
The Turing Way: **A collaborative guide to data science and research**

Vicky Hellon & Anne Lee Steele



Introductions



Vicky Hellon
Community Manager
Turing-Roche Partnership



Anne Lee Steele
Community Manager
The Turing Way

slido



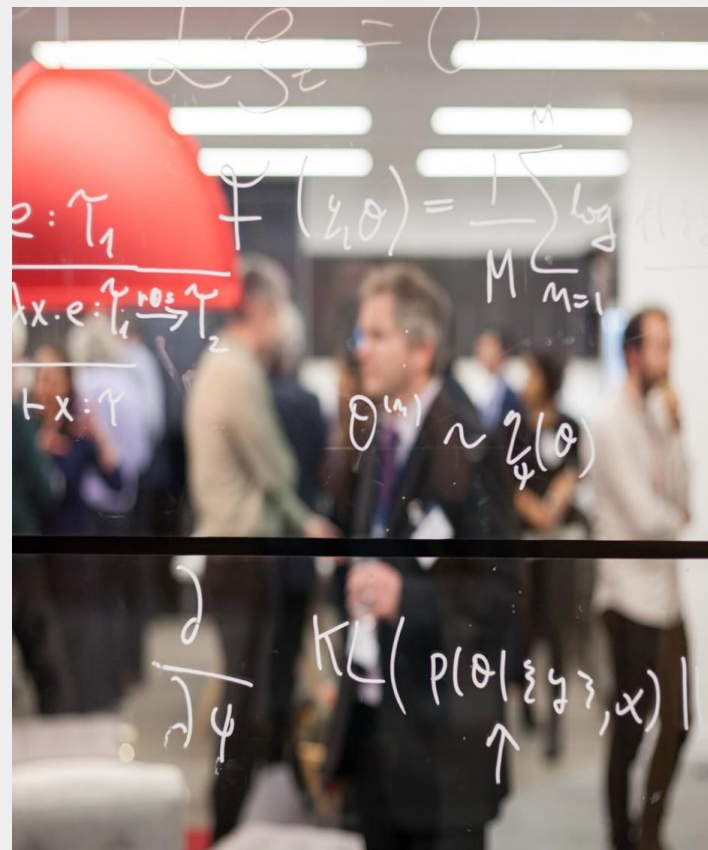
**How familiar are you with
the Turing Way?**

① Start presenting to display the poll results on this slide.

The Alan Turing Institute

Established in 2015 as the National Institute for Data Science and AI

- Advanced world-class research and apply it to real-world problems
- Train the leaders of the future
- Lead the public conversation



Collaboration



accenture

BILL &
MELINDA
GATES
foundation

GCHQ
[dstl]

Ministry
of Defence

LR Lloyd's Register
Foundation

NATS

Office for
National Statistics

Roche

Cardiff University
City, University of London
Creative Computing Institute, University of the Arts, London
Durham University
Goldsmiths, University of London
Imperial College London
Keele University
King's College London
London School of Hygiene and Tropical Medicine
Northumbria University
Nottingham Trent University
Queen's University Belfast
Royal Holloway, University of London
Royal Veterinary College
Teesside University
University of Glasgow
University of Liverpool
University of Nottingham
University of Plymouth
University of Reading
University of Sheffield
University of Strathclyde
University of Surrey
University of the West of England



Research Programmes

Sector Specific

Finance and economics

Defence & security

Data science
for science

Health and
medical sciences

Data-centric
engineering

Urban
analytics

Cross-Cutting

Tools, practices and
systems

Public policy

AI

Data science
at scale

Tools, Practices, Systems (TPS)

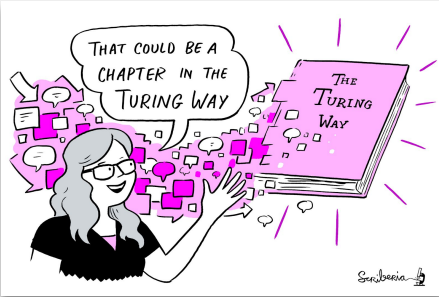
- Trustworthy systems
- Transparent reporting
- Inclusive interoperable design
- Ethical integrity
- Respectful co-creation
- Leadership in open research



The Turing Way

An Open Source guide on Data Science

We involve and support a diverse community to make data science reproducible, ethical, collaborative and inclusive for everyone



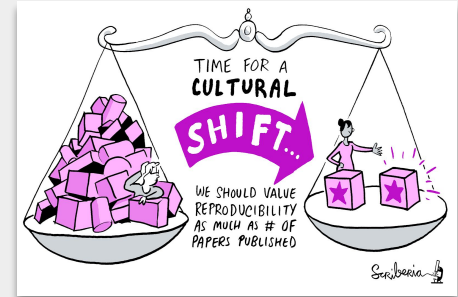
A Book



A Global Community



An Open Source Project



A Culture of Collaboration

slido

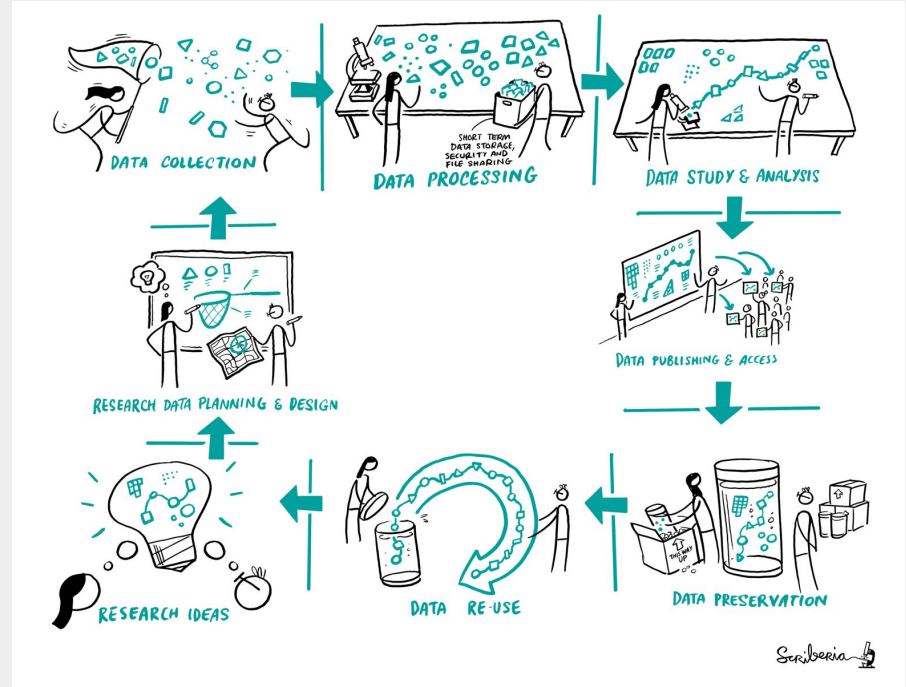


**What does you associate
with the term
reproducibility?**

① Start presenting to display the poll results on this slide.

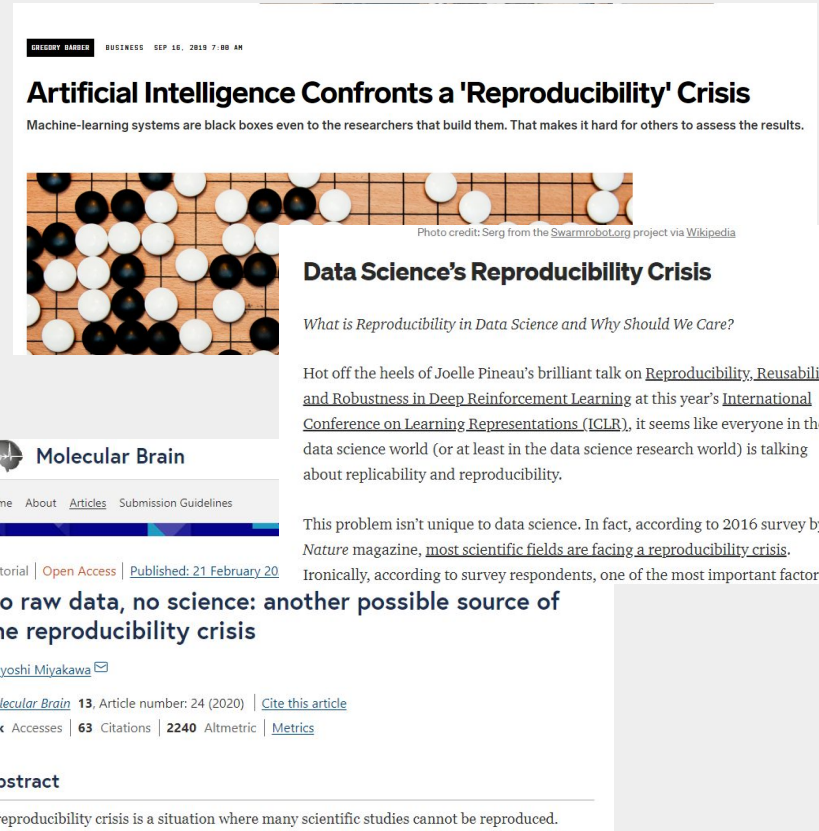
Reproducibility

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable



Reproducibility Crisis

- Many scientific studies are difficult or impossible to reproduce
- Often mentioned in regards to psychology and medicine but is pervasive across many scientific fields
- Range of causes: questionable research practices, 'publish or perish' culture, statistical issues
- Link with open science movement- push for openness and transparency



Turing Way beginnings: Reproducibility



Kirstie Whitaker
Lead of Tools,
Practices & Systems
Programme



Malvika Sharan
Senior Researcher,
Tools, Practices &
Systems

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.



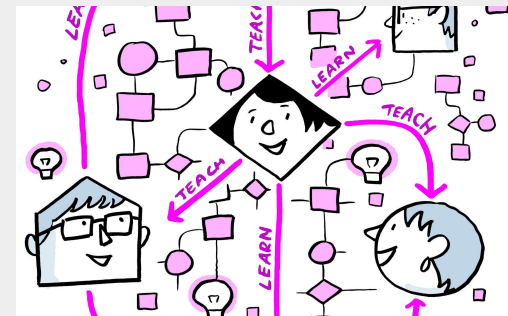
The Turing Way Guides



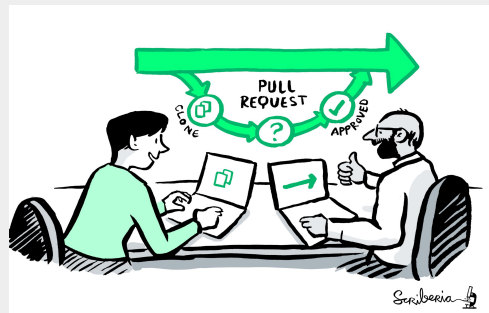
Reproducibility



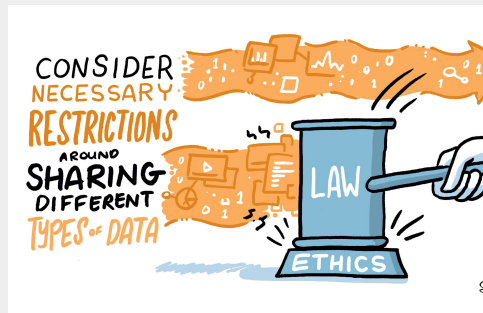
Project Design



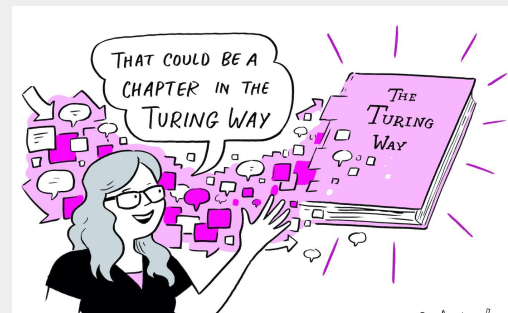
Communication



Collaboration



Ethical Research



Community Handbook

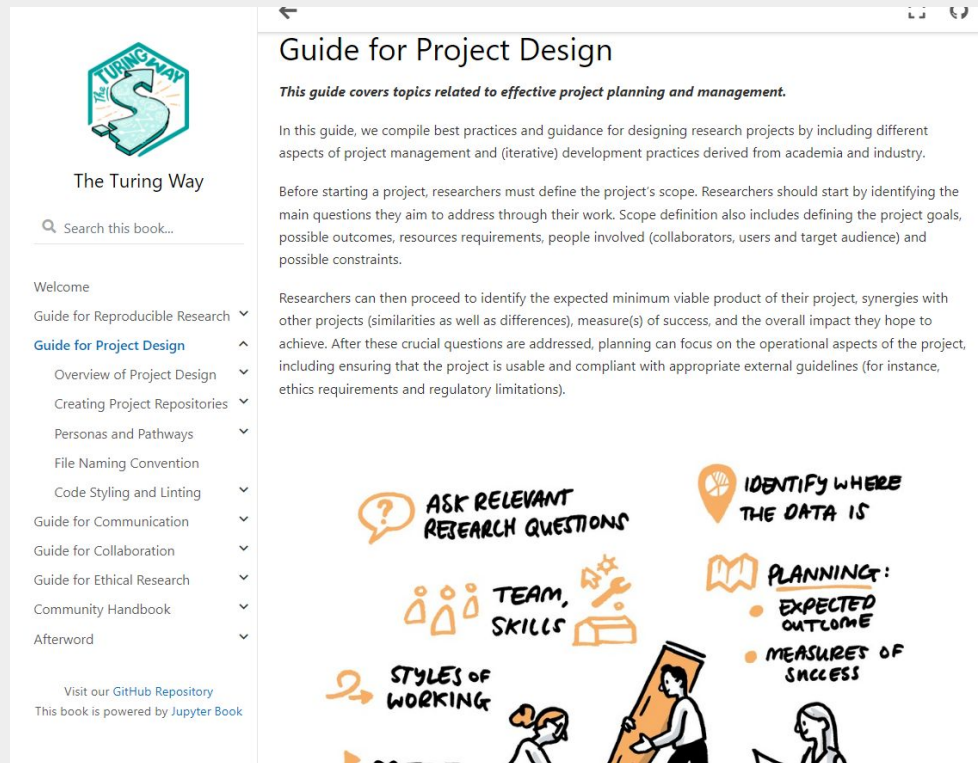
Guide for Reproducible Research

- Open Research- covering open access, open data, open source software, open scholarship and more
- Version Control- comprehensive guide on how to use Git
- Research Data Management
- Reproducible environments and code quality, testing and reviewing



Guide for Project Design

- Planning for Project Design and documentation
- Creating Project Repositories
- Personas and Pathways- considering who can contribute and how
- Code styling and 'hygiene' (linting)



Guide for Communication

- Open Scholarship and Education
- Blogs, Podcasts, Posters, Conference Talks, Social Media
- Making Research Outputs discoverable and publishable
- Authorship and Peer Review
- Communicating in Open Source Projects



Guide for Collaboration

- Getting started with GitHub, reviewing and merging contributions
- Managing a New Community and Team
- Leadership in Data Science
- Research Infrastructure Roles
- Remote Collaboration- organising remote meetings, tools for remote collaboration, managing distributed teams



The Turing Way

Search this book...

Welcome

- Guide for Reproducible Research
- Guide for Project Design
- Guide for Communication
- Guide for Collaboration**
- Getting Started With GitHub
- Maintainers and Reviewers on GitHub
- Managing a New Community and Team
- Leadership in Data Science
- Research Infrastructure Roles
- Remote Collaboration
- Shared Ownership in Open Source Projects
- Guide for Ethical Research

Guide for Collaboration

This guide covers topics related to effective and inclusive collaboration.

Data science is defined by its interdisciplinarity. Our work can only reach its highest potential if there are diverse teams of people involved in designing and delivering the research or product.



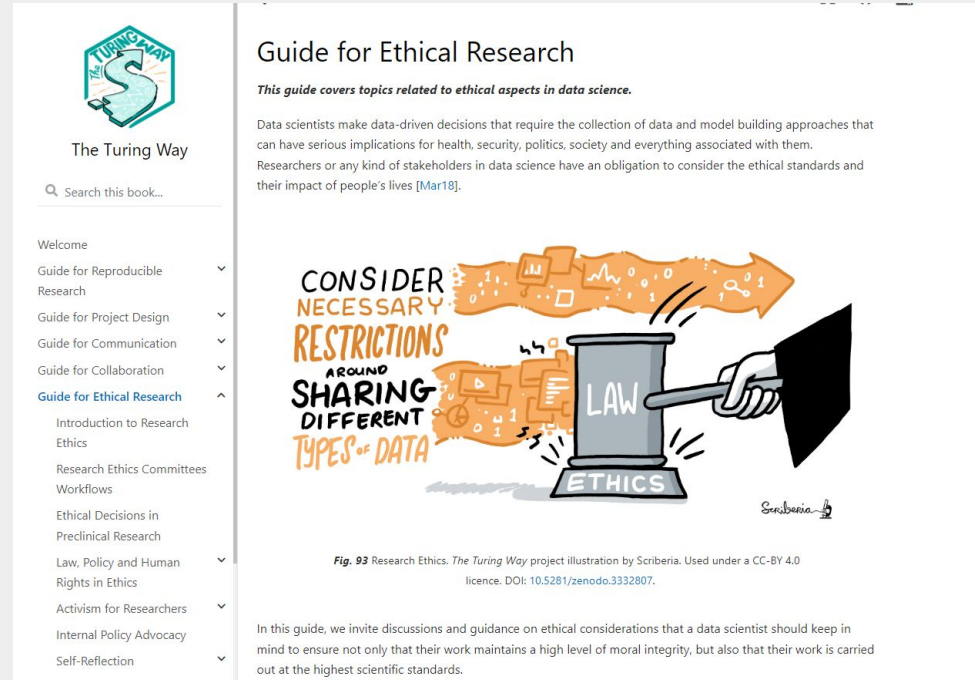
Fig. 80 There is more to collaboration than we see. *The Turing Way* project illustration by Scriberia. Used under a CC-BY 4.0 licence. DOI: 10.5281/zenodo.3332807.

There are many different skills required to work well in groups with a wide range of expertise. In this guide, we welcome contributions in developing guidance on following (but not limited to) topics:

- Designing a project that welcomes contributions
- Distributed collaboration on GitHub
- Reviewing team member's contributions
- Remote working
- Running an inclusive event

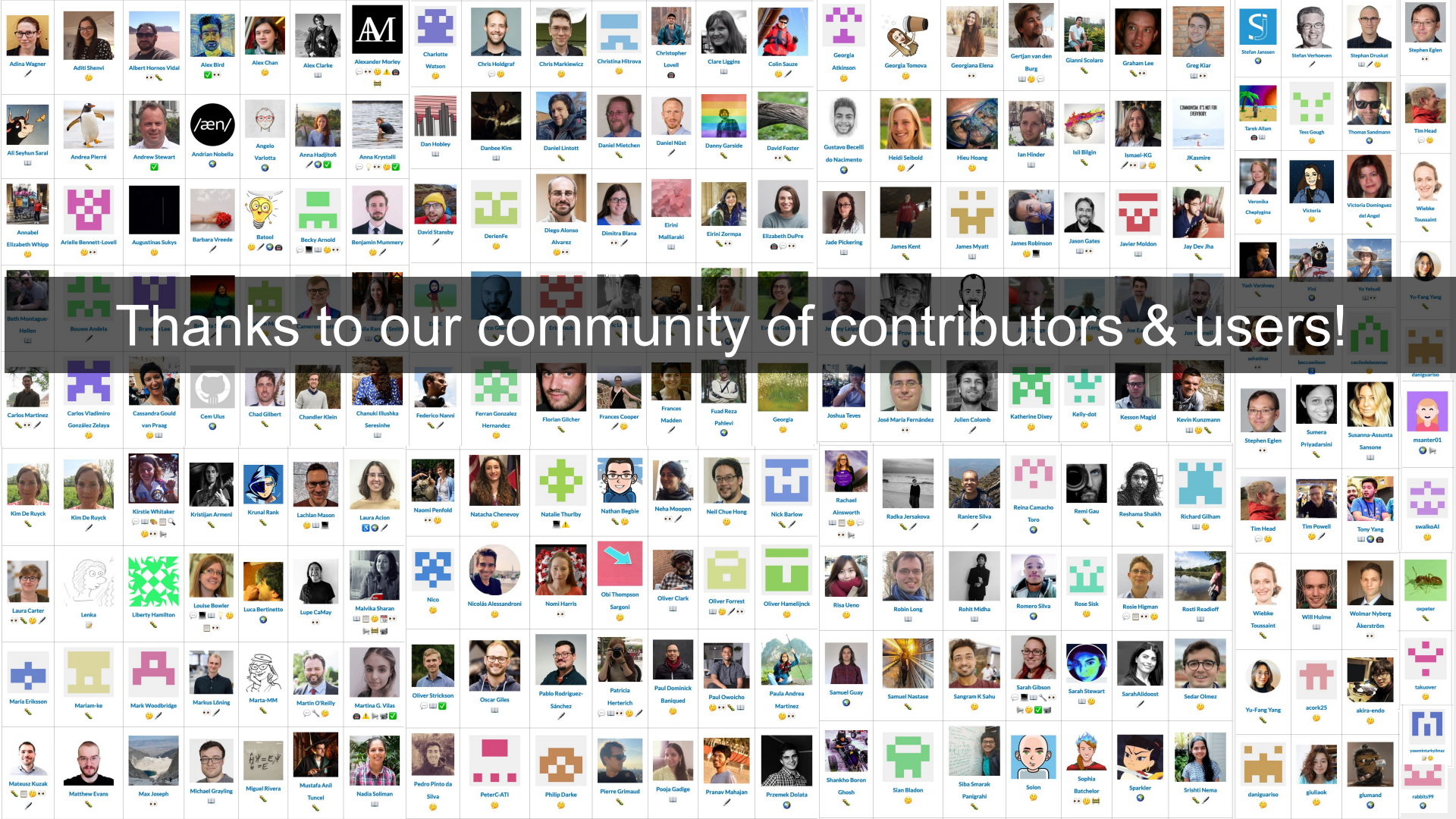
Guide for Ethical Research

- Research Ethics Committees
- Ethical Decisions in Preclinical Research
- Law, Policy and Human Rights in Ethics
- Activism for Researchers
- Self-Reflection



Tying it all together





Thanks to our community of contributors & users!

Open by Design, Collaborative in Practice



Using **open source tools and methods** to collaborate on best practices

Open framework for collaboration, peer production, and project sustainability

slido



**Have you ever used GitHub
or Jupyter Notebooks?**

① Start presenting to display the poll results on this slide.

Reproducible & Open Source Projects



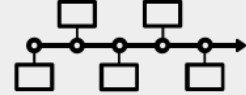
Online Repository



License/Permission



Documentation



Version Control



Dependencies
and data



Review & Verify



Report/Fix Bugs

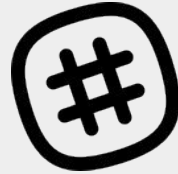


Release & Cite

Open Community Spaces



Github Repository



Slack



Coworking Calls



Collaboration
Cafes



Book Dash



Community-Share
Outs



Fireside Chats



Social Media
Platforms

Remote and In-person events



Collaboration Cafes and
weekly coworking



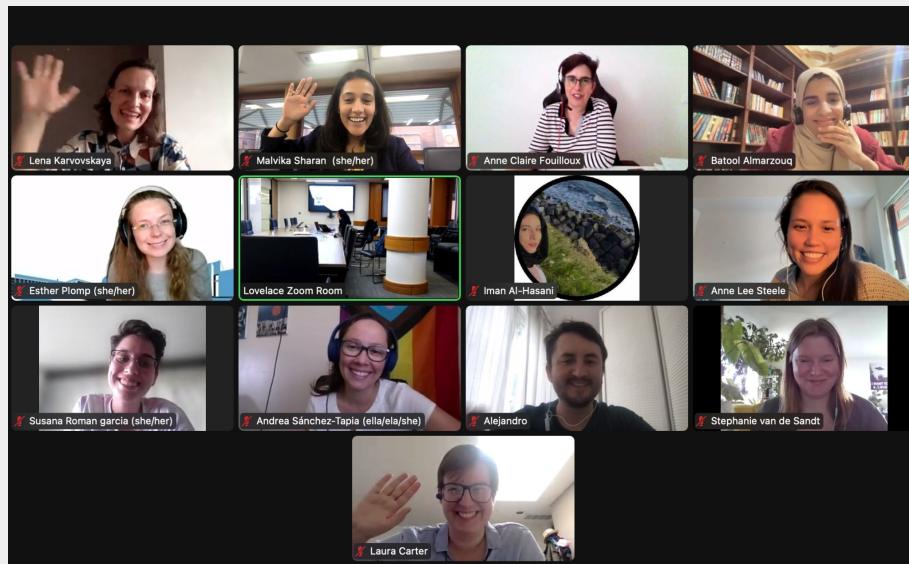
Book Dash Events
(in-person & virtual)

Piloting hybrid-hub events



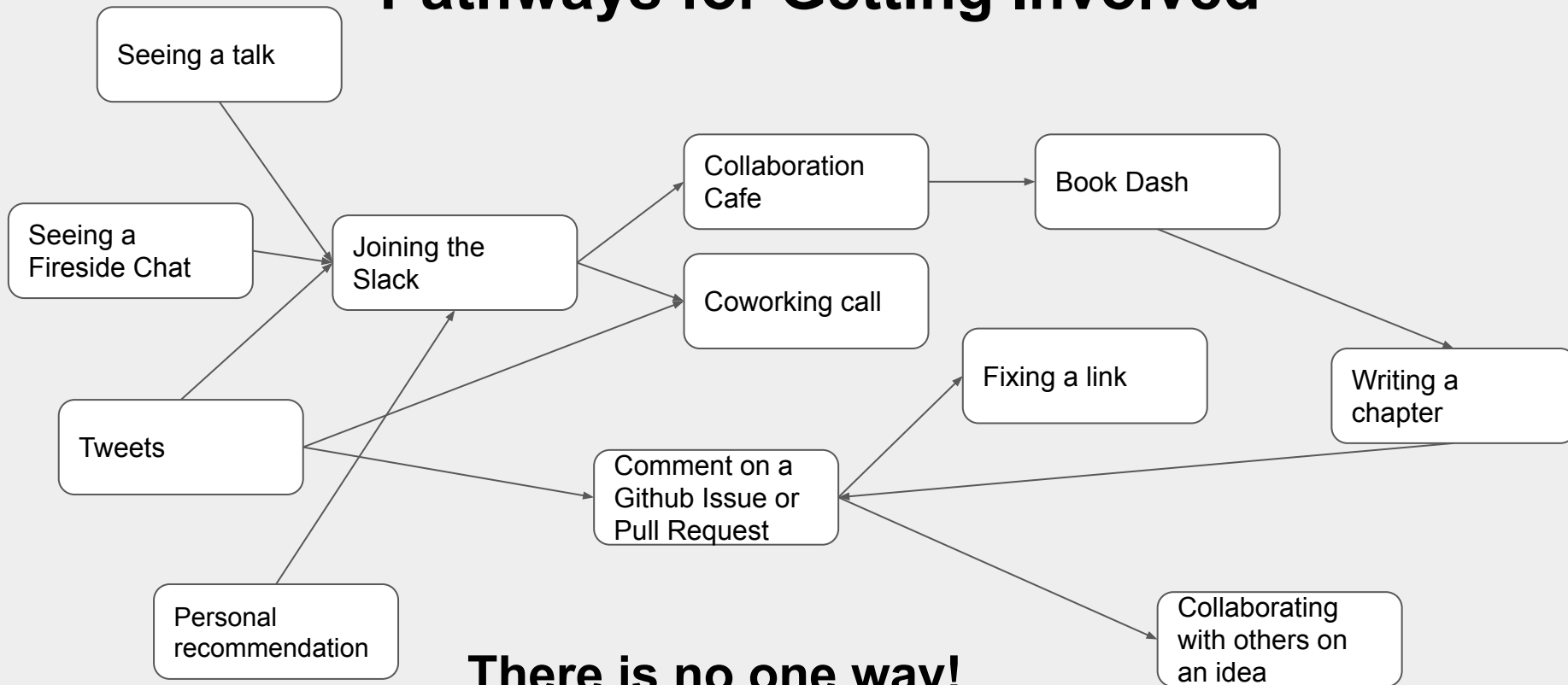
Piloting hubs in Amsterdam (Netherlands),
London and Bristol (UK)

Piloting Hybrid Events



Community Share-out May 2022

Pathways for Getting Involved



There is no one way!

Pathways for Collaboration



Develop & share



Maintain & improve



Make it global



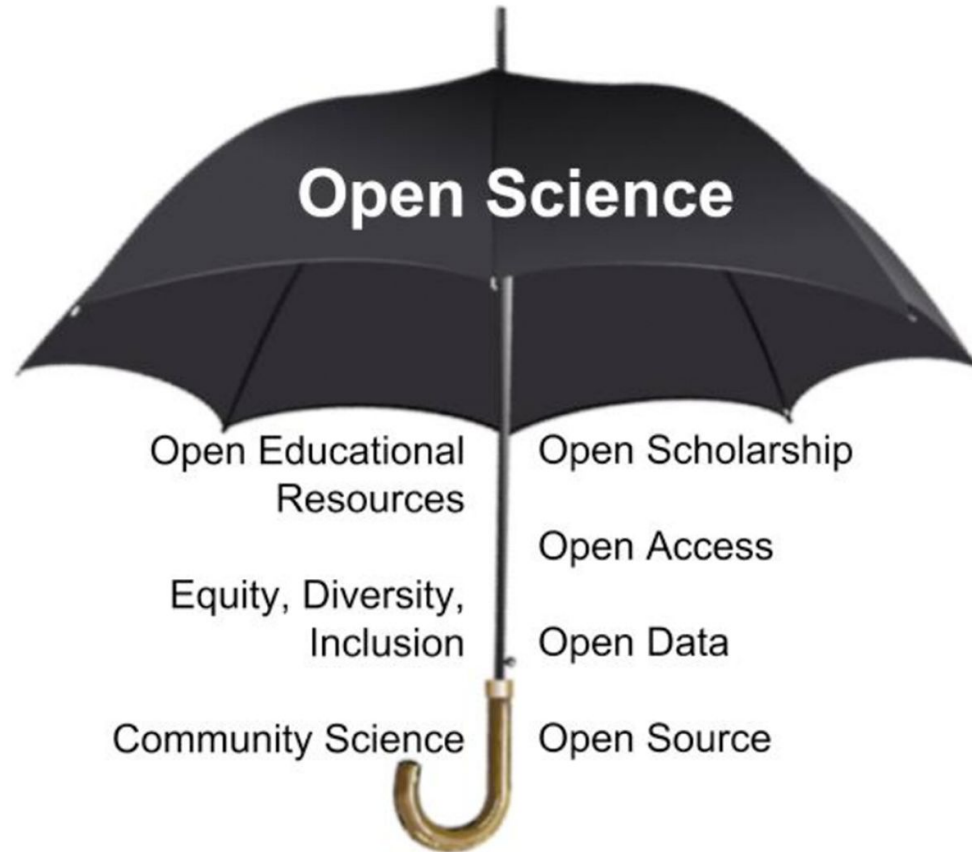
Share best practices

slido



**What you hear the phrase
'open science' what do you
think of?**

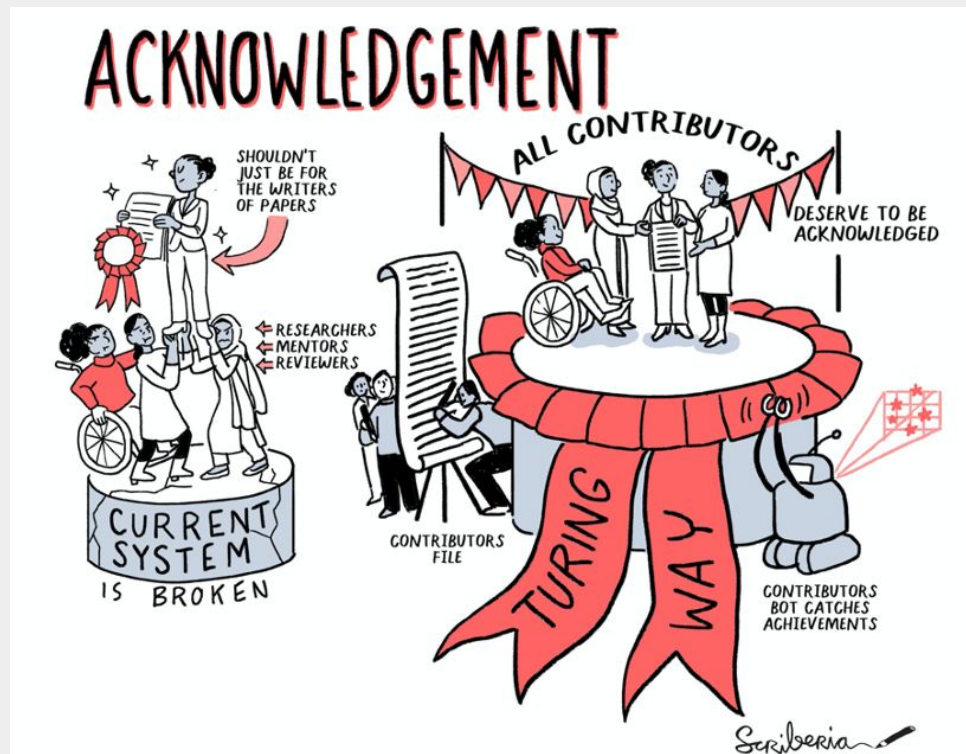
① Start presenting to display the poll results on this slide.



Adapted from: <https://www.meetup.com/Berlin-Open-Science-Meetup/>

Robin Champieux and Danielle Robinson

Contributor record and Community Acknowledgement



Batool Almarzouq

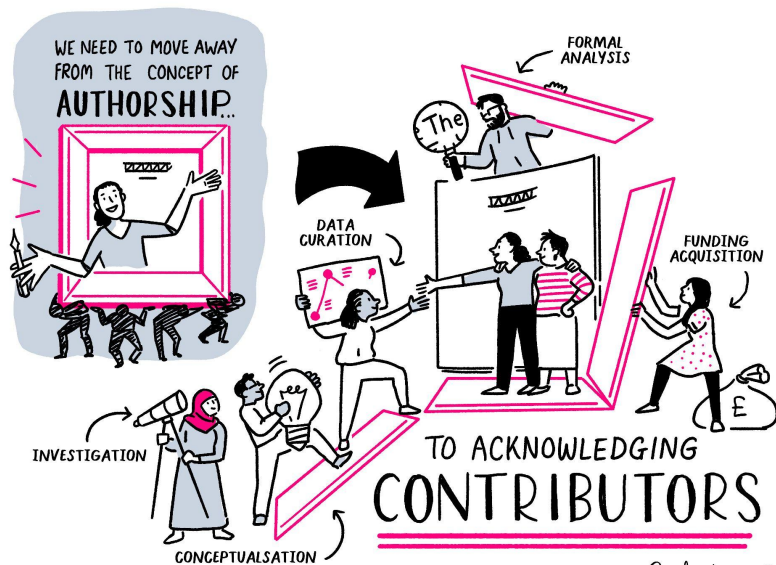
- Role: Book Dash November 2020 Attendee
- ORCID: [0000-0002-3905-2751](https://orcid.org/0000-0002-3905-2751)
- Short bio:

I'm the founder of R-Ladies in Saudi Arabia (Dammam). I initially majored in pharmacology but quickly developed an interest in biochemistry, structural biology, and bioinformatics. I enjoy applying deep learning to answer biological questions.

• Personal highlights:

I am currently co-developing a chapter on "CI services". I have helped upgrade the Jupyter Book Infrastructure and add hypothes.is to enable collaborative annotation of *The Turing Way* chapters. I have also translated the README.me chapter in Arabic. Personal quote: "I find it hard to express my personal thoughts and feelings in words. This was such an amazing experience. It helped me to develop my technical skills. Thank you so much to everyone I met in this Book Dash event :heart:."

A community-oriented collaborative book



All Contributors

Recognize All Contributors

Including those that don't push code



KirstieJane commented on 4 Mar 2019

Member Author ...

@allcontributors[bot] add @LouiseABowler for question code doc examples eventOrganizing review



allcontributors [bot] mentioned this issue on 4 Mar 2019

docs: add LouiseABowler as a contributor #279

Merged



allcontributors [bot] commented on 4 Mar 2019

Contributor ...

@KirstieJane

I've put up a pull request to add @LouiseABowler! 🎉

The Turing Way Community Handbook

Community Handbook

Code of Conduct

Ways to Enter the Project

Style Guide

Maintaining Consistency

Chapter Templates and
Contribution Workflow

Acknowledging Contributors

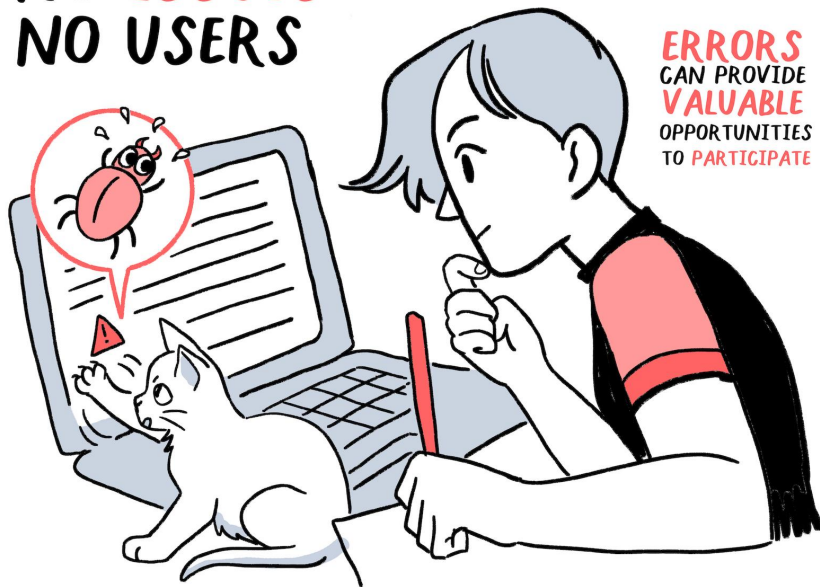
Monthly Newsletters

Book Dash Events

Community Coworking Calls

Template Collection

NO ISSUES
NO USERS

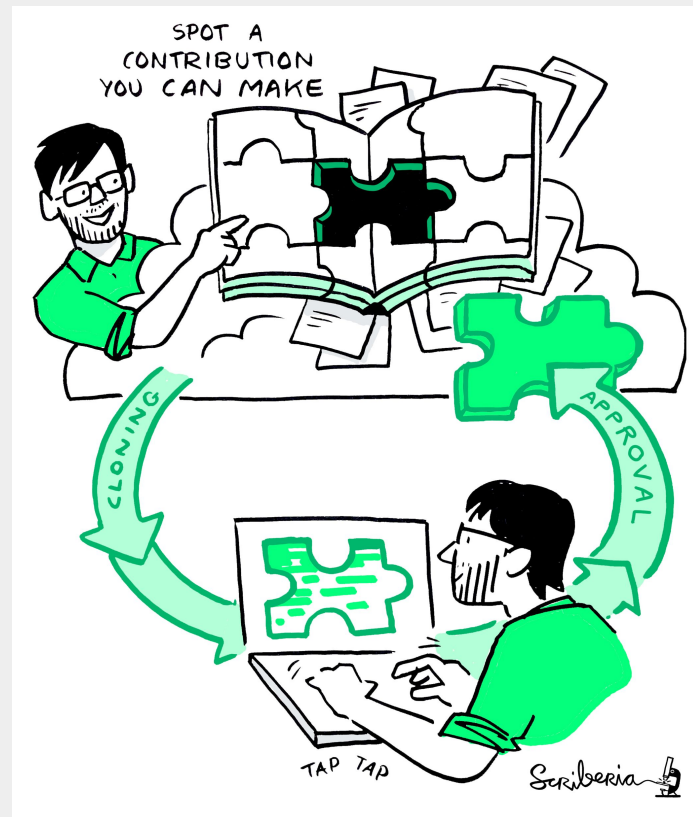


Scriberia

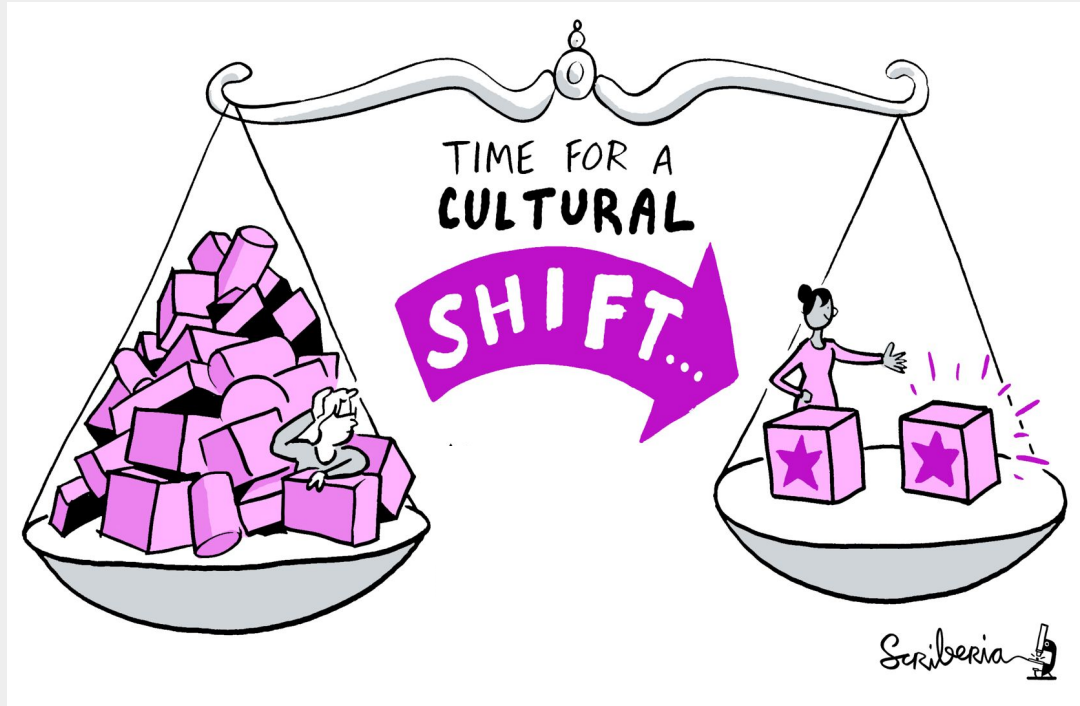
Meeting you where you are!

- 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

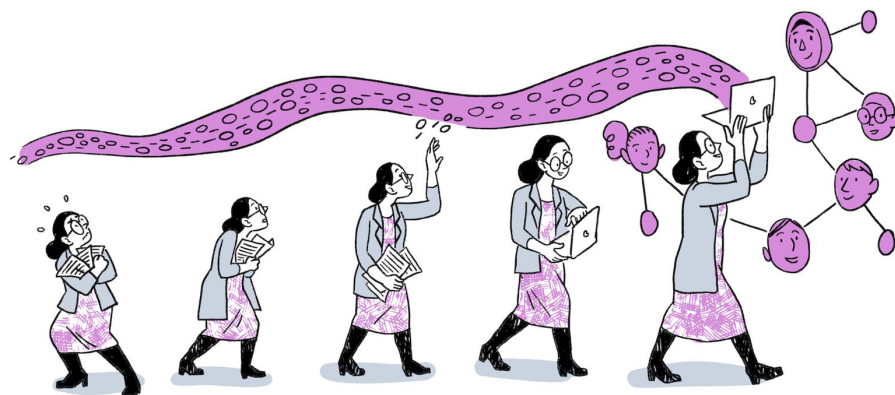
We value your participation!



Barriers to open & reproducible research



- Reward system
- Novel findings
- Publication cost
- Training and skill transfer
- Leadership
- Institutions (often)
- Lack of trust



EVOLVING TOWARDS AN
ERA OF
OPEN RESEARCH

Scriberia

Open Science involves many concepts.

Take possible steps to make your work open.

Ask for feedback and help whenever you can.

Project and Community Growth

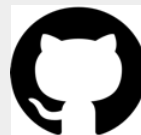
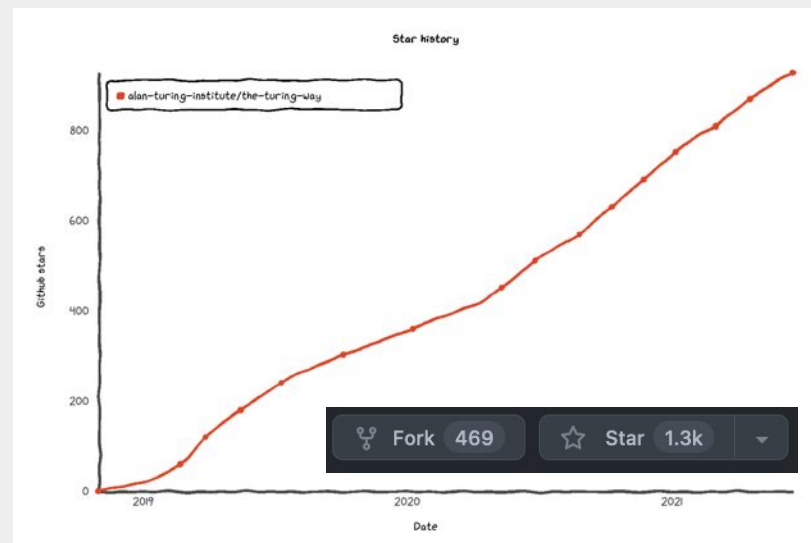
- 3 years, >200 Live Chapters
- Community resources, events, guidance, templates, training
- 325+ direct GitHub contributors and thousands of users

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



8,921
views

11,048
downloads



Notable Impacts Beyond the Project

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

- Featured in **Goldacre Review** (2022)
- Highly-commended project in **The Hidden REF** Competition (2022)
- Reproducibility of scientific results in the **EU 2020 (report)**
- An Emerging Technology Charter by **Mayor of London (policy)**
- Innovation Scholars: **UKRI grant 2020 (funding)**
- CodeRefinery and Library Carpentries (training materials)
- Projects by data scientists including at the Office for National Statistics
- Cited by 20+ peer-reviewed articles & 20+ open source projects

Learn these skills, methods and practices while contributing to the project.



Get involved with the community

- PR / Issue on Github
- Slack workspace
- Community events

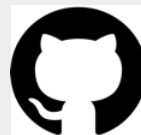
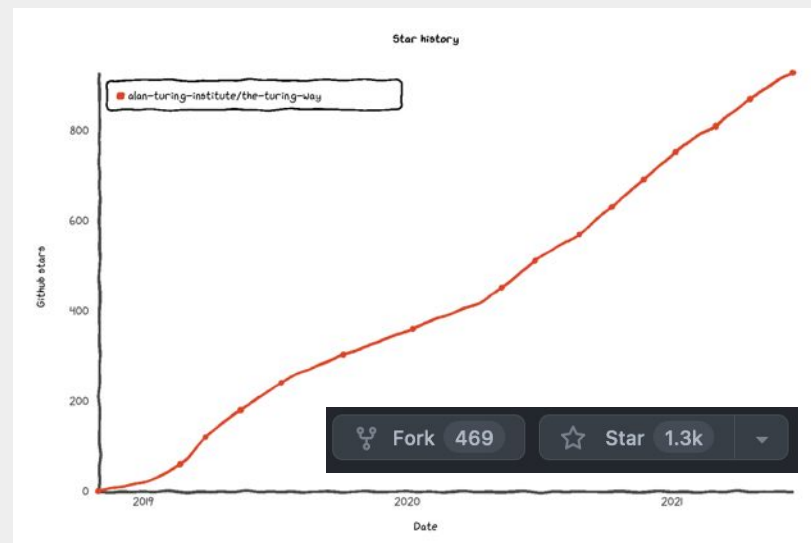
bit.ly/turingway

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



8,921
views

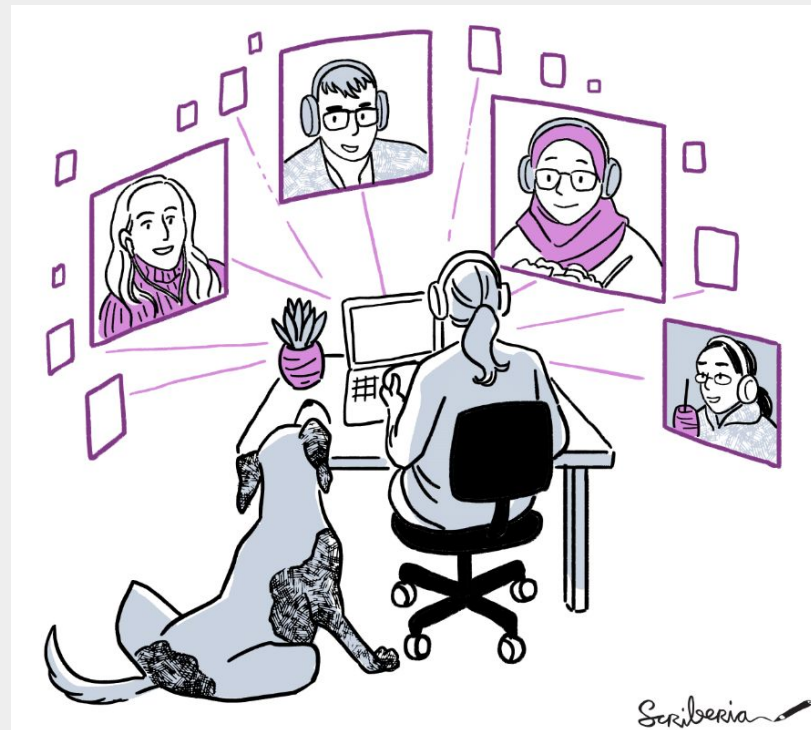
11,048
downloads

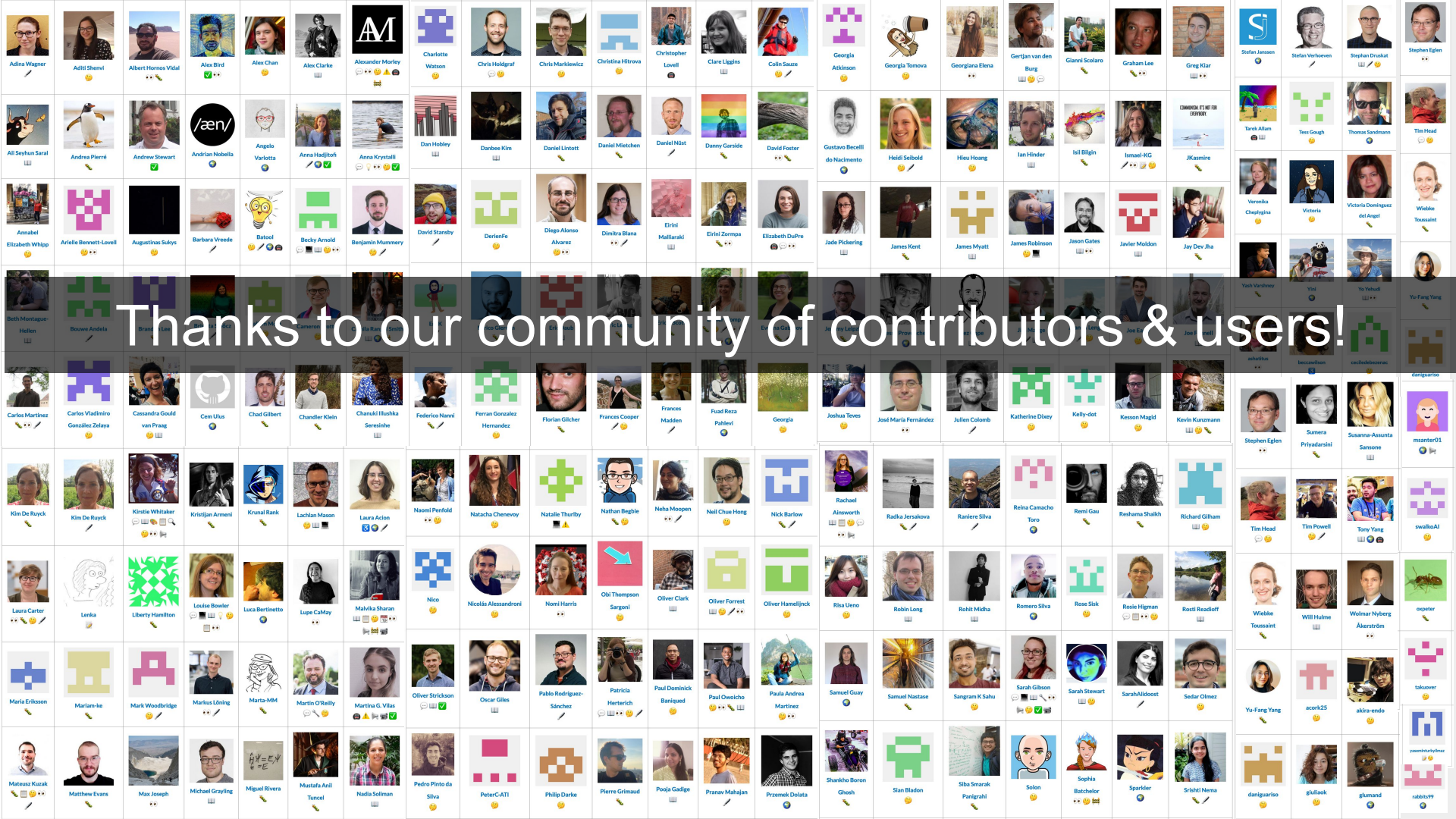


Join a Community Event

- *Collaboration Cafes*
 - 1st & 3rd Wednesdays (14:00 UTC)
- *Coworking Calls*
 - Mondays (10:00 UTC)
- *Fireside Chats*
- *Book Dash*
- *Community Share-out*

bit.ly/turingway





Thanks to our community of contributors & users!

Acknowledgements:

- Kirstie Whitaker, Malvika Sharan, Whitaker Lab Team Members
- *The Turing Way* community, friends & collaborators

Stay in Touch

- Book: the-turing-way.netlify.com
- Join Slack: <https://tinyurl.com/jointuringwayslack>
- Twitter: twitter.com/turingway
- Newsletter: tinyletter.com/TuringWay
- Mastodon: <https://fosstodon.org/@turingway>
- GitHub: github.com/alan-turing-institute/the-turing-way
- Original artwork by Scriberia: <https://doi.org/10.5281/zenodo.3332807>

Contact us: astele@turing.ac.uk or vhellon@turing.ac.uk

