Agenda

Part 1

- What is open science?
- Why practice open science?
- Pillars of open science
- Open science policies
- Open science tools
- Open research data management

Part 2 - EOSC

What is open science

"Open Science is the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods" (FOSTER Open Science Definition).

"Open Science is transparent and accessible knowledge that is shared and developed through collaborative networks" (Vicente-Sáez & Martínez-Fuentes 2018).
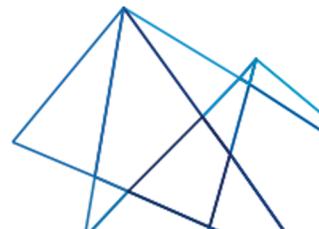
Above and below
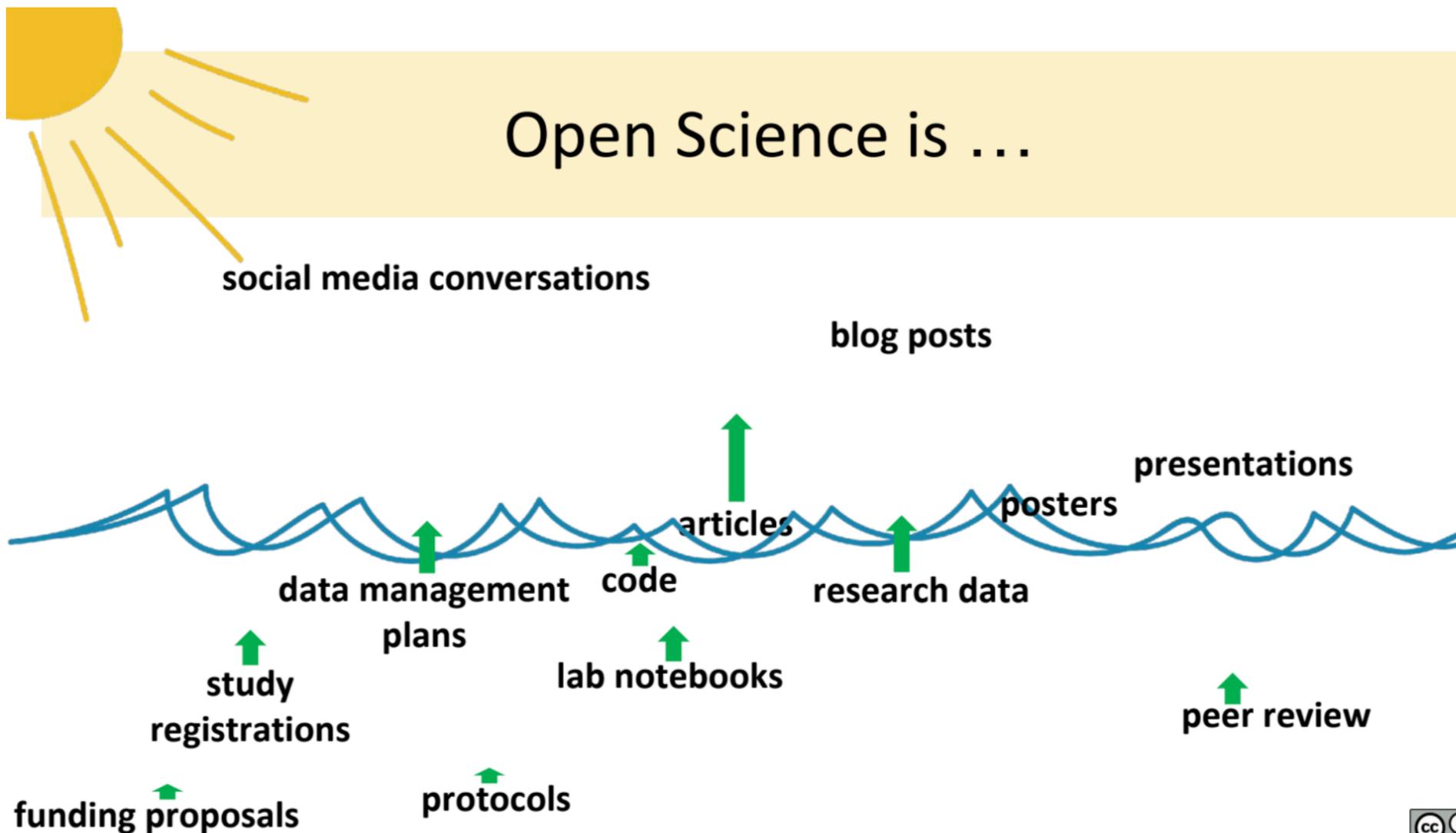
Open Science is …

social media conversations

blog posts

presentations

posters

articles

data management plans

code

research data

lab notebooks

study registrations

peer review

funding proposals

protocols

Open to…
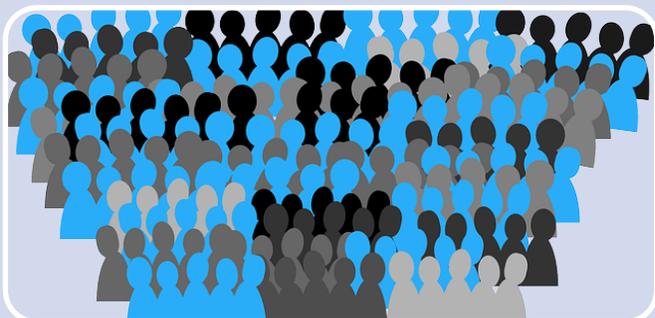
Image by Clker-Free-Vector-Images from Pixabay

Image by Isaac Castillo from Pixabay

Image by Angela Yuriko Smith from Pixabay

Participation

- No discrimination
- Inclusiveness
- Evaluation based on social relevance
- Citizen science
- All knowledge
- Error-friendly

Reuse

- Open access to all (data, code, proposals, …)
- Open standards
- Open licences
- Fully documented processes

To the world

- Multilingual
- Simple explanations
- Reach out of the academia
- Curation
- Public debate

Science was meant to be open

- Since the beginning, science was meant to be openly discussed and criticized

- But, divisions on various basis (national, geo-political, economical, social, ...) influenced the closing of science

- Open science should again be just science, done the right way!


Raphael/Wikimedia Commons/PDArt
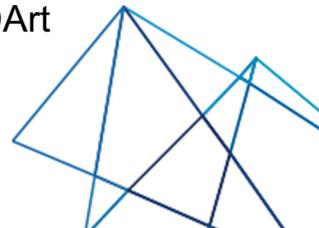
Why

- To extend the availability of research
- To increasing collaboration
- To speed up insights and innovation / efficient science
- To detect errors
- To provide reproducibility of research results
- To get acknowledgement and rewards for more research activities
- To improving the image of science and scholarship
- To producing more accurate science

Research lifecycle …

Preparation
- Define & crowdsource research priorities
- Organize project, team, collaborations
- Get funding / contract

Assessment
- Comment / peer review
- Determine impact of research output
- Determine impact of researchers

Discovery
- Search literature / data / code / …
- Get access
- Get alerts / recommendations
- Read / view
- Annotate

Outreach
- Archive/share posters
- Archive/share presentations
- Tell about research outside academia
- Researcher profiles/networks

Analysis
- Collect, mine, extract data / experiment
- Share protocols / notebooks / workflows
- Analyze

Publication
- Archive / share publications
- Archive / share data & code
- Select journal to submit to
- Publish

Writing
- Write / code
- Visualize
- Cite
- Translate

and how to open it!

Bianca Kramer, Utrecht University Library, Eurodoc Open Science Ambassador training
tinyurl.com/Eurodoc_OS1

Pillars of open science

- Open data

- Open access

- Open methodology

- Open source

- Open evaluation

- Open education

Open data

- Data that can be freely accessed, reused, redistributed for research, teaching and beyond

- No restrictions in the ideal case…

As open as possible, as closed as necessary

- Open data can be inspected, enabling research verification and reproducibility, and fostering wider collaboration

FAIR principle

- Principles defined in 2016

- Community-developed set of guidelines and best practices

- To support the data intensive science

- Ensures that all relevant digital objects are
  Findable, **A**ccessible, **I**nteroperable and **R**e-usable

https://www.force11.org/node/6048
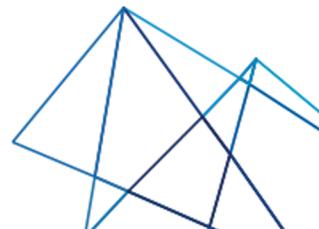
Findable

F1. (meta)data are assigned a globally unique and eternally persistent identifier.

F2. data are described with rich metadata.

F3. (meta)data are registered or indexed in a searchable resource.

F4. metadata specify the data identifier.
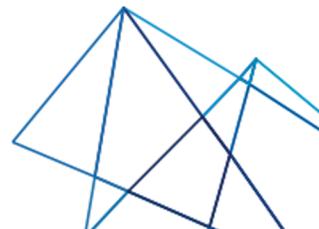
Accessible

A1 (meta)data are retrievable by their identifier using a standardized communications protocol.

A1.1 the protocol is open, free, and universally implementable.

A1.2 the protocol allows for an authentication and authorization procedure, where necessary.

A2 metadata are accessible, even when the data are no longer available.
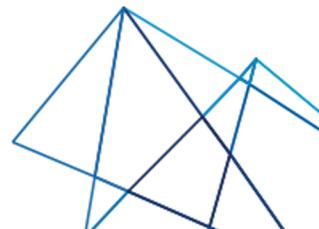
Interoperable

I1. (meta)data use a <u>formal, accessible, shared, and broadly applicable language</u> for knowledge representation.

I2. (meta)data use <u>vocabularies that follow FAIR principles.</u>

I3. (meta)data include <u>qualified references</u> to other (meta)data.
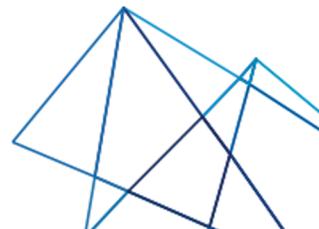
Re-usable

R1. meta(data) have a plurality of accurate and relevant attributes.

R1.1. (meta)data are released with a clear and accessible data usage license.

R1.2. (meta)data are associated with their provenance.

R1.3. (meta)data meet domain-relevant community standards.
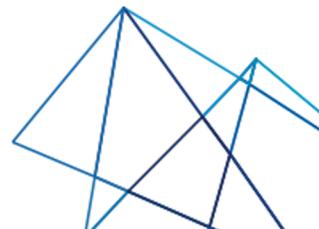
Open access

- Open access to published research
  - (will be covered in another presentation)
- Open access to research data
- Example: COVID-19 research
  - Publishers opened the access to publications (unfortunately, some did it just temporary)
  - Dedicated open data portals were established to share data
  - Results: unprecedented short development cycles for vaccines, medications, protocols, ..
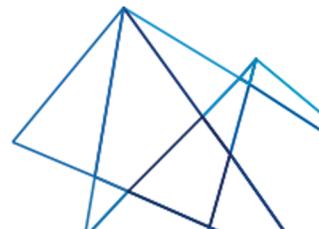
Open methodology

- The transformation of data to conclusions
- "Open methodology" refers to more standardized definitions and reporting of methods across countries and disciplines. (https://www.techlib.cz/en/84233-open-methodology)
- Examples
  - **Open Source,** making programming code publicly available
  - **Open Notebooks**, where daily research work is publicly documented
  - **Open Workflows,** documented and transparent workflows
  - **Open Annotations,** open and collaborative classifications and comments.

Reproducible research

Data dredging
Also known as p-hacking, this involves repeatedly searching a dataset or trying alternative analyses until a 'significant' result is found.

Omitting null results
When scientists or journals decide not to publish studies unless results are statistically significant.

Underpowered study
Statistical power is the ability of an analysis to detect an effect, if the effect exists – an underpowered study is too small to reliably indicate whether or not an effect exists.

Issues

Errors
Technical errors may exist within a study, such as misidentified reagents or computational errors.

Underspecified methods
A study may be very robust, but its methods not shared with other scientists in enough detail, so others cannot precisely replicate it.

Weak experimental design
A study may have one or more methodological flaws that mean it is unlikely to produce reliable or valid results.

Open source

1. **Free Redistribution**
2. **Source Code**
3. **Derived Works**
4. **Integrity of The Author's Source Code**
5. **No Discrimination Against Persons or Groups**
6. **No Discrimination Against Fields of Endeavor**
7. **Distribution of License**
8. **License Must Not Be Specific to a Product**
9. **License Must Not Restrict Other Software**
10. **License Must Be Technology-Neutral**

https://opensource.org/osd
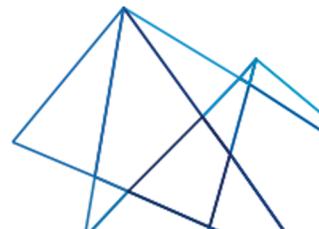
Open evaluation

- **Open Evaluation** means openness in scientific evaluation through peer review and bibliometrics.

- **Open Peer Review** is an umbrella term for a number of overlapping ways that peer review models can be adapted to be in line with the aims of Open Science, including transparency about reviewer and author identities, the publication of review reports and the enablement of greater participation in the peer review process.

- **Open Metrics** means openness of data, methods, and results of bibliometric analyses. The traceability and reusability of evaluation procedures opens up new possibilities in dealing with scientific findings in the fields of research, technology, and innovation.
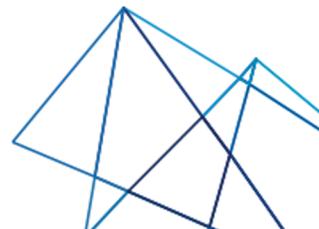
https://www.oana.at/en/about-open-science/

Open education

- Everyone in the world should have access to high-quality educational experiences and resources, and they work to eliminate barriers to this goal.

- Open Education Consortium: "sharing is probably the most basic characteristic of education: education is sharing knowledge, insights and information with others, upon which new knowledge, skills, ideas and understanding can be built." (http://www.oeconsortium.org/about-ocw/)
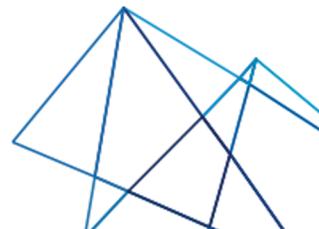
- Examples
  - MOOC

Open science policies

- Strategies and actions aimed at promoting Open Science principles and at acknowledging Open Science practices.

- On multiple levels
  - International
  - State
  - Institutional

- Examples:
  - European University Association reports (https://www.eua.eu/resources/publications/999:open-science-in-university-approaches-to-academic-assessment.html)
 - Still a long way to go!
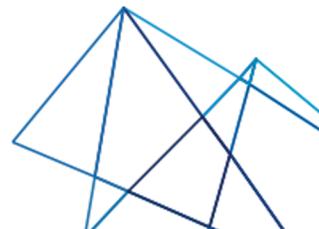
Open science tools

- Collaborative platforms
  - Virtual research environments
    - VI-SEEM - VRE for regional Interdisciplinary communities in Southeast Europe and the Eastern Mediterranean
  - Collaborative writing platforms
    - Overleaf, Authorea, Fidus Writer, ShareLaTeX and Google Docs
- Repositories (more about this later)
  - Zenodo, https://zenodo.org
  - re3data, https://www.re3data.org
- Open workflow tools
  - Apache Airflow, https://airflow.apache.org
  - Apache Taverna, https://incubator.apache.org/projects/taverna.html
 - Camunda, https://camunda.com

ORDM
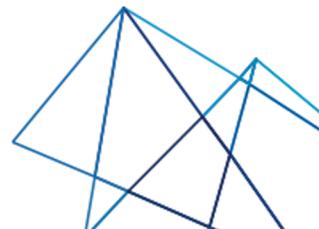
- Research Data management comprises the disciplines related to managing research data as a valuable resource. These can include, but are not limited to, research data governance, security and quality.

Research Data Management Plans

Research Data Management Policies

Research Data Management Services

Research Data Management Standards

Research Data Management Tools

Research Data Management

https://www.fosteropenscience.eu/foster-taxonomy/research-data-management
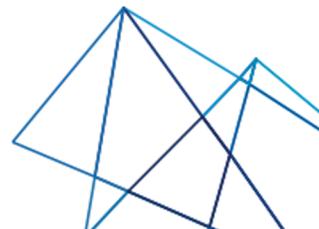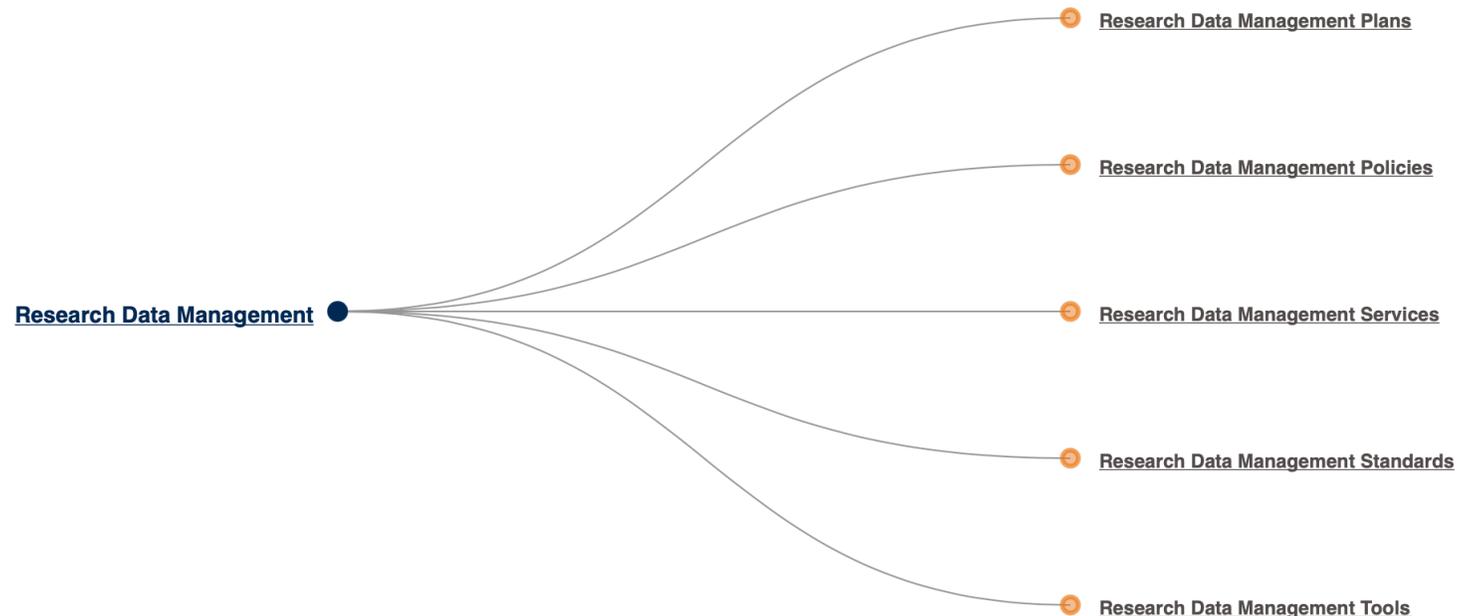
RDM Plan

- A research data management plan (DMP) is a formal document that outlines how to handle research data both during your research and after the research project is completed.

- Common elements of an DMP
  - Data Description
  - Standards and Metadata
  - Data Sharing
  - Archiving and preservation

- Tools to support DMP creation
  - https://dmponline.dcc.ac.uk

Plan to make data work for you

Data Management Plans that meet institutional funder requirements.

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It is provided by the Digital Curation Centre (DCC).

https://www.fosteropenscience.eu/foster-taxonomy/research-data-management

Useful links

- https://training.ni4os.eu/
- https://open-science-training-handbook.github.io/
- https://force11.org/info/the-fair-data-principles/
- https://www.fosteropenscience.eu/

THANK YOU

Any Questions ?