You are free to

- share, adapt or re-mix
- photograph, video or broadcast
- blog, live-blog or post-video

this presentation

Provided that

you attribute the work to its author and respect the rights and licenses associated with its components

Data as first-class citizens in the Science of the 21st century

#OAI12 - 7th September 2021

pcmasuzzo | paola.masuzzo@gmail.com

IGDORE
Institute for Globally Distributed Open Research and Education

CC BY

Photo by Ameen Fahmy on Unsplash

we can't talk about *research* the way we did years and years ago

Photo by Ameen Fahmy on Unsplash

we can't talk about *research* the way we did years and years ago

we'll be able to *innovate* only when we'll feel *free* to experiment

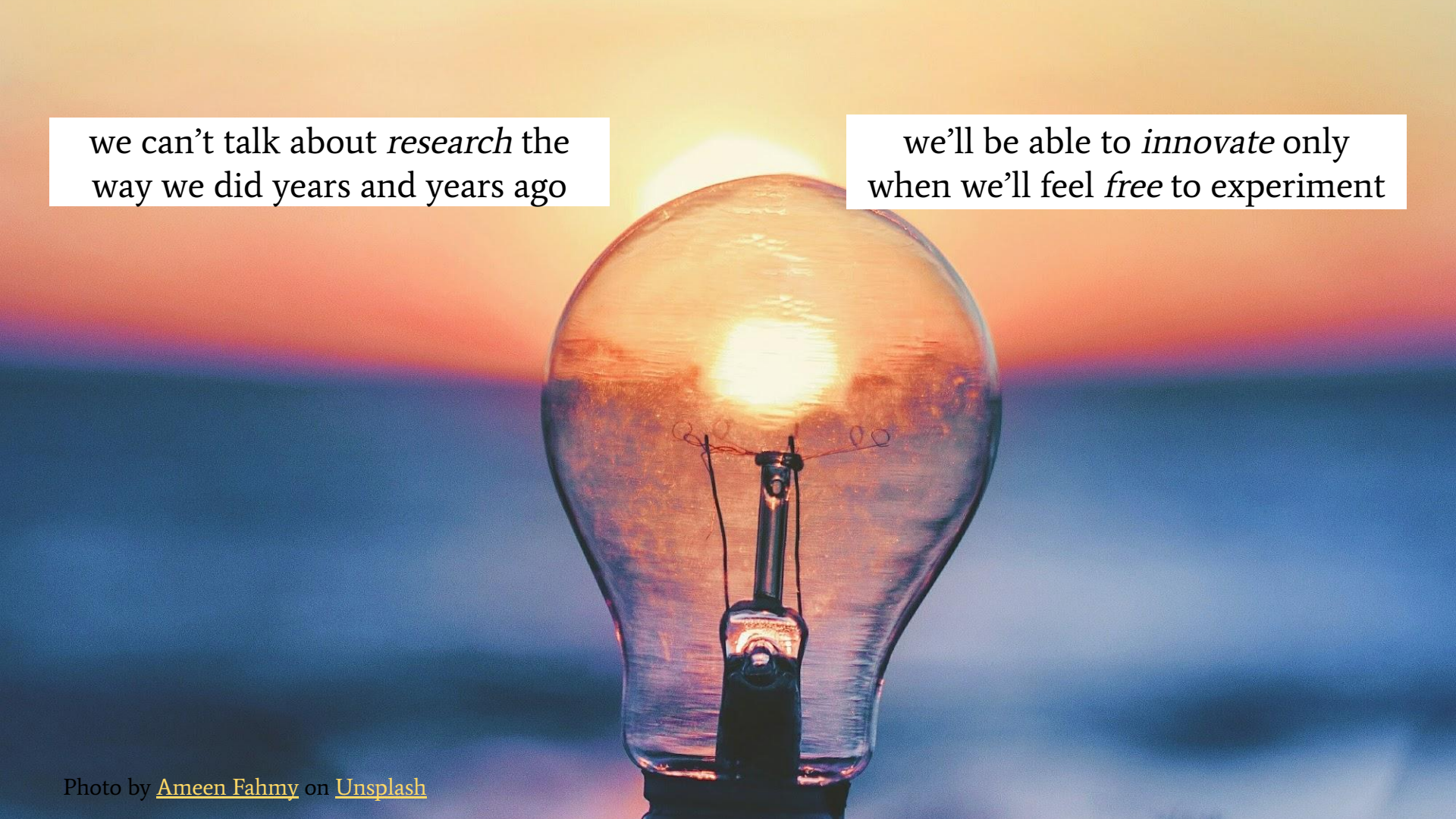
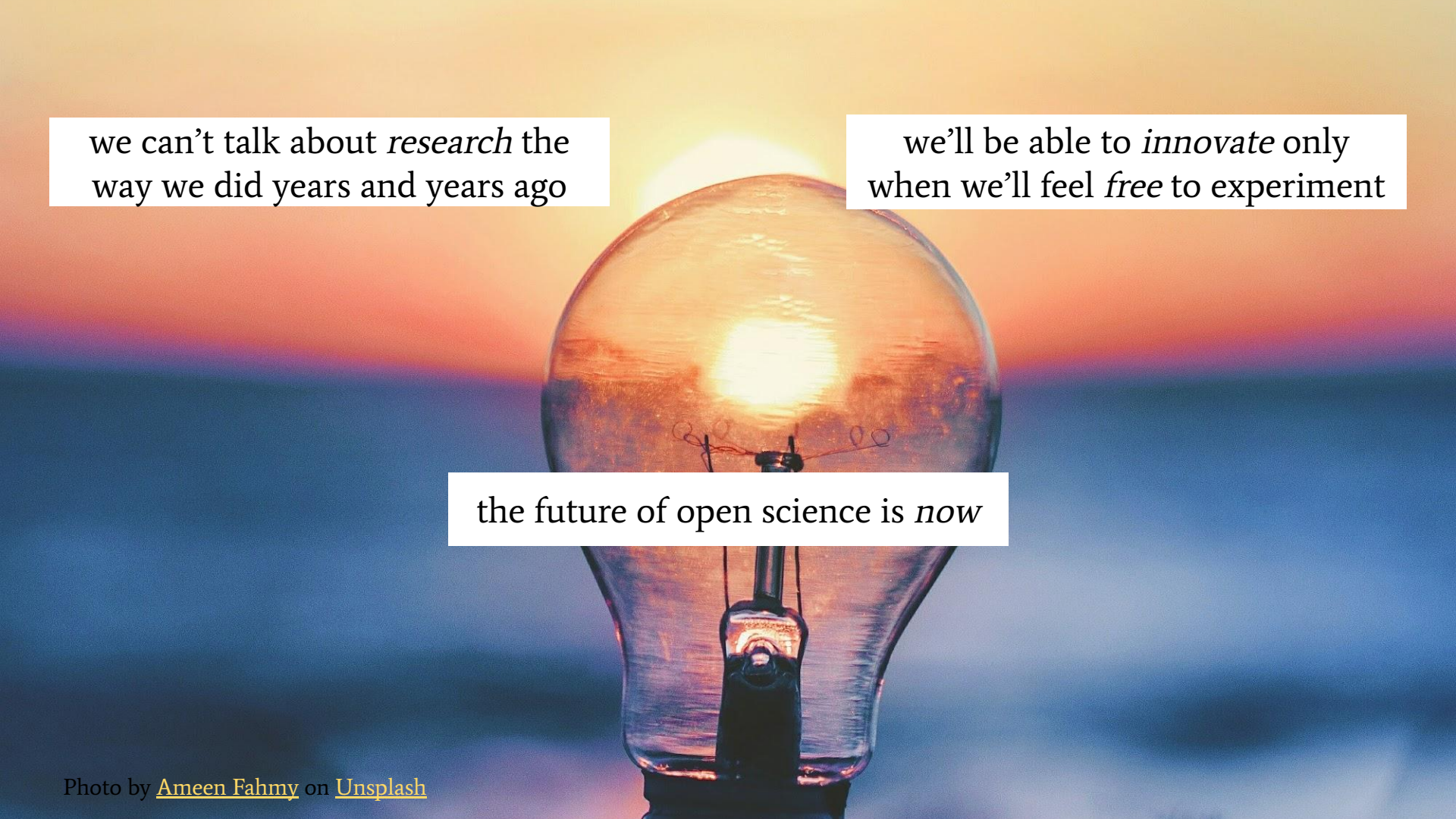


Photo by Ameen Fahmy on Unsplash



we can't talk about *research* the way we did years and years ago

we'll be able to *innovate* only when we'll feel *free* to experiment

the future of open science is *now*

Photo by Ameen Fahmy on Unsplash

open science and digital research data

Photo by Tim Mossholder on Unsplash

the Open Science movement *encourages* researchers to *share* research output *beyond* the contents of a published academic article

CC BY

data as first-class citizens

barriers to data sharing and how to overcome them

open data outside the academic bubble

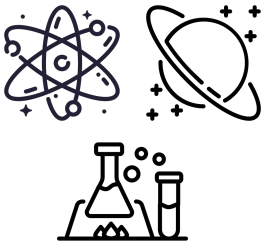
data as first-class citizens

barriers to data sharing and how to overcome them

open data outside the academic bubble

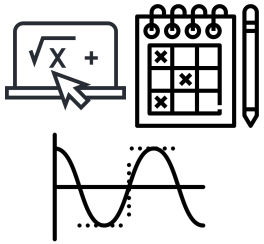
data-informed science

experiments:
empirical science



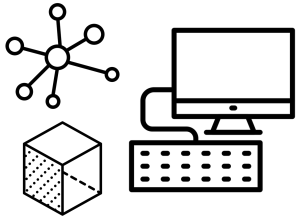
thousands
of years ago

laws:
theoretical science



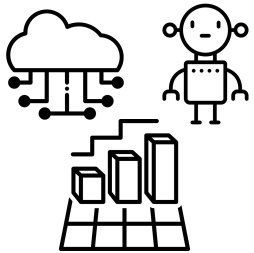
last few
hundreds years

simulations:
computational
science



last few
decades

data:
eScience



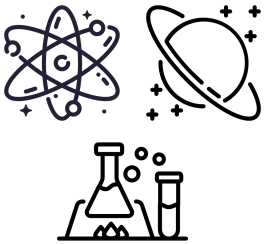
today

CC BY

data-informed science

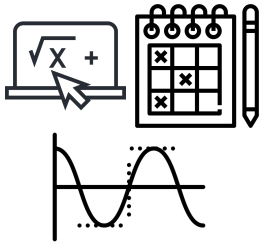
- machine-readable data as substrate for knowledge discovery
- reproducibility
- research assessment

experiments:
empirical science



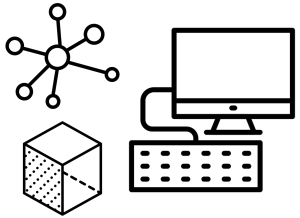
thousands
of years ago

laws:
theoretical science



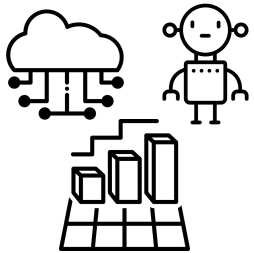
last few
hundreds years

simulations:
computational
science



last few
decades

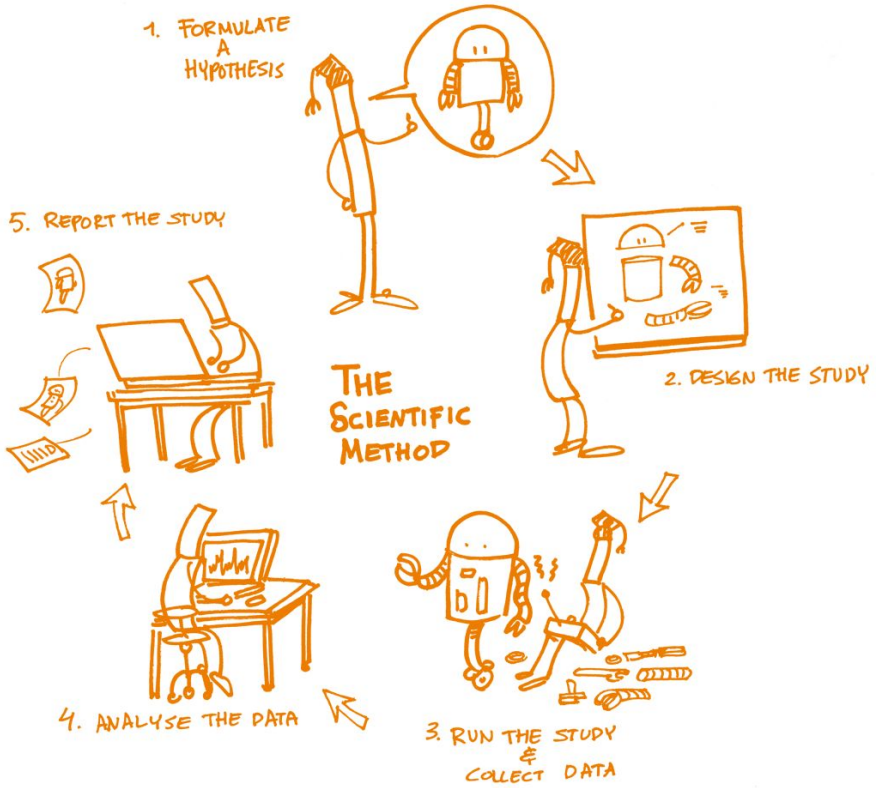
data:
eScience



today

CC BY

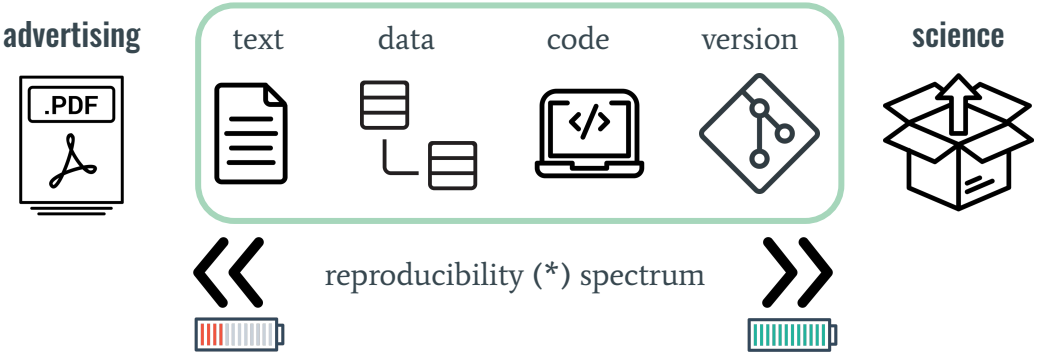
the research method



Open Science Training Handbook - Foster

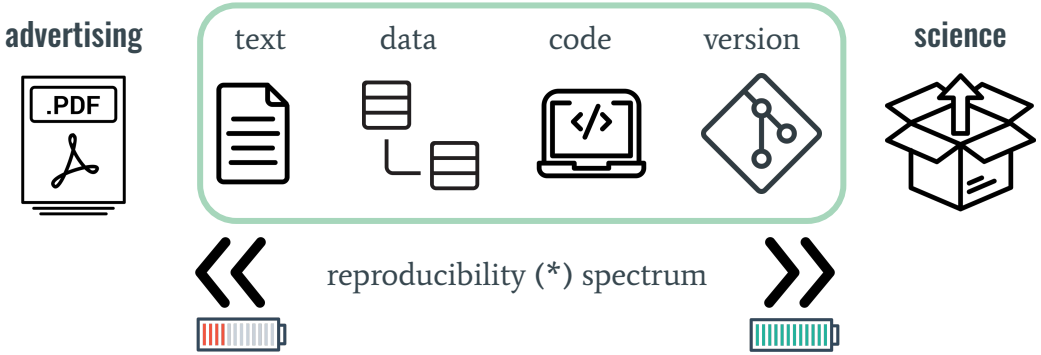
CC BY

research produces so much more than a PDF

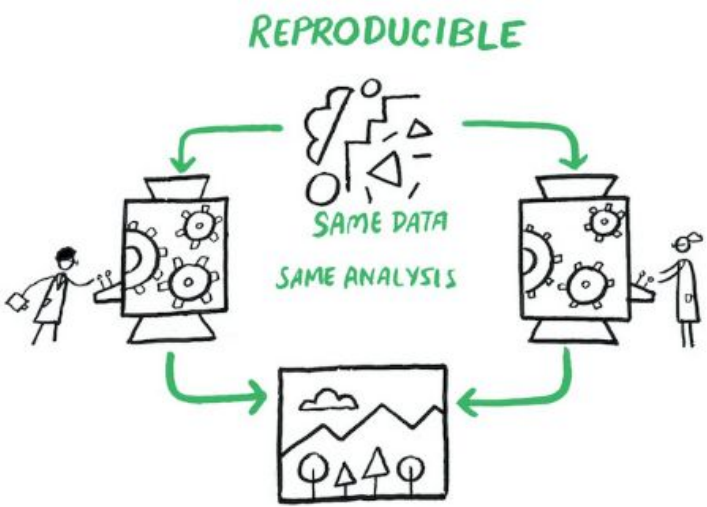


CC BY

research produces so much more than a PDF



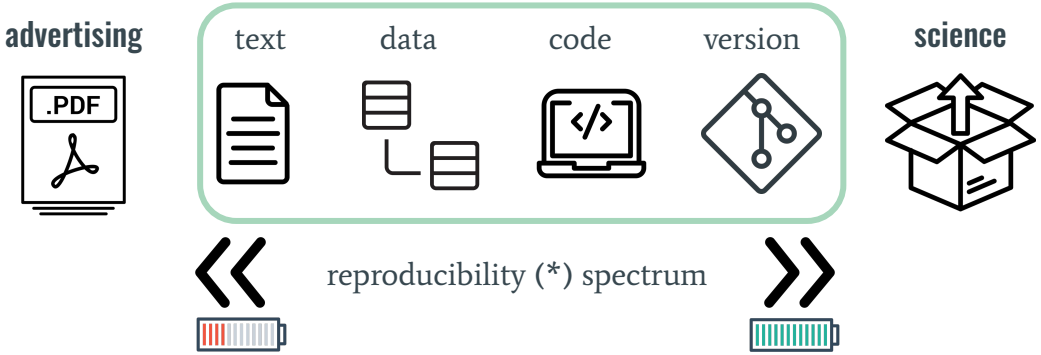
reproducibility is the minimum standard for research validity



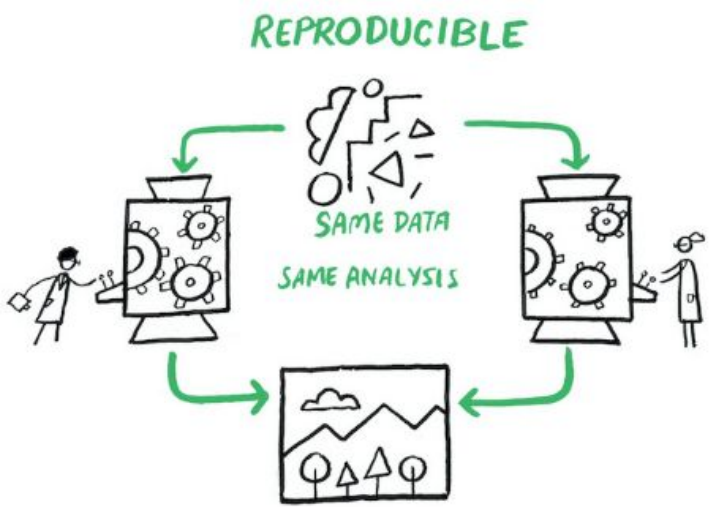
The Turing Way Community, & Scriberia. (2019, July 11). Illustrations from the Turing Way book dashes. Zenodo. http://doi.org/10.5281/zenodo.3332808

CC BY

research produces so much more than a PDF



reproducibility is the minimum standard for research validity



the idea is that objects that belong together should be linked to each other (and other objects), so that they can be discovered on the web

The Turing Way Community, & Scriberia. (2019, July 11). Illustrations from the Turing Way book dashes. Zenodo. http://doi.org/10.5281/zenodo.3332808

CC BY

the FAIR principles: guidance for data stewardship

the FAIR principles have been designed to assist discovery and reuse of research objects through the web

FAIR != OPEN

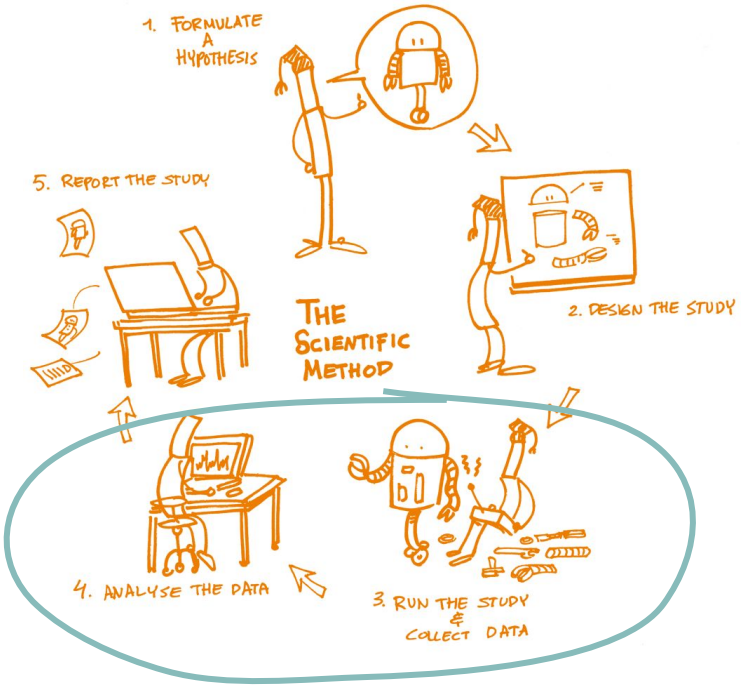
FAIR comes in degrees

FAIR = agnostic of technical implementations

The Turing Way project illustration by Scriberia. Original version on Zenodo. http://doi.org/10.5281/zenodo.3695300

CC BY

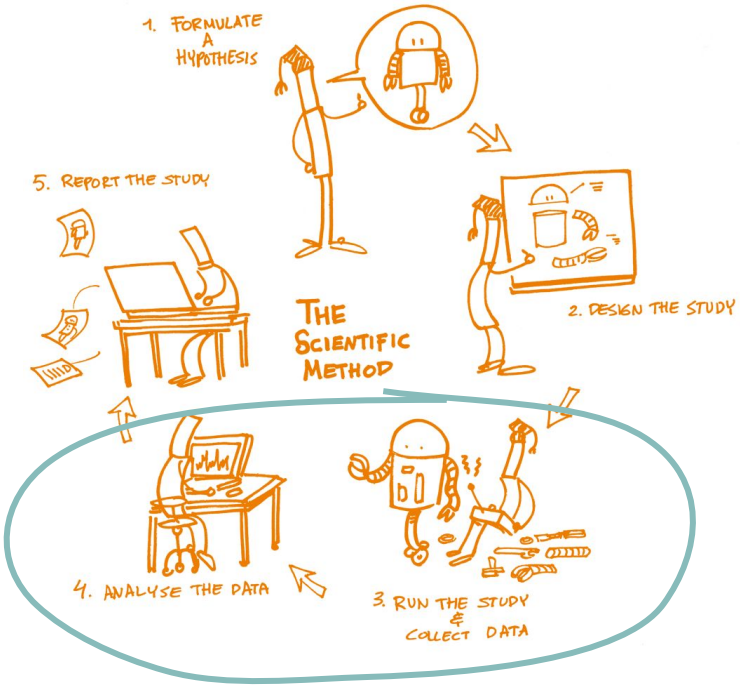
data have a life cycle on their own

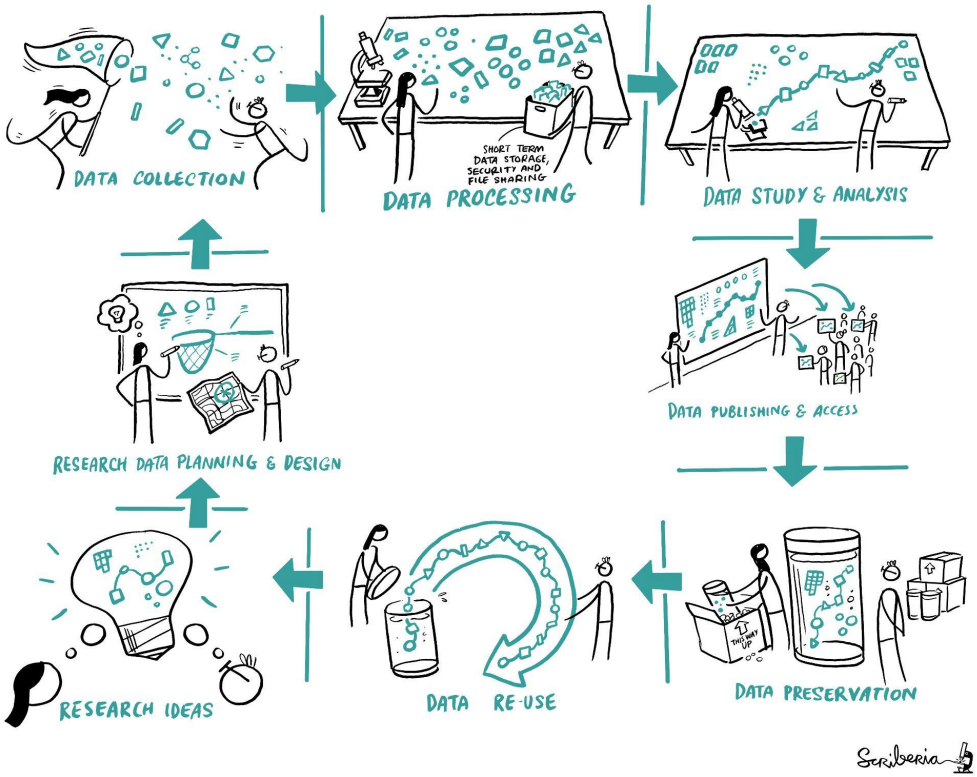


Open Science Training Handbook - Foster

CC BY

data have a life cycle on their own





Open Science Training Handbook - Foster | http://doi.org/10.5281/zenodo.3695300

CC BY

data as first-class citizens

barriers to data sharing and how to overcome them

open data outside the academic bubble

why is data sharing so hard?

it requires additional skills

it takes time (on already very busy agendas)

it's not rewarding (enough?)

we tend to hold ourselves to higher standard than others

Photo by Tim Mossholder on Unsplash

CC BY

a FAIR journey

http://doi.org/10.5281/zenodo.3695300

CC BY

turning FAIR into practice

FAIR Cookbook

Created by researchers and data managers professionals, the FAIR Cookbook is an online resource for the Life Sciences with recipes that help you to make and keep data Findable, Accessible, Interoperable and Reusable (FAIR).

Turning FAIR into practice

The FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. However, the FAIR Principles are aspirational and generic. The FAIR Cookbook guides *researchers* and *data stewards* in their FAIRification journey; and also provides *policy makers* and *trainers* with practical examples to recommend in their guidance and use in their educational material.

https://fairplus.github.io/the-fair-cookbook/content/home.html

CC BY

too easy not to do

F Findability

Exemplar recipes:

- Unique, persistent identifiers
- Search engine optimization

→ More about Findability

A Accessibility

Exemplar recipes:

- Transferring data with SFTP
- Downloading data with Aspera

→ More about Accessibility

I Interoperability

Exemplar recipes:

- Selecting terminologies and ontologies
- Creating a metadata profile

→ More about Interoperability

R Reusability

Exemplar recipes:

- Data licenses
- Declaring data's permitted uses

→ More about Reusability

CC BY

we need to invest more in data stewardship

WORLD VIEW | 25 February 2020

Invest 5% of research funds in ensuring data are reusable

It is irresponsible to support research but not data stewardship, says Barend Mons.

European Commission

cost of not having FAIR research data estimated at least €10 billion each year in the academic sector alone, and €16 billion in *lost* innovation opportunities

https://www.nature.com/articles/d41586-020-00505-7 - Cost of not having FAIR research data

CC BY

we need to invest more in data stewardship

WORLD VIEW | 25 February 2020

Invest 5% of research funds in ensuring data are reusable

It is irresponsible to support research but not data stewardship, says Barend Mons.

Cultural obstacles to research data management and sharing at TU Delft

Authors: Esther Plomp ✉, Nicolas Dintzner, Marta Teperek, Alastair Dunning

https://www.nature.com/articles/d41586-020-00505-7 - https://insights.uksg.org/articles/10.1629/uksg.484/

CC BY

a FAIR data world? not without people

address data management needs across disciplines

data stewards and managers

assist with coding tasks and issues

software developers

teach, support, and inspire other researchers

(open) champions

head of research data services

people who believe in people

CC BY

a FAIR data world? not without people

address data management needs across disciplines

data stewards and managers

assist with coding tasks and issues

software developers

teach, support, and inspire other researchers

(open) champions

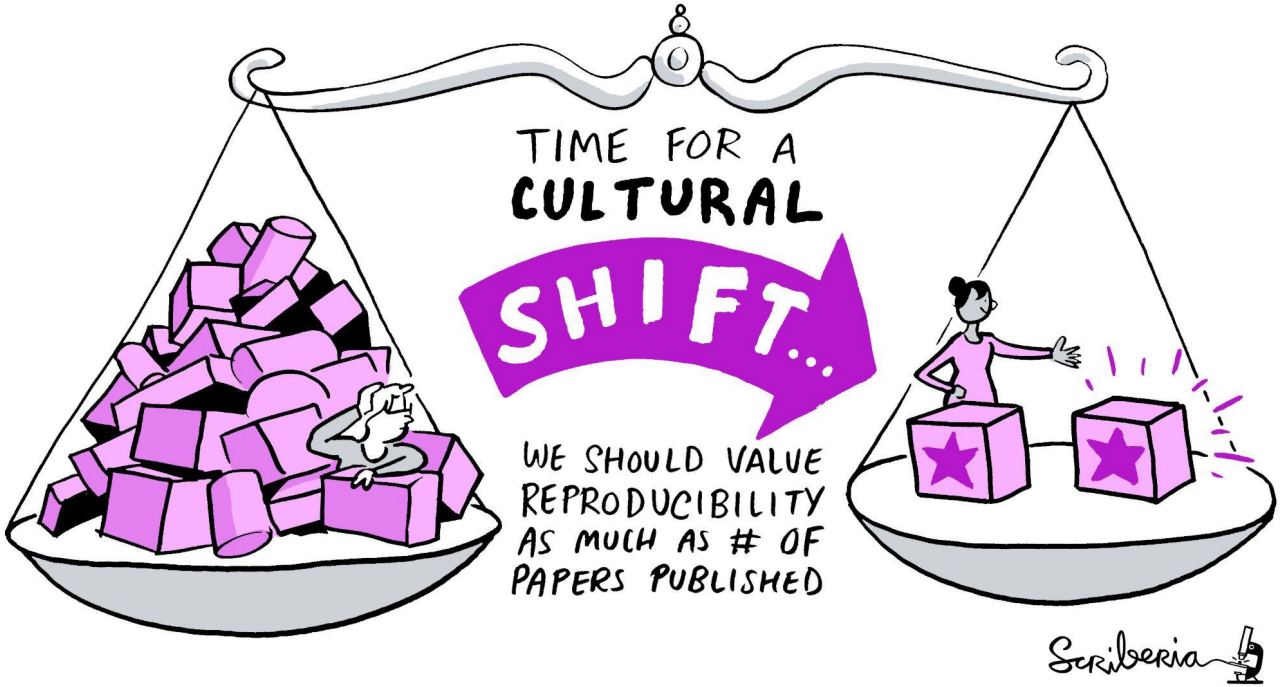
credit
reward
assessment

head of research data services

people who believe in people

CC BY

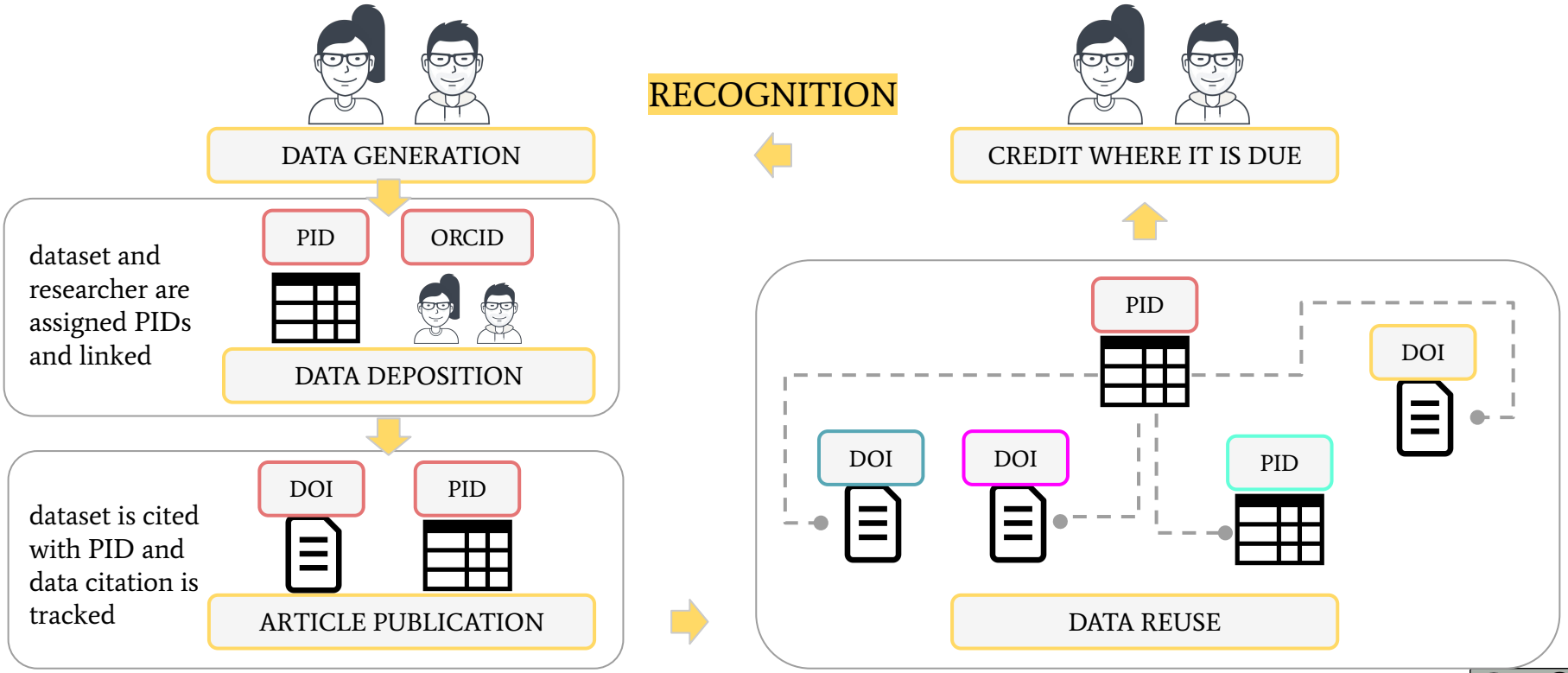
value reproducibility and transparency



The Turing Way project illustration by Scriberia. Original version on Zenodo. http://doi.org/10.5281/zenodo.3695300

CC BY

give credit where credit's due



CC BY

make it ok to be wrong

one of the biggest obstacles might still be *cultural*

https://twitter.com/springernature/status/963382291924963328 - #LoveData18

CC BY

make it ok to go slow(er)

Scientific Life

Fast Lane to Slow Science

Uta Frith[1,*]

slow science in belgium

an interuniversity platform for discussion on academia's future

https://www.sciencedirect.com/science/article/abs/pii/S1364661319302426 - https://slowscience.be/

data as first-class citizens

barriers to data sharing and how to overcome them

open data outside the academic bubble

science *and* society, or science *for* society?

there are times when science can seem to lose its connection to society and its needs

this connection, made of *dialogue* and *collaboration*, is critical to address challenges and build better realities

availability of *data* and *digital products* plays a crucial role in this *synergy*

CC BY

science *and* society, or science *for* society?

there are times when science can seem to lose its connection to society and its needs

this connection, made of *dialogue* and *collaboration*, is critical to address challenges and build better realities

availability of *data* and *digital products* plays a crucial role in this *synergy*

science *with* and *for* society

CC BY

data and democracy

dati bene comune

public trust in COVID data is key to ensure correct risk perception and citizens behavior

sharing and accessing data is key to build this public trust

PDF files and dashboards are not (FAIR) data!

nData

TRANSPARENCY INTERNATIONAL ITALIA

Associazione contro la corruzione

actionaid

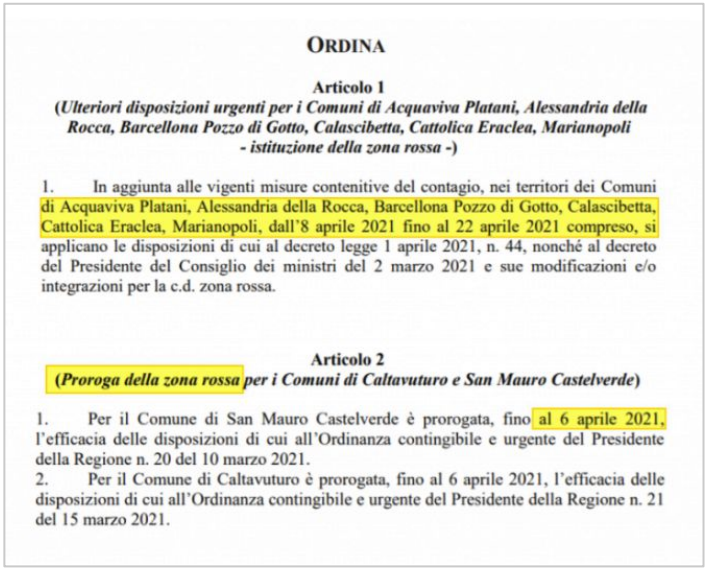
REALIZZA IL CAMBIAMENTO

https://www.datibenecomune.it/

CC BY

data and citizen science

“Red zones” in Sicily: a story of civic hacking



ORDINA

Articolo 1

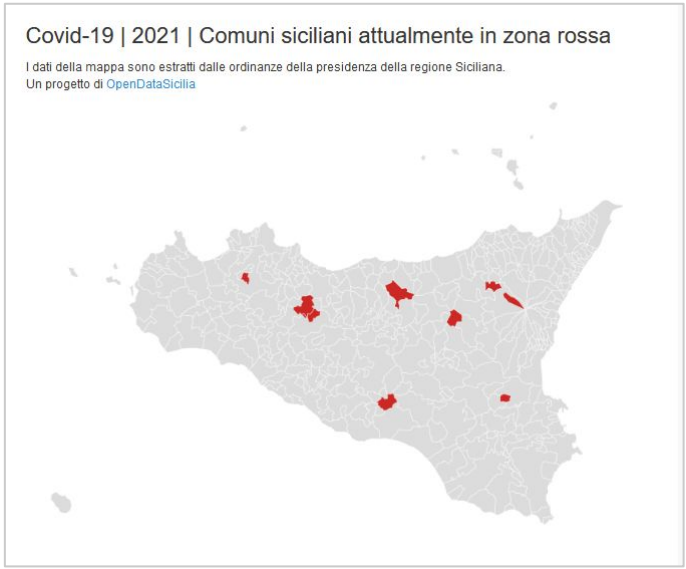
(Ulteriori disposizioni urgenti per i Comuni di Acquaviva Platani, Alessandria della Rocca, Barcellona Pozzo di Gotto, Calascibetta, Cattolica Eraclea, Marianopoli - istituzione della zona rossa -)

1. In aggiunta alle vigenti misure contenitive del contagio, nei territori dei Comuni di Acquaviva Platani, Alessandria della Rocca, Barcellona Pozzo di Gotto, Calascibetta, Cattolica Eraclea, Marianopoli, dall’8 aprile 2021 fino al 22 aprile 2021 compreso, si applicano le disposizioni di cui al decreto legge 1 aprile 2021, n. 44, nonché al decreto del Presidente del Consiglio dei ministri del 2 marzo 2021 e sue modificazioni e/o integrazioni per la c.d. zona rossa.

Articolo 2

(Proroga della zona rossa per i Comuni di Caltavuturo e San Mauro Castelverde)

1. Per il Comune di San Mauro Castelverde è prorogata, fino al 6 aprile 2021, l’efficacia delle disposizioni di cui all’Ordinanza contingibile e urgente del Presidente della Regione n. 20 del 10 marzo 2021.
2. Per il Comune di Caltavuturo è prorogata, fino al 6 aprile 2021, l’efficacia delle disposizioni di cui all’Ordinanza contingibile e urgente del Presidente della Regione n. 21 del 15 marzo 2021.



http://opendatasicilia.it/2021/04/22/red-zones-in-sicily-a-story-of-civic-hacking/

data and industry

https://fairtoolkit.pistoiaalliance.org/

the most relevant benefit of FAIR data to any organization is much greater reusability beyond the initial and primary purpose

the linked metadata provide value as the permanent scientific record, even when the original data has lost immediate value, if it is archived or eventually deleted as part of the data life cycle

CC BY

Thank you

#OAI12 - 9th September 2021

pcmasuzzo | paola.masuzzo@gmail.com

CC BY