


**The
Alan Turing
Institute**



**How can we get
better at coding in
the open?**

Kirstie Whitaker
Pronouns: she/her



@TuringWay @kirstie_j
<https://doi.org/10.5281/zenodo.10215177>

**Research
Programme
Director for
Tools,
Practices &
Systems
The Alan
Turing Institute**



@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

“

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**.

”

@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

The Turing Way is:

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



@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

Thank you to our 400+ contributors



@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

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.”







Personal experience of:

- File drawer effect
- Lack of reproducibility
- Imposter syndrome around coding
- Lonely
- Wasted time
- Hypocrisy of academia

@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

<https://www.turing.ac.uk/about-us/impact/asg>

VERSION CONTROL



Personal experience of:

- File drawer effect
- Lack of reproducibility
- Imposter syndrome around coding
- Lonely
- Wasted time
- Hypocrisy of academia

#bropenscience is broken science

All elements of open scholarship should be in service of equity, diversity and inclusion.

Faster, more accessible
→ Better quality insights
→ A more just world



Whitaker & Guest, 2020.

<https://www.bps.org.uk/psychologist/bropenscience-broken-science>

<https://hdl.handle.net/2066/226397>

@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

“Opening science involves opening up the processes of creating, evaluating, sharing, exploring and storing scientific knowledge, practices and perspectives.”

“All scientific disciplines and aspects of scholarly practices, including basic and applied sciences, natural and social sciences and the humanities, are included.”



A photograph of a forest path with sunlight streaming through the trees, creating a warm and bright atmosphere. The sun is visible in the upper right corner, casting rays of light across the scene. The trees are lush green, and the path is covered in fallen leaves and grass.

**“Sunlight is the
best disinfectant”**

Louis Brandeis

https://en.wiktionary.org/wiki/sunlight_is_the_best_disinfectant

@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

Open practices are ready to deliver safe and ethical data science and AI

- Safety and sustainability
- Accountability
- Fairness
- Explainability
- Data stewardship

Open practices are ready to deliver safe and ethical data science and AI*

(*We just have a little more work to do to promote and normalise them!)

- Safety and sustainability
- Accountability
- Fairness
- Explainability
- Data stewardship

Who is building open source AI?

21 September, 14:00 - 15:30 UTC+1

Register on Eventbrite

Defining



Arielle Bennett

Programme Manager
The Alan Turing Institute



Mophat Okinyi

Union Representative
African Content Moderators Union



Marzieh Fadaee

Senior Research Scientist
Cohere for AI



Abinaya Mahendiran

CTO
Nunnari Labs



David Gray Widder

Postdoctoral Fellow
Cornell Tech



Jennifer Ding

Senior Researcher
The Alan Turing Institute



Hosted
by:



open source
initiative*

Open Source Initiative Deep Dive: Defining Open Source AI



Prof David Leslie
Director of Ethics and
Responsible Research
and Innovation



Victoria Kwan
Corporate Governance
Research Ethics
Manager



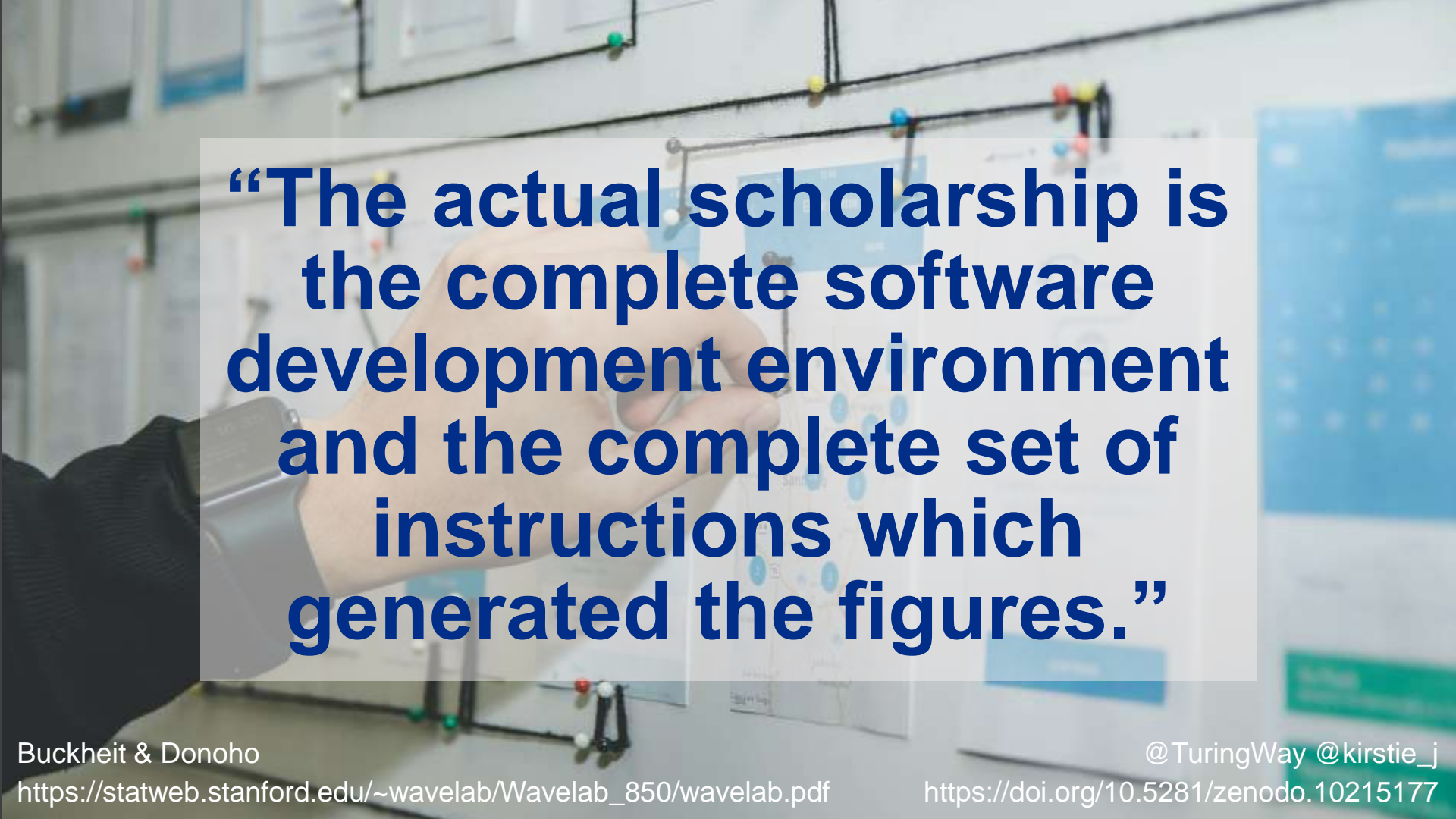
Dr Kirstie Whitaker
Director of Tools,
Practices and Systems
Research Programme

[https://opensource.org/deepdive/webinars/
operationalising-the-safe-d-principles-for-open-source-ai](https://opensource.org/deepdive/webinars/operationalising-the-safe-d-principles-for-open-source-ai)

<https://doi.org/10.5281/zenodo.10215177>
@TuringWay @kirstie_j

Open practices are ready to deliver safe and ethical data science and AI

- Safety and sustainability
- Accountability
- Fairness
- Explainability
- Data stewardship

A hand is pointing at a whiteboard. The whiteboard has several diagrams and charts pinned to it. The diagrams include flowcharts and data visualizations. The charts include a bar chart and a line graph. The background is a whiteboard with various papers and diagrams pinned to it. The text is overlaid on a semi-transparent white box.

“The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.”

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

Is your code doing what
you think its doing?



Is your code doing what
you think its doing?



Is your code doing what
you think its doing?

```
Assert.AreEqual(  
    GetTimeOfDay(),  
    "Morning" )
```



Is your code doing what
you think its doing?

```
Assert.AreEqual(  
    GetTimeOfDay(),  
    "Morning" )
```



Is your code doing what you think its doing?

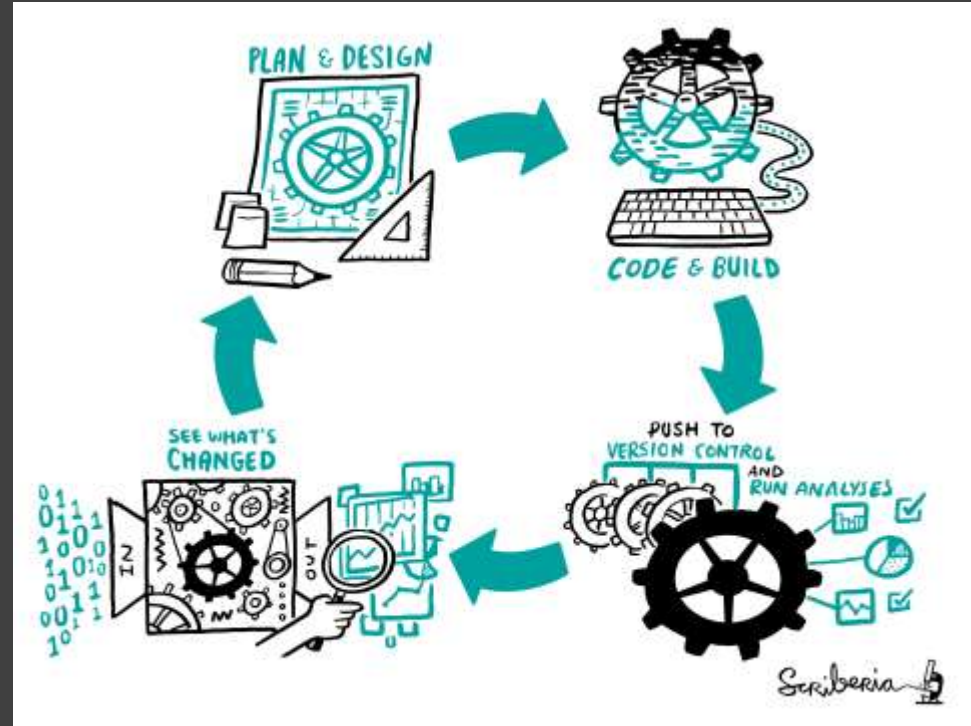
Testing sub-chapters on:

- Smoke
- Unit
- Integration
- System
- Acceptance and regression
- Runtime

And continuous integration...

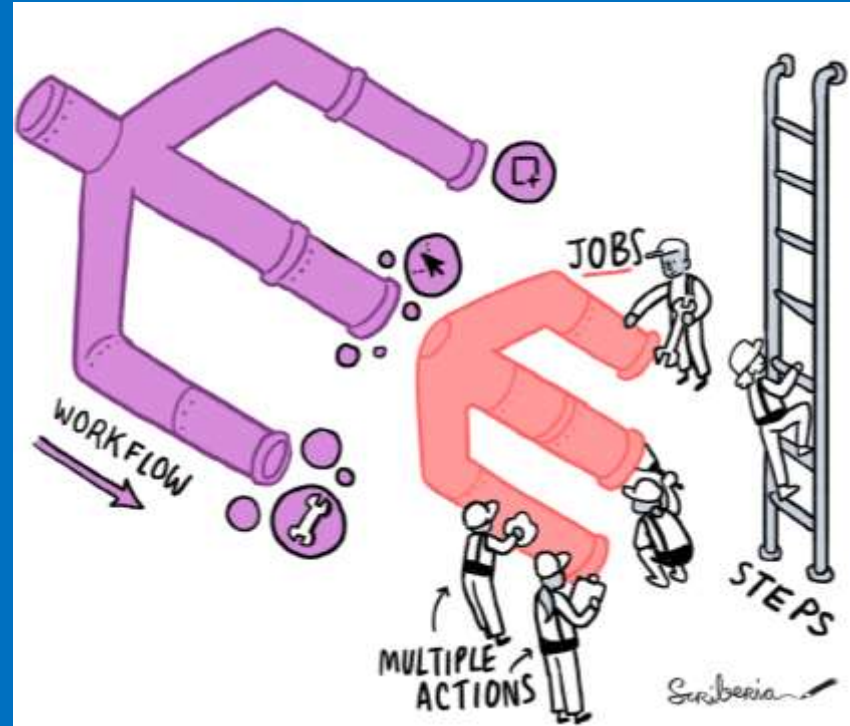


- Plan and design your experiment
- Write down those steps in code
- Push to version control and run the analyses
 - Traditionally done on the cloud, but the important part is that all steps are run every time
- Test to see what's changed



How to deliver explainable results?

- Code your data processing, analysis & visualisations. Share protocols for manual steps.
- Test to catch changes (planned and unplanned).
- Interrogate the analysis / model to understand how it behaves under difference circumstances.



<https://the-turing-way.netlify.app/reproducible-research>
<https://the-turing-way.netlify.app/reproducible-research/testing>
<https://the-turing-way.netlify.com/reproducible-research/ci.html>

@TuringWay @kirstie_j
<https://doi.org/10.5281/zenodo.10215177>

Open practices are ready to deliver safe and ethical data science and AI

- Safety and sustainability
- Accountability
- Fairness
- Explainability
- Data stewardship

A large field of glowing jack-o'-lanterns at night, with a semi-transparent text box in the center. The pumpkins are carved with various faces and are illuminated from within, creating a warm orange glow. Some pumpkins have blue or purple lights around them. The background is dark with some trees and foliage.

“Given enough eyeballs, all bugs are shallow”

Eric Steven Raymond, *The Cathedral and the Bazaar*
Linus' Law: https://en.wikipedia.org/wiki/Linus%27s_law

@TuringWay @kirstie_
<https://doi.org/10.5281/zenodo.10215177>

Our human needs
are one of the major
reasons that coding
in the open is so
challenging



@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

THERE'S MORE TO COLLABORATION

THAN YOU MIGHT THINK!



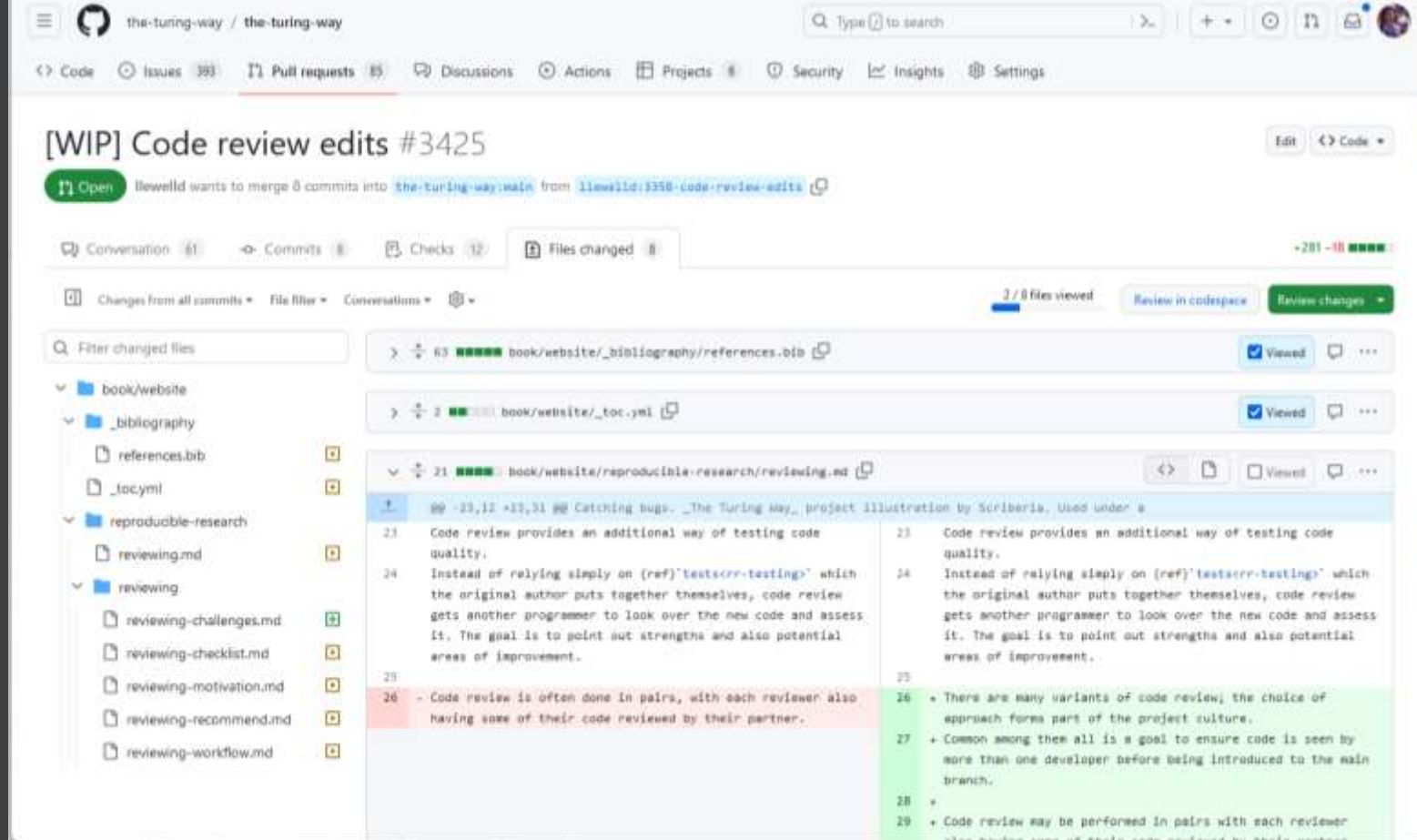
Sensibaia

@TuringWay @kirstie_j

How to get better at (code) review?

- Accept that someone else finding the bug is a success!
- Use a style guide so code is easy to review and understand.
- Build community, explain why code runs the way it does.
- Stay collaborative, work to a shared goal.
- Be kind.





<https://the-turing-way.netlify.app/reproducible-research/reviewing>
<https://github.com/the-turing-way/the-turing-way/pull/3425>

@TuringWay @kirstie_
<https://doi.org/10.5281/zenodo.10215177>

[WIP] Code review edits #3425

Open Changes from all commits File filter Conversations 2 / 8 files viewed Review in codespace Review changes

Filter changed files

- book/website
 - _bibliography
 - references.bib
 - _toc.yml
 - reproducible-research
 - reviewing.md
 - reviewing
 - reviewing-challenges.md
 - reviewing-checklist.md
 - reviewing-motivation.md
 - reviewing-recommend.md
 - reviewing-workflow.md

122 book/website/reproducible-research/reviewing/reviewing-challenges.md

(rr-reviewing-challenges)=

When Code Review Goes Wrong

As is discussed in the (ref) `Motivation` section, code reviews are important and beneficial for the overall health of a project. Nonetheless individual reviews can be challenging and are often a source of tension. Ensuring code review is effective and enriching for those involved relies on the author and reviewers approaching it in a certain way, and avoiding some of the many pitfalls involved.

When author and reviewers are working together and in collaboration, the exercise of reviewing can be both enjoyable and beneficial. When reviews become antagonistic they can cause immense harm.

```
---
width: 630px
name: code-review-good-and-bad
alt: Two panes, both show two rowers in a boat. In the top pane both row in the same direction and the boat makes
---
PLACEHOLDER IMAGE. Code review flow and frustration. _The Turing Way_ project illustration by Scriberia. Used under
```

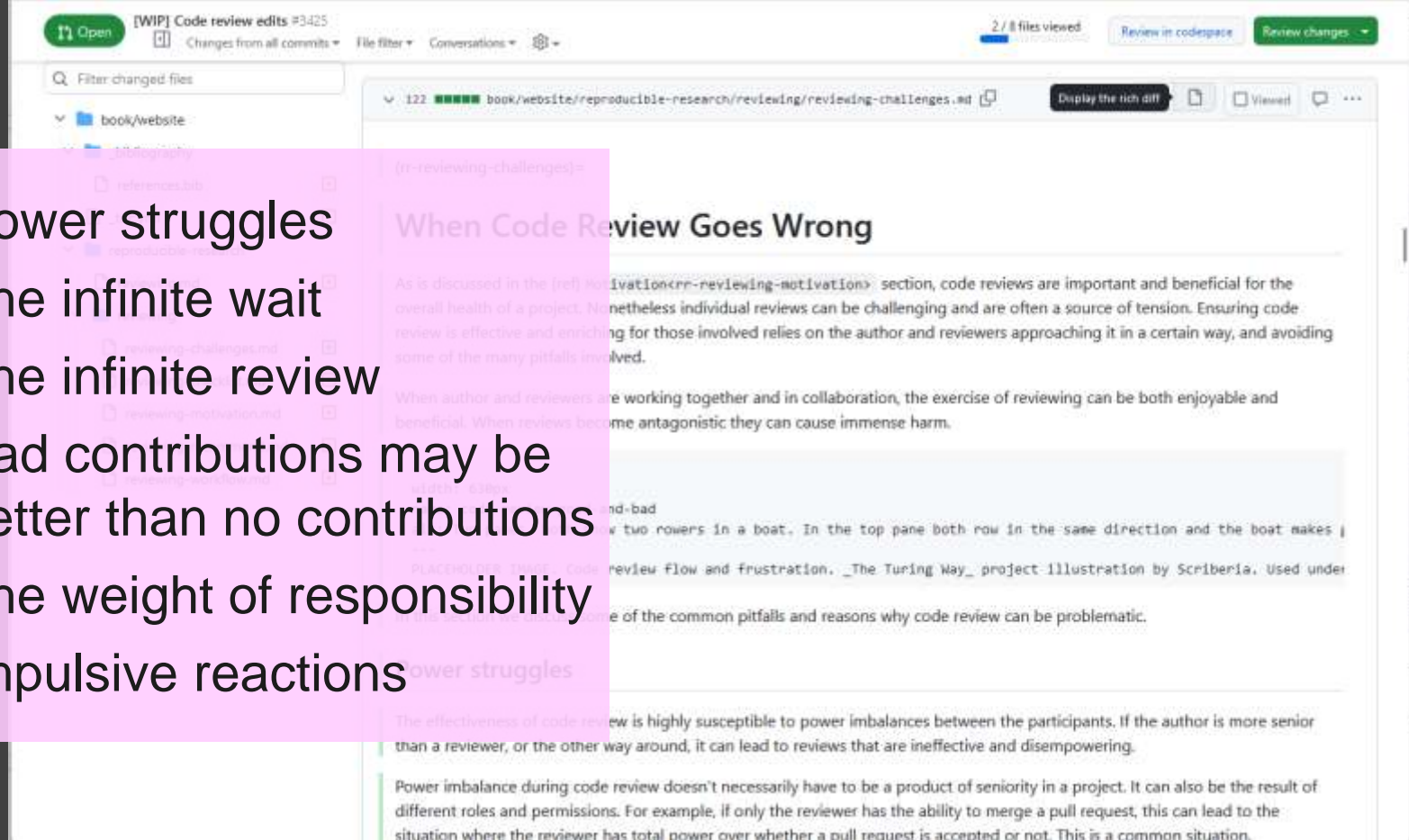
In this section we discuss some of the common pitfalls and reasons why code review can be problematic.

Power struggles

The effectiveness of code review is highly susceptible to power imbalances between the participants. If the author is more senior than a reviewer, or the other way around, it can lead to reviews that are ineffective and disempowering.


Power imbalance during code review doesn't necessarily have to be a product of seniority in a project. It can also be the result of different roles and permissions. For example, if only the reviewer has the ability to merge a pull request, this can lead to the situation where the reviewer has total power over whether a pull request is accepted or not. This is a common situation.

- Power struggles
- The infinite wait
- The infinite review
- Bad contributions may be better than no contributions
- The weight of responsibility
- Impulsive reactions

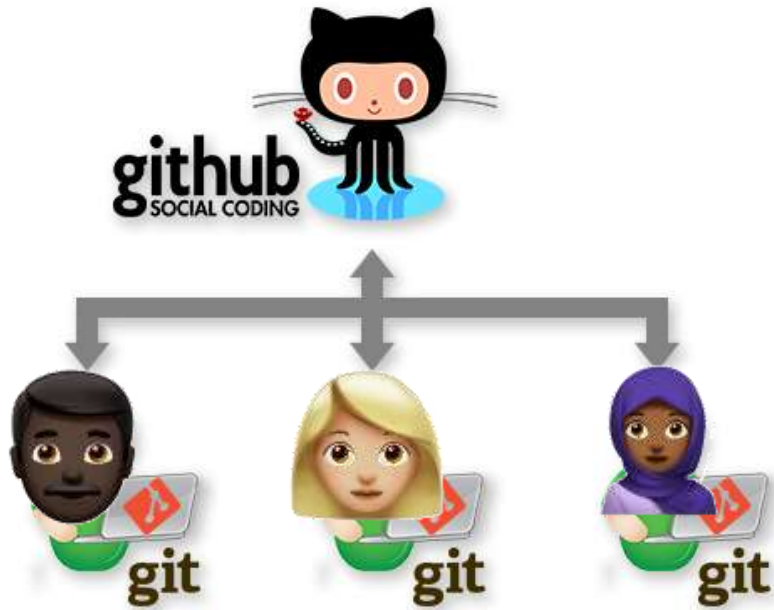


Open practices are ready to deliver safe and ethical data science and AI

- Safety and sustainability
- **Accountability**
- Fairness
- Explainability
- Data stewardship

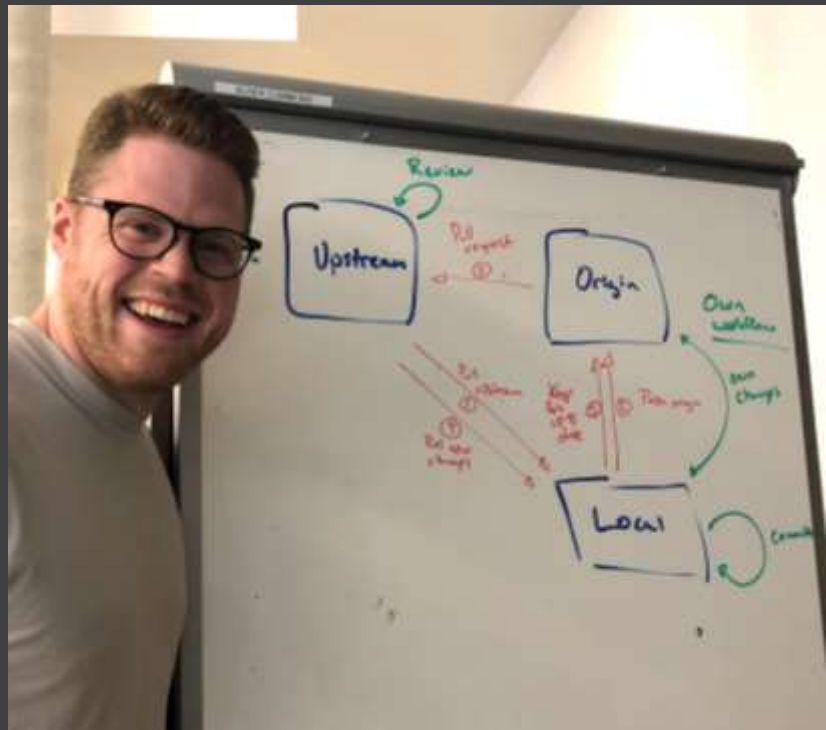
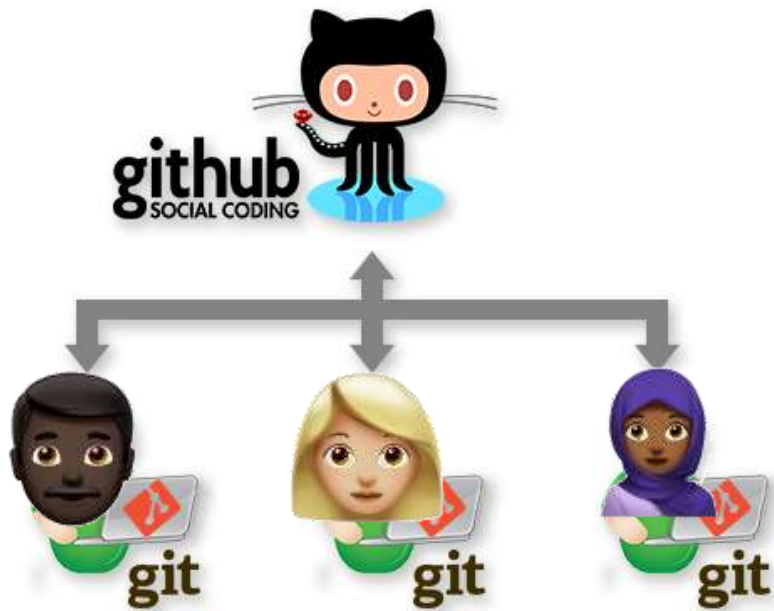
The background is a dark-themed Git commit history. It shows a vertical timeline of commits with colored markers (blue, pink, green) and commit messages. Some visible commit messages include "feat: dynamic directive arguments for v-on, v-bind and custom directives", "perf: improve scoped slots change detection accuracy", and "fix: fix checkbox event edge case in Firefox". A semi-transparent grey box is overlaid on the center of the image, containing white text.

“Anyone could say anything...nothing was official. And to emphasize the point, I labelled the notes ‘Request for Comments.’”



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

@TuringWay @kirstie_j
<https://doi.org/10.5281/zenodo.10215177>



<https://the-turing-way.netlify.app/collaboration/github-novice.html>
<https://the-turing-way.netlify.com/reproducible-research/vcs.html>
<https://neurohackademy.org>

@TuringWay @kirstie_j
<https://doi.org/10.5281/zenodo.10215177>

How to get better at co-creation?

- Document requests for input in issues, and decisions somewhere easy to find.
- Be clear about communication channels, “don’t make decisions in the kitchen”.
- Plan for misunderstandings: have a code of conduct and an enforcement plan.



Open practices are ready to deliver safe and ethical data science and AI

- Safety and sustainability
- Accountability
- **Fairness**
- Explainability
- Data stewardship



“You don’t need a diversity committee if you have a diverse committee”

@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

Teon Brooks, <https://teonbrooks.com>

About AutSPACES

A **research collaboration** between
The Turing, Autistica, & Open Humans

A **community** of autistic people, their
supporters, researchers and open
source developers

A **citizen science** platform where
people can share their experiences of
sensory processing.

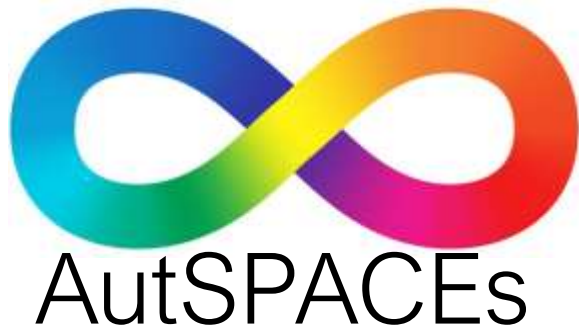
The
Alan Turing
Institute



@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

<https://doi.org/10.5281/zenodo.8164375>



Collect qualitative data to improve our understanding of sensory processing in people's daily lives

Share people's stories and adaptive techniques with others who have similar experiences

Educate neurotypical people to better support their friends, family and colleagues

Advise organisations on how they can design and adapt spaces to improve people's lives

@TuringWay @kirstie_j

Moderation Goals



James Scott



Susanna
Fantoni

A welcoming,
inclusive, and
safe online
space for
autistic people

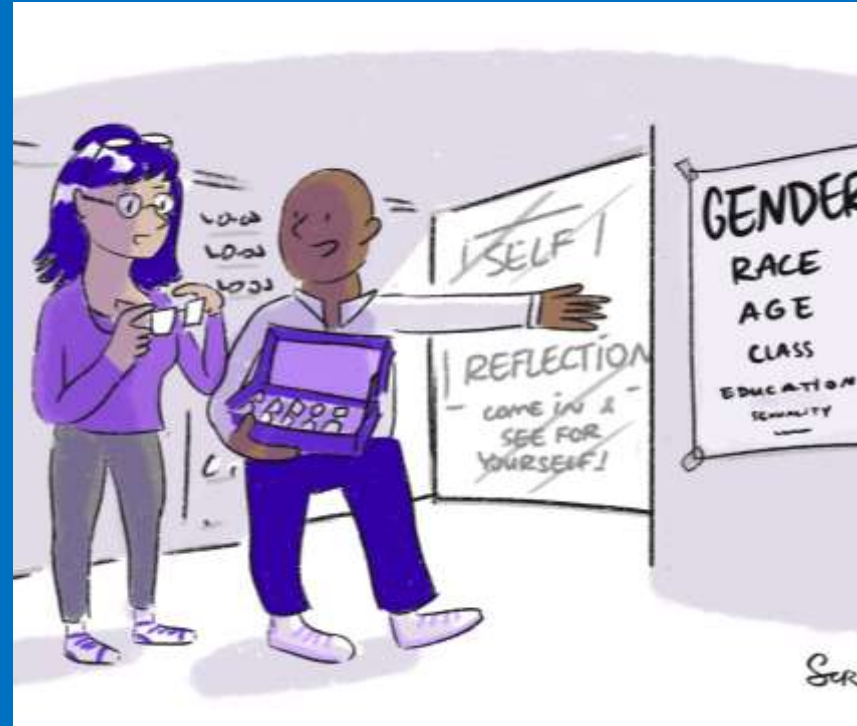
A space where
autistic people
can share
challenging
experiences

Decisions
which balance
the complex
needs of the
community

How to build more useful outputs?

- Consider identity and positionality.
- Consider power and privilege.
- Consider intersectionality.
- Consider your incentives.

**Fairness to all under capitalism
may not be achievable.
Make the compromise with
open eyes.**



@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

<https://the-turing-way.netlify.app/ethical-research/self-reflection>

Open scholarship practices are ready to deliver safe and ethical AI

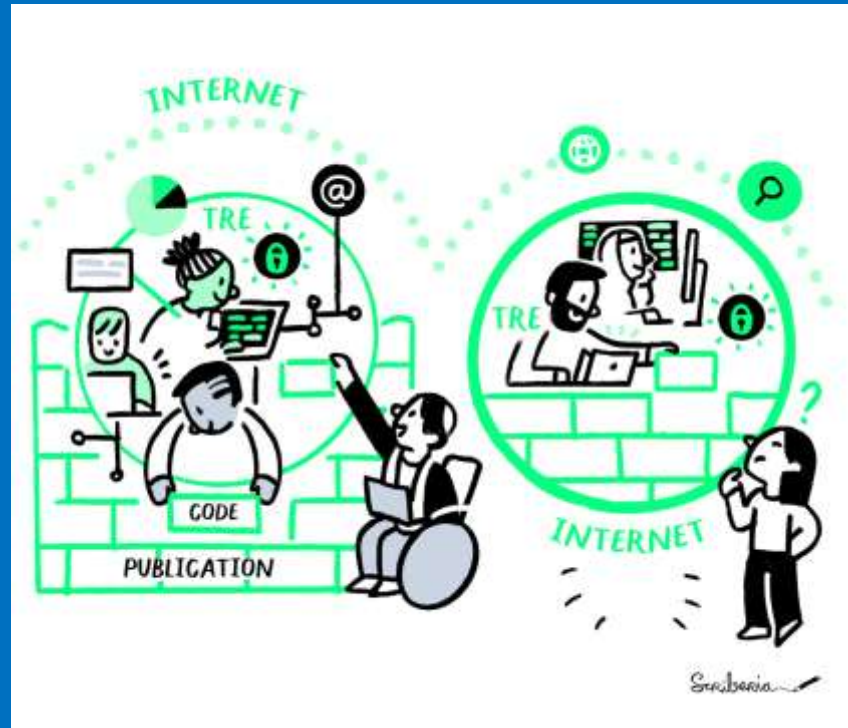
- Safety and sustainability
- Accountability
- Fairness
- Explainability
- **Data stewardship**



“Digital assets should be findable, accessible, interoperable, and reusable”

How to code in the open with sensitive data?

- Make the metadata open.
- Consider sharing “fake” data (simulated, synthetic or random).
- Share pathways to getting access to the data.
- Make the code easy to run on a local machine.



@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

<https://the-turing-way.netlify.app/reproducible-research/rdm>

“

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**.

”

@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

Please join us in The Turing Way!

- 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>
- Collaboration cafes on the 1st and 3rd Wednesday of each month. Next one on **6 December, 3pm – 5pm UTC**.
- 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 @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>

Thank you

- **TPS Senior Researchers:**

Malvika Sharan, Aida Mehonic,
Alden Conner, Jennifer Ding,
Bastian Greshake Tzovaras &
Christopher Burr

- **Research Community**

Managers: Emma Karoune,
Vicky Hellon, Anne Lee Steele,
Sophia Batchelor, Eirini Zormpa,
Gabin Kayumbi, Cassandra
Gould Van Praag

- **GC Research Ethics Manager:**

Victoria Kwan

- **Research Application**

Managers: Hari Sood, Shakir
Laher, Kalle Westerling, Sophie
Arana & Francisco Gomez
Medina

- **Programme management:**

Arielle Bennett, Dave Chapman,
Davide Sarmiento Perez, Batool
Almarzouq, Alexandra Araujo
Alvarez, Sydney Ambrose

- **Research associates and**

fellows: Georgia Aitkenhead,
Giulia Tomba, Priscilla Canizares

@TuringWay @kirstie_j

<https://doi.org/10.5281/zenodo.10215177>



Scriberia 

@TuringWay @kirstie_j