Strategies for transition to an open science culture: What works, what doesn't?

Researcher to Reader Conference

Tuesday 25 February 2025

© Shutterstock / Aphelleon

Researcher to Reader

SPRINGER NATURE

Open data challenges

€10.2 billion annual cost to European economy through lack of FAIR data[1]

$28 billion annual cost to US of irreproducible biology research[2]

75%+ failure rates reported by Pharma replicating conclusions of peer-reviewed papers[3, 4]

MISCONDUCT RETRACTIONS

The number of biomedical research papers retracted because of reasons related to misconduct has risen since 2000.

— Authorship and/or affiliation problems — Duplication
— Ethical and legal problems — Falsification/fabrication
— Plagiarism — Unreliable data — Unreliable results

©nature

1. European Commission, Directorate-General for Research and Innovation, Publications Office, 2019, https://data.europa.eu/doi/10.2777/02999
2. Freedman, L. P., Cockburn, I. M. & Simcoe, T. S. *PLoS Biol.* **13**, e1002165 (2015) http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002165
3. Begley, C. G. & Ellis, L. M. *Nature* **483**, 531–533 (2012), 4. Prinz, F., Schlange, T. & Asadullah, K. *Nature Rev. Drug Discov.* **10**, 712 (2011)
5. https://doi.org/10.1038/d41586-024-01609-0
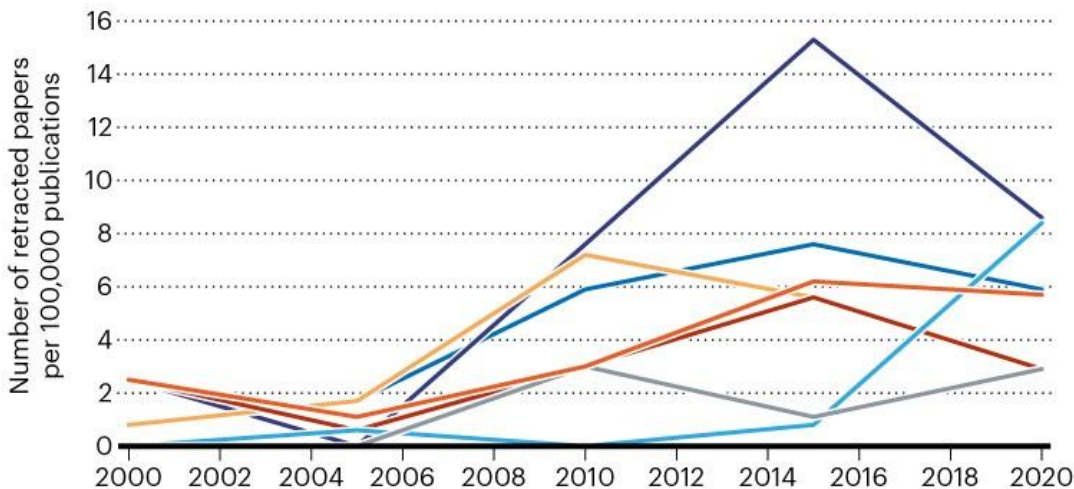
Policy response

NIH issues a seismic mandate: share data publicly

The data-sharing policy could set a global standard for biomedical research, scientists say, but they have questions about logistics and equity.

THE WHITE HOUSE

Breakthroughs for All: Delivering Equitable Access to America's Research

AUGUST 25, 2022 • OSTP BLOG

"**Scientific data** underlying peer-reviewed scholarly publications resulting from federally funded research should be made **freely available and publicly accessible** by default at the time of publication..."

wellcome

As a minimum, the data underpinning research papers should be made available to other researchers at the time of publication, as well as any original software that is required to view datasets or to replicate analyses.

Researcher to Reader

SPRINGER NATURE

Percentage of papers with a link to a dataset in the CZI corpus, by country
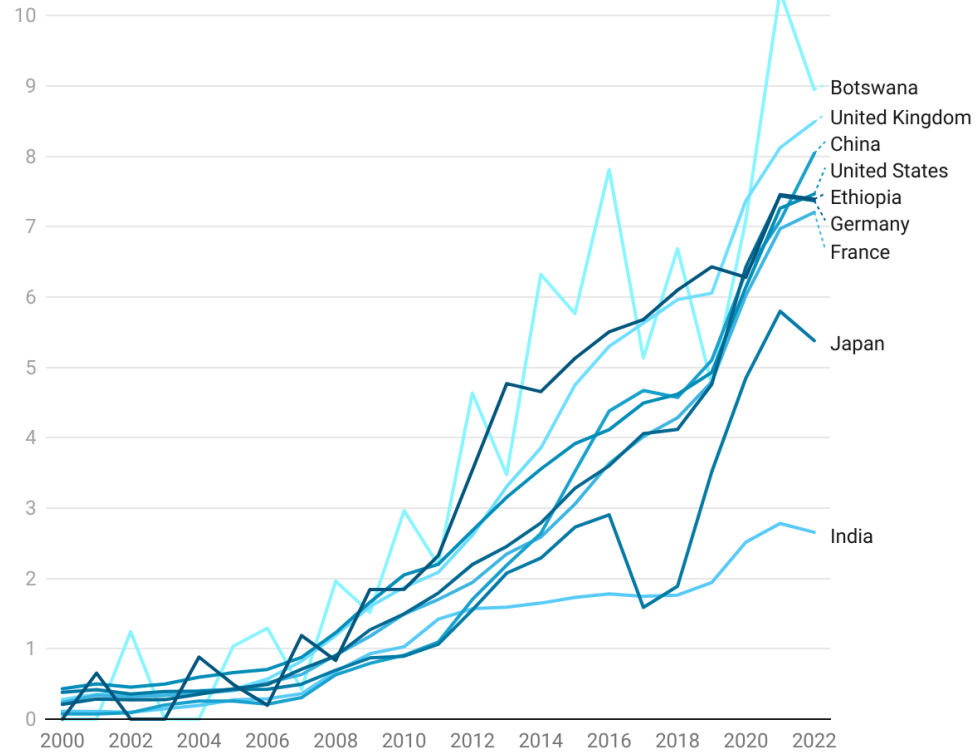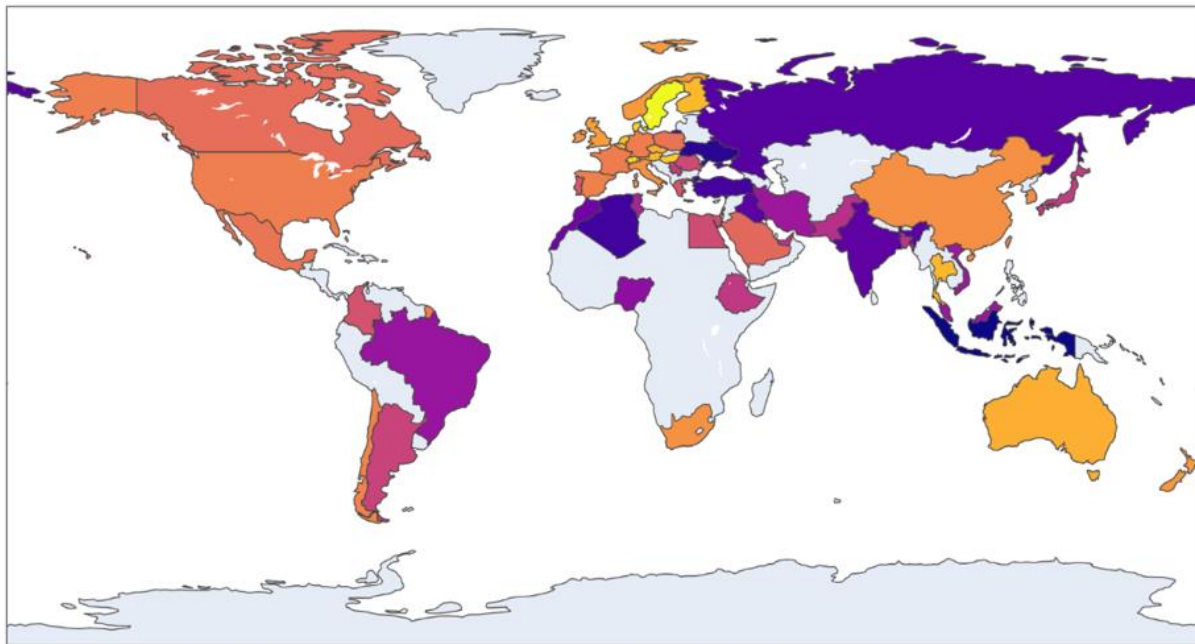
Citing percentage CZI DCC

A **Digital Science** Solution

Panellists

Beth Montague-Hellen

Head of Library & Information Services

The Francis Crick Institute

Riana Bahl

Public Health MSc student London School of Hygiene and Tropical Medicine

Biomedical Science BSc graduate King's College London

Mark Hahnel

VP Open Research at Digital Science

Founder of Figshare

Researcher to Reader

WRITE · REVIEW · PUBLISH · DISTRIBUTION · READ · RESEARCH

SPRINGER NATURE

What doesn't work?

Policies…or at least policies on their own

- Do people know about it?
- Does anyone read it?
- Or remember it at the right time?
- Is it being enforced?
- Are there clear benefits?

When it's just a policy, the usual answers to these questions are all NO!

What does work?

Specific Guidance from:

Institutions

Journals

Funders

→

Guidance

- Where?
- When?
- How?
- If?

Researcher to Reader

SPRINGER NATURE

What does work
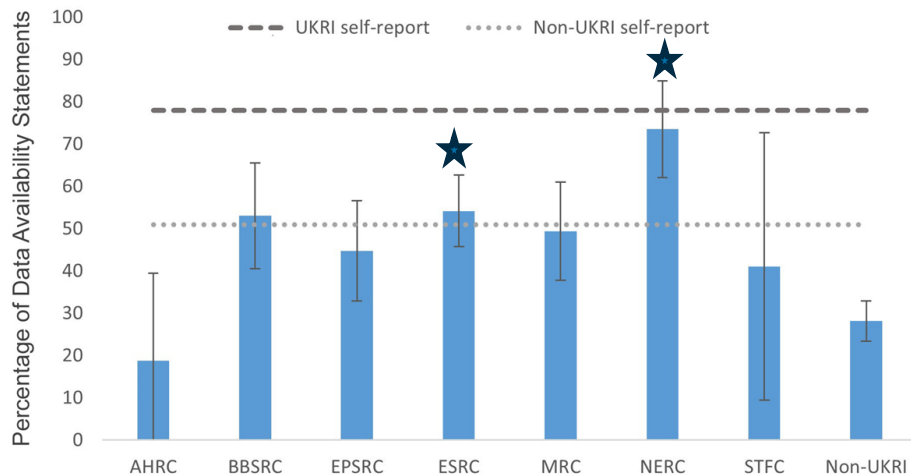
Montague-Hellen, B., & Montague-Hellen, K. (2023). Publishers, funders and institutions: who is supporting UKRI-funded researchers to share data?. *Insights, 36*(1).

How supportive would you be of a national mandate for making research data openly available?

Ethiopia 76.96%
United States 61.22%
Japan 41.86%
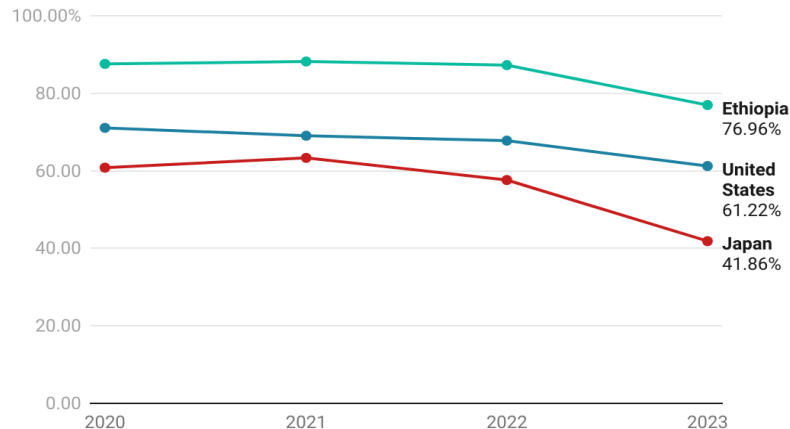
How familiar are you with the FAIR data principles in relation to Open Data?

■ I am familiar with the FAIR data principles　■ I have previously heard of the FAIR data principles but I am not familiar with them　■ I have never heard of the FAIR data principles

A. Ethiopia

	2020	2021	2022	2023
Never heard	54.23%	49.69%	35.87%	20.80%
Previously heard	27.97%	32.30%	33.65%	42.70%
Familiar	17.80%	18.01%	30.48%	36.50%

B. United States

	2020	2021	2022	2023
Never heard	52.02%	46.47%	44.03%	28.30%
Previously heard	26.28%	29.29%	28.85%	34.40%
Familiar	21.69%	24.25%	27.12%	37.60%

C. Japan

	2020	2021	2022	2023
Never heard	53.77%	51.54%	51.79%	55.90%
Previously heard	36.01%	38.00%	37.07%	33.90%
Familiar	10.22%	10.46%	11.14%	10.20%
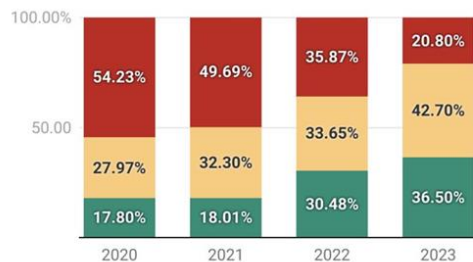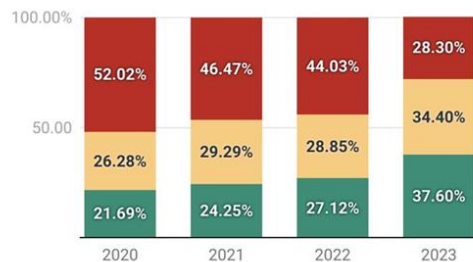
Which one of these circumstances would motivate you the most to share your data?

■ Citation of my research papers   ■ Funder requirement

	Citation of my research papers	Funder requirement
United States	9.98%	7.84%
Ethiopia	21.14%	1.71%
Japan	34.43%	0.66%
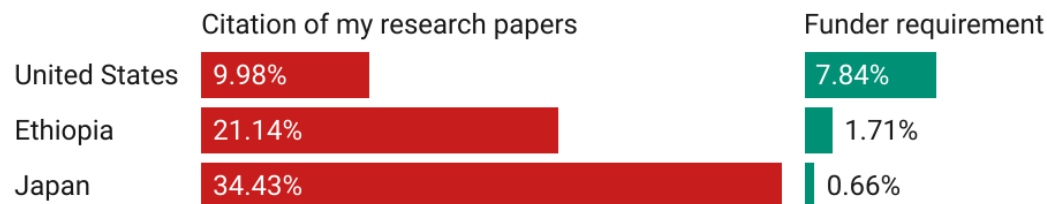
as a proportion of total responses to survey question

Created with Datawrapper

Links to datasets in Make Data Count Data Citation Corpus from countries with more than 50,000 publications per year

Highest percentage links to datasets in the "Make Data Count Data Citation Corpus" from countries with more than 1000 publications per year

Country specific focus

Key trends by country

United States

- Second largest volume of papers behind China. Both countries have a more significant number of researchers to reach with messages for new policies
- The US has a much lower rate of retractions for data related reasons than China
- The biggest percentage reduction in Data Availability Statements declaring that data is available 'on request'
- Policies around open research and open data have been evolving since the late 1990s and the National Institutes of Health (NIH) introduced their landmark Data Management and Sharing Policy in 2023
- Sustained policy developments seem to have had a direct impact on US-based researchers, with 'Funder requirement' a consistently leading motivating factor for data sharing

United Kingdom

- Clear policy frameworks from major research funders
- Data linking practice in published research originating in the UK is above average
- Data Availability Statements are less common in research papers with some form of international collaboration
- Highest awareness of the FAIR data principles (State of Open Data survey 2023)
- The UK's National Data Strategy underscores the importance of open data. Significant investment in key infrastructures that promote data sharing, interoperability and best practices in data stewardship

Germany

- Complex open academic data policy landscape, with a lack of one national open science policy and different approaches from institutions
- Above average of the top ten countries (by volume of publications) for linking to datasets in research papers originating there
- The Federal Government has set goals to strengthen open science and released a Research Data Action Plan that focuses on data sovereignty, data infrastructures and data based innovations

Ethiopia

- Significant growth in research activity and outputs, demonstrating rapid improvement in data sharing best practices
- From 2017-2022, research output has increased threefold to 11,000 publications. The percentage of publications with a DAS has increased from approximately 20% to 65%
- A high percentage of open data linked to papers compared to other countries, above average of the top ten countries (by volume of publications) for linking to datasets
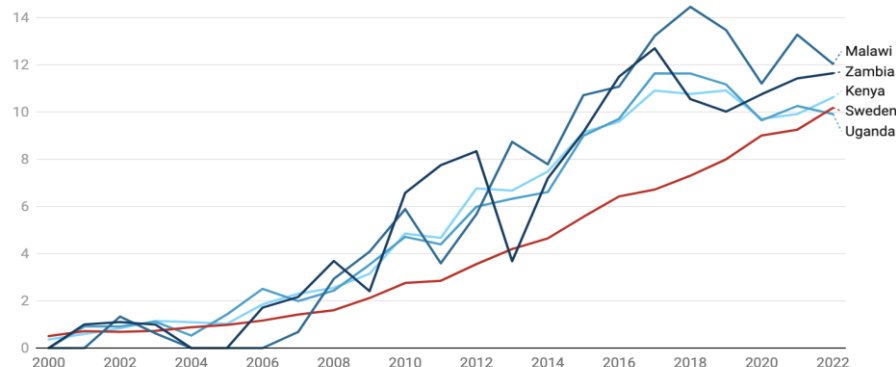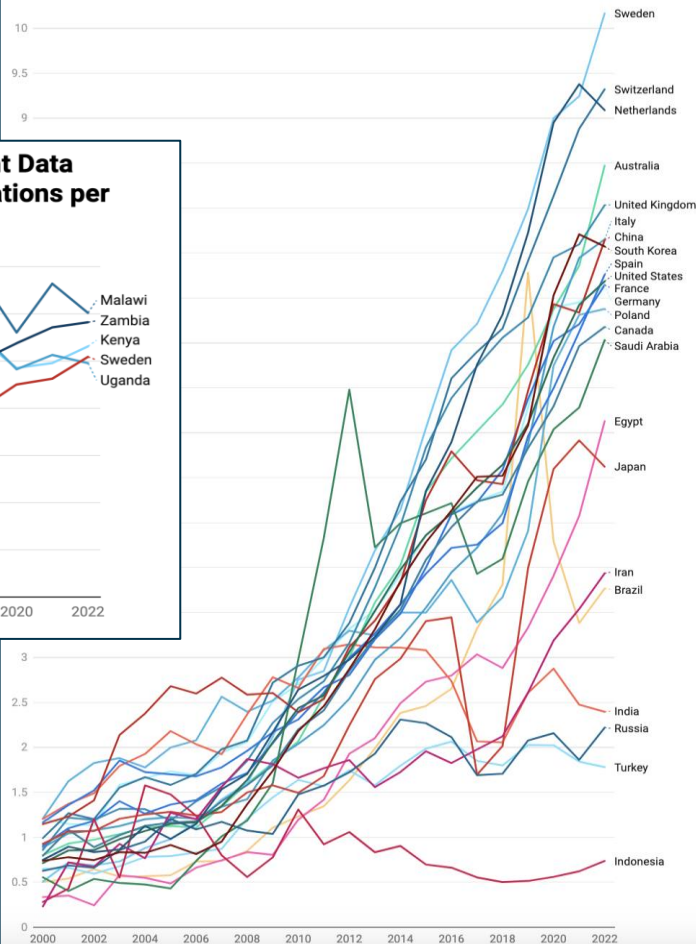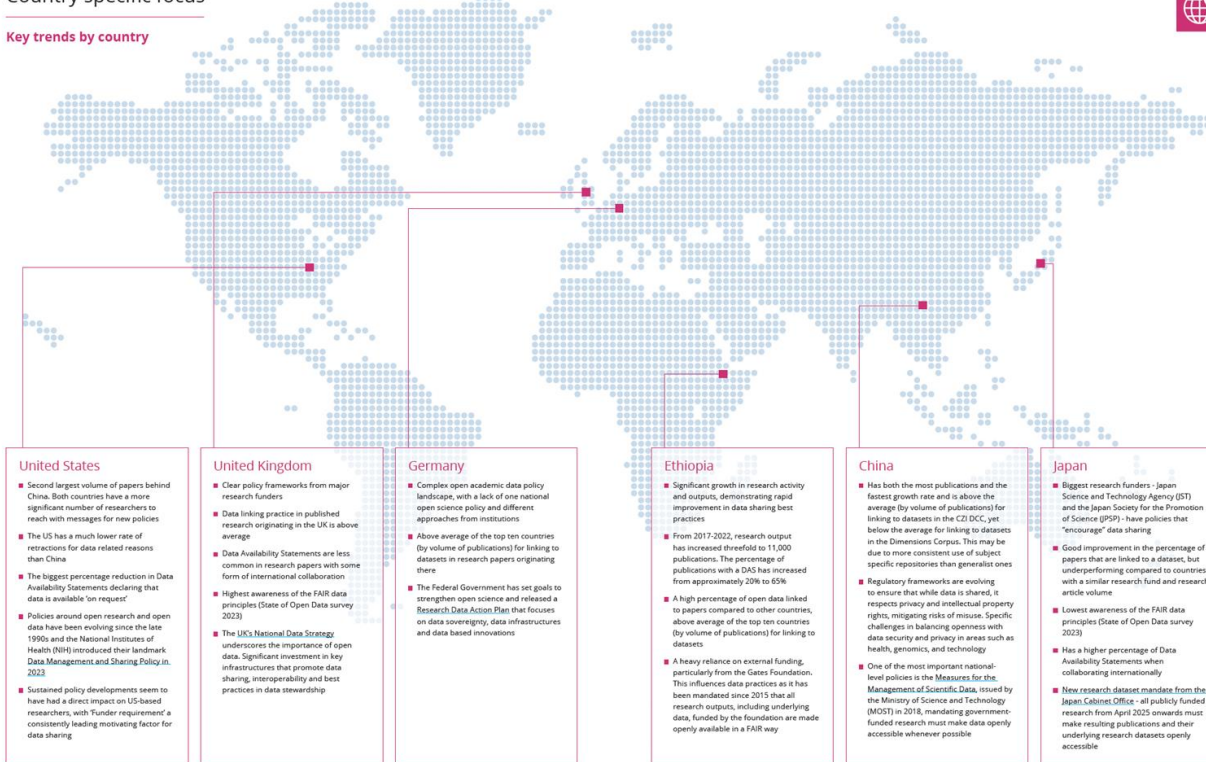- A heavy reliance on external funding, particularly from the Gates Foundation. This influences data practices as it has been mandated since 2015 that all research outputs, including underlying data, funded by the foundation are made openly available in a FAIR way

China

- Has both the most publications and the fastest growth rate and is above the average (by volume of publications) for linking to datasets in the CZI DCC, yet below the average for linking to datasets in the Dimensions Corpus. This may be due to more consistent use of subject specific repositories than generalist ones
- Regulatory frameworks are evolving to ensure that while data is shared, it respects privacy and intellectual property rights, mitigating risks of misuse. Specific challenges in balancing openness with data security and privacy in areas such as health, genomics, and technology
- One of the most important national-level policies is the Measures for the Management of Scientific Data, issued by the Ministry of Science and Technology (MOST) in 2018, mandating government-funded research must make data openly accessible whenever possible

Japan

- Biggest research funders - Japan Science and Technology Agency (JST) and the Japan Society for the Promotion of Science (JPSP) - have policies that "encourage" data sharing
- Good improvement in the percentage of papers that are linked to a dataset, but underperforming compared to countries with a similar research fund and research article volume
- Lowest awareness of the FAIR data principles (State of Open Data survey 2023)
- Has a higher percentage of Data Availability Statements when collaborating internationally
- New research dataset mandate from the Japan Cabinet Office - all publicly funded research from April 2025 onwards must make resulting publications and their underlying research datasets openly accessible

figshare
A **Digital Science** Solution

SPRINGER NATURE

The State of Open Data 2024:

Special report

Bridging policy and

practice in data sharing

#StateOfOpenData