Plan upravljanja istraživačkim podacima

|  |  |  |
| --- | --- | --- |
| Opće informacije | | |
|  | Ime i prezime predlagatelja | Kristian Đokić |
|  | Matična organizacija | Fakultet turizma i ruralnog razvoja |
|  | Naziv projekta | Primjena strojnog učenja u obradi podataka dobivenih senzorima plina za analizu i klasifikaciju namirnica |
|  | Upravitelj podacima | Kristian Đokić |
| 1. | Podaci i dokumentacija | |
|  | Koje ćete podatke prikupljati, obrađivati, stvarati ili se ponovno njima koristiti? (navedite format, vrstu i opseg podataka) | Prikupljaju se podaci dobiveni mjerenjem sa razvojnim kitom Bosch Sensortec BME688. Na razvojnom kitu/pločici nalazi se osam BME688 senzora, a radi se o plinskom senzoru s umjetnom inteligencijom (AI) i integriranim visokolinearnim i visokopreciznim senzorima tlaka, vlažnosti i temperature. Podaci će biti u JSON formatu, a osim navedenih vrijednosti (otpor plinskog senzora, tlak, vlažnost i temperatura) u datotekama je dostupno još 9 stupaca opisanih u JSON datoteci.  Senzor će biti postavljen u staklenu posudu od 0,5 l, a u posudama će biti:   1. Zrak 2. Svježe mlijeko termički neobrađeno (0,2 l) 3. Svježe mlijeko pasterizirano (0,2 l) 4. UHT mlijeko sterilizirano (0,2 l) |
|  | Kako će se podaci prikupljati, obrađivati ili stvarati? (ukratko navedite metodologiju i procese osiguranja kvalitete te načine organiziranja podataka) | Mjerenje se obavlja u trajanju od 20 minuta u svakoj posudi. Mjerenja se ponavljaju svakih 24 sata, još tri puta nakon prvog mjerenja. U periodu između dva mjerenja uzorci su pohranjeni na sobnoj temperaturi (20° C).  Kvaliteta podataka osigurat će se prethodnom pripremom razvojnog kita na način da bude spojen na napajanje 24 sata. S obzirom da se ne radi o laboratorijskoj opremi, nego o senzoru koji je namijenjen za primjenu u kućanskim aparatima, greška kod mjerenja hlapljivih spojeva je ±15%.  Podaci će biti pohranjeni u četiri mape pri čemu su nazivi mapa: DAN0, DAN1, DAN2 i DAN3. Samo ime mape govori koliko je prošlo dana od inicijalnog mjerenja. |
|  | 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) | S obzirom da su podaci pohranjeni u JSON formatu, uključeni su i meta podaci kojima su definirana značenja prikupljenih sirovih podataka. Datoteke sa sirovim podacima imaju ekstenziju „*bmerawdata*“. Za dalju upotrebu vjerojatno će biti potrebna konverzija podataka iz JSON formata u CSV ili XLSX format.  Osim sirovih podataka dobivenih mjerenjem, dostupne su i JSON datoteke u kojima su podaci o konfiguraciji razvojnih kitova/pločica, a prepoznaju se po ekstenziji „*rmeconfig*“. Priložena je i datoteka u kojoj su definirane oznake (labels) za četiri mjerenja po danu, a ta datoteka je također u JSON formatu i ima nastavak „.*bmelabelinfo*“. Navedene datoteke mogu se koristiti u programu Bosch AI-Studio. |
| 2. | Etika, pravo i sigurnost | |
|  | 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)? | Mjere se karakteristike mlijeka za koju ne vrijedi opisana regulativa. Količina mlijeka koja se koristi u mjerenjima manja je od 1 litre. |
|  | 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 nisu osjetljivi i nema potrebe posebno ih štititi. Za vrijeme prikupljanja podataka, računalo na kojem se privremeno pohranjuju zaštićeno je lozinkom. |
|  | Kako ćete upravljati zaštitom autorskih prava i intelektualnog vlasništva? Tko će biti vlasnik podataka? Koje će se licencije primjenjivati na podatke? Koja će se ograničenja primjenjivati na ponovnu uporabu osobnih podataka? | Ne očekuje se da će rezultat istraživanja dovesti do patenta. Ostali problemi intelektualnog vlasništva će se u rješavati prema preporukama institucije [Fakultet turizma i ruralnog razvoja]. Budući da podaci nisu podvrgnuti ugovoru, te se neće patentirati, objavit će se kao otvoreni podaci pod licencom Creative Commons CC0. |
| 3. | Pohrana i čuvanje podataka | |
|  | Kako će podaci biti pohranjeni i kako će biti napravljena sigurnosna kopija podataka (*backup*) tijekom istraživanja? Koji su kapaciteti čuvanja podataka kojim raspolažete? Kojim se procedurama koristite za sigurnosnu kopiju (*backup*)? | Elektronički podaci pohranit će se na računalu glavnog istraživača, pri čemu se sigurnosne kopije izrađuju nakon svake izmjene datoteke ili pojave nove datoteke. Za izradu sigurnosnih kopija koristi se sustav Microsoft OneDrive. Kapacitet za pohranu iznosi 1TB, ali sirovi podaci snimljeni u jednom danu ne prelaze veličinu od nekoliko MB. |
|  | Koji je vaš plan čuvanja podataka? U kojim će se formatima čuvati? | Podaci će se čuvati najmanje pet godina nakon završetka projekta u izvornom formatu dobivenom od strane razvojnog kita/pločice Bosch Sensortec BME688. |
| 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? | Konačnu verziju skupa podatka voditelj projekta podijelit će putem institucijskog repozitorija [Fakultet turizma i ruralnog razvoja] uspostavljenog u nacionalnom sustavu Dabar. Podaci će biti objavljeni pod CC0 licencom. Institucijski repozitorij u sustavu Dabar odabran je jer podržava FAIR principe: skupovima dodjeljuje trajni identifikator URN:NBN, osigurava vidljivost podataka putem OpenAIRE portala i Google Scholara te tražilice dabar.srce.hr, a ujedno doprinosi vidljivosti i transparentnosti rada. |
|  | Ako postoje podaci koji se ne smiju dijeliti (prijavitelji vezani zakonskim, etičkim, autorskim pravila, povjerljivošću i sl.), pojasnite razloge ograničenja. | Svi podaci se smiju dijeliti i ne postoji ograničenje za dijeljenje podataka, osim onih definiranih samom licencom. |
|  | Potvrdite da ćete se koristiti digitalnim repozitorijem koji je u skladu s načelima *FAIR-a*. | Potvrđujem |
|  | 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). | Potvrđujem |

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)](https://repozitorij.srce.unizg.hr/islandora/object/srce:327)