

Plan upravljanja istraživačkim podacima projekta DeSPERADo

Bradshaw, Nicholas James

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2023**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:193:133474>

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

Download date / Datum preuzimanja: **2024-12-24**



Repository / Repozitorij:

[Repository of the University of Rijeka, Faculty of Biotechnology and Drug Development - BIOTECHRI Repository](#)



PLAN UPRAVLJANJA ISTRAŽIVAČKIM PODACIMA (PUP)

Opće informacije		
	Ime i prezime predlagatelja	Nicholas J. Bradshaw
	Matična organizacija	Sveučilište u Rijeci, Odjel za biotehnologiju
	Naziv projekta	Depresija, samoubojstvo i proteinopatija: rasvjetljavanje odnosa između agregacije i patološkog razvoja (DeSPERADo)
	Upravitelj podacima	Nicholas J. Bradshaw
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)	<p>Podaci u Rijeci će se prikupljati u sljedećim formatima, ovisno o vrsti eksperimenta:</p> <ol style="list-style-type: none"> Podaci elektroforeze i Western blota. Slike će biti spremljene u .scn formatu koji koristi program ImageLab (obično do 1 MB), a također i kao .tiff ili .jpg slike (obično do 20 MB) za općenitije gledanje. Ukupno očekujemo 200 – 500 snimljenih slika. Sve membrane koje predstavljaju istraživačke eksperimente bit će potvrđene tri puta u zasebnim eksperimentima prije objavljivanja. Slike imunofluorescentne mikroskopije. Slike će biti spremljene u .vsi formatu koji koristi softver mikroskopa CellSens (50-200 KB), a također i kao .tiff ili .jpg (do 20 MB svaka) slike za općenitije gledanje. Ukupno očekujemo 1000 – 3000 snimljenih slika. Svi kvalitativni rezultati moraju se potvrditi tri puta u zasebnim eksperimentima prije objave. Dodatno, neki rezultati bit će kvantificirani u eksperimentima u kojima je istraživač zaslijepljen za status svakog skupa stanica (npr. koji se protein ondje eksprimira) i mora brojati stanice koje pokazuju specifične fenotipove. Ovi se podaci potom skupljaju prije dekodiranja i analize. Takvi podaci sortirani su u .xls formatu (5-20 takvih eksperimenata, 10-50 KB po datoteci) Podaci iz biofizičkih eksperimenata (u Rijeci i Jülichu) obično su u obliku .xls ili .txt datoteka. Očekuje se da će se generirati 100-300 takvih datoteka, obično manje od 10 MB svaka. Svi takvi eksperimenti koji predstavljaju istraživačke eksperimente bit će potvrđeni tri puta u zasebnim eksperimentima prije objave. Podaci o <i>Drosophili</i> prikupljaju se iz monitora aktivnosti te spremaju se kao datoteke u .txt formatu (do 900 kb svaki). Sirovi podaci su kasnije analizirani od strane naših suradnika iz grupe Rozi Andretic Waldowski i Fakulteta informatike i digitalnih tehnologija. Mi rezultate analize dobivamo u .csv formatu (do 20 kb svaki) koji se obrađuju u Excelu (i pretvaraju u .xls datoteke, do 200 kb svaki). Nadalje, svaka statistička analiza radi se u softveru GraphPad (sprema se kao izvorni format datoteke .pzfx, oko 900 kb; dok se rezultati analize mogu eksportirati u .txt format ili .jpg format za grafove, do 4000 kb). Iznimke su testovi poput geotaksije gdje se cijeli test snima kamerom mobilnog telefona (.jpg datoteke, do 4000 kb), no ti se podaci također analiziraju u Excelu i GraphPadu (isti formati datoteka kao što je prethodno navedeno).

<p>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)</p>	<p>Standardizirani formati naslova datoteka (s nazivima projekata, datumima itd.) Praćenje verzija s datumom (GGGGMMDD) Npr. 20221292_Labexpr_</p>
<p>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>	<p>Obučavat ćemo sve istraživače Koristit ćemo standardni format datoteke Stvari ćemo označiti verzijom ili datumom Upotreba <i>readme.txt</i> (kod) ili sličnog kako bi se sve lako pratilo – e.g, datumi itd.</p>
<p>2. Pravna i sigurnosna pitanja</p>	
<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, navesti metode anonimizacije podataka)?</p>	<p>Ne postoje nikakvi ugovori o povjerljivosti koji bi utjecali na ovaj rad, niti bilo koji drugi čimbenici koji bi nas spriječili u dijeljenju podataka.</p> <p>Samo jedno područje rada u Rijeci odnosi se na osobne podatke: analizu <i>post mortem</i> uzoraka mozga kontrolnih osoba, žrtava samoubojstva i pacijenata. Ovi uzorci dopremaju se sa Sveučilišta Semmelweis, Budimpešta, Mađarska (https://semmelweis.hu/hbtb/) na zajedničkoj, neprofitnoj osnovi. Uzorci su označeni samo jedinstvenim kodom za uzorak, jedinstvenim kodom za pojedinca, dijagnozu pojedinca i regiju mozga. Dodatno smo dobili .xls datoteku koja sadrži ove informacije, dob svake osobe u trenutku smrti, vrijeme između smrti i prikupljanja uzorka i pH vrijednosti uzorka. Uzorci se mogu uzeti s bilo kojeg mjesta u Mađarskoj.</p> <p>Jedina neizravna neizravna identifikacija kojoj imamo pristup je spol, dob u trenutku smrti, dijagnoza (ili jesu li počinili samoubojstvo) i činjenica da su umrli u Mađarskoj, a ne možemo identificirati donatore uzoraka. Stoga ne postoji zabrinutost u vezi GDPR-a, a s podacima koji se tiču ovih uzoraka postupat će se u skladu s istim mjerama opreza kao i s drugim podacima (sljedeće pitanje).</p> <p>Sveučilište Semmelweis je dobilo informirani pristanak od najbližih rođaka za prikupljanje uzoraka, pohranu i pohranu podataka. Ovo je odobrio Odbor za etiku znanosti i istraživanja Ministarstva zdravstva Mađarske (br. 6008/8/2002/ETT) i Regionalni odbor za etiku znanosti i istraživanja Sveučilišta Semmelweis (br. 32/1992/TUKEB).</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>Lokalno, podaci će biti pohranjeni na više razina:</p> <ol style="list-style-type: none"> 1. Svi podaci bit će pohranjeni na sveučilišnom računalu istraživača koji ih je generirao. Oni su zaštićeni lozinkom i zahtijevaju specifičnu prijavu korisnika. Antivirusni softver i vatrozid su prisutni i redovito se ažuriraju 2. Svi se podaci također učitavaju u mapu Google Drive koju koristi istraživačka grupa Bradshaw (ukupna veličina: 2 TB, trenutno se koristi oko 500 GB). Oni sadrže podmape za svakog aktivnog istraživača u laboratoriju. Općenito, svi takvi podaci dostupni su svim članovima istraživačke skupine, iako se pristup pojedinačnim podacima može

		<p>ograničiti samo na generatora podataka i dr. Bradshawa ako će biti potrebno (zbog osjetljivosti, no ne očekuje se da će to biti potrebno za ovaj projekt). Podaci na Google disku su šifrirani i dostupni su samo putem Google prijave, koje zahtijevaju periodičnu dvofaktorsku autentifikaciju. Pristup je ograničen na aktivne članove i studente u Bradshaw grupi, pri čemu prethodni članovi i diplomirani studenti gube pristup.</p> <p>3. Svaki mjesec podaci s Google diska sigurnosno se kopiraju na vanjski tvrdi disk koji se čuva u (zaključanom) uredu dr. Bradshawa.</p> <p>Svi novi članovi grupe obučeni su u ovim postupcima najbolje prakse.</p>
	<p>Kako ćete upravljati zaštitom autorskih prava i drugog intelektualnog vlasništva? Tko će biti vlasnik podataka? Koje će se licencije primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka?</p>	<p>Podaci generirani u projektu bit će u zajedničkom vlasništvu istraživača koji su izravno uključeni u eksperiment i voditelja projekta u instituciji u kojoj je rad obavljen (dr. Bradshaw ili Kouter).</p> <p>Za objavljivanje će svi pojedinci koji ispunjavaju uvjete za autorstvo prema COPE smjernicama biti uključeni kao autori.</p> <p>Najvažnije publikacije iz ovog projekta (one navedene u planu rada) bit će objavljene u časopisima s otvorenim pristupom, pri čemu će autorska prava kolektivno pripadati autorima. Koristit će se licenca CC-BY-4.0, osim ako predmetni časopis ne zahtijeva alternativnu licencu. Za sve dodatne publikacije koje mogu proizaći iz projekta (npr. recenzije, vanjske suradnje, dodatna istraživanja koja proizlaze) publikacija može biti u otvorenom pristupu časopisa s modelom pretplate. U potonjem slučaju, autorsko pravo se može prenijeti na izdavača, prema potrebi.</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>Sve izvorne podatkovne datoteke uključivat će datum, navodeći kada su generirane. Kada su zabilješke ili analize potrebne unutar datoteke, tada će se generirati duplikat datoteke, koji će nositi datum uređivanja. Ako budu potrebna dodatna uređivanja kasnijeg dana, napraviti će se još jedna kopija datoteke s novim datumom i tako dalje. Ako se u jednom danu generiraju dvije kopije, tada će se koristiti slovo za označavanje redoslijeda verzija (npr. 20230921a, 20230921b za prve dvije verzije generirane 21. rujna 2023.). Ovi će datumi stoga pokazati napredak datoteke, uz zadržavanje starijih verzija.</p> <p>Za radne verzije publikacije i druge "opisne" datoteke koristit će se sličan sustav datiranja za praćenje verzija, s novom verzijom (datumski duplikat datoteke) koja se generira i pohranjuje kad god se dokument šalje novom autoru ili povremeno kada jedan autor radi na njemu znatnu količinu vremena.</p> <p>Kao što je opisano u pitanju 2.2, kopije podatkovnih datoteka bit će pohranjene i sigurnosno kopirane na više razina: računalo koje koristi pojedinačni istraživač, Google disk istraživačke grupe i sigurnosna kopija vanjskog tvrdog diska. Podaci će se staviti na računalo(a) i Google Drive u roku od 24 sata od prikupljanja, dok će se vanjska pohrana podataka raditi svaki mjesec.</p> <p>Predviđamo generiranje 200 – 600 GB podataka u vezi s ovim projektom, većinom u obliku mikroskopskih slika i/ili video zapisa iz eksperimenta s <i>Drosophilom</i>.</p>
	<p>Kako će se završne verzije podataka dugotrajno pohraniti i čuvati (i nakon završetka projekta)? U kojim će se formatima čuvati podaci? Koja je očekivana količina podataka koja će se trajno pohraniti (izraženo u MB/GB/TB)?</p>	<p>Svi podaci generirani u projektu čuvat će se u laboratoriju Google Drive, te na vanjskom hard disku backup najmanje 3 godine nakon završetka projekta.</p> <p>Neobrađeni podaci bit će pohranjeni u obliku datoteka .tiff, .jpg, .txt, .meg4, .mp4 ili .xlsx. Podatkovne datoteke specifične za instrumente (kao što su .scn ili .vsi) također će biti pohranjene uz, ali nikada umjesto jednog od ovih drugih formata. Opisne datoteke (kao što su radne verzije rukopisa) bit će pohranjene kao .docx i/ili .pdf datoteke.</p> <p>Završni objavljeni radovi dodatno će biti pohranjeni kao .pdf datoteke na DABAR-u.</p> <p>Očekujemo da ćemo pohraniti 200 – 600 GB podataka</p>

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?	Sve publikacije bit će pohranjene na DABAR stranici Odsjeka, uz maksimalno dopušteni otvoreni pristup (potpuni pristup za radove objavljene u časopisima s otvorenim pristupom). Radi dodatne vidljivosti, poveznice na ove članke također će se pojaviti u CROSB-I-u, na Google Scholar-u i na ResearchGate-u. Svi će radovi biti objavljeni u časopisima koje arhiviraju PubMed i Scopus, radi maksimalne fleksibilnosti, većina generiranih podataka je dovoljno mala da se može uključiti u članke u časopisima, ali veći skupovi podataka bit će arhivirani u PUH.
	Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.	Ne postoje ograničenja.
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima FAIR-a.	Potvrđujem ovo.
	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).	Potvrđujem ovo.