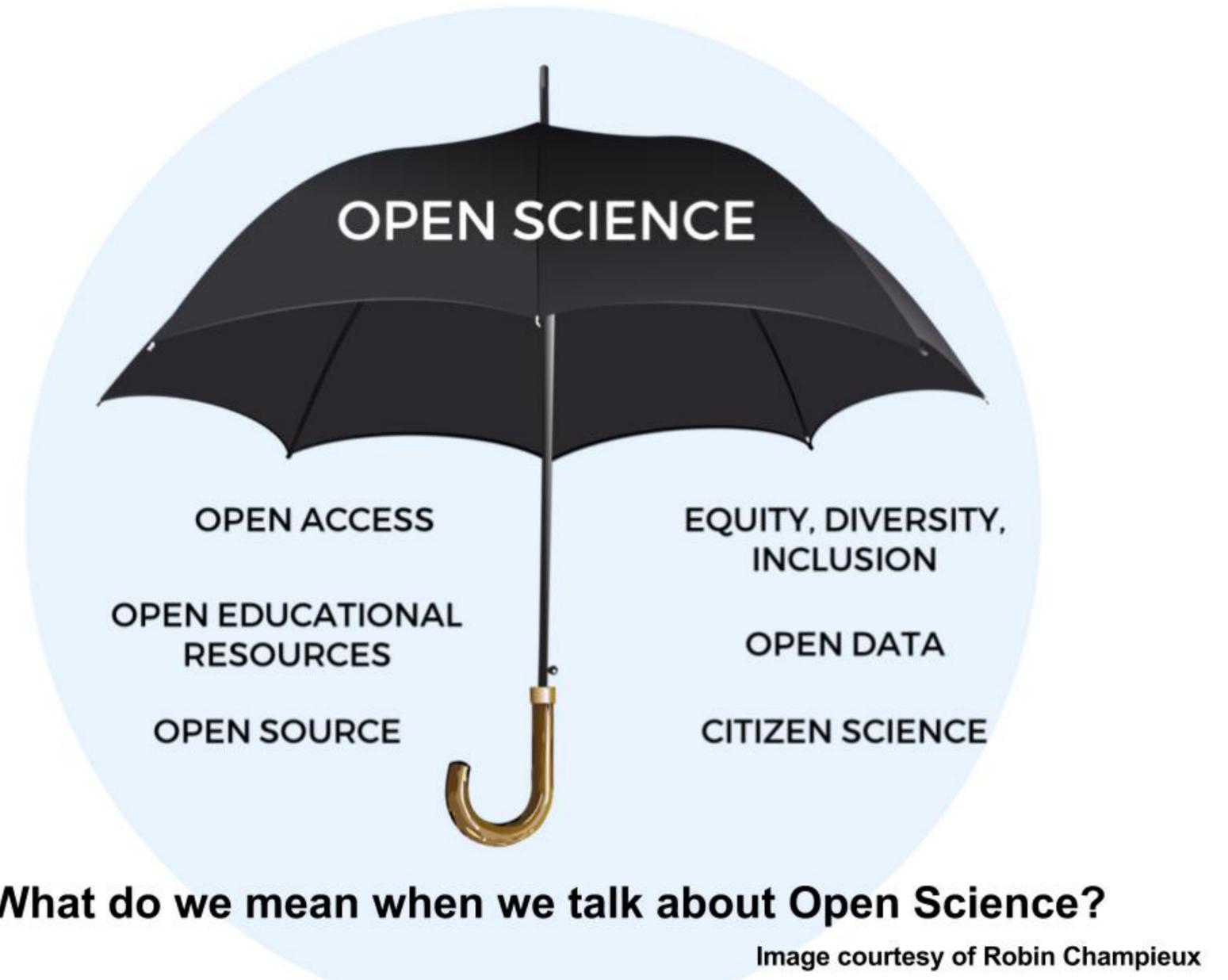
Open Science is the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.

... but isn't this just science?

Rephrase to Open Research or Scholarship to be inclusive of all research domains.

(FOSTER, Open Science Definition:

<https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>)



Barriers to Open Research



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>



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



SPRINGER NATURE

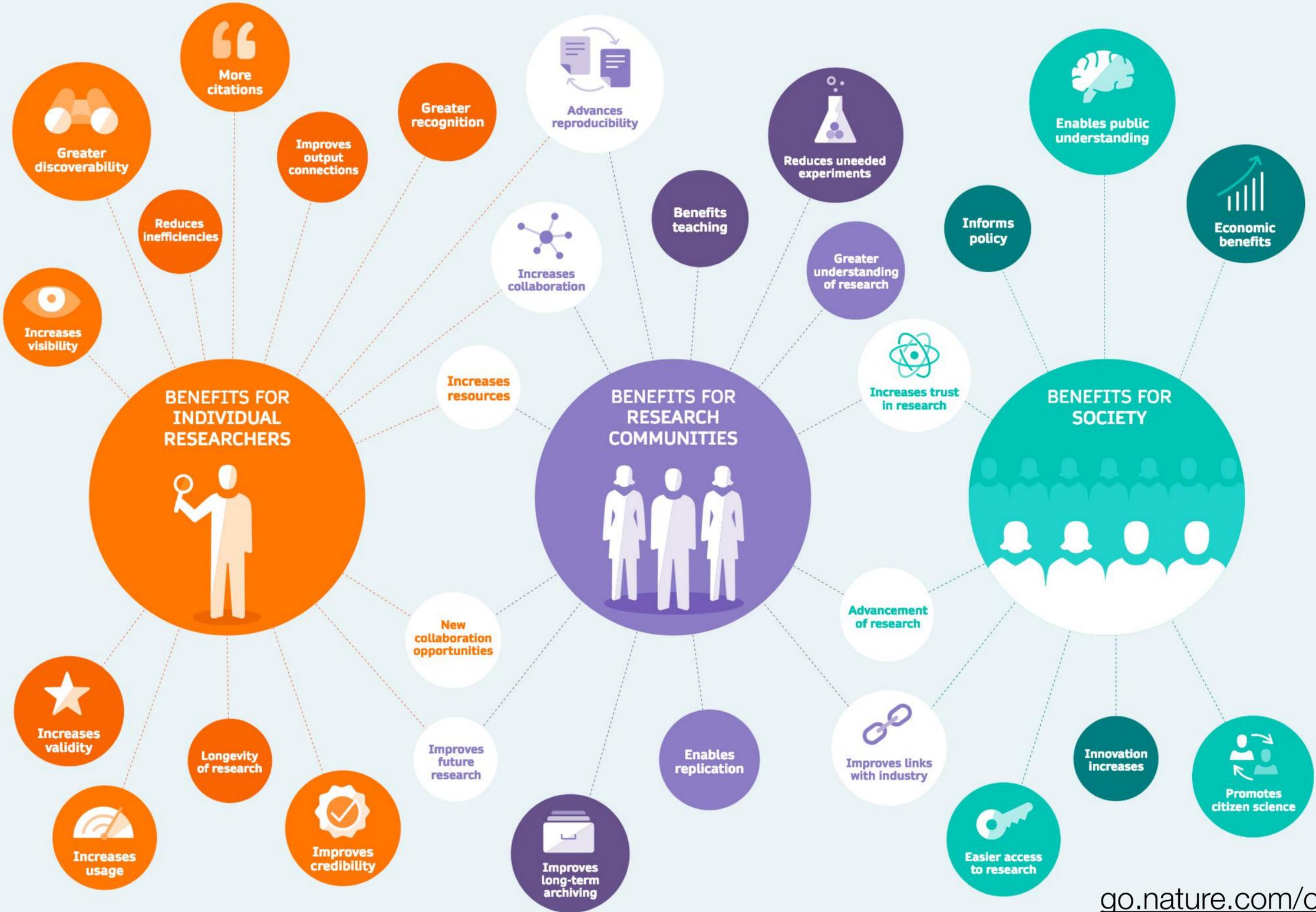
Tennant (2017) <https://doi.org/10.6084/m9.figshare.5383711.v1>

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

Why research openly?



BENEFITS TO SHARING RESEARCH DATA



go.nature.com/opendata



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

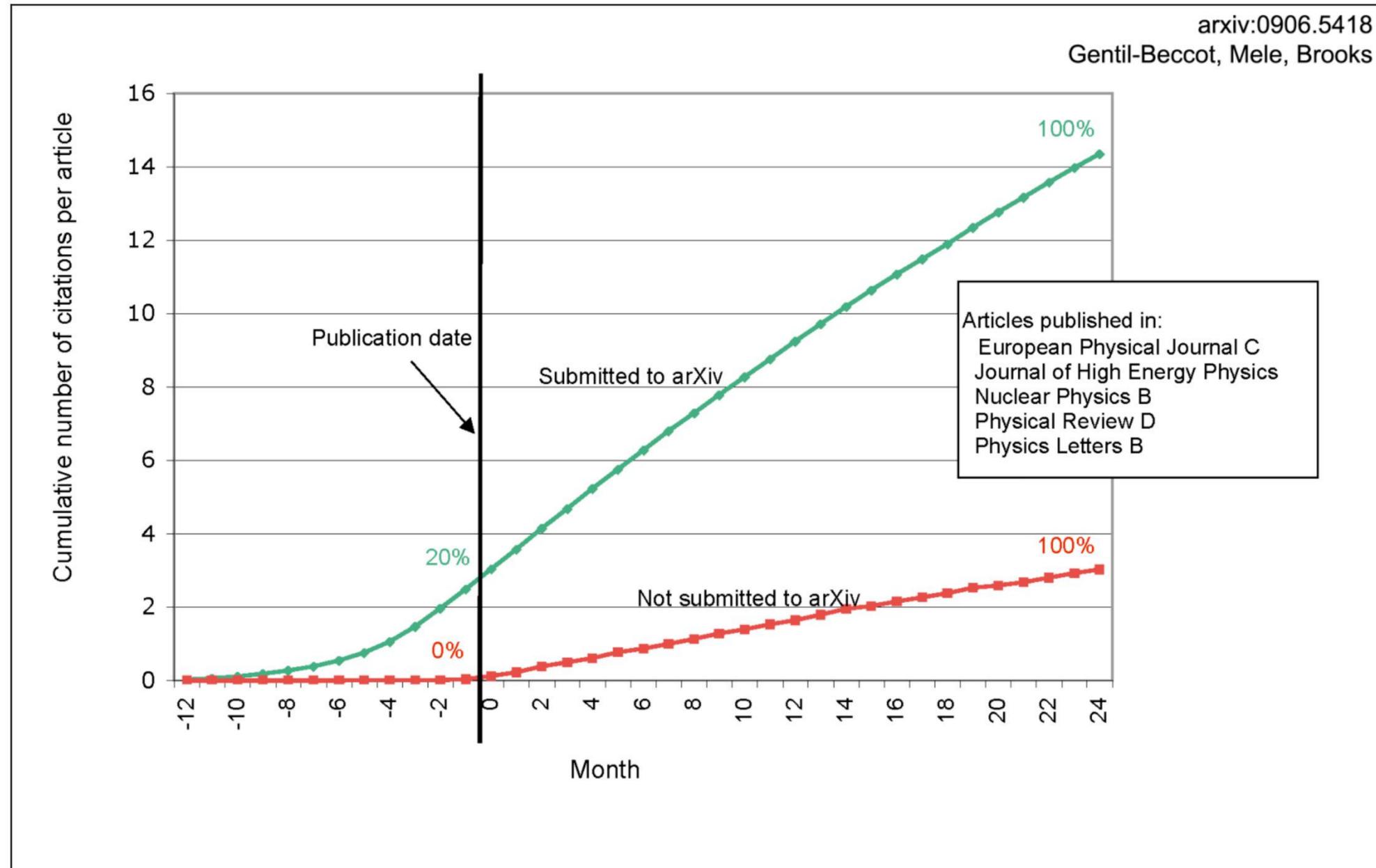
In the open

PODCAST

[Further reading »](#)



Benefit to sharing preprints: more citations!



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



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/>





European
Commission

EU BUDGET FOR THE FUTURE

HORIZON EUROPE

#EUBudget #HorizonEU



Open Science will become the modus operandi of Horizon Europe. It will go beyond the open access policy of Horizon 2020 and require open access to publications, data, and to research data management plans.

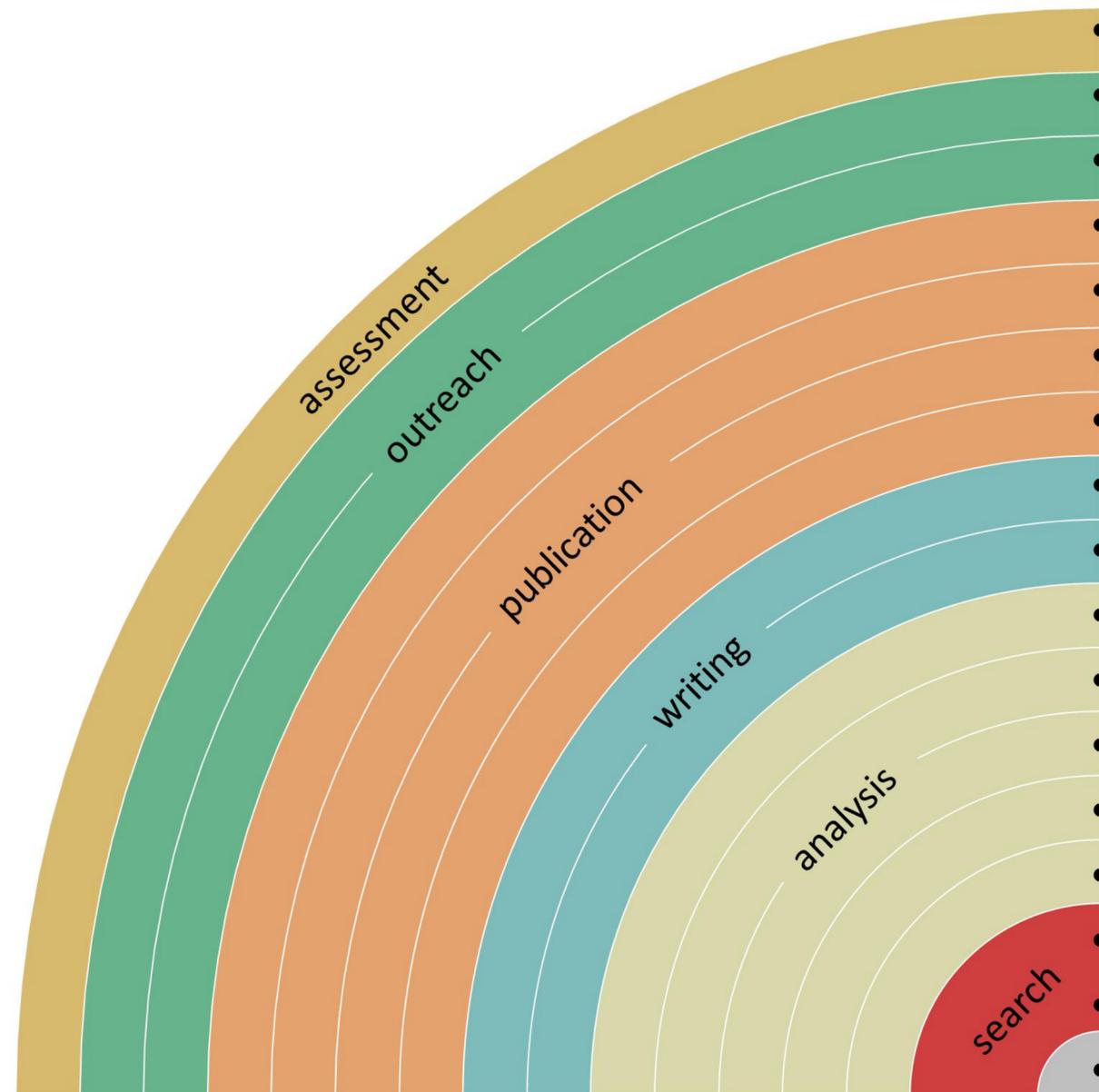
https://ec.europa.eu/commission/sites/beta-political/files/budget-may2018-research-innovation_en.pdf



How to open up your research workflow



You can make your workflow more open by...



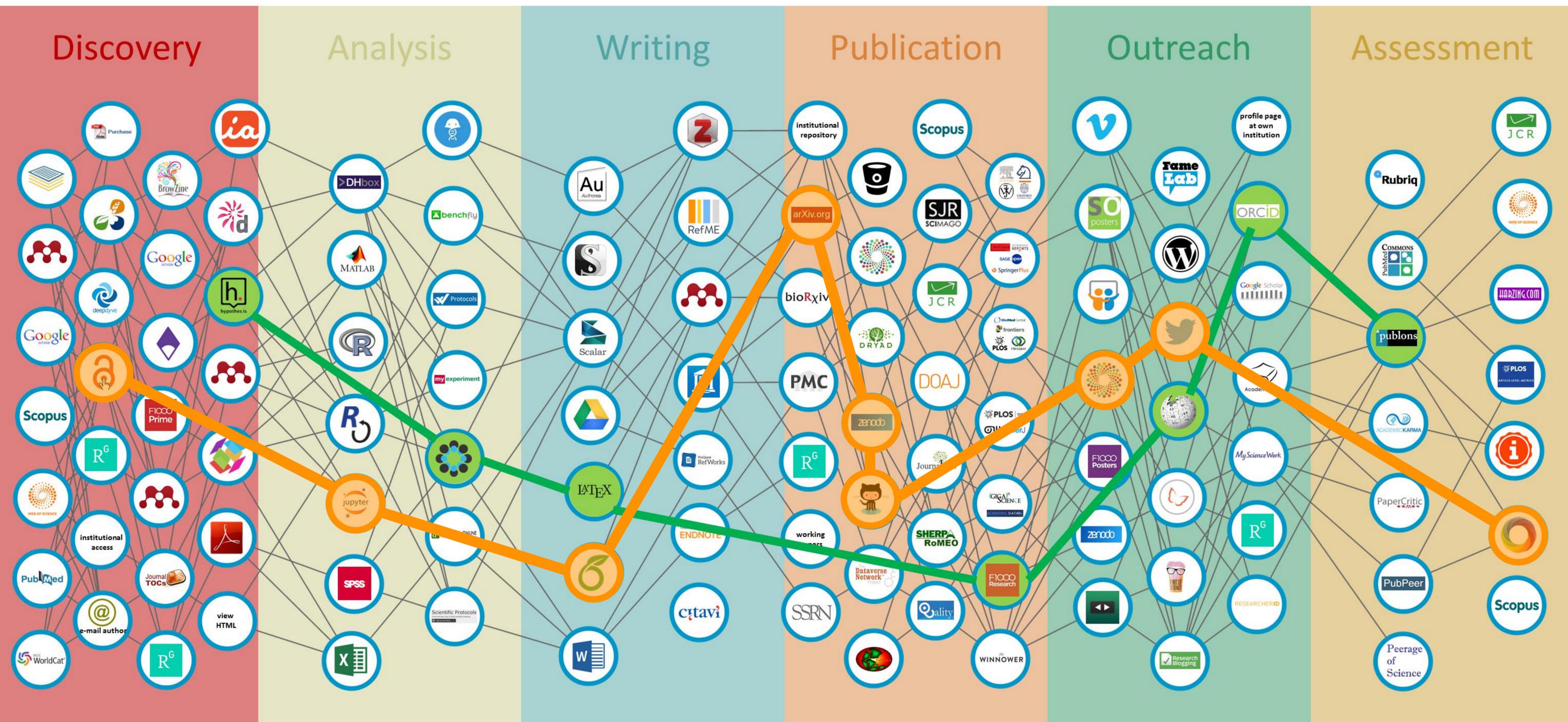
- adding alternative evaluation, e.g. with altmetrics
- communicating through social media, e.g. Twitter
- sharing posters & presentations, e.g. at FigShare
- using open licenses, e.g. CC0 or CC-BY
- publishing open access, 'green' or 'gold'
- using open peer review, e.g. at journals or PubPeer
- sharing 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](https://doi.org/10.5281/zenodo.1147025)

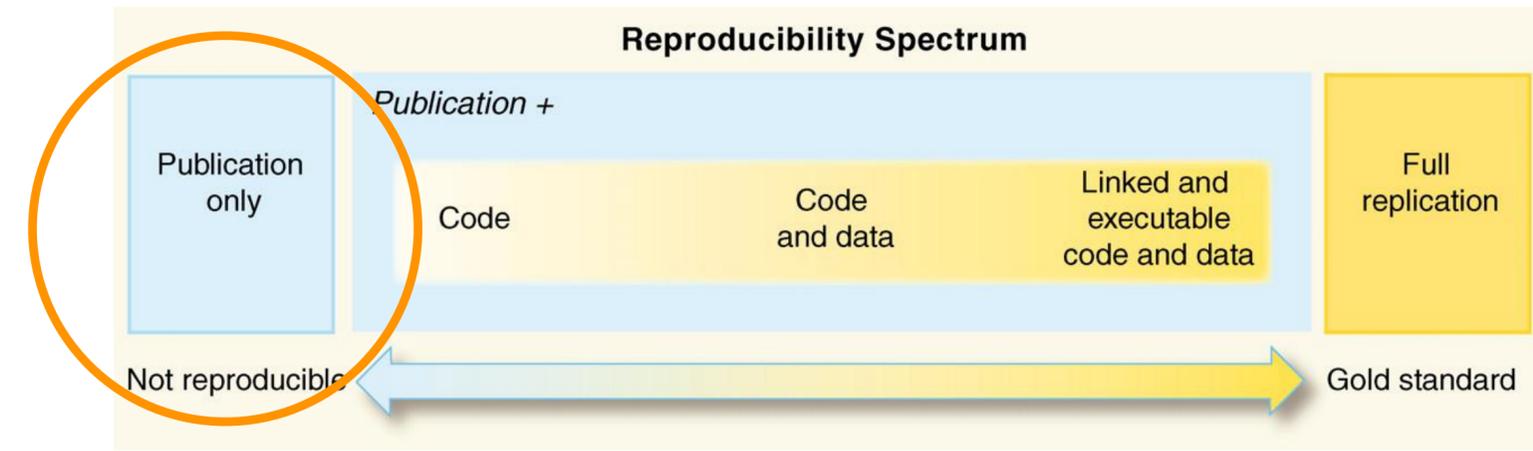




Jeroen Bosman and Bianca Kramer - <https://101innovations.wordpress.com/workflows/>



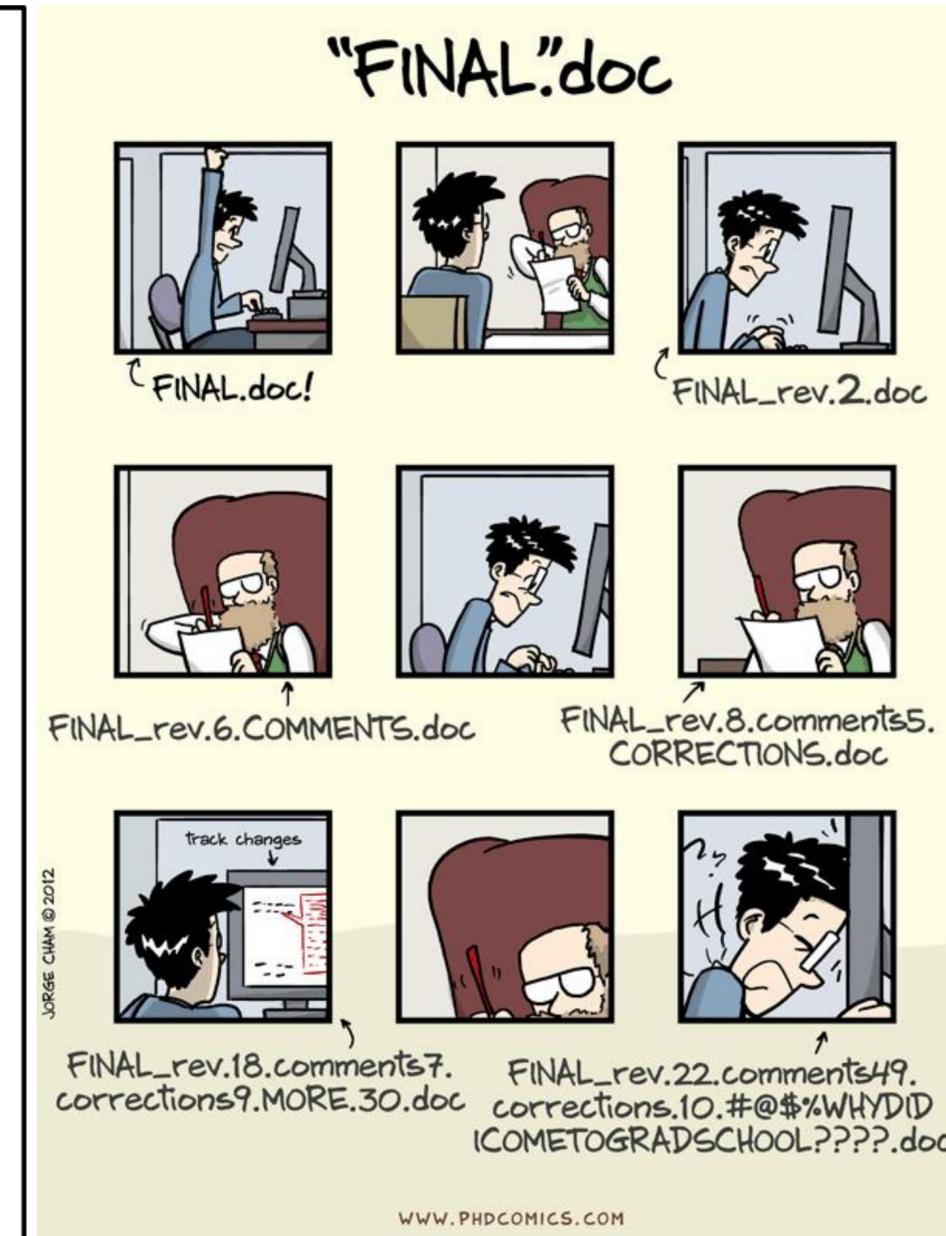
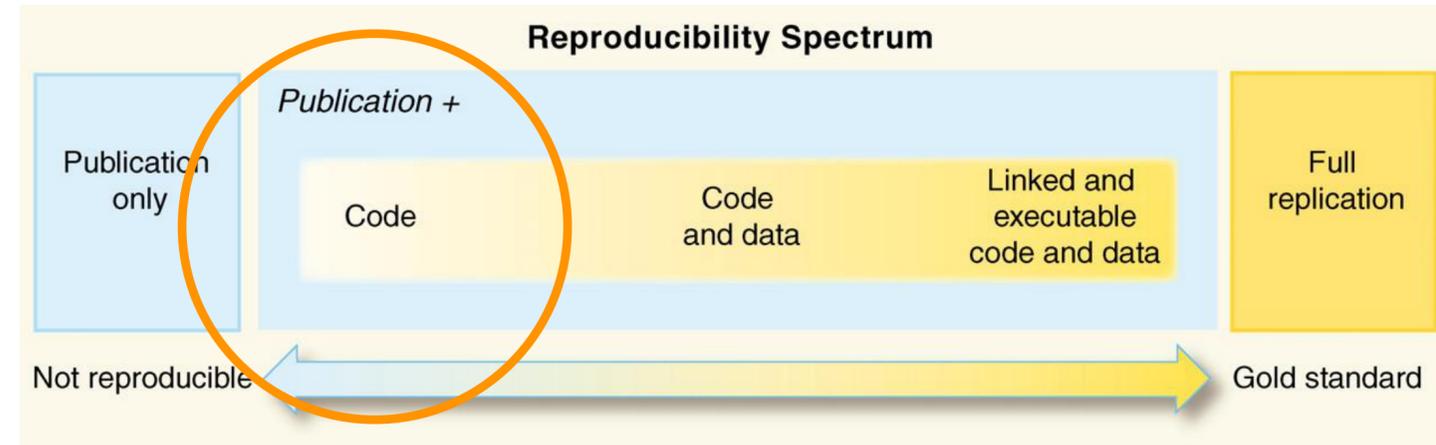
Make your articles Open Access via:
 Gold route by publishing OA, or
 Green route by self-archiving



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

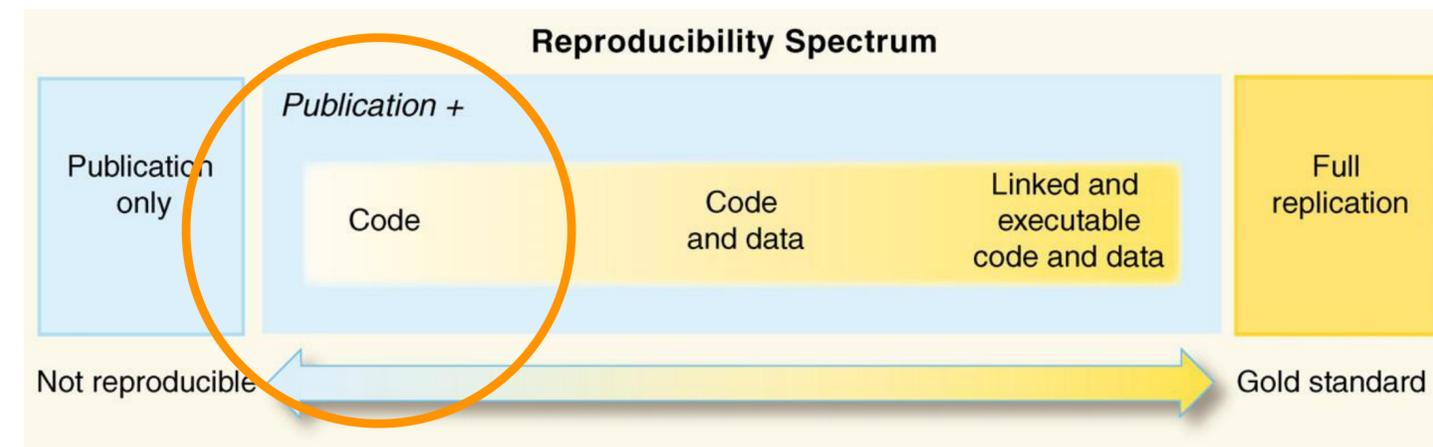
Share code & maintain version control using platforms such as Bitbucket, GitLab & GitHub

- Git is an open source program for tracking changes in text files (version control)
- GitHub is a code hosting platform for version control & collaboration. It lets you & others work together on projects from anywhere
- Facilitates open & reproducible science/code/research!
- Online portfolio & webpage for your research



Share code & maintain version control using platforms such as Bitbucket, GitLab & GitHub

- Git is an open source program for tracking changes in text files (version control)
- GitHub is a code hosting platform for version control & collaboration. It lets you & others work together on projects from anywhere
- Facilitates open & reproducible science/code/research!
- Online portfolio & webpage for your research



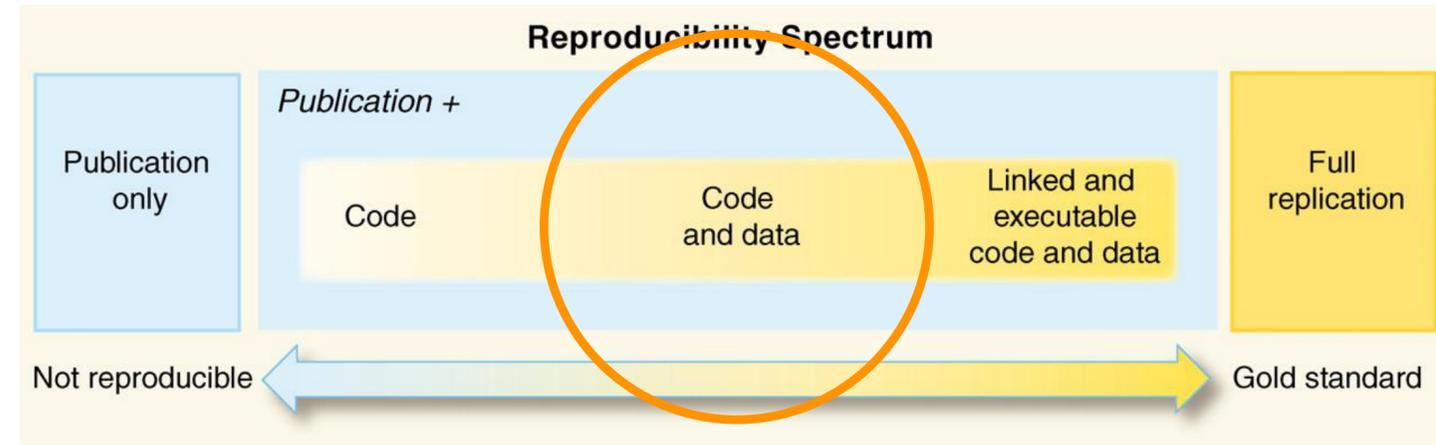
The screenshot shows the GitHub profile of Rachael Ainsworth (username: rainsworth). The profile includes a profile picture, a bio identifying her as a Radio Astronomer & Open Science Champion at the Jodrell Bank Centre for Astrophysics, and her affiliation with the University of Manchester. It also displays statistics for her repositories (11), stars (65), followers (17), and following (35). A section of popular repositories is shown, including "ROSA", "rainsworth.github.io", "GMRT-TAU_catalogue", "Spectral-Energy-Distributions", "awesomeCV", and "paper_scripts". At the bottom, there is a contribution activity heatmap for the last year (2017) showing 317 contributions.



Share research outputs in Open Repositories such as Figshare, Zenodo & the OSF

Catch-all repositories that enable researchers, scientists, projects & institutions to:

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



The top screenshot shows a Zenodo page for the paper 'Constraining the nature of DG Tau A's thermal and non-thermal radio emission' by Simon John Derek Purser et al., dated July 26, 2018. It features 78 views and 79 downloads. Below the title is a list of files for FITS data, including 'DGTau_2012_Robust-2.8.5GHz.fits' (4.7 MB) and 'DGTau_2016_Robust-2.8.5GHz.fits' (4.7 MB).

The bottom screenshot shows a Zenodo page for 'Reproducibility and Open Science' by Rachael Ainsworth, dated October 17, 2018. It has 1,426 views and 646 downloads. The page includes a presentation slide from The Alan Turing Institute, a DOI of 10.5281/zenodo.1464853, and keywords like 'Reproducibility', 'Open Science', and 'Data Science'.



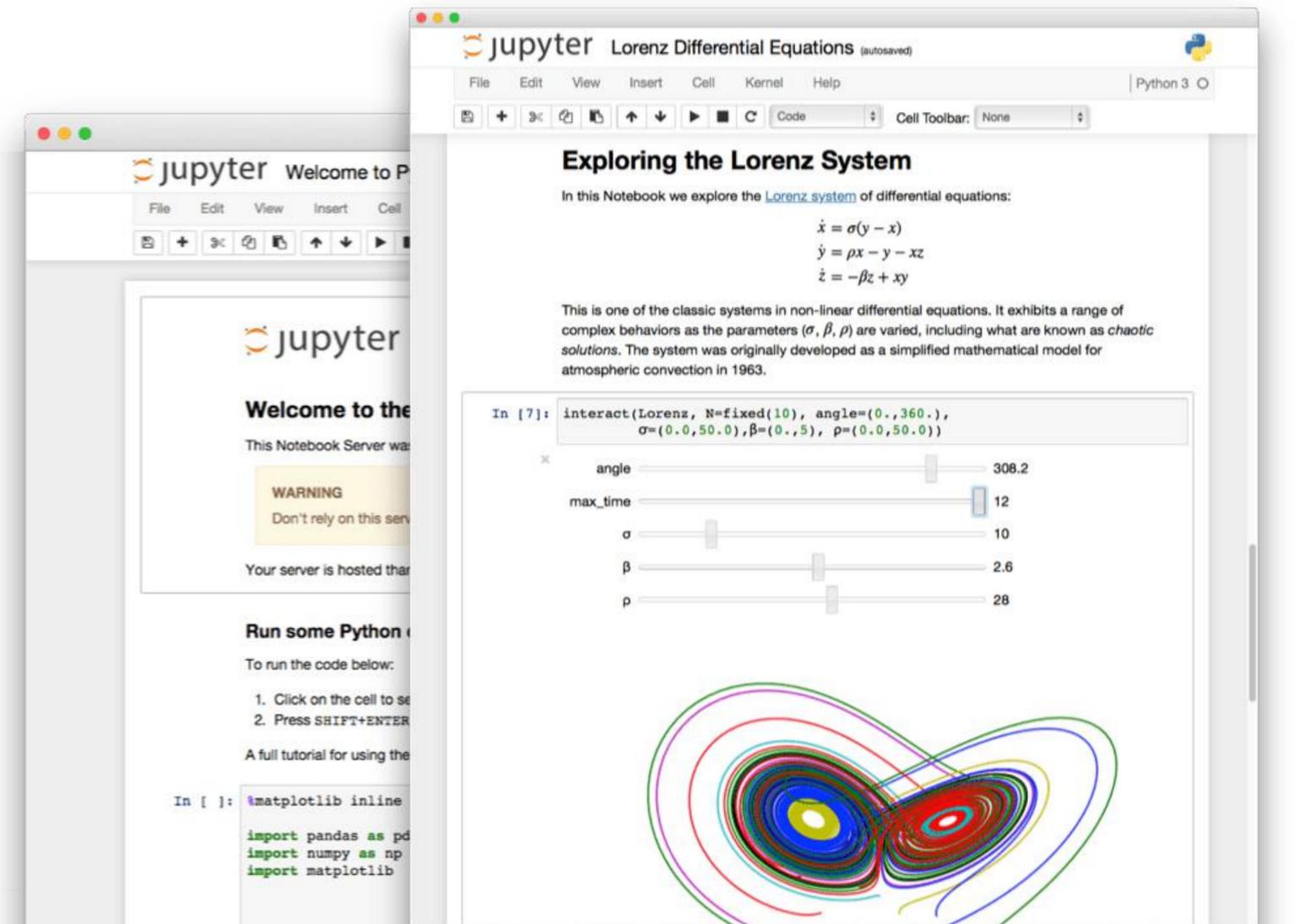
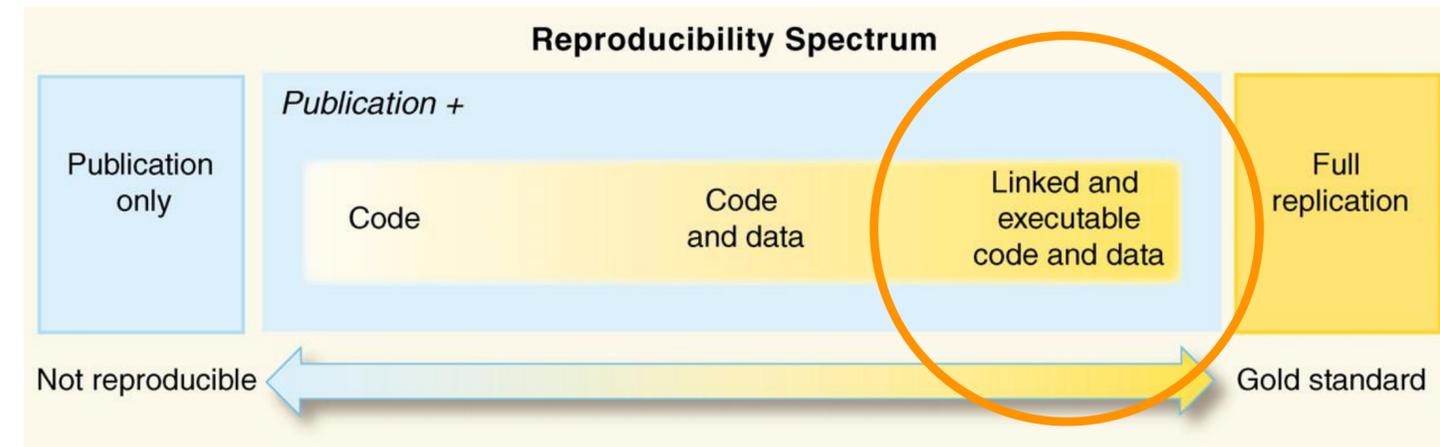
Share analyses using Open Notebooks such as Jupyter & RStudio

Open Notebooks 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.

They 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/>



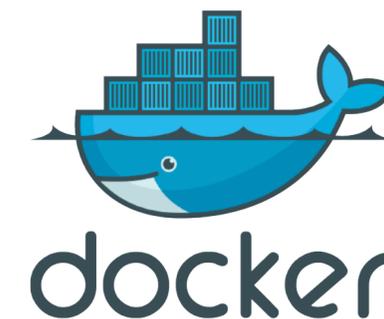
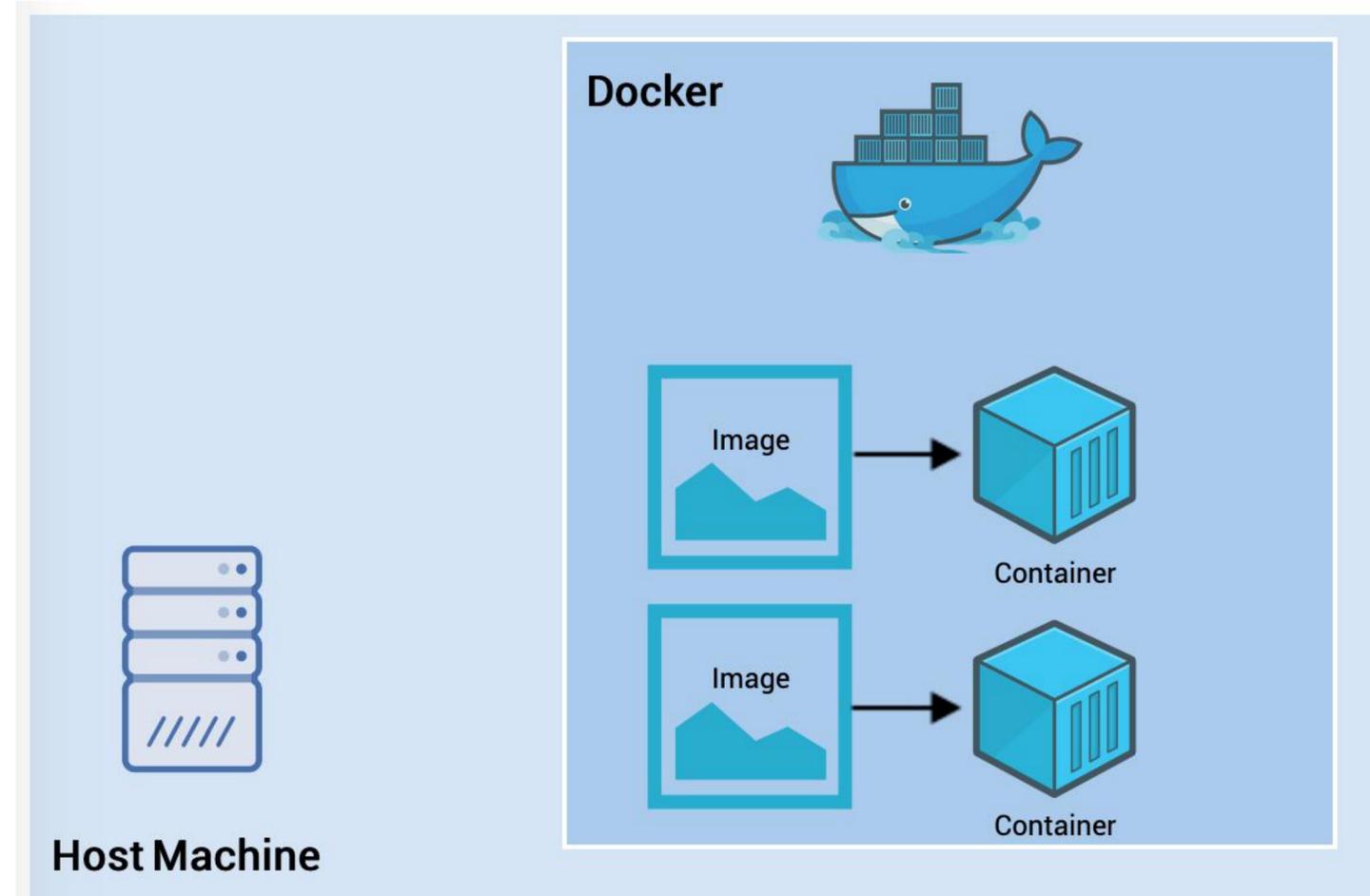
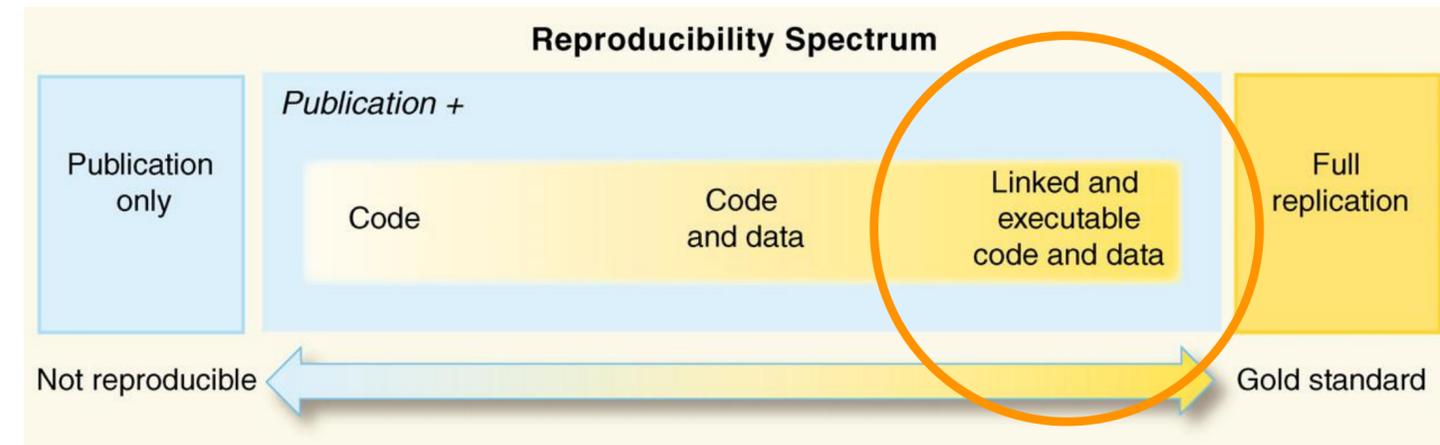
Package data, code & analyses through Containerisation such as with Docker & Singularity

A container image is a lightweight, standalone, executable package of software that includes everything needed to run an application: code, runtime, system tools, system libraries and settings.

Containers can be used to package entire scientific workflows, software and libraries, and even data. This means that you don't have to ask your cluster admin to install anything for you - you can put it in a container and run.

Need to share your code? Put it in a container and your collaborator won't have to go through the pain of installing missing dependencies.

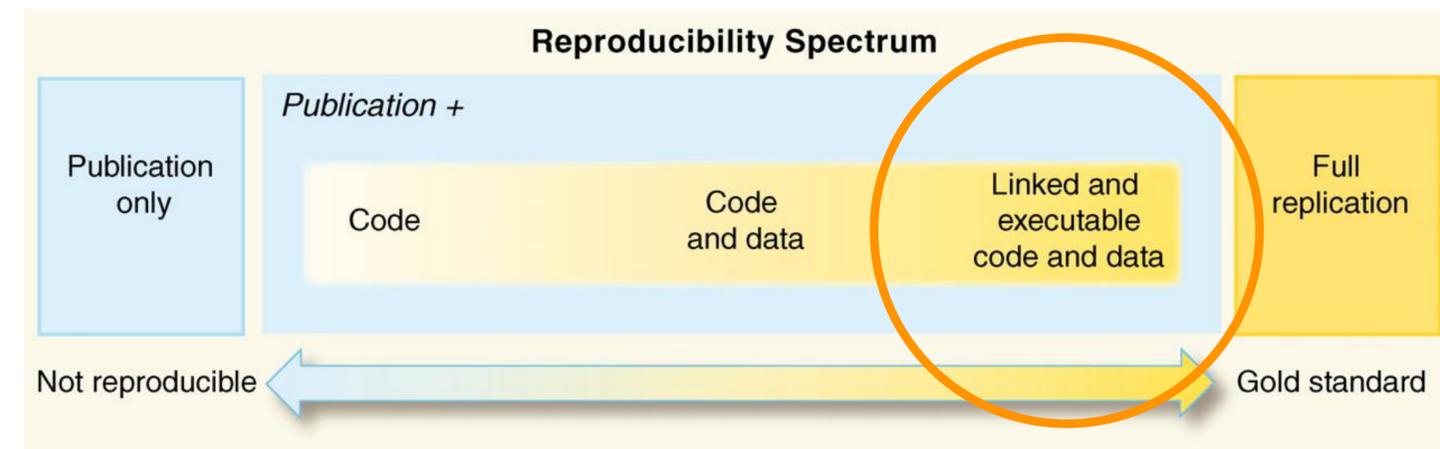
Avoids the “but it worked on *my* laptop...” problem.



Don't want to build your own container? Use Binder!

- 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 that can immediately begin interacting with this environment via the cloud.
- Binder's goal is to enable as many analytic workflows as possible.

mybinder.org



Turn a Git repo into a collection of interactive notebooks

Have a repository full of Jupyter notebooks? With Binder, open those notebooks in an executable environment, making your code immediately reproducible by anyone, anywhere.

Build and launch a repository

GitHub repository name or URL

 GitHub ▾

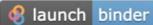
Git branch, tag, or commit

Path to a notebook file (optional)

 File ▾ launch

Copy the URL below and share your Binder with others:

 📄

Copy the text below, then paste into your README to show a binder badge:  ▶



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>



The Turing Way Community and Scriberia
<http://doi.org/10.5281/zenodo.3332808>

Scriberia 



Open Science in Astronomy & a case study



Open Science in Astronomy

Open Access:

- arXiv! Started in August 1991 and provides open access to 1,517,000+ e-prints in (Astro)Physics and many other fields

Open Data:

- 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

Citizen Science:

- Pulsar Hunters
- Galaxy Zoo



Open access to 1,517,000 e-prints in the fields of physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, electrical engineering and systems science, and economics. Submissions to arXiv should conform to Cornell University academic standards. arXiv is owned and operated by Cornell University, a private not-for-profit educational institution. arXiv is funded by Cornell University, the Simons Foundation and by the member institutions.

Subject search and browse:

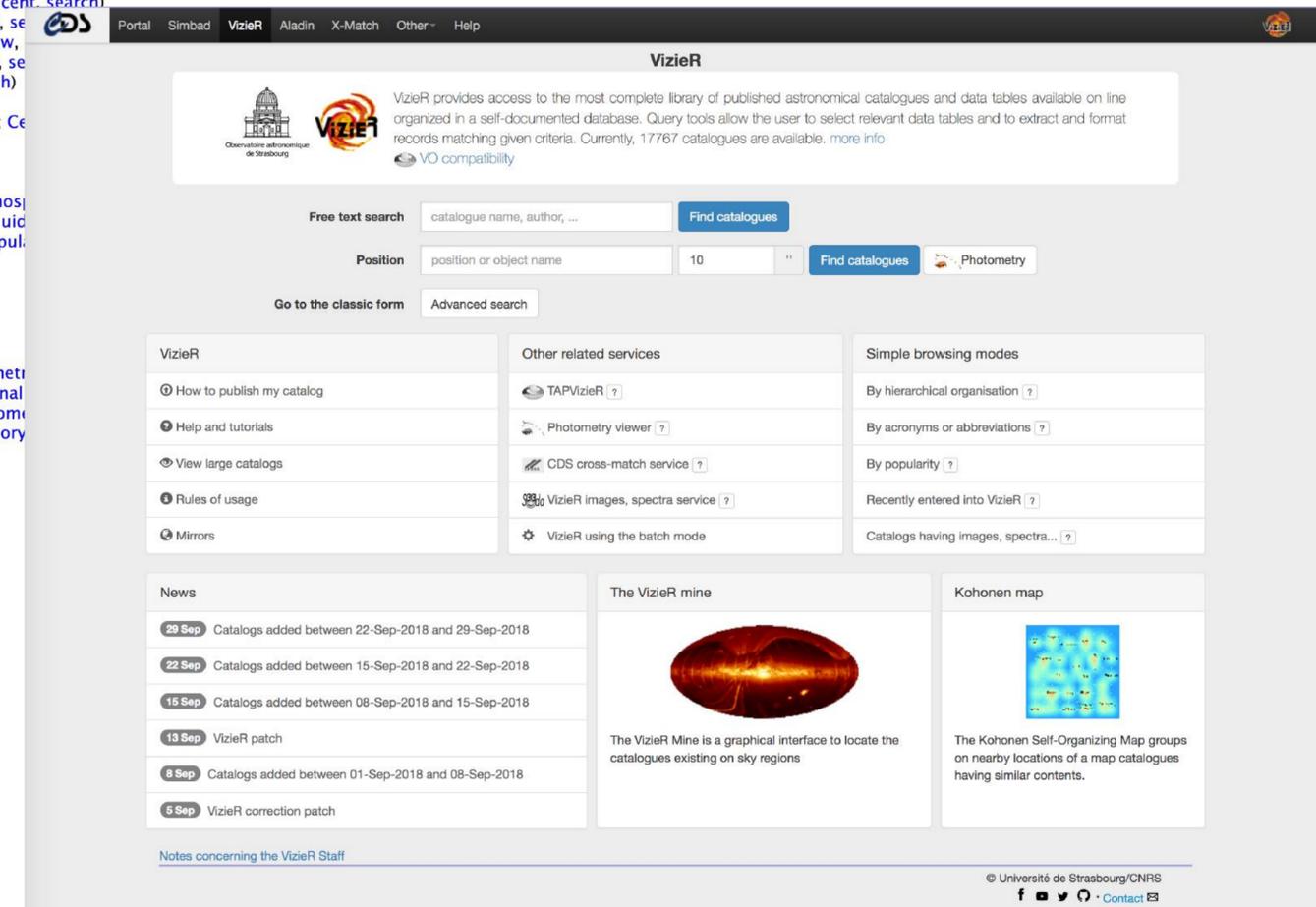
14 Jan 2019: [The annual update from the arXiv team is now available](#)
5 Sept 2018: [arXiv looks to the future with move to Cornell CIS](#)
See cumulative "What's New" pages. Read [robots beware](#) before attempting any automated download

Physics

- **Astrophysics (astro-ph new, recent, search)**
includes: Astrophysics of Galaxies; Cosmology and Nongalactic Astrophysics; Earth and Planetary Astrophysics; High Energy Astrophysical Phenomena; Instrumentation and Methods for Astrophysics; Solar and Stellar Astrophysics
- **Condensed Matter (cond-mat new, recent, search)**
includes: Disordered Systems and Neural Networks; Materials Science; Mesoscale and Nanoscale Physics; Other Condensed Matter; Quantum Gases; Soft Condensed Matter; Statistical Mechanics; Strongly Correlated Electrons; Superconductivity
- **General Relativity and Quantum Cosmology (gr-qc new, recent, search)**
- **High Energy Physics - Experiment (hep-ex new, recent, search)**
- **High Energy Physics - Lattice (hep-lat new, recent, search)**
- **High Energy Physics - Phenomenology (hep-ph new, recent, search)**
- **High Energy Physics - Theory (hep-th new, recent, search)**
- **Mathematical Physics (math-ph new, recent, search)**
- **Nonlinear Sciences (nlin new, recent, search)**
includes: Adaptation and Self-Organizing Systems; Chaos, Dynamics and Bifurcations
- **Nuclear Experiment (nucl-ex new, recent, search)**
- **Nuclear Theory (nucl-th new, recent, search)**
- **Physics (physics new, recent, search)**
includes: Accelerator Physics; Applied Physics; Atmospheric and Oceanic Physics; Data Analysis, Statistics and Probability; Fluid Dynamics; High Energy Physics; Plasma Physics; Population and Earth System Science; Quantum Physics
- **Quantum Physics (quant-ph new, recent, search)**

Mathematics

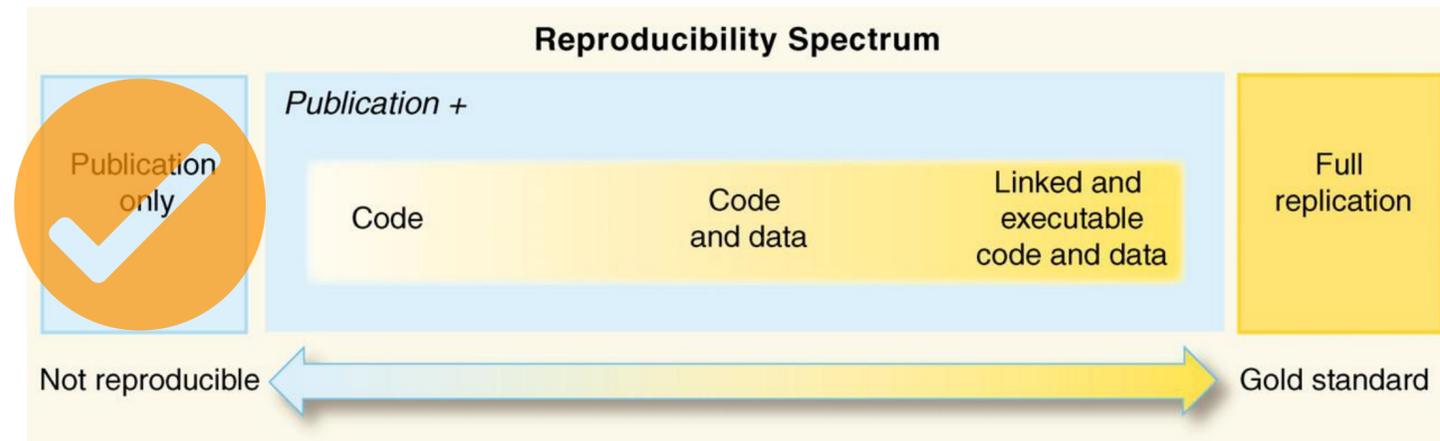
- **Mathematics (math new, recent, search)**
includes (see detailed description): Algebraic Geometry; Algebraic Combinatorics; Algebraic Topology; Differential Geometry; Dynamical Systems; Functional Analysis; Homology; Logic; Mathematical Physics; Metric Geometry; Rings and Algebras; Spectral Theory; Statistics Theory



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



Open Access to 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

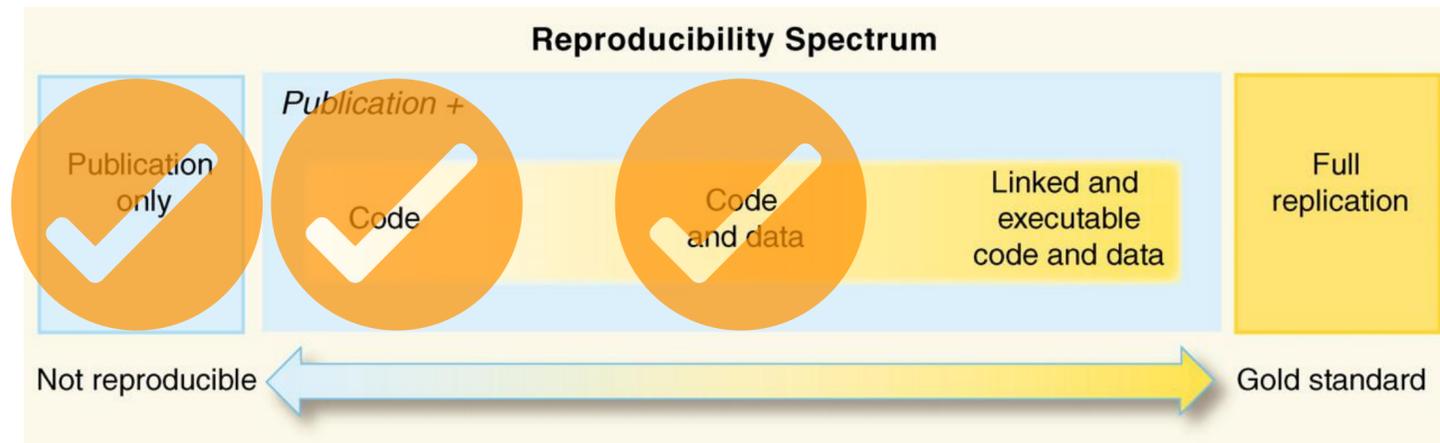
We present a framework for exploration of the statistical relationship between FRB redshifts and dispersion measures, which provides the basis for:

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 ($> 400 \text{ pc cm}^{-3}$) in the rest frame, and as expected, that 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 $P(\text{DM}|z_s)$ profiles out to $z_s = 3$.
2. More rigorous uncertainties to be placed on FRB redshifts than are currently standard practice. By consulting $P(z_s|\text{DM})$ 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 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 DM_{IGM} and DM_{host} using their respective redshift dependences.

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



Open Access to code & data



The screenshot shows the GitHub repository page for 'mbcxqcw2/EEModel'. It includes the repository name, commit history, and a 'launch binder' button. The commit history shows several updates to files like 'host_galaxies', 'lin_mp_files', and 'ExcessElectronLib.py'.

The screenshot shows the Zenodo record for 'mbcxqcw2/EEModel: Master DOI release'. It includes the record title, DOI (10.5281/zenodo.1211089), publication date (April 2, 2018), and a list of files with their sizes. The files include 'EEModel-v1.03.zip' and various data files like 'planck_lin_mp0.0.dat'.

<https://zenodo.org/record/1211089>

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



Linked & executable code & data



Starting repository: mbcxqcw2/EEModel/master
New to Binder? Check out the [Binder Documentation](#) for more information.

Build logs

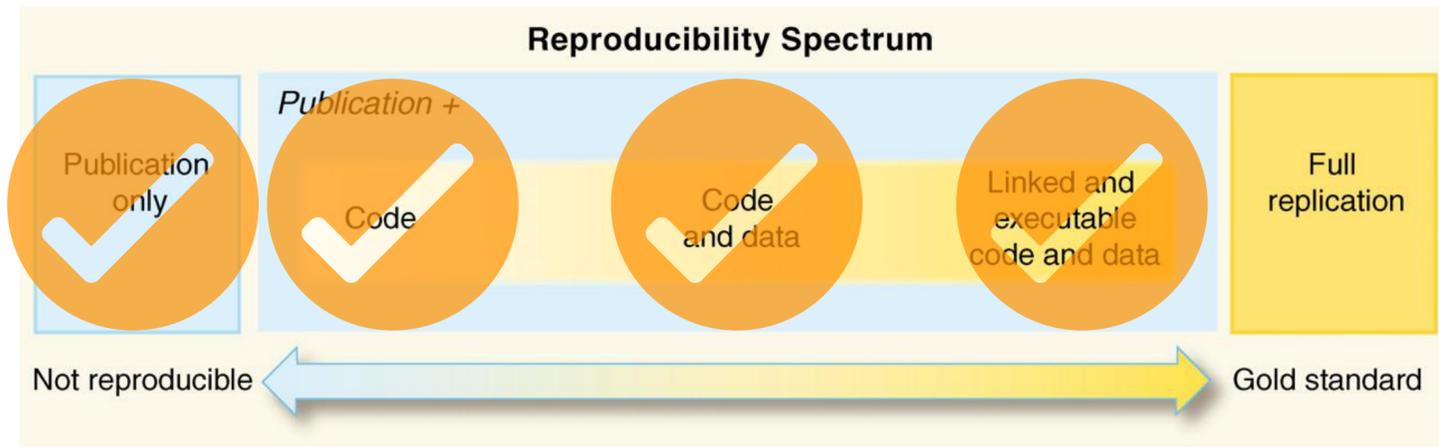
show

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 FAQ

| 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 C 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 building 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#
#####
```



Impact

- 4 April 2018:
 - Submitted manuscript to journal
 - Deposited preprint to arXiv
- 9 April 2018: Received referee report!
- 10 citations to date even though it is not officially published by the journal yet

<https://ui.adsabs.harvard.edu/abs/2018arXiv180401548W/citations>

Papers that cite

Constraining Redshifts of Unlocalised Fast Radio Bursts

[Q view this list in a search results page](#)

- 1 2019MNRAS.488.4220H 2019/09   
Fast radio burst dispersion measures and rotation measures and the origin of intergalactic magnetic fields
Hackstein, S.; Brügger, M.; Vazza, F. *and 2 more*
- 2 2019arXiv190902821W 2019/09   
Probing Diffuse Gas with Fast Radio Bursts
Walters, Anthony; Ma, Yin-Zhe; Sievers, Jonathan *and 1 more*
- 3 2019arXiv190706440R 2019/07   
A Roadmap for Astrophysics and Cosmology with High-Redshift 21 cm Intensity Mapping
Reionization Array, The Hydrogen Epoch of; Collaboration; Aguirre, James E. *and 29 more*
- 4 2019NatAs...3..928R 2019/07   
The prevalence of repeating fast radio bursts
Ravi, Vikram
- 5 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*
- 6 2019BAAS...51c.420R 2019/05   
Fast Radio Burst Tomography of the Unseen Universe
Ravi, Vikram; Battaglia, Nicholas; Burke-Spolaor, Sarah *and 13 more*
- 7 2019ApJ...872...88R 2019/02   
Measuring the Circumgalactic and Intergalactic Baryon Contents with Fast Radio Bursts
Ravi, Vikram
- 8 2018MNRAS.480.3907V 2018/11   
Probing the origin of extragalactic magnetic fields with Fast Radio Bursts
Vazza, F.; Brügger, M.; Hinz, P. M. *and 3 more*
- 9 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*
- 10 2018PhRvD..98j3518M 2018/11   
Finding the missing baryons with fast radio bursts and Sunyaev-Zeldovich maps
Muñoz, Julian B.; Loeb, Abraham



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 (<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 & linking to the DOI in your publications.
- Check out *The Turing Way* - a handbook on reproducible research/ data science openly developed at <https://github.com/alan-turing-institute/the-turing-way/>
- TEDx talk: Research Culture is Broken; Open Science can [help] fix it <https://youtu.be/c-bemNZ-lqA>
- Get in touch! Email: rachael.ainsworth@manchester.ac.uk

