

Perzistentné identifikátory digitálnych informačných entít – udržateľnosť, príklady dobrej praxe a ich potenciál



Minisériu článkov o perzistentných identifikátoroch digitálnych informačných entít uzatvárame pohľadom na udržateľnosť tohto komplexného a technologicky agnostického systému, ktorý je dôverne spätý s internetom, ale zároveň reprezentuje úplne novú kvalitu komunikácie. Pre obidve skupiny však platí to, že nevyhnutnou podmienkou existencie je ich skutočné udržiavanie... Hovoríme o procese neustáleho vývoja technických aj formálnych štandardov, vzájomnej komunikácii jednotlivých prvkov systému, kvalitnej dokumentácii a „výchove“ používateľov. Tak ako web nie je len „abstraktný virtuálny systém“ bez zodpovednosti a pravidiel, tak aj všetky existujúce perzistentné identifikátory digitálnych informačných entít reprezentujú konkrétne subjekty, resp. štruktúry. Príspevok prináša výberový pohľad na príklady dobrej praxe a ukážky efektívneho spájania rôznych typov trvalých identifikátorov vo vedeckom publikačnom výstupe.

Čo bolo príčinou toho, že začali, popri štandardne budovanom a udržiavanom systéme trvalých identifikátorov v prostredí webu, vznikajú ďalšie systémy trvalej identifikácie informačných objektov? Kedy dosiahol systém URL svoje hranice? Bolo potrebné hľadať nové spôsoby saturácie požiadavky trvalej identifikácie informačného objektu v sieťovom prostredí? Kvalita manažmentu a využívania digitálnych informačných objektov sú výrazne determinované komplexnosťou a polyfunkčnosťou používaných identifikačných systémov. Systémov, ktoré majú vlastný systém správy a registrácie identifikátorov. URL požiadavku, perzistencia nespĺňa (lebo to nie je ani jeho účelom, tým je globálna jednoznačnosť v konkrétnom čase) – internetové adresy sa vyznačujú vysokou nestabilitou. Základnou podmienkou existencie a fungovania každého identifikačného systému je udržiavanie súboru pravidiel a definovanie rolí jednotlivých prvkov/subjektov. Okrem toho, do celého procesu vstupuje naplnenie podmienky, že identifikátor personifikuje digitálny dokument/objekt a nie lokalitu/sieťovú adresu, na ktorej je dokument publikovaný/sprístupňovaný. Podmienkou je teda funkčná špeciálna technická infraštruktúra. Základom tejto infraštruktúry je tzv. *resolver*, služba, ktorá funguje nad spoločnou databázou a je schopná na základe konkrétneho (registrovaného) perzistentného identifikátora presmerovať používateľa na URL, kde sa konkrétny dokument nachádza. Nakoľko systém funguje prostredníctvom protokolu HTTP (HTTPS), výsledkom je „webová stránka“. Každá takáto infraštruktúra musí mať funkčné mechanizmy, ktoré sú schopné získať vždy aktuálne platnú adresu informačnej digitálnej entity a smerovať používa-

teľa na obsah. Efektívnosť každého systému je podmienená kvalitou „údržby“; perzistencia každého identifikátora je závislá od životnosti identifikačného systému, ktorý ho pridelil a „trvalosť identifikátorov je skôr otázkou toho, ako dobre je identifikačný systém administratívne zaistený, čo predpokladá zodpovedajúce organizačné, personálne a finančné zdroje“ (Cubr et al., 2016). Na rovnakom princípe funguje aj celý komplex digitálnych knižníc, ktoré, rešpektujúc platné štandardy, budujú vlastné systémy identifikátorov.

V predchádzajúcich príspevkoch sme sústredili pozornosť najmä na systémy DOI a ORCID, dnes si v rámci príkladu dobrej praxi predstavíme aj niektoré ďalšie. Kto zabezpečuje budovanie a udržateľnosť týchto systémov perzistentných identifikátorov, aké mechanizmy relevantných organizačných, personálnych a finančných prvkov tu fungujú? Aké authority stoja za týmito systémami? Čo ponúkli súčasnej vedeckej komunikácii, resp. publikačnému prostrediu, a stali sa preto dôležitými partnermi vydavateľov vedeckej literatúry?

Perzistentné identifikátory – udržateľnosť príklady dobrej praxe

Za udržiavanie konkrétneho systému perzistentných identifikátorov (vrátane permanentného vývoja štandardov, modernizácie infraštruktúry, podporných služieb a servisu, zabezpečovania finančných zdrojov, udržateľnosti...) nesie zodpovednosť vytvorená štruktúra organizačných súčastí (registračné authority). Za dostupnosť registrovaného obsahu zodpovedá, samozrejme, jeho registrátor, nakoľko iba ten má priamy dosah na zverejnenie/umiestnenie tohto obsahu. Rešpektovanie a uplatňovanie prijatých pravidiel v praxi „stojí a padá“ v konečnom dôsledku na tvorcach/producentoch obsahu.

Nielen na úrovni osobnej (autori), ale aj na úrovni korporácií/akademických inštitúcií mobilizuje akademické prostredie svoje úsilie v rovine zlepšovania vizibility vlastného vedeckého výskumu a diseminácie jeho výsledkov, ale aj hľadania nových kontaktov /partnerov/ možností spolupráce. Systémy perzistentných identifikátorov získavajú v tomto kontexte špecifické postavenie a dostávajú novú pridanú hodnotu. Stávajú sa súčasťou „interných“ štandardov, ktoré na korporátnej úrovni formalizujú spracovanie výskumných dát a výstupov, vrátane publikačných. Ďalším z dôvodov, prečo sa takéto dokumenty pripravujú je šírenie povedomia o trvalých identifikátoroch v odbornej komunite (edukačné ciele). Spomínané formálne rámce dostávajú podobu tzv. *odporúčaní dobrej praxe (best practices)*. V zahraničí sa s takouto praxou stretávame v prostredí výskumných inštitúcií, ale aj rôznych profesionálnych formálnych aj neformálnych združení a aktivít. Na Slovensku zatiaľ takáto prax nemá oporu. Prinášame niekoľko príkladov vyššie spomínaných dokumentov.

- CEOS (Committee on Earth Observation Satellites) v rámci kolekcie Best Practices and Guidelines sprístupňuje aj dokument *CEOS Persistent Identifier Best Practices* schválený vo verzii 1.2 v januári 2017. Dokument vysvetľuje podstatu identifikátorov (všeobecne, ale konkrétne na príklade DOI), prípady a spôsoby správneho používania, význam a odporúčania (CEOS, 2017).
- CESSDA ERIC (Consortium of European Social Science Data Archives) v rámci dokumentu *CESSDA ERIC Persistent Identifier Policy Best Practice Guidelines* spracúva pokyny pre poskytovateľov služieb CESSDA s detailnou dokumentáciou. Informácie sú zamerané na podporu poskytovateľov služieb CESSDA pri implementácii globálnych perzistentných identifikátorov (CESSDA ERIC, 2017).
- Research Data Canada (RDC) v rámci Best Practices Series v roku 2016 publikovala vo verzii 1.0 dokument *Unique Identifiers: Current Landscape and Future Trends* (RDC, 2016).
- EOSC (European Open Science Cloud) v rámci spracovanej špecifikácie *PID Information Types WG final deliverable* z roku 2015 predstavuje štandardné základné typy perzistentných identifikátorov, ktoré podľa spracovateľov umožňujú jednoduché overovanie identity a integrity údajov.
- GDIF (Global Diversity Information Facility) prezentuje v dokumente pod názvom *A Beginner's Guide to Persistent Identifiers* z roku 2011 vo verzii 1.0 význam, typy a správne použitie perzistentných identifikátorov pre komunikáciu dát a výsledkov výskumu v oblasti biodiverzity (Richard set al., 2011).

Výberový prehľad približuje rôzne spôsoby prezentácie významu a podstaty perzistentných identifikátorov v rôznych kontextoch. Spracovatelia poukazujú na ich význam v kontexte otvorenej vedy, zlepšenia viditeľnosti a overiteľnosti výsledkov vedeckého výskumu, zlepšenia komunikácie a spolupráce akademickej komunity v globálnom rámci.

Ako môže vyzeráť použitie trvalých identifikátorov vo vedeckom článku?

Perzistentné identifikátory DOI a ORCID v publikačnej praxi (Prečo? Lebo...)

1. Pri registrácii DOI, bez ohľadu na druh obsahu/dokumentu, sa otvára cesta k jeho jednoduchšiemu (a teda širšiemu) citovaniu, zdieľaniu v systémoch, ktoré umožňujú využívanie rôznych metrík a v neposlednom rade tento identifikátor umožňuje indexovanie a vyhľadávanie obsahu/dokumentu na úrovni limity.
2. Pri práci s literárnymi zdrojmi/vedeckými dátami/štatistickými dátami umožňuje DOI registrovaný pre spomínaný obsah autorovi pri tvorbe rukopisu jednoduchú identifikáciu informačného prameňa alebo ďalších použitých obsahových entít (vrátane napr. aj vedeckej metódy, špecifického softvéru, ak boli popísané v rámci špecializovanej databázy a pod.).
3. Samotný autor môže ako prvý krok pred finalizáciou svojho publikačného výstupu využiť možnosť registrácie DOI pre vlastné dáta/datasety, ktoré v rámci výskumu spracoval a použil v upravenej prezentačnej forme v texte rukopisu. Následne môže takto registrovaný obsah prostredníctvom DOI zapojiť jednoducho do publikovaného materiálu. Dataset teda nemusí predstavovať iba informačný produkt druhej strany, môže to byť vlastný obsah autora.

4. Pri použití identifikátora ORCID v kontaktných údajoch autora sa čitateľ dostáva do profilu autora, kde vidí jeho odbornú/vedeckú profiláciu, publikácie, projekty (samozrejme, keď si autor tieto informácie v profile uviedol – Obr. 1a, b).

Landscape Ecol (2019) 34:911–923
<https://doi.org/10.1007/s10980-019-00815-w>



RESEARCH ARTICLE

Using soundscapes to assess biodiversity in Neotropical oil palm landscapes

Paul R. Furumo · T. Mitchell Aide

Received: 6 September 2018 / Accepted: 25 March 2019 / Published online: 6 April 2019
 © Springer Nature B.V. 2019

Abstract

Context Expanding oil palm plantations have caused widespread deforestation and biodiversity loss in Southeast Asia, stigmatizing the industry around the

Using remnant forest fragments as reference sites, we sampled land use classes that are commonly replaced by commercial oil palm including cattle pastures, rice fields, and banana plantations.

Paul Furumo

ORCID iD
<https://orcid.org/0000-0001-6499-1745>

Print view

Other IDs
 Loop profile: 596064

Works (4 of 4) Sort

Correction to: Using soundscapes to assess biodiversity in Neotropical oil palm landscapes

Landscape Ecology
 2019-05-18 | journal-article
 DOI: 10.1007/s10980-019-00833-8

Source: Crossref ★ Preferred source

Using soundscapes to assess biodiversity in Neotropical oil palm landscapes

Landscape Ecology
 2019-04-06 | journal-article
 DOI: 10.1007/s10980-019-00815-w

Source: Crossref ★ Preferred source

Telecoupling Research: The First Five Years

Sustainability
 2019-02 | journal-article
 DOI: 10.3390/su11041033

Source: Multidisciplinary Digital Publishing Institute ★ Preferred source

Obr. 1a, b Prepojenie afiličných údajov autora v článku časopisu s profilom v databáze ORCID

5. Pri použití identifikátora ORCID autorom v článku/dokumente, ktorý je výstupom projektu je v prípade potreby k dispozícii nástroj na kontrolu online prepojením identifikátora v článku s profilom na stránkach orcid.org.

Záver

Každý systém je funkčný v takom rozsahu, v akom sú funkčné jeho jednotlivé prvky. Nie je to *perpetum mobile* fungujúce pod správou *vis major*. Systémy peizistentných identifikátorov ako systémy distribuované sú existenčne závislé na serióznom prí-

stupe všetkých zúčastnených strán. Fungujú na princípe dobrovoľnej participácie. Participácie, ktorá štandardne však znamená jednotu práv a povinností. Tak, ako každý „ľudský produkt“ nesú so sebou ťarchu možných (pochybení) chýb. Pri pohľade do správ zo serverov Crossref reportujúcich registráciu obsahu (pridelovanie DOI) je jasne viditeľná chybovosť aj na strane veľkých vydavateľov. Dokonca možno povedať, čím väčší vydavateľ/väčšia produkcia, tým väčší počet chýb. Sú však mechanizmy, ktoré umožňujú tieto chyby opraviť. To však musí urobiť človek. Komunita, ktorá zabezpečuje vývoj a udržateľnosť systému DOI (štandardy, infraštruktúra...) ponúka nástroje na minimalizáciu chybovosti. Výkon (a zodpovednosť) je však na konkrétnom členovi konzorcia. Podpísaním zmluvy sa člen Crossref zaväzuje dodržiavať všetky podmienky členstva. Ale... A čo ORCID? Nástroj vo vedeckej komunikácii, ktorý má skutočný potenciál. Aká je realita? Ako ho využíva vedecká komunita? Pri vyhľadávaní v databáze možno nájsť autora/akademika, ktorý má záznamov niekoľko („rekord“ pri náhodnom vyhľadávaní je cca 11...). Je to o ľuďoch a uplatňovaní „mnohých“ metrov, ako sa s obľubou konštatuje. Autori striktné žiadajú jednoznačnú identifikáciu svojej osoby (svojho autorstva) v systémoch evidencie publikačnej činnosti a ďalších bibliografických databázach. V prípade, že majú možnosť aktívne vstúpiť do procesu vlastnej vizibility a diseminácie výsledkov svojej práce, vytráca sa motivácia. Vyžaduje ORCID vydavateľ, tak sa zaregistrujem...

Príklady dobrej praxe (nielen tých zopár uvedených v článku) naznačujú, akým smerom by sa mohli „pohnúť ľady“ aj u nás. Nemusia to byť rozsiahle kolaboráty – jednoducho, stručne, zrozumiteľne, a najmä čím skôr.

Vedecká komunikácia dostala v posledných desaťročiach k dispozícii nástroje, ktorých efektívnosť je merateľná. Systémy perzistentných identifikátorov informačných entít (najmä digitálnych) posunuli a stále majú taktiež potenciál posúvať túto komunikáciu k novej kvalite. Spolu s otvoreným prístupom umožňujú zdieľanie výsledkov vedeckého výskumu v globálnom rámci. Eticky, transparentne a možno aj „demokraticky“. (Dúfajme...).

Použitá literatúra

CEOS. 2017. *Best Practices and Guidelines* [online]. [cit. 2019-02-10]. Dostupné na <http://ceos.org/ourwork/lessons-learned-best-practices/>

CESSDA ERIC. 2017. *CESSDA ERIC Persistent Identifier Policy Best Practice Guidelines* [online]. [cit. 2019-02-10]. Dostupné na <http://dx.doi.org/10.18448/16.0041>

CUBR, Ladislav et al. 2016. Srovnání vybraných národních identifikačních systémů užívajících identifikátory URN:NBN. In *Pro-inflow: časopis pro informační vědy* [online], roč. 8, číslo 1, s. 13-53 [cit. 2019-01-20]. ISSN: 1804-2406. Dostupné na <https://doi.org/10.5817/ProIn2016-1-3>

CROSSREF. 2019. Crossref [online]. [cit. 2019-02-10]. Dostupné na <https://www.crossref.org/>

DATA CITE. 2019. *DataCite: Locate, identify, and cite research data with the leading global provider of DOIs for research data* [online]. [cit. 2019-02-10]. Dostupné na <https://datacite.org/>

DOI FOUNDATION. 2019. *The DOI® System* [online]. [cit. 2019-02-10]. Dostupné na <http://www.doi.org/>

EO. 2015 [online]. *PID Information Types WG final deliverable* [online]. [cit. 2019-02-10]. Dostupné na <https://doi.org/10.15497/FDAA09D5-5ED0-403D-B97A-2675E1EBE786>

ORCID. 2019. *ORCID: Connecting reserach and researches* [online]. [cit. 2019-02-10]. Dostupné na <https://orcid.org/>

RDC. 2016. *Unique Identifiers: Current Landscape and Future Trends* [online]. [cit. 2019-02-10]. Dostupné na <https://doi.org/10.5281/zenodo.557106>

Richards K, White R, Nicolson N & Pyle R (2011). A Beginner's Guide to Persistent Identifiers. Global Biodiversity Information Facility. Accessed at <https://www.gbif.org/document/80575>

Príspevok vznikol v rámci projektu Informačný systém výskumu a vývoja – prístupy do databáz pre potreby výskumných inštitúcií (Európsky fond regionálneho rozvoja, Operačný program Výskum a inovácie, ITMS2014+ 313011I407).

PhDr. Ľubica Jedličková, PhD.

lubica.jedlickova@uniag.sk

(Slovenská poľnohospodárska knižnica pri SPU v Nitre)