Istraživački podaci - što s njima? : priručnik o upravljanju istraživačkim podacima

Celjak, Draženko; Dorotić Malič, Ivana; Matijević, Marta; Poljak, Ljiljana; Posavec, Kristina; Turk, Ivana

Educational content / Obrazovni sadržaj

Publication status / Verzija rada: **Accepted version / Završna verzija rukopisa prihvaćena za objavljivanje (postprint)**

Publication year / Godina izdavanja: **2020**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:102:810090

Rights / Prava: Attribution-NonCommercial-ShareAlike 4.0 International/Imenovanje-Nekomercijalno-Dijeli pod istim uvjetima 4.0 međunarodna

Download date / Datum preuzimanja: **2024-03-29**

srce
Sveučilište u Zagrebu
Sveučilišni računski centar

Repository / Repozitorij:

Digital repository of the University Computing Centre (SRCE)

dabar
DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

ISTRAŽIVAČKI PODACI – ŠTO S NJIMA?

Priručnik o upravljanju istraživačkim podacima



This work was developed as part of the Research Data Alliance (RDA) within the activities of Croatian national node "*SRCE - University of Zagreb, University Computing Centre"* and we acknowledge the support provided by the RDA community and structures.

Impressum

Autori:

Draženko Celjak, Sveučilište u Zagrebu, Sveučilišni računski centar

Ivana Dorotić Malič, Sveučilišna knjižnica Rijeka

Marta Matijević, Nacionalna i sveučilišna knjižnica u Zagrebu

Ljiljana Poljak, Sveučilišna knjižnica u Splitu

Kristina Posavec, Sveučilište u Zagrebu, Sveučilišni računski centar

Ivana Turk, Gradska i sveučilišna knjižnica Osijek

Lektorica:

Mia Kožul

Izdavač:

Sveučilište u Zagrebu, Sveučilišni računski centar

Josipa Marohnića 5

10000 Zagreb

https://www.srce.unizg.hr/

Verzija: 1.0

Mjesto i godina izdavanja: Zagreb, 14. rujna 2020.

ISBN: 978-953-8172-56-4

Ovo djelo dano je na korištenje pod licencom *Creative Commons Imenovanje-Nekomercijalno-Dijeli pod istim uvjetima 4.0 međunarodna*. Licenca je dostupna na stranici: http://creativecommons.org/licenses/by-nc-sa/4.0/deed.hr.

Projekt RDA Europe 4.0 –The European Plug-in to the global Research Data Alliance financira se u okviru programa EU za istraživanje i inovacije Obzor 2020 pod ugovorom o dodjeli bespovratnih sredstava br. 777388. Sadržaj ovoga dokumenta ne predstavlja mišljenje Europske komisije te Europska komisija nije odgovorna za bilo kakvu upotrebu takvoga sadržaja.

Sadržaj

1. Uvod 4
2. Što su istraživački podaci? 4
2.1. Zašto dijeliti istraživačke podatke? 4
2.2. Što su FAIR principi i kako dijeliti istraživačke podatke? 6
3. Životni ciklus istraživačkih podataka 8
4. Prije početka istraživanja *(životni ciklus: planiranje)* 9
4.1 Plan upravljanja istraživačkim podacima 9
5.Tijekom istraživanja *(životni ciklus: prikupljanje, obrada, analiza)* 11
5.1 Imenovanje i organiziranje istraživačkih podataka 11
5.2 Verzioniranje istraživačkih podataka 14
5.3 Smjernice za odabir formata za trajnu pohranu podataka 15
5.4 Dokumentacija i metapodaci 16
5.5 Pouzdano čuvanje podataka tijekom istraživanja 17
6. Nakon istraživanja *(životni ciklus: trajna pohrana, dijeljenje, korištenje)* 21
6.1 Objava osobnih i osjetljivih podataka 21
6.2 Informirana privola 22
6.3 Osobni identifikatori sudionika istraživanja 22
6.4 Anonimizacija i pseudonimizacija 23
6.5 Definiranje prava korištenja skupa podataka 26
6.5.1 Što podliježe autorskom pravu 27
6.5.2 Vrste licencija 27
6.5.3 Koje licencije koristiti za istraživačke podatke? 28
6.6 Trajna pohrana 29
6.6.1 Dabar — nacionalna infrastruktura za digitalne repozitorije 29
6.6.2 Preporuke za odabir repozitorija 31
6.6.3 Istraživački podaci kao (dodatak) rad(u) u časopisu 31
6.7 Citiranje istraživačkih podataka 31
6.7.1 Smještanje citata 31
6.7.2 Elementi i primjeri citata 32
7. Bibliografija 34
Pojmovnik 38

Popis tablica 40

Popis slikovnih materija 40

1. Uvod

Research Data Alliance (RDA) je globalna organizacija usmjerena na razvoj društvenih i tehničkih uvjeta za otvoreno dijeljenje i ponovnu upotrebu podataka te poticanje inovacije temeljene na podacima. U okviru projekta RDA Europe 4.0 u Srcu je 2019. uspostavljen nacionalni RDA čvor (https://www.srce.unizg.hr/rda) te je u suradnji s djelatnicima Sveučilišne knjižnice Rijeka, Sveučilišne knjižnice u Splitu, Gradske i sveučilišne knjižnice Osijek te Nacionalne i sveučilišne knjižnice u Zagrebu izrađen ovaj priručnik za istraživače koji donosi najbolje prakse upravljanja istraživačkim podacima, a u svrhu njihovog učinkovitog dijeljenja te ponovne upotrebe.

2. Što su istraživački podaci?

Istraživački podaci su svi prikupljeni, zabilježeni ili generirani podaci u svrhu analize kako bi se došlo do novih, originalnih rezultata istraživanja.[1] Istraživački podaci mogu biti:

1. sirovi (inicijalno prikupljeni)
2. očišćeni (pripremljeni za analizu)
3. obrađeni (podaci koji su rezultat provedene analize)
4. prezentacijski (verzija podataka prilagođena prezentaciji).

2.1. Zašto dijeliti istraživačke podatke?

Istraživački ekosustav postaje sve kompleksniji iz čega proizlazi potreba za boljom suradnjom unutar pojedinih istraživačkih zajednica, interdisciplinarno, ali i uključivanje javnosti (engl. *citizen science*). Sve to nameće potrebu za dijeljenjem istraživačkih podataka o čemu treba voditi brigu ne samo na kraju već i tijekom cijelog istraživačkog procesa. Dijeljenje istraživačkih podataka:

- pridonosi reproducibilnosti i transparentnosti znanstvenih istraživanja
    - Pohranjeni, očuvani i objavljeni podaci mogu se koristiti za validaciju rezultata istraživanja, sekundarne analize ili postavljanje novih hipoteza.
- povećava utjecaj znanstveno-istraživačkoga rada i citiranost

[1] University of Edinburgh, Research Data Service: Our definitions. URL: https://www.ed.ac.uk/information-services/research-support/research-data-service/after/data-repository/definitions (2020-08-28)

- Prema nekim istraživanjima, publikacije koje imaju objavljene istraživačke podatke ostvaruju i do 30 % više citata od onih kojima istraživački podaci nisu objavljeni[2].
- Otvara se mogućnost suradnje na drugim istraživačkim projektima.

- ispunjava zahtjeve financijera istraživanja
  - Financijeri znanstvenih istraživanja sve češće zastupaju stav kako bi publikacije i istraživački podaci u istraživanjima financiranim javnim novcem trebali biti objavljeni i dostupni znanstvenoj zajednici i javnosti.
  - Europska komisija za istraživanja financirana u okviru programa Obzor 2020 propisuje *Open Data Pilot*[3] za objavu istraživačkih podataka u otvorenom pristupu.
  - U okviru novoga programa Horizon Europe (2021. – 2027.)[4] otvorena znanost bit će standard na razini cijeloga programa što znači obavezan otvoreni pristup publikacijama te osiguran otvoreni pristup istraživačkim podacima u skladu s načelom „otvoreni koliko je to moguće, zatvoreni koliko je to potrebno" te obavezan plan upravljanja podacima u skladu s FAIR principima.
  - Hrvatska zaklada za znanost najavljuje uvođenje obaveznoga Plana upravljanja istraživačkim podacima u skladu s FAIR principima kao dio projektne prijave za istraživačke i uspostavne istraživačke projekte.
- omogućava njihovo korištenje u obrazovne svrhe za podučavanje studenata
- potiče napredak u znanosti, doprinosi bržoj primjeni znanosti u gospodarstvu te razvoju društva u cjelini.

[2] Piwowar HA, Vision TJ. 2013. Data reuse and the open data citation advantage. PeerJ 1:e175 https://doi.org/10.7717/peerj.175 (2020-08-28)

[3] H2020 Programme: Guidelines to the rules on open access to scientific publications and open access to research data in Horizon 2020 URL: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf (2020-08-28)

[4] Horizon Europe – sljedeći program EU-a za istraživanje i inovacije (2021. — 2027.) URL: https://ec.europa.eu/info/sites/info/files/research_and_innovation/strategy_on_research_and_innovation/presentations/horizon_europe_hr_oblikovanje_nase_buducnosti.pdf (2020-08-28)

2.2. Što su FAIR principi i kako dijeliti istraživačke podatke?

FAIR je akronim četiriju temeljnih principa:

- ***F**indable* (pronalažljivost)
- ***A**ccessible* (dostupnost)
- ***I**nteroperable* (interoperabilnost)
- ***R**eusable* (ponovna upotreba).

Navedeni principi osigurat će lakše pronalaženje istraživačkih podataka, nesmetan pristup podacima i metapodacima, korištenje uobičajenih standarda i formata kako bi se osigurala mogućnost razmjene podataka među raznim sustavima te pripremu podataka kako bi bili razumljivi i iskoristivi za ponovnu upotrebu.

Tablica 1. FAIR podaci ili Kako svoje istraživačke podatke uskladiti s FAIR principima?

Pronalažljivost (engl. *Findable*): lakše pronalaženje istraživačkih podataka	
	(Meta)podacima se dodjeljuju jedinstveni trajni identifikatori, npr. DOI, URN:NBN (koristi se u repozitorijima u Dabru), ORCID.
	Podaci su opisani bogatim metapodacima koji olakšavaju pronalaženje i citiranje: npr. podaci o autorima, istraživačkom projektu, korištenim protokolima, mjernim uređajima. U repozitorijima u Dabru za metapodatkovni opis koristi se univerzalna shema metapodataka DataCite.
	Podaci su pohranjeni u repozitorij koji ima sučelje za pretraživanje.
	(Meta)podaci su indeksirani u tražilicama – repozitoriji u Dabru indeksirani su u tražilicama, npr. Google Scholar, OpenAIRE, re3data.
Dostupnost (engl. *Accessible*): nesmetan pristup podacima i metapodacima5	
	Metapodaci su dohvatljivi putem identifikatora korištenjem standardnih i otvorenih protokola koji između ostalog omogućuju autentikaciju i autorizaciju.
	Metapodaci su dostupni čak i kad istraživački podaci nisu.
Interoperabilnost (engl. *Interoperable*): korištenje standarda kako bi se osigurala mogućnost razmjene podataka među raznim sustavima	
	Podaci moraju biti u standardnim i preporučenim formatima za dugotrajno čuvanje koji omogućuju kombiniranje i ponovnu upotrebu.
	Za opisivanje podataka koriste se standardni skupovi metapodataka karakteristični za disciplinu.

[5] FAIR, odnosno nesmetan pristup ne znači da podaci moraju biti otvoreno dostupni.

	Gdje god je moguće koriste se kontrolirani rječnici, ontologije i tezaurusi s jedinstvenim trajnim identifikatorima.
Ponovna upotreba (engl. *Reusable*): omogućiti razumijevanje i ponovno korištenje skupa podataka	
	Potrebno je osigurati detaljnu dokumentaciju o prikupljanju, analizi i obradi podataka (laboratorijske uvjete, naziv i verziju korištenog softvera...) kako bi ih drugi mogli razumjeti.
	Potrebno je slijediti najbolje prakse i preporuke za dijeljenje i čuvanje podataka za pojedina područja znanosti.
	Podacima je potrebno dodijeliti licenciju koja jasno određuje uvjete korištenja podataka. Trebalo bi koristiti ljudski i strojno čitljive licencije, kao što su npr. *Creative Commons* licencije.

3. Životni ciklus istraživačkih podataka

Slika prikazuje model životnoga ciklusa istraživačkih podataka. Da bi podaci ostali razumljivi i ponovno upotrebljivi kako tijekom istraživanja tako i nakon istraživanja, potrebno je njima upravljati tijekom svih faza životnoga ciklusa.



Slika 1. Grafički prikaz životnoga ciklusa podataka

Priručnik je tematski podijeljen prema životnom ciklusu istraživačkih podataka koji prati tijek istraživačkoga procesa u tri dijela: prije, tijekom i nakon istraživanja. Treće poglavlje obuhvaća dio planiranja upravljanja istraživačkim podacima. U četvrtom poglavlju, tijekom istraživanja, prikazane su najbolje prakse za rad s podacima u fazama prikupljanja, obrade i analize. Posljednji dio priručnika, nakon istraživanja, donosi pregled najvažnijih preporuka vezanih za trajnu pohranu, dijeljenje i korištenje istraživačkih podataka.

4. Prije početka istraživanja *(životni ciklus: planiranje)*

Prije početka istraživanja potrebno je planirati kako će se upravljati istraživačkim podacima (engl. *Research Data Managment*, RDM). RDM se odnosi na upravljanje podacima koji su nastali tijekom životnog ciklusa istraživanja te je sastavni dio istraživačkog procesa koji istraživačima pomaže da podaci budu organizirani, opisani, pohranjeni i dijeljeni na odgovarajući način.[6]

U okviru planiranja istraživanja potrebno je planirati i upravljanje istraživačkim podacima što uvelike olakšava izrada dokumenta Plan upravljanja istraživačkim podacima (engl. *Data Managment Plan*, DMP).

4.1 Plan upravljanja istraživačkim podacima

DMP obuhvaća definiranje važnih aspekata upravljanja istraživačkim podacima koji osiguravaju da će podaci biti očuvani i iskoristivi i nakon završetka projekta. Izrada DMP-a odnosi se na proces organizacije, planiranja procesa istraživanja, prikupljanja, obrade, pohrane i distribucije podataka u skladu s FAIR principima.

Iako je DMP potrebno napisati prije početka istraživanja, važno je znati kako je DMP "živi" dokument koji će se tijekom cijelog životnog ciklusa istraživačkih podataka mijenjati i nadopunjavati ovisno o tijeku i provedbi istraživanja. Tablica 2. prikazuje smjernice o kojima istraživač treba voditi računa prilikom izrade DMP-a, a nastala je na temelju preporuka Švicarske nacionalne zaklade za znanost, Europske komisije i UK Data Service.

Primjere DMP-a moguće je pronaći na mrežnim stranicama *Digital Curation Centrea* i švicarskoga *Data Life-Cycle Managementa* (DLCM).

[6] Research Data Management Definition available at: https://libguides.depaul.edu/c.php?g=620925&p=4324498

Tablica 2. Elementi DMP-a

Podaci i dokumentacija
• Koji podaci će biti prikupljani, obrađeni i izrađeni (vrsta podataka, osjetljivi i osobni podaci). • Koja metodologija i standardi će se primjenjivati. • Kako će podaci biti analizirani. • Sustavno revidiranje plana upravljanja tijekom cijelog istraživanja kao sastavni dio praćenja napretka istraživanja.
Etika, pravo i sigurnost
• Poznavanje zakonskih, etičkih i drugih obveza vezanih za istraživačke podatke, sudionike u istraživanju, ispitanike, institucije i financijere. • Usklađenost istraživanja i istraživačkih interesa s financijerima projekta te zahtjevima organizacije u sklopu koje se projekt provodi. • Postojanje potrebe za provođenjem postupaka anonimizacije i/ili pseudonimizacije. • Izrada privole o suglasnosti za sudjelovanjem u istraživanju. • Izrada GDPR izjave.
Pohrana i očuvanje
• Način na koji će se podaci pohranjivati i dijeliti tijekom projekta (lokalno na računalu, USB memoriji ili u računalnom oblaku, enkripcija podataka, izrada sigurnosnih kopija i sl.). • Mjesto na kojem će podaci biti trajno pohranjeni i objavljeni nakon završetka istraživanja/projekta (odabir odgovarajućega repozitorija za trajnu pohranu podataka).
Dijeljenje i ponovna upotreba
• Kako će se upravljati istraživačkim podacima tijekom i po završetku projekta. • Hoće li se podaci dijeliti i pod kojim uvjetima (otvoreni pristup, ograničeni pristup, embargo itd.). • Na koji način će se podaci moći koristiti, odnosno pod kojom licencijom će biti objavljeni.

POPIS DOSTUPNIH ALATA

- **Argos** (https://argos.openaire.eu/) — online alat za izradu DMP-a
- **Data Management Skillbuilding Hub** (https://dataoneorg.github.io/Education/bestpractices/) — baza podataka koja pruža preporuke za rad s podacima te pokriva svaki stadij životnog ciklusa podataka
- **DMPonline** (https://dmponline.dcc.ac.uk/) — online alat za izradu DMP-a.

5.Tijekom istraživanja *(životni ciklus: prikupljanje, obrada, analiza)*

Poglavlje se bavi fazama životnoga ciklusa istraživanja koje se odnose na prikupljanje, obradu, analizu i zaštitu istraživačkih podataka. Unutar navedenih faza posebna pozornost posvetit će se pravilima vezanim za imenovanje i organizaciju istraživačkih podataka, kontroliranje različitih verzija istraživačkih podataka, poželjne formate datoteka za pohranu istraživačkih podataka, izradu dokumentacije i metapodatkovnog opisa te zaštitu podataka.

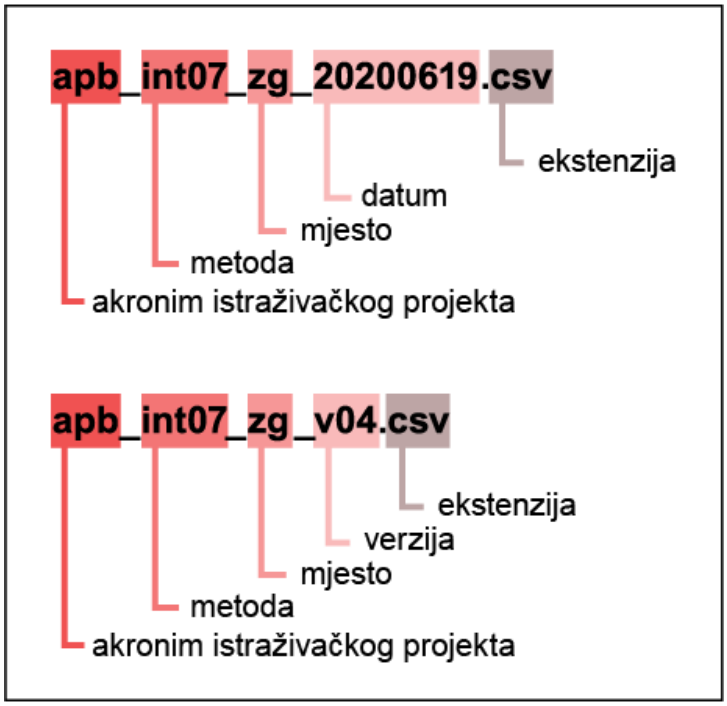
5.1 Imenovanje i organiziranje istraživačkih podataka

Prikupljanjem podataka potrebnih za istraživanje raste broj datoteka i mapa u koje su smješteni istraživački podaci. Pronalazak željenih istraživačkih podataka među mnoštvom datoteka pohranjenih na lokalnom sustavu pohrane ili u računalnom oblaku može biti otežano ako datoteke nisu dosljedno imenovane ili ne postoji logička struktura mapa.

Kako bi proces upravljanja istraživačkim podacima bio učinkovitiji, važno je stvoriti konvenciju imenovanja istraživačkih podataka te odgovarajuću strukturu mapa prije nego što istraživanje krene. Konvencija imenovanja istraživačkih podataka je način dodjeljivanja imena datotekama i mapama vezanim za istraživačke podatke. Evidentira se u dokumentaciji istraživačkih podataka, primjerice *readme* datoteci (više u poglavlju 4.4 Dokumentacija i metapodaci).

Preporuke za imenovanje istraživačkih podataka:

- razviti konvenciju imenovanja istraživačkih podataka koja se sastoji od logičkih elemenata. Relevantne informacije za imenovanje datoteka su naziv projekta ili akronim, ime istraživača ili inicijali, tip podataka, metoda istraživanja, mjesto i datum istraživanja, broj verzije datoteke

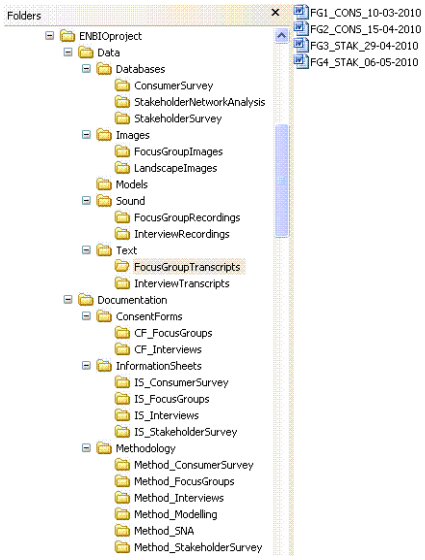


Slika 2. Primjer konvencije imenovanja istraživačkih podataka

- konzistentno koristiti konvencije imenovanja istraživačkih podataka
- koristiti konvencije imenovanja od strane svih članova istraživačkoga tima
- izbjegavati korištenje sličnog imena za više datoteka
- imena datoteka ne smiju biti predugačka, ne više od 32 znaka
- koristiti slova i brojeve ASCII koda (a — z, A — Z i 0 — 9). Izbjegavati korištenje točke ili posebnih znakova u imenu datoteke, kao & , *, %, #, ;, ( ), !, @, $, ^, ~, ‘, { }, [ ], ?, < >
- koristiti donju crtu, minus ili *camel case* umjesto razmaka
- koristiti datum u ISO 8601 standardu (GGGGMMDD)
- identificirati različite verzije datoteka s brojčanim oznakama (više u poglavlju 4.2 Verzioniranje istraživačkih podataka)
- automatska promjena imena datoteka moguća je uz pomoć programskih rješenja, primjerice Bulk Rename Utility.

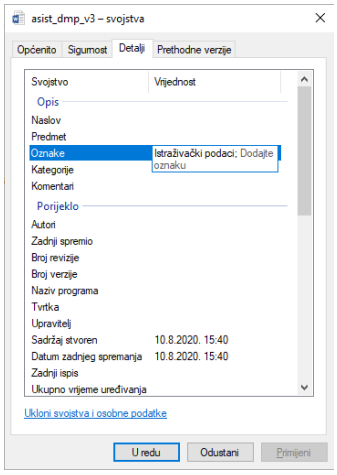
Preporuke za organiziranje istraživačkih podataka:

- stvaranje strukture mapa koja će odgovarati potrebama istraživačkoga projekta. Pri tome je važno obratiti pozornost o vrsti istraživačkih podataka te način na koji će biti organizirani sirovi i analizirani podaci, metode, dokumentacija i ostale popratne datoteke

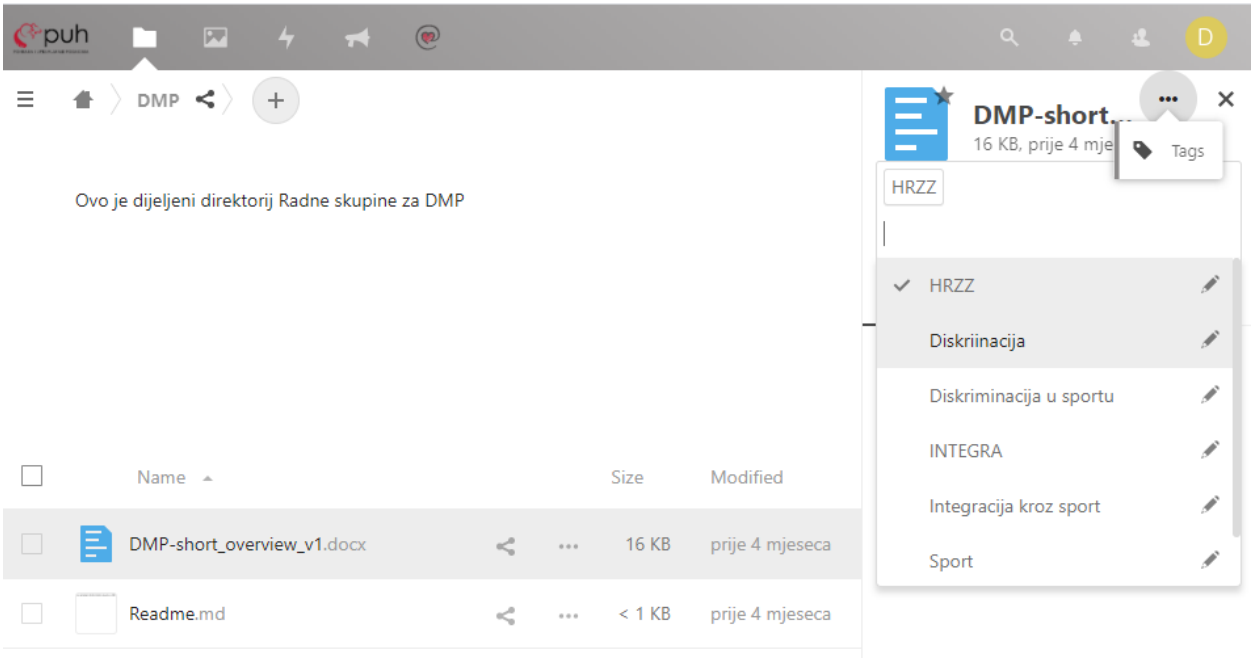


Slika 3. Struktura mapa. Izvor: https://www.ukdataservice.ac.uk/manage-data/format/organising.aspx

- davati jedinstvena imena mapama koja odgovaraju istraživačkom projektu
- postići odgovarajuću dubinu hijerarhije mapa. Pretjerano duboka hijerarhija rezultirat će velikim brojem klikova do željene datoteke, dok će mape pri preplitkoj hijerarhiji sadržavati previše datoteka
- tagiranje (engl. *tagging*) datoteka pomaže u pronalaženju datoteka unutar mapa istraživačkoga projekta. Tagovi su ključne riječi dodijeljene datotekama koje trebaju biti jednostavne i konzistentno korištene. Tagirati i pretraživati lokalne datoteke moguće je npr. u okviru operativnih sustava Windows i macOS (upute u poglavlju 6. Bibliografija) ili korištenjem programa, primjerice Adobe Acrobat ili Adobe Bridge te TagSpaces koji je kompatibilan operativnim sustavima Windows, macOs i Linux.



Slika 4. Tagiranje datoteke u operativnom sustavu Windows



Slika 5. Tagiranje datoteke u sustavu Puh

5.2 Verzioniranje istraživačkih podataka

Nakon nekog vremenskog odmaka teško je razlikovati brojne verzije datoteka, posebno ako na istim datotekama radi više ljudi. Verzioniranje ili kontrola verzija istraživačkih podataka je proces upravljanja promjenama datoteke tijekom istraživanja ili projekta. Kako bi se razumjele promjene, poželjno je i dokumentiranje povijesti promjena — tko, kada i zašto je napravio promjene. U nastavku su navedene preporuke nastale na temelju smjernica Sveučilišta Utrecht.

Preporuke za verzioniranje istraživačkih podataka:

- potrebno je voditi računa o zaštiti sirovih podataka. Budući da predstavljaju temelj cjelokupne analize, važno je da ostanu nepromijenjeni. Potrebno ih je pohraniti na odvojenu i zaštićenu lokaciju od lokacije radnih verzija istraživačkih podataka, primjerice u posebnu mapu koja se može samo čitati (engl. *read only*). Također, kopije sirovih podataka potrebno je pohraniti na više lokacija — na računalu, prijenosnom mediju i računalnom oblaku (više o aspektima očuvanja podataka u poglavlju 4.5 Pouzdano čuvanje podataka tijekom istraživanja)
- potrebno je odlučiti koliko i koje verzije datoteka zadržati, koliko dugo i kako ih organizirati
- na jedinstven način identificirati različite verzije datoteka koristeći vlastitu konvenciju imenovanja. Najjednostavniji način za identificiranje pojedine verzije jest dodati brojčanu oznaku poput *v1*, *v2_6* ili datuma (GGGGMMDD)
- automatska kontrola verzija datoteka moguća je korištenjem programskih rješenja poput Githuba.

5.3 Smjernice za odabir formata za trajnu pohranu podataka

Prilikom pohrane istraživačkih podataka potrebno je voditi računa o tome u kojim formatima će podaci biti pohranjeni i dostupni. Važno je odabrati otvorene i dobro dokumentirane formate koji jesu ili mogu biti podržani od više proizvođača programa što olakšava eventualnu migraciju na nove formate u budućnosti.

Odabirom odgovarajućega poželjnog ili prihvatljivog formata osigurat će se dugoročna pohrana i dostupnost podataka. Pri odabiru odgovarajuće vrste formata za pohranu podataka preporuča se korištenje otvorenih i široko primjenjivih formata poput običnoga teksta (ASCII, .txt), teksta razdvojenoga zarezom (engl. *comma-separated, .csv*) ili XML-a (*Extensible Markup Language*) umjesto zatvorenih, tzv. vlasničkih formata kao što su *Microsoft Office*, *Statistical Package for the Social Sciences* (SPSS) i drugi. U Tablici 3 nalazi se nekoliko primjera otvorenih i zatvorenih formata.

Tablica 3. Primjeri otvorenih i zatvorenih formata

Vrsta datoteke	Otvoreni formati	Zatvoreni formati
Tekstualne	.docx .txt .odf .pdf	.doc
Tablične	.xlsx .csv .ods	.xls
Slikovne	.tiff .jpg .png	.psd .bmp
Audio	.mp3 .flac .wav	.wma
Video	.mp4 .mpeg4 .mpeg	.wmv

Preporuke pri odabiru formata za trajnu pohranu skupa podataka:

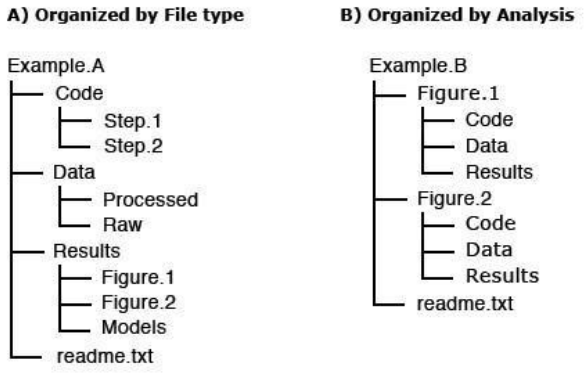
- koristiti otvorene i dobro dokumentirane formate
- koristiti formate za koje postoji podrška na raznim (svim) platformama (operativnim sustavima)
- koristiti tekstualne formate umjesto binarnih (.txt umjesto .pdf)

- izbjegavati enkripciju ili zaštite datoteka lozinkom kad nije nužno jer to otežava njihovu migraciju u druge formate
- izbjegavati sažimanje datoteka s gubitcima (postupak smanjivanja izvorne veličine datoteke kojim se kvaliteta sadržaja smanjuje).

5.4 Dokumentacija i metapodaci

Dokumentiranje je proces opisivanja istraživačkih podataka kojim se tijekom istraživanja bilježi što su istraživački podaci, kako su prikupljeni, analizirani i verzionirani. Budući da dokumentacija daje kontekst istraživačkim podacima, od velike je važnosti javnosti i samom istraživačkom timu za njihovo razumijevanje, dijeljenje i ponovno korištenje.

Metapodacima se opisuje skup istraživačkih podataka (podaci o podacima). Preporuča se zapisivanje metapodataka u tzv. *readme* datoteku i njihovo ažuriranje tijekom istraživačkog projekta. *Readme* datoteka treba biti u tekstualnom formatu (*.txt*) što je od ključne važnosti za trajnu zaštitu i čitljivost.



Slika 6. Readme datoteka smješta se u vršni direktorij skupa podataka
Izvor: https://www.helsinki.fi/en/research/guide-for-data-documentation

Preporuke za metapodatke u *readme* datoteci:

- naslov istraživačkih podataka i/ili istraživačkoga projekta
- autori, ustanova, kontakt
- opis podataka i pregled mapa
- ključne riječi koje opisuju predmet ili sadržaj podataka (koristiti standarde područja znanosti)
- lokacija na koju se istraživački podaci odnose
- metodologija stvaranja i eventualne analize podataka, korištena oprema i/ili program
- trajni identifikatori autora, podataka i rada
- prava korištenja (licencije)
- datumi (GGGGMMDD) povezani s podacima, uključujući datum početka i završetka istraživanja, datum objavljivanja izmjene podataka i vremensko razdoblje obuhvaćeno podacima

- formati istraživačkih podataka
- konvencija imenovanja datoteka
    - Primjer navođenja konvencije imenovanja:
        - *<akronim projekta><metoda><mjesto><verzija ili datum>.<ekstenzija>*
- organizacijska struktura mapa istraživačkoga projekta.

Dokumentacija može sadržavati rječnike istraživačkih podataka i kodne knjige (engl. *codebooks*). Rječnik istraživačkih podataka bilježi i objašnjava varijable, mjerne jedinice, raspon vrijednosti i dopuštene vrijednosti, vrijednosti koje nedostaju, metode, skale, formate datoteka, bilješke i slično.

Metapodaci se mogu formatirati koristeći metapodatkovne standarde područja znanosti koji se mogu pronaći u mrežnom direktoriju Metadata Standards Directory. Osim metapodatkovnih standarda u direktoriju se nalaze i alati za kreiranje metapodataka, validiranje i slično.

Prilikom trajne pohrane skupa podataka u repozitorij u Dabru unose se odgovarajući metapodaci kao što su naslov, opis, vlasnik, prava korištenja (licencija), metodologija, trajni identifikator, format datoteke, dokumentacija, licencija i slično.

5.5 Pouzdano čuvanje podataka tijekom istraživanja

Potpoglavlje se bavi aspektima očuvanja i pohrane istraživačkih podataka tijekom istraživanja kako ne bi došlo do gubitka podataka ili njihovoga neovlaštenog korištenja. Teme zaštite osobnih podataka i dugoročne pohrane konačnih rezultata istraživanja obrađene su u poglavlju 5. Nakon istraživanja.

Prilikom odabira mjesta za pohranu podataka potrebno je razmotriti dostupne mogućnosti i funkcionalnosti poput pohranjivanja s automatskom mogućnošću izrade sigurnosnih kopija, dijeljenja i razmjene podataka s drugim osobama te enkripcije podataka. Tablica 4. prikazuje smjernice za pouzdano čuvanje podataka tijekom istraživanja.

Tablica 4. Popis smjernica za pouzdano čuvanje podataka tijekom istraživanja

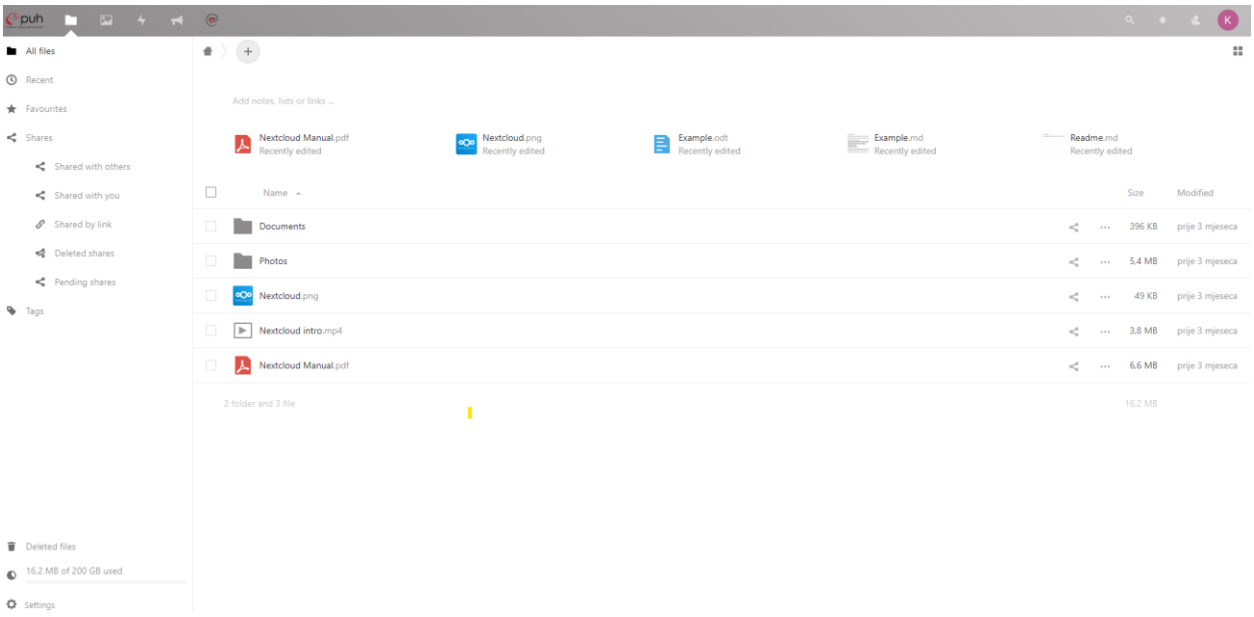
Fizički aspekti zaštite podatka
Gdje se podaci čuvaju? • Ako se podaci čuvaju na uređaju (prijenosno računalo, USB memorija i sl.), postoji li rizik od kvara, gubitka ili otuđenja uređaja? Je li pristup uređaju na kojem su pohranjeni podaci zaštićen na odgovarajući način? • Uređaj na kojem su pohranjeni podaci potrebno je zaštiti odgovarajućom lozinkom koja ne bi smjela biti kraća od 8 znakova koji uključuju brojeve, slova i simbole ispisane kombinacijom malih i velikih slova. Je li uređaj koji se koristi siguran? • Uređaj na kojem se nalaze podaci potrebno je zaštiti antivirusnim programom koji se redovito nadograđuje, uključenim vatrozidom te redovitom nadogradnjom operativnoga sustava.
Izrada sigurnosnih kopija
Izgrađuju li se sigurnosne kopije (engl. *backup*) podataka? • Izraditi najmanje jednu lokalnu kopiju materijala koji nisu pohranjeni u računalnom oblaku. • Izraditi najmanje jednu kopiju materijala pohranjenu na različitoj geografskoj lokaciji. • Pohraniti povijest izmjena (verzije) na najmanje jednu lokaciju. • Dokumentirati sve postupke vezane za pohranu i očuvanje podataka u *readme* datoteci, na primjer na kojim su sve mjestima podaci pohranjeni, koliko kopija, u kojim formatima i slično. Jesu li sve sigurnosne kopije pohranjene na jednom mjestu? • Preporuča se izraditi najmanje 3 u potpunosti odvojene kopije podataka (npr. računalo, prijenosi medij, oblak).
Rad s osobnim i osjetljivim podacima
Štite li se osjetljivi podaci na odgovarajući način? • Osobne i osjetljive podatke potrebno je anonimizirati, a one koje nije moguće anonimizirati potrebno je enkriptirati. Moguća je enkripcija pojedine datoteke, mape ili cijeloga diska. Alati za enkripciju BitLocker (za Windows OS), FileVault2 (za macOS), PGP, VeraCrypt, Axcrypt, SafeHouse. Brišu li se osjetljivi podaci na odgovarajući način? • Obično brisanje datoteka i mapa korištenjem neće trajno ukloniti podatke. Uz pomoć specijaliziranih alata moguće ih je povratiti. Postoje programska rješenja koja omogućuju trajno brisanje: Eraser, WipeFile, FreeRaiser.
Sigurna razmjena podataka
Razmjenjuju li se podaci među istraživačkim timom na siguran način?

- Treba izbjegavati slanje osjetljivih podataka elektroničkom poštom ili nekriptiranim kanalom (FTP).
- Preporuča se koristiti enkriptirane kanale (SFTP, HTTPS) ili servise u oblaku (PUH, FileSender).

Srce istraživačima pruža uslugu za pouzdano čuvanje i razmjenu podataka tijekom istraživanja pod nazivom Puh — Pohrana i upravljanje podacima. Sustavu se pristupa korištenjem AAI@EduHr elektroničkog identiteta te je omogućeno pohranjivanje materijala u oblaku. Sustav Puh omogućuje pohranu u oblaku na tri načina:

- putem *web*-sučelja na adresi https://puh.srce.hr
- korištenjem desktop klijenta odnosno aplikacije koja kontinuirano sinkronizira direktorij na računalu sa sadržajem u Puh-u
- spajanjem prostora na Puh-u kao mrežnoga diska na računalo ili server.

Osim pouzdane pohrane, Puh omogućava funkcionalnost kolaboracije i dijeljenje podataka s članovima istraživačkoga projekta ili drugim istraživačima kojima se dozvoli pristup.



Slika 7. Sučelje sustava Puh

POPIS DOSTUPNIH ALATA

- **Adobe Acrobat** (https://acrobat.adobe.com/us/en/acrobat/pdf-reader.html) — program za stvaranje, pregledavanje i upravljanje PDF dokumentima
- **Adobe Bridge** (https://www.adobe.com/products/bridge.html) — program za upravljanje digitalnim objektima
- **Bulk Rename Utility** (https://www.bulkrenameutility.co.uk/) — alat za automatska promjenu imena datoteke putem programskih rješenja
- **Bura** (https://cnrm.uniri.hr/hpc-bura/) — resurs za visokoučinkovito (engl. *high performance computing*, HPC) modeliranje, simuliranje i obradu podataka
- **FileSender** (https://www.srce.unizg.hr/filesender) — alat za prijenos datoteka
- **Github** (https://github.com/) — online alat za automatsku kontrolu verzija podataka
- **HTC Cloud** (https://www.cro-ngi.hr) — resurs za fleksibilno pokretanje proizvoljnih programskih sustava i aplikacija koji zahtijevaju značajne računalne i spremišne resurse
- **Isabella** (https://www.srce.unizg.hr/isabella/) — resurs za visokoučinkovito (engl. *high performance computing*, HPC) modeliranje, simuliranje i obradu podataka
- **JupyterLab servisi** (https://wiki.srce.hr/display/CRONGI/JupyterLab+servisi) — platforma za interaktivno pisanje i izvršavanje programa i analizu podataka u popularnim programskim jezicima poput Pythona, R i Julie
- **LimeSurvey** (https://www.srce.unizg.hr/limesurvey) — online alat za provođenje anketa, odnosno prikupljanje podataka
- **Microsoft File Checksum Integrity Verifier** (https://www.microsoft.com/en-us/download/details.aspx?id=11533) — alat za provjeru integriteta digitalnih datoteka
- Programi za enkripciju datoteka, mapa, prijenosnih medija i tvrdih diskova:
  - **Axcrypt** (https://www.axcrypt.net/)
  - **BitLocker** (https://support.apple.com/en-gb/HT204837) za Windows OS
  - **FileVault2** (https://support.apple.com/en-gb/HT204837) za macOS
  - **PGP** (https://medium.freecodecamp.org/how-does-pretty-good-privacy-work-3f5f75ecea97)
  - **SafeHouse** (http://safehousesoftware.com/)
  - **VeraCrypt** (https://www.veracrypt.fr/en/Home.html)
- Programi za trajno brisanje datoteka:
  - **Eraser** (https://eraser.heidi.ie/)
  - **FreeRaiser** (https://freeraser.en.softonic.com/)
  - **WipeFile** (https://www.gaijin.at/en/software/wipefile)
- **Puh** (https://www.srce.unizg.hr/puh) — sustav za pouzdano čuvanje i razmjenu podataka tijekom istraživanja

- **TagSpaces** (https://www.tagspaces.org/) — programsko rješenje za upravljanje datotekama.

6. Nakon istraživanja *(životni ciklus: trajna pohrana, dijeljenje, korištenje)*

Poglavlje donosi pregled posljednjih triju faza u životnom ciklusu istraživačkih podataka koje se odnose na trajnu pohranu, dijeljenje i korištenje podataka koji su prikupljeni i obrađeni tijekom projekta. Nadalje, teme kojima se poglavlje bavi odnose se na načine citiranja istraživačkih podataka, vrste licencija i postupak odabira odgovarajuće licencije te upravljanje i objavu osjetljivih podataka koji najčešće uključuju provedbu postupka anonimizacije i/ili pseudonimizacije.

6.1 Objava osobnih i osjetljivih podataka

Ovo poglavlje pruža smjernice za pravilno upravljanje osobnim i osjetljivim podacima. Agencija za zaštitu osobnih podataka (AZOP) u Republici Hrvatskoj osobni podatak definira kao svaku informaciju koja se odnosi na fizičku osobu koja je identificirana ili se može identificirati. Prema AZOP-u osobni podaci obuhvaćaju sljedeće podatke:

- adresu fizičke osobe
- broj telefona
- adresu elektroničke pošte
- osobnu fotografiju
- identifikacijski broj (npr. OIB)
- biometrijske podatke
- podatke o obrazovanju i stručnoj spremi
- podatke o plaći
- podatke o kreditnom zaduženju te
- računima u banci.

Osjetljivi podaci definiraju se kao posebna skupina osobnih podataka koji trebaju biti zaštićeni kako ne bi otkrili identitet osobe, a otkrivaju:

- rasno ili etično podrijetlo
- politička stajališta
- vjerska i druga uvjerenja
- sindikalno članstvo
- zdravlje ili spolni život
- osobne podatke o kaznenom i prekršajnom postupku.

Važno je naglasiti kako se osobni i osjetljivi podaci mogu prikupljati i obrađivati isključivo ako postoji potpisana privola osobe čiji će se podaci prikupljati i naposljetku objaviti. Osobni i

osjetljivi podaci prije objave moraju proći postupak anonimizacije i/ili pseudonimizacije kako bi se zaštitio identitet ispitanika, odnosno sudionika.

6.2 Informirana privola

Prije samog sudjelovanja u istraživanju ispitanicima se dostavlja informirana privola (engl. *Informed Consent*) koju potpisuju te na taj način daju istraživaču dozvolu za obradu i objavu podataka prikupljenih u svrhu istraživanja. U privoli istraživač mora ispitanike upoznati sa svrhom istraživanja, načinima prikupljanja i analize podataka, oblicima diseminacije i objave podataka te o upravljanju podacima nakon završetka projekta. Da bi se omogućilo ponovno korištenje istraživačkih podataka, istraživač treba jasno navesti mogućnost buduće uporabe. Važno je i navesti kako će se osigurati zaštita podataka, primjerice provođenjem postupka anonimizacije i slično. Osim toga, potrebno je izdvojiti pravo ispitanika na povlačenje iz istraživanja.
Nije preporučeno koristiti fraze poput “u potpunosti anonimno” ili “strogo povjerljivo” jer je to u praksi gotovo nemoguće postići. Također, treba izbjegavati obećanja o uništavanju podataka ili da će uvid u podatke imati samo istraživački tim. Privola treba biti pisana jasnim i razumljivim jezikom. Prema Općoj uredbi o zaštiti podataka[7], da bi privola bila valjana, ispitanik dobrovoljno mora posebno, informirano i nedvosmislenom izjavom (usmena privola) ili potpisom (pismena privola) dati pristanak za obradu podataka. Primjer privole za prikupljanje, obradu i objavu osobnih i osjetljivih podataka u svrhu znanstvenog istraživanja moguće je vidjeti na Elsevierovom predlošku.
Privolom se štite sve strane uključene u istraživanje (sudionici/ispitanici i istraživači) te se njome oblikuje proces istraživanja. Uz privolu istraživači trebaju ispitanicima dostaviti i GDPR (engl. *General Data Protection Regulation*) izjavu ako će tijekom istraživanja prikupljati osobne podatke ispitanika.

6.3 Osobni identifikatori sudionika istraživanja

Sveučilište Stanford na svojim mrežnim stranicama pruža klasifikaciju i objašnjenje osobnih identifikatora čija uporaba može razotkriti identitet osoba koje su sudjelovale u istraživanju. Navode se dvije vrste osobnih identifikatora, direktni i indirektni.
Direktni identifikatori su sve informacije koje mogu izravno identificirati osobu te se isključuju iz skupa podataka prije javne objave. Neki od direktnih identifikatora odnose se na:

- imena
- prezimena
- inicijale
- adrese elektroničke pošte
- brojeve telefona
- jedinstvene identifikacijske brojeve (OIB, JMBG i sl.)
- registracijske oznake automobila

[7] https://www.zakon.hr/z/1021/Op%C4%87a-uredba-o-za%C5%A1titi-podataka---Uredba-%28EU%29-2016-679

- mrežne i IP adrese
- biometrijske podatke
- fotografije
- audio snimke
- imena rođaka
- specifične datume (datumi rođenja, vjenčanja i sl.).

Indirektni identifikatori su sve one informacije i podaci koje u kombinaciji s drugim podacima mogu identificirati osobu. Preporučuje se da bi skup podataka ako sadrži tri ili više indirektnih identifikatora, trebao biti pregledan od strane nezavisnog istraživača ili komisije za etiku kako bi se procijenio potencijalni rizik identifikacije. Neki od indirektnih identifikatora odnose se na:

- ime liječnika ili mjesto liječenja
- spol
- rijetke bolesti ili tretmane
- rizično ponašanje
- mjesto rođenja
- socioekonomske podatke (radno mjesto, zanimanje, godišnji prihodi i sl.)
- geografske podatke (poštanski broj mjesta stanovanja)
- opis kućanstva i njegovih članova
- godina rođenja ili godine života
- verbalne odgovore ili transkripte intervjua.

Svi indirektni identifikatori koji nisu potrebni za analizu trebali bi biti uklonjeni ili prilagođeni na način da ih se rasporedi u određene raspone, primjerice umjesto unosa točne godine rođenja određene osobe sve godine rođenja moguće je podijeliti u kategorije (npr. 20 — 29 godina starosti, 30 — 39 godina starosti, 40 — 49 godina starosti, više od 50 godina starosti).

6.4 Anonimizacija i pseudonimizacija

Osjetljive i osobne podataka potrebno je anonimizirati kako bi se zaštitio identitet ispitanika. Postoje dvije vrste anonimizacije — anonimizacija i pseudonimizacija podataka — te se oba postupka mogu provesti na kvalitativnim i kvantitativnim podacima. Kvantitativni podaci uključuju numeričke i slovne vrijednosti, varijable i atribute dok kvalitativni podaci uključuju tekstualne podatke koji se odnose na transkripte intervjua, audio i video snimke te slikovne materijale. ENISA (*The European Union Agency for Cybersecurity*) pruža definicije anonimizacije i pseudonimizacije. Anominizaciju definira kao ireverzibilni postupak izmjene osobnih podataka nakon kojeg se osoba više ne može direktni ni indirektno identificirati. Postupak anominizacije podrazumijeva uklanjanja direktnih i/ili indirektnih identifikatora iz skupa podataka. Podaci se mogu izmijeniti ili drugačije organizirati kako ne bi otkrili identitet ispitanika. Postupak anonimizacije podataka ne može se poništiti, odnosno trajno anonimizira osobne i osjetljive podatke ispitanika. U tablici 5 nalazi se primjer podataka koji nisu anonimizirani dok tablica 6 prikazuje anonimizirane podatke. U tablici 6 vidljivo je kako su u postupku anonimizacije podataka svi direktni identifikatori u potpunosti skriveni te ne postoji mogućnost ponovne identifikacije ispitanika.

Primjer podataka nad kojima je proveden postupak anonimizacije:

Tablica 5. Podaci prije postupka anonimizacije

Ime	E-mail adresa	Stručna sprema
Maja	maja66@primjer.hr	SSS
Ivan	ivan5@primjer.hr	VSS
Marko	marko44@primjer.hr	VSS



Tablica 6. Podaci nakon postupka anonimizacije

Ime	E-mail adresa	Stručna sprema
******	******	SSS
******	******	VSS
******	******	VSS

Pseudonimizacija je definirana kao postupak reidentifikacije osobe koji je, za razliku od anonimizacije, reverzibilan, odnosno osobu se može povratno identificirati. Riječ je o postupku u kojem se osobni podaci ne mogu više povezati sa specifičnim podacima bez dodatnih informacija, pod uvjetom da su dodatne informacije pohranjene odvojeno kako se ne bi otkrio identitet osobe. U tablici 8 prikazan je primjer provedbe pesudoanomimizacije podataka gdje su podaci šifrirani te ih je moguće putem šifrarnika povezati s identitetom ispitanika. Prilikom provedbe pesudoanonimizacije istraživač mora imati pohranjenu kodnu knjigu ili šifrarnik koji mu omogućava ponovnu identifikaciju sudionika. Kodna knjiga morat biti pohranjena na sigurnom mjestu koje je dislocirano od skupa podataka kako ne bi došlo do otkrivanja identiteta ispitanika.

Primjer podataka nad kojima je proveden postupak pseudonimizacije:

Tablica 7. Podaci prije postupka pseudoaninimizacije

Ime	E-mail adresa	Stručna sprema
Maja	maja66@primjer.hr	SSS
Ivan	ivan5@primjer.hr	VSS
Marko	marko44@primjer.hr	VSS



Tablica 8. Podaci nakon postupka pseudoaninimizacije

Ime	E-mail adresa	Stručna sprema
5GzhT	6zhg@primjer.hr	SSS
8kTR8	U76n@primjer.hr	VSS
Lk7f8	mJ76@primjer.hr	VSS

Besplatan mrežni alat za anonimizaciju kvantitativnih podataka naziva se Amnesia te podržava postupak anonimizacije za .csv i .txt datoteke. Alat je trenutačno fokusiran na k-anonimizaciju koja osigurava da se svaki zapis unutar anonimiziranih podataka ne može razlikovati od ostalih zapisa unutar iste tablice, odnosno baze podataka. U tablici 10 može se vidjeti primjer provedbe k-anonimizacije koju je moguće dobiti korištenjem alata Amnesia. U primjeru su imena sudionika sakrivena dok su godine starosti prebačene u odgovarajuće raspone.

Primjer podataka nad kojima je proveden postupak k-anonimizacije

Tablica 9. Podaci prije provedbe k-anonimizacije

Ime	Godine starosti	Dijagnoza	Liječen/a u bolnici
Maja	25	hipertenzija	KBC Zagreb
Ivan	31	pneumonija	KBCCSM
Marko	28	celijakija	KBC Zagreb
Ivana	33	hipertenzija	KBC Sveti Duh
Darko	52	pneumonija	KBCCSM



Tablica 10. Podaci nakon provedbe k-anonimizacije

Ime	Godine starosti	Dijagnoza	Liječen/a u bolnici
****	20 — 29	hipertenzija	KBC Zagreb
****	30 — 39	pneumonija	KBCCSM
****	20 — 29	celijakija	KBC Zagreb
****	30 — 39	hipertenzija	KBC Sveti Duh
****	> 50	pneumonija	KBCCSM

Osim Amnesie dostupni su savjeti o privatnosti, anonimizaciji i pseudonimizaciji istraživačkih podataka pod nazivom RDM Advice & Tips gdje se istraživačima nudi obilje praktičnih savjeta i smjernica kako što kvalitetnije i bolje upravljati s istraživačkim podacima.
UK Data Service također pruža informacije o anonimizaciji te na koji način i nad kojom vrstom podataka se može provesti. Obuhvaća smjernice i pojašnjenja za anonimizaciju kvantitativnih i kvalitativnih podataka.

6.5 Definiranje prava korištenja skupa podataka

Prije objave podataka važno je definirati kako će se skup istraživačkih podataka moći koristiti. To se uređuje licencijom koja predstavlja dozvolu kojom autor definira uvjete za ponovno korištenje djela, u ovom slučaju skupa istraživačkih podataka. Biranjem odgovarajuće licencije, autor zadržava autorsko pravo i omogućuje drugima da djelo koriste uz jasno definirane uvjete.

Također, u kontekstu otvorene znanosti, mnogi financijeri zahtijevaju objavu istraživačkih podataka uz unaprijed definirane licencije.

6.5.1 Što podliježe autorskom pravu

Prije odabira odgovarajuće licencije potrebno je provjeriti podliježe li djelo autorskom pravu te ustanoviti nositelja autorskoga prava (je li to istraživač, poslodavac i sl.). Prema Zakonu[8], autorsko pravo je pravo autora nad njihovim djelima iz književnog, znanstvenog i umjetničkog područja. Glavni kriterij da neko djelo postane autorsko jest njegova originalnost. Postoje izuzetci koji nisu zaštićeni autorskim pravom: ideje, znanstvena otkrića, postupci, metode rada i matematički koncepti, službeni tekstovi i slično. Istraživački podaci obuhvaćaju puno vrsta podataka, od kojih neki nisu autorska djela, na primjer, činjenični istraživački podaci, poput mjere temperature ili tečaja valute. Ostali istraživački podaci, obrađeni na određen način u djelu, a u čijoj je obradi autor imao slobodu izbora u obradi i odabrao je neki svoj originalni način, tretiraju se kao autorsko djelo.
Potrebno je ustanoviti i kome pripada autorsko pravo. Iako prema Zakonu autorsko pravo uvijek pripada autoru, pravo iskorištavanja i objave ponekad pripada i poslodavcu ako je djelo stvoreno u radnom odnosu, ili financijeru ako je nastalo unutar nekog projekta. Praksa na hrvatskim sveučilištima jest da autorsko djelo pripada autoru, osim u nekim slučajevima, poput računalnoga softvera ili audiovizualnim djelima.

6.5.2 Vrste licencija

Creative Commons (CC) je najpopularniji sustav licencija koji je široko prihvaćen. Postoji od 2001. godine i neprestano se poboljšava verzioniranjem. Postoje četiri glavna uvjeta koja definiraju korištenja sadržaja:

- **Imenovanje/*Attribution* (BY)** — mijenjanje, prerada i dijeljenje autorskoga djela su dopušteni uz imenovanje izvornog autora
- **Nekomercijalno/*NonCommercial* (NC)** — dopušta mijenjanje, preradu i dijeljenje isključivo u nekomercijalne svrhe
- **Bez prerada / *NoDerivatives* (ND)** — djelo se može koristiti, ali u izvornom obliku te se ne dopuštaju nikakve prerade
- **Dijeli pod istim uvjetima / *ShareAlike* (SA)** — dopušteno je mijenjanje, prerada i dijeljenje pod istim uvjetima (licencijom) kao izvorno djelo.

Kombinacijom navedenih uvjeta, *Creative Commons* je razvio šest licencija:

- **CC BY** (Imenovanje)
- **CC BY-SA** (Imenovanje-Dijeli pod istim uvjetima)
- **CC BY-ND** (Imenovanje-Bez prerada)
- **CC BY-NC** (Imenovanje-Nekomercijalno)
- **CC BY-NC-SA** (Imenovanje-Nekomercijalno-Dijeli pod istim uvjetima)
- **CC BY-NC-ND** (Imenovanje-Nekomercijalno-Bez prerada).

[8] Zakon o autorskom pravu i srodnim pravima, NN 167/2003

Osim navedenih licencija, postoji i pravni akt **CC0** *Public Domain Dedicatio*n (*Zero*), koji omogućava prijenos djela u javnu domenu, koje se kasnije može slobodno koristiti, bez ikakvih restrikcija.

6.5.3 Koje licencije koristiti za istraživačke podatke?

Općenito, što je licencija manje restriktivna, veća je i mogućnost ponovnog i ispravnog korištenja.



CC0, odnosno prenošenje djela u javnu domenu, čime omogućujemo drugima nesmetano ponovno korištenje tih podataka. Pri korištenju podataka koji su objavljeni uz licenciju CC0 zakonski nije obavezno imenovanje, ali citiranje je norma nametnuta znanstvenim integritetom.



CC BY 4.0, kojom dopuštamo mijenjanje, preradu i dijeljenje uz uvjet imenovanja. Potrebno je navesti autora, poveznicu na izvorno djelo (skup podataka) i naznačiti sve izmjene. U kontekstu podataka, pri korištenju CC BY licencije treba biti oprezan zbog mogućnosti nagomilavanja imenovanja. Kad se podaci višekratno koriste ili u kombinaciji različitih skupova podataka, naznačavanje svih potrebnih imenovanja može stvoriti veliko opterećenje.

Najbolja opcija jest korištenje CC0 uz zamolbu za imenovanjem. Uz objavljene podatke može postojati autorska napomena "*Please attribute my data*" ili slična fraza. Korištenje licencija CC0 i CC BY 4.0 preporuča se od strane Europske komisije za Obzor2020 projekte i RDA.

Kako kombinirati više licencija?
Jednom dodijeljena CC licencija se ne može više mijenjati, tako da treba biti oprezan pri odabiru. Ako na novom skupu podataka, koji je nastao korištenjem drugih skupova podataka s različitim licencijama, ne možemo jasno razdvojiti za što smo koristili koji skup podataka, na njega moramo primijeniti najrestriktivniju licenciju.

Naznačavanje licencije
Prilikom objave skupa podataka potrebno je jasno naznačiti licenciju pod kojom se podaci objavljuju. U repozitorijima i arhivima licenciju je potrebno odabrati tijekom postupka pohrane podataka.
Također, poželjno je licenciju naznačiti i u *readme* datoteci (više u poglavlju 2.4 Dokumentacija i metapodaci).

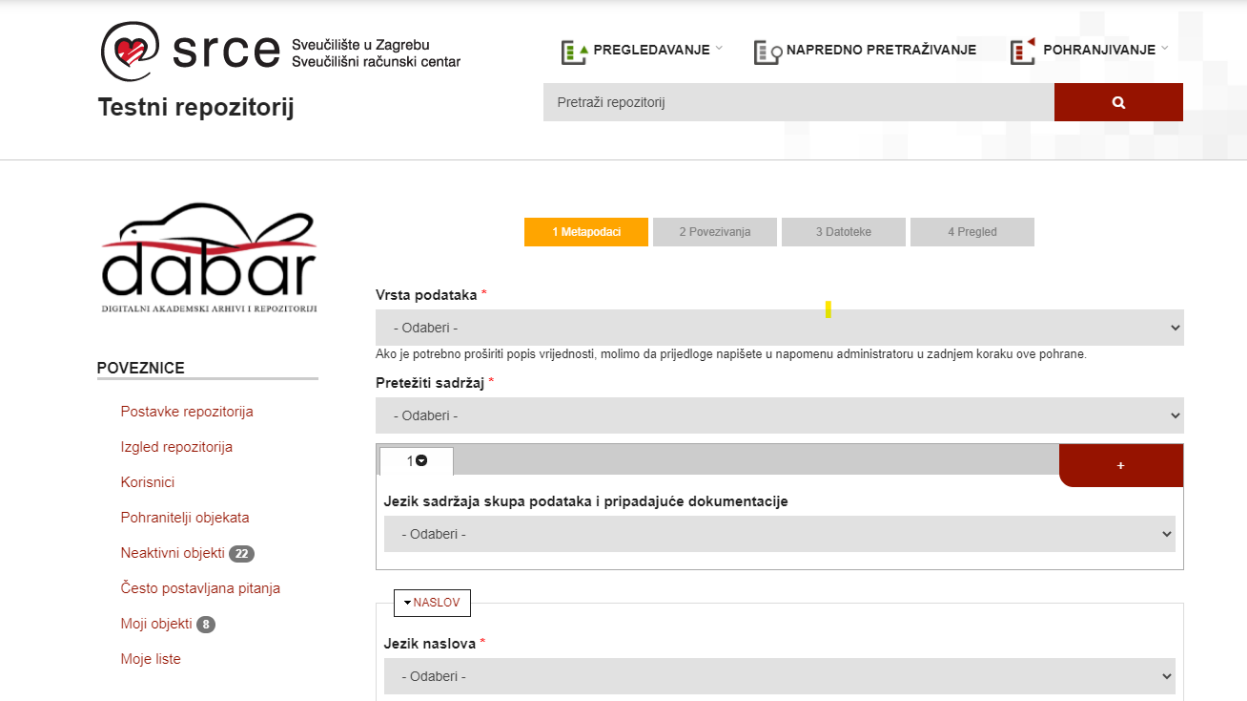
6.6 Trajna pohrana

Nakon završetka istraživanja, podaci se mogu dugoročno pohraniti u repozitoriju autorove ustanove uspostavljenom na nacionalnoj infrastrukturi Dabar. Podršku za pohranu istraživačkih podataka može pružiti knjižničar ili urednik institucijskoga repozitorija, a podatke je moguće i samostalno pohraniti u repozitorij (samoarhiviranje) ako je ta opcija uključena. Osim u repozitoriju, istraživački podaci mogu se objaviti i kao dodatak radu u časopisu.
Osim u institucijski repozitorij, podaci se mogu pohraniti u tematske ili disciplinske repozitorije kakav je, na primjer, Archaeology Data Service za znanstveno polje arheologije, te u repozitorije opće namjene poput Zenodoa.
Pri odabiru repozitorija za trajnu pohranu istraživačkih podataka, važno je da je repozitorij usklađen s FAIR principima (dodjeljuje trajne identifikatore, omogućava odabir licencije, ima podršku za bogat metapodatkovni opis).
Dodatne smjernice za odabir odgovarajućega repozitorija:

- Podržava li formate podataka koji se pohranjuju, osigurava li potrebnu količinu prostora za pohranu, izrađuju li se sigurnosne kopije? Odgovori na ta pitanja trebali bi se nalaziti u politikama repozitorija.
- Je li repozitorij pouzdan i održiv (hardverski, softverski, podrška)? Ima li certifikat, na primjer, *CoreTrustSeal* kojim to dokazuje?
- Omogućava li repozitorij praćenje statistika korištenja podataka (broj pregleda metapodataka, preuzimanje podataka)?

6.6.1 Dabar — nacionalna infrastruktura za digitalne repozitorije

Digitalni repozitoriji uspostavljeni na nacionalnoj infrastrukturi Digitalni akademski arhivi i repozitoriji — Dabar omogućavaju pohranu, dugoročno čuvanje i diseminaciju digitalnih sadržaja uključujući i skupove istraživačkih podataka. Namjena repozitorija uspostavljenih u Dabru jest dugoročno čuvanje i dijeljenje konačne verzije podataka, dok je za čuvanje i dijeljenje radnih verzija tijekom istraživanja namijenjen sustav Puh (više u poglavlju 4.5. Pouzdano čuvanje podataka tijekom istraživanja).



Slika 8. Sučelje za unos skupa podataka sustava Dabar

Osim unosa samoga skupa podataka i pripadajuće dokumentacije, prilikom pohrane skupa podataka potrebno je unijeti i odgovarajuće metapodatke koji omogućavaju pronalaženje i citiranje skupa podataka poput naziva skupa podataka, opisa korištene metodologije, autora skupa podataka, podataka o projektu na kojem je skup podataka nastao, povezanih radova i skupova podataka, licencije i slično. Kod pohrane je također potrebno definirati prava pristupa skupu podataka pri čemu su dostupne razne opcije poput otvorenoga pristupa, embarga, pristupa isključivo djelatnicima i studentima ustanove.
Pohranu u repozitorij može napraviti autor podataka ili urednik repozitorija prema uputama za objavu skupa podataka u repozitorij u Dabru.
Pohrana skupova podataka u Dabru usklađena je s FAIR principima:

- svaki skup podataka opisuje se standardnim skupom metapodataka
- dodjeljuje mu se trajni identifikator (URN:NBN)
- postoje sučelja za pretraživanje
- metapodaci se prosljeđuju u druge servise poput OpenAIRE portala ili Google Scholara
- omogućeno je definiranje prava korištenja.

Osim navedenih stavki, za usklađenost s FAIR principima, potrebno je dokumentirati i strukturirati skup podatak prema standardima znanstvenoga područja ili grane kako bi ga drugi istraživači mogli razumjeti i koristiti. Pohranom istraživačkih podataka u repozitorij u Dabru omogućeno je ispunjavanje obveze otvorene objave istraživačkih podataka nastalih u okviru projekata financiranih od Europske komisije.
Osim istraživačkih podataka, Dabar omogućava pohranu i ostalih sadržaja koji nastaju kroz istraživačke i obrazovne procese, kao što su znanstveni i stručni radovi objavljeni u časopisima ili zbornicima, knjige, poglavlja knjiga, prezentacije, studije, elaborati, projektni prijedlozi i

izvještaji, tehnička dokumentacija, disertacije, diplomski i završni radovi, obrazovni sadržaji, multimedijalni sadržaji (umjetničke fotografije, audio zapisi, video zapisi) i drugo.
U Dabru je uspostavljeno više od 140 repozitorija ustanova iz sustava znanosti i visokog obrazovanja u Republici Hrvatskoj (https://dabar.srce.hr/repozitoriji), a podršku pri pohrani i korištenju pojedinog repozitorija pruža urednik repozitorija na ustanovi.

6.6.2 Preporuke za odabir repozitorija

Važno je provjeriti jesu li financijeri već unaprijed odredili repozitorij za trajnu pohranu podataka ili izdvojili kriterije kojima repozitorij mora udovoljavati.
Preporuka je da se istraživački podaci pohrane u institucijski repozitorij u Dabru, koji podržava FAIR principe i omogućava objavu sukladno zahtjevima koje nalaže Europska komisija.
Pretraživanje tematskih repozitorija moguće je preko re3data, registra koji okuplja više od 2000 repozitorija. Ako se podaci pohranjuju u opći repozitorij, preporuka je da to bude Zenodo.

6.6.3 Istraživački podaci kao (dodatak) rad(u) u časopisu

Istraživačke podatke moguće je objaviti i kao popratni materijal radu u časopisu. Neki znanstveni izdavači, kao što su Elsevier (Scopus) i Clarivate Analytics (Web of Science), imaju detaljno razrađene politike objave istraživačkih podataka uz rad u časopisu, ali i opcije naknadne objave istraživačkih podataka koji se povezuju s prethodno objavljenim radom.
Postoji mogućnost objave istraživačkih podataka kao samostalnog elementa istraživanja u publikacijama koji objavljuju isključivo podatke, tzv. *data journal.* Zbog svoje specifičnosti, takvi časopisi nisu orijentirani na analizu podataka i teorijske prikaze, pa se rad u *data journalu* najčešće sastoji od sažetka, uvoda, opisa podataka, metoda i materijala, zaključka i mogućnosti ponovne uporabe podataka.

6.7 Citiranje istraživačkih podataka

Citiranje istraživačkih podataka važan je aspekt istraživanja — ono omogućuje praćenje rada istraživača te omogućuje ponavljanje istraživanja u istim ili sličnim uvjetima. Olakšava pristup znanstvenim radovima te osigurava valjano pripisivanje autorstva. Osim toga, omogućuje dokazivost teza u istraživačkim radovima.
Stoga je važno prikupljati istraživačke podatke tijekom istraživanja i potom ih objaviti u repozitoriju i pravilno citirati u radu.

6.7.1 Smještanje citata

Unutar rada:

- treba pružiti dovoljno informacija da bi se na popisu referenci mogla pronaći referenca na citirane podatke
- navođenje podataka trebalo bi biti u neposrednoj blizini tvrdnji koje se oslanjaju na navod
- mogu se uključiti dodatne informacije koje identificiraju određene dijelove podataka koji podupiru neku tvrdnju (primjer: Grafički podaci na slici x pokazuju raspodjelu odabranih mjera u podacima [autor(i), godina, odlomak korištenih podataka]).

Puni citat:

- citati se mogu razlikovati po stilu, ali trebali bi biti uključeni u popis literature, uključujući i citate za drugu vrstu djela. Neki od primjera su:
    - autor(i), godina, naslov članka, časopis, izdavač, DOI
    - autor(i), godina, naslov skupa podataka, repozitorij ili arhiv, verzija, *Global Persistent Identifier* (npr. DOI, *handle* – trajni identifikator koji omogućuje pristup i metapodatke)
    - autor(i), godina, naslov, izdavač, ISBN.

6.7.2 Elementi i primjeri citata

Citiranje podataka vrlo je slično citiranju publikacija te je preporučljivo uključiti sljedeće elemente u citat:

- autor(i) ili suradnik(/ci)
- datum objave
- naslov skupa podataka
- izdavač
- identifikator (npr. DOI, Handle) ili URL izvora
- verzija (kad je primjenjiv)
- datum pristupa (kad je primjenjivo).

Redoslijed ovih elemenata nije posebno važan koliko je važna dostatna količina informacija koja bi omogućila pronalazak korištenoga skupa podataka.

Primjer citiranja istraživačkih podataka

Bosančić, Boris. 2011. *Istraživački podaci vezani uz istraživanje označavanja teksta starih knjiga na hrvatskom jeziku pomoću TEI standarda.* Sveučilište u Zadru. URL: urn:nbn:hr:142:123845 Retrieved: 09 57, Aug 28, 2020 (GMT)

POPIS DOSTUPNIH ALATA

- **Amnesia** (https://amnesia.openaire.eu/) — mrežni alat za anonimizaciju kvantitativnih podataka koji su pohranjeni u .csv i .txt obliku
- **COPTR** (*Community Owned Digital Preservation Tool Registry* https://coptr.digipres.org/Main_Page) — mrežni alat koji pomaže istraživačima pronaći alate za očuvanje podataka odnosno registar alata za pohranu podataka
- ***Creative Commons* (CC)** (https://creativecommons.org/) — sustav licencija
- **Dabar** (https://dabar.srce.hr/repozitoriji) — repozitoriji za trajnu pohranu i dijeljenje konačne verzije skupova podatka i ostalih rezultata istraživačkoga projekta, uglavnom po završetku istraživanja
- **RDM Advice & Tips** (https://www.lcrdm.nl/en/rdm-advice-tips) — savjeti o privatnosti, anonimizaciji i pseudonimizaciji istraživačkih podataka
- **Re3data** (https://www.re3data.org/) — omogućuje pretraživanje tematskih repozitorija, a okuplja više od 2000 repozitorija
- **UK Data Service** (https://www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation/quantitative.aspx) — savjeti za provedbu anonimizacije nad kvantitativnim i kvalitativnim podacima
- **Zenodo** (https://zenodo.org/) — opći repozitorij za pohranu podataka.

7. Bibliografija

1. Agencija za zaštitu osobnih podataka. Zaštita osobnih podataka u RH. URL: https://azop.hr/images/dokumenti/217/zastita_op_rh.pdf
2. CARNET. Anonimizacija i pseudonimizacija podataka, 2018. URL: https://www.cert.hr/wp-content/uploads/2018/08/anonimizacija_i_pseudonimizacija_podataka.pdf
3. Cornell University Libraray. Recommended File Formats. URL: https://guides.library.cornell.edu/ecommons/formats
4. Cornell University. Research Data Managment Srevice Group. URL: https://data.research.cornell.edu/content/file-formats
5. Corti, L., Van den Eynden, V., Bishop, L. & Woollard, M. Managing and Sharing Research Data: A Guide to Good Practice. SAGE Publications (2020).
6. Creative Commons. URL: https://creativecommons.org/
7. Dabar. URL: https://dabar.srce.hr/dabar
8. Data Archiving and Networked Services. Preferred formats. URL: https://dans.knaw.nl/en/about/services/easy/information-about-depositing-data/DANSpreferredformatsUK.pdf
9. Deene, Joris. The legal status of research data (copyright, database right). URL: https://www.fosteropenscience.eu/sites/default/files/pdf/1366.pdf
10. DePaul University Library. Research Data Management (A How-to Guide). URL: https://libguides.depaul.edu/c.php?g=620925&p=4324498
11. Digital Curation Centre. Primjer DMP-a. URL: https://www.dcc.ac.uk/resources/data-management-plans/guidance-examples
12. Educopia Institute. Preservation and Curation of ETD Research Data and Complex Digital Objects. URL: https://educopia.org/wp-content/uploads/2018/04/etdplus_storage_guidancebrief.pdf
13. Europska komisija. Horizon 2020 online manual. URL:
14. Force11 — The FAIR Data Principles. https://www.force11.org/group/fairgroup/fairprinciples

15. GoFAIR — FAIR Principles https://www.go-fair.org/fair-principles/

16. Guide for data documentation, 2020. URL: https://www.helsinki.fi/en/research/guide-for-data-documentation

17. Guide to writing "readme" style metadata. URL: https://data.research.cornell.edu/content/readmehttps://learn.canvas.net/courses/2719/pages/exercise-2-readme-file-faculty-follow-up?module_item_id=241426

18. H2020 Programme: Guidelines to the rules on open access to scientific publications and open access to research data in Horizon 2020. URL: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

19. Horizon Europe – sljedeći program EU-a za istraživanje i inovacije (2021. — 2027.) https://ec.europa.eu/info/sites/info/files/research_and_innovation/strategy_on_research_and_innovation/presentations/horizon_europe_hr_oblikovanje_nase_buducnosti.pdf

    https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

    https://study.sagepub.com/corti2e

    https://www.dcc.ac.uk/resources/data-management-plans/guidance-examples

    https://www.dlcm.ch/resources/dlcm-dmp

    https://www.ukdataservice.ac.uk/manage-data/plan/planning.aspx

20. Jisc. Research Data Management Toolkit. URL: https://rdmtoolkit.jisc.ac.uk/manage-store-and-preserve/security/

21. Jones, Sarah; Grootveld, Marjan. (2017, November). How FAIR are your data?. Zenodo. http://doi.org/10.5281/zenodo.1065991

22. Kranjec, I.; M. Glavica; A. Vodopijevec. Istraživački podaci i visokoškolske knjižnice. // Vjesnik bibliotekara Hrvatske 61, 1(2018) , 611-626. doi:10.30754/vbh.61.1.635

23. Krishna, Vasmi. How to tag files in windows for easy retrieval, 2018. URL: https://www.maketecheasier.com/tag-files-in-windows/

24. Matanovac Vučković, Romana. Pravni status autorskih djela stvorenih u radnom odnosu na sveučilištu. Zbornik Hrvatskog društva za autorsko pravo, 11-12 (2011), str. 1 — 26.

URL: http://hdap-alai.hr/romana-matanovac-vuckovic-pravni-status-autorskih-djela-stvorenih-u-radnom-odnosu-na-sveucilistu/

25. OpenAIRE. Amnesia. URL: https://amnesia.openaire.eu/
26. OpenAire. How to find a trustworthy repository for your data? URL: https://www.openaire.eu/find-trustworthy-data-repository
27. Piwowar HA, Vision TJ. 2013. Data reuse and the open data citation advantage. PeerJ 1:e175 https://doi.org/10.7717/peerj.175
28. Rehberger, Dean; Coates, Brendan. File naming in the digital age. URL: http://ohda.matrix.msu.edu/2012/08/file-naming-in-the-digital-age/
29. Research data management support : storing and preserving data. URL: https://www.uu.nl/en/research/research-data-management/guides/storing-and-preserving-data
30. Scherle, Ryan. Filename requirements for digital objects, 2006. URL: https://wiki.dlib.indiana.edu/display/INF/Filename+Requirements+for+Digital+Objects
31. Stanford University. Sharing sensitive data. URL: https://library.stanford.edu/research/data-management-services/share-and-preserve-research-data/sharing-sensitive-data
32. Swiss National Science Foundation. Data Management Plan (DMP) — Guidelines for researchers. URL: http://www.snf.ch/en/theSNSF/research-policies/open_research_data/Pages/data-management-plan-dmp-guidelines-for-researchers.aspx
33. The European Union Agency for Cybersecurity. Pseudonymisation techniques and best practices. URL: file:///C:/Users/Administrator/Downloads/Guidelines%20on%20shaping%20technology%20according%20to%20GDPR%20provisions.pdf
34. UK Data Service. Anonymisation. URL: https://www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation/quantitative.aspx
35. UK Data Service. Data management planning. URL: https://www.ukdataservice.ac.uk/manage-data/plan/planning.aspx

36. UK Data Service. File formats and software. URL: https://www.ukdataservice.ac.uk/manage-data/format/file-formats.aspx
37. UK Data Service. Store your data. URL: https://www.ukdataservice.ac.uk/manage-data/store.aspx
38. University of Edinburgh, Research Data Service: Our definitions. https://www.ed.ac.uk/information-services/research-support/research-data-service/after/data-repository/definitions
39. Use tags to organize files on Mac. URL: https://support.apple.com/lt-lt/guide/mac-help/mchlp15236/mac
40. Utrecht University Guides. Storing and preserving data. URL: https://www.uu.nl/en/research/research-data-management/guides/storing-and-preserving-data
41. Utrecht University. Storing and preserving data. URL: https://www.uu.nl/en/research/research-data-management/guides/storing-and-preserving-data
42. Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

Pojmovnik

- **Anonimizacija**
  - Ireverzibilni postupak izmjene osobnih podataka nakon kojeg se osoba više ne može direktno ni indirektno identificirati.
- **Dabar (Digitalni akademski arhivi i repozitoriji)**
  - Nacionalna infrastruktura za digitalne repozitorije koja ustanovama i drugim pravnim osobama iz sustava znanosti i visokog obrazovanja RH omogućuje uspostavu repozitorija za pohranu, dugoročno čuvanje i diseminaciju digitalnih sadržaja uključujući i skupove istraživačkih podataka.
- **DMP** (*Data Management Plan*)
  - Službeni dokument koji opisuje načine upravljanja podacima tijekom i nakon završetka projekta, a obuhvaća proces organizacije, planiranja procesa istraživanja, prikupljanja, obrade, pohrane i distribucije podataka.
- **DOI** (*Digital Object Identifier*)
  - Jedinstveni trajni digitalni identifikator koji se dodjeljuje objektu bilo koje vrste.
- **Enkripcija**
  - Postupak šifriranja digitalnih informacija tako da ih može razumjeti i interpretirati samo ovlaštena osoba.
- **FAIR** (*Findable, Accessible, Interoperable, Reusable*)
  - Smjernice čija je svrha učiniti podatke vidljivima, dostupnima, interoperabilnim te ponovno upotrebljivima, a namijenjene su istraživačima, knjižničarima i izdavačima te svima koji se bave proizvodnjom i organizacijom podataka.
- **GDPR** (*General Data Protection Regulation*)
  - Izjava koju sudionici istraživanja potpisuju ako će se tijekom istraživanja prikupljati osobni podaci ispitanika.
- **Informirana privola** (engl. *Informed Consent*)
  - Dokument koji informira sudionika istraživanja o istraživanju te njegovim potpisom sudionik daje istraživaču dozvolu za obradu i objavu podataka prikupljenih u svrhu istraživanja.
- **Istraživački podaci**
  - Kvantitativni i kvalitativni podaci prikupljeni, zabilježeni ili generirani tijekom istraživanja.
- **Kodna knjiga / šifrarnik** (engl. *codebook*)
  - Dokumentacija koja omogućava ponovnu identifikaciju sudionika istraživanja.
- **Konvencija imenovanja**
  - Način dodjeljivanja imena datotekama vezanim za istraživačke podatke.
- **Licencija**
  - Dozvola kojom autor definira uvjete za ponovno korištenje djela.
- **Metapodaci**
  - Predstavlja opis skupa istraživačkih podataka (podaci o podacima).
- **ORCID** (*Open Researcher and Contributor ID*)
  - Jedinstveni i trajni identifikator istraživača i suradnika čije će korištenje omogućiti bolju vidljivost autora i interoperabilnost širokoga kruga informacijskih sustava.

- **Osjetljivi podaci**
  - Definiraju se kao posebna skupina osobnih podataka koji trebaju biti zaštićeni kako ne bi otkrili identitet osobe.
- **Osobni podaci**
  - Svaka informacija koja se odnosi na fizičku osobu koja je identificirana ili se može identificirati.
- **PID** (*Persistent Identifiers*)
  - Trajne poveznice na različite oblike digitalnih objekata, publikacija, skupove podataka ili osobe. Mogu biti dodijeljeni objektima (URN:NBN, DOI, itd.) ili osobama (ORCID).
- **Pseudonimizacija**
  - Postupak reidentifikacije osobe koji je, za razliku od anonimizacije, reverzibilan, odnosno osobu se može povratno identificirati.
- **RDM** (*Research Data Management*)
  - Upravljanje podacima koji su nastali tijekom životnog ciklusa istraživanja te je sastavni dio istraživačkoga procesa koji istraživačima pomaže da su podaci organizirani, opisani, pohranjeni i dijeljeni na odgovarajući način.
- **Trajna pohrana**
  - Odnosi se na niz aktivnosti koje se provode tijekom životnog ciklusa istraživanja kako bi se osigurao kontinuirani pristup digitalnim materijalima koliko god to bilo potrebno.
- **URN:NBN** (*Uniform Resource Name:National Bibliography Number*)
  - Jedinstveni trajni identifikator digitalnog objekta koji se dodjeljuje objektima u repozitorijima u Dabru.
- **Verzioniranje**
  - Kontrola verzija istraživačkih podataka je proces upravljanja promjenama datoteke tijekom istraživanja ili projekta.

Popis tablica

Tablica 1. FAIR podaci ili Kako svoje istraživačke podatke uskladiti s FAIR principima? 6
Tablica 2. Elementi DMP-a 10
Tablica 3. Primjeri otvorenih i zatvorenih formata 15
Tablica 4. Popis smjernica za pouzdano čuvanje podataka tijekom istraživanja 18
Tablica 5. Podaci prije postupka anonimizacije 24
Tablica 6. Podaci nakon postupka anonimizacije 24
Tablica 7. Podaci prije postupka pseudoaninimizacije 25
Tablica 8. Podaci nakon postupka pseudoaninimizacije 25
Tablica 9. Podaci prije provedbe k-anonimizacije 26
Tablica 10. Podaci nakon provedbe k-anonimizacije 26

Popis slikovnih materija

Slika 1. Grafički prikaz životnoga ciklusa podataka 8
Slika 2. Primjer konvencije imenovanja istraživačkih podataka 12
Slika 3. Struktura mapa. Izvor: https://www.ukdataservice.ac.uk/manage-data/format/organising.aspx 13
Slika 4. Tagiranje datoteke u operativnom sustavu *Windows* 13
Slika 5. Tagiranje datoteke u sustavu Puh 14
Slika 6. *Readme* datoteka smješta se u vršni direktorij skupa podataka. Izvor: https://www.helsinki.fi/en/research/guide-for-data-documentation 16
Slika 7. Sučelje sustava Puh 19
Slika 8. Sučelje za unos skupa podataka sustava Dabar 30