

Tomáš Havel

Týmový/projektový data steward



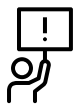
„Mám rád, když jsou data opravdu kvalitní. Když vědec otevře dataset a všechno sedí – formáty, metadata, kód.“



Čelím: nedostatečné institucionální uznání odborné role, roztříštěnost praxe, časová zátěž.



Můj cíl: dělat data prakticky použitelná, zvyšovat jejich kvalitu a přehlednost.



Potřebuji: lepší propojení výzkumníků s IT a prostor pro sdílení dobré praxe.

Náplň práce

Pomáhám výzkumníkům, aby jejich data byla kvalitní, přehledná a znovu použitelná. Pracuji přímo s týmy na fakultě nebo v projektu, kde zajišťuji technickou podporu FAIR principů – od čištění dat po nastavení repozitářů. Spolupracuji s IT a knihovnou, aby vše fungovalo efektivně a podle standardů.

Příklad z praxe

Validace a čištění dat ve výzkumném projektu

Podílím se na zpracování rozsáhlých datasetů v rámci oborového výzkumu.

Pomáhám výzkumníkům připravit data pro sdílení – kontroluji formáty, doplňuji metadata a vytvářím jednoduché skripty pro automatizaci.

Výsledek: dataset, který odpovídá principům FAIR, snadno publikovatelný a znovu využitelný pro další analýzy.

Hlavní činnosti

- Příprava a validace dat
- Čištění datasetů
- Nastavení a správa repozitářů
- Automatizace workflow
- Spolupráce s týmy na kvality dat

Nástroje a kompetence

Python
R
GitHub
OpenRefine
Power BI

- FAIR principy
- Validace formátů a tvorba pipeline
- Práce s oborovými repozitáři
- Týmová spolupráce a komunikace s IT

MŮJ PROFIL

BLÍZKOST VÝZKUMU

Administrativa

Výzkum

ZNALOST IT

MS Office

Programování

ROZSAH PŮSOBNOSTI

Tým

Instituce

ZAMĚŘENÍ ROLE

Práce s daty

Školení

Přínos / Dopad

Pro vědce:

Pomáhám připravit data, aby byla kvalitní, transformovaná, čitelná a snadno sdílená.

Pro instituci:

Podporuji automatizaci, která zlepšuje efektivitu a udržitelnost dat. Zároveň podporuji FAIR standardy.

Specialista na výzkumná data - technická a praktická podpora výzkumných týmů (patří mezi 21 % DS v ČR).

Příklad z praxe 2

Týmový/projektový data steward

Role data stewarda

- **Pomáhá** výzkumným týmům s přípravou, čištěním a validací dat.
- **Překládá** technické postupy do srozumitelných postupů pro vědce.
- **Automatizuje** opakované úkoly a nastavuje workflow pro správu dat.
- **Zajišťuje**, že data splňují FAIR principy a jsou znovu využitelná.
- **Spolupracuje** s knihovnou, IT a projektovým týmem na správě repozitářů.

Kompetence

- Znalost FAIR principů, metadat a datových formátů.
- Schopnost validovat, transformovat a čistit rozsáhlé datové sady.
- Zkušenost s verzováním dat a automatizací procesů.
- Spolupráce s výzkumníky a podpora při technických problémech.

Nástroje

Python	OpenRefine	OSF
R	Power BI	LINDAT
GitHub	Zenodo	DSpace

Validace, čištění a příprava výzkumných dat pro sdílení a znovuvyužití.

Scénář

Výzkumný tým dokončuje projekt zaměřený na sociodemografická a behaviorální data. Cílem je publikovat výsledný dataset v souladu s FAIR principy a s požadavky poskytovatele. Data však pocházejí z různých zdrojů, mají nejednotné formáty i nekompletní metadata.

Týmový/projektový datasteward pomáhá týmu vyřešit technické i organizační překážky. Pomocí Pythonu a OpenRefine připravuje skripty na čištění a kontrolu dat. Ve spolupráci s vědci nastavuje pracovní workflow v GitHubu pro sdílení kódu a verzí souborů. Zároveň kontroluje, že popisy dat splňují požadavky grantové agentury a že dataset je správně uložen v repozitáři. Tím zajišťuje, že výstupy projektu jsou přehledné, znovu využitelné a připravené pro další výzkum.

Přínos pro instituci

Zvyšuje kvalitu a transparentnost datových výstupů.



Zajišťuje, že projekty splňují FAIR principy a podmínky poskytovatelů.

Pomáhá sjednotit přístup ke správě dat napříč fakultami.

Přínos pro vědce

Usnadňuje přípravu a publikaci dat bez technických chyb.



Šetří čas pomocí automatizace a jasných postupů.

Bezpečnost, důvěryhodnost a kvalita dat zajištěna.

Poskytuje technickou i metodickou podporu přímo u výzkumu.