Welcome: In the Open Science Orbit: Introduction to ORRI

*Workshop for experienced researchers
(and other open research enthusiasts)*

04 February 2025

Dr Milena Dobreva, University of Strathclyde, Scotland

A

REINFORCING *Grantee*

SOFIA UNIVERSITY ST. KLIMENT OHRIDSKI
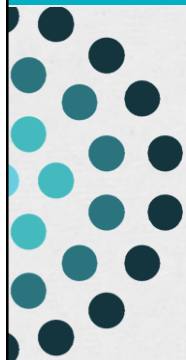
NGO Links

Funded by the European Union

1

REINFORCING

An open science warm-up

2

3

4
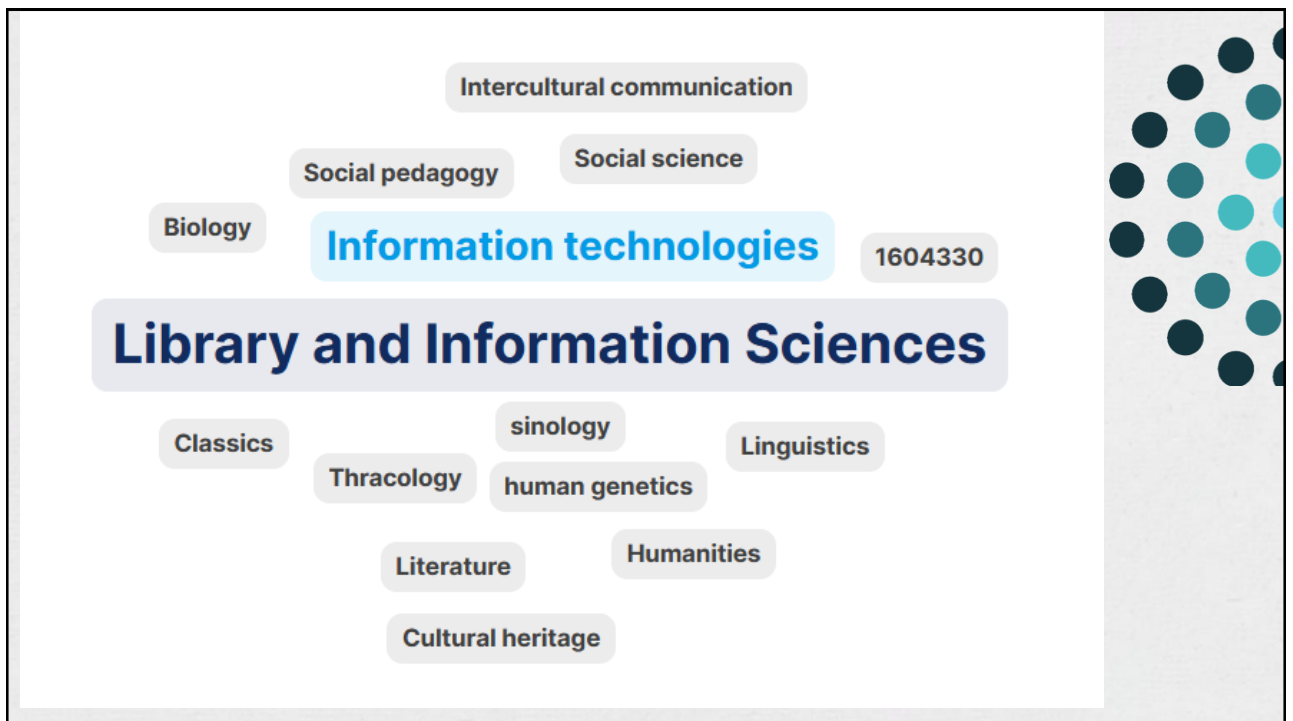
slido

Please download and install the
Slido app on all computers you use

What is the main goal of the Open Science movement?

ⓘ Start presenting to display the poll results on this slide.

5

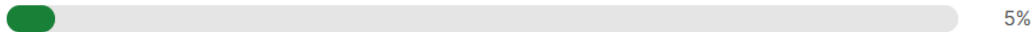**What is the main goal of the Open Science movement?**

⬆ Share ⌄

Multiple Choice Poll   ☑ 21 votes   👥 21 participants

To advise on the best research methodologies as a golden standard for all researchers in a specific domain. - 3 votes

14%

To increase the exclusivity and secrecy surrounding scientific data. - 1 vote

5%

To promote transparency, inclusivity, rigor, and the beneficial impact of scientific research. - 16 votes

76%

To provide access to research findings to a limited group of individuals. - 1 vote

5%

6

slido

Please download and install the Slido app on all computers you use

What is Open Source in the context of Open Science?

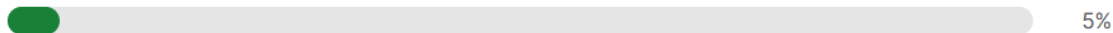ⓘ Start presenting to display the poll results on this slide.

7

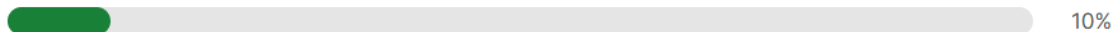**What is Open Source in the context of Open Science?**

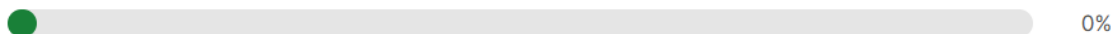Multiple Choice Poll · 21 votes · 21 participants · Share ⌄

A new way of charging for research software. - 1 vote

5%

A way of sharing research software within your department or research group. - 2 votes

10%

A way of keeping research methods secret. - 0 votes

0%

A way of making research methods and software available to everyone. - 18 votes

86%

8

9
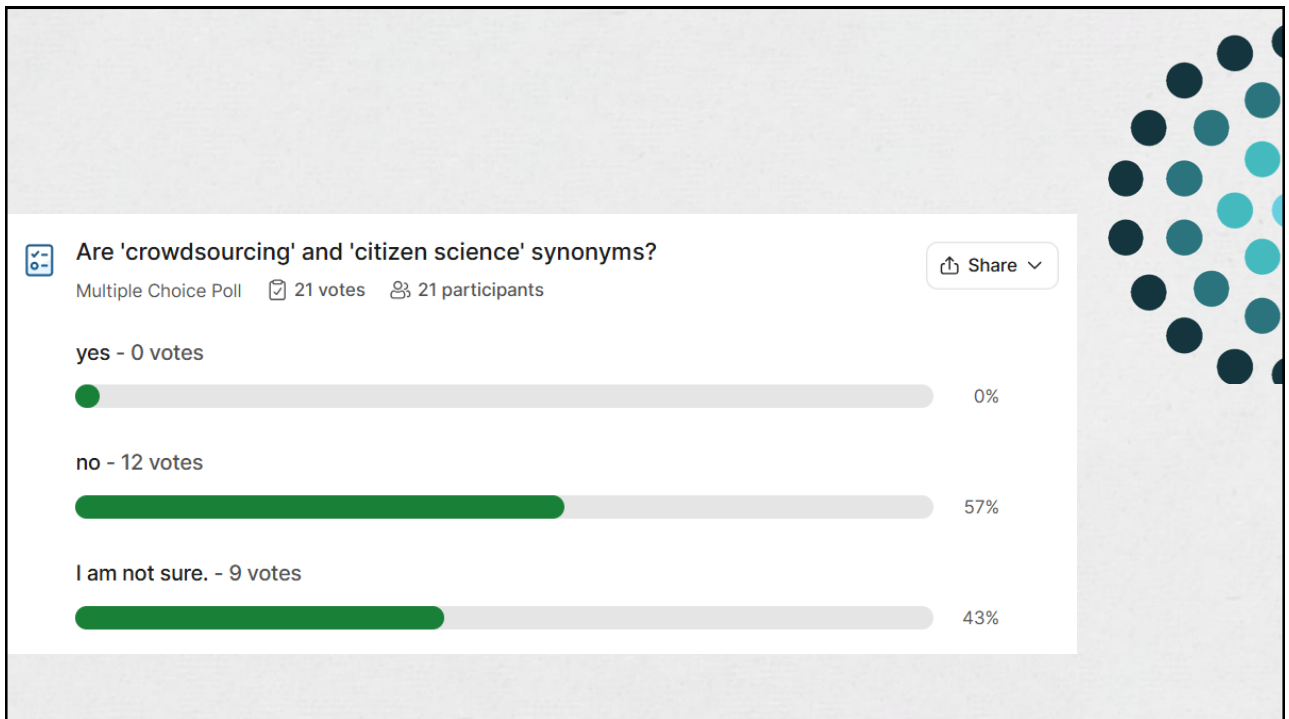
10

slido

Please download and install the
Slido app on all computers you use
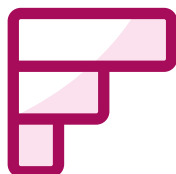
How would you rank the experience of Bulgarian researchers with these ORRI domains?

ⓘ Start presenting to display the poll results on this slide.

11

How would you rank the experience of Bulgarian researchers with these ORRI domains?

⬆ Share ⌄

Ranking Poll   ☑ 19 votes   👥 19 participants

1. science education

2.16
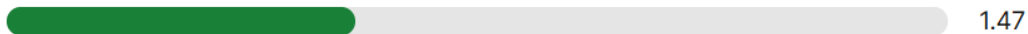
2. open access

2

3. governance

1.47

4. public engagement
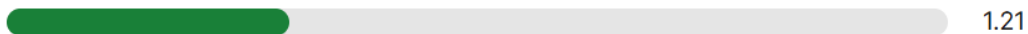
1.21

12

13

14

slido

Please download and install the
Slido app on all computers you use

Could you suggest one skill which can increase considerably the use of ORRI in Bulgaria?

ⓘ Start presenting to display the poll results on this slide.

15

---

17 responses

- Openness
- Sharing the fair data
- Increasing open data literacy
- IPR issues
- Data management
- Sharing
- being academically fair
- digital skills
- Trust
- Knowledge
- Responsible sharing and mentoring
- Only one?
- Experts working with researchers
- Digital skills
- Being data literate
- Openness
- To disseminate findings

16

OPEN DATA

17

ORRI

Research/development/innovation processes that consider the potential impacts of technologies on the environment and society.

18

ORRI

The six pillars of Responsible Research & Innovation, from RRI-tools.eu

19

UNESCO (2021) Pillars of Open Science
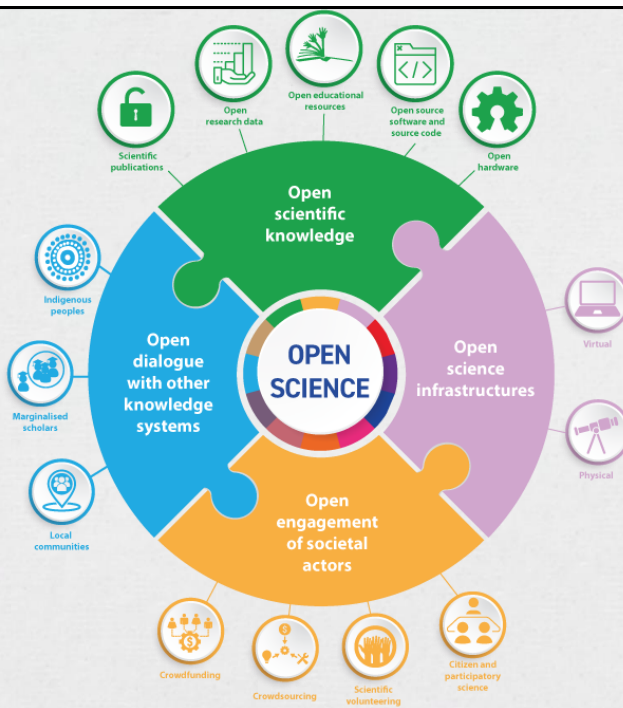
20

What is research data?

"Factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research and that are commonly accepted in the scientific community as necessary to validate research findings. A research data set constitutes a systematic, partial representation of the subject being investigated."

Recommendation of the Council concerning Access to Research Data from Public Funding (OECD, 2006)

21

But there is more (especially in the Humanities but not restricted): collections as data

Image by Mahendra Mahey
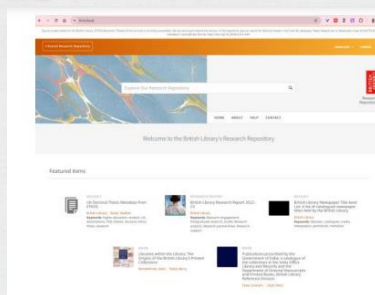
'Collections as Data'

https://bl.iro.bl.uk/
Repository

Download collections as zips
Each dataset has a unique
Identifier (DOI)
can be referenced for research

Over 700 Digitisation / Capturing Born Digital Projects
20 %* Openly Licensed – most online
80 %* Available onsite only

22

Create Process Analyse Store Share Re-use

23

Data and research lifecycle

- Data are created
  in a lab, through fieldwork, measurement, on a computer,
  IoT devices, extraction from primary sources,
  synthetically produced...
- Data are processed
  cleaned up, sampled, anonymised, converted, enriched, tagged ...
- Data are analysed
  statistics, comparison, interpretation, visualisation...
- Data are stored
  for long-term preservation – which version?
- Data are shared
  "as open as possible, as closed as necessary"
- Data are re-used
  - *Citizen contributions?*
  - *Importance for reproducibility*

24

Data and research lifecycle

- DATA PLANNING (BEFORE)

  What is my data like? What policies do I need
  to follow? Who do they belong to? Can I share/what?
  Nature, format, volume, source, collection, access...

- DATA HANDLING/CURATION (DURING RESEARCH)

  How should I store them? Safety, size, security, backups,
  documentation, duplication (green considerations)...

  How am I using my data? Methodology, quality/version control

- DATA SHARING (TOWARDS THE END)

  What happens to my data after my project is over?

  How should I share my data?
  Open and FAIR data, licences, data sustainability and re-use
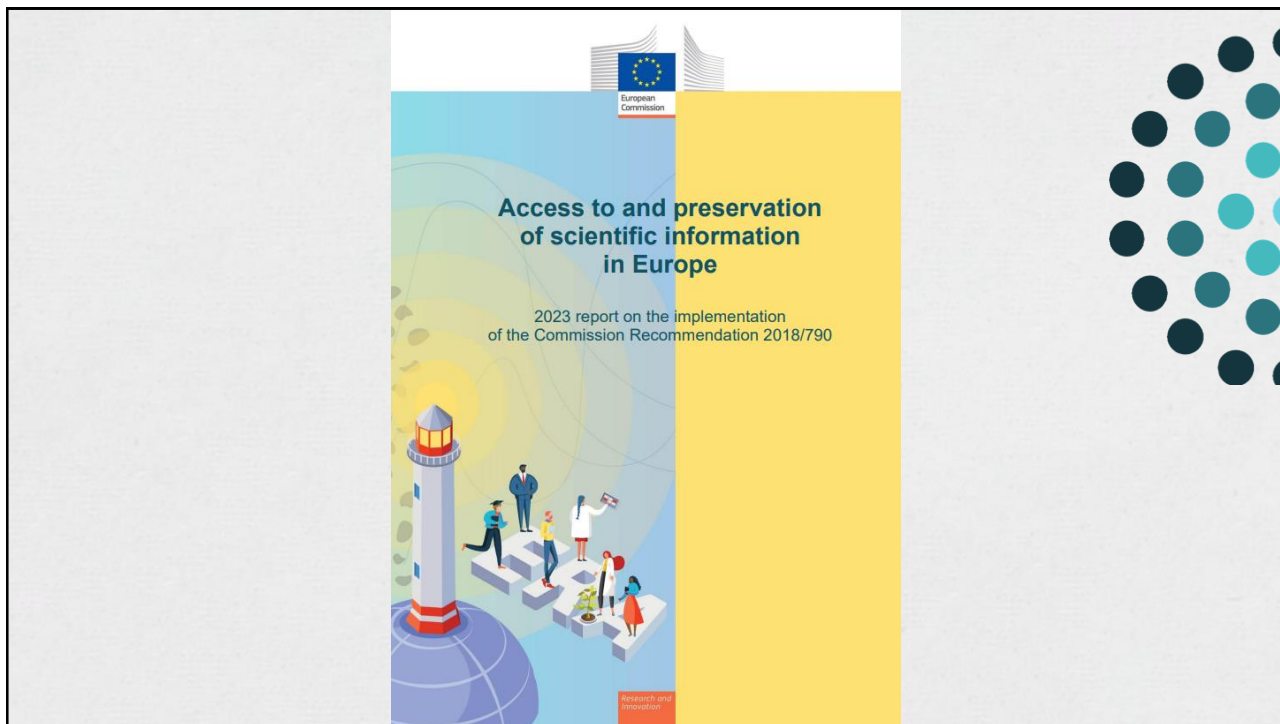
25

Practicalities

- What do I want to share?
- How to provide the context for it?
- Where to share it?
- How to check its use?

Easier if the institution has clear policy
and repository.

26

27

Access to and preservation of scientific information in Europe

2023 report on the implementation of the Commission Recommendation 2018/790

	Share of RPOs with a policy on data management (EOSC 2022)	Share of RPOs with a policy on FAIR data (EOSC 2022)	Share of RPOs with a policy on open data (EOSC 20232
Austria			
Belgium			
Bulgaria			
Croatia			
Cyprus			
Czechia			
Denmark			
Estonia			
Finland			
France			
Germany			
Greece			
Hungary			
Ireland			
Italy			
Latvia			
Lithuania			
Luxembourg			
Malta			
Netherlands			
Poland			
Portugal			
Romania			
Slovakia			
Slovenia			
Spain			
Sweden			
Bosnia and Herzegovina			
North Macedonia			
Norway			
Serbia			
Turkey			

Legend

All
None
Half or more
Less than half
no answer provided

Data: EOSC 2022
Data treatment: Technopolis Group
Note: (1) Missing countries: BE, IT, RO, MK

28

|	Share of RPOs with a policy on data management (EOSC 2022)	Share of RPOs with a policy on FAIR data (EOSC 2022)	Share of RPOs with a policy on open data (EOSC 20232
Austria			
Belgium			
Bulgaria			
Croatia			
Cyprus			
Czechia			
Denmark			
Estonia			
Finland			
France			
Germany			
Greece			
Hungary			
Ireland			
Italy			
Latvia			
Lithuania			
Luxembourg			
Malta			
Netherlands			
Poland			

29

How open/close?

Data that can not/should not be shared

- Legal reasons: copyright, NDA, GDPR
- Ethical reasons – responsible research
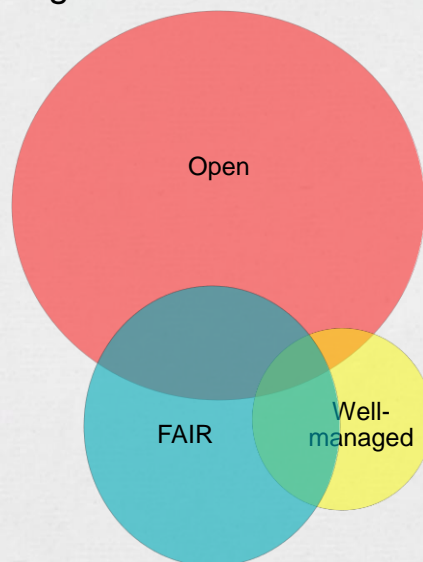- Commercialisation – patent-related data.

Open data

- What is their quality?
- Is re-use possible at all?
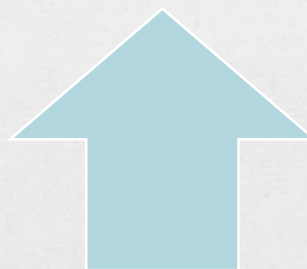
30

Misunderstandings

- Open, FAIR, well-managed

Open

FAIR

Well-managed

31

Repositories – too much or too little choice?

Some considerations on good repository
- Lets you keep all your rights on data
- Provides persistent identifiers (PID)
- Allows to record a right statement
- Would make it clear if your dataset can be used to train AI
- Has metadata standards with controlled vocabularies
  (therefore discipline-specific is usually better)
  - Paradata!
  - Example – Datasheets

- Is popular within the scholars in your discipline
- Has a certification such as CoreTrustSeal
- General / institutional /national / discipline-specific

32

Food for thought…

- Update of the Study on the readiness of research data and literature repositories to facilitate compliance with the Open Science Horizon Europe MGA requirements

ERCEA - The European Research Council Executive
DOI: 10.5281/zenodo.13919643
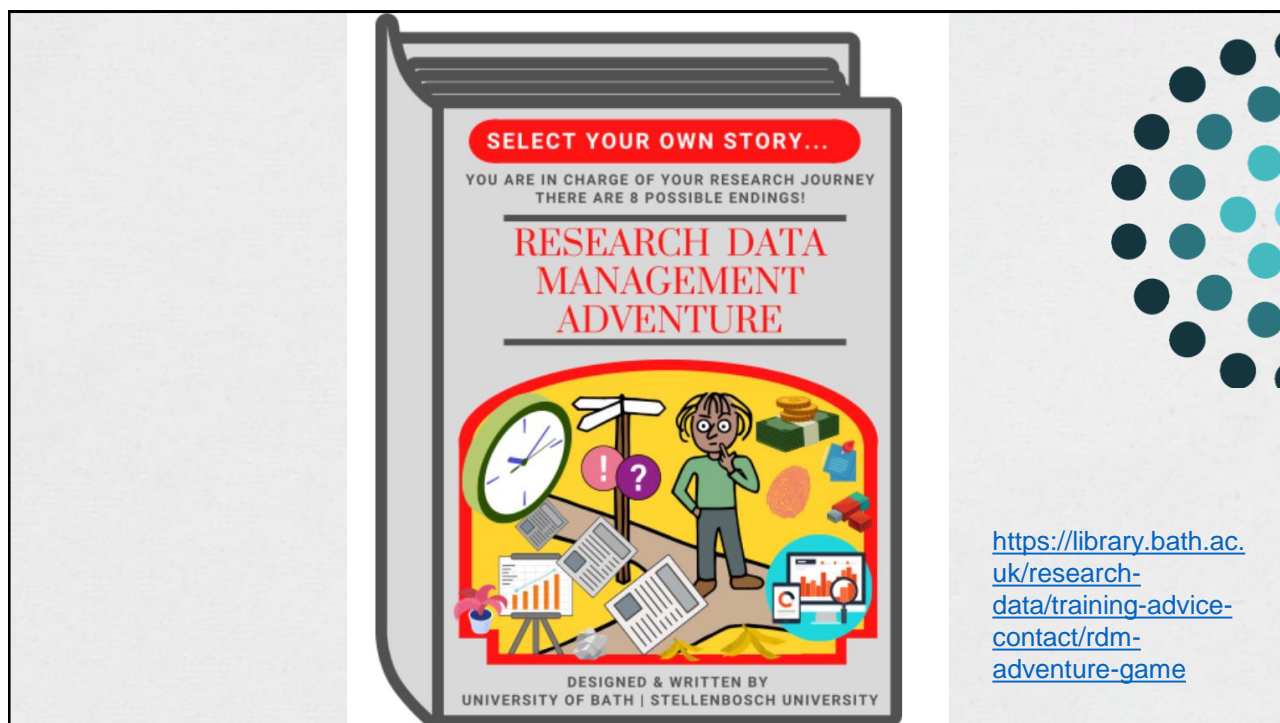2024

Emma Lazzeri

33

Food for thought…

- differences between funder requirements and current repository systems
- additional challenges arising from unclear terminology, such as "international recognition" and "community endorsement."
- Out of the 241 repositories investigated, 186 were found to fulfil the HE MGA definition of a trusted repository, including 92 certified repositories, 116 endorsed by specific research communities, and 99 meeting the essential characteristics identified in the HE AGA.
- Only two repositories, HAL and AUSSDA, demonstrate an exemplary readiness level for respectively literature and data, and data deposition.

34

Some of the inevitable futures?

- Responsible research – and especially security aspects.
- AI-related matters – use of datasets, AI tools supporting research.
- Scientific mis/disinformation.
- Duplication in data infrastructures – green aspects.
- Re-use of data – trusted research.
- Metrics on creating/sharing/reusing datasets in research assessment.
- Attention to specific needs of different research domains.

35

https://library.bath.ac.uk/research-data/training-advice-contact/rdm-adventure-game

36