

Načrt ravnanja z raziskovalnimi podatki (NRRP)

Izdelava informacijskega sistema evidence skladiščenja vzorcev (VzorSIS)

Datum zadnjega urejanja: 24.03.2025

Verzija: NRRP_VzorSIS_v01

0. Splošne informacije

Naziv projekta: Izdelava informacijskega sistema evidence skladiščenja vzorcev (VzorSIS)

Vodja projekta: Petra Gostinčar

Institucija: Geološki zavod Slovenije

Program financiranja: ARIS – RSF

Številka projekta: /

Začetek projekta: 01.05.2024

Konec projekta: 30.04.2026

Podatkovni skrbnik: Petra Gostinčar

Povzetek projekta:

Pri izvajanju temeljnih in aplikativnih raziskav s področja geoznanosti se zbirajo vzorci kamnin in tal, ki predstavljajo fizične dokaze za geološke raziskave. Ti vzorci omogočajo ponovno preverjanje rezultatov in nadaljnje študije, zato je njihovo ustrezno shranjevanje ključnega pomena. Trenutno so vzorci na GeoZS shranjeni v neorganiziranem skladišču, kar otežuje njihovo identifikacijo in uporabo. Poleg tega manjkajo drugi ključni deli sistema, kot so ustrezna organizacija, evidenca skladiščenja in digitalna infrastruktura za vnos, upravljanje in iskanje vzorcev.

V okviru projekta bomo vzpostavili delovno skupino, ki bo na podlagi dobrih praks razvila konceptualni model za celostno obravnavo vzorcev na GeoZS. Pripravili bomo načrt upravljanja zbirke, ki bo vključeval vse potrebne korake za učinkovito organizacijo in shranjevanje vzorcev. Poleg tega bomo vzpostavili informacijski sistem za evidenco skladiščenja vzorcev, ki bo omogočal enostaven vnos, upravljanje in iskanje podatkov. S temi ukrepi bomo postavili temelje za napreden geološki repozitorij, ki bo omogočal učinkovito in trajno shranjevanje ter uporabo vzorcev za prihodnje raziskave.

Izdelava informacijskega sistema evidence skladiščenja vzorcev (VzorSIS)

1. POVZETEK IN OPIS RAZISKOVALNIH PODATKOV

1.1. Ali boste pri projektu ponovno uporabili že obstoječe podatke predhodnih raziskav?

- ☒ Da
☐ Ne

Pri projektu bomo uporabili naslednje že obstoječe podatke predhodnih raziskav:

- Zbirka podatkov o vrtinah
- Evidenca izpirkov in vrtin
- Osnovne geološke karte in gradivo, ki je bilo producirano v postopku priprave osnovnih geoloških kart

1.2. Katere vrste podatkov boste ustvarili ali ponovno uporabili in v katerih formatih bodo shranjeni?

Ustvarili bomo naslednje podatke: besedilni dokumenti (format .docx), podatkovna baza (predvidoma PostgreSQL), preglednice (format .xlsx), shematski prikazi (format .drawio) in vektorski prikazi v .pdf formatu).

1.3. Kakšen je namen ustvarjanja, zbiranja oz. ponovne uporabe podatkov in njihova povezava s cilji projekta?

Namen ustvarjanja, zbiranja in ponovne uporabe podatkov je vzpostavitev informacijskega sistema evidence skladiščenja vzorcev in podatkovne baze, v kateri bodo shranjeni podatki o vzorcih. S tem bomo postavili temelje za vzpostavitev celostnega geološkega repozitorija na Geološkem zavodu Slovenije.

1.4. Na kakšen način boste zbrali ali ustvarili nove podatke in/ali na kakšen način boste uporabili že obstoječe podatke (nanaša se na ponovno rabo)?

Obstoječe podatke iz podatkovnih baz kot sta na primer zbirka podatkov o vrtinah in evidenca izpirkov vrtin bomo povezali s podatki, ki se nahajajo v podatkovni bazi geoloških vzorcev.

Nove podatke bomo ustvarili s popisom vzorcev, ki se že nahajajo v centralnem skladišču GeoZS na Letališki cesti v Ljubljani in vzorcev, ki se nahajajo v različnih prostorih na GeoZS-ju. Podatke bomo zbirali tudi v okviru obiskov primere dobrih praks v tujini, in sicer z neposrednim komuniciranjem s sorodnimi institucijami preko spletne pošte in drugih komunikacijskih kanalov, kot tudi s pridobivanjem podatkov o že vzpostavljeni infrastrukturi in smernicah preko spleta.

1.5. Kakšna je pričakovana velikost podatkov, ki jih nameravate ustvariti ali ponovno uporabiti?

- ☒ 0–10 GB
- ☐ 10–100 GB
- ☐ 100–1000 GB
- ☐ >1000 GB

2. SHRANJEVANJE IN VARNOSTNO KOPIRANJE PODATKOV

2.1. Kje boste hranili podatke, metapodatke in varnostne kopije tekom raziskovanja?

Podatkovna baza geoloških vzorcev bo shranjena na strežniški infrastrukturi GeoZS-ja, prav tako bodo tam hranjene varnostne kopije tekom izvajanja projekta. Za podatke, ki se nahajajo na strežniški infrastrukturi GeoZS, je zagotovljeno redno sistemsko varnostno kopiranje.

Za hrambo in izmenjavo ostalih podatkov in informacij v okviru projekta uporabljamo MS OneDrive, ki omogoča učinkovito delitev informacij v realnem času. Varnostna kopija podatkov se nahaja na GeoZS disku N: Varnostno kopiranje izvajamo na tedenski ravni.

2.2. Kako boste skrbeli za varovanje in zaščito občutljivih podatkov med raziskovanjem?

Dostop do podatkov je omogočen z dodeljevanjem dostopa v MS OneDrive v skladu z GeoZS smernicami. Podatki so dostopni vsem članom projektne ekipe; vsi člani projektne ekipe so iz GeoZS.

2.3. Ali boste izbirali podatke za dolgoročno hrambo?

Ne, za dolgoročno hrambo so predvideni vsi podatki, pridobljeni tekom projekta.

2.4. Kje bodo podatki hranjeni na dolgi rok (npr. zaupanja vreden repozitorij ali arhiv)?

Na dolgi rok bodo podatki in morebitna razvita programska oprema hranjeni in vzdževani v okviru GeoZS strežniške infrastrukture. Po koncu projekta ni predvidena objava podatkov podatkovne baze, saj se bo le-ta po koncu projekta aktivno dopolnjevala. Na dolgi rok je predvidena tudi hramba in objava podatkov v institucionalnem repozitoriju eGeologija.

Elaborat konceptualnega modela bo hranjen v skladu s smernicami oddaje elaboratov, ki so vzpostavljene na GeoZS. Elaborat se bo hranil v okviru GeoLib-a.

3. ZAGOTAVLJANJE NAJDLJIVOSTI PODATKOV (F)

3.1. Ali bodo podatki označeni s trajnim identifikatorjem (npr. DOI)?

☒ Da

☐ Ne

Vzorci, ki se bodo nahajali v bazi vzorcev, bodo tekom projekta pridobili enolične identifikatorje. Vzorci, ki so bili uporabljeni pri izvajanju raziskav, katerih izsledki so objavljeni v znanstvenih, strokovnih in drugih publikacijah, bodo pridobili trajni identifikator IGSN (*International Generic Sample Number* oz. Mednarodna generična številka vzorca).

3.2. Kateri metapodatki bodo ustvarjeni in kateri metapodatkovni standardi bodo pri tem upoštevani?

Podatkovni standard, ki ga uporablja IGSN, se imenuje Mednarodna generična številka vzorca (International Generic Sample Number - IGSN). Ta standard zagotavlja globalno edinstvene, trajne identifikatorje za fizične vzorce, kar omogoča natančno identifikacijo, citiranje in lociranje vzorcev v različnih raziskovalnih publikacijah in bazah podatkov.

Pri zbiranju podatkov bodo ustvarjeni naslednji metapodatki: ID vzorca, ime vzorca, tip vzorca, število vzorcev, ime in priimek skrbnika, ime in priimek namestnika, tip zbirke, način odvzema vzorcev, ID vrtine, ime popisanega zaporedja oz. profila, lokacija vzorčenja (država, občina, kraj, koordinate), način zajema lokacije, namen vzorčenja (vrsta projekta, navedba projekta), vzorčevalec (ime in priimek), datum odvzema, geološki podatki (eratem, sistem, serija, stopnja; litologija, formacija, prisotnost makrofosilov, vrsta fosila), lokacija skladiščenja vzorca, izvedene analize (vrsta analize, analitska metoda, ali ima vzorec zbrusek, ali je vzorec zmlet, ali je vzorec presejan, ali je vzorec izpirek), podatki o vnašalcu (ime in priimek, čas vnosa), podatek o objavi (tip publikacije, PID publikacije), status vzorca, tip dostopa, kvaliteta vzorca, lastništvo vzorca, navedba pogodbe o lastništvu, čas hrambe.

Za gradivo, ki bo oddano v sistem GeoLib bodo zagotovljeni metapodatki v skladu z zahtevami GeoLib-a.

3.3. Ali bodo metapodatki vsebovali ključne besede za izboljšanje najdljivosti in možnosti ponovne uporabe?

☒ Da

☐ Ne

V podatkovni bazi bo mogoče iskanje po vseh zgoraj naštetih metapodatkih geoloških vzorcev.

4. ZAGOTAVLJANJE DOSTOPNOSTI PODAKOV (A)

4.1. Ali bodo vsi podatki odprto dostopni?

- ☐ Da
☒ Ne

Podatki bodo na voljo zaposlenim na GeoZS, po potrebi pa tudi drugim, ki bodo za podatke zaprosili po uradni poti. Ob koncu projekta podatkovna baza geoloških vzorcev še ne bo vsebovala vseh podatkov vzorcev, ki jih skladiščimo v centralnem skladišču, zato je pred končanjem ne bomo delili.

4.2. Pojasnite, kdaj boste podatke delili oz. kdaj bodo odprto dostopni in za koliko časa?

V času trajanja projekta ni predvideno, da bodo podatki o geoloških vzorcih, ki bodo vneseni v podatkovno bazo, deljeni izven organizacije. Podatki bodo javno objavljeni, ko bo vzpostavljena celotna digitalna infrastruktura za upravljanje s podatki geološkega repozitorija. S tem bomo zagotovili sistematično in nadzorovano deljenje podatkov.

4.3. Ali pričakujete omejitve glede deljenja podatkov ali so prisotni razlogi za določitev embarga?

Ko bo vzpostavljen celotna digitalna infrastruktura geološkega repozitorija, bo omogočen dostop do podatkov z avtentikacijo. Z določitvijo različnih tipov uporabnikov (t.i. user roles) bo nadzorovan dostop do podatkov o vzorcih oz. o sklopih podatkov o vzorcih.

Deljenje podatkov o vzorcih, za katere bo določeno, da so zaupne narave, bo omejeno v skladu z načelom »odprti, kolikor je mogoče in zaprti kolikor je nujno«. Omejen bo dostop do osebnih podatkov, ki se hranijo. Uvedli bomo evidenco dostopa do osebnih podatkov; uporabniki bodo obveščeni o beleženju evidence.

4.4. Ali bo za dostop do podatkov oz. njihovo branje potrebna dodatna dokumentacija oz. ustrezna programska oprema?

- ☐ Da
☒ Ne

Za dostop do podatkov bodo lahko uporabljena široko uporabljana MS Office orodja. Za dostop do podatkov in njihovo branje iz PostgreSQL baze podatkov ni potrebna dodatna dokumentacija ali posebna programska oprema. PostgreSQL je odprtokodna relacijska podatkovna baza, ki omogoča enostaven dostop do podatkov preko standardnih SQL poizvedb.

5. ZAGOTAVLJANJE INTEROPERABILNOSTI PODAKOV (I)

5.1. Katere geslovnike ali šifrantne boste uporabili pri pripravi podatkov in metapodatkov?

Pri projektu ne bomo uporabili geslovnikov ali šifrantov.

6. ZAGOTAVLJANJE PONOVNE UPORABE (R)

6.1. Kakšno dokumentacijo, potrebno za ponovno uporabo podatkov, boste zagotovili?

V konceptualnem modelu celostne obravnave vzorcev bo prikazana shema podatkovne baze. Zapisane bodo tudi podrobnosti o strukturi podatkovne baze, kot so tabele, polja, relacije med tabelami in druge pomembne informacije.

Vse datoteke, ki jih bomo generirali tekom projekta, bodo jasno poimenovane z vsebinsko logičnim, a kratkim imenom, brez strešic, z navedbo datuma izdelave verzije dokumenta.

6.2. Opišite, s katerimi odprtimi licencami boste opremili svoje podatke.

Predvideno je, da bomo podatke opremili z CC BY-NC-SA licenco.

6.3. Kakšne postopke zagotavljanja kakovosti podatkov boste uporabili?

Zajem podatkov, ki se bodo vnašali v podatkovno bazo o geoloških vzorcih bo standardiziran; to bomo izvedli s pomočjo prednastavljenih vnosnih form. S pomočjo vnosnih form bomo določili format podatkov, ki jih je možno vnašati in določili pravila za vnos (npr. obvezna polja, nabor vrednosti, transformacija koordinat ipd.).

7. ETIČNE IN PRAVNE ZAHTEVE, KODEKSI RAVNANJA

7.1. Ali obstajajo etična ali pravna vprašanja, ki bi lahko vplivala na deljenje podatkov?

- ☐ Da
☒ Ne

7.2. Ali boste med izvajanjem projekta obdelovali in hranili osebne podatke?

- ☒ Da
☐ Ne

Hranili bomo osebne podatke (ime in priimek) skrbnika vzorca, namestnik skrbnika, vzorčevalca, in vnašalca. Glede hrambe podatkov v skladu s Splošno uredbo o varstvu osebnih podatkov bomo ravnali v skladu z navodili zato pristojne osebe na naši instituciji.

Podatki ne bodo anonimizirani.

7.3. Kako boste uredili lastništvo avtorskih pravic in pravic intelektualne lastnine podatkov, ki jih boste ustvarili ali ponovno uporabili?

Osebe, ki bodo sodelovale pri projektu, bodo na vseh avtorskih delih, ustvarjenih tekom projekta, na delodajalca trajno prenesle vse materialne avtorske pravice, navedene v 22. členu ZASP. Prenos

navedenih avtorskih pravic je izključen, neomejen, velja za območje celega sveta za ves čas trajanja pravic in vključuje pravico do nadaljnega prenosa pridobljenih pravic na tretje osebe. Pred prenosom pridobljenih avtorskih pravic na tretje osebe bo delodajalec osebe, ki sodelujejo na projektu, obvestil.

8. DRUGI RAZISKOVALNI REZULTATI

8.1. Ali boste poleg podatkov ustvarili ali ponovno uporabili tudi druge raziskovalne rezultate?

- ☒ Da
☐ Ne

Za vnos podatkov in avtomatizacijo bomo uporabljali Microsoft Power Automate. To orodje omogoča enostavno ustvarjanje in upravljanje avtomatiziranih procesov brez potrebe po obsežnem programiranju. Programska koda bo zapisana v JSON formatu, ki je pogosto uporabljan v IT okolju.

9. FINANČNA SREDSTVA

9.1. Kakšni bodo stroški ravnanja s podatki in drugimi rezultati projekta po načelih FAIR in katere (finančne in časovne) boste namenili kritju teh stroškov?

Stroški strojne opreme bodo financirani iz sredstev, namenjenim vzdrževanju infrastrukturne opreme GeoZS.

10. ODGOVORNOSTI ZA RAVNANJE S PODATKI IN POTREBNI VIRI

10.1. Kdo (npr. vloga, položaj in institucija) bo odgovoren za ravnanje s podatki (npr. podatkovni svetovalci)?

Za ravnanje z raziskovalnimi podatki bodo odgovorni podatkovni skrbniki na GIC-u. Ker je služba podatkovnega skrbništva še v fazi vzpostavljanja, ne moremo natančno navesti, katera oseba bo pri projektu opravljala vlogo podatkovnega skrbnika.