

Open Science

Hands-On

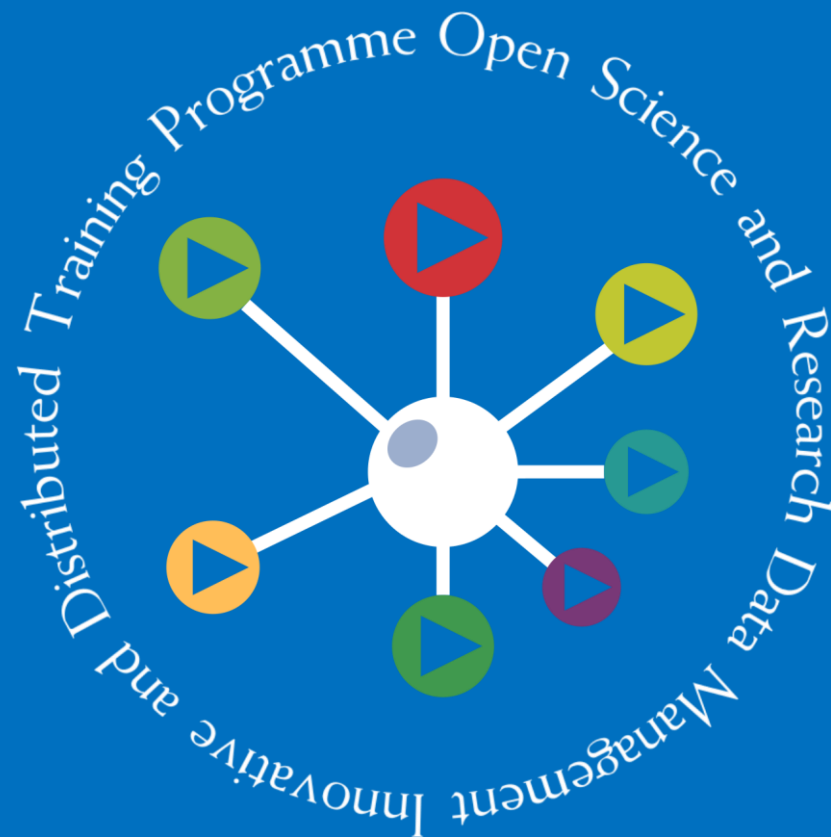
This module is part of the Intensive programme for HE learner „Student spring School” within the project TrainRDM - component of the *Open Science* course

Grant Agreement Number: 2020-1-RO01-KA203-080170

University POLITEHNICA of Bucharest
27-31 March 2023

Based on: <https://open-science-training-handbook.github.io/>

This resource is released under [CC-BY licence](#)



Open Concepts and Principles



Open Science discussion topics



Exercise description

 Form groups of four

 Discuss the topics (e.g., the ones from the next slide).from your own perspectives.

 Summarize most important points that came up for the whole group

Open Science discussion topics (2)



Discussion topics



Topic 1
Making research
more fun

Topic 5
Open peer review

Topic 9
Data sharing

Topic 2
Scooping

Topic 6
Citizen science

Topic 10
ResearchGate
for outreach

Topic 3
APCs

Topic 7
Impact factors

Topic 11
PI demands

Topic 4
Support from
funders & government

Topic 8
Preprints

Topic 12
Scholarly commons

Open Access to Published Research Publications



Finding a Repository



🌐 Search the Registry of Research Data Repositories (<https://www.re3data.org/>) for a suitable repository for your research data.

10 minutes

🌐 How many repositories do you find for your subject?

🌐 Use the filter criteria (e.g., repository type, country, data upload restrictions) to further limit your result.

🌐 Take a look at your favorites on the detail page. What do you find regarding metadata standards, data upload and data access? Which persistent identifier system is used?

🌐 Discussion: **What did you find? What** do you look for in particular when selecting a repository?

10 minutes

How do I find a suitable repository?



Registry of Research Data Repositories <https://www.re3data.org/search>

re3data.org

Search Browse ▾ Suggest Resources ▾ Contact

Filter

Subjects ⊕

Content Types ⊕

Countries ⊕

AID systems ⊕

API ⊕

Certificates ⊕

Data access ⊕

Data access restrictions ⊕

Database access ⊕

Database access restrictions ⊕

Database licenses ⊕

Data licenses ⊕

Data upload ⊕

Data upload restrictions ⊕

Enhanced publication ⊕

Institution responsibility type ⊕

Institution type ⊕

Keywords ⊕

Metadata standards ⊕

PID systems ⊕

Provider types ⊕

Search...

Search

Toogle short help

← Previous

1

2

3

4

5

6

7

...

84

Next →

Sort by ▾

Found 2090 result(s)

UniProtKB/Swiss-Prot

UniProt Knowledgebase

Subject(s) Basic Biological and Medical Research General Genetics Biology Life Sciences

Content type(s) Networkbased data Structured graphics Plain text other

Country Switzerland United Kingdom

UniProtKB/Swiss-Prot is the manually annotated and reviewed section of the UniProt Knowledgebase (UniProtKB). It is a high quality annotated and non-redundant protein sequence database, which brings together experimental results, computed features and scientific conclusions. Since 2002, it is maintained by the UniProt consortium and is accessible via the UniProt website.

Khazar University Institutional Repository

KUIR

Subject(s) Humanities and Social Sciences Life Sciences Natural Sciences Engineering Sciences

Choose the right version for the repository (1)



🌐 Discuss with your neighbor:

🌐 What criteria would you use to select a repository? What is particularly important to you?

🌐 Consider criteria such as

🌐 costs,

🌐 storage location,

🌐 reliability,

🌐 visibility,

🌐 citability, storage duration,

🌐 support offered, etc.

Set up OSF project & link to other platforms



Exercise description

-  Create an OSF collaborative environment from data to publication.
-  Connect your OSF project to GitHub.
-  Upload any raw code, images, data, tables to project.
-  Obtain a DOI and ARK identifier for your project.

Set up OSF project & link to other platforms (2)



🌐 Lou is a first year graduate student working on a project in a biomedical research laboratory. He's trying to decipher data left by a former post doc as a start for his thesis project.

🌐 **Download** Lou's files:

🌐 <http://tinyurl.com/hvna4mg>

🌐 Create an OSF account

🌐 Create a project called "Lou's project"

Open Research Data and Materials

Bucharest 27-31 September 2023



What is research data for me?



Exercise description

-  Think about the last articles they wrote/read.
-  Was there supplementary material (e.g., tables, images)? Write down examples and types of research data in their field of work. What information or data would you need in order to reanalyse the study?
-  What would be needed for your own dissertation/article to be understood properly?
-  Present your results either

Why not share data?



- **Exercise description**

- In pairs or small groups,
 - You have five minutes to make a list as long as possible of all the reasons why researchers might not wish to share their data.
- Report back on your reasons:
 - Are valid reasons or not?
 - Provide strategies for how to overcome legitimate concerns.

Establishing a Reproducible Data Analysis Workflow



- ✿ Format, time needed
 - ✿ Individually and as a group, 4–8 hours (example [here](#))
- ✿ Topic
 - ✿ [Reproducible Research and Data Analysis](#)
- ✿ Learning objectives
 - ✿ Use a (small) computational task relevant to your discipline/background, and establish it as an open and reproducible workflow.
 - ✿ Understand the key concepts, tools and services that are useful in the context of reproducibility.
- ✿ Exercise description
 - ✿ Each participant selects a dataset and corresponding data analysis process that is relevant to their field.
 - ✿ Both dataset and the analysis process should be short enough that it concludes within a few minutes.
 - ✿ Moreover, for the purposes of this exercise, the programming language should be Python or R, but other languages can be accommodated with slight changes in the underlying tools.
 - ✿ The participant initially runs the process in the traditional form, and then asks one of the other participants to re-run it with no external help. Identify both the time required for another person to run this, as well as the obstacles encountered.
 - ✿ Apply the same process using the Jupyter / Git / MyBinder approach;
 - ✿ write the process as a Jupyter notebook, upload dataset and notebook to a repository on GitHub, and then connect the repository to mybinder.
 - ✿ After than, ask again the same person to re-run this. Identify the change in time and accessibility.

THANK YOU!



Follow us

