

Zapisnik sestanka RDA vozlišča Slovenija

Datum sestanka: 15. november 2021

Udeleženci: Milan Ojsteršek (UM), Jernej Pliberšek (Geološki zavod Slovenije), Uroš Kunaver (CTK), Mojca Rupar Korošec, Maja Vihar (CTK), Katja Meden (IJS), Tomaž Erjavec (CLARIN-SI), Sonja Bezjak (ADP), Andrej Pančur (INZ/DARIAH), Mojca Kotar (UL), Janez Štebe (ADP), Mitja Vovk Iskrič (CTK), Jasna Šinigoj (Geološki zavod Slovenije), Matija Krivic (Geološki zavod Slovenije), Damjana Drobne, Ana Inkret

Dnevni red sestanka:

1. Vzpostavljene področne pristope pri vrednotenju in izboru raziskovalnih podatkov (5 minut CLARIN.SI in ADP)
2. Vprašanja in želje za poenotenje (razprava cca 20 minut, vodi Milan Ojsteršek)
3. Interes za sodelovanje in izmenjavo izkušenj ter pristopov (razprava cca 20 minut)

K 1.

Janez: ADP želi spodbuditi in nagraditi raziskovalce, ki želijo deliti podatke. Podatki morajo biti celoviti, imajo dolgoročno vrednost za znanost, zagotavljajo ponovljivost. BIBLIO-D 2.20 kategorija je nastala na pobudo ADP, implementirana je v knjižnični sistem. Ta sistem ima večletno tradicijo, iz katere je potrebno izhajati: pri poenotenju pristopov je smiselno upoštevati uveljavljene pristope na nacionalni ravni, uskladiti nove pristope z uveljavljenimi in nato nasloviti primanjkljaje. ADP opravi recenzijo podatkov – pregleda formate, prisotnost metapodatkov in dokumentacije, ovrednoti potencial ponovne rabe preko kriterijev pomena za znanost (npr. edinstvenost pridobljenih podatkov, napor pri pripravi, kakovost in celovitost podatkov in podobno) – tudi vidike kakovosti in znanstvene relevance kar v FAIR načelih sicer ni zajeto.

Tomaž: CLARIN repozitorij (<https://www.clarin.si/repository/xmlui/>) je primerljiv ADP – dajalci sami izpolnijo metapodatke, ki jih pregleda uredniška ekipa glede na pravilnik. Pri pridobivanju BIBLIO-D so se naslonili na ADP (glej <https://docs.google.com/document/d/1KMj2IU3nl-GsAP-HqO8gq0X2ISXPTXWnlApxNJaE58w/edit>), pridobitev pomeni dodatno motivacijo za dajalce, da podatke objavijo v repozitoriju.

K 2.

Milan: subjektivnost metrik, ki jih uporabljata ADP in CLARIN. Potrebno bi bilo dodati tudi kakšne kvantitativne metrike. Definirati bi bilo potrebno npr. minimalne metodološke kriterije, kriterije znanstvene odličnosti, obseg dela (količina, porabljene ure dela, število spremenljivk...). Težko je tudi oceniti uporabno vrednost za znanost in ali gre pri podatkih za raziskovalno praznino.

Viri za druge modele:

- FAIR Data Maturity Model. Specification and Guidelines: <https://zenodo.org/record/3909563#.YYztYmDMKU>
- Recommendations on FAIR metrics for EOSC: <https://op.europa.eu/sl/publication-detail/-/publication/ced147c9-53c0-11eb-b59f-01aa75ed71a1>
- F-UJI Automated FAIR data assessment tool: <https://www.f-uji.net/>
- FAIRsFAIR D2.7 Framework for assessing FAIR Services: <https://doi.org/10.5281/zenodo.5336233>.
- FAIRsFAIR Data Object Assessment Metrics (0.4). Zenodo. <https://doi.org/10.5281/zenodo.4081213>

Tomaž: FAIR so manj lastnosti podatkov kot lastnosti repozitorija. Repozitorij določa formate podatkov, ki jih objavlja. Subjektivnost kriterijev o znanstveni odličnosti drži, a se ji je težko izogniti, je v domeni strokovnjaka – primer SAZU, ki vključuje kriterij vloženega dela.

Uroš: Pomen podatkov se izredno razlikuje po panogah. Razviti skupne kriterije, ki bi jih lahko uporabil ARRS, je zelo težko. Za nekatera področja so že uveljavljeni tuji repozitoriji, kako bi to ocenjevanje potekalo za podatke, ki so objavljeni tam?

Dilema je v poudarjanju točkovanja in nagrajevanju, s katerim se znova vzpostavi tekmovanje med področji. Rešitev je morda v recenziranju, ki ga opravljajo področni strokovnjaki, ki naj obravnavajo tudi podatke. Morda je primernejša simbolna nagrada za sodelovanje oz. deljenje.

Milan: Pomen objave surovih podatkov.

Tomaž: Če ni točkovanja, je mogoče podatke že zdaj zavesti v Cobiss pod kategorijo 2.20.

Damjana: Kompromis pri vrednotenju, najti ustrezne kriterije. Deljenje podatkov se uveljavlja kot standard v znanosti, zato se moramo skupaj dogovoriti za minimalne standarde. Deljenje surovih podatkov je bilo do nedavna samo dodatno delo, z odprto znanostjo pa se spreminja njihova vloga – raziskovalci morajo zato sodelovati pri oblikovanju metapodatkov, ki bodo omogočali ponovno rabo.

Maja: previdno pri točkovanju s podatki, saj se vrednotenje znanosti svetovno odmika od kvantitativnega nagrajevanja. Primer Nizozemske, kjer odprte (in kakovostne) raziskave nagrajujejo z več sredstvi (glej <https://www.nwo.nl/en/news/26-projects-stimulate-open-science>). Tipologija 2.20 govori o zaključeni znanstveni zbirki podatkov, pojavljajo pa se primeri, ko raziskovalci sproti objavljajo vmesne podatke, ki po mnenju OSIC niso enako pripravljene.

Nov akcijski načrt vključuje spodbudo, da finančni načrt vključuje delo s podatki. Razlikovanje med področnimi repozitoriji s certifikati, ekipami, ter tistimi za samoarhiviranje. Vprašljiva je smiselnost podvajanja raziskovalnih podatkov na več mestih, razen v primerih zagotavljanje specifičnih storitev.

Uroš: Kako ukrepati v primerih, ko ni specializiranega repozitorija, podatki pa morda zelo kakovostni? Pomislek o gradnji nacionalnih repozitorijev, ko že obstajajo uveljavljeni področni repozitoriji.

K 3.

Naslednji koraki: oblikovanje sklepov, naslavljanje vprašanj in predlogov na druge inštitucije, sodelovanje v prihodnosti. Eno od možnosti: vzpostavitev podatkovne revije z meddisciplinarnim uredniškim odborom.

Uroš: Nacionalni sistem točkovanja bo določil ARRS oz. OSiCi, lahko pa bi oblikovali minimalne zahteve za raziskovalne podatke, da so primerni, jih je mogoče hraniti. Druga stvar je izobraževanje uporabnikov.

Sonja: Določiti vloge različnih vrst repozitorijev (institucionalni, nacionalni, področni), ki naj sodelujejo med sabo. Repozitoriji v vlogi »založnikov« podatkov, zato se je mogoče nasloniti za obstoječe politike.

Sklep: Nadaljujemo z aktivnosti – se po potrebi posvetujemo in usklajujemo med seboj; nadaljujemo s pogovori.

