

Správa výzkumných dat

Jak podporovat open science



CENTRUM PRO PODPORU
OPEN SCIENCE
Univerzita Karlova

Dagmar Hanzlíková & Milan Janíček

Projekt RKV II

27. 5. 2021, v 09:15

Centrum pro podporu open science

openscience@cuni.cz

Toto dílo podléhá licenci [Creative Commons Uved'te původ 4.0 Mezinárodní Licence](https://creativecommons.org/licenses/by/4.0/).



Toto školení bylo podpořeno projektem OP VVV Rozvoj kapacit pro výzkum a
vývoj UK II, CZ.02.2.69/0.0/0.0/18_054/0015222



UNIVERZITA
KARLOVA



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



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



HR EXCELLENCE IN RESEARCH

Struktura

- Open science
- Správa výzkumných dat
 - Co jsou výzkumná data a FAIR data
 - Data management plan (DMP)
 - Nástroje pro tvorbu DMP
- Podmínky poskytovatelů financí
- Podpora na UK



Open Science

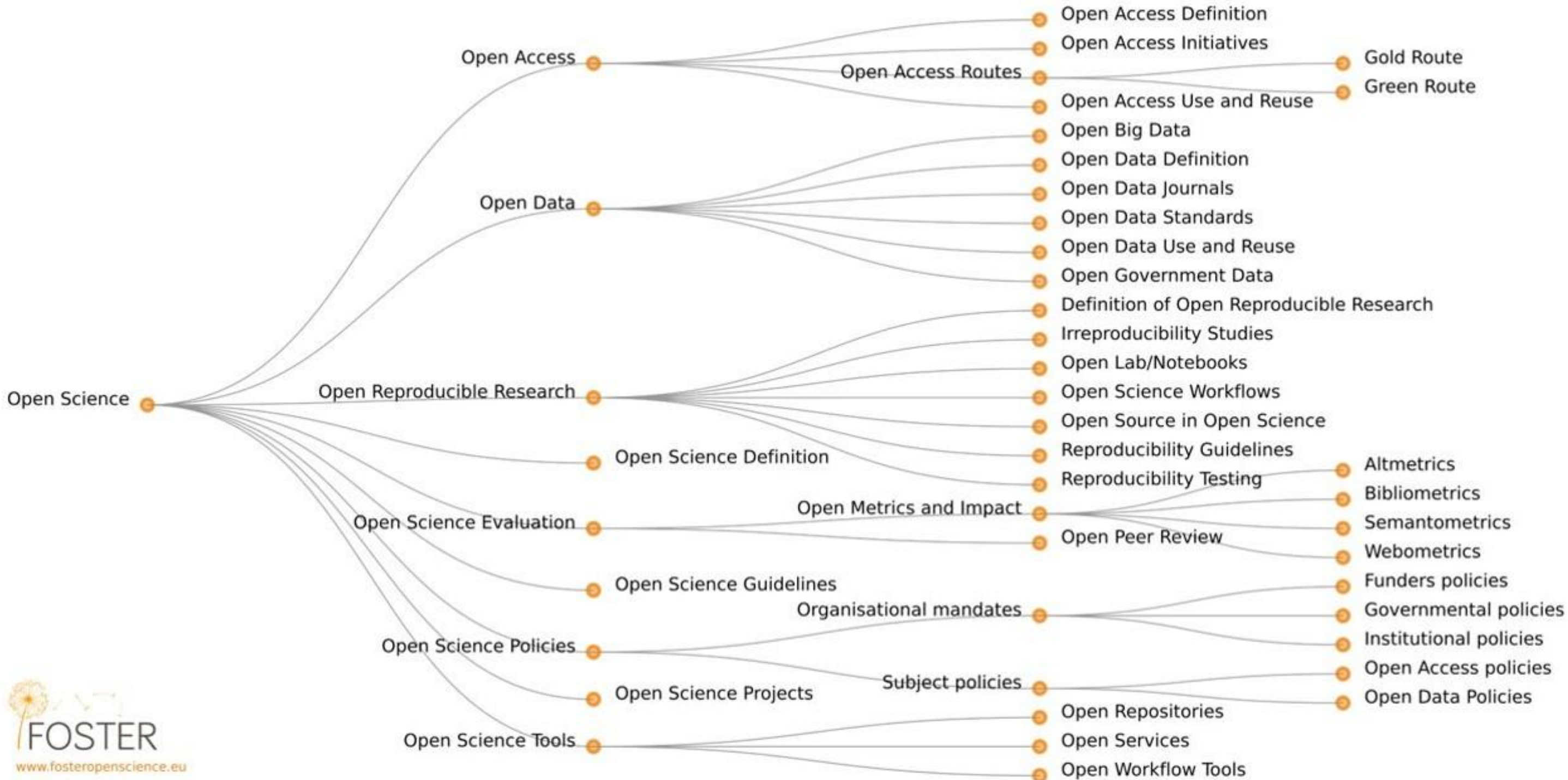
Open Science Definition

- **Open Science** is the practice of science in such a way that others can collaborate and contribute, where research data, lab notes and other research processes are freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods.

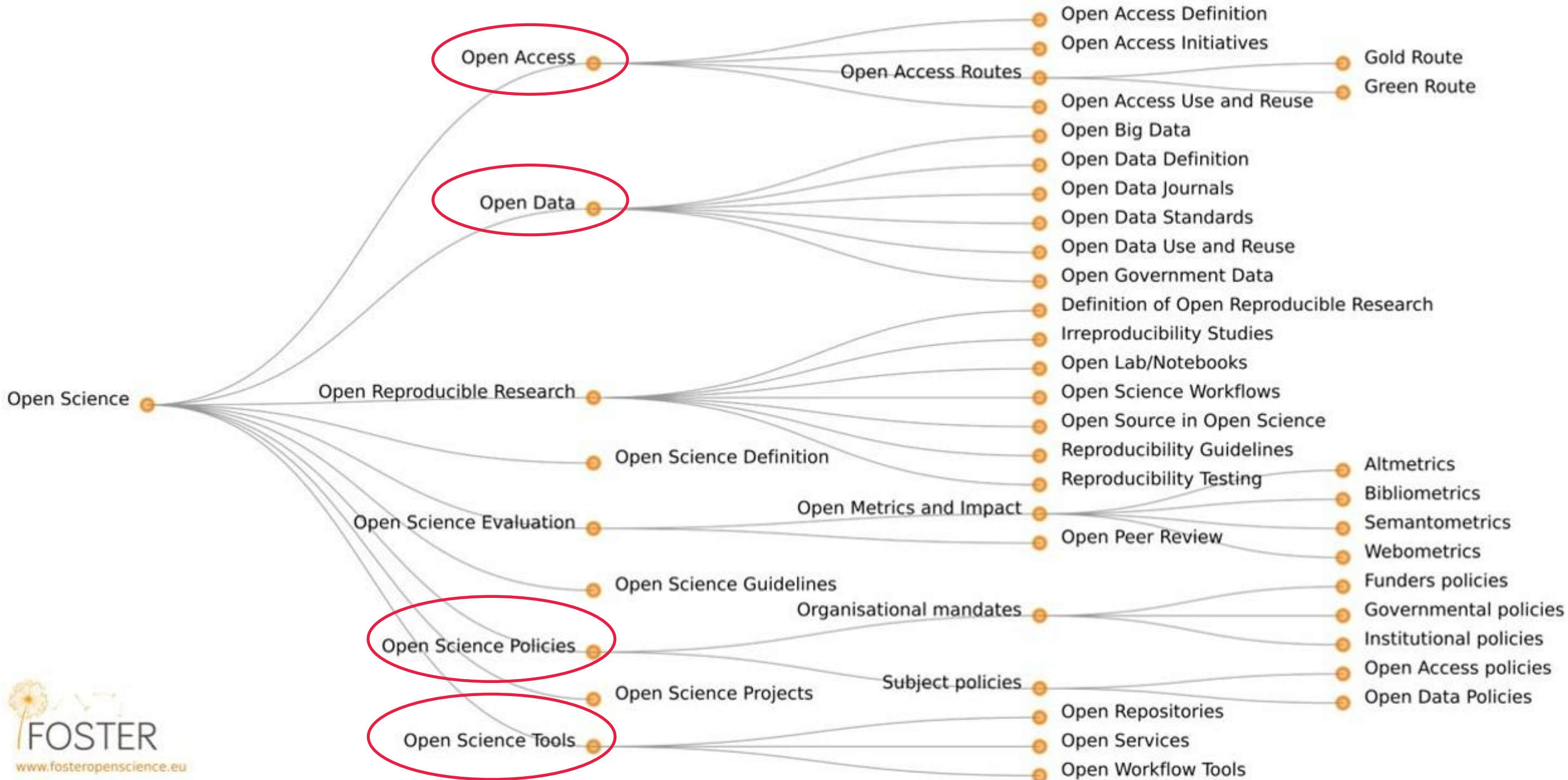
<https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>



Open Science Taxonomy



Open Science Taxonomy



Správa výzkumných dat

Co jsou výzkumná data?

= informace, které byly vytvořeny za účelem zjištění či reprodukování výsledků výzkumu

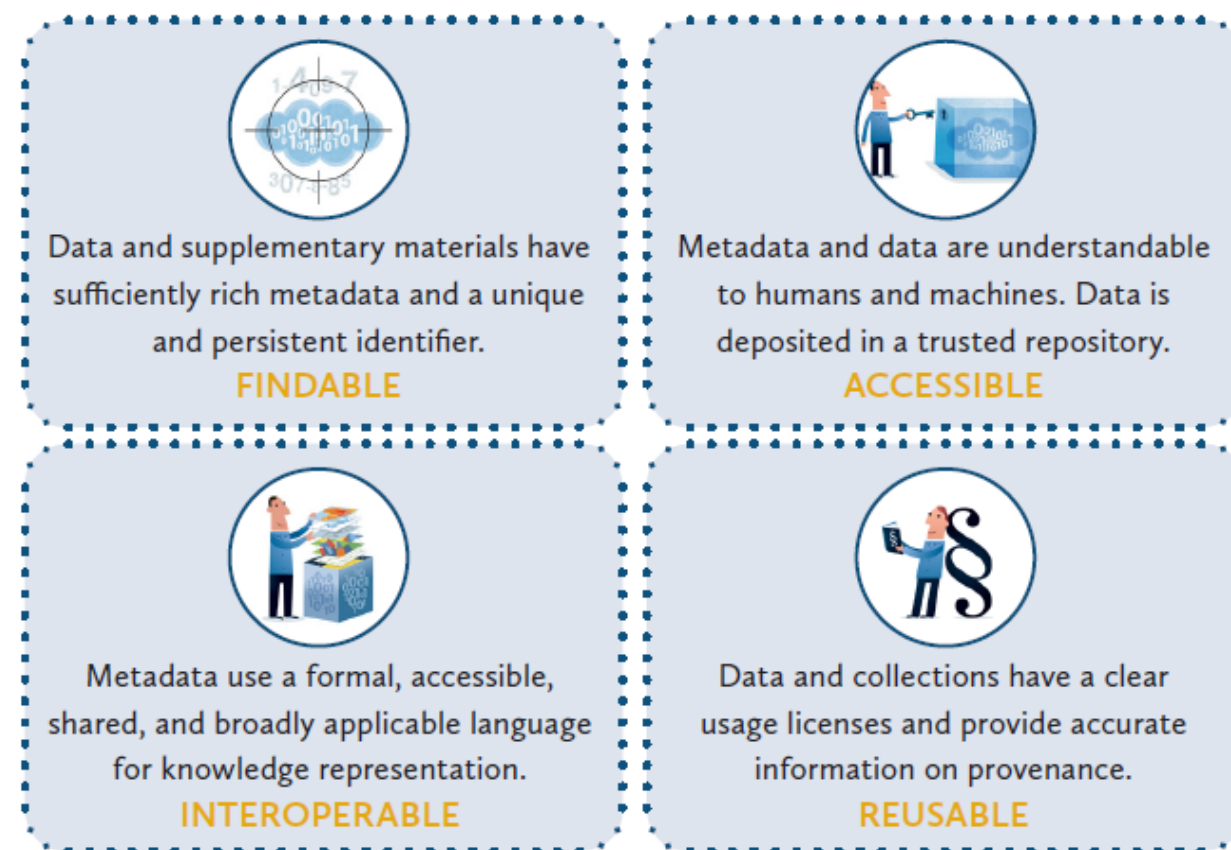
- Výzkumná data mohou mít různou podobu (digitální i nedigitální)
 - Tabulky, dokumenty
 - Audio a video nahrávky, obrázky, fotografie
 - Software, skript
 - Laboratorní deníky, terénní poznámky
 - Vzorky, exempláře
- Raw data vs. zpracovaná data



FAIR data

- **F**indable (vyhledatelná)
 - Perzistentní identifikátory
 - Strojově čitelná metadata
- **A**ccessible (dostupná)
 - Repozitáře
- **I**nteroperable (interoperabilní)
 - Standardizované formáty a slovníky
- **R**eusable (opětovně využitelná)

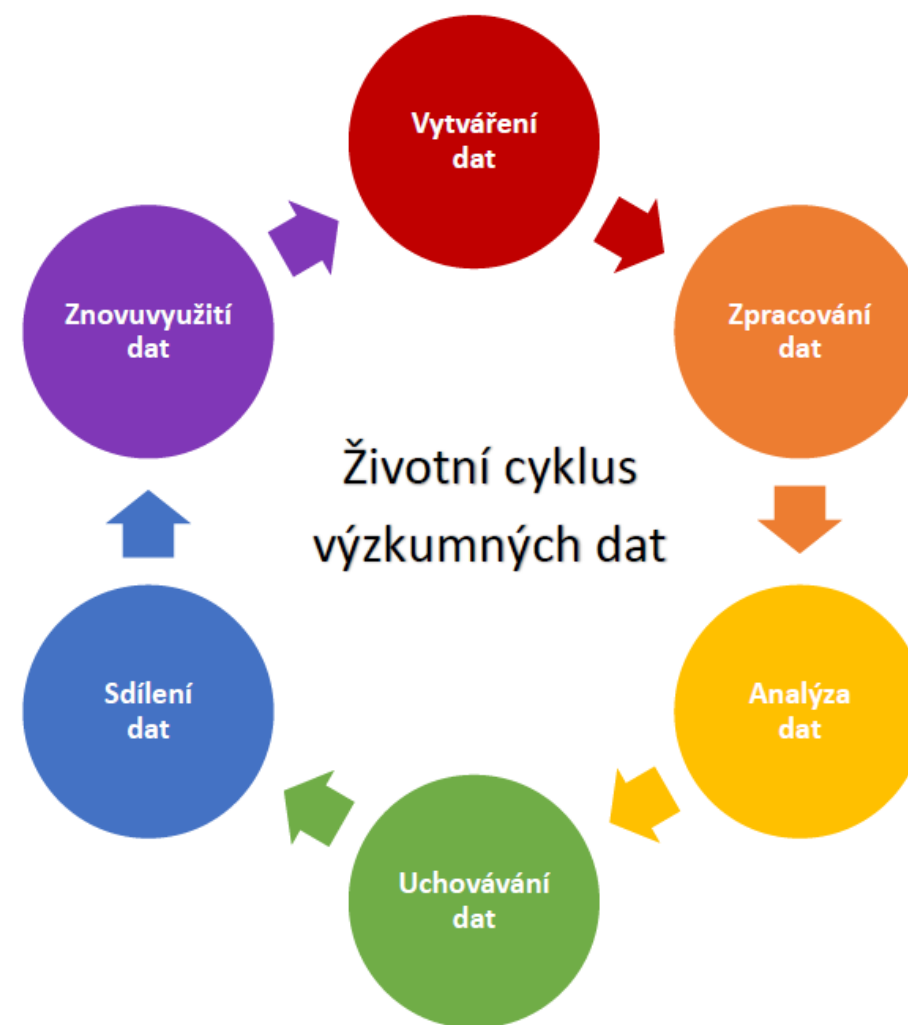
What is FAIR DATA?



Správa výzkumných dat

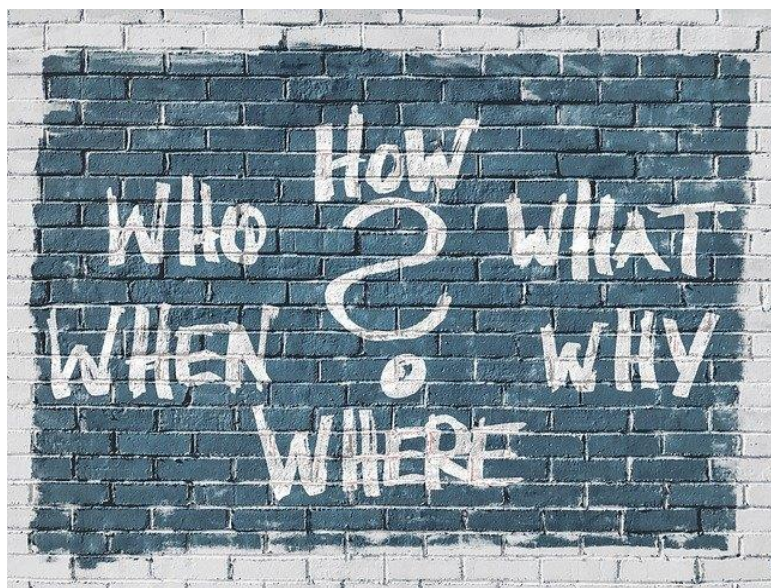
= organizování, skladování a dlouhodobé uchovávání

- Snazší psaní a revize článků
- Lepší zabezpečení dat, nižší riziko jejich ztráty
- Integrita výzkumu, lepší reputace
- Pokrok v celosvětovém výzkumu díky možnosti sdílení a opakovaného využití dat



Data management plan (DMP)

= dokument, který specifikuje, *jaká data a jakým způsobem budou během výzkumu vytvářena*, a obsahuje informace o jejich *dostupnosti a využití*



Obsah DMP

1. Administrativní data
2. Sběr dat
3. Dokumentace a metadata
4. Etické a právní otázky
5. Skladování a zálohování
6. Dlouhodobá ochrana
7. Sdílení dat
8. Zodpovědnost a zdroje

Proč vytvářet DMP?

- Pomáhá předvídat **potenciální problémy**
- Zajištění, že data jsou **přesná, úplná a spolehlivá**
- **Kontinuita** a **konzistence** dlouhodobých projektů
- Pomoc při **sdílení dat**
- Splnění podmínek **poskytovatelů financí** (např. H2020, Horizon Europe)

• version B: only changing the article, eg. OIP_a-WN1 → OIP_the-WN1

↓ sentence ↓ speaker ↓ article ↓ white noise ↓ context 1

DMP: Administrativní údaje

*Uved'te **základní informace**, které poskytnou kontext k danému DMP*

- Základní informace o výzkumu, např. název projektu, jméno, kontaktní údaje, ID hlavního řešitele, informaci o poskytovateli financí
- Krátký popis výzkumu, ke kterému se data vztahují (abstrakt)
- Nařízení, opatření nebo směrnice, které upravují uchovávání dat na instituci (pokud existují)

DMP: Sběr dat

*Definujte, **jaká data** budou vytvářena a **jakým způsobem***

- Existují data, která lze znovu využít?
- Jaké standardy nebo metodiky budete k vytváření dat používat?
- V jakém formátu budou data uchovávána?
- Jak budou složky a soubory strukturovány a pojmenovány?
- Zajistíte kontrolu kvality dat?



DMP: Dokumentace a metadata

*Poskytněte informace, které jsou potřeba pro **čtení a interpretaci dat** v budoucnosti*

- Bude vytvořena doplňující příloha s metadaty a dokumentací?
- Jak bude tato příloha vytvořena?
- Jaké metadatové standardy budou využity a proč?



DMP: Etické a právní otázky

Zvažte etické a právní záležitosti

- Máte souhlas pro uchovávání a sdílení dat?
- Jak budete chránit osobní údaje účastníků výzkumu, pokud to bude třeba?
- Budou data sdílena, nebo bude sdílení omezeno?
- Budou data licencována?



DMP: Skladování a zálohování

*Zvažte, kde budou data **skladována** a jak budou **zálohována**, včetně **přístupu** k nim a **zajištění bezpečnosti***

- Máte dostatek místa pro ukládání dat, nebo budete muset započítat náklady na úložné prostory?
- Kdo bude zodpovědný za zálohování a obnovení dat?
- Jaké je potenciální riziko pro zabezpečení dat a jak bude řešeno?
- Jak zajistíte, že se spoluřešitelé dostanou bezpečně k datům?



DMP: Dlouhodobá ochrana

*Určete, která data jsou vhodná pro **dlouhodobou ochranu** a jak nejlépe je **uchovat***

- Jaká data musí být uchována nebo zničena ze smluvních/právních důvodů?
- Jak vybrat data pro dlouhodobé uchování?
- Vyčíslili jste časové a finanční náklady na přípravu dat pro dlouhodobé uchování a sdílení?



DMP: Sdílení dat

*Rozmyslete, **která data budete sdílet a jak**, a také jak vám uživatelé vašich dat budou moci **přiznat autorství***

- S kým budete data sdílet a za jakých podmínek?
- Jakým způsobem budete data sdílet?
- Kdy budou data dostupná?
- Je potřeba nastavit nějaká omezení ve sdílení dat?
- Jak se o datech dozví vaši potenciální uživatelé?



DMP: Zodpovědnost a zdroje

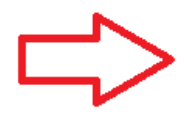
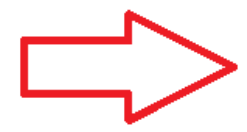
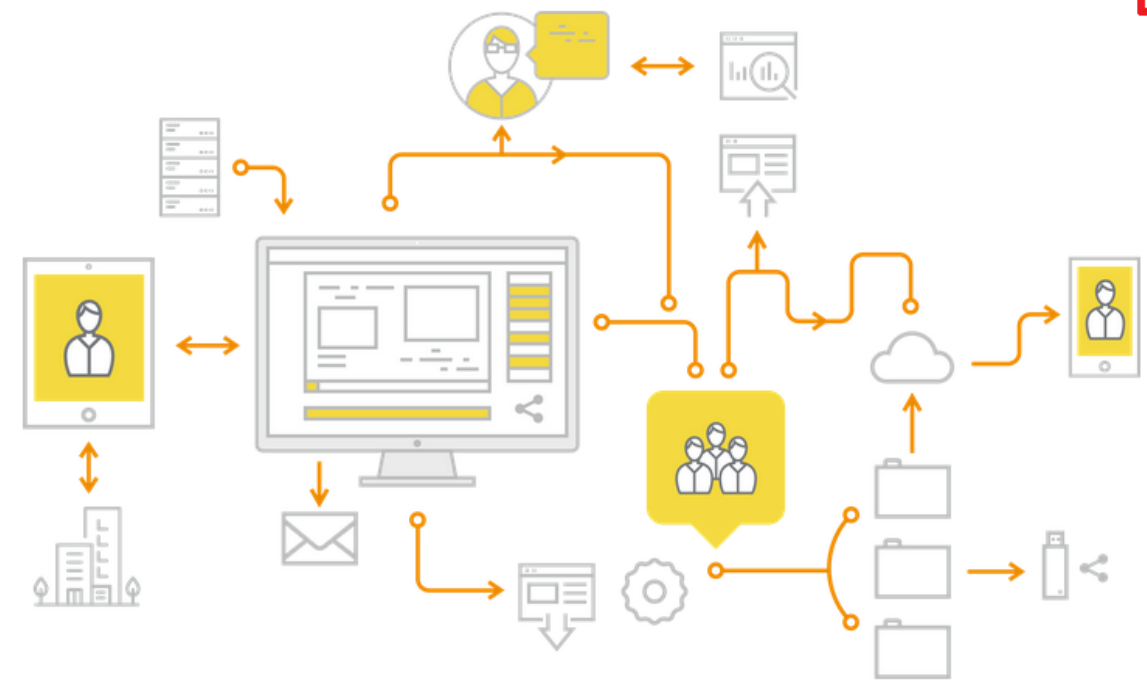
Určete *role a odpovědnost* pro všechny aktivity s daty v rámci výzkumu a zvažte případné *náklady s vytvářením a uchováváním dat*

- Kdo bude odpovědný za vedení DMP?
- Jak bude odpovědnost rozdělena mezi řešitele projektu?
- Potřebujete nějaké další vybavení nebo expertizu?
- Potřebujete nějaké další finanční zdroje pro sběr, zpracování a uchování dat?

Nástroje pro tvorbu DMP

Plan to make data work for you

Data Management Plans that meet institutional funder requirements.



Sign in Create account

* Email

* Password

Forgot password?

☐ Remember email

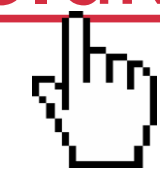
Sign in

- or -

Sign in with your institutional credentials

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It is provided by the Digital Curation Centre (DCC).

dmponline.dcc.ac.uk



A role of striatal M4 muscarinic acetylcholine receptors in learning of stereotypic motor sequences

Project Details

Plan overview

Write Plan

Share

Download

expand all | collapse all

13/13 answered

Data Collection (2 / 2)

What data will you collect or create?

B

I

Following kinds of data will be created:

- „master sheet“
 - contains the information about the experiment - which mouse is in which part of the experiment, dates of specific sessions
 - master sheet will be updated manually by researchers
 - master sheet will be stored in XSL (MS Excel) format
- operant conditioning box program
 - defines sequence of operations that is requested in respective parts of mice's training
 - the program will be written in device specific programming language (MedAssociates)
- operant conditioning box logs
 - logs are stored automatically by the operant box
 - all mouse's interactions with operant box are recorded into log as two values
 - timestamp
 - operation code (eg. correct press, wrong press, issuing of reinforcement)
 - the data are in plain text structure of the data (semantics) is described in

Guidance

Comments

DCC

FAIRsFAIR

Questions to consider:

- What type, format and volume of data?
- Do your chosen formats and software enable sharing and long-term access to the data?
- Are there any existing data that you can reuse?

Guidance:

Give a brief description of the data, including any existing data or third-party sources that will be used, in each case noting its content, type and coverage. Outline and justify your choice of format and consider the implications of data format and data volumes in terms of



ds-wizard.org/



About

Features

Resources

Get Started



Get Started Using the Data Stewardship Wizard

Demo

For exploring the DSW features

- Easy to sign up and use
- A shared instance with other users
- **Not for serious usage**

- ✓ Questionnaires & DMPs
- ✓ Knowledge Models Management

Try the demo

Researchers

For individual researchers

- Easy to sign up and use
- Ready to use Knowledge Models
- Privacy and stability

- ✓ Questionnaires & DMPs



Start planning

Self-Managed

For organizations

- All the DSW features available
- Your own instance
- You need to host and run the instance by yourself

- ✓ Questionnaires & DMPs
- ✓ Knowledge Models Management
- ✓ User Management & Organization Settings
- ✓ Features configuration & visual style customizations

Contact us


DSW Cloud


For organizations


We offer managing the DS Wizard instance for interesting projects that want to use it seriously but don't want to run it by themselves. *


- ✓ Questionnaires & DMPs
- ✓ Knowledge Models Management
- ✓ User Management & Organization Settings
- ✓ Features configuration & visual style customizations


Contact us

 DS Wizard


 Knowledge Models


 Projects


 Help >


 Dagmar Hanzlíková >


<< Collapse sidebar


Vliv ochytřovacího zaklínadla na výsledky testů 

 Questionnaire

 Metrics


 Preview

 Documents


 Settings

View

TODOs

Version history 

Current Phase

Before Submitting the Proposal 

Chapters

I. Administrative information1

II. Re-using data1

III. Creating and collecting data6

IV. Processing data3

V. Interpreting data2

VI. Preserving data5

VII. Giving access to data2

2 What existing encodings/terminologies/vocabularies/ontologies will you be using?

Horizon 2020 DMP


+ Add

3 Will you be using new types of data?

Horizon 2020 DMP


Sometimes the type of data you collect can not be stored in a commonly used data format. In such cases you may need to make your own, keeping interoperability as high as possible.

☒ Desirable: Before Submitting the Proposal

 Data Stewardship for Open Science: [ikk](#)

☐ a. No, all of my data will fit in common formats

Interoperability

☐ b. Yes, I will need to use custom formats for some of my data 

4 How will you be collecting and keeping your metadata?

Horizon 2020 DMP

Science Europe DMP

Setkali jste se s pojmem Data Management Plan?

<u>Ano, už jsem ho vytvářel/a</u>	9%
<u>Ano, ale ještě jsem ho nevytvářel/a</u>	47%
<u>Ne.</u>	44%

Počet hlasů: 98

DMP

<u>... bych určitě zvládl/a vytvořit bez další podpory.</u>	1%
<u>... většinu bych zvládl/a vytvořit bez další podpory.</u>	56%
<u>... menší část bych zvládl/a vytvořit bez další podpory.</u>	32%
<u>... bych nezvládl/a vytvořit.</u>	10%

Počet hlasů: 71

Jak vnímáte to, jak zpracováváte svá data?

<u>Nemusím nic zlepšovat! :-)</u>	1%
<u>S daty nemám problém, ale něco by zlepšit šlo.</u>	28%
<u>Mám velký prostor pro zlepšení.</u>	62%
<u>Nemám data.</u>	9%

Počet hlasů: 86



Podmínky poskytovatelů financí

Horizon 2020 & Horizon Europe



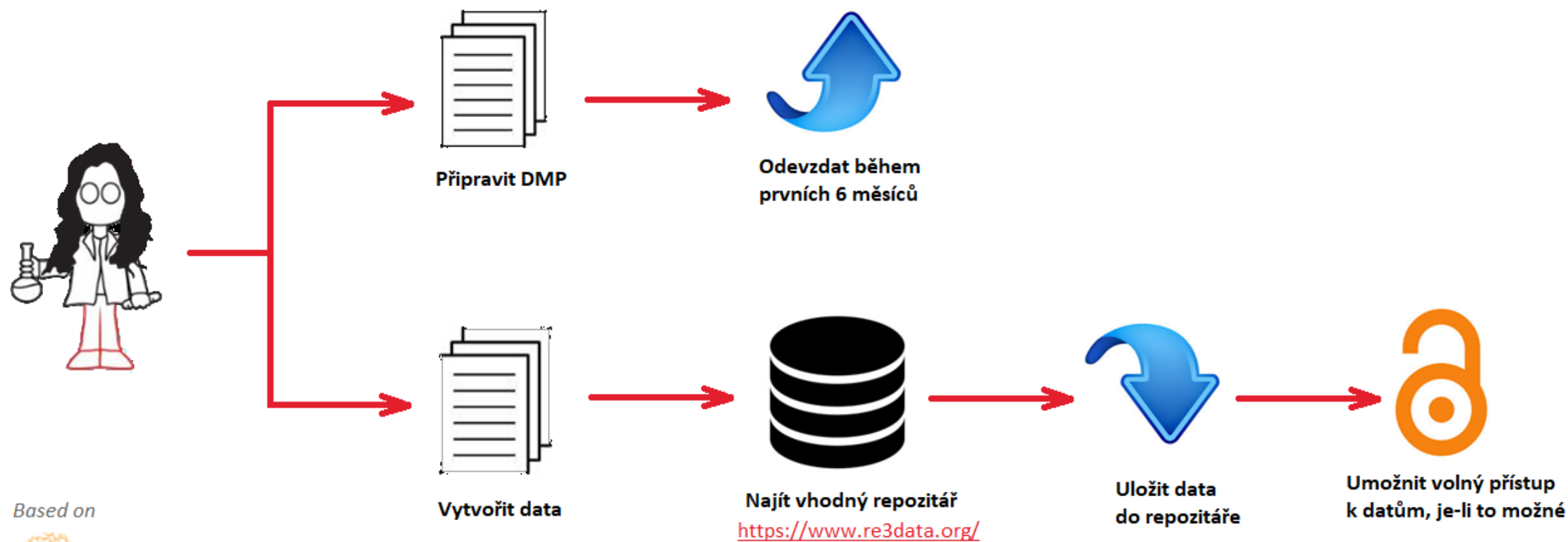
Horizon 2020: ORD Pilot

Povinnosti

1. Vytvořit (a aktualizovat) **plán managementu dat**
2. **Uložit data** a související metadata do datového repozitáře
3. Zajistit třetím stranám **volný přístup k datům** a umožnit jejich bezplatné vytěžování, využívání, reprodukování a šíření
4. Upřesnit, jaké **nástroje** jsou potřeba pro použití dat nutných k validaci vědeckých závěrů, v ideálním případě tyto nástroje také poskytnout



Horizon 2020: ORD Pilot



Based on



www.fosteropenscience.eu



Horizon Europe

*Příjemce se zavazuje zodpovědně **spravovat výzkumná data**, která v rámci projektu vznikla, v souladu s principy **FAIR**.*

Povinnosti

1. Vytvořit (a aktualizovat) **plán managementu dat**
2. **Uložit data** a související metadata do datového repozitáře
3. Zajistit (prostřednictvím repozitáře) **otevřený přístup** k uloženým datům v souladu s principem “**as open as possible, as closed as necessary**”
4. Upřesnit, jaké **nástroje** jsou potřeba pro použití dat nutných k validaci vědeckých závěrů, v ideálním případě tyto nástroje také poskytnout



České grantové agentury

TA ČR: KAPPA

*Příjemce se zavazuje vytvořit **plán managementu dat** a zodpovědně spravovat výzkumná data v souladu s principy **FAIR**.*

Povinnosti

1. Během prvních 6 měsíců odevzdat **plán managementu dat**
2. **Uložit data** a související metadata do datového repozitáře
3. Zajistit **otevřený přístup** k uloženým datům v souladu s principem “**as open as possible, as closed as necessary**”
4. Upřesnit, jaké **nástroje** jsou potřeba pro použití dat nutných k validaci vědeckých závěrů, v ideálním případě tyto nástroje také poskytnout

Evropská směrnice o otevřených datech

Evropská směrnice 2019/1024 o otevřených datech a opakovaném použití informací veřejného sektoru

- Podpora výzkumu z veřejných prostředků
- Smlouva o poskytnutí podpory
 - Způsob správy výzkumných dat (DMP)
 - Informace o dostupnosti a způsobu šíření výsledků výzkumu a výzkumných dat
- Poskytování informací
- Přístup k datům

Soutěž Primus

- Zásady soutěže Primus
 - Řešitel i fakulta se zavazují spolupracovat na vytvoření *Data management plánu daného projektu* (čl. VI, odst. 3, písm. f)
 - Spolupráce s DMP skupinou
 - Konkrétní podoba této povinnosti ošetřena v jednotlivých smlouvách

dmp@cuni.cz



Uznatelnost nákladů

- Myslet na finance při psaní projektu
- Náklady na práci s daty
 - Hardware, software
 - Zabezpečení, úložiště
 - Personální zajištění

Náklady uznatelné
po dobu trvání projektu!





Podpora na UK

Centrum pro podporu open science

- Tvorba metodických a vzdělávacích materiálů
 - Informační webové stránky
 - E-learningové kurzy k Open access a Správě výzkumných dat
 - Informační letáky, factsheety a návody v Zenodo komunitě
- Podpora vědecké komunity v oblastech otevřené vědy
 - Výběr vhodného OA časopisu, prověřování pochybných časopisů...
 - Nakládání s výzkumnými daty, tvorba DMP...
 - Právní podpora v oblasti autorsko-právní problematiky
- Školení, přednášky

Pracovní skupina pro RDM

dmp@cuni.cz

- Podpora vědců
 - Tvorba data management plánů
 - Práce s výzkumnými daty
- Tvorba metodických a vzdělávacích materiálů
- Přednášky a školení
- Strategický rozvoj správy dat a DMP na UK
 - Začlenění DMP do interních grantových soutěží





Prostor pro diskuzi

Užitečné odkazy

- [Směrnice Evropského parlamentu a Rady \(EU\) 2019/1024](#) o otevřených datech a opakovaném použití informací veřejného sektoru
- [DMPonline](#), [DSW](#): Nástroje pro tvorbu DMP
- [GO FAIR](#): Iniciativa věnující se principům FAIR
- [FAIRsharing](#): Nástroj pro vyhledání oborových metadatových standardů a databází
- [Re3data.org](#): Registr repozitářů výzkumných dat
- Obecné repozitáře: [Zenodo](#), [Figshare](#), [Dryad](#)
- [DCC How-to Guides](#): Návody a checklisty ke správě výzkumných dat

Použité obrázky

- [6,7] FOSTER Open Science Taxonomy [online]. Dostupné z https://www.fosteropenscience.eu/themes/fosterstrap/images/taxonomies/os_taxonomy.png Podléhá licenci CC-BY
- [10] LIBER. Implementing FAIR Data Principles: The Role of Libraries. Dostupné z <https://libereurope.eu/wp-content/uploads/2017/12/LIBER-FAIR-Data.pdf>
- [12] Geralt. questions-qui-quoi-comment. *Pixabay* [online]. Dostupné z <https://pixabay.com/images/id-2110967/> Podléhá licenci [Pixabay](#).
- [23,24] Screenshot <https://dmponline.dcc.ac.uk/> 31. 1. 2021
- [25,26] Screenshot <https://ds-wizard.org/get-started.html> 27. 4. 2021
- [37] SCR3AMFR3AK. Money. *Pixabay* [online]. Dostupné z <https://pixabay.com/images/id-2082383/> Podléhá licenci [Pixabay](#).



Děkujeme za pozornost.

Centrum pro podporu open science
Ústřední knihovna UK
openscience@cuni.cz