The Alan Turing Institute

The Turing Way Book Dash

The #TuringWay team

University of Manchester: 17 May 2019

http://bit.ly/book-dash-mcr

Agenda

http://bit.ly/book-dash-mcr

8:30	Arrive at dashing venue
9:00	Intro to the book dash
9:30	Team forming and brain storming
10:30	Coffee break
11:00	Group working
12:30	Lunch
13:30	Group working, coffee available at 15:00
16:00	Celebrate successes (group share outs)
17:00	Close

Founding the Institute

"We will found The Alan Turing Institute to ensure Britain leads the way again in the use of big data and algorithm research"

George Osborne, Chancellor of the Exchequer Budget Speech, March 2014

The Alan Turing Institute



Network of industry, charity, government partners

Network of university members

Strategic government investment

The Institute's partners and collaborators

























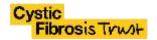






















16/05/2019 The Alan Turing Institute

4

Our university network



















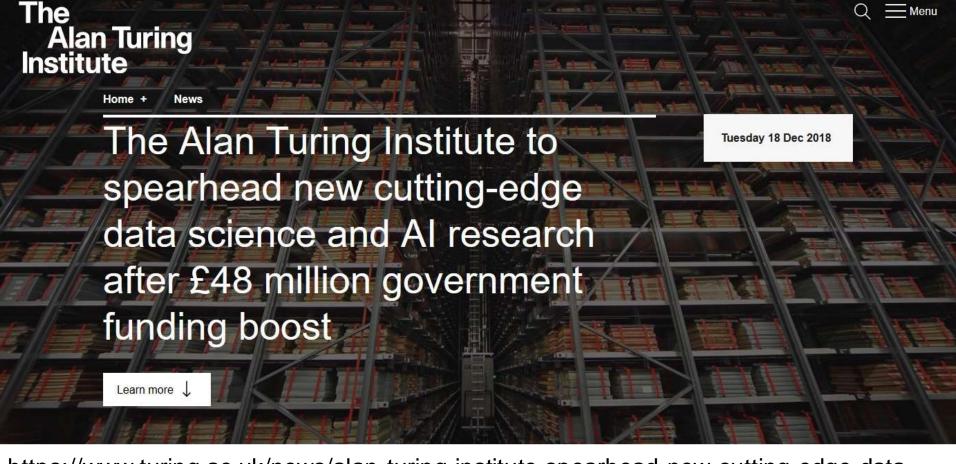












https://www.turing.ac.uk/news/alan-turing-institute-spearhead-new-cutting-edge-data-science-and-artificial-intelligence



Urban analytics

Developing data science and AI focused on the process, structure, interactions and evolution of agents, technology and infrastructure within and between cities.



Data-centric engineering

Bringing together world-leading academic institutions and major industrial partners from across the engineering sector, to address new challenges in data-centric engineering.



Data science for science

Ensuring that research across science and the humanities can make effective use of state of the art methods in artificial intelligence and data science.

Cross cutting theme: Tools, systems and practices



Health

Accelerating the scientific understanding of human disease and improving human health through data-driven innovation in Al and statistical science.



Public policy

Working with policy makers on data-driven public services and innovation to solve policy problems, and developing ethical foundations for data science and Al policy-making.



Research Engineering

Connecting research to applications, helping create usable and sustainable tools, practices and systems.

https://www.turing.ac.uk/research/ai-science-engineering-health-and-government



The Alan Turing Institute

The Turing Way

A lightly opinionated handbook for reproducible data science

https://github.com/alan-turing-institute/the-turing-way

What does reproducible mean?

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

https://dx.doi.org/10.6084/m9.figshare.7140050

Why don't people do this already?

Is not considered for promotion

Takes time

Barriers to reproducible research

Publication bias towards novel findings

Requires additional skills Plead the 5th

Support additional users

Held to higher standards than others

https://dx.doi.org/10.6084/m9.figshare.7140050

How can the Turing Way help?

Make reproducibility, "too easy not to do"

Share the responsibility of reproducibility

The Turing Way

- 1. Introduction
- 2. Reproducibility
- 3. Open Research
- 4. Version Control
- 5. Collaborating on GitHub/GitLab
- 6. Research Data Management
- 7. Reproducible Environments
- 8. Testing
- 9. Reviewing
- 10. Continous Integration
- Reproducible Research with Make
- 12. Risk Assessment

Powered by Jupyter Book

Welcome to the Turing Way

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.

A bit more background

Reproducible research is necessary to ensure that scientific work can be trusted. Funders and publishers are beginning to require that publications include access to the underlying data and the analysis code. The goal is to ensure that all results can be independently verified and built upon in future work. This is sometimes easier said than done. Sharing these research outputs means understanding data management, library sciences, sofware development, and continuous integration techniques: skills that are not widely taught or expected of academic researchers and data scientists.

The Turing Way is a handbook to support students, their supervisors, funders and journal editors in ensuring that reproducible data science is "too easy not to do". It will include training material on version control, analysis testing, and open and transparent communication with future users, and build on Turing Institute case studies and workshops. This project is openly developed and any and all questions, comments and recommendations are welcome at our github repository: https://github.com/alan-turing-institute/the-turing-way.

https://the-turing-way.netlify.com

I ON THIS PAGE

A BIT MORE BACKGROUND

THE BOOK ITSELF THE TURING WAY COMMUNITY

Checklists for researcher, PI and admin team



- Researcher
 - Version control
 - Capturing compute environment
 - Writing and running the code
- PI
 - Results presented are those from the final run of the analysis
 - Check that another researcher can run the code
- Admin
 - Version control
 - Data and code archive
 - Open access publication

What is Jupyter Book?

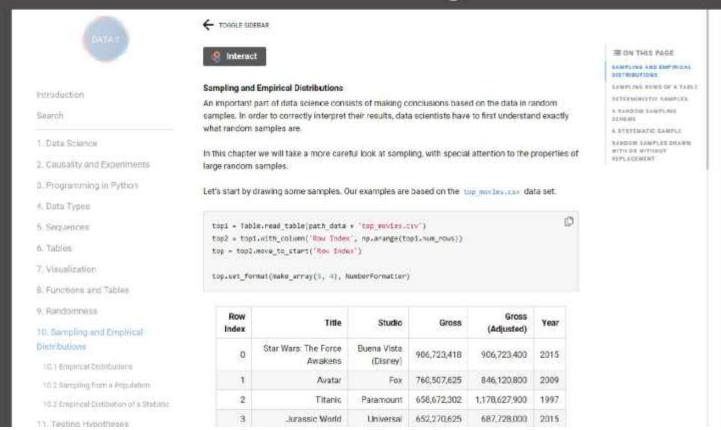
Build an online book with Jupyter Notebooks and Markdown



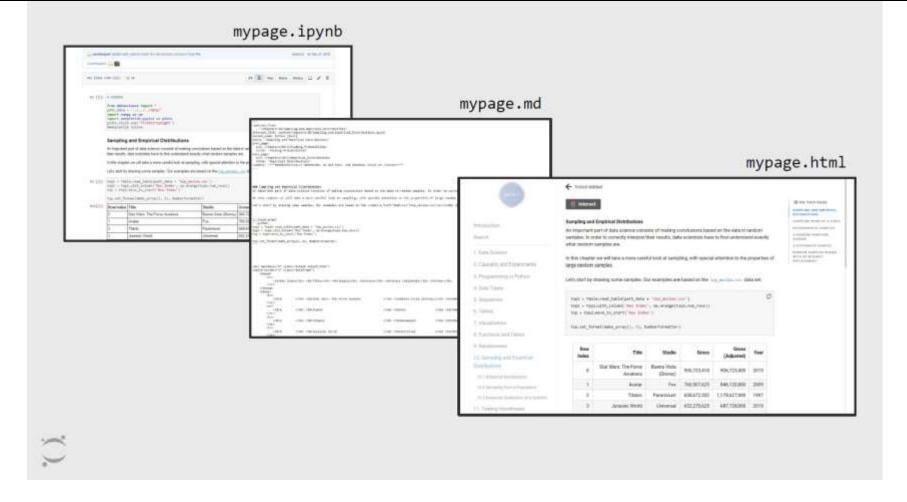
jupyter.org/jupyter-book



inferentialthinking.com



https://speakerdeck.com/choldgraf/jupyter-book-interactive-books-running-in-the-cloud



https://speakerdeck.com/choldgraf/jupyter-book-interactive-books-running-in-the-cloud

Interactive buttons let readers explore

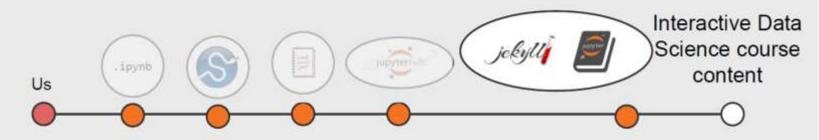
- A page built from a notebook gets an interact button
- Use JupyterHub/Binder to spawn a Jupyter server
- git pull` the underlying notebook for the page
- Initialize an interactive environment





In summary

Jupyter Book builds on tools in the Jupyter ecosystem to create interactive, beautiful books.



jupyter.org/jupyter-book



https://speakerdeck.com/choldgraf/jupyter-book-interactive-books-running-in-the-cloud



The Alan Turing Institute

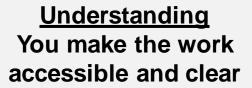
Contributing to The Turing Way

https://github.com/alan-turing-institute/the-turing-way



Open Leadership Principles





Read more



Sharing
You make the work
easy to adapt,
reproduce, and spread



You build shared ownership and agency to make the work inviting and sustainable for all.

https://mozilla.github.io/olm-whitepaper

stie j moz

Built by a team....and you!

- Rachael Ainsworth
- Becky Arnold
- Louise Bowler
- Sarah Gibson
- Patricia Herterich
- Rosie Higman
- Anna Krystalli
- Alex Morley
- Martin O'Reilly
- Kirstie Whitaker
- . . .



Code of Conduct

"The Turing Way team are dedicated to providing a welcoming and supportive environment for all people...we do not tolerate behaviour that is disrespectful to our community members or that excludes, intimidates, or causes discomfort to others."

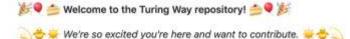
https://github.com/alan-turing-institute/the-turing-way/blob/master/CODE OF CONDUCT.md

Code of Conduct

- Be respectful of different viewpoints and experiences.
- Use welcoming and inclusive language.
- Do not harass people.
- Respect the privacy and safety of others stickers available if you don't want your photo taken
- Be considerate of others' participation.
- Don't be a bystander.

https://github.com/alan-turing-institute/the-turing-way/blob/master/CODE OF CONDUCT.md

Contributing to the Turing Way



The point of this guide is to welcome new users and contributors to the Turing Way community. We hope that these guidelines are designed to make it as easy as possible to get involved. Don't let trying to be perfect get in the way of being good - we welcome all contributions and would love it if you could follow these guidelines to make sure your contributions can be easily integrated but exciting ideas are more important than perfect pull requests.

If you have any questions that aren't discussed below, please let us know through one of the many ways to get in touch.

Table of contents

Been here before? Already know what you're looking for in this guide? Jump to the following sections:

- · Joining the community
- Inclusivity
- · Get in touch
- . Contributing through GitHub
- · Writing in Markdown
- . Where to start: issues
- · Making a change with a pull request
- . The process of writing chapters
- Style Guide
- Recognising Contributions

Joining the community

The Turing Way is a community-oriented and -led project. We therefore require that all contributions adhere to our Code of Conduct.

Inclusivity

This project aims to be inclusive to people from all walks of life and to all research fields. This should be taken into account in contributions.

The following are examples of inclusive actions that we encourage from contributors to the Turing Way:

- Refer to "open research" rather than "open science" so that we do not exclude members of the humanities and social sciences from our community.
- Make sure colour pallettes are accessible to colour-blind readers and contributors. Here's a useful blog post on tips for designing scientific figures for color blind readers by Luk at Somersulat 1824.

Get in touch

There are many ways to get un touch with the Turing Way team!

- · Ping us in our gitter channel,
 - This is our preferred method of open communication and discussion! We'd love for you to swing by to say hello.
- . Join the discussion in our issues and pull requests.
 - Can't find your idea being discussed anywhere? Open a new issue! (See our Where to start: issues section below.)
- Subscribe to our mailing list with which we send monthly project updates.
- . Check out the #TuringWay hashtag on Twitter.
- You can contact the PI of the Turing Way project Kirstie Whitaker by email at kwhitaker@turing.ac.uk.
- . You can also contact members of the Turing Way team through their preferred ways of communication here.

Contributing through GitHub

Git is a really useful tool for version control. GitHub sits on top of Git and supports collaborative and distributed working.

We know that it can be daunting to start using Git and GitHub if you haven't worked with them in the past, but the Turing Way maintainers are here to help you figure out any of the jargon or confusing instructions you encounter!

In order to contribute via GitHub you'll need to set up a free account and sign in. Here are some instructions to help you get going. Remember that you can ask us any questions you need to along the way.

Writing in Markdown

GitHub has a helpful page on getting started with writing and formatting on GitHub.

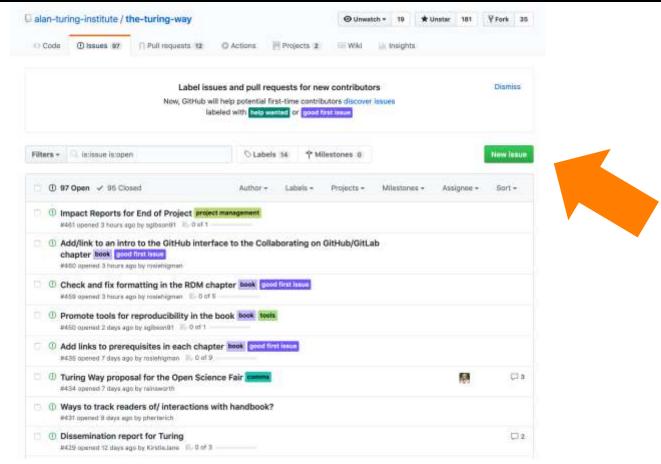
Most of the writing that you'll do will be in Markdown. You can think of Markdown as a few little symbols around your text that will allow GitHub to render the text with a little bit of formatting. For example you could write words as **bold** (**bold**), or in italics (*italics*), or as a link ([link](https://https://youtu.be/dQw4w9WgXcQ)) to another webpage.

Also when writing in Markdown, please start each new sentence on a new line. While this formats in the same way as if the new line wasn't included, it makes the diffs produced during the pull request review easier to read!

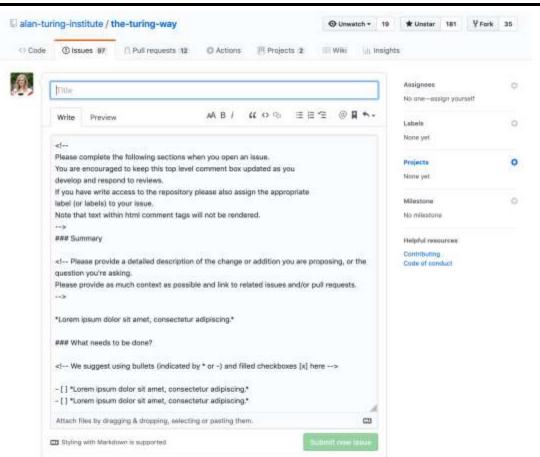
1. Comment on an existing issue or open a new issue referencing your addition

This allows other members of the Turing Way team to confirm that you aren't overlapping with work that's currently underway and that everyone is on the same page with the goal of the work you're going to carry out.

If you open a new issue, please follow the basic guidelines laid out in our issue template. The issue template will automatically be rendered in the comment section of the new issue page so all you need to do is edit the "Lorem ipsum" sections.



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



- 1. Add descriptive title.
- 2. Fill in the issue template by replacing the *Lorem ipsum* sections with what needs to be done or fixed.
- 3. Add assignees and labels where relevant (ignore projects and milestones).

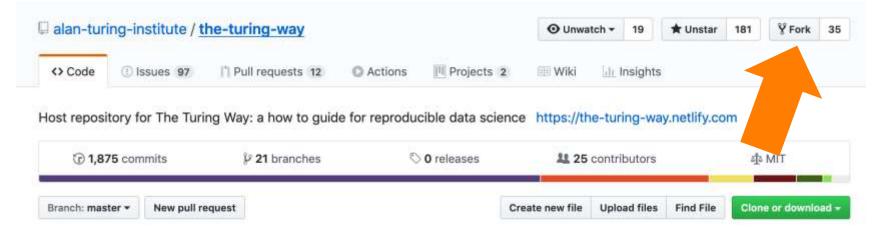
book-dash-mcr

4. Submit issue and off you go!

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

2. Fork the Turing Way repository to your profile

This is now your own unique copy of the Turing Way. Changes here won't affect anyone else's work, so it's a safe space to explore edits to the code!



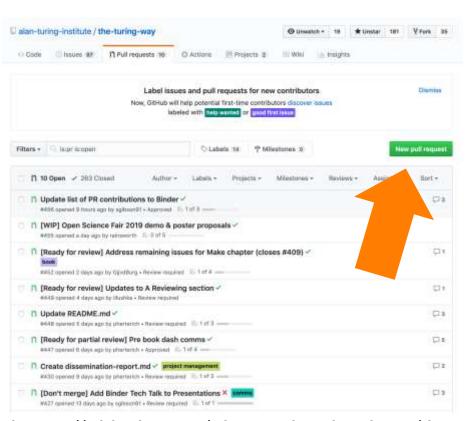
3. Make the changes you've discussed

Try to keep the changes focused. If you submit a large amount of work all in one go it will be much more work for whomever is reviewing your pull request.

While making your changes, commit often and write good, detailed commit messages.

If you feel tempted to "branch out" then please make a new branch and a new issue to go with it – for example, if writing a new chapter, create a new branch in the Turing Way repository to pull to, named with the chapter title.

4. Submit a pull request

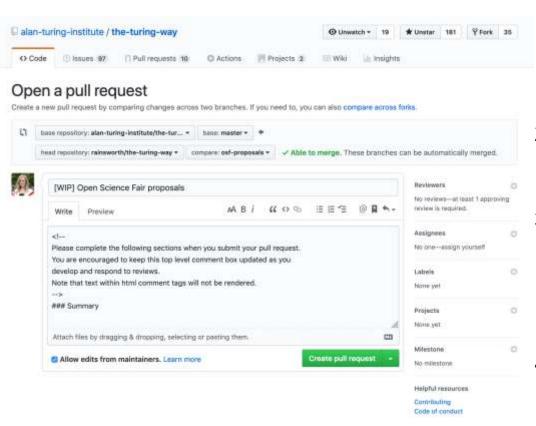


We encourage you to open a pull request as early in your contributing process as possible.

This allows everyone to see what is currently being worked on.

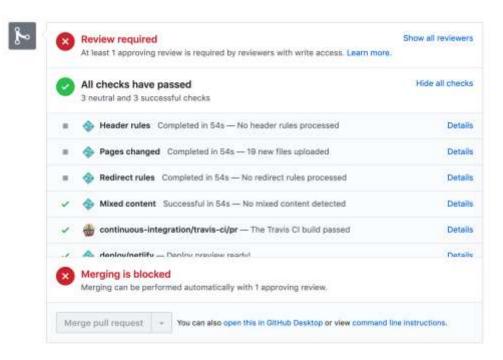
It also provides you, the contributor, feedback in real time from both the community and the continuous integration as you make commits.

You can also submit pull requests to other contributors' branches!



- Select which branches you want to merge: the-turing-way branch ← your branch
- 2. Add descriptive title with [WIP] or [Ready for review] tag.
- Fill in the pull request template, replacing the Lorem ipsum sections with what has been done & reference relevant issue.
- 4. Add book-dash-mcr label & create pull request!

https://github.com/alan-turing-institute/the-turing-way/pulls



A member of the Turing Way team will then review your changes to confirm that they can be merged into the main repository.

You don't need to submit a new pull request when you make a change in response to a review.

You can update your fork of the Turing Way repository and the pull request will automatically update with those changes and the CI will feedback.

The Turing Way

- 1. Introduction
- 2. Reproducibility
- 3. Open Research
- 4. Version Control
- 5. Collaborating on GitHub/GitLab
- 6. Research Data Management
- 7. Reproducible Environments
- 8. Testing
- 9. Reviewing
- 10. Continous Integration
- Reproducible Research with Make
- 12. Risk Assessment

Powered by Jupyter Book

Welcome to the Turing Way

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 so means making sure PhD students, postdocs, PIs and funding teams know which parts of reproducibility" they can affect, and what they should do to nudge data e efficient, effective and understandable.

A bit more background

Reproducible research is necessary to ensure that scientific work can be trusted. Funders and publishers are beginning to require that publications include access to the underlying data and the analysis code. The goal is to ensure that all results can be independently verified and built upon in future work. This is sometimes easier said than done. Sharing these research outputs means understanding data management, library sciences, sofware development, and continuous integration techniques: skills that are not widely taught or expected of academic researchers and data scientists.

The Turing Way is a handbook to support students, their supervisors, funders and journal editors in ensuring that reproducible data science is "too easy not to do". It will include training material on version control, analysis testing, and open and transparent communication with future users, and build on Turing Institute case studies and workshops. This project is openly developed and any and all questions, comments and recommendations are welcome at our github repository: https://github.com/alan-turing-institute/the-turing-way.

https://the-turing-way.netlify.com

I ON THIS PAGE

A BIT MORE BACKGROUN

THE BOOK ITSELF THE TURING WAY COMMUNITY

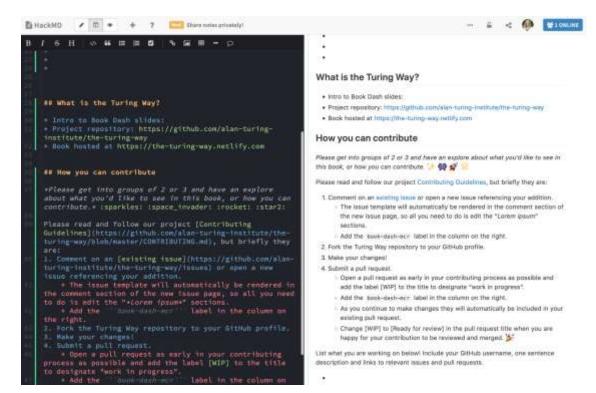
What types of contributions can you make?

- Proofread existing chapters (check for typos, grammar, links, etc.)
- Add further information to an existing chapter
- Suggest topics for new chapters (particularly if you can write them!)
- Review [Ready to review] pull requests
- Contribute to discussion [WIP] pull requests and issues
- Submit a case study or your tips and tricks for reproducible research via our Google submission form (link in README).
- Checklists
- Restructuring the chapter sections (like the reproducibility chapter)
- Let us know if you are struggling with contributing in any way so that we can improve our Contributing Guidelines!

What types of contributions can you make?

- Proofread existing chapters (check for typos, grammar line)
- Add further information to an existing chapter
- Review [Ready to review] pull rec.
- Contribute to discussion "\"
- Submit a case studion form (link in README).
- Jonapter sections (like the reproducibility chapter)
 Jow if you are struggling with contributing in any way so
 Jat we can improve our Contributing Guidelines!

Update the HackMD: http://bit.ly/book-dash-mcr



Add what you're working on, any questions you have or anything you've learned!

Link to your issues and pull requests.

This will help us write up our Book Dash Report.

The emoji key to celebrate our contributors



Emoji	Represents
0	Answering Questions (on gitter, GitHub, or in person)
	Bug reports
2	Biogposts
=	Code
ш	Documentation and specification
0	Design
Q	Examples
	Event Organizers
1943	Financial Support
Q	Funding/Grant Finders
9	Ideas & Planning
0	Infrastructure (Hosting, Build-Tools, etc)
4	Plugir/utility libraries
••	Reviewed Pull Requests
1	Tools
0	Translation
4	Tests
2	Tutorials
lie	Talks
wi	Videos

https://github.com/alan-turing-institute/the-turing-way/blob/master/README.md

Neurohackademy

"Every hackathon should have a gong that you can ring when you complete your first pull request."



https://neurohackademy.org #csvconf#TuringWay @kirstie_j https://doi.org/10.5281/zenodo.2669548

Agenda

http://bit.ly/book-dash-mcr

Arrive at dashing venue
Intro to the book dash
Team forming and brain storming
Coffee break
Group working
Lunch
Group working, coffee available at 15:00
Celebrate successes (group share outs)
Close

16/05/2019

The Alan Turing Institute

Agenda

http://bit.ly/book-dash-mcr

8:30	Arrive at dashing venue
9:00	Intro to the book dash
9:30	Team forming and brain storming
10:30	Coffee break
11:00	Group working
12:30	Lunch
13:30	Group working, coffee available at 15:00
16:00	Celebrate successes (group share outs)
17:00	Close

16/05/2019

The Alan Turing Institute



- #TuringWay
- https://github.com/alan-turing-institute/the-turing-way
- gitter.im/alan-turing-institute/the-turing-way