

Getting Started with
Research Data
Management Support



Kaunas



30/10/2019

Who are we?



Santosh Ilamparuthi

Data Steward

TU Delft



Shalini Kurapati

Open Science & RDM
Fellow

Politecnico di Torino



Marta Teperek

Data Stewardship
Coordinator

TU Delft



Who are you?



Go to

www.menti.com

Use the code:

56 72 63



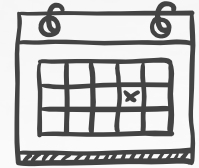


Aims of today's workshop

Aims of today's workshop

- ❑ Discuss the key elements of RDM support
- ❑ Provide concrete case studies on what can be done and tactics for researcher engagement
- ❑ Discover tools and resources which can be used/adapted
- ❑ Start discussions and collaborations

Agenda for today



- 10.00 - 10.15: Introduction
- 10.15 - 11.30: Getting started with RDM support
- 11.30 - 12.00: *Coffee break and snacks*
- 12.00 - 13.15: How to engage with researchers?
- 13.15 - 13.30: Summary

This workshop is for you - make it yours

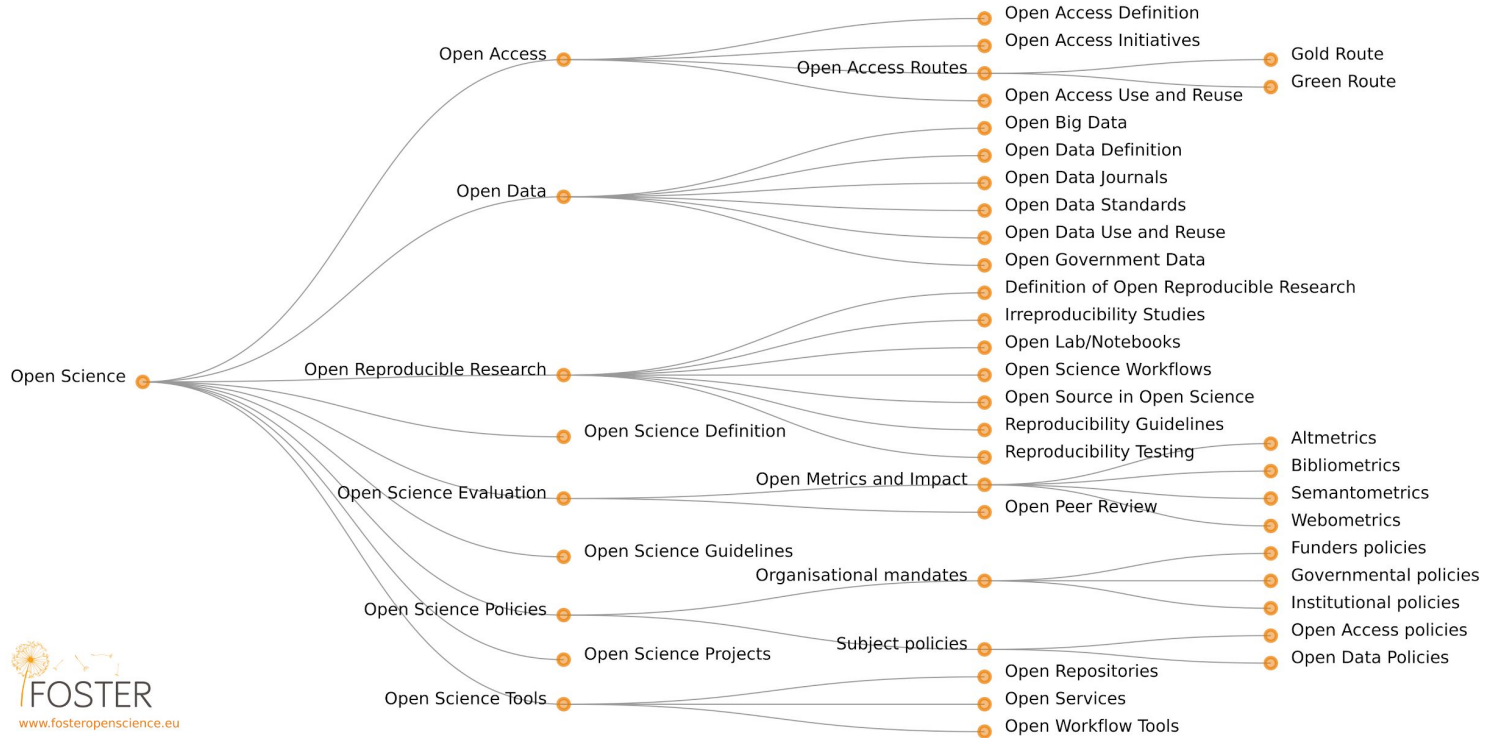
- Ask any questions you might have at any time
- Add your suggestions / solutions - we learn from each other
- We co-create this workshop

Setting the context

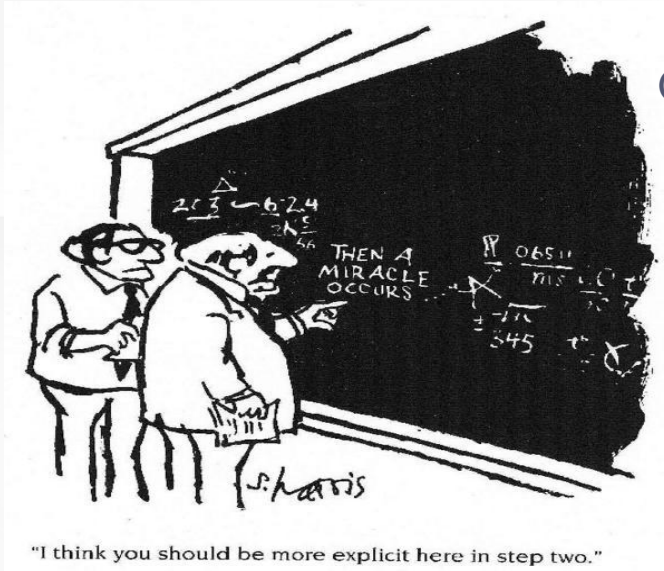
Why Open Science and RDM

What comes to your mind when you hear the term 'Open Science'?

Open Science Taxonomy



Science \neq Miracles



10.5281/zenodo.3472587

Open Science means:

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



**OPEN SCIENCE:
JUST
SCIENCE
DONE RIGHT**

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

If Open Science is science done right..

Why is everyone talking about RDM and Open Science now



THE EVOLUTION OF ACADEMIA

PUBLISH



PUBLISH
OR
PERISH



PUBLISH
IN HIGH IMPACT
JOURNALS
OR
PERISH



PUBLISH
FREQUENTLY IN
HIGH IMPACT
JOURNALS
AND
MAYBE
YOU WON'T
PERISH

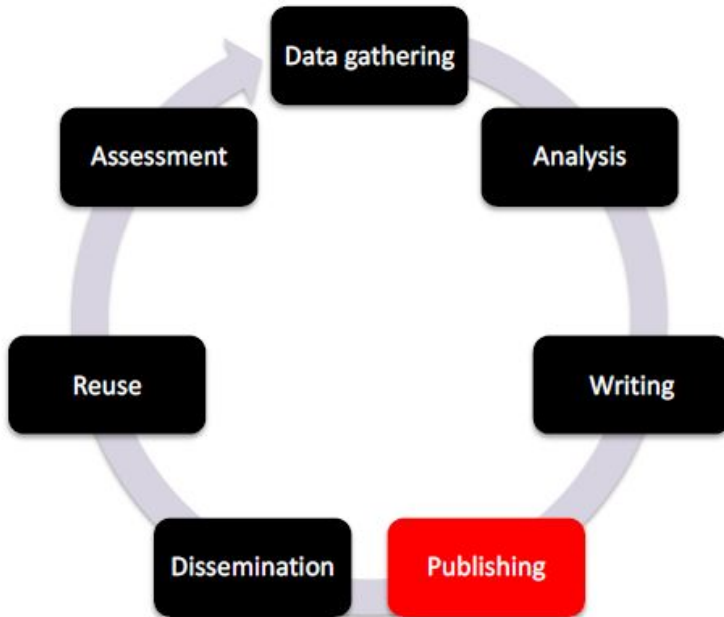


facebook.com/pedromics

Copyright:

<https://velica.deviantart.com/art/Publish-or-Perish-645355248>

How did we get here?



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>

Pressure to publish positive results...

- ❑ Design the study
- ❑ Collect data
- ❑ Analyze data as prespecified
- ❑ Oops! $P > 0.05$?
- ❑ Torture data until it confesses
- ❑ Then, and only then... write the manuscript



Adapted from:
Turner, Erick (2016):
<https://doi.org/10.6084/m9.figshare.3381379.v1> c

Some consequences

Report finds massive fraud at Dutch universities

Investigation claims dozens of social-psychology papers contain faked data.

Ewen Callaway

When colleagues called the work of Dutch psychologist Diederik Stapel too good to be true, they meant it as a compliment. But a preliminary investigative report (go.nature.com/tqmp5c) released on 31 October gives literal meaning to the phrase, detailing years of data manipulation and blatant fabrication by the prominent Tilburg University researcher.



It's not all about fraud, but mainly sloppy science

The reproducibility crisis

IS THERE A REPRODUCIBILITY CRISIS?



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

The main 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].

Funding bodies are pushing for open science, FAIR data



UK Research
and Innovation

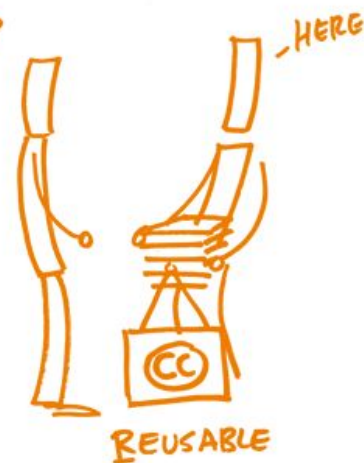
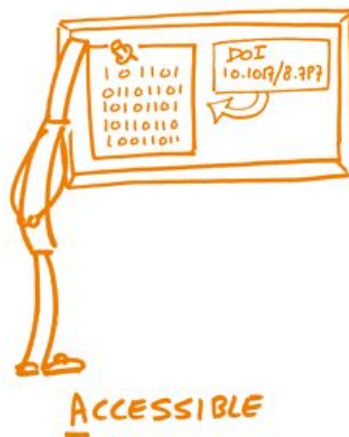


More will follow!

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

What is FAIR?

FAIR DATA PRINCIPLES



Hochstenbach, P. (2018). *Open Research Data Material - FAIR data principles*. [image] Available at: <https://hochstenbach.wordpress.com/> [Accessed 26 Apr. 2018].

Research Data Management 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

Overview of the RDM essentials in everyday research

Backup, Storage, File naming and organisation

What is your reaction if..

- Your laptop got stolen
- You lost your USB stick
- Your portable hard drive got damaged
- Your stuff on Dropbox / Google Drive disappeared



NOOOO
S**f***!!!



It hurts,
But I can
manage
with
some
work



Easy peasy, I
have it
covered

These things happen

CASH REWARD
for returning my lost backpack



200774ventolare.com

- Black [AK] Burton Rucksack
- Lost on Friday 15. July at 8 pm in the Panton Arms pub 43, Panton St. Cambridge
- Containing a laptop (white MacBook), a black external hard drive and scientific research documents

The external hard drive is **VERY** important to me as it contains 5 years of research data which are crucial for my PhD thesis!!!

If you found it, I would be extremely grateful if you could return it to the Panton Arms or contact me on: [REDACTED] (@cam.ac.uk)

Thank you!!



These things happen

Visiting Professor Robbed At Gunpoint, Years Of Research Data Stolen -



Visiting Professor Robbed At Gunpoint In Davis, Years Of Research ...

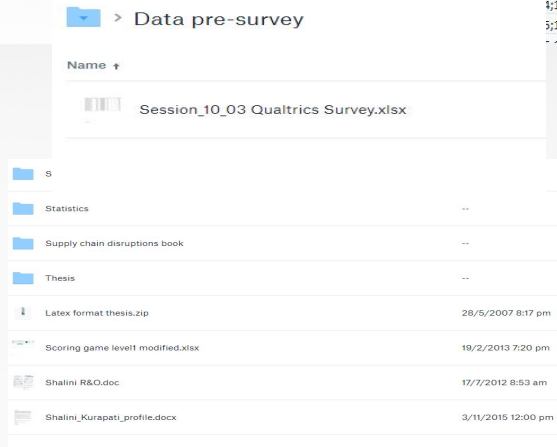
The search is on for the suspects who robbed a UC Davis professor at gunpoint - and now years of valuable scientific research data have been s...

sacramento.cbslocal.com

<https://sacramento.cbslocal.com/2019/09/26/uc-davis-visiting-professor-robbed-gunpoint-research-stolen/>

My own experience

Game data: Online, no local storage, sometimes data loss due to network issues, servers in NL, sessions in US, DE and NL



```
gsid;"ScorePunten";"ScoreContainers";"ScoreCranes";"ScoreVessels";"StarsTotalScore";"StarsContainers";"StarsCranes";"StarsVe  
15088;1000;500;0;19300;3;1;5;5;0;0;0;0;0;330.787;99.8424;500;0;0;0;207.89;0;122.896;2;3;0;0;0;0;0;0;0;1;1;1;193;193  
15091;1000;0;0;0;2;1;5;1;0;0;0;0;0;50.8116;10.1623;0;0;0;0;50.8116;0;2;3;0;0;0;0;0;0;0;0;0;1;1;0;0;0;0;0;0;15091;4;1;0;  
15092;1000;500;0;14400;3;1;5;5;0;0;0;0;0;397.737;146.826;500;0;0;0;361.75;0;35.9879;2;3;0;0;0;0;0;0;0;0;1;1;1;144;144  
15093;1000;500;0;14700;3;1;5;5;0;0;0;0;0;725.656;154.203;500;0;0;0;360.599;0;365.053;2;3;0;0;0;0;0;0;0;0;1;1;1;147;14  
15095;1000;500;0;9000;3;1;5;4;0;0;0;0;0;688.82;210.05;500;0;0;0;580.781;0;108.036;2;3;0;0;0;0;0;0;0;0;1;1;1;90;90;0;0;  
15096;1000;0;0;0;2;1;5;1;0;0;0;0;0;17.6702;3.53402;0;0;0;0;17.6702;0;2;3;0;0;0;0;0;0;0;0;1;1;0;0;0;0;0;0;15091;4;1;0;  
15098;1000;500;0;19400;3;1;5;5;0;0;0;0;0;337.66;92.8853;500;0;0;0;196.953;0;140.707;2;3;0;0;0;0;0;0;0;0;1;1;1;194;194  
15099;1000;500;0;21200;3;1;5;5;0;0;0;0;0;223.915;79.2127;500;0;0;0;176.193;0;47.7217;2;3;0;0;0;0;0;0;0;0;1;1;1;212;21  
15100;1000;500;0;15400;3;1;5;5;0;0;0;0;0;522.019;150.016;500;0;0;0;411.868;0;110.151;2;3;0;0;0;0;0;0;0;0;1;1;1;154;15  
15101;1000;500;0;11300;3;1;5;5;0;0;0;0;0;741.69;196.303;500;0;0;0;618.976;0;122.705;2;3;0;0;0;0;0;0;0;0;1;1;1;113;113  
15103;1000;500;0;14100;3;1;5;5;0;0;0;0;0;530.218;173.987;500;0;0;0;427.559;0;102.66;2;3;0;0;0;0;0;0;0;0;1;1;1;141;141  
1;1000;300;0;0;2;1;5;1;0;0;0;0;0;508.793;145.745;300;0;0;0;445.318;0;63.476;2;3;0;0;0;0;0;0;0;0;1;1;0;0;0;0;0;0;0;  
5;1000;500;0;7900;3;1;5;4;0;0;0;0;0;754.93;220.39;500;0;0;0;674.076;0;80.8513;2;3;0;0;0;0;0;0;0;0;1;1;1;79;79;0;0;
```

Local drives Post-game survey embedded in the game platform, limited access

MATB data local: collected on pendrive, sometimes computer crashed, data loss

- COMM_2014_10031139.txt
- MATB_2014_10031139.txt
- RMAN_2014_10031139.csv
- RMAN_2014_10031139.txt
- SYSM_2014_10031139.txt
- TRCK_2014_10031139.csv
- TRCK_2014_10031139.txt
- WRS_2014_10031139.csv
- WRS_2014_10031139.txt

My own experience

The image shows a screenshot of the SurveyMonkey dashboard. In the background, there's a survey titled "Disruption mitigation board-game usability survey" with various tabs like "OVERVIEW", "DESIGN SURVEY", "PREVIEW & SCORE", "COLLECT RESPONSES", and "ANALYZE RESULTS". A modal window titled "Put Your Data to Work" is open, stating "Export in the formats you want. Upgrade to a paid plan to export your survey results as PowerPoint, spreadsheets, and more. Download your data into any of these formats: CSV, PNG, PDF, PPT, SPSS and XLS." In the foreground, a "Plans & Pricing" table is displayed, comparing four plans: Basic (€0), Standard (€420/year), Best Value Advantage (€456/year), and Premier (€1,188/year). The Best Value Advantage plan is highlighted in green.

Plans & Pricing
Choose your plan today. Get started now with the survey platform that adapts to your needs—even if they change over time.

	BASIC	STANDARD	BEST VALUE ADVANTAGE	PREMIER
With SurveyMonkey, you can:	€ 0 Always free	€ 420 / year See monthly plan	€ 456 / year	€ 1,188 / year
	YOUR PLAN	UPGRADE	UPGRADE	UPGRADE
Number of surveys @	UNLIMITED	UNLIMITED	UNLIMITED	UNLIMITED
Questions per survey @	10	UNLIMITED	UNLIMITED	UNLIMITED
Number of responses @	100 / survey	UNLIMITED	UNLIMITED	UNLIMITED
24/7 customer support via email @	✓	Priority	Expedited	Expedited & Phone
NEW! Team accounts & tools @ Collaboration tools for survey projects	1 user total with commenting only	1 user total with commenting only	2 users total with all team tools	2 users total with all team tools
NEW! Quizzes @ Create assessments with automatic scoring	✓	✓	✓	✓
Customized surveys @ Logos, colors, survey URL		✓	✓	✓
Data exports @ CSV, PDF, PPT, XLS		✓	✓	✓
Advanced data exports @ SPSS			✓	✓
Customized survey experience @ Skip logic, question & answer piping		Skip logic only	✓	✓

Used survey monkey in 2012. They changed policy for exporting my own data!

Don't leave it to Google!

Google services Terms of Use:

When you upload, submit, store, send or receive content to or through our Services, you give Google (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works (such as those resulting from translations, adaptations or other changes we make so that your content works better with our Services), communicate, publish, publicly perform, publicly display and distribute such content. The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones. This license continues even if you stop using our Services (for example, for a business listing you have added to

<https://policies.google.com/terms?hl=en>

And this..



new.psd



newfinal.psd



newfinalfinal.psd



newfinalestfinal.psd



newfinalestfinal
forsure.psd



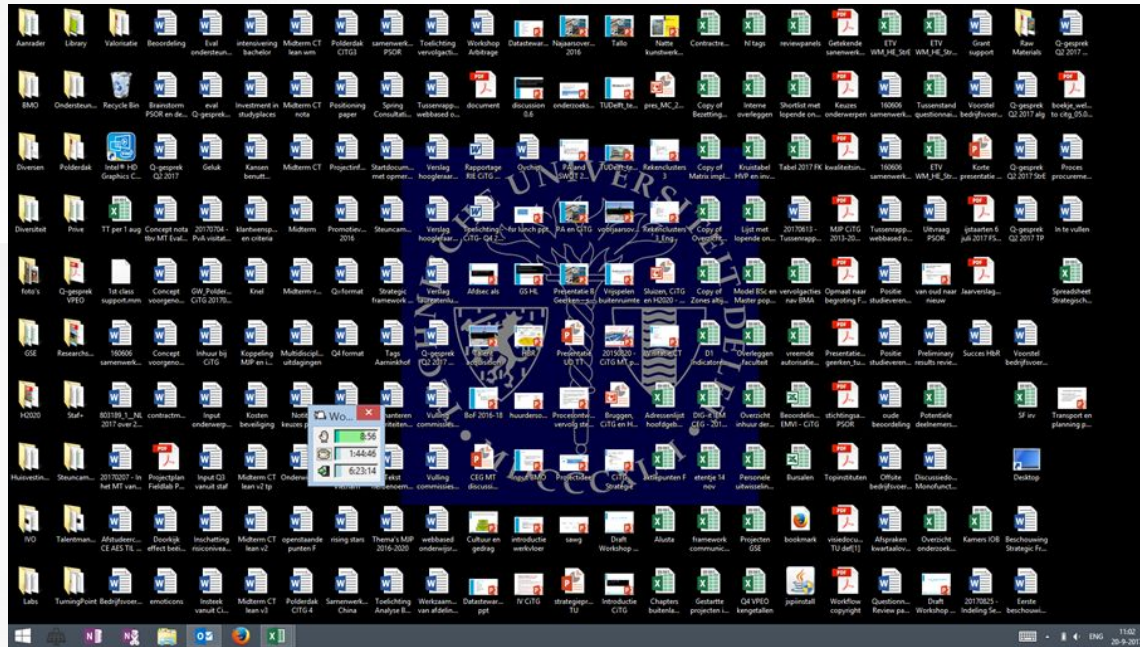
newfinalestfuckthis
shitfinal.psd

@10PM.COM

Copyright:

<http://10pm.com/>

Recognise this?



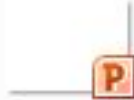
And this..



Important



Experiment
1



Lab
meeting
FINAL



Meeting
notes



My talk



Paper
submission



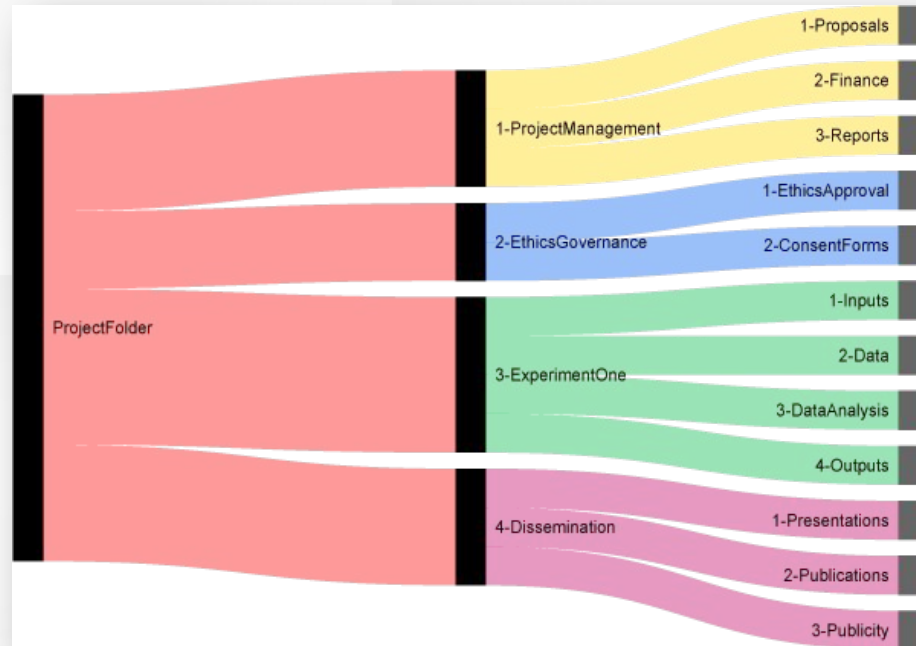
PhD revised

In three years would you recognise these?

Data organisation: Tips

- ▣ Consistent
- ▣ Meaningful to you and your colleagues
- ▣ Think about the continuity of your research
- ▣ Would you be able to easily find your own data files?
- ▣ Would your colleague be able to easily find and understand your data files?
- ▣ Critically look at your files (and clean them if needed!) once a week

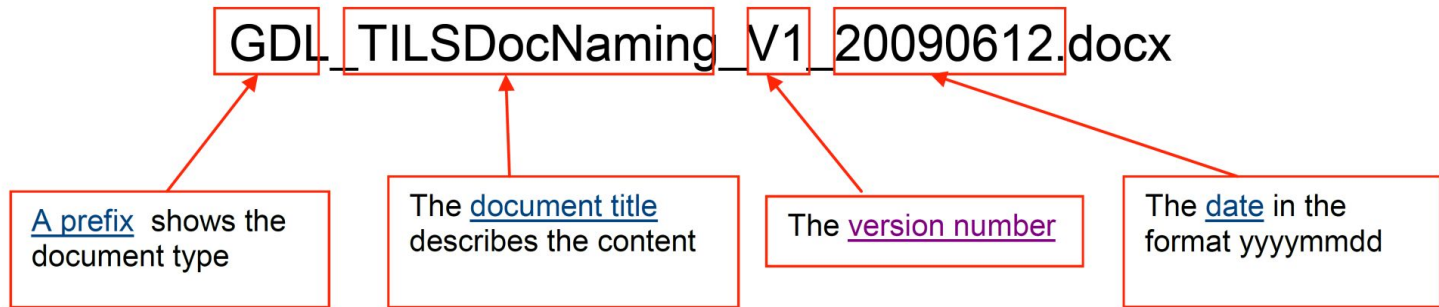
Folder organisation



File naming: Human and machine friendly

TILS Document Naming Convention

Document naming for the TILS Division should follow this convention:



http://www.data.cam.ac.uk/files/gdl_tilsdocnaming_v1_20090612.pdf

Putting this into practise with a dedicated tool

Data Management Plans

Data Management Plans

During
Research



After
Research



DMP for
Research
Project

Data Management Plan - During Research

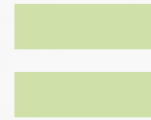


Data Management Plan - After Research

Archiving



Access



After
Research

How to for DMPs

Funder

- General Template
- Initial Vs Final

Self

- University Template
- Specific Needs

Creating a Plan:

Go to dmponline.dcc.ac.uk

Login Page

Welcome

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It is provided by the Digital Curation Centre (DCC).

Join the growing international community that have adopted DMPonline:



17,622 Users



203 Organisations



23,083 Plans



89 Countries

Some funders mandate the use of DMPonline, while others point to it as a useful option. You can [download funder templates](#) without logging in, but the tool provides tailored guidance and example answers from the DCC and many research organisations. Why not sign up for an account and try it out?

Sign in

Create account

* Email

* Password

[Forgot password?](#)

Remember email

Sign in

- or -

Sign in with your institutional credentials

Creating the Plan

Create a new plan

Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.

* What research project are you planning?

mock project for testing, practice, or educational purposes

* Select your research organisation

- or - No research organisation associated with this plan or my research organisation is not listed

* Select the primary funding organisation

- or - No funder associated with this plan or my funder is not listed

Create plan

Cancel

DMPs

- ▣ Great as a planning and tracking tool for RDM
- ▣ Funders expect it
- ▣ Use [dmponline](#) as a resource with excellent guidelines and exportable DMPs

Time for a break!



11.30 - 12.00



Open Science: What's in for me, selfish benefits
for researchers

We all know Open Science is good for the society



Increasing Impact

Who looks into your data set? Google does!

A search input field with a vertical cursor on the left and a microphone icon on the right.

Google Search

I'm Feeling Lucky

Use Orcid to your advantage

 Impactstory



Ethan White   

University of Florida Associate Professor

OVERVIEW

ACHIEVEMENTS

TIMELINE

PUBLICATIONS

ACHIEVEMENTS

[view all](#)



Open Hero Top 10%

Every single one of your papers is free to read online. Wow! That's a level of access only 2% of other researchers achieve. Open access [helps real people](#), and that's pretty heroic.



Global Reach Top 25%

Your research has been saved and shared in 35 countries. That's high: only 18% of researchers get that much international attention.



Greatest Hit

Your top publication has been saved and shared 3 times.

TIMELINE

3 Online mentions
over 1 years



2 1

PUBLICATIONS

 Ten simple rules for a successful remote postdoc
2019

3 

 Individual Tree-Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural Networks
2019 *Remote Sensing*

 [datacarpentry/organization-geospatial](#): Data Carpentry: Introduction to Geospatial Concepts, June 2019
2019

Publish your presentations

zenodo



Upload

Communities

Log in

Sign up

September 4, 2018

Presentation Open Access

Open science in a nutshell: What's in for me?

Kurapati, Shalini; de Jong, Michiel

"Open science in a nutshell" is a presentation given by Shalini Kurapati and Michiel de Jong to PhD students at the Faculty of Technology, Policy and Management at TU Delft on 7 September 2018.

Abstract

Shalini will talk about the benefits of data management and the motivation behind data sharing during the PhD research cycle and offer some practical tips on these topics. She will also throw some light on the available resources, training and support available for PhD students with regards to disciplinary specific data management.

Michiel will elaborate on how to create impact with open science and how open science can help increase your web presence. He will introduce several platforms which can help you showcase your work (including social media tools).

2,142

views

410

downloads

[See more details...](#)



39

[See more details](#)

- Tweeted by 49
- On 1 Facebook pages
- Mentioned in 1 Google+ posts
- 1 readers on Mendeley

<https://doi.org/10.5281/zenodo.1409027>

Publish your data in a certified repository

4TU.Centre for Research Data

[Contact](#) | [Terms of use](#) | [Login](#)

Dataset | **Data presented in the Paper "Behavioral self-organization underlies the resilience of a coastal ecosystem" by De Paoli et al (PNAS 2017)**

▶▶▶▶▶ Link/cite as <https://doi.org/10.4121/uuid:9e1752fd-cce7-408e-a02c-16d1f663ec5f> | [full citation](#)

▼ go to DATA section ▼

title	?	Data presented in the Paper "Behavioral self-organization underlies the resilience of a coastal ecosystem" by De Paoli et al (PNAS 2017)
creator	?	orcid van de Koppel, J. (Johan)
creator	?	orcid de Paoli, H. (Hélène)
contributor	?	orcid Herman, P.M.J. (Peter)
contributor	?	NIOZ Royal Netherlands Institute for Sea Research, Department of Estuarine and Delta Systems, and Utrecht University
contributor	?	orcid Silliman, B.R. (Brian)
contributor	?	van den Berg, A. (Aniek)
contributor	?	van der Heide, T. (Tjisse)
date accepted	?	2017-06-12
date created	?	2017-04-11
date published	?	2017
description	?	This dataset contains the experimental data that is published in the paper "Behavioral self-organization underlies the resilience of a coastal ecosystem", with the above contributors as authors, published in the Proceedings of the National Academy of Science of the USA. The data was obtained during the Waddensleutels and klosselwad projects. The current dataset contains the raw data and R scripts for analysis and for the figures in that manuscript.
language	?	en
publisher	?	NIOZ Royal Netherlands Institute for Sea Research
spatial coverage	?	Intertidal flats inc. oyster/mussel reefs south of Schiermonnikoog, Netherlands ▶ The experiments were conducted on mussel beds situated at land-bound intertidal flats to the south of Schiermonnikoog, the Netherlands (location: 53.455, 6.147 DD)
subject	?	Mussel beds • Mussels • Restoration • Spatial ecology • Spatial patterns
▲ in collection	?	General collection of datasets
time coverage	?	years 2012-2013
related publication	?	Behavioral self-organization underlies the resilience of a coastal ecosystem (article, 2017)
licence	?	General terms of use

DATA

[?](#) Dataset files (80.9 KiB) >> [download complete dataset \(zip\)](#) | [download separate files](#)

- bag-info
- contents of this dataset, 13 files

<< more info...

Home

Upload datasets

Personal page

🔍

- » Search in Data
- » Search in "info"

U Delft

U/e

IIVERSITEIT
VENTE.

<https://doi.org/10.4121/uuid:9e1752fd-cce7-408e-a02c-16d1f663ec5f>

Track the impact

Behavioral self-organization underlies the resilience of a coastal ecosystem

H de Paoli, T van der Heide... - Proceedings of the ..., 2017 - National Acad Sciences

Abstract Self-organized spatial patterns occur in many terrestrial, aquatic, and marine ecosystems. Theoretical models and observational studies suggest self-organization, the formation of patterns due to ecological interactions, is critical for enhanced ecosystem resilience. However, experimental tests of this cross-ecosystem theory are lacking. In this study, we experimentally test the hypothesis that self-organized pattern formation ...

Cited by 2 All 6 versions Web of Science: 1 Cite Save

 Create alert

Publish your code

The screenshot shows the Zenodo website interface. At the top is a blue navigation bar with the Zenodo logo on the left, a search bar in the center, and 'Upload' and 'Communities' links on the right. Below the navigation bar, the page content includes the date 'November 23, 2017' and two tags: 'Software' and 'Open Access'. The main title is 'Pyerls - Python modelling of dislocations using the Peierls analysis'. Below the title, the author is listed as 'Thompson, Robert Peter' with a small ID icon. Underneath, it says 'Thesis supervisor(s) Clegg, William J'. A description follows: 'Supplementary Code to sit alongside a thesis submission, University of Cambridge, 2017'. At the bottom of the page, there is a 'Preview' section with a dropdown arrow. It contains a file entry 'pyerls-thesis_supplement.zip' with a file icon and a download icon, and a folder entry 'pyerls-thesis_supplement' with a folder icon.

zenodo Search Upload Communities

November 23, 2017 Software Open Access

Pyerls - Python modelling of dislocations using the Peierls analysis

Thompson, Robert Peter

Thesis supervisor(s)
Clegg, William J

Supplementary Code to sit alongside a thesis submission, University of Cambridge, 2017

Preview

- pyerls-thesis_supplement.zip
- pyerls-thesis_supplement

<https://doi.org/10.5281/zenodo.1065538>

understanding RDM costs

You can find many many free and certified repositories upto a certain data size: Zenodo, 4TU centre for research data etc.

If you need to understand/advise on how to budget for RDM costs in general:

4TU RDM Costing Tool



What's Zingtree?

Data Management Costing Tool

Data Management costing tool



Welcome to the Data Management Costing Tool. This is for TU Delft researchers and staff to help determine costs and staffing requirements in project proposals. Let's start with some questions about your project which will help us estimate the data management needs of your project.



[bit.ly/RDM costing](https://bit.ly/RDMcosting)

OpenAire RDM costing tool



Guides for Researchers

How to identify and assess Research Data Management (RDM) costs

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

THE COST OF DATA MANAGEMENT

HOW TO CALCULATE COSTS?

Estimating costs RDM tool

COMMENTS (1)

Increase your funding chances

Winning Horizon 2020 with Open Science

Grigorov, Ivo; Davidson, Joy; Donnelly, Martin; Jones, Sarah; Elbaek, Mikael

Investing in research towards solutions for Societal Challenges is a key priority of the EU Innovation Union[1], and the EC has placed focus on "Open Innovation, Open Science and Open to the World"[2] as a strategy to make the EU leader in Research & Innovation.

Open Science is about removing all barriers to full sphere basic research knowledge and outputs (research data, research code, publications, policy briefs), and thus feeding Open Innovation and the knowledge-based economy. The concept is central to EU's Responsible Research and Innovation philosophy[3], and removing barriers to basic research measurably contributes to the EU's Blue/Green Growth Agenda[4]. Despite the potential of the internet age to deliver on that promise only 50% of today's basic research is freely available[5].

The workshop will demonstrate how and why Open Science can be a key tool for research managers formulating and managing grant proposals.

Learning Outcomes follow FP7 FOSTER Open Science Learning Objectives (<http://dx.doi.org/10.5281/zenodo.15603>) :

- Understand EC's Open Science Agenda priorities;
- Gain insight into what is, and will be, mandated on access to research, and data management;
- Be capable to integrate Open Science in the research concept, and its impact;
- Understand Open Science's impact on EC Evaluation process based on examples from 2014-2015.

<https://zenodo.org/record/46602#.XbCnKegzZPY>

Helps you while writing papers

The screenshot shows a web browser displaying a PLOS Medicine article. The browser's address bar shows the URL: journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001789. The page header includes the PLOS MEDICINE logo, navigation links (Browse, Publish, About), a search bar, and user options (create account, sign in). Below the header, there are links for OPEN ACCESS and PEER-REVIEWED. The main content area is divided into two columns. The left column contains navigation options: Discussion, Supporting Information (highlighted), Acknowledgments, Author Contributions, References, Reader Comments (0), Media Coverage (1), and Figures. The right column lists supplementary materials: S1 Protocol. Sweave file for survival analysis. (PDF), S1 Table. Distribution of samples. (PDF), and S2 Table. Digital PCR results. (PDF). Below this, a 'Subject Areas' section lists various topics with radio buttons: Genome evolution, Phylogenetic analysis, Chemotherapy, Evolutionary genetics, Genomic medicine, Cancer treatment, and Neurofibromatosis t... The bottom of the page shows the start of the article's abstract and 'Methods and Findings' section.

Discussion

Supporting Information

Acknowledgments

Author Contributions

References

Reader Comments (0)

Media Coverage (1)

Figures

DOI: 10.1371/journal.pmed.1001789.s019 (PDF)

S1 Protocol. Sweave file for survival analysis.

doi:10.1371/journal.pmed.1001789.s020 (RNW)

S1 Table. Distribution of samples.

doi:10.1371/journal.pmed.1001789.s021 (PDF)

S2 Table. Digital PCR results.

doi:10.1371/journal.pmed.1001789.s022 (PDF)

Subject Areas

- Genome evolution
- Phylogenetic analysis
- Chemotherapy
- Evolutionary genetics
- Genomic medicine
- Cancer treatment
- Neurofibromatosis t...

Discussion

Supporting Information

Acknowledgments

Author Contributions

References

Reader Comments (0)

Media Coverage (1)

Figures

Methods and Findings

Evolutionary inference and phylogenetic quantification of heterogeneity was performed using the MEDICC algorithm on high-resolution whole genome copy number profiles and selected genome-wide sequencing of 135 spatially and temporally separated samples from 14 patients with HGSC who received platinum-based chemotherapy. Samples were obtained from the clinical CTCC QV0304 studies, and patients were enrolled between 20 July 2007 and 22 October 2009. Median follow-up of the cohort was 31 mo (interquartile range 22–48 mo), censored after 26 October 2013. Outcome measures were overall survival (OS) and progression-free survival (PFS). There were marked differences in the degree of clonal expansion (CE) between patients (median 0.74, interquartile range 0.66–1.15), and dicholimitation by median CE showed worse survival in CE-high cases (PFS 12.7 versus 10.1 mo, $p = 0.009$; OS 42.6 versus 23.5 mo, $p = 0.003$). Bootstrap analysis with resampling showed

Capture all the complexity of your research beyond just text and pdf

But.. what about my career



Things are changing.. slowly but surely!

The 'reputed' journals need it already!

nature > nature research > editorial policies > reporting standards and availability of data, materials, code and protocols

nature research

Search

Login

Editorial policies

Authorship

Competing interests

Confidentiality

Plagiarism and duplicate publication

Image integrity and standards

Preprints & Conference Proceedings

Peer-review policy

Reporting standards and availability
of data, materials, code and

Reporting standards and availability of data, materials, code and protocols

An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. A condition of publication in a Nature Research journal is that **authors are required to make materials, data, code, and associated protocols promptly available to readers without undue qualifications.** Any restrictions on the availability of materials or information must be disclosed to the editors at the time of submission. Any restrictions must also be disclosed in the submitted manuscript.

Decision makers are pushing for change



Evaluation of Research Careers fully acknowledging Open Science Practices

Rewards, incentives and/or recognition for researchers
practicing Open Science

[https://cdn1.euraxess.org/sites/default/files/policy_library/os-rewards-wgreport-final_integrated_](https://cdn1.euraxess.org/sites/default/files/policy_library/os-rewards-wgreport-final_integrated_0.pdf)


[0.pdf](https://cdn1.euraxess.org/sites/default/files/policy_library/os-rewards-wgreport-final_integrated_0.pdf)

Funders might prefer those who practise Open Science..

~~SOON~~

Open Science Career Assessment Matrix (OS-CAM)	
<i>Open Science activities</i>	<i>Possible evaluation criteria</i>
RESEARCH OUTPUT	
Research activity	Pushing forward the boundaries of open science as a research topic
Publications	Publishing in open access journals Self-archiving in open access repositories
Datasets and research results	Using the FAIR data principles Adopting quality standards in open data management and open datasets Making use of open data from other researchers
Open source	Using open source software and other open tools Developing new software and tools that are open to other users
Funding	Securing funding for open science activities

There's a whole collection of academic jobs mentioning OS














OSFHOME ▼

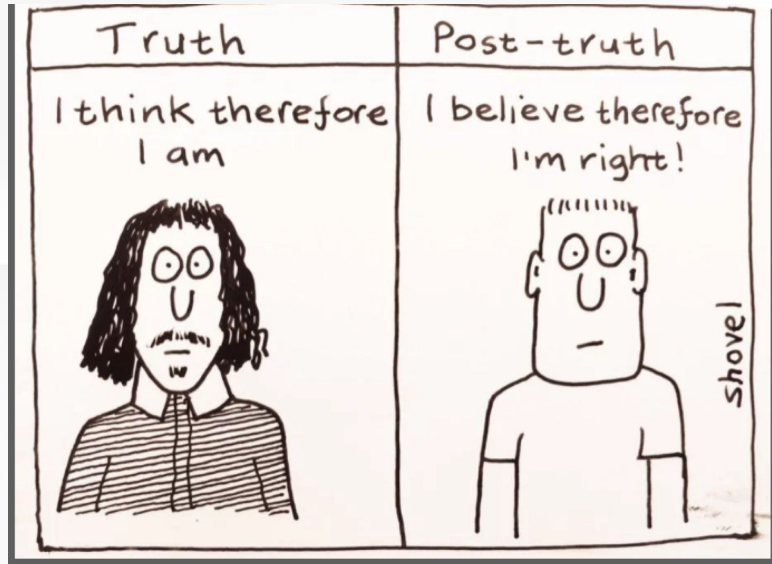
Academic job offers that mentioned op...

Files

<https://osf.io/7jbnt/>

Name ^ v	Modified ^ v
Academic job offers that mentioned open sci...	
- OSF Storage (United States)	
 180911_OTTT_Sensory_Sciences_deuts...	2019-03-01 08:24 AM
 180911_OTTT_Sensory_Sciences_englis...	2019-03-01 08:24 AM
 2019_HH-Uni_W2-3_Sportpsychologie.p...	2019-05-21 08:26 AM
 Aachen_W3_Prof_Angewandte_comput...	2019-04-10 11:52 AM
 AllgemeinePsychologieUHH.pdf	2018-10-12 08:51 AM
 Announcement_W3_Professur_SocialPs...	2018-01-18 02:25 PM
 Bielefeld_Psychologie_Sportwiss_W2_P...	2019-04-23 11:53 AM
 Charite_Berlin.png	2018-03-10 09:08 AM
 Dresden_WiMi.pdf	2018-07-13 12:37 PM
 Keele_Lecturer_2018.pdf	2018-07-11 01:14 PM
 Leipzig-W2_Professorship_in_Psycholog...	2019-05-17 06:11 AM

There is no looking back really



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

Comic by
<https://thenorwichradical.com/2017/01/12/post-truth-politics-and-the-war-on-intellect/>

TU Delft Data Stewardship Team



Coordinator located at the library
A dedicated Data Steward at every faculty:
all have a PhD (or equivalent)

<https://www.tudelft.nl/library/datastewardship/>



What do the Data Stewards do?



Personal approach:

- They are there for any data questions
- Help researchers who “don’t have data” and who “don’t have problems”
- Places: coffee machine, canteen, lift, doorway...
- Every occasion is good!

Change happens one person at a time

Objective: Improve daily practice not enforcement



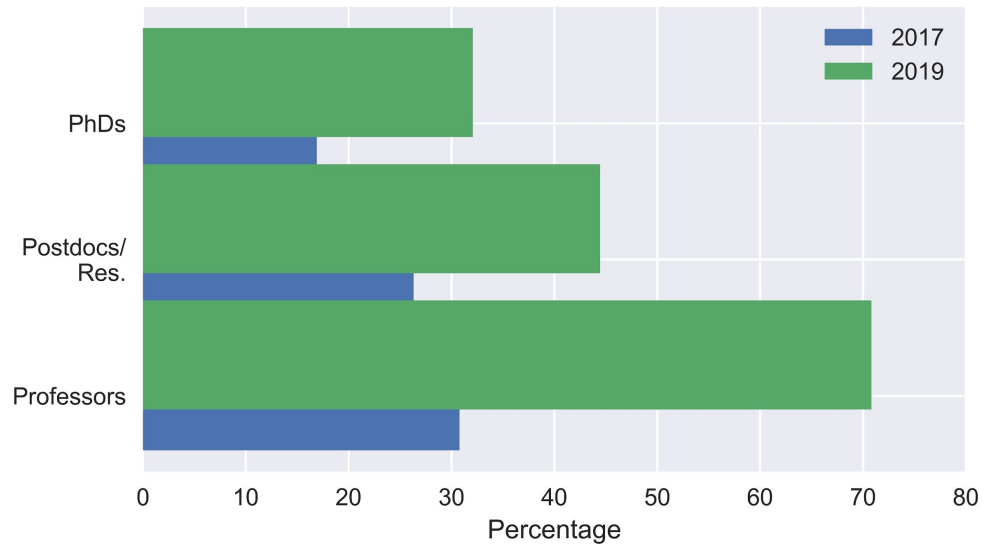
Does it make a difference to have a Data Steward?



Heather Andrews

Heather Andrews, Data
Steward at Faculty of
Aerospace Engineering,
TU Delft

Are you aware of funders' expectations for Findable, Accessible,
Interoperable and Reusable (FAIR) data?
Results 2017 & 2019



"Change happens one person at a time"

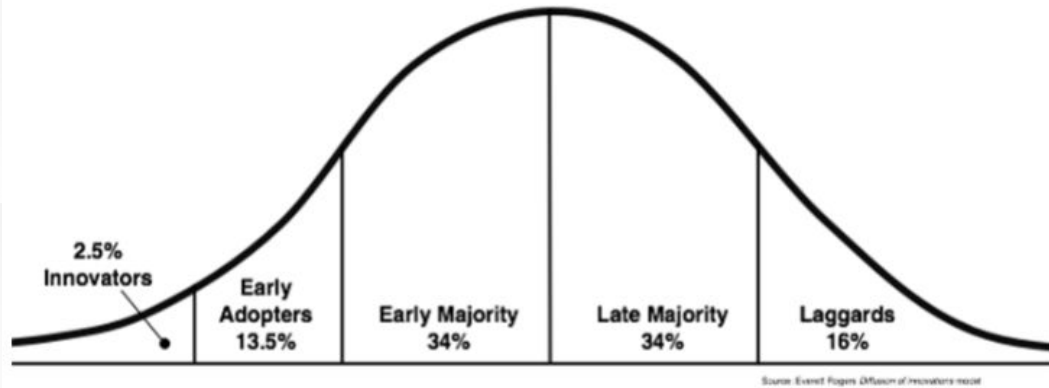


Jonathan Petters,
Virginia Tech

"To a large extent, researchers are not interested in data management planning. Apart from the fact that they have to do it for grant proposals, they're generally not interested in sharing data. They want their problems to be solved (...) So our aim (...) is to help researchers where they want help. While helping them we have the opportunity to talk to them about data management planning and data sharing"

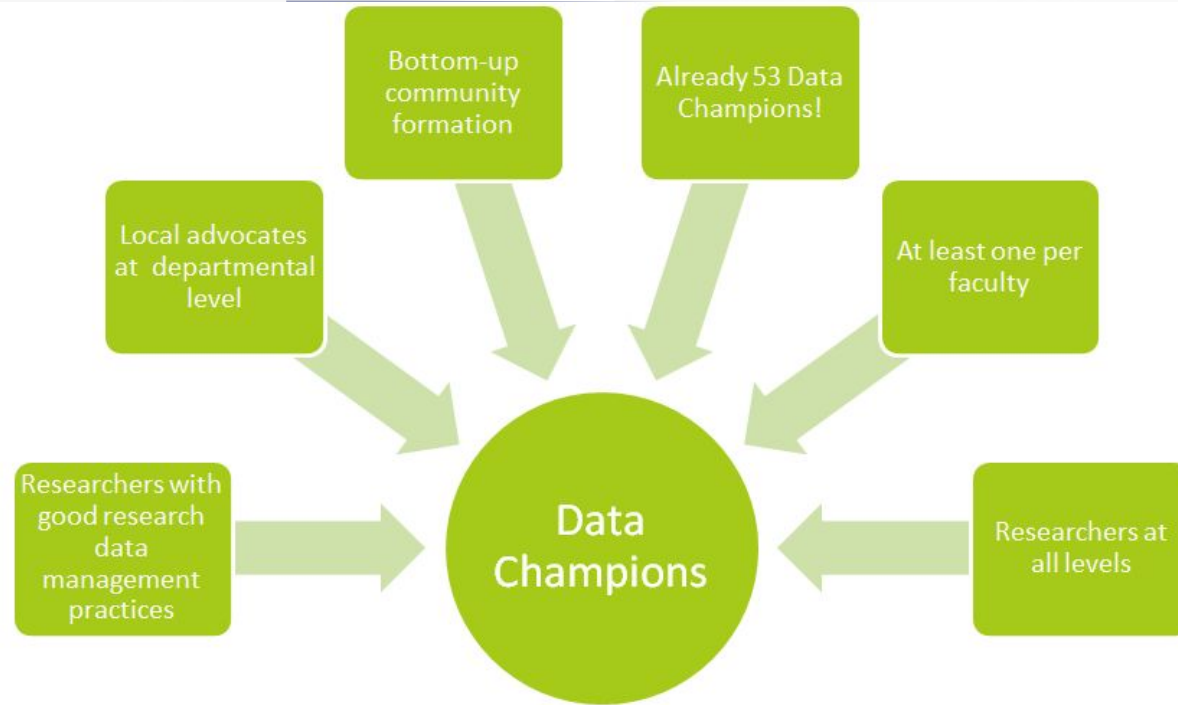
Engaging Researchers with Data Management: The Cookbook,
DOI: 10.11647/OBP.0185

Small incremental (realistic!) steps can make a big difference



Diffusion of Innovations; Rogers, 1963

Data Champions



Data Champions - The Rewards

Increased impact & visibility

Get involved in policy development

Networking opportunities

Travel grants



Learning new skills

Credited during evaluation

What do the Data Champions do?
















Launch of TU Delft Open Hardware Community

 **DelftOpenHardware** 
Group ID: 6229781

People at TU Delft interested and working on open hardware projects.

Subgroups and projects Shared projects Archived projects Last created ▼

  Open-source-Centrifuge-for-WetLab 	★ 0	4 days ago
 F Fume Sensor for Precious Plastic Machines  A project to give an indication of the amount of fumes in the air and v	★ 0	4 days ago
  community-website 	★ 0	3 days ago
  hardocs 	★ 0	4 days ago
 B Biogas  Mesothermic bio-digester for small farms	★ 0	2 days ago

<https://gitlab.com/go-commons/delftopenhardware>

What do the Data Champions do?

Champion the use of new tools



**Keep calm and go paperless:
Electronic lab notebooks can improve
your research.**



<https://openworking.wordpress.com/2019/07/05/keep-calm-and-go-paperless-electronic-lab-notebooks-can-improve-your-research/>

Data Stewardship and Data Champions

- ▣ Infrastructure is important but at the end of the day people do the research
- ▣ In person interaction with researchers is important to get their buy in
- ▣ Peer exemplars as examples is great for advocacy and motivation for the good RDM practice

Case Study from Italy

RDM and OpenScience at PoliTechnico di Torino



35 000	Students enrolled in Bachelor's and Master's degree programmes
565	Students enrolled in I and II level Specializing Masters programmes Lifelong Learning courses Training courses for businesses
684	PhD Doctoral Candidates a.y. 2018/2019
72%	Male Students
28%	Female Students
68%	Students coming from outside Piedmont 52% Italians living outside Piedmont 16% international students
5 302	First-year students a.y. 2018/2019



PhD programme

Out of the 771 PhD candidates

SCHOLARSHIPS

85% of PhD candidates receive a scholarship (616 in total)

The PhD grant is **30%** higher than national average level (**18,855 Euros** per year gross)

25% of PhD scholarships are funded by external non-academic bodies

10 PhD programmes offered with the apprenticeship scheme

PhD PROGRAMMES DURATION

3 years

1/3 learning activity

2/3 research activity

EXCELLENCE OF PhD CANDIDATES

80% of PhD candidates have a Master's degree with final grade of **105/110**

60% with **110/110**



International Research Grants



246 European and International Projects

76 Million EUR
--- since 2014

191 EU Horizon 2020 Projects

46 coordinated
70 Mio EUR | 16,55% success rate
among which

12 ERC
European Research Council

36 MSCA
Marie Skłodowska Curie Actions

118 IL-SC-JTI
Industrial Leadership Societal Challenges e Joint Technology Initiative

15,5 M EUR

11,5 M EUR

36 M EUR

3 EIT Communities

European Institute of Innovation & Technology

Inno Energy

EIT Digital

EIT Manufacturing

Collaborations with Companies and Banking Foundations



FIAT CHRYSLER AUTOMOBILES



49 R&D projects

In 2014-2018

45 R&D projects

In 2016-2018

9 R&D projects

In 2014-2017

Collaboration with
big industrial
partners



Research Infrastructures

IAM@Polito | Integrated Additive Manufacturing@Politecnico di Torino

DynLab4JMat | Laboratorio per l'Integrità Strutturale di Giunzioni e Materiali Sottoposti a Carichi Dinamici

TEST-eDrive | Nuova infrastruttura di test di e-DRIVE per applicazioni automobilistiche

FIP | Tecnologie Fotoniche per l'Industria 4.0 in Piemonte

PiqueT | Piemonte Quantum Enabling Technology

SAX | Strumentazioni avanzate per sistemi complessi

11
Open Research Infrastructures

HPC4AI | Centro di Competenza Calcolo ad Alte Prestazioni e Intelligenza Artificiale Torino

CCL | CO2 Circle LAB

PAsTISs | PARco per le Tecnologie Innovative per la Salute

CARS-HEV@Polito
Centre for Automotive Research and Sustainable mobility-Hybrid/Electric powertrain and Vehicle Laboratory

BRISK2 | Biofuels Research Infrastructure

Well equipped IT infrastructure wise

- The Information Technology area provides services to research: computing and data backup needs.
- Researchers can request secure backup (for regular data) in two different locations for upto 6 months of retrieval
- Applications: Web server, database, computing power, software, and virtual infrastructure (table)

Virtual Machines	Virtual CPUs	RAM [GB]	Storage [TB]	Backup [TB]
30	98	250	11	40

Cloud Act

Some push for data sovereignty

CLOUD Act

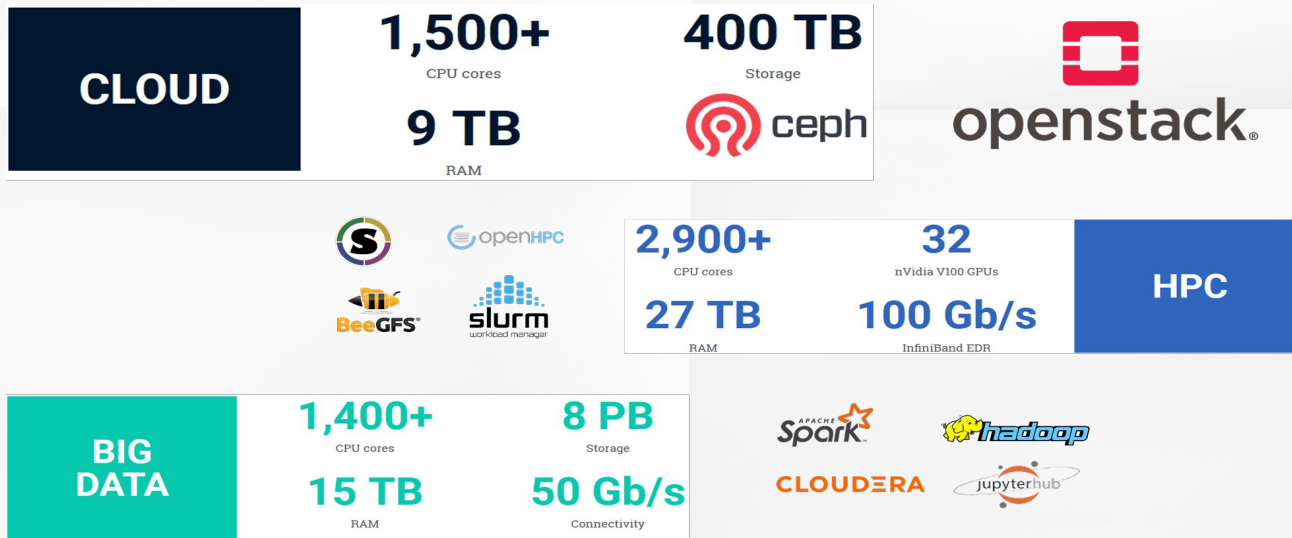
From Wikipedia, the free encyclopedia

The **Clarifying Lawful Overseas Use of Data Act** or **CLOUD Act** ([H.R. 4943](#)) is a United States **federal law** enacted in 2018 by the passing of the [Consolidated Appropriations Act, 2018](#), PL 115-141, section 105 [executive agreements](#) on access to data by foreign governments. Primarily the CLOUD Act amends the [Stored Communications Act](#) (SCA) of 1986 to allow federal law enforcement to compel U.S. based technology companies via warrant or subpoena to provide requested data stored on servers regardless of whether the data are stored in the U.S. or on foreign soil.

[Contents](#) [[hide](#)]

Big data, HPC and cloud at PolTo

HPC, Private cloud and Big Data services in collaboration with SmartData@Polito e HPC@POLITO.



My role at PoliTo

Helping Dr. Federica Cappelluti, rector's advisor on open science on:

- ▣ Training and Awareness on OS and RDM
- ▣ Demonstrating value and driving cultural change towards open science
- ▣ Collaborating with university stakeholders
- ▣ Design of a policy roadmap
- ▣ Advice (as much as I can) on all data related matters of Open Science
- ▣ Researcher oriented

Future plans

- ▣ RDM policy development and implementation
- ▣ Training and awareness FAIR data and Data Management Plans
- ▣ Researcher engagement
- ▣ Industry collaborations and Open Science
- ▣ Data and Software Carpentries
- ▣ PhD students and graduate schools
- ▣ Long term plan: To increase support and keep it sustainable

We have the technical infrastructure, we need
human capa

Engaging Researchers with Data Management: What Works?

So what works?

How to engage researchers with RDM?

All researchers?

And how to avoid THIS?

(and make them excited instead?)

*How to re-use and stop re-inventing
the wheel?*



Word of mouth (and conferences) help...



**...but is it really the
best way?**

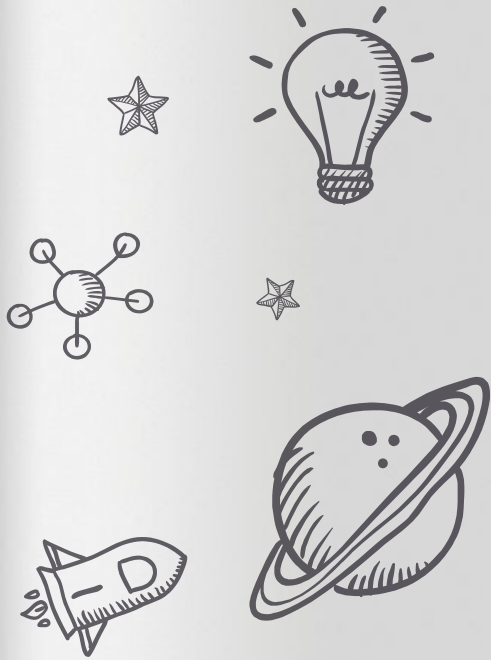


Libraries for Research Data Interest Group



RESEARCH DATA ALLIANCE

<https://www.rd-alliance.org/groups/libraries-research-data.html>



Quest

*to find innovative ways of engaging
researchers with research data*

&

to make these findings accessible to others

Engaging researchers with research data - project in numbers

- ◆ 30 + project members from 3 continents
- ◆ Created a survey to collect case studies
- ◆ Contacted 60 funders, 80 scientific institutions, and used 28 mailing lists
- ◆ 216 responses, 88 described their activities
- ◆ 50 unique scientific institutions
- ◆ 24 cases selected for in-depth interviews and write up

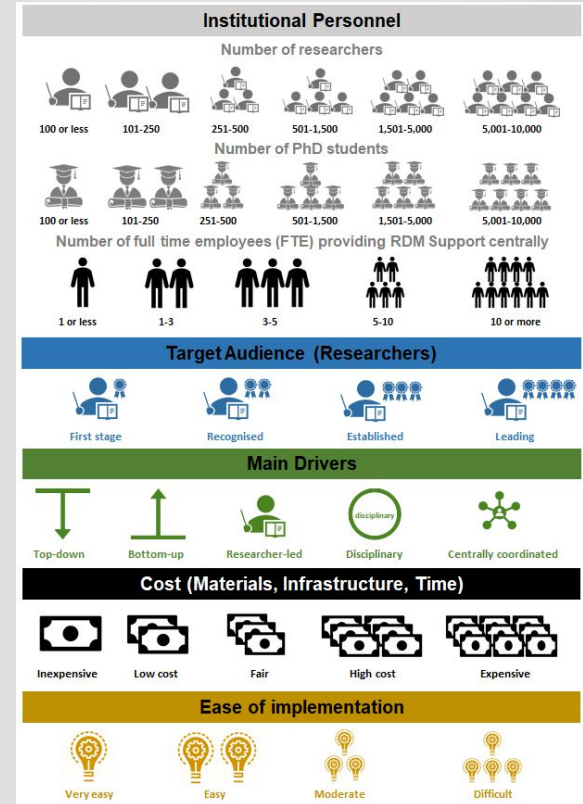
Book sprint - to write up these case studies



<https://www.rd-alliance.org/blogs/book-sprint-success-team-writing-exercise-win.html>

Like in a cookbook:

We have identified the key necessary ingredients for each recipe



Each case has its own ingredient list

No of researchers



No of PhDs



Target audience



Main drivers



Ease of implementation



FTEs for RDM



Costs Materials



Costs Infrastructure



Costs People



Quick preview

“Data Conversations” at Lancaster University

Little time? Little money?... but want a community of researchers talking with passion about data?

Have it all with Data Conversations!

Data managers at Utrecht University

Pool of ‘data managers’ that can be flexibly hired and embedded within research teams to provide highly targeted RDM support.

Engaging with senior researchers at Harvard Business School

Transition to digital data management practices helps engage with researchers who are about to leave the institution.



Engaging Researchers with Data Management: The Cookbook

October 2019

DOI: 10.11647/OBP.0185

<https://www.openbookpublishers.com/product/1080>



That's a lot of 'stuff'

But a lot of things which can be re-used

Summary of today - part 1

- ▣ Starting with RDM:
 - Why RDM and Open Science
 - The key elements of RDM support
 - Data Management Plans
 - DMPonline

Summary of today - part II

- How to engage with researchers?
 - Selfish benefits
 - Case study from Torino
 - Case study from TU Delft
 - More examples: *Engaging Researchers with Data Management: The Cookbook*

Discussion and take-home messages



Go to

www.menti.com

Use the code:

56 72 63

Thanks!

Any questions?

- ▣ @san_ilm & s.illamparuthi@tudelft.nl
- ▣ @shalini_kr & shalini.kurapati@polito.it
- ▣ @martateperek & m.teperek@tudelft.nl