OpenAIRE and Open Science

Elli Papadopoulou
OpenAIRE Greek National Open Access Desk
ATHENA Research & Innovation Center

Go to **www.menti.com** and use the code **1755 2241**

MICCAI x OpenAIRE session

Go to

www.menti.com

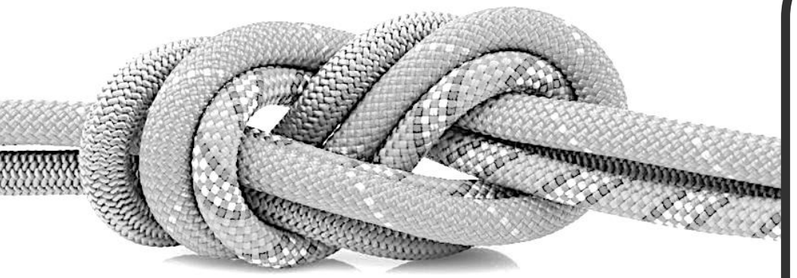
Enter the code

1755 2241

Or use QR code

Open Science

What is Open Science



Spiros Athanasiou et al, 2020. National Plan for Open Science.

Open Science is the **new standard of practices, means and collaboration** for producing and distributing scientific output and research results, with a direct scientific, economic and societal impact

Core principles

1. **Collaboration**
2. **Open Access**
   Read and access scientific information
3. **FAIR* principles**
   Discovery, Interoperability, Reuse
4. **Documentation**
   Transparency, Accuracy

Areas of action

Scientific **outputs**

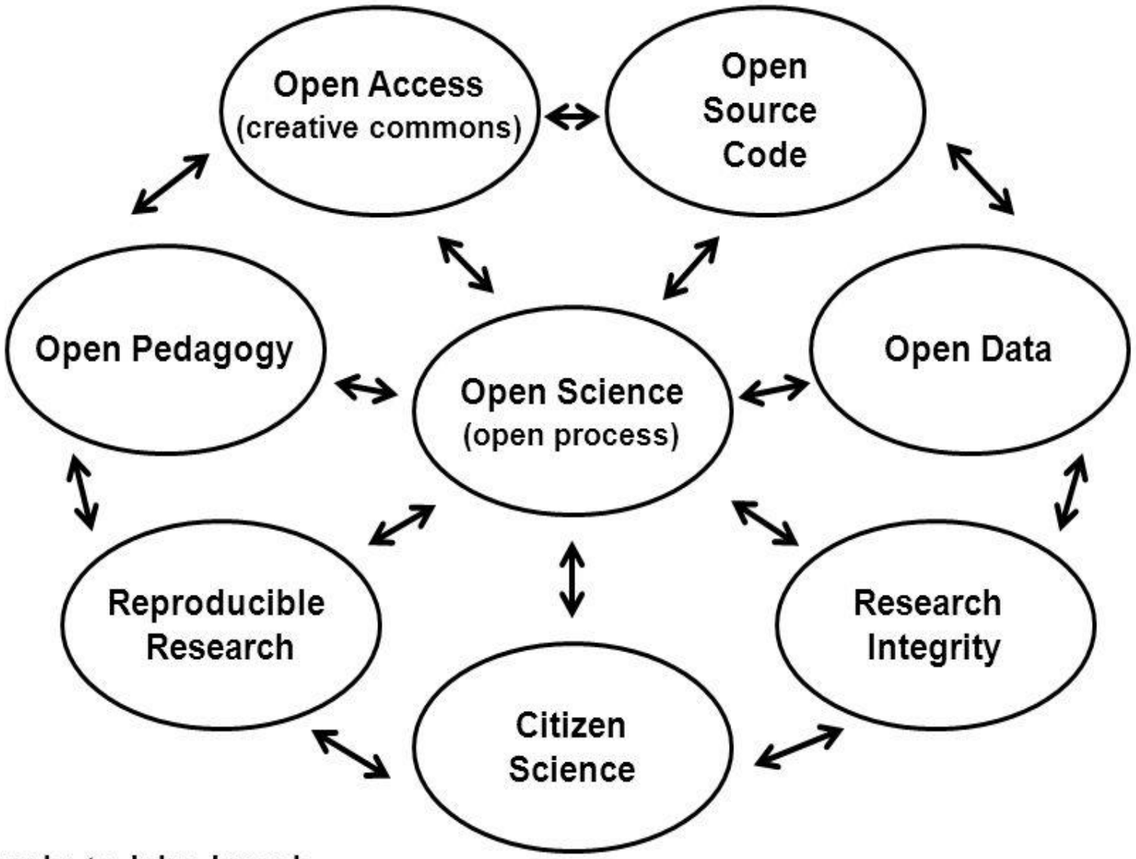
Infrastructures and services for research

Training and new skills

OpenAIRE

ΑΘΗΝΑ
Έρευνα & Καινοτομία
Τεχνολογίες Πληροφορίας

Open Science Ecosystem



With thanks to John Jungck

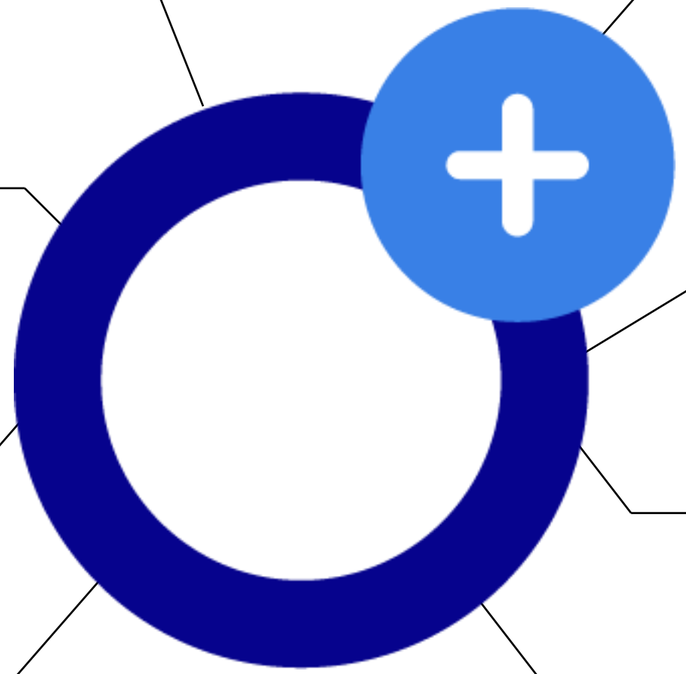
Why choosing Open Science?

Research

- Reach wider audience
- Re-use research outputs
- Validate research
- Prevent information and data loss

Economy

- Stimulate innovation
- Strengthen regional and national markets
- New job openings



Researchers

- Promote integrity
- Increase use of citations and get more credits
- Rewards in the EOSC

Society

- Builds trust <– transparency
- Inclusivity in Science-> Citizen Science
- Collaborations on national and EU level

OpenAIRE

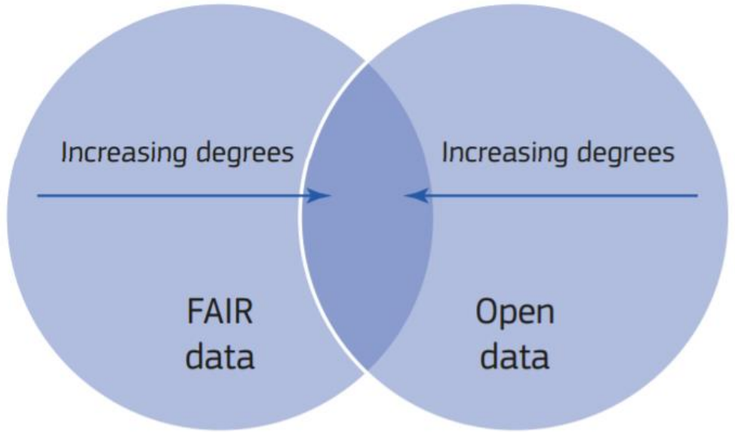
European Commission

ΑΘΗΝΑ
Έρευνα & Καινοτομία
Τεχνολογίες Πληροφορίας

Research Data Management

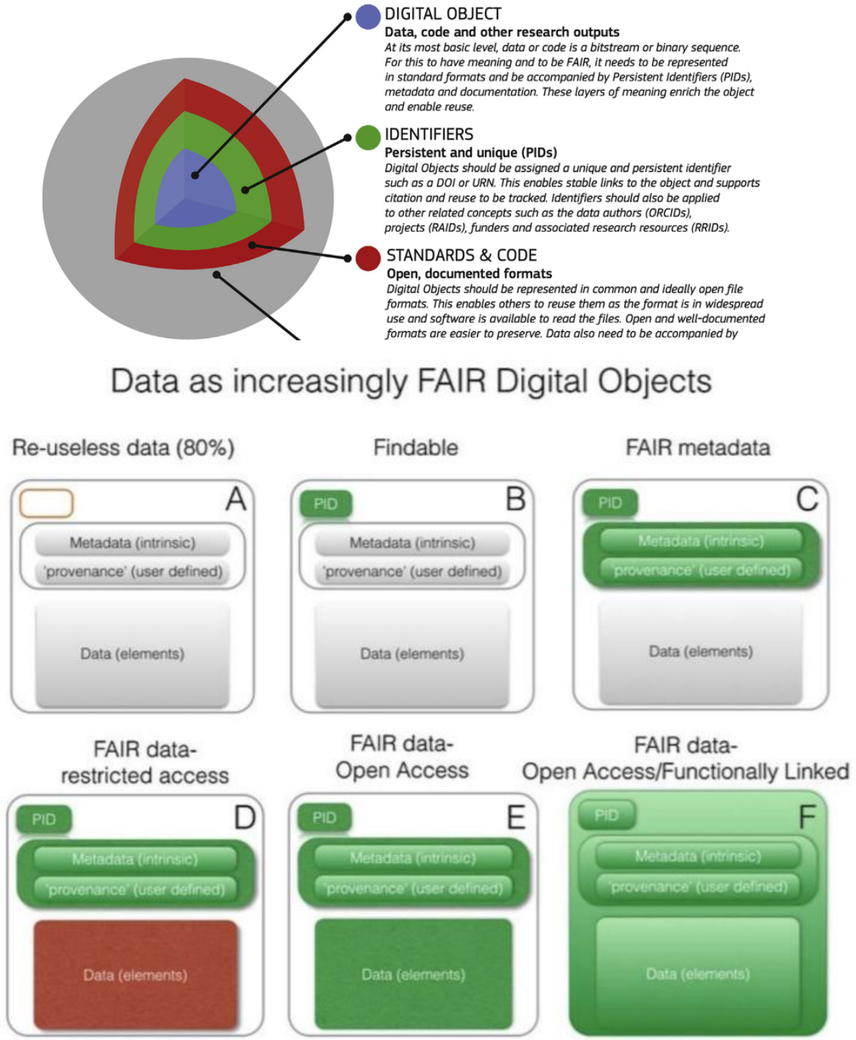
EOSC - The Web of FAIR data

- FAIR-enabling ecosystems



"highly distributed ecosystem requiring technical mechanisms linking resources, and social mechanisms to define specifications, standards and protocols"

FAIR principles for digital objects



https://www.ands.org.au/working-with-data/fairdata/training

RDM lifecycles

https://datamanagement.hms.harvard.edu/about/what-research-data-management/biomedical-data-lifecycle

Table 1
Life cycle stages and identified questions

Stages	Questions
Identifying	▪ What data are available? ▪ What is the current audience for these data? ▪ What potential future audiences exist for these data? ▪ Is this an isolated data set or could it be combined with other sets?
Digitizing	▪ Are the data in digital format? ▪ If no, what would it take to digitize the data? ▪ Are the data in a stable digital format that can be preserved?
Cleaning	▪ How many people have touched or will touch the data? ▪ What rules have been created to ensure consistent data standardization? ▪ What tools am I using to standardize the data?
Describing	▪ Is there a README.txt file outlining the project? ▪ Is there a standard ontology applicable to this data set? ▪ What information would others need to use the data?
Storing and preserving	▪ What access is needed to work with the data now? ▪ Who needs access now? ▪ What are the best storage options for the future? ▪ What is the intended duration of preservation?
Sharing	▪ Are there any privacy concerns about these data? ▪ Who is the owner of this data set? ▪ What institutional policies apply to these data? ▪ How can sharing rights be maximized?
Analyzing	▪ What analysis tools are available? ▪ What are the limitations of the data set?

Goben A, Raszewski R (2015) The data life cycle applied to our own data. J Med Libr Assoc. 103(1): 40–44. doi: 10.3163/1536-5050.103.1.008

Data Management Plans

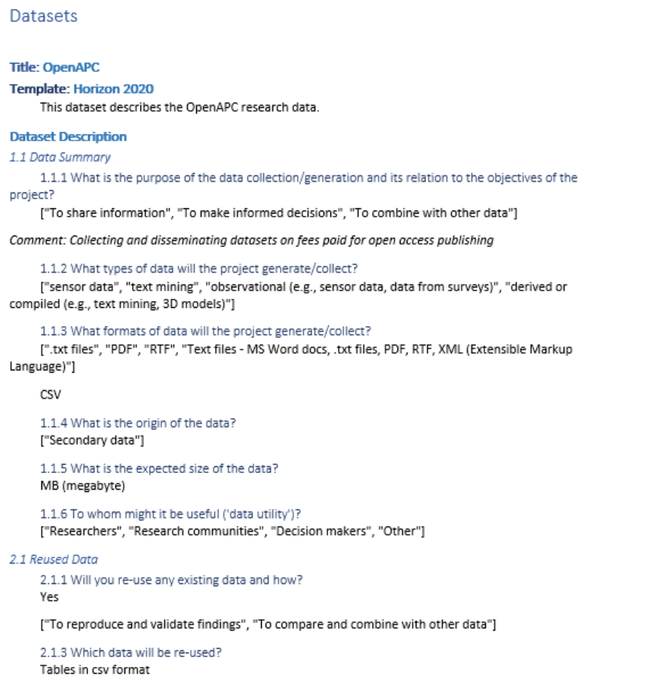
Deliverable and “living” document

- documents processes undertaken throughout data management lifecycle, including costs

≠

What is not a DMP?

Research assessment method



Datasets

Title: OpenAPC

Template: Horizon 2020

This dataset describes the OpenAPC research data.

Dataset Description

1.1 Data Summary

1.1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?

["To share information", "To make informed decisions", "To combine with other data"]

Comment: Collecting and disseminating datasets on fees paid for open access publishing

1.1.2 What types of data will the project generate/collect?

["sensor data", "text mining", "observational (e.g., sensor data, data from surveys)", "derived or compiled (e.g., text mining, 3D models)"]

1.1.3 What formats of data will the project generate/collect?

[".txt files", "PDF", "RTF", "Text files - MS Word docs, .txt files, PDF, RTF, XML (Extensible Markup Language)"]

CSV

1.1.4 What is the origin of the data?

["Secondary data"]

1.1.5 What is the expected size of the data?

MB (megabyte)

1.1.6 To whom might it be useful ('data utility')?

["Researchers", "Research communities", "Decision makers", "Other"]

2.1 Reused Data

2.1.1 Will you re-use any existing data and how?

Yes

["To reproduce and validate findings", "To compare and combine with other data"]

2.1.3 Which data will be re-used?

Tables in csv format

Open Science in Europe

Article 29

Article 29.2 Grant Agreement
- Open Access
eck embargo periods)

Publications

"As open as possible,
as closed as necessary"

Article 29.3 Grant A
- Open & FAIR data
- DMP "living document"

Research Data

Opting out options

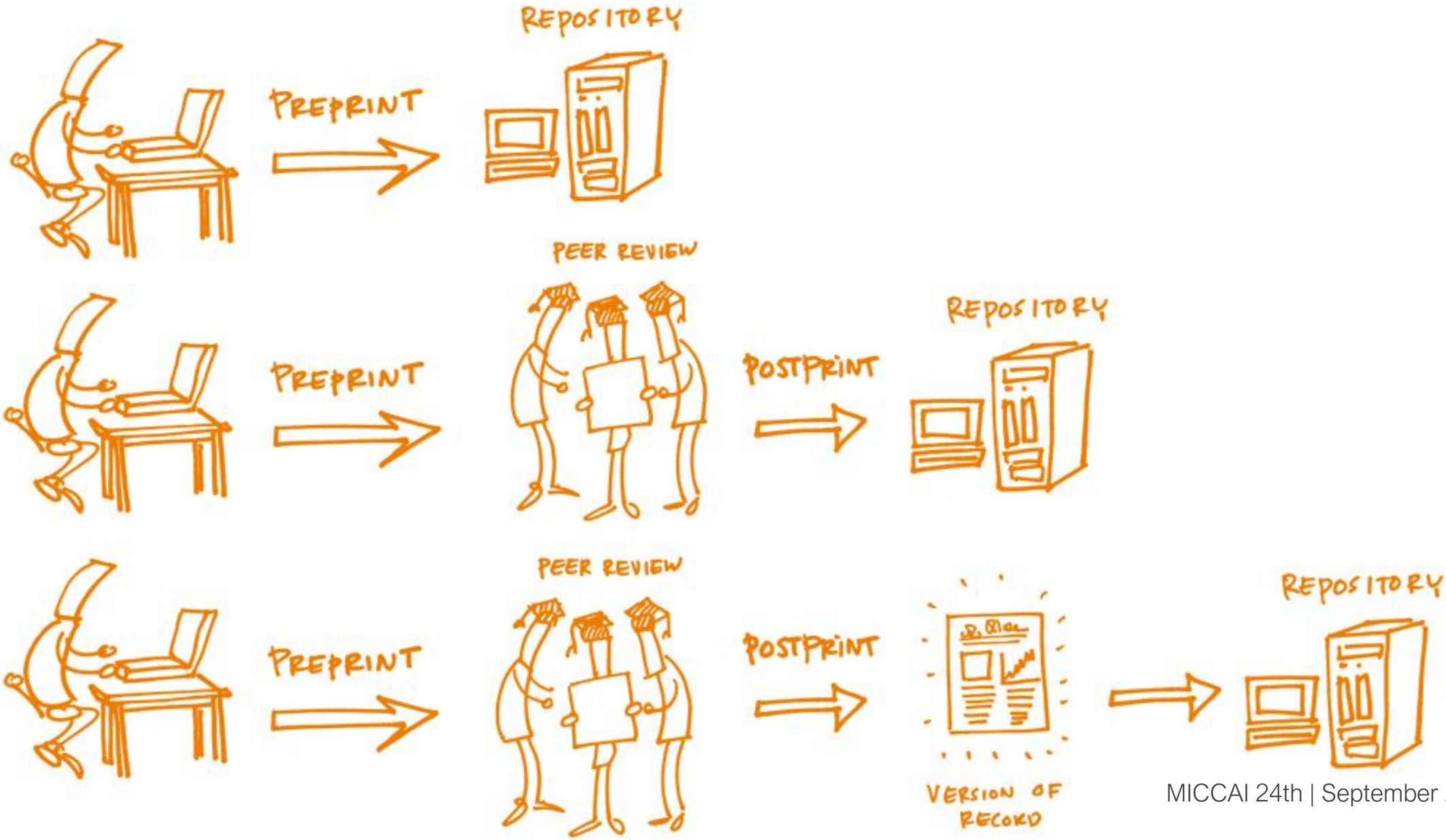
Access and preservation

- **Deposit a version of the publication in a literature repository and ensure open access**
  - Embargoes may apply: 6 months for STEM or 12 months for SSH

- **Publish research in an Open Access journal**
  - Article Processing Charges (APCs)

+ open access to metadata!

MODES OF SELF-ARCHIVING

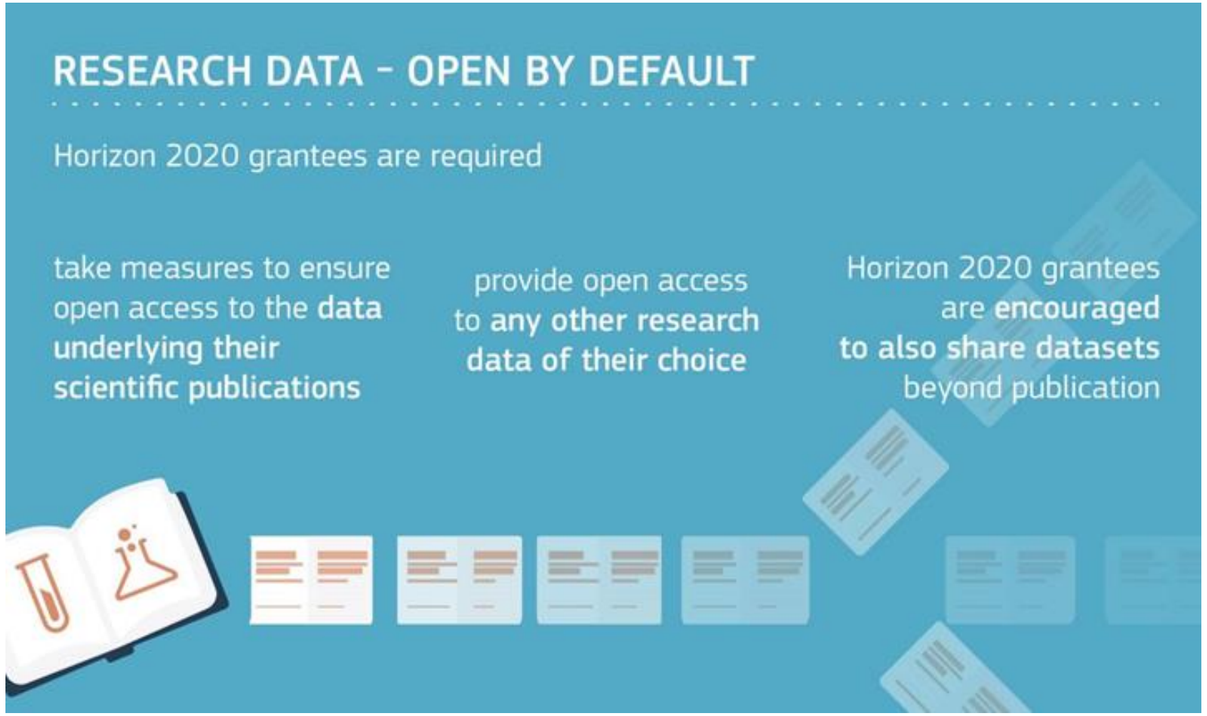


Acknowledge funding

When you deposit, you must also ensure open access to the descriptive metadata that identify the deposited publication. This metadata must be in a standard format and must include all of the following:

- the words ["European Union (EU)" and "Horizon 2020"] ["Euratom" and Euratom research and training programme 2014-2018"];
- the name of the project, acronym and grant number;
- the publication date, and length of embargo period if applicable, and
- a persistent identifier.

Research data



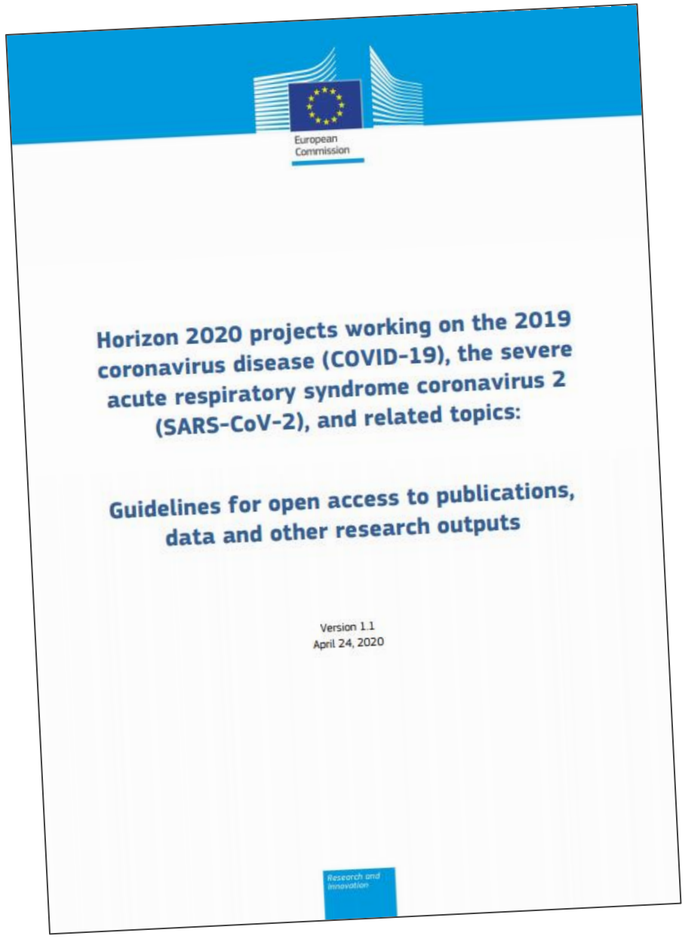
1. Develop a Data Management Plan (DMP)
2. Deposit research data in data repositories
3. Provide open access to research data, if possible.
4. Make data FAIR

COVID-19 research

Open access data benefits millions of scientists around the world and is essential for a rapid response to the COVID-19 pandemic

Artist's impression of open access COVID-19 data sharing. Credit: Spencer Phillips, EMBL-EBI.

Open Access for COVID-19 research


European Commission

Horizon 2020 projects working on the 2019 coronavirus disease (COVID-19), the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and related topics:

Guidelines for open access to publications, data and other research outputs

Version 1.1
April 24, 2020

Research and Innovation

- **All peer-reviewed research publications** relevant to the outbreak are made immediately open access, or freely available at least for the duration of the outbreak;
- Research findings relevant to the outbreak are shared **immediately with the WHO upon journal submission**, by the journal and with author knowledge;
- Research findings are made available via preprint servers before journal publication, or via platforms that make papers openly accessible before peer review, with clear **statements regarding the availability of underlying data**;
- Share interim and final research data, **together with protocols and standards** used to collect the data, as rapidly and widely as possible - **including with public health and research communities and the WHO**;
- Authors are clear that data or preprints shared ahead of submission will not pre-empt its publication in these journals.

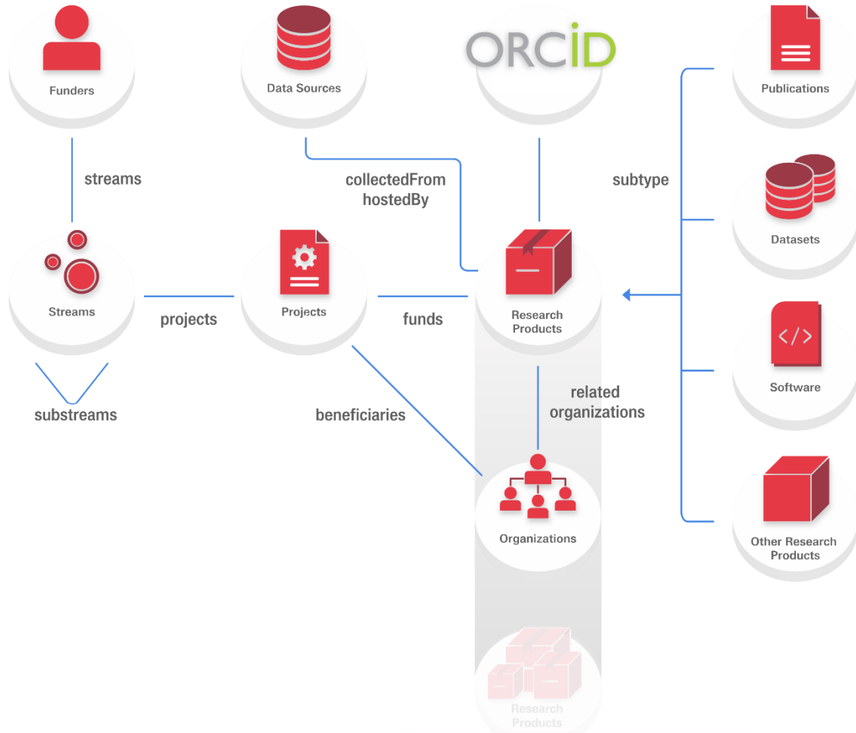
https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-guidelines-oa-covid-19_en.pdf

OpenAIRE

About OpenAIRE

Paneuropean electronic infrastructure: added value horizontal services

- **Distributed** and **participatory** electronic infrastructure
  - Based on existing national infrastructures
  - Interconnection of various actors across Europe
- **Linked** Open Science
  - Support for all types of research entities and outputs
- **Human** network
  - National nodes in more than 34 countries



EXPLORE OpenAIRE RESEARCH GRAPH

OpenAIRE services in EOSC

1

Guidelines

Setting compliance of EOSC data sources

Data Source

OpenAIRE Guidelines for DATA Repositories

OpenAIRE Guidelines for OTHER research products

OpenAIRE Guidelines for CRIS

PROVIDE

Register
Validate

3

Global sources

EOSC scientific product catalogue

OpenAIRE RESEARCH GRAPH

DEVELOP

EXPLORE

MONITOR

5

Funder - Institution – Community OS Monitoring

EUROPEAN OPEN SCIENCE CLOUD

Findable Accessible Interoperable Reusable

2

argos

zenodo

CONNECT

4

Research Community Gateways

amnesia

OpenAIRE

ΑΘΗΝΑ Έρευνα & Καινοτομία Τεχνολογίες Πληροφορίας


Policies
Aligning
Training
Empowering
Services
Connecting
OpenAIRE

OpenAIRE RESEARCH GRAPH

Putting research into context.
Making the connections.

explore.openaire.eu

USAGE ANALYTICS

Make your content count

provide.openaire.eu

zenodo

a catch-all-repository

www.zenodo.org

OpenAIRE | MONITOR

Simplify research tracking & monitoring

monitor.openaire.eu

amnesia

Anonymize your datasets

amnesia.openaire.eu

OpenAIRE | CONNECT

Build an Open Research Gateway for your community

connect.openaire.eu

OA BROKER

Routing information across content providers

provide.openaire.eu

ScholeXplorer

Interlinking data and publications

scholexplorer.openaire.eu

METADATA VALIDATOR

Helping you apply the OpenAIRE Guidelines

provide.openaire.eu

OpenAIRE | EXPLORE

Discover and share research

explore.openaire.eu

OpenAIRE | DEVELOP

APIs to access the OpenAIRE Research Graph

develop.openaire.eu

argos

Plan and follow your data

openaire.eu/argos

Services for all

Resources

How to comply with H2020 mandates - publications

How can identifiers improve the dissemination o your research outputs?

Legal issues

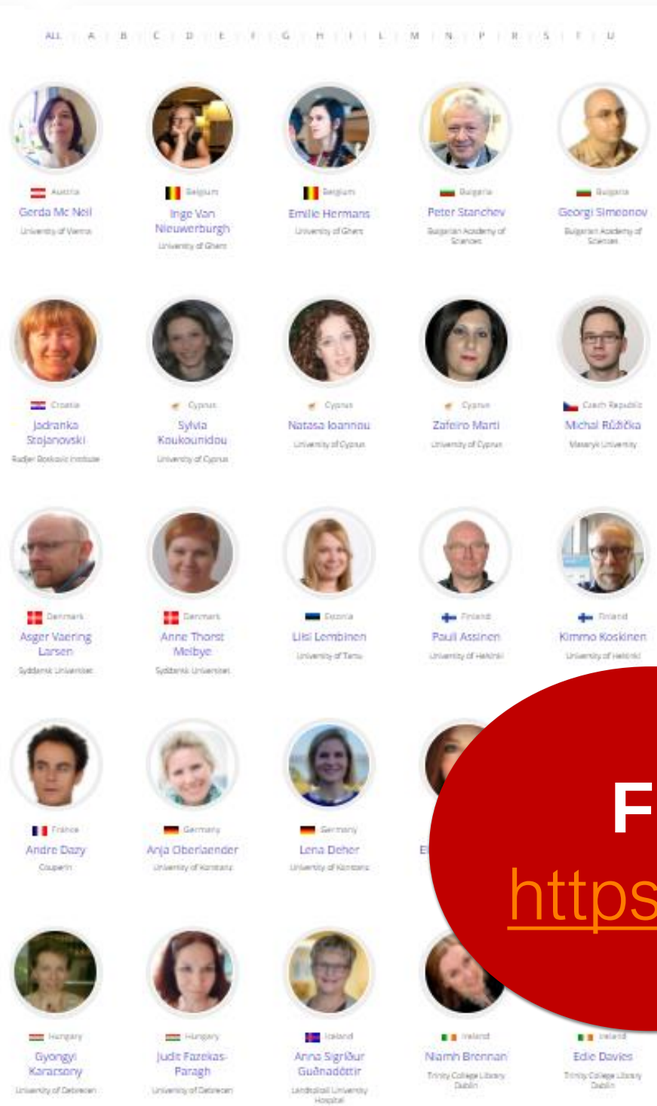
How do I know if my research data is protectec

How do I license my research data

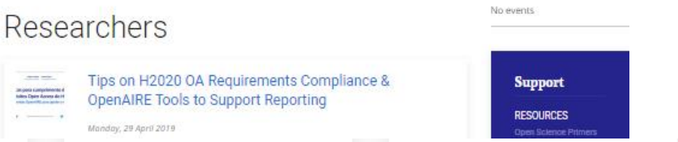
Can I reuse someone else's research data

Data reuse stories & use cases

Series of stories and use cases reporting the process of data reuse, describing experiences of reuse of a variety of research data.



Researchers



ıonymize your data

For researchers

amnesia

Report your publication and data to the EC

For project coordinators

EXPLORE

Claim a publication or data to your funding

For researchers

EXPLORE

Register & Validate your repository

Content Enrichment

For Repository Managers

Track the usage activity of your repository

Find a NOAD in your country!

https://www.openaire.eu/contact-noac

OpenAIRE

ΑΘΗΝΑ
Έρευνα & Καινοτομία
Τεχνολογίες Πληροφορίας

Useful Links

- **OpenAIRE Helpdesk: https://www.openaire.eu/support/helpdesk**
- **OpenAIRE NOADs: https://www.openaire.eu/contact-noads**

- **OpenAIRE Factsheets: https://www.openaire.eu/openaire-h2020-factsheets**
- **OpenAIRE Research Community Dashboard: https://connect.openaire.eu/**
- **OpenAIRE webinars:**
- **https://www.openaire.eu/frontp**

Don’t miss out!
Sign up to OpenAIRE and subscribe to the Newsletter to get informed about news and updates

Thank you!

Elli Papadopoulou

elli.p@athenarc.gr

ORCID 0000-0002-0893-8509

@elli_lib

Benefits and concerns of biomedical data sharing

Reasons to share individual-level data	Concerns about sharing individual-level data
To improve science	**May hamper science**
• Enable verification, replication, and expansion of research results • Address biases, deficiencies, and dishonesty in research • Enable novel analyses and increase study power • Improve meta-analyses • Maximize data use, particularly for datasets that cannot be replicated • Inform research design and research funding • Improve teaching resources • Increase primary data producers' academic profiles and collaboration opportunities	• Reputational harms of critical secondary analyses • Consequences of flawed/poor quality secondary analyses • Reduction of incentives for primary research • Increased incentives to conduct short-term research rather than long-term research • Opportunity costs of curating and sharing data
To improve health	**May hamper health**
• Inform health care planning and allocation • Inform regulatory review • Improve evidence base for clinical decision making • Improve use of health care resources • Improve patient care	• Effects of flawed secondary analyses on scientific evidence base • Burden of evaluating validity of secondary analyses • Effects of second-guessing regulatory procedures, policies, and processes
Explicit moral claims	**Explicit ethical issues**
• Importance of maximizing the value and utility of data • Promotion of scientific values • Promotion of best practices in research conduct, analysis, and reporting • Demonstration of respect for research participants • Promotion of the public good	• Protection of participants' privacy and confidentiality • Validity of consent, including broad consent • Potential harms of secondary research for research participants including discrimination and stigma • Researchers' ability to fulfill commitments made to research participants during data collection • Effects of moral distance and limited awareness of the context in which data were collected • Potential impacts on public trust and confidence of conflicting analyses • Balancing the interests of differing stakeholders in data sharing • Making best use of limited research resources
	Barriers to sharing
	• Costs of developing and maintaining appropriate expertise and infrastructure • Curation costs • Ownership, intellectual property rights, and commercial confidentiality • Lack of policies and processes

Bull S, et al. (2015) Views of Ethical Best Practices in Sharing Individual-Level Data from Medical and Public Health Research. Journal of Empirical Research on Human Research Ethics. http://dx.doi.org/10.1177/1556264615594767