

# 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

### Navodilo za izpolnjevanje

Obrazec izpolnite tako, da vsebino vnašate v celice v skrajnem desnem stolpcu oz. tam označite eno od ponujenih možnosti. V teh celicah so sedaj v sivi barvi pisave navedeni razlage oz. navodila za vnos ustreznih podatkov in opisov. To pomožno besedilo lahko po vnosu vsebine izbrišete.

0		Splošne informacije
0.1	Šifra projekta	V1-2566
0.2	Naziv projekta	Analiza obremenitve zunanjega zraka z alergeni cvetnim prahom
0.3	Šifra vodje projekta	34480
0.4	Ime in priimek vodje projekta	doc. dr. Rene Markovič
0.5	Ime in priimek osebe, ki je v RO zadolžena za podporo pri ravnanju z raziskovalnimi podatki	Če je v RO oseba, zadolžena za podporo pri ravnanju z raziskovalnimi podatki (t. i. podatkovni svetovalec oz. angl. <i>data steward</i> ), se pričakuje, da se načrt pripravi v sodelovanju z njo. Zaželeno je, da poleg imena in priimka te osebe navedete tudi datum sestanka pripravljavca načrta s to osebo.
0.6	Interna pravila RO za ravnanje z raziskovalnimi podatki	Če ta obstajajo, navedite sklic na interna pravila o ravnanju z raziskovalnimi podatki, npr. povezavo na spletno stran, kjer so ta pravila objavljena.
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?	<input type="checkbox"/> Da Ponovno bomo uporabili podatke o dnevni koncentraciji cvetnega prahu za obdobje 2002–2023 (vir: NLZOH) in meteorološke podatke (vir: ARSO/NLZOH). Namen je analiza dolgoročnih trendov, prostorska primerjava in modeliranje vplivov podnebnih sprememb.
1.2	Katere vrste podatkov boste ustvarili oz. ponovno uporabili in v katerih formatih bodo shranjeni?	V sklopu projekta niso bili zajeti novi surovi podatki. Projekt je temeljil na razvoju analitičnih metod in algoritmov za obdelavo obstoječih arhivskih podatkov. Ustvarjena sta bila dva tipa rezultatov: (1) programska koda (Python skripte) in (2) izsledki analiz v obliki agregiranih podatkovnih nizov.
1.3	Kakšen je namen ustvarjanja, zbiranja oz. ponovne uporabe podatkov in njihova povezava s cilji projekta?	Podatke zbiramo in urejamo z namenom izvedbe statistične analize in modeliranja dolgoročnih trendov obremenitve zraka. Ti podatki so ključni za uresničevanje ciljev projekta: identifikacijo kritičnih obdobjev cvetenja, povezovanje s podnebnimi dejavniki in pripravo na dokazih temelječih priporočil za javno zdravje ter urejanje zelenih površin.
1.4	Kakšna je pričakovana velikost podatkov, ki jih nameravate ustvariti oz. ponovno uporabiti?	<input type="checkbox"/> 0–10 GB
2		Shranjevanje in varnostno kopiranje podatkov
2.1	Kje bodo podatki med izvajanjem projekta shranjeni in varnostno kopirani?	Med izvajanjem projekta bodo podatki in koda shranjeni v repozitoriju na platformi GitHub, kar omogoča verzioniranje (nadzor sprememb) in sodelovanje med partnerji. Občutljivi ali surovi podatki bodo hranjeni na varnem strežniku z omejenim dostopom. Za varnostne kopije bo skrbel vodja projekta z rednim lokalnim kopiranjem in sinhronizacijo repozitorija na institucionalni strežnik.
2.2	Kako boste izbrali podatke za dolgoročno hrambo?	Dolgoročno bomo hranili končne prečiščene nabore podatkov (final datasets), ki so podlaga za objavljene rezultate v znanstvenih člankih, ter pripadajočo programsko kodo, nujno za replikacijo analiz. Surovi vhodni podatki ostajajo v arhivu NLZOH.
2.3	Ali bodo podatki shranjeni v zaupanja vrednem repozitoriju?	<input type="checkbox"/> Da

		Končni podatki bodo trajno shranjeni v repozitoriju Zenodo ( <a href="https://zenodo.org">https://zenodo.org</a> ). Zenodo je zaupanja vreden repozitorij, skladen z zahtevami OpenAIRE, ki zagotavlja trajno hrambo, dodelitev digitalnega identifikatorja objekta (DOI) in dolgoročno dostopnost.
<b>3.</b>	<b>Zagotovitev podatkov na način FAIR</b>	
<b>3.1</b>	<b>Zagotavljanje najdljivosti podatkov (F)</b>	
<b>3.1.1</b>	<b>Ali bodo podatki označeni s trajnim identifikatorjem (PID)?</b>	<input type="checkbox"/> Da  Podatkovni nabori in koda bodo označeni z DOI (Digital Object Identifier), ki ga avtomatsko dodeli repozitorij Zenodo ob objavi.
<b>3.1.2</b>	<b>Kateri metapodatki bodo ustvarjeni in kateri metapodatkovni standardi bodo pri tem upoštevani?</b>	Ustvarili bomo metapodatke, ki bodo vključevali: ime kraja, kjer se merilno mesto nahaja, datum meritve, taksonomsko klasifikacijo (vrsta/rod cvetnega prahu), metodo vzorčenja (Hirstova metoda) in merske enote (število zrn/m <sup>3</sup> ). Sledili bomo standardom, ki se uporabljajo v Evropski aerobiološki mreži (EAN).
<b>3.1.3</b>	<b>Ali bodo metapodatki vsebovali ključne besede za izboljšanje najdljivosti in možnosti ponovne uporabe?</b>	<input type="checkbox"/> Da  Metapodatki bodo vključevali ključne besede: alergeni cvetni prah, kakovost zraka, aerobiologija, Slovenija, časovne vrste, Hirstova metoda. To bo izboljšalo najdljivost v iskalnikih.
<b>3.2</b>	<b>Zagotavljanje dostopnosti podatkov (A)</b>	
<b>3.2.1</b>	<b>Ali bodo vsi podatki odprto dostopni?</b>	Surovi podatki o koncentracijah peloda v zraku niso bili ustvarjeni v okviru tega projekta – ampak so bili pridobljeni iz arhiva meritev, zbranih v okviru nacionalnega monitoringa. Osrednji prispevek projekta so metode in algoritmi za obdelavo aerobioloških podatkov, katerih rezultati (izračunani trendi, sezonske analize, modelirani vplivi podnebnih sprememb) bodo javno dostopni po objavi znanstvenih publikacij.
<b>3.2.2</b>	<b>Kdaj bodo podatki odprto dostopni in za koliko časa?</b>	Podatki bodo postali odprto dostopni po objavi znanstvenih publikacij, ki temeljijo na teh podatkih, oziroma najkasneje ob zaključku projekta. Dostopni bodo trajno (minimalno 20 let skladno s politiko Zenodo).
<b>3.2.3</b>	<b>Na kakšen način bo v primeru omejitev pri uporabi omogočen dostop do podatkov med izvajanjem projekta in po njegovem zaključku?</b>	Med izvajanjem projekta bo dostop do delovnih verzij podatkov in kode na GitHubu omejen na člane projektne skupine (dostop z avtentikacijo). Po zaključku bo repozitorij (ali njegova arhivska kopija na Zenodo) postal javen.
<b>3.2.4</b>	<b>Ali bo za dostop do podatkov oz. njihovo branje potrebna dodatna dokumentacija oz. informacija o ustrezni programski opremi?</b>	<input type="checkbox"/> Da Podatki bodo v formatu .csv, ki ne zahteva posebne opreme. Za ponovno uporabo in zagon priložene analitske kode pa bo potreben interpreter za programski jezik Python (ki je odprtokoden in prosto dostopen) ter standardne knjižnice (pandas, numpy), ki bodo navedene v dokumentaciji.
<b>3.3</b>	<b>Zagotavljanje interoperabilnosti podatkov (I)</b>	
<b>3.3.1</b>	<b>Katere geslovnike oz. šifrate boste uporabili pri pripravi podatkov in metapodatkov?</b>	Uporabili bomo standardno biološko taksonomijo za poimenovanje rastlinskih rodov in vrst (npr. Ambrosia artemisiifolia, Betula, Poaceae), kar zagotavlja interoperabilnost z mednarodnimi aerobiološkimi bazami.
<b>3.3.2</b>	<b>Ali boste primorani uporabiti manj poznane ali</b>	<input type="checkbox"/> Da <input type="checkbox"/> Ne

	lastne geslovnike oz. šifrante?	
<b>3.4</b>	<b>Zagotavljanje ponovne uporabe podatkov (R)</b>	
<b>3.4.1</b>	<b>Na kakšen način boste zagotovili dokumentacijo, potrebno za ponovno uporabo podatkov?</b>	V repozitoriju bo priložena datoteka README (v formatu .md ali .txt). Vsebovala bo opis strukture podatkov, definicije spremenljivk (podatkovni slovar), opis metodologije čiščenja podatkov in navodila za zagon Python skript.
<b>3.4.2</b>	<b>Ali bodo vaši podatki javno dostopni in licencirani v skladu z odprto licenco CC0, da bo s tem omogočena čim širša ponovna uporaba?</b>	<input type="checkbox"/> Ne  Podatkov ne bomo opremili z licenco CC0, temveč z licenco CC BY 4.0 (Priznanje avtorstva). Ta licenca omogoča prosto deljenje in predelavo, vendar zahteva, da uporabnik ustrezno navede vir (avtorje raziskave in NLZOH kot vir meritev). Programska koda bo objavljena pod licenco MIT.
<b>3.4.3</b>	<b>Kakšne postopke zagotavljanja kakovosti podatkov boste uporabili?</b>	Kakovost izvornih podatkov je zagotovljena s standardizirano Hirstovo metodo vzorčenja. V okviru projekta bomo uporabili avtomatizirane Python skripte za detekcijo napak (outlierjev) in konsistentnost časovnih vrst pred končno analizo.
<b>4.</b>	<b>Etični in pravni vidiki</b>	
<b>4.1</b>	<b>Ali obstajajo etična ali pravna vprašanja, ki bi lahko vplivala na deljenje podatkov?</b>	<input type="checkbox"/> Ne Raziskava temelji izključno na okoljskih meritvah (cvetni prah v zraku) in ne vključuje dela z ljudmi ali osebnimi podatki.
<b>4.2</b>	<b>Ali boste med izvajanjem projekta obdelovali oz. hranili osebne podatke?</b>	<input type="checkbox"/> Ne
<b>4.3</b>	<b>Ali bodo med projektom ustvarjene oz. ponovno uporabljene posebne vrste osebnih podatkov?</b>	<input type="checkbox"/> Ne
<b>4.4</b>	<b>Kako boste uredili lastništvo avtorskih pravic in pravic intelektualne lastnine podatkov, ki jih boste ustvarili ali ponovno uporabili?</b>	Lastnik surovih merskih podatkov je NLZOH (in deloma ARSO za vremenske podatke). Rezultati analiz, izvedeni trendi in razvita programska koda bodo v solastništvu konzorcijskih partnerjev (UM FNM in NLZOH). Skladno s pogodbo se rezultati prenesejo v javno domeno oziroma odprti dostop.
<b>5.</b>	<b>Drugi raziskovalni rezultati</b>	
<b>5.1</b>	<b>Ali boste poleg podatkov ustvarili ali ponovno uporabili tudi druge raziskovalne rezultate?</b>	<input type="checkbox"/> Da Poleg podatkov bomo razvili in objavili programske kodo (Python skripte) za avtomatizirano obdelavo aerobioloških podatkov. Koda bo prosto dostopna na GitHubu in arhivirana na Zenodo, kar bo omogočilo drugim raziskovalcem uporabo naših metodologij na lastnih podatkih.
<b>6.</b>	<b>Finančna sredstva</b>	
<b>6.1</b>	<b>Kakšni bodo stroški ravnanja s podatki in drugimi rezultati projekta</b>	Stroški ravnanja s podatki (urejanje, čiščenje, priprava metapodatkov) so vključeni v redne stroške dela raziskovalcev (cenovna kategorija A).

	<b>po načelih FAIR in kako bodo kriti?</b>	Uporaba repozitorijev GitHub in Zenodo je brezplačna, zato dodatni materialni stroški niso predvideni.
<b>6.2</b>	<b>Kdo bo odgovorna oseba za ravnanje z raziskovalnimi podatki pri projektu?</b>	Odgovorna oseba za izvajanje NRRP je vodja projekta, doc. dr. Rene Markovič.

#### 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