

Plan upravljanja istraživačkim podacima

Opće informacije		
	Ime i prezime predlagatelja	Daniela Marić Pfannkuchen
	Matična organizacija	Institut Ruđer Bošković
	Naziv projekta	Molekularne životne strategije fitoplanktona u Jadranu– Adri LIFE
	Upravitelj podacima	Daniela Maric Pfannkuchen e mail:Daniela.maric.pfannkuchen@irb.hr
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 ćemo četiri različita tipa sirovih podataka:</p> <ol style="list-style-type: none"> <li>1. oceanografska mjerenja</li> <li>2. slike fitoplanktonske zajednice sa svjetlosnog mikroskopa</li> <li>3. fiziološka mjerenja</li> <li>4. sekvenciranja</li> </ol> <p>Svi podaci pohranit će se u digitalnom obliku u formatu koji se dobije izravno s instrumenata (primjerice, LM files za svjetlosni mikroskop u TIFF obliku; Tecan files za fiziološka mjerenja s rezultatima u CSV obliku, CTD files oceanografskih mjerenja u CSV obliku te NGS files sekvenciranja u fasta i fastq obliku). Podaci sekvenciranja zauzet će otprilike 40 GB, a za ostale podatke (mjerenja i slike) očekuje se zauzet će otprilike 2 GB.</p>
	Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka)	<p>Svi uzorci, na kojima će se prikupljati podaci, pripremit će se prema preporučenim protokolima koji su za planirana istraživanja dostupni u literaturi. Datoteke će se nazvati prema prije dogovorenim normama. Skup podataka pratit će i INFO.txt datoteka u kojoj će se opisati korišteni eksperimentalni protokol za sva četiri tipa stvorenih sirovih podataka. Također, zapisat će se i odstupanja od protokola i ostale korisne informacije.</p> <p>Za svaki tip sirovih podataka pohranit će se i set metapodataka (primjerice, za slike svjetlosne mikroskopije: veličina polja, povećanje, faza, uvećanje, snaga, promjer otvora itd.) čime se omogućuje bolje razumijevanje dobivenih podataka unutar radne grupe i povećat će se vrijednost skupa podataka pri ponovnoj uporabi podataka.</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>Dva tipa metapodataka uzet će se u razmatranje unutar okvira projekta AdriLife: oni koji odgovaraju projektnim publikacijama i objavljenim podacima istraživanja.</p> <p>U kontekstu upravljanja podacima, metapodaci tvorit će podskup podataka koji objašnjava svrhu, podrijetlo, opis, vremensku referencu, stvaratelja podataka, uvjete pristupa i uporabe zbirke podataka.</p> <p>Metapodaci koji najbolje opisuju podatke ovise o prirodi podataka. Za istraživačke podatke nastale u projektu AdriLife teško je uspostaviti općeniti kriterij za sve podatke, jer je priroda prvotno razmatranih podataka različita. Zbog toga metapodaci temeljit će se na općenitoj shemi koju koristi Zendo, a koja uključuje sljedeće elemente:</p> <ul style="list-style-type: none"> <li>• naziv: slobodni tekst</li> <li>• stvaratelj: prezime, ime</li> </ul>

		<ul style="list-style-type: none"> <li>• datum</li> <li>• tema: izbor ključnih riječi i klasifikacija</li> <li>• opis: tekst koji opisuje sadržaj podataka i ostale dodatne informacije</li> <li>• opis: što je potrebno za interpretaciju podataka</li> <li>• format: detalji formata</li> <li>• vrsta izvora: skup podataka, slike, audio itd.</li> <li>• identifikator: DOI</li> <li>• pravo pristupa: zatvoreni pristup, pristup pod embargom, ograničen pristup, otvoreni pristup.</li> </ul> <p>Datoteka README.txt može se koristiti kao ugodan način za sve datoteke i mape koje obuhvaćaju projekt objašnjavajući kako su svi skupovi datoteka međusobno povezani, u kojem su formatu, te jesu li određene datoteke namijenjene zamjeni ostalih datoteka itd.</p> <p>Kako organizirati kompletan skup metapodataka trenutno je tema na mnogim europskim projektima. Trenuto se s time bave na europskom projektu JERICO-S3 (H2020, <a href="https://www.jerico-ri.eu">https://www.jerico-ri.eu</a>) s kojim ovaj projekt surađuje i u skladu je s načelima FAIR (<a href="https://www.go-fair.org/fair-principles">https://www.go-fair.org/fair-principles</a>) principa. Strukture metapodataka će biti u skladu sa standardima EMODNET (<a href="https://emodnet.ec.europa.eu/en/emodnet-web-service-documentation#inline-nav-1">https://emodnet.ec.europa.eu/en/emodnet-web-service-documentation#inline-nav-1</a>). Metapodaci o bioraznolikosti pridržavat će se temeljnih Darwinovih standarda (<a href="https://github.com/gbif/ipt/wiki/GMPHowToGuide">https://github.com/gbif/ipt/wiki/GMPHowToGuide</a>). Koraci i postupci kontrole kvalitete za skupove metapodataka trenutno su u razvoju u okviru projekta JERICO-S3.</p>
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)?	Pri izvedbi ovog projekta neće se kršiti etička načela. Ovaj projekt generirat će podatke opisa morske zajednice fitoplanktona i oceanografskih mjerenja te projekt nije ograničen sporazumom o povjerljivosti. Prikupljanja, obrada, čuvanje i dijeljenje podataka provodit će se u skladu sa znanstveno-istraživačkim radnim načelima Instituta Ruđer Bošković. Trenutno ne planiramo sve podatke djeliti javno no u slučaju promjene plana čuvanje i dijeljenje podataka provodit će se u skladu sa znanstveno-istraživačkim radnim načelima Instituta Ruđer Bošković.
	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?	Podaci će se pohraniti u centraliziranom sustavu (CIM baza podataka) za pohranu kojim upravlja Centar za istraživanje mora Instituta Ruđer Bošković. Pristup podacima upravlja se preko identiteta znanstvenika koji je siguran sustav i slijedi najbolje prakse u pogledu upravljanja identitetom. Naš centralni sustav pohranjivanja podataka ima dostatnu zalihost, vrši se zrcaljenje i stalno se nadzire.
	Kako ćete upravljati zaštitom autorskih prava i intelektualnog	Podaci su prikladni za dijeljenje. Podaci su dobiveni promatranjem (dakle, jedinstveni su) i mogli bi se koristiti za druge analize ili za usporedbe fitoplanktonskih zajednica. Prilike za novu

	vlasništva? Tko će biti vlasnik podataka? Koje će se licencije primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka?	uporabu su velike. Naš je omogućiti najširu moguću uporabu podataka njihovim objavljivanjem u znanstvenim časopisima i pohranom u javno dostupnim bazama podataka (National Center for Biotechnology Information (NCBI) i European Nucleotide Archive (ENA)). Proizvođači podataka imaju autorska prava nad podacima i osiguravaju mogućnost ponovnog korištenja tih podataka pod tim autorskim pravama.
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> )?	Podaci će se tijekom istraživanja pohraniti na računala istraživača na projektu a zatim kopirati u centralizirani sustavu (CIM baza podataka) za pohranu kojim upravlja Centar za istraživanje mora Instituta Ruđer Bošković 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. Dodatna automatizirana (RSYNC) sigurnosna kopija datoteka sa servera i s osobnih računala uspostaviti će se s PUH-om.
	Koji je vaš plan čuvanja podataka? U kojim će se formatima čuvati?	Podatke ćemo na neograničen vremenski rok čuvati u centraliziranom sustavu (CIM baza podataka) i također pohraniti u odgovarajući arhiv za podatke na kraju projekta (primjerice, Zenodo) kao i u javno dostupne baze podataka (National Center for Biotechnology Information (NCBI) i European Nucleotide Archive (ENA)) tijekom projekta kroz realizirane znanstvene publikacije. Datoteke ćemo pohraniti u formatima koji se dobiju izravno s instrumenta ako u takvom formatu ne trebaju specijalizirani softver za korištenje podataka. Gdje bude potrebno, datoteke ćemo pretvoriti u jednostavne i dostupne formate (primjerice, u CSV, PDF ili u kodirane jednostavne tekstualne datoteke).
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?	Objava članaka tijekom i nakon završetka projekta ima za cilj smanjiti rizik od gubitka podataka, a omogućit će i dijeljenje podataka pohranom u javno dostupne baze podataka (National Center for Biotechnology Information (NCBI) i European Nucleotide Archive (ENA)) kao i njihovu vidljivost potencijalnim korisnicima. Također, dijeljenje podataka ostvarit će se kroz arhiv za podatke Zenodo koji primjenjuje značajke dugoročne pohrane i podržava načela FAIR-a ( <a href="http://about.zenodo.org/principles/">http://about.zenodo.org/principles/</a> ). Katalog podataka će biti pokazan na web stranicama projekta i IRB-a u sklopu oceanograskim podacima Centra za istraživanje mora i odgovarajućim web sučeljima.
	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 12 mjeseci od završetka projekta.
	Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima FAIR-a.	Potvrdjemo da ćemo se koristiti digitalnim repozitorijem koji je u skladu s načelima FAIR-a. ( <a href="https://www.go-fair.ntraorg/fair-principles">https://www.go-fair.ntraorg/fair-principles</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).	Potvrmo da ćemo se koristiti digitalnom repozitorijuma Centra za istraživanje mora (IRB) koja je neprofitna organizacija
--	--

Ref:

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