**PLAN UPRAVLJANJA ISTRAŽIVAČKIM PODACIMA (PUP)**

|  |  |  |
| --- | --- | --- |
| Opće informacije | | |
|  | Ime i prezime predlagatelja | **Ivana Ivančić Baće** |
|  | Matična organizacija | Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet |
|  | Naziv projekta | Kako je CRISPR adaptacija prilagodljiva na uvjete okoliša? Utjecaj DNA helikaza. |
|  | Upravitelj podacima | Ivana Ivančić Baće |
| 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) | Tijekom trajanja projekta prikupljat ćemo sljedeće sirove podatke:  1. Slike agaroznih gelova nakon elektroforeze generirati će se u jpg i/ili tif formatu (oko 100njak datoteka, svaka veličine 200 -500 KB)  2. Nukleotidne sekvence dobivene nakon Sangerovog sekvenciranja (potvrda kloniranja) od servisa Macrogen dobit će se u FASTAQ, PDF i Text formatu (manje od 50 datoteka, svaka veličine oko 250 KB)  3. Nukleotidne sekvence dobivene Illumina MiSeq sekvenciranjem (trebat će još doradu i analizu) od servisa Macrogen dobit ćemo u Excel formatu (10tak Excel tablica, svaka veličine 300 KB)  4. Slike proteinskih gelova nakon SDS-PAGE elektroforeze generirati će se u jpg ili tif formatu (oko 20 datoteka, veličine 500 KB svaka)  5. Tekstualni podaci generirani izvješćima, člancima, prezentacijama, u docx i pdf formatu, manje od 100 datoteka veličine oko 500 KB svaka  6. Posteri, u ppt i pdf formatu, manje od 50 datoteka oko 5 MB svaki |
|  | 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) | 1. Primarni podaci prikupljat će se s računala koja su povezana s kamerom koja ih snima i analizira programom Kodak 1D Image Analysis Software v.3.6.0. Taj program generira slike u jpg ili tif formatu koje se mogu dodatno analizirati u samom programu ili na osobnom računalu. Oni će biti organizirani u posebnoj podmapi „ slike gelova"  2. Svi podaci o nukleotidnim sekvencama dobit će se putem elektroničke pošte i pohraniti kao datoteke u posebnim podmapama “sekvenciranje” grupirane u podmape prema tipu analize (kloniranje ili potvrda genotipa)  3. Svi podaci od MiSeq sekvenciranja dobit će se putem elektroničke pošte i pohraniti u podmapu „NGS“, a potom nakon analiza posebnim računalnim programima pakta R i statistikom. Analizirani podaci bit će organizirani u posebnim podmapama “analizirani podaci”, “grafovi”.  4 4. Slike proteinskih gelova također se pohranjuju u računalo u programu Image Studio Lite u formatu za slike jpg ili tif. Slike gelova pohranit će se u potom u osobna računala u posebnoj podmapi (SDS-PAGE). Obrađene i analizirane slike bit će u posebnim podmapama “analizirani podaci”.  5. Izvještaji i bilješke o istraživanju bit će generirani u Wordu i organizirani u posebnoj podmapi “izvješća” za svako razdoblje.  6. Posteri konferencije bit će generirani u Power Pointu i organizirani u posebnoj podmapi „konferencije“, a potom u podmape prema nazivu konferencije.  7. Nacrti znanstvenih radova bit će generirani u Wordu i organizirani u posebnoj podmapi “radovi”  8. Svi će podaci biti ispravno imenovani i organizirani u glavnoj mapi projekta s relevantnim podmapama. |
|  | 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, *ReadMe* datoteke i sl.) | Izradit ćemo datoteku ReadMe s dovoljno informacija za razumijevanje projekta i osnovnu navigaciju kroz podmape. Navest će se naziv projekta, datum, opis, podaci o financiranju i kontaktu, kao i opis organizacije datoteke. Svi podaci bit će organizirani u mapu projekta sa sadržajem organiziranim u podmape: red podataka, analizirani podaci, grafikoni, konferencije, izvješća, radovi. |
| 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 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)? | Nismo ograničeni ugovorom o povjerljivosti jer se projekt provodi na laboratorijskim nepatogenim bakterijskim sojevima u kontroliranim uvjetima. |
|  | 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? | Tijekom trajanja projekta pristup podacima bit će dopušten samo ovlaštenim osobama i suradnicima na projektu. Postavit će se lozinke na datoteke i računala. Vatrozid, sigurnosne nadogradnje i zakrpe za operativne sustave, kao i zaštita od virusa bit će redovito ažurirani. Redovito će se stvarati nekoliko sigurnosnih kopija (kompjuteri glavnog istraživača i suradnika, vanjski tvrdi disk i on-line oblak za pohranu). Zbog prirode samog istraživanja, nećemo imati osjetljive podatke. |
|  | 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? | Svi suradnici na projektu i njihove institucije u potpunosti podržavaju politiku projekta o dijeljenju podataka pod open-source licencama. Rezultati javno financiranog istraživanja bit će dostupni besplatno, kako unutar znanstvene zajednice tako i široj javnosti što je prije moguće, nakon što se intelektualno vlasništvo ocijeni i osigura. |
| 3. | Pohrana i čuvanje podataka | |
|  | Kako će radne verzije podataka biti pohranjene tijekom projekta?  Kako će se napraviti sigurnosne kopije tih podataka (*backup*)?  Koja je očekivana količina podataka koja će se prikupiti i čuvati tijekom projekta (izraženo u MB/GB/TB)? | Tijekom projekta različite verzije skupova podataka bit će pohranjene kao datoteke u memoriji analitičkog instrumenta i na osobnom računalu svakog istraživača koji je napravio analize tijekom projekta do objave svih podataka. Podatkovne datoteke će se sigurnosno kopirati na vanjski tvrdi disk. Svi podaci će se periodički organizirati i pohranjivati na računalu glavnog istraživača i svakodnevno sigurnosno kopirati na vanjski tvrdi disk. Sigurnosna kopija također će biti pohranjena u online bazi podataka u oblaku PUH. Tijekom projekta očekujemo prikupljanje i pohranjivanje 3 GB podataka. |
|  | 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)? | Podaci će biti sačuvani povremenom provjerom integriteta podataka i kopiranjem podatkovnih datoteka na novi kompjuter ili vanjski disk svake dvije do pet godina. Podaci će biti arhivirani u formatima s dugoročnom dostupnošću, za tekstualne datoteke i prezentacije to će biti format DOCX (Office Open XML) i PDF (.pdf), za Excel proračunske tablice i statističke podatke to će biti format „vrijednosti odvojene zarezom ili tabulatorom“ (.csv, .tsv, .txt). Očekujemo trajno pohranjivanje 5 GB podataka. |
| 4. | Dijeljenje i ponovna uporaba podataka | |
|  | Kako i gdje će se podaci dijeliti? Koji repozitorij će se koristit za dijeljenje podataka? Kako će potencijalni korisnici doznati za podatke? | Podaci će biti dostupni kao datoteke s dodatnim informacijama o člancima u časopisu ili javnim repozitorijima poput Research gate i bit će popraćeni odgovarajućim citatom.  Institucijski repozitorij u sustavu Dabar osigurava vidljivost podataka putem tražilica Google Scholar te OpenAIRE, a na naslovnoj strani dabar.srce.hr dostupna je tražilica za pretraživanje sadržaja svih repozitorija u sustava Dabar. |
|  | Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja. | Svi podaci će se moći dijeliti. |
|  | Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima *FAIR-a*. | Da, Institucionalni digitalni repozitorij u sustavu Dabar u skladu je s FAIR načelima. |
|  | 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 da ćemo koristiti digitalni repozitorij koji održava neprofitna organizacija (Prirodoslovno-matematički fakultet Sveučilišta u Zagrebu). |