

DELAVNICA, DEL 1

RAVNANJE Z RAZISKOVALNIMI PODATKI

Irena Vipavc Brvar

Sonja Bezjak

Maja Ojsteršek

ADP, Univerza v Ljubljani

Vsebina:

1. odpiranje raziskovalnih podatkov:
 - na mednarodni ravni
 - v Sloveniji
2. življenjski krog podatkov
3. načrtovanje ravnanja s podatki

Zakaj je treba odpreti raziskovalne podatke?

Vse več pobud in smernic za odprti dostop
do raziskovalnih podatkov!

Nekaj razlogov za odpiranje raziskovalnih podatkov:

- preverjanje analiz,
- testiranje hipotez,
- primerjava z drugimi raziskavami,
- zmanjšanje tveganja napak in goljufij,
- zmanjšanje potrebe po podvajanju dela,
- medpodročno raziskovanje,
- priznanje raziskovalnega dosežka (citiranje),
- zahteve financerjev in revij,
- ...

([BioMed Central's position statement on open data](#))

Zakaj odprti dostop do raziskovalnih podatkov?

Večja učinkovitost in boljši izkoristek nacionalnih in globalnega znanstvenega sistema.



Načela in usmeritve OECD za dostop do javno financiranih raziskovalnih podatkov (2007):

- pomaga maksimizirati raziskovalni potencial novih digitalnih tehnologij in omrežij,
- zagotavlja večji izkoristek javnih sredstev vloženih v raziskovanje,
- pomeni potencialni vir znanja, ki je potrebno pri soočanju z izzivi, ki pretijo človeštvu. (Načela OECD 2007).

Priporočilo Evropske komisije o dostopu do znanstvenih informacij in njihovem arhiviranju (7.7.2012):

„Odprt dostop je ključni dejavnik politik držav članic pri zagotavljanju odgovornih raziskav in inovacij, saj omogoča dostop do rezultatov raziskav za vse, poleg tega pa tudi spodbuja angažiranje družbe.“

O odpiranju raziskovalnih podatkov tudi v Okvirnem programu EU za raziskave in inovacije **Obzorje 2020/Horizon 2020!**



Ker večji in širši dostop do **znanstvenih objav in podatkov** pomaga:

- graditi na obstoječih raziskovalnih rezultatih,
- spodbuja sodelovanje in zmanjšuje podvojevanje truda,
- pospeši inovacije (hitreje na trgu = hitrejša rast),
- vključi državljane in družbo (transparentnost znanstvenega postopka).

(Horizon 2020, [Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020](#))

Podatkovne storitve:

- GenBank,
- UniProt,
- Protein Data Bank,
- Amadeus,
- ZACAT
- ESS,
- Eurostat,
- UK Data Service
- ...

Revije, ki s politikami zahtevajo dostop do podatkov:

Nature, PLOS Biology, PLOS Medicine, PLOS Genetics, GigaScience, MycoKeys, BioRisk, NeoBiota, Nature Conservation, Comparative Cytogenetics ...

Podatkovne revije:

CODATA Data Science Journal, Earth System Science Data, Geoscience Data Journal, Journal of Open Archeology Data, Journal of Open Public Health Data, Journal of Open Psychology Data ...

,Odpri podatki` začetek aktivnosti v Sloveniji

Ministrstvo za znanost, izobraževanje in šport financiralo CRP

ODPRTI PODATKI

priprava akcijskega načrta za vzpostavitev sistema odprtega dostopa do podatkov iz javno financiranih raziskav v Sloveniji



Cilj projekta priprava *predloga strategij in politik*, ki bo uredil področje dostopa do raziskovalnih podatkov financiranih z javnimi sredstvi



Vprašanje kdaj, v katerih primerih in kako odpreti raziskovalne podatke.

Metoda: pristop ‚od spodaj navzgor‘

22 polstrukturiranih intervjujev:

- raziskovalci, knjižničarji in predstojniki raziskovalnih ustanov oziroma oddelkov,
- fizika, biologija, gradbeništvo, zgodovina, glasba, jezikoslovje, arheologija, ekonomija, socialno delo, antropologija, medicina ...

10 inštitutov

17 raziskovalnih področij

6 fakultet

Pestrost, heterogenost, raznolikost...

Kakšne podatke ustvarjate / uporabljate?

- ✓ numerični podatki, meritve,
- ✓ rezultati numeričnih modelov,
- ✓ ekonomski modeli,
- ✓ anketni podatki,
- ✓ intervjuji,
- ✓ dnevniški zapisi,
- ✓ delovna poročila,
- ✓ slike,
- ✓ grafi,
- ✓ risbe,
- ✓ besede (besedilni korpus),
- ✓ pisna gradiva,
- ✓ statistični letopisi,
- ✓ popisi prebivalstva,
- ✓ bibliografske podatkovne zbirke,
- ✓ zvočni zapisi,
- ✓ video,
- ✓ programska oprema,
- ✓ strojna oprema,
- ✓ aplikacije,
- ✓ spektri,
- ✓ senzogrami,
- ✓ lidarski podatki,
- ✓ herbariji,
- ✓ zbirke živali,
- ✓ vzorci tkiv.

Več o tem: [Opis stanja na področju raziskovalnih podatkov v Sloveniji](#) (Štebe et. al. 2012)

Ugotovitve:

Glede dokumentiranja, hrambe in dostopa do podatkov imajo:

- raziskovalci različne navade in poglede,
- ustanove neenotna pravila in prakse.



Toda pogosto imajo **identične probleme.**

1. Pomembne ovire:

Ni politik in strategij glede ravnanja s podatki

Raziskovalci in ustanove pogosto nimajo standardiziranih postopkov dokumentiranja.



Glede **shranjevanja oziroma arhiviranja** digitalnega gradiva ravnamo po lastni presoji, vendar bi si želeli uradnih priporočil in navodil, ki bi urejala to področje.
(raziskovalec s področja arheologije, NO14)



Kot hiša nimamo nekega enotnega predpisa o tem, na kakšen način bi moral vsak raziskovalec **hraniti in dokumentirati svoje izvirne podatke**. To je bolj stvar neke metodike, metodologije dela, kot se je je raziskovalec navadil v času sodelovanja s svojim mentorjem in priprave doktorata.
(raziskovalec s področja naravoslovnih ved, NO3)

2. Pomembne ovire:

Ni kulture deljenja podatkov

Raziskovalci samovoljno odločajo kdaj, komu in zakaj bodo dali podatke, če sploh. Po principu lastništva.

Ja, majčkeno smo **ljubosumni**. To smo mi delali, za to smo porabili toliko in toliko ur in svojega dragocenega razuma in se nam zdi škoda to nekomu zastonj dati. Čeprav po drugi strani smo pa dobili plačano, to je le končni produkt projektov, ki so nam jih plačali. (znanstvenik s področja favnistike jamskih nevretenčarjev, NO22)



V obstoječih razmerah raziskovalci niso naklonjeni deljenju podatkov.



Če želimo dobiti podatke od SURS-a, ki je plačan, da samo to dela, moramo plačati. Sedaj imam ravno en CRP, kjer delamo s podatki s področja trga dela. Vsi materialni stroški bodo šli za to, da plačamo dostop do mikro podatkov na SURS-u, pa upam, da bo zadostovalo. In če se ne morejo zmeniti na državni ravni, da bi državni uslužbenci to počeli, ki so že tako ali tako plačani, ne vem, zakaj bi jaz to dala, ki porabim 150 % svojega prostega časa, da te stvari urejam in delam. **Problem je v sistemu**, sistem je narobe postavljen. (Polona Domadenik, znanstvenica s področja ekonomije)

3. Pomembne ovire:

Pravni vidiki

Odpri dostop odpira vrsto pravnih vprašanj:

- avtorske pravice,
- zasebnost, anonimnost,
- etika,
- zlorabe (kraja, napačna interpretacija) ...

Primer s področja zgodovine:

Z ARRS-jem ne pridemo skupaj, ker v bistvu ARRS spodbuja k digitalizaciji in seveda k spletni objavi, ne vidi pa problemov v ozadju, ne nazadnje tistega fizičnega dela, zbiranja soglasij. Imamo nek manjši nabor stvari, ki so **avtorsko-pravno** še **sporne** in to urejamo, ampak ARRS bi moral tudi tukaj narediti korak naproti k temu. Ali pa JAK, da z nekim uredništvom sklene pogodbe na ta način, da se potem na primer avtor ali odreče ali kako drugače. (znanstvenik s področja zgodovine, NO10)

V strahu pred kršitvami ali zlorabami se raziskovalci največkrat vedejo restriktivno.

4. Pomembne ovire:

Neenako razvita podatkovna infrastruktura in storitve

Naravoslovci in družboslovci so pogosteje vključeni v mednarodne podatkovne kroge, poznajo repozitorije, protokole, omejitve in prednosti dostopa.

- Zlasti na področjih, ki so nacionalno zamejena, se **zametki** podatkovnih storitev **še** pojavljajo:
 - Arzenal, Virtualna zakladnica nacionalne dediščine, ZRC SAZU
 - DEDI, Enciklopedija naravne in kulturne dediščine na Slovenskem,
 - SI-story, Zgodovina Slovenije,
 - Slovensko leposlovje na spletu,
 - SIDIH, Digitalna infrastruktura za humanistiko in umetnost,
 - Inštitut za biomedicinsko informatiko (IBMI)
 - Glasbenonarodopisni inštitut,
 - ARKAS, Arheološki kataster Slovenije
 - ...

Raziskovalci **nimajo časa, sredstev in znanja**
za ravnanje s podatki



*„Pomembno je, da začnemo dovolj zgodaj v življenjskem ciklu **načrtovati ravnanje s podatki**. Na ta način načrtovanje postane integralni del raziskovalne prakse, zmanjša vložek časa in finančnih stroškov in zelo poveča kakovost podatkov, ne le za druge, tudi za nas same.“*

(UK DA, Sharing Your Data)

Podatkov ne bi delil z drugimi, ker...



Poiščite **argumente za deljenje podatkov** in jih zapišite...



Razprava...

Življenjski krog podatkov



- načrtovanje raziskovanja
- **načrt ravnanja s podatki**

Življenjski krog podatkov



- zbiranje podatkov (eksperiment, opazovanje, meritev, simulacija...)
- **zajem in priprava metapodatkov**

Življenjski krog podatkov



- vnos podatkov, digitalizacija, transkript
- preverjanje veljavnosti, čiščenje podatkov
- anonimizacija
- opis podatkov

Življenjski krog podatkov



- hramba podatkov v priporočenih formatih
- na priporočenih medijih
- priprava varnostne kopije
- priprava metapodatkov in dokumentacije
- arhiviranje podatkov

Življenjski krog podatkov



- interpretiranje podatkov / izvedeni podatki
- produciranje rezultatov / publikacije

Življenjski krog podatkov



Načrt ravnanja z raziskovalnimi podatki (NRRP)

(DATA MANAGEMENT PLAN (DMP))

- načrtovati kdo, kdaj in kako bo ravnal s podatki,
- skrb za učinkovitost in kakovost v vseh fazah,
- zagotoviti etično in pravno skladnost,
- opredeliti pomen za dolgotrajno ohranjanje,
- načrtovati mesto oddaje podatkov,
- zagotoviti, da načrt ravnanja upošteva zahteve raziskovalnega področja in krovnih ter institucionalnih politik.

Načrt ravnanja z raziskovalnimi podatki (NRRP)

z NRRP so med seboj povezani različni akterji (raziskovalci, financerji, knjižničarji, podatkovna središča) in opredeljene njihove odgovornosti



NRRP je temelj kakovostnega ravnanja
z raziskovalnimi podatki

Načrt ravnanja z raziskovalnimi podatki (NRRP)

V okviru projekta Odprti podatki sta nastala predloga za dva strateška dokumenta:

- 1. Določila politike odprtega dostopa do raziskovalnih podatkov v Sloveniji*
- 2. Akcijski načrt za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov v Sloveniji*



NRRP je opredeljen kot:

- **obvezni del projektne dokumentacije**, ki jo kandidat odda, kadar zaprosi za raziskovalna sredstva iz javnih virov,
- **zagotovi tehnično in pravno ustreznost** podatkov in gradiva, pred oddajo v pooblaščenno podatkovno središče

Prednosti deljenja podatkov za raziskovalce

- 1. HRAMBA:** Podatke bomo pregledali in **varno spravili**. Dajalci podatkov lahko vselej dobijo izročeno gradivo, denimo iz oddaljenih lokacij ali če ga založijo.
- 2. DOSTOP:** Standarden opis raziskave bo **dostopen preko tiskanih in računalniških katalogov**, tako da bo vaše delo postalo znano širšemu občinstvu. Podatki iz raziskave bodo na voljo za sekundarno analizo drugim raziskovalcem, s čimer se bo povečal njihov izkoristek.
- 3. OVREDNOTENJE:** Podatki bodo skladno z merili Arhiva družboslovnih podatkov znanstveno ovrednoteni. Podatki sprejeti v hrambo bodo zavedeni v bibliografski sistem COBISS in bodo šteli k **znanstveni uspešnosti raziskovalca**. Točke oziroma reference potrebne npr. za habilitacije, napredovanje ali sodelovanje pri projektih.
- 4. CITIRANJE:** Uporabniki so dolžni **citirati avtorja** podatkov v svojih objavah, o katerih vodimo evidenco na podlagi uporabnikovega sporočila o objavi.
- 5. OMEJITEV DOSTOPA:** Avtor ima pravico za določen čas in za določene kategorije uporabnikov **omejiti razpoložljivost gradiva**, če se mu zdi to potrebno, npr. če želi prvi opraviti dodatne analize.

Povezave na dokumente:

Vse o projektu Odprti podatki

<http://www.adp.fdv.uni-lj.si/blog/projekt-odprti-podatki/#axzz2nGLVB7PA>

Predlog politik

http://www.adp.fdv.uni-lj.si/media/gradiva/politike_financerja_raziskovalni_podatki_v2_1_osnutek.docx

Predlog akcijskega načrta

http://www.adp.fdv.uni-lj.si/o_arhivu/publikacije/odpp10_akcijski_nacrt/

Poročilo Opis stanja na področju raziskovalnih podatkov v Sloveniji

Http://www.adp.fdv.uni-lj.si/o_arhivu/publikacije/odpp10_opis_stanja/

Delavnice projekta Odprti podatki

<http://www.adp.fdv.uni-lj.si/blog/projekt-odprti-podatki/delavnice/#axzz2nGLVB7PA>