

UKLÁDÁNÍ
VÝZKUMNÝCH
DAT –
SEMINÁŘ
PRO
MENDELOVU
UNIVERZITU

AMUNI

2023-03-16



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

MUNI
ICS

Seminář pro Mendelovu univerzitu
Ukládání výzkumných dat

Michal Růžička, ÚVT MU
ruzicka@ics.muni.cz

2023-03-16

Obsah přednášky

1. Kontext
2. Životní cyklus výzkumných dat
3. FAIR principy
4. DMPlanning vs. DMPlan
5. Nástroje a zdroje pro podporu FAIR principů
6. Úložiště pro výzkumná data v instituci a eInfrastruktura
 - Příklad Masarykovy univerzity
 - Příklad CESNET
 - Služby Mendelu
7. Repozitáře pro výzkumná data praktičtěji
 - Typy repositářů
 - Jak najít vhodný repositář
 - Příklad MAFIL CEITEC MU
8. EOSC
 - Implementace EOSC v ČR
9. Kam dál
 - Kurz Bezpečná správa výzkumných dat
 - Kybernetická bezpečnost

Kontext

Open Science a projekty

– Podmínky:

- Horizon Europe
- Exceles
- OP JAK – Špičkový výzkum

– Povinné/nepovinné principy

- **Povinné**: Open Access, Data Management (DMP)
- **Nepovinné**: Citizen Science, Open Source, Preprints, ...

Životní cyklus výzkumných dat

Životní cyklus výzkumných dat



- Jaká data (znovu) používáte
 - včetně licencí, které vám to umožňují,
- jaká data generujete a jakým způsobem,
- kde je ukládáte, zálohujete, dlouhodobě uchováváte,
- jak je trvale a jedinečně identifikujete,
- jak je zpracováváte,
- jak je analyzujete,
- kde je zveřejňujete a sdílíte,
- kdo tohle všechno zaplatí;
- o čem data skutečně jsou,
- k čemu jsou data (ne)vhodná,
- kdo může data znovu použít,
- jaká konkrétní data podporují vaše výsledky,
- jak je použít pro opakování vašich experimentů atd.

Zdroj: ELIXIR RDMkit, <https://rdmkit.elixir-europe.org/>

MUNI
ICS

FAIR principy

FAIR principy – Findable

The first step in (re)using data is to find them. Metadata and data should be **easy to find** for both **humans** and **computers**. **Machine-readable metadata** are essential for automatic discovery of datasets and services, so this is an essential component of the [FAIRification process](#).

- *F1. (Meta)data are assigned a **globally unique and persistent identifier**.*
- *F2. Data are described with **rich metadata** (defined by R1 below).*
- *F3. **Metadata** clearly and **explicitly include the identifier of the data they describe**.*
- *F4. (Meta)data are **registered or indexed in a searchable resource**.*

Zdroj: GO FAIR, <https://www.go-fair.org/fair-principles/>

FAIR principy – Accessible

Once the user finds the required data, she/he/they need to know how can they be accessed, possibly including authentication and authorisation.

- A1. (Meta)data are *retrievable by their identifier using a standardised communications protocol*.
 - A1.1 The *protocol is open, free, and universally implementable*.
 - A1.2 The *protocol allows for an authentication and authorisation procedure, where necessary*.
- A2. *Metadata are accessible, even when the data are no longer available*.

Zdroj: GO FAIR, <https://www.go-fair.org/fair-principles/>

FAIR principy – Interoperable

The data usually need to be **integrated with other data**. In addition, the data need to **interoperate with applications** or **workflows** for analysis, storage, and processing.

- I1. (Meta)data use a **formal, accessible, shared, and broadly applicable language for knowledge representation**.
- I2. (Meta)data use **vocabularies that follow FAIR principles**.
- I3. (Meta)data **include qualified references to other (meta)data**.

Zdroj: GO FAIR, <https://www.go-fair.org/fair-principles/>

FAIR principy – Reusable

The ultimate goal of FAIR is to optimise the **reuse of data**. To achieve this, metadata and data should be **well-described** so that they can be replicated and/or combined in different settings.

- *R1. (Meta)data are richly **described with a plurality of accurate and relevant attributes**.*
 - *R1.1. (Meta)data are released with a clear and accessible **data usage license**.*
 - *R1.2. (Meta)data are associated with **detailed provenance**.*
 - *R1.3. (Meta)data meet **domain-relevant community standards**.*

Zdroj: GO FAIR, <https://www.go-fair.org/fair-principles/>

FAIR data – časté otázky

- „*I když mám zkušenost s DMP, stále mi není zcela jasný jeho skutečný účel a jaký bude mít přínos pro naši skupinu.*“
 - Uvědomění si, že data jsou „občané první kategorie“.
 - Data vyžadují péči – váš výzkum kriticky závisí na podkladových datech.
 - Jednou vyprodukovaná užitečná data by měla sloužit dlouhodobě, neměla by zmizet / stát se nesrozumitelnými s koncem výzkumného projektu.
 - DMP by měl být zejména onen proces, který pomáhá plánovat a provádět správu dat; neměl by to být primárně DMP dokument pro grantovou agenturu.

FAIR data – časté otázky (2)

- „*V jaké ‚kvalitě‘ by měla být sdílena výzkumná data? Raw data? Vyčištěná data? Analyzovaná data? Agregovaná data?*“
 - Pokud byste byli **vy v roli příjemců** těchto dat, **jaká data byste chtěli dostat**, aby pro vás byla **srozumitelná** a **užitečná** pro **další použití**?
 - Posuzujte **individuálně pro každý dataset** a sdílejte v této úrovni „kvality“.

DMPlanning vs. DMPlan

DMPlanning vs. DMPlan

1. Data Management Planing

„**Proces** plánování, popisu a informování o životním cyklu dat a **činností** spojených s jejich správou v průběhu výzkumu.“

2. Data Management Plan (DMP)

„**Dokument**, který **popisuje** tyto **činnosti** (dokumenty často vyžadují poskytovatelé dotací či grantů).“

Zdroj: Petra Dědičová, Data management a jak psát data management plan
<https://www.slideshare.net/butlibrary/data-management-a-jak-psat-data-management-plan-41441697>

DMP jako projektová podmínka

- DMP je stále častěji vyžadován jako standardní součást podání/řešení/finalizace projektů.
- Např. Horizon Europe požaduje
 - iniciální DMP,
 - aktualizace DMP uprostřed projektu,
 - finální revize DMP na konci projektu.
- DMP je živý dokument – měl by být průběžně aktualizován po celou dobu projektu.
- Obdobně v OP JAK, Exceles; nově i v TAČR a očekáváme i GAČR od roku 2023, nejpozději 2024.

– **Tip:** *Horizon Europe Open Science Requirements in Practice*

– OpenAIRE webinar, 13. 3. 2023

– Slajdy:

<https://doi.org/10.5281/zenodo.7324363>




Nástroje pro tvorbu DMP

- **Sdílený dokument** (Google Docs, Office 365 Word, Overleaf, ...).
 - Podporuje zejména přípravu DMP dokumentu.
- **Specializovaný software**, typicky webová aplikace umožňující spolupráci.
 - Dle míry sofistikovanosti může podporovat nejen tvorbu DMP dokumentu, ale také samotný proces správy výzkumných dat.


Data Stewardship Wizard

<https://ds-wizard.org/>

DS Wizard

My Experiment   

[Questionnaire](#) [Metrics](#) [Preview](#) [Documents](#) [Settings](#)

View [Import answers](#) [Comments](#) 4 [TODOs](#) 1 [Version history](#) 

Projects

- List
- Importers

Documents

Settings

Current Phase

Before Submitting the Proposal



Chapters

- I. Administrative information 1
- II. Re-using data 4
- III. Creating and collecting data 7
- IV. Processing data 4**
- V. Interpreting data 2
- VI. Preserving data 4
- VII. Giving access to data 3

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.


1 Will you be using a shared working space to work with your data?   1 comments

[Horizon 2020 DMP](#) [Horizon Europe DMP](#) [Science 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 5 minutes ago by Albert Einstein.

1.b.1 Will this work space be run by dedicated specialists? TODO 


If your work space is run and maintained by specialists, e.g. the ICT department of one of the institutes involved in the projects, this means that backup and restore as well as access management is properly addressed.

Desirable: Before Submitting the DMP

Named versions only

October 2022


▼ 4. 10.


9:38 

Current **1.0.0**

Will this work space be run by dedicated specialists?


Yes


 Albert Einstein

9:38 


Will you be using a shared working space to work with your data?


Yes


 Albert Einstein

9:37 

Data format/type

 XML Query Language

 Albert Einstein

9:34 

Cleared reply of Data set:

“In preparing for battle
I have always found that
plans are useless,
but planning is indispensable” ...

— Dwight D. Eisenhower

Ukázkové DMP

- University of Vienna's Phaidra repository – mnoho stovek veřejně dostupných Horizon 2020 DMP: <https://hdl.handle.net/11353/10.1140797>
- DMPonline: https://dmponline.dcc.ac.uk/public_plans
- Argos:
 - DMP: <https://argos.openaire.eu/explore-plans>
 - Datové sady: <https://argos.openaire.eu/explore>

University of Vienna's Phaidra Repository – DMP Collection Overview

Phaidra_link	document_title	document_description	project_id	project_acronym	project_title	cordis_project_link	project_start
https://phaidra.univie.ac.at/o:1139130	Data management plan	Data management plan providing a detailed outline of APPLICATE data management strategy, includi	727862	APPLICATE	Advanced Prediction in Polar regions and beyond: Modelling, observing system design and Linkages	https://cordis.europa.eu/project/id/727862	2016
https://phaidra.univie.ac.at/o:1139131	D7.1 Data Management Plan & Handbook	This deliverable describes internal quality assurance and communication procedures & will also include	780298	Made4You	Open and Inclusive Healthcare for Citizens Based on Digital Fabrication	https://cordis.europa.eu/project/id/780298	2018
https://phaidra.univie.ac.at/o:1139132	Data management plan	This deliverable comprises the project data management plan.	842018	GREEN-WIN	Green growth and win-win strategies for sustainable climate action	https://cordis.europa.eu/project/id/842018	2015
https://phaidra.univie.ac.at/o:1139133	Data Management Plan (DPM) V1	The DPM comprises the provision for making the project data findable, accessible, interoperable and i	769255	SAFEWAY	GIS-BASED INFRASTRUCTURE MANAGEMENT SYSTEM FOR OPTIMIZED RESPONSE TO EXTREME WEATHER	https://cordis.europa.eu/project/id/769255	2018
https://phaidra.univie.ac.at/o:1139134	Data Management Plan	Data Management Plan for the data generated by SlideWiki platform. This plan will include the suitable	688095	SlideWiki	Large-scale pilots for collaborative OpenCourseWare authoring, multiplatform delivery and Learning A	https://cordis.europa.eu/project/id/688095	2016
https://phaidra.univie.ac.at/o:1139135	Data Management Plan	Data Management Plan	723852	COROMA	Cognitively enhanced robot for flexible manufacturing of metal and composite parts	https://cordis.europa.eu/project/id/723852	2016
https://phaidra.univie.ac.at/o:1139136	Data management plan (DMP) based on data policies of I	A first version of the data management plan for CALIPSOplus is available	730872	CALIPSOplus	Convenient Access to Light Sources Open to Innovation, Science and to the World	https://cordis.europa.eu/project/id/730872	2017
https://phaidra.univie.ac.at/o:1139137	Data Management Plan	This Deliverable will describe how to cope with the Data Management Plan Pilot as described in Articl	688188	MURAB	MRI and Ultrasound Robotic Assisted Biopsy	https://cordis.europa.eu/project/id/688188	2016
https://phaidra.univie.ac.at/o:1139138	Data Management Plan	Data Management Plan will be submitted in M6 and checked and updated for M36 and as appropriate	688865	BREAKBEN	Breaking the Nonuniqueness Barrier in Electromagnetic Neuroimaging	https://cordis.europa.eu/project/id/688865	2016
https://phaidra.univie.ac.at/o:1139139	Data Management Plan	WP1 will produce a Data Management Plan in M6.	675451	CompBioMed	A Centre of Excellence in Computational Biomedicine	https://cordis.europa.eu/project/id/675451	2016
https://phaidra.univie.ac.at/o:1139140	ORDP: Data Management Plan	This Deliverable will describe how to cope with the Data Management Plan Pilot as described in Articl	688188	MURAB	MRI and Ultrasound Robotic Assisted Biopsy	https://cordis.europa.eu/project/id/688188	2016
https://phaidra.univie.ac.at/o:1139141	A data management plan for the Icelandic RIF station in c	A data management plan for the Icelandic RIF station in connection with ABDS for the selected focal e	730938	INTERACT	International Network for Terrestrial Research and Monitoring in the Arctic	https://cordis.europa.eu/project/id/730938	2016
https://phaidra.univie.ac.at/o:1139142	Data management plan	The data management plan describes how existing and newly generated data are processed, manage	733032	HBM4EU	European Human Biomonitoring Initiative	https://cordis.europa.eu/project/id/733032	2017
https://phaidra.univie.ac.at/o:1139143	First version of the Data Management Plan	Report on First version of the Data Management PlanThe progress of the implementation of Data Mar	642154	FISSAC	FOSTERING INDUSTRIAL SYMBIOSIS FOR A SUSTAINABLE RESOURCE INTENSIVE INDUSTRY	https://cordis.europa.eu/project/id/642154	2015
https://phaidra.univie.ac.at/o:1139144	Data Management plan - M12	Data Management plan - yearly updateThis is a document outlining how the research data collected o	654850	Residue2Heat	Renewable residential heating with fast pyrolysis bio-oil	https://cordis.europa.eu/project/id/654850	2016
https://phaidra.univie.ac.at/o:1139145	Data management plan	This report exposes the Open Data Management Plan and Open Research Data Pilot preparation and	731148	INVADE	Smart system of renewable energy storage based on INtegrated EVs and bAtteries to empower mob	https://cordis.europa.eu/project/id/731148	2017
https://phaidra.univie.ac.at/o:1139146	Data Management plan - M24	Data Management plan - yearly updateThis is a document outlining how the research data collected o	654850	Residue2Heat	Renewable residential heating with fast pyrolysis bio-oil	https://cordis.europa.eu/project/id/654850	2016
https://phaidra.univie.ac.at/o:1139147	Project Management, Quality and Risk Plan	Includes plans for project management(also a detailed data management plan), quality and risk mana	740698	MARISA	Maritime Integrated Surveillance Awareness	https://cordis.europa.eu/project/id/740698	2017
https://phaidra.univie.ac.at/o:1139148	Data management plan	Data Management plan, including publication policy and Intellectual Property Rights (IPR).	641762	ECOPOTENTIAL	ECOPOTENTIAL: IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATI	https://cordis.europa.eu/project/id/641762	2015
https://phaidra.univie.ac.at/o:1139149	INTERACT Data Management Plan	INTERACT Data Management Plan	730938	INTERACT	International Network for Terrestrial Research and Monitoring in the Arctic	https://cordis.europa.eu/project/id/730938	2016
https://phaidra.univie.ac.at/o:1139150	Data management plan	This deliverable will deliver the data management plan produced within DataBio, following the EC rec	732064	DataBio	Data-Driven Bioeconomy	https://cordis.europa.eu/project/id/732064	2017
https://phaidra.univie.ac.at/o:1139151	Data Management Plan (v1)	First version of the Data Management Plan	740712	COMPACT	Competitive Methods to protect local Public Administration from Cyber security Threats	https://cordis.europa.eu/project/id/740712	2017
https://phaidra.univie.ac.at/o:1139152	Data management plan	Data management plan	730403	INNOPATHS	Innovation pathways, strategies and policies for the Low-Carbon Transition in Europe	https://cordis.europa.eu/project/id/730403	2016
https://phaidra.univie.ac.at/o:1139153	Admin: Data management plan	Data management plan (3)	722346	EUROPAPH	The Extensive and Ubiquitous Role of Polycyclic Aromatic Hydrocarbons (PAHs) in Space	https://cordis.europa.eu/project/id/722346	2016
https://phaidra.univie.ac.at/o:1139154	Data Management Plan	"The Data Management Plan is developed as part of the "Clearing House" (Task 2.4). The purpose o	731289	InterFlex	Interactions between automated energy systems and Flexibilities brought by energy market players	https://cordis.europa.eu/project/id/731289	2017
https://phaidra.univie.ac.at/o:1139155	FAIR Data Management Plan	FAIR (findable, accessible, interoperable and reusable) Data Management Plan defining collection, tre	778465	RURITAGE	Rural regeneration through systemic heritage-led strategies	https://cordis.europa.eu/project/id/778465	2018
https://phaidra.univie.ac.at/o:1139156	Data Management plan	Data Management plan	766840	COSY-BIO	Control Engineering of Biological Systems for Reliable Synthetic Biology Applications	https://cordis.europa.eu/project/id/766840	2017
https://phaidra.univie.ac.at/o:1139157	First Updated Data Management Plan	The project data management plan will be developed during the first six months of the project and sho	730944	RINGO	Readiness of ICOS for Necessities of Integrated Global Observations	https://cordis.europa.eu/project/id/730944	2017
https://phaidra.univie.ac.at/o:1139158	Open Data Management Plan	Open Data Management Plan	770143	ReFreeDrive	Rare Earth Free e-Drives featuring low cost manufacturing	https://cordis.europa.eu/project/id/770143	2017
https://phaidra.univie.ac.at/o:1139159	Initial Data Management Plan (DMP).	The project data management plan will be developed during the first six months of the project and sho	730944	RINGO	Readiness of ICOS for Necessities of Integrated Global Observations	https://cordis.europa.eu/project/id/730944	2017
https://phaidra.univie.ac.at/o:1139160	Open Research Data Pilot Management Plan	Report on the Open Research Data Management Plan for WEARPLEX	825339	WEARPLEX	Wearable multiplexed biomedical electrodes	https://cordis.europa.eu/project/id/825339	2019
https://phaidra.univie.ac.at/o:1139161	Data Management Plan	Development (and regular updating) of a Data Management Plan (DMP) outlining the project's policy t	838335	Net4Society5	National Contact Points (NCPs) Network of Societal Challenge 6 'Europe in a changing world – Includ	https://cordis.europa.eu/project/id/838335	2019
https://phaidra.univie.ac.at/o:1139162	Data Management Plan (DMP)	Data Management Plan (DMP)	780839	MOLOKO	Multiplex pH/Otic sensor for pLAsmonic-based Online detection of contaminants in milk	https://cordis.europa.eu/project/id/780839	2018
https://phaidra.univie.ac.at/o:1139163	Data Management Plan	DB 6 : Data Management plan	736937	M-CUBE	MetaMaterials antenna for ultra-high field MRI	https://cordis.europa.eu/project/id/736937	2017
https://phaidra.univie.ac.at/o:1139164	Data Management Plan	Data Management Plan: Report on consortium Data Management Plan finished, detailing what data th	738899	MagnaPharm	Magnetic Control of Polymorphism in Pharmaceutical Compounds	https://cordis.europa.eu/project/id/738899	2017

<https://hdl.handle.net/11353/10.1159821>

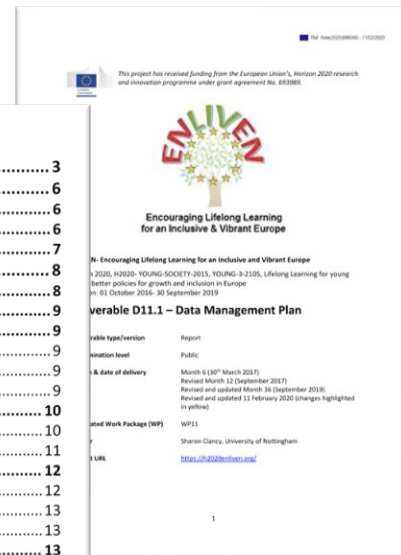
MUNICI
ICS

ENLIVEN ('Encouraging Lifelong Learning for an Inclusive and Vibrant Europe') Data Management Plan

- <https://hdl.handle.net/11353/10.1139743>
- Oblast **sociálních věd**.
 - Audio a video data, přepisy **rozhovorů**.
 - Omezení **přístupu** k datům.
- Využití **existujících dat** + sběr **vlastních**.
- **Průběžná aktualizace DMP**
 - Přehled změn v tabulce.
 - Migrace (některých) dat do UK.
 - Ukládání v MS O365 Teams (šifrovaně).
- **Etika** popsána v **samostatném dokumentu**, z DMP jen **odkázána**.
- Spolupráce více institucí → **rozdělení odpovědnosti**, explicitní **vymezení**, kdo je za co **odpovědný**.
 - Odkaz na využívání institucionální podpory pro DMP, pokud si nebudou jisti.
- **Pěkný popis ochrany dat**.
 - Informace o anonymizaci v příloze.
- Dokumentace a výstupní **publikace** (včetně webových stránek) **součástí DMP**.
- Popis hardware a software by mohl být **podrobnější**.
 - Ale popis zálohování je přítomen.
- **Vyjasněna autorská práva**, proces **QA**.

Table of Contents

1	Changes to Data Management Plan (D11.1) – September 2019.....	3
2	Introduction	6
2.1	Project Summary.....	6
2.2	Types of Data	6
2.3	Organisation of the ENLIVEN project.....	7
3	Data Management.....	8
3.1	Principles	8
3.2	Organisation	9
3.3	Types of Data	9
3.3.a	Aggregated secondary statistical data	9
3.3.b	Secondary data in form of anonymised micro data sets:	9
3.3.c	Primary data collection	9
3.4	Data Use and Protection.....	10
3.4.a	Secondary analysis of EUROSTAT and related scientific-use micro data files	10
3.4.b	Interview Transcription and Data Protection	11
3.5	Documentation	12
3.5.a	Documentation Reports.....	12
3.5.b	Project Management Documents	13
3.5.c	Website.....	13
3.6	Hardware and Software.....	13
3.6.a	Intelligent Decision Support System	13
3.6.b	Data Backup and Recovery	13
3.7	Intellectual Property and Ownership	13
3.7.a	Intellectual Property	13
3.7.b	Joint ownership.....	14
3.8	Open Access.....	14
3.9	Quality Assurance	14
4	Work-Package-specific Data Issues.....	15
4.1	WP1: Mapping European and national policies and programmes, and their contribution to economic and social inclusion	15
4.2	WP2: Constraints and facilitators of access and participation	15
4.3	WP3: The role of European governance in adult education & learning policy	15
4.4	WP4: Improving our understanding of the effect of system characteristics by building stronger data and adding a longitudinal, regional & sectoral focus	16
4.5	WPs5–7: Studying the role of workplace learning and patterns of work organisations for early career structuration; qualitative interviews on learning biographies	16
4.6	WP8: Knowledge discovery on evidence-based policy making in participating countries; & WP9: Establishment of Intelligent Decision Support System for evidence-based policy making	16
4.7	WP10-11: Dissemination and Project Management & Integration.....	17
5	Appendices.....	18
5.1	Table 1: Key elements of the framework to ensure anonymization within the ENLIVEN research process (for storage/use within the project): (to be refined within the research project)	18
5.2	Table 2: Processing of data in the qualitative research implemented by the ENLIVEN project.....	19



RECETOX MU DMP

- Generický dokument popisující RECETOX infrastrukturu jako celek.
 - Sazba jako LaTeXový dokument přes Overleaf
- DMP pro projekty generovány jako vybrané podčásti z generického dokumentu.

```
Source Rich Text
1 %\program=xe $\LaTeX$ 
2 %\encoding=UTF-8 Unicode
3
4 \newif\iftodos\todostrue
5 \newif\ifgenericDMPtemplate\genericDMPtemplatetrue
6 \newif\ifurbanx\urbanxfalse
7 \newif\ifurbanxfinal\urbanxfinalfalse
8 \newif\iferachair\erachairfalse
9 \newif\iferachairfinal\erachairfinalfalse
10 \newif\ifce\cefalse
11 \newif\ifcefinal\cefinalfalse
12
13 %%%
14 %%% vvv Set Options vvv
15 %%%
16
17 %
18 %% Typeset final version of CETOCOEN Excellence (a.k.a. CE) DMP, i.e. without generic sections, TODOS etc.?
19 %%
20 %\cefinaltrue % Uncomment to typeset final version of CETOCOEN Excellence (a.k.a. CE) DMP
21
22 %
23 %% Typeset CETOCOEN Excellence (a.k.a. CE) DMP?
24 %%
25 %\cetruetrue % Uncomment to typeset CETOCOEN Excellence (a.k.a. CE) DMP
26
27 %
28 %% Typeset final version of ERA-Chair DMP, i.e. without generic sections, TODOS etc.?
29 %%
30 %\erachairfinaltrue % Uncomment to typeset final version of ERA-Chair DMP
31
32 %
33 %% Typeset ERA-Chair DMP?
34 %%
35 %\erachairtrue % Uncomment to typeset ERA-Chair DMP
36
37 %
38 %% Typeset final version of URBAN_X DMP, i.e. without generic sections, TODOS etc.?
39 %%
40 %\urbanxfinaltrue % Uncomment to typeset final version of URBAN_X DMP
41
42 %
43 %% Typeset URBAN_X DMP?
44 %%
45 %\urbanxtrue % Uncomment to typeset URBAN_X DMP
46
47 %
48 %% Hide TODOS in the document?
49 %%
50 %\todosfalse % Uncomment to hide TODOS in the document
51
52 %%%
53 %%% ^^^ Set Options ^^^
54 %%%
55 --
```


MUNI
ICS

Nástroje a zdroje pro podporu FAIR principů

Nástroje a zdroje pro podporu FAIR principů

– The ELIXIR Research Data Management Kit (RDMkit)

- <https://rdmkit.elixir-europe.org/>
- Online průvodce dobrou praxí správy dat v přírodních vědách.
- Rozcestník na řadu externích zdrojů.

Tematické dělení dle fází životního cyklu dat,

role osoby v procesu práce s daty,

úlohy s daty (analýza, publikování, kontrola kvality, ...).

Včetně odkazů na národní zdroje, relevantní software a vzdělávací materiály.



Data management
Data life cycle
Your role
Your domain
Your tasks
Compliance monitoring
Data analysis
Data management plan
Data organisation
Data protection
Data publication
Data quality
Data storage
Data transfer
Documentation and metadata
Existing data
Identifiers
Licensing
Machine actionability
Sensitive data
Tool assembly
National resources
All tools and resources
All training resources

Nástroje a zdroje pro podporu FAIR principů (2)

– UK Data Service – Research data management

- <https://www.ukdataservice.ac.uk/manage-data.aspx>
- Stručný průvodce s odkazy na externí zdroje k různým úlohám při správě výzkumných dat.
- Obsahuje např.
 - checklist pro DMP,
 - odkazy na legislativu relevantní pro správu výzkumných dat,
 - doporučení k určování nákladů na správu dat apod.



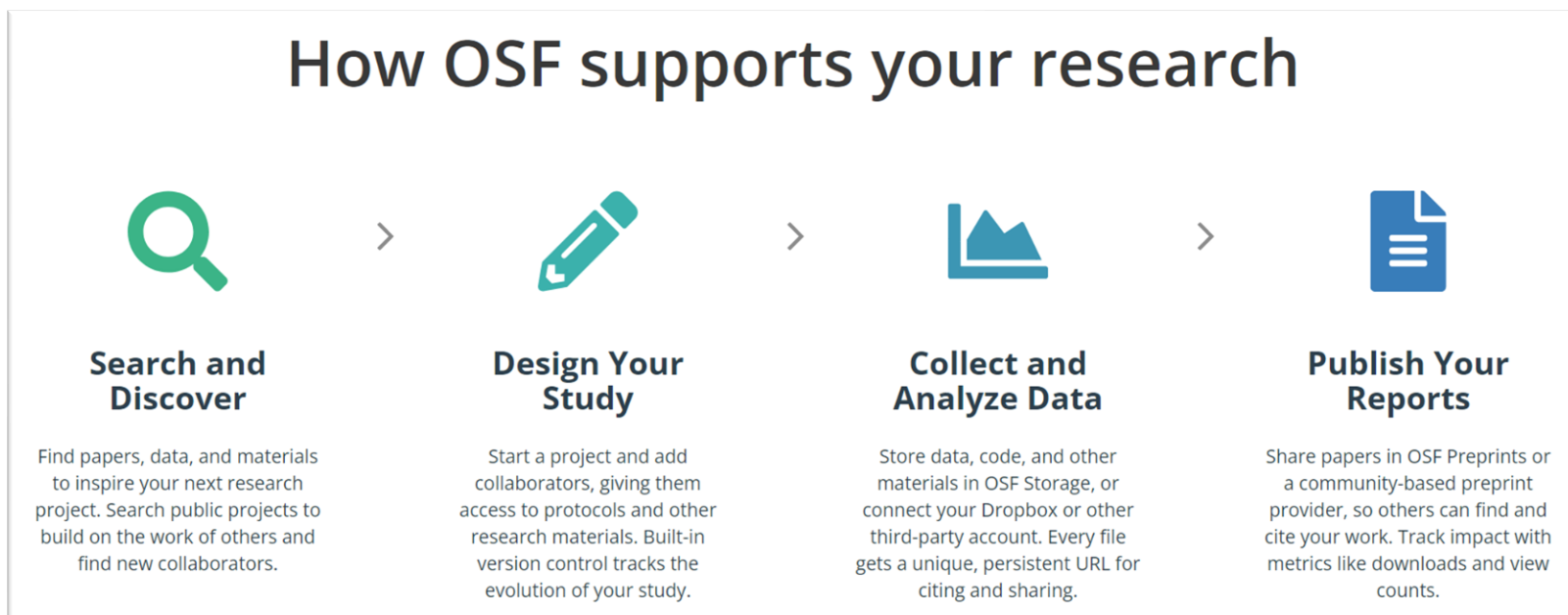
Rights in data
Document your data
Format your data
Store your data
Collaborative research
Data lifecycle
Plan to share
Data Protection
Ethical issues
Anonymisation

Nástroje a zdroje pro podporu FAIR principů (3)

– OSF

– <https://osf.io/>

– Platforma pro podporu workflow vědeckého výzkumu.

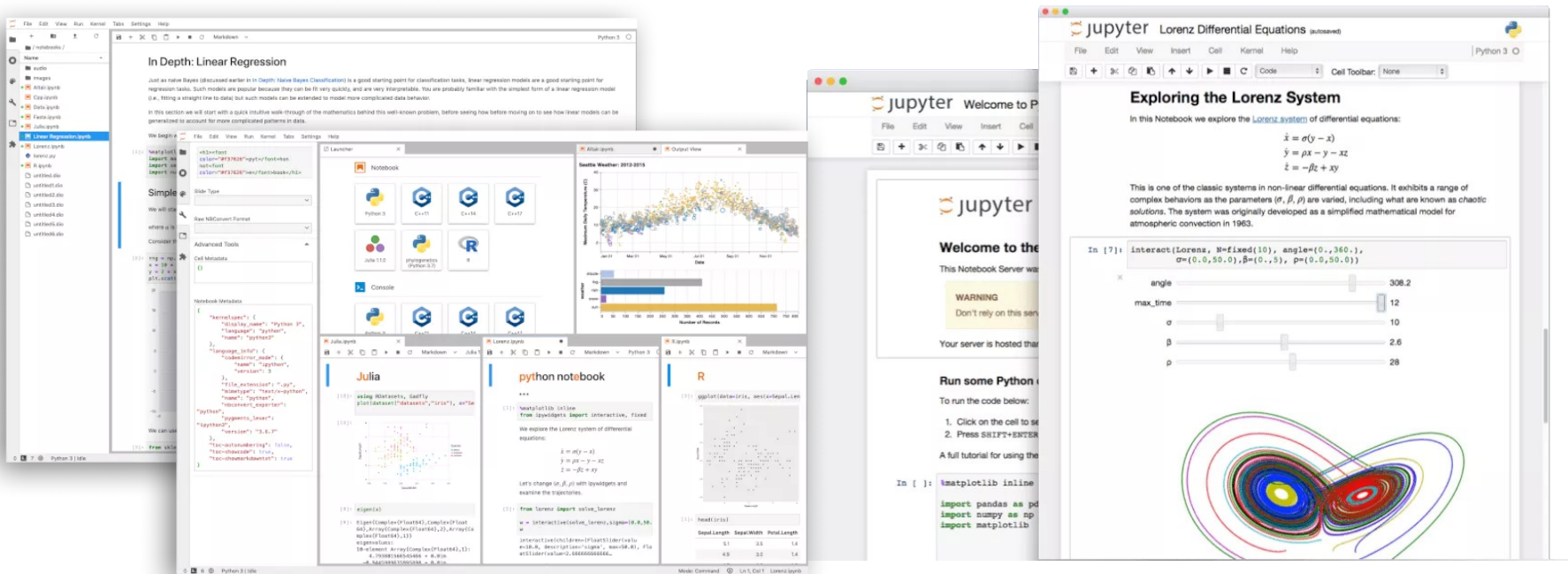


Nástroje a zdroje pro podporu FAIR principů (4)

– JupyterLab

– <https://jupyter.org/>

– Dokumentované zpracování dat umožňující sdílení.

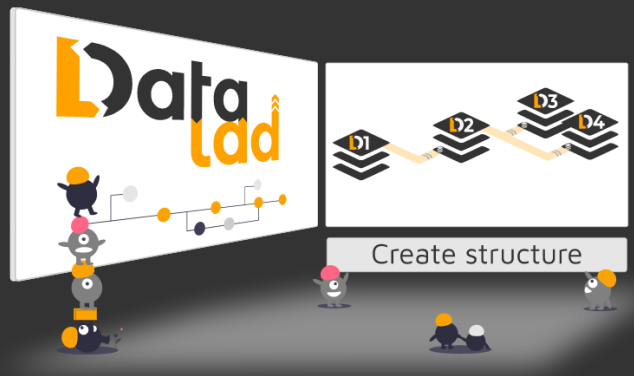


Nástroje a zdroje pro podporu FAIR principů (5)

– DataLad

– <https://www.datalad.org/>

– *DataLad is a free and open-source distributed **data management** system that **keeps track** of your data, creates structure, **ensures reproducibility**, supports collaboration, and integrates with widely used data infrastructure.*



The image shows a promotional graphic for DataLad. On the left, the DataLad logo is displayed in a stylized font. Below it, several small, colorful cartoon characters are arranged in a line. In the center, a diagram illustrates the 'Create structure' process, showing four data blocks labeled D1, D2, D3, and D4, each represented by a stack of three horizontal lines. A hand icon is shown pointing to each block. To the right of the diagram, the text 'distributed data management' is written in white. Below that, 'free and open source' is written in white. At the bottom right, there is a yellow button with a rocket icon and the text 'Get DataLad'. Below the button, there is a 'Star' icon and the number '398'.

distributed data management

free and open source

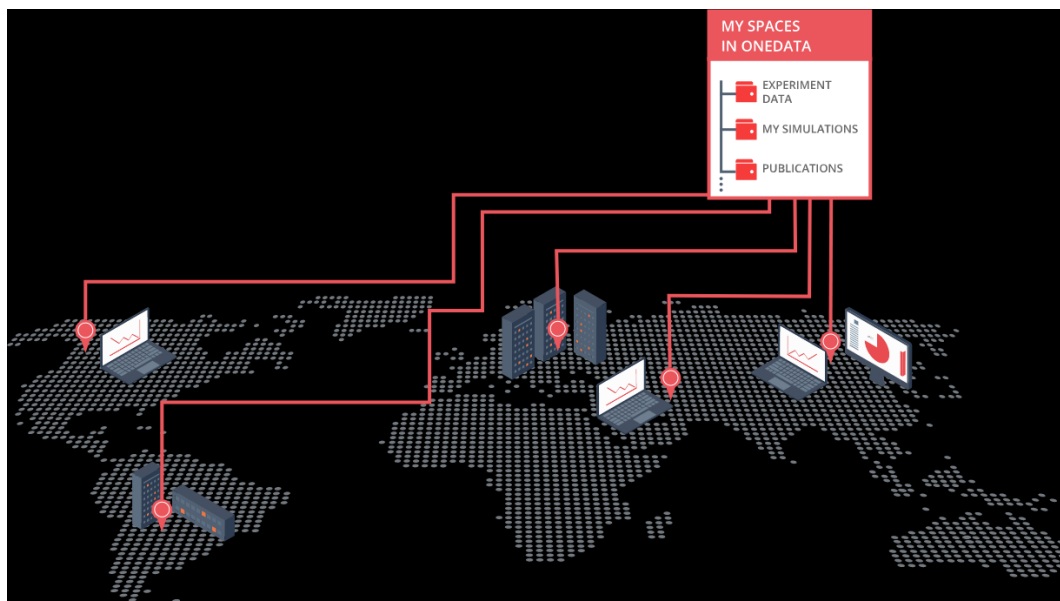
Get DataLad

Star 398

Nástroje a zdroje pro podporu FAIR principů (6)

– Onedata

- <https://onedata.org/>
- Perform *heavy computations on huge datasets*. Access your data in a dropbox-like fashion regardless of its location. Publish and *share* your results with public or closed communities.



Úložiště pro výzkumná data v instituci a české eInfrastruktuře

Příklad Masarykova univerzita, CESNET; služby Mendelu

Kategorizace úložišť na MU a doporučení pro jejich využívání

- Datová úložiště na MU: <https://it.muni.cz/kategorie/datova-uloziste>
- Doporučení pro užívání úložišť: <https://it.muni.cz/prehledy/doporuzeni-pro-uzivani-ulozist>
 - Kategorizace dat.
 - Kategorizace úložišť.
 - Vhodnost úložišť pro různé typy dat.

Úložiště na MU – kategorizace dat

KATEGORIE	POPIS	PŘÍKLADY
Veřejná data	<ul style="list-style-type: none">• Data zpřístupnitelná komukoliv bez jakýchkoliv omezení, např. veřejně vystavená na internetu.• Jejich zveřejnění nepředstavuje žádné ohrožení pro MUNI nebo jiné instituce či osoby.	<ul style="list-style-type: none">• Prezentace z veřejných přednášek;• veřejně přístupné výzkumné zprávy;• open-source software;• veřejná výzkumná data;• propagace, veřejné informace o službách.
Interní data	<ul style="list-style-type: none">• Data určená jen pro vnitřní potřebu obecně definované skupiny osob (např. spolupracovníci projektu, pracovníci instituce apod.).• Nevyžadují však zvláštní regulaci nebo ochranu (ze zákona, dle smlouvy apod.).• Zpřístupnění mimo danou skupinu nezpůsobí přímou škodu (finanční, morální, právní apod.).	<ul style="list-style-type: none">• Interní korespondence;• zápisy z jednání;• vnitřní regulace a předpisy;• vnitřní plány práce, poznámky apod.;• nedokončené/nepublikované výzkumné zprávy.

Úložiště na MU – kategorizace dat (2)

KATEGORIE	POPIS	PŘÍKLADY
Diskrétní data	<ul style="list-style-type: none">• Data určená výhradně pro vnitřní potřebu přesně definované skupiny osob (např. zaměstnanec a jeho přímý nadřízený, pracovník HR oddělení a uchazeč o zaměstnání, skupina správců IT systému s administrátorskými právy k němu).• Vyžadují ze své povahy regulaci nebo ochranu, typicky jsou data chráněná ze zákona nebo na základě nějaké smlouvy/licence (jedná se např. o osobní údaje osob, data spadající pod obchodní tajemství apod.).• Zpřístupnění mimo danou skupinu osob velmi pravděpodobně způsobí škodu (finanční, morální, právní apod.).	<ul style="list-style-type: none">• Ekonomické a personální údaje osobní povahy;• osobní údaje studentů / zaměstnanců / spolupracovníků...;• čísla identifikačních průkazů, rodná čísla apod.;• čísla kreditních karet;• cenná výzkumná data (poskytující např. konkurenční výhodu) nebo data obsahující jinak citlivé informace;• rozsáhlé kolekce interních dat;• přístupové údaje (např. hesla či šifrovací klíče) k málo významným systémům a interním datům.

Úložiště na MU – kategorizace dat (3)

KATEGORIE	POPIS	PŘÍKLADY
Citlivá data	<ul style="list-style-type: none">• Data určená striktně jen pro vnitřní potřebu přesně definované skupiny osob (např. zdravotník a jeho pacient, řešitelé projektu s bezpečnostním prověřením určité úrovně apod.).• Vyžadují ze své povahy zvláštní regulaci nebo obzvláštní ochranu, typicky jsou data přísně chráněná ze zákona nebo na základě smlouvy/licence (jedná se např. o velmi cenná data spadající pod obchodní tajemství, citlivé osobní údaje apod.).• Zpřístupnění mimo danou skupinu oprávněných osob velmi pravděpodobně způsobí škodu (finanční, morální, právní apod.) velkého rozsahu se závažnými/nevratnými následky.• <i>V praxi bude do této kategorie spadat jen málo dat, většina bude spadat nejvýše do kategorie diskrétních dat.</i>	<ul style="list-style-type: none">• Zdravotní data, citlivé osobní údaje;• velmi cenná výzkumná data (poskytující např. unikátní a těžko opakovatelnou konkurenční výhodu) nebo výzkumná data obsahující vysoce důvěrné údaje;• rozsáhlé kolekce diskrétních dat;• přístupové údaje (např. hesla či šifrovací klíče) k důležitým systémům a datům kategorie diskrétní nebo citlivá.

Úložiště na MU – kategorizace úložišť

TYP ÚLOŽIŠTĚ	POPIS
Přenosná média	Např. flash disky, paměťové karty, externí HDD/SSD, CD, DVD, ... Tj. externí paměťová média, která nejsou pevnou součástí žádného zařízení a uživatelé je používají k přenášení informací mezi zařízeními nebo pro off-line uložení dat.
Lokální úložiště	V počítačích Datová úložiště pevně zabudovaná ve stolních počítačích/noteboocích (typicky interní HDD/SSD apod.) v kancelářích zaměstnanců, ve studovnách apod. V mobilních zařízeních Datová úložiště pevně zabudovaná v mobilních zařízeních, tj. mobilních telefonech, tabletech apod. (typicky interní nevyjímatelná paměť, v zařízení instalovaná paměťová karta apod.) v použití zaměstnanců/studentů.
Síťová a cloudová úložiště ÚVT	<u>Datová úložiště provozovaná ÚVT</u> a zpřístupněná koncovým uživatelům přes datovou síť – tzv. <u>standardní</u> a <u>střední</u> síťová úložiště. Do této kategorie spadají také <u>datová úložiště CERIT-SC</u> pro velkoobjemová výzkumná data.
Úložiště IS MUNI	Dokumentový server, Úschovna a podobné úložné kapacity integrované v systému <u>IS MUNI</u> .
Úložiště CESNET	Datová úložiště provozována <u>Oddělením datových úložišť sdružení CESNET</u> . Do této kategorie spadají i služby, které tato úložiště využívají pro fyzické uložení dat, např. <u>CESNET OwnCloud</u> , <u>CESNET FileSender</u> apod.

Úložiště na MU – kategorizace úložišť (2)

TYP ÚLOŽIŠTĚ	POPIS
Externí úložiště <i>Datová úložiště provozovaná externími subjekty, tj. mimo MUNI a CESNET.</i>	Se smlouvou s MUNI
	MUNI Microsoft O365 Cloudová datová úložiště poskytovaná v rámci služby Microsoft Office 365 pro Masarykovu univerzitu . Zejména se jedná o osobní úložiště OneDrive a dokumentové knihovny služby SharePoint a Skupin O365. Patří sem ale také další data uložená v MUNI O365 cloudu, jako např. elektronická pošta v MUNI O365 Outlook, soubory sdílené v sociální síti MUNI Yammer apod.
	MUNI Google Workspace for Education Cloudová datová úložiště poskytovaná v rámci služby Google Workspace for Education pro Masarykovu univerzitu . Zejména se jedná o datové kapacity MUNI Disk Google, patří sem ale i další data uložená v MUNI Workspace for Education cloudu, např. elektronická pošta v MUNI Google Mail, poznámky v MUNI Google Keep, kalendářová data v MUNI Google Calendar apod.
	Bez smlouvy s MUNI
	Veřejná Google/Microsoft/Dropbox/... úložiště Do této kategorie spadají zejména veřejné cloudové služby (zřízené typicky zdarma soukromým koncovým uživatelem jen proti elektronické registraci přes web) jako Disk Google, Microsoft OneDrive, Dropbox, Amazon úložiště, repositáře na GitHub apod. Zásadním rozdílem a „poznávacím znamením“ této kategorie cloudových úložišť oproti cloudovým službám uvedeným výše je, že MUNI nemá žádný (právní) vztah s provozovateli těchto externích služeb, a proto není schopna garantovat jakékoliv záruky ohledně bezpečnosti/důvěrnosti uložených dat nebo politiky nakládání s nimi.

Úložiště na MU – Vhodnost úložišť pro různé typy dat

TYP ÚLOŽIŠTĚ	POUŽITÍ			
	ZELENÁ: VEŘEJNÁ DATA	MODRÁ: INTERNÍ DATA	ORANŽOVÁ: DISKRÉTNÍ DATA	ČERVENÁ: CITLIVÁ DATA
PŘENOSNÁ MÉDIA (FLASH DISKY, EXTERNÍ HDD, CD, DVD, ...)	Vhodné	Možné doporučeno použití šifrování	Nevhodné možné při použití šifrování	Nevhodné
LOKÁLNÍ ÚLOŽIŠTĚ				
V POČÍTAČÍCH (STOLNÍ, NOTEBOOKY)	Vhodné	Vhodné	Vhodné doporučeno použití šifrování	Nevhodné možné v dobře odůvodněných případech, při provedení individuální analýzy, použití šifrování a aplikaci dalších bezpečnostních opatření, která z analýzy vyplynou
V MOBILNÍCH ZAŘÍZENÍCH (MOBILNÍ TELEFONY, TABLETY, ...)	Vhodné	Vhodné nutný zámek obrazovky (vzor, čtečka otisků prstů, PIN, heslo)	Možné nutné použití šifrování nutný silný zámek obrazovky (čtečka otisků prstů, PIN, heslo)	Nevhodné možné v dobře odůvodněných případech, při provedení individuální analýzy, použití šifrování a aplikaci dalších bezpečnostních opatření, která z analýzy vyplynou
SÍŤOVÁ A CLOUDOVÁ ÚLOŽIŠTĚ ÚVT (TZV. STANDARDNÍ A STŘEDNÍ SÍŤOVÉ ÚLOŽITĚ, VIZ KATALOG IT, ÚLOŽIŠTĚ CERIT-SC)	Vhodné	Vhodné	Vhodné	Vhodné, doporučeno provedení individuální analýzy, použití šifrování a aplikaci dalších bezpečnostních opatření, která z analýzy vyplynou

Úložiště na MU – Vhodnost úložišť pro různé typy dat (2)

TYP ÚLOŽIŠTĚ	POUŽITÍ			
	ZELENÁ: VEŘEJNÁ DATA	MODRÁ: INTERNÍ DATA	ORANŽOVÁ: DISKRÉTNÍ DATA	ČERVENÁ: CITLIVÁ DATA
ÚLOŽIŠTĚ IS MUNI (NAPŘ. DOKUMENTOVÝ SERVER, ÚSCHOVNA APOD.)	Vhodné	Vhodné	Vhodné	Vhodné, doporučeno provedení individuální analýzy, použití šifrování a aplikaci dalších bezpečnostních opatření, která z analýzy vyplynou
ÚLOŽIŠTĚ CESNET (NAPŘ. CESNET ARCHIVNÍ ÚLOŽIŠTĚ, OWNCLOUD, FILESENDER, ..., VIZ ODDĚLENÍ DATOVÝCH ÚLOŽIŠŤ CESNET)	Vhodné	Vhodné	Vhodné	Vhodné, doporučeno provedení individuální analýzy, použití šifrování a aplikaci dalších bezpečnostních opatření, která z analýzy vyplynou
EXTERNÍ ÚLOŽIŠTĚ				
SE SMLOUVOU S MUNI				
MUNI MICROSOFT O365 (MUNI O365 ONEDRIVE, SHAREPOINT, ..., VIZ MUNI O365)	Vhodné	Vhodné	Vhodné doporučeno použití šifrování	Možné výhradně s adekvátním procesním pokrytí dané situace na základě individuální analýzy a aplikaci bezpečnostních opatření, která z analýzy vyplynou
MUNI GOOGLE G SUITE FOR EDUCATION (VIZ MUNI GOOGLE APPS)	Vhodné	Vhodné	Nevhodné možné při použití šifrování	Nevhodné
GRAMMARLY	Vhodné	Vhodné	Nevhodné	Nevhodné
BEZ SMLOUVY S MUNI				
VEŘEJNÁ GOOGLE, MICROSOFT, DROPBOX, ... ÚLOŽIŠTĚ	Vhodné	Nevhodné	Nevhodné	Nevhodné

SensitiveCloud – infrastruktura pro práci s citlivými daty

- Budováno CERIT-SC.

- <https://www.cerit-sc.cz/infrastructure-services/sensitivecloud>

- Dvě komponenty:

- SensitiveCloud **Computer**.

- Poskytovaný jako *Platform as a Service (PaaS)*, postaveno na **Kubernetes**.

- Uživatel spravuje **jen aplikaci**, nikoliv celý virtuální stroj.

- Ideálně *Software as a Service (SaaS)* – R Studio, Jupyter Notebook apod.

- SensitiveCloud **Storage**.

- Úložiště pro zpracovávaná **data**.

- Integrace v rámci compute.

- Oddělená síť, WireGuard VPN, Perun AAI s vícefaktorovým přihlášením, fyzické zabezpečení hardware, ...

- **Formalizované** postupy správy, neustálé **vylepšování**, budováno dle požadavků **ISO 27k**.

Oddělení datových úložišť CESNET

- <https://du.cesnet.cz/>
- Mohou být využívána pro účely výzkumu, a to akademickými pracovníky a studenty výzkumných institucí v ČR.
- Datová úložiště jsou certifikována podle normy pro systém managementu bezpečnosti informací ČSN EN ISO/IEC 27001:2014.
- Případy použití:
 - Datový prostor pro zálohování.
 - Datový prostor pro archivaci.
 - Datový prostor pro sdílení dat.

Datová úložiště CESNET – FileSender

- Webová služba pro **jednorázové předávání** (dočasné uschování) **velkých** souborů, které poskytuje sdružení CESNET.
- Uživatel **přes webové rozhraní nahraje soubor** a připojí e-mail adresáta nebo adresátů, případně zprávu.
- Adresáti **dostanou e-mailem odkaz** ke stažení souboru.
- Pro nahrání souboru je třeba se **přihlásit přes federaci identit eduID.cz**, služba je tedy dostupná **všem zaměstnancům a studentům MU a také většiny akademických institucí v ČR**.
 - Pokud potřebujete dostat velký soubor od někoho, kdo **identitu** ve federaci **nemá**, je možné ve FileSenderu po přihlášení **vytvořit pozvánku**. Druhá strana komunikace pak obdrží webový odkaz, na který může nahrát soubor. Mailem pak dostanete odkaz, kde si tento soubor můžete stáhnout.
- Soubor je v systému **uložen maximálně měsíc**, poté je automaticky smazán. Lze nastavit i kratší dobu.
- Velikost jednoho souboru může být **až 500 GB**.

Datová úložiště CESNET – OwnCloud

- Nástroj pro **synchronizaci souborů** a podporuje synchronizaci dat na **platformách Windows, Mac a Linux** a na mobilních **zařízeních s Android a iOS**.
 - Obdoba Google Drive, Microsoft OneDrive apod., ale provozovaná na CESNET infrastruktuře s vyšší standardně dostupnou kapacitou pro „neplatící“ uživatele.
- **Přihlášení přes federaci identit eduID.cz**.
 - Pro vytvoření účtu postačuje první přihlášení do systému.
- Standardně poskytuje **100 GB kapacity na uživatele**.
- Umožňuje sdílení dat ve skupinách, které si uživatel může nastavit
- Dovoluje soubory **sdílet webovým odkazem komukoli**, komu odkaz sdělíte
- **Data zůstávají v ČR**, sdružení CESNET nezískává žádná práva na uložený obsah, ten stále náleží vlastníkům dat.

Webové a cloudové služby Mendelovy univerzity

– <https://tech.mendelu.cz/26227-webove-a-cloudove-sluzby>

The screenshot shows the Mendel University Helpdesk interface. At the top left, the logo for Mendelova univerzita v Brně is displayed. The main header includes the text 'Helpdesk MENDELU' and a search bar with the placeholder 'hledaný výraz...'. Below the header is a navigation menu with tabs for 'Aktuální informace', 'Časté dotazy', 'Návody a postupy', 'Helpdesk', and 'Kontakty'. The main content area features a large aerial photograph of the university campus. Below the photo, the title 'Webové a Cloudové služby' is prominently displayed, followed by a breadcrumb trail: 'Nacházíte se zde: Návody a postupy > Webové a Cloudové služby'. On the left side, there is a vertical menu with categories: 'UIS', 'SAP', 'E-MAIL A SLUŽBY O365', 'SÍTĚ', 'WEBOVÉ A CLOUDOVÉ SLUŽBY', and 'DALŠÍ SYSTÉMY'. The 'WEBOVÉ A CLOUDOVÉ SLUŽBY' category is expanded, listing various services such as 'Osobní stránky uživatelů', 'Stránky pracovišť a projektů', 'Připojení projektových stránek', 'Připojení k serveru Disk', 'Emersion CMS', 'ownCloud - datové úložiště', 'FileSender - úschovna', 'OneDrive - cloudové úložiště O365', and 'Online schůzky a semináře'. On the right side, there is a green box with a white icon of a globe and clouds, labeled 'Web & Cloud služby'.

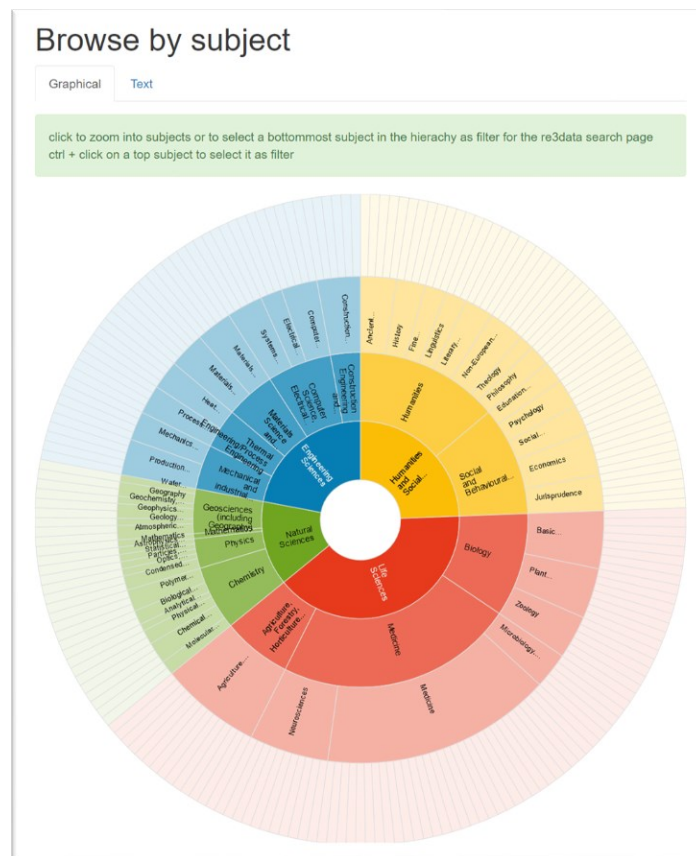
Repozitáře pro výzkumná data

Typy repozitářů

- **Oborové, národní, institucionální, catch-the-all, ...**
- Běžný postup práce s výzkumnými daty je **dělení po oborech, nikoliv po státech či institucích**.
 - Nezohledňujeme národní/institucionální/... specifika, ale oborová specifika.
- Pokud řešíme projektový/institucionální/národní **standard**, musíme být **kompatibilní s mezinárodními oborovými zvyklostmi**.
 - Oborové standardy řeší např. EOSC Task Forces: <https://www.eosc.eu/task-force-faq>
- Pokud chceme shromažďovat **informace napříč obory**, musí to být **formát dostatečně obecný**, který možná přesáhne rámec výzkumných dat.
 - Např. aplikační profil formátu DCAT ([DCAT-AP](#)), který se používá pro popis otevřených dat např. ze státní správy.

Jak najít/vybrat datový repozitář

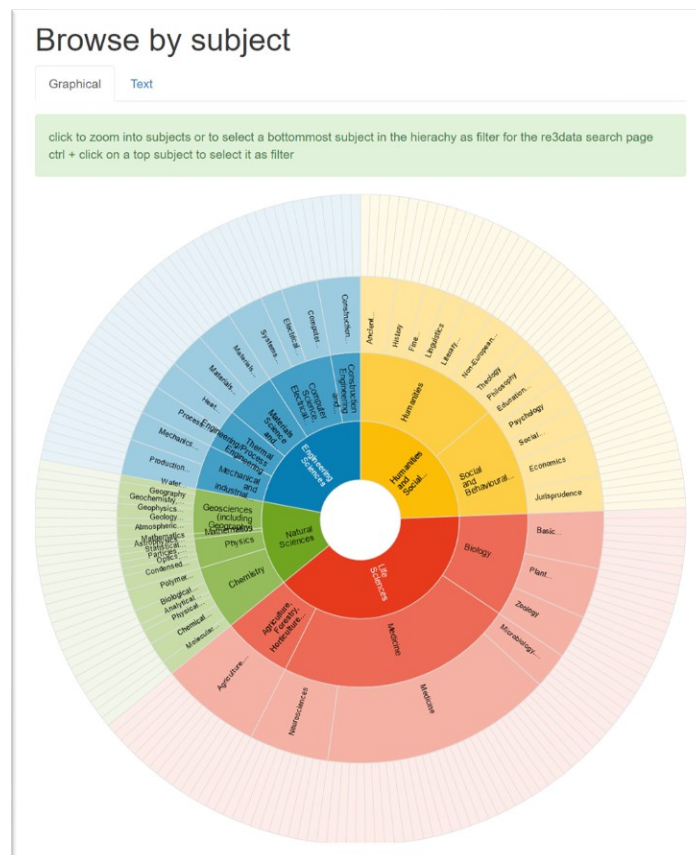
- OpenAIRE: [Jak najít důvěryhodný repozitář pro vaše data](#)
 - Preferovány jsou důvěryhodné certifikované repozitáře.
 - [CoreTrustSeal](#) (seznam certifikovaná repozitářů).
 - [Nestor Seal](#) (verifikace dle DIN 31644).
 - [ISO 16363](#).
 - Např. ale známé [Zenodo](#) žádnou certifikaci nemá...
- Mezi nejpoužívanější obecné repozitáře patří
 - [Zenodo](#),
 - [Figshare](#) nebo
 - [Dryad](#).



Zdroj: <https://www.re3data.org/browse/by-subject/>

Jak najít/vybrat datový repozitář (2)

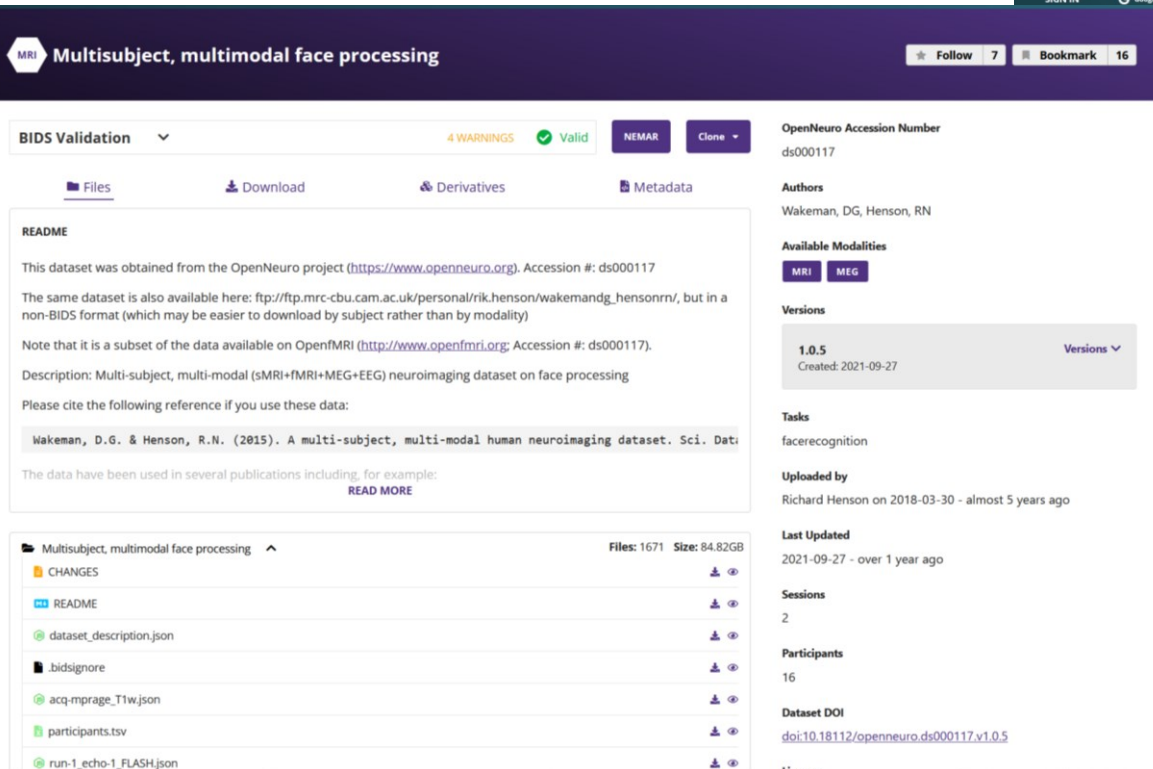
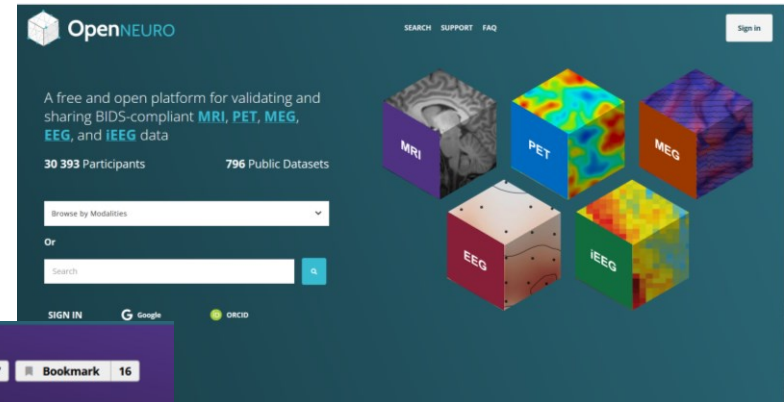
- Adresáře repozitářů:
 - Open Access repozitáře: [OpenDOAR](#)
 - Datové repozitáře: [re3data.org](#)
- [ELIXIR RDMkit | Data publication](#)
 - [EMBL-EBI's data submission wizard](#).
 - [ELIXIR Deposition Databases](#).
 - [Scientific Data journal's recommended repositories](#).



Zdroj: <https://www.re3data.org/browse/by-subject/>

Oborový datový repozitář prakticky

- OpenNeuro
(<https://openneuro.org/>)
 - Repozitář pro neurovizuální data.



- Viz *Workshop pro neurovědní uživatele MAFIL CEITEC MU* (22. 2. 2023):

- Slajdy: <https://openscience.muni.cz/openscience-na-mu/realizovane-akceprednasky/workshop-pro-neurovedni-uzivatele-cf-mafil-2023-02-22>
- Videozáznam na YouTube [Open Science kanálu MU](https://youtu.be/j6ifh88D3h0): <https://youtu.be/j6ifh88D3h0>

Oborový datový repozitář prakticky (2)

– Upload do repozitáře

Upload Dataset

Step 1: Select Files | Step 2: Validation | **Step 3: Metadata** | Step 4: Accept Terms

To protect the privacy of the individuals who have been scanned, we require that all scan data be defaced before publishing a dataset.

Select a [BIDS dataset](#) to upload

Select folder

close

Upload Dataset

Step 1: Select Files | Step 2: Validation | **Step 3: Metadata** | Step 4: Accept Terms

Incomplete fields in this form will make it more difficult for users to search for your dataset.

We recommend completing the applicable fields to improve your search results.

DOI of papers from the source data lab

10.1002/hbm.25698

Species

Human

Study Type

Longitudinal

Domain Studied

Brain

Number of Trials (if applicable)

1

Study Design

Block visual-motor tasks with repetitions over various acquisition parameters

Papers published from this dataset

DX status(es)

Healthy / Control

Grant Funder Name

Ministry of Education, Youth, and Sports of the Czech Republic

Grant Identifier

LM2018129

Oborový datový repozitář prakticky (3)

– Ověření záznamu DOI

The screenshot shows the OpenNeuro interface for a dataset. At the top, there's a navigation bar with 'SEARCH', 'SUPPORT', 'FAQ', 'UPLOAD', and 'My Account'. The dataset title is 'Multi-echo simultaneous multislice fMRI dataset: Effect of acquisition parameters on fMRI data'. A yellow banner states: 'This dataset has not been published! Before it can be published, please [create a version](#)'. Below this, there's a 'BIDS Validation' section showing 'Valid' with a green checkmark and a 'Clone' button. A menu includes 'Files', 'Publish', 'Share', 'Versioning', 'Download', 'Metadata', and 'Delete'. The 'README' section contains a detailed description of the dataset's purpose, acquisition parameters, and location. On the right, there's an 'OpenNeuro Accession Number' (ds004499), 'Authors' (Mikl Michal, Kovarova Anezka, Gajdos Martin, Marecek Radek, Novakova Marie, Slavicek Tomas), 'Available Modalities' (MRI), 'Version' (Draft, Updated: 2023-02-17, with a 'Create Version' button), 'Tasks' (FCN1, FCN2, FCN3, FCN4, FCN5, FCN6, FCN7), 'Uploaded by' (Michal Mikl on 2023-02-16 - about 24 hours ago), 'Sessions' (1), and 'Participants' (50). At the bottom, a file list shows 'dataset_description.json'.

Datový repozitář CESNET

- Pilotní provoz: <https://data.narodni-repozitar.cz/>
- Návody: https://du.cesnet.cz/cs/navody/narodni_repozitar/start
- Pro ukládání a řízené zveřejňování datových sad.
 - Vytvoření metadatového záznamu, který popisuje datovou sadu,
 - připojení jednotlivých souborů, které tvoří vlastní datovou sadu.
 - Každý veřejný záznam může být na požádání opatřen persistentním identifikátorem DOI.
 - Možnost vazby na publikace založená na dané datové sadě přes DOI publikace v metadatech.
- Pro nahrávání dat nutná registrace.
 - Je schvalovaná, není okamžitý přístup.
- Možnost odložené zveřejnění dat.
 - Časové embargo.
 - Metadata jsou viditelná okamžitě pro dohledatelnost.

Na co dát pozor při výběru repozitáře

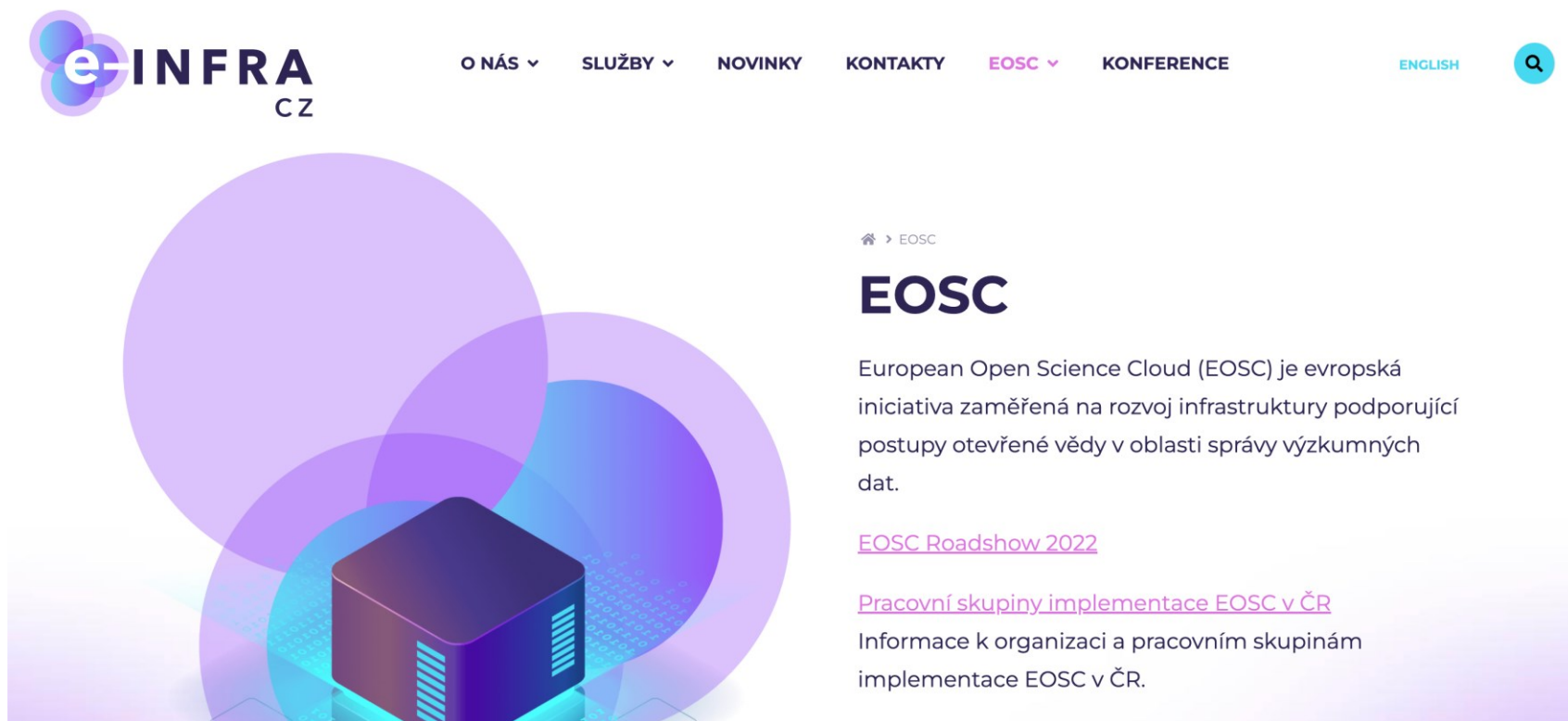
- Přidělí repozitář vašim datům **perzistentní identifikátor** (např. DOI)?
Díky perzistentním identifikátorům jsou vaše data snáze vyhledatelná a citovatelná.
- Je repozitář **důvěryhodný/certifikovaný**?
U certifikovaných repozitářů máte větší jistotu, že je o vaše data dobře postaráno.
- Poskytuje vybraný repozitář **otevřený přístup** k uloženým datům?
Pokud chcete svá data sdílet otevřeně, pak je toto klíčová informace.
- Opatří repozitář vaše data **licencí**? Uvádí jasné podmínky, za jakých lze data uložená v repozitáři využívat?
Je důležité, aby uživatelé vašich dat věděli, jak s nimi mohou nakládat.
- Poskytne repozitář pro vaše data **úvodní stránku s metadaty** (tzv. landing page)?
Metadata pomohou ostatním data najít, zjistit, co je to za data, a také jak je citovat.
- Umožňuje repozitář **verzování**?
Pokud svůj dataset aktualizujete, můžete aktualizovanou verzi nahrát jako novou verzi k původnímu datasetu. Nový dataset dostane přidělený svůj vlastní perzistentní identifikátor a uživatelé snadno zjistí, jaká je nejnovější verze nebo jaká verze byla použita v původní studii.

MUNI
ICS

EOSC

Implementace European Open Science Cloud (EOSC) v ČR

<https://www.e-infra.cz/eosc/>



e-INFRA
CZ

O NÁS ▾ SLUŽBY ▾ NOVINKY KONTAKTY **EOSC ▾** KONFERENCE

ENGLISH 🔍

🏠 > EOSC

EOSC

European Open Science Cloud (EOSC) je evropská iniciativa zaměřená na rozvoj infrastruktury podporující postupy otevřené vědy v oblasti správy výzkumných dat.

[EOSC Roadshow 2022](#)

[Pracovní skupiny implementace EOSC v ČR](#)

Informace k organizaci a pracovním skupinám implementace EOSC v ČR.

EOSC v ČR – očekávané služby Národní datové infrastruktury (NDI)

- Služby očekávané od NDI vychází z výstupů pracovní skupiny Základní služby.
 - Pracovní skupiny pro implementaci EOSC v ČR jsou stále otevřené novým členům, viz Pracovní skupiny implementace EOSC v ČR.
- Navrhované služby se dají rozdělit do několika základních kategorií:
 - Důvěryhodné repozitáře
 - Služby pro podporu práce s persistentními identifikátory (PID)
 - Služby pro řízení přístupu k datům, autentizaci a autorizaci uživatelů (AAI) a podporu jejich spolupráce
 - Služby pro podporu a plánování správy výzkumných dat
 - Machine actionability
 - Služby pro FAIRifikaci výzkumných dat
 - Služby transferu dat
 - Kvalita služeb
 - Kompatibilita s existujícími platformami
 - Statistiky a přehledy
 - Národní příspěvek k budování EOSC

Důvěryhodné repozitáře

- Národní repozitářová platforma (NRP) nad NDI.
 - Součástí i Národní metadatový adresář (NMA).
- Místo, kam můžete (i velká) data nejen uložit, ale také
 - spolehnout se, že tam budou v následujících letech dostupná a
 - budete moci s daty pracovat.
- Určeno i pro „živá“ data, tj. podpora aktivní práce s daty, ne jen uložení archivních dat.
- Poskytování služeb dle vhodných standardů/certifikací pro provoz důvěryhodných repozitářů.
 - Důvěryhodné repozitáře dle definic EU, EOSC, grantových agentur apod.
 - Pravděpodobně CoreTrustSeal.
 - Rodina standardů ISO 27 tisíc apod.?
- Nestaráte se o vlastní infrastrukturu, zabezpečená úložiště apod., ale
 - využíváte služby.
 - NDI umožní postavení vlastních instancí repozitářů (oborových, institucionálních, ...).

Precizní kontrola vlastníka dat nad řízením přístupu k datům i metadatům

- Data uložená v NDI jsou pod plnou kontrolou jejich vlastníků.
 - Zpřístupnění dalším osobám se děje na pokyn jejich vlastníka.
- Už samotná informace o existenci dat či metadata může být citlivá informace.
 - Řízení přístupu až na úrovni řízení viditelnosti existence samotných metadat či jejich částí apod.
- Federované AAI služby
 - Autentizace uživatelů přes federované služby přihlašování (eduID.cz apod.),
 - tj. s využitím existující identity z instituce, bez nutnosti vytváření dalšího účtu, hesla atd.

Služby pro sdílení dat s jinými uživateli

- Sdílení dat by mělo být možné jednak s konkrétními osobami a skupinami
 - výběr např. z institucionálních identit daných federovaným přihlášením apod.,
 - případně sdílením „odkazem“.
- Může být řešeno i formou „data visiting“.
 - Řízené povolením přístupu vybraným uživatelů do „domovského“ úložiště s danými daty,
 - bez nutnosti přenosu do jiné lokality.
- Možnost anonymních view-only přístupů k datům.
 - Pro podporu předání dat pro zaslepené recenzní řízení apod.
 - Stále častější povinnost sdílení podkladových dat k publikacím.
 - Někdy již v okamžiku recenzního řízení.
 - Ukázka existující implementace např. v platformě OSF: <https://help.osf.io/hc/en-us/articles/360042097853-Create-a-View-only-Link-for-a-Registration>

Služby pro podporu a plánování správy výzkumných dat

- Definované (strojově zpracovatelné) politiky práce s daty v repozitářích v NDI.
- Podpora tvorby DMP.
 - Vygenerování DMP
 - nebo vybraných částí
 - na základě politik použitého repozitáře, znalosti licencí, formátů, velikostí dat, cílených retencí dat (maximální či minimální doba uchovávání např. na základě zákonných požadavků) apod.
 - Možnost napojení/integrace existujících nástrojů pro DMP.
 - Např. Data Stewardship Wizard.

Metriky a měření plnění FAIR atributů pro výzkumná data

- Pro datasey může být zajímavé nějakým způsobem **měřit** plnění jednotlivých **FAIR atributů**.
 - NRP EOSC CZ by měla **definovat metriky** hodnocení a **poskytovat nástroje** pro **měření** plnění FAIR atributů.
- Hodnocení na základě informací o
 - datech,
 - metadatech,
 - nastavení a politikách konkrétního repozitáře apod.
- Příkladem implementace podobného nástroje může být např. FAIR-Checker (<https://fair-checker.france-bioinformatique.fr/>) nebo FAIR metriky pro DMP v Data Stewardship Wizard.



Zdroj: <https://researchers.ds-wizard.org/>

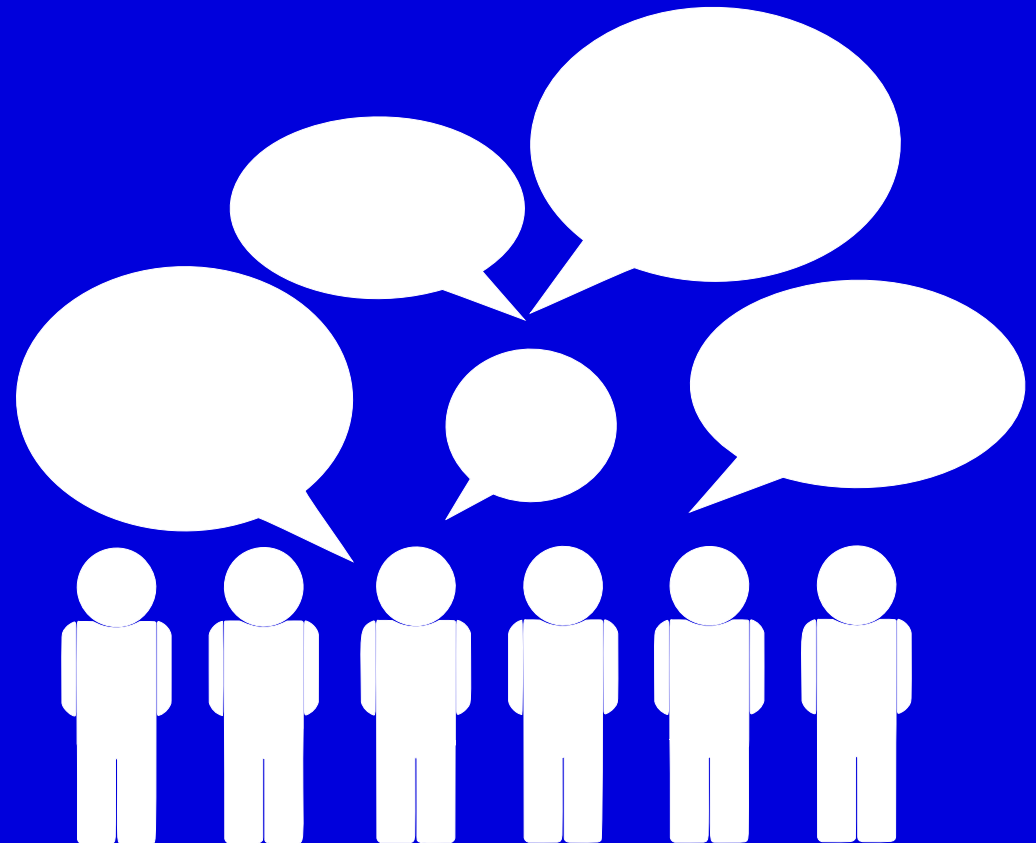
MUNI
ICS

Kam dál

Kam další – další zdroje informací

- Kurz *Bezpečná správa výzkumných dat* od Open Science týmu MU.
 - <https://www.ics.muni.cz/o-nas/celozivotni-vzdelavani>
 - Druhý běh skončil v prosinci 2022, třetí běh v přípravě.
- Kyberbezpečnostní doporučení týmu CSIRT-MU
 - <https://csirt.muni.cz/>
 - Bezpečnostní portál MUNI
 - <https://security.muni.cz/>
 - **Kurz Kyberkompas**
 - <https://security.muni.cz/cybercompass>
 - Kurz GDPR
 - <https://security.muni.cz/gdpr-kurz>
 - Techniky sociálního inženýrství
 - https://security.muni.cz/socialni_inzenyrstvi

Otázky?



Zdroj: [Communicate_communication_conference_2028004](#) od [OpenClipart-Vectors](#) z [Pixabay](#)

UKLÁDÁNÍ
VÝZKUMNÝCH
DAT –
SEMINÁŘ
PRO
MENDELOVU
UNIVERZITU

AMUNI

2023-03-16



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY