

The Alan Turing Institute

The Turing Way

Reproducible, ethical, inclusive
and collaborative science

Dr Emma Karoune

On behalf of *The Turing Way* community

Pronouns: she/her



@TuringWay @ekaroune

<https://doi.org/10.5281/zenodo.6533251>

Dr Emma Karoune

Open Researcher

- Tools, practices and systems programme
- Turing Way
- DECOVID
- Turing-RSS Lab
- SSI Fellow - SSI Fellows page
- Elixir-UK - FAIR Data Fellow

Archaeobotanist

- FAIR Phytoliths Project
- Open reference collections



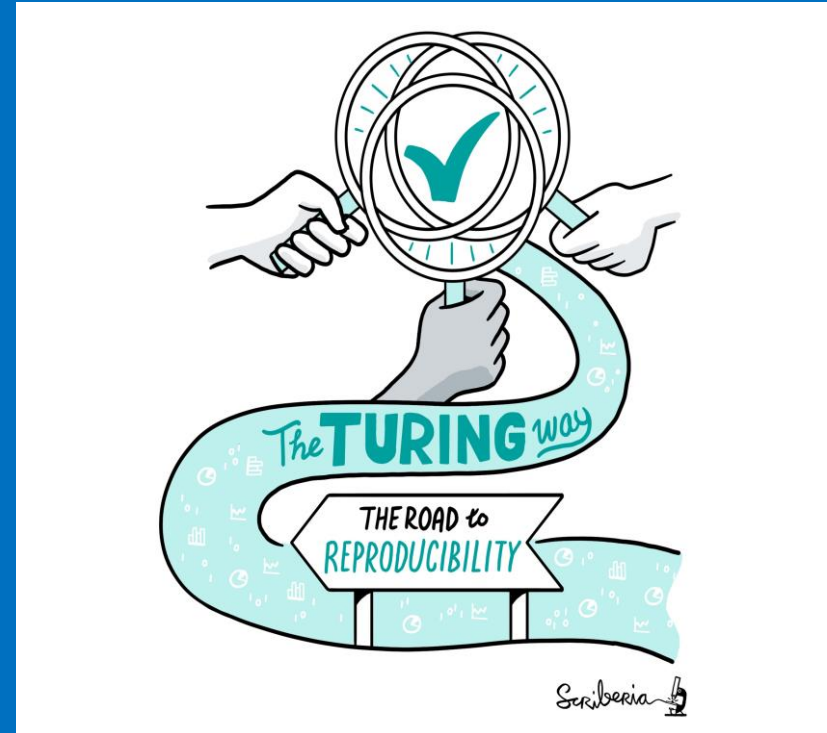


The Turing Way is an **open source book project** that involves and supports a **diverse research community** in ensuring that reproducible and ethical **data science** is accessible and comprehensible **for everyone**.



The Turing Way is:

- a book
- a community
- a global collaboration
- a whole tonne of work



@TuringWay @ekaroune

<https://doi.org/10.5281/zenodo.6533251>



Picture credit: Chris Gorgolew
@TuringWay @ekarou
<https://doi.org/10.5281/zenodo.65332>

Malvika Sharan

“No one can change research culture on their own. Scaling our community by empowering YOU to participate is how we will change the world.”



<https://doi.org/10.5281/zenodo.4536335>

<https://foss-backstage.de/session/>

[building-culture-collaboration-open-source-communities-turing-way](https://foss-backstage.de/session/building-culture-collaboration-open-source-communities-turing-way)

@TuringWay @ekaroune

<https://doi.org/10.5281/zenodo.6533251>

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

Is not considered
for promotion

Held to higher
standards than
others

Publication bias
towards novel
findings

Requires
additional
skills and a
change of
mindset

Barriers to reproducible research

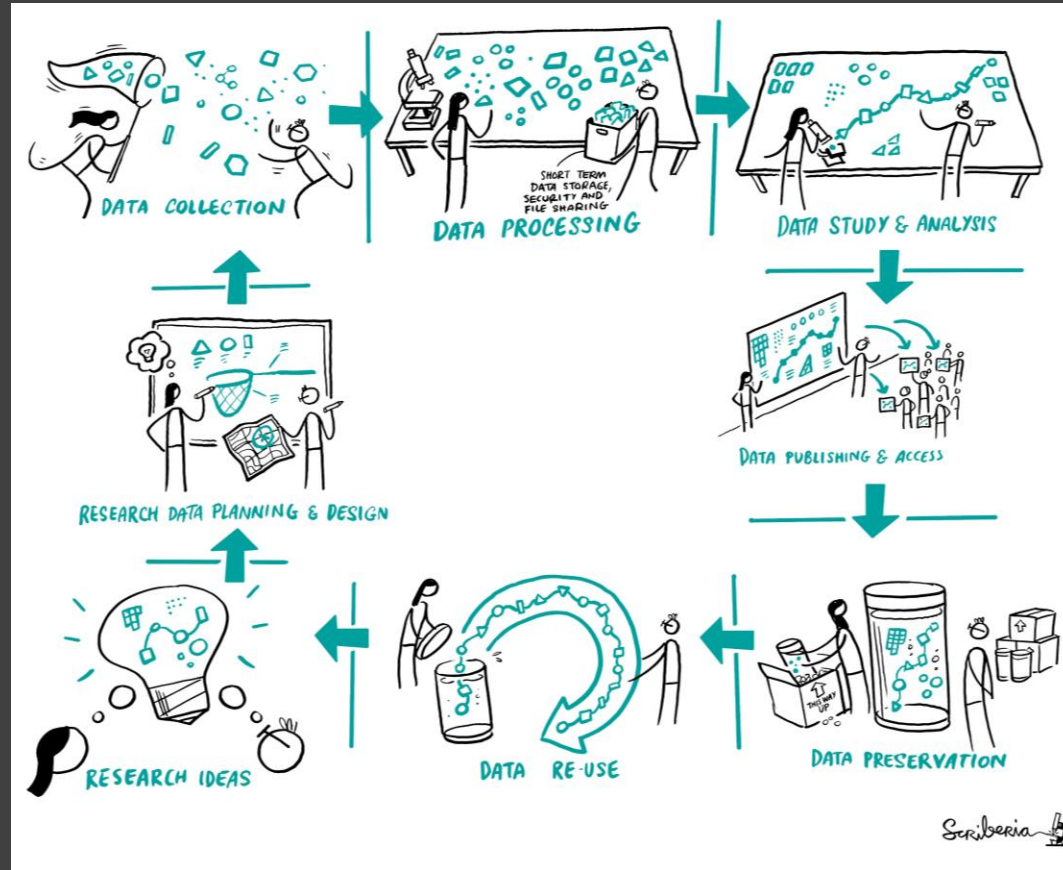
Plead the 5th

Support additional
users

Takes time


To be fully reproducible we have to cover all the steps of the research cycle

And that is super overwhelming...but we're here to help



The Book contents

- Five guides:
 - Reproducible research
 - Project Design
 - Communication
 - Collaboration
 - Ethical Research
- Community Handbook



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


Welcome

The Turing Way is an open source community-driven guide to reproducible, ethical, inclusive and collaborative data science.

Our goal is to provide all the information that data scientists in academia, industry, government and the third sector need at the start of their projects to ensure that they are easy to reproduce and reuse at the end.

The book started as a guide for reproducibility, covering version control, testing, and continuous integration. However, technical skills are just one aspect of making data science research "open for all".

In February 2020, *The Turing Way* expanded to a series of books covering reproducible research, project design, communication, collaboration, and ethical research.



@ekaroune, @turingway, #TuringWay

<https://the-turing-way.netlify.app/welcome>



The Turing Way

Search this book...

Welcome

Guide for Reproducible Research

Overview

Open Research

Version Control

Licensing

Research Data Management

Reproducible Environments

BinderHub

Code quality

Code Testing

Code Reviewing Process

Continuous Integration

Reproducible Research with Make

Research Compendia

Credit for Reproducible Research

Risk Assessment

Case Studies

Version control

Research data management

Reproducible environments

Code testing

Continuous integration

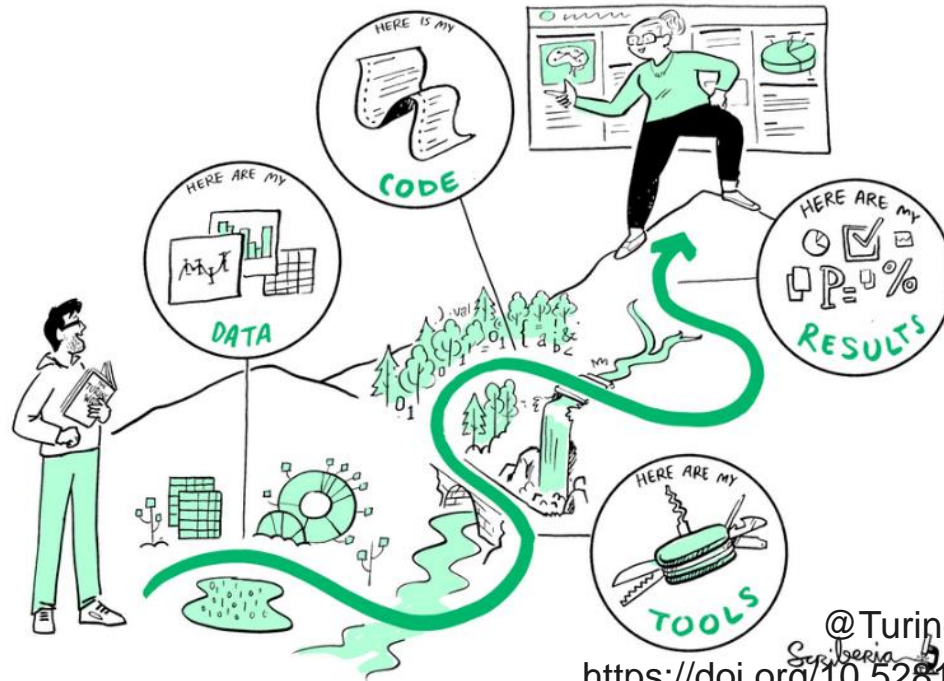


Guide for Reproducible Research

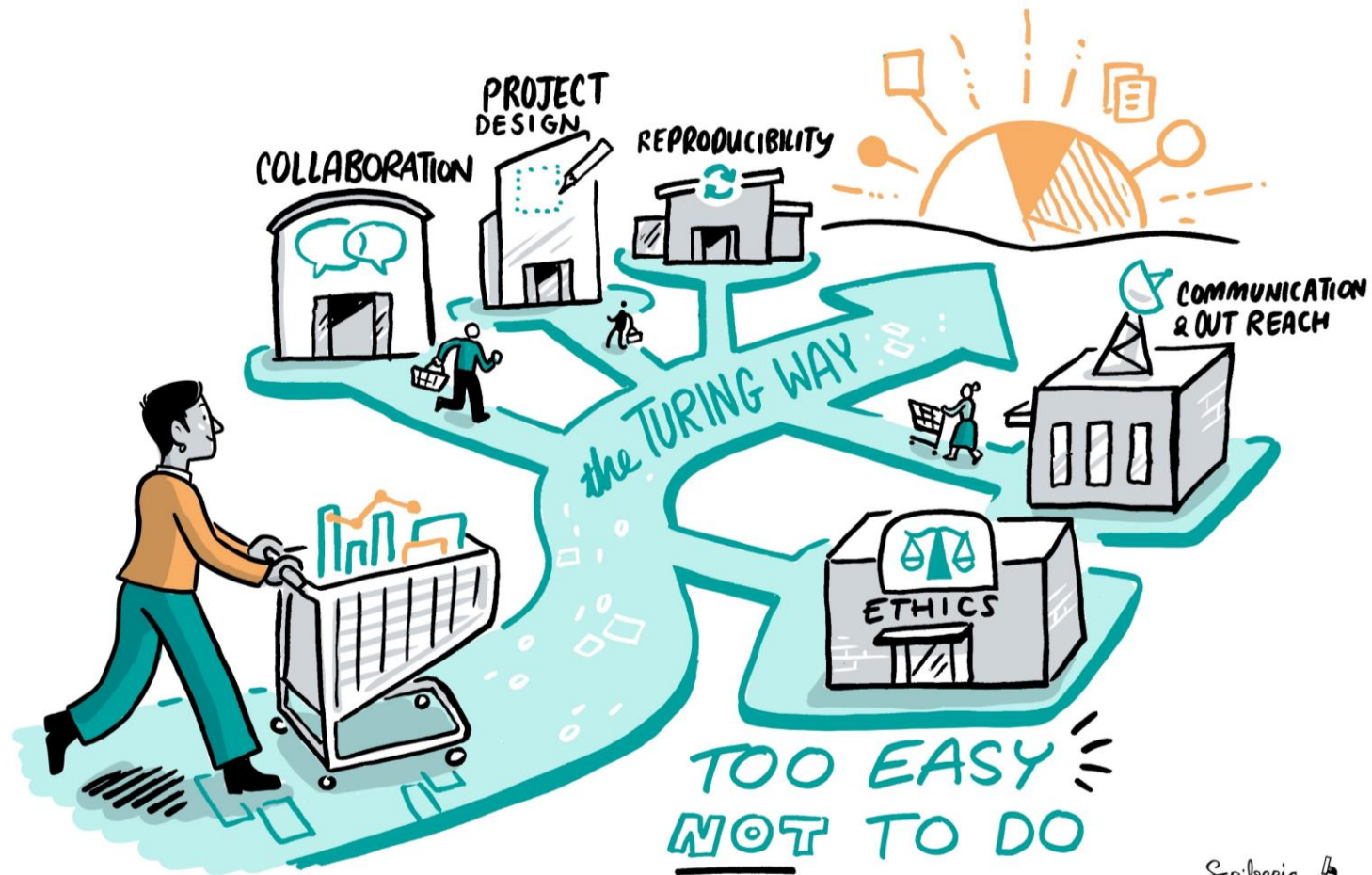
This guide covers topics related to skills, tools and best practices for research reproducibility.

The Turing Way defines reproducibility in data research as data and code being available to fully rerun the analysis.

There are several definitions of reproducibility in use, and we discuss these in more detail in the [Definitions of Reproducibility](#) section of this chapter. While it is absolutely fine for us each to use different words, it will be useful for you to know how *The Turing Way* defines *reproducibility* to avoid misunderstandings when reading the rest of the handbook.



@TuringWay @ekaroune
Siberia
<https://doi.org/10.5281/zenodo.6533251>



@TuringWay @ekaroune

<https://doi.org/10.5281/zenodo.6533251>

Emma Karoune

“People who use GitHub regularly don’t realise how many words there are that people like me don’t know! It can be really demoralising.”



@TuringWay @ekaroune

<https://doi.org/10.5281/zenodo.6533251>

<https://the-turing-way.netlify.app/collaboration/github-novice.html>

An open source project

- Hosted on Github.
- And built with other Open Source projects (Git, Jupyter Book, netlify).

alan-turing-institute / the-turing-way

Unwatch 38 Unstar 623 Fork 202

<> Code Issues 232 Pull requests 65 Actions Projects 2 Wiki Security Insights

master 86 branches 4 tags Go to file Add file Code

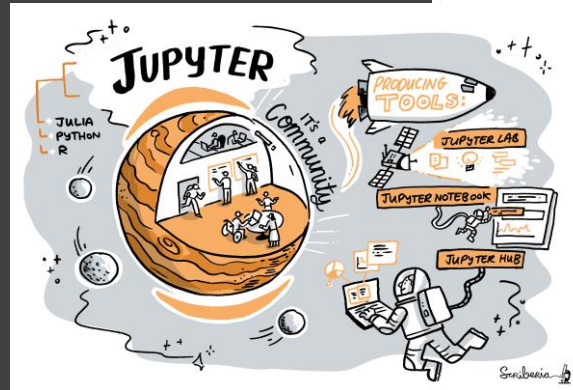
malvikasharan Merge pull request #1279 from alan-turing-institute/malvika... 455458 4 days ago 5,105 commits

.github	Merge pull request #985 from alan-turing-institute/refine-tests	4 months ago
book	Merge pull request #1279 from alan-turing-institute/malvikasharan-r...	4 days ago
communications	Update README.md	2 months ago
conferences	Update README.md	2 months ago
open-life-science-mentoring	Updated OLS-2 Ethics README.md	11 days ago
project_management	split acknowledgement file into two subchapters	3 months ago
templates	Updating Github templates	17 months ago
tests	Update tests/no-bad-latin.py	11 days ago
workshops	minor update	last month
.all-contributorsrc	docs: update .all-contributorsrc	2 months ago
.gitignore	Merge pull request #985 from alan-turing-institute/refine-tests	4 months ago
	Merge pull request #1130 from srishti-nema/add-label	4 months ago
	Update CONTRIBUTING.md	2 months ago

About: Host repository for The Turing Way: a how to guide for reproducible data science. the-turing-way.netlify.app

Releases: v0.0.4 new chapter: Conti... on 28 May 2019

Packages: No packages published. Publish your first package



@ekaroune, @turingway, #TuringWay

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



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 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.



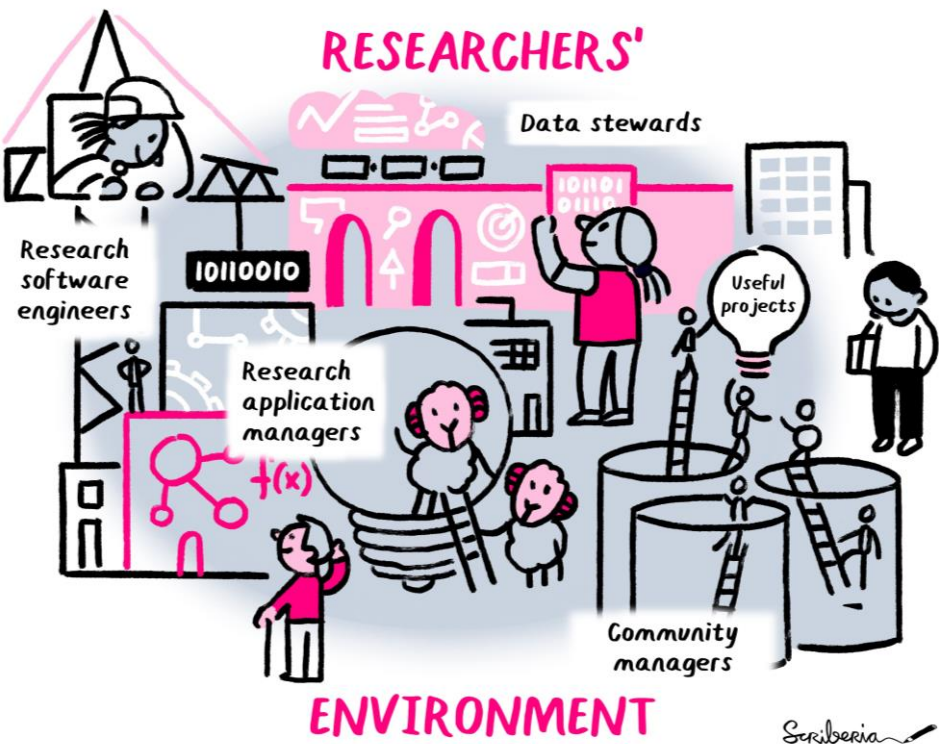
Fig. 80 There is more to collaboration than we see. *The Turing Way* project illustration by Scriberia. Used under a CC-BY 4.0 licence. DOI: 10.5281/zenodo.3332807.

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

@TuringWay @ekaroune

<https://doi.org/10.5281/zenodo.6533251>



Research infrastructure roles

- Community Managers
- Data Stewards
- RSEs
- Project Managers and more!

Led by Arielle Bennett & Esther Plomp

Project design

Guide for Project Design

Overview of Project Design

Creating Project Repositories

Personas and Pathways

File Naming Convention

Code Styling and Linting



The Turing Way

Search this book...

Welcome

Guide for Project Design

Overview of Project Design

Creating Project Repositories

Personas and Pathways

File Naming Convention

Code Styling and Linting

Guide for Communication

Guide for Collaboration

Guide for Ethical Research

Community Handbook

Afterword

Visit our GitHub Repository

This book is powered by Jupyter Book

Guide for Project Design

This guide covers topics related to effective project planning and management.

In this guide, we compile best practices and guidance for designing research projects by including different aspects of project management and (iterative) development practices derived from academia and industry.

Before starting a project, researchers must define the project's scope. Researchers should start by identifying the main questions they aim to address through their work. Scope definition also includes defining the project goals, possible outcomes, resources requirements, people involved (collaborators, users and target audience) and possible constraints.

Researchers can then proceed to identify the expected minimum viable product of their project, synergies with other projects (similarities as well as differences), measure(s) of success, and the overall impact they hope to achieve. After these crucial questions are addressed, planning can focus on the operational aspects of the project, including ensuring that the project is usable and compliant with appropriate external guidelines (for instance, ethics requirements and regulatory limitations).

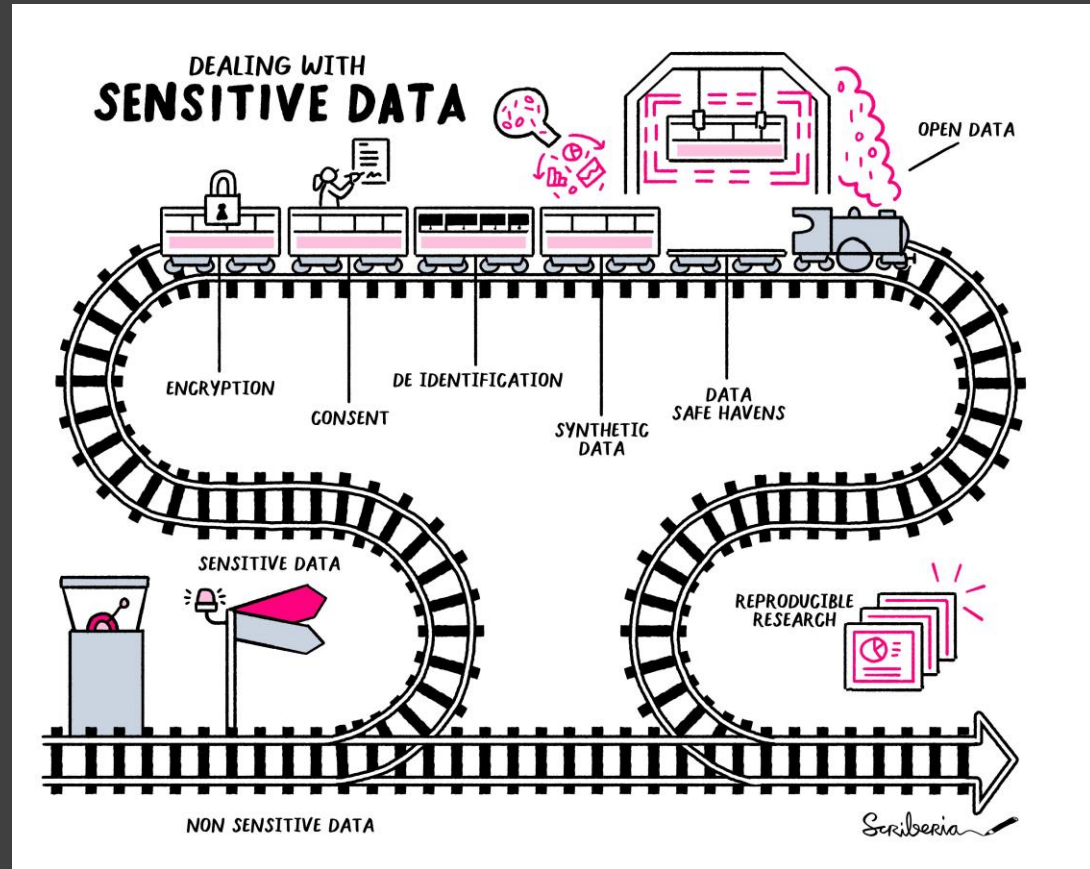


@TuringWay @ekaroune

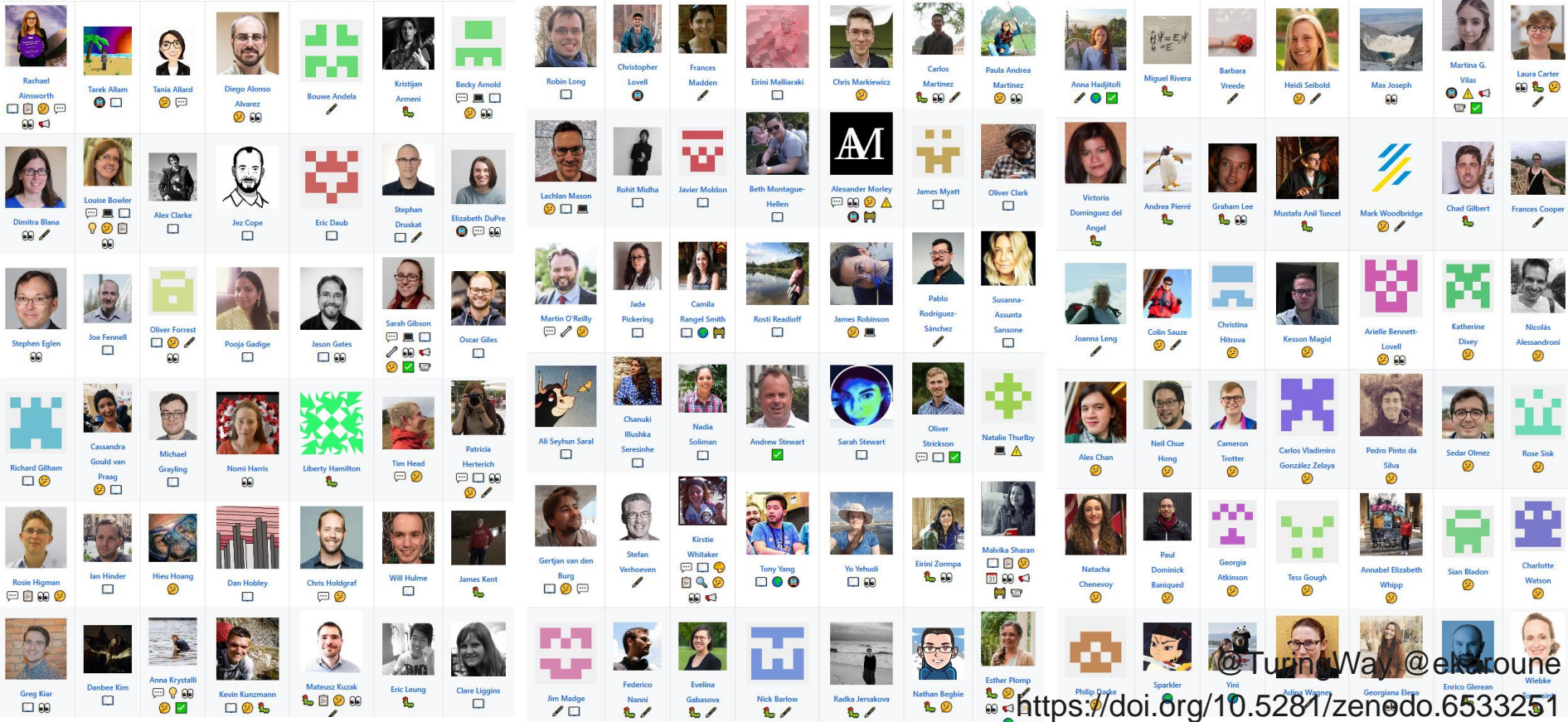
<https://doi.org/10.5281/zenodo.6533251>

Guide for Project design – Sensitive Data

- Can still work reproducibly (transparently)
- Not one way to do this
- Can publish openly (maybe not data?!)

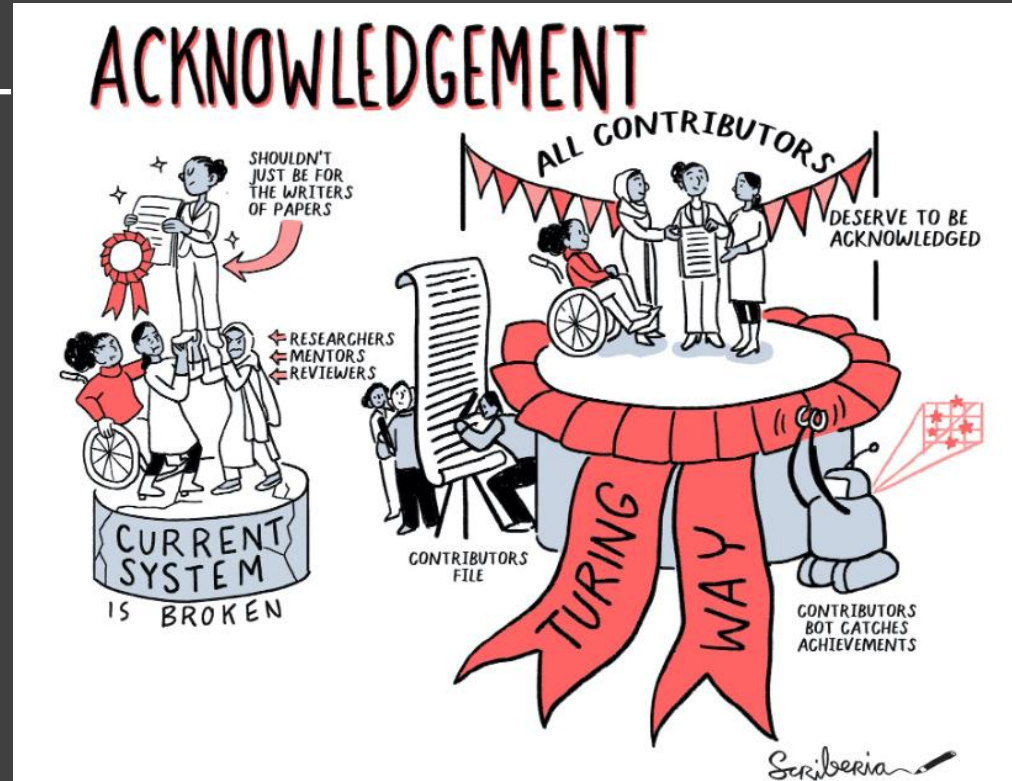


Thank you to our 300+ contributors



Acknowledgement

- The Turing Way is collectively written, and attributed to the entire community
- We record contributions in free text to customise the acknowledgement to community members' individual needs



Mentored contributions



Collaboration cafes
Co-working sessions



Book Dash

@TuringWay @ekaroune
<https://doi.org/10.5281/zenodo.6533251>

Martina Vilas

“The co-working hours are friendly for introverts who might be intimidated working with new people. These personal interactions are also crucial for staying motivated!”



<https://martinagvilas.github.io>

Collaboration cafes

- 1st and 3rd Wednesdays of each month
- All remote participation
 - Zoom call
 - Pomodoro technique
 - Breakout rooms for mentored contributions
- Everyone welcome

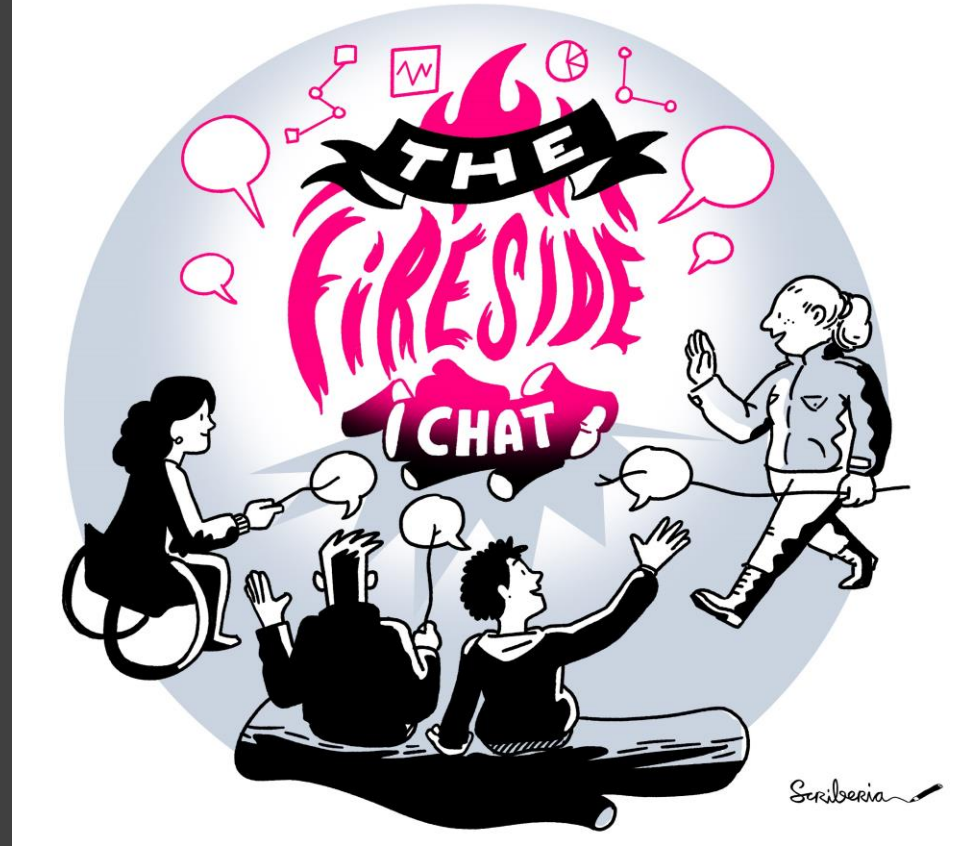


Fireside chat

- Next one is in June.
- Watch the previous ones on YouTube

[Enabling Inclusive Research Conferencing in the "Post-Pandemic Era"](#)

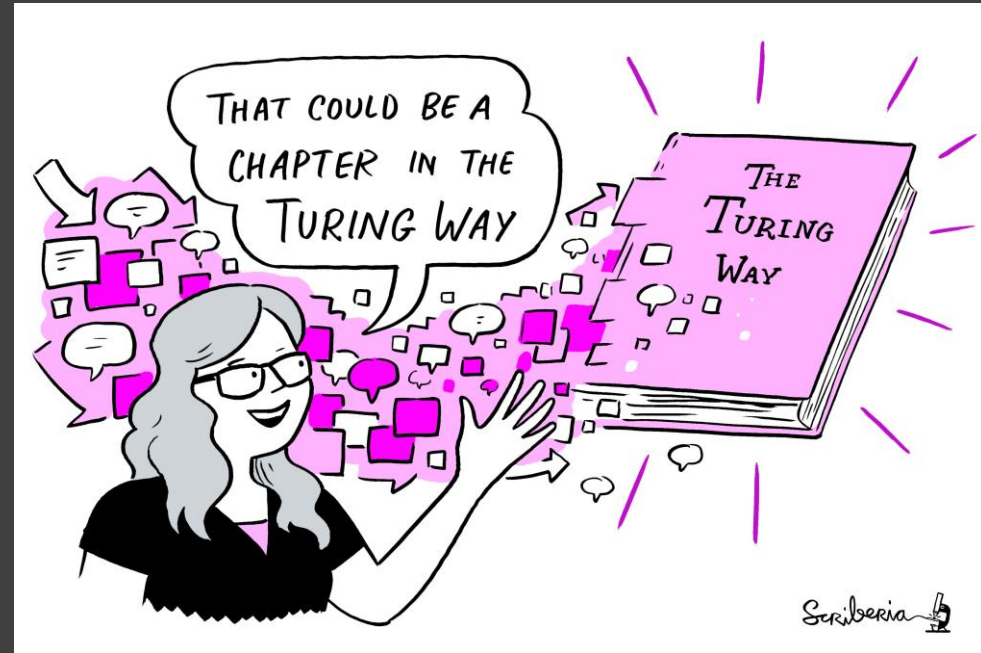
<https://www.youtube.com/watch?v=ImwJqo1zbMI>



@TuringWay @ekaroune
<https://doi.org/10.5281/zenodo.6533251>

Join us!

- You can help in so many ways!
 - Editing
 - Writing
 - Linking
 - Automating
 - Curating
 - Translating
 - Promoting



@TuringWay @ekaroune

<https://doi.org/10.5281/zenodo.6533251>

<https://github.com/alan-turing-institute/the-turing-way/issues/1712>

Thank you

The
Alan Turing
Institute

- Book: <https://the-turing-way.netlify.com>
- Newsletter: <https://tinyletter.com/TuringWay>
- GitHub: <https://github.com/alan-turing-institute/the-turing-way>
- Slack: <https://tinyurl.com/jointuringwayslack>
- Next fireside chat: June – look out for the date on twitter!
- Next Collaboration Café: 1 June at 3pm UK, 4pm CEST, 10am EDT
- Original artwork by Scriberia: <https://doi.org/10.5281/zenodo.3332807>
- This work was supported by The UKRI Strategic Priorities Fund under the EPSRC Grant EP/T001569/1, particularly the "Tools, Practices and Systems" theme within that grant, and by The Alan Turing Institute under the EPSRC grant EP/N510129/1.

@TuringWay @ekaroune

<https://doi.org/10.5281/zenodo.6533251>