

Plan upravljanja istraživačkim podacima na projektu Hrvatske zaklade za znanost HELA

Begičević, Nina

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2023**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:211:113015>

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

Download date / Datum preuzimanja: **2024-06-30**



Repository / Repozitorij:

[Faculty of Organization and Informatics - Digital Repository](#)



Plan upravljanja istraživačkim podacima

Opće informacije		
	Ime i prezime predlagatelja	Nina Begičević Ređep
	Matična organizacija	Fakultet organizacije i informatike
	Naziv projekta	Podizanje zrelosti visokih učilišta za implementaciju analitika učenja – HELA
	Upravitelj podacima	Laura Strupar
1.	Prikupljanje podataka i dokumentacija	
	Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka)	<p>Prikupljat ćemo i generirati nekoliko vrsta podataka.</p> <ol style="list-style-type: none"> 1. Tablica s referencama znanstvenih radova (preuzeto iz znanstvenih baza podataka, dopunjeno metapodacima generiranim tijekom analize) (format CSV, max 1 GB) 2. Cjeloviti tekstovi znanstvenih radova – preuzeti iz izvora i predobrađeni za analizu metodama analize prirodnog jezika (neg. <i>natural language processing</i>) (u TXT formatu, UTF-8 encoding, max 2 GB) 3. Matrice frekvencije riječi po dokumentima na temelju modela vreća riječi i modela udpipe (po vrstama riječi i funkciji) (RData format, max 3GB) 4. Tablica revizija slučajeva upotrebe LA (format XLSX, max 1 MB) 5. Podaci prikupljeni od eksperata, kroz npr. fokus grupe, AHP, Delphi, sortiranje karata, upitnike (format XLSX, max 20 MB) 6. Podaci prikupljeni u okviru slučajeva primjene analitike učenja (CSV ili XLSX, max 100 MB)
	Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka)	<p>Podaci pod 1.) i 2) će nastati preuzimanjem podataka iz baza podataka, odnosno relevantnih publikacija. Valjanost izbora radova će se osigurati kroz uključenost najmanje dvoje istraživača u selekciju radova po zadanim kriterijima. Metapodaci će nastati temeljem analize tekstova radova. Predobrada dokumenata za podatke pod 3) će se provesti po zadanom protokolu. Matrica frekvencija će se generirati standardnim metodama predobrade teksta i analize prirodnog jezika u jeziku R. Tablica revizija slučajeva pod 4) će se provesti na način da će dvoje istraživača nezavisno primijeniti zadani protokol na ekstrakciju metapodataka iz izabranog skupa radova. Podaci pod 5) će se prikupiti putem upitnika, fokus grupa i drugih metoda elicitacije ekspertnog znanja, a pod 6) u okviru primjene analitike učenja iz sustava za upravljanje učenjem i administrativnih sustava te putem anketnih upitnika.</p> <p>Podaci će biti organizirani u direktorije koji prate strukturu ciljeva i isporuka projekta, uz dogovorenu shemu imenovanja datoteka i opis njihovog sadržaja.</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)	Uz svaki skup podataka kreirati će se i skup metapodataka sukladno DataCite Metadata V4.4. Metapodaci će sadržavati obvezne elemente: Stvaratelj: prezime, ime Naslov: slobodni tekst Nakladnik: FOI Godina objave Vrsta izvora: skup podataka Opis: slobodni tekst sa svim informacijama potrebnim za interpretaciju podataka Identifikator: DOI (nakon pohrane u vanjski repozitorij podataka) Pravo pristupa: ovisno o pravima intelektualnog vlasništva i suglasnosti ispitanika neki podaci će biti dostupni u otvorenom pristupu, a neki u zatvorenom.
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 pohranjuju informirani o tome i jesu li dali privolu? Kojim ćete se metodama koristiti u svrhu zaštite osjetljivih podataka (GDPR - posebne kategorije osobnih podataka)?	Dio podataka se odnosi na cjelovite tekstove znanstvenih radova od kojih su neki zaštićeni autorskim pravima, a neki dostupni u otvorenom pristupu. Podatke pod 2) ne možemo dijeliti zbog zaštite autorskog prava. Podaci pod 5) i 6) će biti anonimizirani kako bi se zaštili osobni podaci. Institucijsko etičko povjerenstvo dalo je pozitivno mišljenje o prikupljanju podataka na projektu. Sudjelovanje će biti dobrovoljno, a Ispitanici će dati informiranu privolu za istraživanje i ponovnu upotrebu podataka.
	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 za vrijeme trajanja projekta biti pohranjeni u sustavu Google drive. U okviru servisa je osigurano redovito pohranjivanje sigurnosnih kopija podataka. Na osobnim računalima istraživača neće se pohranjivati osobni podaci ispitanika, nego samo anonimizirani podaci koji su potrebni pojedinom istraživaču za provedbu istraživanja.
	Kako ćete upravljati zaštitom autorskih prava i intelektualnog vlasništva? Tko će biti vlasnik podataka? Koje će se licencije	Vlasnik podataka je FOI. Autori znanstvenih radova zadržavaju autorsko pravo, osim ako se rad objavljuje u časopisu koji traži prijenos autorskog prava na nakladnika. U okviru odobrenih financija na projektu dio radova će biti objavljen u "golden" otvorenom pristupu. Neće se dijeliti osobni podaci, nego samo anonimizirani podaci ili podaci koji se ne odnose na pojedine osobe. Na podatke koji će se dijeliti primijeniti će se Creative Commons licencije.

	primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka?	
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 za vrijeme istraživanja biti pohranjeni u sustavu Google Drive i MS Teams Share Point. Davatelji usluga redovito izrađuje sigurnosne kopije. Kapacitet čuvanja podataka višestruko nadilazi potrebe projekta. Kopija pojedinih podataka će se čuvati na računalima istraživača koji rade na pojedinim isporukama.
	Koji je vaš plan čuvanja podataka? U kojim će se formatima čuvati?	Podatke ćemo čuvati najmanje tri godine nakon završetka projekta u institucijskoj riznici dokumenata u nacionalnom sustavu Dabar. Ovisno o vrsti podataka podaci će se čuvati u formatima CSV, TXT, XLSX i RData
4.	Dijeljenje i ponovna uporaba podataka	
	Kako i gdje će se podaci dijeliti? Na kojem repozitoriju planirate dijeliti podatke? Kako će potencijalni korisnici doznati za podatke?	Konačnu verziju podataka koje je moguće dijeliti ćemo podijeliti putem repozitorija otvorenih istraživačkih podataka Harvard DataVerse, a kopije ćemo postaviti i u institucijski repozitorij FOI u nacionalnom sustavu Dabar i Harvard DataVerse. Oba sustava osiguravaju FAIR načela pohrane istraživačkih podataka.
	Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.	Nije moguće dijeliti podatke koji su nastali predobradom teksta znanstvenih radova prije primjene metoda analize prirodnog jezika, zato što je sadržaj djelomično zaštićen autorskim pravima nakladnika.
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima <i>FAIR-a</i> .	DA
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji održava neprofitna organizacija	DA

(ako ne, objasnite zašto ne možete dijeliti podatke na digitalnom repozitoriju koji nije komercijalan).	
---	--

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 \(unizg.hr\)](#)