

Plan upravljanja istraživačkim podacima, MeRID

Palmović, Marijan

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2024**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:158:316473>

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

Download date / Datum preuzimanja: **2024-09-04**



Repository / Repozitorij:

[Faculty of Education and Rehabilitation Sciences -
Digital Repository](#)



Plan upravljanja istraživačkim podacima

Opće informacije		
	Ime i prezime predlagatelja	Marijan Palmović
	Matična organizacija	Sveučilište u Zagrebu, Edukacijsko-rehabilitacijski fakultet
	Naziv projekta	Pouzdanost mjerenja individualnih razlika u rečeničnoj obradi: Međujezična perspektiva (MeRID)
	Upravitelj podacima	Marijan Palmović, marijan.palmovic@erf.unizg.hr
1.	Prikupljanje podataka i dokumentacija	
	Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka)	Prikupljat će se podaci mjerenja pokreta očiju na korpusu tekstova različitih registara. Uz to, prikupljat će se bihevioralni podaci (mjerenje brzine čitanja vlastitim tempom, engl. 'self-paced reading') na tim istim tekstovima. Konačno, prikupit će se podaci dobiveni eksperimentom u kojem ispitanici čitaju pojedinačne rečenice. Sirovi podaci su u specifičnom formatu uređaja (sada koristimo SMI-IView500 eyetracker koji sprema podatke u svoj format, tj. .idf datoteke). Bihevioralni se podaci iz softwera za prikazivanje (Paradigm Software) eksportiraju u .xls formatu. Veličina datoteke za pojedinog ispitanika je cca 5MB, a planiramo uključiti 100 ispitanika, dakle, ukupno oko 500MB podataka. Sama obrada podataka uključuje stvaranje baze podataka s različitim prikazima podataka (fiksacije, sakade, zone fokusa) i, prema procjeni, za ovaj broj sudionika može biti oko 2GB. Veličina datoteka na čitanju vlastitim tempom je oko 2MB i zanemariva je za probleme s pohranom. Sudionici će ispunjavati upitnik s pitanjima relevantnim za kriterije uključivanja (npr. dob, spol, obrazovanje, dioptrija).
	Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka)	Podaci će se prikupljati u Laboratoriju za psiholingvistička istraživanja, Sveučilišni kampus Borongaj. Laboratorij je opremljen uređajem za mjerenje pokreta očiju SMI iView500 koji je već zastario tako da se projektom planira nabava odgovarajućeg uređaja (tj. istog onog koji ima švicarska strana projekta), a koji je danas standard za takvu vrstu istraživanja (Eyelink Duo) i ima dovoljnu rezoluciju (2kHz). Svaki istraživački uređaj za mjerenje pokreta oka (i postojeći i planirani) ima odgovarajući software za obradu prikupljenih podataka, od procjene kvalitete (u kojoj je mjeri sudionikov pogled ostao "u kalibraciji", koji je postotak praćenja oka, ispravak artefakata (npr. treptanje) do obrade kvantitativnih podataka (broj fiksacija, duljina fiksacija, fiksacije u zonama interesa, sakade itd.). Taj software automatski integrira sve prikupljene pojedinačne podatke svakog sudionika u jedinstvenu bazu podataka. Numerički se podaci eksportiraju u .txt ili .csf radi statističke obrade u programu R. Na isti se način analiziraju i bihevioralni podaci.
	Koju ćete dokumentaciju i metapodatke ustupiti osim podataka? (navedite koje su informacije potrebne korisnicima kako bi mogli čitati i interpretirati	Podatke mjerenja pokreta očiju nije moguće koristiti bez odgovarajućih programskih licenci. Slikovne podatke moguće je eksportirati radi prikaza u publikacijama ili ih jednostavno učiniti dostupnima kao ilustraciju prikupljenih podataka. Za istraživačke paradigme koje namjeravamo koristiti (paradigma čitanja) to je u pravilu tzv. scanpath, tj. putanja kretanja oka od fiksacije do fiksacije s vrijednostima pojedinih fiksacija ili sakada u milisekundama ili stupnjevima. Kvantificirane podatke moguće je ustupiti kao R Markdown datoteke u kojima se nalaze i kodovi u R-u, što je najjednostavniji način za replikaciju istraživanja i rezultata za druge istraživače. U takve je datoteke moguće (i

	podatke u budućnosti te koji će se standardi koristiti pri tumačenju podataka)	planirano) ugraditi osnovne podatke o projektu, skupne podatke o sudionicima i njihovom odazivu i opis mjerenih varijabli.
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)?	Nismo ograničeni nikakvim sporazumom o povjerljivosti. Istraživanje je dobilo dopuštenje Etičkog povjerenstva ERF-a i Sveučilišta u Zürichu. Projekt poštuje sva ograničenja i zahtjeve kako je utvrđeno Općom uredbom o zaštiti osobnih podataka (GDPR). Od svakog će se sudionika tražiti da potpiše obaviješteni pristanak, ali se ime sudionika nigdje drugdje ne pojavljuje tj. svi su podaci u istraživanju anonimni i ne mogu dovesti do identifikacije sudionika. Formular obaviještenog pristanka sastavljen je prema uzusima u području psiholingvistike ili primijenjene lingvistike (Cergol, K. 2021. Etika istraživanja u primijenjenoj lingvistici, Zagreb: Srednja Europa) i sadrži osnovne obavijesti i kontakt istraživača (radi pristupa svojim podacima ili dodatnim obavijestima), kratak opis istraživanja i metode, opis upotrebe podataka te obavijesti o mogućnosti odustajanja bez posljedica. Same metode koje se upotrebljavaju u istraživanju smatraju se posve neinvazivnim i neškodljivim.
	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?	Sve datoteke s prikupljenim podacima čuvaju se u formatima koji su nečitljivi bez softverskog ključa (USB dongle). Sami podaci su na diskovima računala šifrirani. Dodatno, sva oprema nalazi se u Laboratoriju u koji se ne može ući bez elektronske propusnice i ključa, a zgrada ima zaštitara 24 sata. Pristup podacima imaju oni istraživači koji sudjeluju na projektu. Istraživači koji ne sudjeluju na projektu, a imaju pristup opremi, neće imati lozinku za podatke ovog istraživanja u samom softveru uređaja. Uređaj za praćenje pokreta oka nije na mreži, tj. nije spojen na Internet.
	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?	Podaci su prikladni za dijeljenje. Radi se o rezultatima mjerenja pokreta oka ili rezultatima vremena čitanja vlastitim tempom i mogu se koristiti u drugim analizama ili u metaanalizama. Svrha stvaranja korpusa tekstova označenih s psiholingvističkim parametrima kao što je brzina čitanja od riječi do riječi ili duljina fiksacija, regresija i sl. i jest mogućnost da se ti podaci koriste za različite svrhe i da ga koriste drugi istraživači u svijetu.
3.	Pohrana i čuvanje podataka	
	Kako će podaci biti pohranjeni i kako će biti napravljena sigurnosna	Podaci će se tijekom istraživanja pohranjivati na diskovima računala koje upravlja uređajem za mjerenje pokreta oka i računala na kojem će se prikupljati bihevioralni podaci. Sam software uređaja za mjerenje pokreta oka ima modul

	kopija podataka (<i>backup</i>) tijekom istraživanja? Koji su kapaciteti čuvanja podataka kojim raspolazete? Kojim se procedurama koristite za sigurnosnu kopiju (<i>backup</i>)?	koji radi backup čitave prikupljene baze i sprema je u komprimiranu datoteku koja se preseljava na drugi disk. Ta se sigurnosna kopija ne može "raspakirati" bez odgovarajućeg programa i softverskog ključa.
	Koji je vaš plan čuvanja podataka? U kojim će se formatima čuvati?	Kao što je navedeno, sami uređaji generiraju datoteke u svojim specifičnim formatima koji nisu čitljivi bez licenciranog softwera i ključa. Statistički će se podaci čuvati u .txt ili .csv, te .html formatima (R Markdown datoteke). Članci će se čuvati u .pdf formatu. Kao objavljeni korpus podaci će biti trajno dostupni na internetu.
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?	Podaci će se dijeliti preko odgovarajućeg repozitorija Sveučilišta u Zürichu. Odgovarajući podaci za pristup bit će navedeni u publikacijama i na internetskoj stranici projekta. Budući da su svi istraživači na ovome projektu uključeni u "networking" projekt COST sa sličnom temom, podaci o ovome projektu dijelit će se i putem tog projekta koji, sam, ima odgovarajuću web stranicu i internetske foruma (Slack).
	Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.	Svi podaci potrebni za publiciranje znanstvenih radova bit će dostupni u trenutku objavljivanja s obzirom da se radi o vrsti jezičnog korpusa. Svi su podaci anonimizirani i poštivat će se sva načela GDPR. Dodatno, tijekom projekta razvit će se sustav kodiranja psiholingvističkih podataka koji čine označavanje jezičnoga korpusa zbirno, ne za pojedinačne sudionike (npr. postotci fiksiranja ili prosječne duljine fiksacija). Same baze podataka na uređajima ne mogu biti dostupne zbog softverskih ograničenja.
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima <i>FAIR-a</i> .	Digitalnim podacima koristi ćemo se u skladu s načelima FAIR-a.
	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).	Sam jezični korpus, označen bihevioralnim, tj. psiholingvističkim podacima (eyetracking+self-paced reading) bit će javno dostupan, za što je zaduženo Sveučilište u Zürichu.

Ref:

[1] Celjak, D., Dorotić Malić, 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 \(unizg.hr\)](#)