Open and reproducible science

WHY AND HOW

Sara Villa

Postdoctoral researcher @Denk's group
The Turing Way Contributor
OLS Open Seeds programme mentor

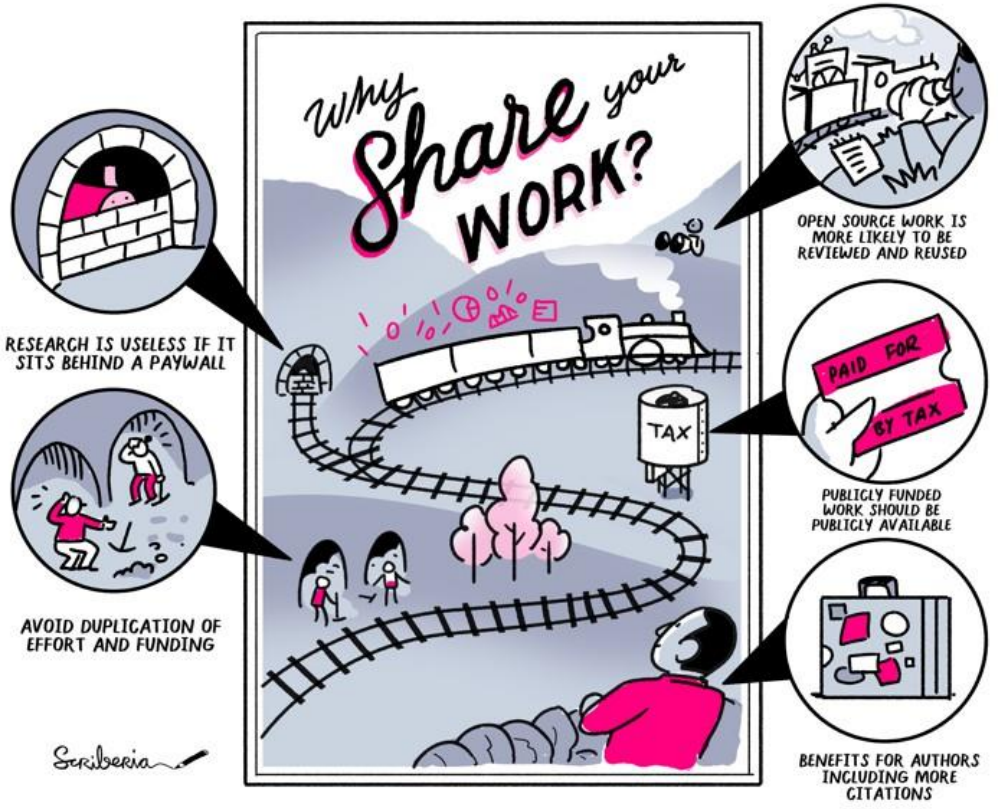
sara.villa@kcl.ac.uk

@VillaScience


OPEN RESEARCH
OPPORTUNITY?
SENSITIVITY?
INCLUSIVITY?
FUNDING?
CLOSED RESEARCH
Scriberia

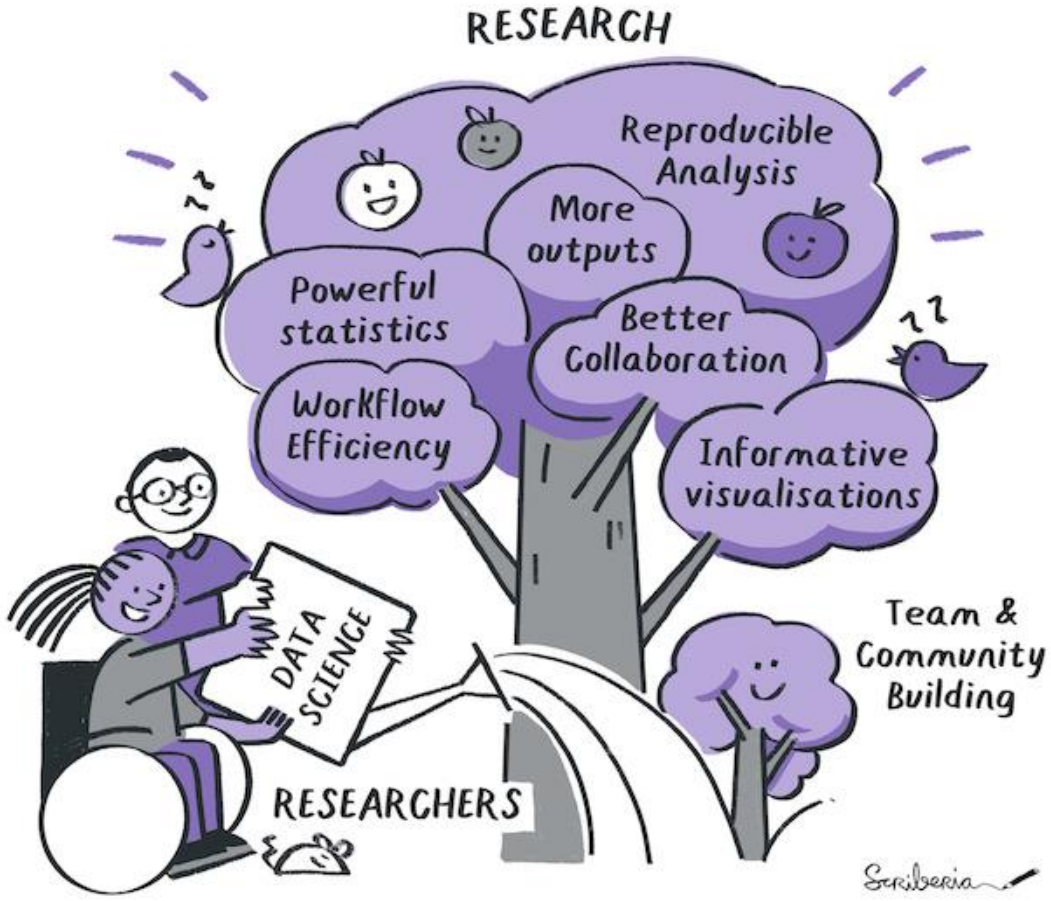
WHY Open science

School	Belief	Aim
Infrastructure	Efficient research depends on the available tools and applications.	Creating openly available platforms, tools, and services for researchers.
Pragmatic	Knowledge-creation could be more efficient if researchers worked together.	Opening up the process of knowledge creation.
Measurement	Academic contributions today need alternative impact measurements.	Developing an alternative metric system for research impact.
Democratic	The access to knowledge is unequally distributed.	Making knowledge freely available for everyone.
Public	Research needs to be made accessible to the public.	Making research accessible for citizens.

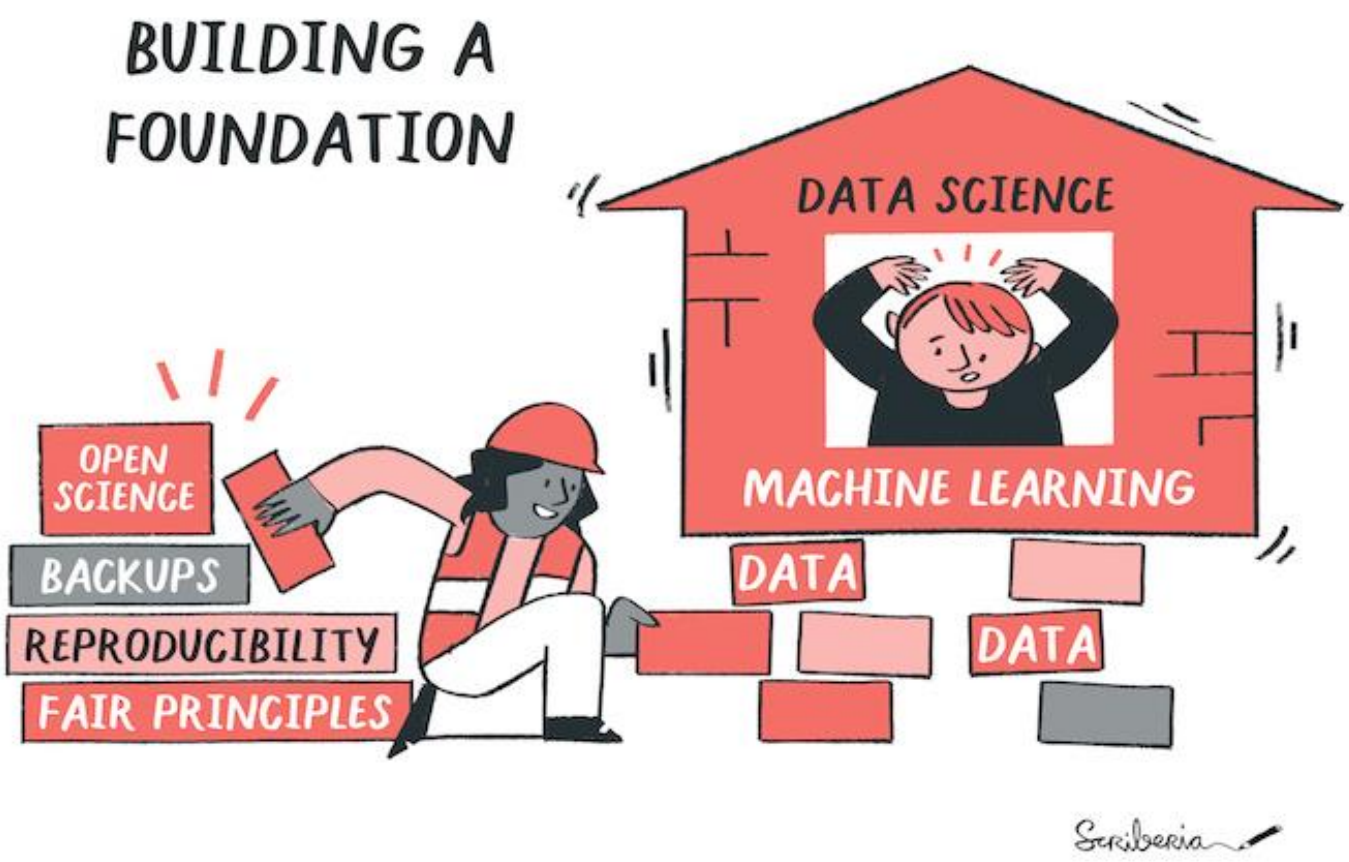


Reinterpreted from https://www.cos.io/blog/cos-celebrates-10-years

Every research project is a data science project



OS principles can be applied to ANY research field

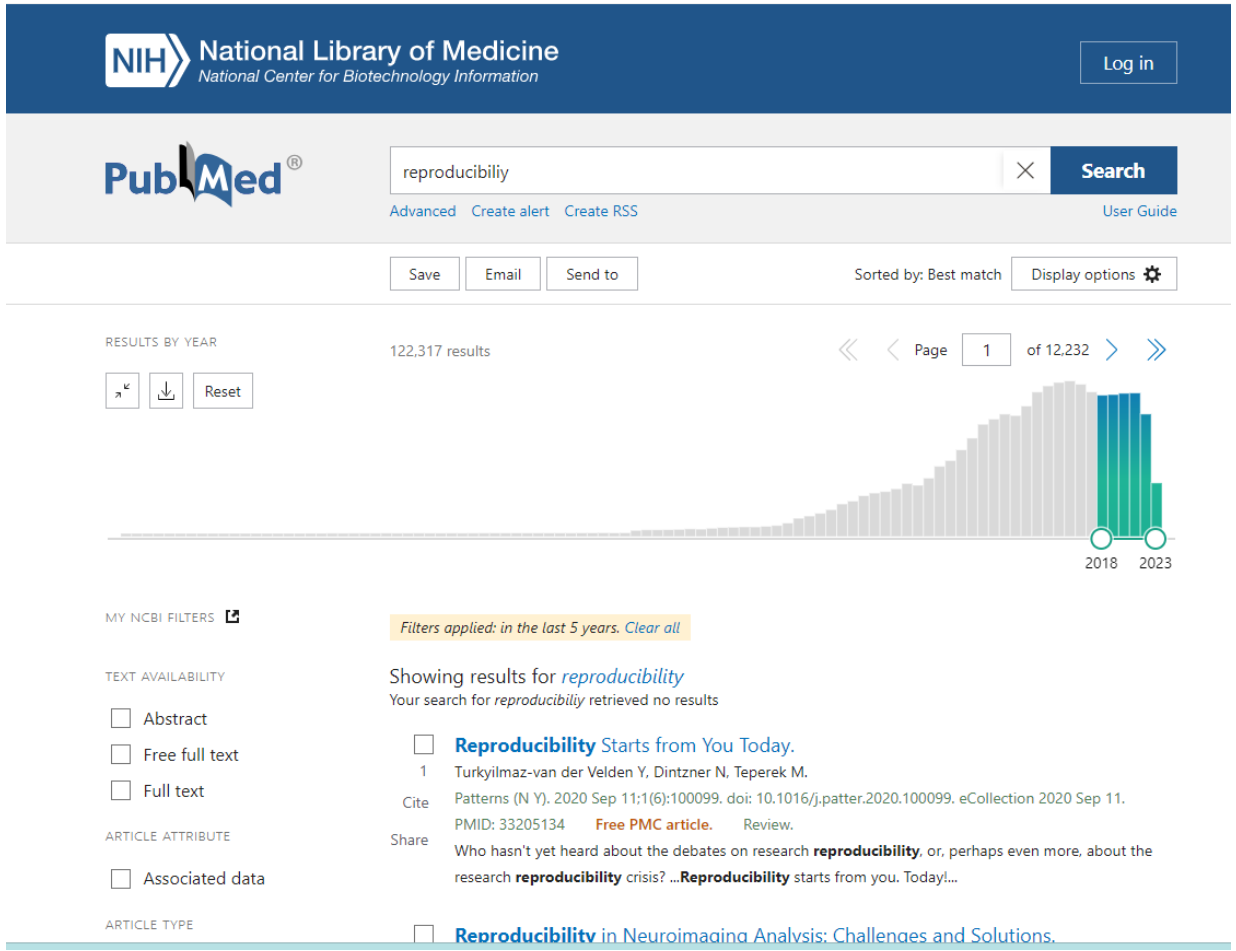


https://rdmkit.elixir-europe.org/

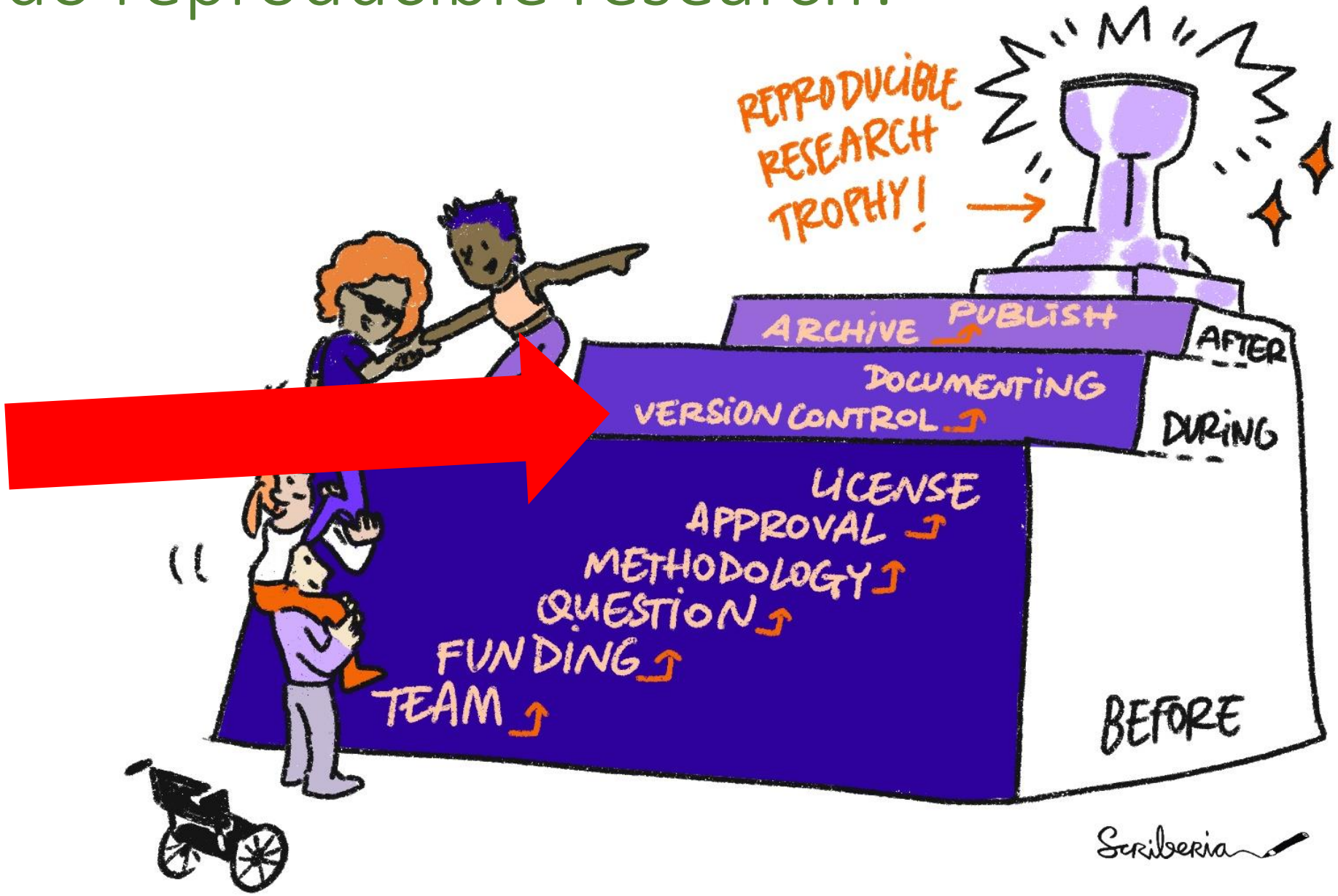
https://osf.io/

https://www.go-fair.org/fair-principles/i1-metadata-use-formal-accessible-shared-broadly-applicable-language-knowledge-representation/

Is reproducibility becoming a buzz word?

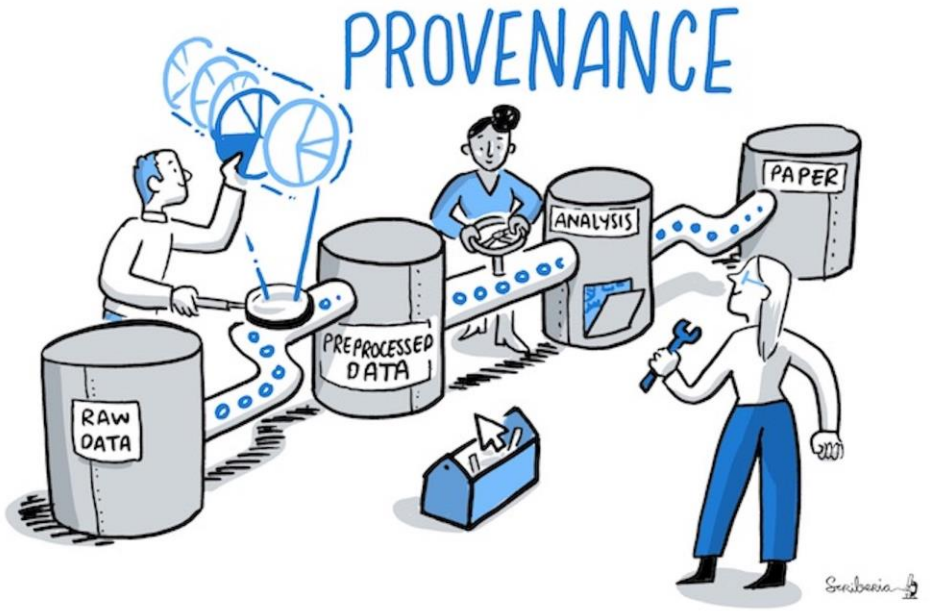


How to do reproducible research?



SPECTRUM OF DATA SCIENCE KNOWLEDGE

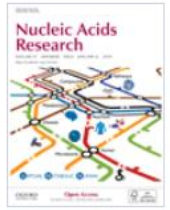
What is reproduced? \ Interaction style	Graphical	Command line
Software & versions	binder	CONDA
Entire system		docker



OXFORD ACADEMIC | Journals | Books

Nucleic Acids Research

Issues | Section browse | More Content | Submit | Purchase | About | Nucleic Acids Research | Advanced Search



Volume 47, Issue D1
08 January 2019

Article Contents

Abstract
INTRODUCTION
INTERACTIVE ANALYSIS REPORTS
DATA MINING
INTEGRATIVE DATA ANALYSIS TOOLS
CONCLUSION
DATA AVAILABILITY
COMPUTATIONAL RESOURCES
SUPPLEMENTARY DATA
ACKNOWLEDGEMENTS

JOURNAL ARTICLE

Stemformatics: visualize and download curated stem cell data

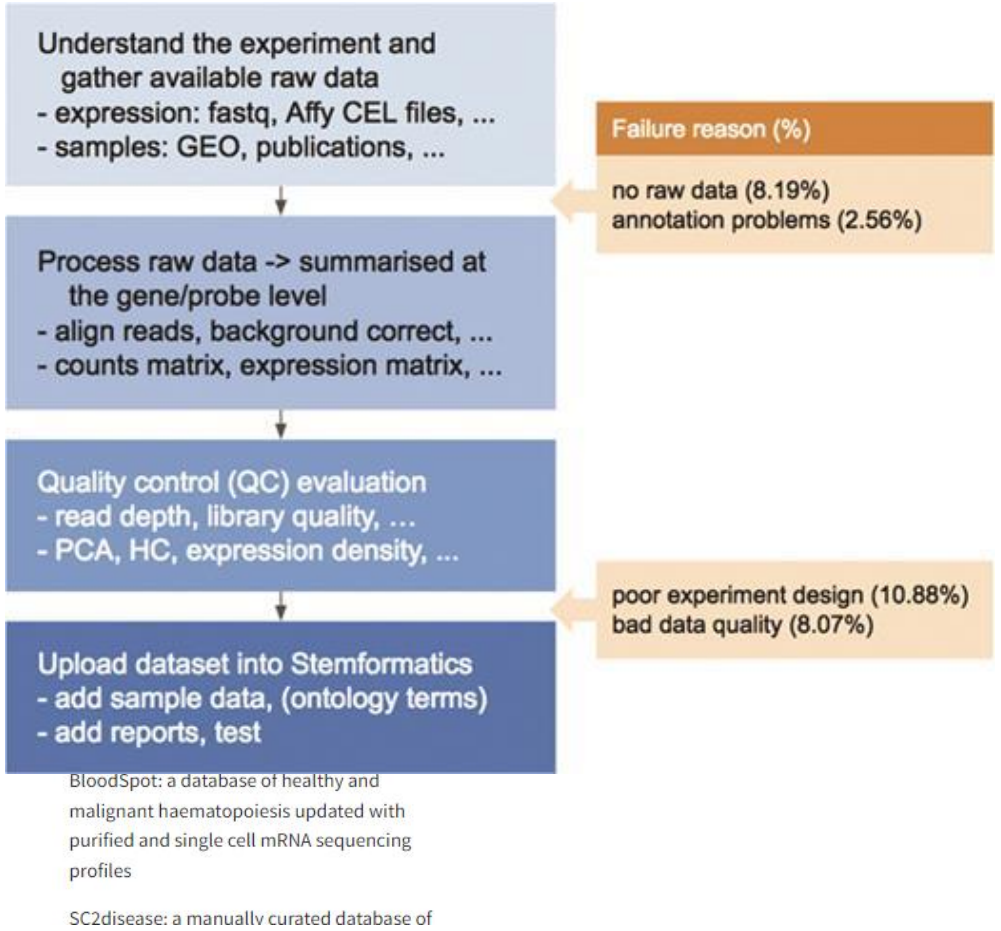
Jarny Choi, Chris M Pacheco, Rowland Mosbergen, Othmar Korn, Tyrone Chen, Isha Nagpal, Steve Englart, Paul W Angel, Christine A Wells

Nucleic Acids Research, Volume 47, Issue D1, 08 January 2019, Pages D841–D846,
https://doi.org/10.1093/nar/gky1064
Published: 08 November 2018 **Article history**

PDF | Split View | Cite | Permissions | Share

Abstract

Stemformatics is an established gene expression data portal containing over 420 public gene expression datasets derived from microarray, RNA sequencing and single cell profiling technologies. Developed for the stem cell community, it has a major focus on pluripotency, tissue stem cells, and staged differentiation. Stemformatics includes curated 'collections' of data relevant to cell reprogramming, as well as hematopoiesis and leukaemia. Rather than simply rehosting datasets as they appear in public repositories, Stemformatics uses a stringent set of quality control metrics and its own pipelines to process handpicked datasets from raw files. This means that about 30% of datasets processed by Stemformatics fail the quality control metrics and never make it to the portal, ensuring that Stemformatics data are of high quality and have been processed in a consistent manner. Stemformatics provides easy-to-use and intuitive tools for biologists to visually explore the data, including



Insufficient sample size, power calculations… are basics often driving to a poor quality study

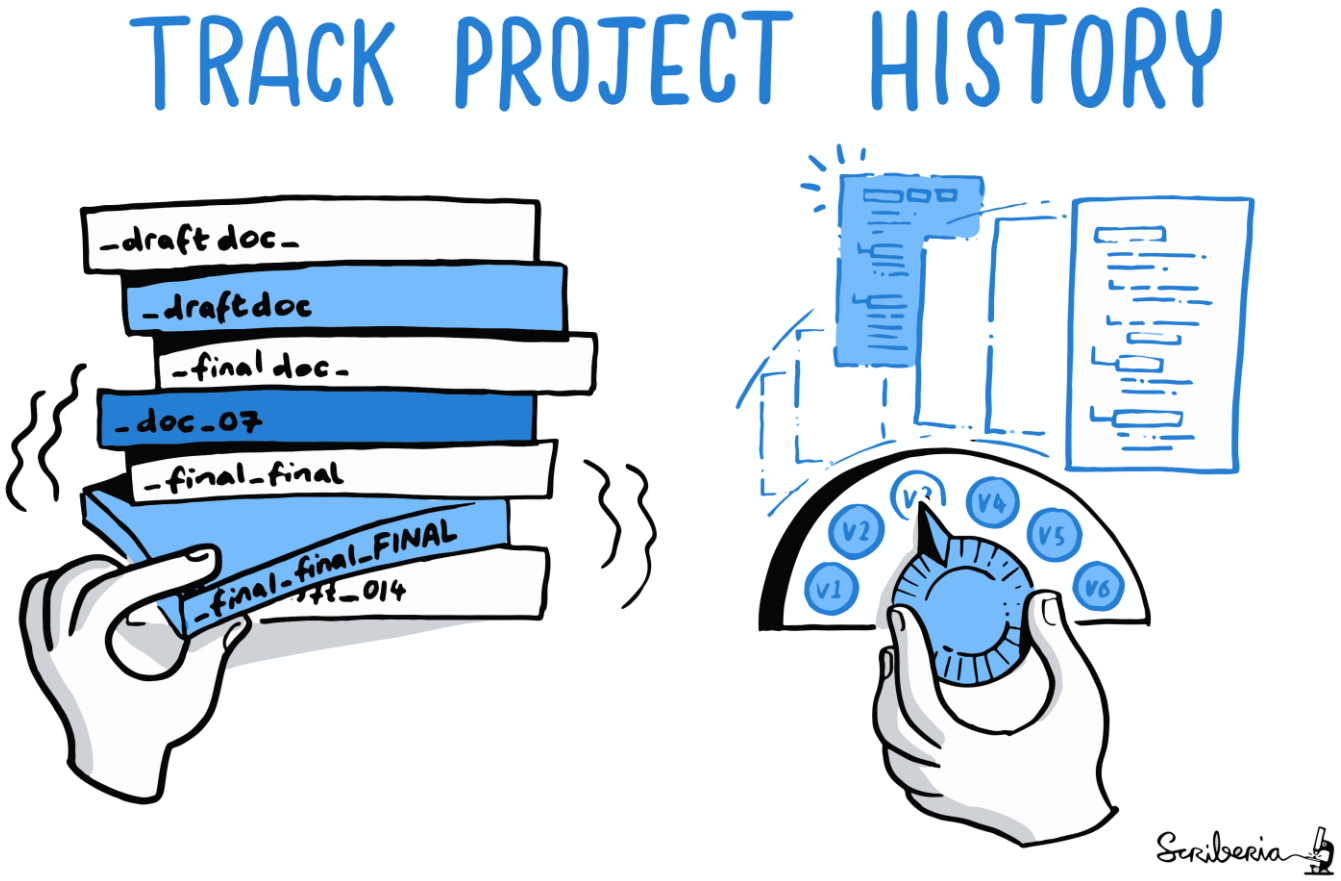
https://lakens.github.io/statistical_inferences/

Feedback from project design and publish negative results too!

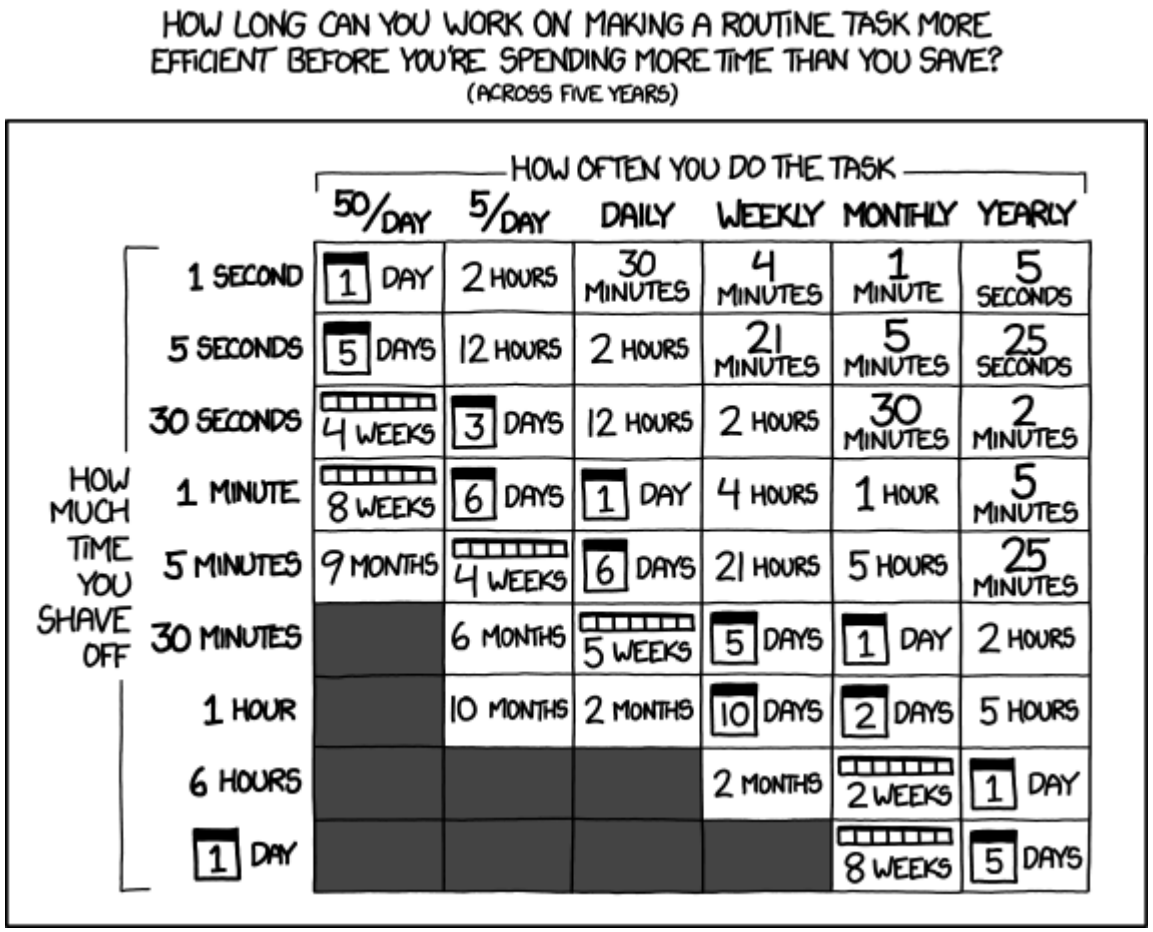
https://www.bnacredibility.org.uk/academia

https://www.cos.io/initiatives/registered-reports

Version control



Learning new tools is worthy for your research most of the times…



https://xkcd.com/1205/

What does a quantitative researcher do?

Read literature

Write papers

Bring in funding

Design experiments

Collect / access data

Set strategy

Mentor colleagues

Run statistical analyses

Promote work to decision makers

Manage a team

Interpret results

@TuringWay @kirstie_j

https://doi.org/10.5281/zenodo.7749650

What does a quantitative researcher do?

Peer review

Read literature

Write papers

Bring in funding

Design experiments

Collect / access data

Testing & quality control

Set strategy

Mentor colleagues

Manage budgets

Run statistical analyses

Promote work to decision makers

Report on progress

Manage a team

Interpret results

@TuringWay @kirstie_j

https://doi.org/10.5281/zenodo.7749650

What does a quantitative researcher do?

Manage (cloud) compute resources

Peer review

Read literature

Deliver trainings

Write papers

Bring in funding

Collect / access data

Design experiments

Set strategy

Testing & quality control

Mentor colleagues

Manage budgets

Run statistical analyses

Promote work to decision makers

Scope user requirements

Generalise code to software

Visualisation

Report on progress

Manage a team

Interpret results

@TuringWay @kirstie_j

https://doi.org/10.5281/zenodo.7749650

What does a quantitative researcher do?

Manage (cloud) compute resources

Peer review

Read literature

Deliver trainings

Archive data & code

Write papers

Bring in funding

Design experiments

Collect / access data

Testing & quality control

Set strategy

Mentor colleagues

Manage budgets

Run statistical analyses

Promote work to decision makers

Scope user requirements

Generalise code to software

Visualisation

Report on progress

Manage a team

Interpret results

@TuringWay @kirstie_j

https://doi.org/10.5281/zenodo.7749650

What does a quantitative researcher do?

Manage (cloud) compute resources

Commercialise a product

Peer review

Read literature

Deliver trainings

Archive data & code

Write papers

Bring in funding

Collect / access data

Design experiments

Set strategy

Testing & quality control

Mentor colleagues

Track impact

Manage budgets

Run statistical analyses

Promote work to decision makers

Scope user requirements

Generalise code to software

Visualisation

Report on progress

Manage a team

Interpret results

@TuringWay @kirstie_j

https://doi.org/10.5281/zenodo.7749650

What does a quantitative researcher do?

Attend / chair meetings

Manage (cloud) compute resources

Commercialise a product

Peer review

Promote EDI and work-life balance

Read literature

Deliver trainings

Archive data & code

Write papers

Bring in funding

Give talks

Design experiments

Share work with the public

Collect / access data

Testing & quality control

Set strategy

Maintain ethical & legal principles

Mentor colleagues

Track impact

Manage budgets

Run statistical analyses

Promote work to decision makers

Host events

Scope user requirements

Generalise code to software

Visualisation

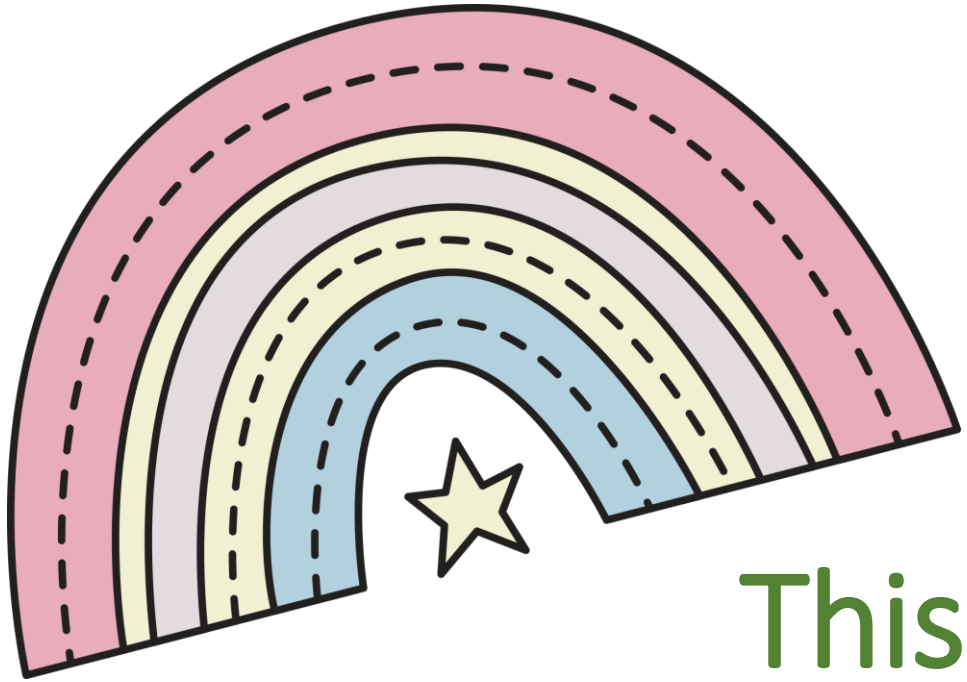
Report on progress

Manage a team

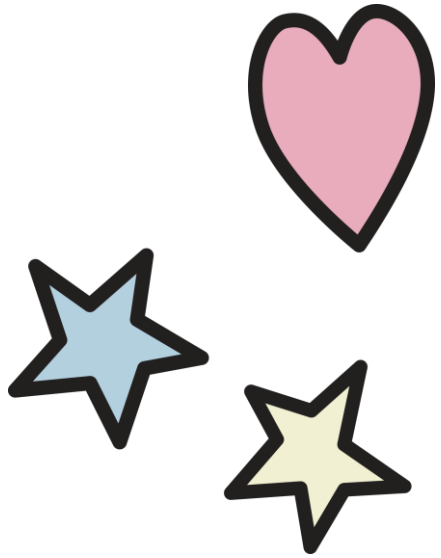
Interpret results

@TuringWay @kirstie_j

https://doi.org/10.5281/zenodo.7749650



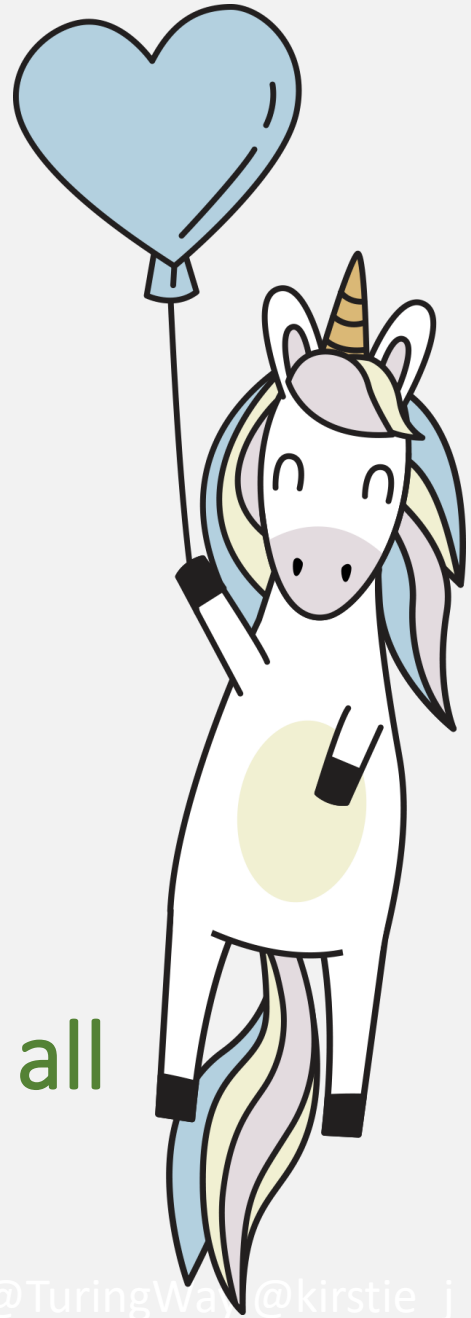
This person is a unicorn



They are delightful

But they don’t exist

We can’t expect individuals to be able to do all these tasks to the highest standard



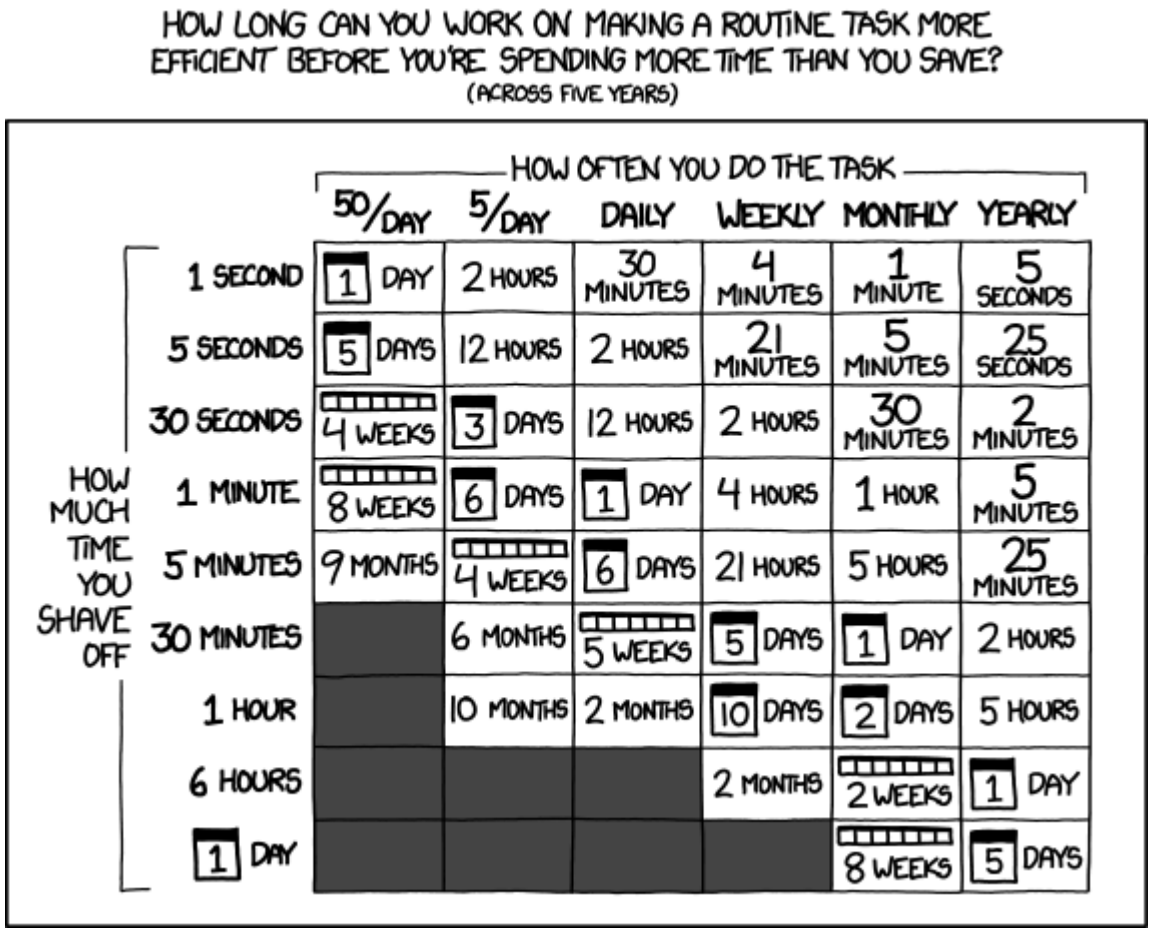
https://doi.org/10.5281/zenodo.7749650

@TuringWay @kirstie_j

https://doi.org/10.5281/zenodo.7749650

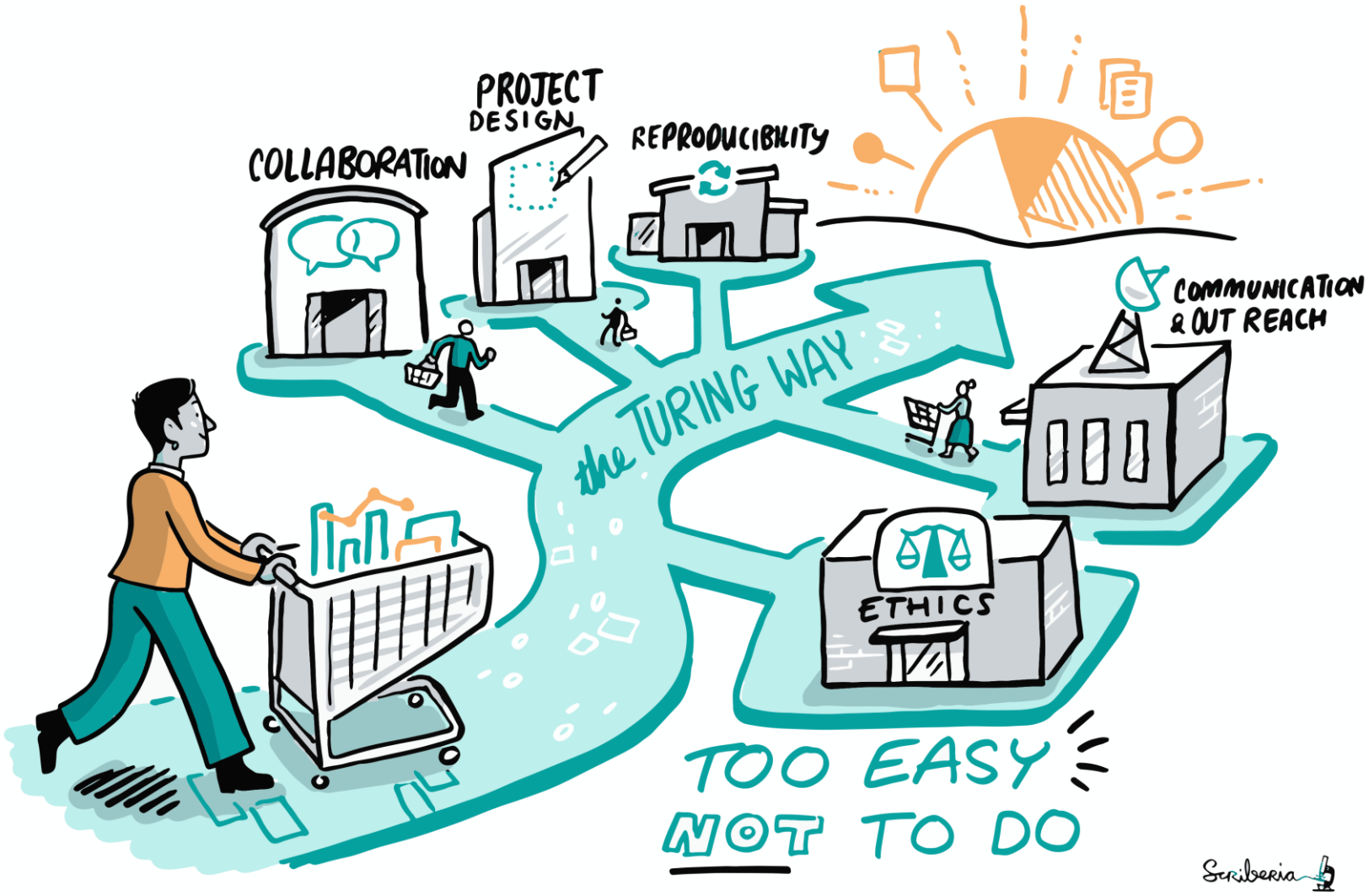
There is one way…

Learning new tools might be worthy… or it might not sometimes

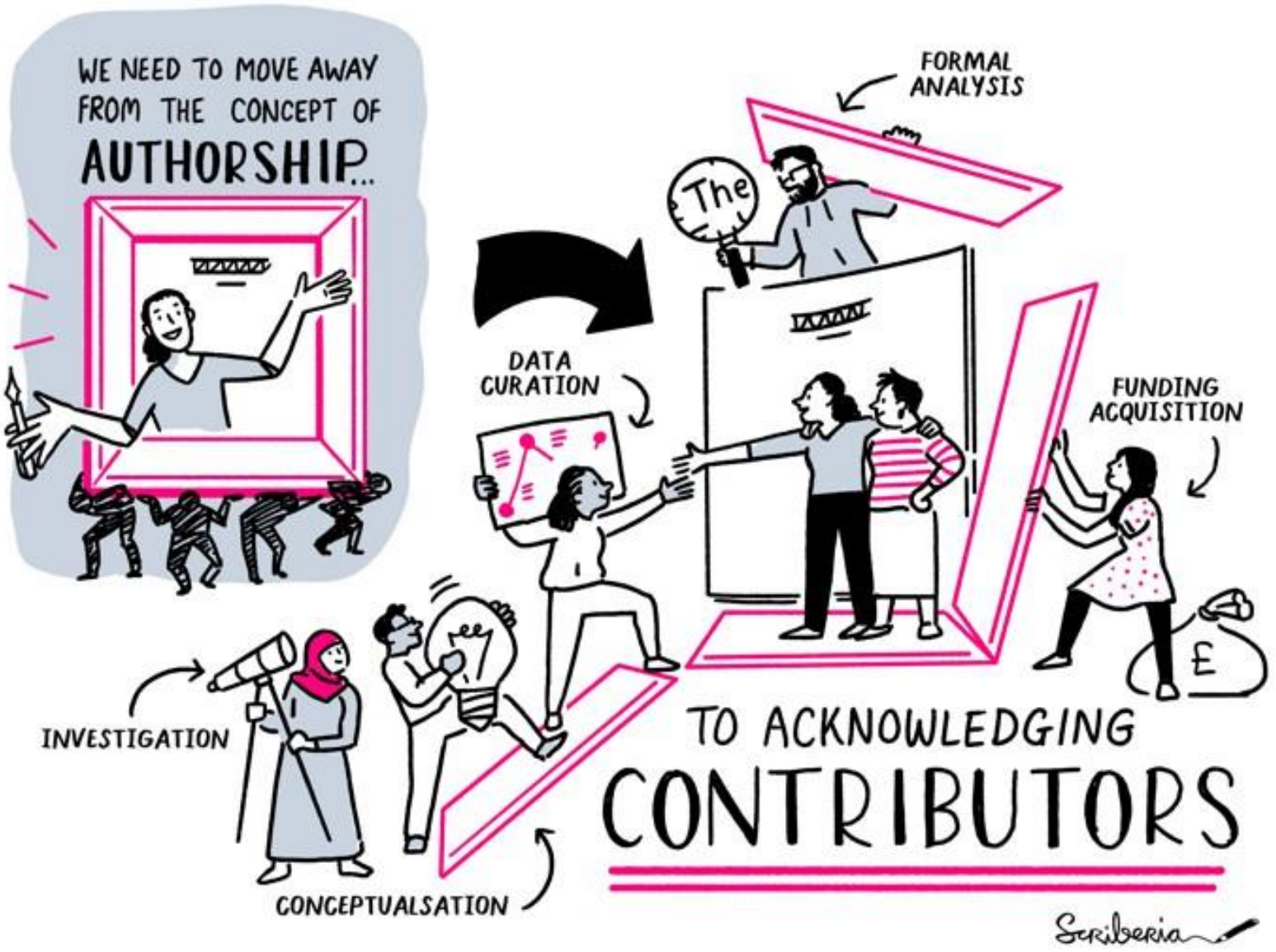


https://xkcd.com/1205/

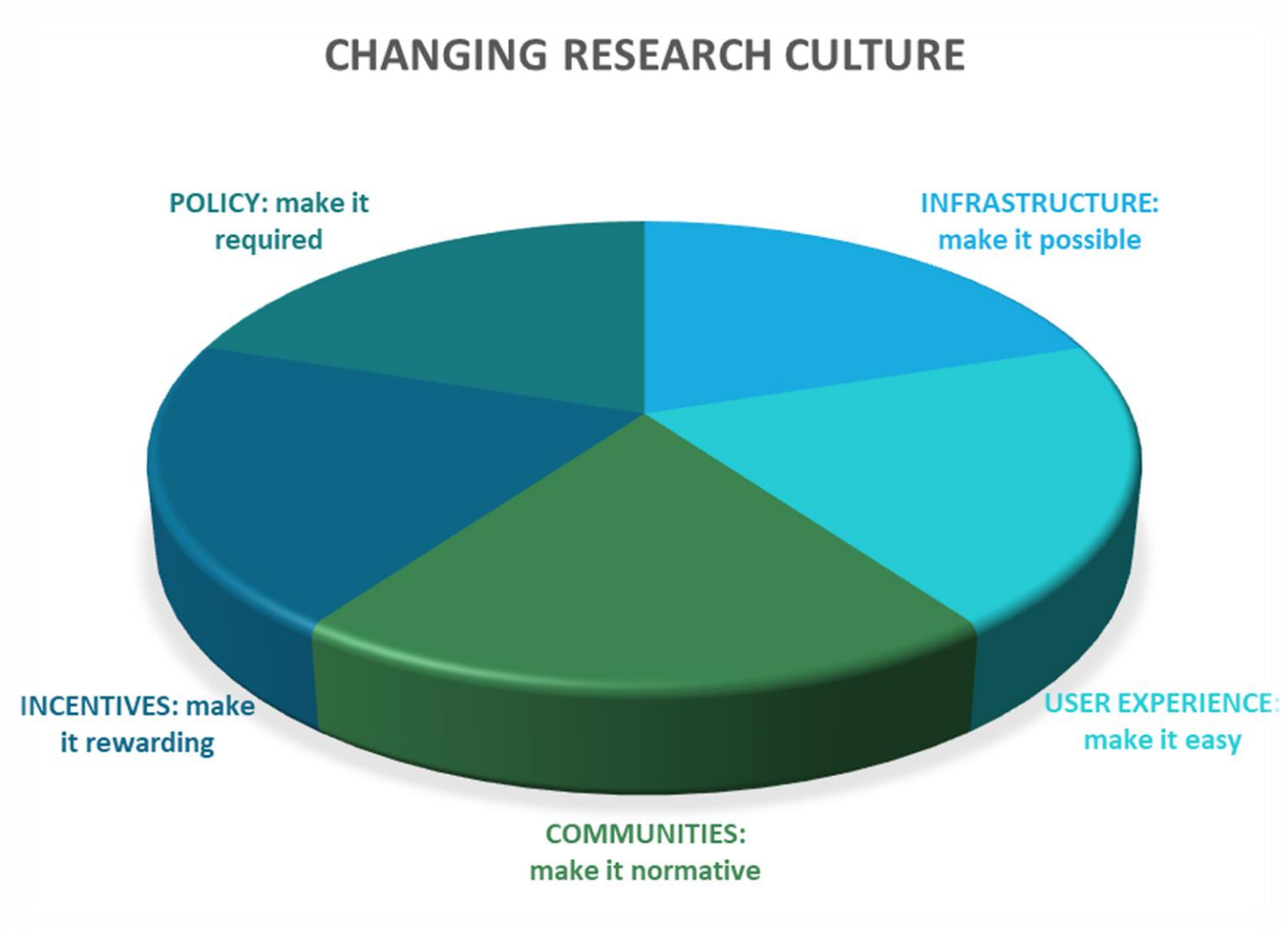
There is another way… The Turing way!



Key aspect for collaboration in neuroscience



It’s never THAT easy

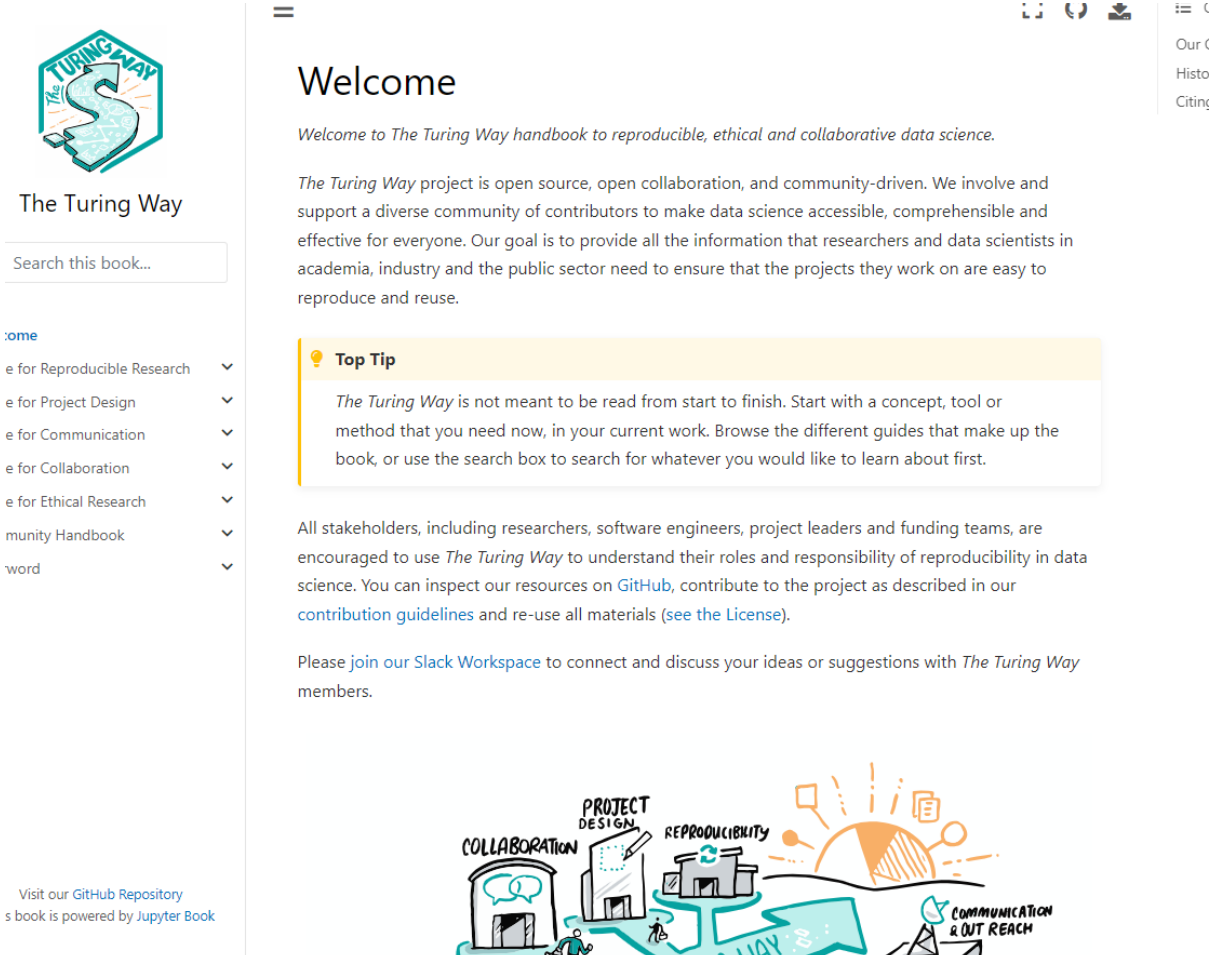


Reinterpreted from https://www.cos.io/blog/cos-celebrates-10-years

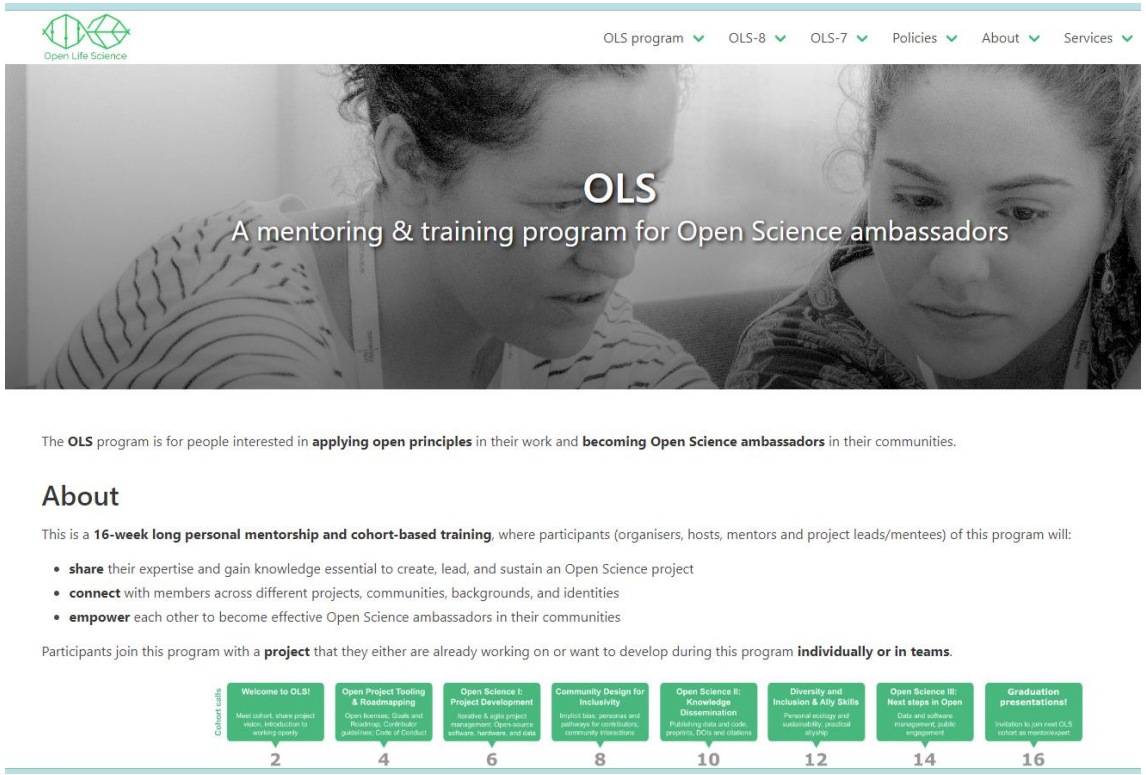
FOR YOU TO START… plenty of resources/tools out there… HAPPY TO HELP!

https://the-turing-way.netlify.app/index.html

https://software.ac.uk/



https://openlifesci.org/



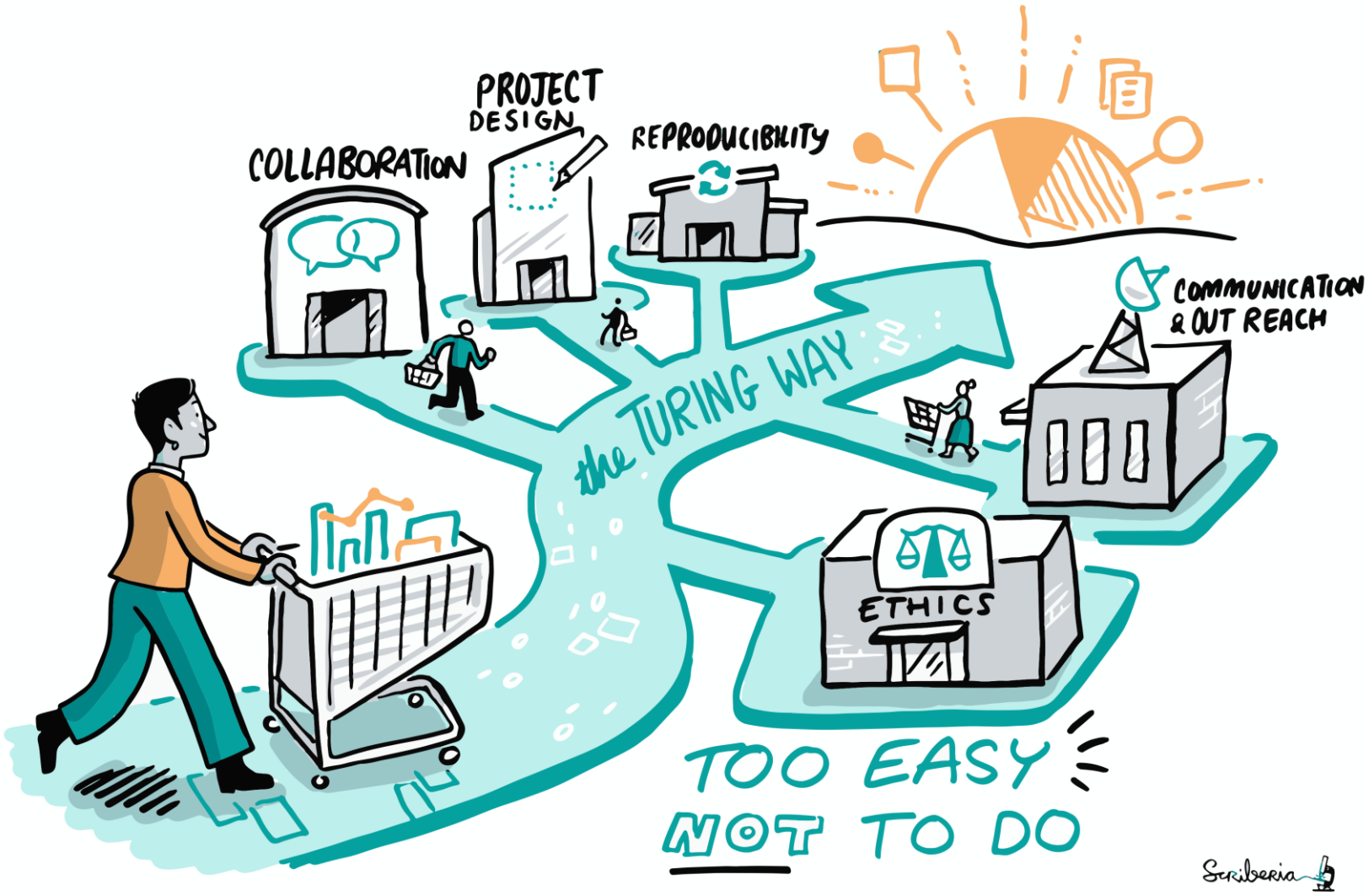
Don’t be overwhelmed… just choose from the Open science buffet!

Don't try to stuff yourself on everything, select what works for this study and let's steadily improve our fields..

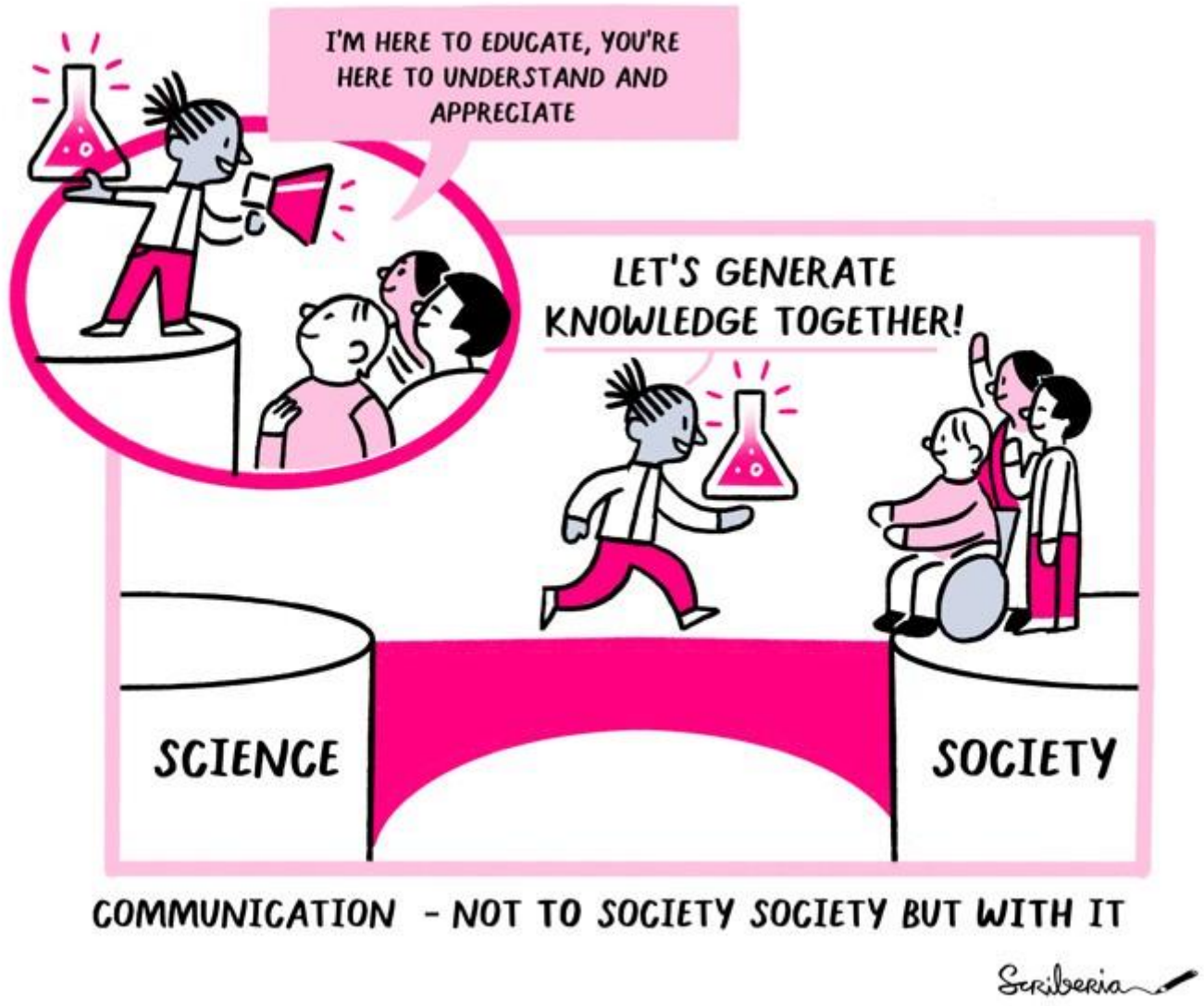
Take your pick from the ‘buffet’ of open science practices from transparency, statistics, preregistration, multi-lab collaborations…

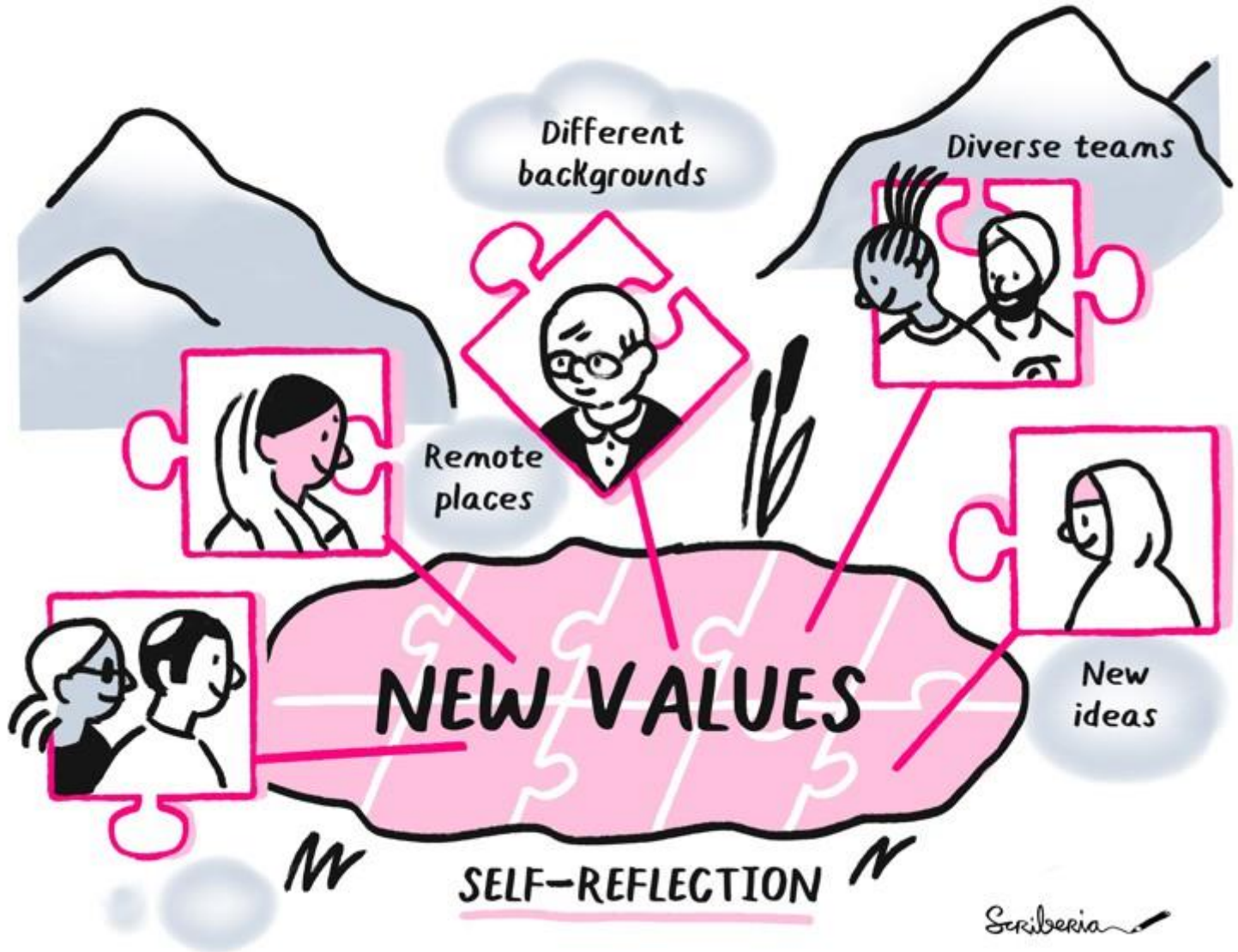
open science buffet" metaphor (h/t Christina Bergmann)

There is another way… The Turing way!



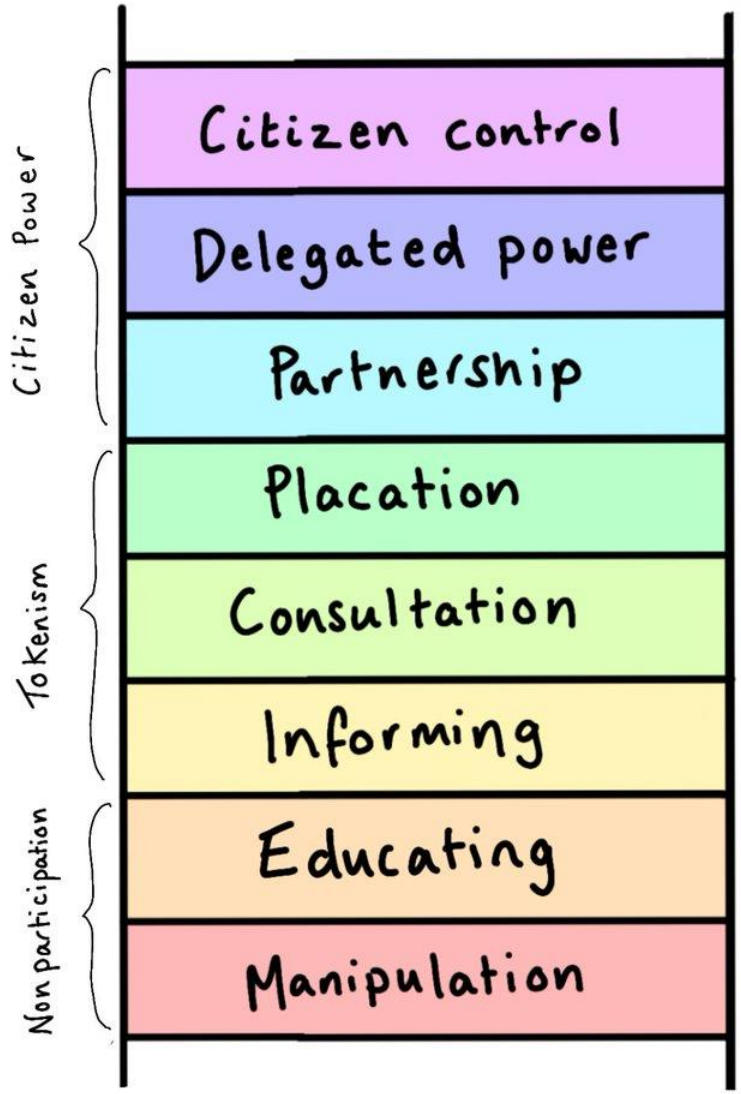
Self-reflection and change of mindset




Different backgrounds
Diverse teams
Remote places
New ideas
NEW VALUES
SELF-REFLECTION
Scriberia

Ladder of Participation

(Arnstein, 1969)



Stakeholders have the idea and set up the project

Goal created by a facilitator but resources and responsibility given to citizens

Stakeholders have direct involvement in decision making

Stakeholders shape ideas, but final decision sits with facilitators

Stakeholder views are sought but decisions made by facilitators

Stakeholders are informed on decisions but no opportunity to contribute

Assumption that the stakeholders are passive receipients

The illusion of participation when actually power is denied

Drawn by Juliet Young