

# Dijeljenje podataka: Početnica Mreže za ponovljivu znanost Ujedinjenog Kraljevstva (*United Kingdom Reproducibility Network (UKRN)*)



Autori (Z-A): [John Towse](#), [Sally Rumsey](#), [Nicholas Owen](#), [Patrick Langford](#), [Matt Jaquier](#) [glavni autor], [Cylcia Bolibaugh](#).

Prijevod (Z-A): Dora Pejdo

## Što je dijeljenje podataka?

Dijeljenje podataka je proces stavljanja [podataka istraživanja](#) na raspolaganje za širu diseminaciju. *Findable, Accessible, Interoperable, Reusable (FAIR)* [principi](#) nalažu da podaci i meta-podaci trebaju biti: dostupni (laki za pronalaženje, i za ljude i za računala); pristupačni (popraćeni jasnim uputama za pristup i autorizaciju); interoperabilni (kompatibilan s drugim podacima i/ili alatima); i ponovno iskoristivi (prikladno opisani kako bi se omogućila daljnja upotreba).

## Zašto bismo trebali dijeliti podatke?

Mnogo je pozitivnih razloga za dijeljenje podataka. Ono otvara mogućnosti za znanstvena istraživanja promicanjem inovacija putem novih upotreba podataka i novih suradnji. Povećava transparentnost i pomaže u osiguravanju pouzdanosti znanstvenog zapisa.

### ***Mogu li ljudi pronaći greške u mom radu?***

Ponekad se greške pronađu u javno dijeljenom radu, ali pronalaženje grešaka je dobro – to znači da više ne vjerujemo u pogrešne stvari. Svatko griješi, a [stvarne pogreške vjerojatno neće naštetiti vašem radu ili karijeri](#).

### ***Mogu li ljudi koristiti otvorene podatke za „ukrasti“ moje istraživanje koje provodim?***

„Krađa“ istraživanja se gotovo nikad ne događa, ali ako vas to brine, možete staviti *embargo* na svoje podatke tako da im samo vi možete pristupiti na nekoliko mjeseci ili godina.

## Kako mogu dijeliti podatke?

### ***Što se može dijeliti s kim?***

Svako polje ima svoje običaje, izazove i rješenja. Savjeti u ovom odjeljku su općeniti i usredotočeni su na ljudske podatke (ali također mogu biti relevantni kada, na primjer, postoje komercijalne, sigurnosne ili pravne implikacije). Važno je planirati dijeljenje istraživačkih podataka od početka vašeg projekta. Većina podataka o ljudskim istraživanjima može se dijeliti etički i zakonito uz odgovarajuću upotrebu informiranog **pristanka, anonimizacije, upravljanja pravima i kontrole pristupa**. Unutar jednog istraživačkog projekta, neki aspekti podataka mogu se potpuno otvoreno dijeliti, dok će druge možda trebati zaštititi (npr. ugovorom s krajnjim korisnikom).

### **Ljudski podaci**

Zakoni o zaštiti podataka (npr. GDPR) definiraju obradu [osobnih podataka](#), ali se ne primjenjuju na anonimizirane podatke, a postoje i neki [izuzeci](#) za znanstvene studije. Imajte

na umu da se neki naizgled [anonimizirani podaci](#) mogu [ponovno identificirati](#). Vaša ustanova ili financijer može imati dodatna pravila u vezi s upravljanjem podacima (npr. zahtijevati plan [upravljanja podacima](#))

### Suglasnost

Obavijestite sudionike o tome koji će podaci biti pohranjeni, kako će se čuvati, kako će se vjerojatno dugoročno koristiti i kako će se zaštititi povjerljivost. Podatkovna služba Ujedinjenog Kraljevstva ima banku [primjera obrazaca pogodnih za dijeljenje](#) različitih vrsta podataka. Osigurajte da obrasci pristanka **ne obećavaju** da ćete uništiti podatke ili da će se podaci vidjeti ili biti pristupljeni samo od strane istraživačkog tima.

### Anonimizacija

Razmotrite postupke potrebne za [anonimizaciju](#) na početku vašeg projekta, posebno u vidu vrste pristanka koju ćete tražiti od sudionika.

Radite s anonimiziranim podacima kad god je to moguće. To smanjuje rizik od curenja osobnih podataka i osigurava da anonimizacija ne „iskrivi“ podatke. Uklonite izravne i neizravne identifikatore iz [kvantitativnih](#) i [kvalitativnih](#) podataka.

Gdje se podaci ne mogu dovoljno de-identificirati, razmislite o nametanju kontrole pristupa.

### Kontrola pristupa

Prilikom upravljanja osjetljivim podacima, osiguranje ili zaštita podskupova podataka može biti prikladna. Većina velikih repozitorija dopušta upravljanje pristupom kroz [Licencu za krajnjeg korisnika](#) koja propisuje uvjete za pristup te podrazumijeva dogovor da se ne pokušava ponovno identificirati pojedince.

### Sami podaci

Kako bi se zajednički podaci mogli učinkovito koristiti, oni bi trebali biti [razumljivi ljudima i strojevima](#), što podrazumijeva usklađenost sa određenim standardima. Standardi [metapodataka](#) olakšavaju dugoročnu upotrebu i integraciju; neki se također primjenjuju [na određene vrste podataka](#). *Digital Curation Centre (DCC)* vodi [popis standarda](#) koje možete provjeriti. [Roche i suradnici \(2015.\)](#) daju nekoliko općenito primjenjivih preporuka za izradu podataka koji su potpuni i ponovljivi, uključujući:

- Koristite jasne i detaljne opise [podataka i metapodataka](#) kako bi povećali upotrebljivost. Posljedično, tamo gdje se podaci odnose na popratni istraživački rad, trebalo bi transparentno izvijestiti koji se dijelovi skupa podataka odnose na koje tvrdnje u radu.
- Pokušajte podatke učiniti dostupnima prema vlastitim uvjetima, neovisno o drugom radu, posebno kada se mogu pružiti dodatni podaci.
- Osigurajte da podaci koji nedostaju ili koji su izuzeti budu u potpunosti anotirani.
- Pružite što više sirovih podataka kako bi korisnici mogli "ponovno izgraditi" informacije.
- Dijelite kod za analizu i skripte za obradu podataka tamo gdje je to potrebno/moguće.
- Izbjegavajte postavljanje formata/vrste datoteka u vlasništvu tvrtki, organizacija ili pojedinaca kao jedinog načina pristupa podacima. Gdje je to prikladno, pretvorite datoteke u otvorene ili standardne formate prije pohranjivanja podataka u repozitorij.
- Razmislite o tome da zamolite prijatelje i kolege da pregledaju predaju podataka koje pripremate kako biste osigurali kontrolu kvalitete, osobito ako skup podataka neće biti pregledan od strane kolega.
- Razmislite o tome da zamolite prijatelje i kolege da pregledaju predaju podataka koje pripremate kako biste osigurali kontrolu kvalitete, osobito ako skup podataka neće prolaziti proces recenzije.

### Dijeljenje podataka

1. [Pronađite repozitorij](#) koji odgovara vašem skupu podataka. Provjerite ima li repozitorij mogućnosti koje su vam potrebne: Je li prikladniji opći ili specifični repozitorij? Koja su ograničenja veličine datoteke? Koliko će dugo podaci biti dostupni? Trebate li upravljati dopuštenjima za dijeljenje ili postaviti razdoblje embarga? Hoćete li dobiti [DOI](#) ili sličan

[trajni identifikator](#)? Provjerite uvjete i odredbe o obvezi zadržavanja vaših podataka. Možete pretraživati popis repozitorija po disciplinama na <https://www.re3data.org/>.

2. Pružite smjernice za ponovnu upotrebu: [odlučite o odgovarajućoj licenci](#) za svoje podatke; budite jasni o tome kako biste željeli citirati navedeni skup podataka. Neki repozitorij će to učiniti automatski, dok će vam drugi omogućiti da sve prilagodite.
3. Podijelite trajni URL! Obavijestite ljude o resursu koji ste upravo stvorili – u radu, na konferenciji, na društvenim medijima.

*Podaci trebaju biti što je moguće otvoreniji i koliko je potrebno zatvoreni.*

## Drugi resursi

- Provjerite dostupnost institucijske pomoći (npr. knjižnica) za savjet, smjernice i iskustva.
- **Digital Curation Centre**: resursi i informacije <http://www.dcc.ac.uk/resources>.
- **FAIRsharing**: resursi o standardima za podatke i metapodatek <https://fairsharing.org/>.
- **MANTRA**: besplatni online tečaj na temu digitalnih podataka za istraživače <https://mantra.edina.ac.uk/>.
- **UK Data Archive**: socijalni, ekonomski i populacijski podaci <https://www.dataarchive.ac.uk/>.
- **Caldicott Principles**: za dijeljenje biomedicinskih podataka <https://www.ukcgc.uk/manual/principles>.
- **OpenAIRE**: resursi o etičkim i pravnim pitanjima <https://www.openaire.eu/guides>
- **FOSTER Open Science**: besplatni tečaj o zaštiti podataka i etici <https://www.fosteropenscience.eu/learning/data-protection-and-ethics/#/>.
- **Cessda Training**: vodič za upravljanje podacima <https://www.cessda.eu/Training/TrainingResources/Library/Data-Management-Expert-Guide>.

Recenzirali Chris Gorgolewski (0000-0003-3321-7583), Olivier Klein (0000-0003-2737-8049), Frederick Verbruggen (0000-0002-7958-0719), i Tony Weidberg (0000-0002-5158-307X). Uredile Emma L Henderson (0000-0002-5396-2321) i Jackie Thompson (0000-0003-2851-3636). Prevela Dora Pejdo.

Izjava o sukobu interesa: Recenzent Chris Gorgolewski radi za [Google Dataset Search](#).



Ovaj rad je licenciran pod [Creative Commons Attribution-ShareAlike 4.0 međunarodnom licencom](#).