National Czech Programme | eosc

Data Stewardship Wizard: a STEM use-case

Georgia Koutentaki
FIT CTU in Prague
& UCT Prague

&

Jan Vališ
UCT Prague
& NTK & IMC CAS

Co-funded by the European Union

Ministry of Education, Youth and Sports of the Czech Republic

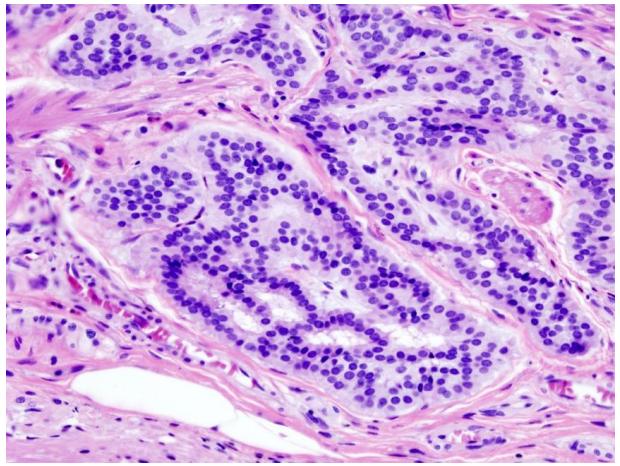
Unless otherwise noted, this presentation is licensed under CC BY 4.0. Cropped screenshots of the Zenodo interface and Data Stewardship Wizard (Apache 2.0) are included for educational illustration; third-party trademarks and logos remain the property of their respective owners.

The use-case: Raman analysis of rat thymus

Jan Vališ

Histopathology

- Slow (hours – days)
- Sensitive & specific

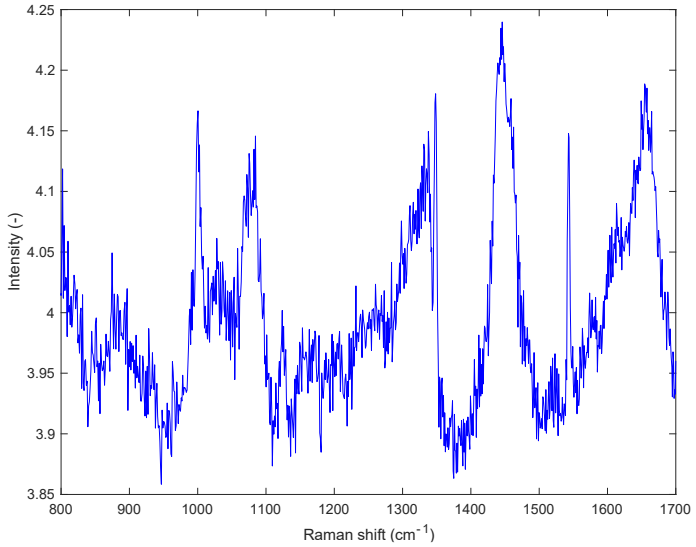


KGH, Wikimedia.org

CC BY SA

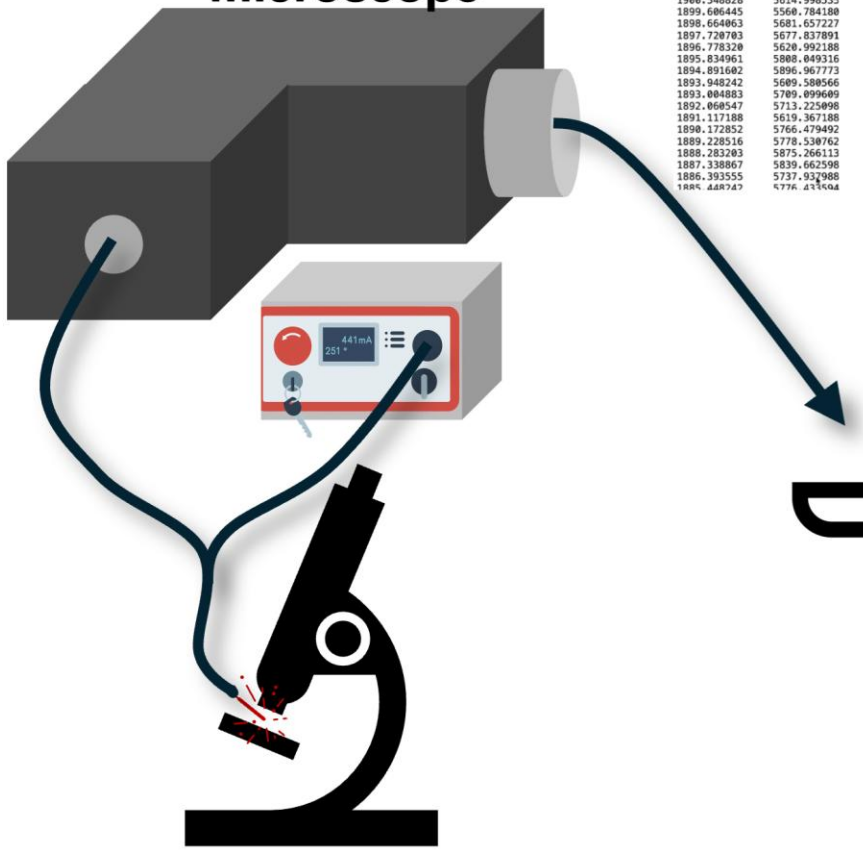
Raman spectroscopy

- Fast (seconds)
- Objective



The experiment at a glance

Raman spectrometer with microscope

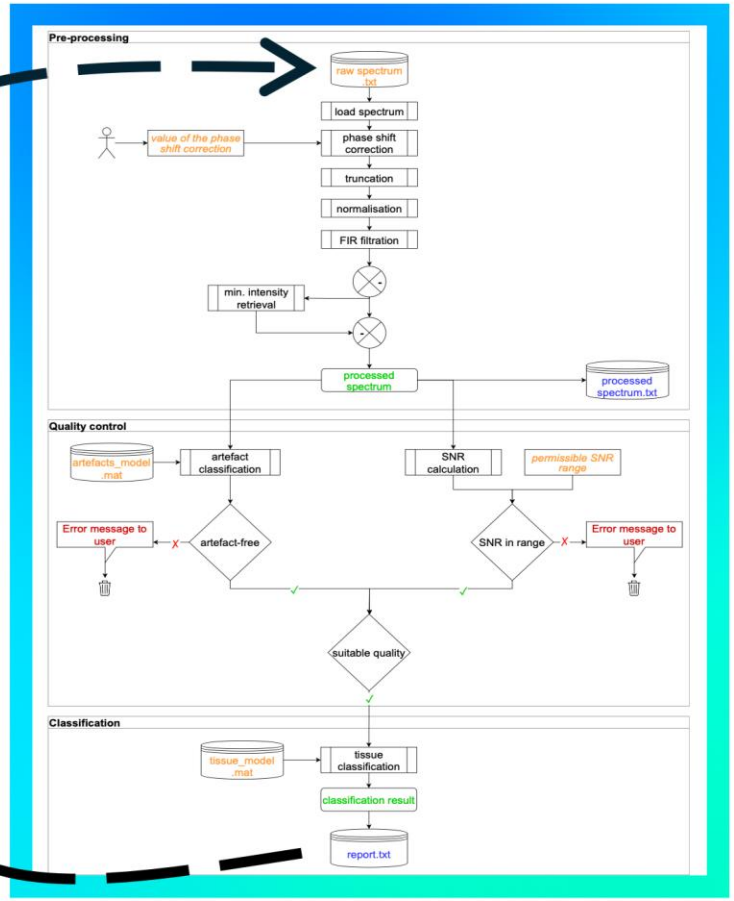


Computer





Automated classification pipeline



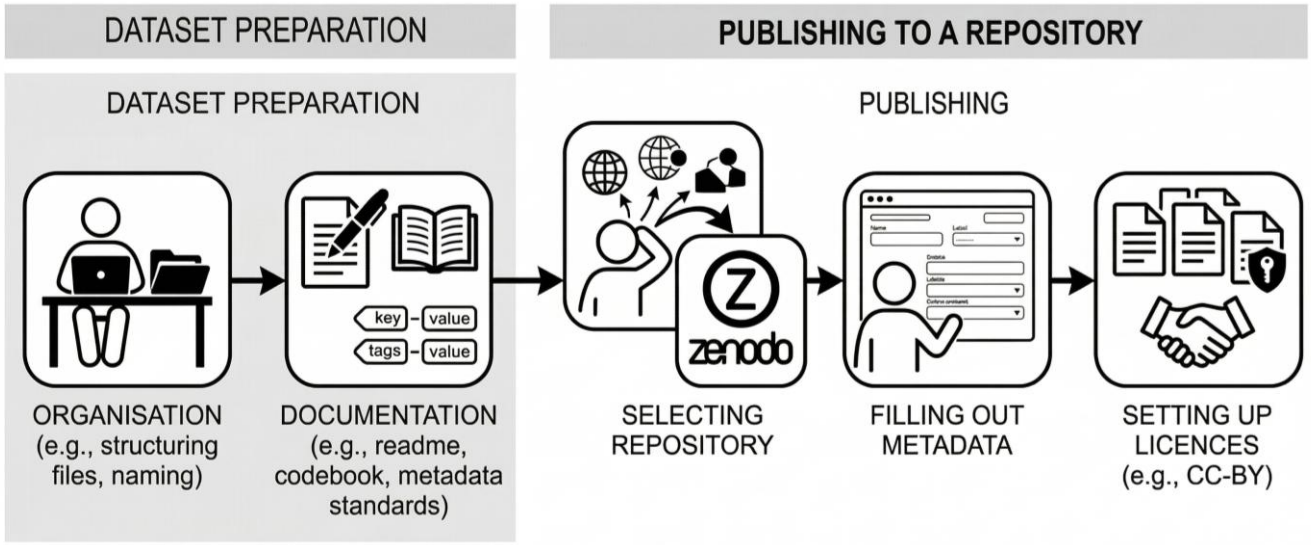
ReadMe

- Many forms – the goal is the same: documentation
- One example:

  Czech Data Steward Community README_FILES_EXAMPLES (10.5281/zenodo.21835978)
- Read more & see the webinar on READMEs here

Want to learn more?

- **Learn how to prepare and publish your dataset**
- Nov 24, 2026 (10:00 – 11:30)
- Register
- Topics
  - Dataset preparation
    - Organization
    - Documentation
    - Preferred formats
  - Repository selection
  - Licensing
  - Dataset deposition on Zenodo



Geminy (3 Flash, 2026-04-07), prompt: Create a flowchart depicting: dataset preparation (organization, documentation), and publishing to a repository (selecting a repository, filling out metadata, setting up licenses). Use Zenodo repository. The flowchart should be horizontal (left to right; 16:9) with minimal or no text - mainly pictograms. Use only black-and-white; simple lines. Style is professional/academic/technical.

The DMP

Georgia Koutentaki

Don’t know where to start?

We can show you!

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



Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

What will we see?



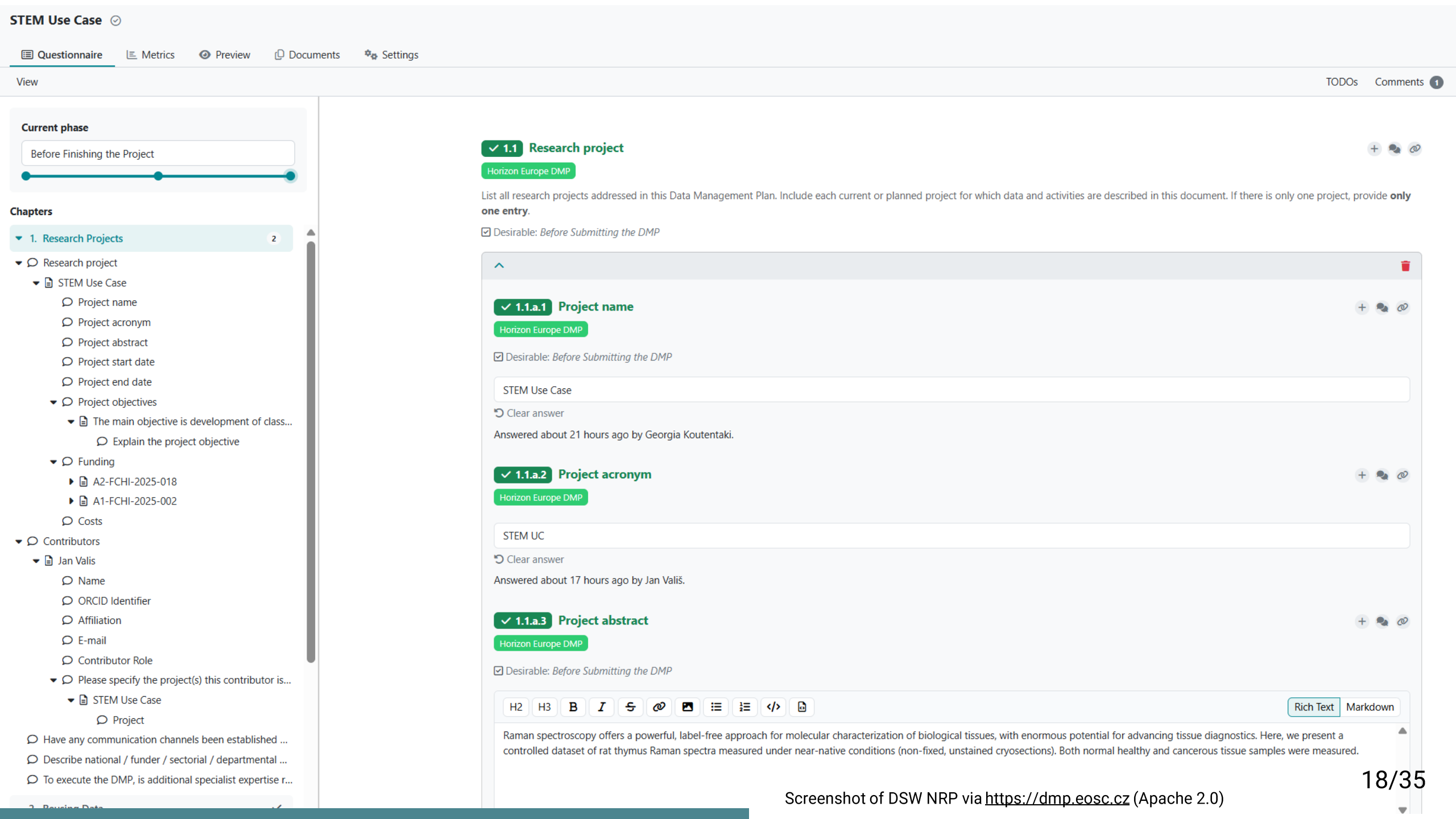
How to transfer the information from the README to the DMP.



Why you have both documents and what do they serve?

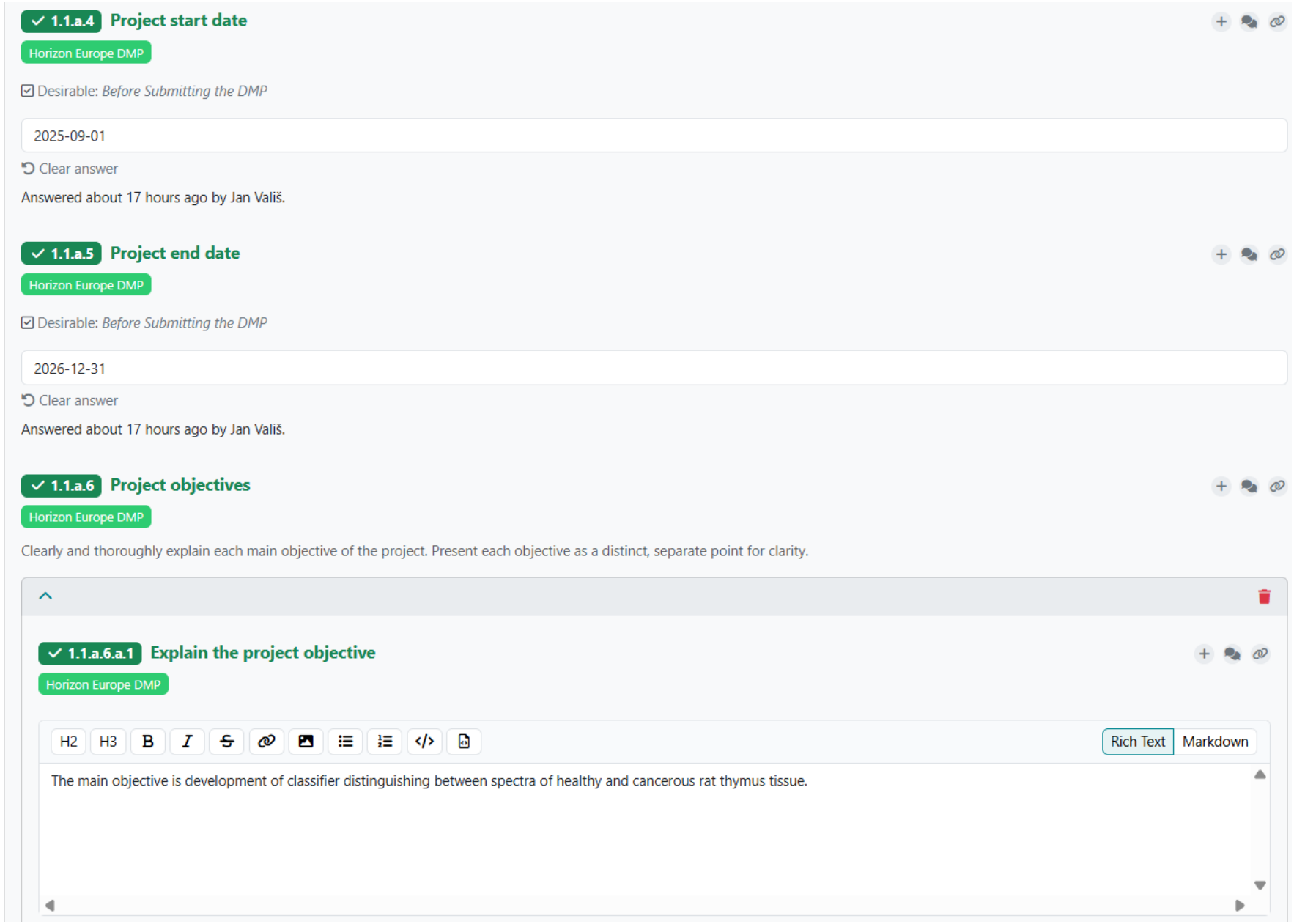


Functionalities of the NRP DSW that make your life easier.



Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

National Czech Programme | eosc



Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

✓ 1.2 Contributors

Horizon Europe DMP

Each person or institution contributing to creating or executing the data management plan should be added as a contributor. **A project probably should have a Contact Person, and a Data Curator.**

☑ Desirable: *Before Submitting the DMP*

✓ 1.2.a.1 Name

Horizon Europe DMP

☑ Desirable: *Before Submitting the DMP*

Jan Valis

Clear answer

Answered about 21 hours ago by Georgia Koutentaki.

✓ 1.2.a.2 ORCID Identifier

Horizon Europe DMP

Please provide only the ORCID identifier (for example, `0000-0002-1825-0097`).

Jan Vališ ORCID: **0000-0003-0349-6868**

valisj@vscht.cz

jan.valis@techlib.cz

valis@imc.cas.cz

Clear answer

Answered about 21 hours ago by Georgia Koutentaki.

✓ 1.2.a.3 Affiliation

Horizon Europe DMP

University of Chemistry and Technology, Prague (VŠCHT) type education type funder

Clear answer

Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

✓ 1.2.a.5 **Contributor Role**

Horizon Europe DMP

Roles in a project should be given as they are defined by datacite.

You should specify at least one "Contact Person". If your project has a work package for data management, identify the leader of that work package as "Data Curator". Additionally, at least one "Creator of DMP" must be identified to ensure they are listed as a creator in the final DMP.

☑ Desirable: *Before Submitting the DMP*

- ☐ **a. Contact Person**
- ☐ **b. Data Collector**
- ☑ **c. Data Curator**
- ☐ **d. Data Manager**
- ☐ **e. Data Protection Officer**
- ☑ **f. Data Steward**
- ☐ **g. Distributor**
- ☐ **h. Editor**
- ☐ **i. Producer**
- ☐ **j. Project Leader**
- ☐ **k. Project Manager**
- ☑ **l. Project Member**
- ☑ **m. Researcher**
- ☐ **n. Rights Holder**
- ☐ **o. Sponsor**
- ☑ **p. Supervisor**
- ☐ **q. Work Package Leader**
- ☑ **r. Creator of DMP**
- ☐ **s. Other**

Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

In DSW

3.1.a.1 Dataset name

Horizon Europe DMP

Give the name of the dataset.

Desirable: *Before Submitting the DMP*

Raman spectra of rat thymus tissue

Clear answer

Answered about 21 hours ago by Georgia Koutentaki.

3.1.a.2 Identifier of the dataset

Horizon Europe DMP

Please add all "formal" identifiers you have for this data set: these can be handles or DOIs or any other type. One important purpose of these identifiers is to be able to refer back to the dataset.

A good identifier is *persistent* (i.e. it does not change, and also the same identifier will never be used for another data set), *globally unique* (nobody else uses the same identifier for a different data set) and *resolvable* (you can actually locate the data set if you only know the identifier).

Please note that if you are publishing the dataset in a repository, the repository may create a new persistent identifier, or you may provide an existing one that has already been created. You will be able to specify such additional identifiers under the "publication" questions. Specify one main identifier here.

Desirable: *Before Finishing the Project*

3.1.a.2.a.1 Dataset Identifier Type

Horizon Europe DMP

Provide the type of identifier. Potential identifiers may include a DOI, Handle, or URL. You will be presented with suggestions from FAIRsharing, but you also have the option to manually enter the identifier type that is not listed there.

FAIRsharing.org **Digital Object Identifier** (https://www.doi.org/the-identifier/resources/handbook/, DOI: 10.25504/FAIRsharing.hFLKCn)

Clear answer

Answered about 21 hours ago by Georgia Koutentaki.

3.1.a.2.a.2 Dataset Identifier

Horizon Europe DMP

To enter the specific **value** of the identifier. If possible, refer exactly to the version that you are using.

For identifiers based on a type available in FAIRsharing, provide only the identifier itself (for example, a DOI, accession number, or ORCID identifier).

For custom or non-standard identifier types that are not available in FAIRsharing, and where a publicly accessible web page exists, provide the full URL instead of only the identifier value. This allows the identifier to be displayed as a clickable link in the generated document.

Examples

- DOI: `10.5281/zenodo.12345`
- ORCID: `0000-0002-1825-0097`
- PRIDE project: `https://www.ebi.ac.uk/pride/archive/projects/PXD034237`
- Custom repository record: `https://repository.example.org/record/12345`

Desirable: *Before Finishing the Project*

https://doi.org/10.5281/zenodo.20533749

In README

```
# DATASET NAME: Raman spectra of rat thymus tissue

## BASIC INFORMATION
Dataset version: 1.1 (2026-03-02)
README version: 1.2 (2026-10-19)
PID: [https://doi.org/10.5281/zenodo.18989675]
(https://doi.org/10.5281/zenodo.18989675)
Main contact: [janstovd@vscht.cz](mailto:janstovd@vscht.cz)
Project website: N/A

### Creators:
- Matyas Garnol, 0009-0006-8397-1364 (Investigation, Methodology);
- Daniela Janstova, 0000-0002-0300-3878 (Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project Administration, Software, Visualization);
- Jan Valis, 0000-0003-0349-6868 (Conceptualization, Data curation, Investigation, Methodology);
- Marcela Dendisova, 0000-0002-5895-243X (Resources - HW, Supervision);
- Jan Mares, 0000-0003-4693-2519 (Resources - HW&SW, Supervision)
### Contributors:
- Veronika Spisska, 0000-0003-2339-4906 (Resources - samples).


---

## DESCRIPTION

### Brief Description

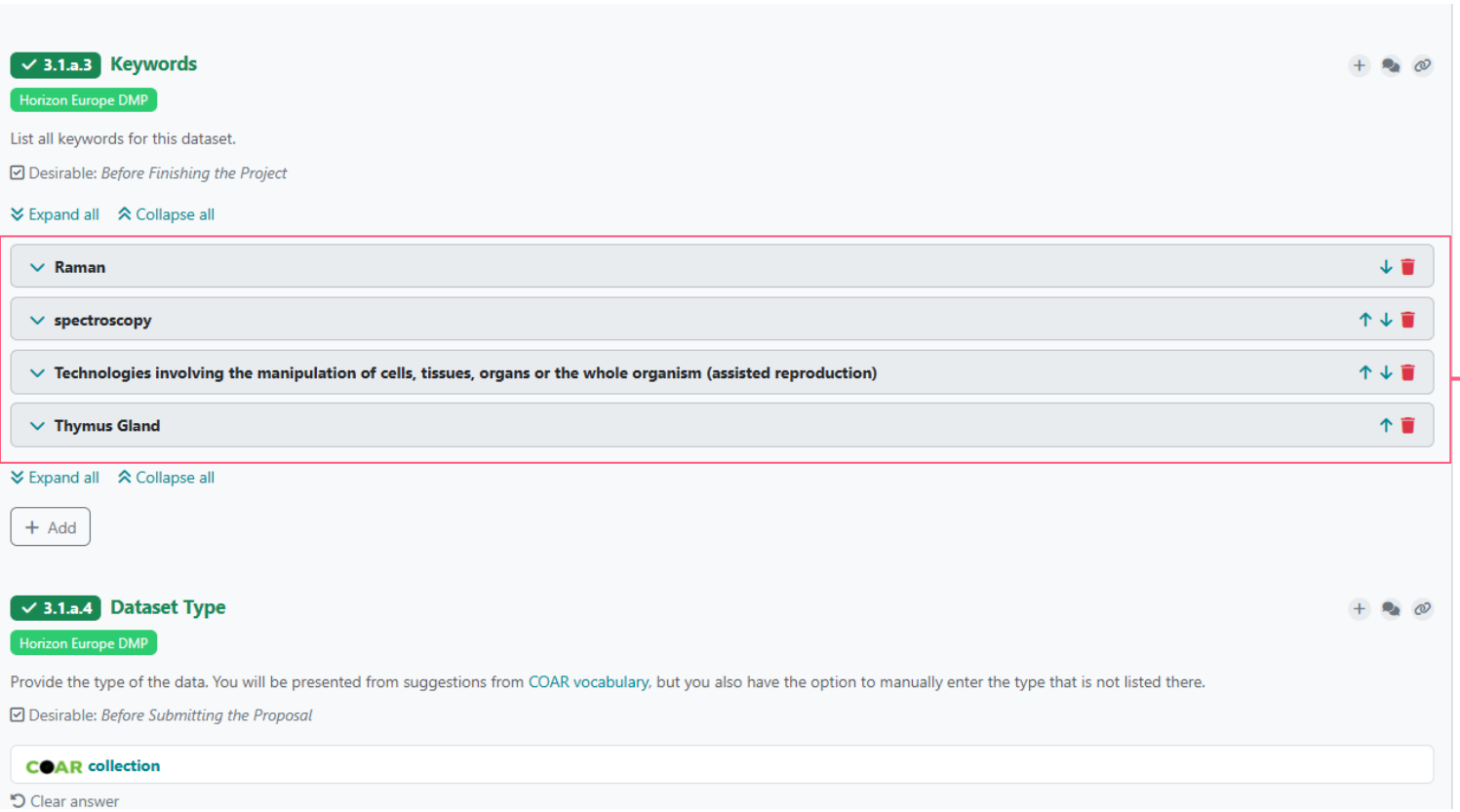
Raman spectroscopy offers a powerful, label-free approach for molecular characterization of biological tissues, with enormous potential for advancing tissue diagnostics. Here, we present a controlled dataset of rat thymus Raman spectra measured under near-native conditions (non-fixed, unstained cryosections). Both normal healthy and cancerous tissue samples were measured.

### Research Context

This labeled dataset can be used to train and evaluate pre-processi methods as well as machine-learning methods for classification of Raman spectra of biological samples.
```

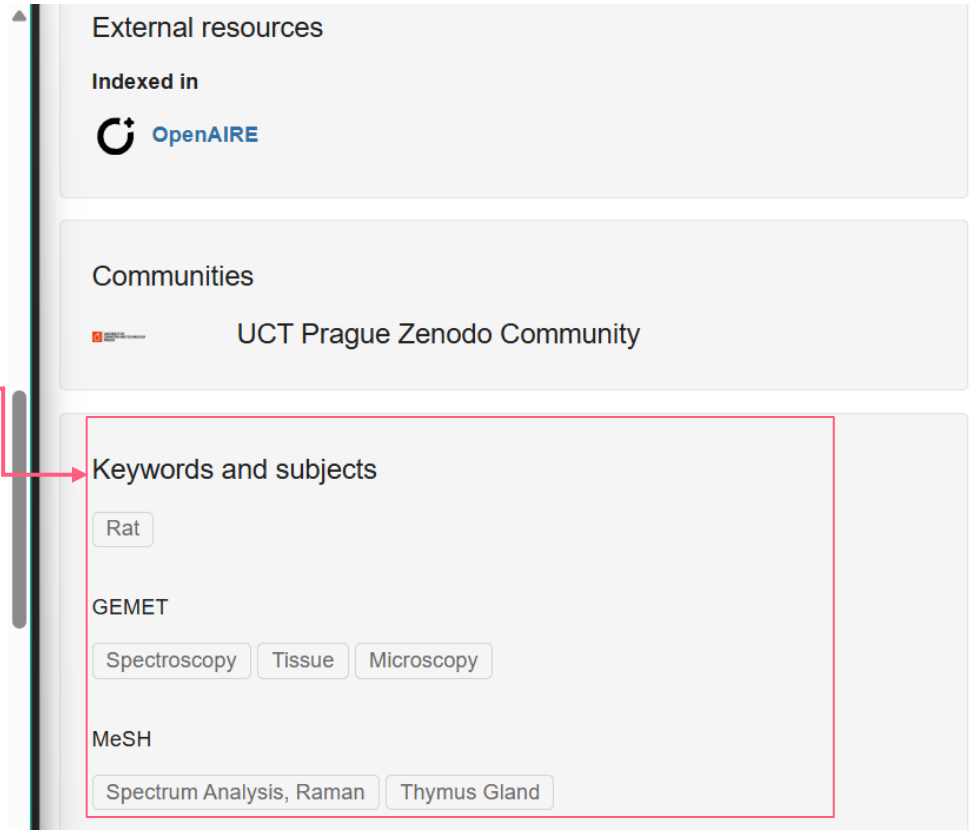
Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

In DSW



Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

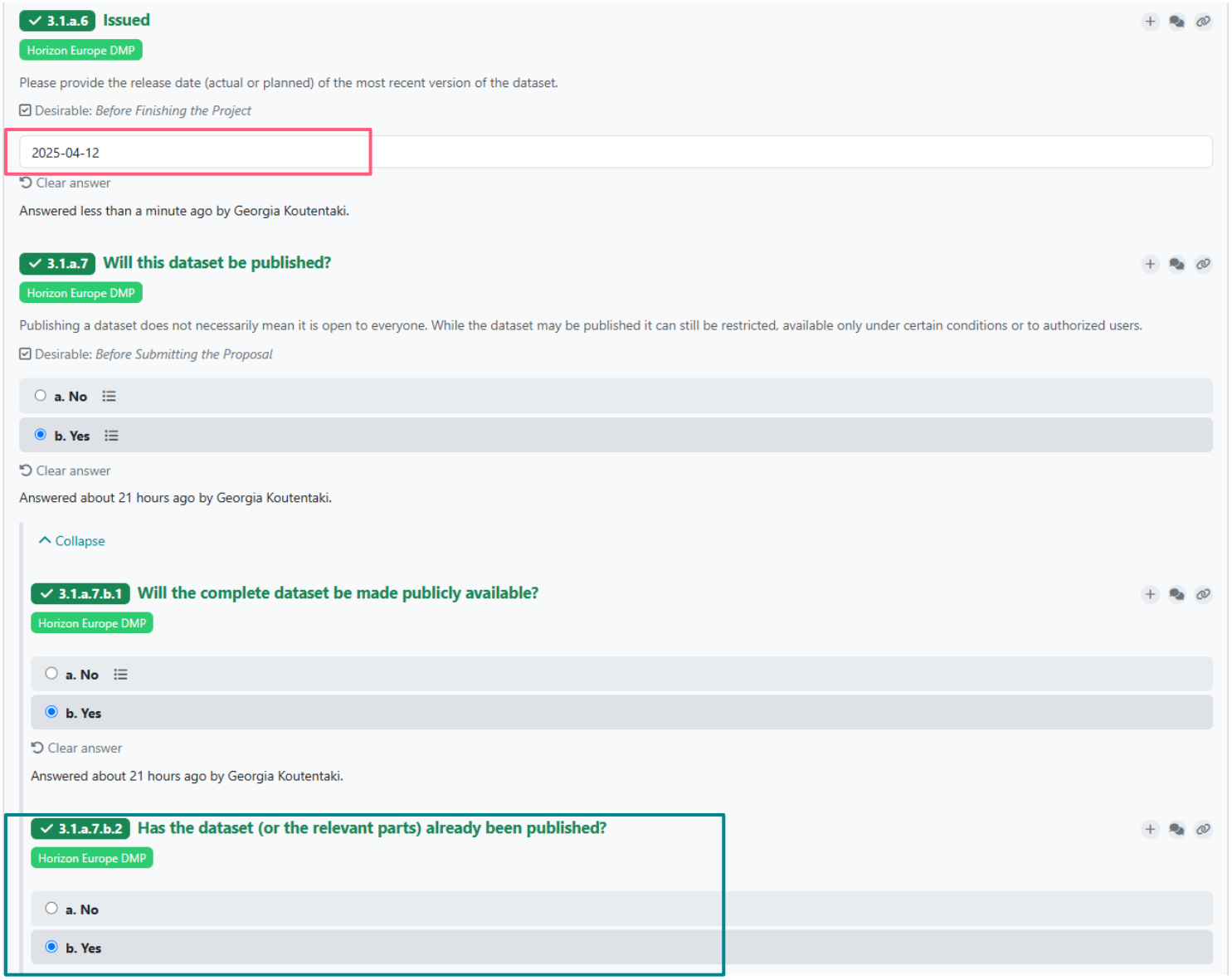
In Zenodo



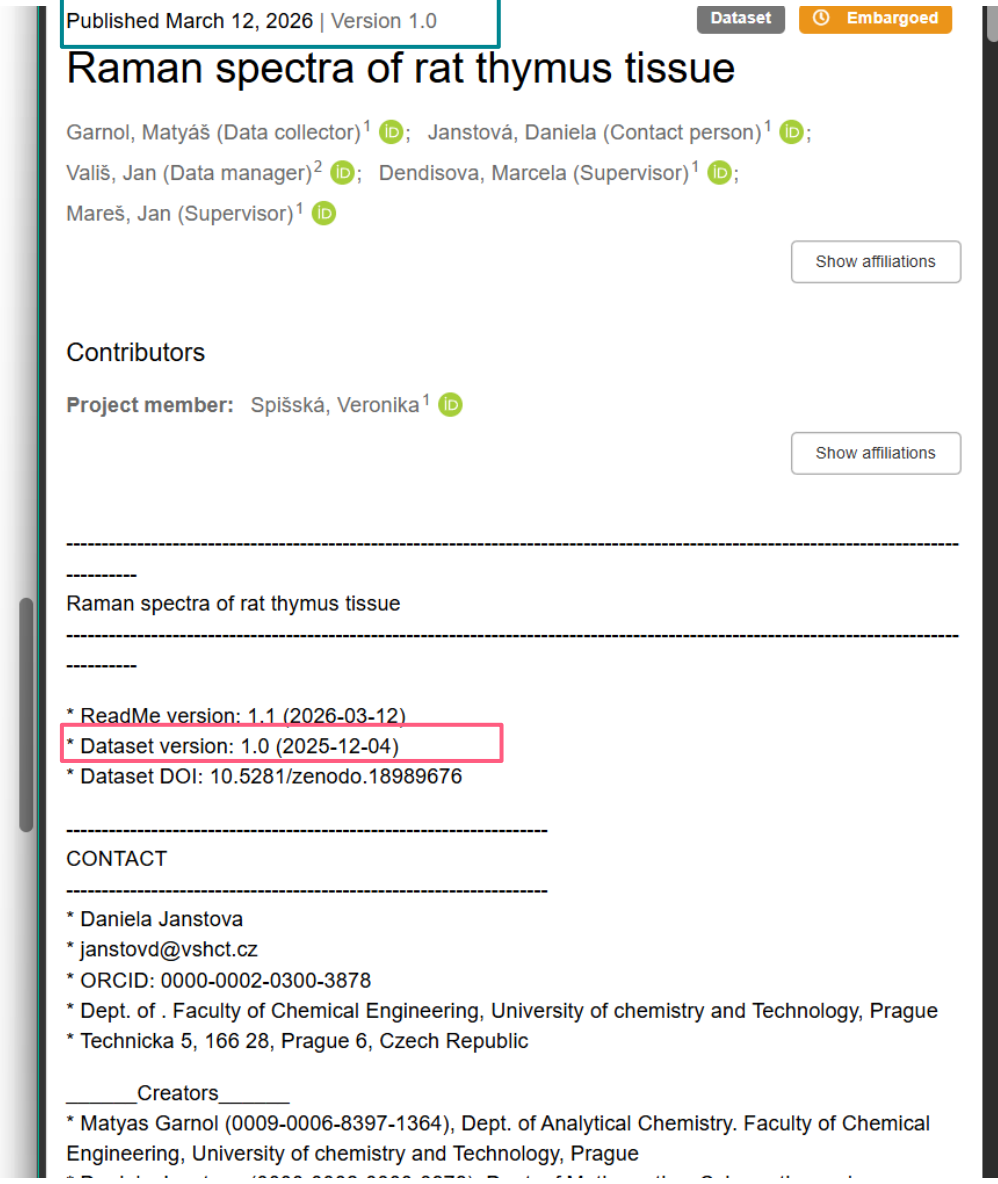
Screenshot adapted from the Zenodo interface: Garnol et al. (2026), *Raman spectra of rat thymus tissue* (Version 1.1) [Dataset]. https://doi.org/10.5281/zenodo.20533749

National Czech Programme | eosc

In DSW



In Zenodo



Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

Screenshot adapted from the Zenodo interface: Garnol et al. (2026), *Raman spectra of rat thymus tissue* (Version 1.1) [Dataset]. https://doi.org/10.5281/zenodo.20533749

In DSW

3.1.a.9 Can you list any qualified references to other data (e.g., from your project or previous research), and explain their relevance or relationship to your current data?

Horizon Europe DMP

Please list any references to other data sets, e.g. from your project or from earlier research.

A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is erefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextua

For example: *Dataset X was published in 2020 and assigned a DOI. In 2025, you build upon this dataset to produce Dataset Y, with the relationship between the two datasets clearly documented.*

Desirable: *Before Finishing the Project*

Attention! ONLY DATA
Not publications of articles

3.1.a.9.a.1 Identifier type of the data that will be referrend to

Horizon Europe DMP

Provide the type of identifier. You will be presented with suggestions from FAIRsharing, but you also have the option to manually enter the identifier type that is not listed there.

FAIRsharing.org **Digital Object Identifier** (https://www.doi.org/the-identifier/resources/handbook/, DOI: 10.25504/FAIRsharing.hFLKCn)

Clear answer

Answered about 16 hours ago by Jan Vališ.

3.1.a.9.a.2 Identifier of the data that will be referred to

Horizon Europe DMP

To enter the specific **value** of the identifier. If possible, refer exactly to the version that you are using.

For identifiers based on a type available in FAIRsharing, provide only the identifier itself (for example, a DOI, accession number, or ORCID identifier).

For custom or non-standard identifier types that are not available in FAIRsharing, and where a publicly accessible web page exists, provide the full URL instead of only the identifier value. This allows the identifier to be displayed as a clickable link in the generated document.

Examples

- DOI: `10.5281/zenodo.12345`
- ORCID: `0000-0002-1825-0097`
- PRIDE project: `https://www.ebi.ac.uk/pride/archive/projects/PXD034237`
- Custom repository record: `https://repository.example.org/record/12345`

10.5281/zenodo.18989676

Clear answer

Answered about 16 hours ago by Jan Vališ.

3.1.a.9.a.3 The relationship between this dataset and the referred data

Horizon Europe DMP

IsVersionOf

Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

In Zenodo

No citations found

183 VIEWS

11 DOWNLOADS

Show more details

Versions

Version 1.1 — Jun 3, 2026
10.5281/zenodo.20533749

Version 1.0 — Mar 12, 2026
10.5281/zenodo.18989676

View all 2 versions

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.18989675. This DOI represents all versions, and will always resolve to the latest one. Read more.

External resources

Indexed in

OpenAIRE

Communities

UCT Prague Zenodo Community

Screenshot adapted from the Zenodo interface: Garnol et al. (2026), *Raman spectra of rat thymus tissue* (Version 1.1) [Dataset]. https://doi.org/10.5281/zenodo.20533749

In DSW

4.1.a.10.b.1 Ethical committee name

Horizon Europe DMP

Name of the ethics committee or institutional review board that approved the research project.

Ministry of Education, Youth and Sports of the Czech Republic

Clear answer

Answered about 21 hours ago by Georgia Koutentaki.

4.1.a.10.b.2 Ethical committee approval number

Horizon Europe DMP

Official reference or approval number assigned by the ethics committee for the approved research project.

MSMT-31592/2019-4

Clear answer

Answered about 21 hours ago by Georgia Koutentaki.

4.1.a.10.b.3 Ethical issues report link

Horizon Europe DMP

Please specify where the protocol from the meeting with the ethics committee can be accessed.

4.1.a.10.b.4 Ethical issues report

Horizon Europe DMP

Optionally, insert the Ethical Issues Report itself.

Upload file

4.1.a.10.b.5 Is a new approval or extension of any existing approval under research ethics laws needed for re-use of the data?

Horizon Europe DMP

If the data that you will re-use is coupled to people, the ethical approval that was originally obtained may not be covering your current research. In that case an extension of the original approval, or a new approval, may be necessary.

Desirable: *Before Submitting the DMP*

- a. Not applicable
- b. The existing ethical approval under research ethics laws covers our reuse
- c. An extension of existing ethical approval under research ethics laws will be needed
- d. New ethical approval under research ethics laws will be needed to cover our usage of the data

In README

```
10.5281/zenodo.18989675.

---

## ETHICS AND DATA PROTECTION

### Personal and Sensitive Data Protection

N/A (No human personal data involved; animal models only).

### Ethical Approval

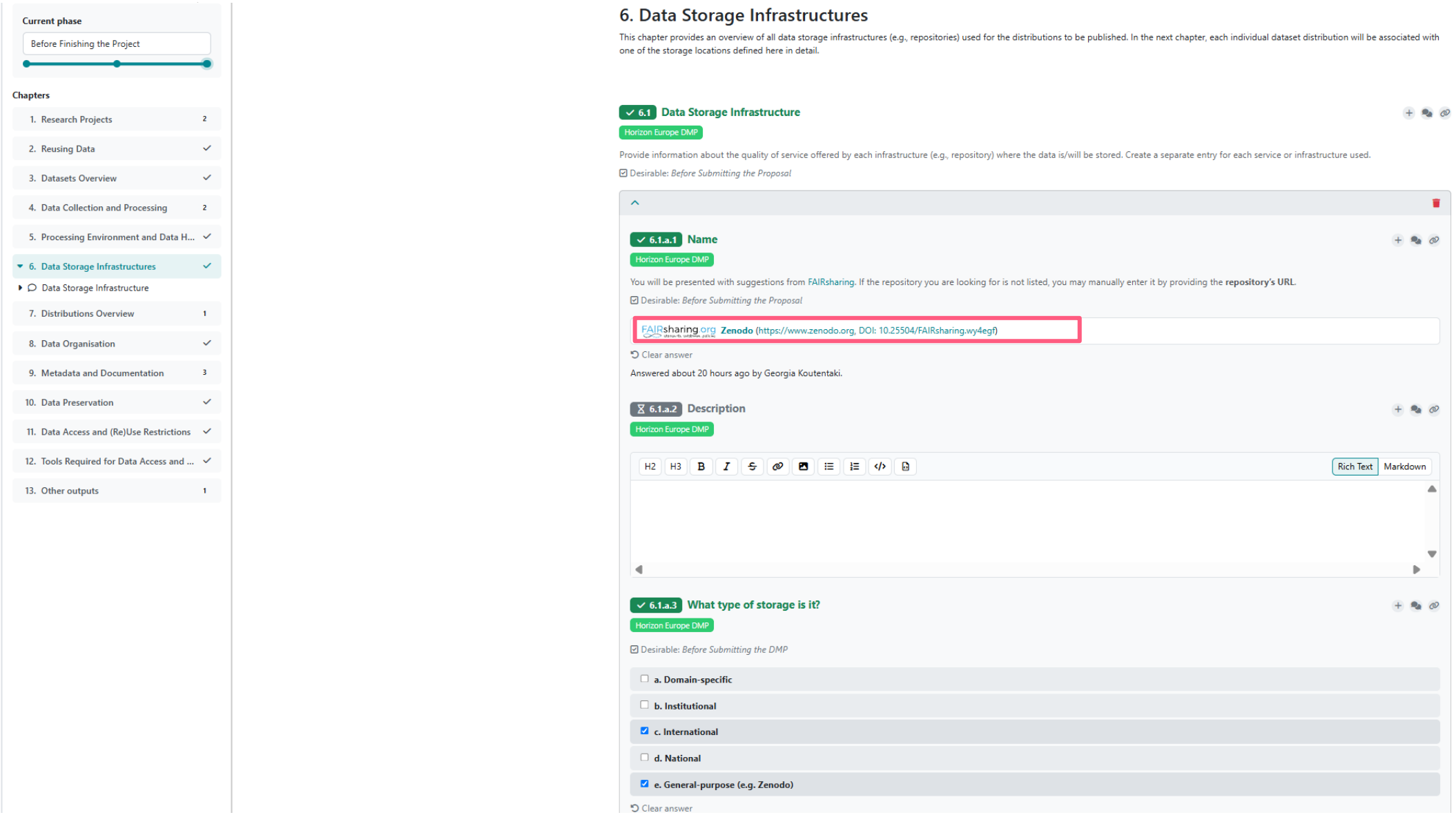
All experimental procedures were conducted in strict accordance with the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and complied with the Animal Protection Law of the Czech Republic. The experimental protocol was approved by the Ministry of Education, Youth and Sports of the Czech Republic (approval number: MSMT-31592/2019-4, December 2, 2019).

### Informed Consent

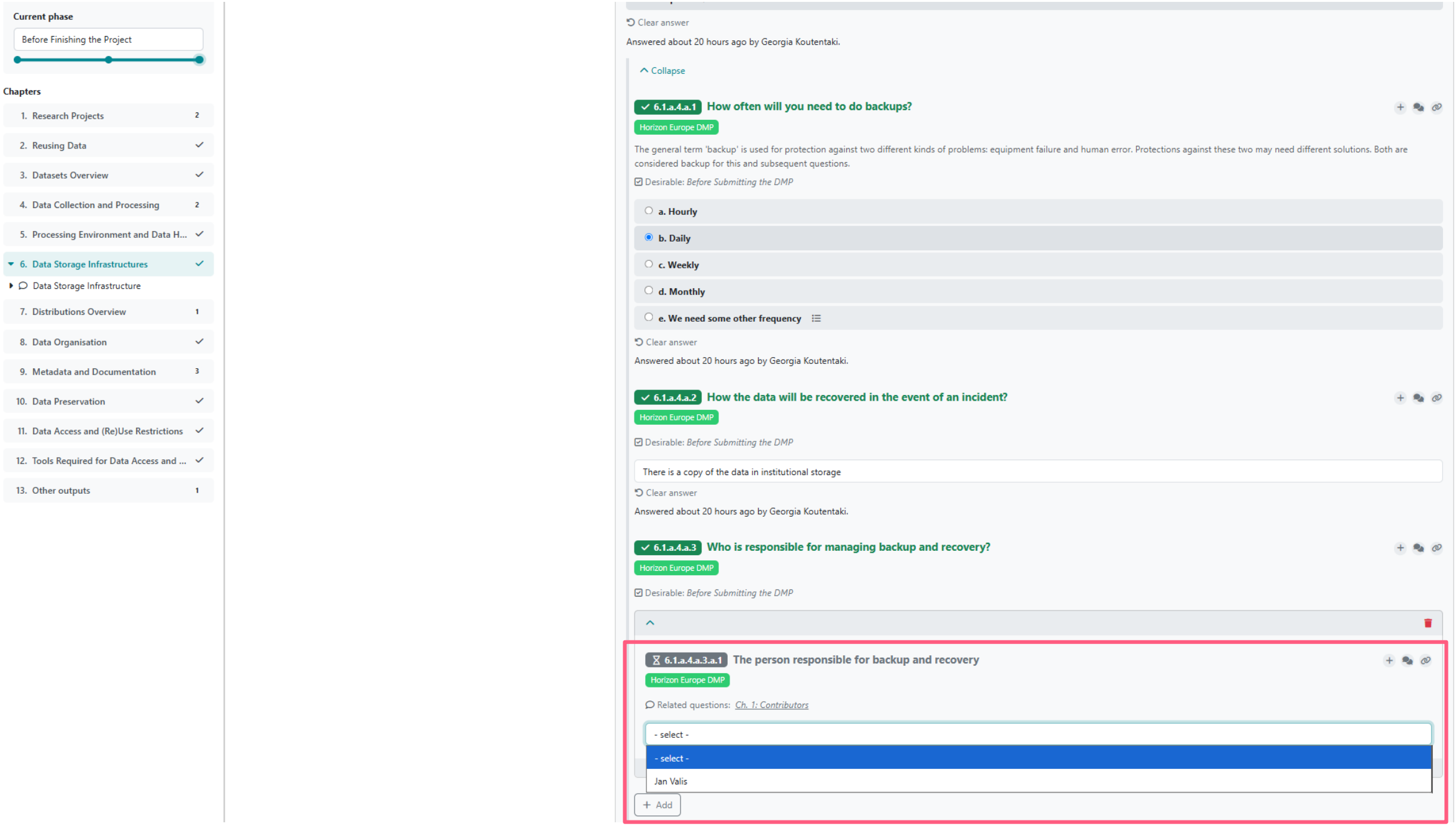
N/A

---
```

Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)



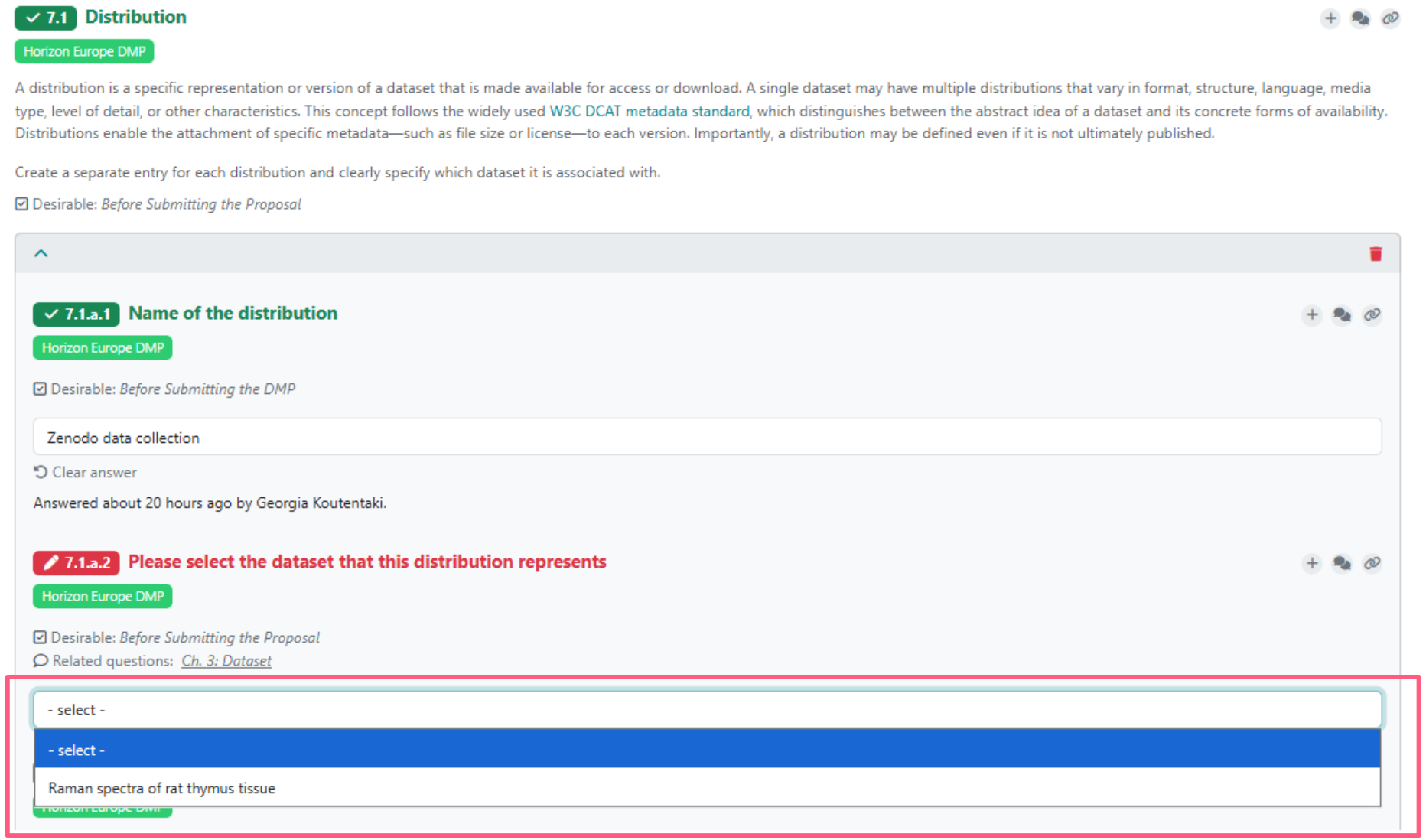
Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)



Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

7. Distributions Overview

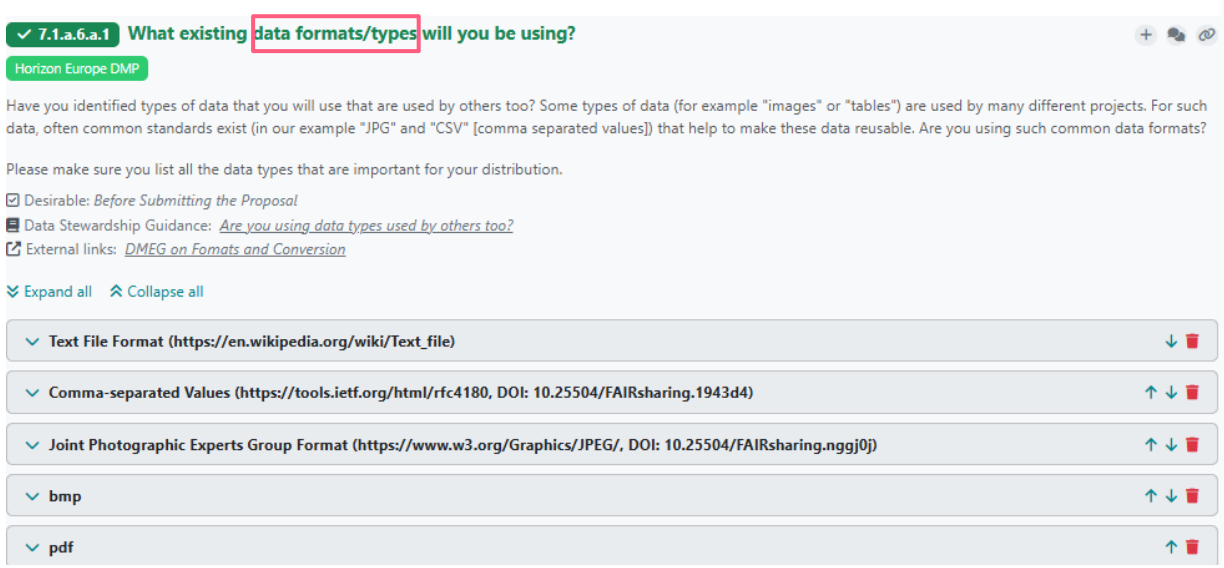
This chapter provides an overview of all distributions created from the datasets produced by your project(s).



Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

National Czech Programme | eosc

In DSW



In README

```
34
35
36
37
38
39
40
41 ## DATASET/COLLECTION CONTENTS
42
43 ### Directory Structure
44
45 - 000_ReadMe.md
46 - 000_Documentation/ : Ethical Board Applications/Approvals, DMPs, SOPs, and Metadata
47 - 001_Data-Raw_Spectra/ : Raw Raman spectra (.txt files)
48 - 003_Data-Raw_Pictures/ : Raw pictures from the microscope (.bmp, .jpg files)
49 - 004_Code/ : Python scripts for PreProcessing, QC, and Classification
50 - 005_Publications/ : Related diploma thesis, conference materials, and journal papers
```

Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

National Czech Programme | eosc

In DSW



Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

In Zenodo

Embargoed

The files will be made publicly available on December 31, 2026

Reason: Until journal article is accepted. For access to the dataset before this date, plese, get in touch with the contact person listed in the description

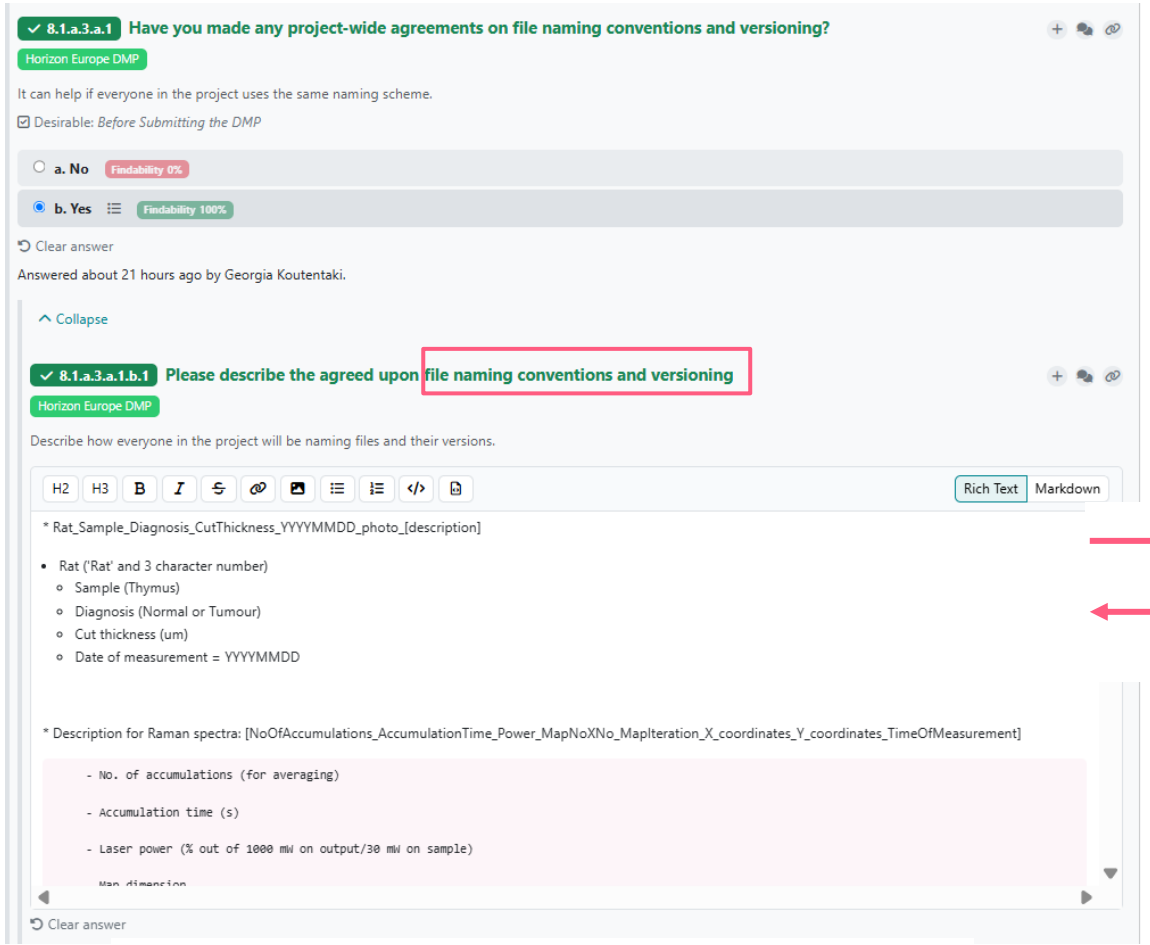
Additional details

Funding	**University of Chemistry and Technology** A2-FCHI-2025-018 **University of Chemistry and Technology** A1-FCHI-2025-002
Dates	**Created** 2025-03-25/2025-12-04 Data collection

Screenshot adapted from the Zenodo interface: Garnol et al. (2026), *Raman spectra of rat thymus tissue* (Version 1.1) [Dataset]. https://doi.org/10.5281/zenodo.20533749

In DSW

In README





Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

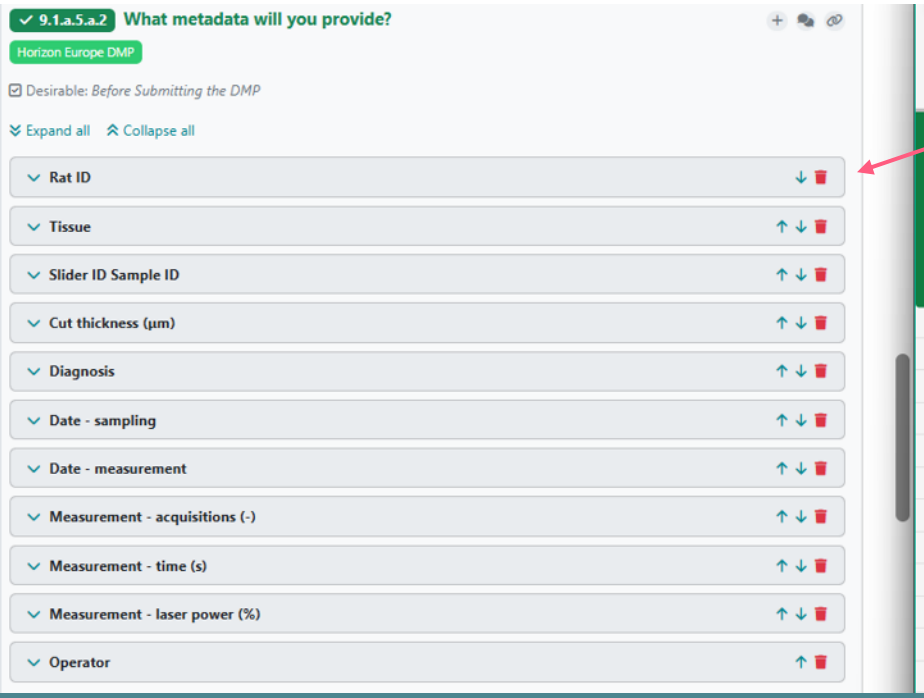
National Czech Programme | eosc

Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

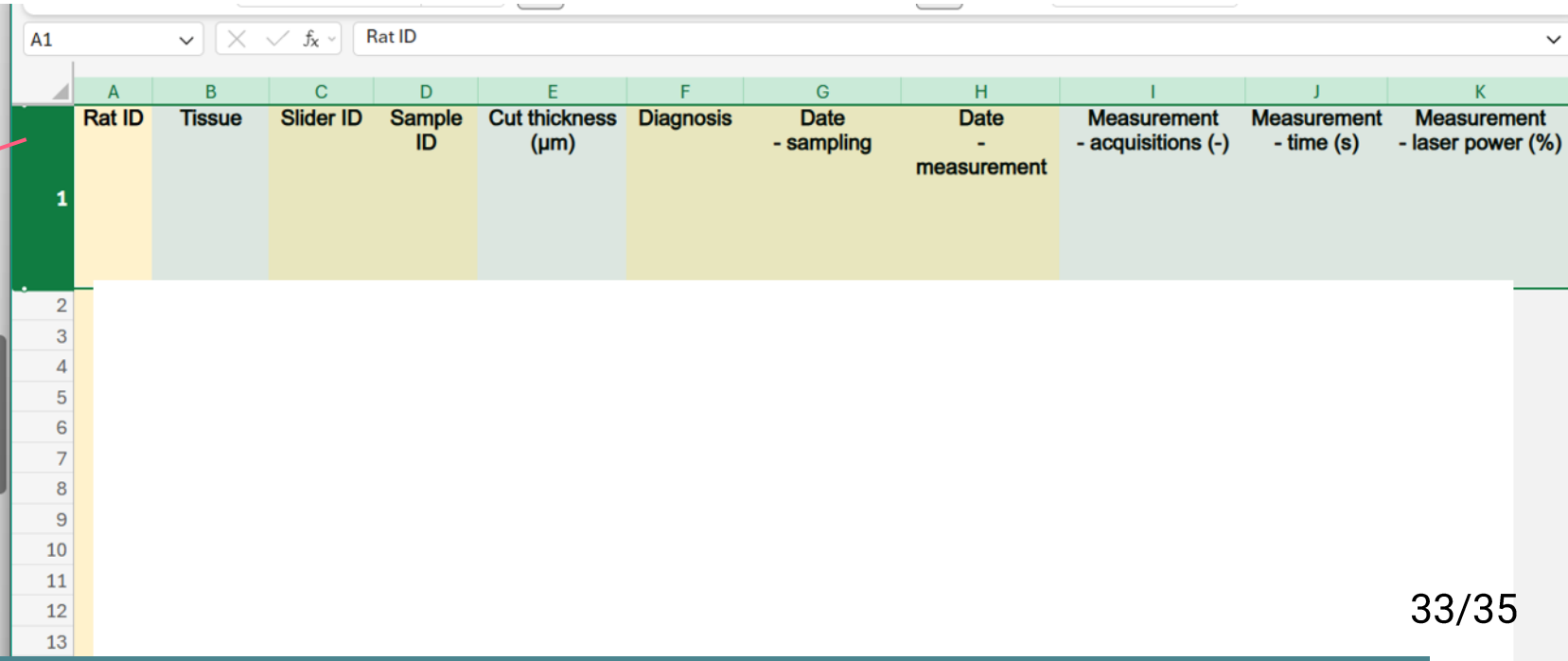
Documentation Folder

- Thymus_Doc_20240102_EthicalBoardApplication.docx
- Thymus_Doc_20240102_EthicalBoardApplication.pdf
- Thymus_Doc_20240115_EthicalBoardApproval.pdf
- Thymus_Doc_20240116_SOP-v1.docx
- Thymus_Doc_20240116_SOP-v1.pdf
- Thymus_Doc_20260413_DMP-v2.pdf
- Thymus_Metadata.xlsx

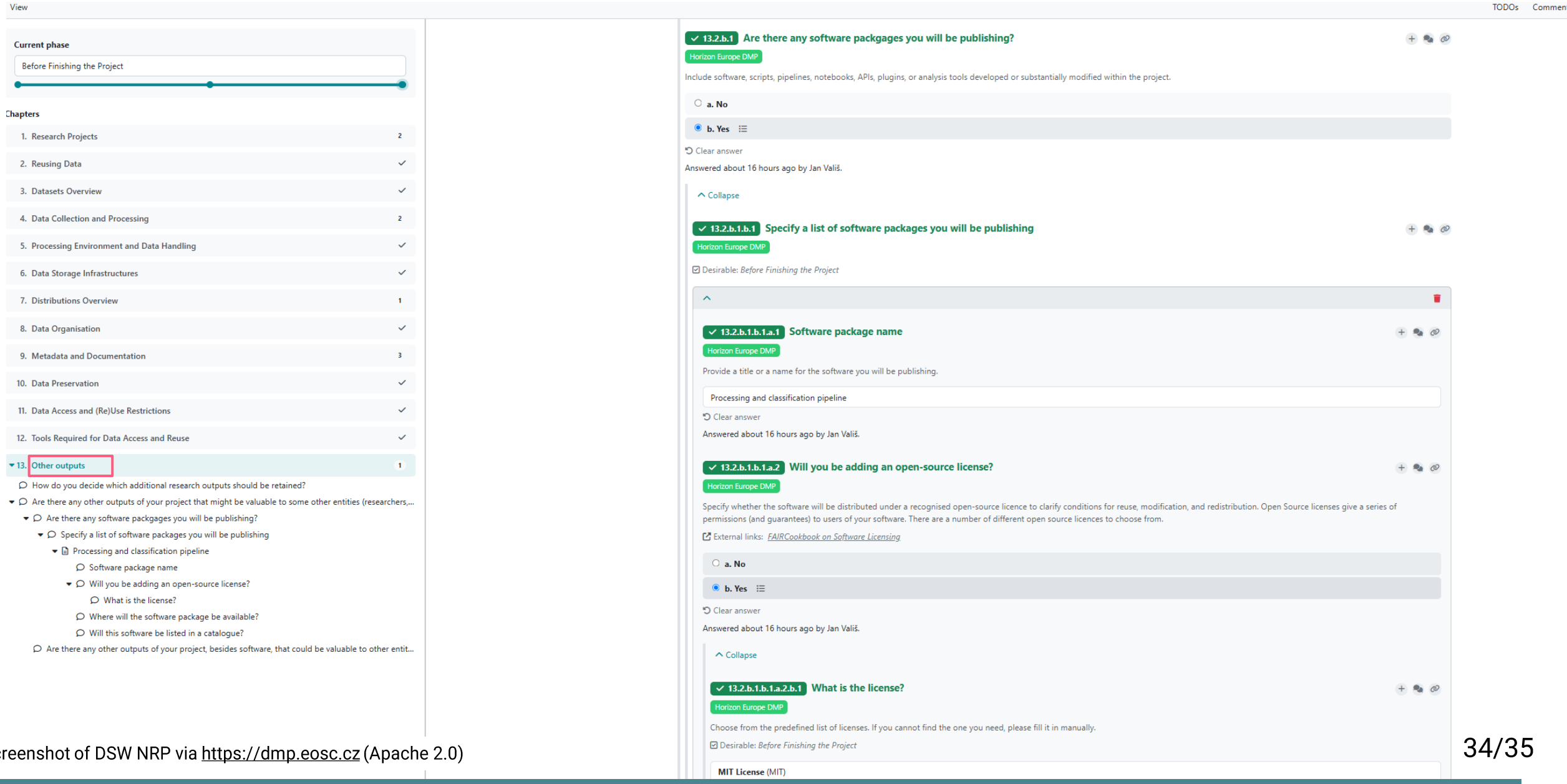
In DSW



In Metadata file



National Czech Programme | eosc



Screenshot of DSW NRP via https://dmp.eosc.cz (Apache 2.0)

References and sources

- Data Stewardship Wizard (dmp.eosc.cz). Knowledge Model: dmp.eosc.cz:czech-nrp-km:1.0.2. Used under Apache-2.0. Document template: Horizon Europe DMP (1.0.2) for the Czech National Repository Platform.
- Garnol, M., Janstová, D., Vališ, J., Dendisova, M., & Mareš, J. (2026). Raman spectra of rat thymus tissue (Version 1.1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.20533749
- European Organization For Nuclear Research, & OpenAIRE. (2013). *Zenodo*. CERN. https://doi.org/10.25495/7GXK-RD71
- Pergl, R., Hooft, R., Suchánek, M., Knaisl, V., & Slifka, J. (2019). "Data Stewardship Wizard": A Tool Bringing Together Researchers, Data Stewards, and Data Experts around Data Management Planning. *Data Science Journal*, 18(1), 59. https://doi.org/10.5334/dsj-2019-059

National Czech Programme | eosc

Questions?

georgia.koutentaki@cvut.cz

jan.valis@techlib.cz

Co-funded by the European Union

Ministry of Education, Youth and Sports of the Czech Republic

CC BY