The Alan Turing Institute

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

Alejandro Coca-Castro

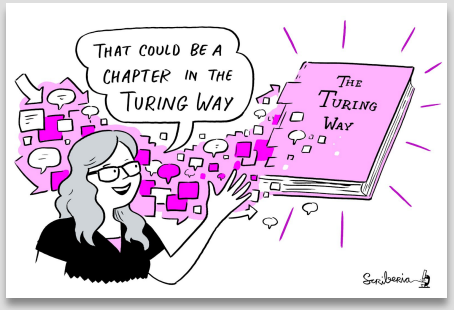
Pronouns: He/him

https://github.com/alan-turing-institute/the-turing-way, @alejo_coca, @turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6866185

The Turing Way

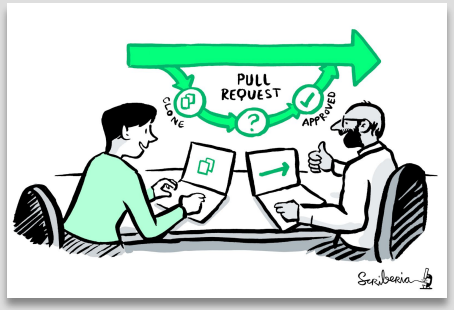
An Open Science and community lead guide on Data Science.

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

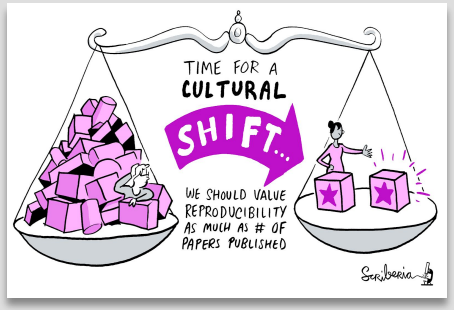


A Book

A Community



An Open Source Project



A Culture of Collaboration

https://github.com/alan-turing-institute/the-turing-way, @alejo_coca, @turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6866185

The Alan Turing Institute

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

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

Sector Specific

Cross-Cutting

Guide for Reproducible Research

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

Welcome

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

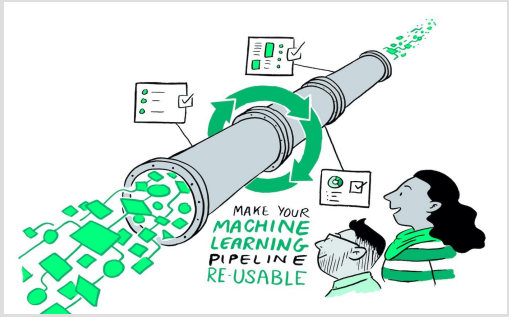
Our goal is to provide all the information that data scientists in academia, industry, government and the third sector need at the start of their projects to ensure that they are easy to reproduce and reuse at the end.

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.

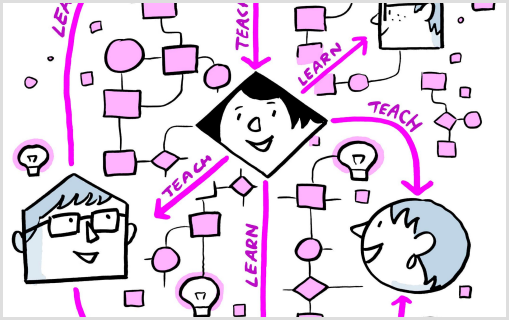
https://github.com/alan-turing-institute/the-turing-way, @alejo_coca, @turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6866185

The Turing Way Guides

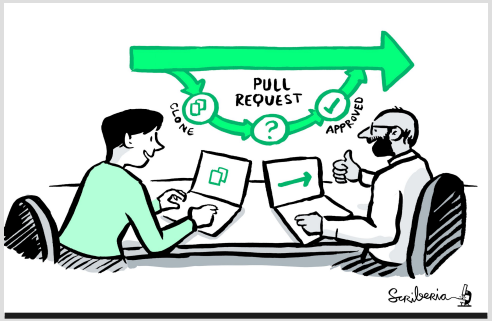


Reproducibility

Project Design

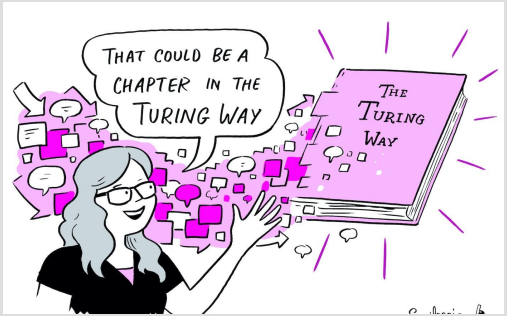


Communication



Collaboration

Ethical Research



Community Handbook

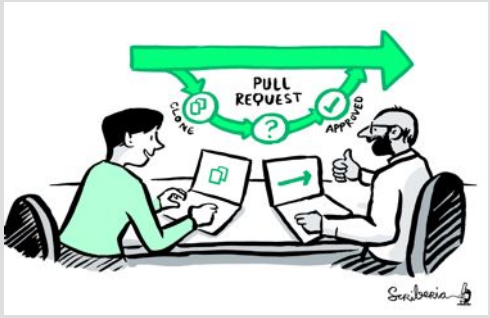
https://github.com/alan-turing-institute/the-turing-way, @alejo_coca, @turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6866185

Pathways for Collaboration

Develop & share

Maintain & improve

Share resources



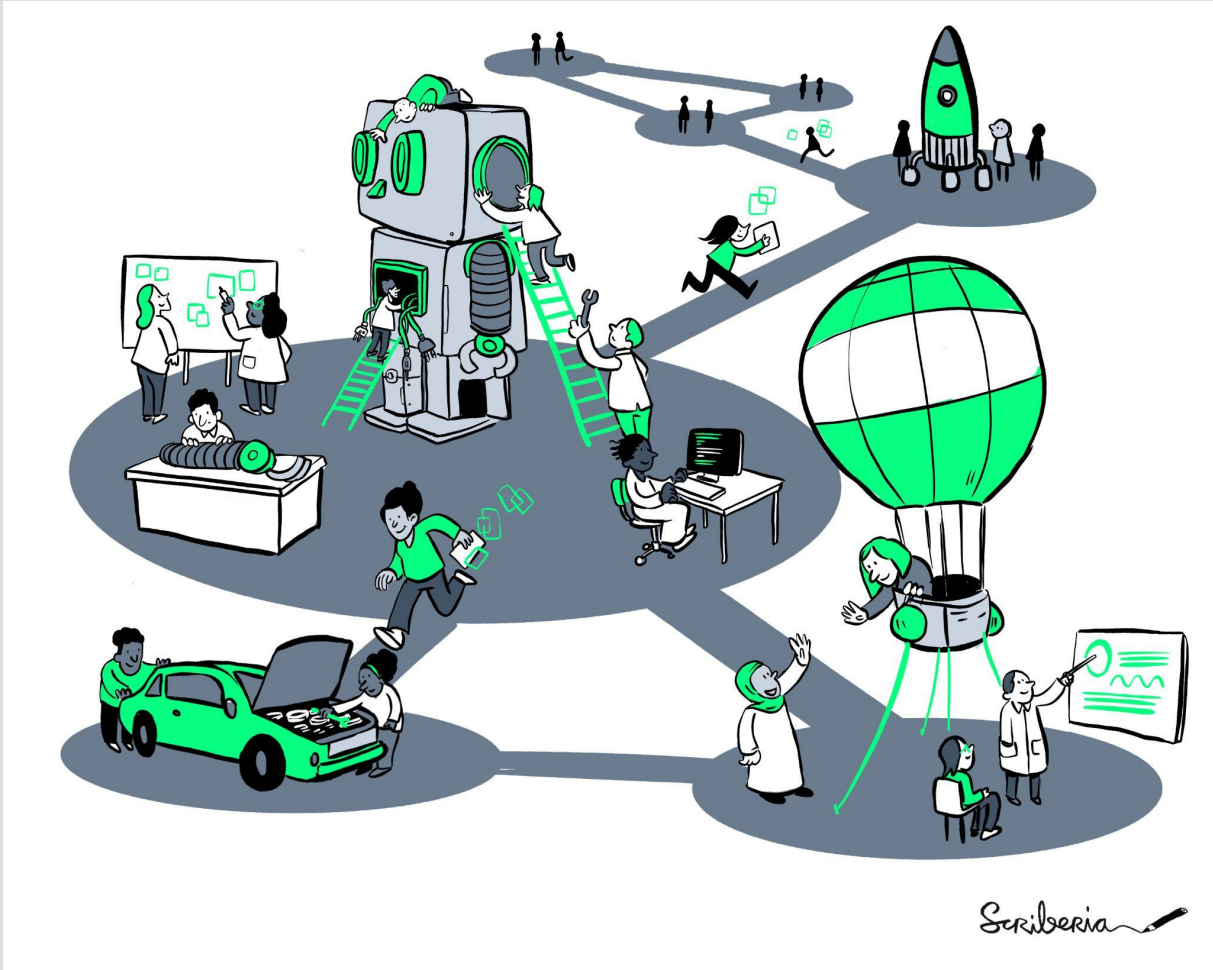
Review and update

Make it global

Share best practices

https://github.com/alan-turing-institute/the-turing-way, @alejo_coca, @turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6866185

Use cases



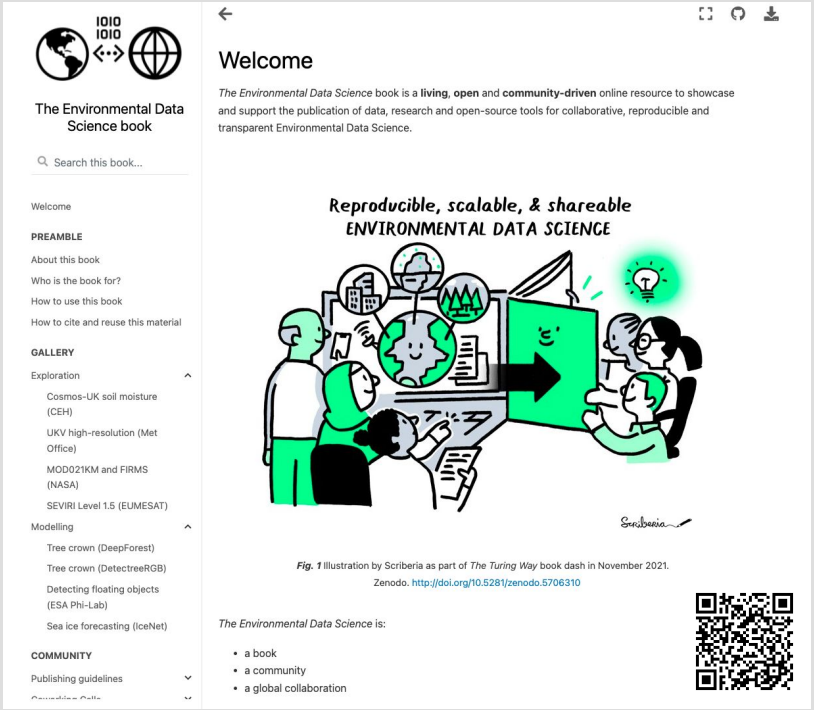
Environmental Data Science book

Mission

Support the publication of data, research and open-source tools for collaborative, reproducible and transparent Environmental Science

Features

- **Community-based model**
- Collaborative publication of executable and reproducible jupyter notebooks
- Promoting scalable tools in *python* e.g. Pangeo (current), R & Julia (future) using interactive computing resources



Learnings from *The Turing Way*

Community & translation of Open Science concepts

https://github.com/alan-turing-institute/the-turing-way, @alejo_coca, @turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6866185

Environmental Data Science Book resources: https://the-environmental-ds-book.netlify.app/

2i2c

Mission

Make interactive computing more impactful through community-centered open infrastructure services

Features

- **Non-profit based model**
- Open source operations/dev-ops as a service
- Working with the tech industry (cloud providers) to democratise interoperable interactive computing

Learnings from *The Turing Way*

Emerging (engineers, product manager, community) roles & transparent communication (all documents)

https://github.com/alan-turing-institute/the-turing-way, @alejo_coca, @turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6866185

2i2c resources: https://2i2c.org/, https://infrastructure.2i2c.org/en/latest/, https://docs.2i2c.org/en/latest/

The Turing-Roche Strategic Partnership

Mission

5-year strategic partnership for exploring patient heterogeneity through advanced analytics

Features

- **Research/industry based model**
- Ensuring outputs are shared for reuse
- Open knowledge sharing through events and communication channels
- Encouraging Roche researchers/practitioners to share industry practices in *The Turing Way*

Learnings from *The Turing Way*

Community building, translate and exchange Open Science fundamentals within the health industry

The Alan Turing Institute

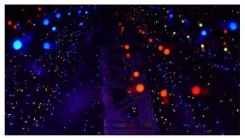
Roche

The Turing-Roche knowledge share series: Synthetic data →
Monday 25 Jul 2022
Time: 15:00 - 16:00

Florimond Houssiau I Marta Batlle I Finn Janson

The Turing-Roche knowledge share series: Predictive modelling in healthcare →
Monday 30 May 2022
Time: 15:00 - 16:00

Jasmina Bogojeska I Matthew Sperrin I Elif Ozkirimli I Brieuc Lehmann



The Turing-Roche knowledge share series: Modern topics in missing data →
Monday 25 Apr 2022
Time: 14:00 - 15:00

Robin Mitra I Chris Harbron

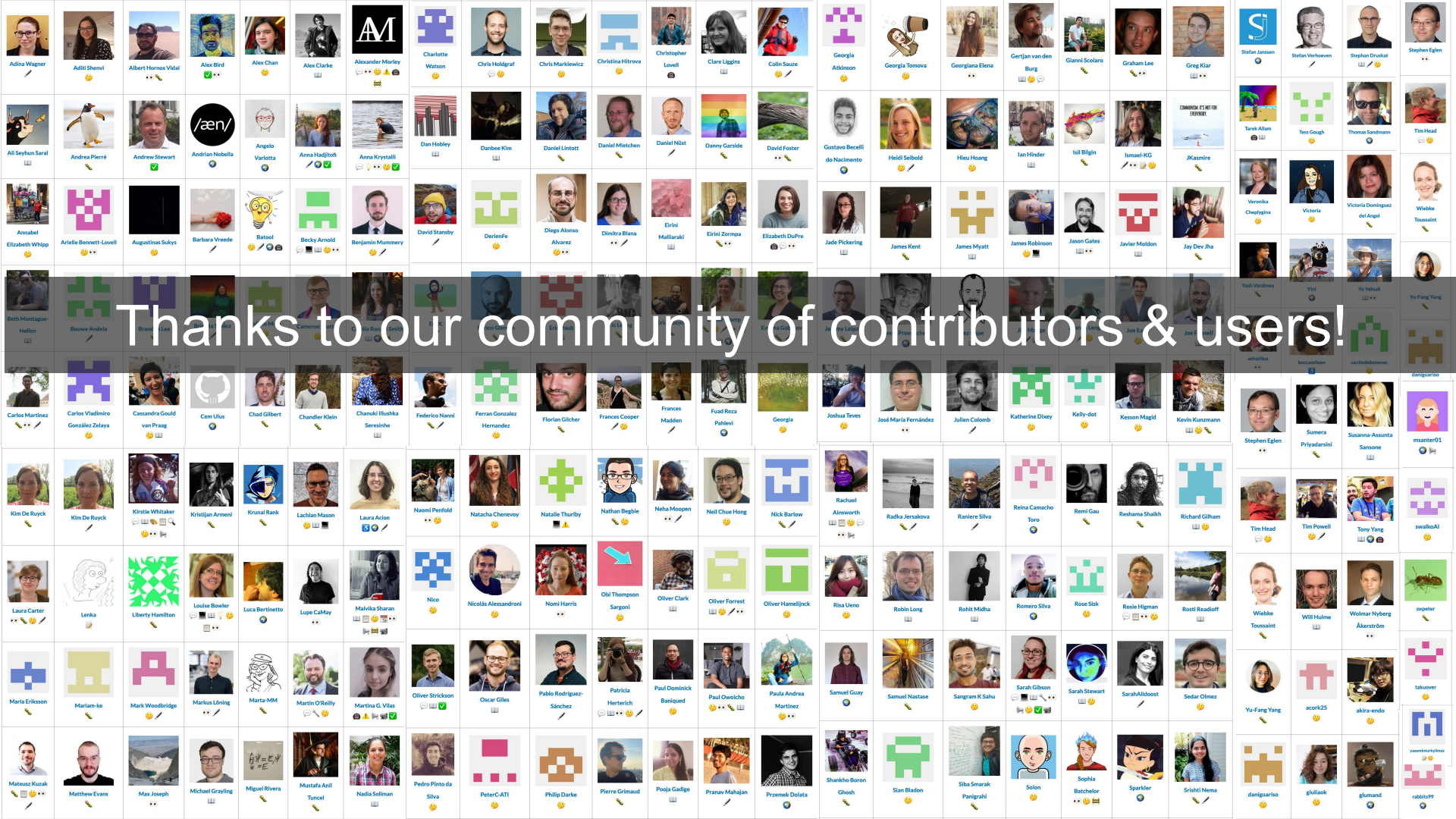
Knowledge share series

Project info: https://www.turing.ac.uk/research/research-projects/alan-turing-institute-roche-strategic-partnership

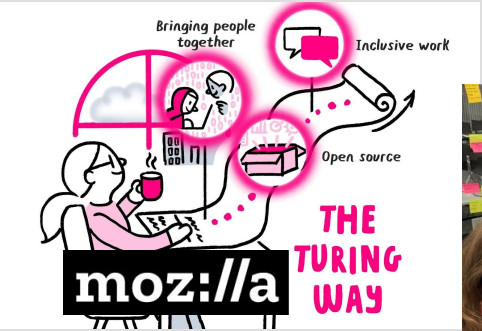
Notable Impacts Beyond the Project

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

- 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)**
- CodeRefinary and Library Carpentries (training materials)
- Projects by data scientists including at the Office for National Statistics
- Cited by 30+ peer-reviewed articles & 20+ open source projects

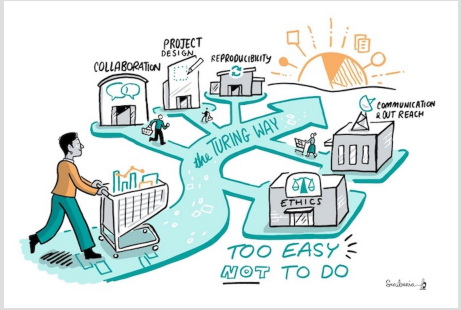

Thanks to our community of contributors & users!

The Turing Way



Proposal by **Kirstie Whitaker** approved

Create README.md
KirstieJane committed on 1 Nov 2018



250+ live chapters, 340+ contributors, 3000+ Twitter followers, 3000+ monthly visitors of the five guides.

2 project leads, a community manager, 10+ part paid staff, volunteer members, 4 translation and localisation leads.

2018

The Alan Turing Institute

2019

Book launched with **10 members**, **3 chapters**, many workshops, **2 Book Dashes**. Year ended with Several chapters by **67 contributors**

2020

Book Dash 2020,
Project expanded to **five guides**, Community Handbook, social media, Translation started. **238 contributors**.

Logo was designed with Scriberia

2021

Project growth, **134+ chapters** citation in peer reviewed articles, government reports and policies, new projects emerged as extension. **309 contributors**

Highly commended: *The Turing Way*, The Turing Way Community, The Alan Turing Institute

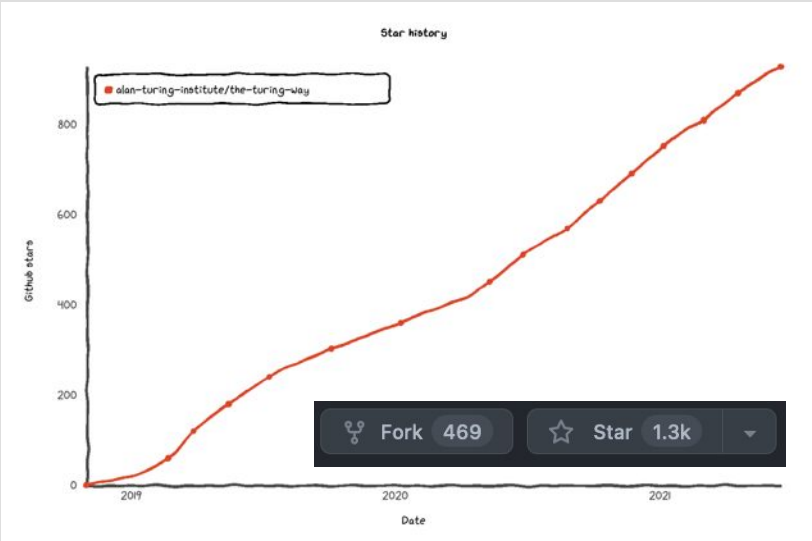
hidden REF

2022

https://github.com/alan-turing-institute/the-turing-way, @alejo_coca, @turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6866185

How to get involved

- Pull Request / Issue on Github
- Slack workspace
- Community events
  - Collaboration Cafes
  - Coworking Calls
  - Fireside Chats
  - Book Dash



bit.ly/turingway

https://zenodo.org/record/3332807

zenodo 8,921 views 11,048 downloads

https://github.com/alan-turing-institute/the-turing-way, @alejo_coca, @turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6866185

Acknowledgements:

– Kirstie Whitaker, Malvika Sharan, Anne Lee Steele

– *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 me: **acoca@turing.ac.uk**