

Plan upravljanja istraživačkim podacima

Opće informacije		
	Ime predlagatelja i broj projekta	Upišite ime glavnog istraživača
	Matična organizacija	
	Naziv i šifra projekta	
	Upravitelj podacima	upišite ime prezime te e-adresu osobe koja je odgovorna za upravljanje podacima i <i>Planom upravljanja istraživačkih podataka</i>
I	Prikupljanje podataka i dokumentacija	
	Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka)	Primjer: Podaci koji će se prikupljati su odgovori na upitnik kreiran u <i>LimeSurvey</i> aplikaciji. Podaci koji se prikupljaju su: dob, spol, županija u kojoj je osoba boravila većinu vremena u proteklih 6 mjeseci, te odgovori na IPIP-50 upitnik. Podaci će biti pohranjeni u CSV datoteci. S obzirom na to da planiramo prikupiti podatke od oko 1000 sudionika, veličina datoteke trebala bi iznositi maksimalno 10 MB.
	Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka)	Primjer: Podaci će se prikupljati ciljanim slanjem upitnika zainteresiranim skupinama sudionika. Nakon preuzimanja podataka s <i>LimeSurvey</i> servisa, sirovi podaci će se staviti u poseban direktorij, te će se s njih ukloniti pravo pisanja (eng. <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. jesu li vrijednosti dobi valjan broj manji od 100; odgovaraju li rasponi vrijednosti na česticama <i>iPIP</i> upitnika uporišnim točkama koje su bile zadane u aplikaciji <i>LimeSurvey</i>). Osim toga, provest će se ručna usporedba vrijednosti u datoteci s pročišćenim podacima i vrijednosti u datoteci sa sirovim podacima za 50 nasumično odabranih unosa. Istraživači će voditi <i>README.md</i> 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? (navedite koje su informacije potrebne korisnicima kako bi mogli čitati i interpretirati podatke u budućnosti te koji će se standardi koristiti pri tumačenju podataka)	Primjer: Dokumentacija uz podatke bit će upitnik korišten za prikupljanje podataka, pozivna pisma poslana sudionicima, upute sudionicima te upute anketarima. Dokumentacija će sadržavati opis projekta, ciljeve, ciljanu skupinu te informacije o uzorkovanju, jedinici analize, načinu prikupljanja podataka, stupnju odaziva, vremenskom te prostornom obuhvatu. Izradit ćemo kodnu knjigu s opisom varijabli i oznaka vrijednosti. Opisat ćemo upotrebu varijabli otežavanja te izvedenih varijabli nastalih nakon procesa prikupljanja podataka.

II	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 obrađuju 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>Primjer 1. Molim provjeriti uključuje li Vaš projekt jedno od ovih etičkih problema.</p> <ul style="list-style-type: none"> • sudionici (uključujući bilo koji oblik sudjelovanja osoba, uključujući i nemedicinska istraživanja, primjerice ankete, promatranja, praćenje lokacije ljudi i slično) • humane stanice/tkiva • humane embrionalne matične stanice • klinička istraživanja • zbirku osobnih/privatnih podataka • istraživanje na životinjama • treće zemlje (pristup i korist dijeljenja)/pravni problemi oko izvoza podataka. • okolišni i/ili zdravstveni i sigurnosni problemi (primjerice negativan utjecaj na okoliš i/ili na zdravlje i sigurnost istraživača) • potencijalna vojna primjena (tehnologija dvostruke namjene) <p>Ako smatrate da nema etičkih problema u Vašem projektu, tada se možete koristiti sljedećom izjavom: Pri izvedbi ovog projekta neće se kršiti etička načela.</p> <p>Primjer 2. Projekt poštuje sva ograničenja i zahtjeve kako je utvrđeno Zakonom o provedbi Opće uredbe o zaštiti podataka [navesti zakone]. Kraj projekta nije vezan za sudionike, te objavljeni rezultati ne smiju dovesti do njihove identifikacije. Zbog toga će svi sudionici biti obaviješteni o osnovnim informacijama:</p> <ul style="list-style-type: none"> • o autoru/odgovornoj osobi • vrsti i opsegu prikupljenih/obrađenih podataka • ciljevima obrade • svaka komunikacija ostvarena s trećim stranama/primateljima/planirana prekogranična komunikacija bit će u skladu s postojećim zakonima i smjernicama [navesti koje]. • o mogućnosti odustajanja sudjelovanja na ovom projektu u bilo koje vrijeme, bez ikakvih posljedica, kao i u slučaju odbijanja sudjelovanja. • pravo pristupa i ispravka prava
	<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>Primjer 1. Podaci će se obrađivati i njima upravljati u zaštićenom nemrežnom okruženju koristeći se virtualnom desktop tehnologijom.</p> <p>Primjer 2. Svi sugovornici i sudionici fokus grupa potpisat će suglasnost koje je odobrilo [naziv tijela]. Jamčili smo anonimnost u svih naših sudionika. Stoga nećemo pohranjivati <i>wav</i> datoteke, čime bi ugrozili to jamstvo, nego će one biti uništene koristeći softver za temeljito brisanje (<i>shred</i>). No, anonimizirane transkripte intervjua i fokus grupa ćemo pohraniti, za što ćemo osigurati pisanu suglasnost za buduće dijeljenje podataka. Svi podaci o identitetu sudionika čuvat će se u enkriptiranom obliku, a ključ za dekripciju imat će samo uži članovi istraživačkog tima.</p> <p>Primjer 3. Podaci će se pohraniti u centraliziranom sustavu za pohranu kojim upravlja Odjel za informatiku naše organizacije [naziv organizacije]. Pristup podacima se upravlja preko identiteta ustanove, koji je siguran sustav i</p>

		slijedi najbolje prakse u pogledu upravljanja identitetom. Naš centralni sustav pohranjivanja podataka ima dostatnu zalihost, vrši se zrcaljenje i stalno je nadzirano.
	Kako ćete upravljati zaštitom autorskih prava i 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>Primjer 1. Ne očekuje se da će rezultat istraživanja dovesti do patenta. Ostali problemi intelektualnog vlasništva će se u rješavati prema preporukama institucije [naziv institucije]. Budući da podaci nisu podvrgnuti ugovoru, te se neće patentirati obavit će se kao otvoreni podaci pod licencom <i>Creative Commons CC0</i>.</p> <p>Primjer 2. Podaci su prikladni za dijeljenje. Podaci su dobiveni promatranjem (dakle jedinstveni su) i mogli bi se koristiti za druge analize ili za usporedbe učinaka klimatskih promjena među mnogim stvarima. Prilike za novu upotrebu su velike. Zbog toga, naš cilj je da omogućimo najširu moguću upotrebu podataka i obavit ćemo ih pod <i>Creative Commons CC0</i>.</p> <p>Primjer 3. U projektu će se koristiti sekundarni podaci preuzeti iz repozitorija Hrvatskog arhiva podataka za društvene znanosti [navesti koji]. Ti podaci dostupni su samo za korištenje u znanstvene svrhe. Prije preuzimanja podataka registrirat ćemo se u sustav te ćemo se obvezati podatke koristiti samo za dozvoljene svrhe te da ćemo poštivati Opće uvjete korištenja podataka.</p>
III	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>)?	<p>Primjer 1. Podaci će se tijekom istraživanja s računala glavnog istraživača kopirati u nacionalni sustav za pohranu i dijeljenje podatka Puh (https://www.srce.unizg.hr/puh) koji članovima projektnog tima omogućava pristup aktualnoj verziji podataka i na kojem se dnevno automatizirano izrađuje sigurnosna kopija podataka. Uz to, glavni istraživač dnevno radi sigurnosnu kopiju s računala na vanjski disk.</p> <p>Primjer 2. Podatke ćemo pohraniti i izraditi sigurnosnu kopiju na tri mjesta:</p> <ul style="list-style-type: none"> • na prijenosnom računalu [Ime istraživača] • na institucionalnom SAN • na nacionalnom sustavu za pohranu i dijeljenje podatak PUH • na drugom mjestu [navesti gdje] <p>[Ime istraživača] bit će odgovoran za pohranu i sigurnosne kopije, koje će se raditi tjedno. Sigurnosne kopije na institucionalnoj infrastrukturi automatizirane su pomoću <i>RSYNC</i> alata.</p> <p>Primjer 3. Laboratorijski dnevnic i tiskane kopije <i>NMR</i>-a i masene spektroskopije čuvaju se u laboratoriju glavnog istraživača. Dodatni elektronički podaci pohranit će se na računalu glavnog istraživača, koji izrađuje sigurnosne kopije dnevno. Osim toga, istraživači će koristiti laboratorijski prostor na institucijskom repozitoriju za sekundarnu pohranu podataka. Laboratorij glavnog istraživača ima na raspolaganju 1TB prostora za pohranu podataka, koji može i dodatno proširiti. Svi podaci o projektu pohranit će se na institucionalnom repozitoriju koji redovito izrađuje sigurnosne kopije.</p>
	Koji je vaš plan čuvanja podataka nakon projekta? U kojim će se formatima čuvati?	Primjer 1. Podatke ćemo čuvati trajno u institucijskom repozitoriju [ustanova] uspostavljenom na sustavu Dabar. Tablične podatke čuvat ćemo u CSV obliku, a tekstualne u DOCX (<i>Office Open XML</i>) te PDF-a obliku. DOC oblik obavezno će se konvertirati u DOCX oblik.

		<p>Primjer 2. Podaci će se čuvati najmanje tri godine nakon završetka projekta, prema smjernicama donatora. Ako su nastali izumi ili nove tehnologije pristup podacima bit će ograničen do otkrivanja izuma i/ili patentne prijave, koja će se izvršiti preko Ureda za transfer tehnologija.</p> <p>Primjer 3. Podatke ćemo čuvati 10 godina na virtualnom poslužitelju kojeg [ustanova] ima u oblaku Srca i također pohraniti u odgovarajući repozitorij za podatke na kraju projekta (npr. <i>Zenodo</i>). Gdje bude moguće, datoteke ćemo pohraniti u otvorenim arhivskim formatima primjerice <i>word</i> dokumenti pretvorit će se u <i>PDF</i> ili kodirane u jednostavne tekstualne datoteke. <i>Excel</i> datoteke pretvorit će se u CSV oblik. U slučaju kad je to moguće uključit ćemo i informacije o korištenom softveru i broju njegove verzije.</p> <p>Primjer 4. Podatke ćemo trajno pohraniti u Hrvatskom arhivu 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. Skup podataka sadržavat će dovoljno dokumentacije za kasniju interpretaciju podataka i mogućnost ponovne upotrebe.</p>
IV	Dijeljenje i ponovna uporaba podataka	
	<p>Kako i gdje će se podaci dijeliti? Na kojem repozitoriju planirate dijeliti podatke? Kako će potencijalni korisnici doznati za podatke?</p>	<p>Primjer 1. Podaci i dokumentacija potrebna za interpretaciju podataka bit će dostupni u Hrvatskom arhivu podataka za društvene znanosti (CROSSDA; data.crossda.hr). Bit će dostupne dvije verzije podataka: 1) potpuno anonimizirana verzija bit će dostupna pod licencom <i>CC BY 4.0</i> Međunarodna. 2) verzija podataka koja sadrži rizik od deidentifikacije sudionika bit će dostupna uz posebnu dozvolu, samo za znanstvene svrhe, te će joj moći pristupiti samo registrirani istraživači uz prihvaćanje obveze o korištenju podataka isključivo u dozvoljene i Općih uvjeta korištenja podataka.</p> <p>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 <i>Data Catalogue</i> (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 <i>DOI</i>.</p> <p>Primjer 2. Konačnu verziju skupa podataka voditelj projekta podijelit će putem institucijskog repozitorija [ustanova] uspostavljenog u nacionalnom sustavu Dabar gdje će biti pohranjene i publikacije i ostala projektna dokumentacija. Podaci će biti objavljeni pod <i>CC0</i> licencom. Institucijski repozitorij u sustavu Dabar odabrali smo jer podržava FAIR načela: skupovima dodjeljuje trajni identifikator URN:NBN, osigurava vidljivost podataka putem OpenAIRE portala i Google Scholar a tražilice dabar.srce.hr, a ujedno doprinosi vidljivosti i transparentnosti rada [ustanova].</p> <p>Primjer 3. Neki od podataka koji se trenutno prikupljaju podijelit će [Ime istraživača] na <i>Github</i> repozitoriju (rezultati, broj projekta, istraživanje s <i>Twittera</i>). Glavna revizija ovih stranica bit će dovršene korištenjem <i>Github-Zenodo</i> poveznice (vidi: https://guides.github.com/activities/citable-code/). Svi ostali podaci objavit će se na <i>Zenodo</i> pod <i>CC0</i> licencom. <i>Zenodo</i> smo odabrali jer podržava FAIR načela (http://about.zenodo.org/principles/). Objava članka neposredno nakon završetka projekta ima za cilj smanjiti rizik od gubitka podataka, dok dvogodišnji embargo nam jamči prvenstvo korištenja podataka. <i>Zenodo</i> primjenjuje značajke dugoročne pohrane.</p>

<p>Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.</p>	<p>Primjer 1. Podaci neophodni za bilo koju publikaciju bit će dostupni u trenutku objavljivanja. Svi neobjavljeni podaci pohranit će se u repozitorij na 12 mjeseci od završetka projekta. Primjer 2. Astronomske podaci bit će distribuirani, ali pod embargom na godinu dana, zbog prioriteta upotrebe. Primjer 3. Svi osobni podaci bit će anonimizirani u skladu s profesionalnim standardima i Zakonom o provedbi Opće uredbe o zaštiti podataka. Paket <i>SDCMicro</i> (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.</p>
<p>Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima <i>FAIR data</i>.</p>	<p>Arhiv CROSSDA postupa s podacima prema načelima FAIR.</p>
<p>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).</p>	<p>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 ulozu pružatelja usluga za Konzorcij europskih arhiva za društvene znanosti (CESSDA ERIC).</p>