

Workshop: Zenodo Community

Jan Vališ ^{ID}

Ústav analytické chemie, Fakulta chemicko-inženýrská, VŠCHT Praha
& Akademické služby, NTK



Spolufinancováno
Evropskou unií



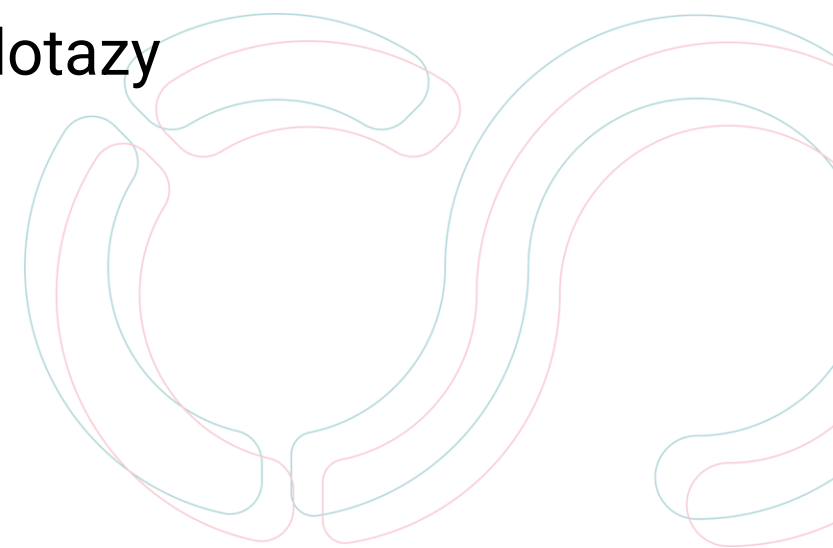
M U N I
I C S

Registrační číslo IPs EOSC-CZ

CZ.02.01.01/00/22_004/0007682

Program

- 1) Představení a rozdělení do skupin
- 2) Úvod k Zenodo
- 3) Založení a publikace záznamu, časté problémy a dotazy
Přestávka
- 4) Založení a správa komunity
- 5) Shrnutí



Zenodo

- Zadarmo
- Dublin Core & Data Cite metadata
- Implementace PID (DOI, ORCID)
- Důvěryhodné (CERN & OpenAIRE)
- Nabídka licencí
- Možnost tvorby komunit

Existují pro můj obor relevantní:
- metadatová schémata?
- repozitáře?

Jaké licence
potřebuji?

Zenodo

Je 50 GB dost?

- Zadarmo
- Dublin Core & Data Cite metadata
- Implementace PID (DOI, ORCiD)
- Důvěryhodné (CERN & OpenAIRE)
- Nabídka licencí
- Možnost tvorby komunit
- 50 GB/100 souborů na dataset
- Generická metadata
- Limitované řízení přístupu
- Generický repozitář
- Omezené kurátorství dat

Publikuje se
v mém oboru
v Zenodo?

Co, když Zenodo není řešení?

- Kdy hledat jinde?
 - Citlivá data, potřeba pokročilého řízení přístupu.
 - Specializovaná (meta)data
 - Oborové zvyklosti
 - Požadavky poskytovatele podpory
- Kde hledat?
 - re3data.org
 - FAIRSharing.org



Týká se mě to?

Dobrá praxe

Je to v souladu
s mou
zkušeností/praxí?

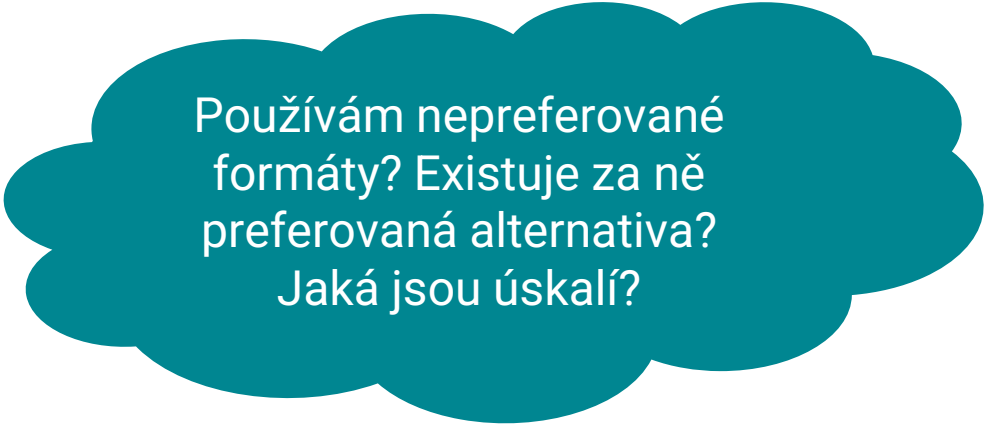
- Pojmenování souborů – jednotné, bez speciálních znaků, YYYYMMDD
- Organizace sady – jasná, mělká a dobře popsaná
- Preferované formáty
- Metadata
 - generická (název, autoři, popis, klíčová slova, související díla, finance, datum, licence)
 - oborově specifická (stacionární a mobilní fáze, teplota, ionizace; rod a druh; žánr a oblast...)
- ReadMe File
 - název a DOI, kontakt, licence, metody, pojmenování souborů, doplňující informace
 - TXT, ev. PDF/A
 - UTF-8

Časté problémy datasetů

- Absence/kvalita Read Me File
- Struktura
- Formáty
 - RAR
 - Nepreferované



S jakými problémy
se setkávám u mých
datasetů?



Používám nepreferované
formáty? Existuje za ně
preferovaná alternativa?
Jaká jsou úskalí?

Časté problémy záznamů

- Chybějící metadata
- Nevyužití/chybné využití PID (DOI, ORCiD)
- *Creator vs Contributor*
- Chybějící/nevhodná licence
- Nepropojení se souvisejícími dokumenty
- Publikace nevhodných dat (autoarchivace)
- (ne)využití embarga/omezeného přístupu
- (ne)přiřazení komunitě

S jakými problémy
se setkávám
u záznamů?



Diskuze

- Existují pro můj obor relevantní
 - metadatová schémata?
 - repozitáře - publikuje se v mém oboru v Zenodo?
- Jaké licence potřebuji?
- Je 50 GB, resp. 100 souborů dost?
- Mám citlivá data, požadavky poskytovatele/oboru, kde Zenodo nestačí?
- Používám nepreferované formáty? Existuje za ně preferovaná alternativa?
Jaká jsou úskalí?

Přestávka

Zenodo komunita

- Sdružování příbuzného obsahu
- QC/QA datasetů
- Výkaznictví
 - projekty
 - instituce
- Zvýšení dosahu/dopadu

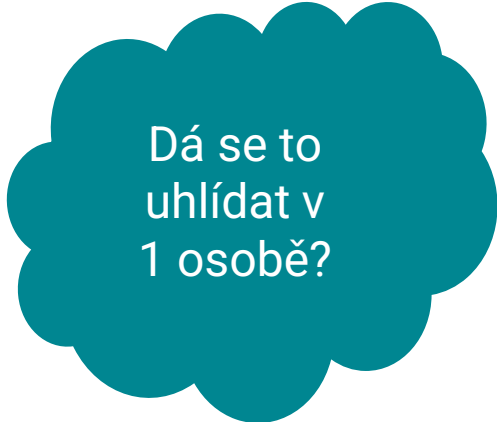
Komunita
projektu, nebo
celé instituce?

Jak moc
kontrolovat/
zasahovat
do datasetů?

*Jen data availability
statement ve článku
nebo i jiné šíření?*

Zenodo komunita

- Sdružování příbuzného obsahu
 - QC/QA datasetů
 - Výkaznictví
 - projekty
 - instituce
 - Zvýšení dosahu/dopadu
- Nelze dále dělit
 - Ruční správa
 - Omezená statistika



Dá se to
uhlídat v
1 osobě?

Zenodo komunita - problémy

- Těžko dohledatelná komunita
- Nejasné zaměření
- Dodržování/vymáhání standardů

Ví o nás někdo?

Máme jasné zaměření?
Není v tom zmatek?

Máme páku na
QC?

Diskuze

- Komunita projektu, nebo celé instituce?
- Máme jasné zaměření?
- Není v tom zmatek?
- Jak moc kontrolovat/zasahovat do datasetů?
 - Má to smysl?
 - Je to u nás přípustné? Máme na to páku?
 - Kolik nás na to je? Dá se to v 1 osobě?
- Ví o nás někdo? Jen *data availability statement* ve článku nebo i jiné šíření?

Zenodo Sandbox

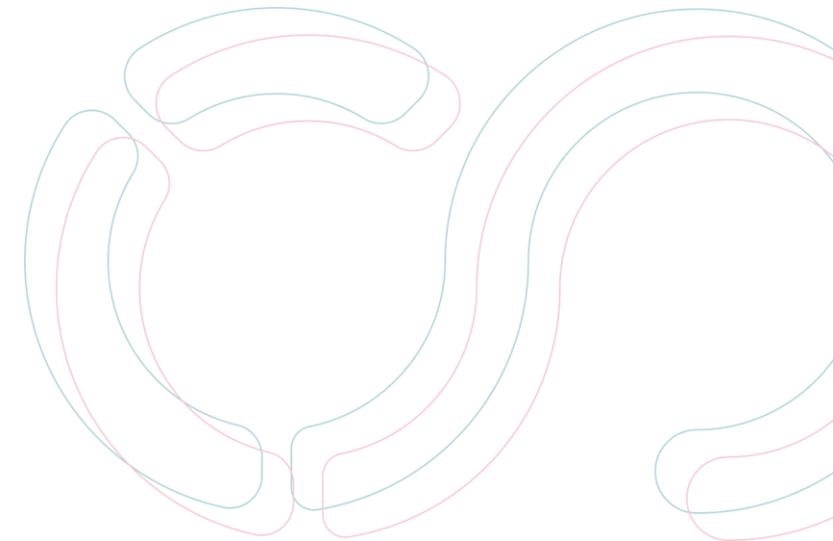
- <https://sandbox.zenodo.org>
- Přihlášení přes GitHub účet
- Nezávazná tvorba datasetů a/nebo komunit

Kde se dozvědět více?

- [RDM na webu NTK](#)
- [Manuál](#) pro přípravu datasetů pro uživatele Dataverse UIT (Norsko)
- Schätz, M., Vališ, J., Koutentaki, G., & Sklenář, A. (2024). [Paper Archiving in the Zenodo Repository and in the UCT Prague Zenodo Community](#) (v 0).
- Schätz, M., Vališ, J., Koutentaki, G., & Sklenář, A. (2024). [Datasets archiving in the Zenodo repository and in the UCT Prague Zenodo Community](#) (v 0).
- **DS Discord**

Co si odnést

- Zenodo má své limitace
- Uživatelé (a jejich data) vyžadují podporu
- Komunita může
 - vyžadovat péči
 - zvýšit dopad/dosah
 - usnadnit výkaznictví



Dotazy?

jan.valis@techlib.cz ^{id}



Spolufinancováno
Evropskou unií



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

M U N I
I C S

cesnet
.....

VŠB TECHNICKÁ
UNIVERZITA
OSTRAVA

IT4INNOVATIONS
NÁRODNÍ SUPERPOČÍTAČOVÉ
CENTRUM

Registrační číslo IPs EOSC-CZ
CZ.02.01.01/00/22_004/0007682