

Plan upravljanja istraživačkim podacima za DeShPet

Kalafatović, Daniela

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2024**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:193:133179>

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

Download date / Datum preuzimanja: **2024-09-28**



Repository / Repozitorij:

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



Plan upravljanja istraživačkim podacima

Opće informacije		
	Ime i prezime predlagatelja	Daniela Kalafatović
	Matična organizacija	Sveučilište u Rijeci, Odjel za biotehnologiju
	Naziv projekta	Dizajn katalitički aktivnih peptida i peptidnih nanostrukture (UIP-2019-04-7999)
	Upravitelj podacima	Daniela Kalafatović
1.	Prikupljanje podataka i dokumentacija	
	Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka)	<p>Tijekom projekta generirat će se, prikupljati i obrađivati različite vrste sirovih podataka:</p> <ol style="list-style-type: none"> 1. HPLC – za potvrđivanje mase dobivenih peptida, 2. MALDI-MS za potvrđivanje mase dobivenih peptida, 3. Spektroskopska karakterizacija (FTIR, UV-vis, fluorescencija)– za snimanje samo-sastavljanja peptida, 4. Mikroskopija atomskih sila (AFM) za snimanje morfologije peptidnih nanostrukture, 5. Simulacijama molekularne dinamike peptida i proteina, 6. Primjene algoritama strojnog učenja i generativnih modela. <p>Svi podaci pohranit će se u digitalnom obliku u formatu koji se dobije izravno s instrumenta, primjerice:</p> <ul style="list-style-type: none"> • TIFF za slike s instrumenta AFM; • CSV format za sirove podatke sa HPLCa, FTIRa, UV-VISa i fluorometra; • mzXML i CSV format za masenu spektroskopiju; • PDB, ENT i PDBx/mmCIF datoteke za analizu simulacijom molekularne dinamike; • CSV format za sirove podatke iz baza podataka Uniprot, M-CSA i SabioRK; • arff ili csv podaci za ulazne i izlazne podatke računalnih modela strojnog učenja. <p>Datoteke mzXML biti će obrađene u softveru mMass, a datoteke u formatu CSV biti će grafički obrađene u Excel-u ili Origin-u za daljnju uporabu. PDB datoteke unesene u online alat CHARMM-GUI proizvedu podatke za pripremu simulacije molekularne dinamike u sljedećim formatima: PY, STR, YML, DAT, PSF, CRD, PDB, TOP, OUT, INP, PMR, RTF, MDP, GRO, NDX, i ITP, komprimirani u formatu TGZ. Podaci pripremljeni u CHARMM-GUI se izmjene programima za uređivanje teksta poput Notepad ili WordPad, i zatim se sa GROMACS programom za molekularnu dinamiku iz njih izvedu idući formati: XTC, GRO, TRR, TPR, CPT, EDR, TXT, TOP, OUT, ERR, XPM, XVG i NDX. Procjenjuje se da će simulacije molekularne dinamike proizvesti 1TB podataka u periodu od godine dana.</p>

<p>Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka)</p>	<p>Kvaliteta analitičkih podataka osigurat će se umjeravanjem instrumenata, ponavljanjem laboratorijskih i računalnih eksperimenata, usporedbom s literaturnim i prije dobivenim podacima, recenziranjem.</p> <p>Eksperimentalni podaci pohraniti će se u lokalni repozitorij na prijenosnoj memoriji iz mjernog uređaja. Metode i materijali snimit će se u elektronsku laboratorijsku bilježnicu.</p> <p>Dodatno, opažanja i zapisi eksperimenata prikupljaju se skeniranjem iz ručno pisanih zapisa (laboratorijskog dnevnika) te se digitaliziraju jednom godišnje.</p> <p>Analitički podaci prikupljaju se s instrumenata koji ih generira i obrađuje u matičnom programu:</p> <ol style="list-style-type: none"> 1. FTIR: MicroLab PC (format datoteke: .a2r) i Agilent Resolutions Pro (verzija: 5.2.0, CD 846), format datoteke: .bsp) 2. HPLC: OpenLAB (verzija: A.02.02 (1.3.4)), format datoteke: .D 3. MALDI: FlexControl (verzija: 3.4.135.0) i FlexAnalysis (verzija:3.4.76.0), format datoteke: Analysis directory name (*) 4. UV-Vis spektrometar: Cary WinUV Scan Application (verzija: 5.0.0.999), format datoteke: .BSW 5. Fluorescentni spektrometar: Cary Eclipse Scan Application (verzija:1.2 (147)), format datoteke: .FBSW <p>Podaci dobiveni računalnim programima pohraniti će se na privatnom dijeljenom git repozitoriju, dok će podaci generirani simulacijama molekularne dinamike biti pohranjeni na hard-drive disku za dugoročnu pohranu. Rezultati istraživanja u obliku "progress reporta" pohranjeni su na dijeljenom repozitoriju u oblaku.</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>Dokumenti i mape nazvat će se prema interno dogovorenoj konvenciji koja za svaki skup podataka uključuje identifikaciju istraživača, datum i vrstu podataka.</p> <p>Završni skup podataka pohranit će se u odabranom repozitoriju, popraćen s README dokumentom sa sadržajem svih datoteka kao i koja je konvencija korištena za njihovo imenovanje.</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 pohranjuju</p>	<p>Pri izvedbi ovog projekta neće se kršiti etička načela.</p> <p>Sigurnost na radu: glavni istraživač osigurava odgovarajuće sigurnosne postupke u skladu s relevantnim lokalnim/nacionalnim smjernicama/zakonskim propisima. Sigurnost svih sudionika na projektu (istraživača, osoba uključenih u projekt ili trećih osoba) te pravilno odlaganje otpada prioritet je svih istraživanja.</p>

	informirani o tome i jesu li dali privolu? Kojim ćete se metodama koristiti u svrhu zaštite osjetljivih podataka (GDPR - posebne kategorije osobnih podataka)?	Projekt će biti proveden unutar Zakona o zaštiti na radu (Klasa: 011-01/14-01/88, Urbroj: 71-05-03/1-14-2). Trenutne smjernice osiguravaju sigurnost i zdravlje na radnom mjestu, a dužni su ih se pridržavati svi sudionici na projektu.
	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?	Dio podataka će se obrađivati i njima upravljati u lokalnom nemrežnom okruženju, a za dio podataka koji je dijeljen na mrežnim repozitorijima pristup će biti ograničen korisničkim imenom i lozinkom. Projektom nije planirano generirati osjetljive podatke.
	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 Sveučilišta u Rijeci. Podatke prikupljene tijekom istraživanja u najvećoj mogućoj mjeri objavit će se kroz znanstvene publikacije ili kao otvoreni podaci pod licencijom Creative Commons CC0.
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>)?	Laboratorijski dnevnic i tiskane kopije spektroskopskih analiza karakterizacije peptida čuvaju se u laboratoriju glavnog istraživača. Dodatni elektronički podaci pohranit će se na računalu glavnog istraživača, koji godišnje izrađuje sigurnosne kopije. Osim toga, istraživači će se koristiti dijeljenim mrežnim prostorom na repozitoriju koji je besplatno dostupan djelatnicima institucije (npr. OneDrive, PUH i sl.) za sekundarnu pohranu podataka. Laboratorij glavnog istraživača ima na raspolaganju 1 TB prostora na sustavu OneDrive i 200GB na sustavu PUH po osobi za pohranu podataka koji se može i dodatno proširiti. Korišteni sustav za dijeljeni mrežni prostor u oblaku ima vlastite procedure osiguranja sigurnosnih kopija podataka.
	Koji je vaš plan čuvanja podataka? U kojim će se formatima čuvati?	Podatke ćemo čuvati minimalno 5 godina od završetka projekta na virtualnom poslužitelju kojeg djelatnici s aktivnim AAI@EduHr korisničkim računima imaju na raspolaganju u oblaku (npr. OneDrive, PUH i sl.). Gdje bude moguće, datoteke ćemo pohraniti u formatima otvorenim za uređivanje, primjerice, .docx ili .txt za tekstualne dokumente,

		.xlsx ili .csv za tablične datoteke. Kada je to moguće uključit ćemo i informacije o softveru korištenom za stvaranje predmetnih podataka.
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?	Završni podaci se prvenstveno planiraju dijeliti putem objavljenih znanstvenih članaka u renomiranim časopisima. Objava članka za vrijeme te neposredno nakon završetka projekta ima za cilj smanjiti rizik od gubitka podataka te povećati vidljivost podataka prema široj znanstvenoj zajednici. Dodatno, razmotriti će se objave na Github repozitoriju ili drugim otvorenim repozitorijima (npr. UCI)
	Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja.	Podaci neophodni za bilo koju publikaciju bit će dostupni u trenutku objavljivanja. Svi neobjavljeni podaci pohranit će se u repozitoriju na 5 godina od završetka projekta. Ograničenja po pitanju zakonskih i etičkih regulativa se ne očekuju za podatke prikupljene ovim projektom.
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima <i>FAIR-a</i> .	Potvrđujem.
	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.