

Establishing a career in human-data interaction

Malvika Sharan

Connect @malvikasharan

27 March 2024



This work is licensed under a [Creative Commons Attribution 4.0 Generic License](https://creativecommons.org/licenses/by/4.0/)

Find this talk on Zenodo: <https://doi.org/10.5281/zenodo.10888503>, DOI: 10.5281/zenodo.10888503

A career in the intersection of data science and community building

Malvika Sharan

Connect @malvikasharan

27 March 2024



This work is licensed under a [Creative Commons Attribution 4.0 Generic License](https://creativecommons.org/licenses/by/4.0/)

Find this talk on Zenodo: <https://doi.org/10.5281/zenodo.10888503>, DOI: 10.5281/zenodo.10888503

London, UK

Co-Founded Open Life Science

Senior Researcher - The Alan Turing Institute

Co-lead of The Turing Way

Lead of Community Management Team

Germany

M.Sc, PhD, Community Building experience

Worked in European Molecular Biology Laboratory,
connected with Open Science projects

Bangalore,
Karnataka

**Bachelor's degree
and worked briefly**

Ranchi,
Jharkhand

Hometown

**The
Alan Turing
Institute**



Open Life Science



**Software
Sustainability
Institute**



**THE
CARPENTRIES**



**SOCIETY OF RESEARCH
SOFTWARE ENGINEERING**

CS&S

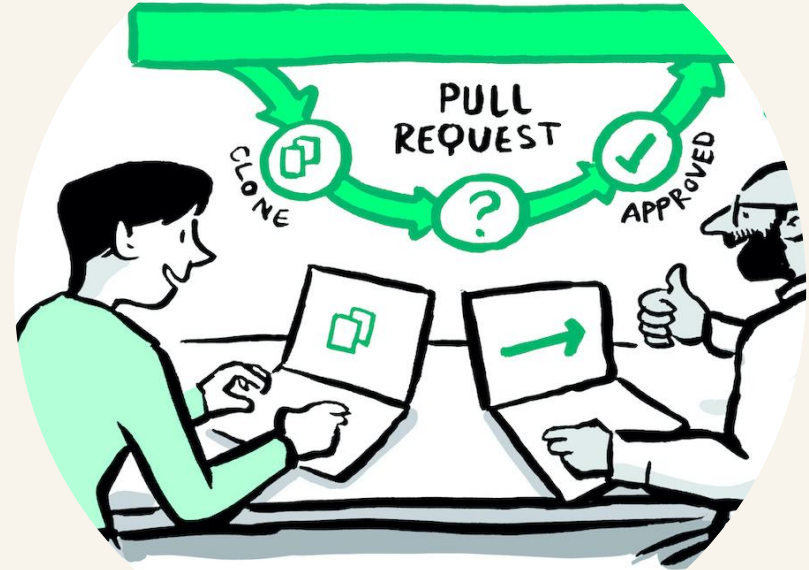
Code for
Science &
Society



**Meta
Docencia**

Early years as a researcher

- Theoretical understanding of data science
Evolving landscape of data and technology
- ML/AI were not new to biomedical domain
- Most researchers are not experts; most teams work in the open
- Lack of structured support/mentorship





Open science for equity


Combines many forms of practices to make scientific knowledge 'open' (by design).

Accessible for everyone

No price barriers, no technical barriers!

Inclusive for everyone

Enables everyone to participate and benefit from scientific advancements.



*Openness and access are **not binary conditions**, but highly situational, contingent and dependent on the privilege and context of **knowledge makers and gatekeepers**.*

Contextualizing Openness: Situating Open Science, Leslie Chan, Angela Okune, Rebecca Hillyer, Denisse Albornoz, and Alejandro Posada

Open science for equity

Combines many forms of practices to make scientific knowledge 'open' (by design).

Accessible for everyone

No price barriers, no technical barriers!

Inclusive for everyone

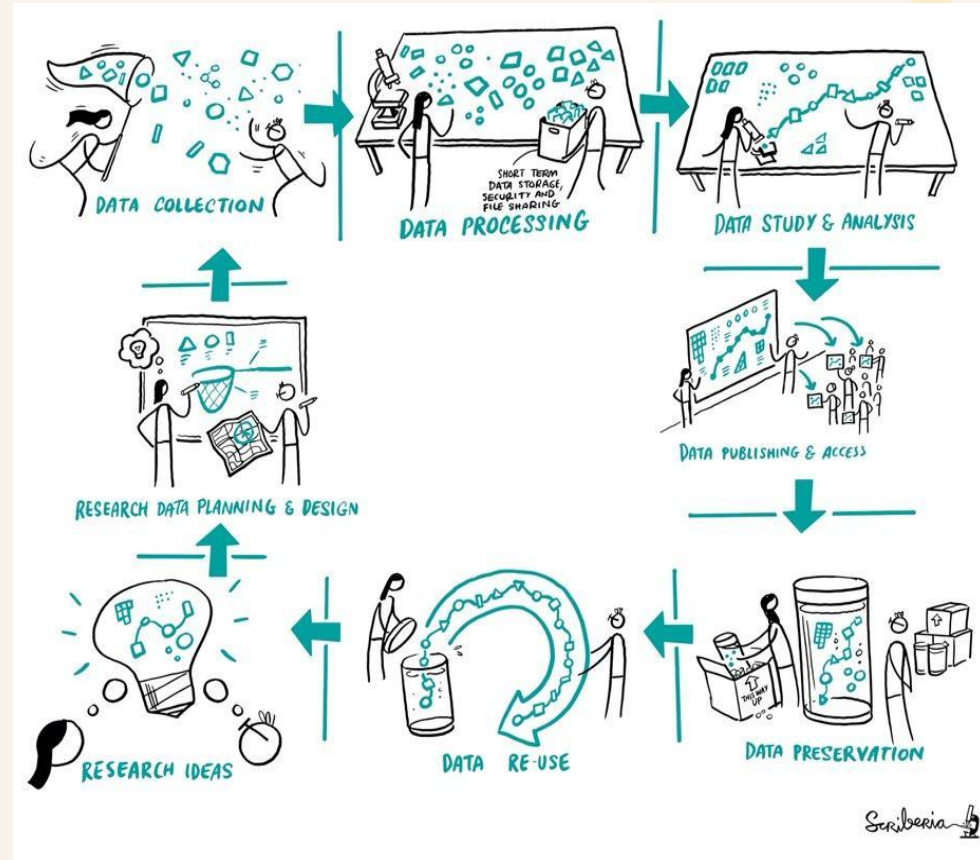
Enables everyone to participate and benefit from scientific advancements.



Data science is socio-technical

Human agency is intertwined with technology.

Tools, practices, and infrastructures we build, use and are impacted by - form complex socio-technical systems.



Applying socio-technical lens

- Close collaboration with lab researchers
- Open science and reproducible workflow
- Open and community-based tech skills
- Documentation & collaboration
- Policy advocacy and community building



Germany
PhD

<https://www.open-bio.org>, BOSC: <https://www.open-bio.org/2019/08/27/tips-for-informal-discussions/>,
<https://www.software.ac.uk/blog/2020-12-17-ten-arguments-against-open-science-you-can-win>,
<https://www.uni-wuerzburg.de/en/university>, <https://carpentries.org/>
FOSTER-UNESCO Open Science for Doctoral Schools,. (2023, March 29). Retrieved from
<https://www.fosteropenscience.eu/event/foster-unesco-open-science-doctoral-schools>

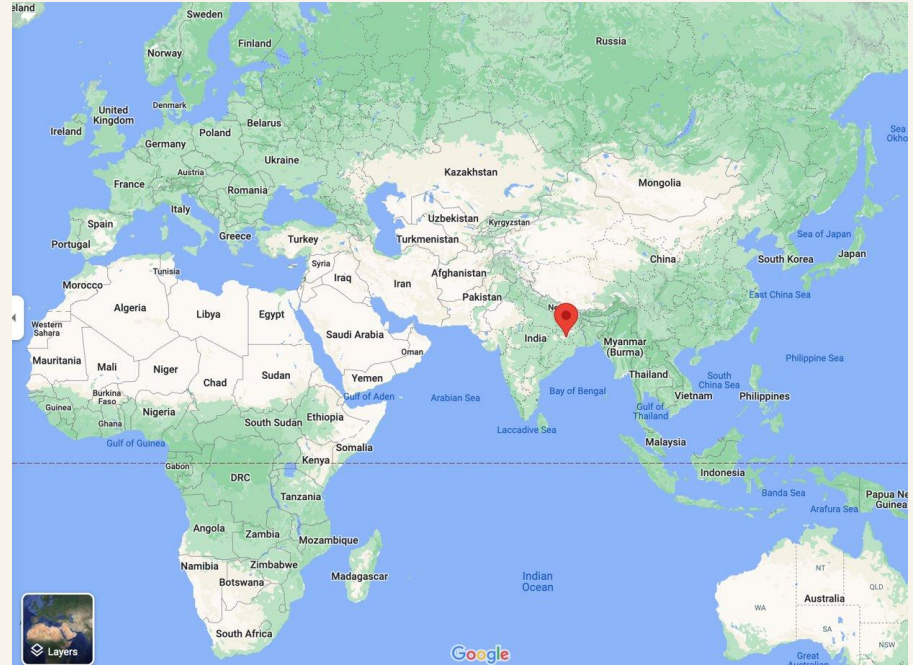
My positionality

Jharkhand - “Land of Forests”
Independent state since 2000

Also a place that shaped my curiosity for learning about nature, biology, social dynamics and my deep understanding of community

Bangalore
Undergrad in
Biotech

Jharkhand
Hometown



Community building

Process of sharing access to skills and resources that an individual or a group needs to participate in a community, and influence decision-making process.

Opportunity to connect, belong and contribute towards a shared purpose.

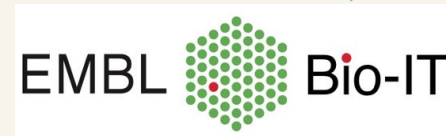
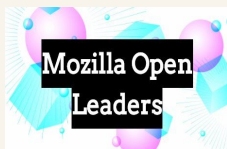


Digital skills & practices through community building

Spaces to listen, ask question, explore solutions, collaborate, build shared infrastructure, celebrate and recognise.

Opportunities for open leadership.

Germany
EMBL



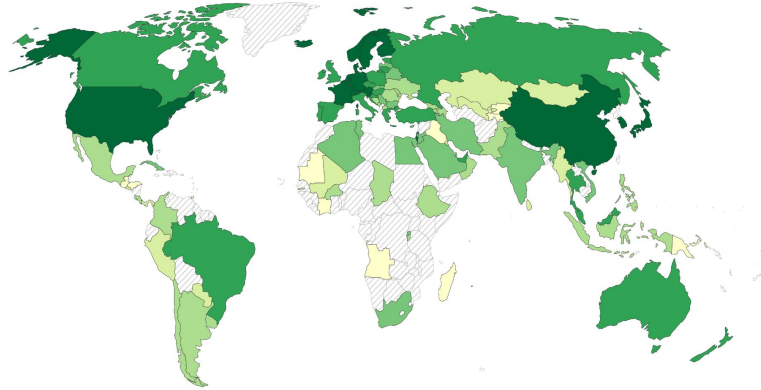
@malvikasharan, Logos from respective websites, Shared under CC-BY 4.0, DOI: 10.5281/zenodo.5518163, The de.NBI/ELIXIR-DE training platform - Bioinformatics training in Germany and across Europe within ELIXIR. F1000Research, 8. <https://doi.org/10.12688/f1000research.20244.2>

Economic and knowledge divide

Research & development spending as a share of GDP, 2021

Includes basic research, applied research, and experimental development.

Our World
in Data



No data 0% 0.1% 0.2% 0.5% 1% 2%

Source: UNESCO (via World Bank)

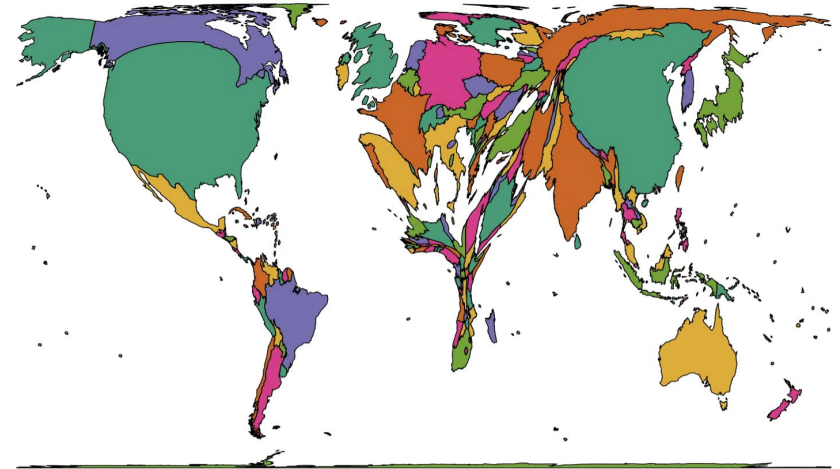
OurWorldInData.org/research-and-development • CC BY

Note: Spending includes current and capital expenditures (public and private) on research.

Distribution of AI Research Publications

2014-2023

Geographic distribution



□ = 50 000

Total: 3.51 million

Mozilla Open
Leaders



Software
Sustainability
Institute

OECD, Research and development (R&D) - Gross domestic spending on R&D - OECD Data. (2023). <https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm>. ourworldindata.org/grapher/research-spending-gdp | "Research on Artificial Intelligence – the global divides - TL:DR - Digital Science." Digital Science, 2024. PricewaterhouseCoopers. [Research and methodology. PwC Global AI study](#), 2024.

Economic and knowledge divide

Research & development spending as a share of GDP, 2021
Includes basic research, applied research, and experimental development.

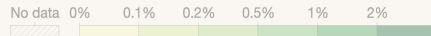
Our World
in Data

Distribution of AI Research Publications
Geographic distribution

2014-2023

Total global R&D expenditures is \$2.47 trillion in 2022 (\$726 billion in 2000) - with 70% countries account for less than 1% of the world's average spending.

AI could contribute up to \$15.7 trillion to the global economy by 2030, with only \$1.7 trillion of this economic impact is expected to impact the Global South.



Source: UNESCO (via World Bank)

OurWorldInData.org/research-and-development • CC BY

Note: Spending includes current and capital expenditures (public and private) on research.

= 50 000

Total: 3.51 million

Mozilla Open
Leaders



Software
Sustainability
Institute

OECD, Research and development (R&D) - Gross domestic spending on R&D - OECD Data. (2023). <https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm>. ourworldindata.org/grapher/research-spending-gdp | "Research on Artificial Intelligence – the global divides - TL:DR - Digital Science." Digital Science, 2024. PricewaterhouseCoopers. [Research and methodology. PwC Global AI study](#), 2024.

Recognising global collaborators in data science



“What requires recognition is [...] the status of individual group members as full partners in social interaction.

Misrecognition, accordingly, [is being] prevented from participating as a peer in social life.”

– Nancy Fraser, Rethinking Recognition

Co-creating ethical and responsible systems for collaborations

- Who makes decisions?
- Who gets to participate in knowledge/tech production?
- Who controls and protects this system?
- Who enjoys and benefits from the outcome?
- Who is responsible for addressing challenges and harm?
- How can open source/science further knowledge equity?
- How can you/I take a leadership role in service of others?

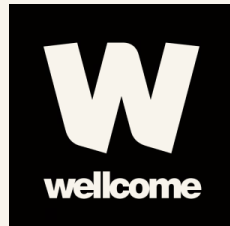


The **OLS** (formerly Open Life Science) is an open science capacity building organisation. Known for our 16-week long project-based mentoring and training programme, Open Seeds.

8 Cohorts, 265+ projects. Fellowship for leadership development.

Delivering 5 cohorts of 6 week long cohorts with **NASA-TOPS - Nebula cohorts**.

Chan
Zuckerberg
Initiative 



The
Alan Turing
Institute

CS&S Code for Science & Society
Event Fund

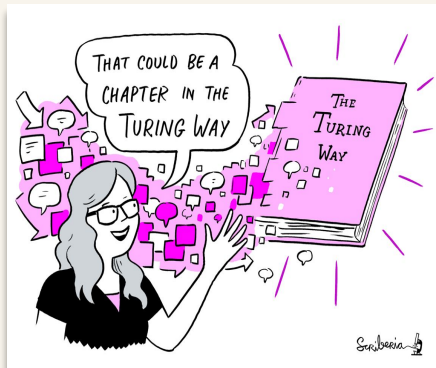


The Turing Way



A community-led handbook to best practices in Data Science.

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



Book



Community



Global Collaboration



Culture Change



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



The project belongs to the community

- **5+ years**, >300 Chapters, community infrastructure
- **450+ contributors**, **6000+ monthly users** - community-led/informed governance
- **Translation** in 5+ languages.
- **Collaboration** with 50+ organisations.
- **Replication/extension** by 12+ projects
- **Data Science Without Borders** to contextualise practices in African Countries



19,098
views

16,710
downloads



Open Post Academics



The Alan Turing Institute

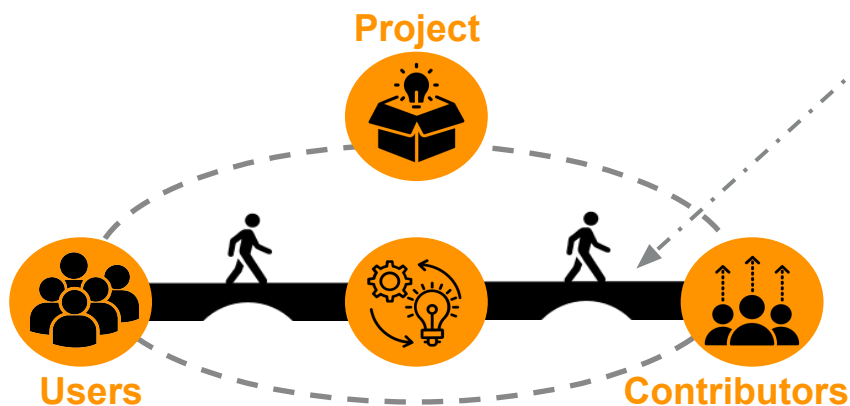
UK's National Institute for Data Science and AI

- Advance world-class research and apply it to national and global challenges
- Build skills for the future
- Drive an informed public conversation



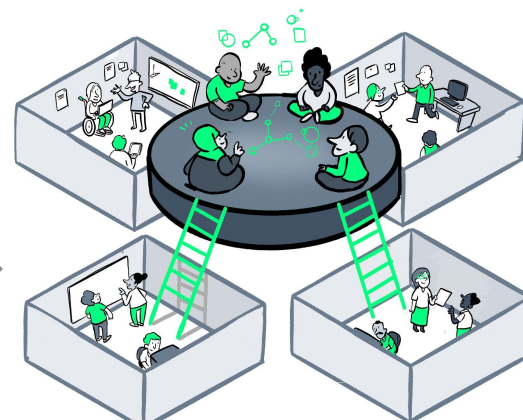
Tools, practices and systems

Building open source infrastructure to empower a global, decentralised network of people who connect data with domain experts



Strategic and technical support for
Stakeholder Engagement and Participation

Community Management



Researchers and collaborators



Dr Emma Karoune
Senior Research Community
Manager, Health | Tools,
Practices and Systems



**Dr Cassandra
Gould Van Praag**
Senior Research Community
Manager for the Environment
and Sustainability Grand
Challenge



Dr Gabin Kayumbi
Senior Research Community
Manager, Data Centric
Engineering (DCE) | Tools,
Practices and Systems



Anne Lee Steele
Research Community
Manager, The Turing Way |
Tools, Practices and Systems



Vicky Hellon
Research Community
Manager, Turing-Roche
Partnership | Tools, Practices
and Systems



Ayesha Magill
Widening Participation Officer



Sophia Batchelor
Research Community
Manager



Denise Bianco
Skills Officer



**Alexandra
Araujo Alvarez**
Research Project Manager

Better Recognition of Infrastructure Roles

- **The Turing Way Practitioners Hub** in close collaboration with UK orgs (government, industry, public sector): www.turing.ac.uk/turing-way-practitioners-hub
- **The Turing's Skill Policy Award** for engaging with policy ecosystem - led by **Dr. Emma Karoune**



BridgeAI



British Antarctic Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL



London, UK

Co-Founded Open Life Science

Senior Researcher - The Alan Turing Institute

Co-lead of The Turing Way

Lead of Community Management Team

Germany

M.Sc, PhD, Community Building experience

Worked in European Molecular Biology Laboratory,
connected with Open Science projects

Bangalore

**Bachelor's degree
and worked briefly**

Ranchi,
Jharkhand

Hometown

**The
Alan Turing
Institute**



Open Life Science



Software
Sustainability
Institute



**THE
CARPENTRIES**



**Meta
Docencia**



SOCIETY OF RESEARCH
SOFTWARE ENGINEERING

CS&S Code for
Science &
Society

My career in data science & community building



Ranchi,
Jharkhand

@malvikasharan, Image of Patratu Valley in Ranching, Is Patratu Valley Jharkhand's hidden gem? [Precious Rongmei](#), iStock, The Times of India

The
Alan Turing
Institute



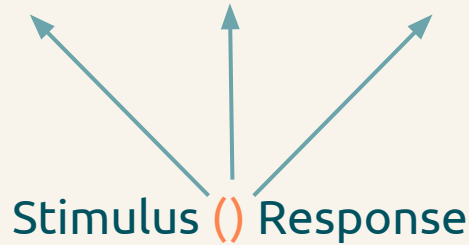
Open Life Science



Freedom, Authenticity, Balance



Freedom, Authenticity, Balance



“Between stimulus and response there is a space. In that space is our power to choose our response. In our response lies our growth and our freedom.”

Viktor Frankl*

Freedom, Authenticity, Balance

Stimulus () Response

Human-centered

Community building

Courage

Joy

Creativity

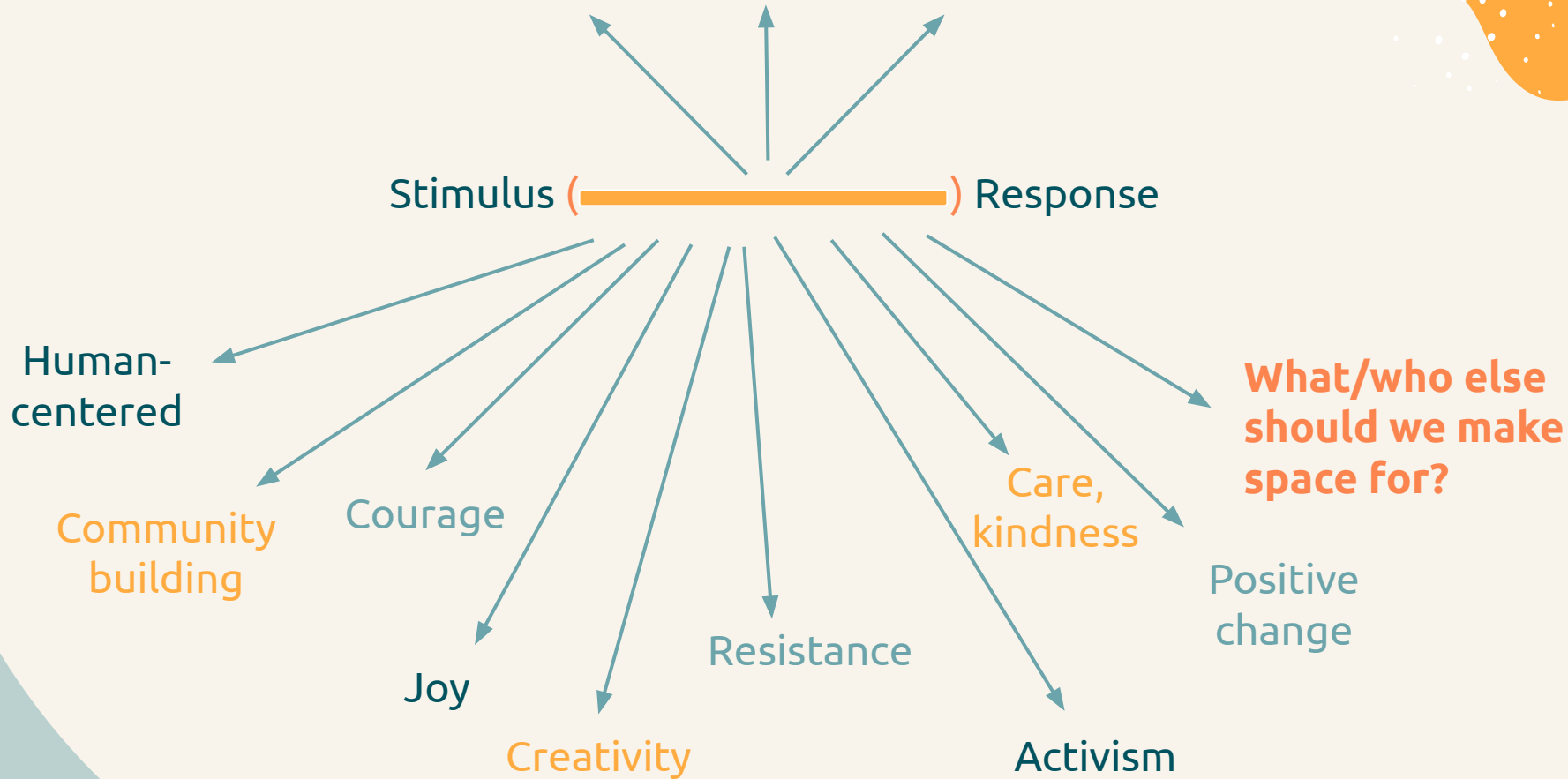
Resistance

Care, kindness

Activism

What/who else should we make space for?

Positive change



Acknowledgements & Links

The Turing Way

- Kirstie Whitaker, The Turing Way team and community
- Book: the-turing-way.netlify.com, Twitter: twitter.com/turingway
- GitHub: github.com/alan-turing-institute/the-turing-way, Links: <https://bit.ly/turingway>

OLS; Open Life Science

- Yo Yehudi, Bérénice Batut, Emmy Tsang, OLS team and community
- Website: openlifesci.org, Twitter: twitter.com/openlifesci
- GitHub: github.com/open-life-science/open-life-science.github.io

Connect with me

- Website: malvikasharan.github.io, *@malvikasharan on most platforms*