

Plan upravljanja istraživačkim podacima projekta EnergoDem

Ivanda, Krešimir; Strmota, Marin

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2022**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:148:208174>

Rights / Prava: [Public Domain Dedication](#)/[Prenošenje u javno dobro](#)

Download date / Datum preuzimanja: **2024-07-22**



Repository / Repozitorij:

[REPEFZG - Digital Repository - Faculty of Economics
& Business Zagreb](#)



Plan upravljanja istraživačkim podacima

Opće informacije		
	Ime i prezime predlagatelja	Marin Strmota
	Matična organizacija	Ekonomski fakultet, Sveučilište u Zagrebu
	Naziv projekta	Sociodemografska odrednice potrošnje energije u Hrvatskoj (EnergoDem)
	Upravitelj podacima	Marin Strmota, mstrmota@efzg.hr
1.	Prikupljanje podataka i dokumentacija	
	Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka)	<p>Podaci koji će se prikupljati su mikropodaci (podaci na razini pojedinca i kućanstva) koji će biti prikupljeni na više načina i iz više izvora.</p> <p>1.) Anonimizirani mikropodaci kupljeni od Državnog zavoda za statistiku i to Anketa o dohotku stanovništva (ADS), Anketa o potrošnji kućanstva (APK) i Anketa o radnoj snazi (ARS). Podaci će biti prikupljeni u formatu kojeg isporuči Državni zavod za statistiku, a u dosadašnjoj praksi to su bile CSV, TXT i SAS7BDAT datoteke. Podaci će biti nabavljani na godišnjoj razini odnosno u intervalima u kojima budu dostupni obzirom na vrijeme provođenja terenske ankete i obrade od strane DZS-a. Svaka pojedinačna baza podataka (ADK, APK, ARS) varira u veličini uzorka, a kreće se u rasponu od 8 000 do 38 000 te shodno tome i veličina datoteka varira od 3 do 41 MB.</p> <p>2.) Podaci prikupljeni kroz terensko istraživanje odnosno anketni upitnik. Terensko istraživanje će, po radnom planu, provoditi agencija specijalizirana za provedbu anketnog istraživanja. Podaci koji se prikupljaju su demografski i socioekonomski podaci o članovima kućanstva (broj članova, dob, spol, obrazovanje, prihodi) te podaci o potrošnji energije (izvori energije, troškovi, trošila u kućanstvu, obrasci potrošnje). Podaci će biti isporučeni u strojno čitljivom obliku u CSV datoteci. Očekivana veličina uzorka je do 1000 osoba, a broj varijabli do 40 pa će i očekivana veličina datoteke biti do 20 MB.</p>
	Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka)	<p>Podaci su kupljeni i nastavit će biti periodično nabavljani od Državnog zavoda za statistiku. Podaci se istraživačkom timu dostavljaju emailom ili na cd-u. Podaci DZS-a su anonimizirani (bez mogućnosti identifikacije pojedinca ili kućanstva). Podaci će se obrađivati i analizirati u programskom jeziku R, a finalni output analize bit će tablice, grafikoni, slike i tekst objavljeni u znanstvenim publikacijama.</p> <p>Obrada podataka u jeziku R radit će se kroz projektnu datoteku tj. sirovi podaci, R skripta i output analize nalazit će se u istom direktoriju. U obradi podataka koristit će se različite metode statističke analize (deskriptivna i inferencijalna statistika, konkretno modeli linearne i logističke regresije). Podaci</p>

		<p>DZS-a uz sirove podatke dolaze i sa popratnim datotekama (XLSX formata) u kojima se nalaze opisi varijabli iz sirovih podataka.</p> <p>1) Podaci prikupljeni kroz terensko istraživanje će se prikupljati kroz anketno istraživanje CAPI tipa koje će provoditi specijalizirana agencija za provedbu anketnih/terenskih istraživanja. Istraživački tim će kreirati sami upitnik odnosno set varijabli od interesa. Od agencije očekujemo isporuku konačnih rezultata u strojno čitljivom formatu (preferiramo CSV datoteku) te dodatnu datoteku s opisom varijabli (u TXT ili XLSX formatu). Istraživači će potom podatke analizirati u R programskom jeziku. Obrada podataka u jeziku R radit će se kroz projektnu datoteku tj. sirovi podaci, R skripta i output analize nalazit će se u istom direktoriju. U obradi podataka koristit će se različite metode statističke analize (deskriptivna i inferencijalna statistika, konkretno modeli linearne i logističke regresije).</p>
	<p>Koju ćete dokumentaciju i metapodatke ustupiti osim podataka? (navedite koje su informacije potrebne korisnicima kako bi mogli čitati i interpretirati podatke u budućnosti te koji će se standardi koristiti pri tumačenju podataka)</p>	<p>Osim sirovih podataka, istraživački tim izradit će upitnik koji će se koristiti u terenskom/anketnom istraživanju, pozivna pisma poslana sudionicima terenskog istraživanja te upute sudionicima i anketarima. Dokumentacija će sadržavati opis i ciljeve projekta EnergoDem, ulogu terenskog istraživanja, informacije o anonimnosti ispitanika i informacije o mjestu objave rezultata koji će uslijediti po analizi anketnog istraživanja.</p>
2.	Pravna i sigurnosna pitanja	
	<p>Jeste li ograničeni sporazumom o povjerljivosti? Imate li potrebna dopuštenja za prikupljanje, obradu, čuvanje i dijeljenje podataka? Jesu li osobe čiji se podaci pohranjuju informirani o tome i jesu li dali privolu? Kojim ćete se metodama koristiti u svrhu zaštite osjetljivih podataka (GDPR - posebne kategorije osobnih podataka)?</p>	<p>Pri izvedbi ovog projekta neće se kršiti etička načela.</p> <p>Podaci kupljeni od DZS-a su ograničeni sporazumom o povjerljivosti koji su potpisali voditelj projekta Marin Strmota kao zaposlenik institucije Ekonomski fakultet Sveučilišta u Zagrebu i voditelj projekta EnergoDem te Državni zavod za statistiku. Prije same nabave podataka od DZS-a istraživački tim je predao prijedlog istraživanja odnosno zahtjev za korištenjem povjerljivih podataka u kojem je definirao način korištenja i zaštite podataka. DZS je potom na svojim internim tijelima i procedurama odobrio zahtjev te omogućio i konačno isporučio podatke istraživačkom timu. Kupljeni podaci su u potpunosti anonimizirani te istraživači ne mogu identificirati pojedince ili kućanstva u anketama koje su predmet analize (ARS, APK, ADS).</p> <p>Inicijalni postupak zaštite osjetljivih podataka i anonimizacije je u domeni Državnog zavoda za statistiku te je detaljno opisan u metodologiji svake pojedine nabavljene ankete te internim aktima DZS-a. Podatke</p>

		kupljene od DZS-a članovi istraživačkog tima ne smiju dijeliti s trećim osobama, davati u javnost u izvornom obliku ili koristiti na bilo koji način koji nije propisan sporazumom o povjerljivosti.
	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?	Podaci će se pohraniti u centraliziranom sustavu za pohranu kojim upravlja Odjel za informatiku naše organizacije Ekonomski fakultet Sveučilišta u Zagrebu putem @net.efzg.hr individualnog identiteta povezanog s AAI identitetom. Pristup podacima se upravlja preko identiteta ustanove, koji je siguran sustav i slijedi najbolje prakse u pogledu upravljanja identitetom. Naš centralni sustav pohranjivanja podataka ima dostatnu zalihost, vrši se zrcaljenje i stalno je nadziran.
	Kako ćete upravljati zaštitom autorskih prava i intelektualnog vlasništva? Tko će biti vlasnik podataka? Koje će se licencije primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka?	Ne očekuje se da će rezultat istraživanja dovesti do patenta. Ostali problemi intelektualnog vlasništva će se u rješavati prema preporukama Ekonomskog fakulteta Sveučilišta u Zagrebu. Kupljeni podaci od DZS-a dostupni su samo istraživačkom timu i to isključivo u znanstvene svrhe prema sporazumu o povjerljivosti.
3.	Pohrana i čuvanje podataka	
	Kako će podaci biti pohranjeni i kako će biti napravljena sigurnosna kopija podataka (<i>backup</i>) tijekom istraživanja? Koji su kapaciteti čuvanja podataka kojim raspolazete? Kojim se procedurama koristite za sigurnosnu kopiju (<i>backup</i>)?	Podaci će se pohraniti u dijeljenom direktoriju na centraliziranom sustavu institucionalne infrastrukture Google Workspace for Education kojim upravlja Služba za informatiku Ekonomskog fakulteta u Zagrebu kojim će pristup imati samo članovi istraživačkog tima. Sigurnosne kopije sirovih podataka prije bilo kakve analize ili transformacije same baze ćemo pohraniti na računalu istraživača: Marin Strmota Krešimir Ivanda Inicijalna baza podataka neće se mijenjati već će promjene same baze podataka proizlaziti kroz R skriptu kojom će se transformirati podaci kako bi se provodila analiza.
	Koji je vaš plan čuvanja podataka? U kojim će se formatima čuvati?	Podatke ćemo čuvati u institucijskom repozitoriju uspostavljenom na sustavu Google Workspace for Education (konkretno Google Disk). Tablične podatke čuvat ćemo u CSV, TXT i XLSX obliku, a tekstualne u DOXC i PDF obliku.
4.	Dijeljenje i ponovna uporaba podataka	
	Kako i gdje će se podaci dijeliti? Na kojem repozitoriju planirate dijeliti	Podaci se neće dijeliti jer prema sporazumu o povjerljivosti s DZS-om podaci neće biti dostupni javnosti ni drugim istraživačima koji nisu dio istraživačkog tima HRZZ EnergoDem projekta.

	podatke? Kako će potencijalni korisnici doznati za podatke?	
	Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.	Sporazum o povjerljivosti s DZS-om koji je inicijalni kreator podataka ne dopušta dijeljenje podataka.
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima <i>FAIR-a</i> .	Podaci se neće dijeliti jer prema sporazumu o povjerljivosti s DZS-om podaci neće biti dostupni javnosti ni drugim istraživačima koji nisu dio istraživačkog tima HRZZ EnergoDem projekta.
	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).	Podaci se neće dijeliti jer prema sporazumu o povjerljivosti s DZS-om podaci neće biti dostupni javnosti ni drugim istraživačima koji nisu dio istraživačkog tima HRZZ EnergoDem projekta.

Ref:

[1] Celjak, D., Dorotić Malič, I., Matijević, M., Poljak, Lj., Posavec K. i Turk, I.: „Istraživački podaci - što s njima?“ [Istraživački podaci - što s njima?: priručnik o upravljanju istraživačkim podacima | Digitalni repozitorij Srca \(unizg.hr\)](#)