Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije

NAČRT ZA RAVNANJE Z RAZISKOVALNIMI PODATKI

OBRAZEC ARIS

Ta obrazec je namenjen pripravi načrta za ravnanje z raziskovalnimi podatki (NRRP) za raziskovalne projekte, ki jih (so)financira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS), kot je določeno v 4. členu Uredbe o izvajanju znanstvenoraziskovalnega dela v skladu z načeli odprte znanosti (Uradni list RS, št. 59/23).

Raziskovalni podatki so opredeljeni kot zapisi o dejstvih (številčni podatki, besedilni, zvočni in slikovni zapisi), ki predstavljajo osnovno podlago za znanstveno raziskovanje in ki v okviru znanstvene skupnosti veljajo kot ustrezno sredstvo za preverjanje veljavnosti raziskovalnih spoznanj.

Prosimo vas, da izpolnite spodnji obrazec NRRP in ga posredujete ARIS **najkasneje v šestih mesecih od začetka izvajanja raziskovalnega projekta**. Priporočljivo je, da NRRP med izvajanjem raziskovalnega projekta po potrebi redno pregledujete in posodabljate. V primeru sprememb posodobljen NRRP priložite vmesnemu in zaključnem poročilu o rezultatih raziskovalnega projekta.

Pregled vsebine NRRP:

0. Splošne informacije
1. Povzetek in opis raziskovalnih podatkov
2. Shranjevanje in varnostno kopiranje podatkov
3. Zagotovitev podatkov na način FAIR
   - 3.1 Zagotavljanje najdljivosti podatkov (F)
   - 3.2 Zagotavljanje dostopnosti podatkov (A)
   - 3.3 Zagotavljanje interoperabilnosti podatkov (I)
   - 3.4 Zagotavljanje ponovne uporabe podatkov (R)
4. Etični in pravni vidiki
5. Drugi raziskovalni rezultati
6. Finančna sredstva

Uporabljene kratice:

- ADP – Arhiv družboslovnih podatkov
- GDPR – Splošna uredba o varstvu podatkov
- IT – informacijska tehnologija
- RO – raziskovalna organizacija
- ZVDAGA – Zakon o varstvu dokumentarnega in arhivskega gradiva ter arhivih
- ZVOP-2 – Zakon o varstvu osebnih podatkov

0	**Splošne informacije**	
0.1	**Šifra projekta**	J3-70136
0.2	**Naziv projekta**	Ustrahovani um: Neželen dolgoročni vpliv medvrstniškega nasilja na razvoj duševnih težav in nevrobioloških sprememb / BULLIEDMIND: Bullying's Unwanted Long-Lasting Impact Encompassing the Development of Mental health Issues and Neurobiological Deviations
0.3	**Šifra vodje projekta**	54046
0.4	**Ime in priimek vodje projekta**	Satja Mulej Bratec
0.5	**Ime in priimek osebe, ki je v RO zadolžena za podporo pri ravnanju z raziskovalnimi podatki**	Na Filozofski fakulteti UM nimamo določene osebe za podporo pri ravnanju z raziskovalnimi podatki.
0.6	**Interna pravila RO za ravnanje z raziskovalnimi podatki**	Na Filozofski fakulteti UM ne obstajajo interna pravila za ravnanje z raziskovalnimi podatki.
0.7	**Verzija NRRP**	1.0
1	**Povzetek in opis raziskovalnih podatkov**	
1.1	**Ali boste pri projektu ponovno uporabili že obstoječe podatke predhodnih raziskav?**	☑ Da ☐ Ne Ponovno bomo uporabili podatke dveh obsežnih prospektivnih študij: Bavarian Longitudinal Study (BLS) in Avon Longitudinal Study of Parents and Children (ALSPAC). To nam bo omogočilo neposredno preučiti dolgoročne nevrobiološke posledice medvrstniškega nasilja v otroštvu / mladostništvu na različne izide v odraslosti, brez potrebe po dolgotrajnem in dragem zbiranju novih prospektivnih podatkov.
1.2	**Katere vrste podatkov boste ustvarili oz. ponovno uporabili in v katerih formatih bodo shranjeni?**	**Vrste podatkov**: Ponovno bomo uporabili številčne in besedilne podatke, natančneje vedenjske in fenotipske podatke (npr. vprašalniki in intervjuji o medvrstniškem nasilju, okoljskih faktorjih, ocene duševnega zdravja, demografske spremenljivke), ter nevroslikovne podatke (strukturni MRI/T1, resting-state fMRI, DWI/DTI). V okviru projekta bomo ustvarili tudi nove harmonizirane sestavljene rezultate (npr. za primerljivo oceno izpostavljenosti medvrstniškemu nasilju med študijama) in gruče (ang. »clusters«), pridobljene z algoritmi strojnega učenja. **Formati**: Fenotipski in vedenjski podatki bodo shranjeni v formatih .csv ali .sav. Nevroslikovni podatki bodo zaradi zahtev izbrane programske opreme (npr. FreeSurfer, SPM, R) shranjeni v standardnih odprtih slikovnih formatih (npr. NIfTI - .nii ali .nii.gz).
1.3	**Kakšen je namen ustvarjanja, zbiranja oz. ponovne uporabe podatkov in njihova povezava s cilji projekta?**	Namen je sistematično raziskati večmodalne nevrobiološke mehanizme, ki povezujejo medvrstniško nasilje z izidi na področju duševnega zdravja v odraslosti. Podatki bodo uporabljeni za preučevanje sprememb v morfologiji možganov ter sprememb v funkcijski in strukturni

		povezljivosti možganov. Te spremembe bomo povezali z izidi na področju duševnega zdravja z uporabo mediacijskih analiz. Kot zadnje, bodo podatki uporabljeni za prepoznavanje individualnih profilov ter zaščitnih mehanizmov z uporabo strojnega učenja in razložljive umetne inteligence (XAI).
1.4	**Kakšna je pričakovana velikost podatkov, ki jih nameravate ustvariti oz. ponovno uporabiti?**	☐ 0–10 GB ☐ 10–100 GB ☐ 100–1000 GB ☑ >1000 GB
2	**Shranjevanje in varnostno kopiranje podatkov**	
2.1	**Kje bodo podatki med izvajanjem projekta shranjeni in varnostno kopirani?**	Podatki bodo med izvajanjem projekta shranjeni in varnostno kopirani na varno zaklenjenih diskih na Filozofski fakulteti UM, v zaklenjenem prostoru, z uporabo enkripcijskega standarda AES-256 (kot zahtevajo lastniki podatkov). Osnovne analize bodo izvedene lokalno. Napredne analize, ki vključujejo strojno učenje, bodo izvedene na superračunalniku HPC RIVR ter v Podatkovnem centru in »AI analysis« Centru na UM FERI.
2.2	**Kako boste izbrali podatke za dolgoročno hrambo?**	Surovi podatki so dolgoročno hranjeni na sedežih primarnih študij (BLS in ALSPAC). Na naši strani dolgoročne hrambe teh podatkov ne bo. V skladu s pogodbami o souporabi podatkov bomo po zaključku naših analiz oziroma ob izteku projekta vse pridobljene kopije podatkov varno in trajno izbrisali ali jih vrnili lastnikom podatkov. Dolgoročno bomo hranili samo neprepoznavne in neobčutljive podatke, povezane z izpeljanimi analizami (npr. beta koeficiente regresijskih analiz) in podatke/rezultate analiz na skupinski ravni.
2.3	**Ali bodo podatki shranjeni v zaupanja vrednem repozitoriju?**	☐ Da ☑ Ne Dostop in prenos podatkov sta strogo omejena zgolj na pooblaščene člane projektne skupine v skladu z določili GDPR in podpisanimi pogodbami. Ker gre za posebne vrste zdravstvenih in osebnih podatkov, nam jih lastniki posredujejo psevdonimizirane. V repozitoriju lahko hranimo samo neprepoznavne in neobčutljive podatke, povezane z izpeljanimi analizami (npr. beta koeficiente regresijskih analiz) in podatke/rezultate analiz na skupinski ravni.
3.	**Zagotovitev podatkov na način FAIR**	
3.1	**Zagotavljanje najdljivosti podatkov (F)**	
3.1.1	**Ali bodo podatki označeni s trajnim identifikatorjem (PID)?**	☑ Da ☐ Ne Surovi in psevdonimizirani podatki zaradi pogodbenih omejitev ne bodo javno objavljeni in ne bodo prejeli javnega PID. Predregistracije analiz, vse objave in vsa programska koda, uporabljena za analize, pa bodo odprto dostopni in označeni s trajnim identifikatorjem (DOI).

3.1.2	**Kateri metapodatki bodo ustvarjeni in kateri metapodatkovni standardi bodo pri tem upoštevani?**	Za nevroslikovne podatke (sMRI, fMRI, DTI) bomo ustvarili metapodatke v skladu s standardom BIDS (Brain Imaging Data Structure). Za fenotipske in vedenjske podatke se bomo opirali na smernice že uveljavljenih mednarodnih konzorcijev.
3.1.3	**Ali bodo metapodatki vsebovali ključne besede za izboljšanje najdljivosti in možnosti ponovne uporabe?**	☑ Da ☐ Ne Javno dostopna koda in predregistracije bodo opremljeni s ključnimi besedami za izboljšanje najdljivosti. Uporabljene bodo ključne besede, kot so: peer victimization, bullying, resilience, development, machine learning, ALSPAC, BLS, MRI.
3.2	**Zagotavljanje dostopnosti podatkov (A)**	
3.2.1	**Ali bodo vsi podatki odprto dostopni?**	☐ Da ☑ Ne Čeprav podatki ne morejo biti odprto dostopni zaradi pravnih omejitev in sporazumov o souporabi podatkov, so ti podatki skladno s pravili ALSPAC in BLS na voljo vsakemu raziskovalcu, ki pri lastnikih vloži uradno prošnjo za dostop in podpiše ustrezen sporazum o souporabi podatkov.
3.2.2	**Kdaj bodo podatki odprto dostopni in za koliko časa?**	Čeprav podatki ne morejo biti odprto dostopni, bodo trajno odprto dostopni predregistracije, analitična in programska koda, rezultati in znanstvene publikacije, v revijah z odprtim dostopom ter relevantnih repozitorijih (GitHub, protocols.io, OSF ipd.).
3.2.3	**Na kakšen način bo v primeru omejitev pri uporabi omogočen dostop do podatkov med izvajanjem projekta in po njegovem zaključku?**	Primarni podatki bodo ves čas dostopni izključno pooblaščenim članom projektne skupine v nadzorovanem in varnem lokalnem okolju. Tretjim osebam so podatki dostopni le prek uradne prošnje izvornim institucijam, medtem ko bodo vsi rezultati, programska koda in publikacije odprto dostopni vsem preko repozitorijev (glej točko 3.2.2).
3.2.4	**Ali bo za dostop do podatkov oz. njihovo branje potrebna dodatna dokumentacija oz. informacija o ustrezni programski opremi?**	☑ Da ☐ Ne Za branje in obdelavo slikovnih podatkov so potrebna ustrezna programska okolja (npr. FreeSurfer, SPM, fMRIPrep). Celotna programska koda, ustvarjena tekom projekta, bo objavljena v obliki odprte izvorne kode in bo opremljena z ustrezno dokumentacijo, ki bo omogočala ponovljivost analiz in uporabo podatkov.
3.3	**Zagotavljanje interoperabilnosti podatkov (I)**	
3.3.1	**Katere geslovnike oz. šifrante boste uporabili pri pripravi podatkov in metapodatkov?**	Pri harmonizaciji fenotipskih podatkov obeh študij se bomo opirali na smernice že uveljavljenih mednarodnih konzorcijev. Uporabljali bomo standardizirana imena spremenljivk, ki bodo zbrana v podatkovnih slovarjih. Kjer so bili uporabljeni različni instrumenti, bo strokovna ekipa določila optimalne strategije za harmonizacijo. Za nevroslikovne podatke bomo dosledno uporabljali standarde za strukturiranje podatkov (kot je BIDS, opredeljen v točki 3.1.2).

3.3.2	**Ali boste primorani uporabiti manj poznane ali lastne geslovnike oz. šifrante?**	☑ Da ☐ Ne Zaradi združevanja in harmonizacije dveh specifičnih vzdolžnih kohort bo v določenih primerih neizogibna uporaba projektno specifičnih spremenljivk, sestavljenih rezultatov in šifrantov. V teh primerih bomo transparentno zagotovili dokumentacijo za boljše razumevanje in kasnejšo interoperabilnost.
3.4	**Zagotavljanje ponovne uporabe podatkov (R)**	
3.4.1	**Na kakšen način boste zagotovili dokumentacijo, potrebno za ponovno uporabo podatkov?**	Vsa ustrezna dokumentacija bo zagotovljena v publikacijah in repozitorijih navedenih v točki 3.2.2. Zagotavljanje dokumentacije bo izvedeno v strogem skladu s pogodbami o podatkih in omejitvami s strani lastnikov primarnih kohort (ALSPAC in BLS).
3.4.2	**Ali bodo vaši podatki javno dostopni in licencirani v skladu z odprto licenco CC0, da bo s tem omogočena čim širša ponovna uporaba?**	☐ Da ☑ Ne Primarni podatki niso in ne bodo v naši lasti ter zaradi pogodbenih omejitev ne morejo biti javno dostopni pod licenco CC0. Vendar pa bodo vse znanstvene publikacije in ustvarjena koda v največji možni meri objavljeni pod najmanj restriktivno odprto licenco CC-BY.
3.4.3	**Kakšne postopke zagotavljanja kakovosti podatkov boste uporabili?**	V projektu bomo uporabljali izključno že pridobljene in preverjene podatke, katerih kakovost zagotavljajo institucije, ki so jih pridobile (Univerza Bristol, Univerza Warwick, Tehniška Univerza v Münchnu). Predhodno obdelavo in kontrolo kakovosti nevroslikovnih podatkov bomo izvajali z uveljavljenimi programskimi paketi (FreeSurfer, SPM, fMRIPrep, DIPY) in po enotnih protokolih za harmonizacijo med obema študijama (BLS in ALSPAC). Fenotipski podatki bodo prav tako natančno harmonizirani in preverjeni glede na mere nasilja, izide duševnega zdravja in demografske kovariate.
4.	**Etični in pravni vidiki**	
4.1	**Ali obstajajo etična ali pravna vprašanja, ki bi lahko vplivala na deljenje podatkov?**	☑ Da ☐ Ne Deljenje podatkov je strogo omejeno s podpisanimi sporazumi o prenosu in souporabi podatkov (Data Transfer/Sharing Agreements). Kar zadeva pridobivanje izvornih podatkov, so bila predhodno pridobljena vsa ustrezna etična dovoljenja primarnih študij. Etična odobritev za BLS je bila izdana s strani etičnih komisij Univerzitetne otroške bolnišnice v Münchnu, Bavarskega zdravstvenega sveta (Landesärztekammer Bayern) in Univerzitetne bolnišnice v Bonnu (za podatke iz odraslosti). Etično odobritev za ALSPAC sta izdala »ALSPAC Ethics and Law Committee« in lokalne raziskovalne etične komisije.
4.2	**Ali boste med izvajanjem projekta obdelovali oz. hranili osebne podatke?**	☑ Da ☐ Ne

		Projekt obdeluje izključno psevdonimizirane podatke, ki nam jih posredujejo lastniki podatkov študij BLS in ALSPAC. Raziskovalna ekipa ne bo imela dostopa do imen ali neposrednih identifikatorjev posameznikov. Obdelava bo izvedena v celoti skladno s Splošno uredbo o varstvu podatkov (GDPR) in ZVOP-2 ter pod pogoji zapisanimi v obstoječih sporazumih o souporabi podatkov.
4.3	**Ali bodo med projektom ustvarjene oz. ponovno uporabljene posebne vrste osebnih podatkov?**	☑ Da ☐ Ne Obdelovali bomo zdravstvene podatke (mere duševnega zdravja), demografske podatke, in nevroslikovne podatke, ki spadajo med posebne vrste osebnih podatkov skladno z 9. členom GDPR. Ker raziskovalci nimajo dostopa do identifikatorjev, so ti podatki dosledno psevdonimizirani in se bodo uporabljali strogo kot spremenljivke v naših modelih ob upoštevanju določil GDPR.
4.4	**Kako boste uredili lastništvo avtorskih pravic in pravic intelektualne lastnine podatkov, ki jih boste ustvarili ali ponovno uporabili?**	Lastništvo nad primarnimi raziskovalnimi podatki in s tem povezano intelektualno lastnino v celoti ostane prvotnima izvornima institucijama. Pravice intelektualne lastnine, ki bodo nastale na novo iz naslova naših obdelav in analiz v okviru projekta (npr. ustvarjeni modeli in izpeljana koda), pa bodo urejene skladno z zakonodajo, pravili financerja, pravili Univerze v Mariboru oziroma Filozofske fakultete UM ter morebitnimi dogovori med projektnimi partnerji.
5.	**Drugi raziskovalni rezultati**	
5.1	**Ali boste poleg podatkov ustvarili ali ponovno uporabili tudi druge raziskovalne rezultate?**	☐ Da ☑ Ne
6.	**Finančna sredstva**	
6.1	**Kakšni bodo stroški ravnanja s podatki in drugimi rezultati projekta po načelih FAIR in kako bodo kriti?**	Glavni predvideni stroški pri ravnanju s podatki so povezani s plačilom dostopa do primarnih raziskovalnih podatkov (administrativni stroški in stroški priprave podatkov, ki jih zaračunajo izvorni lastniki podatkov). Ti stroški bodo v celoti kriti iz dodeljenih finančnih sredstev raziskovalnega projekta (ARIS projekt J3-70136). Ostali stroški, povezani z varno hrambo in obdelavo (npr. uporaba superračunalnika HPC RIVR in lokalnih strežnikov), so kriti v okviru obstoječe infrastrukturne in institucionalne podpore raziskovalne organizacije (Univerze v Mariboru).
6.2	**Kdo bo odgovorna oseba za ravnanje z raziskovalnimi podatki pri projektu?**	Odgovorna oseba za ravnanje z raziskovalnimi podatki pri projektu bo vodja projekta, doc. dr. Satja Mulej Bratec.

Uporabljeni viri:

- *Anotirana predloga načrta za ravnanje s raziskovalnimi podatki za projekte Obzorja Evropa.* CTK UL. Dostopno na: https://dirrosdata.ctk.uni-lj.si/raziskovalni-podatki/nacrt-ravnanja-z-raziskovalnimi-podatki/.
- Bezjak, Sonja (ur.) (2024). *Spoznaj FAIR: Priročnik o odprti znanosti v Sloveniji*. Univerza na Primorskem. Dostopno na: https://www.hippocampus.si/ISBN/978-961-293-328-9.pdf.
- *Horizon Europe Data management plan template*. Dostopno na: https://www.openaire.eu/images/Guides/HORIZON_EUROPE_Data-Management-Plan-Template.pdf.
- *NWO Template Data management plan*. Dostopno na: https://www.nwo.nl/en/research-data-management.

Verzija dokumenta: 1.0
Številka: 6311-94/2024-1
Datum: 15. 10. 2024