NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

Introduction to Open Science

6th International PhD Summer School

13 Jul 2022 | KTU | Lithuania

eva.hnatkova@techlib.cz

CC BY

Overview

- What is Open Science?
- Why is it important?
- How to practice it?
- Discussion on Open Science

Science

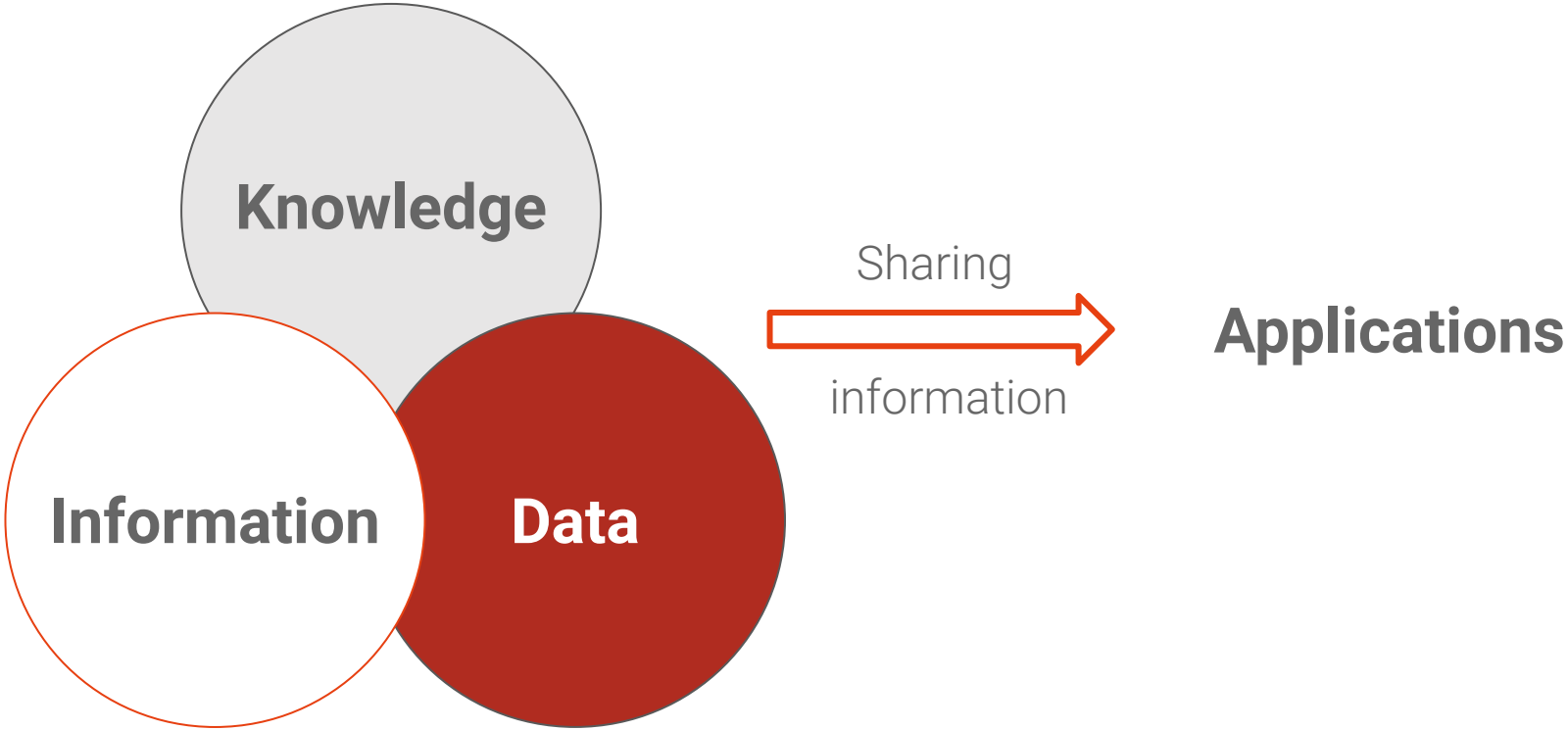
Science (from the Latin word scientia, meaning "***knowledge***") is one of the most **important** channels of knowledge, which help us to understand the world around us.

It has a specific role, as well as a variety of functions

for the benefit of our society:

- creating new knowledge
- improving education
- increasing the quality of our lives

Why do we publish?



Four major scientific “crisis”

1) Access: most research is paywalled to most people

- 75% of all scientific production in history must be paid

2) Increasing prices

- In the last 30 years, the subscription fees to scientific journals increased 300 times the inflation rate
- In 2016, 7.6 billion € have been spent globally to pay articles already paid by public funding
- Scientific publishers' net profit is >35% per year (for Google is 25%)

https://paywallthemovie.com/

Four major scientific “crisis”

3) **Evaluation:** driven by metrics controlled by publishers

4) Reproducibility crisis

- **50% of research is not reproducible!!**

Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

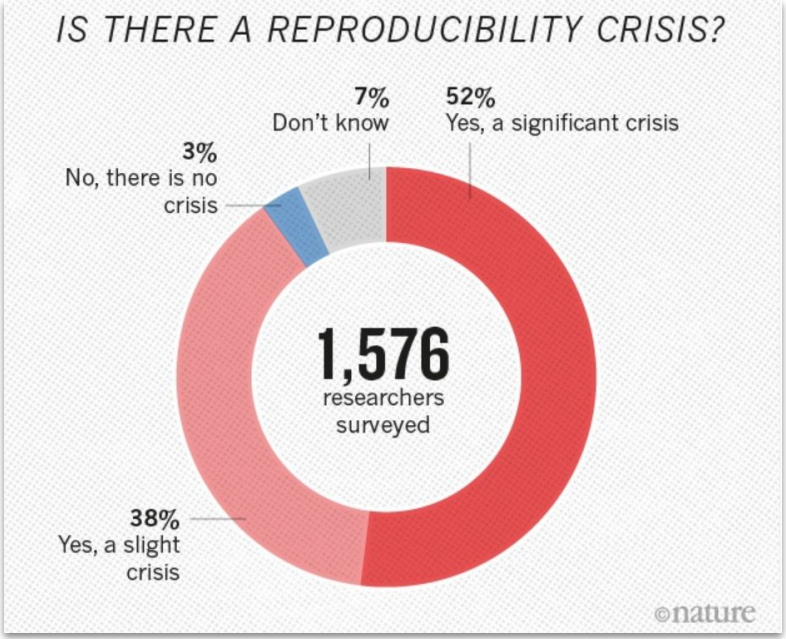
Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the factors that influence this problem and some corollaries thereof.

Modeling the Framework for False Positive Findings

Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there

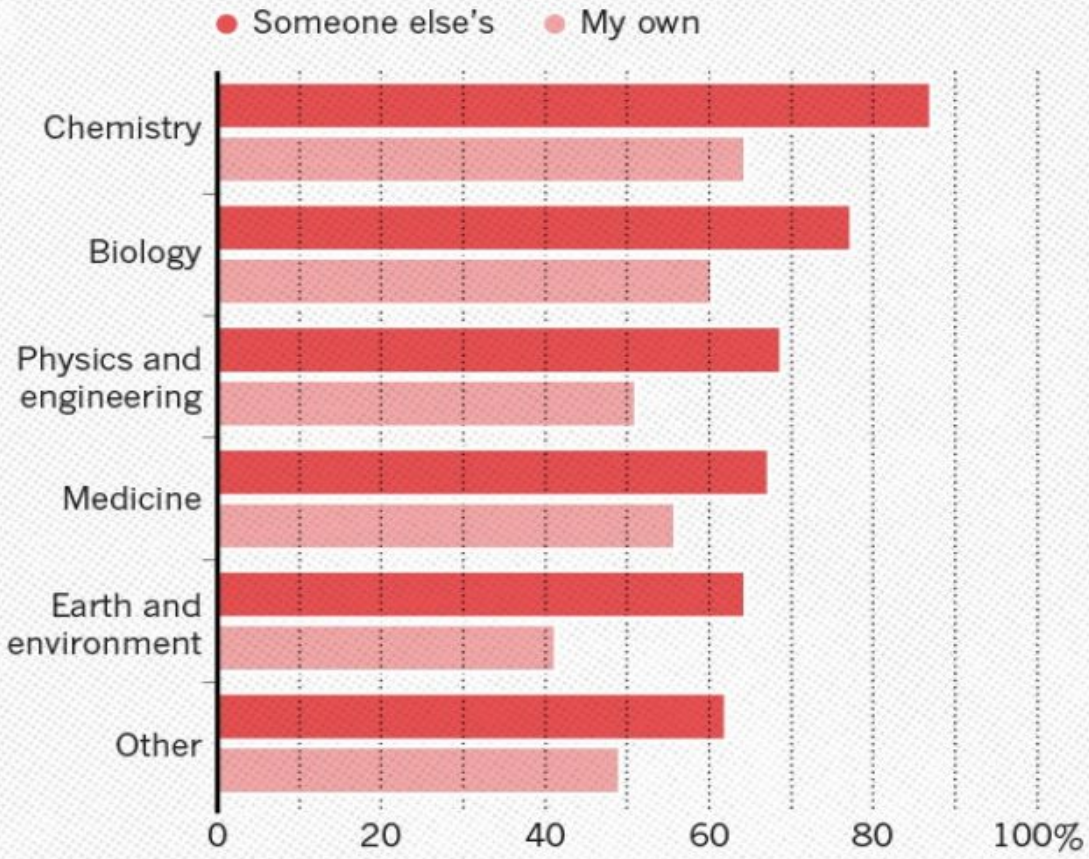
John P. A. Ioannidis (Plos Med, 2005)



Monya Baker. (Nature, 2016)

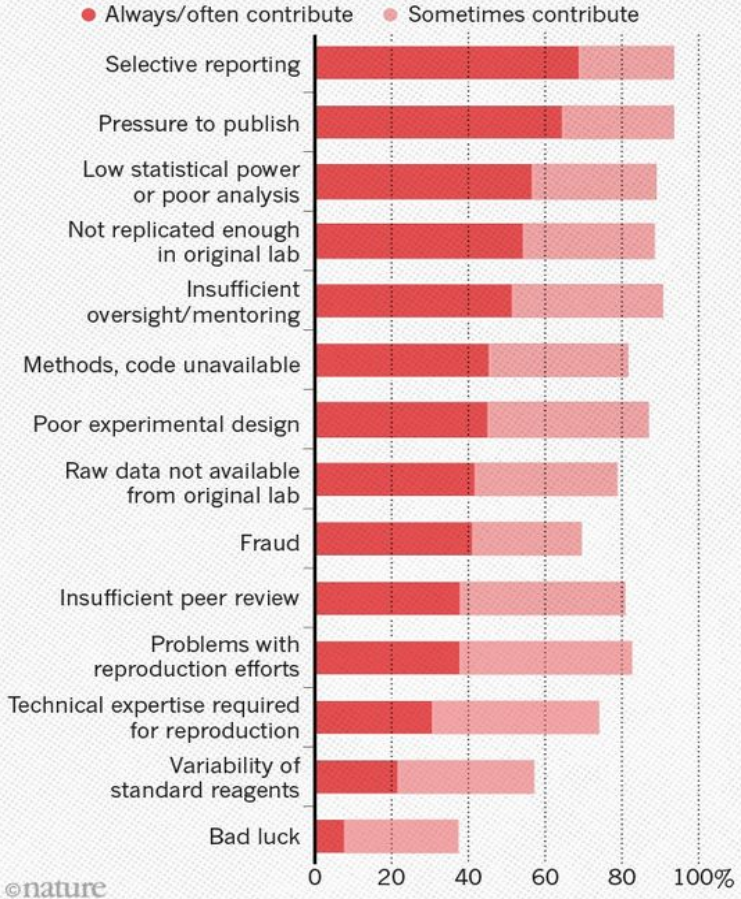
HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



WHAT FACTORS CONTRIBUTE TO IRREPRODUCIBLE RESEARCH?

Many top-rated factors relate to intense competition and time pressure.



©nature

Other adaptive behaviors

Increase of **self-citations**

- +179 % in 12 years in Italy

other types of manipulation of indicators

- abuse of co-authorship
- fake peer-reviews

ELSEVIER

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Self-citations as strategic response to the use of metrics for career decisions

Marco Seeber[a,*], Mattia Cattaneo[b], Michele Meoli[b], Paolo Malighetti[b]

[a] Department of Sociology, Ghent University, Korte Meer 3, 9000 Ghent, Belgium
[b] Department of Engineering, University of Bergamo, Via Pasubio 7b, 24044 Dalmine, BG, Italy

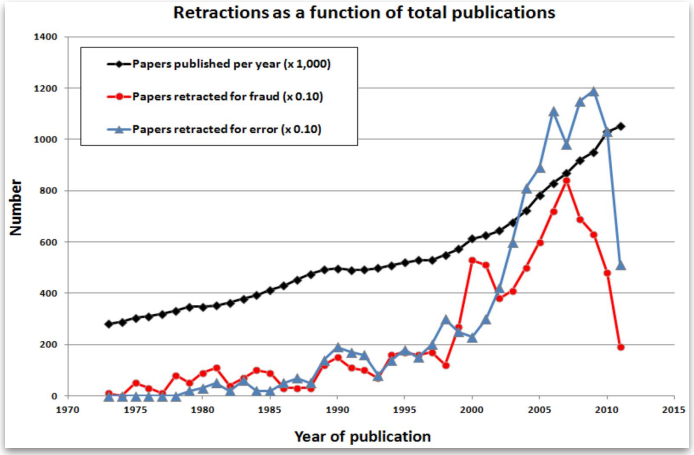
Check for updates

Marco Seebera et al. (Elsevier, 2019)

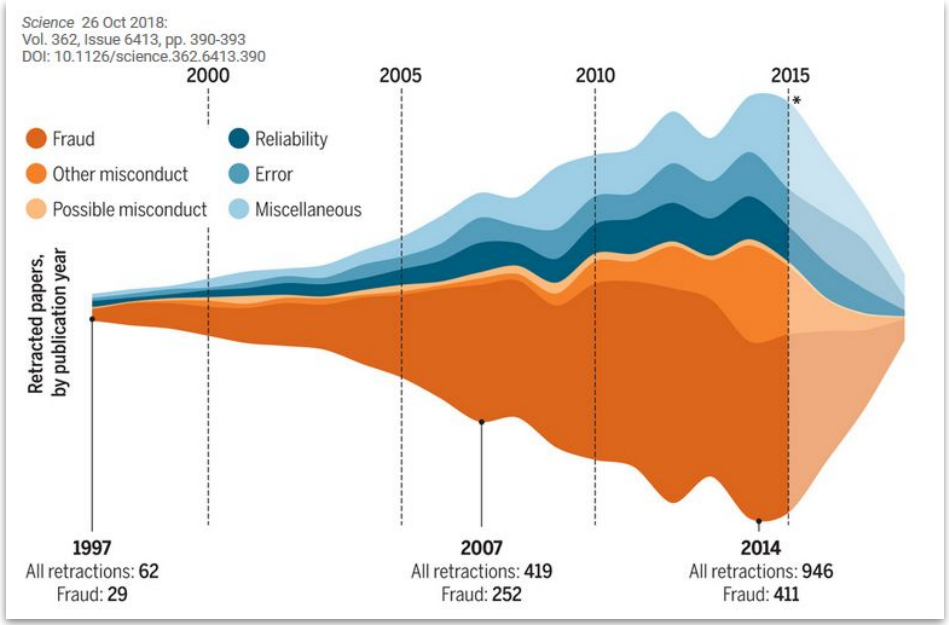
Other adaptive behaviors

Increase in **retract rate** ⇨ Particularly in **highly ranked journals**

- errors
- misconduct
- fraud



R. Grant Steen et al. (Plos, 2013)



Jeffrey Brainard: Rethinking retractions (Science, 2018)

Science in the time of COVID-19

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

OECD TACKLING CORONAVIRUS (COVID-19)
CONTRIBUTING TO A GLOBAL EFFORT

OECD Policy Responses to Coronavirus (COVID-19)

Why open science is critical to combatting COVID-19

Updated 12 May 2020

Disclaimer

Quick search

Key messages

Achievements of open science

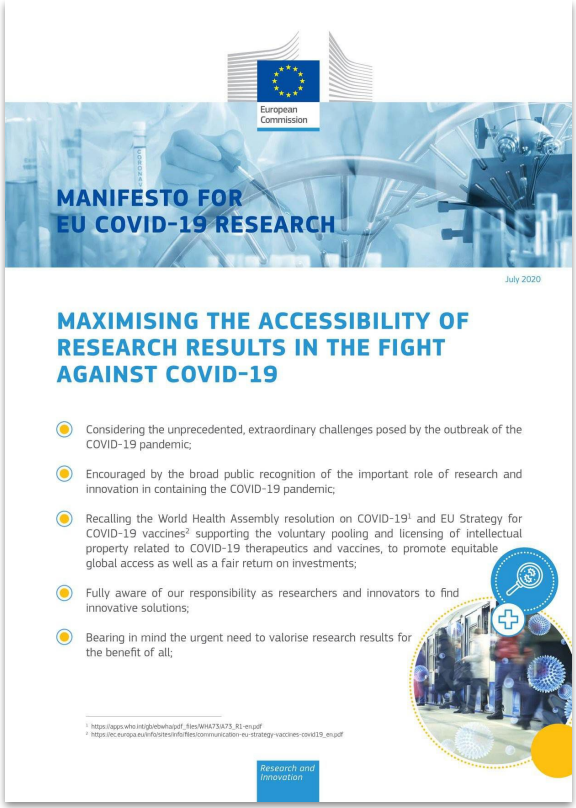
OPEN

Key messages

- In global emergencies like the coronavirus (COVID-19) pandemic, open science policies can remove obstacles to the free flow of research data and ideas, and thus accelerate the pace of research critical to combating the disease.

Source: OECD

Science in the time of COVID-19



European Commission

MANIFESTO FOR EU COVID-19 RESEARCH

July 2020

MAXIMISING THE ACCESSIBILITY OF RESEARCH RESULTS IN THE FIGHT AGAINST COVID-19

- Considering the unprecedented, extraordinary challenges posed by the outbreak of the COVID-19 pandemic;
- Encouraged by the broad public recognition of the important role of research and innovation in containing the COVID-19 pandemic;
- Recalling the World Health Assembly resolution on COVID-19[1] and EU Strategy for COVID-19 vaccines[2] supporting the voluntary pooling and licensing of intellectual property related to COVID-19 therapeutics and vaccines, to promote equitable global access as well as a fair return on investments;
- Fully aware of our responsibility as researchers and innovators to find innovative solutions;
- Bearing in mind the urgent need to valorise research results for the benefit of all;

[1] https://apps.who.int/gb/ebwha/pdf_files/WHA73/A73_R1-en.pdf
[2] https://ec.europa.eu/info/sites/info/files/communication-eu-strategy-vaccines-covid19_en.pdf

Research and Innovation

By endorsing the manifesto you commit to

Make the generated results, whether tangible or intangible, public and accessible without delay, for instance on the Horizon Results Platform, on an existing IP sharing platform, or through an existing patent pool.

Make scientific papers and research data available in open access without delay and following the FAIR principles via preprint servers or public repositories, with rights for others to build upon the publications and data and with access to the tools needed for their validation.In particular, make COVID-19 research data available through the European COVID-19 Data Platform

Where possible, grant for a limited time, non-exclusive royalty free licences on the intellectual property resulting from EU-funded research. These non-exclusive royalty free licenses shall be given in exchange for the licensees' commitment to rapidly and broadly distribute the resulting products and services under fair and reasonable conditions to prevent, diagnose, treat and contain COVID-19.

Manifesto for EU COVID-19 Research (EC, July 2020)

Science in the time of COVID-19

NTK
Národní technická knihovna
National Library of Technology

COVID-19 linked data

DEPOSIT LINK SEARCH ABOUT SIGN IN

Corona Virus Disease

COVID-19

Research outcomes | Search by title, author, abstract, DOI, orcid... | SEARCH

SUMMARY

PUBLICATIONS	RESEARCH DATA	SOFTWARE	OTHER RESEARCH
238,391	8,551	857	11,019

This portal provides access to publications, research data, projects and software that may be relevant to the Corona Virus Disease (COVID-19). The OpenAIRE COVID-19 Gateway aggregates COVID-19 related records, links them and provides a single access point for discovery and navigation. We tag content from the OpenAIRE Research Graph (10,000+ data sources) and additional sources. All COVID-19 related research results are linked to people, organizations and projects, providing a contextualized navigation.

Curated by: Alessia Bardi, Iryna Kuchma, Evgeny Bobrov, Ivana Truccolo, Elizabete Monteiro ...

View all 10 curators

Created: 16-Mar-2020 Members: 50

Projects: 130 Content Providers: 15

Linked to 7 Zenodo Communities

Subjects

COVID19, SARS-CoV, HCoV-19, mesh:C000657245, MERS-CoV, Síndrome Respiratorio Agudo Severo, mesh:COVID-19, COVID2019, COVID-19, SARS-CoV-2, 2019 novel coronavirus, severe acute respiratory syndrome coronavirus 2, Orthocoronavirinae, Coronaviridae, mesh:D045169, coronavirus, SARS, coronaviruses, coronavirus disease-19, sars cov 2

View all

OpenAIRE portal: Corona Virus Disease (COVID-19)

Science in the time of COVID-19

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

COVID-19 Data Portal

About ▾ Partners Related resources Bulk downloads Submit data

Viral Sequences Host Sequences Expression Proteins Biochemistry Imaging Literature

Accelerating research through data sharing

Read and sign our letter in support of open COVID-19 data >

Viral sequences

Raw and assembled sequence and analysis of SARS-CoV-2 and other coronaviruses.

391,141 records >

Host sequences

Raw and assembled sequence and analysis of human and other hosts.

11,654 records >

Expression

Gene and protein expression data of human genes implicated in the virus

Proteins

Curated functional and classification data on the SARS-CoV-2 protein entries and

Our letter in support of open COVID-19 data

We need to support the submission of open data, especially SARS-CoV-2

[The European COVID-19 Data Platform]

What about the other global challenges?

Access to scientific information is considered important to help solving these problems

Sustainable Development Goals for 2015-2030 (United Nations)

A strategic vision for Europe

European Commission (2016)

Open Innovation

- Involve more stakeholders in the innovation process

Open Science

- Facilitate more open research via digital tools

Open to the World

- Support more societal impact & collaboration both in and outside Europe

Key drivers of Open Science

- Availability of digital technologies allowing e.g. for data-intensive research
- Increase of the global scientific population
- Public demand for better and more efficient science
- Demand for accountable, responsive and transparent science
- Need to address faster societal challenges
- Need to contribute to economic growth

Melanie Imming, & Jon Tennant. (2018, June 8). Sticker open science: just science done right

What is Open Science?

Some definitions

Open Science represents a new approach to the scientific **process** based on **cooperative work** and new ways of diffusing knowledge by using digital technologies and new **collaborative tools**.

Open Science is the effort to make the primary outputs of publicly funded research results - **publications and the research data** – publicly accessible in digital format with no or minimal restriction, to the scientific community, the business sector, or society more generally, to enhance transparency and collaboration, and fostering innovation.

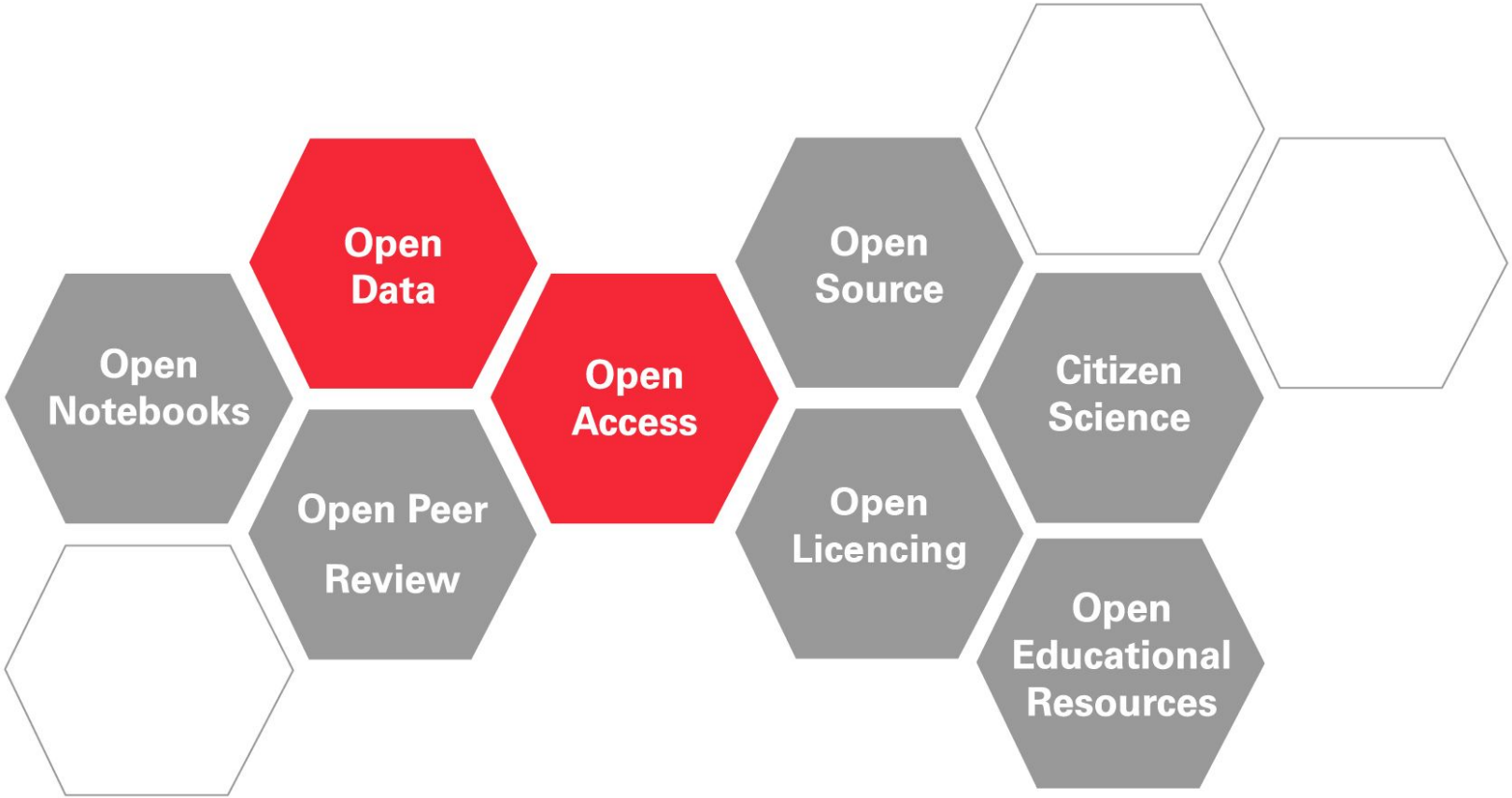
UNESCO Definition of Open Science

Open Science is an inclusive construct that **combines various movements** and **practices** aiming:

- **to make** multilingual **scientific knowledge** openly available, **accessible** and **reusable** for everyone;
- to **increase** scientific **collaborations** and **sharing of information** for the benefits of science and society;
- and to **open the processes** of scientific knowledge creation, evaluation, and communication to societal actors beyond the traditional scientific community.

UNESCO, Nov 2021

Open Science: an umbrella term



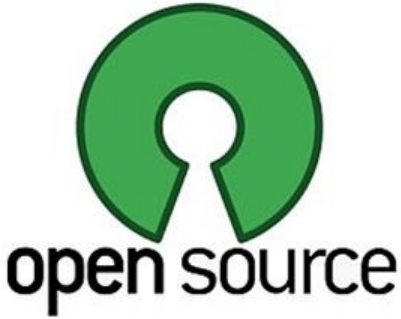
Benefits - Why do Open Science?

- opens access to research
- increases discoverability
- increases (social) impact
- facilitates reproducibility
- shares & saves resources
- speeds & helps innovation

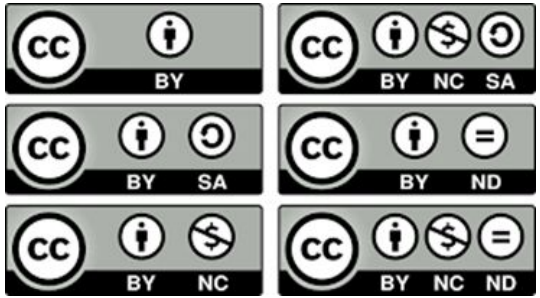
SGDs 2015 -2030 (United Nations)

Access to research is considered important to help solving global challenges

Open Science practices



OPEN LICENCING

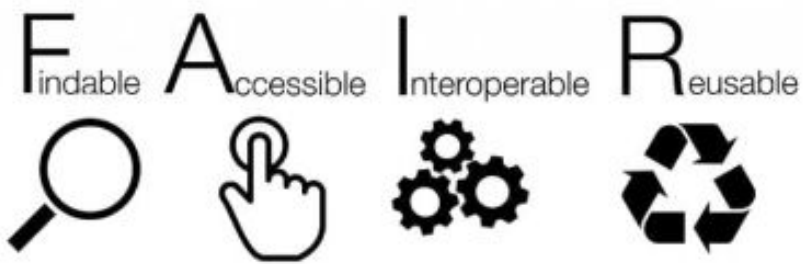


FOSTER Taxonomy of Open Science | Masuzzo & Martens (2017)

Closed versus Open Access

Traditional model	Open Access
• **closed access** only for institutions (library/consortium) which pay **subscriptions to journals** • paid access for all others	• **free access for all to read** • AAM version in repository ( ) • VoR version in the journal ( )
PAYWALL The Business of Scholarship	OPEN ACCESS

Open versus FAIR Data



≠

OPEN DATA

FAIR principles

- **Findable** - rich metadata, keywords, persistent identifiers
- **Accessible** - data in trusted repositories
- **Interoperable** - standards
- **Re-usable** - licenses

Go FAIR | Hansen & al (2018)

FAIR Data – open by default

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

FAIR data: "As open as possible, as closed as necessary"

OPEN	RESTRICTED/CONTROLLED	CLOSED
Freely used, modified & shared by anyone for any purpose	Limits on who can access & use data, how, or for what purpose	Under embargo, unable to share

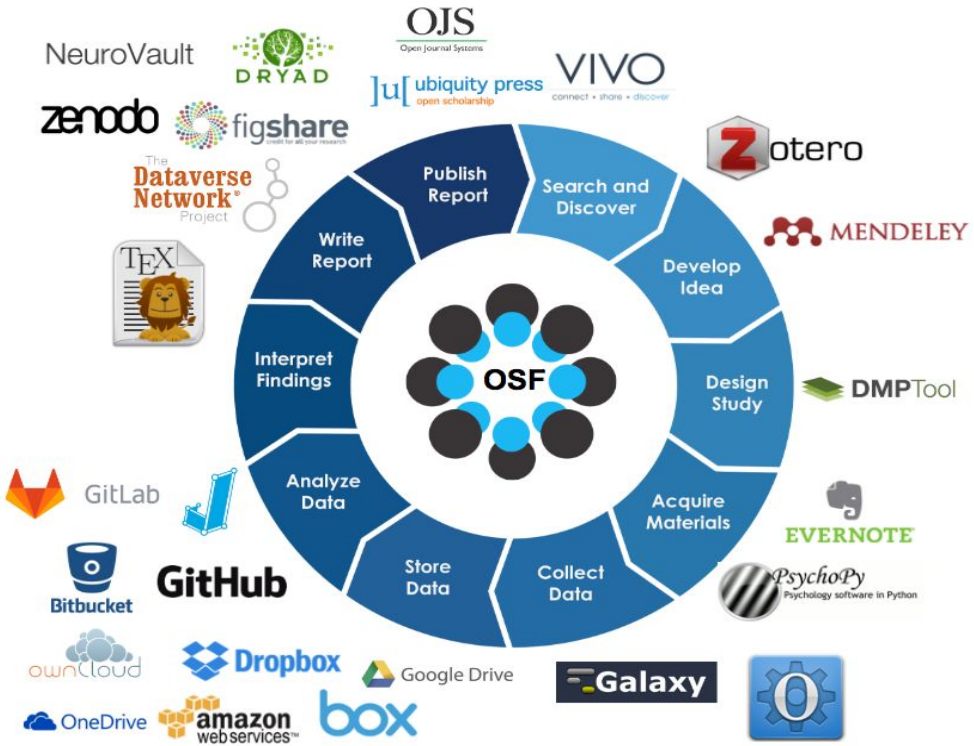
Opening up your data does not mean you are losing control!

Research Data Management

Research data workflow

- collect data
- curate data
- analyse data
- report data
- archive data

> **Data Management Plan!**



Variety of tools | Open Science Framework

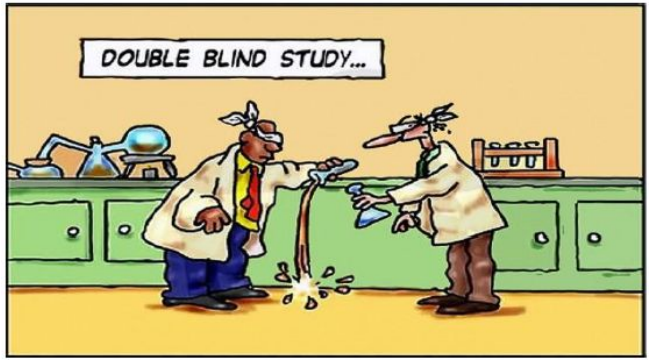
European Open Science Cloud (EOSC)

EOSC - a virtual research environment for all researchers

https://eosc-portal.eu/

- Researchers at the core
- Store/share own data
- Find/re-use other data
- Combine different data
- Deploy AI tools on data
- Go beyond current science
- Engage in developing EOSC

Open Peer Review



Tennant (2018) | Ross-Hellauer (2017)

Standard Peer Review

- authors submit to editors
- authors/reviewers ‘blind’
- versions/reviews closed

Open Peer Review

- authors submit to platform
- authors/reviewers public
- versions/reviews public

Creative Commons licences for reuse

Opening the article is not sufficient, you have to specify what could be done with your work. By a **CC** license you give other people the **right to share, use, and build upon** a work that you have created.

BY - Attribution to the creator

SA - Share-alike => not more restrictive

CC Zero / waiver license

NC - Non-commercial

ND - No derivatives

Public Domain Mark

The notion of Creative Commons was coined by Lawrence Lessig, the Net lawyer in 2001

PUBLIC DOMAIN — Free content, no restrictions

BY — Attribution alone

BY SA — Attribution + ShareAlike

BY ND — Attribution + NoDerivatives

BY NC — Attribution + Noncommercial

BY NC SA — Attribution + Noncommercial + ShareAlike

BY NC ND — Attribution + Noncommercial + NoDerivatives

Source: https://en.wikipedia.org/wiki/Creative_Commons_license

Citizen Science

- Reaching out to citizens
- Citizens deciding research
- Citizens helping research
- Citizens doing research

Examples of Citizen Science

Monitoring local water resources around the globe

- A programme promoting citizen participation
- Volunteers use a pre-built testing kit to collect local data on four markers of water quality—temperature, turbidity, dissolved oxygen levels and acidity

Discussion on Open Science

- Why should we do Open Science?
- Why should we **NOT** do Open Science?
- What research outputs should be open?