



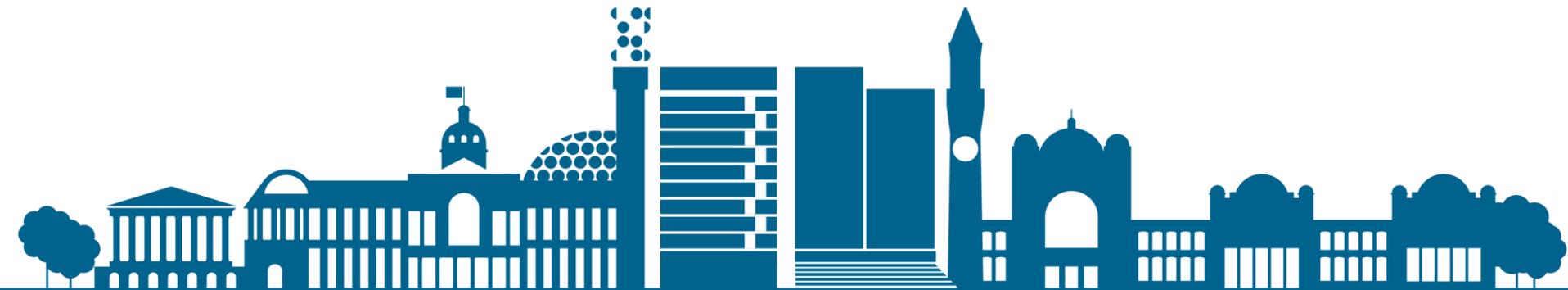
UNIVERSITY OF  
BIRMINGHAM

# The Turing Way - supporting reproducibility and Open Research

Patricia Herterich, Library Services

The Hacker Within

18.02.2019



# Join us on...

 #TuringWay or @pherterich

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

 [gitter.im/alan-turing-institute/the-turing-way](https://gitter.im/alan-turing-institute/the-turing-way)

This talk is available on zenodo at:

<https://doi.org/10.5281/zenodo.2566430>



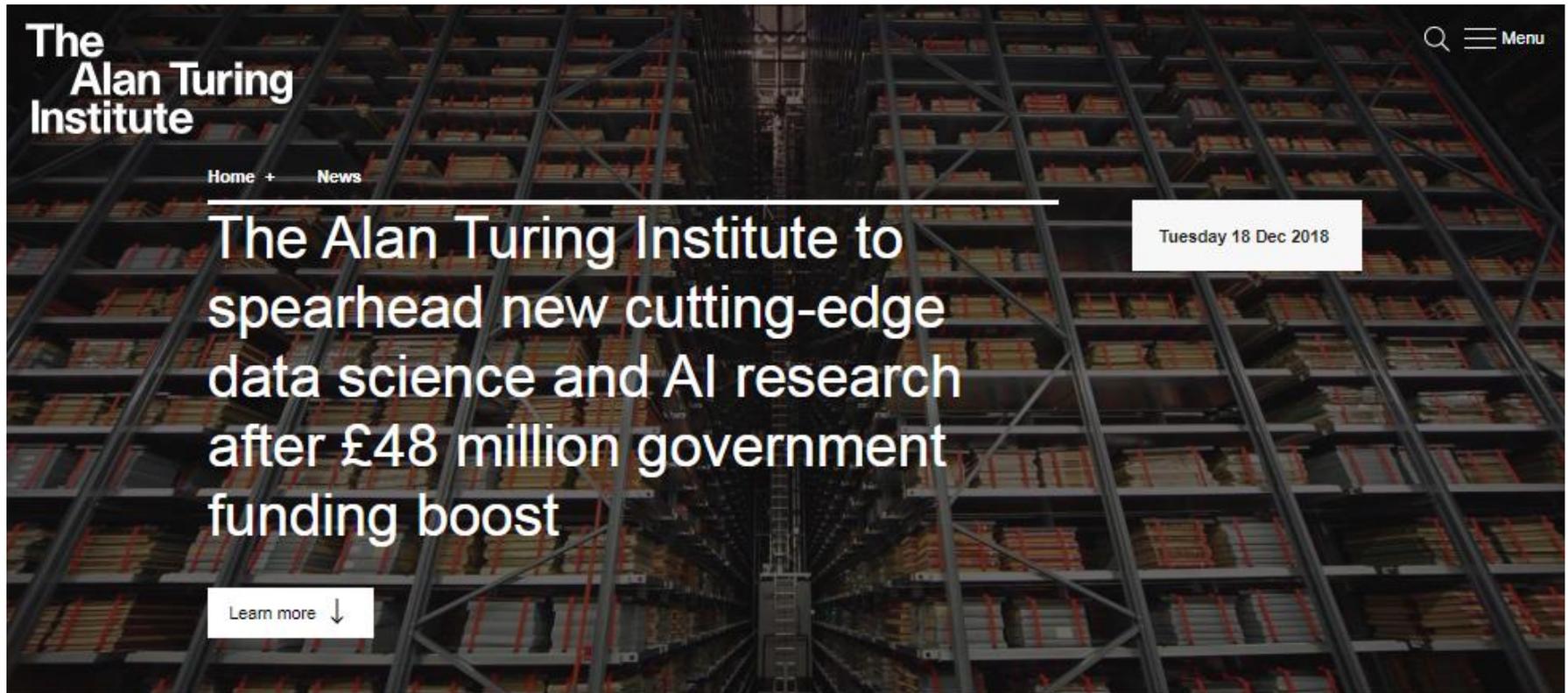
# What is the Turing Way?

- A lightly opinionated handbook for reproducible data science
- Openly and collaboratively developed at

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



# Who pays for this?

A screenshot of the The Alan Turing Institute website. The background is a dark, high-angle photograph of a server room with rows of server racks. In the top left corner, the logo for 'The Alan Turing Institute' is displayed. In the top right corner, there is a search icon and a 'Menu' button. Below the logo, there are navigation links for 'Home' and 'News'. The main headline reads: 'The Alan Turing Institute to spearhead new cutting-edge data science and AI research after £48 million government funding boost'. To the right of the headline, a date box indicates 'Tuesday 18 Dec 2018'. Below the headline, there is a 'Learn more' button with a downward arrow.

The Alan Turing Institute

Home + News

The Alan Turing Institute to spearhead new cutting-edge data science and AI research after £48 million government funding boost

Tuesday 18 Dec 2018

Learn more ↓

UK Research and Innovation (UKRI) has **today announced** funding for two major new research programmes in artificial intelligence, both led in collaboration with The Alan Turing Institute.



UNIVERSITY OF  
BIRMINGHAM

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



# Tools, practices and systems for AI

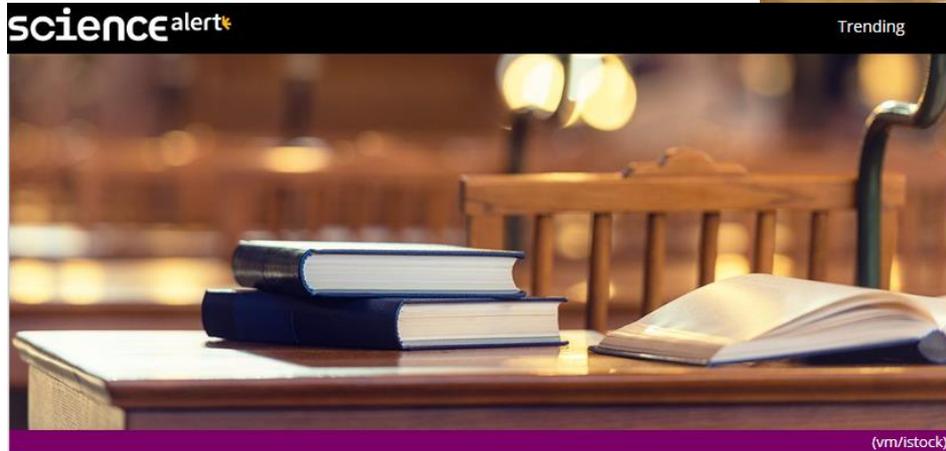


UNIVERSITY OF  
BIRMINGHAM

# Is there a reproducibility crisis?

## The replication crisis has engulfed economics

November 2, 2015 7:31pm EST



Trending

in Shutterstock.com

HUMANS

### Science's 'Replication Crisis' Has Reached Even The Most Respectable Journals, Report Shows

MIKE MCRAE 27 AUG 2018

An attempt to replicate the findings of 21 social science experiments published in two high-profile science journals has thrown up a red flag for reliability in research.

The question of just how much trust we should put into one-off studies isn't new. But there's a silver lining to the so-called [Reproducibility Crisis](#) we need



UNIVERSITY OF BIRMINGHAM

A sense of crisis is developing in economics after [two Federal Reserve economists](#) came to the alarming conclusion that economics research is usually not replicable.

The [economists took 67 empirical papers](#) from 13 reputable academic journals. Without assistance from the original researchers they were only able to get the same result in a third of cases.



VIEWS 23,107

The Reproducibility Project: Cancer Biology is an initiative to independently replicate selected results from a number of [high-profile papers](#) in the field of cancer biology. For each paper a Registered Report detailing the proposed experimental designs and protocols for the replications is peer reviewed and published prior to data collection; the results of these experiments are then published as a Replication Study. [The project](#) is a collaboration between the Center for Open Science and Science Exchange.

<https://web.archive.org/web/20190104143016/https://www.sciencealert.com/replication-results-reproducibility-crisis-science-nature-journals>

<https://web.archive.org/web/20181116064359/http://theconversation.com/the-replication-crisis-has-engulfed-economics-49202>

<https://elifesciences.org/collections/9b1e83d1/reproducibility-project-cancer-biology>

# What do we mean by reproducible?

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable



# 5 selfish reasons to work reproducibly

- Avoiding disaster
- Writing papers easier
- Convincing reviewers
- Facilitating continuity of work
- Building your reputation

Markowetz, F. (2015). Five selfish reasons to work reproducibly. *Genome Biology*, 16(1). <https://doi.org/10.1186/s13059-015-0850-7>



# Why don't people do this already?

Is not considered  
for promotion

Takes  
time

Publication bias  
towards novel  
findings

## Barriers to reproducible research

*Requires  
additional  
skills*

Plead the  
5th

Support additional  
users

*Held to higher  
standards than others*



UNIVERSITY OF  
BIRMINGHAM

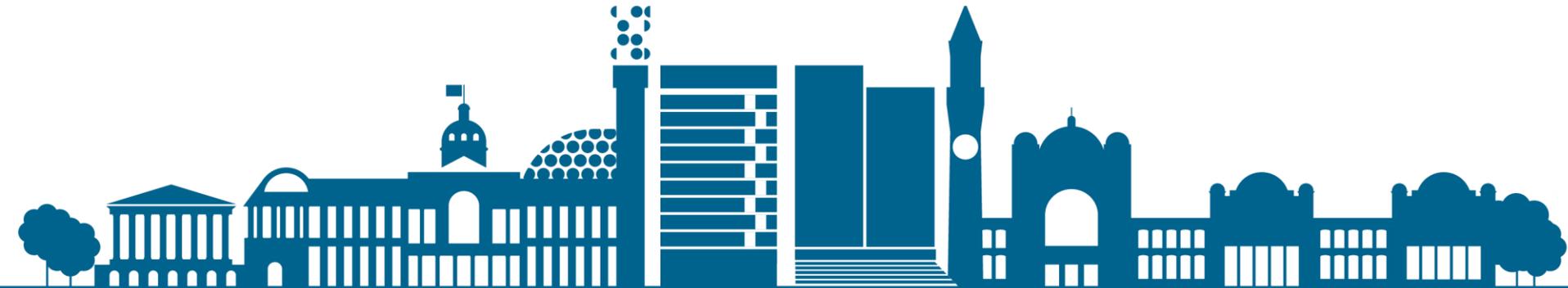
<https://doi.org/10.6084/m9.figshare.7140050>



UNIVERSITY OF  
BIRMINGHAM

# Can we reduce the barriers?

The Turing Way handbook to the rescue!



# Content

## □ Chapters will include:

- Research data management
- Open science
- Reproducibility
- Version control with git
- Capturing your compute environment
- Testing for research
- Continuous integration
- Collaborating through GitHub/GitLab

[https://github.com/alan-turing-institute/the-turing-way/blob/master/book\\_skeleton.md](https://github.com/alan-turing-institute/the-turing-way/blob/master/book_skeleton.md)



UNIVERSITY OF  
BIRMINGHAM

# Desirable content includes...

- ❑ Working Environment and notebooks
- ❑ Coding Styles and Linting
- ❑ Deep Learning
- ❑ Ethics
- ❑ Credit for Reproducible Research
- ❑ Scoping a data project
- ❑ Similarities and differences across data science disciplines
- ❑ Visualisation

<https://github.com/alan-turing-institute/the-turing-way/blob/master/CONTRIBUTING.md>



UNIVERSITY OF  
BIRMINGHAM

# 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



# Tools for reproducibility

## □ Binder to the rescue!

<https://mybinder.readthedocs.io/en/latest>



UNIVERSITY OF  
BIRMINGHAM

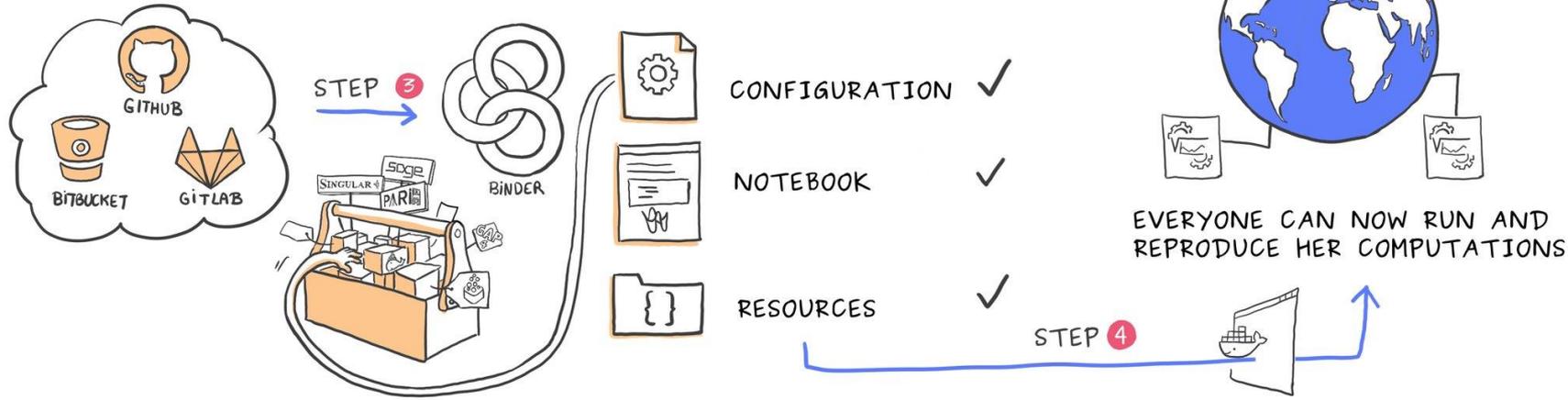
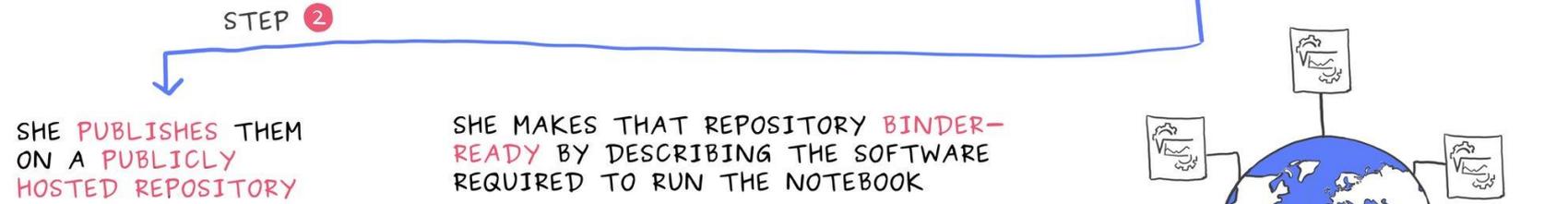
## Binder Team

Binder's governance and team structure is defined in the Binder Project Governance page. Below we list the current team members of Binder.

(listed alphabetically, with affiliation, and main areas of contribution)

 <p><b>Jessica Forde</b> UC Berkeley team red</p> 	 <p><b>Tim Head</b> Wild Tree Tech team red</p>	 <p><b>Lindsey Heagy</b> UC Berkeley team blue</p> 	 <p><b>Chris Holdgraf</b> Berkeley Institute for Data Science team red</p> 
 <p><b>M Pacer</b> Netflix team blue</p>	 <p><b>Yuvi Panda</b> UC Berkeley team red</p> 	 <p><b>Min Ragan-Kelley</b> Simula team lead data,</p> 	 <p><b>Zach Sailer</b> Project Jupyter team blue</p> 
 <p><b>Erik Sundell</b> Sandvik CODE team blue</p> 	 <p><b>Carol Willing</b> Project Jupyter team red Python, Community</p>		

<https://jupyterhub-team-compass.readthedocs.io/en/latest/team.html#binder-team>



UNIVERSITY OF BIRMINGHAM

Courtesy of Juliette Belin:

<https://twitter.com/JulietteTaka/status/1082735653929000960>

# Live demo

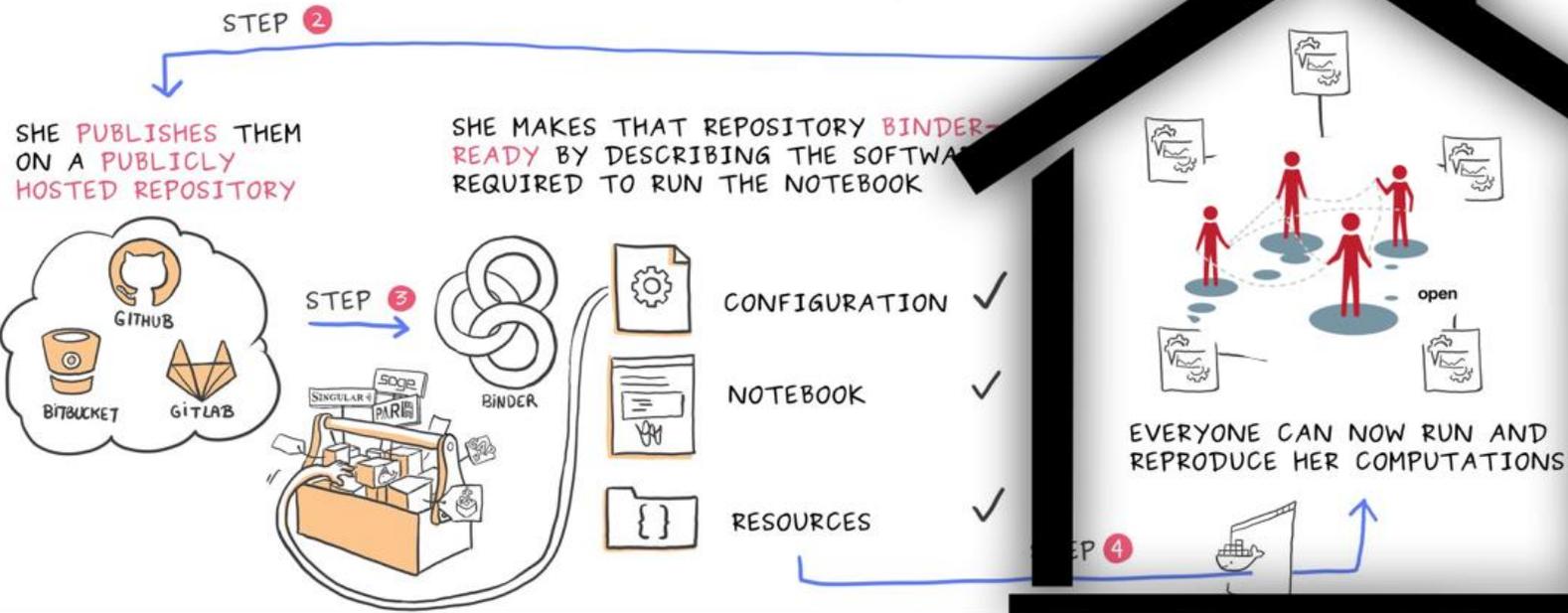
□ [https://github.com/sgibson91/ukri\\_demo](https://github.com/sgibson91/ukri_demo)

□ Back-up:

The screenshot shows a web browser window displaying the GitHub repository page for 'sgibson91/ukri\_demo'. The browser's address bar shows the URL 'https://github.com/sgibson91/ukri\_demo'. The repository page includes a header with the repository name, 'Watch' (0), 'Star' (0), and 'Fork' (0) buttons. Below the header, there are navigation tabs for 'Code', 'Issues' (0), 'Pull requests' (0), 'Projects' (0), 'Wiki', 'Insights', and 'Settings'. The main content area features a description: 'Demo repo presented at the UKRI Cloud Workshop (12th Feb 2019)'. It also displays statistics: '24 commits', '3 branches', '0 releases', and '1 contributor'. A 'Branch: master' dropdown and a 'New pull request' button are visible. A list of recent commits is shown, with the most recent being 'sgibson91 Update README to explain the repo structured' from 21 hours ago. Below the commit list, the 'README.md' file is open, showing the title 'UKRI Cloud Workshop Demo' and a description of the repository's purpose and structure.



UNIVERSITY OF  
BIRMINGHAM



UNIVERSITY OF BIRMINGHAM

Courtesy of Juliette Belin:

<https://twitter.com/JulietteTaka/status/1082735653929000960>

and Sarah Gibson

# Built by a team...and you!

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

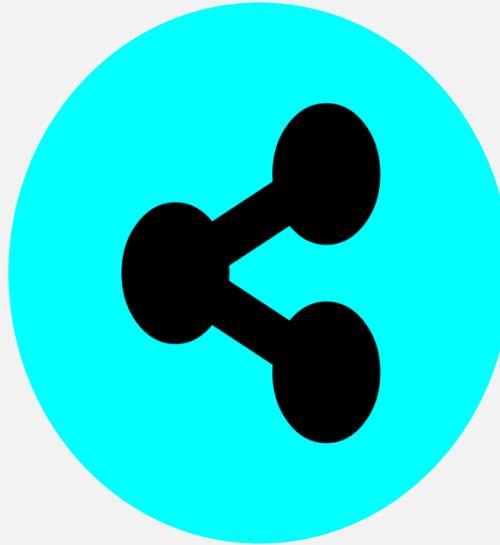


## Open Leadership Principles



### Understanding

You make the work accessible and clear



### Sharing

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



### Participation & Inclusion

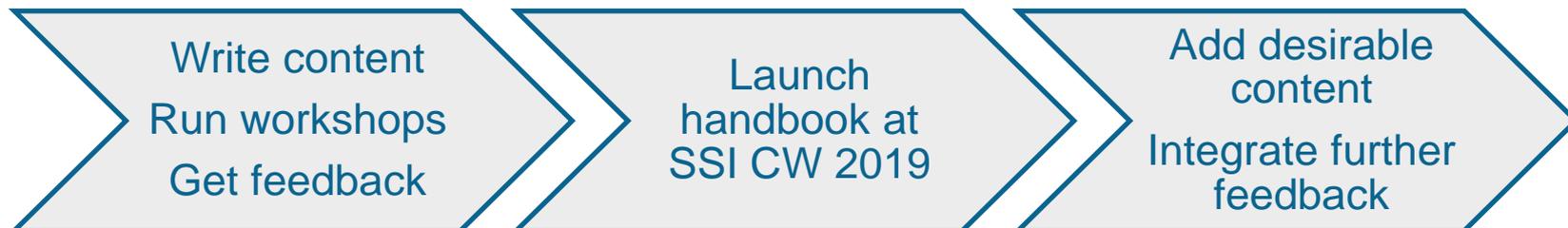
You build shared ownership and agency to make the work inviting and sustainable for all.

**Read more**

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

<https://doi.org/10.6084/m9.figshare.7564682>

# Timeline



#CollabW19  
1 – 3 April 2019,  
Loughborough



About Programmes and Ever

## Collaborations Workshop 2019 (CW19)

Twitter: #CollabW19

The Software Sustainability Institute's Collaborations Workshop series brings together researchers, developers, innovators, managers, funders, publishers, leaders and educators to explore best practices and the future of research software. Collaborations Workshop 2019 (CW19) will take place from Monday 1st to Wednesday 3rd April 2019 at **Loughborough University** in the **West Park Teaching Hub**, Loughborough.



Image by Loughborough University.

The themes of the workshop are interoperability, documentation, training and sustainability.

Register at:

<https://www.software.ac.uk/cw19>



UNIVERSITY OF  
BIRMINGHAM

Register on the CW19 Eventbrite page

# So... how can you get involved?

□ Attend one of our workshops:

## **Boost your research reproducibility with Binder**

During this free workshop we will discuss reproducible computing environments, show examples of others' projects in Binder and help you learn how to prepare a Binder-ready project. At the end of the workshop you will be able to take some of your own content (in a R or Jupyter notebook, or scripts that can be run in the terminal) and prepare it so that it can be used by others on [mybinder.org](https://mybinder.org).

Friday 1st March, [University of Manchester](https://www.manchester.ac.uk)

Tuesday 12th March, [The Alan Turing Institute](https://www.turing.ac.uk)



UNIVERSITY OF  
BIRMINGHAM

# Workshops for RSEs

## **Build a BinderHub**

During this free workshop we will demonstrate how to build your own BinderHub on Microsoft Azure cloud computing resources. We will help you get started with building a BinderHub on your institution's computing platform and discuss the challenges of maintaining a BinderHub. At the end of the workshop you will know why this would be a useful resource for your team, and will know where to look for help and support building your institution's BinderHub.

Monday 18th March: [University of Sheffield](#)

**We can cover travel & accommodation costs!**



UNIVERSITY OF  
BIRMINGHAM

# Submit your tips and tricks!

On GitHub directly or by filling our quick submission form at

<https://goo.gl/forms/akFqZEIy2kxAjfZW2>

or even contribute to chapters and conversations in our GitHub issues and through pull requests!

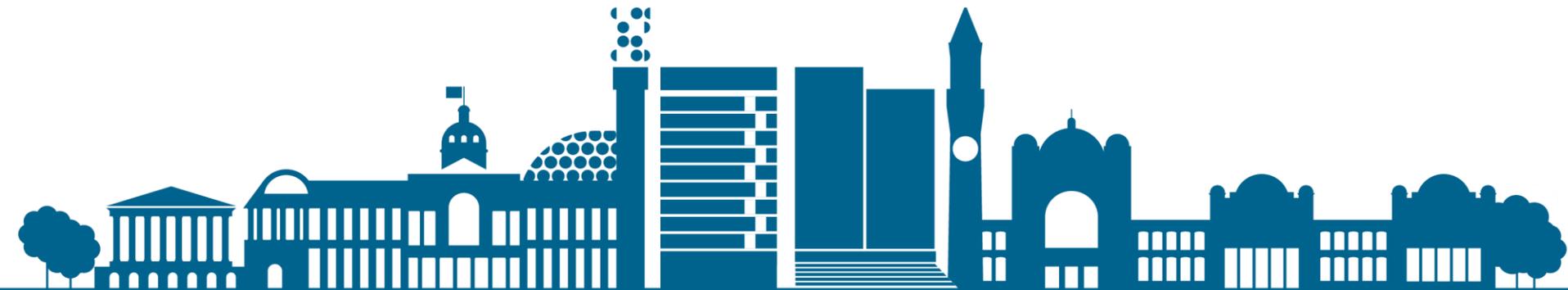


UNIVERSITY OF  
BIRMINGHAM

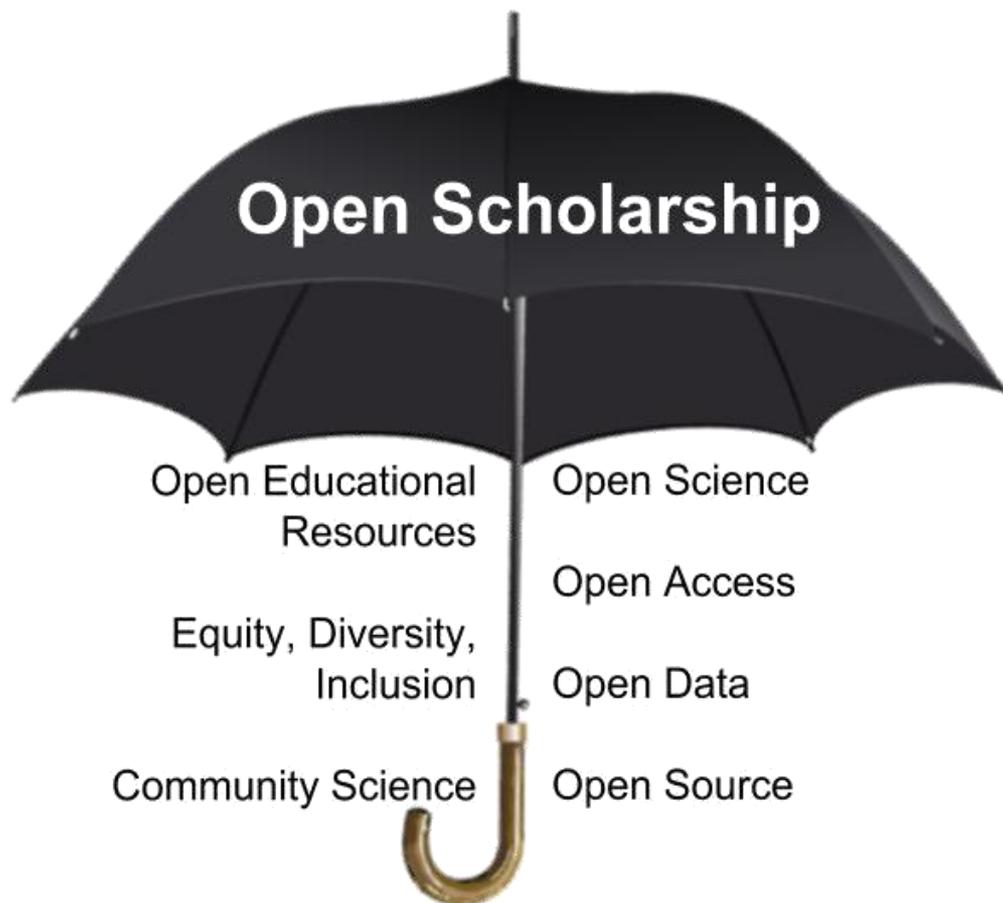


UNIVERSITY OF  
BIRMINGHAM

# Why is someone from Library Services working on this?



# We support Open Scholarship



Adapted from: <https://www.meetup.com/Berlin-Open-Science-Meetup/>

Robin Champieux and Danielle Robinson



# With Scholarly Communications Services on...

## □ Open Access

- Immediate Gold Open Access – funder budgets & internal Open Access fund
- Green OA through PURE
- Monograph chapters
- Theses

<http://intranet.birmingham.ac.uk/openaccess>



# With Scholarly Communications Services on...

- Open (Research) Data
  - Review of Data Management Plans
  - Institutional data repository:  
<https://edata.bham.ac.uk>
  - Guidance on data sharing, documentation, anonymization etc...

<https://intranet.birmingham.ac.uk/rdm>



# Want to talk other aspects?

- Pre-prints
- Open peer review
- Electronic lab notebooks
- Open Educational Resources
- ...
- Get in touch:
  - [openaccesspublications@contacts.bham.ac.uk](mailto:openaccesspublications@contacts.bham.ac.uk)
  - [research-data@contacts.bham.ac.uk](mailto:research-data@contacts.bham.ac.uk)
  - [copyright@contacts.bham.ac.uk](mailto:copyright@contacts.bham.ac.uk)



# How else can you get involved in reproducible research?



UK Reproducibility Network (UKRN) – peer-led consortium to support reproducible research

West Midlands Network led by Charlotte Hartwright (lecturer at Aston University)



# Set up an Open Science Working Group

- ... and join the UK Network
- More info at <https://osf.io/vgt3x/>
- Inaugural meeting of UK Network of Open Science Working Groups

Aston University – 11 April 2019

Register at:

<https://www.eventbrite.co.uk/e/ukoswg19-tickets-55582448600>



Find what works for you!

*Every little helps*



# Thank you!

## Join us!

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

 [gitter.im/alan-turing-institute/the-turing-way](https://gitter.im/alan-turing-institute/the-turing-way)



**The  
Alan Turing  
Institute**



UNIVERSITY OF  
BIRMINGHAM

LIBRARY  
SERVICES

