

# The Alan Turing Institute

---

**Open Science & Reproducibility:**  
*Ethically-led computational research*

Malvika Sharan  
Pronouns: she/her/hers



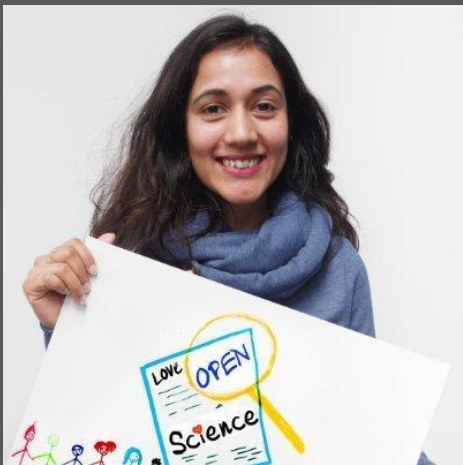
---

# Objectives

- **Setting the tone:** *Scientific errors have real world effect!*
- Define reproducible and open research
- How reproducible and open principles relate to ethics
- Reflect on your own work (current or future)

*Apply best practices & share resources from your own work!*

# The Alan Turing Institute



Senior Researcher

Tools, Practices and Systems  
The Alan Turing Institute, UK

- PhD in Bioinformatics
- Open Access & Open Source research publications
- Computational and Open Science skill training (2015-)
- Community Building in Open Science (2016-) - **EMBL-HD**
- Co-lead of *The Turing Way* & Open Life Science (2019-)



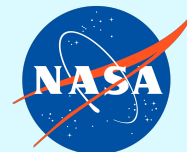
Open Life Science



Software  
Sustainability  
Institute



THE  
CARPENTRIES



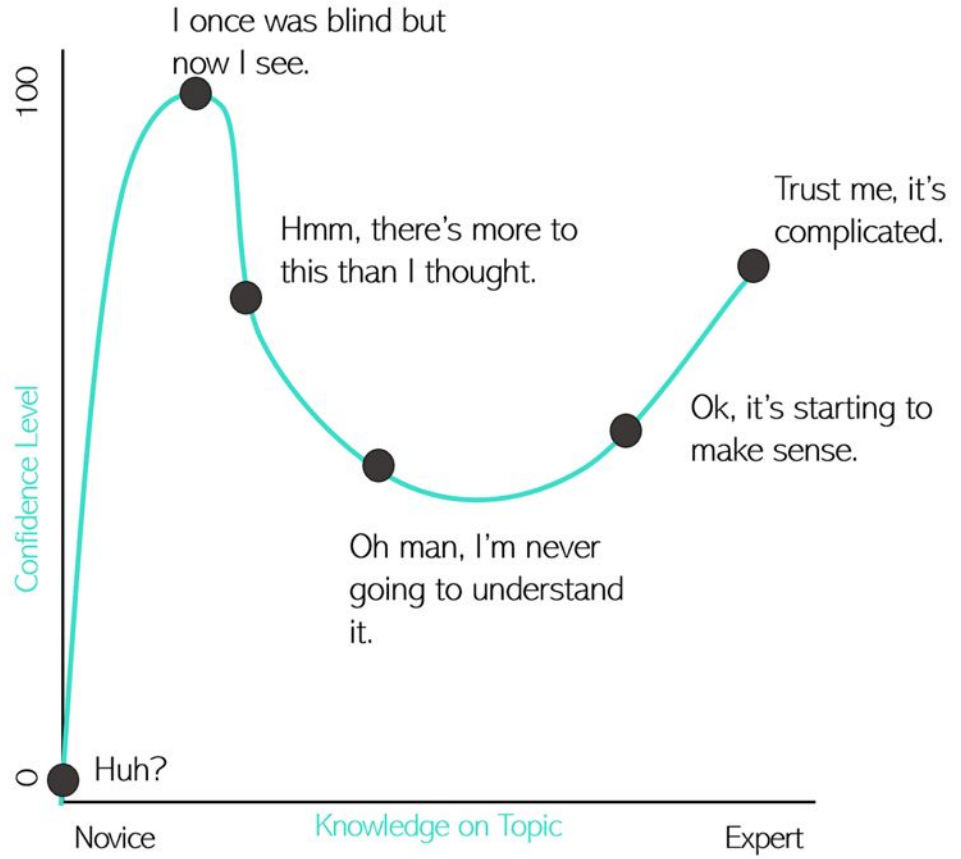
Meta  
Docencia

CS&S

Code for  
Science &  
Society

# Disclaimer:

You probably already know all about it!



Kaylee Somerville, The Hidden Power of Intellectual Humility - The Decision Lab. 2020. <https://thedecisionlab.com/insights/society/the-hidden-power-of-intellectual-humility>

Adapted from: Squad. (2018, December 13). Dunning-Kruger Effect: Definition, Test, Examples & Quiz. Science Terms. <https://scienceterms.net/psychology/dunning-kruger-effect/>



## Scientific errors have real world effects

---

*Researchers have an obligation to consider the ethical standards (right actions) and their impact on society.*



# Scientific errors have real world effects

	B	C	I	J	K	L	M
2	Real GDP growth						
3	Debt/GDP						
4	Country	Coverage	30 or less	30 to 60	60 to 90	90 or above	30 or less
26			3.7	3.0	3.5	1.7	5.5
27	Minimum		1.6	0.3	1.3	-1.8	0.8
28	Maximum		5.4	4.9	10.2	3.6	13.3
29							
30	US	1946-2009	n.a.	3.4	3.3	-2.0	n.a.
31	UK	1946-2009	n.a.	2.4	2.5	2.4	n.a.
32	Sweden	1946-2009	3.6	2.9	2.7	n.a.	6.3
33	Spain	1946-2009	1.5	3.4	4.2	n.a.	9.9
34	Portugal	1952-2009	4.8	2.5	0.3	n.a.	7.9
35	New Zealand	1948-2009	2.5	2.9	3.9	-7.9	2.6
36	Netherlands	1956-2009	4.1	2.7	1.1	n.a.	6.4
37	Norway	1947-2009	3.4	5.1	n.a.	n.a.	5.4
38	Japan	1946-2009	7.0	4.0	1.0	0.7	7.0
39	Italy	1951-2009	5.4	2.1	1.8	1.0	5.6
40	Ireland	1948-2009	4.4	4.5	4.0	2.4	2.9
41	Greece	1970-2009	4.0	0.3	2.7	2.9	13.3
42	Germany	1946-2009	3.9	0.9	n.a.	n.a.	3.2
43	France	1949-2009	4.9	2.7	3.0	n.a.	5.2
44	Finland	1946-2009	3.8	2.4	5.5	n.a.	7.0
45	Denmark	1950-2009	3.5	1.7	2.4	n.a.	5.6
46	Canada	1951-2009	1.9	3.6	4.1	n.a.	2.2
47	Belgium	1947-2009	n.a.	4.2	3.1	2.6	n.a.
48	Austria	1948-2009	5.2	3.3	-3.8	n.a.	5.7
49	Australia	1951-2009	3.2	4.9	4.0	n.a.	5.9
50							
51			4.1	2.8	2.8	=AVERAGE(L30:L44)	

BBC Sign in News Sport Weather iPlayer Sounds

## NEWS

Home UK World Business Politics Tech Science Health Family & Education

### Magazine

# Reinhart, Rogoff... and Herndon: The student who caught out the profs

By Ruth Alexander  
BBC News

© 20 April 2013

f t e Share

This week, economists have been astonished to find that a famous academic paper often used to make the case for austerity cuts contains major errors. Another surprise is that the mistakes, by two eminent Harvard professors, were spotted by a student doing his homework.



It's 4 January 2010, the Marriott Hotel in Atlanta. At the annual meeting of the American Economic Association, Professor Carmen Reinhart and the former chief economist of the International Monetary Fund, Ken Rogoff, are presenting a research paper called Growth in a Time of Debt.

<https://statmodeling.stat.columbia.edu/2013/04/16/memo-to-reinhart-and-rogoff-i-think-its-best-to-admit-your-errors-and-go-on-from-there>  
<https://www.bbc.co.uk/news/magazine-22223190>

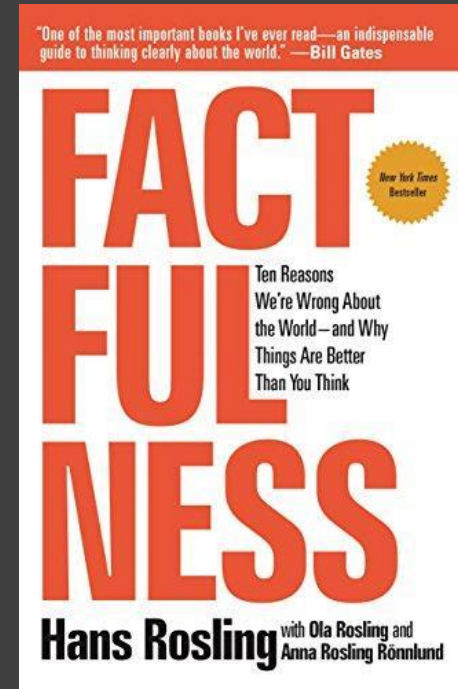
@turingway, CC-BY 4.0,  
The Turing Way, DOI: 10.5281/zenodo.7428708

---

# Scientific errors have real world effects

“We will redouble our efforts to avoid such errors in the future ...” - Reinhart & Rogoff

“Our problem is that we don’t know what we don’t know, and even our guesses are informed by unconscious and predictable biases.” - Hans Rosling



<https://www.gapminder.org/>, Rosling, H., Rönnlund, A. R., & Rosling, O. (2018). Factfulness: Ten Reasons We're Wrong About the World--and Why Things Are Better Than You Think. <https://www.bbc.co.uk/news/magazine-22223190>

@turingway, CC-BY 4.0,  
The Turing Way, DOI: 10.5281/zenodo.7428708

---

# Price of Popularity

Researchers in Australia publishes unreliable COVID-19 data with their preprint and misreported that an anti-parasite drug could stop the virus from replicating in cells.



*In a desperate attempt to save dying patients, doctors began justifying the drug's use against COVID-19 as the virus spread aggressively throughout Latin America.*

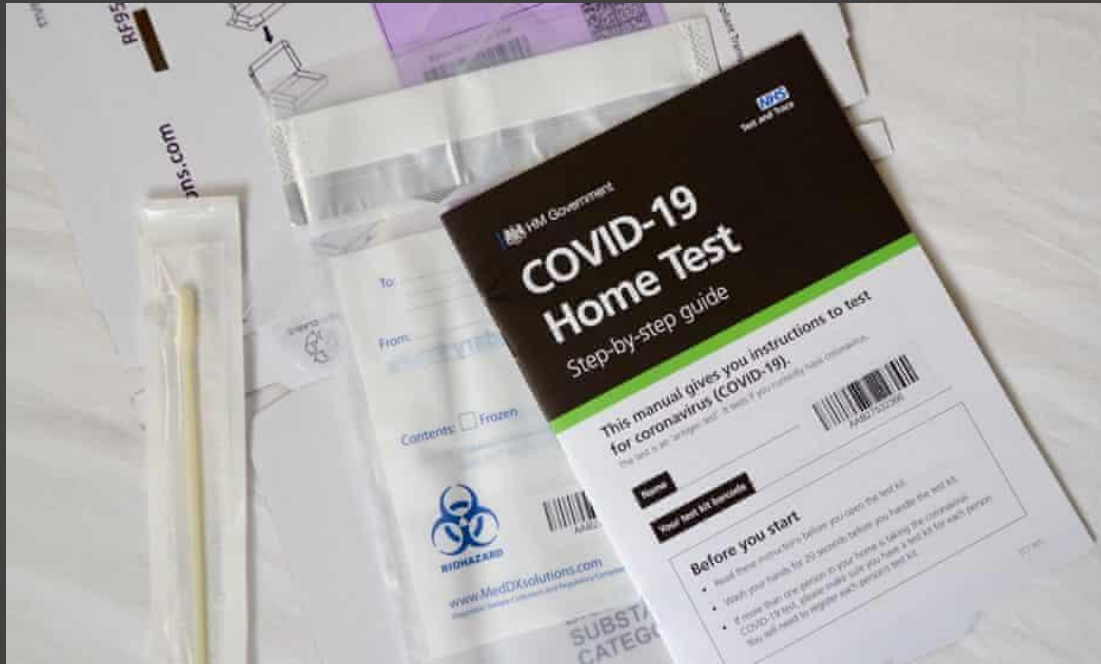
Reardon, S. (2021). Flawed ivermectin preprint highlights challenges of COVID drug studies. *Nature*, 596, 173–174. doi: 10.1038/d41586-021-02081-w. Mega, E. R. (2020). Latin America's embrace of unproven COVID treatment hinders drug trials. *Nature*, 586, 481–482. doi: 10.1038/d41586-020-02958-2, Slides under DOI: 10.5281/zenodo.7428708

# Post-publication Peer Review is Important, But ...

*To consult the [experts] after an experiment is finished is often merely to ask to conduct a **post mortem examination**. [...] can perhaps say what the experiment died of. - Ronald Fisher*



# Public Scrutiny is Important for Accountability



## Under-reported figures

From 25 Sept to 2 Oct

**50,786**

Cases initially reported by PHE

**15,841**

Unreported cases, missed due to IT error

**8 days** of incomplete data

**1,980** cases per day, on average, were missed in that time

**48 hours** Ideal time limit for tracing contacts after positive test

Source: PHE and gov.uk [↗](#)



# Bias and ‘That’s what you would expect there’

*Data about location, or ‘geospatial’ data, is the record of what we do, and where we do it. - GOV*

*There is a danger that [...] how it is analysed, can reflect inherent biases about a location and, equally, can influence how the findings of a project are interpreted. - UKSA*



9 Location Data Opportunities

- Infrastructure
- Transport
- Housing
- Environment
- Public health
- Emergency response
- Ocean economy
- Retail
- Finance

**Mission 1:** Promote and safeguard the use of location data.

**Mission 2:** Improve access to better location data.

**Mission 3:** Enhance capabilities, skills and awareness.

**Mission 4:** Enable Innovation.

*Unlocking the power of location: The UK’s geospatial strategy. (2020) GOV.*

*<https://www.gov.uk/government/publications/unlocking-the-power-of-locationthe-uks-geospatial-strategy>*

*Ethical considerations in the use of geospatial data for research and statistics. (2022, March 01). UK Statistics Authority.*

*[uksa.statisticsauthority.gov.uk/publication/ethical-considerations-in-the-use-of-geospatial-data-for-research-and-statistics/pages/5](https://uksa.statisticsauthority.gov.uk/publication/ethical-considerations-in-the-use-of-geospatial-data-for-research-and-statistics/pages/5). Office, C.*

---

# Misconduct in Biomedical Research: “The Dark Side of Science”

*Systematic scan of 20,000 papers in 40 journals published over 20 years shows inappropriately duplicated images in about 4% of them.*

*“We need to rely on science, but we need to be better able to distinguish good science from bad. We need to make science more self-correcting.”*



**Elisabeth Bik**

*Microbiologist by training,  
Scientific misconduct hunter by  
passion, professional scientific  
integrity consultant*

*Bik, E. M., Fang, F. C., Kullas, A. L., Davis, R. J., & Casadevall, A. (2018). Analysis and Correction of Inappropriate Image Duplication: the Molecular and Cellular Biology Experience. *Molecular and Cellular Biology*, 30037982. <https://pubmed.ncbi.nlm.nih.gov/30037982>  
Genetics has an estimated retraction rate 0.15%. Fabrication/falsification and plagiarism were almost mutually exclusive reasons for article retraction of entire retracted papers in life science. Dal-Ré, R., & Ayuso, C. (2019). Reasons for and time to retraction of genetics articles published between 1970 and 2018. *Journal of Medical Genetics*, 56(11), 734–740. doi: 10.1136/jmedgenet-2019-106137*



---

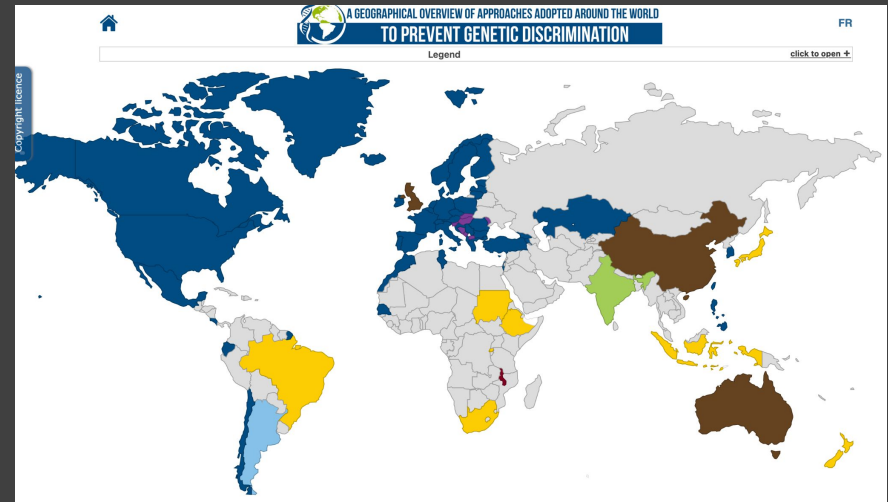
# Avoid Errors Before the Harm Occurs

1. Mistaken research design or analysis processes
  - Design for **reproducible and collaborative** research
2. Lack of diverse tools, methods and expertise to draw from
  - **Openly shared community-developed** data-handling tools
3. Inappropriate baseline comparison and assumption of future use
  - **Lack of technical understanding and ethical implications**

# Tracking discrimination in personalised medicine

*“No one shall be subjected to discrimination based on genetic characteristics”*

*Universal Declaration on the Human Genome & Human Rights, adopted by 77 nations in 1997*



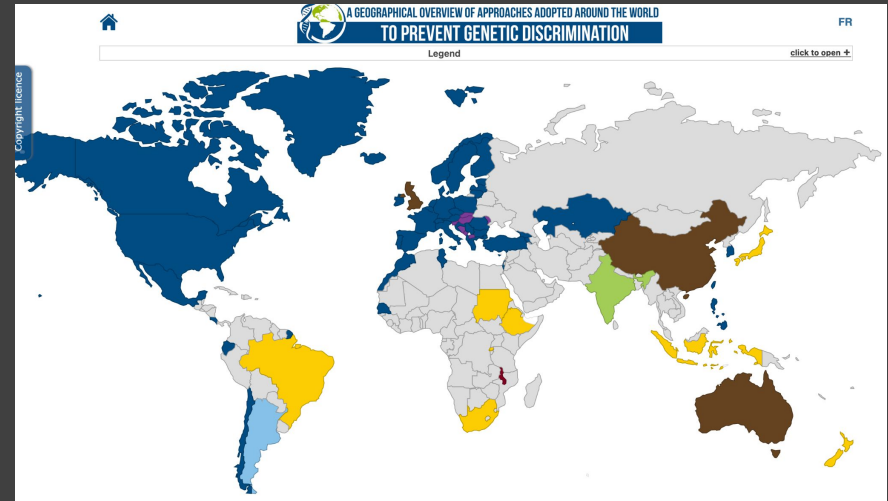
Genetic Discrimination Observatory. <https://gdo.global>  
Joly, Y., Salman, S., Ngueng Feze, I., Granados Moreno, P., Stanton-Jean, M., Lacey, J., ...Love, R. (2017). DNA Testing for Family Reunification in Canada: Points to Consider. *Journal of International Migration and Integration*, 18(2), 391–404. doi: 10.1007/s12134-016-0496-7,

# Tracking discrimination in personalised medicine

*“No one shall be subjected to discrimination based on genetic characteristics”*

*Universal Declaration on the Human Genome & Human Rights, adopted by 77 nations in 1997*

*There have been concerns on the increasing use of DNA tests in family reunification processes of nationals in Canada from certain regions including Africa, Asia, and Latin America. [1]*



Genetic Discrimination Observatory. <https://gdo.global>

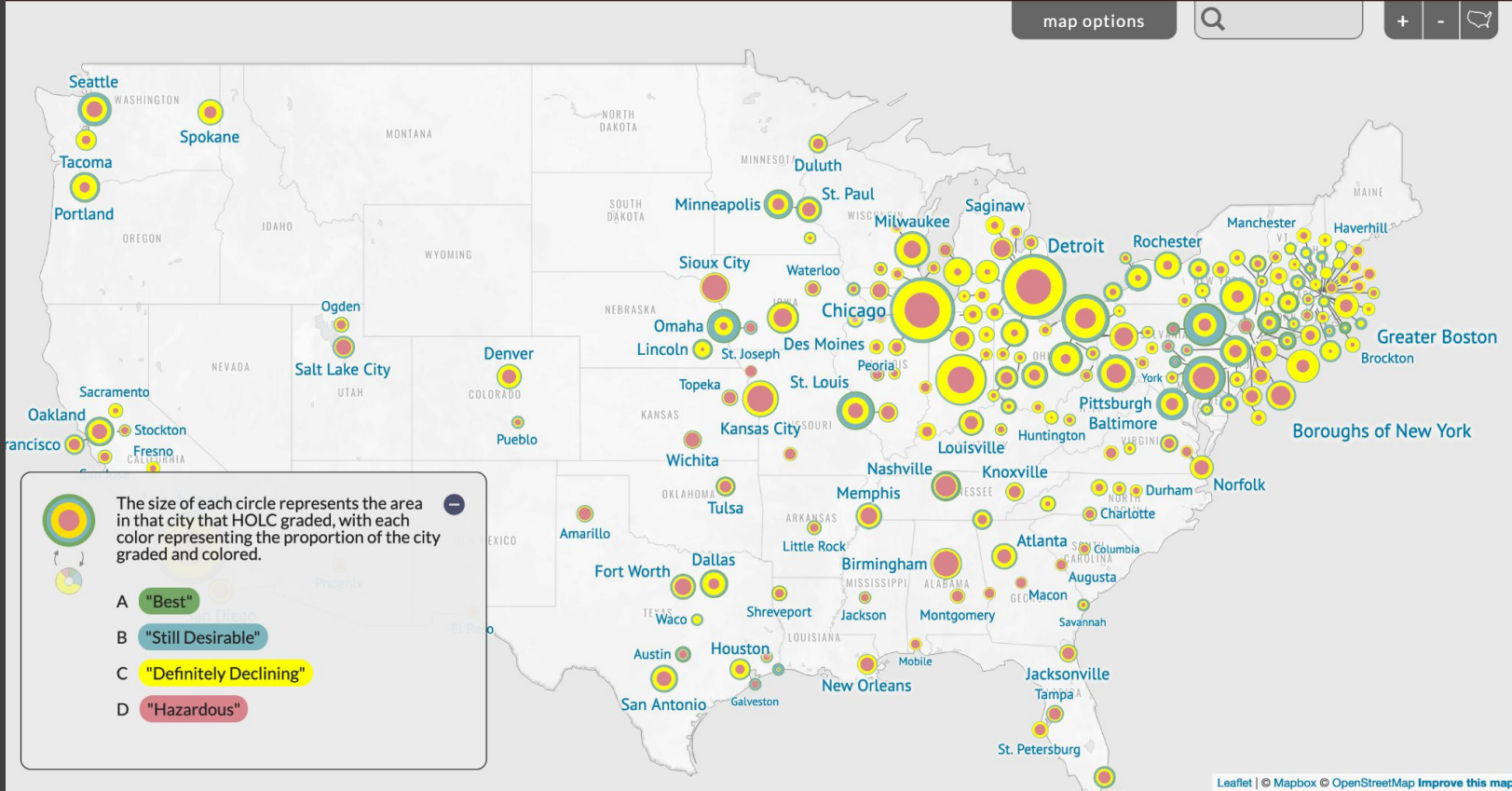
[1] Joly, Y., Salman, S., Ngueng Feze, I., Granados Moreno, P., Stanton-Jean, M., Lacey, J., ...Love, R. (2017). DNA Testing for Family Reunification in Canada: Points to Consider. *Journal of International Migration and Integration*, 18(2), 391–404. doi: 10.1007/s12134-016-0496-7,

---

# Tracking discrimination in personalised medicine

*Citizenship and Immigration Canada has adopted a policy of suggesting DNA testing only as a last resort [...], in practice, there have been concerns on the increasing use of DNA tests in family reunification processes of nationals from certain regions including Africa, Asia, and Latin America. [1]*

*Increasing availability of genetic information and the patenting of human genes may lead, respectively, to a loss of reproductive autonomy and a reduction in equitable access to medical care; hence new regulations and/or legislation may be required to ensure appropriate control over genetic information and use of intellectual property rights in human genes.*



The term 'redlining' was coined by the sociologist John McKnight in the 1960s, to define the discriminatory practice of avoiding investment in communities with unfavourable or high-risk demographics, typically with large minority and immigrant populations. Interactive map: <https://dsl.richmond.edu/panorama/redlining/#loc=5/39.1/-94.58>



Formerly redlined areas have less **tree cover** today than areas that weren't redlined.

0% Percentage tree cover 100%



Redlining was outlawed by the 1970s - but the practice has left lasting marks.

Low access to housing finance, less investment, more paved surfaces, reduced tree cover → influencing extreme heat and worse impact of climate change.

They have more **paved surfaces**, like roads and parking lots, that absorb and radiate heat.

# *The Turing Way*

---



# The Turing Way

---

An Open Science and community lead guide on Data Science.

We involve and support a **diverse community** to make research **reproducible, ethical, and collaborative** for everyone.





# The Alan Turing Institute

The national  

---

institute for data  
science and  
artificial  
intelligence



# Research Programmes

**Sector specific  
research**



**Finance and economics**



**Defence & security**



**Data science  
for science**



**AI**



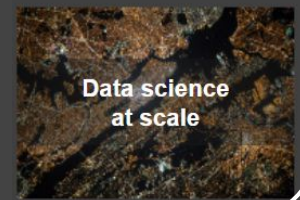
**Health and  
medical sciences**



**Data-centric  
engineering**



**Urban  
analytics**



**Data science  
at scale**

**Integration of  
best practices**



**Tools, practices and  
systems**



**Public policy**

---

# Tools, Practices and Systems

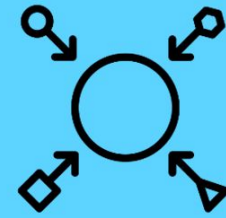
- Trustworthy systems
- Transparent reporting
- Inclusive interoperable design
- Ethical integrity
- Respectful co-creation
- Leadership in open research



Trust



Transparency



Inclusivity



Integrity



Respect



Leadership

# The Turing Way Book on Reproducibility



Kirstie Whitaker

Director: Tools, Practices & Systems Programme (TPS)



Malvika Sharan

TPS Senior Researcher:  
Open Research & Community

*The Turing Way* is a lightly opinionated guide to reproducible data science.

Our goal is to provide all the information that researchers need at the start of their projects to ensure that they are easy to reproduce at the end.

This also means making sure PhD students, postdocs, PIs, and funding teams know which parts of the "responsibility of reproducibility" they can affect, and what they should do to nudge data science to being more efficient, effective, and understandable.



---

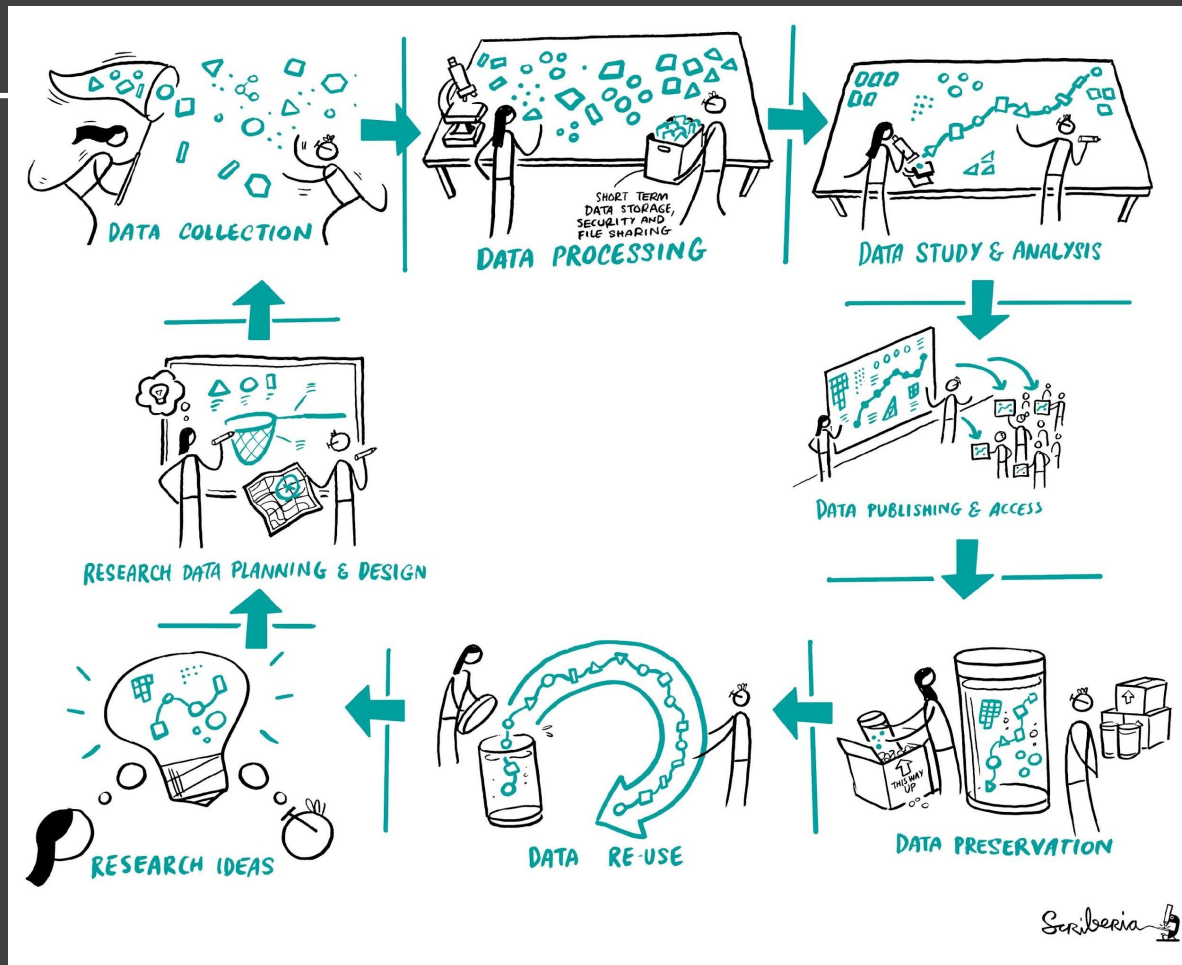
# Reproducibility

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable



# Reproducibility

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable



Scriberia

---

# Why do reproducible research?



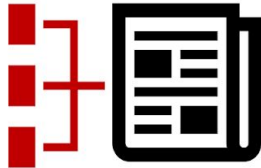
Track Project History



Collaborate & Review



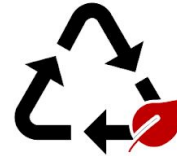
Avoid Misinformation



Write Paper Efficiently



Get Credits Fairly



Ensure Continuity

Is not considered  
for promotion

Held to higher  
standards than  
others

Publication bias  
towards novel  
findings

# Barriers to reproducible research

Requires  
additional  
skills

Plead the 5th

Support additional  
users

Takes time



Held to higher standards than

Is not co  
for pro

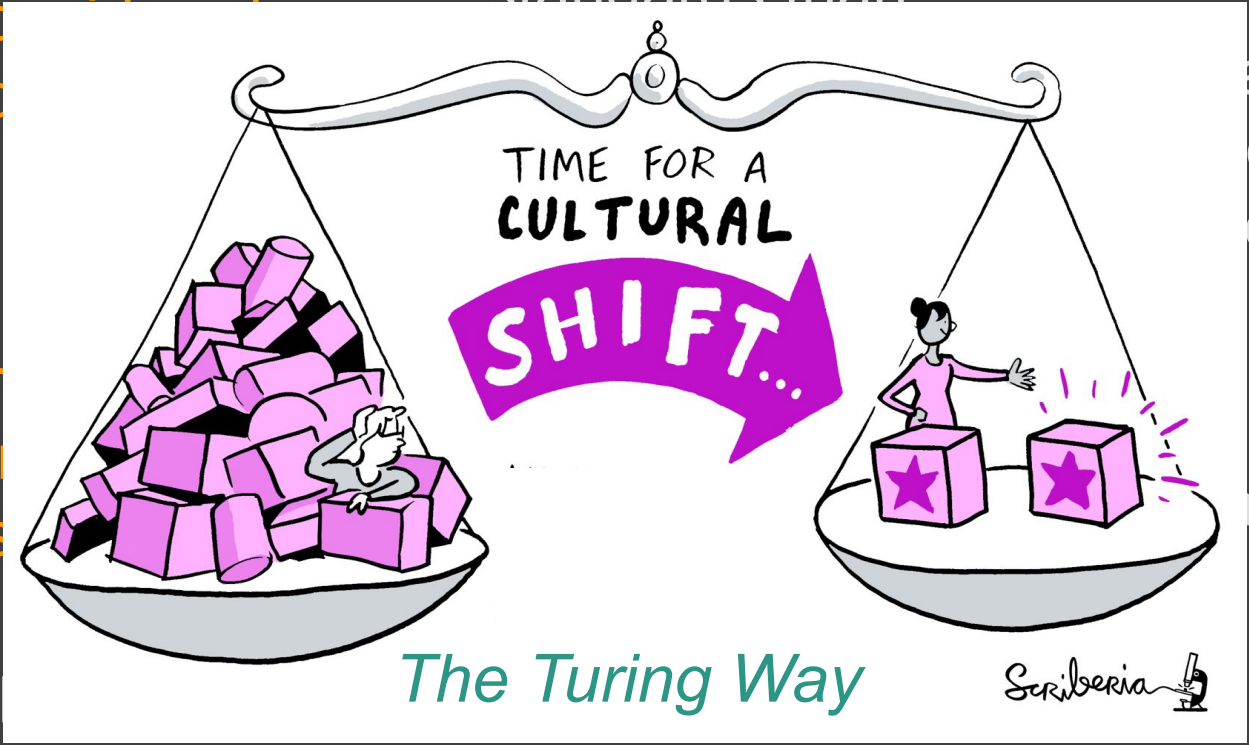
ation bias  
ds novel  
dings

Require  
addition  
skills

and the 5th

Sup

users



*The Turing Way*

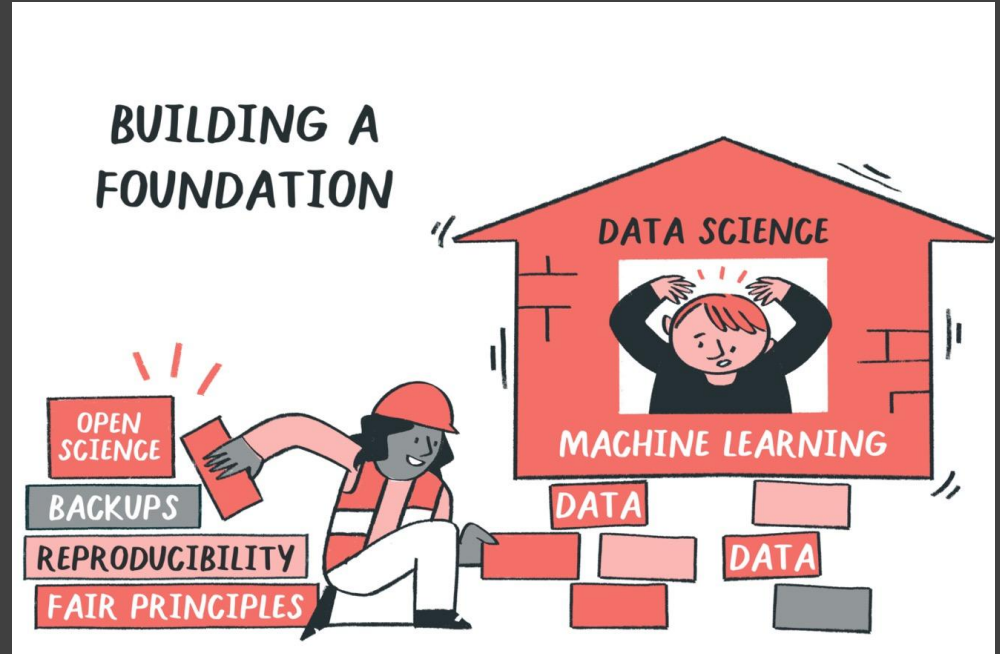
Scriberia 

---

## Open Science as foundational skill-set

Purpose of openness is to ensure **reproducible, ethical and collaborative research** process and outcome.

*Not all data can or should be published.*



*“An article about computational science in **a scientific publication** is not the scholarship itself, it **is merely advertising of the scholarship**. The actual scholarship is the **complete software development environment and the complete set of instructions which generated the figures.**”*

— *Buckheit and Donoho (paraphrasing John Claerbout)*

*WaveLab and Reproducible Research, 1995*

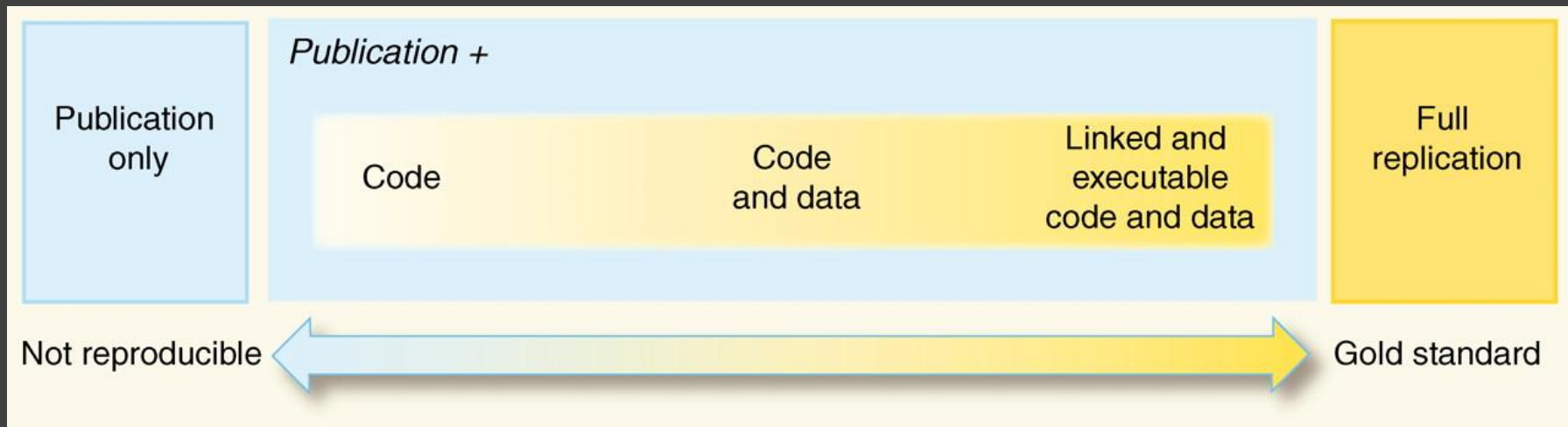
---

## Take home message

Sharing your code and data isn't enough



# We need research environment and guidelines too!



# We need research environment and guidelines too!



Peng, 2011, doi: [10.1126/science.1213847](https://doi.org/10.1126/science.1213847)

@turingway, CC-BY 4.0, The Turing Way, Slides' DOI: [10.5281/zenodo.7428708](https://doi.org/10.5281/zenodo.7428708)

## Guide for Reproducible Research

Overview

Open Research

Version Control

Licensing

Research Data Management

Reproducible Environments

BinderHub

Code quality

Code Testing

Code Reviewing Process

Reusable Code

Continuous Integration (CI)

Reproducible Research with

Make

Research Compendia



**Fig. 3** The Turing Way project illustration by Scriberia. Used under a CC-BY 4.0 licence. DOI: [10.5281/zenodo.3332807](https://doi.org/10.5281/zenodo.3332807).

# Computational Reproducibility

- Track changes to your code/resource (version control)
- Write clean, understandable and error free code
- Save and share your workflow and computational environment
- Make your code open for others to test, use and collaborate on



# Projects are for people!

*Publicly accessible research that anyone can (re)use, modify, and distribute for any purpose.*



---

# Open Science

Everyone can **freely access, reuse, reproduce, build upon and distribute** research objects.



---

# Open Science

Everyone can **freely access, reuse, reproduce, build upon and distribute** research objects.

Community framework for **collaboration, peer production and sustainability.**



elly, Anthony, "With Head, Heart and Hand: Dimensions of Community Building" (Boolarong Press) ISBN 978-0-86439-076-9,

Wenger, McDermott, Snyder, A Guide to Managing Knowledge: Cultivating Communities of Practice. HBS (2002)

@malvikasharan, Image by Scriberia for The Turing Way, Shared under CC-BY 4.0, DOI: 10.5281/zenodo.7428708

Open Science aims to transform research by making it more  
transparent → accessible → reliable → reproducible →  
reusable → collaborative → beneficial to society.

To achieve this openness in research, we need to make  
each element of the research process ...

- *publicly available,*
- *with permission to view, use, modify and distribute, and*
- *description for how one can collaborate.*

# Truly Open Licenses: common elements



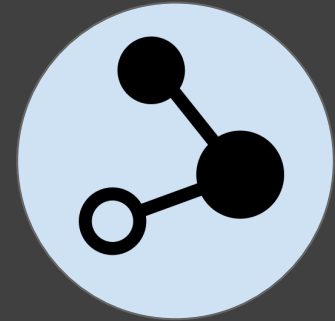
use

Anyone can use the work for any purpose



modify

Anyone can modify the work



share

Anyone can redistribute both the original and modified work

*Open source software is software that can be freely used, modified, and shared (in both modified and unmodified form) by anyone.*

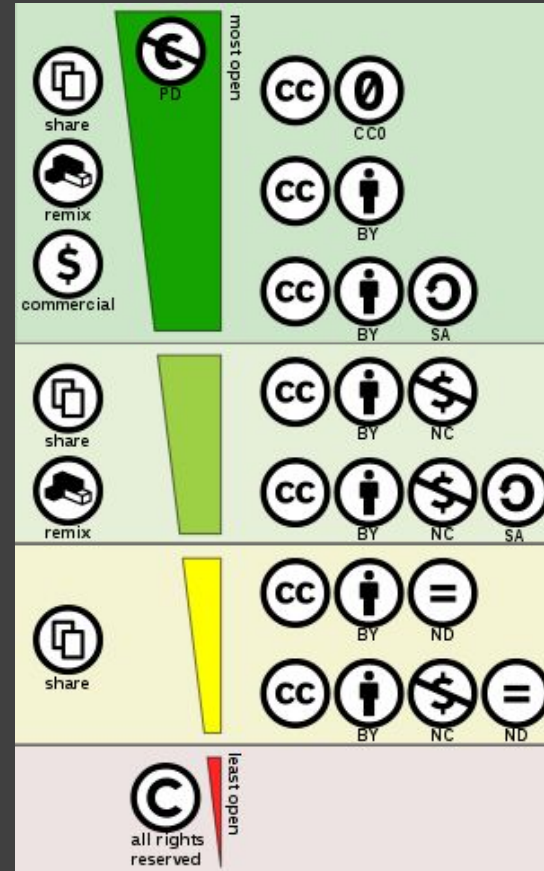
- [GitHub Glossary, Open Source](#)

@openlifesci, CC-BY 4.0, The Turing Way, DOI: 10.5281/zenodo.7428708

# Assignment: Add a License

Use a **different license** for content, code and data. For example

- Writing/docs/images/metadata: **CC-BY**
- Code: **MIT License**
- Data will also have different license based on what you can/can't share!





*UNESCO regards science as **a human right** -- the greatest **collective endeavour** -- that has significant **value as a common good**.*





*UNESCO regards science as **a human right** -- the greatest **collective endeavour** -- that has **significant value as a common good**.*

... aims to **turn right to Open Science into duty** to be exercised by researchers, institutions, educators, publishers, employers, libraries, and policymakers, facilitated by Member States.

Open Science

Open Research

Open Scholarship

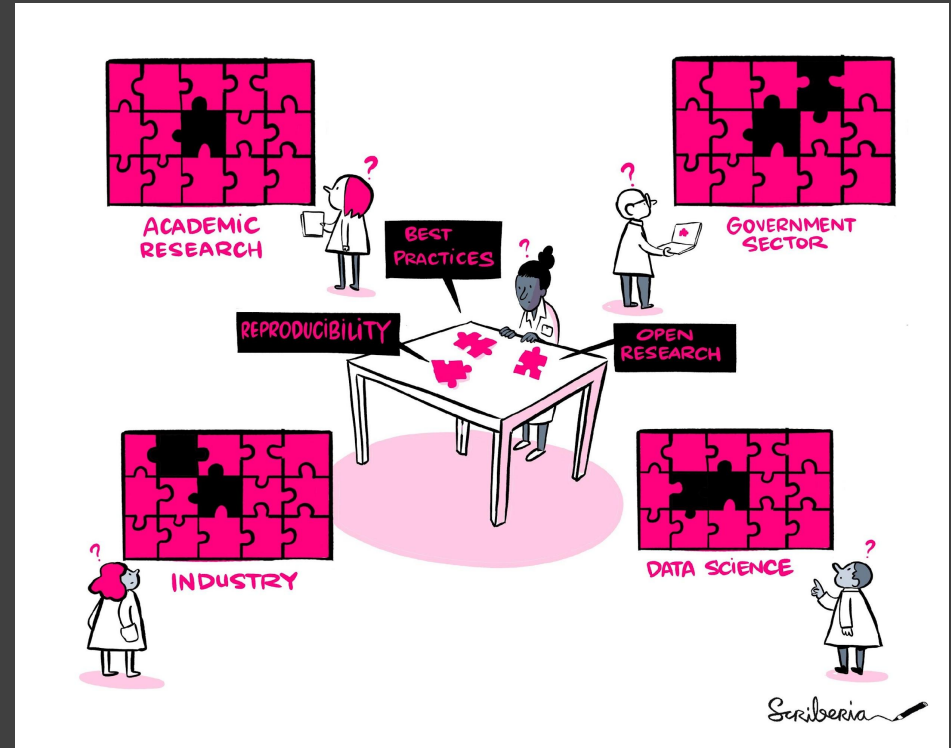


We often use these terms interchangeably - all mean the same open practices and considerations for you!

# Integrating open science

Open Science means different things in different contexts, and hence, there is no fixed “best practice”.

Many barriers to openness.

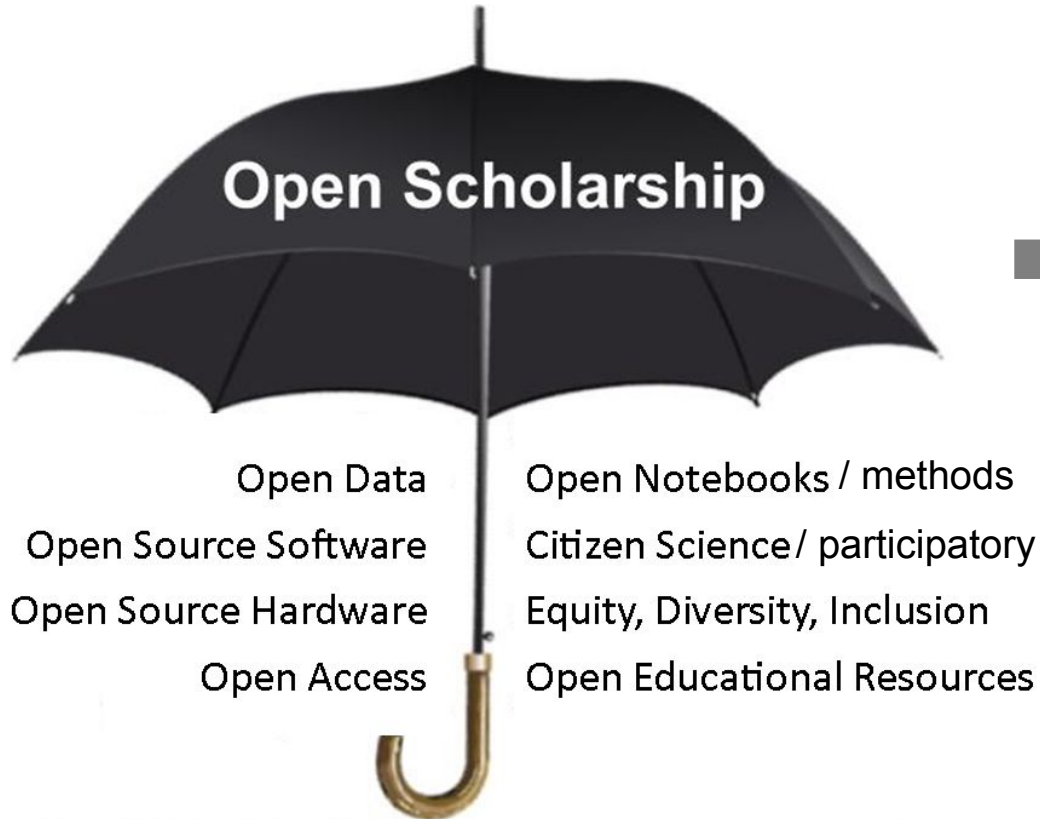


@malvikasharan, Image by Scriberia for The Turing Way, Shared under CC-BY 4.0, DOI: 10.5281/zenodo.6587260

Ten arguments against Open Science that you can win. (2022, May 26).

<https://www.software.ac.uk/blog/2020-12-17-ten-arguments-against-open-science-you-can-win>

# What are some barriers we should work to remove?



Transparency →  
Reproducibility →  
Research Quality →  
Sustainability

+

Collaboration →  
Inclusive Research →  
Equity and Diversity →  
Global Accessibility

Applying best practices in our research requires intention, resources, time and collaboration, which can be overwhelming.





**Book:**  
[the-turing-way.netlify.app/](https://the-turing-way.netlify.app/)

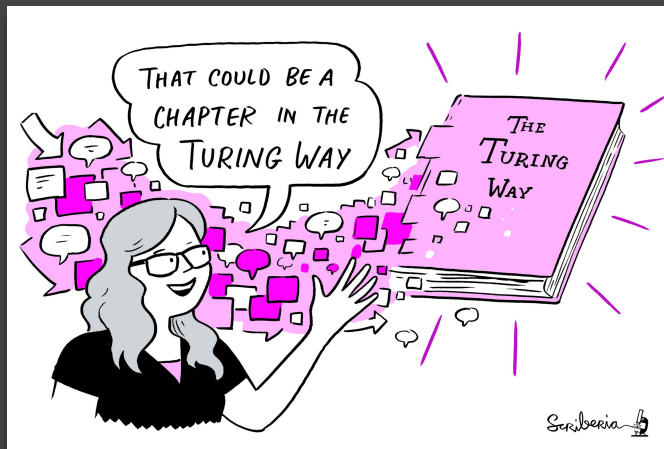
**GitHub:**  
[github.com/alan-turing-institute/the-turing-way](https://github.com/alan-turing-institute/the-turing-way)

**Twitter:**  
[twitter.com/turingway](https://twitter.com/turingway)

**Email:**  
[theturingway@gmail.com](mailto:theturingway@gmail.com)

CC-BY 4.0, *The Turing Way*

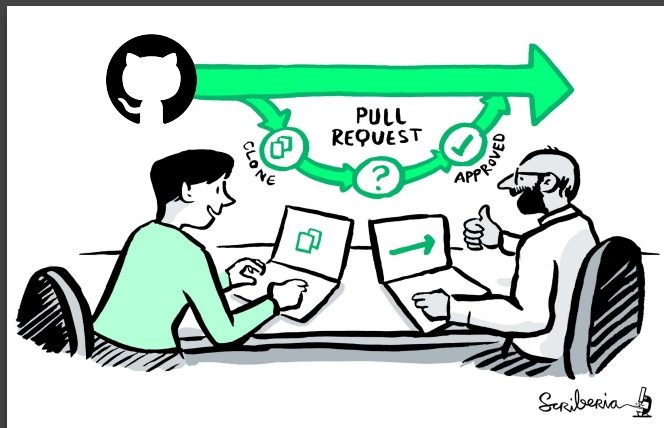
## A Book



## A Community



## An Open Source Project



## A Culture of Collaboration





# A book: “Work in Progress”



## The Turing Way

🔍 Search this book...

### Welcome

Guide for Reproducible Research

Guide for Project Design

Guide for Communication

Guide for Collaboration

Guide for Ethical Research

Community Handbook

Afterword

Visit our [GitHub Repository](#)

This book is powered by [Jupyter Book](#)



Scriberia



## Overview of Project Design

Planning for Project Design

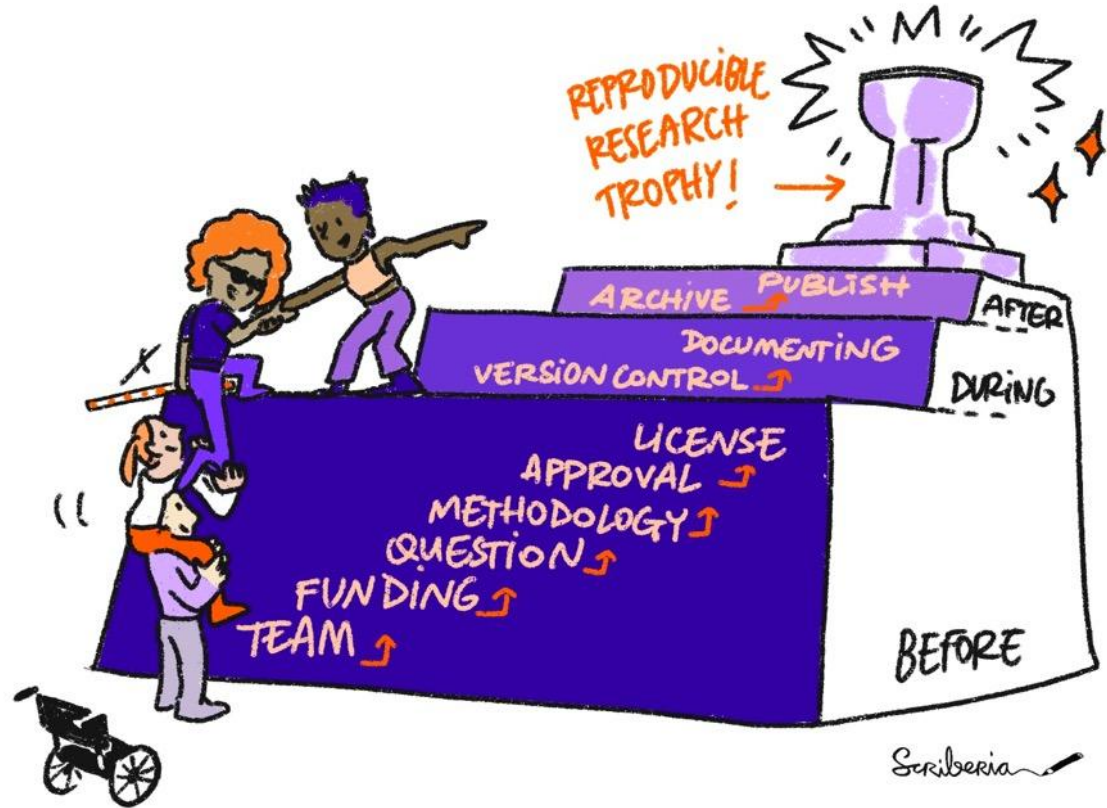
Communication and  
Collaboration

Reproducibility Methods

Version Control and  
Documentation

Sharing Your Research  
Work

Creating Project Repositories



# Guide for Communication

---



## Guide for Communication

Overview of Guide for Communication

Blogs for Research Communication

Lay Summaries

Podcasts for Research Communication

Presenting Posters and Conference Talks

Social Media for Research Communications

Making Research Objects Citable

Communications in Open Source Projects

# Guide for Collaboration

## Guide for Collaboration

Getting Started With GitHub

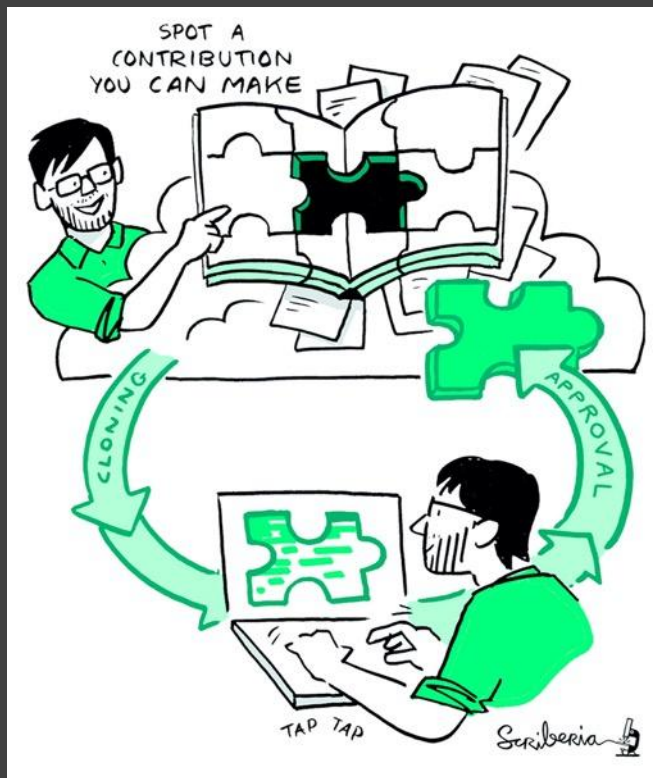
Maintainers and Reviewers on  
GitHub

Managing a New Community  
and Team

Leadership in Data Science

Remote Collaboration

Shared Ownership in Open  
Source Projects



# Guide for Ethical Research

## Guide for Ethical Research

Introduction to Research Ethics

Research Ethics Committees

Workflows

Ethical Decisions in Preclinical  
Research

Law, Policy and Human Rights  
in Ethics

Activism for Researchers

Internal Policy Advocacy





# Ethical and Collaborative Research

Respect - Protect - Fulfil

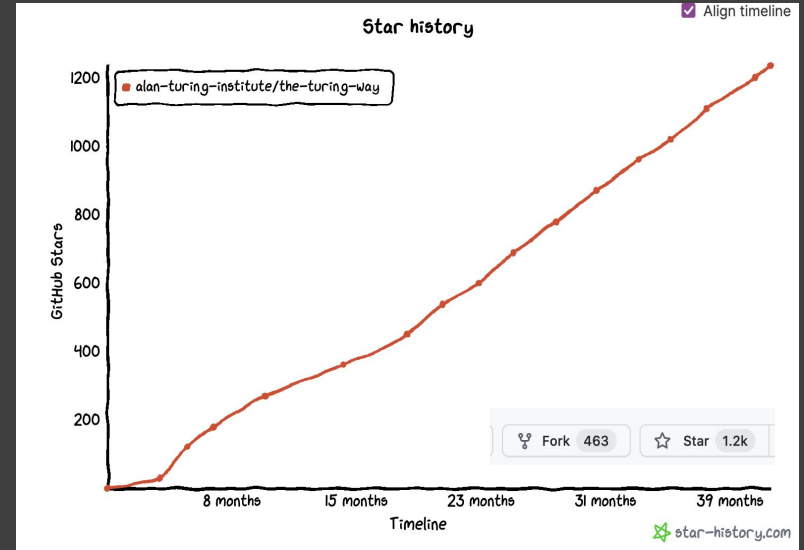


1. Project Design
2. Data and Methodology
3. Equity, diversity, accessibility
4. Consent, privacy and security
5. Further Societal Consequences

Factor	RRI	Research Integrity
<b>Perception of ethical questions</b>	Outward (for example: how does the public perceive scientific output?)	Inward (for example: how do researchers conduct themselves?)
<b>Who researchers interact with</b>	With research subjects (for example: do we treat them with respect?)	With fellow researchers (for example: do we adhere to a code of conduct?)
<b>What shapes research</b>	Shaped by society (for example: how is AI as a solution generally understood?)	Shaped by institutional norms (for example: what research questions are being funded?)

# Project and community growth

- >3 years, >260 Live Chapters
- Community resources, events, guidance, templates, training
- 400+ direct GitHub contributors and thousands of users



<https://zenodo.org/record/3332807>

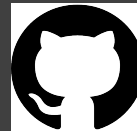


13,940

views

13,856

downloads



# Qualitative/notable impacts

---

Resources are being used by learners, educators, community builders, policy makers and researchers globally

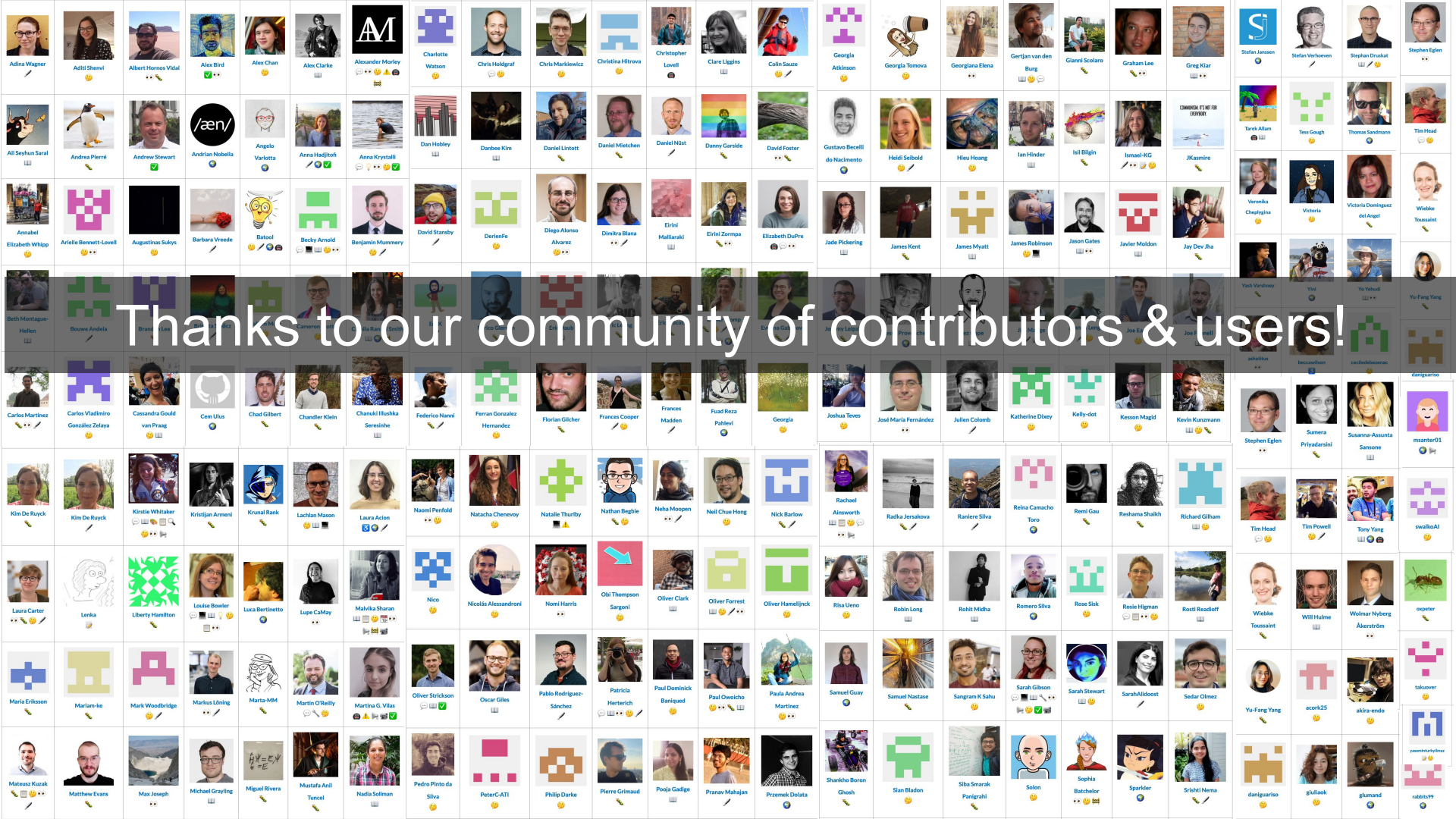
- Reproducibility of scientific results in the [EU 2020 \(report\)](#)
- An Emerging Technology Charter by [Mayor of London \(policy\)](#)
- Innovation Scholars: [UKRI grant 2020](#) (funding call)
- NASA TOPS, CodeRefinery and Library Carpentries [training materials](#)
- Recommended in [Goldacre Review for modern open working](#) into health data
- Cited by [30+ peer-reviewed articles](#) & 100+ online publications
- [Highly commended](#) by HiddenRef and Royal Society Athena Awards



# Evidence of successful replication or extension

---

- [Quality Assurance of Code for Analysis and Research](#), Office for National Statistics , UK
- [Turing Data Stories](#), [Turing Commons](#) and [The Environmental Data Science book](#), The Alan Turing Institute
- [FAIR Cookbook](#) by researchers at University of Oxford, Novartis, FAIRplus Consortia
- [UCL Institute of Health Informatics Coding Club Handbook](#), University College London
- [A Citizen Science Guide for Research Libraries](#), LIBER Citizen Science Working Group



Thanks to our community of contributors & users!

# Community Support for Open Collaboration

## Community Handbook

Code of Conduct

Style Guide

Maintaining Consistency

Contributing and Developing

Chapters - Templates and

Workflow

Acknowledging Contributors

Monthly Newsletters

Book Dash Events

Community Coworking Calls

Fireside Chat Series

Giving a Turing Way Talk

Template Collection

Afterword

Thank you for joining the incredible journey of *The Turing Way*!



# Addressing Common Concerns

- Sharing something on openly does **NOT** automatically make it reusable
- Work shared with an open license **NOT** to be used without attribution
- Sharing with a license does **NOT** give away your rights to publish or sell
- Reproducibility for you does **NOT** mean the reproducibility for others
- Open science does **NOT** mean publishing sensitive and private data
- Not sharing data openly does **NOT** remove your responsibility for reproducibility

# Reproducibility

- Is my code/data **correct**?
- Can others **read and test** it?
- Is my **workflow robust**?
- Have I provided **guidance**?
- Is my work **citable**?

*Reproducibility should ensure higher scientific standards allowing others to test and reuse your work ...*

# Open Science

- Is my code/data **freely available**?
- Can others **modify and share** it?
- Is my **workflow reusable**?
- Have I provided **permission**?
- Is my work **open for collaboration**?

*... and Open Science should allow anyone to reuse, report errors, fix issues, build on and collaborate*

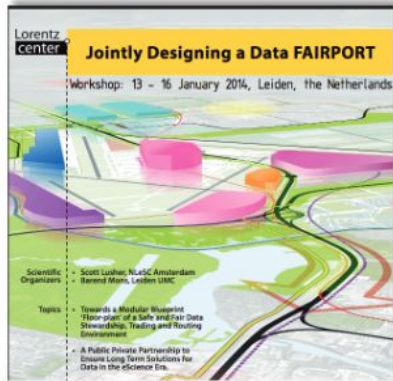


So many things to consider: *Where can I start?*

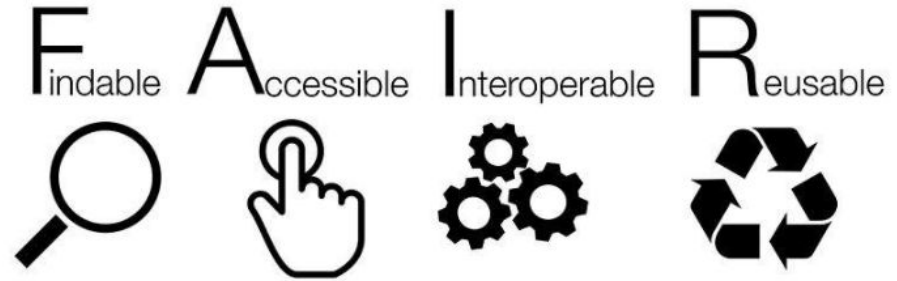
## FAIR principles

- Findable,
- Accessible
- Interoperable
- Reusable





2014



2016

SCIENTIFIC DATA 

SPRINGER NATURE

OPEN

**Comment: The FAIR Guiding Principles for scientific data management and stewardship**

SUBJECT CATEGORIES

- » Research data
- » Publication characteristics

Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, Jildau Bouwman, Anthony J Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J G Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A.C. 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Momis A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons

SCIENTIFIC DATA | 3:160018 | DOI: 10.1038/sdata.2016.18

A set of principles to enhance the value of all digital resources

Developed and endorsed by *researchers, service providers, publishers, funding agencies and industry partners*

FAIR principles from **Wilkinson *et al.* (2016)**  
**DOI: 10.1038/sdata.2016.18**



# FAIR data analogy: *You would not buy food with no labels!*



Annotation makes it easier to find important things



FAIR doesn't require data to be open, but needs Metadata information along with detailed research process.

Metadata: “data descriptors” that facilitate cataloguing data and data discovery



Adapted from talk by Philippe Rocca-Serra (2020)

# Open and reproducibility enable ethical research

- Reproducibility can be facilitated by open, but **open is a choice**
- Best practices should be considered at all stages for ethical integrity
- As open as possible, **as closed as necessary**
  - Open principles should be applied when you can
  - **NEVER** for private, confidential or sensitive data (ethically)
- Start with FAIR (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable)

---

## Meeting you where you are!

- i. Join the community
- ii. Learn a new skill
- iii. Share your skills
- iv. Collaborate with others
- v. Mentor others' contributions
- vi. Represent this community

*We value your participation!*

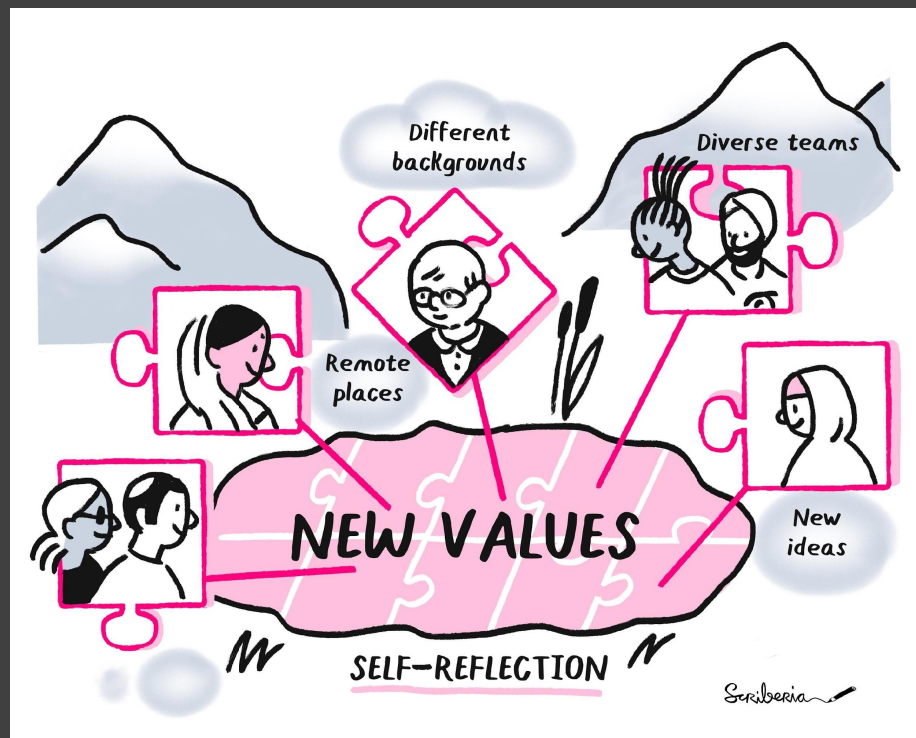


---

# The Turing Way Events

- *Collaboration Cafes*
  - 1<sup>st</sup> & 3<sup>rd</sup> Wednesdays (14:00 UTC)
- *Coworking Calls*
  - Mondays (10:00 UTC)
- *Fireside Chats*
  - Monthly on Friday
- *More ways to connect*

[bit.ly/turingway](https://bit.ly/turingway)



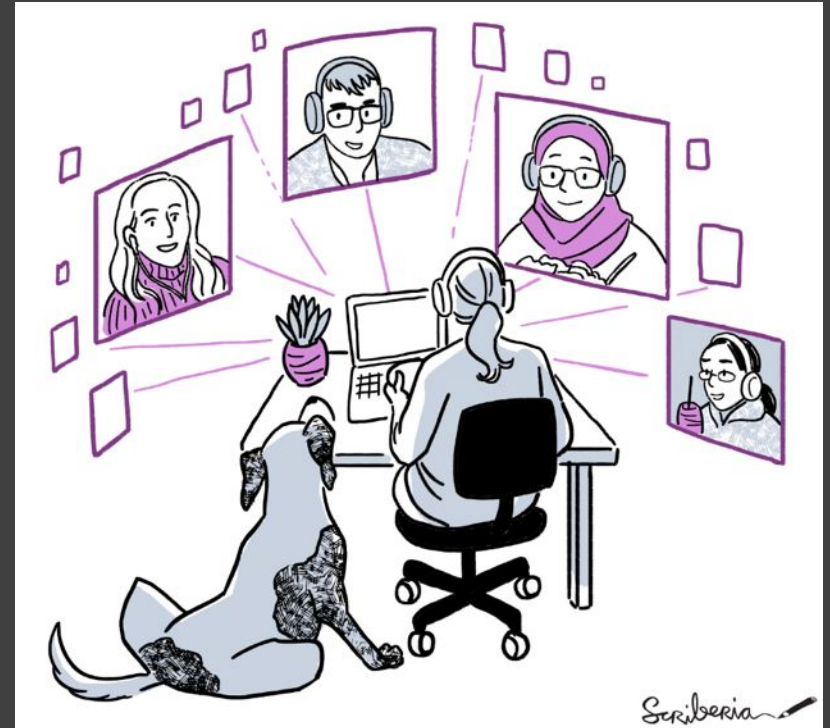
---

## Attend a Book Dash

*Virtual week-long Hackathon like event to work with others on The Turing Way: **14 - 18 November 2022***

Attend the community share out tomorrow at 11:00 and 17:00 UTC

[bit.ly/turingway](https://bit.ly/turingway)



---

# Reflection Exercise

What are the some technical or ethical concerns you have about your work?

- Who are impacted by your work?
- What resources or expertise you need but don't have access to?
- What resources and expertise other need to use your work?
- What tools and practices you have yet to learn about?



---

# Acknowledgements & Links

The  
Alan Turing  
Institute

- The Turing Way community members
- **Book:** [the-turing-way.netlify.com](https://the-turing-way.netlify.com)
- **Twitter:** [twitter.com/turingway](https://twitter.com/turingway)
- **Mastodon:** [fosstodon.org/@turingway](https://fosstodon.org/@turingway)
- **GitHub:** [github.com/alan-turing-institute/the-turing-way](https://github.com/alan-turing-institute/the-turing-way)



Useful links & opportunities are listed here: <https://bit.ly/turingway>

Illustrations by Scriberia for The Turing Way community: <https://zenodo.org/record/3332807>



13,940	13,856
<small>👁 views</small>	<small>📄 downloads</small>