Better Software, Better Research:
How the Software Sustainability Institute is helping to promote reproducible research

Dr. Rachael Ainsworth
Research Software Community Manager
Software Sustainability Institute, University of Manchester

@rachaelevelyn
@rainsworth
https://doi.org/10.6084/m9.figshare.11898450

Software Sustainability Institute

THE UNIVERSITY of EDINBURGH
UNIVERSITY OF Southampton
UNIVERSITY OF OXFORD
MANCHESTER 1824 The University of Manchester

UKRI Arts and Humanities Research Council
UKRI Medical Research Council
UKRI Biotechnology and Biological Sciences Research Council
UKRI Natural Environment Research Council
UKRI Economic and Social Research Council
UKRI Science and Technology Facilities Council
UKRI Engineering and Physical Sciences Research Council

About me

- Community Manager for the Software Sustainability Institute at the University of Manchester (and 2019 Fellow!)
- Research background in Astrophysics
- Open Research Advocate: passionate about promoting openness, transparency, reproducibility, wellbeing and inclusion in STEM/research
- TEDx speaker: youtu.be/c-bemNZ-IqA
- Was a cartoon in the UK's National Science and Media Museum Hello Universe exhibition
- Founded the Manchester women in data meetup group HER+Data MCR meetup.com/HER-Data-MCR

HER+Data MCR

@herplusdatamcr

About the Software Sustainability Institute (SSI)

- A national facility promoting the advancement of software in research since 2010, by cultivating better, more sustainable, research software to enable world-class research:
  "Better software, better research"

- A collaboration between the universities of Edinburgh, Manchester, Oxford and Southampton.

- Teams: Software, Training, Policy, Community and Outreach

- In 2018, SSI was awarded funding from all seven research councils.

  - Supported by the UK Research Councils through grants EP/H043160/1, EP/N006410/1 and EP/S021779/1, with additional project funding from Jisc.

- https://www.software.ac.uk/

THE UNIVERSITY of EDINBURGH
UNIVERSITY OF Southampton
UNIVERSITY OF OXFORD
MANCHESTER 1824 The University of Manchester

UKRI Arts and Humanities Research Council
UKRI Medical Research Council
UKRI Biotechnology and Biological Sciences Research Council
UKRI Natural Environment Research Council
UKRI Economic and Social Research Council
UKRI Science and Technology Facilities Council
UKRI Engineering and Physical Sciences Research Council

SSI Teams

- **Software**: Helping the community to develop software that meets the needs of reliable, reproducible, and reusable research.
- **Policy**: Collecting evidence on and promoting the place of software in research and sharing with stakeholders.
- **Communications**: Exploiting our platform to enable engagement, delivery and uptake.
- **Training**: Delivering essential software skills to researchers, partnering with institutions, doctoral schools and the community.
- **Community**: Developing Communities of Practice by supporting the right people to understand and address topical issues.

Director

Director

Neil Chue Hong

Deputy Director

Simon Hettrick

Research Director

Caroline Jay

Co-Investigators

Co-investigator

Les Carr

Co-Investigator

David De Roure

Co-Investigator

Carole Goble

Co-Investigator

Mark Parsons

Team Leads

Communications

Selina Aragon

Software

Steve Crouch

Training

Aleksandra Nenadic

Community

Shoaib Sufi

Consultants

Community Manager

Rachael Ainsworth

Project Officer

Giacomo Peru

Community Officer / Research Software Engineer

Mario Antonioletti

Senior Software Consultant

John Robinson

Research Software Engineer

James Graham

Events Coordinator

Graeme Smith

Event Manager

Clem Hadfield

Community Manager for UK Research Software Engineering

Claire Wyatt

SSI in Numbers

- 4 founding universities
- 70+ software consultancy projects across all UKRI domains
- 150 Fellows from 70+ domains
- 270+ workshops and events
- 1500+ Research Software Engineers engaged
- 6000+ researchers trained and 250+ new trainers created
- 25k unique web visitors each month, 170+ external contributors, and 7.1k followers on Twitter


Andrew Stewart


Anna Krystalli


Benjamin Krikler


Catherine Smith


Daniel Hobley


Frances Cooper


Gergana Daskalova


Isla Myers-Smith


Jessica Ward


Laura James


Leandro Liborio


Lucy Whalley

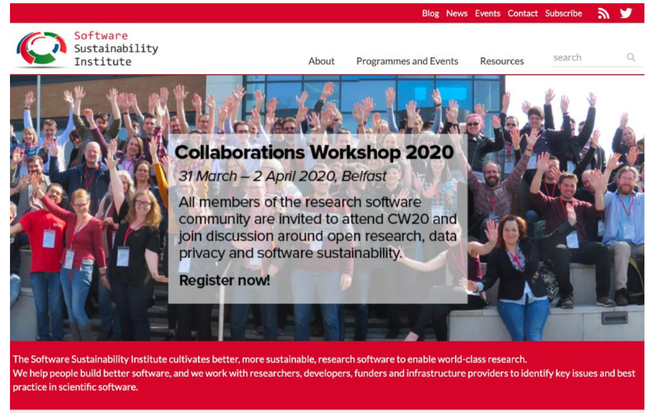

Malvika Sharan


Mozhgan Kabiri chimeh


Patricia Herterich


Rachael Ainsworth


Sorrel Harriet



THE CARPENTRIES

RSE
SOCIETY OF RESEARCH SOFTWARE ENGINEERING

THE UNIVERSITY of EDINBURGH
UNIVERSITY OF Southampton
UNIVERSITY OF OXFORD
MANCHESTER 1824
The University of Manchester

UKRI Arts and Humanities Research Council
UKRI Medical Research Council
UKRI Biotechnology and Biological Sciences Research Council
UKRI Natural Environment Research Council
UKRI Economic and Social Research Council
UKRI Science and Technology Facilities Council
UKRI Engineering and Physical Sciences Research Council

SSI and UKRN

- SSI Fellow Andrew Stewart is the Local UKRN Lead, leads the Open Research Working Group, and sits on the Open Research Strategy Group at the University of Manchester.

- SSI formally partnered with UKRN as a stakeholder in January 2020, and we look forward to collaborating on a range of joint initiatives.

Andrew Stewart

Blog News Events Contact Subscribe

Software Sustainability Institute

About Programmes and Events Resources search

Software Sustainability Institute joins forces with the UK Reproducibility Network

Posted by n.chuehong on 21 January 2020 - 7:40am

The Institute is delighted to announce that it will be partnering with the UK Reproducibility Network (UKRN) as a stakeholder. The UKRN is a peer-led consortium that aims to ensure the UK retains its place as a centre for world-leading research by investigating the factors that contribute to robust research, promoting training activities, disseminating best practice, and working with stakeholders to ensure coordination of efforts across the sector.

Like the Institute, the UKRN works across disciplines ranging from the arts and humanities to the physical sciences, and encourages grass-roots practice. It is led by Marcus Munafò (Bristol), Chris Chambers (Cardiff), Laura Fortunato (Oxford), Alexandra Collins (Imperial), and Malcolm Macleod (Edinburgh). We look forward to working with them on a range of joint initiatives, including training in software development and data management skills.

UKRN website: http://www.ukrn.org/

SSI and NISO

- The National Information Standards Organization (NISO) identifies, develops, maintains, and publishes technical standards to manage information in today's continually changing digital environment.
- SSI Director Neil Chue Hong has been involved in the NISO work to define standardised badges for reproducibility.
  - Working across the publishers and learned professional societies to get agreement on the definition of reproducibility as well as creating a set of “levels” relating to outputs that can be badged to promote reproducibility.
  - There will be a draft recommendation for comment coming out shortly.

Neil Chue Hong

NISO

Home / Standards Committees

Taxonomy, Definitions, and Recognition Badging Scheme Working Group

Overview

This project is beginning in March, 2019.

Publishers and researchers are placing greater emphasis on the practice of reproducibility as an essential ingredient of the scientific research process. Critical to the issue of reproducibility is the taxonomy used to define the various levels of reproducibility, and agreement on a standardized badging scheme that can be applied in the publishing process (and perhaps used as a currency in the academic rewards culture). As reproducibility begins to spread across the scholarly publishing landscape, recommended badging schemes and the related taxonomies are developing on an ad hoc basis—creating a need for some standardization.

Recognizing that reproducibility standards can vary across disciplines, this effort will focus on standardization across the Computational and Computing Sciences, although adoption by other disciplines would be encouraged.

https://www.niso.org/standards-committees/reproducibility-badging

SSI and The Carpentries

- Coordination of UK workshops and supporting the growing UK Carpentries community.

- Library Carpentry was started by SSI Fellow James Baker in 2015. James used his Fellowship funds to launch initial Library Carpentry workshops, which in the first year attracted 59 participants from 14 institutions in London and reached 200-250 librarians. Since then, a number of workshops have run in various countries across four continents.

- SSI Training Lead Aleksandra Nenadic organises Instructor Training events, focussing on developing teaching skills that are broadly useful across all of the Carpentry curricula and teaching in general - with emphasis on the “best practices” of teaching computational skills based on key educational research findings and how they can be used to help people learn better and faster.

James Baker

Aleksandra Nenadic

THE CARPENTRIES

software carpentry

DATA CARPENTRY

library carpentry

HPC CARPENTRY

https://carpentries.org/

Community

Collaborations Workshop

- Bringing together researchers, developers, innovators, managers, funders, publishers, leaders and educators to explore best practices and the future of research software
- Unconference: keynote presentations, mini-workshop/demo sessions, discussion groups, lightning talks, collaborative ideas, speed blogging and a hack day
- CW20 takes place March 31 - April 2, 2020 at Queen's University Belfast, Northern Ireland
- Themes: Open Research, Data Privacy and Software Sustainability (also Scholarly Communication and Community)
- http://bit.ly/ssi-cw20

Fellowship Programme

- Engaging with natural ambassadors of better software practice from the research community
  - The promotion of better more sustainable software is best done by those in the domains
- Supporting those involved with improving software practice; helping form communities of practice
- Fellows help SSI discover important information about software in different research domains
- Fellows help guide SSI training, policies, community work and consultancy engagements

ReproHacks

Anna Krystalli

- The aim isn't to undermine or discredit researchers or their work. Those organising and participating in ReproHacks recognise that reproducibility is a beneficial scientific activity in itself. The feedback that the events generate help researchers to improve their research, how it's presented and ultimately aiding reproducibility and accelerating research.
- Some other individual benefits for research software engineers and data scientists include:
  - Practical experience of reproducibility
  - Experience of working with other people's code and data
  - Helping to improve reproducibility in future projects
  - Increase the opportunities for the reuse of reproducible experiments
  - Experience of working with 'real' data

12 MAR
2020

ReproHack - The University of Manchester

A ReproHack is a one-day reproducibility hackathon. These events bring together researchers, RSEs and data scientists to reproduce papers from published scientific code and data.

Room F37, Sackville Street Building, University of Manchester

12 Mar 2020 10 a.m. — 5 p.m.

Reproducibility Initiative - International Semantic Web Conference 2019 (ISWC 2019)

Alejandra Gonzalez-Beltran (STFC, UK) & Michael Cochez (VU Amsterdam, the Netherlands)

SW Reproducibility Initiative

Alejandra Gonzalez-Beltran

- The **International Semantic Web Conference** (https://iswc2019.semanticweb.org/) is the premier international forum for the Semantic Web and Linked Data community; 18th edition was in New Zealand in October 2019
- Topics: knowledge graphs, linked data, linked schemas and AI on the Web
- For the first time, we ran a **Reproducibility Initiative** for ISWC submissions
  - Authors of accepted research track papers that include a significant quantitative experimental evaluation were invited to participate in a **'reproducibility certification'**
  - Dedicated Programme Committee of '**reproducers'**
  - The evaluation implies trying to replicate and/or reproduce the submission's results with the help of the authors
  - **'Reproducers'** were encouraged to interact with the authors to evaluate their submissions (via repositories in bitbucket)
  - Objectives
    - Encourage a culture of **data and code sharing**
    - Encourage a culture of **rigorous and transparent results**
    - Highlight the **impact** of semantic web research
    - Enable easy **dissemination of results**
  - 11 papers submitted
    - 8 papers reproduced
      - **only 3** papers did **not** need interaction with the authors to be reproduced
    - 3 not reproduced (yet)
  - We awarded the Best Reproduced paper

http://repro.semanticweb.org/

#PsyTeachR

- **Goal**: To develop undergraduate teaching ‘aimed at instilling the principles of transparency and scientific rigor’ from day one of the course.
- **Skill-Based Teaching**: Our research methods teaching now emphasises essential ‘data science’ graduate skills that have been overlooked in traditional approaches to teaching, including programming skills, data visualisation, data wrangling and reproducible reports.
- **The How**: Interactive R/Rstudio assignments to enhance analytical & communication skills, and instils a better understanding of Open Science. R is the Tool. Skills are the Focus.
- **Fellowship Aim**: To develop pedagogical collaborations that improve research methods teaching across the UK through workshops and seminars on how to integrate these skills into teaching, admin, and research.
- Better Software Better Teaching @softwaresaved

Phil McAleer

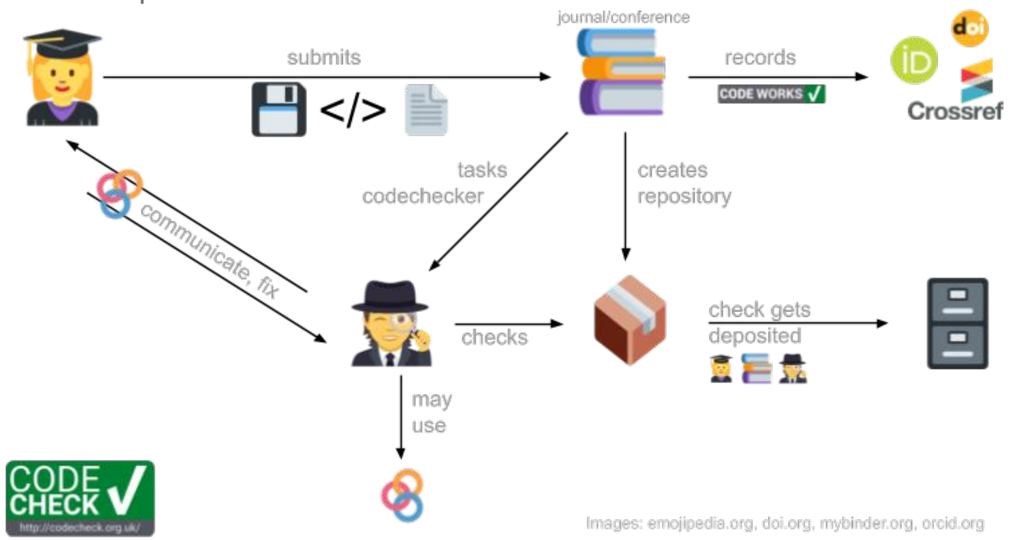
https://psyteachr.github.io

Phil McAleer
@mcaleerp @uofgpsychology
philip.mcaleer@glasgow.ac.uk
School of Psychology, University of Glasgow

https://doi.org/10.6084/m9.figshare.11898450

CodeCheck

- Supports codecheckers with a workflow, guidelines and tools to evaluate computer programs underlying scientific papers.
- The independent time-stamped runs conducted by codecheckers will award a “certificate of reproducible computation” and increase availability, discovery and reproducibility of crucial artefacts for computational sciences.



nature neuroscience

Commentary | Published: 25 May 2017

Toward standard practices for sharing computer code and programs in neuroscience

Stephen J Eglen, Ben Marwick, Yaroslav O Halchenko, Michael Hanke, Shoaib Sufi, Padraig Gleeson, R Angus Silver, Andrew P Davison, Linda Lanyon, Mathew Abrams, Thomas Wachtler, David J Willshaw, Christophe Pouzat & Jean-Baptiste Poline

Nature Neuroscience **20**, 770–773(2017) | Cite this article

1369 Accesses | **29** Citations | **323** Altmetric | Metrics

Computational techniques are central in many areas of neuroscience and are relatively easy to share. This paper describes why computer programs underlying scientific publications should be shared and lists simple steps for sharing. Together with ongoing efforts in data sharing, this should aid reproducibility of research.

https://www.nature.com/articles/nn.4550

Stephen Eglen

Open Life Science

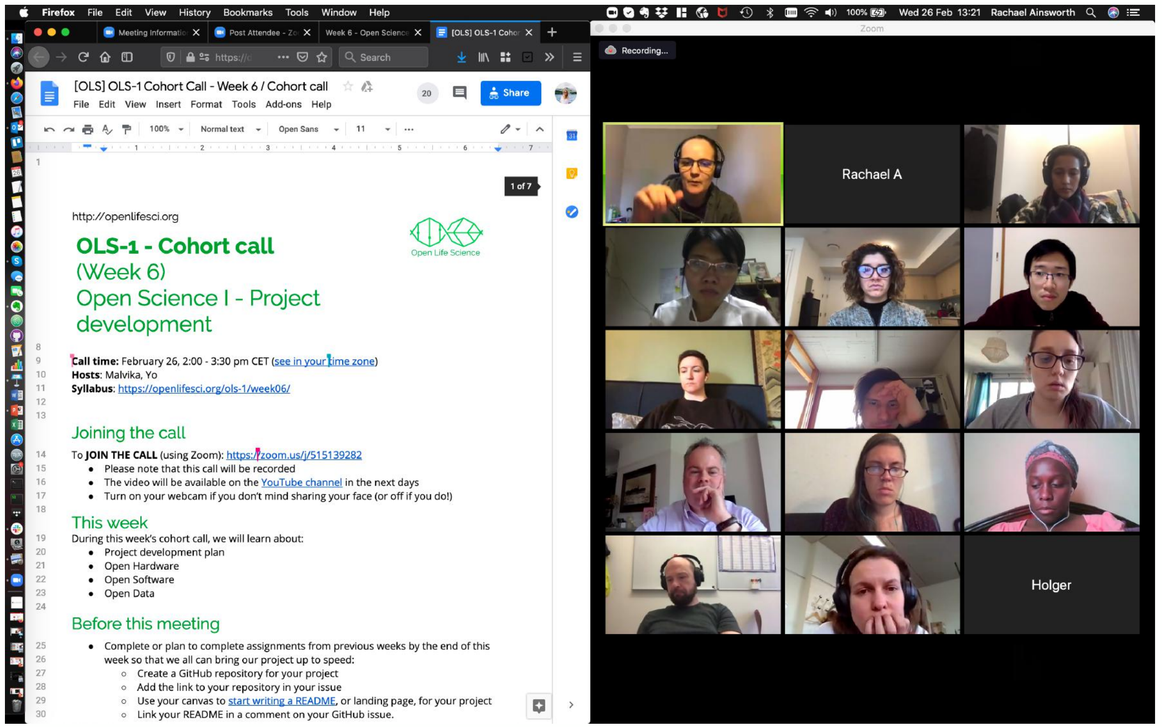
- A mentoring & training program for Open Science ambassadors in Life Science
- **Purpose**: Training for early stage researchers and young leaders interested in furthering their Open Science skills
- **Outcome**: Ambassadors for Open Science practice, training and education across multiple European and international bioinformatics communities.
- **Process**: A 15-week mentoring & training program, based on the Mozilla Open Leader program, helping participants in becoming Open Science ambassadors by using three principles:
  - Sharing essential knowledge required to create, lead, and sustain an Open Science project.
  - Connecting members across different communities, backgrounds, and identities by creating space in this program for them to share their experiences and expertise.
  - Empowering them to become effective Open Science ambassadors in their communities.

Yo Yehudi

Malvika Sharan

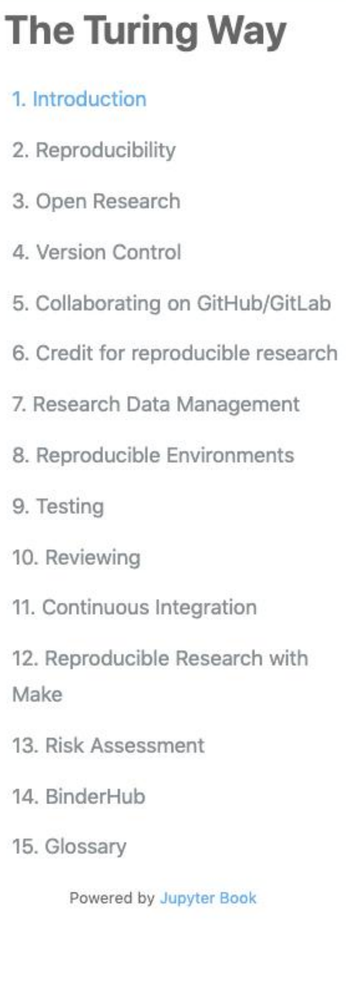
Open Life Science

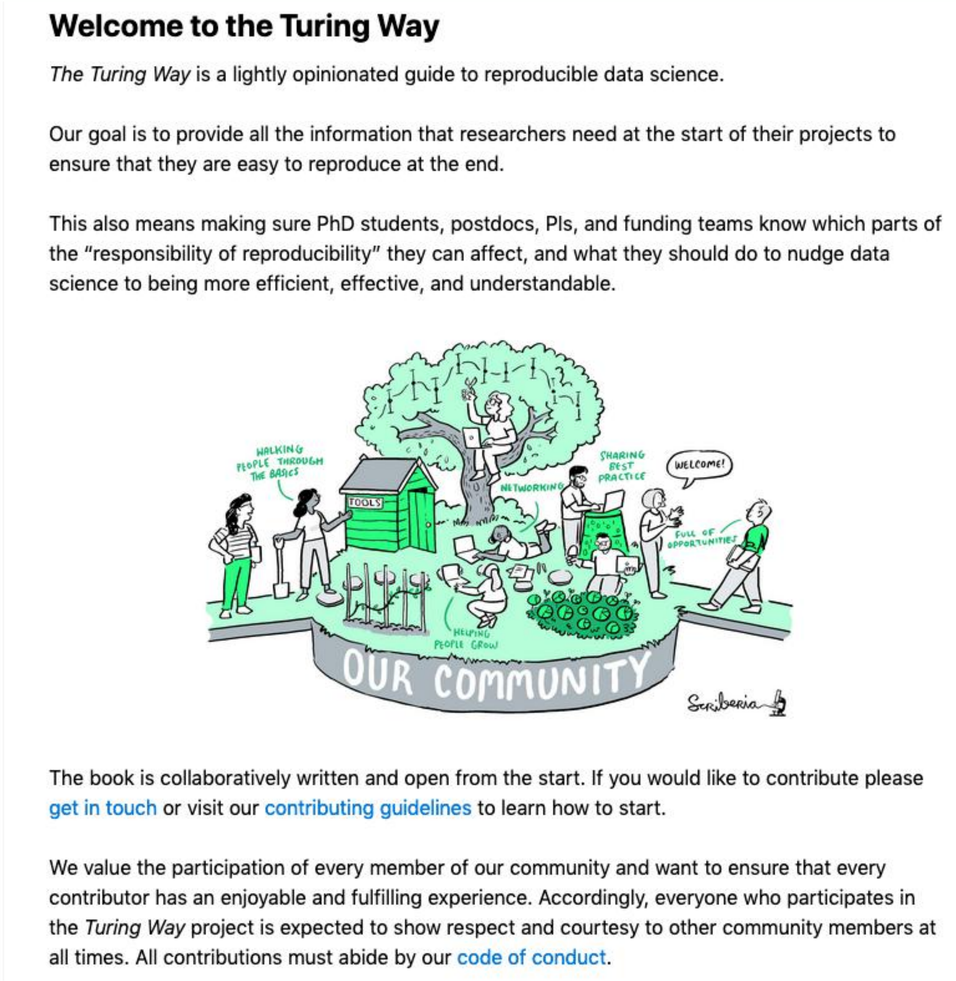
https://openlifesci.org/



The Turing Way Project

- Project led by Kirstie Whitaker at The Alan Turing Institute to make reproducible research "too easy not to do"
- In short: *The Turing Way* encompasses a handbook, community, collaboration, workshops and training
- Team of researchers, research software engineers, librarians and YOU!
- Demonstrates open and transparent project management and communication with future users, as it is openly developed at our GitHub repository: https://github.com/alan-turing-institute/the-turing-way

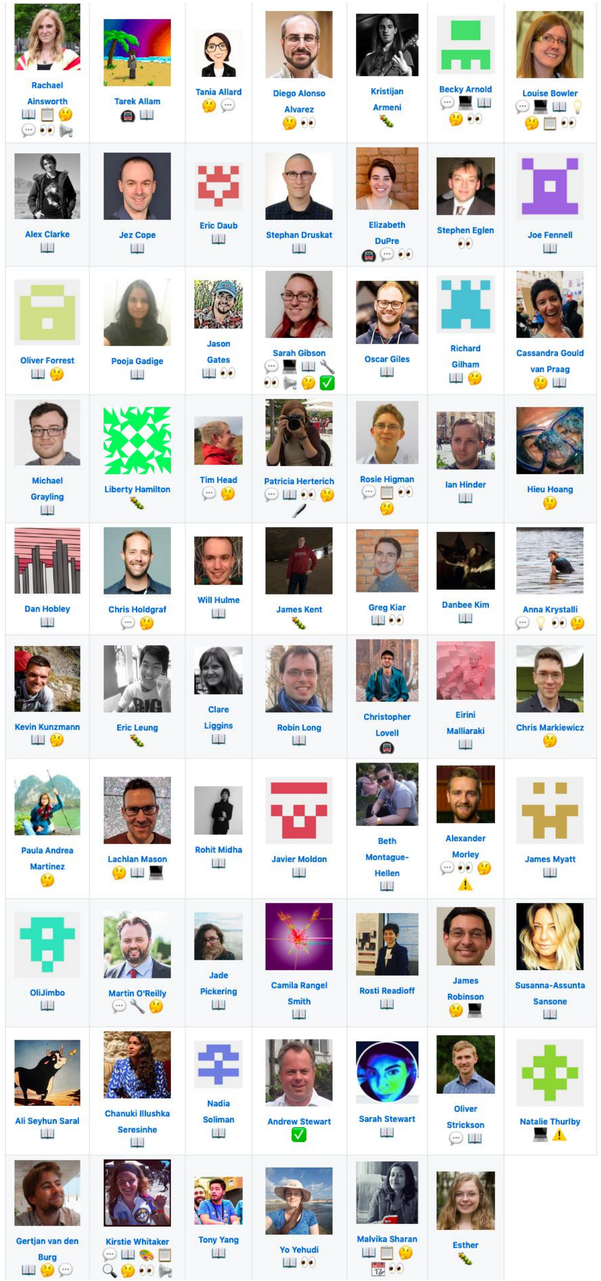




Handbook at: https://the-turing-way.netlify.com

The Turing Way Project

- Project led by Kirstie Whitaker at The Alan Turing Institute to make reproducible research "too easy not to do"
- In short: *The Turing Way* encompasses a handbook, community, collaboration, workshops and training
- Team of researchers, research software engineers, librarians and YOU!
- Demonstrates open and transparent project management and communication with future users, as it is openly developed at our GitHub repository: https://github.com/alan-turing-institute/the-turing-way



ata science.

at the start of their projects to

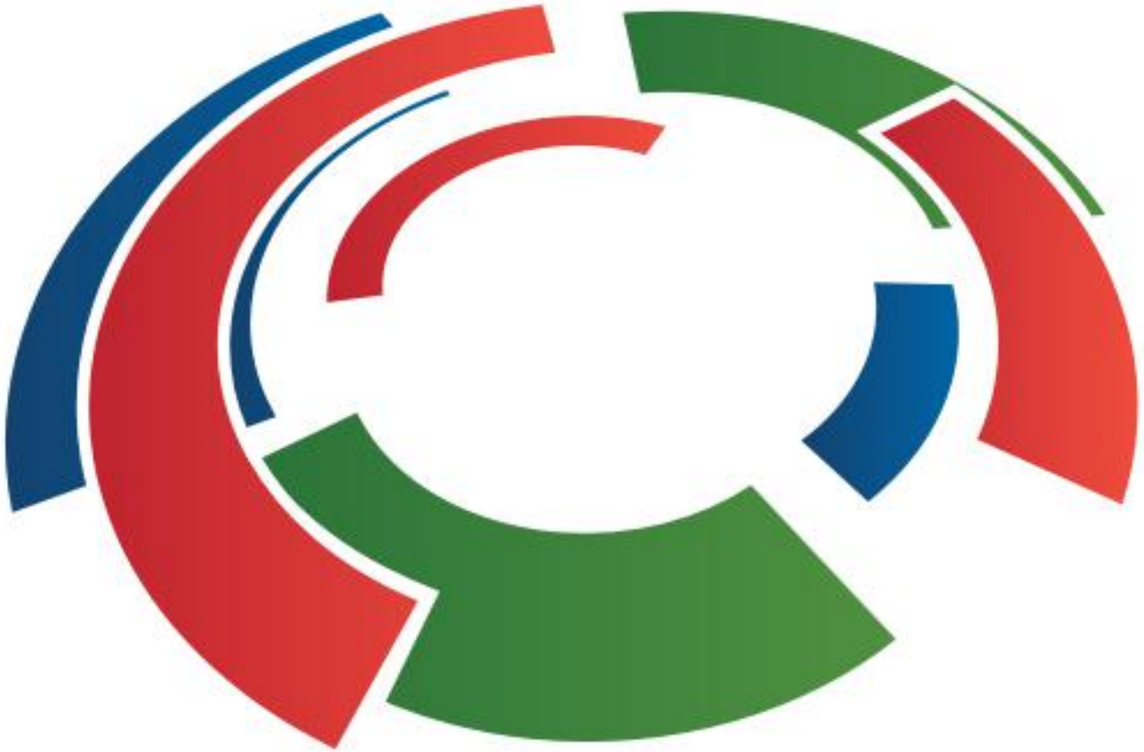
funding teams know which parts of
they should do to nudge data

ou would like to contribute please
start.

and want to ensure that every
gly, everyone who participates in
sy to other community members at

Ha

etlify.com



www.software.ac.uk