



The role of Data Stewards in realising Open Science objectives

Dr. ir. Shalini Kurapati
22 Nov 2022, *Uni Pisa*

About



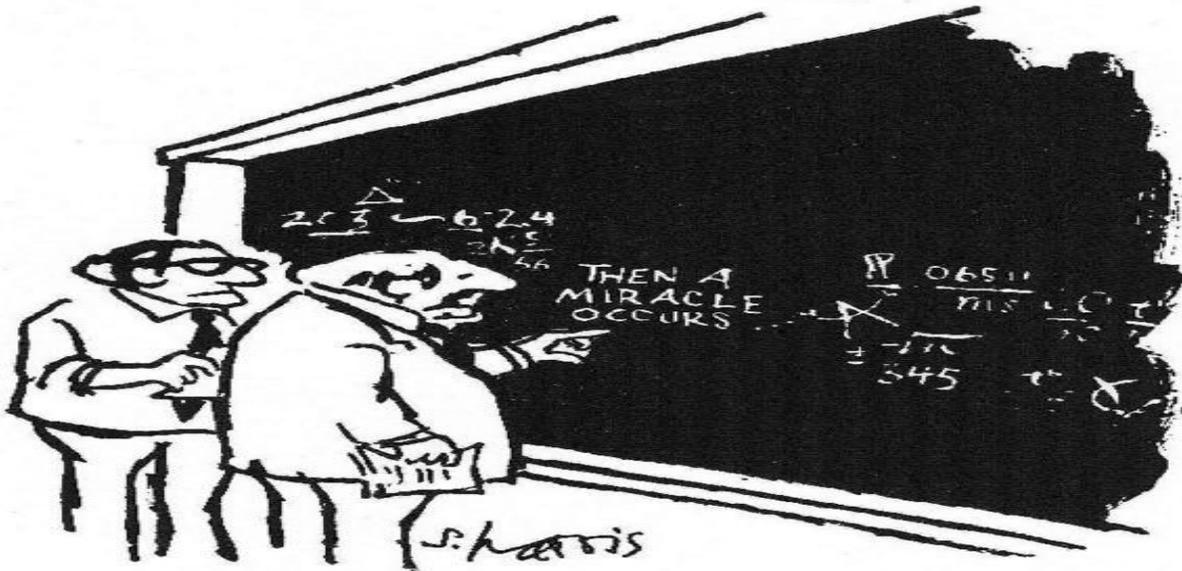
Dr. ir. Shalini Kurapati



<https://orcid.org/0000-0003-1492-4100>

- Co-founder and CEO clearbox.ai
- Past: RDM at TU Delft, advised researchers on data matters
- Open Science and RDM fellow at PoliTo
- 9 years of research experience at the intersection of technology, policy and management
- Certified Information Privacy Professional/Europe

Open Science



"I think you should be more explicit here in step two."

Open Science means:

- Evidence based results,
- Transparency, reproducibility, research rigour
- Validation and verification
- Dissemination and access
- And all other things that basically define science.

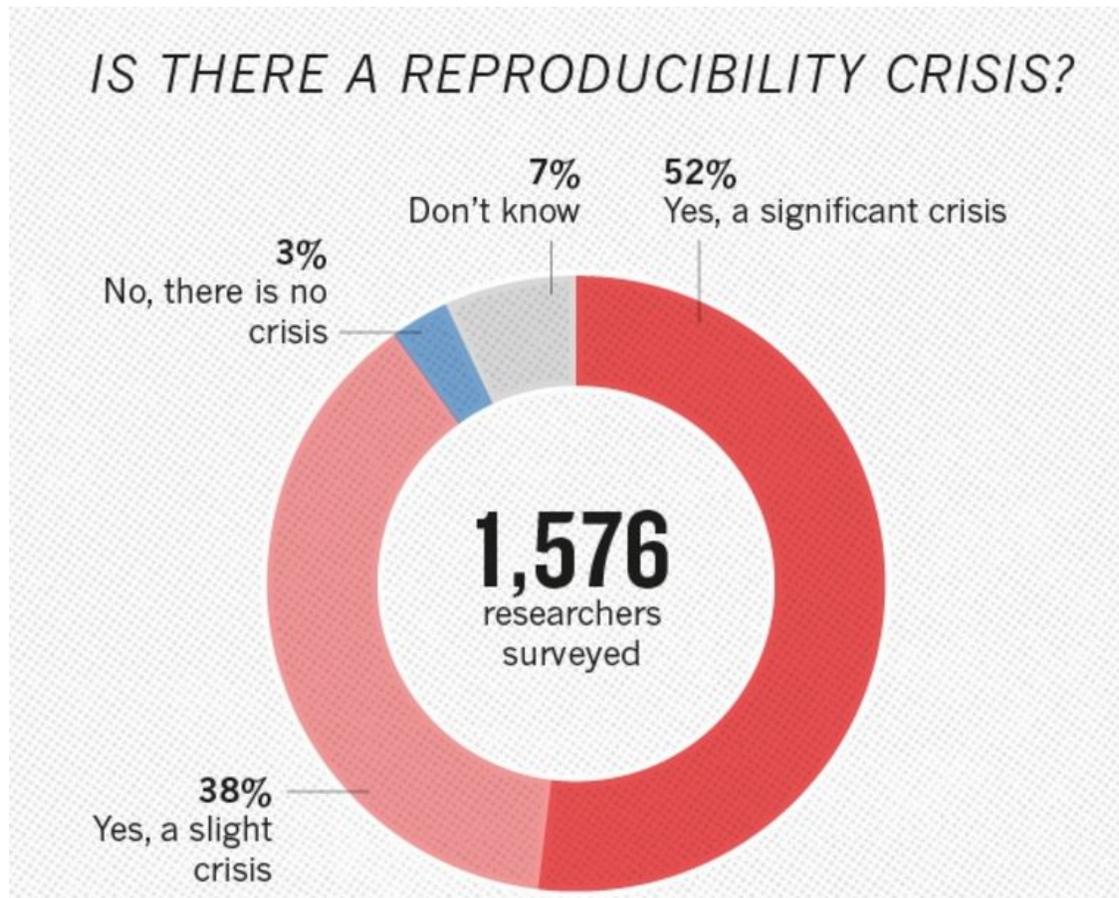


It's nothing new

**OPEN SCIENCE:
JUST
SCIENCE
DONE RIGHT**

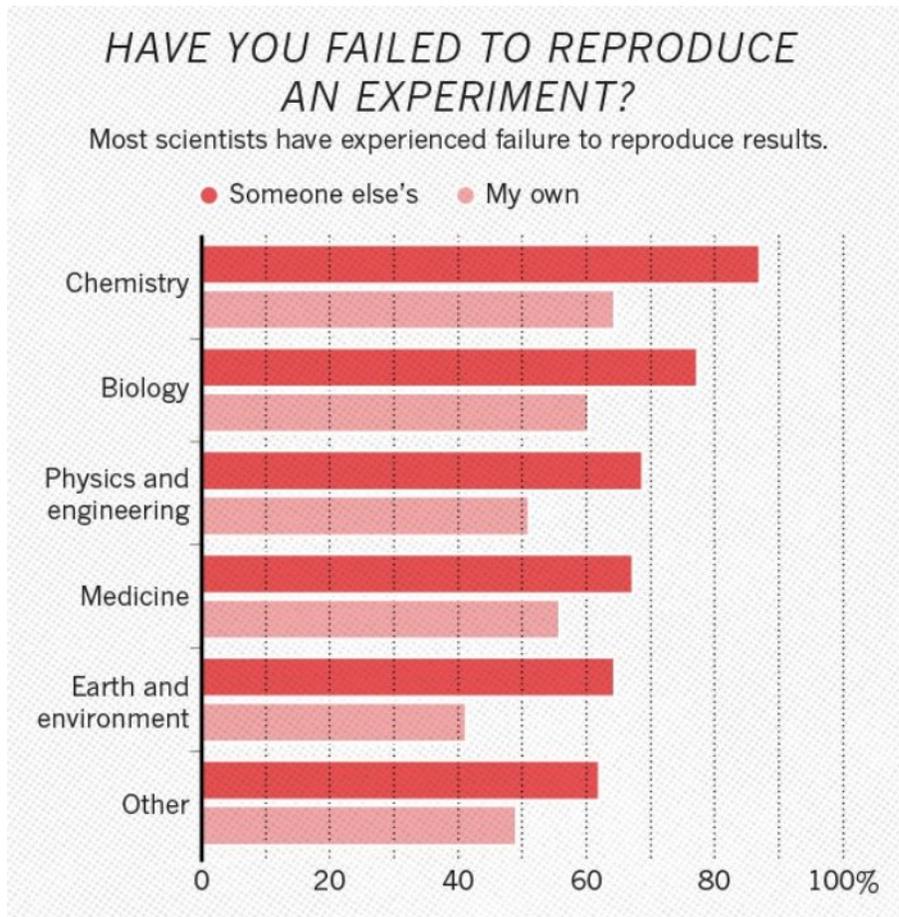
<https://zenodo.org/record/1285575>

The reproducibility crisis



<https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970>

In the engineering and physics disciplines



About 70% cannot reproduce others' experiments

and more than 50% cannot reproduce their own experiments!

Nature 533, 452–454 doi:10.1038/533452a

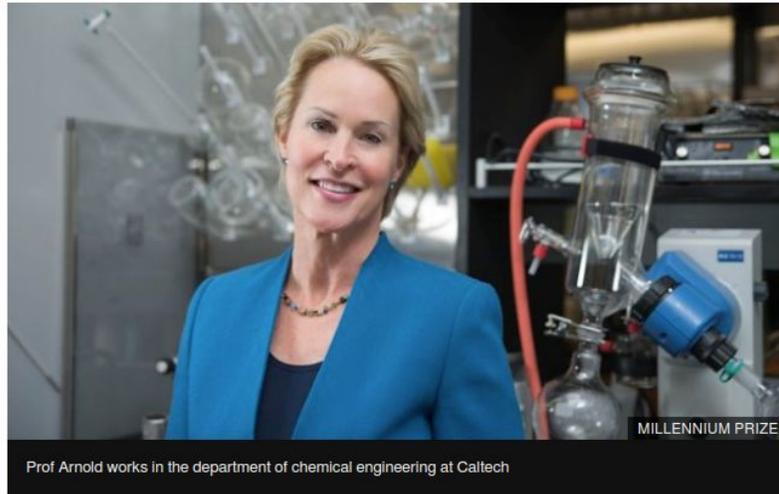
Happens even to the best of scientists

Nobel Prize-winning scientist Frances Arnold retracts paper

3 January 2020

f     Share

Nobel Prize



“It has been retracted because the results were not reproducible, and the authors found data missing from a lab notebook.

Reproduction is an essential part of validating scientific experiments. If an experiment is a success, one would expect to get the same results every time it was conducted.”

<https://www.bbc.com/news/world-us-canada-50989423>

Reasons for the crisis



- Selective reporting
- Pressure to publish
- Insufficient supervision and training
- **Supporting data / methods / code not available**

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, [online] 533(7604), pp.452-454. Available at: <https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970> [Accessed 26 Apr. 2018].

You see the problem?

19th century
scientist

I must find the
explanation for this
phenomenon in order
to truly understand
Nature...



21st centurt
~~scientist~~
academic

I must get the
result that fits my
narrative so I can
get my paper into
Nature..



facebook.com/pedromics

<https://www.euroscientist.com/open-scientists-in-the-shoes-of-frustrated-academics-part-i-open-minded-scepticism/>

Funding bodies are pushing for open science, focus on OA publications and FAIR data



More are following, nationally and regionally

<https://www.ukri.org/funding/information-for-award-holders/data-policy/common-principles-on-data-policy/>

RESEARCH DATA - OPEN BY DEFAULT



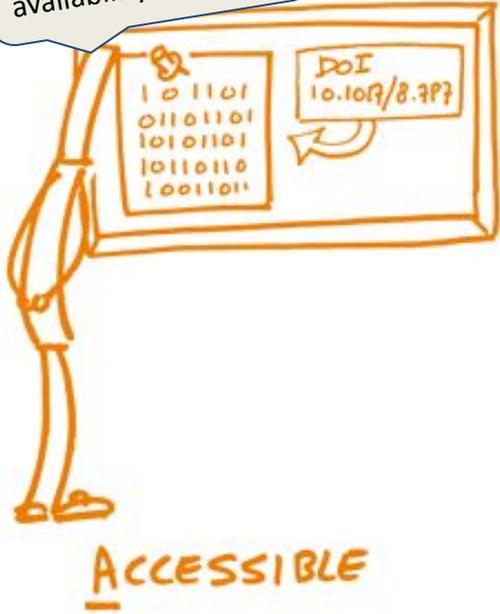
HORIZON 2020 GRANTEEES ARE REQUIRED

FAIR DATA PRINCIPLES

Metadata, persistent identifier (doi), indexed



Certified repository, doi, citable, record availability



Standards for data exchange



Terms of use, licenses, access control





Scriberia 

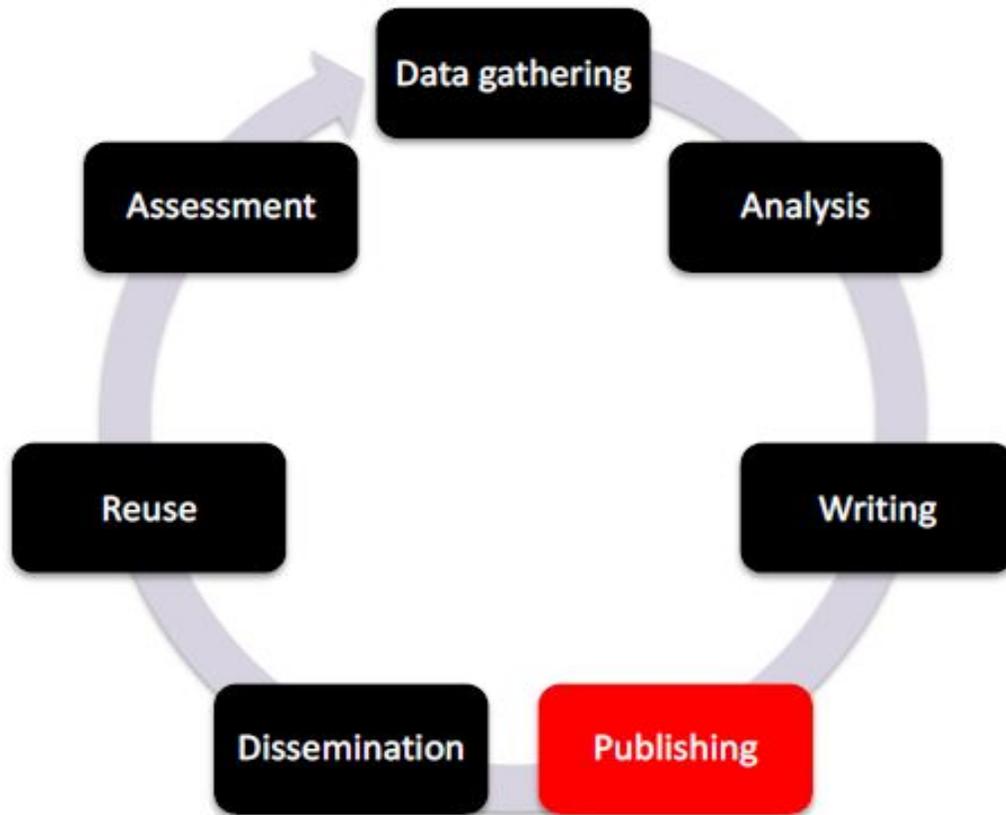
It's highly discipline specific.. FAIR is not same for all





It's a lot of work!! And new- not just researchers

<https://www.openaire.eu/how-to-comply-to-h2020-mandates-rdm-costs>



The only thing that counts in academia is publication of novel results in high impact journals

<https://www.repository.cam.ac.uk/handle/1810/276106>

Most importantly



Culture shift.. engaging with researchers.. showing the practical ropes

In any context, RDM is key!



We cannot achieve open science or FAIR data without RDM.

It is not simple but crucial!

Solutions at the intersection of technology, culture and awareness and policy



Immediate reactions to “data” requirements:

- It would take me 5 years to find all my data!
- The PhD/postdoc who had the data left the lab
- Should we write down all protocols?
- Data management is a waste of time
- Nobody will understand my data
- People can just ask me for it when they need it



In my experience



- They are there for any data questions
- Help researchers who “don’t have data” and who “don’t have problems”
- First point of contact, general practitioners
- Every occasion is good!
- Change happens one person at a time

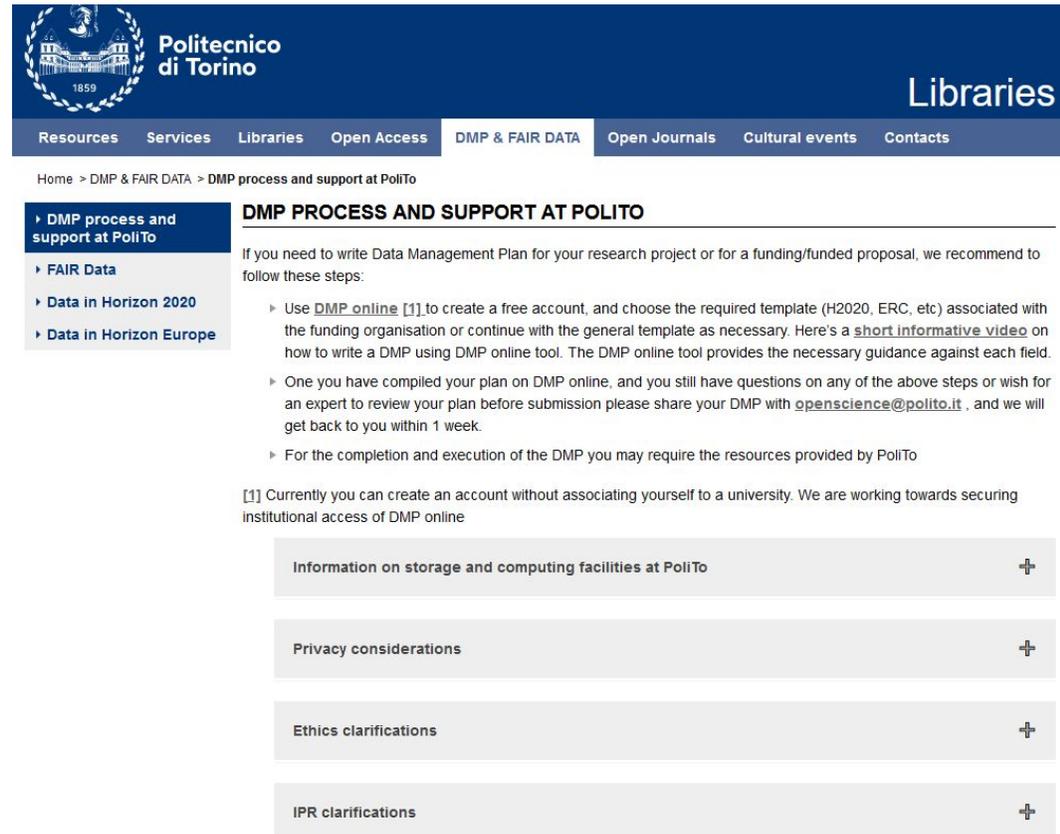
Objective: Improve daily practice not enforcement



TU Delft Experience



In POLITICO



The screenshot shows the website for Politecnico di Torino Libraries. The header includes the university logo and name, and a navigation menu with items like Resources, Services, Libraries, Open Access, DMP & FAIR DATA, Open Journals, Cultural events, and Contacts. The main content area is titled "DMP PROCESS AND SUPPORT AT POLITO" and provides instructions on how to create a Data Management Plan (DMP) using the DMP online tool. It includes a list of steps and a list of resources for storage, privacy, ethics, and IPR clarifications.

Politecnico di Torino Libraries

Resources Services Libraries Open Access **DMP & FAIR DATA** Open Journals Cultural events Contacts

Home > DMP & FAIR DATA > DMP process and support at PoliTo

DMP PROCESS AND SUPPORT AT POLITO

If you need to write Data Management Plan for your research project or for a funding/funded proposal, we recommend to follow these steps:

- ▶ Use [DMP online \[1\]](#) to create a free account, and choose the required template (H2020, ERC, etc) associated with the funding organisation or continue with the general template as necessary. Here's a [short informative video](#) on how to write a DMP using DMP online tool. The DMP online tool provides the necessary guidance against each field.
- ▶ One you have compiled your plan on DMP online, and you still have questions on any of the above steps or wish for an expert to review your plan before submission please share your DMP with openscience@polito.it, and we will get back to you within 1 week.
- ▶ For the completion and execution of the DMP you may require the resources provided by PoliTo

[1] Currently you can create an account without associating yourself to a university. We are working towards securing institutional access of DMP online

- Information on storage and computing facilities at PoliTo
- Privacy considerations
- Ethics clarifications
- IPR clarifications

https://www.biblio.polito.it/en/dmp_fair_data/dmp_process_and_support_at_polito2

HE Guidance

September 24, 2021

Other **Open Access**

Guidance on Open Science and Research Data Management in Horizon Europe proposals

 Shalini Kurapati;  Federica Cappelluti

This guidance document has been developed for POLITO researchers, support staff and anyone who is getting started with OS practices, to guide them in writing data and open science paragraphs and plans for Marie Curie actions and other Horizon Europe proposals.

Edited by Politecnico di Torino

Written by Shalini Kurapati, Federica Cappelluti

in collaboration with:

Research Support Department (ARI) - Sara Rollino, Francesca Spinazzola, Silvia Tabotta, Chiara Biglia

Library - Open Access Group - Anna Benenti, Donella Cioffi, Paola Ferrero, Maria Girard, Monica Margara, Anna Perotti, Sandra Sacco, Lorenzo Torriani, Ivana Viglione

Acknowledgements for scientific support and revisions: Elena Giglia (University of Turin)

Last update – 24 September 2021

911

 views

757

 downloads

[See more details...](#)

Indexed in

OpenAIRE

Publication date:

September 24, 2021

DOI:

DOI [10.5281/zenodo.5527043](https://doi.org/10.5281/zenodo.5527043)

License (for files):

[CC BY](#) Creative Commons Attribution 4.0 International

Data stewardship: Is new and evolving



Figure 2.5 ZonMw/ELIXIR data stewardship roles in the data stewardship landscape⁵⁸

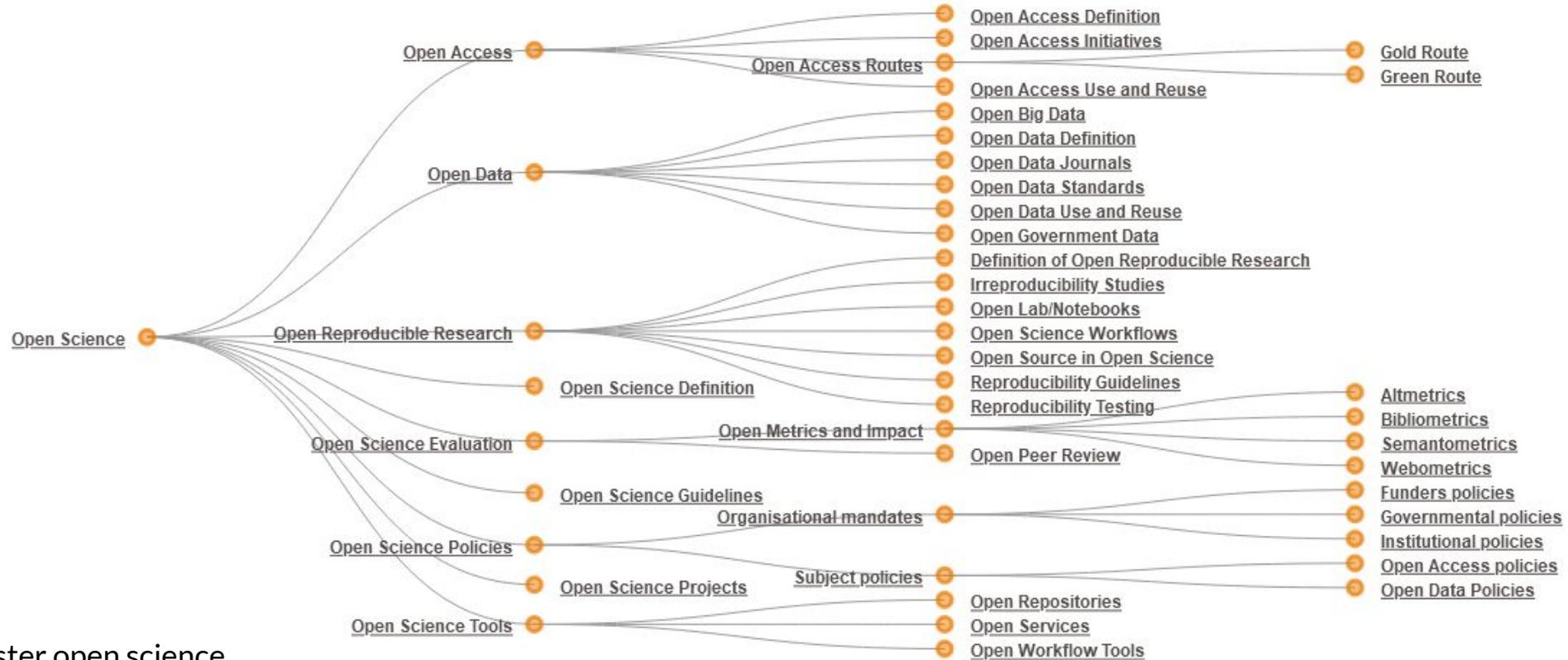


Study in Denmark
<https://zenodo.org/record/3609516#YpZm31RByUk>

TABLE 3. Titles given to the role of a DS.

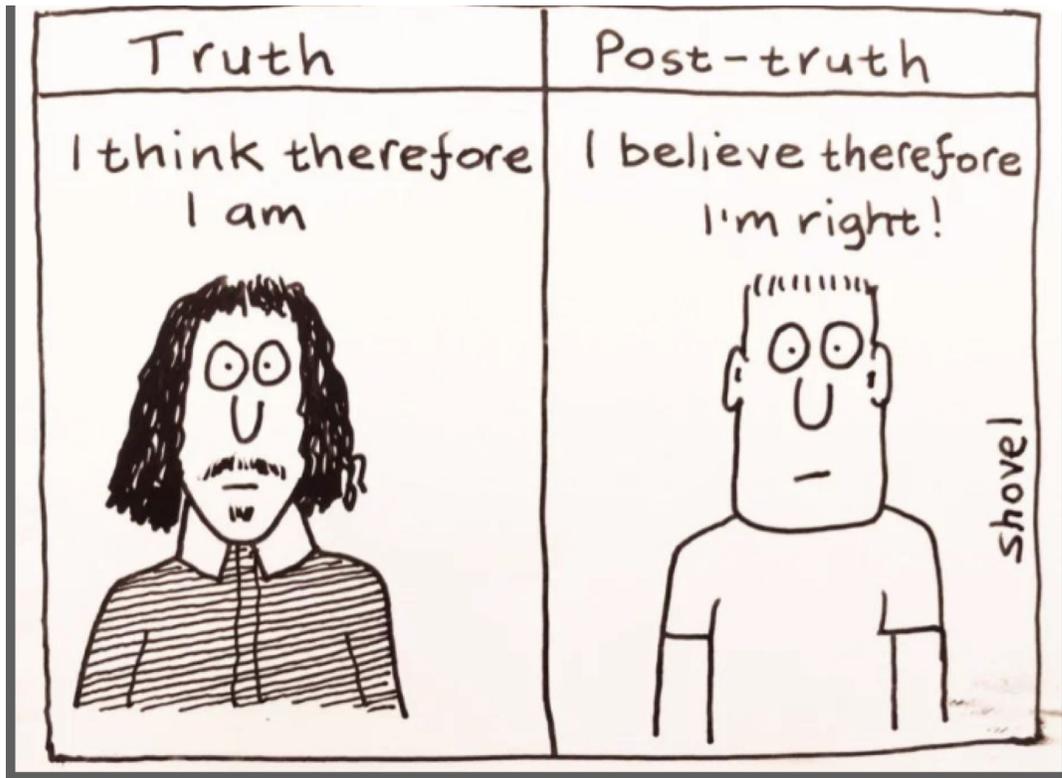
Preferred term	<i>n</i>
Don't know	37
Data manager	29
DPO	13
Data Steward	10
Data librarian	6
Information Management Specialist	4
Arkivar / Digitaliseringskonsulent	3
Data scientist	2
Dataansvarlig	2
IT specialists	2
Records manager	2
Consultant	2
Colleague	2
Administrative medarbejder - IT	2
Project managers / dept. coordinator	2
Business intelligence analyst	1
Data Engineers	1
Data handler	1
Data koordinator (coordinator)	1
Data specialists	1
Data / system owners	1
Datadisciple	1
IT Risk Management	1
IT-koordinator	1
Systemansvarlig og digitaliseringskonsulent	1
Projektmedarbejder	1
Research Data Officer	1
Research supporter	1

Open Science, more changes to follow



foster open science

There is no looking back really



Open Science is the only way forward in the post truth world



Thank you!

Twitter @shalini_kr



Published under a CC BY 4.0 license.

<https://creativecommons.org/licenses/by/4.0/>