

- MENDELU
- Ústav
- vědecko-pedagogických
- informací a služeb



Data Management Plan

= Plán správy dat

Naděžda Fasurová, Soňa Kehmová, Martina Šmardová

Open Science centrum, leden 2024

Co je to Plán správy dat?

- Plán správy dat je dokument, který:
 - popisuje nakládání s výzkumnými daty během projektu a po jeho skončení.**
- Ekvivalent anglického pojmu „data management plan“.
- Podmínka v projektech





Proč je potřeba řešit správu výzkumných dat?



Výstupem výzkumu jsou nejen publikace ale i **výzkumná data**



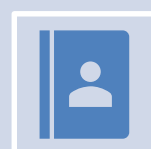
K ověření výsledků výzkumu, pro **znovuvyužitelnost** výzkumných dat



Navazuje na **Open Access**: náklady na výzkum jsou vysoké - požadavek, aby **veřejnost měla také přístup k datům** vzniklým z veřejných prostředků

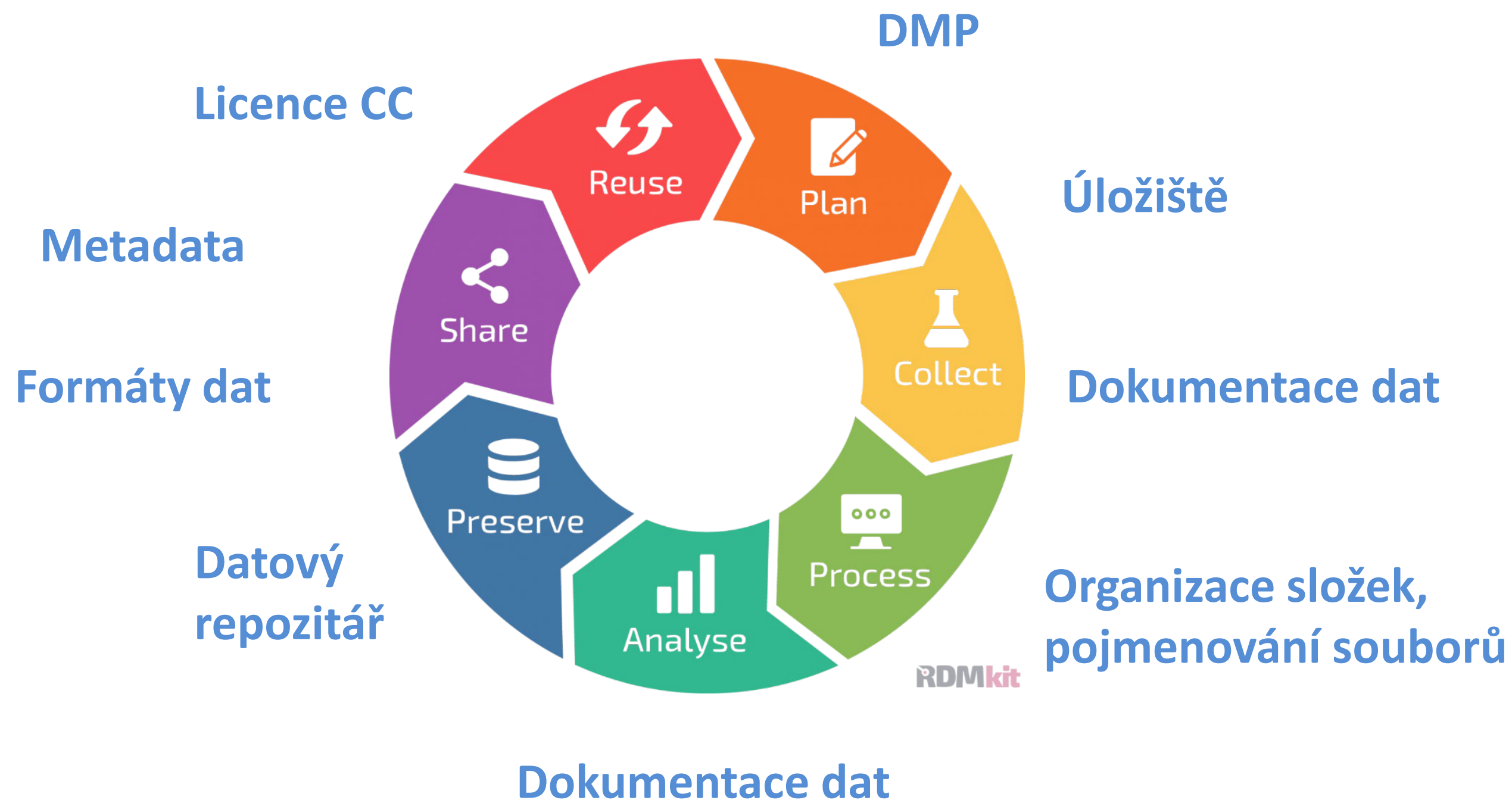


Projektové podmínky: popis správy výzkumných dat v projektu, Plán správy dat



Poskytnutí výzkumných dat pro recenzní řízení

Správa výzkumných dat: životní cyklus dat



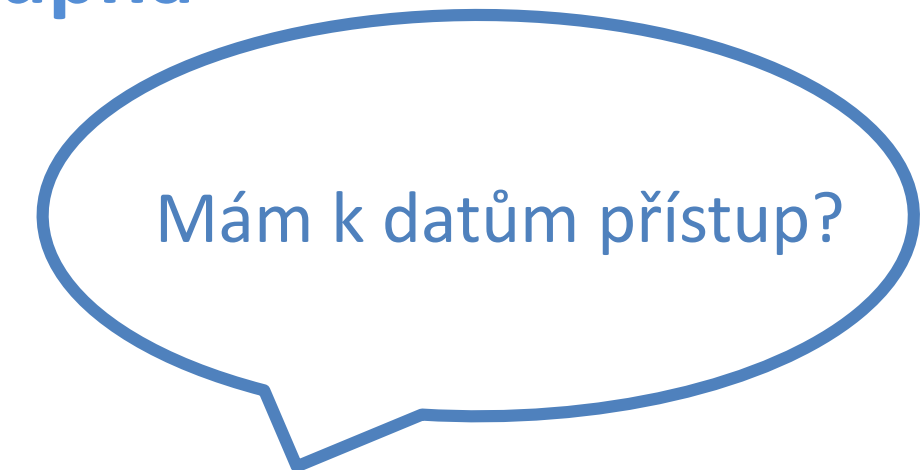
PRINCIPY FAIR = dobrá práce s daty

Findable = vyhledatelná



- identifikátor DOI / Handle, metadata

Accessible = přístupná



- důvěryhodný repozitář, informace o přístupnosti dat

Interoperable = interoperabilní



- oborové hesláře a popis dat (metadata), dlouhodobě udržitelné formáty souborů, kvalita výzkumných dat

Reusable = využitelná



- dokumentace dat, licence (CC BY, CC BY-NC-ND)

Podmínky poskytovatelů financí



Kdy odevzdat DMP:

Nejčastěji je potřeba odevzdat DMP společně s **první průběžnou zprávou** o řešení a na konci projektu.

- ❑ **GAČR** - aktuálně není povinnost odevzdat DMP, je povinný pouze stručný popis správy výzkumných dat při podání projektu; [\(LINK\)](#)
- ❑ **TAČR** - DMP odevzdat s první průběžnou zprávou, stručný popis správy výzkumných dat při podání projektu [\(LINK\)](#)

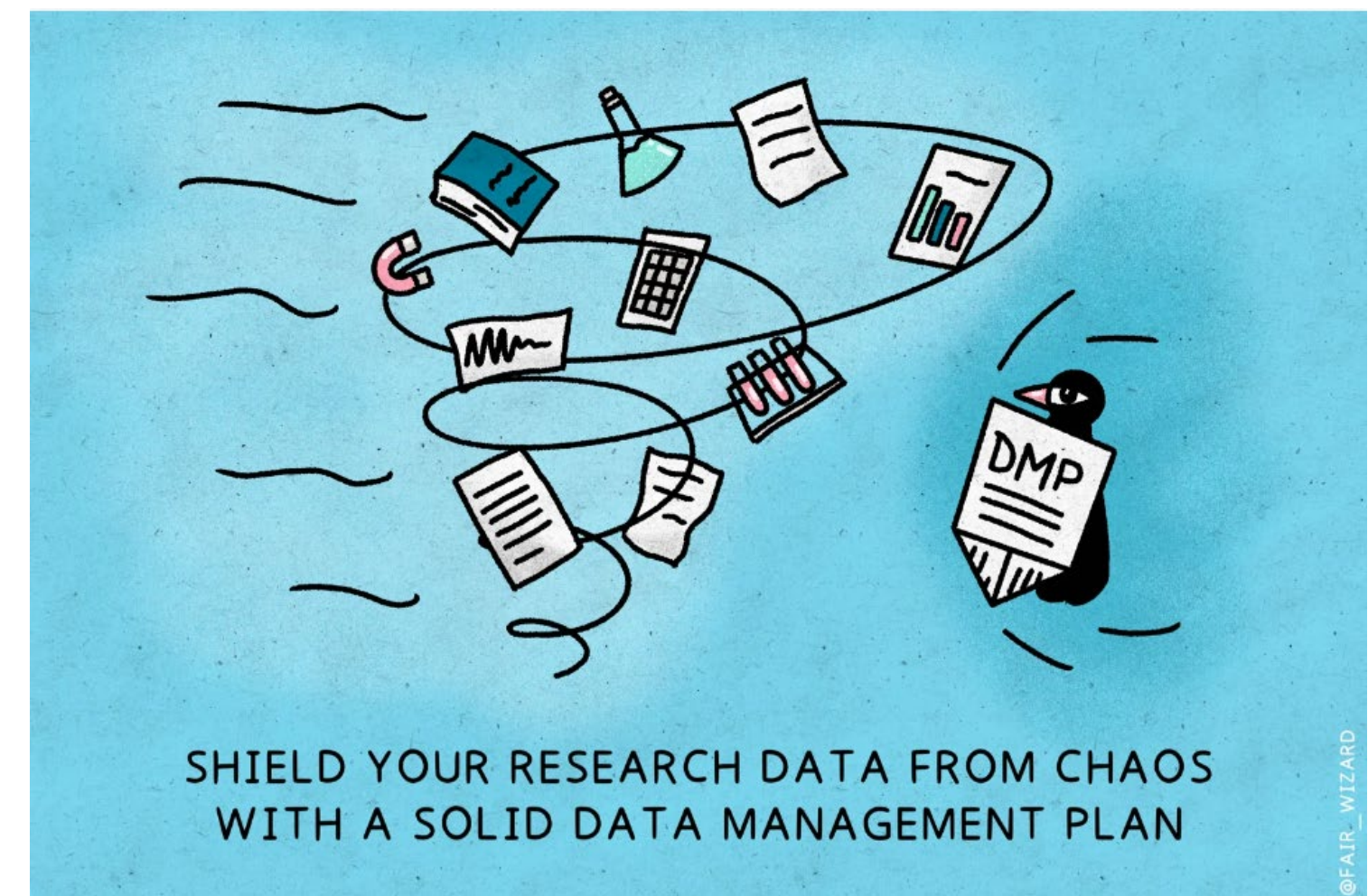
- ❑ **Horizon Europe** - DMP odevzdat do 6 měsíců od začátku řešení projektu [\(LINK\)](#)
- ❑ **OP JAK** – DMP společně s první průběžnou zprávou, s dalšími průběžnými zprávami a na konci projektu [\(LINK\)](#)

<https://uvis.mendelu.cz/projekty-a-open-science>

Plán správy dat – Data Management Plan

Data Management Planning (proces) x Data Management Plan (dokument)

- ❑ DMP je **živý dokument** - je pravidelně aktualizován během výzkumu
- ❑ Reflektuje skutečnou práci s daty
- ❑ **DMP píšete pro sebe:** kontinuita a konzistence dlouhodobých projektů - nastavení správy dat, pomáhá předvídat potenciální problémy i výdaje projektu



Zdroj: FAIR Wizard - <https://fair-wizard.com/blog/about-data-management-planning>

T A
Č R

PLÁN SPRÁVY DAT

verze2020_01

Informace o projektu:

Kód projektu:

Název projektu:

Datum:

Verze Plánu správy dat:

Zpracovatel Plánu správy dat:

(Zadejte prosím kód projektu)

(Vyplňte prosím název projektu)

(Zadejte prosím datum ve formátu DD/MM/RRRR)

(Vyplňte prosím číslo verze Plánu správy dat)

(Vyplňte prosím jméno zpracovatele a případně kontakt)



Horizon Europe

Data Management Plan template
(vzor)

Šablony DMP

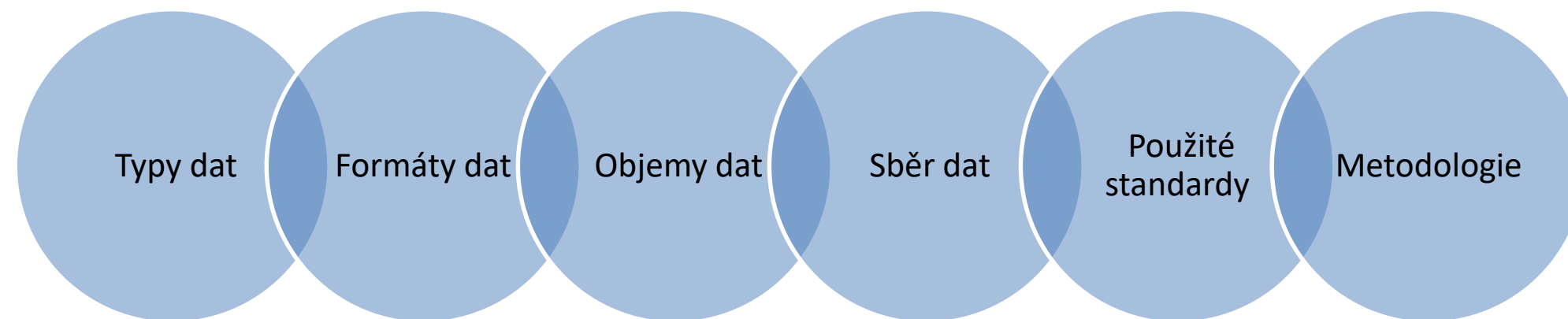
- Horizon Europe** - v angličtině [LINK](#)
- NTK** - [přeložená šablona HE](#)
- TAČR** - [šablona DMP](#)

DMP: šablona Horizon Europe

1. Data Summary / Shrnutí dat
2. FAIR data / FAIR data (přístup k datům, sdílení a znovuvyužití dat)
3. Other research outputs / Další výstupy výzkumu
4. Allocation of resources / Alokace zdrojů
5. Data security / Bezpečnost dat
6. Ethics / Etika
7. Other issues / Ostatní

Generated data x Reused data

Popište:



Popište, jak budou primární data sbírána/generována.
Při využití sekundárních dat odkažte na jejich zdroj.

TYPY DAT	DOPORUČENÉ FORMÁTY	PŘIJATELNÉ FORMÁTY
Tabulková data	csv, .tab, .por, .xml	.txt, xls, .dbf, .ods, .sav, .dta, .mdb
Geoprostorová data	.shp, .shx, .dbf, .prj, .sbx, .sbn, .tif, .tfw, .dwg, .gml	.mdb, .mif, .kml, .ai, dxf, .svg
Textová data	.rtf, .txt, .xml	.html, .doc
Obrazová data	.tif	.jpg, .gif, .tif, .tiff, .raw, .psd, .bmp, .png, .pdf
Audio data	.flac	.mp3, .aif, .wav
Video data	.mp4, .ogv, .ogg, .mj2	.avchd
Dokumentace a skripty	.rtf, .pdf, .xhtml, .htm, .odt	.txt, .doc, .xls, .xml

FAIR



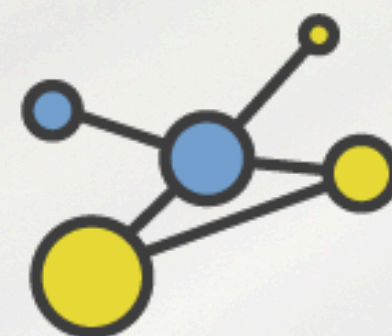
Findable

Vyhledatelná



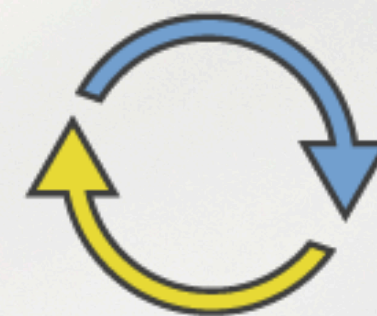
Accessible

Přístupná



Interoperable

Interoperabilní



Reusable

Znovupoužitelná

2. FAIR data

FINDABLE - making data findable, including provisions for metadata / zajištění dohledatelnosti dat (findability), včetně ustanovení pro metadata

- **Will data be identified by a persistent identifier?**
 - *When data is published in repository XY, identifier DOI/Handle will be given to the data.*

DOI (Digital Object Identifier)



trvalý identifikátor pro publikační výstupy a také pro výzkumná data

Příklad: <https://doi.org/10.5281/zenodo.4008297>

Handle

[https://en.wikipedia.org/wiki/Handle System](https://en.wikipedia.org/wiki/Handle_System)

Metadata: Data o datech

- Metadata je **dokumentace, která popisuje data** ve standardizovaném formátu
 - Určeno pro strojové čtení
 - Správný popis dat umožňuje výzkumná data najít a znovu využít

- *Příklad: název, jména autorů, abstrakt, klíčová slova, metody...*

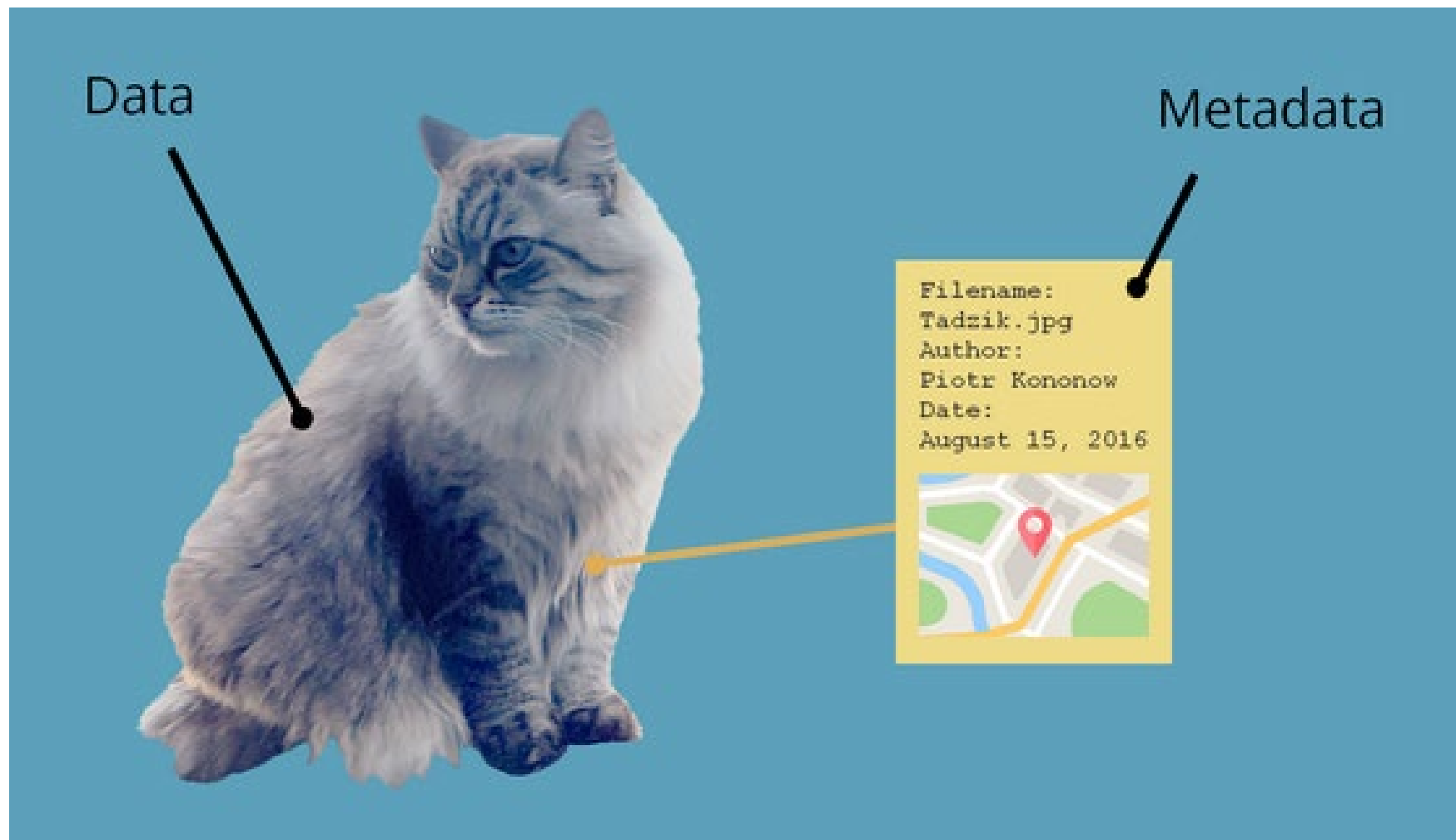


- V repozitáři při ukládání dat vyplňujete základní metadata (<https://zenodo.org/uploads/new>)

- **Videa o metadatech:**

- Ghent University [VIDEO](#), Utrecht university: [VIDEO](#)

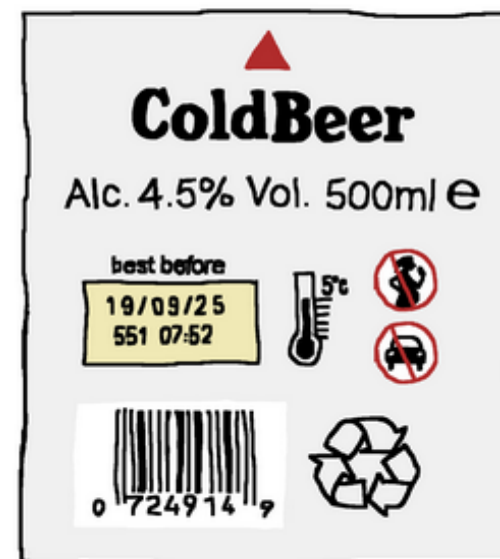
- Doporučení pro metadatový popis NTK



DATA



METADATA

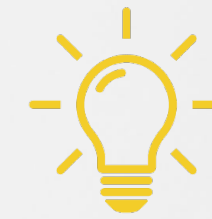


Datové repozitáře

Pro dlouhodobé ukládání **finálních dat**

- Data po skončení výzkumu, podkladová data k článku

Musí být důvěryhodné (Core Trust Seal, provozován důvěryhodnou institucí).



Co umožňuje datový repozitář

- Uložení dat/zveřejnění dat
- Metadatový popis
- Verzování
- Identifikátor DOI/Handle
- Vložení licence Creative Commons
- Odkaz na další výstupy
- Citace

Oborové: Český sociálněvědní datový archiv; ForestPlos; HydroShare

Obecné: Zenodo, Národní datový repozitář, Harvard Dataverse, Dryad

Institucionální datový repozitář

Zjistěte si podmínky repozitáře. Více na <https://uvis.mendelu.cz/datove-repozitare>

<https://zenodo.org/record/4008297#.ZGXUKnZByUk>



- **Will all data be made openly available?** If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions.
- Vždy je nutné popsat např.: *Výzkumná data budou zveřejněna v datovém repozitáři Zenodo pod licencí CC BY, datům bude přidělen identifikátor DOI a budou popsána metadaty XY. Data YZ nebudou zveřejněna kvůli jejich ochraně dokud nebude podána patentová přihláška.*



Otevřená data - Open data

“as open as possible, as closed as necessary” - co nejotevřenější, ale uzavřené podle potřeby s ohledem na zájmy příjemce.

Výjimky z povinnosti zpřístupnit data:

- práva na ochranu duševního vlastnictví (např. patent, užitný vzor);
- práva na ochranu soukromí a osobních údajů (např. audia, videa, dotazníky)
- práva na ochranu obchodního tajemství, bezpečnosti státu nebo jiných oprávněných zájmů příjemce (např. v případě možnosti obchodního využití)

- Odůvodnit v Plánu správy dat a doporučuje se zveřejnit alespoň metadata.

INTEROPERABLE

- Making data interoperable / Interoperabilita dat

- Pro znovupoužití dat je důležité, aby data mohla být kombinována s dalšími jinými daty, aby data bylo možné otevřít, přečíst, aby data byla v takovém formátu, aby data mohla jít použít a kombinovat.
- Při mezioborovém výzkumu- využití vašich dat nebo dat někoho jiného
- Kombinování dat v různých aplikacích pro analýzu dat

- Zeptejte se kolegů, jaké oborové standardy využívají.

- Příklady ze života - PDF formát, není strojově čitelný (scan obrázku, scan ručně psaného textu)**

REUSABLE

Increase data re-use / Zvýšení opakovaného použití dat

Dokumentace

Dokumentace pomáhá pochopit strukturu a obsah samotných dat a jejich **kontext**.

dokumentace na úrovni projektu x na úrovni dat

- Jak data interpretovat (soubor README)
- Použité metadatové standardy (oborové standardy)
- Zpracování dat – popsat analýzu dat, případně metody

Příklady dokumentace:

- Pojmenovávání souborů a složek, organizace složek
- Vysvětlení zkratk, pojmů, proměnných, jednotek
- Ukládání souborů README.txt s výzkumnými daty
- Jednotné verzování souborů a dokumentace procesu manipulace s daty
- Používání Elektronických laboratorních deníků
- Přístrojová metadata (typ a nastavení přístroje, kroky měření, datum měření)

Některé z uvedených činností už běžně děláte, tak je napište do DMP.

3. Other research outputs/Další výstupy výzkumu



Další výstupy výzkumu

Popište, jaké další výstupy v rámci projektu vzniknou:

software, video, workflows, protocols, models, new materials, samples, posters, educational materials, presentations, etc.

Budou výstupy otevřené (repozitář, případně webové stránky)?

Budou uložené v repozitáři (např. Zenodo nebo v repozitáři MENDELU)?

Budete je dokumentovat a popisovat?

Budete další výstupy sdílet?

4. Allocation of resources/ Alokace zdrojů

- ❑ Pokud budete žádat o finanční podporu na Open Access a správu dat
- ❑ APC poplatky (poplatky za Open Access publikování)
- ❑ Možné náklady: úložiště (navýšení prostoru), zakoupení software ke správě dat, náklady na FAIR data (dokumentace, příprava na otevření dat, metadata)
- ❑ **Data Steward (pokud to projekt dovoluje)**
 - = Správce výzkumných dat v rámci týmu
- ❑ **jeden z členů výzkumného týmu**
 - ❑ **Role Data Stewarda v projektu:**
 - Nastavení správy dat v týmu
 - DMP
 - Dokumentace dat
 - Data quality control processes
 - Vkládání dat do repozitáře



Bezpečnost a ukládání dat během projektu -volba úložiště dat



Úložiště pro výzkumná data slouží pro tzv. „živá data“, se kterými se během výzkumu pracuje.

- Vyberte vhodné úložiště dat: <https://uvis.mendelu.cz/uloziste>
- Popište zálohování dat (úložiště univerzity, Microsoft, CESNET)
- Soubory s citlivými daty (data o pacientech) šifrujte, (osobní údaje) anonymizujte.
- Posílejte data přes **FileSender** nebo sdílejte například přes OneDrive.

Neposílejte data e-mailem!

- **Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).**
- *Before publication, the data will be pseudonymized to protect respondents' privacy according to the GDPR.*
- *Intellectual property rights: We have an industrial contact collaborating with us on a related project (ID XY). We may share data with them under our license agreement.*

Ochrana dat a ochrana soukromí subjektů.

Možné etické a právní otázky:

- Dotazníky - osobní údaje - anonymizace dat, GDPR
- Rozhovory – audia, videa - přepisy aj.
- Vzorky - lidské/zvířecí, pokusy na zvířatech (+ Oprávnění k používání pokusných zvířat od MZE)
- Geneticky modifikované organismy GMO (+ Oprávnění k nakládání s GMO od MŽP)
- Závazky vůči partnerům (komerční a průmyslové využití dat)....



7. Other issues

Ostatní

- ❑ **Strategie Open Access a Open Science na MENDELU**
- ❑ **EOSC (European Open Science Cloud)**

Příklad:

Mendel University in Brno has the Open Access and Open Science Strategy (<https://doi.mendelu.cz/pdfs/doi/9900/02/9300.pdf>), which was approved in 2022. Main goals are support for Open Access publishing and Research Data management at the university.

Mendel University in Brno has an institutional repository (<https://uvis.mendelu.cz/repozitar>), publication records are harvested to the OpenAIRE. Open Science centre at the university was established in 2023.

Research Data management will use data infrastructure within the National data infrastructure EOSC CZ when it is possible.

Co **nedělat**

NE: V projektu vzniknou **různé druhy dat** jako například tabulky, dotazníky a fotografie.

ANO: V projektu vzniknou data o velikosti 100 GB. Raw data budou zahrnovat textové soubory, také osobní data, statistická data ve formátech *.docx, *.txt a *.csv.

NE: Výzkumná data budou uložena na **sdíleném úložišti**, **zálohování** bude mít **na starosti výzkumný tým**.

ANO: Výzkumná data budou zálohována na **OneDrive od O365** (smluvně zajištěno s MENDELU). **Zálohování** má na starosti **Microsoft**. Výzkumný tým bude pravidelně zálohovat data na externím disku. Data budou na **OneDrive** uložena po dobu 10 let od ukončení projektu.

NE: Výzkumná data budou zveřejněna stejně jako publikační výstupy na **webové stránce projektu**

ANO: Výzkumná data budou zveřejněna v datovém repozitáři Zenodo pod licencí CC BY, datům bude přidělen identifikátor DOI a budou popsána metadaty XY. Data YZ nebudou zveřejněna kvůli jejich ochraně, dokud nebude podána patentová přihláška.

Jak Vám může Open Science centrum pomoci?

- Konzultace k Data Management plan
- Konzultace k Open Science v žádostech o projekt
- Open Access publikování - APC, podmínky vydavatelů
- Výběr úložiště a datového repozitáře
- Open Access fond + slevy na OA publikování
- Vložení publikací do repozitáře MENDELU

Open 
Science

open-science@mendelu.cz

Děkujeme za pozornost

Open Science centrum

Open 
Science

uvis.mendelu.cz/osc-aktuality
open-science@mendelu.cz