

Plan upravljanja istraživačkim podacima projekta RAPTOVAX

Vodopijevec, Alen

Data management plan / Plan upravljanja istraživačkim podacima

Publication year / Godina izdavanja: **2021**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:241:152984>

Rights / Prava: [Attribution-ShareAlike 4.0 International/Imenovanje-Dijeli pod istim uvjetima 4.0 međunarodna](#)

Download date / Datum preuzimanja: **2024-10-02**



Repository / Repozitorij:

[Fulir DATA - Ruđer Bošković Institute Research Data Repository](#)

Plan upravljanja istraživačkim podacima projekta RAPTOVAX

Inicijalna verzija: v1.0

Autor: Alen Vodopijevec <alen@irb.hr>

Uvod

U ovom dokumentu opisan je inicijalni plan upravljanja istraživačkim podacima koji se prikupljaju i generiraju tijekom projekta RAPTOVAX. Osim osnovnih karakteristika podataka, njihove količine i formata, opisani su i procesi i pravila pohrane i dijeljenja podataka za vrijeme trajanja kao i po završetku projekta. Prilikom izrade ovog dokumenta primijenjene su "FAIR Data" smjernice¹ i OAIS referentni model² za arhiviranje digitalne građe. Inicijalni plan će se revidirati po potrebi, a po završetku projekta objavit će se finalna verzija.

O projektu

Osnovni podaci

Naziv projekta: Robusne i adaptabilne biološke platforme za nova cjepiva (RAPTOVAX)

Šifra projekta: KK.01.1.1.04.0099

Trajanje: 20-12-2019 - 20-12-2022

Nositelj: Institut Ruđer Bošković

Voditelj projekta: Saša Kazazić (MBZ: 229585, CROSB: 17020)

¹ Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

² "Reference Model For An Open Archival Information System (OAIS)" (PDF). CCSDS Secretariat. Lipanj 2012. <https://public.ccsds.org/Pubs/650x0m2.pdf> Pristupljeno 29.11.2021.

Ciljevi projekta

Današnja se tehnologija u proizvodnji virusnih cjepiva uglavnom odnosi na atenuirane viruse (s oslabljenim, ali živim patogenima). Iako vrlo efikasna, cjepiva sa živim atenuiranim virusom imaju brojne nedostatke: I) pojava nuspojava izazvanih replikacijom virusa uslijed cijepljenja, što dovodi do nekih simptoma bolesti; II) razvojni, proizvodni i analitički postupci za ova cjepiva su složeni, a proizvodnja relativno nižih prinosa. Konkretno, za rad s visoko virulentnim patogenima potrebno je osigurati kapitalno intenzivne laboratorijske uvjete razina BSL3/4, što predstavlja radikalno visoke barijere uopće za ulazak u IRI; III) velik broj virusnih bolesti ostaje bez mogućnosti prevencije cijepljenjem unatoč višegodišnjim pokušajima razvoja cjepiva na tradicionalan način.

Kako bi se adresirali te efektno riješili navedeni problemi i odgovorilo tržišnim potrebama, projektom se predlaže razvoj novog oblika cjepiva, odnosno kreiranje univerzalnih nanoplatformi temeljenih na rekombinantnom proteinu virusa humane parainfluence tipa 2 (HPIV2) koje će biti prilagodljivi proteinski neinfektivni nosači antigenskih determinanti drugih patogenih virusa. Ova rješenja posjeduju superiorne tehnološke karakteristike spram potencijalnih supstituta i danas konkurentskih proizvoda, prvenstveno s aspekta učinkovitosti (znatno bolje aktiviraju imunološki odgovor) te sigurnosti i neškodljivosti (nisu infektivni jer ne sadrže genetsku informaciju patogena). Očekivani rezultati nakon provedbe predloženih projektnih aktivnosti su serija nanočestica i virusu sličnih čestica (virus-like particles, VLP) kandidata za cjepiva protiv virusa ospica, gripe, RSV-a i HCV-a. Kandidati će biti proizvedeni u laboratorijskim količinama, a njihova imunogeničnost testirana na pokusnim životinjama. Dizajn antigena i postupak ekspresije bit će patentno zaštićen, a relevantni izvorni istraživački podaci prikupljeni tijekom projekta i korišteni za analizu, bit će javno objavljeni.

Karakteristike istraživačkih podataka

Za prikupljanje podataka potrebnih za analizu koristi se spektrometar masa "LCMS-9030 Quadrupole Time-of-Flight Liquid Chromatograph Mass Spectrometer"³ proizvođača Shimadzu. Bit će izvršen niz mjerenja koja su sastavni dio eksperimenata kako bi se došlo do relevantnih

³ [Poveznica na stranicu proizvođača - instrument LCMS-9030](#)

podataka. Svaki eksperiment sastoji se od tri mjerenja, a softver LabSolutions⁴ sprema podatke u sirovom (raw) obliku, lokalno, na računalu koje upravlja instrumentom. Nakon inicijalne analize rezultata mjerenja, podaci onih eksperimenata koji su se pokazali relevantnima za daljnje analize spremaju se na mrežnoj platformi za pohranu podataka. Inicijalna kopija i podaci neuspjelih eksperimenata brišu se s lokalnog računala.

Svaki eksperiment sastoji se od 3 mjerenja, a tijekom cijelog projekta predviđa se oko 200 eksperimenata. Instrument LCMS-9030 generira podatke u vlasničkom formatu koji je čitljiv samo pomoću softvera LabSolutions. Svi prikupljeni podaci, njihove verzije nastale u različitim fazama obrade i analize te optimirani trodimenzionalni strukturni modeli proteina HPIV2, nanočestica NP-a kao i VLP čestica dobiveni homološkim računalnim modeliranjem bit će pohranjeni na lokalnoj platformi za sigurnu pohranu datoteka.

Sustav za pohranu podataka

Sustav za pohranu istraživačkih podataka sastoji se od tri fizička poslužitelja. Dva poslužitelja bit će konfigurirani kao virtualizacijska platforma za smještaj virtualnih poslužitelja aplikacija i baza podataka. Predviđen je njihov smještaj na odvojenim lokacijama unutar Instituta kako bi se povećala dostupnost i sigurnost podataka u slučaju problema s infrastrukturom. Dodatno, treći poslužitelj imat će ulogu diskovnog prostora i prostora za pohranu sigurnosnih kopija podataka. Predviđen ukupan kapacitet diskovnog prostora za pohranu podataka iznosi između 10 i 20TB. Točan kapacitet ovisit će o konfiguraciji RAID polja sustava za pohranu.

Podaci (SIP⁵ i AIP⁶) i pristup

Pristup resursima sustava za pohranu podataka bit će konfiguriran na nekoliko načina kako bi se zadovoljile moguće potrebe krajnjih korisnika i softverskih alata koje koriste:

- Mapiranje mrežnog datotečnog sustava (NFS, SMB, SSHFS i sl.)
- Prijenos putem nekog od standardnih protokola (npr. sFTP)
- Platforma za pohranu “u oblaku” - Nextcloud⁷
 - Web sučelje za upravljanje datotekama
 - Desktop klijent za sinhronizaciju lokalnih i udaljenih direktorija

⁴ [Poveznica na stranicu proizvođača - softver LabSolutions](#)

⁵ SIP - Submission information package (OAIS) - izvorna (raw) verzija podataka

⁶ AIP - Archival information package (OAIS) - podaci pripremljeni za arhiviranje

⁷ <https://nextcloud.com/>

- Mogućnost dijeljenja datoteka s autenticiranim korisnicima

S obzirom na razinu obrade pohranjivat će se nekoliko verzija podataka:

- Sirovi (raw) podaci dobiveni direktno sa LCMS-9030 u vlasničkom (eng. proprietary) formatu
- Set podataka pripremljenih za obradu u otvorenom formatu za što će se koristiti softverska aplikacija HDEaminer⁸
- trajektorije dobivene simulacijama molekulskom dinamikom (tekstualni format)

Pohranjeni podaci prikupljeni i generirani tijekom projekta bit će dostupni autenticiranim korisnicima s dozvolom pristupa putem lokalne mreže Instituta Ruđer Bošković, a isto tako bit će moguće i dijeljenje podskupova podataka s članovima tima i vanjskim suradnicima s vanjske internetske mreže po potrebi. Članovi istraživačkog tima i vanjski suradnici koristit će dodijeljene korisničke podatke za pristup pohranjenim datotekama.

Objava istraživačkih podataka (DIP⁹) u otvorenom pristupu

Svi prikupljeni istraživački podaci pripremljeni za analizu bit će javno objavljeni u Repozitoriju istraživačkih podataka "FULIR Data"¹⁰ i putem javnog dijela gore opisane platforme za pohranu podataka. FULIR Data je institucijski repozitorij Instituta Ruđer Bošković razvijen na nacionalnoj platformi za smještaj institucijskih repozitorija DABAR¹¹. Repozitorij uređuje interdisciplinarni tim Centra za znanstvene informacije (CZI) IRB-a koji se sastoji od knjižničara, informacijskih stručnjaka i informatičara. DABAR se nalazi na hardverskoj infrastrukturi Sveučilišnog računskog centra (SRCE) koji tu platformu razvija i održava u suradnji s više ustanova hrvatske akademske i istraživačke zajednice u što je uključen i CZI tim (predsjednik koordinacijskog odbora, članovi radnih skupina - za istraživačke podatke, za interoperabilnost i za korisničke funkcionalnosti).

FULIR Data je platforma koja omogućava pohranu i objavu podataka prema FAIR¹² principima što podrazumijeva pronalazljivost i dobru vidljivost (promocija rezultata istraživanja), neometanu dostupnost, interoperabilnost te mogućnost ponovnog korištenja i reproducibilnost istraživanja.

⁸ <http://massspec.com/hdexaminer/>

⁹ DIP - Dissemination information package (OAIS) - Podaci pripremljeni za objavu

¹⁰ <https://data.fulir.irb.hr/>

¹¹ <https://dabar.srce.hr>

¹² <https://www.go-fair.org/fair-principles/>

Skupovi podataka objavljeni u FULIR Data repozitoriju dobivaju stalni i jedinstveni identifikator URN:NBN što omogućava i jednoznačno citiranje skupa podataka ukoliko se koristi od strane drugih istraživača u njihovim analizama i radovima.

FULIR Data interoperabilan je s vanjskim globalnim platformama (npr. OpenAIRE Portal¹³, Re3Data¹⁴ katalog) koje preuzimaju metapodatke o pohranjenim skupovima podataka i na taj način povećavaju vidljivost kako objavljenih podataka tako i samog istraživanja i njegovih autora.

Završne napomene

Etičkih problema vezano uz prikupljanje i objavu istraživačkih podataka nema. Postoji mogućnost postavljanja embarga¹⁵ ili nemogućnosti objave podskupova podataka ako se pokaže da bi taj postupak mogao ugroziti proces prijave patenta.

Inicijalna verzija Plana upravljanja istraživačkim podacima prikupljenih i generiranih tijekom RAPTOVAX projekta mijenjat će se s obzirom na razvoj lokalne platforme za pohranu podataka i s time vezanih procesa. Planirane su još najmanje dvije revizije ovog dokumenta do kraja projekta.

¹³ <https://explore.openaire.eu/>

¹⁴ <https://www.re3data.org/repository/r3d100013529>

¹⁵ Ogdoda otvaranja pristupa podacima na unaprijed određeni vremenski period.