Open and Responsible Research 3
Being Open and Responsible at Home

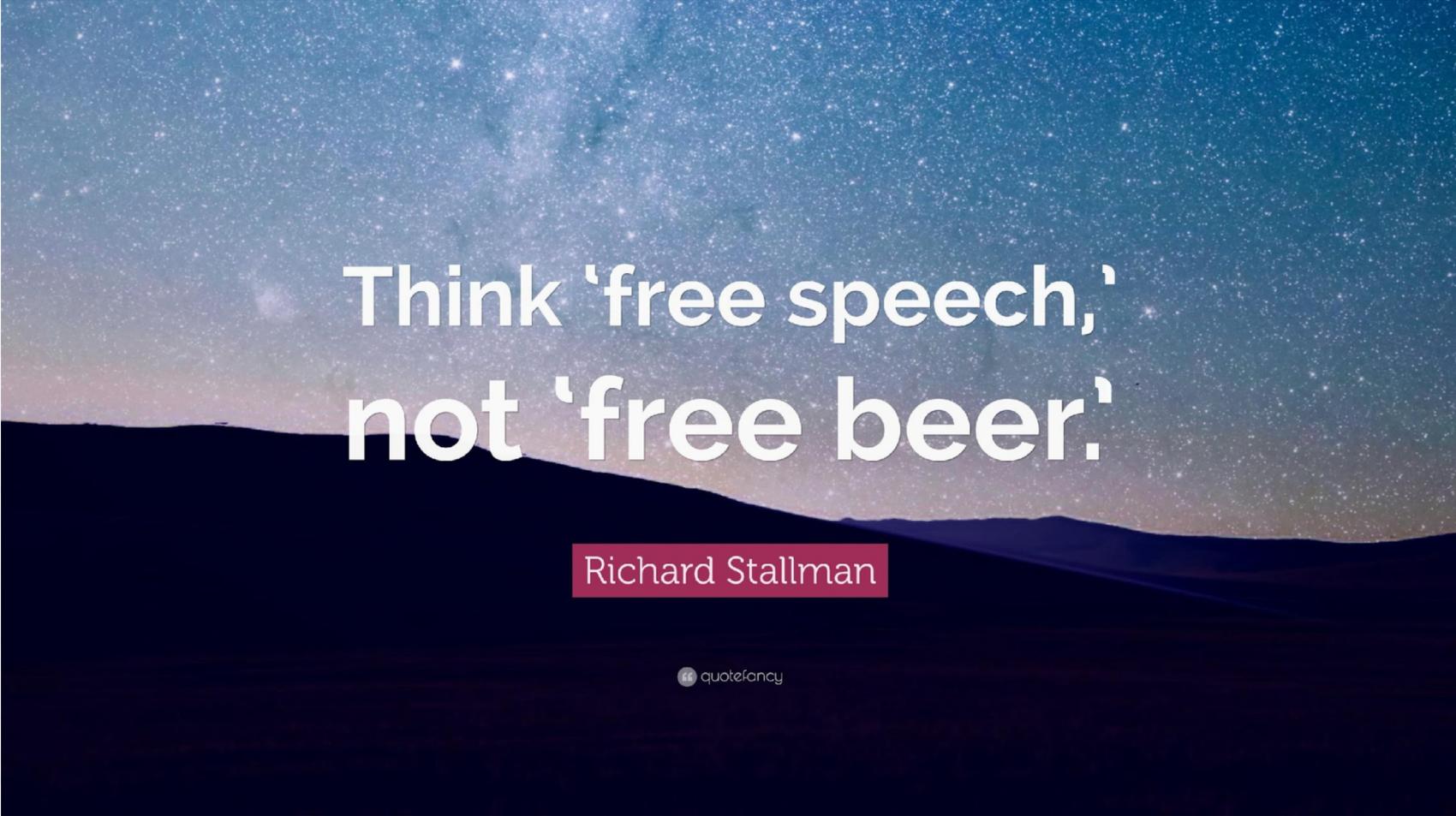Louise Bezuidenhout
Emanuele Ratti

CODATA-RDA Data Science School

Plan

- Personal concerns and implementing OS practices at home

  - What challenges do we have about implementing these in our research environments?

  - What kinds of assistance can we get?

Recap From Week 1

Recap From Monday

Think 'free speech,' not 'free beer.'
Richard Stallman

Open Science

- **The products of scientific research should be freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control**

  - Transparency in experimental methodology, observation, and collection of data

  - Public availability and reusability of scientific data

  - Public accessibility and transparency of scientific communication

 - Using web-based tools to facilitate scientific collaboration

    http://www.openscience.org/blog/?p=269]

5

Openness and Responsible Conduct of Research

Open Lab Books: Transparency in research practices
Sharing and openness: enhance transmission of values

Open Peer Review: Transparency in peer review leads to better dialogue and collegial behaviour

Open Access: Improves availability of research outputs
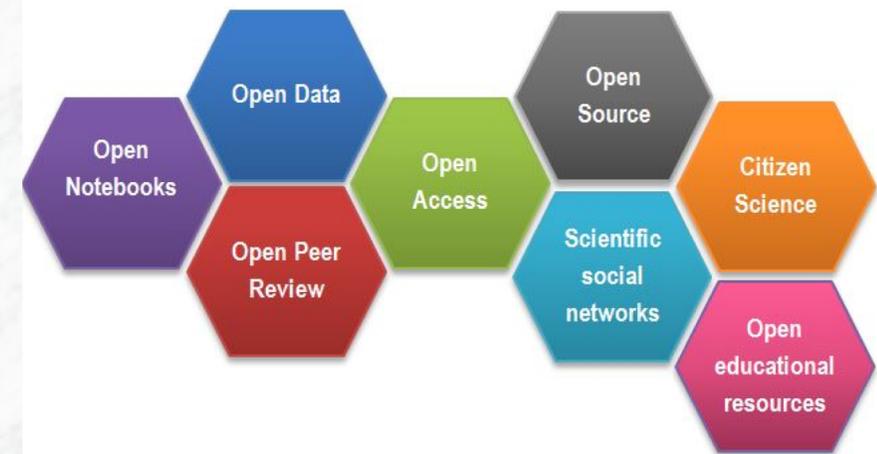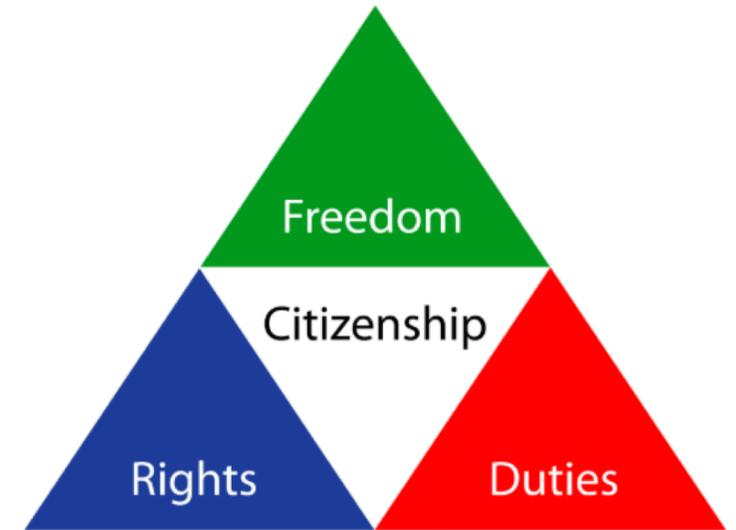*Open publishing:* leads to improved citations, credit and collaboration
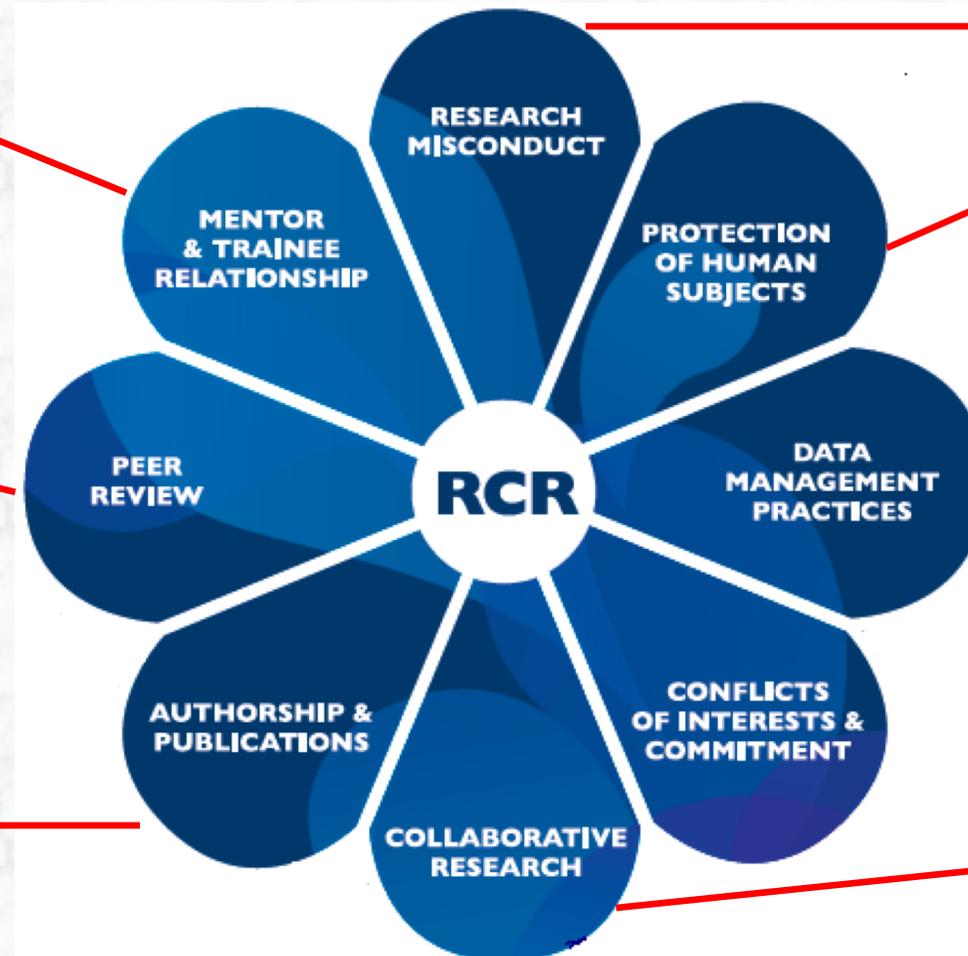
Open Data and Open Methodologies: Improve transparency and reproducibility of research

Open Science Tools: Improve collaboration

RESEARCH MISCONDUCT

MENTOR & TRAINEE RELATIONSHIP

PROTECTION OF HUMAN SUBJECTS

PEER REVIEW

RCR

DATA MANAGEMENT PRACTICES

AUTHORSHIP & PUBLICATIONS

CONFLICTS OF INTERESTS & COMMITMENT
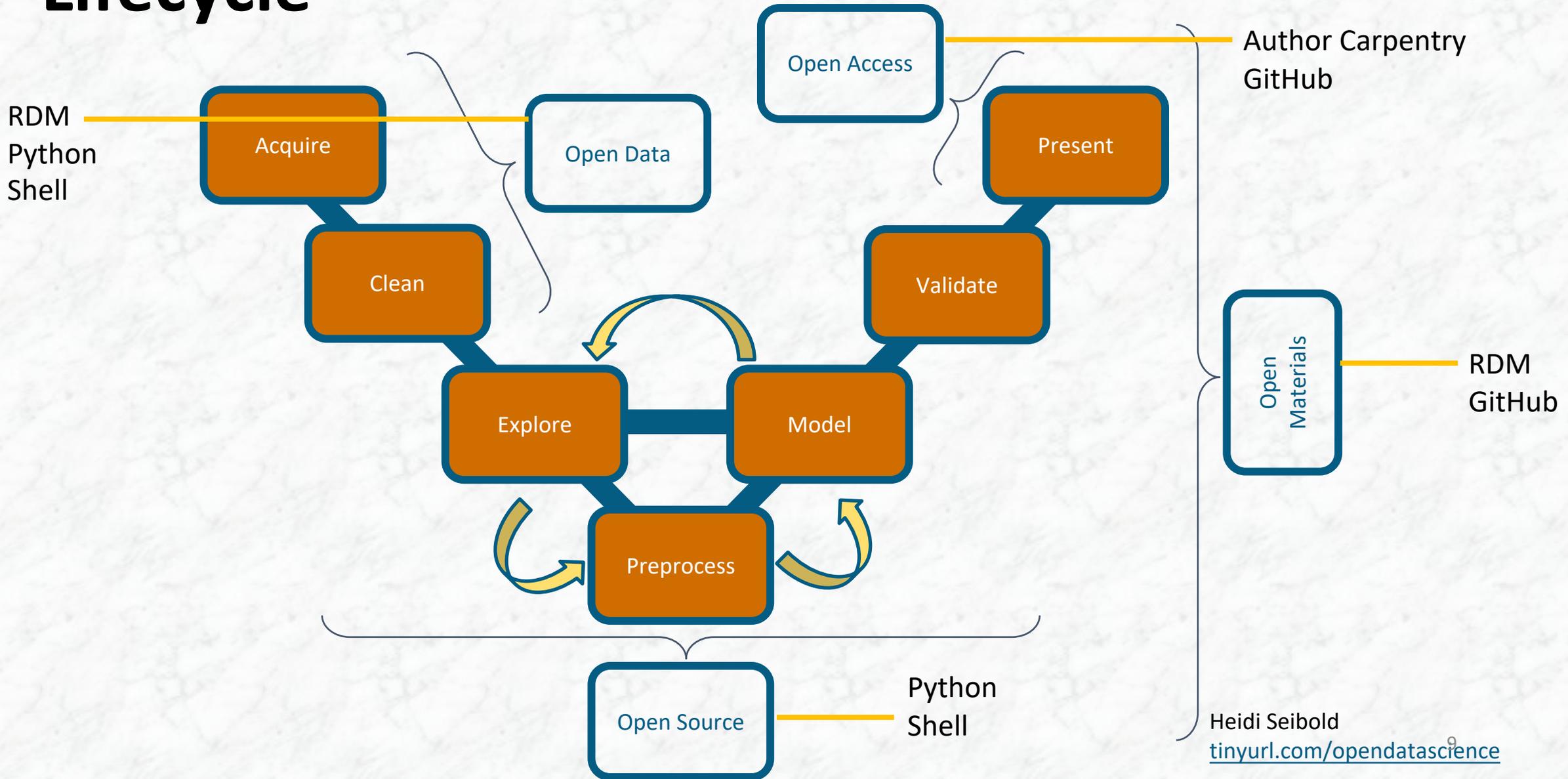
COLLABORATIVE RESEARCH

Open and Responsible Science Citizenship

- Consider ethics in daily activities

- Engage in discussions about the "big picture"

- Contribute, safeguard and curate community resources

- Contribute to community-building activities
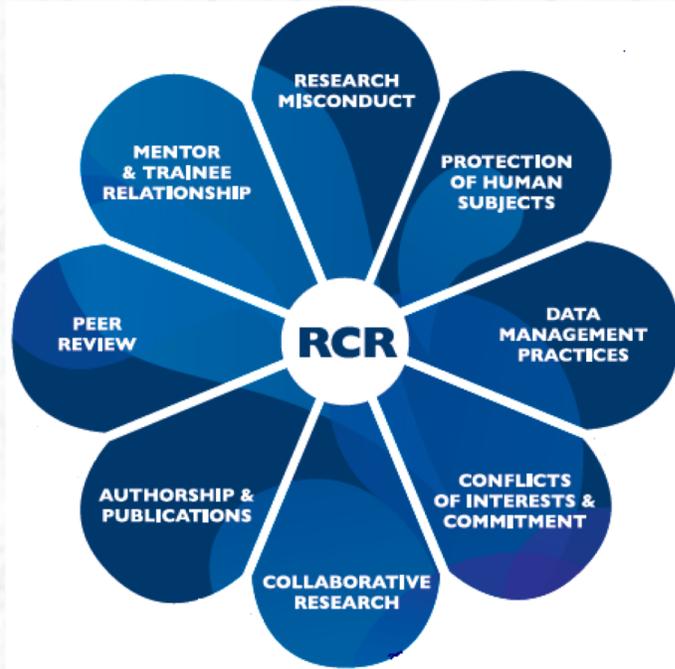
- Uphold and promote community values

https://www.fosteropenscience.eu/content/what-open-science-introduction

Practical Tools, Ethical Impact

Open Science Throughout The Research Lifecycle

RDM
Python
Shell

Acquire

Open Data

Clean

Open Access

Present

Validate

Explore

Model

Preprocess

Open Source

Python
Shell

Author Carpentry
GitHub

Open Materials

RDM
GitHub

Heidi Seibold
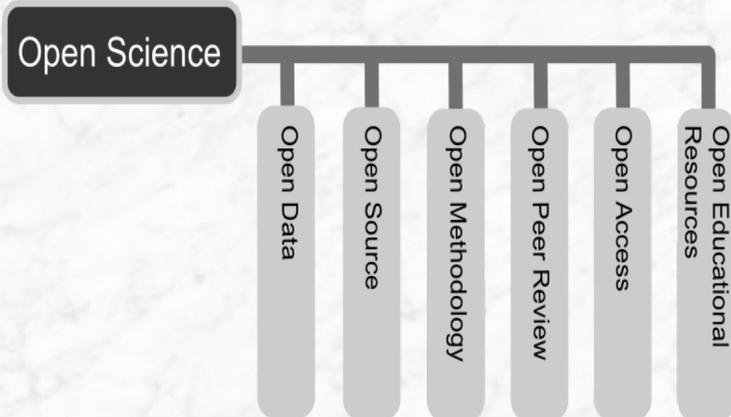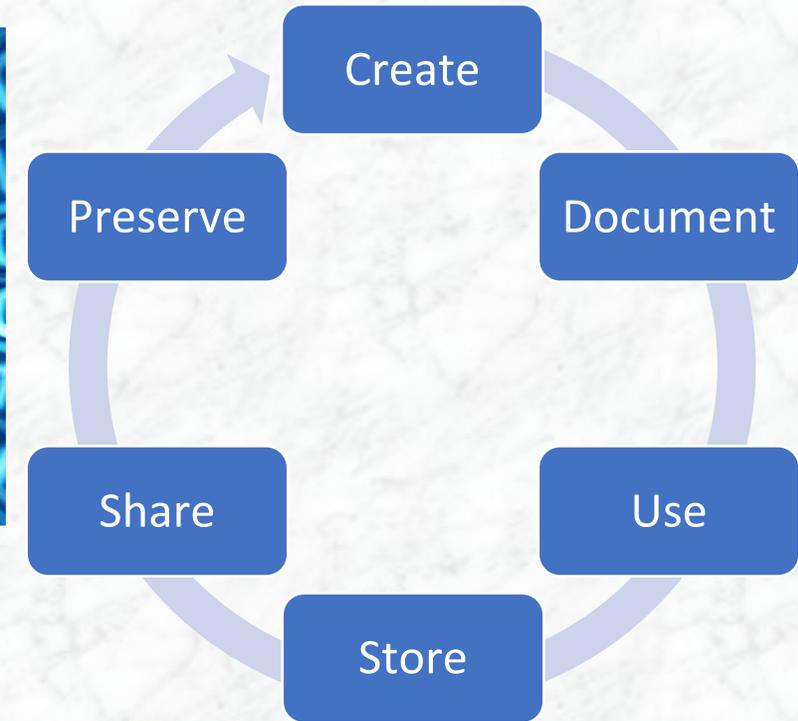tinyurl.com/opendatascience

9

How to be a responsible and open data scientist

Tools for responsible and open data science
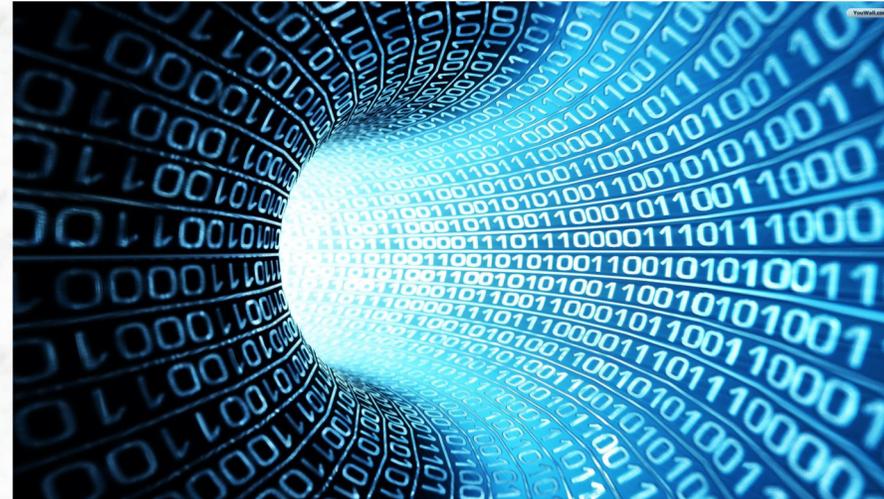
Doing responsible and open data science research

Everything Has Ethical Considerations

- Research practices and the tools that underpin them have **ethical implications**

- Being aware of their characteristics/specifications can help you avoid building in **ethical complications** into your work

  - Legal requirements – ie. GitHub, repository requirements

  - Marginalization of learners/users – using CLI instead of GUI

  - Re-use of data – sharing data, not respecting restrictions

- Even the smallest tools **can have** significant ethical consequences

Challenges of Being Open at Home

Implementing Open and Responsible Practices in Your Own Research

- Challenges are **common** – everyone has them

- Physical, social and regulatory contexts **influence perceptions** of Open Science and **ability to engage** in Open Science activities

SA1/3: I think it leads to better science

KY1/1: I won't release data unless I first of all publish

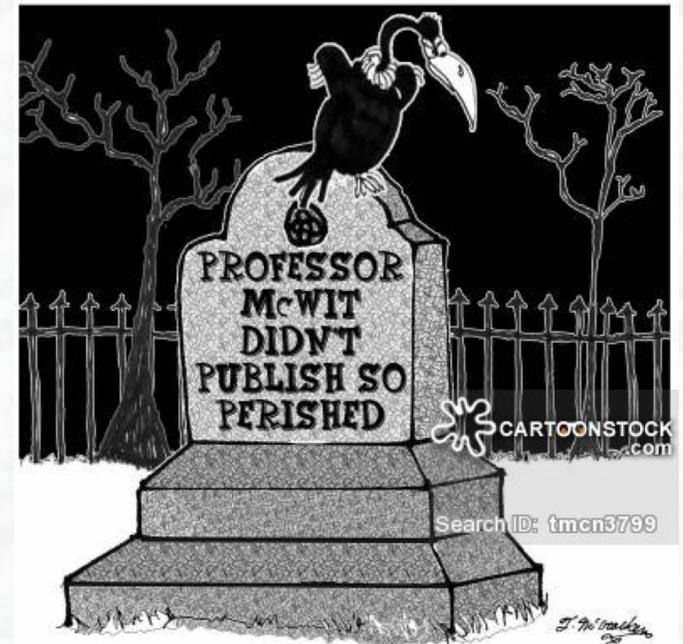| Institutional/cultural issues | Supervisor is not supportive of open access and data sharing

Institute may not allow to share data due insecurity and misuse

The Research Office at our university alerted us that due to lack of funding they are reconsidering allocation of $$ to open access publishing

Lack of policy on open science; institutional repository initiated by has since 2017 not been outdoored; there seems to be a lack of capacity of the library to champion open science. | Personal concerns | Data literacy Skills

What strategies can one use to promote the culture of open science in an Institution/ Country where research funding is poor

How open can you be with respect to your data and methods when trying to publish your findings

When you say about "Open Data" are we focusing on primary data or also secondary data?

if you don't have budget allocation to publish in open Science as journals have high APC

How open can you be with respect to your data and methods when trying to publish your findings?

How can we encourage postgraduate researchers with the intention of solving indigeneous problems using newly collected data (especially African-generated data) to follow explicitly open research conduct. To the best of my knowledge, newly gathered data are not encouraged when it comes to publishing your research findings?

Researcher ambiguity on data sharing and what level of data can be shared?

Which is better: collecting data involving human subjects virtually (e.g. Google Forms, Amazon Mechanical Turk) or in person (e.g. interviews, focus groups)?

Skills gap |
| Infrastructural/resource issues | Big data storage and permanency in the system (limited technical capacity of institutions)

Ethical permits indicating the data / material should be destroyed at the end of the study goes against reproducibility and cooperation

Data security and restriction (embargo) on data access

Many institutions especially in developing countries are lacking data storage facilities

I mentioned inability to get an IR initiative implemented; the fact is since the initiative was launched in 2017 till now, only a desktop computer was provided for the IR project, no server has been procured and the librarian assigned to manage the project has no clue what IR is all about. | | |

It's OK To Have Challenges

- Challenges can be categorized into a number of **different areas**:

  1. Cultural resistance and lack of institutional/peer support

  2. Resource limitations

  3. Personal concerns

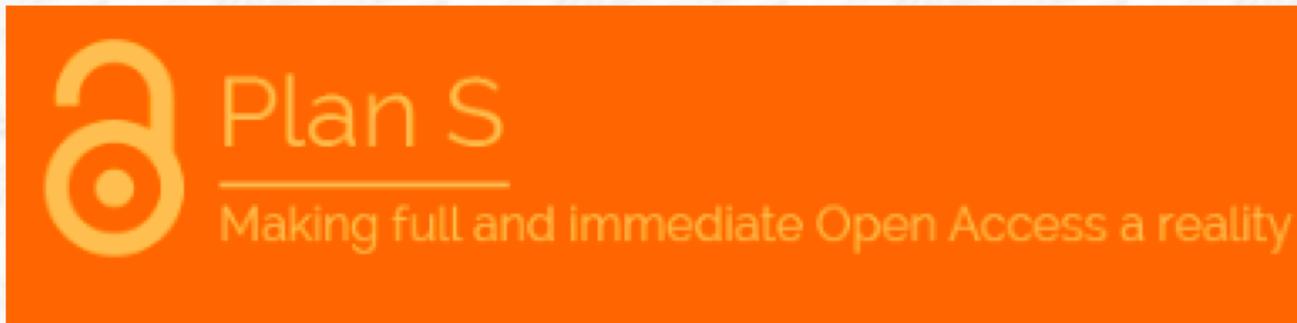- Challenges are **not insurmountable** – many resources can help address them
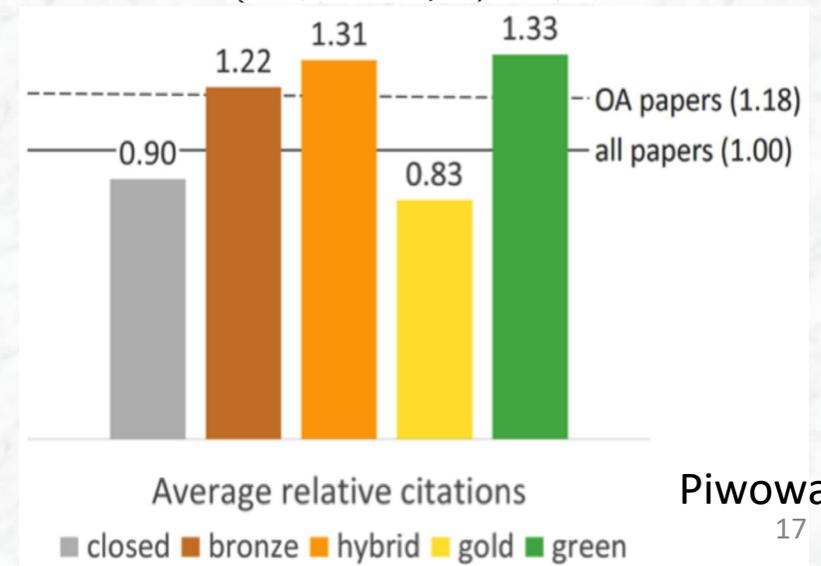
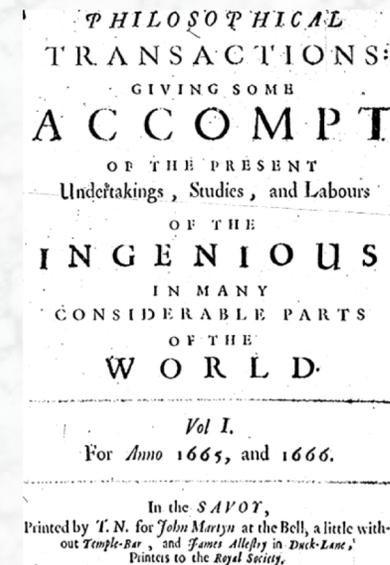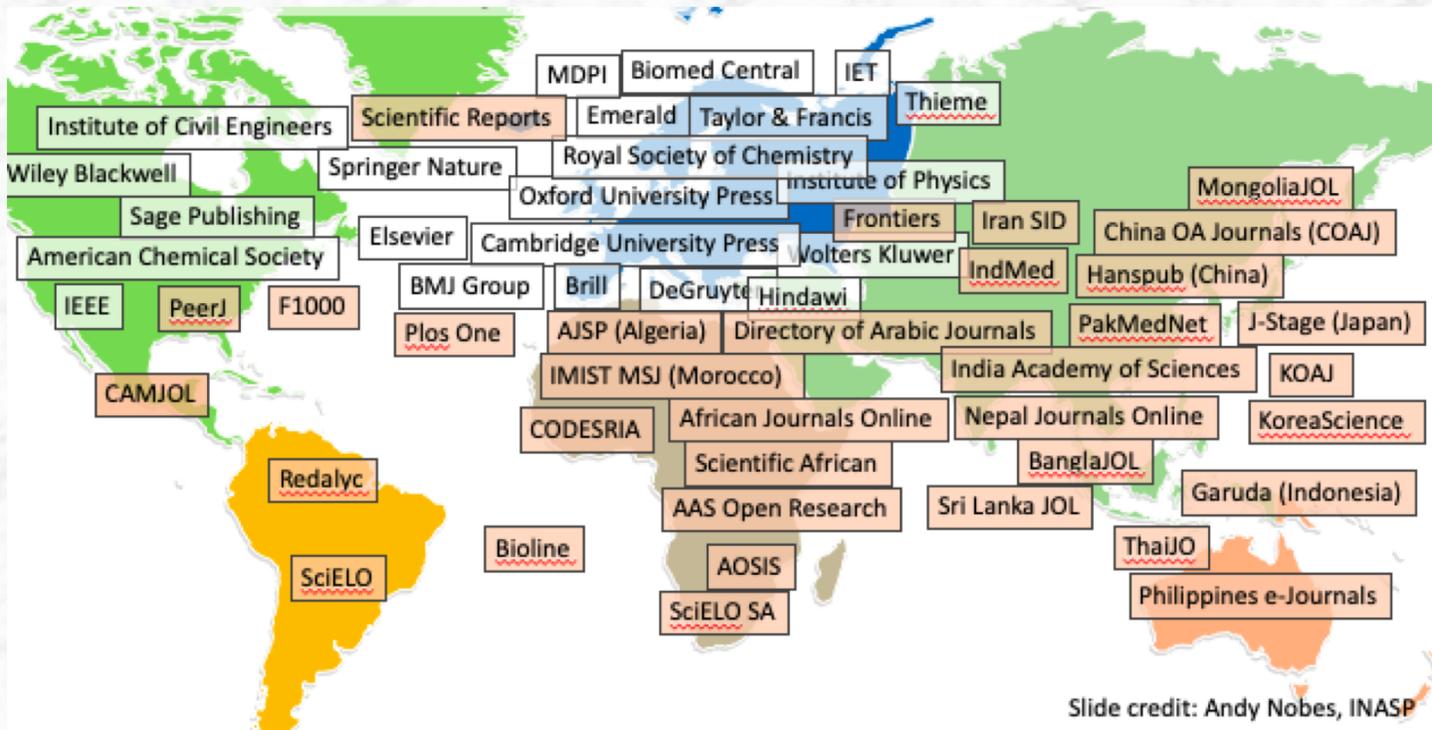# 1. Cultural Resistance and Lack of Support

- A quick background:

  - Inherited colonial academic systems

  - Historic lack of funding and resources limiting research scope

  - "Parachute research"

- Problems include

  - Lack of institutional support

  - Lack of regulations/guidance

  - Lack of trust

PROFESSOR McWIT DIDN'T PUBLISH SO PERISHED

CARTOONSTOCK.com

Search ID: tmcn3799

1a. Getting Your Institution On Board

Slide credit: Andy Nobes, INASP

PHILOSOPHICAL TRANSACTIONS: GIVING SOME ACCOMPT OF THE PRESENT Undertakings, Studies, and Labours OF THE INGENIOUS IN MANY CONSIDERABLE PARTS OF THE WORLD.

Vol I.
For Anno 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little without Temple-Bar, and James Allestry in Duck-Lane, Printers to the Royal Society.

Plan S
Making full and immediate Open Access a reality

| | 1.22 | 1.31 | | 1.33 |
OA papers (1.18)
all papers (1.00)
0.90
0.83
Average relative citations
closed | bronze | hybrid | gold | green

Piwowar 2018

1b. Getting Better Protection and Guidance
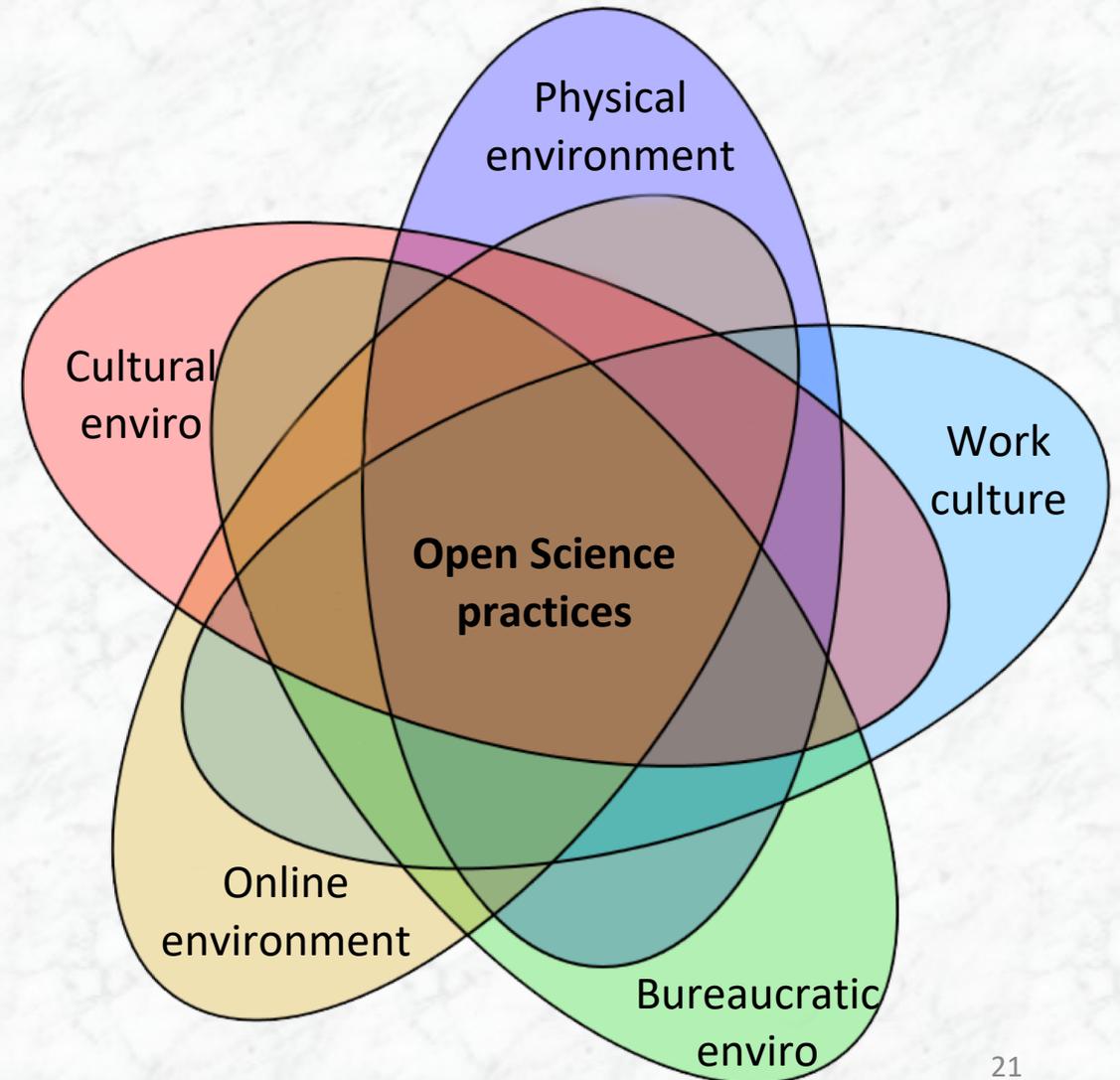
1c. Getting Over Issues of Trust

Still Needed: Positive Examples

- Need for **more positive examples** to dispel "urban myths" and lurking ghosts

- Need enthusiastic **champions and mentors**

- Effective personal networks

- What else can help foster open research cultures and maximize their support?
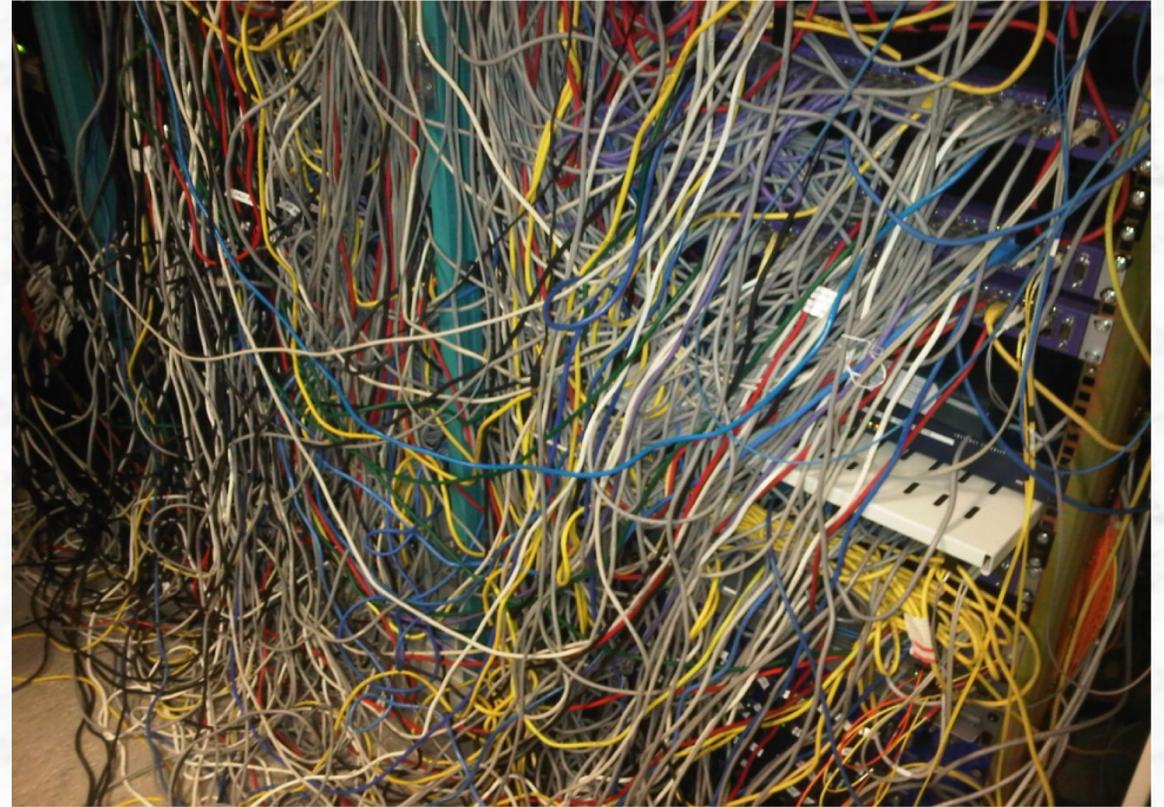
www.bbc.com

2. Infrastructures that Support Openness

- Many **institutions struggle** with legacies of low-resourcing

- Strategic resource distribution often means that OS activities are **under-funded**
  - Lack of finances to fund Open Science practices
  - Lack of ICT infrastructures
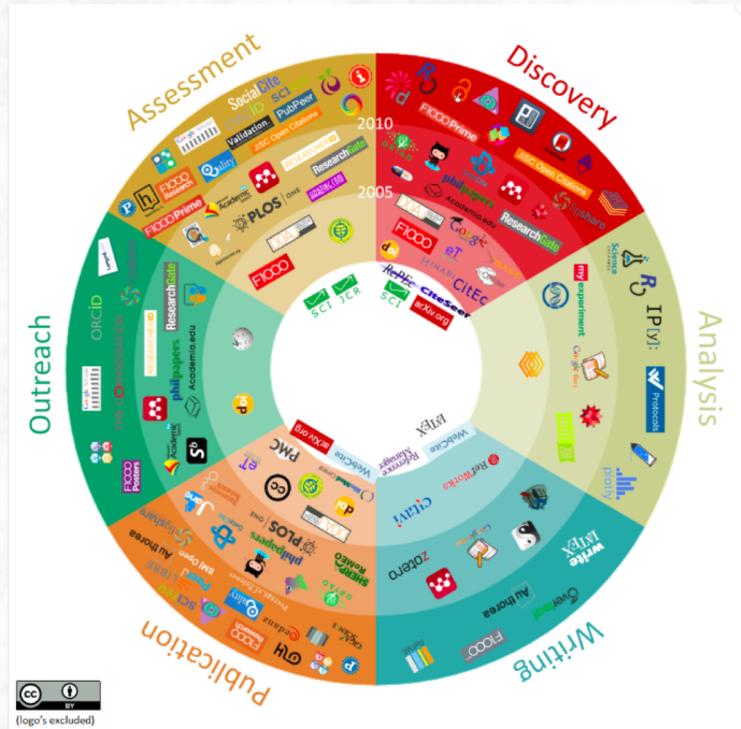 - Lack of technical support
  - Lack of guidance

Physical environment

Cultural enviro

Work culture

Open Science practices

Online environment

Bureaucratic enviro

Just Because the Resources Are Online ...

…. doesn't mean they're accessible

2a. Lack of Resources

DOAJ

SEARCH ⌄   DOCUMENTATION ⌄

THE DIRECTORY OF OPEN ACCESS JOURNALS

## Find open access journals & articles.

APC waivers and discounts

BMC offers waivers and discounts for article processing charges (APCs) for papers whose corresponding authors are based in low-income countries.

BMC offers APC waivers to papers whose corresponding authors are based in countries classified by the World Bank as low-income economics as of July 2017.

zenodo

You can make your workflow more open by ...
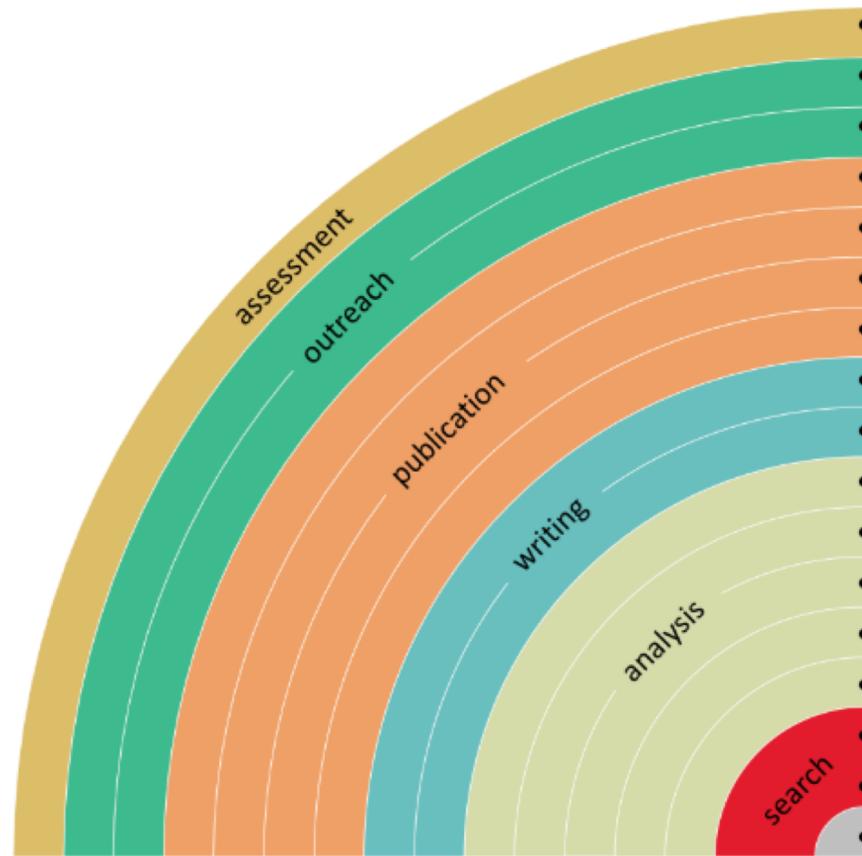
adding alternative evaluation, e.g. with altmetrics
communicating through social media, e.g. Twitter
sharing posters & presentations, e.g. at FigShare
using open licenses, e.g. CC0 or CC-BY
publishing open access, 'green' or 'gold'
using open peer review, e.g. at journals or PubPeer
sharing preprints, e.g. at OSF, arXiv or bioRxiv
using actionable formats, e.g. with Jupyter or CoCalc
open XML-drafting, e.g. at Overleaf or Authorea
sharing protocols & workfl., e.g. at Protocols.io
sharing notebooks, e.g. at OpenNotebookScience
sharing code, e.g. at GitHub with GNU/MIT license
sharing data, e.g. at Dryad, Zenodo or Dataverse
pre-registering, e.g. at OSF or AsPredicted
commenting openly, e.g. with Hypothes.is
using shared reference libraries, e.g. with Zotero
sharing (grant) proposals, e.g. at RIO

Knowing Where to Look

Research Databases and data sources

There is a wealth of research data in various databases around the world – much of it publicly available. Here are a few examples of where to look:

- Global Partnership for Sustainable Development Data www.data4sdgs.org/
- Flowminder: http://www.flowminder.org/
- Worldpop: http://www.worldpop.org.uk/
- University of Connecticut Research Database Locator: http://rdl.lib.uconn.edu/byTitle.php
- Listing of Open Access Databases (LOADB): http://www.loadb.org/
- Research4Life programme:
    - **AGORA** - Access to Global Online Research in Agriculture http://www.fao.org/agora/en/
    - **HINARI** - Access to Research for Health programme http://www.who.int/hinari/en/
    - **OARE** - Online Access to Research in the Environment http://web.unep.org/oare/
    - **ARDI** - Access to Research for Development and Innovation http://www.wipo.int/ardi/en/

African databases:
- OpenAFRICA: https://africaopendata.org/
- African Development Bank Statistical Data Portal http://dataportal.opendataforafrica.org/
- Directory of Data Repositories in Africa (DODRIA) https://researchdatadirectoryonafrica.com/
- FAO Agricultural databases http://www.fao.org/statistics/databases/en/

Offline databases:
- TEEAL (The Essential Electronic Agricultural Library) https://teeal.org/
- eGranary Digital Library https://www.widernet.org/eGranary/
- **Wiki Project Med Foundation** http://medbox.iiab.me/home/
- See also the Wikipedia list of academic databases and search engines

Thanks to Andy Nobes, INASP

2b. Lack of Expertise and Training

OPEN SCIENCE MOOC — FREE | OPEN | LEARNING

THE CARPENTRIES

CODATA-RDA Research Data Science Schools

R Studio Community

Forums

If you're looking for a forum in your native language, please check out the local user groups page at the Python Wiki.

- Python Forum (English)
- Python-Forum.de (German)
- /r/learnpython (English)

Support Networks

Academic support networks - organisations and NGOs

There are many international organisations and NGOs providing support to academics, ranging from free resources and access, training, Networking and subject-specific advice. Some useful organisations are listed below

AuthorAID www.authoraid.info

Eifl (Electronic Information for Libraries) www.eifl.net

Equator Network www.equator-network.org

CoDATA (Committee on Data of the International Council for Science) www.codata.org

Global Health Network https://tghn.org/

Global Young Academy https://globalyoungacademy.net/

Healthcare Information for All www.hifa.org

INASP www.inasp.info

Mendeley network https://www.mendeley.com/research-network/community

MedicineAfrica http://medicineafrica.com/

OWSD (Organisation for Women in Science in the Developing World) www.owsd.net

Scholars at Risk Network https://www.scholarsatrisk.org/

ResearchGate https://www.researchgate.net/

Research4Life http://www.research4life.org/

TWAS (The World Academy of Sciences for the advancement of science in developing countries) https://twas.org/

Indepth Network http://www.indepth-network.org/

International Health Policies http://www.internationalhealthpolicies.org/

Wessex Global Health Network http://www.wessexghnetwork.org.uk/

Thanks to Andy Nobes, INASP

27

Still Needed

- Local investment in Open Science infrastructures may **take time**

- However, **global infrastructures** and practices are **changing rapidly**

- Need more **LMIC voices** in these discussions to make sure that they work for researchers in lower-resourced contexts

What else can help researchers overcome resource limitations?

3. Personal Concerns

- As researchers we have **concerns** about implementing some Open Science practices

- These concerns are **legitimate**, and often relate to **cultural and regulatory challenges**

  - Concerns about being scooped

  - Concerns about scrutiny of data and methods

  - Misuse of data

  - Unintended harms

REASONS WHY RESEARCHERS ARE HESITANT TO SHARE THEIR DATA

|	
42%	Intellectual property or confidentiality issues
36%	My funder/institution does not require data sharing
26%	I am concerned that my research will be scooped
26%	I am concerned about misinterpretation or misuse
23%	Ethical concerns
22%	I am concerned about being given proper citation credit or attribution
21%	I did not know where to share my data
20%	Insufficient time and/or resources
16%	I did not know how to share my data
12%	I don't think it is my responsibility
12%	I did not consider the data to be relevant
11%	Lack of funding
7%	Other

3a. Knowing Your Rights/Responsibilities

3b. Openness as a Continuum

"Your primary collaborator is yourself 6 months from now, and your past self doesn't answer emails" (Russ Poldrack)
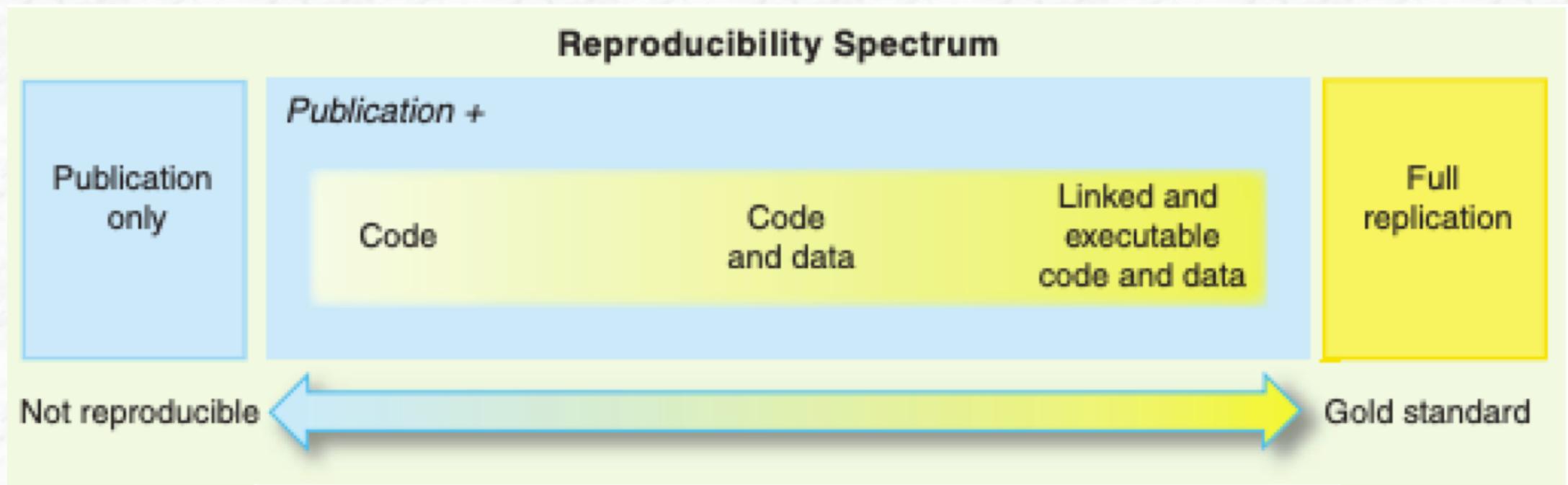
3c. Managing Risk

- Unintended harms are an **unavoidable** element of research

- Using **trusted infrastructures** can offset some concern as they set requirements on users and contributors

- Discuss concerns with peers – often they will have good advice

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Still Needed: More Evidence

- Lack of evidence of LMIC concerns

- Tendency to treat LMICs concerns as **"same as HIC but more"**

- Need **more evidence** about what is working, what is preferred and what is still needed

- Creating, joining and interlinking **networks of support** is key to fostering Open Science

What else can help researchers overcome concerns about being open?
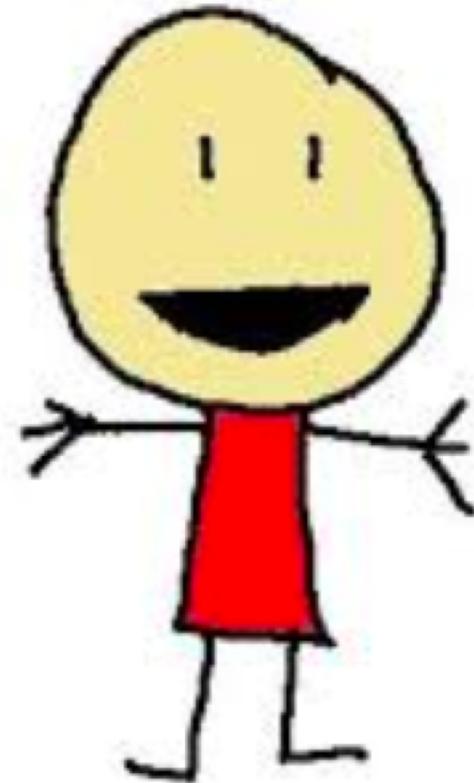
Openness Is A Lifelong Journey

Publish Preprints

FAIRify data

Make code available

Publish Lab-Notebooks

Use version control

Preregister your project

Do science communication

Creating Your Own List of Resources

- Use the document of concerns developed in week 1
- Start adding resources identified in the presentation
- We will discuss this more in the live sessions