

The Uptake of **Open Science** :

Mapping the Results of a Systematic Literature Review

Hanna Shmagun

hanna.shmagun@gmail.com

(presenter)

Charles Oppenheim

c.oppenheim@rgu.ac.uk

Jangsup Shim

sjsll1it@gmail.com

Jaesoo Kim

jaesoo@kisti.re.kr

Content

1

Evolution of Science (Second Open Science revolution)

2

Concept of “Open Science”

3

Research problem. RQs. Methods

4

Open Science Taxonomy based on keyword co-occurrence

Content

5

Fields of knowledge related to Open Science research

6

Open Science data characteristics and categories

7

Open Science data = Open Government data ?

8

Proposals for future research

9

References

1. Evolution of Science

- shift towards amplifying
Collective Intelligence -



Secrecy of science (alchemy)

Scientific Revolution in Western Europe

Competition among noble patrons for reputable scientists

2/2 17 – 18 cent

1st “Open Science” revolution

ICT

Democratic movements: Open Source, Open Access, Open Government

Knowledge Economy

end 20 – 21 cent

2nd Open Science revolution

Scientific journal system emerged

Increases access to knowledge for general public

Enhances ability to extract value from existing data & knowledge

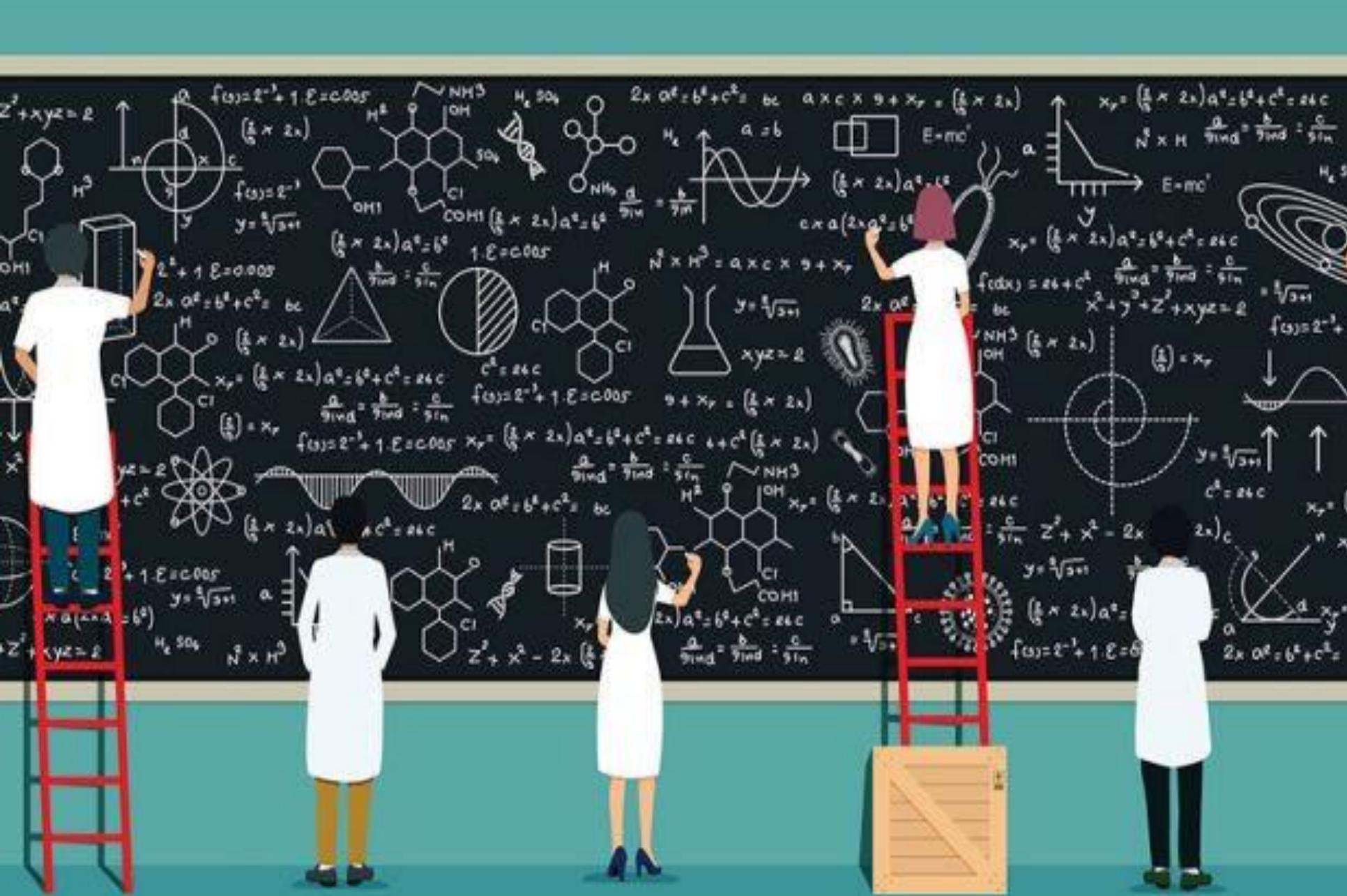
linked data; data intelligence

Speeds up scientific discoveries (networked science)

microexpertise; “Citizen Science”

1665: the world’s first scientific journal by the Royal Society of London





2. Concept of “Open Science”



Paul A David

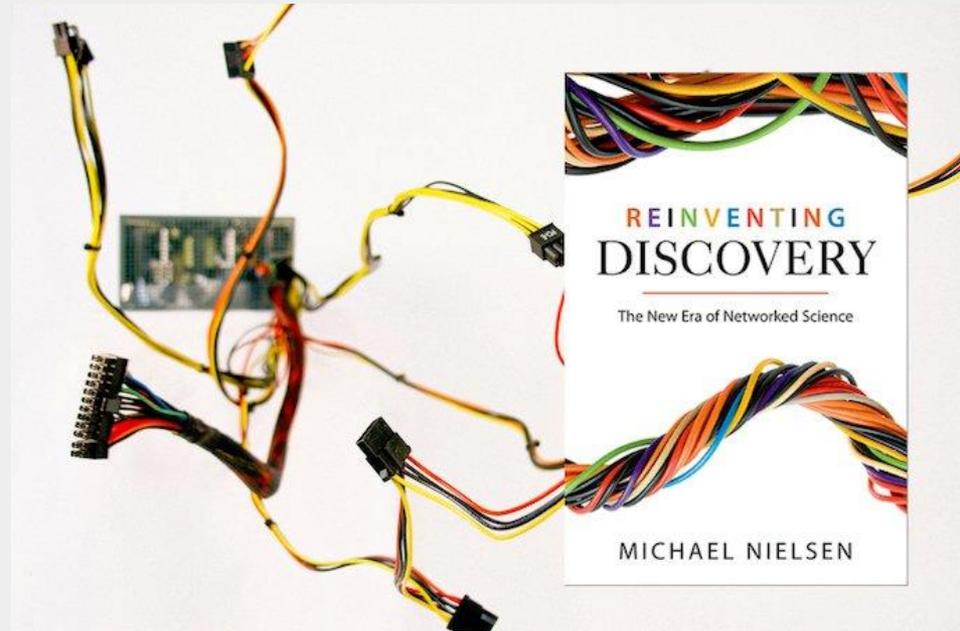
Professor of Economics (Emeritus) & Senior Fellow of SIEPR,
Stanford University
Stanford, California · 124 connections

All Souls College, Oxford

Join to Connect

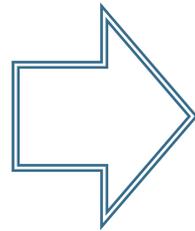
The term was originally coined by
Prof. P.A. **David** in **2003**.

Open Science (OS) is an umbrella notion encompassing various movements aiming to remove the barriers to knowledge creation and dissemination by maximising openness at each stage of the research life cycle thanks to the networking benefits of ICT.



The concept has been popularised by
M. **Nielsen** in **2011**.

5 schools of understanding



Open Science

Democratic school

Access to knowledge is equally distributed to everyone

Pragmatic school

More efficient knowledge creation process based on scientific collaborations

Infrastructure school

scientific web platforms, tools, and services

Public school

Citizens engagement (“citizen science”) and science popularisation (blogs, Twitter etc.)

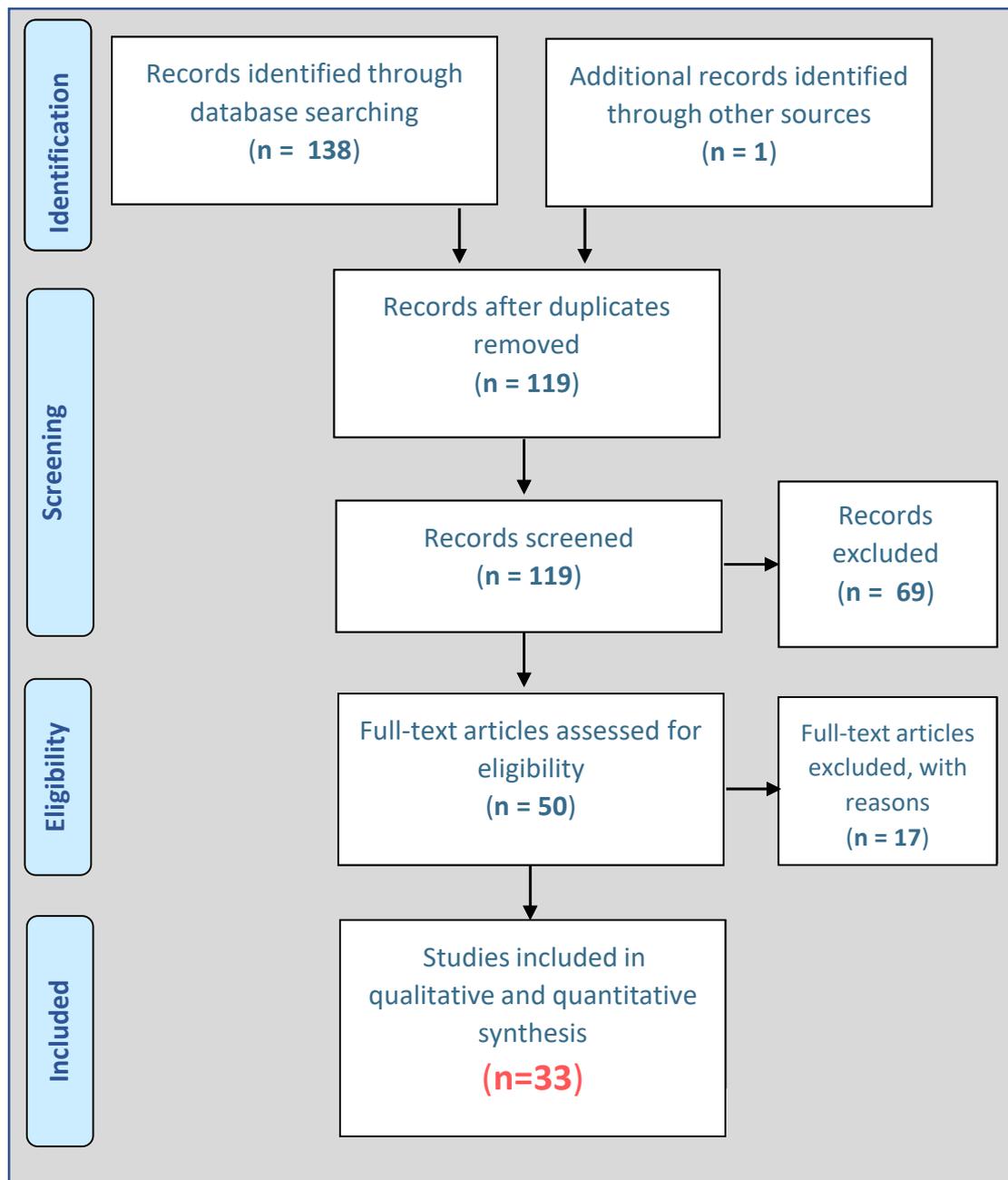
Measurement school

Alternative metric system for scientific impact

3. Research problem Research questions Methods



Mixed-method systematic literature review



PRISMA
Flow Diagram
protocol

WEB OF SCIENCE™

EBSCO

Scopus®



5 years:

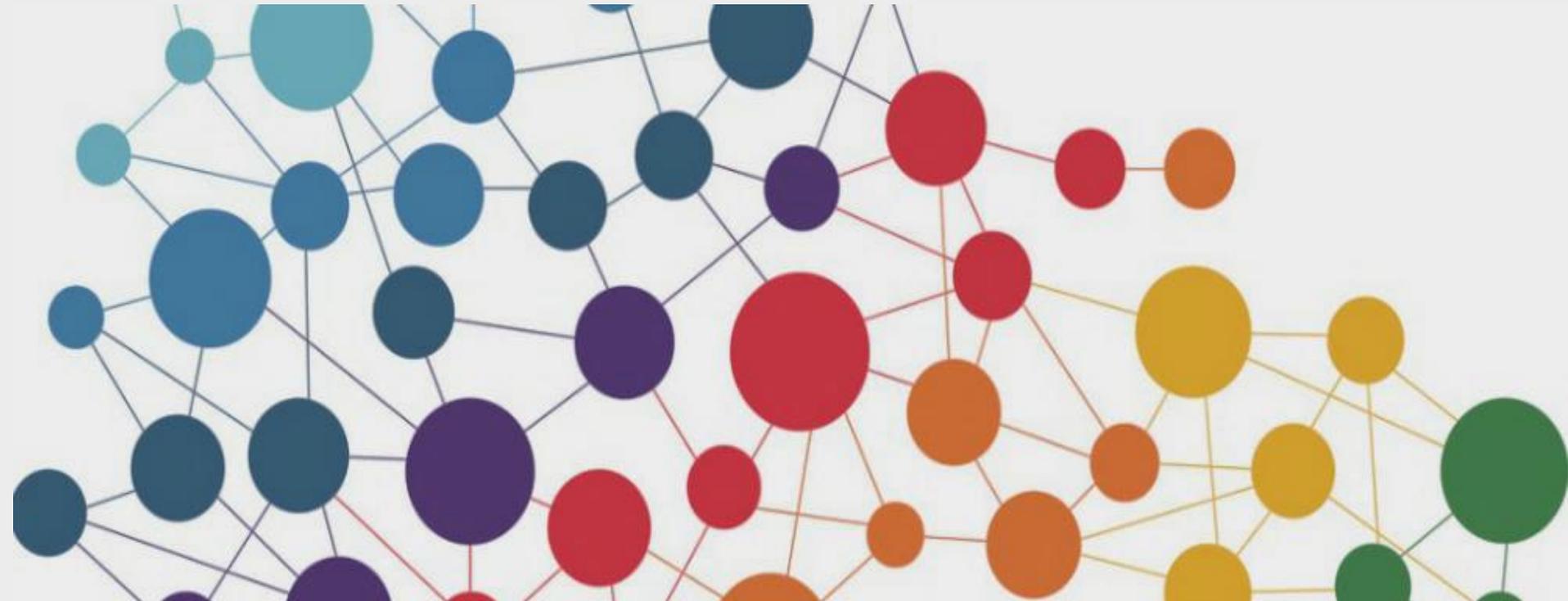
Jan. 2014~

Aug. 2019

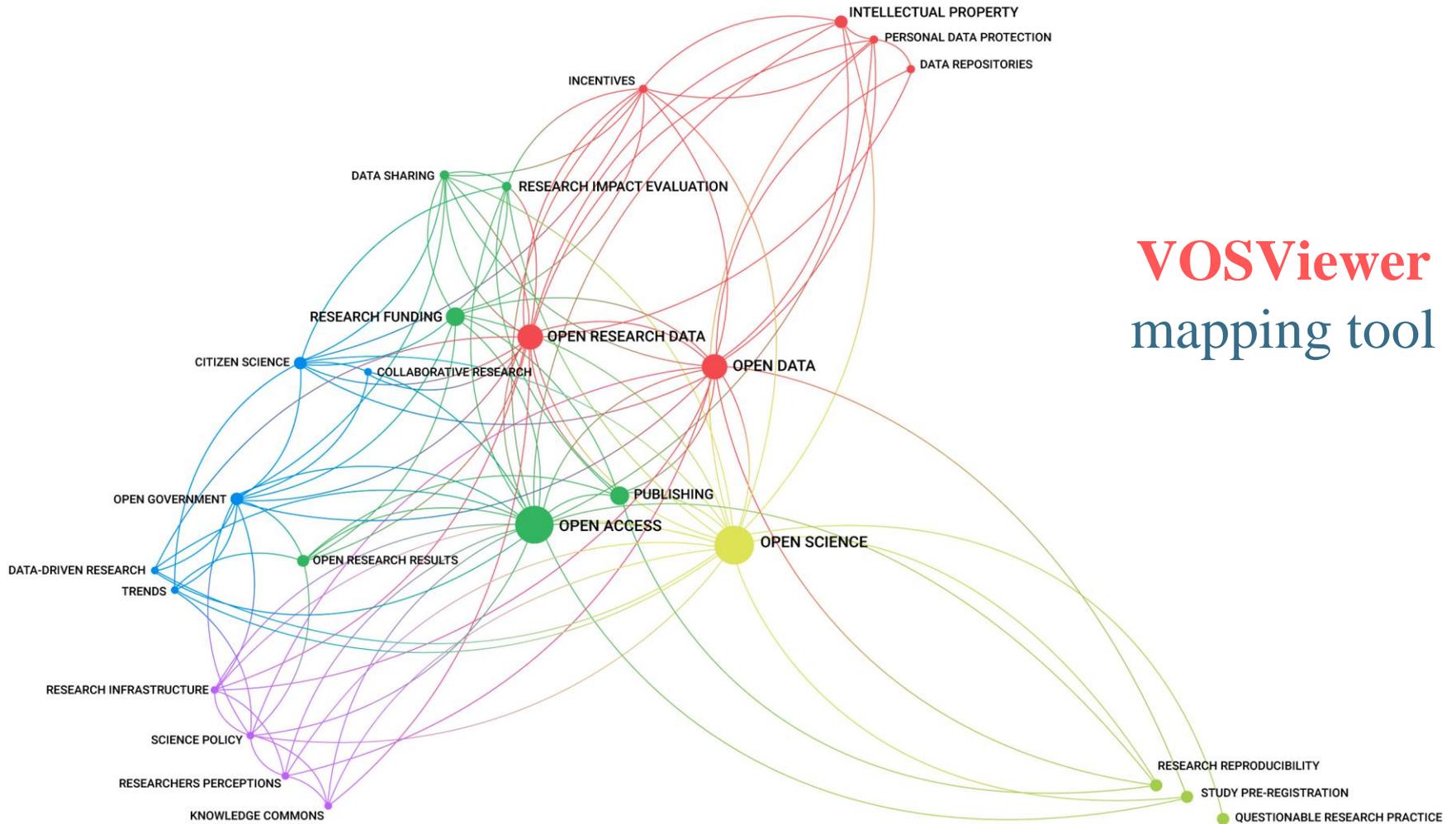
Research questions	Methods	Outcomes
<p>RQ1. What are the most important sub-concepts of OS?</p>	<p>Network analysis of keyword co-occurrence by building an adjacency matrix</p>	<p>Concept map (VOSViewer science mapping tool)</p>
<p>RQ2. What are the main fields of knowledge (disciplines) under which OS is studied?</p>	<p>Identification of each paper's knowledge field(s): international scientific databases and UNESCO standard classifications, authors' affiliations; Frequency calculation</p>	<p>Frequency pie chart</p>
<p>RQ3. What are the key characteristics and categories of OS data?</p>	<p>Qualitative Literature review</p>	<p>Descriptive taxonomy of OS data characteristics and categories</p>
<p>RQ4. What is the relationship, if any, between OS data and OG Data?</p>	<p>Qualitative Literature review</p>	<p>Evidence of the relationship between OS data and OG Data</p>

4. OS Taxonomy

based on keyword co-occurrence
network analysis



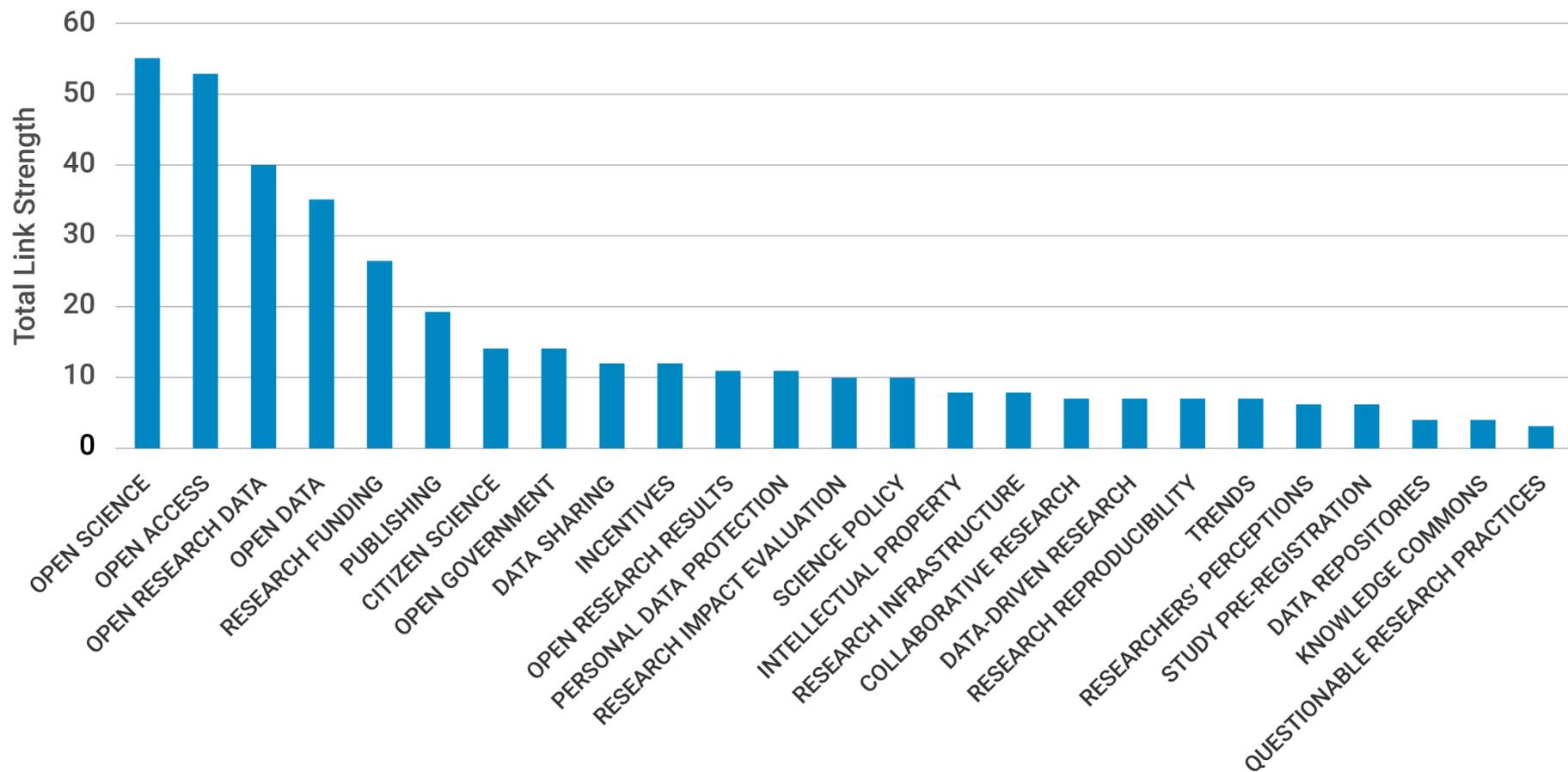
OS Concept Map



VOSViewer
mapping tool

The most important keywords (concepts) related to OS

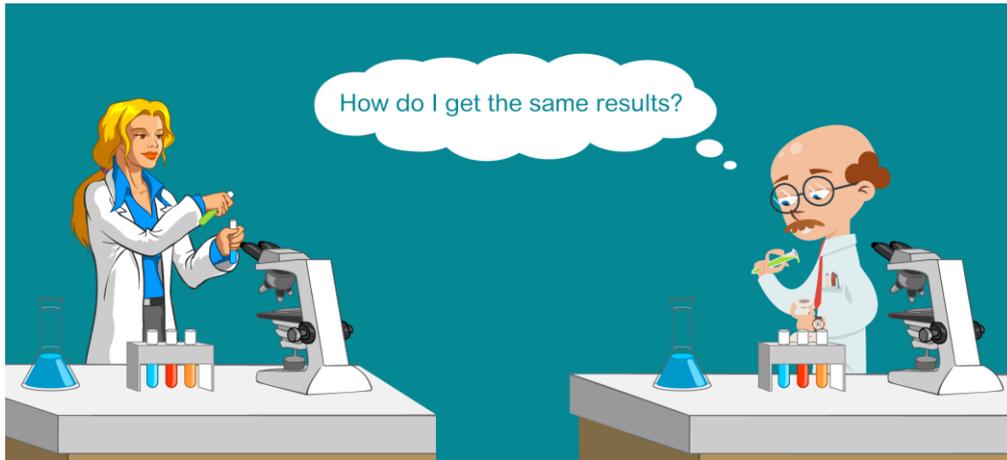
based on nodes' Total Link Strength



The most compatible concepts in OS research (cluster analysis)



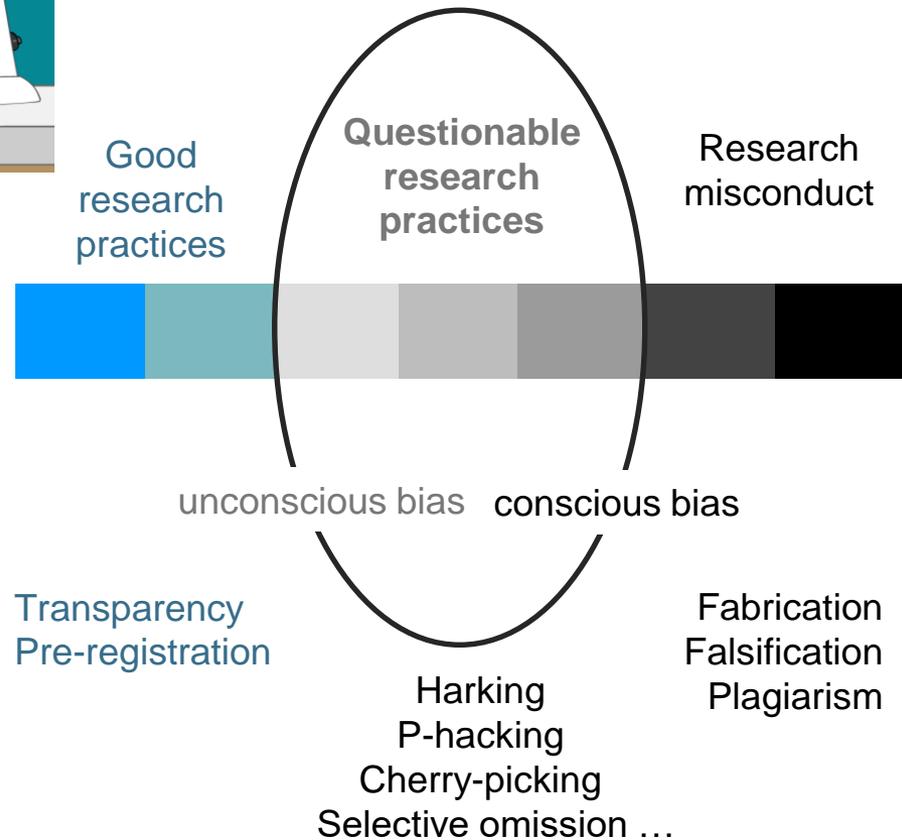
The main idea of Cluster #1 research



Research reproducibility 

Prevent “questionable research practices” 

How? via **Open Science** 





Began in 1990s

The main idea of Cluster #2 research

OPEN ACCESS is a mechanism by which research results (published articles) are distributed *online, free of charge*, and free of barriers to *copying / reuse* by applying a public copyright license, such as Creative Commons.

Major international events in the OA movement:

Budapest OA Initiative
(2002)

Bethesda Statement on OA Publishing
(2003)

Berlin Declaration on OA to knowledge
(2003)

Motivation:

- rising journal subscription cost
- publicly funded research (taxpayers money)
- improve visibility of research impact

The main idea of Cluster #3 research

Open data used as a primary source for scientific research and necessary to validate research findings. For example, genome sequence data, observational or experimental data from clinical trials, survey data, computational models, etc.



IPR protection mechanisms are not well-adapted to research data (if to compare with publications)

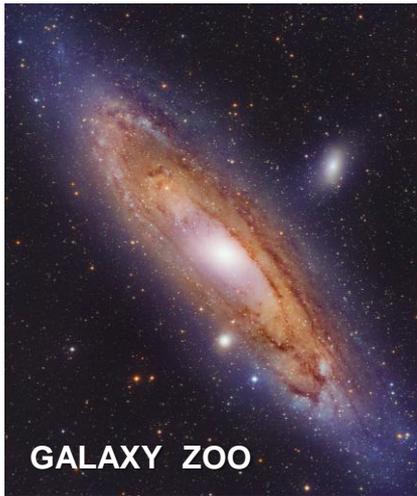
Personal Data protection is critical issue

Need to develop **Incentives** for researchers to share research data



The main idea of Cluster #4 research

Anyone can be a scientist!



[video](#)

CITIZEN SCIENCE is a form of open collaboration where members of the public participate in the scientific process, including identifying research questions, collecting and analyzing the data, interpreting the results, and problem solving [5].



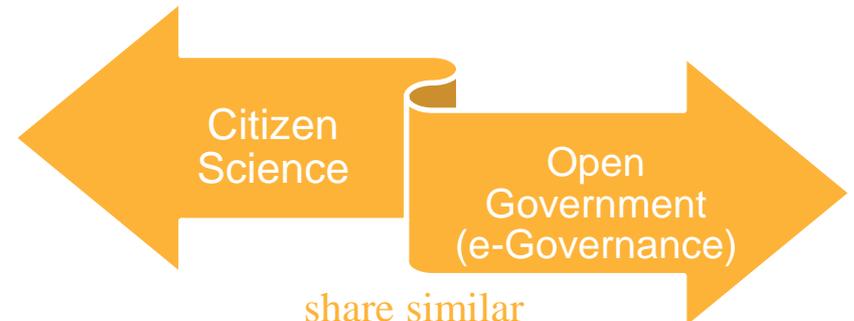
[video](#)



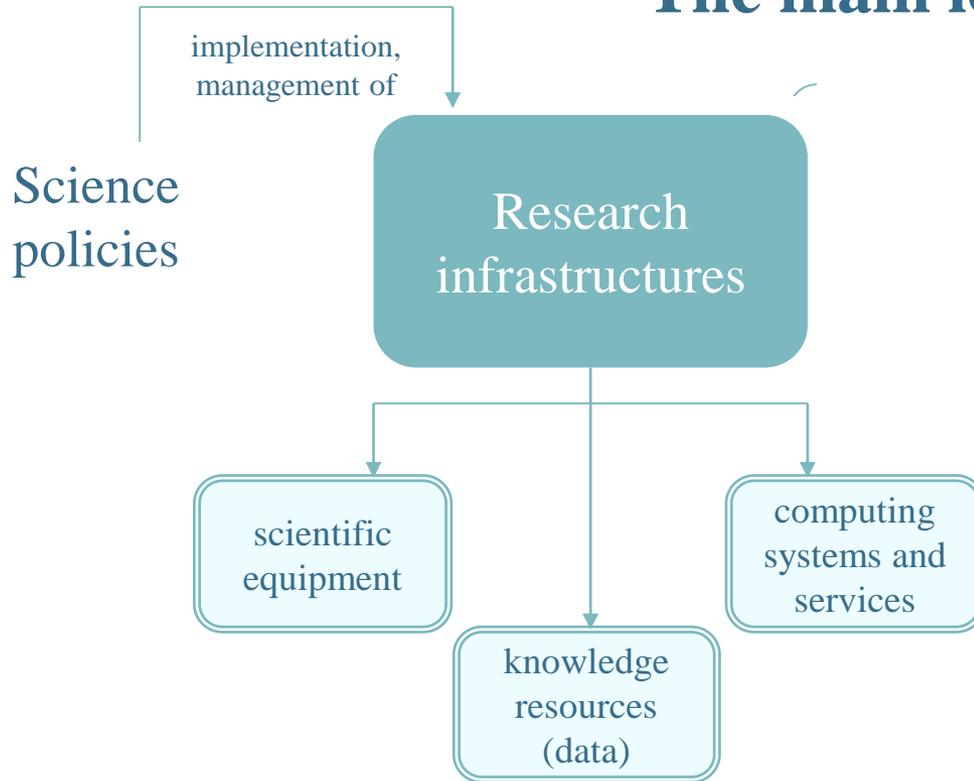
Collaborative research



Public participation



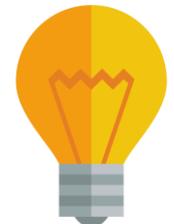
The main idea of Cluster #5 research



resources jointly used and managed by a group of people, meaning that there are no private appropriation and commercial distribution of such resources for the group members.



Traditional Commons
(*G. Hardin*, 1965)



Knowledge as Commons
(*Elinor Ostrom* et al.)



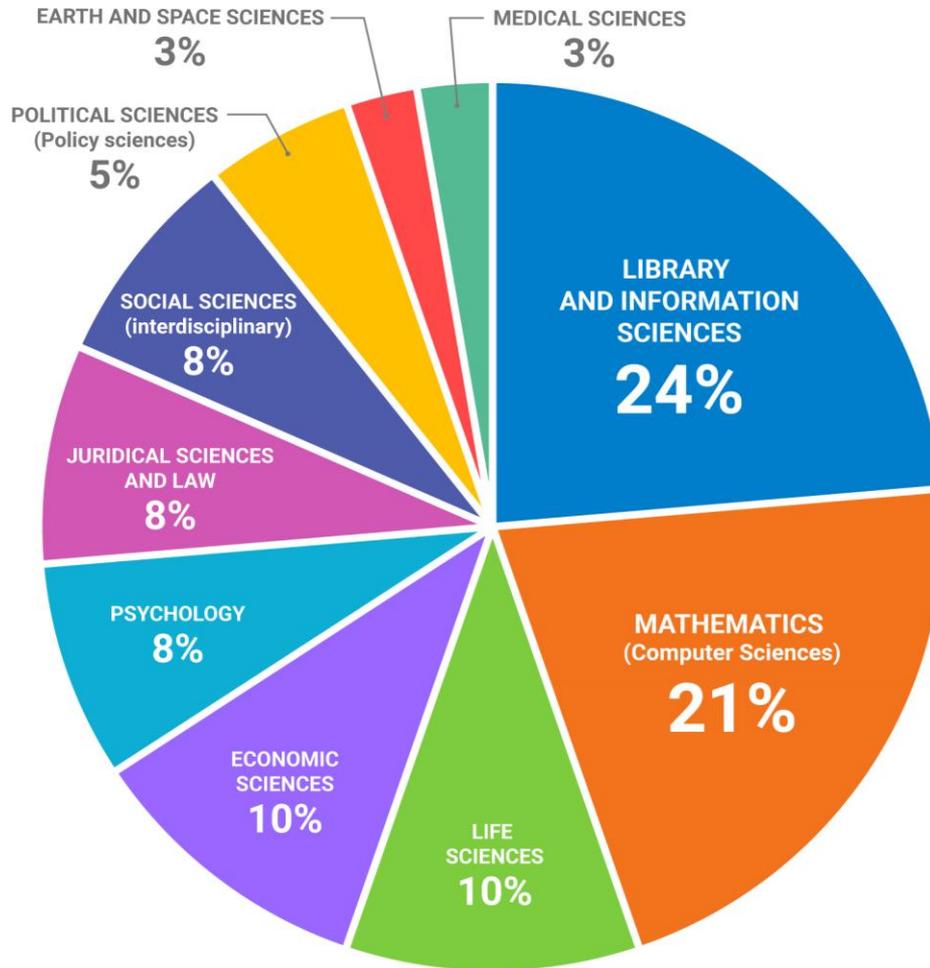
EPOS - pan-European infrastructure for solid Earth science
(<https://www.epos-ip.org>), [video](#)

A close-up photograph of two hands holding two interlocking puzzle pieces. The hands are positioned on the left and right sides of the frame, with fingers gripping the edges of the pieces. The puzzle pieces are a light, warm color, possibly wood or a similar material, and are set against a bright, golden-yellow background that creates a soft, glowing effect. The lighting is warm and directional, highlighting the texture of the hands and the edges of the puzzle pieces.

5. Fields of knowledge

related to OS research

Open Science - a domain for interdisciplinary research



Fields of OS research (frequency distribution)

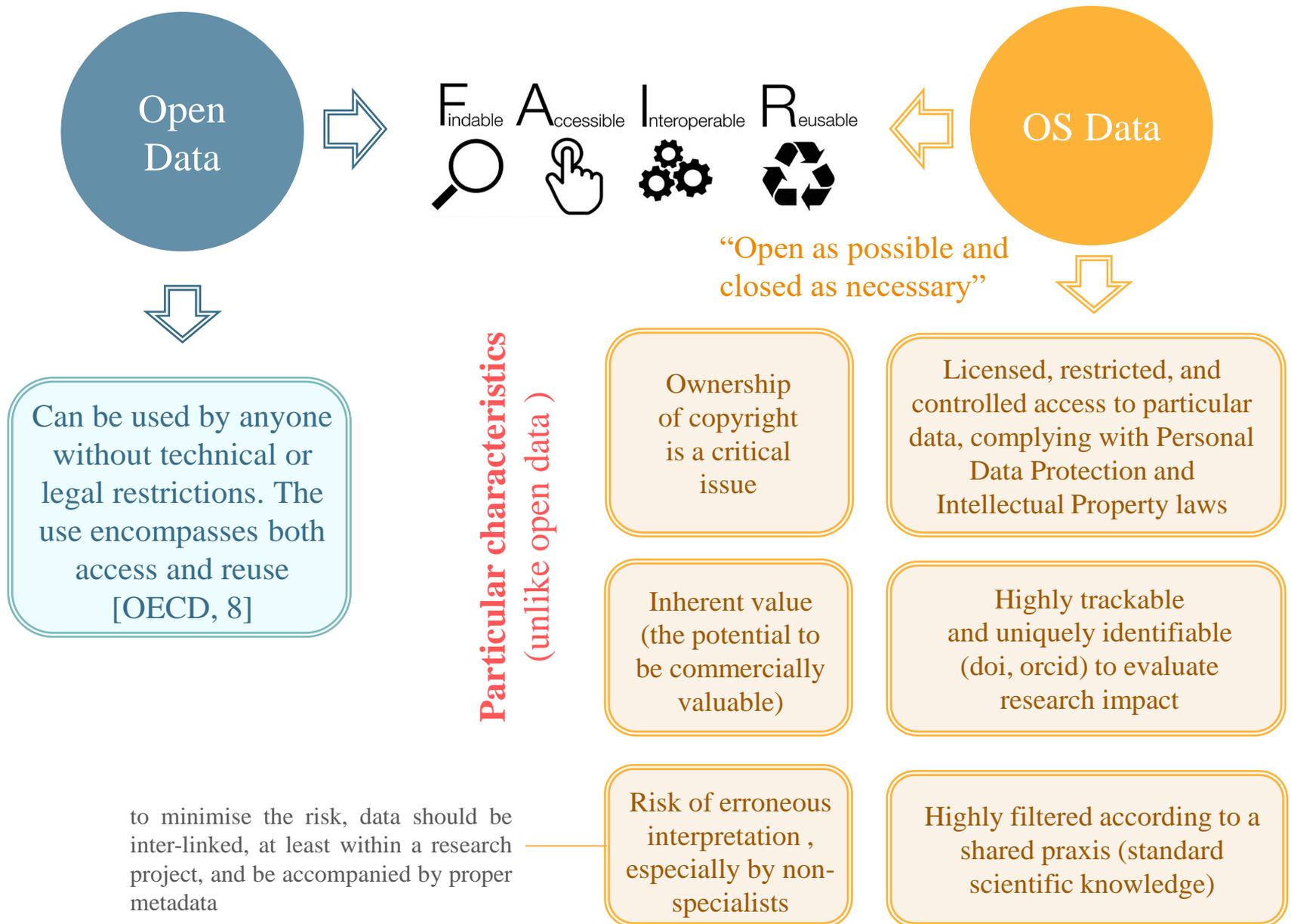
Despite the differences between individual fields in the research focus, the state of development and the context of OS practices, a shared praxis has started to emerge within these fields.

Shared praxis is a common piece of knowledge and its practical application, which might be used for interdisciplinary collaboration on further development of OS domain.



6. OS data characteristics and categories

How is OS data different from open data?



**Study
pre-registration
plans / reports**

**Open
research
data**

**OPEN
SCIENCE
DATA**

**Open
research
results**

OS practices have been mostly associated with publicly-funded research project results

**Posts and
comments
on scientific
network
platforms**

**Open
research
software codes**

Funded research results

Government funders create centralised national repositories to provide public access to the funded research results, such as research reports, peer-reviewed published papers, patents

The U.S: classic.ntis.gov

Japan: kaken.nii.ac.jp

R&D의 모든 것!

공지사항 제1회 NTIS 동영상 콘텐츠 공모전 수상작 발표 2019.09.27

Korea: ntis.go.kr

EU: cordis.europa.eu

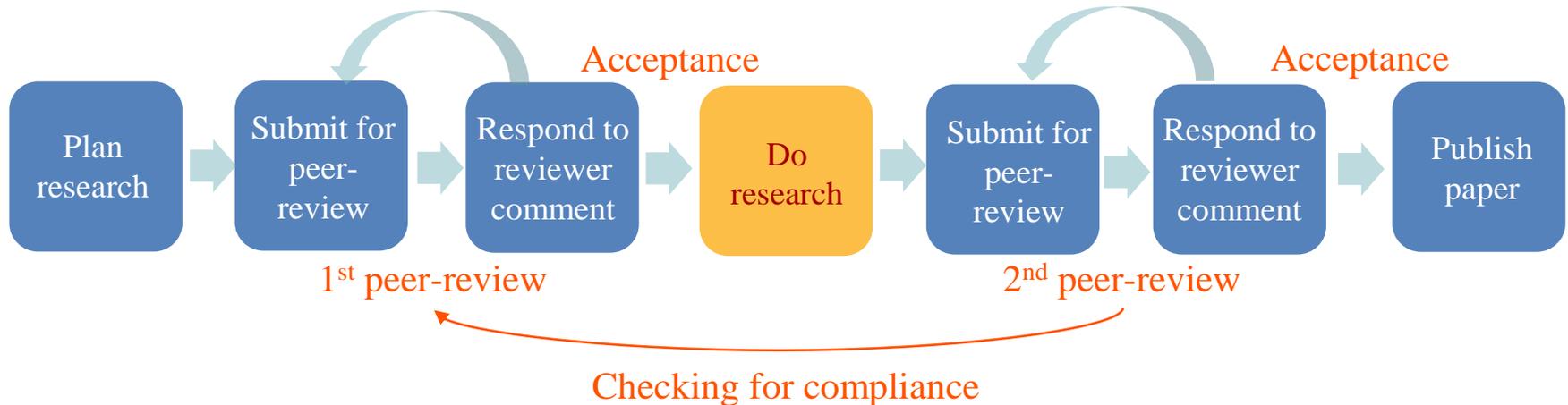
Study pre-registration reports

The formal documentation of the study design, methods, analysis plans, and hypothesis prior to commencing the research, stored in a read-only public repository

Traditional publishing



Registered reports



Study pre-registration reports

NIH U.S. National Library of Medicine

ClinicalTrials.gov

319,012 pre-registered trials
from 209 countries in the world

Find Studies ▾

About Studies ▾

Submit Studies ▾

ClinicalTrials.gov is a database of privately and publicly funded clinical studies conducted around the world.

run by the United States National Library of Medicine

- So far rarely practiced outside **Psychology and Clinical Studies**.
- Pre-registration of clinical trials (controlled clinical investigation) is **required by law in the U.S** and a condition for publication them in relevant journals

provides service for a study pre-registration for research across all disciplines



 OSFHOME ▾

Search

Support

Donate

Sign Up

Sign In



OSF
PREREGISTRATION

Improve your research with preregistration. By writing out specific details such as data collection methods, analysis plans, and rules for data exclusion, you can make important decisions early on and have a clear record of these choices. This can help reduce biases that occur once the data are in front of you.

Use OSF Registries to discover previously registered work.

Preregister

Open research data



Korean Research Data Platform
(dataon.kisti.re.kr) run by KISTI



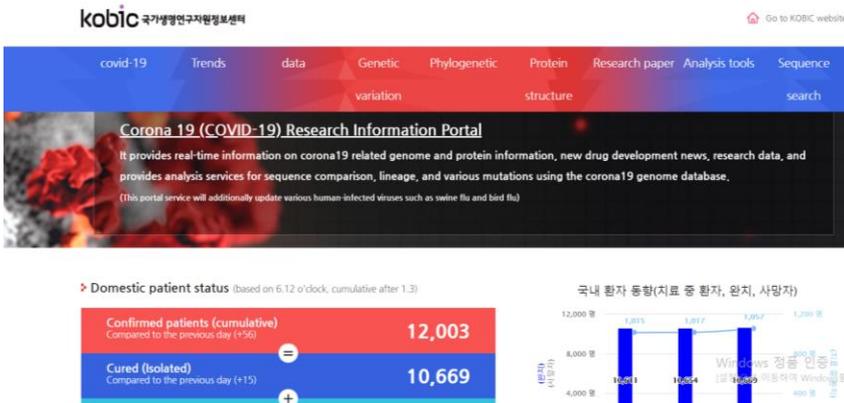
Deposit and share your data. Get academic credit.

Harvard Dataverse is a repository for research data. Deposit data and code here.

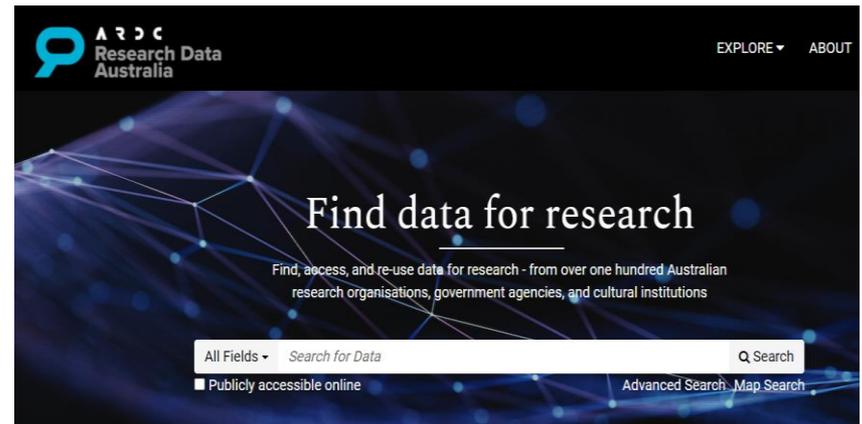
91,532 datasets 9,052,965 downloads

Add a dataset +

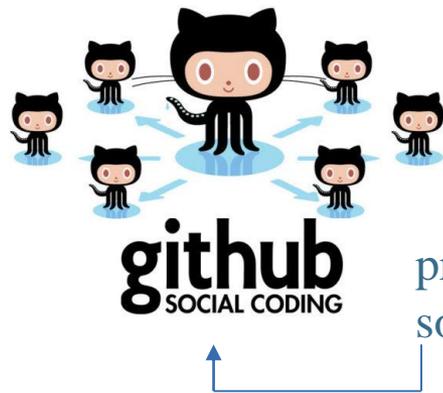
dataverse.harvard.edu



Korean Bioinformatics Center (KOBIC)'s
COVID-19 research data portal
(https://www.kobic.re.kr/covid19/go_data_view)



Research Data Australia portal
(researchdata.edu.au) run by the Australian
Research Data Commons (ARDC)



Open source code & collaborators' inline comments in pull requests

provides services for maintaining, sharing, and collaborating on source code, which can be used in scientific research

A screenshot of a GitHub pull request interface. At the top, the file path is "app/views/site/campaigns/building_the_future.html.erb". The code editor shows several lines of ERB code. Lines 224-227 are highlighted in green, indicating a suggested change. Below the code, a comment from user "kimestoesta" is visible, stating "Sections could be a little wider try:". Below the comment, a "Suggested change" box shows a diff: line 232 is being replaced with a wider column class. A "Give us feedback" link is next to the suggested change box, and an "Apply suggestion" button is at the bottom right of the diff box.

```
14 app/views/site/campaigns/building_the_future.html.erb
```

```
221 221 <% end %>
222 222 </div>
223 223
224 + <% if btf_campaign_tiffani_enabled? && |btf_campaign_jamica_enabled? %>
225 +   <div class="container-xl mx-auto mb-8">
226 +     <div class="d-flex flex-justify-center flex-lg-items-stretch flex-wrap px-2 px-md-4 px-lg-5">
227 +       <div class="col-sm-8 col-md-6 col-lg-4 p-2">
```

kimestoesta a minute ago
Sections could be a little wider try:

Suggested change ⓘ [Give us feedback](#)

```
232 - <div class="col-sm-8 col-md-6 col-lg-4 p-2">
232 + <div class="col-sm-10 col-md-8 col-lg-6 p-2">
```

[Apply suggestion](#) ▼

reproducible
computational
research

Posts and Comments on:

Science blogs



Gowers's Weblog

Mathematics related discussions

gowers.wordpress.com

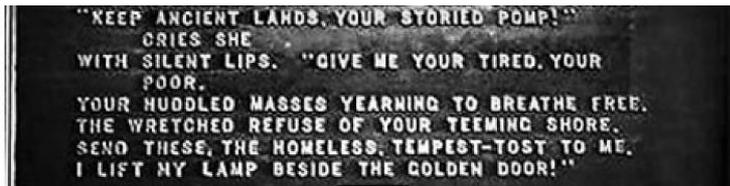
...m solved (probably)

DHJ(3) and related results: 1050-1099 >

project
started
in 2009

Polymath1 and open collaborative mathematics

In this post I want to discuss some general issues that arise naturally in the light of how the polymath experiment has gone so far. First, let me say that for me personally this has been one of the most exciting six weeks of my mathematical life. That is partly because it is always exciting to solve a problem, but a much more important reason is the way this problem was solved, with people chipping in with their thoughts, provoking other people to have other thoughts (sometimes almost accidentally, and sometimes more logically), and ideas gradually emerging as a result. Incidentally, many of these ideas are still to be properly explored: at some point the main collaboration will probably be declared to be over (though I suppose in theory it could just go on and on, since it seems almost impossible to clear up every interesting question that emerges) and then I hope that the comments will be a useful resource for anybody who wants to find some interesting open problems.



What's new

Updates on my research and expository papers, discussion of open problems, and other maths-related topics. By Terence Tao

Subscribe to feed Home About Career advice On writing Books Applets

RECENT COMMENTS

Anonymous on 245A, Notes 5: Differentiation...

A cheap version of nonstandard analysis

2 April, 2012 in expository, math.CA, math.LO | Tags: Bolzano-Weierstrass theorem, nonstandard analysis, sequential compactness

terrytao.wordpress.com

Academic social networks



Citizen science platforms

Galaxy Zoo Talk

Search or enter a #tag

Notes

General comment threads about individual subjects

bluemagi Subject 36258427 *7 minutes ago*

Help

A place to get general help for using the site.

Ine Theunissen DECaLS is not working. *10 hours ago*

Science

A place to discuss the science behind the site, and inspired by the results

bluemagi Norma Jellyfish Galaxy *5 days ago*

Chat

A place for general chat

7. Open Science data = Open Government data ?



Looks quite similar ...

- **Transparency**
- **Participation**
- **Collaboration**



“My administration is committed to creating an unprecedented level of openness in Government.” – Barack Obama

“Memorandum for the Heads of Executive Departments and Agencies – Transparency and Open Government” Jan 2009

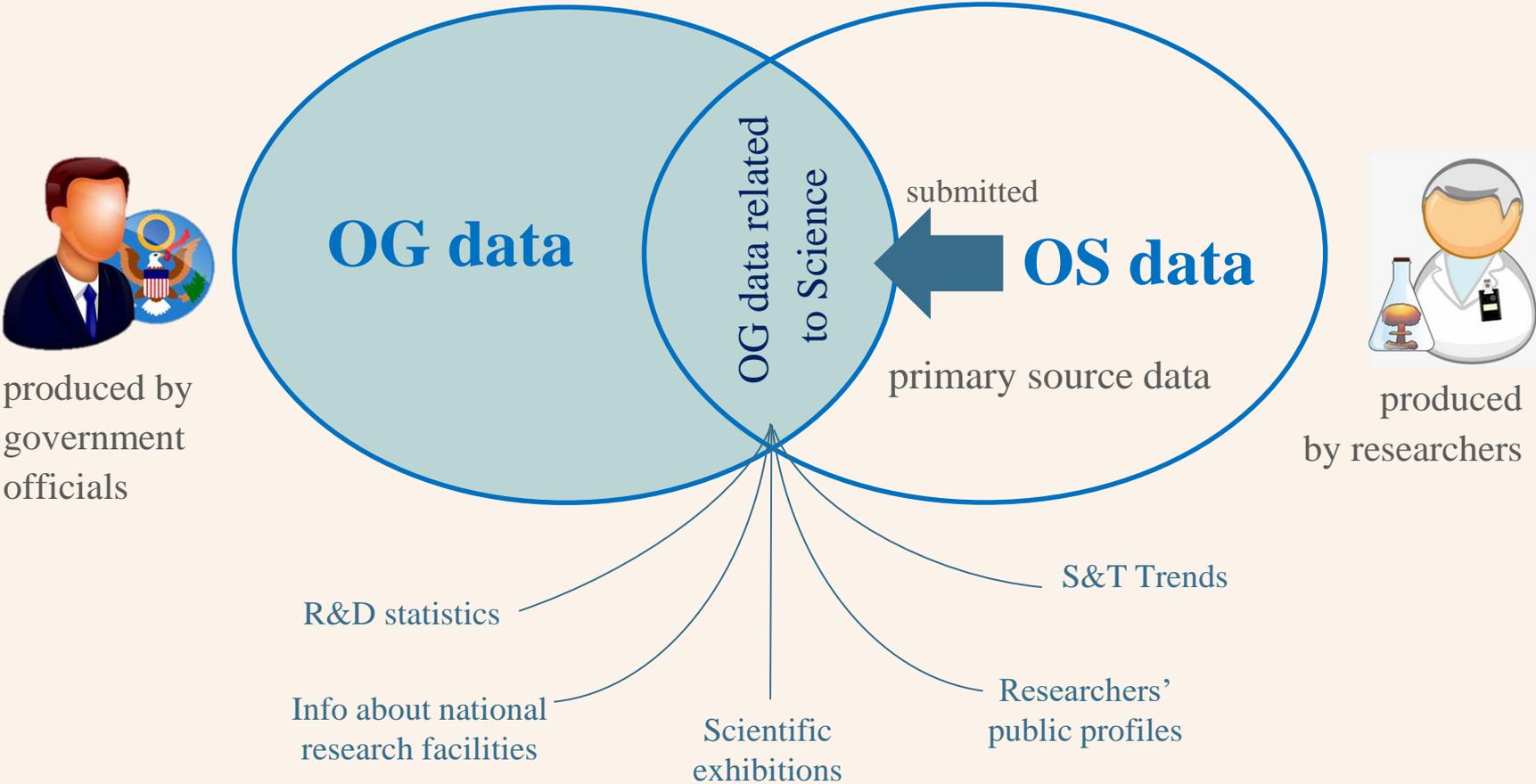
Shared principles

In the U.S., some **Citizen Science** projects and **OS data sharing** initiatives have been developed in line with the Obama Administration’s **Open Government (OG)** Directive of 2009.

Both OG data and OS data are created using taxpayers’ money
(public sector information)

But different in nature

Different management policies mandated by separate laws
Different responsible agencies



8. Future research

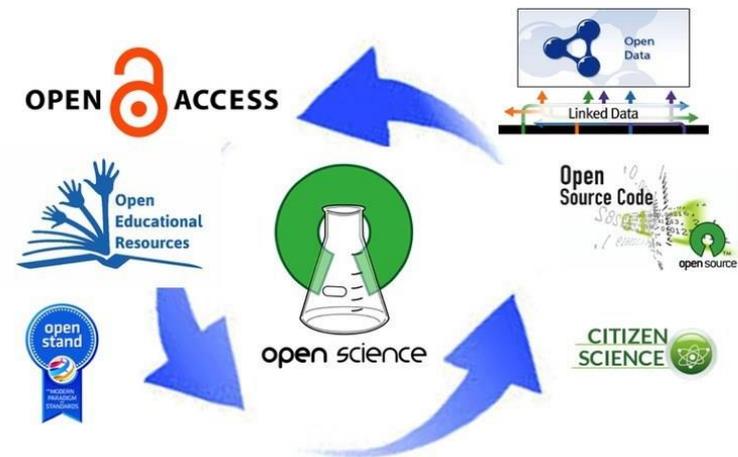


Proposals for Future Research

To develop a conceptual model of the **ideal national OS data ecosystem** that, among other things, improves the quality and effectiveness of scholarly communication during public health emergencies.

To investigate **OS and other data (OG data, private data) fusion opportunities** in the context of public health emergencies.

OS as a tool for combatting public health emergencies



Thank You