

The Alan Turing Institute

Are best practices applicable to your project?

*Contextualising The Turing Way for the
global Community*

Batool Almarzouq, Emma Karoune, & Malvika Sharan

Pronouns: she/her



Who are we?



Malvika Sharan

Senior Researcher, Open Research & Community, The Alan Turing Institute, Co-lead investigator of The Turing Way



Emma Karoune

Senior Community Manager for the Turing RSS Lab and the DECOVID project



Batool Almarzouq

Postdoctoral researcher in KAIMRC and lead of Open Science Community Saudi Arabia (OSCSA).

The Turing Way

An Open Source guide on Data Science.

*We involve and support a **diverse community** to make research **reproducible, ethical, open and inclusive** for everyone.*

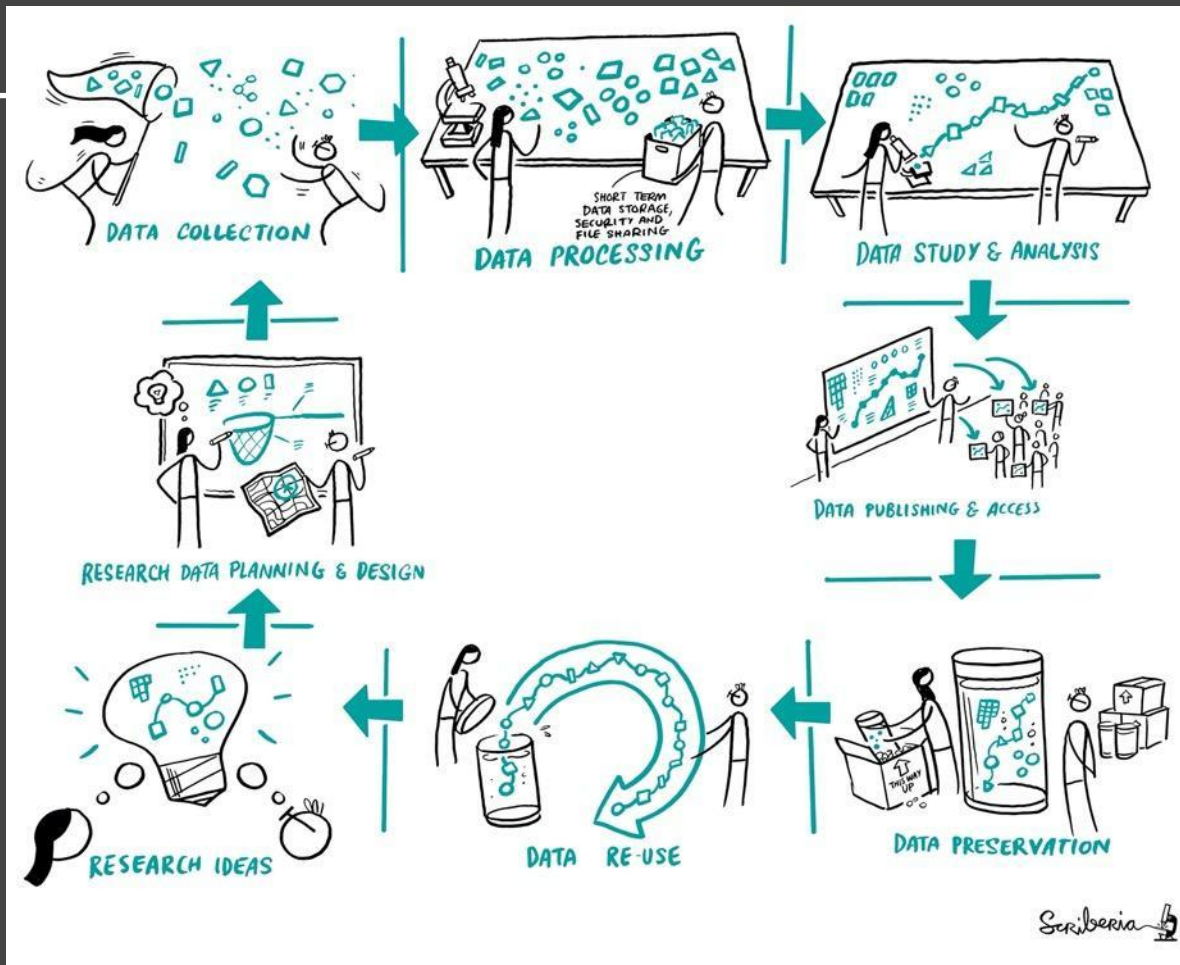


Reproducibility

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

Reproducibility

<https://turingway-arabic.netlify.app/welcome.html>



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



Fig. 3 The Turing Way project illustration by Scriberia. Used under a CC-BY 4.0

licence. DOI: [10.5281/zenodo.3332807](https://doi.org/10.5281/zenodo.3332807).

Moonshot Goal: Reproducibility “too easy not to do”



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

Visit our [GitHub Repository](#)

This book is powered by [Jupyter Book](#)

Welcome

The Turing Way is an open source community-driven guide to reproducible, ethical, inclusive and collaborative data science.

Our goal is to provide all the information that data scientists in academia, industry, government and the third sector need at the start of their projects to ensure that they are easy to reproduce and reuse at the end.

The book started as a guide for reproducibility, covering version control, testing, and continuous integration. However, technical skills are just one aspect of making data science research “open for all”.

In February 2020, *The Turing Way* expanded to a series of books covering reproducible research, project design, communication, collaboration, and ethical research.





Book:
the-turing-way.netlify.app/

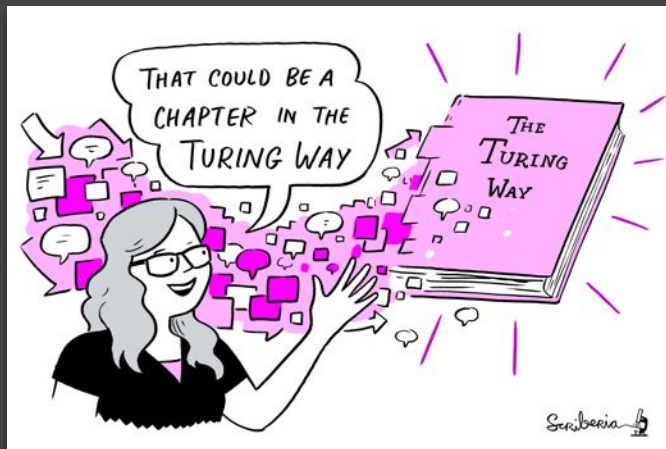
GitHub:
github.com/alan-turing-institute/the-turing-way

Twitter:
twitter.com/turingway

Email:
theturingway@gmail.com

CC-BY 4.0, *The Turing Way*

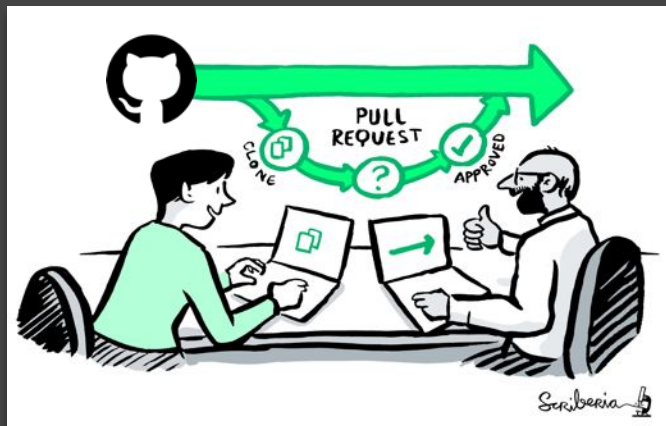
A Book with Five guides



A Community



An Open Source Project



A Culture of Collaboration



📁 .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

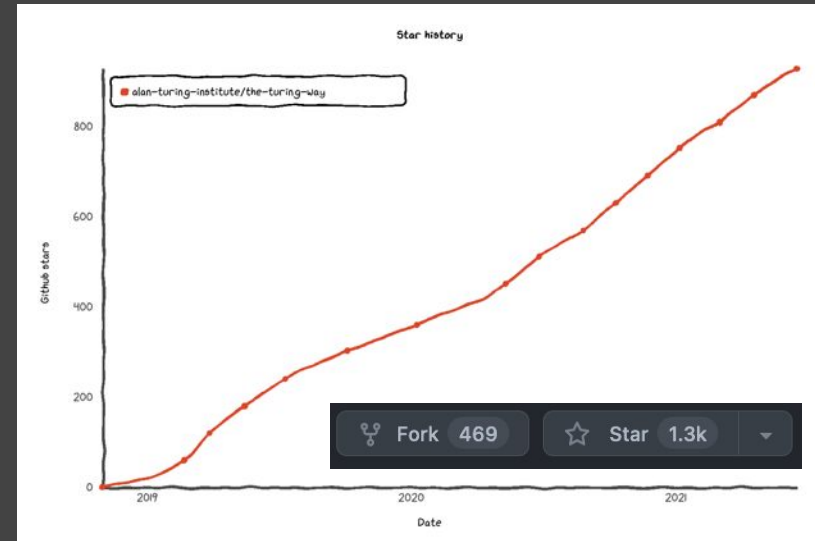


📄 .all-contributorsrc	Merge pull request #991 from alan-turing-institute/all-contributors/a...	5 days ago
📄 .gitignore	ignore pptx in workshop folder	9 months ago
📄 .travis.yml	add html-proof file again	last month
📄 <u>CODE_OF_CONDUCT.md</u>	her -> their	6 months ago
📄 <u>CONTRIBUTING.md</u>	Update CONTRIBUTING.md	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
📄 contributors.md	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

- Hosted on GitHub
- Open & collaborative
- CC-BY 4.0 License

Project and Community Growth

- 3 years, >200 Live Chapters
- Community resources, events, guidance, templates, training
- 300+ direct GitHub contributors and thousands of users



<https://zenodo.org/record/3332807>

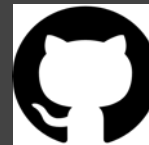


8,921

views

11,048

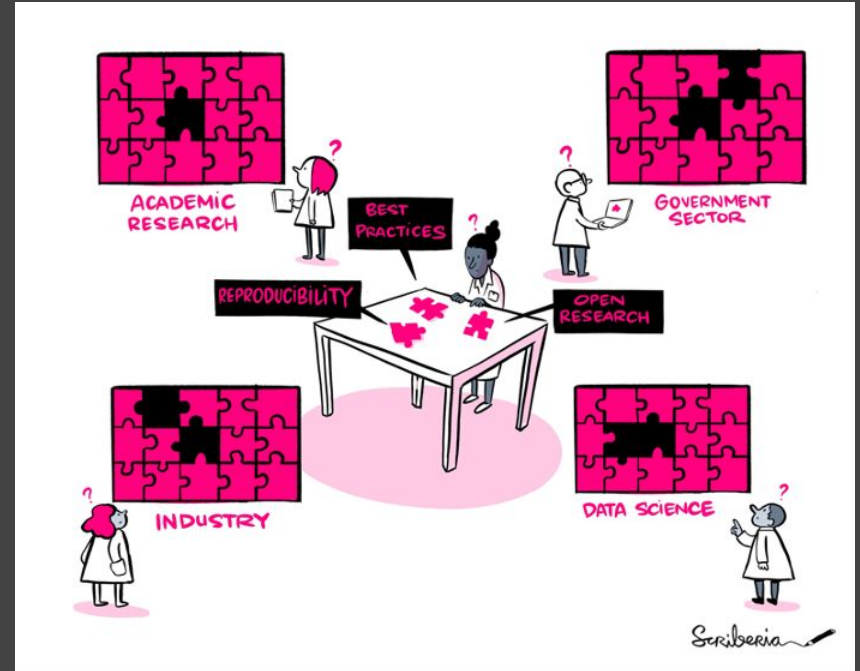
downloads



Contextualising Best Practices

UNESCO regards *science/research as a human right* - the greatest collective endeavour.

Best Practices vary based on the contexts and situations of researchers – *same approaches don't apply to everyone.*

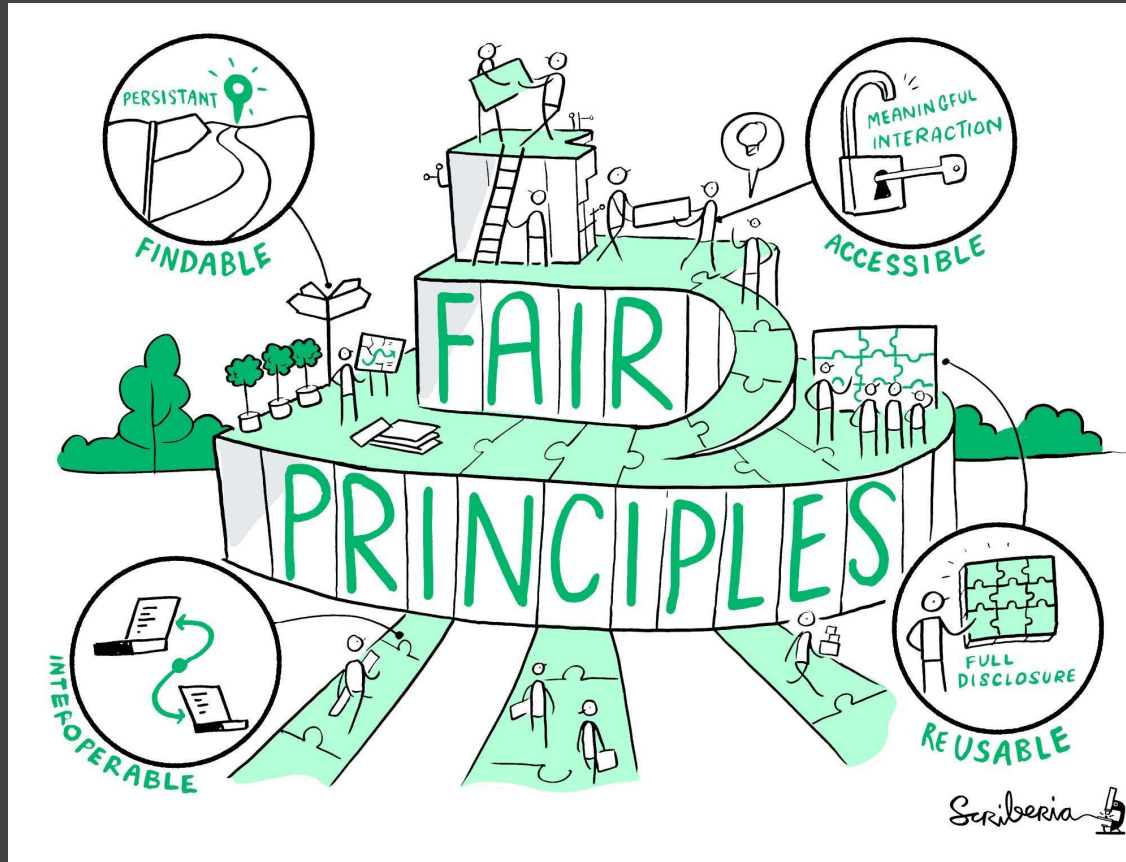


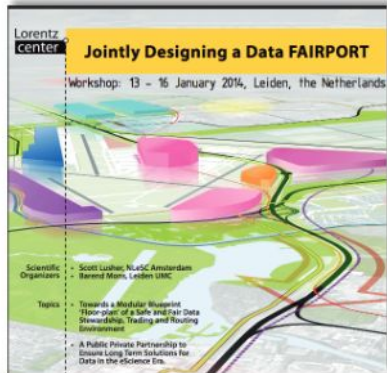
Right to enjoy the benefits of scientific progress and its applications. UNESCO. <https://en.unesco.org/human-rights/science>

@turingway, CC-BY 4.0, DOI: 10.5281/zenodo.6627260

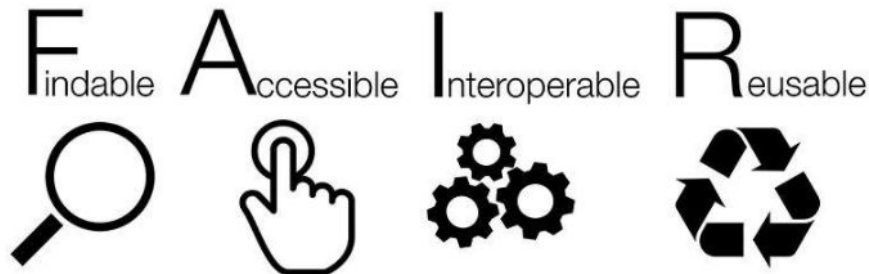
FAIR principles

- Findable
- Accessible
- Interoperable
- Reusable





2014



A set of principles to enhance
the value of all digital resources

Developed and endorsed by *researchers,*
service providers, publishers, funding
agencies and industry partners

SCIENTIFIC DATA 

SPRINGER NATURE

OPEN

**Comment: The FAIR Guiding
Principles for scientific data
management and stewardship**

SUBJECT CATEGORIES

- Research data
- Publication characteristics

Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Anie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Boume, Jildau Bouwman, Anthony J Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J G Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A.C. 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Momis A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons

SCIENTIFIC DATA | 3:160018 | DOI: 10.1038/sdata.2016.18

FAIR principles from **Wilkinson *et al.* (2016)**
DOI: 10.1038/sdata.2016.18

FAIR for

- Software
- Training materials
- Workflows
- And more!

FIVE RECOMMENDATIONS FOR FAIR SOFTWARE



PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS

EDITORIAL

Ten simple rules for making training materials FAIR

Leyla Garcia, Bérénice Batut, Melissa L. Burke, Mateusz Kuzak, Fotis Psomopoulos, Ricardo Arcila, Teresa K. Attwood, Niall Beard, Denise Carvalho-Silva, Alexandros C. Dimopoulos, Victoria Dominguez del Angel, Michel Dumontier, Kim T. Gurwitz, [...], Patricia M. Palagi [view all]

Published: May 21, 2020 • <https://doi.org/10.1371/journal.pcbi.1007854>

FAIR Computational Workflows

Carole Goble^{1†}, Sarah Cohen-Boulakia², Stian Soiland-Reyes^{1,4}, Daniel Garijo³, Yolanda Gil³, Michael R. Crusoe⁴, Kristian Peters⁵ & Daniel Schober⁵

¹Department of Computer Science, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

²Laboratoire de Recherche en Informatique, CNRS, Université Paris-Saclay, Batiment 650, Université Paris-Sud, 91405 ORSAY Cedex, France

³Information Sciences Institute, University of Southern California, Marina Del Rey CA 90292, USA

⁴Common Workflow Language project, Software Freedom Conservancy, Inc. 137 Montague St STE 380, NY 11201-3548, USA

⁵Leibniz Institute of Plant Biochemistry (IPB Halle), Department of Biochemistry of Plant Interactions, Weinberg 3, 06120 Halle (Saale), Germany

Keywords: Computational workflow; Reproducibility; Software; FAIR data; Provenance

Citation: C. Goble, S. Cohen-Boulakia, S. Soiland-Reyes, D. Garijo, Y. Gil, M.R. Crusoe, K. Peters & D. Schober. FAIR computational workflows. *Data Intelligence* 2(2020), 108–121. doi: 10.1162/dint_a_00033

FAIR implementation

- Domain specific implementation needed
- EOSC - European funding for FAIR
 - Pilot implementation projects
 - Infrastructure
 - Training

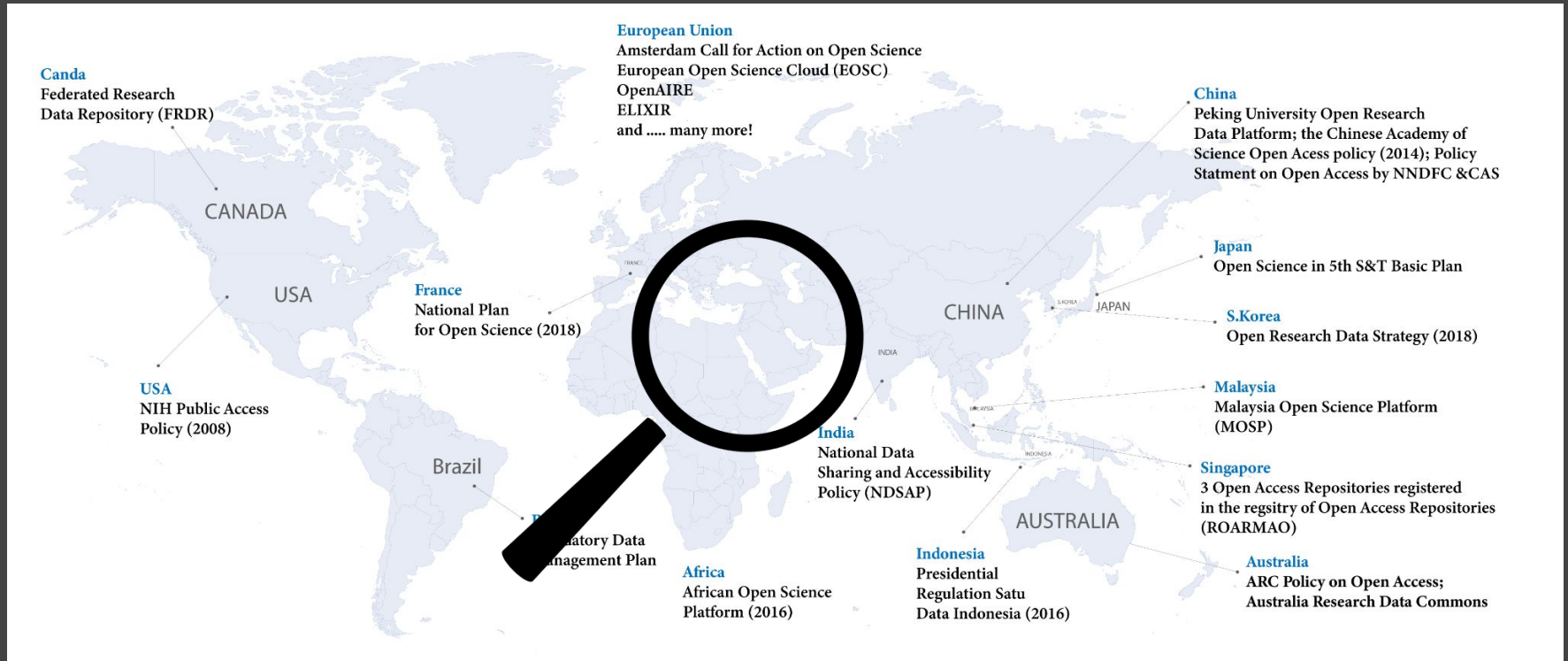


As open as possible, as closed as necessary

C - Collective benefit
A - Authority to control
R - Responsibility
E - Ethics

F - Findable
A - Accessible
I - Interoperable
R - Reusable

Landscape of Open Science



Overview of the current Open Science policies and Infrastructure around the Globe. Image adopted from Malaysian Open Science Platform by Open Science community Saudi Arabia (OSCSA), licensed under CC BY 4



There is **NO existence** for Open Science policies, resources with [minimal OS infrastructure](#) in the Arabic Speaking countries.

Accessing **Local Data** has always been **a struggle!**

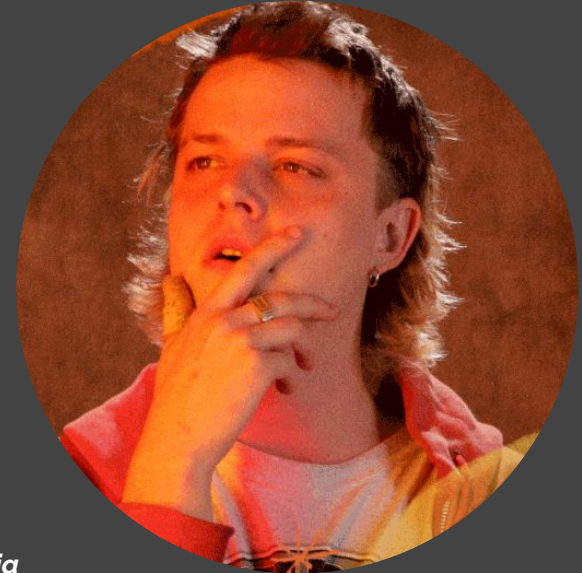
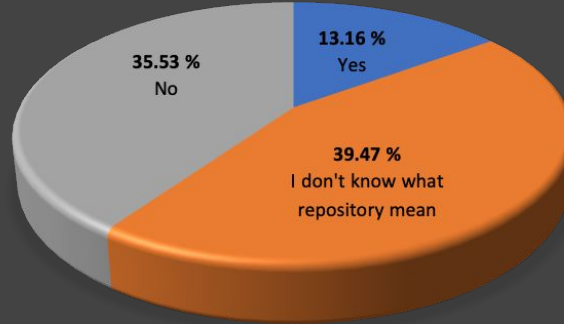
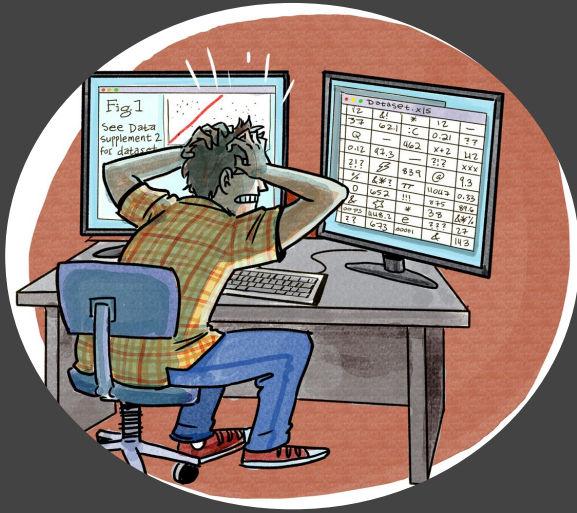


Illustration by [Ainsley Seago](#) (CC-BY 4.0).

A survey by **Open Science Community Saudi Arabia** for 80 individuals from 24-2-2022 until 11-4-2022.

Let's create DOI for my own raw data?

Limited repositories that generate DOI!
There's no DOI Registration Agencies (RAs) in the Arab World.



Let's use our institutional repository?

Oh, we don't have one!

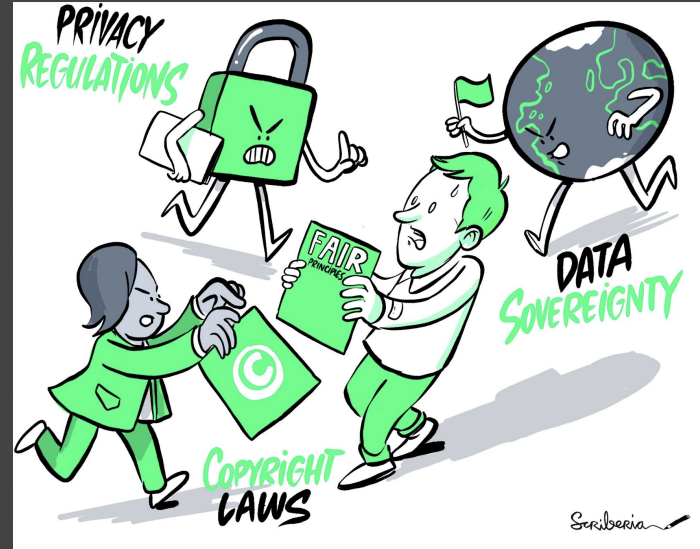
Let's use a general purpose repository?

Data sovereignty in many countries make it illegal to use general purpose repositories hosted outside the country

Let's use an Open License?

There is no open science policy, open license is very likely to be **unallowed** by main research funders

OSCSA Interviewed **individuals** representing different **initiatives/institution** in Arabic-speaking countries



Open science advocates in the MENA face **legal, technical, and cultural** blockers unfamiliar to champions of open science elsewhere.

Reflection

What are the challenges you have experienced in your own community while trying to adopt best practices developed elsewhere?

Write in our [shared document](#)

Please add any links and references.

Reflection

How can we maintain an international platform to create resources or best practices that are globally applicable?

Write in our [shared document](#)

Please add any links and references.

Reflection

How can we create opportunities to champion, capture and highlight local knowledge, practices and views?

Write in our [shared document](#)

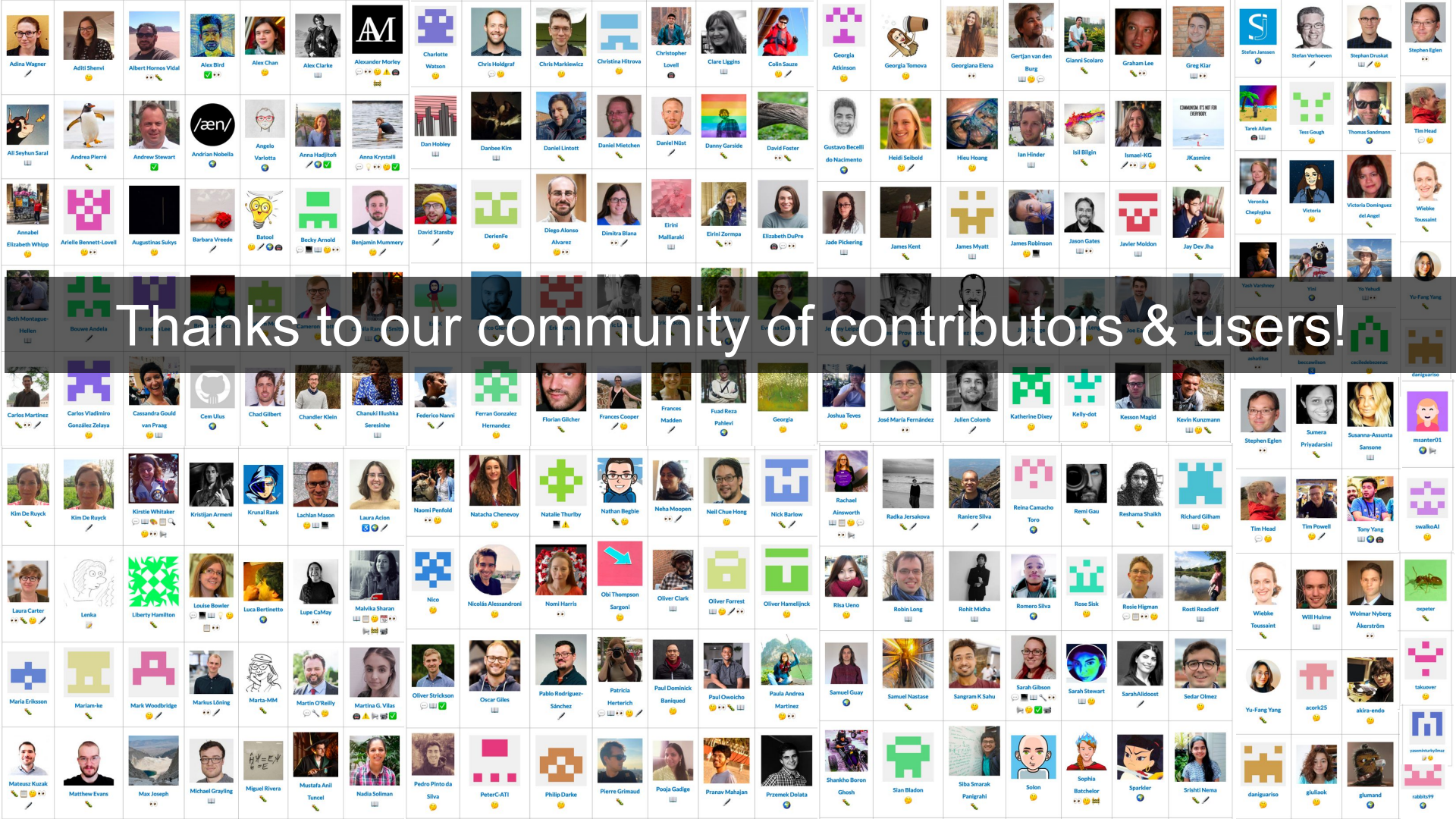
Please add any links and references.

Reflection

What are the different forms of knowledge production and outcomes as well as best practices you have benefitted from but don't see represented in the international community?

Write in our [shared document](#)

Please add any links and references.



Thanks to our community of contributors & users!

Pathways for Collaboration



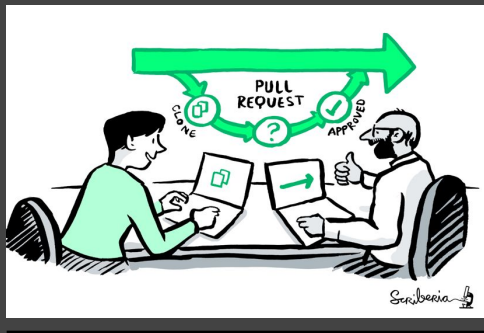
Connect with us



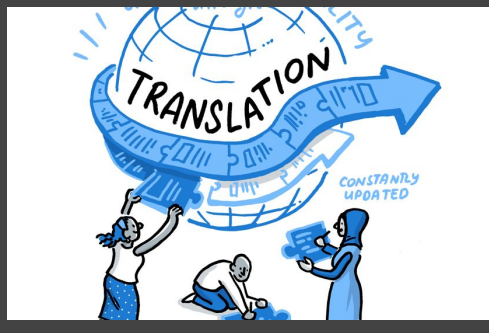
Start where you can



Discuss your ideas



Edit, review, update



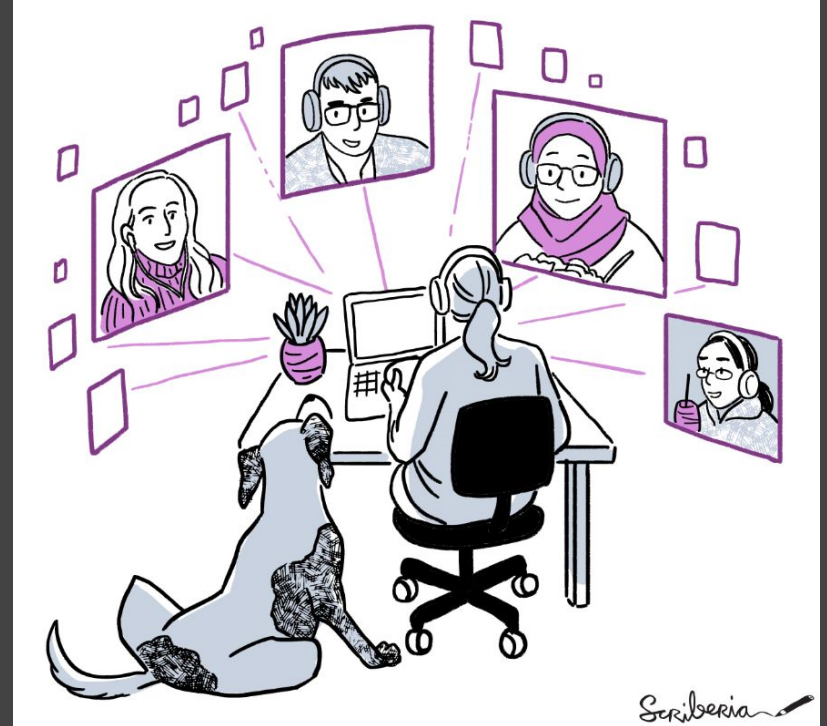
Help make it global



Join the community

Next Book Dash: **November 2022**

- Virtual event
- 4 days, short development sessions (1-2 each day)
- Collaborate, develop resource, add features



Communities to exchange best practices for open and reproducible research



Website: <https://osc-ksa.com/>

GitHub: github.com/Open-Science-Community-Saudi-Arabia



Website: <https://open-phytoliths.github.io/FAIR-phytoliths/>

GitHub: <https://github.com/open-phytoliths>

Open Life Science

A mentoring & training program for Open Science Ambassadors

- Sharing
- Connecting
- Empowering

Join the program

- 1:1 mentoring
- Training calls on Open Science
 - Talks from experts
 - Break out discussions
- Training for mentors

OLS-6 Period: Sep 2022 - Jan 2023
Apply individually or in team to join the cohort

Website: <https://openlifesci.org/>
Application for the next cohort is open!

Connect with us:

- Batool Almarzouq (@batool664), Core Contributor
- Emma Karoune (@ekaroune), Core Contributor
- Kirstie Whitaker (@kirstie_j), Project Co-Lead
- Malvika Sharan (@malvikasharan), Project Co-Lead
- Anne Lee Steele (@aleesteele), Community Manager
- *The Turing Way* community, friends & collaborators



Useful links:

- Twitter: twitter.com/turingway
- Newsletter: tinyletter.com/TuringWay
- GitHub: github.com/alan-turing-institute/the-turing-way
- Slack: <https://tinyurl.com/jointuringwayslack>
- Artwork by Scriberia: <https://doi.org/10.5281/zenodo.3332807>

