

How I Learned to Stop Worrying and Love the DMP

Antonín Zita <zita@ics.muni.cz>,

Masaryk University | CERIT-SC

Long Live Research Data: National EOSC CZ Conference 2025, 2–3 December, Ostrava



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

OPEN DATA

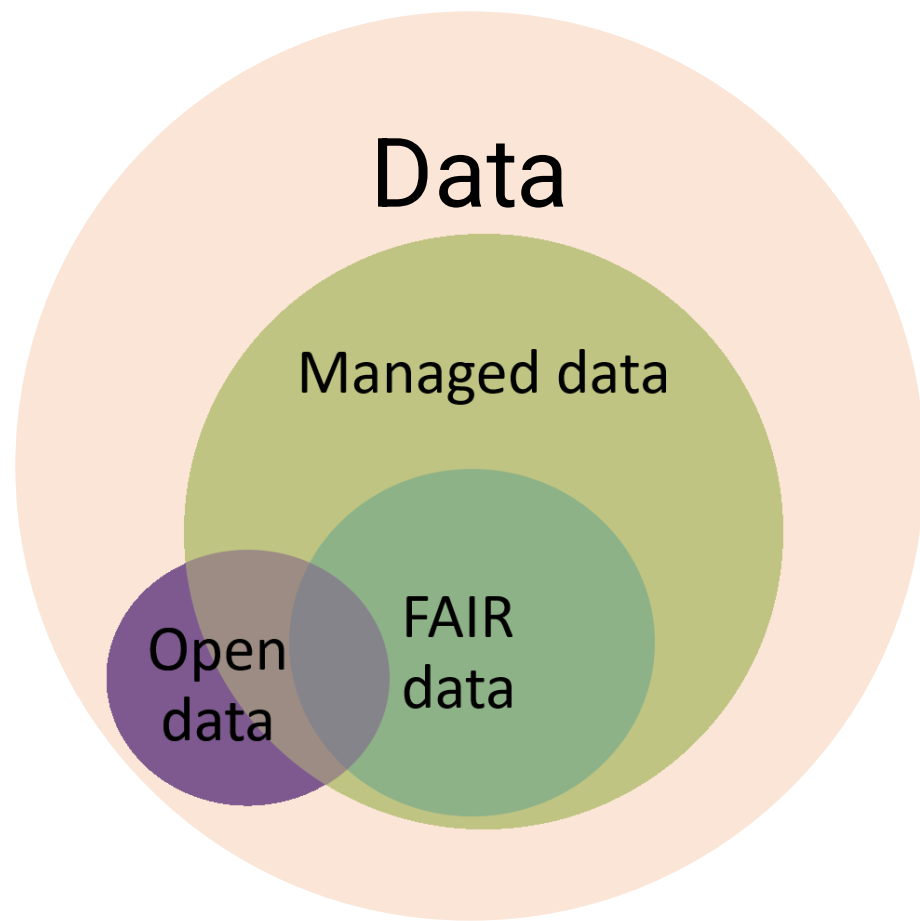
FAIR Data

Research data is necessary to enable verification and further use of research results. In order to comply with its legal obligations in the area of open data, GACR has included research data management in its grant funding policy. The objectives of this policy are:

- Required
 - MEYS, GA CR, ...
 - Open Data vs FAIR data
 - <https://gacr.cz/ga-cr-a-otevrena-veda/>
 - all scientific publications produced within GACR-funded projects have traceable open data (with legal exceptions such as the protection of intellectual property, trade secrets, national security or legitimate commercial interests of the recipient or a third party)
 - GACR project investigators submit a Data Management Plan (DMP) latest along with the submission of the first Interim Project Report, and they update the DMP regularly
 - open data follows the FAIR principles of Findable, Accessible, Interoperable, and Reusable (FAIR)
 - costs related to open data are eligible project costs, GACR recommends the creation of a data steward position within the project
- [Národní agentura pro zemědělský výzkum](#) zabezpečuje národní dotace projektů výzkumu a vývoje financovaných z účelových prostředků Ministerstva zemědělství. Výzvy jsou vyhlašovány v souladu se zákonem č. 130/2002 Sb. a s ohledem na výzkumná data podporují zásady FAIR.

Otevřená věda v ČR. MŠMT.

FAIR Principles



- FAIR principles
 - Findable
 - Accessible
 - Interoperable
 - Reusable

“As open as possible, as closed as necessary.”

Sourcej: FAIR Handbook for the Data Steward Community in the Czech Republic (ENG), <https://doi.org/10.71495/hxhc-6f57>

DMPlanning vs DMP

1. Data Management Plan

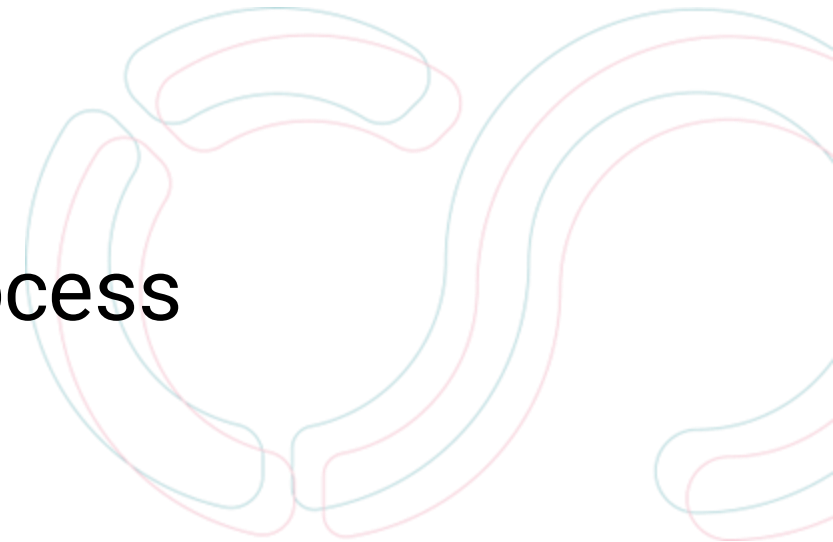
vs.

2. Data Management Planning

Document

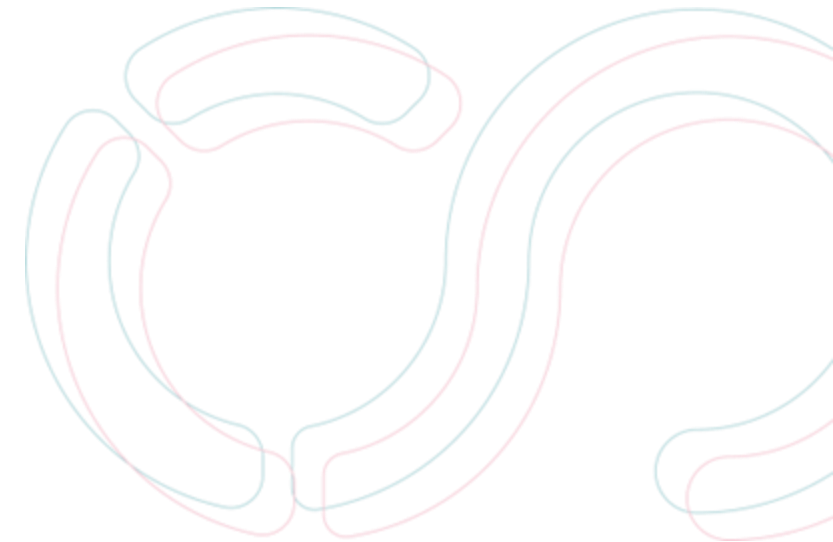
vs.

Process



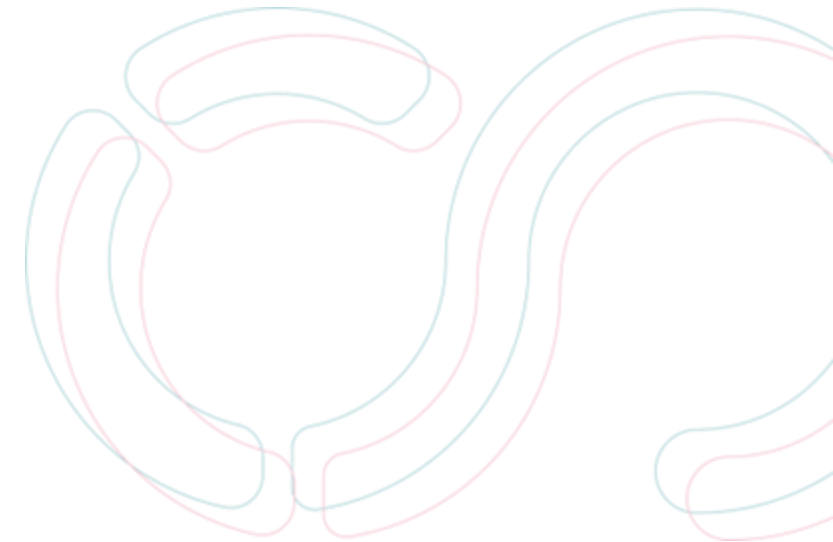
DMP Myth No. 1

- Information in the DMP cannot be changed
- DMP is a **living document**



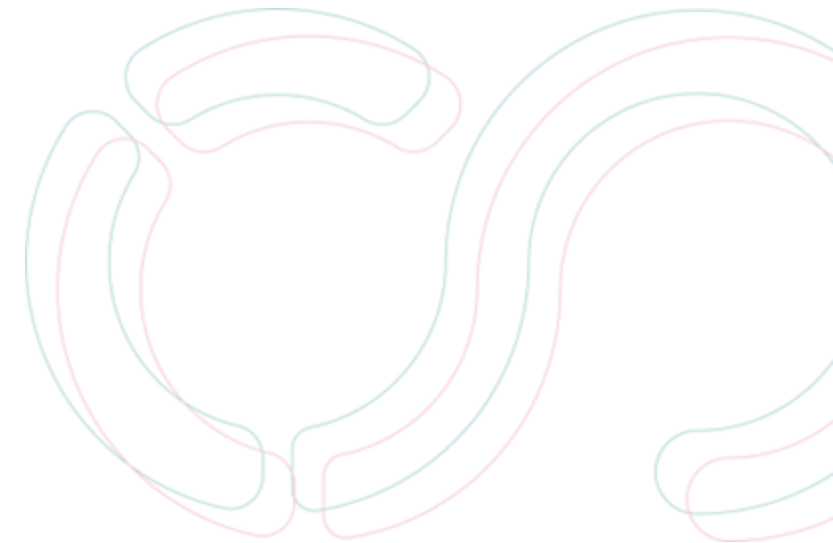
DMP Myth No. 2

- DMP should contain only the “correct” solutions
- DMP should contain the **actual solutions**



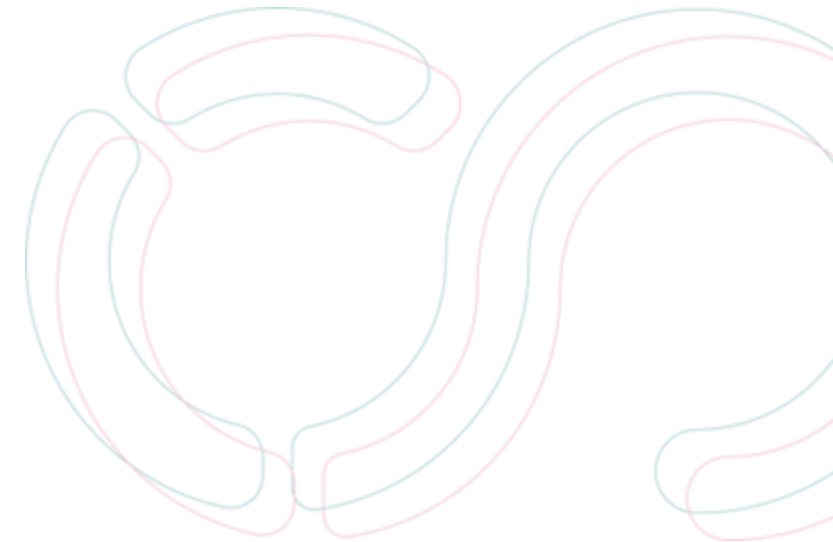
DMP Myth No. 3

- **FAIR data** must always be **open data**
- Option of **embargo/not publishing data**



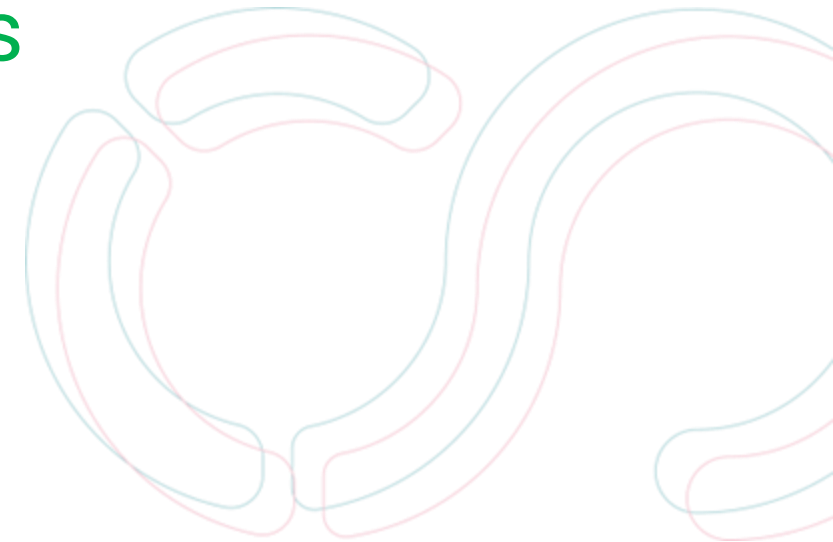
DMP Myth No. 4

- Completing the DMP in MS Word is easier
- Open-ended answers can **lead to issues**



DMP Myth No. 5

- DMP is useful only for the funding agency
- **DMP(lanning)** is useful for the researchers



DMP Put Simply AKA “It ain’t no rocket science!” “

The following must be absolutely clear from the DMP:

Yes, we know how to handle our data.

Yes, we will create data in these formats and quantities, and yes, we want our data to be FAIR because we want to preserve it or increase its value. Yes, we will use secondary data. Yes, we will use repositories. Yes, we will publish our data under a certain license. ...

DMP – Horizon Europe

DATA MANAGEMENT PLAN

Date:	[dd/mm/yyyy]
Version:	[DMP version]

1. Data Summary

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

What types and formats of data will the project generate or re-use?

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

What is the expected size of the data that you intend to generate or re-use?

What is the origin/provenance of the data, either generated or re-used?

To whom might your data be useful ('data utility'), outside your project?

2. FAIR data

2.1. Making data findable, including provisions for metadata

Will data be identified by a persistent identifier?

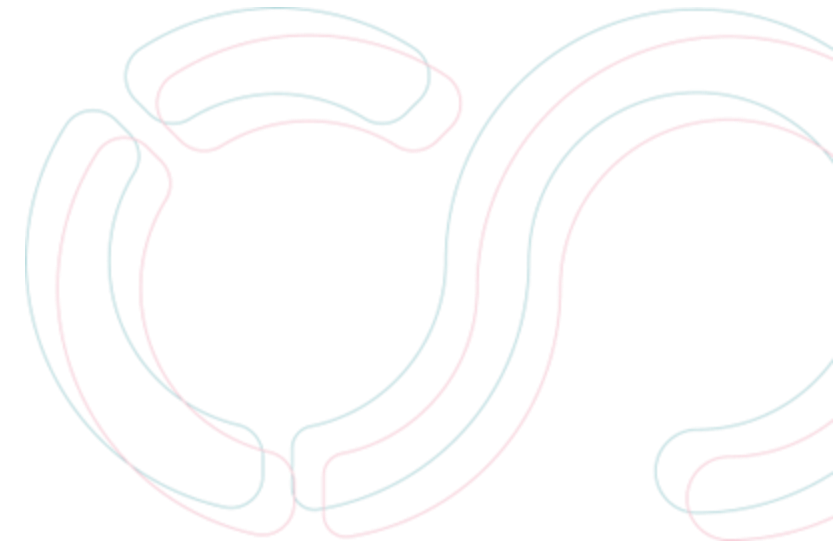
Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Will metadata be offered in such a way that it can be harvested and indexed?

Data Stewardship Wizard – General Information

- Questionnaire open source tool
- Recommended in [Horizon Europe Programme Guide](#)
- Tracks FAIR metrics
- Filters questions by using tags
- External validation (email, ORCID, DOI, ...)
- Templates
- Creating the DMP document



Gaining Access to DSW (I)

ELIXIR DSW

- DSW Cloud within the ELIXIR infrastructure (dsw.elixir-europe.org)

Institutional Instance

- DSW is an open-source tool with publicly available documentation
- Institutions can have their own instances in which they maintain the current **Knowledge Model** and can make further modifications

FAIR Wizard

- Commercial service for institutions outside of ELIXIR or that do not wish to maintain their own instance

Gaining Access to DSW (II)

DSW NRP (dmp.eosc.cz)

- Part of the **National Repository Platform (NRP)**
- KM and other content developed specifically for the needs of the NRP
- KM different from the “regular” KM DSW, but still meeting the needs of the given calls (HE, MEYS, GA CR, ...)
- Still in pilot mode, but already accessible to **everyone*** (***within the e-INFRA infrastructure**)

Data Management Planning within National Repository Platform for Research Data



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

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

Start planning



Pilotní provoz služby

Tato služba pro plánování správy dat je pilotní součástí Národní repozitářové platformy pro výzkumná data. Obsah i funkce průběžně vylepšujeme v rámci projektu a budeme rádi za vaši zpětnou vazbu.

Pilot Service Notice

This data management planning service is a pilot within the National Repository Platform for research data in the Czech Republic. We are continuously improving its content and features over the course of the project and welcome your feedback.



Co-funded by
the European Union



Provoz této služby je podporován projektem Národní repozitářová platforma pro výzkumná data (grant číslo CZ.02.01.01/00/23_014/0008787).

Log In

[Forgot your password?](#)

Log In


Or connect with

 e-INFRA CZ

Pro přihlášení se do této služby je nutné se nejprve v rámci e-INFRA CZ AAI registrovat [ZDE](#).

To log in into this service it is first required to register within e-INFRA CZ AAI organization [HERE](#).


 DMP.eosc.cz

 Dashboard

 Projects

 eosoc.cz

 DSW Guide

 Antonín Zita
Researcher

Welcome, Antonín!

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



You have no projects yet, start by creating some.



Create Project

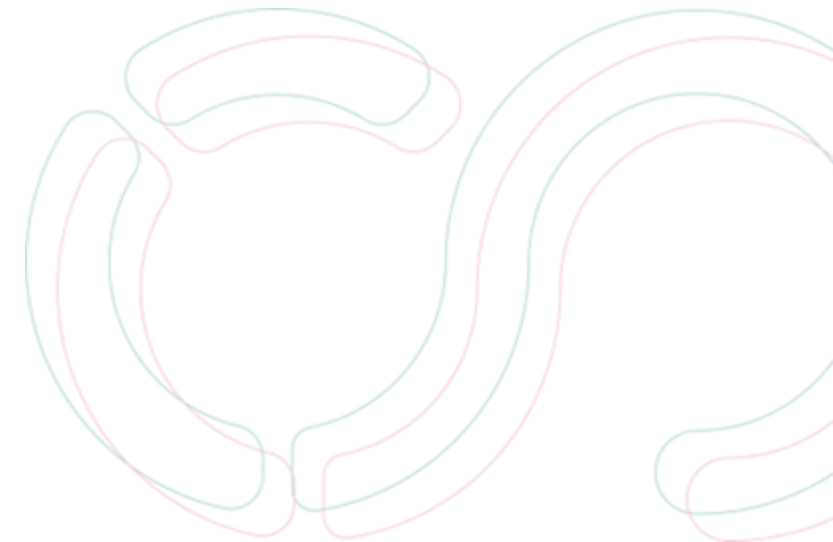
Project is a workspace where you create your DMP. It is based on a knowledge model, which contains knowledge about what should be asked and how based on the research field or organization's needs. You can use document templates to transform the answers into a document. This document can be anything,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

Data Stewardship Wizard – Other Functions

- Project Ownership
 - Management multiple users
 - Sharing, commenting, ...
- Project Phases
 - Required and optional items
 - Visual feedback for questionnaire tracking
- Document Export Different
 - File types and templates
 - Migration between new KM versions
- Versioning



 DS Wizard

Dashboard


Projects

Storage Costs Evaluator

DSW Guide

Antonin Zita
Researcher

⏪ Collapse sidebar

Test 5 

Questionnaire

Metrics

Preview

Documents

Settings

View


Import replies


Share

TODOs

Comments

Version history





Current phase

Before Submitting the DMP

Chapters

I. Administrative information3

II. Re-using data✓

III. Creating and collecting data5

IV. Processing data1

V. Interpreting data✓

VI. Preserving data6

IV. Processing data

In the processing phase, the data will be undergoing the mostly automated steps for processing, before the analysis and interpretation.

In this chapter, many questions are focusing on the compute environment that is used to process the data and make it available for interpretation by project partners. Some of those questions (e.g. on workflow systems and data provenance) are also relevant for the work in the interpretation phase.

✓ IV.1 Will you be using a shared working space to work with your data?


Horizon Europe DMP

Will you be using a working space containing data and software specific to the project that is shared between all the people working on the data in the project? Sometimes such a system is called a *Virtual Research Environment*.

☒ Desirable: Before Submitting the Proposal

☒ a. No, participants in the project each have different collections of data and tools

☐ b. Yes

 Clear answer


Answered 2 minutes ago by Antonin Zita.

✓ IV.2 How will you validate the integrity of the results?

Horizon Europe DMP

☒ Desirable: Before Submitting the DMP

☒ a. Explore

 Clear answer

Answered 2 minutes ago by Antonin Zita.

Collapse

IV.2.1 Will you run a subset of your jobs several times across the different compute infrastructures you are using?

User Interface – Basics

- DMP phases
- DMP chapters
- Question types:
 - Mandatory questions
 - Optional questions
- Answer types:
 - Multiple choice
 - Open text
 - Text completion

Current phase

Before Submitting the DMP

 VI.1.a.6






 VI.1.a.6

 VI.1.a.5

Chapters

I. Administrative information



- ▶  Contributors
- ▶  Research Project(s)
 -  To execute the DMP, is additional specialist e...
 -  Do you require hardware or software in addi...
- ▶  Describe national / funder / sectorial / depar...

II. Re-using data



III. Creating and collecting data



IV. Processing data



V. Interpreting data

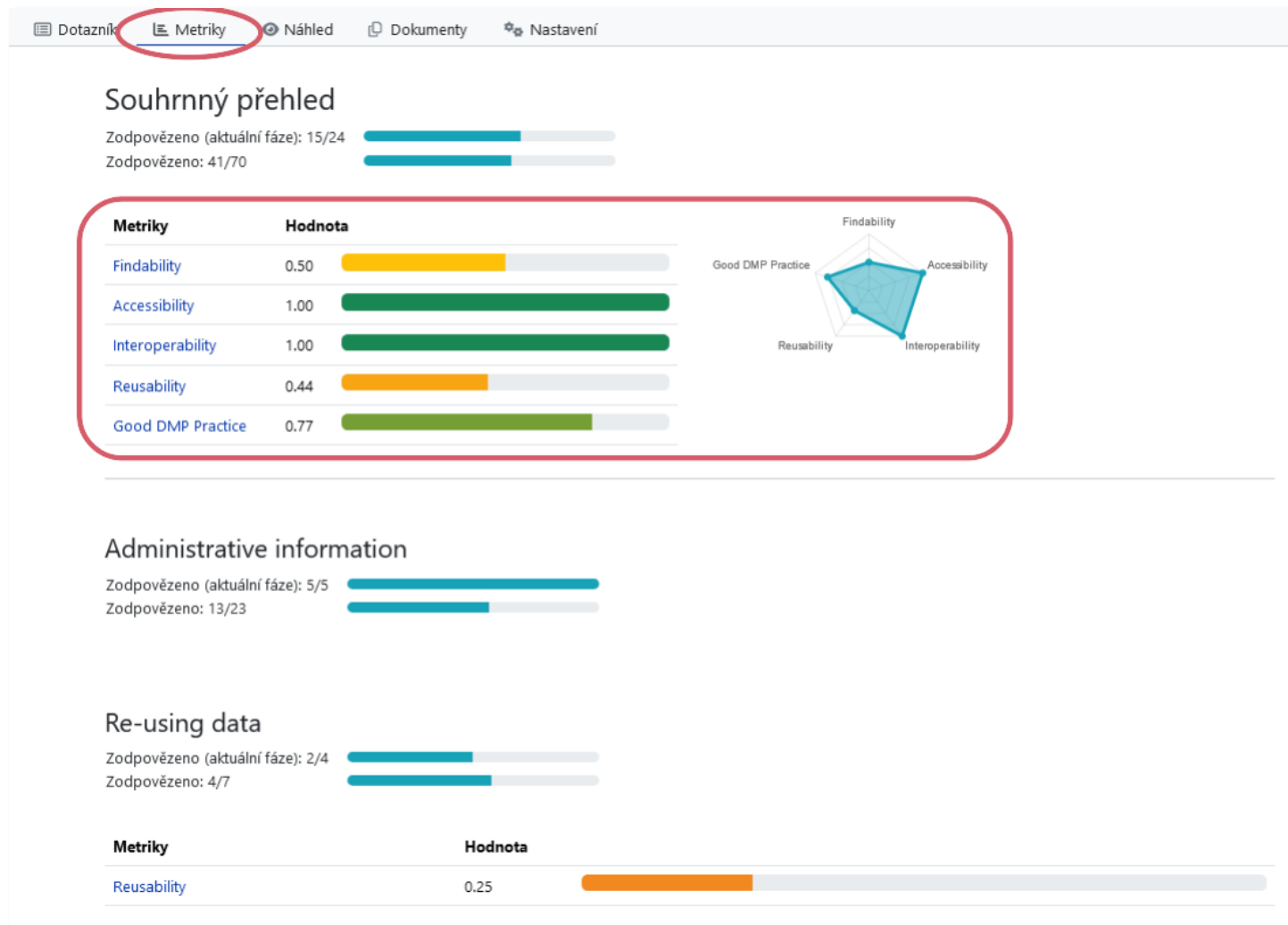


VI. Preserving data

2

VII. Giving access to data





Did you consider the possible impact to the organization if information is lost?

Good DMP Practice

e fázi: *Before Submitting the Proposal*

Good DMP Practice

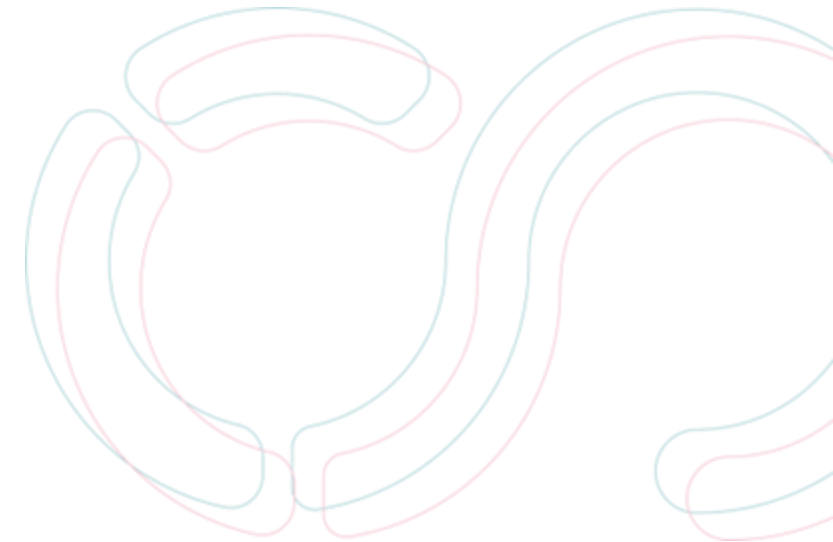
the effect is small Good DMP Practice

the risk is acceptably low Good DMP Practice

we will need to work on this. Good DMP Practice

Future Development of DSW NRP

- Modification of KM NRP
 - Better work with AAI
 - Translations of KM and templates
- Integration with NRP services and repositories
- Search available
- Cross-referencing questions



Future Development of DSW NRP

Questionnaire

Metrics

Preview

Documents

Settings

View

Import replies

Actions

TODOs 1

Comments

Version history

Current phase

Before Submitting the DMP

Chapters

I. Administrative information6

II. Re-using data1

Contributors

Research Project(s)

To execute the DMP, is additional spe...

Do you require hardware or software i...

Describe national / funder / sectorial / ...

I. Administrative information

✓ I.1 Contributors

Horizon 2020 DMPHorizon Europe DMPScience Europe DMPmaDMP

Each person 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

✓ Albert Einstein

+ Add

re-use

Results for: re-use

Question

...rom elsewhere, and facilitate re-use of your own data later, it is...

Chapter

...ctice to check all options to re-use existing available data, eith... ..eful) make them available for re-use by others. And the circle islete if such data is actually re-used.

Chapter

...ion needed by people who will re-use your data, and with the acces...

Future Development of DSW NRP

✓ I.1 What datasets are you collecting?

⌵ Expand all ⌴ Collapse all

✓ **Butterfly specimen dataset**



✓ **Flowers dataset**



+ Add

Future Development of DSW NRP

Reference

 c9baed0a

 Move

 Delete

Reference Type

Cross Reference

Question

What datasets are you collecting?

[Go to related question](#)

Description

Consider all datasets you are collecting



Future Development of DSW NRP

II.1 How big storage do you need (in GBs)?

💬 Related questions: What datasets are you collecting? (Consider all datasets you are collecting)

Thank you for your attention

E: <zita@ics.muni.cz>, W: eosc.cz



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY