
Perzistentní identifikátory, jejich přínos a národní podpora v ČR

Persistent Identifiers, Their Benefits and National Support in the Czech Republic

Mgr. Hana Heringová, <https://orcid.org/0000-0002-2962-3936> / Národní centrum
pro perzistentní identifikátory, Národní technická knihovna (National Centre for
Persistent Identifiers, National Library of Technology), Technická 6, Praha 6

Mgr. Jakub Šindelář, <https://orcid.org/0000-0002-7222-6266> / Národní centrum
pro perzistentní identifikátory, Národní technická knihovna (National Centre for
Persistent Identifiers, National Library of Technology), Technická 6, Praha 6



Spolufinancováno
Evropskou unií



RESUMÉ: Text představuje fungování a přínos perzistentních identifikátorů (PID) a jejich zapojení do výzkumného prostředí v České republice. Zaměřuje se zejména na identifikátory DOI a ORCID a systém podpory jejich zavádění v ČR s podporou Národního centra pro perzistentní identifikátory v Národní technické knihovně. Předkládá také závěry analýzy nákladů a přínosů zavádění PID v českém ekosystému vědy a výzkumu a vyčíslení úspor, které může přinést.

KLÍČOVÁ SLOVA: perzistentní identifikátory, PID, DOI, ORCID, správa výzkumných informací, automatizace, úspory, národní služby, analýza nákladů a přínosů, Česká republika

SUMMARY: The text presents the functioning and benefits of persistent identifiers (PIDs) and their integration into the research environment in the Czech Republic. It focuses in particular on DOI and ORCID identifiers and the system supporting their implementation in the Czech context through the National Centre for Persistent Identifiers at the National Library of Technology. It also presents the conclusions of a cost-benefit analysis of the implementation of PIDs in the Czech science and research ecosystem and quantifies the potential savings.

KEYWORDS: persistent identifiers, PIDs, DOI, ORCID, research information management, automation, savings, national services, cost-benefit analysis, Czech Republic

Perzistentní identifikátory jsou trvalá a jednoznačná označení nejrůznějších typů entit. Mohou identifikovat jak výsledky výzkumné činnosti, jako jsou odborné články, knihy, datové sady či software, tak osoby, které se podílejí na výzkumné činnosti, výzkumné organizace, agentury financující výzkum, přidělené granty, přístroje, konference a další součásti výzkumného cyklu.

V prostředí knihoven se setkáme s identifikátory jako je ISBN (*International Standard Book Number*) anebo ISSN (*International Standard Serial Number*). V posledních dekádách se ale v digitalizovaném prostředí výzkumu objevují a prosazují další typy identifikátorů, které se postupně staly nepostradatelným základem jeho virtuální infrastruktury (de Castro et al., 2023). Soudobé digitální identifikátory ve spojení s příslušným resolverem/proxy serverem poskytují trvale funkční (proto perzistentní) odkaz na označený zdroj nebo údaje o něm. Tak perzistentní identifikátory odstraňují známou bolest internetu, tzv. „zahnívání“ odkazů (Klein, van de Sompel, 2015), způsobenou tím, že odkazovaný obsah je přesunut na jinou adresu anebo se s vývojem webové stránky adresa změní.

Bez perzistentních identifikátorů se neobejde uskutečnění principů FAIR dat – Findability (dohledatelnost), Accessibility (přístupnost), Interoperability (interoperabilita), Reusability (znovupoužitelnost) – a otevřené vědy (Evropská komise, 2018). Jedná se například o identifikátor **ORCID** (*Open Researcher and Contributor ID*) pro výzkumné pracovníky, identifikátor **ROR** (*Research Organization Registry*) **ID** pro organizace působící ve vědě a výzkumu či rozvíjející se **RAID** (*Research Activity Identifier*) pro výzkumné aktivity. Pro trvalou identifikaci nejrůznějších typů výsledků vědy a pro další účely, např. identifikaci grantů, se již zažilo využívání **DOI** (*Digital Object Identifier*). Postupně vznikají i další identifikátory, které často odpovídají na potřebu vědecké komunity přesně identifikovat například přístroje použité při výzkumu (koncept identifikátoru PIDINST) anebo odborné konference (identifikátor ConflDent). Tyto identifikátory jsou doplněny otevřeně dostupnými metadaty popisujícími základní vlastnosti příslušných entit – např. u ROR ID hlavní a variantní názvy organizace, místo, kde organizace sídlí, typ organizace apod. Mohou popisovat jejich obsah, technické či právní vlastnosti (např. licence, pod níž je dostupná datová sada) (de Castro et al., 2023).

Přínosy perzistentních identifikátorů

Meadows et al. (2019) zdůrazňují jako klíčovou vlastnost perzistentních identifikátorů jejich síťovost. Identifikátory a jejich metadata umožňují entity nejen jednoznačně identifikovat a popisovat, ale také mezi nimi vytvářet (důvěryhodná) propojení, například mezi autorem a dílem nebo grantem a výsledkem výzkumu, který byl grantem podpořen. Vzájemná propojení identifikátorů prostřednictvím metadat přináší výhody pro výzkumníky, instituce i vědeckou komunitu celkově. (Tyto přínosy jsou ale podmíněné dostatečnou mírou využití perzistentních identifikátorů v celém ekosystému komunikace vědy.)

Přínosem pro výzkumné pracovníky je zejména transparentní stanovení jejich podílu na vytvoření výsledků výzkumu, ať už se na něm podílejí v jakékoliv roli – nejen jako autoři, ale např. jako data stewardi¹ či recenzenti. Jasná a důvěryhodná propojení mezi výzkumným

1 Data steward je nově vzniklá pracovní pozice, která podporuje správu výzkumných dat. Je odpovědný za zajištění kvality, integrity, dostupnosti a bezpečnosti dat v rámci celého výzkumného cyklu. To zahrnuje správné a odpovědné nakládání s daty od jejich vytvoření a přípravy přes používání a ukládání až po archivaci a sdílení či znovuvyužití. Více viz: <https://www.eosc.cz/komunita/data-stewardi>.

pracovníkem a výsledkem jeho práce poté může usnadnit administrativu spojenou např. s vyplňováním žádostí o granty. Už dnes může výzkumník v některých systémech vyplnit pouze svůj identifikátor ORCID a jeho prostřednictvím se pak do žádosti automaticky načte jeho publikační činnost. Podobné integrace výrazně zkracují čas nutný k vyplnění žádostí či výkazů a omezují chyby, které by vznikly při manuálním přepisování identických informací.

Výzkumným organizacím zase perzistentní identifikátory výrazně usnadňují a zefektivňují správu informací o jejich výsledcích. Agenturám financujícím výzkum umožňují identifikátory mnohem snadněji sledovat a přesně vyhodnocovat dopad jimi poskytnuté finanční podpory. Propojené perzistentní identifikátory a s nimi spojená metadata obecně vytvářejí velmi hodnotný soubor dat, který může sloužit pro vytváření statistik, odhalování nových trendů a zpřesnění podkladů pro strategická rozhodnutí (Jones, 2024).

Národní podpora perzistentních identifikátorů v Národní technické knihovně

V lednu roku 2023 vzniklo jako součást Národní technické knihovny (dále NTK) Národní centrum pro perzistentní identifikátory. Chod Centra je zajišťován projektem CARDS, r. č. CZ.02.01.01/00/22_004/0004342, podpořeným MŠMT z Operačního programu Jan Amos Komenský, jehož je NTK řešitelem. Hlavním cílem Centra je podporovat zavádění perzistentních identifikátorů kompatibilních s evropskými a mezinárodními standardy v českém prostředí vědy a výzkumu. Konkrétně poskytuje finanční, administrativní a metodickou podporu zavádění identifikátorů objektů DOI a ORCID pro výzkumníky, ale i zavádění dalších typů identifikátorů jako je ROR ID pro výzkumné instituce či RAiD pro výzkumné aktivity. Součástí Centra se stalo také národní centrum ISSN, jehož provoz NTK zajišťuje již tradičně. V Centru pracují odborní koordinátoři, kteří shromažďují a organizují znalosti o perzistentních identifikátorech a poskytují metodickou podporu institucím i jednotlivcům se zájmem o jejich využívání. Centrum také slouží jako fórum pro výměnu a sdílení zkušeností institucí zejména na úrovni České republiky. Kromě toho Centrum, resp. NTK hostila v červnu roku 2024 první ročník mezinárodní konference věnující se perzistentním identifikátorům PIDfest, která do ČR přinesla i širokou škálu zkušeností ze zahraničí.

Národní centrum ORCID

Jednou ze služeb Národního centra pro perzistentní identifikátory je podpora identifikátoru ORCID. Tato služba zajišťuje prostřednictvím národního konsorcia, jehož je NTK vedoucím členem, přístup ke službám spojeným s identifikátorem výzkumníků ORCID pro české výzkumné, akademické a další neziskové instituce.

ORCID iD (*Open Researcher and Contributor Identifier*) slouží k trvalé identifikaci výzkumníků a dalších osob, kteří se jako přispěvatelé podílejí na výstupech výzkumu (vedoucí projektu, tvůrci softwaru, laboratorní technici apod.) či jsou zapojeni v publikačních procesech např. jako recenzenti. Samotný identifikátor má tvar šestnáctimístného numerického kódu a ve spojení s proxy v podobě URL – např. <https://orcid.org/0000-0003-1395-0841> – směřuje k profilu výzkumných pracovníků, který může sloužit jako životopis zachycující jejich afiliaci, publikační činnost, získané granty a další údaje o vědecké kariéře vlastníka ORCID iD. Identifikátor umožňuje především jednoznačnou identifikaci konkrétního výzkumníka v průběhu celé jeho kariéry bez ohledu na aktuální afiliaci, zemi působení či změnu jména a tím také jednoznačně přiřazení výsledkům jeho výzkumu. Vždy platí, že profil ORCID může aktualizovat pouze sám jeho vlastník. Aktualizaci profilu lze

do velké míry automatizovat – lze například do profilu načítat publikace, v jejichž meta-datech je uvedený příslušný ORCID iD – ale i s tím musí vlastník profilu vždy dát souhlas. Profil lze propojit s dalšími identifikátory vlastníka jako je *Researcher ID* v systému *Web of Science* (v originále ve tvaru *Web of Science Researcher ID*) či *Author ID* v databázích *Scopus* (*Scopus Author ID*) a uvést v něm také své kontaktní údaje.

Organizace ORCID, která systém zajišťuje, umožňuje vytváření profilů ORCID výzkumníkům zdarma. Kromě toho ale za poplatek poskytuje možnost ORCID iD prostřednictvím rozhraní API integrovat do nejrozličnějších systémů, které shromažďují informace o výzkumné činnosti a je u nich třeba jednoznačně identifikovat jednotlivé výzkumníky. Jedná se typicky například o systémy evidence publikační činnosti. V systémech, kde je ORCID integrován, je možné od sebe jasně odlišit jednotlivé výzkumníky a ze systémů zapisovat do profilů výzkumníků např. údaje o afiliaci či o autorství publikací. Institucím, které služby ORCID využívají v modelu konsorcia, ORCID nabízí tzv. prémiové služby. To znamená integraci až do pěti různých systémů a kromě toho také možnost využívat jednoduchý nástroj pro zápis afiliace do profilů výzkumníků instituce. Informace, které instituce odešle do profilu výzkumníka instituce, je na profilu poté označena jako důvěryhodná.

Tyto prémiové služby mohou využívat instituce, které jsou součástí národního konsorcia ORCID. Národní centrum ORCID zajišťuje konsorcium administrativně, poskytuje členským institucím metodické vedení a pomáhá jim řešit jakékoliv potíže při využívání služeb ORCID. Především ale členským institucím díky financování z projektu CARDS v plné výši hradí poplatky za služby ORCID, a to až do konce projektu na konci roku 2028. Do ukončení projektu CARDS má konsorcium kapacitu podpořit až 110 institucí.

V současnosti má národní konsorcium ORCID 25 členských institucí – 19 univerzit, dále pak mj. CESNET, Grantovou agenturu ČR či Knihovnu AV ČR, která v konsorciu zastupuje ústavy Akademie. Kromě členských institucí, u nichž se předpokládá, že budou služby ORCID integrovat do svých systémů, mohou metodických služeb národního centra ORCID využívat i jakékoliv další organizace. Například v roce 2023 centrum proškolovalo o identifikátoru ORCID postupně kolem stovky knihovníků a pracovníků podpory vědy zejména v univerzitních knihovnách.

Národní centrum DOI

Druhou službou Národního centra pro perzistentní identifikátory je národní centrum DOI, které v českém výzkumném ekosystému podporuje zavádění digitálního identifikátoru objektů (DOI). DOI pomáhá jednoznačně a trvale identifikovat mnoho typů objektů od digitálních (odborné články, datové sady, software) přes abstraktní entity (granty, výzkumné aktivity, odborné konference) po hmotné (výzkumné vzorky a jejich kolekce, přístroje využívané při výzkumu). Umístění objektu, resp. metadat o objektu se může v čase měnit, DOI však vždy směřuje na aktuální lokaci či informaci o ní.

Přidělování DOI mají na starosti tzv. registrační agentury. Dvě nejvýznamnější registrační agentury jsou Crossref a DataCite. Agentura Crossref vznikla v roce 1999 a její služby jsou hojně využívány vydavateli odborných publikací zejména pro články. Oproti tomu registrační agentura DataCite vznikla v roce 2009 jako odpověď na stále rostoucí zájem o výzkumná data a potřebu je jednoznačně identifikovat a odkazovat na ně. Služby DataCite našly uplatnění nejprve hlavně v prostředí digitálních repozitářů,

svůj záběr ale agentura postupně dále rozšiřuje i na přidělování identifikátorů vzorkům, přístrojům či výzkumným aktivitám.

NTK, obdobně jako v případě Národního centra ORCID, založila a vede národní konsorcium pod registrační agenturou DataCite, které na rozdíl od agentury Crossref umožňuje konsorciální čerpání služeb. V tuto chvíli (březen 2025) má konsorcium 13 členských institucí včetně čtyř univerzit, Knihovny AV ČR, sdružení CESNET ad. Veškeré poplatky za využití služeb agentury DataCite prostřednictvím národního konsorcia jsou také hrazeny Národní technickou knihovnou z projektu CARDS. Do ukončení projektu v roce 2028 služba může podpořit až 110 institucí. Národní centrum DOI uvažuje také o podpoře přidělování DOI prostřednictvím agentury Crossref pro české odborné časopisy, které zatím svým článkům DOI nepřidělují.

Podpora dalších identifikátorů

Národní centrum pro perzistentní identifikátory věnuje pozornost i dalším identifikátorům jako je ROR pro výzkumné organizace nebo RAiD pro výzkumné činnosti. Poskytuje metodickou podporu pro jejich využití a jejich širší zavádění. Další informace lze nalézt na webu *identifikatory.cz*.

Přínosy služeb

Jak bylo řečeno výše, perzistentní identifikátory mají tím větší přínos, čím více jsou využívány. Teprve od určité míry rozšíření a využívání mohou zefektivnit procesy ve vědě jako jsou publikační postupy či hodnocení vědy (Jones, 2024). Jak už dříve ukázala analýza NTK týkající se konkrétně ORCID iD (Votavová et al., 2022), v českém prostředí může být výše poplatků za služby spojené s identifikátory a náklady na jejich integraci překážkou jejich zavádění do informačních systémů ve výzkumu. Cílem služeb Národního centra je proto prostřednictvím financování z projektu CARDS přispět k maximálnímu rozšíření perzistentních identifikátorů v českém ekosystému vědy a výzkumu, a to nejen na úrovni výzkumných institucí, ale také na úrovni národních systémů, jako jsou systémy poskytovatelů financí nebo systém IS VaVal².

Analýza společnosti MoreBrains Cooperative (Jones, 2024) zpracovaná pro NTK ukázala, že po necelých 2 letech svého fungování Národní centrum nyní českým výzkumným institucím šetří přes tři miliony Kč ročně. A tato částka s rostoucím počtem institucí, jež budou součástí konsorcií, a s dalšími podporovanými identifikátory v čase ještě výrazně naroste. Díky automatizaci procesů bude odstraněno opakované manuální zadávání metadat pro tři klíčové entity (publikace, projekty a granty). Tak lze v české vědě a výzkumu potenciálně dosáhnout dalších finančních úspor v řádu milionů korun ročně, nemluvě o eliminaci chyb způsobených opakovaným manuálním zadáváním a dalších nepřímých přínosech. Investice do integrací identifikátorů do různých informačních systémů za pomoci podpory na národní úrovni přinesou vyšší a rychlejší návratnost, než kdyby české instituce zapojené do výzkumu zaváděly perzistentní identifikátory samy. Analýza vyčíslila, že časové úspory plynoucí z využívání perzistentních identifikátorů by mohly dosáhnout až 7207 člověkodní ročně, což by v efektivnějším scénáři s centrální národní podporou odpovídalo čistým úsporám přes 13 milionů korun ročně.

2 Informační systém výzkumu, vývoje a inovací (IS VaVal) shromažďuje informace o výzkumu, vývoji a inovacích podporovaných z veřejných rozpočtů v České republice a je jediným autorizovaným, úplným a závazným zdrojem těchto informací. Více viz <https://www.isvavai.cz/>.

Závěr

Národní centrum pro perzistentní identifikátory funguje něco přes dva roky. Zatím dosáhlo dílčích úspěchů v propagaci a podpoře identifikátorů v českém prostředí výzkumu, ale k plné realizaci jejich potenciálu vede ještě dlouhá cesta. Podpora identifikátorů ORCID a DOI se bude v nadcházejících letech rozšiřovat mezi další instituce a systémy. Kromě toho bude Národní centrum sledovat rozvoj dalších typů identifikátorů a podporovat podle potřeby jejich zavádění v ČR. Kromě zmíněných úspor, jichž bude dosaženo díky automatizaci (eliminace zbytečného opakovaného ručního zadávání metadat), může podpora perzistentních identifikátorů přinést také řadu hůře vyčíslitelných hodnot, jako je zvýšená viditelnost českého výzkumu ve světě, lepší přehled o fungování sektoru vědy, výzkumu a inovací a lepší naplňování principů FAIR a open science. Reálné přínosy činnosti Centra a celkový pozitivní dopad identifikátorů na české prostředí vědy a výzkumu na konci projektu CARDS v roce 2028 vyhodnotí druhá fáze analýzy nákladů a přínosů (Jones, 2024).

Národní centrum pro perzistentní identifikátory je financováno projektem CARDS, registrační číslo projektu: CZ.02.01.01/00/22_004/0004342.

Seznam literatury a elektronických zdrojů

DE CASTRO, Pablo; HERB, Ulrich; ROTHFRITZ, Laura a SCHÖPFEL, Joachim, 2023. *Building the plane as we fly it: the promise of Persistent Identifiers*. Online. Zenodo. Dostupné z: <https://doi.org/10.5281/zenodo.7258286>. [cit. 2025-02-25].

Evropská komise: Generální ředitelství pro výzkum a inovace, 2018. *Turning FAIR into reality – Final report and action plan from the European Commission expert group on FAIR data*. Online. Úřad pro publikace Evropské unie. Dostupné z: <https://data.europa.eu/doi/10.2777/1524>. [cit. 2025-02-25].

JONES, Phill; BROWN, Josh; MEADOWS, Alice a MURPHY, Fiona, 2024. *Analýza nákladů a přínosů implementace PID v České republice*. Online. Národní technická knihovna. Dostupné z: <https://doi.org/10.48813/zm93-wy56>. [cit. 2025-02-25].

KLEIN, Martin a VAN DE SOMPEL, Herbert, 2015. *Reference rot in web-based scholarly communications*. Online. LSE Impact Blog. Dostupné z: <http://blogs.lse.ac.uk/impactofsocialsciences/2015/02/05/reference-rot-in-web-based-scholarly-communication>. [cit. 2025-03-06].

MEADOWS, Alice; HAAK, Laurel a BROWN, Josh, 2019. *Persistent identifiers: the building blocks of the research information infrastructure*. Online. Insights: the UKSG journal. Roč. 32, č. 1. Dostupné z: <https://doi.org/10.1629/uksg.457>. [cit. 2025-02-28].

VOTAVOVÁ, Barbora; DIBUSZOVÁ, Eva; HNÁTKOVÁ, Eva a SVOBODA, Martin, 2022. *Analýza účelnosti a zájmu o centralizované zajišťování perzistentního identifikátoru ORCID iD*. Online. Národní technická knihovna. Dostupné z: <https://doi.org/10.48813/6erb-z009>. [cit. 2025-02-25].

HERINGOVÁ, Hana a ŠINDELÁŘ, Jakub. Perzistentní identifikátory, jejich přínos a národní podpora v ČR. Knihovna: knihovnická revue. 2025, roč. 36, č. 1, s 73–78. ISSN 1801-3252.