

Obecné doporučení pro metadatový popis výsledků výzkumu (zejména publikací a dat)

Verze 1.0 (19. 8. 2022)

Národní technická knihovna



Tento dokument vytvořila Národní technická knihovna pro příjemce finančních prostředků z projektů výzkumu a vývoje jako doporučení k vytváření metadatových záznamů při ukládání výsledků výzkumu do repozitářů.

Dokument je dostupný elektronicky: <https://doi.org/10.48813/yt6w-6h15>

Text a jednotlivá doporučení vycházejí z požadavků a doporučení Evropské Komise v [General Model Grant Agreement - contract for the Horizon Europe programme \(HORIZON\)](#), dále ze [schématu DataCite](#) pro metadatový popis výzkumných dat a dalších běžných doporučení a zvyklostí pro metadatový popis repozitářů.

Doporučení mohou být časem aktualizována v návaznosti na standardy v oblasti otevřené vědy, požadavky poskytovatelů podpory, nebo aktualizace metadatových schémat užívaných v repozitářích v ČR i v zahraničí.

Slovníček pojmů

Publikace - publikované i nepublikované textové dokumenty (např. odborné články, konferenční sborníky či příspěvky, certifikované metodiky, výzkumné zprávy aj.)

Výzkumná data - reprezentace informací shromážděných nebo generovaných za účelem zjištění nebo ověření výsledků výzkumu.

Metadata/metadatový záznam - strukturované údaje (data), popisující výzkumná data / publikace.

Repozitář - informační systém určený k dlouhodobému uchovávání a diseminaci digitálních objektů (výsledků výzkumu), tj. zajišťující uložení, ochranu, integritu, autenticitu a zpřístupnění digitálních objektů v dlouhodobém horizontu.

FAIR - definice zásad, podle kterých by (meta)data měla být vyhledatelná (findable), přístupná (accessible), čitelná a strojově zpracovatelná (interoperable) a znovu využitelná (reusable).

Perzistentní identifikátor (PID) - identifikátor používaný k trvalému a jednoznačnému určení nějaké entity: tištěného nebo elektronického zdroje (např. ISBN nebo DOI), osoby, organizace, přístroje apod.

Metadata ukládaných publikací

V repozitáři musí být metadatové záznamy v souladu se zásadami FAIR veřejně dostupné¹ a strojově čitelné a musí obsahovat minimálně tyto údaje:

- název dokumentu,
- celá jména původců (tj. autorů a dalších přispěvatelů),
- datum vydání nebo zveřejnění,
- typ dokumentu (např. článek, kniha, atd.),
- vydavatel,
- jazyk dokumentu,
- informace o dostupnosti publikací (např. časové embargo, licence² a jiné údaje o dostupnosti).

Dále se zvlášť doporučuje uvádět i další údaje jako jsou:

- trvalé identifikátory publikace (např. ISBN, ISSN, DOI, aj.),
- informace o financování (poskytovatel financí a číslo projektu),
- trvalé identifikátory osob (např. ORCID), organizací (např. ROR) apod.

Metadatové záznamy by měly obsahovat také trvalé identifikátory k dalším souvisícím výsledkům výzkumu (např. DOI pro výzkumná data) nebo k nadřazeným celkům (např. ISSN časopisu u odborného článku, ISBN knihy u kapitoly), pokud takové existují. Zároveň by měla obsahovat trvalý identifikátor/odkaz na jakékoli jiné přístroje a nástroje potřebné k ověření závěrů publikace.

Metadata ukládaných výzkumných dat

V repozitáři musí být metadatové záznamy v souladu s principy FAIR veřejně dostupné³, strojově čitelné a musí obsahovat minimálně tyto údaje:

- název datové sady,
- celá jména původců (tj. autorů a přispěvatelů),
- datum (plánovaného) zveřejnění,
- vydavatel⁴,
- popis datové sady,
- informace o dostupnosti dat (př. časové embargo, licence⁵ a jiné údaje o dostupnosti).

¹ Více informací: <https://www.go-fair.org/fair-principles/>

² Užití specifických licenčních podmínek obvykle určuje/doporučuje poskytovatel financí, vydavatel, repozitář nebo jiná autorita.

³ Více informací: <https://www.go-fair.org/fair-principles/>

⁴ Vydavatelem může být autorita odpovědná za zveřejnění dat (např. mateřská instituce výzkumníka, repozitář), nebo vydavatel časopisu, jiných publikací nebo dat.

⁵ Užití specifických licenčních podmínek obvykle určuje/doporučuje poskytovatel financí, vydavatel, repozitář nebo jiná autorita.

Dále se zvláště doporučuje uvádět i další údaje jako jsou:

- trvalé identifikátory datové sady,
- informace o financování (poskytovatel financí a číslo projektu),
- trvalé identifikátory osob, organizací apod.,
- zařazení do vědní klasifikace podle oborů (ideálně dle [FRASCATI](#)⁶, popř. jiné),
- klíčová slova.

Tyto údaje pomohou při specifickém vyhledávání a mohou zvýšit znouvupoužitelnost výzkumných dat.

Metadatové záznamy by měly opět zahrnovat trvalé identifikátory (např. ISBN, ISSN, DOI, aj.) pro souvisící publikace a další výsledky výzkumu.

Vzhledem k oborové specifičnosti a od toho se odvíjející odlišné informační potřebě důrazně doporučujeme k základnímu metadatovému popisu připojit i oborově specifické údaje, pokud to technické řešení využívaných repozitářů umožňuje.

Zdroje pro výběr (oborového) repozitáře:

- <https://www.re3data.org/>
- <https://v2.sherpa.ac.uk/opensoar>
- <https://fairsharing.org/search>

Zdroj/e pro oborové popisy:

- <https://www.dcc.ac.uk/guidance/standards/metadata>
- <https://fairsharing.org/standards/>
- <https://rd-alliance.github.io/metadata-directory/standards/>

Použité zdroje:

DATA CITE METADATA WORKING GROUP. *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs* [online]. 2021 [cit. 2022-02-23]. Version 4.4. <https://doi.org/10.14454/3w3z-sa82>

EUROPEAN COMMISSION. *Horizon Europe (HORIZON) Euratom Research and Training Programme (EURATOM): General Model Grant Agreement EIC Accelerator Contract* [online]. 15 December 2021 [cit. 2022-02-23]. Version 1.1. Dostupné z: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/agr-contr/general-mga_horizon-euratom_en.pdf

KTD: *Česká terminologická databáze knihovnictví a informační vědy (TDKIV)* [online databáze]. Praha: Národní knihovna České republiky, 2003- [cit. 2022-03-04]. Dostupné z: <https://aleph.nkp.cz/cze/ktd>

⁶ Toto třídění využívá IS VaVal, též i [prototyp repozitáře pro Národní repozitářovou platformu](#).