

Plan upravljanja istraživačkim podacima - projekt Motivacije i prepreke za kolektivne akcije (MOTIKA)

Pavlović, Tomislav

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2022**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:284:482298>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom](#).

Download date / Datum preuzimanja: **2025-03-12**



Repository / Repozitorij:

[repository PILAR](#)

PLAN UPRAVLJANJA ISTRAŽIVAČKIM PODACIMA (PUP)

Opće informacije		
	Ime i prezime predlagatelja	Renata Franc
	Matična organizacija	Institut društvenih znanosti Ivo Pilar
	Naziv projekta	Motivacije i prepreke za kolektivne akcije
	Upravitelj podacima	Vanja Dergić (kvalitativni podaci) i Tomislav Pavlović (kvantitativni podaci)
1. Prikupljanje podataka i dokumentacija		
	Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite formate, vrste i opseg svih podataka s kojima ćete raditi, a ne samo krajnji skup podataka koji će biti rezultat istraživanja)	Podaci koji će se prikupljati su odgovori na: 1) Polu-strukturirane intervjuje na uzorku od oko 50 odraslih sudionika (25 na uzorku građana, a 25 na uzorku članova udruga). Audio podaci bit će pohranjeni u MP3 formatu, a transkripti u PDF formatu. Procijenjena veličina snimki i transkripata trebala bi biti manja od 10 GB. 2) <i>Online</i> ankete kreirane u 1KA aplikaciji. Prikupljat će se sociodemografski podaci i odgovori na upitnike na dvama uzorcima, a podaci će biti pohranjeni u .csv formatu. Svaki dokument u .csv formatu bit će praćen kodnim planom u tekstualnom obliku koji će sadržavati: 1) redni broj pitanja, 2) kod varijable u .csv bazi, 3) sadržaj pitanja te 4) moguć raspon odgovora uz opis značenja pojedine vrijednosti odgovora. U sklopu istraživanja bit će pripremljene dvije baze podataka. Planiramo pilot <i>online</i> anketna istraživanja na manjim prigodnim uzorcima sudionika te longitudinalno <i>online</i> anketno prikupljanje u tri vala s ukupnim početnim uzorkom od oko 4120 punoljetnih sudionika (3120 na uzorku građana, a 1000 na uzorku članova udruga) i planiranim otpadom sudionika u svakom idućem valu od po 50% na uzorku građana i 30% na uzorku članova udruga. Temeljem navedenog pretpostavljamo da bi veličina baze podataka prikupljenih na općem uzorku trebala biti manja od 5 MB, a veličina baze podataka prikupljenih na sudionicima iz različitih građanskih udruga manja od 2 MB.
	Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete, načine organiziranja podataka te alate i instrumente kojima ćete se koristiti za prikupljanje i obradu)	Kvalitativni podaci prikupit će se polustrukturiranim intervjuima na: 1) kvotnom uzorku odraslih građana, 2) prigodnom uzorku odraslih članova udruga. Snimljeni audio zapisi intervjuja bit će pohranjeni u poseban direktorij. Ti će podaci biti transkribirani te potom prebačeni u NVivo program u kojem će se analizirati. Anketni podaci će se prikupljati slanjem poveznice na <i>online</i> anketu preko aplikacije 1KA na: 1) prigodnim uzorcima odraslih /studenta, 2) kvotnom uzorku odraslih građana, te 3) članovima udruga. Nakon preuzimanja podataka s 1KA aplikacije, sirovi podaci će se pohraniti u poseban direktorij, te će se s njih ukloniti pravo editiranja (engl. <i>write permission</i>). Podaci će se čistiti koristeći programski jezik R, a pročišćeni podaci će se spremati u posebnu datoteku i zaseban direktorij. Skripta korištena za čišćenje podataka bit će pohranjena zajedno s podacima. Nakon čišćenja, podaci će se validirati kako bi se utvrdilo jesu li unosi valjani (npr. odgovaraju li rasponi vrijednosti na česticama pojedinih upitnika uporišnim točkama koje su bile zadane u aplikaciji 1KA). Čišćenje će biti provedeno u više koraka, a nakon svakog od njih će se inspekcijom tabličnih i grafičkih prikaza obrazaca odgovaranja na čestice utvrđivati koji su daljnji koraci potrebni kako bi prikupljeni rezultati odgovarali izvorno pripremljenom kodnom planu. Nakon završne iteracije, vizualnom inspekcijom provjerit će se podudarnost pročišćenih rezultata s

		izvornima kako bi se osiguralo da su izmjene uvedene samo za one vrijednosti koje budu izlazile iz kodnim planom predviđenih okvira. Istraživači će voditi ReadMe.md datoteku u kojoj će ukratko biti opisana struktura direktorija unutar projekta, te sadržaj pojedinih datoteka ili skupova datoteka.
	Koju ćete dokumentaciju i metapodatke izraditi osim podataka? (dokumentacija mora sadržavati informacije i standarde potrebne korisnicima kako bi mogli samostalno čitati i interpretirati podatke u budućnosti, primjerice, kodne knjige, <i>ReadMe</i> datoteke i sl.)	<p>Dokumentacija uz podatke bit će:</p> <ol style="list-style-type: none"> 1. scenarij za polustrukturirane intervju - pitanja korištena za prikupljanje podataka, 2. upitnici korišteni za prikupljanje podataka. <p>Također, dokumentacija će sadržavati i podatke o: autorima, ustanovi, kontaktu, istraživačkom projektu; pozivna pisma sudionicima, upute sudionicima, korištene protokole, te upute intervjuerima. Dokumentacija u vezi s istraživačkim projektom sadržavat će naslov, ključne riječi, kratki opis projekta, ciljeve, opis metodologije (uzorka sudionika i uzorkovanja, mjernih instrumenata, odnosno tema/pitanja u intervjuima, načinu prikupljanja podataka), stupnju odaziva, vremenskom trajanju istraživanja i lokaciji provođenja istraživanja, kao i prava korištenja. Izradit će se kodna knjiga s opisom varijabli i oznakama vrijednosti kao i <i>ReadMe</i> datoteke.</p>
2.	Pravna i sigurnosna pitanja	
	Jeste li ograničeni sporazumom o povjerljivosti? Imate li potrebna dopuštenja za prikupljanje, obradu, čuvanje i dijeljenje podataka? Jesu li osobe čiji se podaci obrađuju informirani o tome i jesu li dali privolu? Kojim ćete se metodama koristiti u svrhu zaštite osjetljivih podataka (GDPR - posebne kategorije osobnih podataka, navesti metode anonimizacije podataka)?	<p>U projektu će se poštovati sva ograničenja i zahtjevi utvrđeni Zakonom o provedbi Opće uredbe o zaštiti podataka (Uredba (EU) 2016/679 Europskog parlamenta i Vijeća od 27. travnja 2016. o zaštiti pojedinaca u vezi s obradom osobnih podataka i o slobodnom kretanju takvih podataka, te o stavljanju izvan snage Direktive 95/46/EZ), kao i drugi primjenjivi nacionalni zakoni o zaštiti osobnih podataka. Stoga, kako bi donijeli informiranu odluku o sudjelovanju u istraživanju, sudionici će najprije biti upoznati s pravilima privatnosti u cilju zaštite vlastitih prava i interesa, odnosno koji osobni podaci će se prikupljati, na koji način će biti obrađivani i u koje svrhe će se koristiti, kao i svojim pravima u vezi s osobnim podacima. Primjerice, svi sudionici će biti obaviješteni o: nazivu i voditeljici projekta, svrsi istraživanja, vrsti i opsegu prikupljenih/obrađenih podataka, načinima prikupljanja i analize podataka, načinima osiguranja zaštite podataka (npr. anonimizaciji), oblicima diseminacije i objave podataka te o upravljanju podacima nakon završetka projekta, kao i mogućnostima njihove buduće uporabe, o mogućnosti odustajanja sudjelovanja na ovom projektu u bilo koje vrijeme bez ikakvih posljedica, kao i u slučaju odbijanja sudjelovanja, o pravu pristupa podacima i povlačenju prava pristupa. Svaka komunikacija ostvarena s trećim stranama/primateljima/ i planirana prekogranična komunikacija bit će u skladu s postojećim zakonima i smjernicama Opće uredbe o zaštiti podataka. Osobni podaci će se obrađivati pošteno i zakonito, isključivo na temelju valjanih pravnih osnova i u skladu s informiranom suglasnošću sudionika u vezi sa sudjelovanjem u istraživanju te obradom i pohranom podataka.</p> <p>U svrhu zaštite podataka odgovori koji spadaju u posebne kategorije ili koji omogućuju identifikaciju sudionika će biti (pseudo)anonimizirani. (Pseudo)anonimizirane transkripte intervjua ćemo pohraniti na računalu zaštićenom lozinkom, a buduće dijeljenje podataka unutar članova istraživačkog tima bit će omogućeno temeljem pismene suglasnosti sudionika. S obzirom na provođenje pseudoanonimizacije, istraživači će imati pohranjenu kodnu knjigu koja će omogućavati ponovnu identifikaciju sudionika, a bit će pohranjena na sigurnom mjestu, dislociranom od prikupljenih podataka kako ne bi došlo do otkrivanja identiteta sudionika. Svi podaci o identitetu sudionika čuvat će se u enkriptiranom obliku (npr. uz korištenje alata BitLocker, FaulVault2), a ključ za dekripciju imat će samo uži članovi istraživačkog tima.</p>

	<p>Kako će se regulirati pristup podacima i njihova sigurnost? Koji su potencijalni rizici koje treba uzeti u obzir? Kako ćete osigurati sigurnost pohrane osjetljivih podataka?</p>	<p>Svi sudionici kvalitativnog (polustrukturirani intervjui) i kvantitativnih (<i>online</i> ankete) istraživanja dat će suglasnost za sudjelovanje u istraživanju, sadržaj kojeg će prethodno odobriti Etičko povjerenstvo Instituta društvenih znanosti Ivo Pilar. Kvantitativni podaci prikupljat će se putem 1KA aplikacije koja osigurava sigurnost podataka s obzirom da se komunikacija između korisnika i web stranice 1ka.si odvija putem SSL standarda (Secure Socket Layer) koji uz pomoć enkripcije osigurava da treća strana ne može pristupiti podacima. Štoviše, sigurnosno osjetljivi podaci, koji uključuju korisničke lozinke u 1KA bazi podataka, pohranjuju se u šifriranom obliku.</p> <p>Podaci će biti pohranjeni u centraliziranom sustavu za pohranu kojim upravlja informatički administrator Instituta društvenih znanosti Ivo Pilar. Pristup podacima će se upravljati preko identiteta ustanove, koji je siguran sustav i slijedi najbolje prakse u pogledu upravljanja identitetom. Naš centralni sustav pohranjivanja podataka ima dostatnu zalihost, zaštićen je antivirusnim programom koji se redovito nadograđuje, vrši se zrcaljenje i stalno je nadziran. Izradit će se kopije sirovih podataka koje će se pohraniti na više lokacija — na računalu, prijenosnom mediju i računalnom oblaku kako bi se umanjio rizik gubitka podataka zbog kvara ili gubitka uređaja. Usto, dokumentirat će se svi postupci u vezi s pohranom podataka u <i>ReadMe</i> datoteci. Izbjegavat će se slanje osjetljivih podataka elektroničkom poštom te će se umjesto toga koristiti enkriptirani kanali ili servisi u oblaku.</p>
	<p>Kako ćete upravljati zaštitom autorskih prava i drugog intelektualnog vlasništva? Tko će biti vlasnik podataka? Koje će se licence primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka?</p>	<p>Podaci iz triju valova istraživanja, prikupljeni na općem uzorku i uzorku sudionika iz udruga, s pripadnom dokumentacijom bit će pohranjeni u Hrvatskom arhivu podataka za društvene znanosti (CROSSDA; data.crossda.hr), članu Consortium of European Social Science Data Archives, te popraćeni globalnim identifikacijskim brojem (DOI). Podaci će biti pohranjeni pod licencom CC-BY 4.0 (otvoreno dostupni uz uvjet citiranja autora baze podataka), dok će metapodaci biti dostupni pod licencom Creative Commons Public Domain Dedication (CC0) te će sadržavati podatke o financiranju, broj odobrenja, ime projekta, akronim i broj projekta i uvjete licence. Tako dostupne baze podataka neće sadržavati osobne podatke.</p>
3.	Pohrana i čuvanje podataka	
	<p>Kako će radne verzije podataka biti pohranjene tijekom projekta? Kako će se napraviti sigurnosne kopije tih podataka (<i>backup</i>)? Koja je očekivana količina podataka koja će se prikupiti i čuvati tijekom projekta (izraženo u MB/GB/TB)?</p>	<p>Radne verzije podataka pohranit ćemo na više mjesta i izraditi sigurnosne kopije:</p> <ul style="list-style-type: none"> • na prijenosnom računalu Vanje Dergić (kvalitativni podaci) i Tomislava Pavlovića (kvantitativni podaci) • na institucionalnom SAN • na stolnom računalu Renata Franc <p>Vanja Dergić i Tomislav Pavlović bit će odgovorni za pohranu i sigurnosne kopije, koje će se raditi tjedno. Sigurnosne kopije na institucionalnoj infrastrukturi automatizirane su pomoću RSYNC alata.</p> <p>S obzirom na to da planiramo: provođenje polustrukturiranih intervjua s 50 sudionika, pilot <i>online</i> anketiranja na manjim uzorcima te longitudinalno <i>online</i> anketiranje u tri vala s ukupnim početnim uzorkom građana od oko 3120 i predviđenim otpadom sudionika u svakom idućem valu od po 50%, pretpostavljena veličina datoteke bi mogla iznositi do 5 MB. Baza koja bi sadržavala podatke prikupljene na oko 1000 članova udruga u prvom valu, uz predviđeni otpad sudionika u svakom od iduća dva valu po 30%, obujmom bi iznosila do 2 MB.</p>
	<p>Kako će se završne verzije podataka dugotrajno pohraniti i čuvati (i nakon</p>	<p>Kvantitativni podaci glavnog istraživanja (<i>online</i> ankete u 3 vala na uzorku građana i prigodnom uzorku članova udruga) u obliku dvije završne baze (jedne s podacima 3 vala <i>online</i> anketiranja uzorka građana te druge s podacima 3 vala <i>online</i> anketiranja prigodnog uzorka članova udruga) bit će trajno pohranjeni. Obje završne baze zajedno s</p>

	završetka projekta)? U kojim će se formatima čuvati podaci? Koja je očekivana količina podataka koja će se trajno pohraniti (izraženo u MB/GB/TB)?	pripadnim dokumentima trajno ćemo pohraniti u Hrvatski arhiv podataka za društvene znanosti (CROSSDA; Data.crossda.hr) koji se bavi pripremom skupova podataka za dugoročnu pohranu. Skup podataka čuvat će se u otvorenim formatima za dugoročnu pohranu, uz detaljni metapodatkovni opis u DDI standardu, a u mapi će biti prisutan i <i>ReadMe</i> dokument s kratkih opisom drugih dokumenata. Skup kvalitativnih podataka (anonimizirani transkripti) s popratnom dokumentacijom bit će pohranjen i dostupan pet godina od datuma završetka projekta. Očekivana ukupna količina podataka koja će se trajno pohraniti trebala bi iznositi manje od 10 GB.
4.	Dijeljenje i ponovna uporaba podataka	
	Kako i gdje će se podaci dijeliti? Koji repozitorij će se koristiti za dijeljenje podataka? Kako će potencijalni korisnici doznati za podatke?	Podaci i dokumentacija potrebna za interpretaciju podataka bit će dostupni u Hrvatskom arhivu podataka za društvene znanosti (CROSSDA; data.crossda.hr) pod licencom. Podaci će biti pohranjeni pod licencom CC-BY 4.0 (otvoreno dostupni uz uvjet citiranja autora baze podataka), dok će metapodaci biti dostupni pod licencom Creative Commons Public Domain Dedication (CC0). Međunarodna Arhiv CROSSDA je nacionalni pružatelj usluga za Konzorcij europskih arhiva podataka za društvene znanosti (CESSDA ERIC). Time se povećava međunarodna vidljivost skupova podataka relevantnoj međunarodnoj zajednici istraživača. CESSDA Data Catalogue (datacatalogue.cessda.eu) uključen je u EOSC portal. Detaljno pretraživanje skupova podataka omogućeno je upotrebom DDI standarda za metapodatke, uz upotrebu kontroliranih rječnika i tezaurusa ELSST (elsst.cessda.eu) za podatke nastale u društvenim istraživanjima. Skupovima podataka dodjeljuje se trajni identifikator iz sustava DOI.
	Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.	Svi osobni podaci bit će anonimizirani u skladu s profesionalnim standardima i Zakonom o provedbi Opće uredbe o zaštiti podataka. Paket SDCMicro (https://cran.r-project.org/package=sdcMicro) koristit će se za procjenu rizika identifikacije. Osigurat ćemo da svaki skup podataka ima k-anonimnost najmanje 3.
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima FAIR-a.	Arhiv CROSSDA postupa s podacima prema FAIR načelima.
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji održava neprofitna organizacija (ako ne, objasnite zašto ne možete dijeliti podatke na digitalnom repozitoriju koji nije komercijalan).	Arhiv CROSSDA dio je nacionalne javne istraživačke infrastrukture za područje društvenih znanosti. Osnovan je pri Filozofskom fakultetu Sveučilišta u Zagrebu uz financijsku potporu Ministarstva znanosti i obrazovanja u ulozi pružatelja usluga za Konzorcij europskih arhiva za društvene znanosti (CESSDA ERIC).