

The Alan Turing Institute

The Turing Way: A Digital Commons for Open Science and Reproducibility

Cite.Software Pre-Workshop - 18 September

Malvika Sharan

Pronouns: she/her/hers



The Turing Way

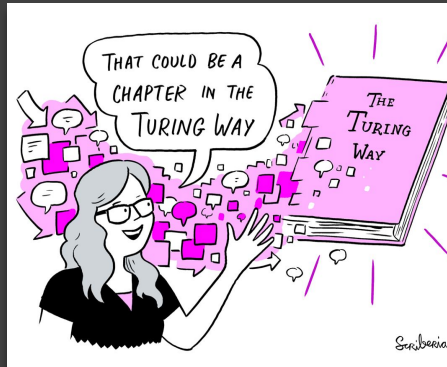
A community-led project on data science.

We are an open source, open collaboration and globally distributed community working together to **make data science accessible, reproducible and beneficial.**

The Turing Way

A community-led project on data science.

We are an open source, open collaboration and globally distributed community working together to **make data science accessible, reproducible and beneficial.**



Co-created Book



Community-based



Participatory process



Culture change

Started in 2019 as a book on Reproducibility

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

Risk Assessment

Case Studies

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](#) 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.



Open and reproducible research along with good collaboration, communication and ethics should be seen as fundamental research methods.

The Turing Way: Guides on Data Science



Kirstie Whitaker

Director: Tools,
Practices & Systems
Programme (TPS)



Malvika Sharan

TPS Senior
Researcher: Open
Research & Community



Anne Lee Steele

Community Manager,
The Turing Way



Alexandra
Araujo Alvarez

Project Manager

4+ years

>300 Chapters

Numerous training
and community
resources.

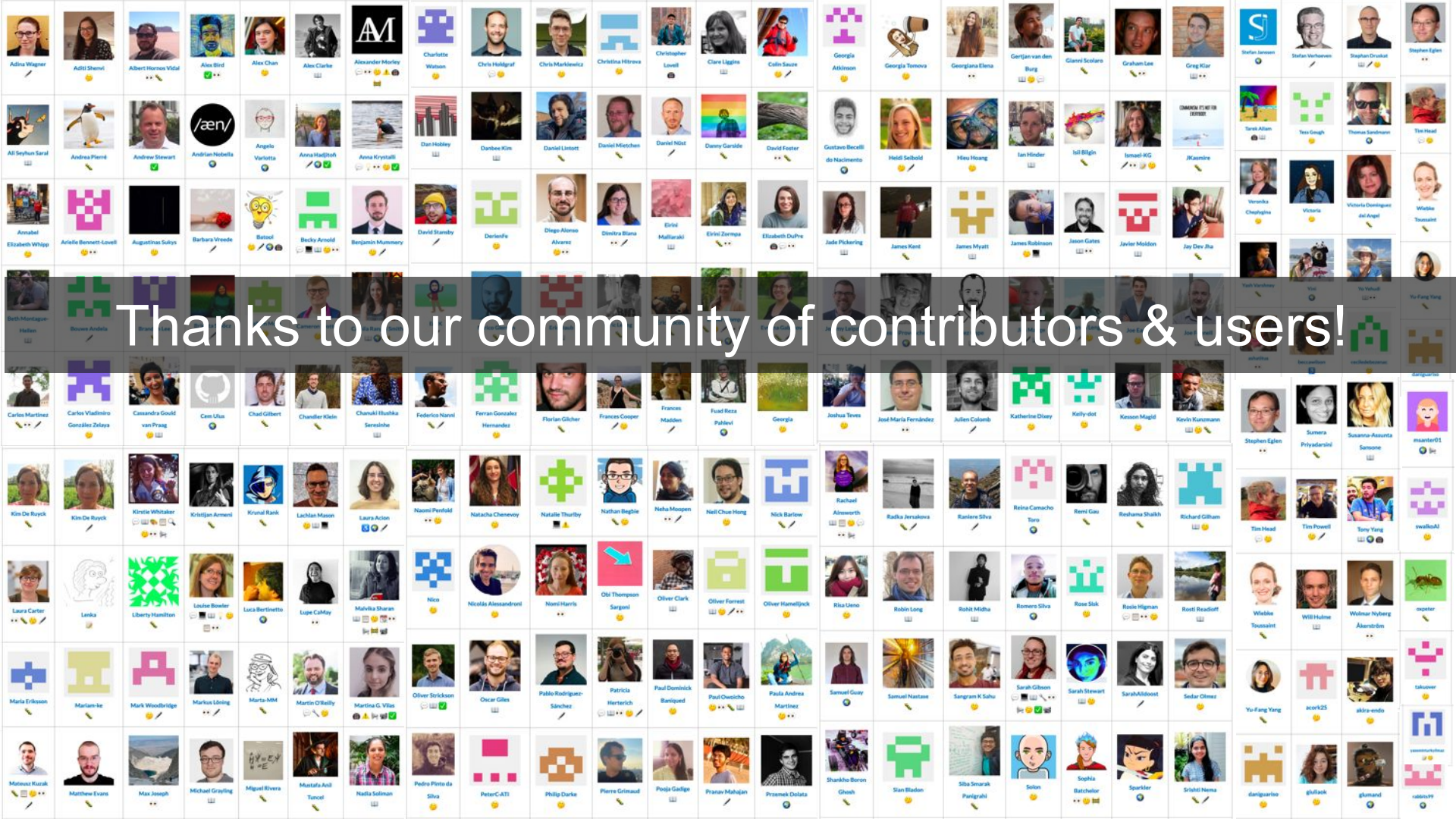


The Turing Way

🔍 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 ▾



Thanks to our community of contributors & users!

Project Growth & Recognitions

- 450+ contributors, 3500+ monthly users - supported by 25+ core members (governance)
- 2 awards and engagement with 35+ orgs
- References in 40+ peer reviewed articles and 100s of publications, reports, policy documents
- Replication/extension by 10+ Turing & external projects

 19,098 views 16,710 downloads



Belongs to the community

This is **not** a Turing project - although it has great support from the Institute.

This is **always a work in progress** - evolving with the needs in the community.

We are creating community resources together -- **it's about the way, the journey, not a set of rules.**



The Turing Way as a Digital Commons



Elinor Ostrom (1933-2012)
2009 Nobel Laureate

Knowledge commons: Information, data, and content that is **collectively owned and managed by a community of users** - without depleting their quantity or quality.

The Turing Way as a Digital Commons



Elinor Ostrom (1933-2012)
2009 Nobel Laureate

Knowledge commons: Information, data, and content that is collectively owned and managed by a community of users - without depleting their quantity or quality.

Digital commons are global digital resources produced and maintained **collectively and in a decentralized manner**. Use licensing, authorship, peer production, governance and **participatory ways to fostering equitable access** to resources.

📁 .github	Remove prettier configuration
📁 book	minor update
📁 communications	Fix typos
📁 conferences	Add KW formatting pedantry
📁 project_management	Update online-collaboration-cafe.md
📁 templates	Updating Github templates
📁 tests	Add "et cetera" as a deprecated Latinism
📁 workshops	Remove mis-pasted text

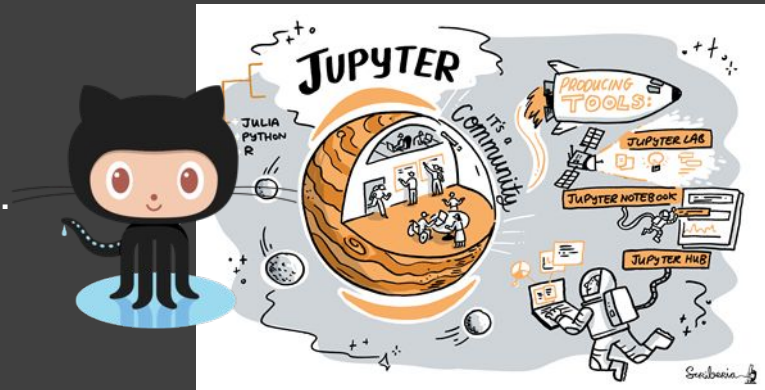
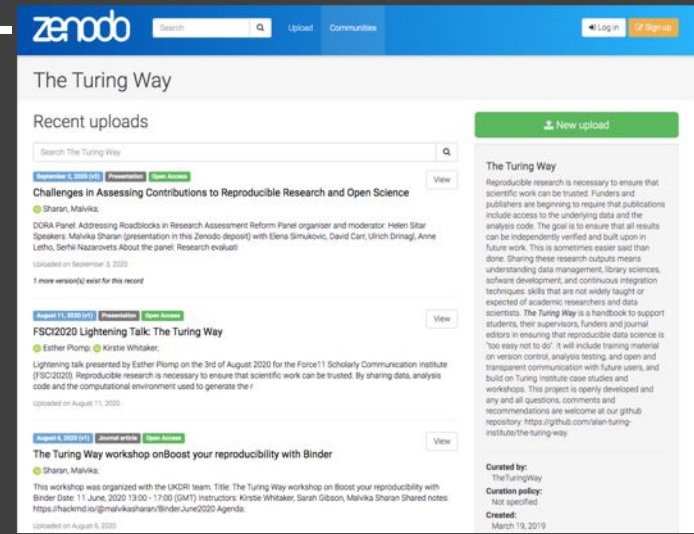


📄 .all-contributorsrc	Merge pull request #991 from alan-turing-institute/all-contributors/a...	5 days ago
📄 .gitignore	ignore pptx in workshop folder	9 months ago
📄 .travis.yml	add html-proof file again	last month
📄 <u>CODE_OF_CONDUCT.md</u>	her -> their	6 months ago
📄 <u>CONTRIBUTING.md</u>	Update CONTRIBUTING.md	2 months ago
📄 GOVERNANCE.md	Read through months later	5 months ago
📄 LICENSE.md	Fix typo in licence	2 months ago
📄 <u>README.md</u>	Merge pull request #991 from alan-turing-institute/all-contributors/a...	5 days ago
📄 book_skeleton.md	Update book_skeleton.md	13 months ago
📄 <u>contributors.md</u>	Add myself to contributors.md	11 months ago
📄 tips_and_tricks_survey.md	Update tips_and_tricks_survey.md	14 months ago
📄 ways_of_working.md	Adjust team contact section	5 months ago

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

An open source project

- Everyone can freely read, reuse, distribute, modify and help develop.
- The project belongs to *The Turing Way* community.
- Built on open source infrastructure:
 - Git, Jupyter Book, Binder, project bots etc.
 - Nelfy: [the-turing-way.netlify.app/welcome](https://zenodo.org/communities/the-turing-way)



The Turing Way Guides



Reproducibility



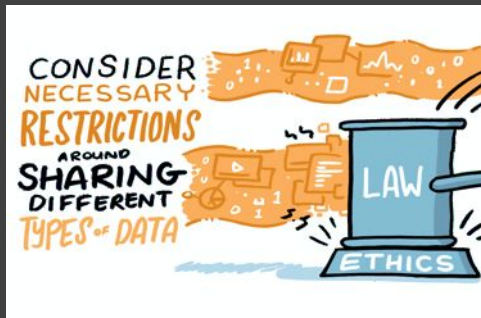
Project Design



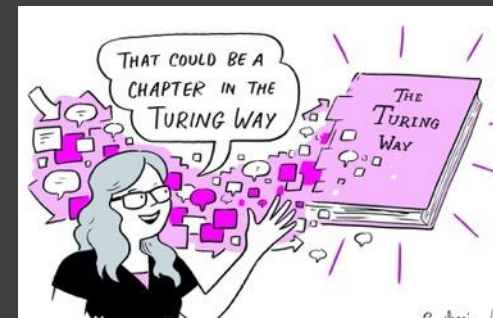
Communication



Collaboration



Ethical Research



Community Handbook

The Turing Way Guides



Top Tip

The Turing Way is not meant to be read from start to finish. Start with a concept, tool or method that you need now, in your current work. Browse the different guides that make up the book, or use the search box to search for whatever you would like to learn about first.



Collaboration



Ethical Research



Community Handbook

The Turing Way Guides



💡 Top Tip

The Turing Way is not meant to be read from start to finish. Start with a concept, tool or method that you need now, in your current work. Browse the different guides that make up the book, or use the search box to search for whatever you would like to learn about first.

Easier said than done! We don't know what we don't know.



Collaboration



Ethical Research



Community Handbook

Dedicated Pathways with Curated Chapters

deploy-preview-3277--the-turing-way.netlify.app/#different-pathways

Different Pathways

Early Career Researchers

- [Guide for Project Design](#)
- [Getting Started With GitHub](#)
- [Creating Project Repositories](#)

And more...

Project Leaders

- [Open Leadership in Data Science](#)
- [Guide for Project Design](#)
- [Creating Project Repositories](#)

And more...

Research Software Engineers

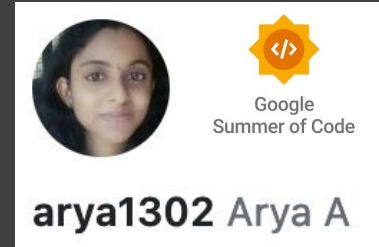
- [Research Infrastructure Roles: Introduction](#)
- [Citing Research Objects](#)
- [Advanced Structure for Data Analysis](#)

And more...

Software Citation

- [Research Object to capture the Research Life Cycle](#)
- [Making Research Objects Citable](#)
- [Steps for Making Research Objects Citable](#)

And more...



Johanna Bayer: GSoC
Co-Mentor, 2023

Research Engineers: Ian
Stanson, Lydia France,
May Yong (2022)

Dedicated Pathway for Software Citation

deploy-preview-3277--the-turing-way.netlify.app/software-citation

Software Citation

- Research Object to capture the Research Life Cycle
- Making Research Objects Citable
- Steps for Making Research Objects Citable

And more...

Software Citation

All sectors, from research to government, strongly rely upon the use of the software. This curated set of chapters guides the sharing and citation of software components from source code to computational workflows and executables while attributing all contributors.

Table Of Contents

These are the pages curated for SOFTWARE CITATION

- Guide for Reproducible Research
 - Overview of Reproducible Research
 - Open Research
 - Open Source Software
 - Licensing
 - Software Licenses
 - Research Data Management
 - Documentation and Metadata
 - Sharing and Archiving Data
 - Reusable Code
 - Overview of Code Reuse
 - Research Compendia
 - Research Object to capture the Research Life Cycle
 - Making Research Objects Citable
 - Steps for Making Research Objects Citable
 - Citing Research Objects
 - Software Citation with CITATION.cff
 - Create a CITATION.cff using cffinit

Software Citation with CITATION.cff

pathway Software Citation

The [Citation File Format](#) lets you provide citation metadata, for software or datasets, in plaintext files that are easy to read by both humans and machines.

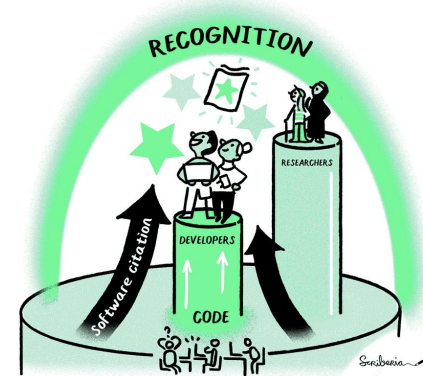
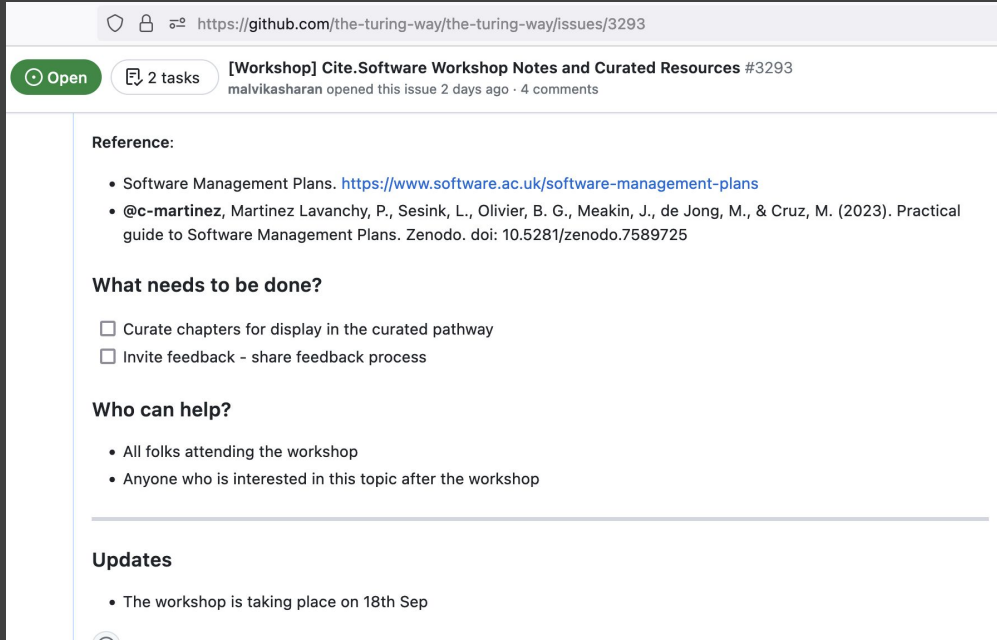


Fig. 71 Research software developers get recognition by making software citable. *The Turing Way* project illustration by Scriberia. Zenodo. <http://doi.org/10.5281/zenodo.3332807>

How can we collaborate?

[github.com/the-turing-way/
the-turing-way/
issues/3293](https://github.com/the-turing-way/the-turing-way/issues/3293)



The screenshot shows a GitHub issue page. At the top, the browser address bar displays the URL: <https://github.com/the-turing-way/the-turing-way/issues/3293>. Below the address bar, there is a navigation bar with a green 'Open' button, a '2 tasks' indicator, and the issue title: '[Workshop] Cite Software Workshop Notes and Curated Resources #3293'. Below the title, it says 'malvikasharan opened this issue 2 days ago · 4 comments'. The main content area is divided into sections: 'Reference:' with two bullet points, 'What needs to be done?' with two unchecked checkboxes, 'Who can help?' with two bullet points, and 'Updates' with one bullet point.

Reference:

- Software Management Plans. <https://www.software.ac.uk/software-management-plans>
- @c-martinez, Martinez Lavanchy, P., Sesink, L., Olivier, B. G., Meakin, J., de Jong, M., & Cruz, M. (2023). Practical guide to Software Management Plans. Zenodo. doi: 10.5281/zenodo.7589725

What needs to be done?

- Curate chapters for display in the curated pathway
- Invite feedback - share feedback process

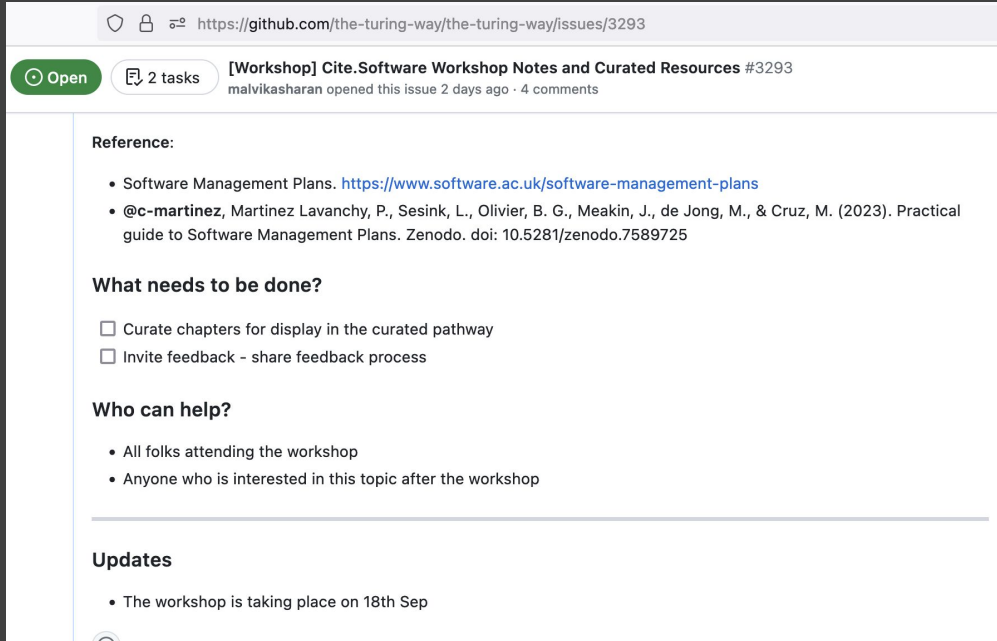
Who can help?

- All folks attending the workshop
- Anyone who is interested in this topic after the workshop

Updates

- The workshop is taking place on 18th Sep

How can we collaborate?



The screenshot shows a GitHub issue page for the repository 'the-turing-way/the-turing-way'. The issue title is '[Workshop] Cite Software Workshop Notes and Curated Resources #3293', opened by malvikasharan 2 days ago with 4 comments. The issue is marked as 'Open' and has '2 tasks'. The content includes a 'Reference' section with a link to 'Software Management Plans' and a citation for '@c-martinez'. A 'What needs to be done?' section contains two unchecked checkboxes: 'Curate chapters for display in the curated pathway' and 'Invite feedback - share feedback process'. A 'Who can help?' section lists 'All folks attending the workshop' and 'Anyone who is interested in this topic after the workshop'. An 'Updates' section notes 'The workshop is taking place on 18th Sep'.

github.com/the-turing-way/the-turing-way/issues/3293

Option 1: We will use a **shared document** to capture all ideas.

bit.ly/2023-cite-software

- Curated practices
- Centralised chapters pointing to dispersed resources
- CC-BY 4.0, hence repurposed with fair attribution.

How can we collaborate?

The screenshot shows a GitHub issue page for "[Workshop] Cite.Software Workshop Notes and Curated Resources #3293". The page is open in a browser with the URL <https://github.com/the-turing-way/the-turing-way/issues/3293>. The issue title is "[Workshop] Cite.Software Workshop Notes and Curated Resources #3293" and it was opened by malvikasharan 2 days ago. The issue has 2 tasks and 4 comments.

Reference:

- Software Management Plans. <https://w>
- @c-martinez, Martinez Lavanchy, P., S guide to Software Management Plans.

What needs to be done?

- Curate chapters for display in the cura
- Invite feedback - share feedback proc

Who can help?

- All folks attending the workshop
- Anyone who is interested in this topic

Updates

- The workshop is taking place on 18th S

Comments:

- c-martinez** commented 44 minutes ago (Collaborator):
Other sections that relate to RSE:
<https://the-turing-way.netlify.app/reproducible-research/testing>
<https://the-turing-way.netlify.app/reproducible-research/reviewing>
<https://the-turing-way.netlify.app/reproducible-research/renv>
<https://the-turing-way.netlify.app/reproducible-research/ci>
<https://the-turing-way.netlify.app/reproducible-research/code-quality>
- malvikasharan** commented 19 minutes ago (Member, Author):
Thank you @c-martinez -- I am going to add these to the curation, which will be currently being displayed as the 4th pathway here: <https://deploy-preview-3277--the-turing-way.netlify.app/#different-pathways>
- c-martinez** commented 9 minutes ago (Collaborator):
Good luck with the Workshop! I am not sure if I will be able to join, but I am sure it will be great!

The bottom of the screenshot shows a text input field for leaving a comment and a rich text editor toolbar with options for bold, italic, underline, link, and other formatting.

github.com/the-turing-way/the-turing-way/issues/3293

Option 1: We will use a **shared document** to capture all ideas.

Option 2: If you have a GitHub (GH) account - please directly comment or **collaborate on GH**.

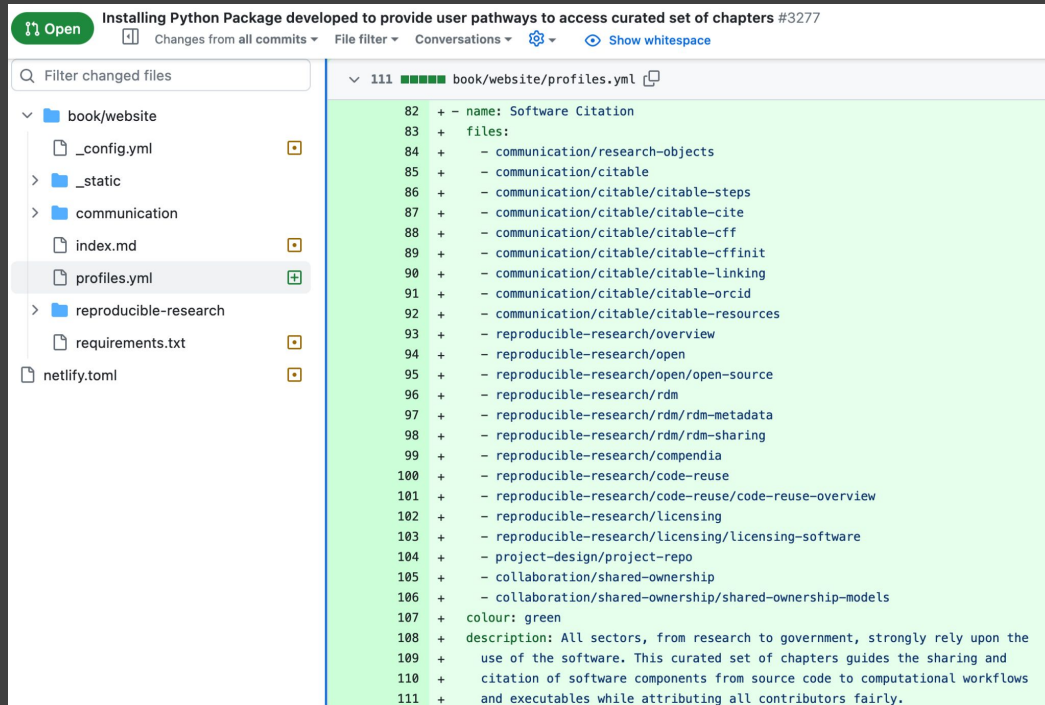
Option 3: Create a GH account and indulge us! :) It will also **allow us to attribute openly!**

Some next steps

- Identify what exists and display them in the Software Citation pathway

Some next steps

- Identify what exists and display them in the Software Citation pathway



```
82 + - name: Software Citation
83 +   files:
84 +     - communication/research-objects
85 +     - communication/citable
86 +     - communication/citable/citable-steps
87 +     - communication/citable/citable-cite
88 +     - communication/citable/citable-cff
89 +     - communication/citable/citable-cffinit
90 +     - communication/citable/citable-linking
91 +     - communication/citable/citable-orcid
92 +     - communication/citable/citable-resources
93 +     - reproducible-research/overview
94 +     - reproducible-research/open
95 +     - reproducible-research/open/open-source
96 +     - reproducible-research/rdm
97 +     - reproducible-research/rdm/rdm-metadata
98 +     - reproducible-research/rdm/rdm-sharing
99 +     - reproducible-research/compedia
100 +     - reproducible-research/code-reuse
101 +     - reproducible-research/code-reuse/code-reuse-overview
102 +     - reproducible-research/licensing
103 +     - reproducible-research/licensing/licensing-software
104 +     - project-design/project-repo
105 +     - collaboration/shared-ownership
106 +     - collaboration/shared-ownership/shared-ownership-models
107 +   colour: green
108 +   description: All sectors, from research to government, strongly rely upon the
109 +     use of the software. This curated set of chapters guides the sharing and
110 +     citation of software components from source code to computational workflows
111 +     and executables while attributing all contributors fairly.
```

github.com/the-turing-way/the-turing-way/pull/3277

File with chapters:
[book/website/profiles.yml](#)

Some next steps

- Identify what exists and display them in the Software Citation pathway
- (Not a blocker) Merge the pathway branch (currently being reviewed)!
- Identify gaps based on resources we have curated
- Create short overview chapters pointing to the dispersed resources
- Include them in appropriate guides, add them to the pathway
- Repurpose them with attribution (CC-BY 4.0) ← with eScience center

Fundamental considerations for the Commons

Information, data, and content that is collectively owned and managed by a community of users - without depleting their quantity or quality.

...participatory ways to fostering equitable access to shared resources.

- To **whom does knowledge/privilege belong?**
- **Who benefits** from the production and circulation of knowledge?
- **Who gets to participate** in the knowledge production processes?
- In what ways can systems be used to **increase the agency of more people** over knowledge production, circulation and benefits?

Recognition for all Contributors

Afterword

Glossary

Bibliography

Contributors Record



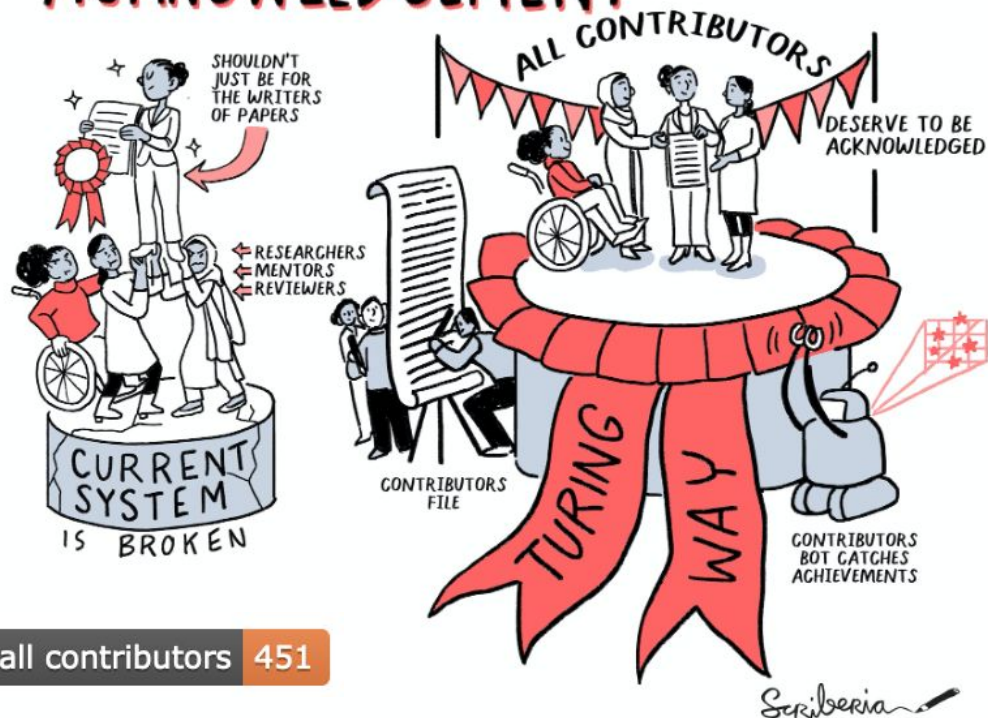
All Contributors

Recognize All Contributors

Including those that don't push code

Cite as: *The Turing Way Community. (2022). The Turing Way: A handbook for reproducible, ethical and collaborative research (1.0.2). Zenodo. <https://doi.org/10.5281/zenodo.7625728>*

ACKNOWLEDGEMENT



all contributors 451

Many Pathways for Contribution



Develop



Maintain & improve



Share resources



Review and update



Make it global

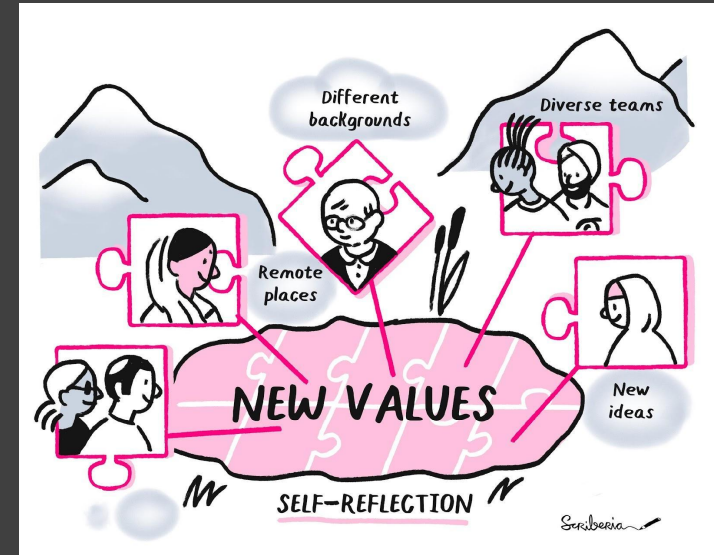


Share best practices

The Turing Way Channels

- *Translation co-working*: Tuesday (15:00 UK)
- *Collaboration Cafes*: 1st & 3rd Wednesdays
- *Infrastructure & accessibility* working groups
- *Book Dash*: May and November yearly
- *Outreachy with Open Source Saudi Arabia, GSoC and GSoD with INCF*

Your contribution is important - *we record and acknowledge* in the README file and in the book.

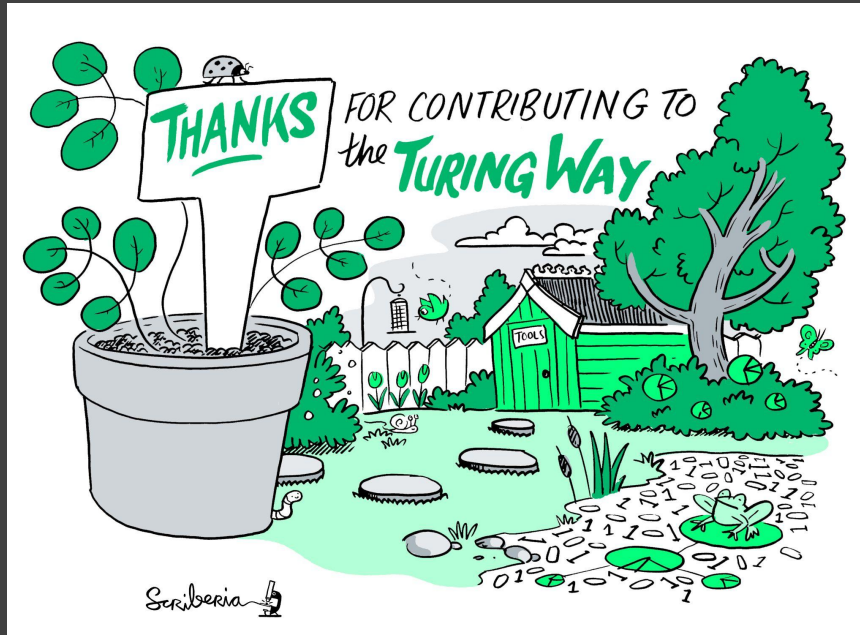


Growing *The Turing Way*: moving at the speed of trust

- Shifting from a grassroots open science project to a **global collaboration takes intentional care**
- Working towards shifting mindset to **sustainability first** in infrastructure as well as environment.

Together maintaining the digital commons





Move fast and break things

Collaborate, move slowly, maintain things

X

Acknowledgements & Links

The
Alan Turing
Institute

The Turing Way Community

- Jupyter Book Team, Infrastructure Maintenance Working Group
- Arya A. Johanna Bayer, an Stenson, Lydia France, May Yong
- Kirstie Whitaker, Anne Lee Steele, Alexandra Araujo Alvarez + many Turing Way Core Team and Turing's TPS team members

Links and resources

- **Book:** the-turing-way.netlify.com, **Twitter:** twitter.com/turingway
- **GitHub:** github.com/alan-turing-institute/the-turing-way
- **Useful links & opportunities are listed here:** <https://bit.ly/turingway>

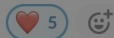


Images used in this talk have been co-developed by Scriberia and The Turing Way community, shared under CC-BY 4.0. <https://zenodo.org/record/7587336>



Anne Lee Steele (she/her) 2 days ago

Just used the phrase "I'm the Community Manager of The Turing Way, an open source project and community that documents, maintains, and translates best practices in data science and AI." to describe The Turing Way to someone (my first time using this particular wording)... and it made me think of the one-liners we use to describe the project. What other one-liners do people use? How have they changed in the past few years?



What other one-liners do people use?

4 replies



Anne Lee Steele (she/her) 1 day ago

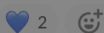
A few ones that come to mind are:

- "The Turing Way is an community-led handbook to best practices in data science"
- "The Turing Way aims to make reproducibility too easy not to do by collecting best practices into a collection of open source guides"
- "The Turing Way involves and supports a diverse community to make data science reproducible, ethical, collaborative, and inclusive for all" (taken from our default slide deck)



Malvika Sharan 1 day ago

Adding one more: The Turing Way is an open source handbook on reproducibility, we are a global community for everyone.



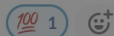
I say something different about The Turing Way every time I mention it! that the key is curating what I say to that particular listener!



KirstieJane 11 hours ago

I've been thinking about this question. I feel like I say something different about The Turing Way every time I mention it!

There's something about it being so many things that the key is curating what I say to that particular listener!



means different things and becomes what they do. #FeatureNotABug?



Malvika Sharan 9 hours ago

Based on the "way" of the Turing Way that people take, it means different things and becomes what they do. #FeatureNotABug? 🤖 (edited)

