National Czech Programme | eosc

DMP – Why me???

Georgia Koutentaki, MEng.
Data Steward for STEM fields, National Repository Platform, FIT CTU
Data Steward, Data Stewardship support team, UCT Prague
International Coordinator of the Czech Data Stewards Community
Coordinator of the STEM data stewards subcommunity

georgia.koutentaki@cvut.cz
https://orcid.org/0000-0002-8099-9963

Co-funded by the European Union

Ministry of Education, Youth and Sports of the Czech Republic

National Czech Programme | eosc

You are in the right place to change your mind

- You don’t understand what DMPs are?
- You don’t understand why DMPs are “suddenly” important?
- You don’t see how they help anyone?
- You don’t know what to add there and what to leave out?
- You want to provide a DMP but don’t have all the info?
- Do you consider DMPs a waste of you time?

Open science and data management vocabulary

- OS: open science
- DS: data steward
- DMP: data management plan
- Metadata: “data about the data”, aka documentation for the data
- FAIR: Findable, Accessible, Interoperable, Reusable
- CC or Creative Commons: open licenses used for data and publications. CC is the abbreviation. *There are several CC licenses and several versions of them.*
- PID: persistent identifier is a long-lasting, unique digital reference to a person, place, or object.

https://openscience.cz/cs/heslar/

National Czech Programme | eosc

Findability:
Ability to find the data easily, by humans or machines

Accessibility:
How can I access the data once I find them? Are the processes described in the metadata? Is there need for authorization/ authentication?

Interoperability:
Can I use/join/ move and re-store the data in different applications without them losing meaning?

Reusability:
How can I reuse the data (terms of use)? Are there metadata available? Is the provenance clear?

Findability **A**ccessibility **I**nteroperability **R**euse

National Czech Programme | eosc

FAIR DATA

Are FAIR data always Open data?

No. Data can be FAIR, even when it is **restricted**. For example, in cases of **sensitive data**, such as NDA-related or IP data, or personal data.

Data for long term preservation? Why?

- Correct data interpretation
- Data formats that can be opened years later
- Non-proprietary data formats = no need for specific software/hardware
- Clean data = no misunderstandings
- Organized data = useful data
- Clear instructions (aka. metadata/readme files) = accessibility & reusability
- Licensing = clear reuse terms

What is a DMP?

A Data Management Plan is a document that includes information about the organization and handling of data (usually during a project)

It contains information about:

- Contributors
- Datasets
- Quality processes for data handling
- Related software
- License of data
- Reused data and more…

National Czech Programme | eosc

Thinking of DMPs

Misconceptions

- × A time-consuming obligation
- × A document where you write few sentences to get the *ok* from the funder
- × A random questionnaire that you do not wish to understand
- × A document that anyone can fill in, even if they do not actively work with the data

Opportunities

- ✓ A way to avoid losing data, time & resources
- ✓ A document that organizes your project, so you can publish and create more projects later- increasing efficiency
- ✓ A method for data preservation and organization for scientists
- ✓ A truthful document to improve scientific documentation

National Czech Programme | eosc

What do we expect to see in a DMP?

Before submission

- What you expect the data to be
- Data sizes
- Software/Hardware
- Ethics
- Storage options
- Policies to follow/ data ownership
- Data management people
- Plan for data retention after the project
- Plan for naming conventions
- How you plan to make your data FAIR

National Czech Programme | eosc

What do we expect to see in a DMP?

Mid-project

- First publications of data
- More specific list of data to be produced
- Chosen repositories
- DOIs or other PIDs
- Licenses for data
- Changes in storage expectations
- Other changes
- How you make your data FAIR

National Czech Programme | eosc

What do we expect to see in a DMP?

Final DMP

- All data published or reasons why data are not published
- Specific storage/repositories where the data are archived
- Longevity of the archive
- Licenses + PIDs
- Changes on the data management
- Final naming conventions
- Data formats for long term archiving
- FAIRness of data

Data management plan - example

National Czech Programme | eosc

Responsibilities and rights
Who is responsible for follow-up and revision of this data management plan?
How will responsibilities be distributed among possible external collaborators? Will there be a separate agreement on this? (A *separate agreement may be necessary* e.g. for commercial collaboration.)
Who has the right to manage the data?
Who can access the data during the project period?
Who has ownership of the data? (Normally the institution, unless ownership has been agreed on differently e.g. with external collaborators.)

Collecting/generating data
What kind of data will be collected/generated (e.g. observations, simulations, interviews)? What are the sources (e.g. corpora or other raw data)?
What standards and methods will be used for data collection/generating?
When will the data be collected/generated?
Is there need for extra hardware or software?
Is there need for special expertise for collecting/generating data?

Documentation and metadata
How will the data be documented so that they are comprehensible and reusable for yourself and others also in the long term? (According to best practice, research data should be documented in a **ReadMe file** which explains column headings, abbreviations etc.)
What kind of metadata standard(s) will be used?
What file format(s) will be used?
What kind of folder structure and filename conventions will be used?
Is special software for reading/interpreting the data necessary?

README file for data

Authors: [names +ORCID]

Contributors: people & affiliations with ORCID preferably

Contact person: name & email

Abstract of the data collection

License/ Terms of use: The data are provided under “X license”.

Data file name: e.g. 2026_Microscopy_im_v2.tiff

Data origin: [..]

Data linked works: The data file is part of collection “ZZZ” of datasets and can be found in [doi].
It is also mentioned in publication “AAA” under [doi].

README file for data

Abbreviations and names used in the data file: “2026_Microscopy_im_v2.tiff”

Abbreviation	Explanation:
sam_i	Sample i, where i is the number of sample, i=1,2,3..
im	image
nu	Numerical data
vj	Version j, where j=1,2,3..

Naming convention of data:

YYYY_method_data-type_version	YYYY=year of collection Method= microscopy, rheometry, spectrometry data-type= image, numerical

EXAMPLE:

2026_Microscopy_im_v2	Year of collection 2026_microscopy method_image_version 2

README file for code/scripts

Licenses and re-use terms:

This list, including its selection and arrangement, is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0), as specified in the LICENSE tab.

Notice on scripts:

The scripts in this repository are licensed under the MIT License (see LICENSE--CODE). While the code was generated with the assistance of AI (ChatGPT, Google Gemini, and Perplexity), the underlying logic, structural design, and iterative refinement were directed and authored by Georgia Koutentaki.

[...]

National Czech Programme | eosc

README file for code/scripts

[...]

Suggested Citations

If you use or reference this work, please use the following citations:

For the Resource List (Content)

Koutentaki, G., Thürlová, H., Vališ, J., & Schätz, M. (2026). *Data Steward Resources CZ: A Curated List*. GitHub Repository. https://github.com/UCT-datastewardship/Data-Steward-Resources-CZ. Licensed under CC BY 4.0.

For the Project Scripts (Code)

Koutentaki, G. (2026). *Data Steward Resources: Project Scripts & Tools* [Software]. GitHub Repository. https://github.com/UCT-datastewardship/Data-Steward-Resources-CZ. Licensed under MIT License.

National Czech Programme | eosc

Storage and preservation during the project
What are the procedures for storage and backup, and where will this be done?
Who is responsible for backup and restoring the data?
What is the expected file size for the data?
Do you have sufficient storage facilities, or do you need extra services?

Archiving and sharing
Which data will be preserved, and which will be destroyed at the end of the project?
Will (a selection of) the data be long-term preserved, and how is this decided?
Will the data be made openly available? If only a selection of the data will be openly available, specify which data. (Research data should be made openly available, unless considerations regarding security, personal privacy, commercial or legal issues demand limitation of access.)
If data will not be shared, what is the reason?
Do the data need processing (e.g. conversion to persistent file format(s), depersonalization) before they can be shared? If yes, how will this be done?

Archiving and sharing
Where will data, metadata, documentation and code associated with the data be archived?
What kind of methods or software are needed to get access to the data? Are the methods/software openly available?
When will the data be made available, and how long will they be stored?
How will the data be licensed for reuse?
Are there other conditions, restrictions or embargo on the use of the data?

National Czech Programme | eosc

Data management plan - examples

Ethics and consent
Does your data include sensitive data?
Are you going to collect informed consent to store and share the data? If so, how?
How are you going to secure confidentiality and identity protection?

Using a designated application for your DMPs

National Czech Programme | eosc

Data Stewardship Wizard NRP - through e-INFRA CZ

(dmp.eosc.cz)

- Knowledge Model (KM) and other functions developed specifically for the needs of the NRP & the Czech RDM infrastructure
- The new KM fulfills the needs of common funder requirements (HE, MŠMT, GACR, …)
- Still in pilot mode, but available to all*

*within e-INFRA infrastructure

Data Management Planning within National Data Infrastructure (NDI)

DSW

Data Stewardship Wizard (DSW) is an open-source tool for effective data management planning. Within the NDI, researchers can use a localized instance of DSW for their projects.

Currently, a **pilot operation is underway**.

Start Planning

DMP.eosc.cz

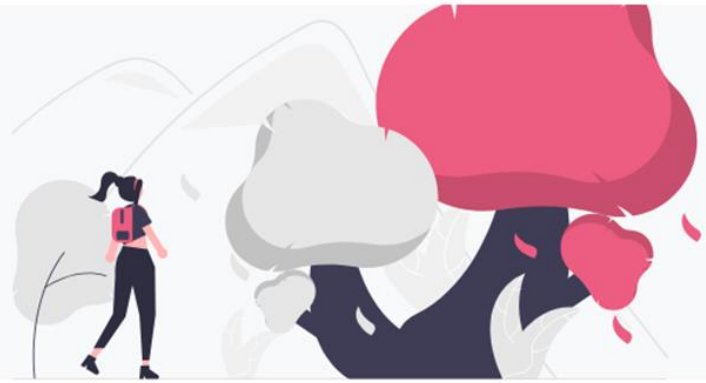
- Dashboard
- Projects
- eosc.cz
- DSW Guide

Researcher

Collapse sidebar

Welcome, !

As a researcher, you create and collaborate on data management plans.



You have no projects yet, start by creating some.

Create Project

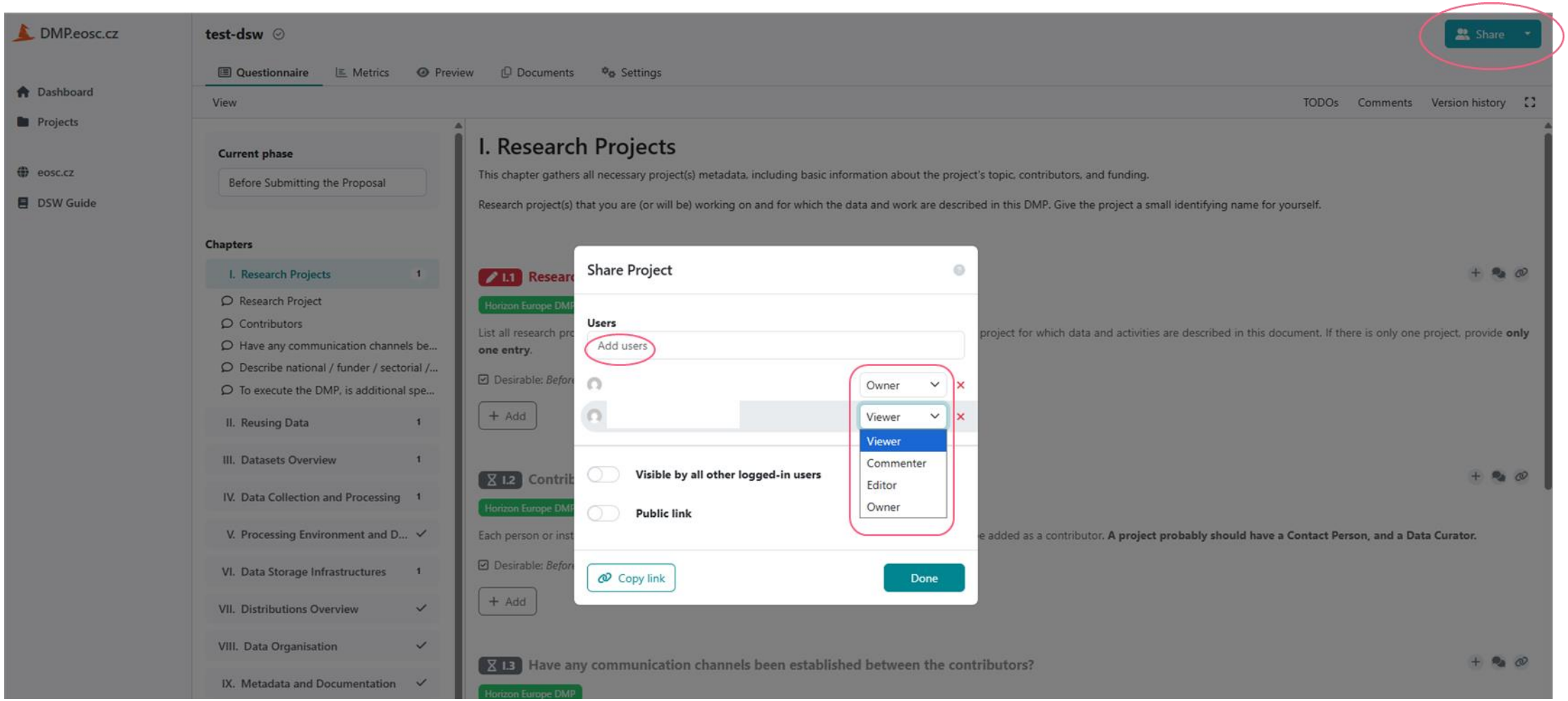
Project is a workspace where you create your DMP. It is based on a knowledge model, which contains knowledge about what should be asked and how based on the research field or organization's needs. You can use document templates to transform the answers into a document. This document can be anything, from PDF to machine-actionable JSON.

You can create a new project from a project template that data stewards prepare for you to have an easier start or from scratch where you set up everything yourself.

Create

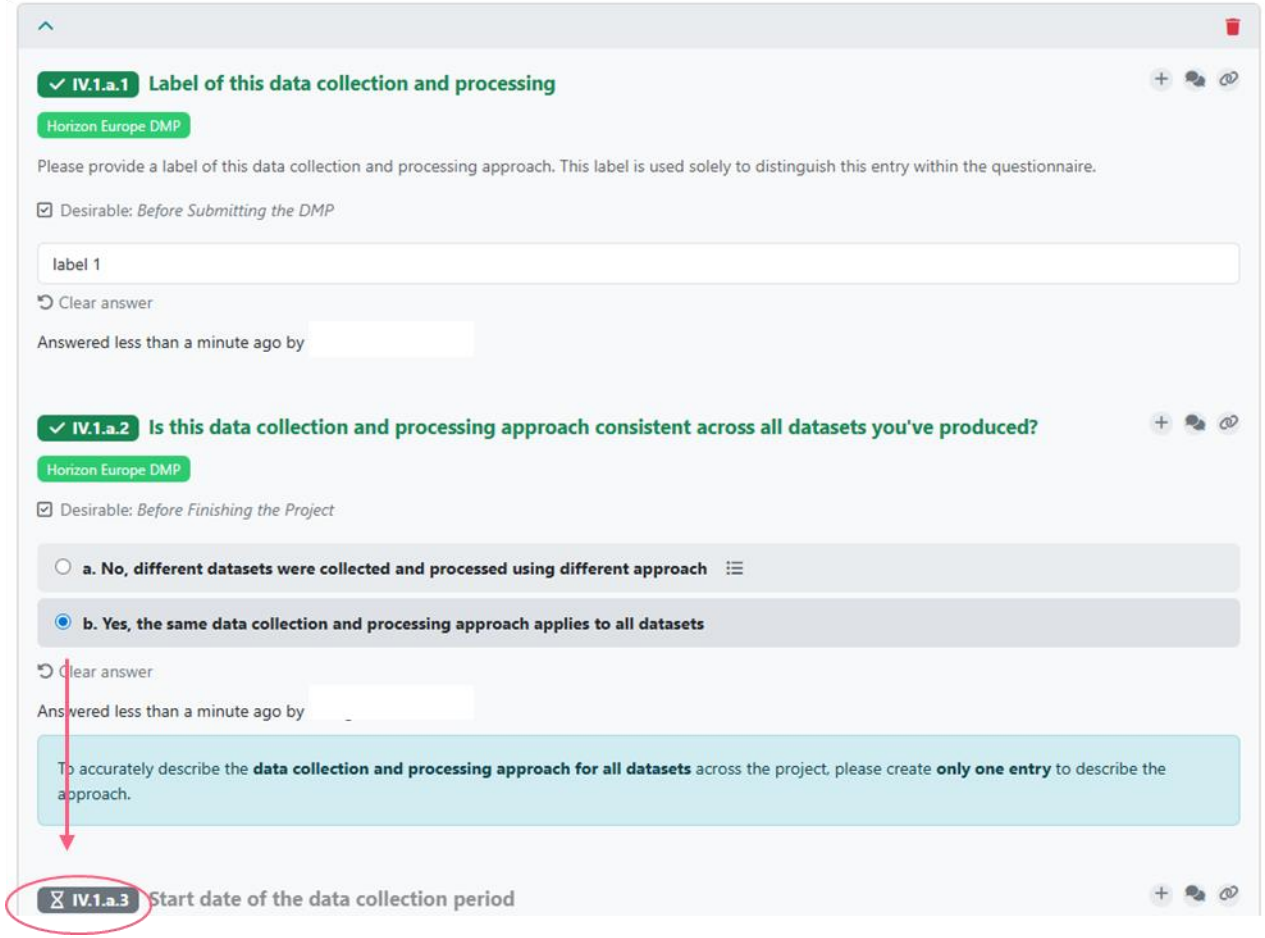
National Czech Programme | eosc

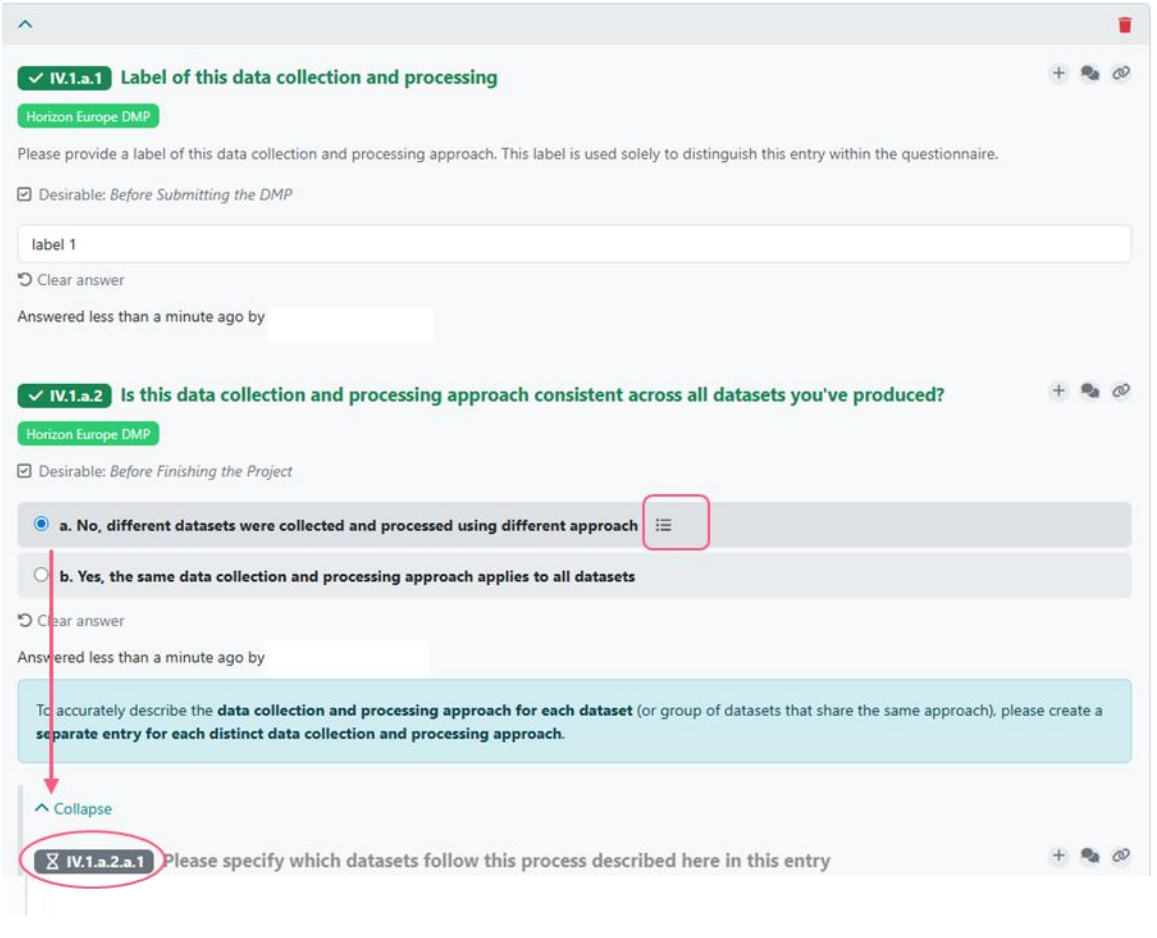
Collaborate and communicate with co-workers



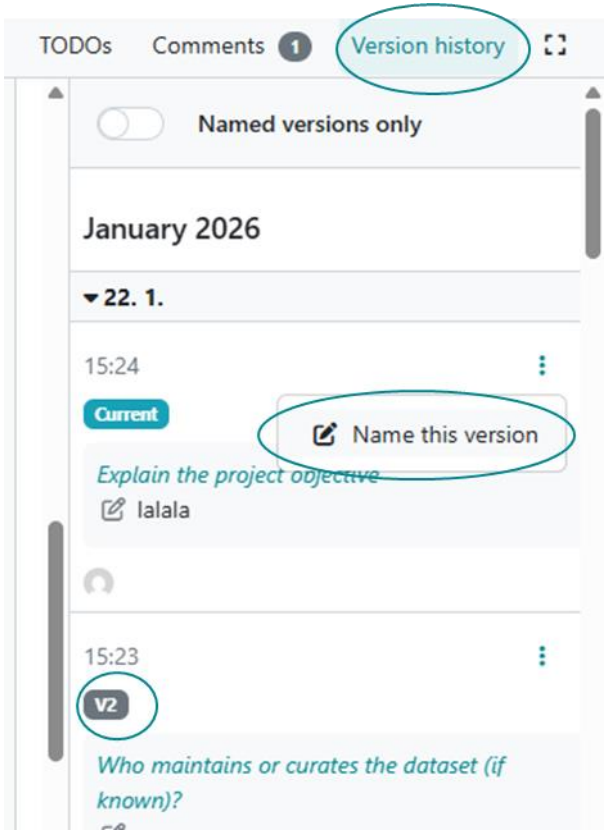
National Czech Programme | eosc

Nested questions





Versioning DMPs



In DSW

History of changes

Version	Publication date	Changes
V2	22 Jan 2026	Addition of existing datasets info
V1_3	22 Jan 2026	Added metadata description
V1_2	22 Jan 2026	Addition of data formats
V1_1	22 Jan 2026	Initial version DMP

In the document

National Czech Programme | eosc

Let's try!

- Dataset https://dmp.eosc.cz/wizard/projects/f18c018f-e5e1-459b-9f6d-289bbf8c8d65?questionPath=b55c4488-6c0e-4c36-88e5-44aa401157db.a7a67bd9-7912-4321-880e-7cf0355ce301

- Data preservation https://dmp.eosc.cz/wizard/projects/f18c018f-e5e1-459b-9f6d-289bbf8c8d65?questionPath=c86ec70c-e5ef-4b0c-b969-772c658ffdb9.ddd721da-cad4-48fa-9794-3bd876b0ef47

National Czech Programme | eosc

Citations, sources & references

- https://creativecommons.org/cc-licenses/
- Data Stewardship Wizard (DSW) NRP. https://dmp.eosc.cz, software & Czech NRP KM under Apache 2.0
- Koutentaki, G. (2026). Data management plan in practice [Presentation]. Doctoral Days 2026, CTU Prague.
- Koutentaki, G. Prepare your team for effective data management II [CC BY 4.0]. https://www.youtube.com/watch?v=D9baGD94v6I
- UiT The Arctic University of Norway. DocEnhance Data Stewardship Course (Module 2, Session 2). [Adapted under CC BY 4.0]. https://courses.docenhance.eu/
- Wilkinson, M. D., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. https://www.nature.com/articles/sdata201618
- https://openscience.cz/cs/heslar/

National Czech Programme | eosc

Thank you

georgia.koutentaki@cvut.cz

Co-funded by the European Union

Ministry of Education, Youth and Sports of the Czech Republic