

Citovanie otvorených vedeckých dát

Matej Harvát

Odbor podpory otvorenej vedy

CVTI SR

Prečo citovať (dáta)?

Výskum a písanie práce v ideálnom prípade zahŕňa proces skúmania a učenia sa.

Uvádzaním zdrojov ukazujete čitateľovi, ako ste dospeli k svojim záverom, a uznávate prácu iných ľudí, ktorá vás k týmto záverom priviedla.

Uvádžanie zdrojov:

- dokumentuje váš výskum a vedeckú činnosť
- uznáva prácu iných, ktorých vedecké poznatky prispeli k vašej práci
- pomáha čitateľovi pochopiť kontext vašich argumentov
- poskytuje informácie, ktoré môže čitateľ použiť na vyhľadanie ďalších informácií o vašej téme.
- zvyšuje dôveryhodnosť vášho vedeckého diela
- poskytuje vám príležitosť preukázať vlastnú integritu a pochopenie akademickej etiky.



https://www.youtube.com/watch?v=jiCZKV-alC0&embeds_referring_euri=https%3A%2F%2Fwww.icpsr.umich.edu%2F&source_ve_path=OTY3MTQ&feature=emb_imp_woyt&ab_channel=ICPSR

Prečo citovať dáta?

Dôvody prečo citovať dáta:

- Zabezpečiť, aby jednotlivec alebo skupina autorov zodpovedná za generovanie údajov získala náležité uznanie.
- Pomôcť čitateľovi rozpoznať a nájsť údaje.
- Podporovať reprodukovateľnosť údajov a umožniť ostatným opakovať a overovať výsledky výskumu.
- Uľahčiť spoluprácu tým, že sa umožní ostatným výskumníkom opätovne používať a stavať na rovnakých údajoch
- Pomôcť sledovať vplyv a opätovné použitie súborov údajov (datasetov)
- Zvýšiť využiteľnosť výskumu vďaka lepšej vyhľadateľnosti súborov údajov

Dáta by sa mali citovať z rovnakých dôvodov, ako sa citujú články v časopisoch:

Vo všetkých vedeckých alebo akademických prácach je potrebné uvádzať zdroje, či už ide o rozsiahly článok alebo stručnú správu.



Citovanie a zvýšenie čítanosti

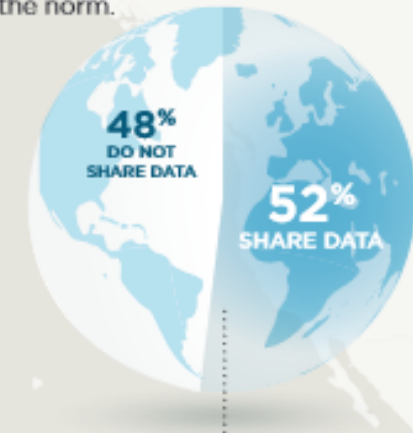
- Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B (2020) **The citation advantage of linking publications to research data.** In PLoS ONE 15(4): e0230416. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0230416>
- Piwowar HA, Vision TJ. 2013. Data reuse and the open data citation advantage. PeerJ 1:e175 <https://doi.org/10.7717/peerj.175>

PLOS, BMC - články odkazujúce na dáta v úložisku, až o 25% vyšší citačný vplyv ako je priemerná citovanosť článku po 3 rokoch od vydania.

- Vydavateľstvo Wiley: [Researcher Data Insights Survey](#), marec 2014, viac ako 2000 vedcov globálne
2/3 autorov zverejňuje dáta prostredníctvom vedeckých časopisov

GLOBAL DATA SHARING TRENDS

Data sharing practices vary widely across research fields and geographic areas. Just over half of researchers report making their data publicly available, though archiving results in repositories is not yet the norm.



WAYS DATA IS SHARED

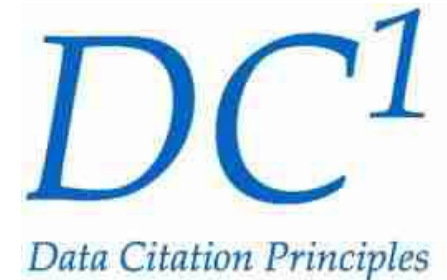
- 📄 67% As supplementary material in a journal
- 🌐 37% Personal, institutional or project webpage
- 🏛️ 26% Institutional data repository (i.e. university or institute-sponsored)
- 🗂️ 19% Discipline-specific data repository
- 🌐 6% General-purpose data repository (e.g. Dryad, figshare)
- ✓ 5% Other

Globally, researchers also report sharing their data in limited and non-permanent ways: 57% are sharing data at a conference while 42% of researchers share their data upon informal request (e.g. email, direct contact, etc.).

Zásady citovania dát

Citácie údajov obsahujú obvyklé komponenty iných citácií :

- **Autor alebo tvorca/ zostavovateľ:** subjekt/subjekty zodpovedné za vytvorenie údajov.
- **Názov:** názov súboru údajov alebo jeho stručný opis, ak názov chýba
Publisher: entity responsible for hosting the data (like a repository or archive)
- **Vydavateľ alebo distribútor:** subjekt, ktorý sprístupňuje, archivuje, publikuje a/alebo distribuuje súbor údajov
- **Dátum uverejnenia:** dátum publikovania alebo iného zverejnenia údajov
- **URL alebo najlepšie DOI:** link, ktorý odkazuje na údaje.
- **Prístup k údajom:** keďže väčšina údajov sa zverejňuje bez verzií vydania, je dôležité zaznamenať čas prístupu k údajom pre prípad, že by sa časom vydali novšie verzie.



[dataset] Authors; Year; Dataset title; Data repository or archive;
Version (if any); Persistent identifier (e.g. DOI)

Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014 <https://doi.org/10.25490/a97f-egykh>

FORCE 11 – Spoločná deklarácia princípov citovania dát:
<https://force11.org/info/joint-declaration-of-data-citation-principles-final/>

Zásady citovania dát

Citovanie dát podľa princípov TOP (Transparency and Openness Promotion): <https://osf.io/9f6gx/wiki/Guidelines/>

Návod/šablóna na 3 úrovne politík spojené s otvorenými dátami, ktoré si môže vybrať vydavateľ

Citácie článkov sú bežné a dobre formulované. Podobné štandardy možno uplatniť aj na citovanie údajov, kódu a materiálov, aby sa uznali a uznali ako pôvodné duševné príspevky.

Úroveň 1 odporúča citačné normy,

Úroveň 2 vyžaduje dodržiavanie citačných noriem a

Úroveň 3 vyžaduje a presadzuje dodržiavanie citačných noriem.

Príklady citovanie dátových súborov (datasetov):

- Barnes, Samuel H. Italian Mass Election Survey, 1968. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1992-02-16. <https://doi.org/10.3886/ICPSR07953.v1>
- Schneider, Barbara, and Waite, Linda J. The 500 Family Study [1998-2000: United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2008-06-03. <https://doi.org/10.3886/ICPSR04549.v1>
- Federal Judicial Center. Judicial District Data Book, 1983: [United States]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2006-01-18. <https://doi.org/10.3886/ICPSR08439.v1>

Trvalé identifikátory

TRVALÝ IDENTIFIKÁTOR /PID

je dlhotrvajúci digitálny odkaz na:

a/prispievateľa,

b/objekt a

c/organizáciu,

„kód, ktorý zostáva konštantný ako prostriedok na identifikáciu digitálneho objektu bez ohľadu na zmeny jeho umiestnenia na internete“

- Dlhodobé zabezpečenie tejto služby trvalých identifikátorov objektov, prispievateľov a organizácií je nevyhnutné s ohľadom na robustné stratégie správy údajov.

- DOI (Digital object identifier)
- Prostredníctvom **DataCite** = medzinárodné konzorcium členov z viac ako 20 krajín – bližšie informácie na <https://www.datacite.org/members.html>. Zahŕňa dátové centrá, knižnice, vládne agentúry, výskumné univerzity a ďalšie subjekty, ktoré majú právo registrovať špecifický vedecký obsah (objekty) a pridelovať týmto dátam perzistenté identifikátory DOI. Členovia DataCite nesú zodpovednosť za správu, uchovávanie a archiváciu dát.
- DataCite priamo DOI neprideluje, toto je úlohou jednotlivých členov.
- <https://www.icpsr.umich.edu/files/ICPSR/emailsletters/iassist.html>



Služby DataCite

- **Prideľovanie DOI**
- **Find data**
Platforma pre integrované vyhľadávanie. Bližšie informácie: <http://search.datacite.org>
- **DataCites Profile**
Poskytuje nástroje tak pre personálne použitie (personálny identifikátor ORCID), ako aj z pohľadu členov (rozhranie na správu informácií o členstve a kontaktné údaje o svojej organizácii). Bližšie informácie: <https://profiles.datacite.org>
- **Citation Formatter**
Poskytuje jednoduché rozhranie na automatické extrahovanie metadát z DOI a vytvorenie úplnej citácie. Podporuje viac ako 5.000 rôznych citačných štýlov v 45 jazykoch. Riešenie používa Citation Style Language (open source). Bližšie informácie: <http://citation.crosscite.org>
- **Štatistické prehľady**
- **A iné...**