

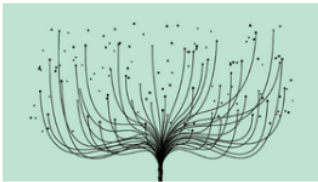
Reproducible research workflows

Agata Bochynska, PhD

Open Research and Digital Scholarship Center
University of Oslo Library

@AgataBochynska

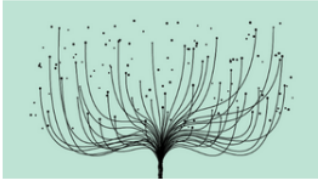
agata.bochynska@ub.uio.no



Time and place: Feb. 3, 2023 10:00 AM–11:30 AM, Zoom

Introduction to open and reproducible research

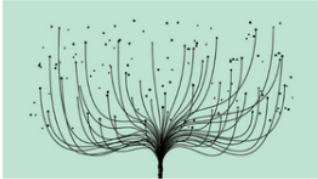
Learn about what open research is and how to make your own research more transparent and reproducible.



Time and place: Feb. 6, 2023 10:00 AM–11:30 AM, Zoom

Preregistration of research studies

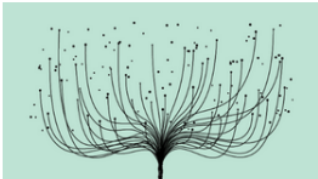
Learn about what preregistration is and how to preregister your own studies.



Time and place: Feb. 8, 2023 10:00 AM–11:30 AM, Zoom

Preregistration on Open Science Framework (OSF)

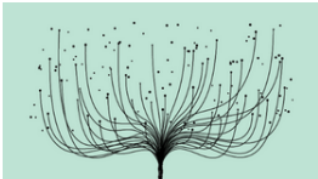
Learn about how to preregister your study on Open Science Framework (OSF) and how to navigate the platform.



Time and place: Feb. 13, 2023 10:00 AM–11:30 AM, Zoom

Reproducible research workflows

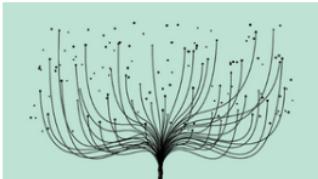
Learn about tools and practices for more reproducible and effective research.



Time and place: Feb. 27, 2023 1:00 PM–2:30 PM, Zoom

Open Access publishing

Learn about Open Access and how can you choose the best way to publish your results openly.



Time and place: Mar. 3, 2023 9:00 AM–10:30 AM, Zoom

Researcher visibility

Learn about different tools, platforms and services to share your research and other contributions, and how you utilise them to make yourself and your work more visible to the academic community and the society at large.

Open and reproducible research courses

Feb 3rd – March 3rd 2023

Roadmap

- Some definitions
- Open research and reproducibility
- Reproducible data acquisition, processing, analyses and reports/publications (with some useful tools)
- Take-aways
- Q&A time!

Reproducibility means
obtaining the same results
with the same data

Replicability means
obtaining similar results
with new data

Reproduced

results are consistent when following the same method and analysis steps with the **same input data**

Replicated

results are consistent across studies that answer the same research question, each of which has obtained **its own data**

Generalized

results apply in **other contexts or populations**
that differ from the original one

		DATA	
		Same	Different
ANALYSIS	Same	Reproduced	Replicated
	Different	Robust	Generalized

Reproducibility means
obtaining the same results
with the same data

Quantitative studies

computational reproducibility

Qualitative studies

process transparency

Quantitative studies

computational reproducibility

“obtaining consistent computational results using the same input data, computational steps, methods, code, and conditions of analysis”

Quantitative studies

computational reproducibility

Re-running analyses/code with the same data

Qualitative studies

process transparency

Arriving at similar (consistent) interpretation by following the same analysis process

Qualitative studies

process transparency

Following the step-by-step reasoning and interpretation process

Reproducibility is strongly
associated with **transparency**

“**Open Science** has the potential of making the scientific process more **transparent, inclusive** and **democratic**. It is (...) a true game changer in bridging the science, technology and innovation gaps and fulfilling the **human right to science**.”



UNESCO Recommendation on Open Science

https://youtu.be/I3Wkvx_ZaFo
<https://www.unesco.org/en/natural-sciences/open-science>

Reproducible does not
(have to) mean fully **open**

As open as possible,
as closed as necessary

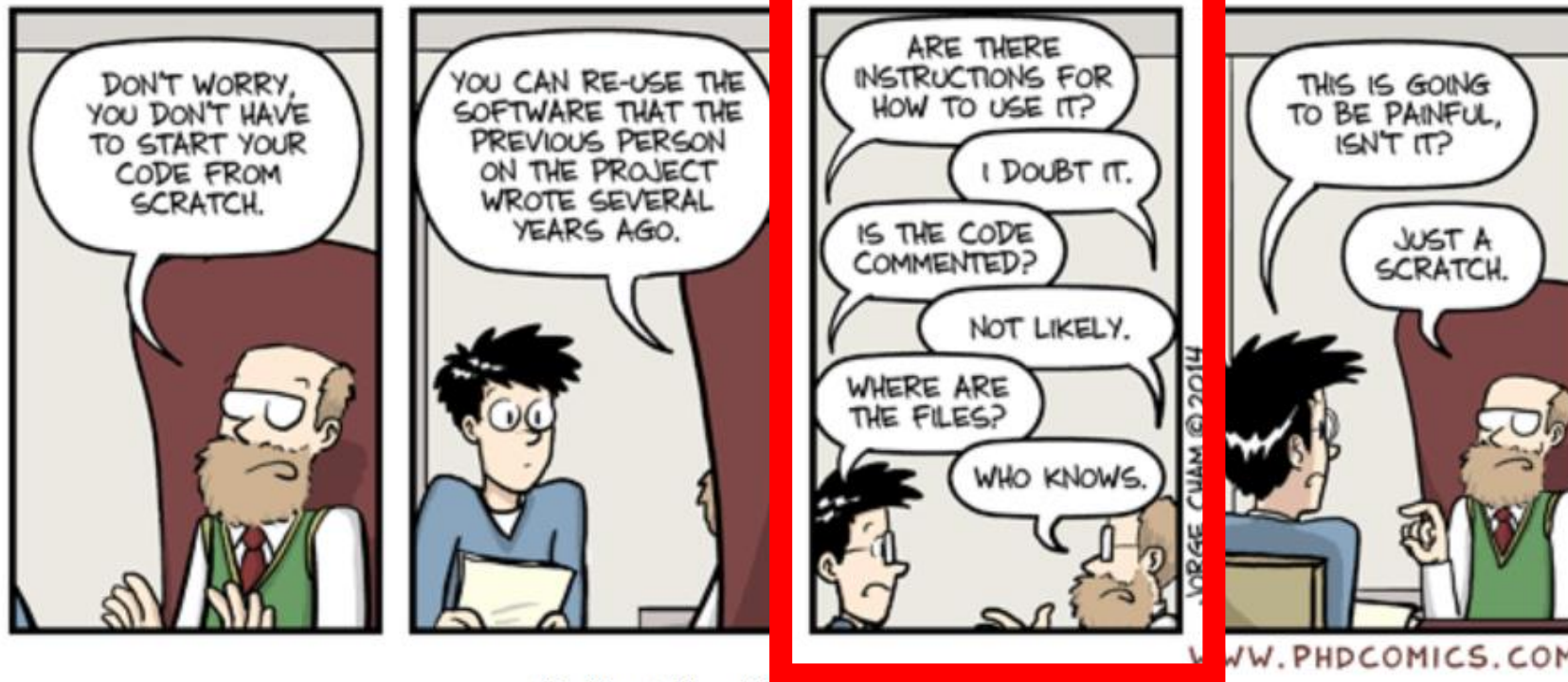
**Making data, tools and
analyses available is
essential for reproducibility**

But not sufficient!

Open but not usable

Piled Higher and Deeper by Jorge Cham

www.phdcomics.com



title: "Scratch" - originally published 3/12/2014

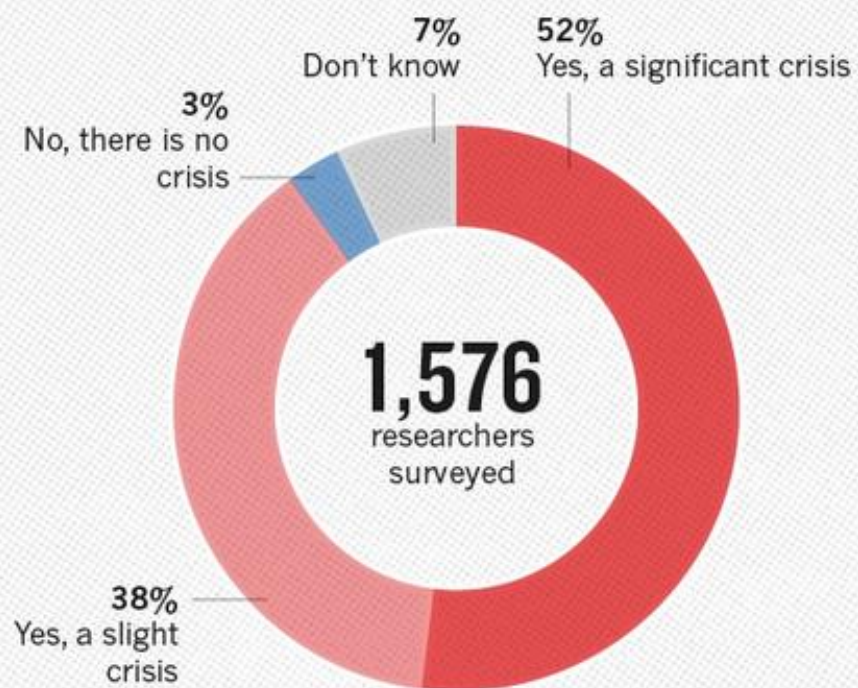
★ Quick access

- Documents
- Downloads
- Pictures
- GEO1010
- Masterseminar 1

Name	Date modified	Type	Size
P-10-01	26.10.2020 16:29	File folder	
SEM 2012	26.10.2020 16:25	File folder	
P-10-03-01.tif	11.06.2012 11:17	TIF File	1 880 KB
P-10-03-02b.tif	11.06.2012 11:25	TIF File	1 880 KB
P-10-03-02c.tif	11.06.2012 11:26	TIF File	1 880 KB
P-10-03-05.tif	11.06.2012 11:38	TIF File	1 880 KB
P-10-03-79b.tif	11.06.2012 12:59	TIF File	1 880 KB
P-10-03-92b.tif	11.06.2012 13:26	TIF File	1 880 KB
P-10-03-109.tif	11.06.2012 13:49	TIF File	1 880 KB
P-10-03-119b.tif	11.06.2012 14:02	TIF File	1 880 KB
P-10-03-193b.tif	11.06.2012 15:45	TIF File	1 880 KB
P-10-03-222b.tif	11.06.2012 16:29	TIF File	1 880 KB
P-10-03-222c.tif	11.06.2012 16:30	TIF File	1 880 KB
P-10-03-226b.tif	12.06.2012 09:33	TIF File	1 880 KB
P-10-03-226c.tif	12.06.2012 09:33	TIF File	1 880 KB
Gasser_2014.pdf	27.01.2015 13:11	Adobe Acrobat D...	5 127 KB
Gernigon_and_Bronner_2012.pdf	30.01.2015 15:07	Adobe Acrobat D...	3 445 KB
Gernigon_et_al_2014.pdf	26.01.2015 16:02	Adobe Acrobat D...	54 447 KB
Glorstad_Clark-2010.pdf	15.09.2014 15:06	Adobe Acrobat D...	10 269 KB
Gronlie et al. 1980 Seismic inversion of Bj...	09.10.2012 13:28	Adobe Acrobat D...	643 KB
Gudlaugsson_et_al_1998.pdf	17.10.2014 12:50	Adobe Acrobat D...	3 411 KB
Harland_and_Gayer_1972.pdf	17.10.2014 11:47	Adobe Acrobat D...	1 474 KB
Høy_and_Lundchien_2011_NBarenstSea...	17.09.2014 14:10	Adobe Acrobat D...	13 600 KB
Isaksen 1996 Organic geocem Bjornoya.p...	09.10.2012 13:24	Adobe Acrobat D...	1 439 KB
Klausen et al 2015 Triassic Snadd in Baren...	30.03.2015 13:15	Adobe Acrobat D...	14 911 KB
Klausen_et_al_2014_Triassic_Snadd_Fm.pdf	08.09.2014 13:20	Adobe Acrobat D...	5 069 KB
Klitzke_et_al_Barents_Sea_Region_2014-pr...	21.01.2015 15:22	Adobe Acrobat D...	2 951 KB
Gabrielsen et al. 1990_Structural_E_BS_NP...	11.03.2013 15:45	Adobe Acrobat D...	9 099 KB
Gac_et_al_2013_ultra_deep_EBB.pdf	20.01.2015 12:15	Adobe Acrobat D...	1 008 KB
Abstract.doc	15.09.2014 16:04	Microsoft Word 9...	234 KB
Copy of NuAge_090214-120115-run1.xls	22.01.2015 13:06	Microsoft Excel 97...	1 390 KB
CPP_Svalbard_Bjornoya_Copy of CUMUL...	17.12.2014 15:35	Microsoft Excel 97...	2 429 KB
2013 Wintermeeting poster-1_Edina_Final...	07.01.2013 15:12	Adobe Acrobat D...	130 KB
ICPMS data	26.10.2020 16:31	File folder	
Franz Josef presentation.pptx	09.01.2013 12:06	Microsoft PowerP...	11 334 KB

Network

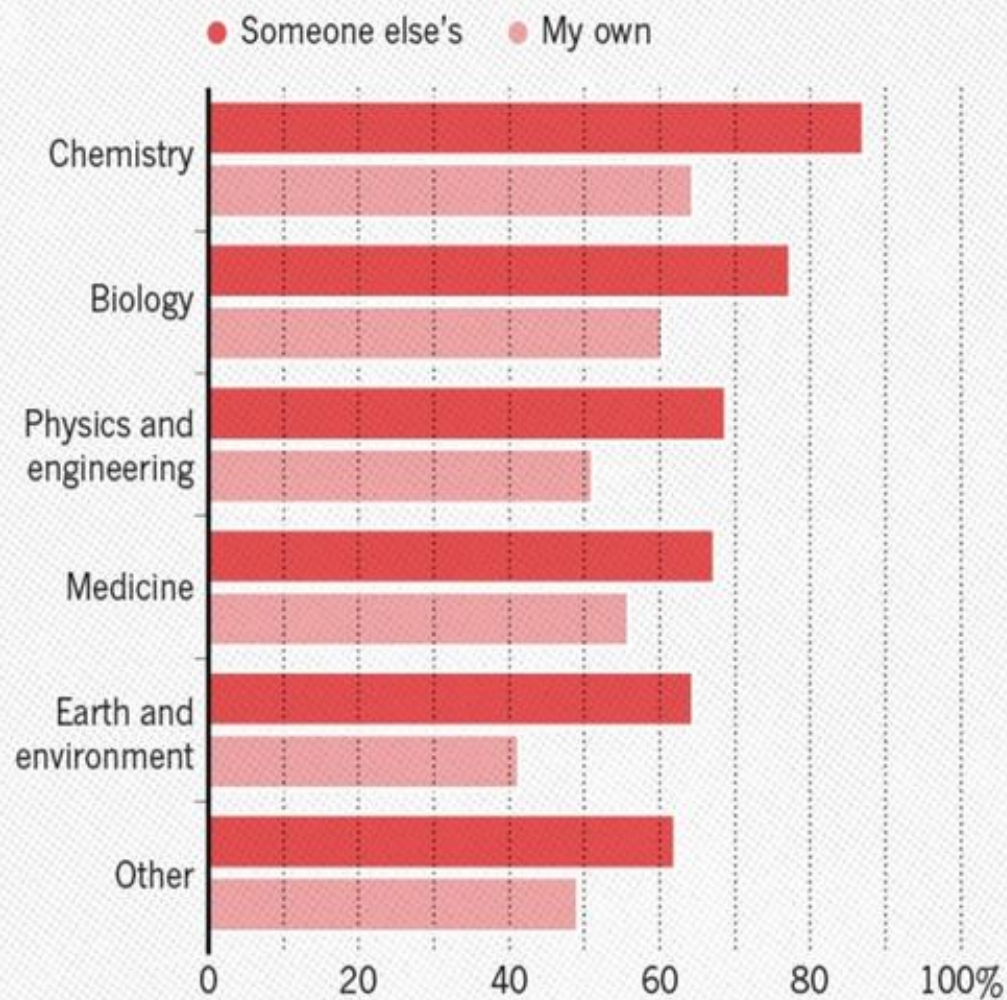
IS THERE A REPRODUCIBILITY CRISIS?



©nature

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



Reasons for irreproducibility:

- Unavailability of materials, data and/or analyses
- Poor data management
- Unclear analysis specification
- Lack of documentation
- Errors in reporting numbers
- Lack of quality checking procedures
- Insufficient peer review

Is science really facing a reproducibility crisis, and do we need it to?

Daniele Fanelli  [Authors Info & Affiliations](#)

Edited by David B. Allison, Indiana University Bloomington, Bloomington, IN, and accepted by Editorial Board Member Susan T. Fiske November 3, 2017 (received for review June 30, 2017)

March 12, 2018 | 115 (11) 2628-2631 | <https://doi.org/10.1073/pnas.1708272114>

Commentary | [Open Access](#) | [Published: 10 February 2022](#)

The reproducibility debate is an opportunity, not a crisis

[Marcus R. Munafò](#) , [Chris Chambers](#), [Alexandra Collins](#), [Laura Fortunato](#) & [Malcolm Macleod](#)

[BMC Research Notes](#) **15**, Article number: 43 (2022) | [Cite this article](#)

2575 Accesses | **6** Citations | **73** Altmetric | [Metrics](#)



The right **mindset** means
you are already halfway
there



Reproducible research workflows

Data acquisition and processing

Data analyses

Data reports (manuscripts)

Data acquisition and processing

Data organization

Data documentation

Version control

Organized data

```
project_name/
├── README.md           # overview of the project
├── data/               # data files used in the project
│   ├── README.md      # describes where data came from
│   └── sub-folder/    # may contain subdirectories
├── processed_data/    # intermediate files from the analysis
├── manuscript/        # manuscript describing the results
├── results/           # results of the analysis (data, tables, figures)
├── src/               # contains all code in the project
│   ├── LICENSE        # license for your code
│   ├── requirements.txt # software requirements and dependencies
│   └── ...
└── doc/               # documentation for your project
    ├── index.rst
    └── ...
```

Research project with a proper data file structure. Image taken from CodeRefinery, Lesson on Reproducible Research. Shared under CC-BY 4.0.

Versioned data

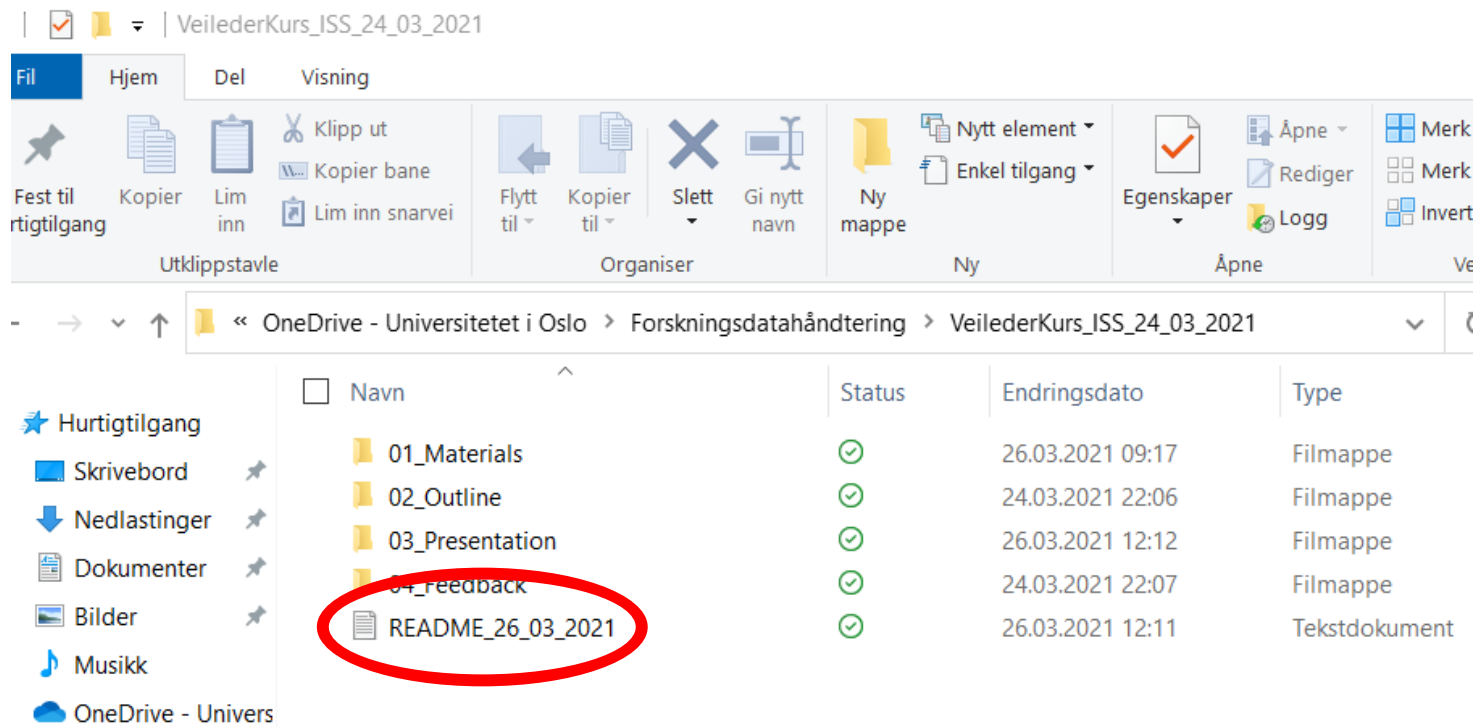
Versioning refers to saving **new copies** of your files when you make changes so that later you can go back and **retrieve** specific **versions** of your files

- DataFileName_1.0 = original document
- DataFileName_1.1 = original document with minor revisions
- DataFileName_2.0 = document with substantial revisions



Documented data: README-files

- The first file to open
- Map for navigating and exploring files and their content
- One README.txt file per folder



Documented data: CODEBOOK

- Explains all variables and their codes in the dataset
- It typically contains:
 - variable names, variable labels, variable codes, variable formats, missing data (in quantitative research)
 - codes, code definitions, examples of what to include with a given code (in qualitative research)
- Can be also called **Data Dictionary**

Example

Sex: biological sex self-reported by the participant (0 – Female, 1 – Male)

Gender: self-reported gender (0 – Female, 1 – Male, 2 – Other)

Age: age in months at the date of test (automatically calculated in the system as the difference between the date of test and the date of birth)

	A	B	C	D	E	F	G
1	subjid	sex	gender	recruitment	condition	group	age
2	10001	1		2	9999	9999	272
3	10002	1	1	1	2	2	187
4	10003	1		1	9999	9999	291
5	10004	1	1	1	1	1	300
6	10005	1	1	1	1	1	288
7	10006	1		6	0	9999	254
8	10007	1	1	4	1	2	233
9	10008	0	2	9999	0	0	266
10	10009	1	1	2	1	1	195
11	10010	1	2	2	1	1	207
12	10011	1	1	1	1	2	234
13	10012	1		2	1	2	231
14	10013	1	1	0	1	2	257
15	10014	0	2	0	0	0	309
16	10015	0	2	1	2	2	254
17	10016	0	2	5	0	9999	161
18	10017	1		6	0	9999	
19	10018	1	1	1	1	2	314
20	10019	1	1	5	0	0	188
21	10020	1		0	1	9999	
22	10021	1	1	4	9999	9999	
23	10022	1		1	1	2	294
24	10023	1	1	0	9999	9999	
25	10024	1	1	1	1	1	291
26	10025	1	1	2	1	1	214
27	10026	1	1	1	2	2	332
28	10027	1	1	0	1	2	348
29	10028	1	1	1	1	2	324
30	10029	0	2	9999	0	0	293

Tools that help: Templates

- Cornell University template and guide to README.txt-files:

<https://data.research.cornell.edu/content/readme>

- README.txt-files from DataverseNO:

[General template](#)

[Example for social sciences](#)

[Example for life sciences](#)

Tools that help: Templates

Take a look, make a copy, edit!

[Template and examples](#)

Tools that help: Nettskjema codebook

[View](#)[Form builder](#)[Codebook](#)[Settings](#)[Collect responses](#)[See results](#)

Codebook

Mapping questions and alternatives to variables is necessary if the results of a survey are going to be processed in an external analysis tool (e.g. SPSS, STATA or R).

[Read more about, and get an introduction to the codebook in Nettskjema](#)

 [Download codebook as text](#)

 [Download codebook as SPSS syntax file](#)

Documentation is
essential for reproducibility

"For my first work-related tweet of 2020, I am totally bummed to announce that we have retracted last year's paper on enzymatic synthesis of beta-lactams. **The work has not been reproducible**"

"It is painful to admit, but important to do so. I apologize to all. I was a bit busy when this was submitted and did not do my job well."

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

BBC Sign in Home News Sport Reel Worklife Tr

NEWS

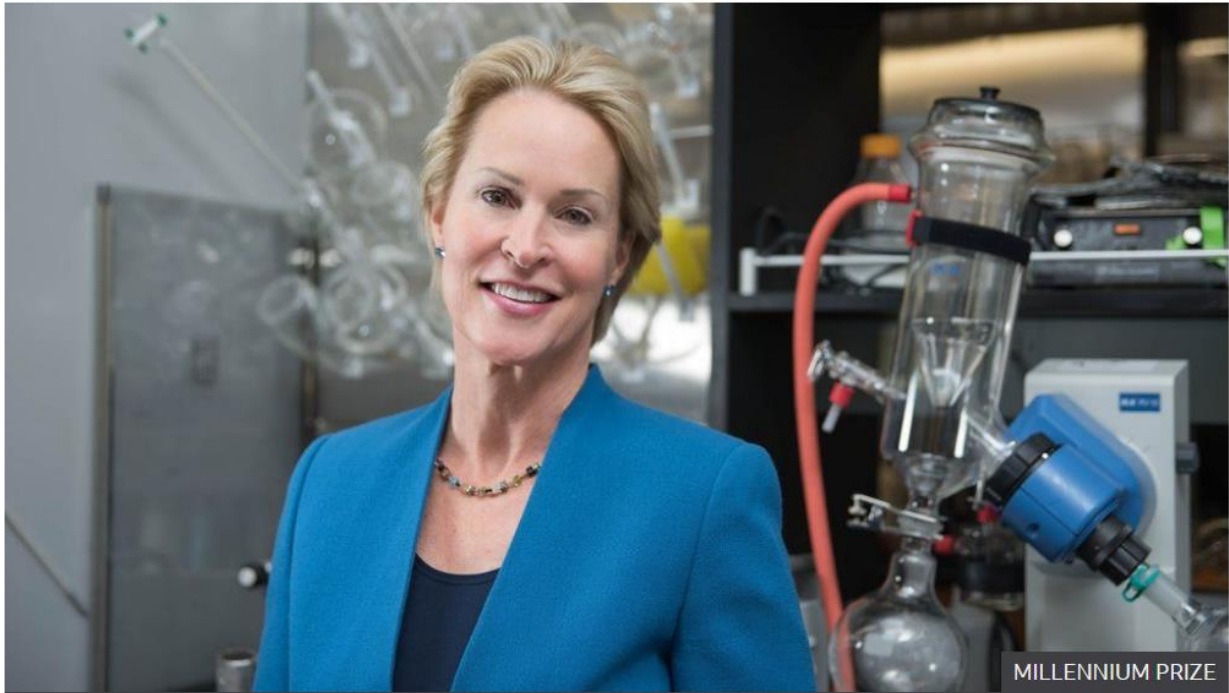
Home Coronavirus Climate Video World UK Business Tech Science Stories Entertainment & Ar

World Africa Asia Australia Europe Latin America Middle East US & Canada

Nobel Prize-winning scientist Frances Arnold retracts paper

🕒 3 January 2020

 Nobel Prize



MILLENNIUM PRIZE

Prof Arnold works in the department of chemical engineering at Caltech

Tools that help: ELN

Electronic Lab Notebooks

help document research, experiments and procedures
performed in laboratories



Courses in research data management and sharing

April 12th – 27rd 2023

Introduction to research data management

Apr. 12, 9:00 AM, Zoom

Data management planning

Apr. 14, 9:00 AM, Zoom

Data organization, metadata, and documentation

Apr. 19, 9:00 AM, Zoom

Data classification and storage selection

Apr. 21, 9:00 AM, Zoom

Copyright and licensing

Apr. 21, 1:00 PM, Zoom

Sharing and archiving research data

Apr. 26, 9:00 AM, Zoom

Finding and reusing research data

Apr. 27, 9:00 AM, Zoom

Data analysis

Step-by-step documentation

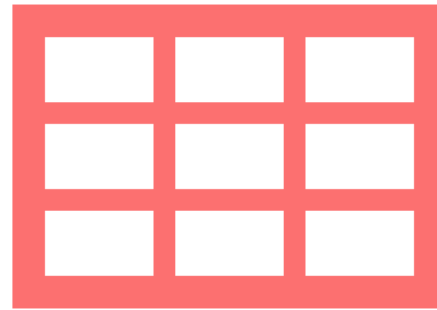
Version control

Cloud computing and/or containers

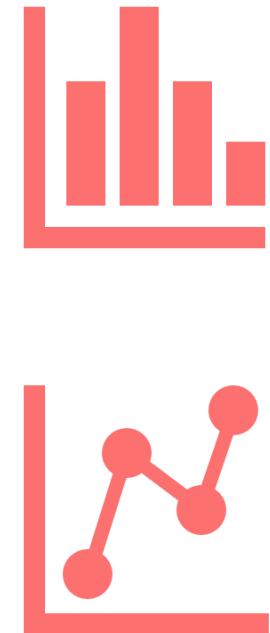
Raw data



Processed data



Analyzed data



Qualitative studies process transparency

Following the step-by-step reasoning and
interpretation process

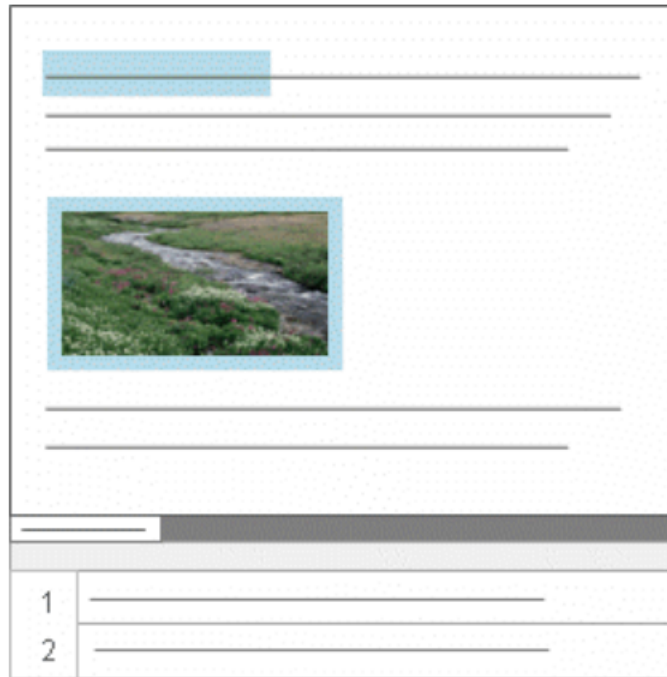
Tools that help: Annotations



Use annotations to comment on selected parts of a source or node

Like scribbled notes in the margin, annotations let you record comments, reminders or observations about specific content in a source or node.

Annotated content is highlighted in blue and the text of the annotation is displayed in the **Annotations** tab at the bottom of the window.



Annotation for Transparent Inquiry (ATI) at a Glance

Annotation for Transparent

▼ Inquiry (ATI)

ATI Models

ATI Instructions

Why ATI?

Empowering Openness in Law-Related Research: A Pilot

Working with Sensitive Research Data (WSRD)

Publications

Annotation for Transparent Inquiry (ATI) facilitates transparency in qualitative research by allowing scholars to “annotate” specific passages in an article. Annotations amplify the text and, when possible, include a link to one or more data sources underlying a claim; data sources are housed in a repository.

Hungary, this was their only stated concern. However, many states conditioned their recognition decision on an action related to Indian troop withdrawal and gave three different types of reasons for doing so. States also differed in the extent of troop withdrawal they required before recognition. See [Table 2](#) for a full list of states, their stated reason for conditioning recognition on withdrawal (if any can be identified), and what recognition was conditioned on (whether actual withdrawal or a proxy).

The first type of reason, opposition to condoning or legitimizing aggression, is labeled as “Non-aggression.” A good example comes from Mexican Foreign Minister Emilio Óscar Rabasa who reported that the Mexican president had decided not to recognize Bangladesh because, “since the Mexicans, like many Latin Americans, refuse to condone territorial aggrandizement as a result of war, they would prefer to wait on the withdrawal of Indian troops as the sign of true independence.”⁹⁰

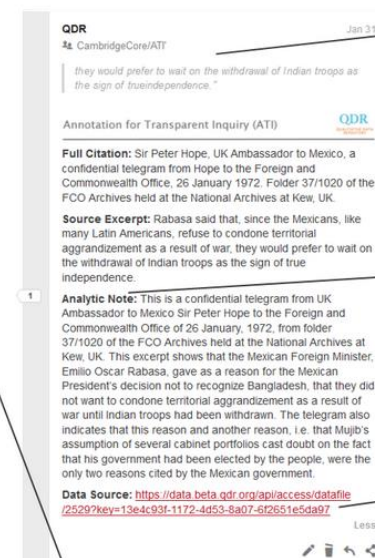
This statement also appeals to “true independence.” Self-determination is another important value expressed by the Mexican representative and is the second type of reason commonly appealed to as justifying recognition as Bangladesh. For

88. See [Figure 2](#).

89. A frequent concern was that states had to recognize in a group, or on the same day as multiple other states. However, even allowing for minor coordination problems, this in and of itself cannot explain the length of time taken to make recognition decisions and declarations.

90. Cable from Hope, 16 January 1972, FCO 37/1020.

Any digitally published manuscript can be annotated using ATI (here: an article in *International Organization* published by Cambridge University Press)



The screenshot displays a QDR interface. At the top, it shows the QDR logo and a date 'Jan 31'. Below this, a text passage is shown: 'they would prefer to wait on the withdrawal of Indian troops as the sign of true independence.' This passage is highlighted in yellow. To the right of the passage, there is a box labeled 'Annotation for Transparent Inquiry (ATI)' which contains the following text: 'Full Citation: Sir Peter Hope, UK Ambassador to Mexico, a confidential telegram from Hope to the Foreign and Commonwealth Office, 26 January 1972, Folder 37/1020 of the FCO Archives held at the National Archives at Kew, UK. Source Excerpt: Rabasa said that, since the Mexicans, like many Latin Americans, refuse to condone territorial aggrandizement as a result of war, they would prefer to wait on the withdrawal of Indian troops as the sign of true independence. Analytic Note: This is a confidential telegram from UK Ambassador to Mexico Sir Peter Hope to the Foreign and Commonwealth Office of 26 January, 1972, from folder 37/1020 of the FCO Archives held at the National Archives at Kew, UK. This excerpt shows that the Mexican Foreign Minister, Emilio Oscar Rabasa, gave as a reason for the Mexican President's decision not to recognize Bangladesh, that they did not want to condone territorial aggrandizement as a result of war until Indian troops had been withdrawn. The telegram also indicates that this reason and another reason, i.e. that Mujib's assumption of several cabinet portfolios cast doubt on the fact that his government had been elected by the people, were the only two reasons cited by the Mexican government. Data Source: https://data.beta.qdr.org/api/access/datafile/r25297key=13e4c93f-1172-4d53-8a07-6f2651ef5da97'. Below the annotation, there is a 'Less' button.

ATI Annotation: Displayed alongside article. Created by author, curated by QDR, hosted and served by Hypothesis, displayed on publisher's web site

Elements of an ATI annotation: One or more of the following:

- Analytic Note
- Source Excerpt
- Source Excerpt Translation
- Link to Data Source

Link to data source housed in QDR

Any passage in the text or in notes of a manuscript can be annotated using ATI

SYMPOSIUM

Active Citation: A Precondition for Replicable Qualitative Research

Andrew Moravcsik, *Princeton University*

Quantitative studies

computational reproducibility

Re-running analyses/code with the same data

Tools that help: analysis via code

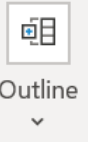
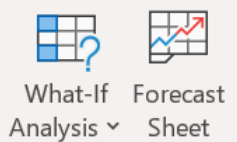
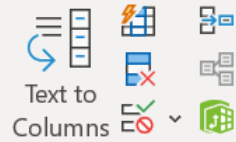
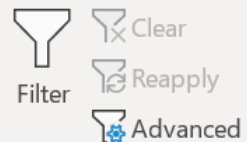
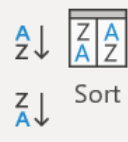
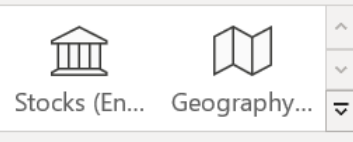


File Home Insert Draw Page Layout Formulas **Data** Review View Help

Share Comments



Refresh
All ▾



Get & Transform Data

Queries & Connections

Data Types

Sort & Filter

Data Tools

Forecast

A1	▼	⋮	✕	✓	f_x	1
----	---	---	---	---	-------	---

[illegible]

```
57 {r}
58 # Main Analysis Data
59
60 #load wide format data and preview
61 sum_data <- read.csv("Data/Experiment2_SumData.csv")
62 head(sum_data)
63
64 #check summary statistics for the dataset
65 describe(sum_data)
66 {r}
```

```
67
68 {r}
69 # Main Analysis
70
71 #ttest on the total proportion looking to shape change against chance (0.5)
72 t.test(sum_data$ShapeProportion, mu=0.5)
73 sd(sum_data$ShapeProportion)
74 se <- sd(sum_data$ShapeProportion)/sqrt(length(sum_data$ShapeProportion))
75 se
76
77 #compute the effect size (Cohen's D)
78 cohensD(sum_data$ShapeProportion, mu=0.5)
79
80
81 # Bayesian ttest on the total proportion looking to shape change against chance (0.5)
82 testMain <- ttestBF(sum_data$ShapeProportion, mu=0.5)
83 testMain
84 sd(sum_data$ShapeProportion)
85 se <- sd(sum_data$ShapeProportion)/sqrt(length(sum_data$ShapeProportion))
86 se
87 {r}
```

Tools that help: version control

- Git: Free and open source version control system
- GitHub: is an internet hosting service for software development and version control using Git





Carpentry@UiO

Carpentry@UiO is a community of people who are passionate about learning, teaching, and sharing best practices and digital skills for making the research process more reproducible and effective. If you want to get involved, or join one of our workshops, check us out!



The Unix Shell

Shell speeds up repetitive and tedious processes. It is also essential skills needed to use high-performance computing (HPC) resources.



Version Control with Git

Git helps you to keep track of what you've done, for a better collaboration and for yourself in future. In the workshop we use GitHub as well.



Programming in Python

Python is now widely used in scientific computing with various powerful packages. Carpentry@UiO runs workshops for participants with no programming experience ("Plotting and Programming in Python" lesson) and for participants at intermediate level ("Programming with Python" lesson, episodes 10-12).



R for Reproducible Scientific Analysis

R is commonly used for statistical analysis, but it is also a powerful programming language. Workshops on R focuses on teaching best practices for scientific computing: breaking down analyses into modular units, task automation, and encapsulation. Workshops on R may use lessons from Data Carpentry instead.



Using Databases and SQL

Databases include powerful tools for search and analysis, and can handle large, complex data sets. The lesson will show how to use a database to explore research data by using SQL.



Carpentry@UiO

Carpentry@UiO is a community of people who are passionate about learning, teaching, and sharing best practices and digital skills for making the research process more reproducible and effective. If you want to get involved, or join one of our workshops, check us out!



Learn, teach, and share digital skills
and best practices

Be a part of an interdisciplinary
community

Make use of and contribute to
community-built teaching materials

Tools that help: shared notebooks



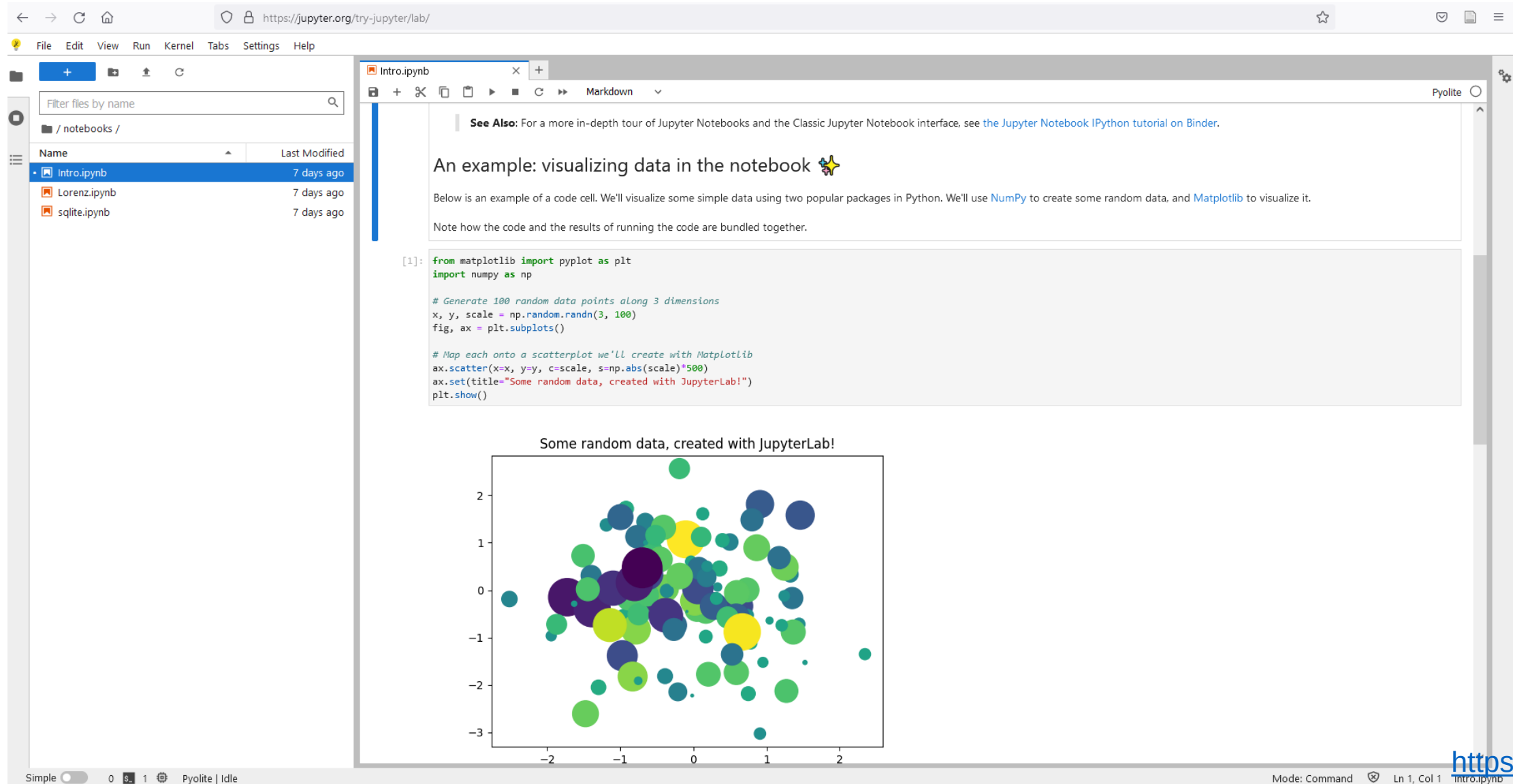
[Try](#) [Install](#) [Get Involved](#) [Documentation](#) [News](#) [Governance](#) [Security](#) [About](#)



Free software, open standards, and web services for interactive computing across all programming languages

<https://jupyter.org/>

Tools that help: shared notebooks



The screenshot displays the JupyterLab web interface in a browser. The address bar shows <https://jupyter.org/try-jupyter/lab/>. The left sidebar contains a file browser with a search bar and a list of notebooks: `Intro.ipynb` (7 days ago), `Lorenz.ipynb` (7 days ago), and `sqlite.ipynb` (7 days ago). The main area shows the `Intro.ipynb` notebook. It begins with a text cell containing a link to a tutorial. Below this is a code cell with the following Python code:

```
[1]: from matplotlib import pyplot as plt
import numpy as np

# Generate 100 random data points along 3 dimensions
x, y, scale = np.random.randn(3, 100)
fig, ax = plt.subplots()

# Map each onto a scatterplot we'll create with Matplotlib
ax.scatter(x=x, y=y, c=scale, s=np.abs(scale)*500)
ax.set(title="Some random data, created with JupyterLab!")
plt.show()
```

The code cell is followed by a scatter plot titled "Some random data, created with JupyterLab!". The plot shows 100 data points in a 2D space, where the color and size of each point represent a third dimension (scale). The points are clustered in the center, with colors ranging from dark purple to bright yellow and sizes varying from small dots to large circles. The x and y axes both range from approximately -2.5 to 2.5.

At the bottom right of the interface, the URL <https://jupyter.org/> is visible.

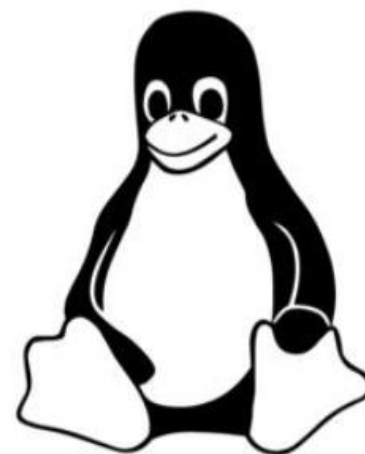
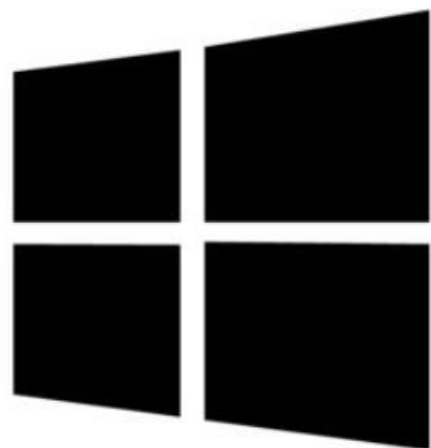
Tools that help: Quality check

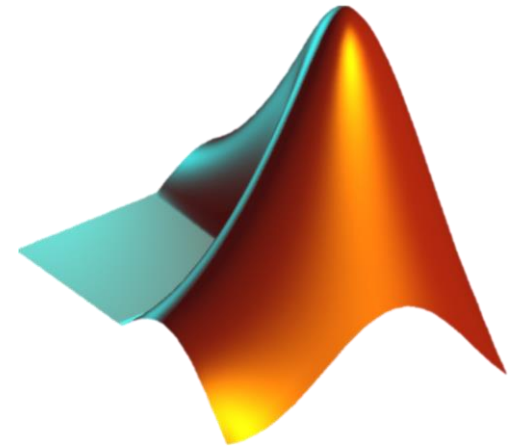
Language	Static code analysis tool	Language	Formatter Tool
Python	Pylint , prospector	Python	Black , yapf
R	lintr	R	formatR
Shell/Bash	shellcheck	Shell/Bash	ShellIndent
		HTML	Tidy

Tools that help: Code review

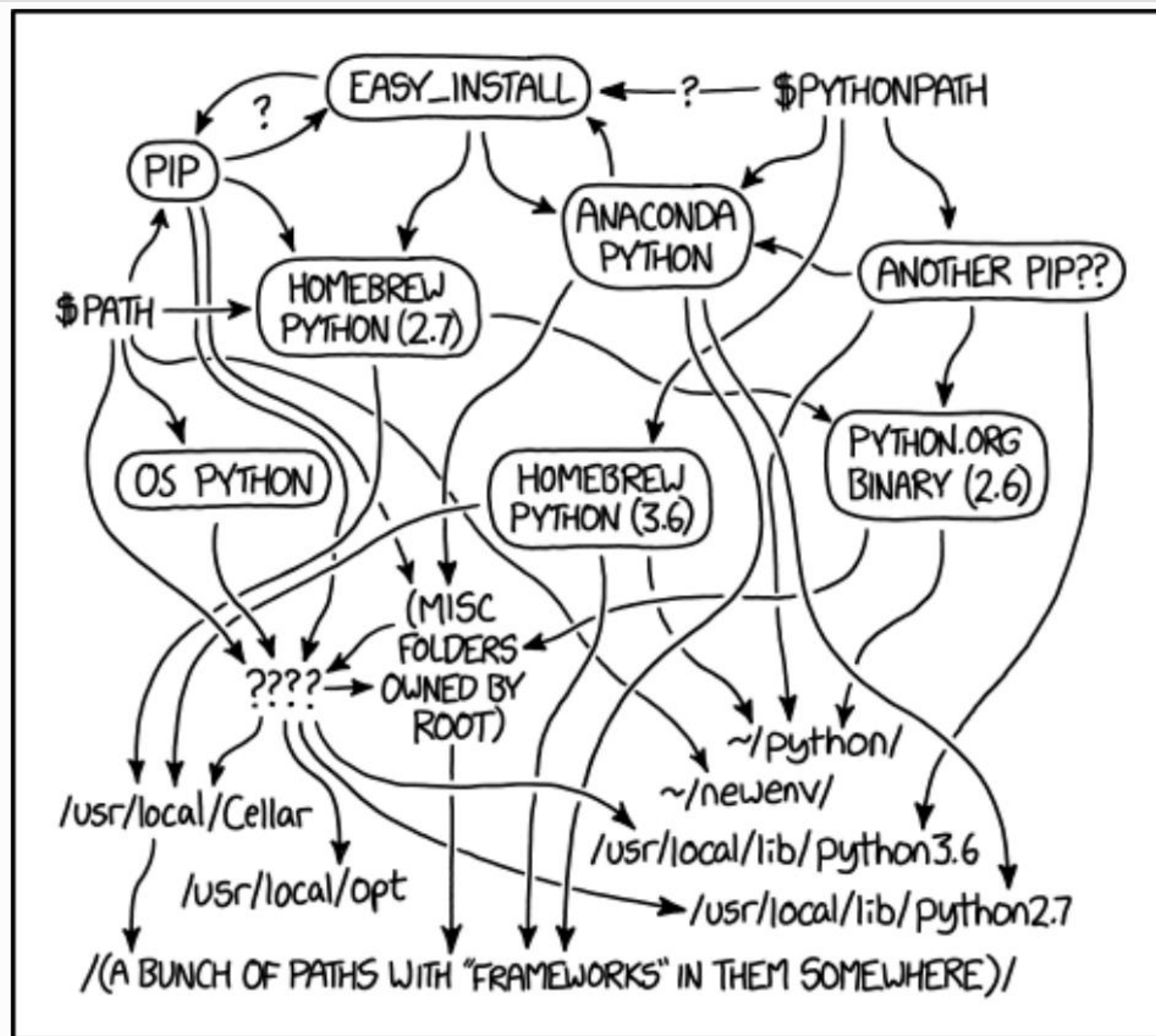


However...









MY PYTHON ENVIRONMENT HAS BECOME SO DEGRADED
THAT MY LAPTOP HAS BEEN DECLARED A SUPERFUND SITE.

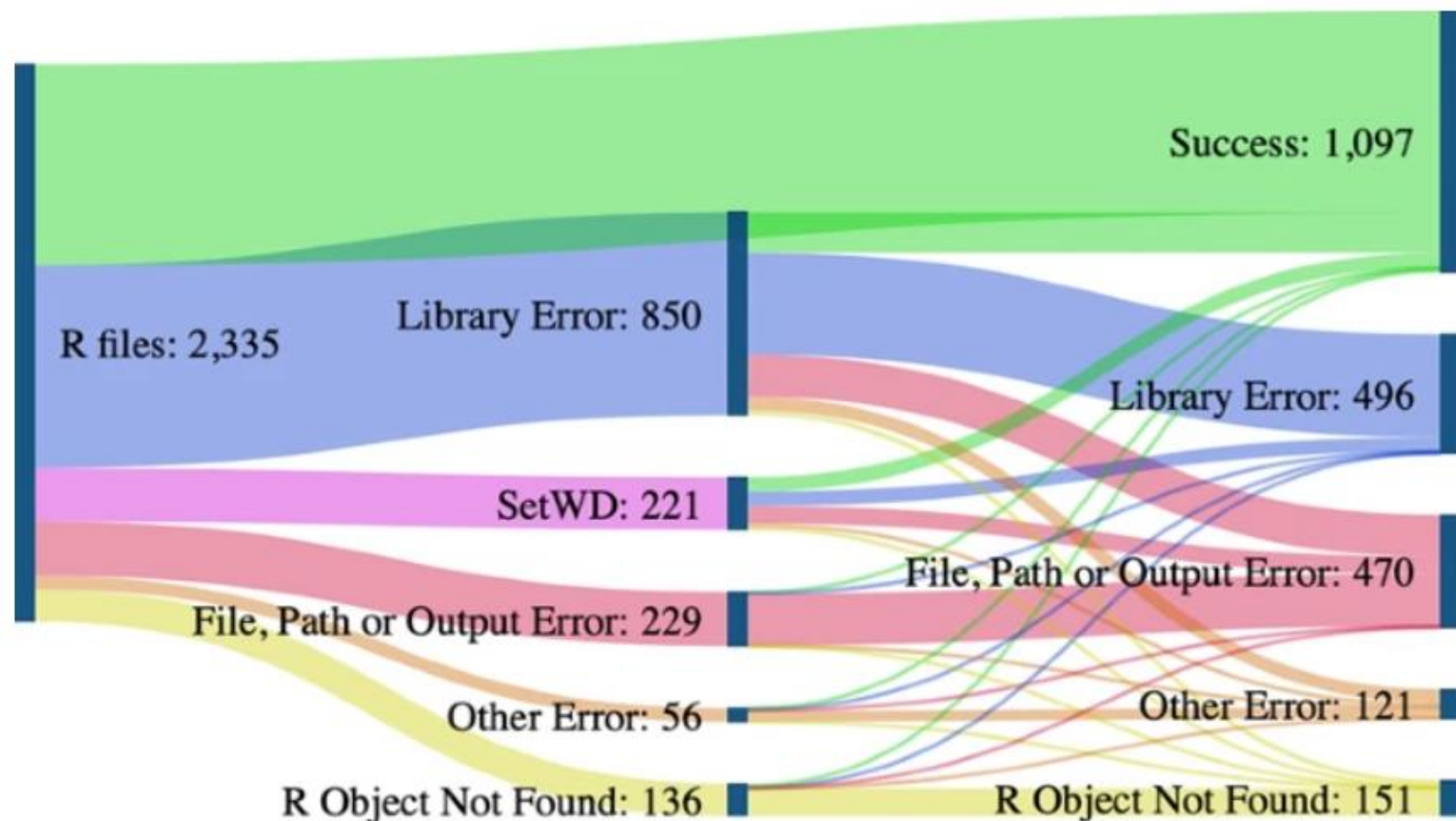
The script throws an
error

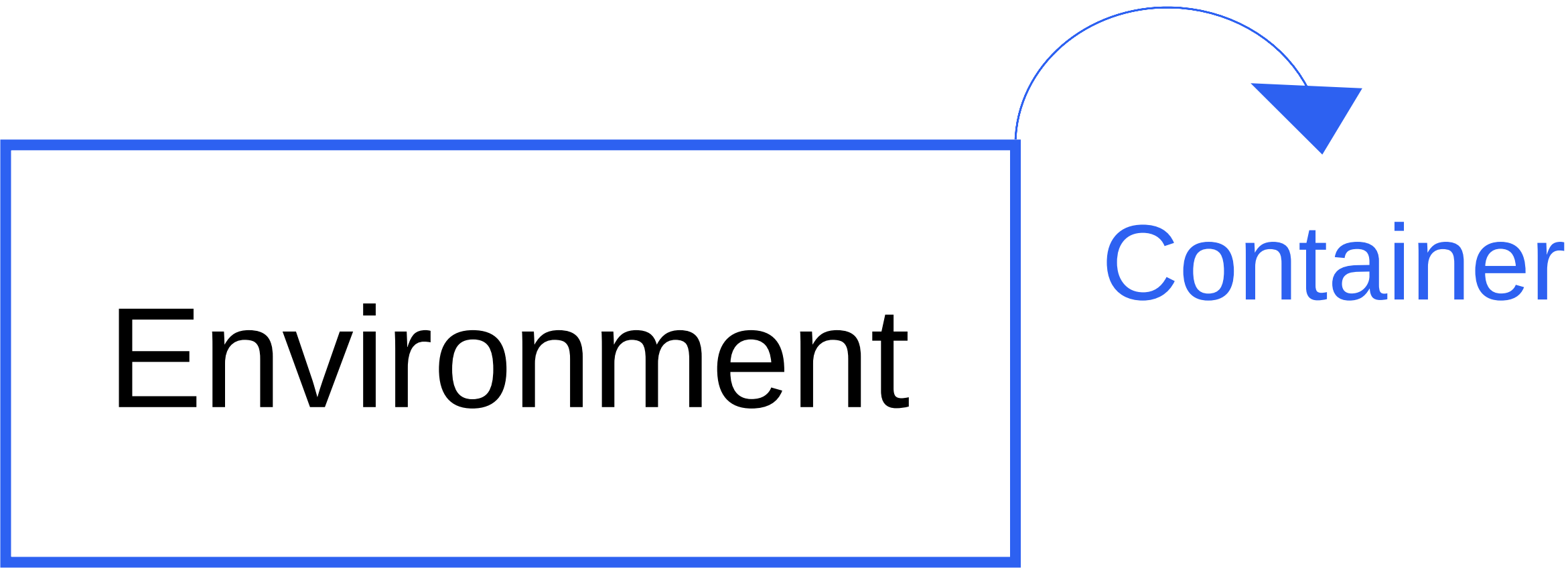
A large-scale study on research code quality and execution

Ana Trisovic , Matthew K. Lau, Thomas Pasquier & Mercè Crosas

Scientific Data **9**, Article number: 60 (2022) |

19k Accesses | **7** Citations | **399** Altmetric





Environment

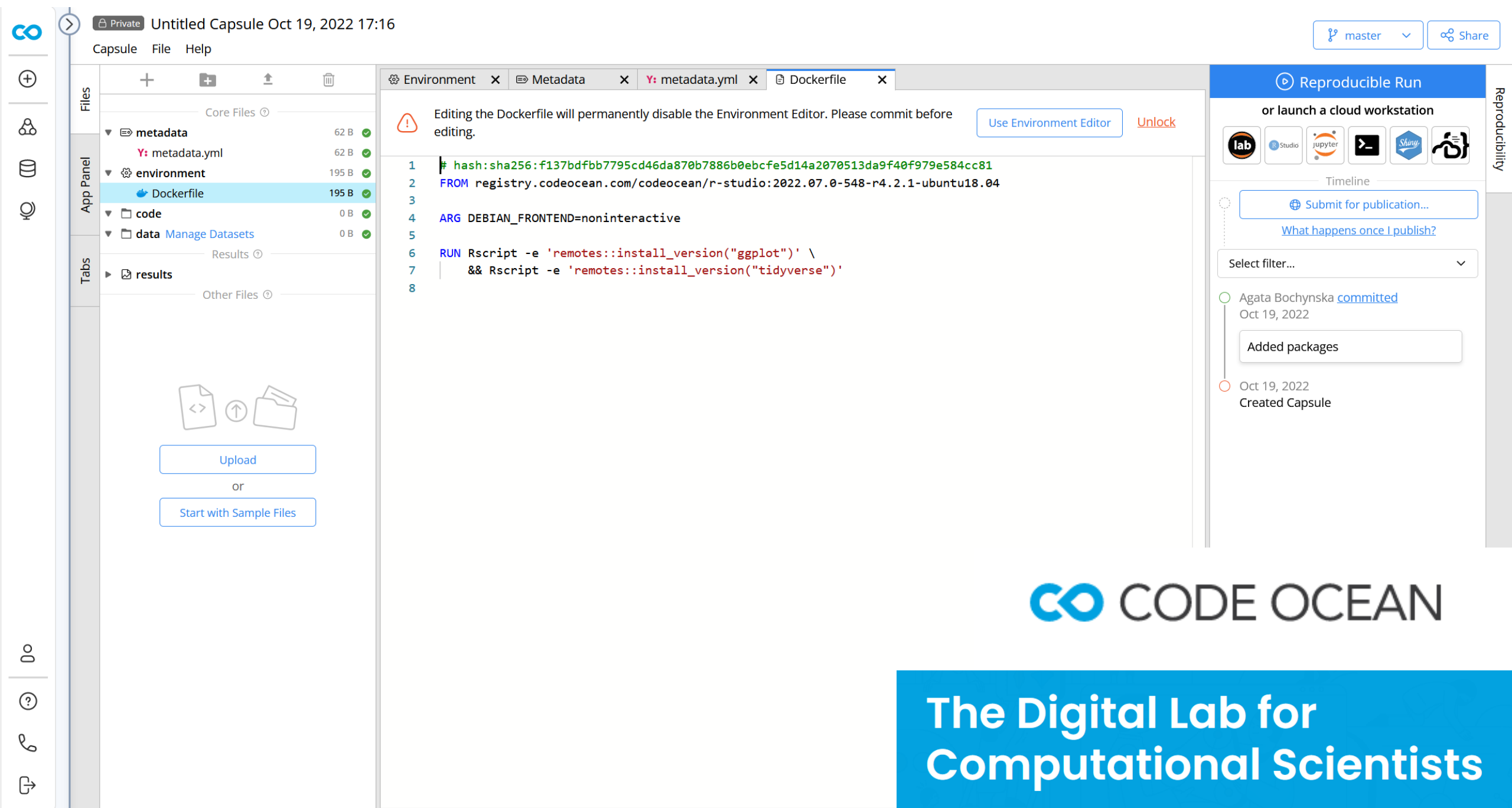
Container

Containers **capture** computational environments

...and make it easier to
reproduce the results

Tools that help: Containers

What is reproduced? Interaction style	Graphical	Command line
Software & versions	 binder	 CONDA
Entire system		 docker



Data reports (manuscripts)

Linked tables and analyses
Version control
Collaboration

Tools that help: R Markdown

R Markdown

from  Studio

[Get Started](#)

[Gallery](#)

[Formats](#)

[Articles](#)

[Book](#)

[References](#)

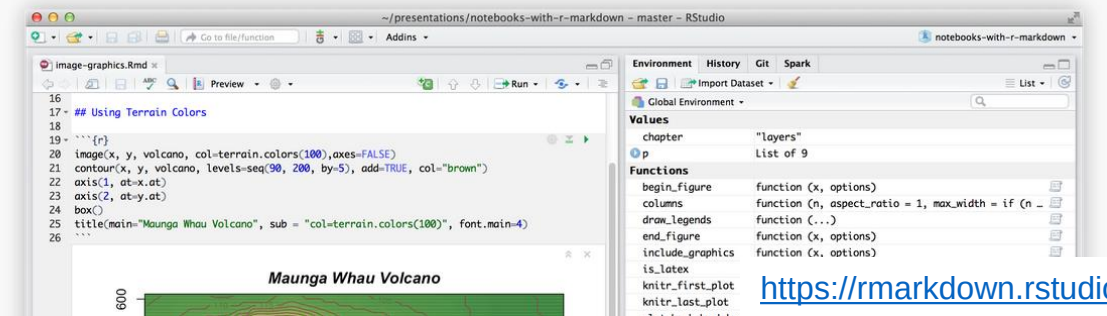


Analyze. Share. Reproduce.

Your data tells a story. Tell it with R Markdown.

Turn your analyses into high quality documents, reports, presentations and dashboards.

R Markdown documents are fully reproducible.
Use a productive [notebook interface](#) to weave together narrative text and code to produce



<https://rmarkdown.rstudio.com/>

Tools that help: R Markdown

Efficient

- everything in one place: write text, cite sources, tidy data, analyse, make tables, and plot it in one programme/screen
- re-do one, more, or all of these easily (e.g. if you need to make changes after peer-review such as excluding subgroups or observations)
- decrease the possibility of making mistakes in the process (the whole manuscript is updated throughout when you make a change in one place)

Flexible

- output to various formats, e.g., HTML, LaTeX, PDF, Word

Open access/source

- use for free (and get help from the user community)
- create documents accessible to anyone with a computer and internet connection

Tools that help: Quarto

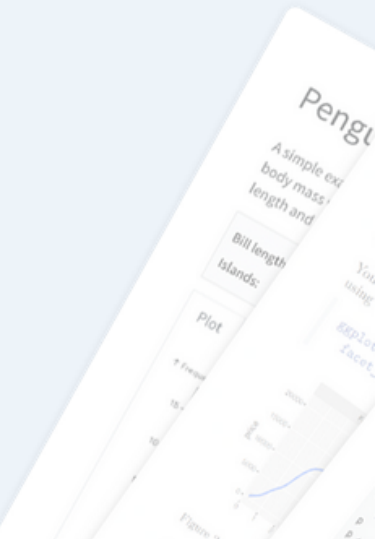


[Overview](#) [Get Started](#) [Guide](#) [Extensions](#) [Reference](#) [Gallery](#) [Blog](#) [Help](#) ▼

Welcome to Quarto

Quarto® is an open-source scientific and technical publishing system built on [Pandoc](#)

- Create dynamic content with [Python](#), [R](#), [Julia](#), and [Observable](#).
- Author documents as plain text markdown or [Jupyter](#) notebooks.
- Publish high-quality articles, reports, presentations, websites, blogs, and books in HTML, PDF, MS Word, ePub, and more.
- Author with scientific markdown, including equations, citations, crossrefs, figure panels, callouts, advanced layout, and more.



Other tools

- [Overleaf](#) (collaborative LaTeX editor)
- [HackMD](#) (a realtime web-based collaborative Markdown editor)
- [Manuscripts.io](#) (a collaborative authoring tool that support scientific content and reproducibility)
- [Rrtools](#) (instructions, templates, and functions for making a basic compendium suitable for writing a reproducible journal article or report with R)
- [Jupyter Notebooks](#) (can be used for supplementary material with journal articles).

Reproducible research workflows

Data acquisition and processing

Data analyses

Data reports (manuscripts)

Take-aways

- Be **transparent** about your full research workflow: research questions, methods, data, step-by-step procedures and analyses
- Make sure you have good **documentation** for all outputs and all stages of your research process
- Keep track of **versions** and do a solid **quality check** of your methods, data and analyses
- **Verify** your own work: try to reproduce your own results and/or have others do it
- Make your methods, data and analyses **open** (if you can)

Thank you!

Questions?

Agata Bochynska, PhD

Open Research and Digital Scholarship Center
University of Oslo Library

@AgataBochynska

agata.bochynska@ub.uio.no

Welcome to Norway's Reproducibility Network

Towards open & reproducible science

JOIN US

Our Mission

The Norwegian Reproducibility Network (NORRN) is a peer-led network that aims **to promote and enable rigorous, robust and transparent research practices in Norway**. We attempt to achieve this goal by establishing appropriate training activities, designing, and evaluating research improvement efforts, disseminating best practices, and working with stakeholders to ensure coordination of efforts across the sector. NORRN's activities span multiple levels, including researchers, librarians, institutions, and other stakeholders (e.g., funders and public authorities).

ReproducibiliTea

Journal Club

**JOIN IN AND DISCUSS WITH FELLOW
STUDENTS AND RESEARCHERS**

**OPEN RESEARCH, REPRODUCIBILITY
and RESEARCH IMPROVEMENT**



Join us

Everyone is welcome to join us - whether you are an enthusiast of open and reproducible research, a skeptic, or a cautious explorer. Currently, all meetings are hybrid with the possibility of joining on-site at Blindern or via Zoom. Grab a cup of tea (coffee?) and join us!

Subscribe to our mailing list





University of Oslo Library

Digital Scholarship Center and Open Research

Open Science Lunch 2023

Each last Thursday of the month at 12.00 we invite you to join us virtually for an online open lunch to hear about how to make your research more open.

February 23rd

Open your data with DataverseNO

March 30th

Public health reporting: an open source approach

April 27th

Open knowledge resources: Store norske leksikon and Wikipedia

May 25th

Keep the rights to your work: UiO's Rights Retention Policy



Upcoming



Time and place: Feb. 23, 2023 12:00 PM–1:00 PM, Hybrid: Georg Sverdrups hus and Zoom

Open your data with DataverseNO

Learn about DataverseNO - a new institutional data archive for open data at UiO.



Time and place: Mar. 30, 2023 12:00 PM–1:00 PM, Hybrid: Georg Sverdrups hus and Zoom

Public health reporting: an open source approach

Join us for a discussion on how open source tools can enable automated, reproducible and scalable public health reporting.



Time and place: Apr. 27, 2023 12:00 PM–1:00 PM, Hybrid: Georg Sverdrups hus and Zoom

Open knowledge resources: Store norske leksikon and Wikipedia

Join us for a discussion on using, maintaining and contributing to freely available open knowledge resources.



Time and place: May 25, 2023 12:00 PM–1:00 PM, Hybrid: Georg Sverdrups hus and Zoom

Keep the rights to your work: UiO's Rights Retention Policy

Learn about how you can retain the rights to your published work with the new Rights Retention Policy at UiO.

[Norwegian
version of this
page](#)

Digital Scholarship Centre

At the Digital Scholarship Centre (DSC) you get guidance on how you can make the best possible use of digital tools and methods in your research and communication activities.

Open Access



Information about open access publishing, publisher agreements, self-archiving, requirements, and guidelines.

Open and reproducible research



Make your research more transparent and reproducible.

Research Data Management



Managing your data both during and after a research project.

Text-mining



Information about digital tools for searching, mining, and analysing textual data.

Systematic search



Information about systematic literature searching, how to get started, and how to get help.

Visualisation



Use of visual methods to explore, communicate and understand data.

Carpentry@UiO



Offers workshops in foundational digital skills such as coding and data management.

Reference management



Styles, tools, and information on reference management.

Open and reproducible research

Learn about how to make your research more open and reproducible and get involved in initiatives and communities that are interested in sharing and improving research at UiO.

More and more researchers and students across disciplines are implementing open research practices, preregistering their hypotheses, methods, and analysis plans and sharing research materials, data and analysis scripts. Digital Scholarship Center can help you learn about and implement these practices in your own research as well as advise on the policies and requirements from funders.

Open Science Lunch



Every last Thursday of the month we meet at noon to discuss topics related to open research.

ReproducibiliTea@UiO



Join us for a Journal Club where we read and discuss papers on open research and meta-science.

Norwegian Reproducibility Network



Join a broader community that aims to promote and enable rigorous, robust and transparent research practices in Norway

Courses and workshops



Click here for the list of upcoming and previous courses and workshops on open and reproducible research at UiO.

[Norwegian
version of this
page](#)

Digital Scholarship Centre

At the Digital Scholarship Centre (DSC) you get guidance on how you can make the best possible use of digital tools and methods in your research and communication activities.

Åpen og reproducerbar forskning | Visualisering | Carpentry@UiO | Åpen tilgang | Tekstutvinning

Senter for Digital Forskerstøtte

Digital Scholarship Centre

DSC NEWS

Desember 2022

Digiforskbloggen | Forskningsdatahåndtering | Systematisk litteratursøk | Referansehåndtering

<https://sympa.uio.no/ub.uio.no/subscribe/dsc-news/subscribe>

More resources:

The Turing Way: Guide for Reproducible Research

<https://the-turing-way.netlify.app/reproducible-research/reproducible-research.html>

CodeRefinery: Reproducible Research

<https://coderefinery.github.io/reproducible-research/motivation/>