

Odprta znanost v Sloveniji in svetu

Odtenki odpiranja

Pred kratkim je bil med novicami revije *Nature* objavljen poziv znanstvenikov za popolnoma odprto deljenje genomskih podatkov koronavirusa (Van Noorden, 2021). Poziv je na prvi pogled nekoliko presenetljiv: ali niso v času korone podatki takoj na voljo vsem, da s skupnimi močmi izvemo kar največ, da se lahko zavarujemo pred trenutno največjo grožnjo človeštvu? Ozadje poziva je zanimivo za uvod v posebno številko s prispevki o odprti znanosti, s posebnim poudarkom na odpiranju raziskovalnih podatkov. Kakšni so odtenki odpiranja, če je do skrajnosti popolnoma odprtega še več odtenkov bolj ali manj omejenega dostopa? Komu lahko odprtost koristi in komu škodi in ali ne bodo tisti, ki so na slabšem, še več izgubili?

Razprava v *Nature* je v precep vzela podatkovno storitev GISAID,¹ ki prevzema genomske podatke z vsega sveta. Podobne storitve ponujajo tudi druga podatkovna središča. Tako se vzpostavlja tudi slovenski COVID-19 podatkovni portal,² del skupne istoimenske evropske pobude, s priporočili, kako raziskovalne podatke s posameznih področij genomike oddati v podatkovne arhive evropske povezave EBML-EBI,³ ki pa se od GISAID razlikujejo v nekaterih podrobnostih.

V svojem pozivu za še večjo odprtost genomskih podatkov se znanstveniki dotikajo posameznih odtenkov odprtosti. Najprej se

1 Glej <https://www.gisaid.org/>.

2 Glej <https://www.covid19dataportal.si/>.

3 Glej <https://www.ebi.ac.uk/submission/>.

zavzemajo za vseh omejitev prosti dostop do podatkov in pri tem navajajo problematičnost zahteve, da se morajo uporabniki podatkov registrirati in da podatkov, do katerih dostopajo, ne smejo deliti naprej. Nasprotna tovrstnemu omejenemu dostopu je popolna odprtost, ki je običajno urejena s pripisovanjem licenc prostega deljenja podatkov CC0.⁴ Ta v svoji skrajnosti tudi na drugih področjih, ne samo na področju genomskih podatkov, dostikrat zbuja pomisleke tako med dajalci podatkov kot pri skrbnikih repozitorijev. Licenc prosta javna raba podatkov CC0 uporabnika razbremeni morebitnih pravnih obveznosti, uporaba podatkov pa temelji na zaupanju in ne na pravnem okviru. Denimo, dosledno citiranje uporabljenih podatkov je pričakovano, ne pa zapovedano.

V omenjeni razpravi zagovorniki bolj reguliranega dostopa do genomskih podatkov navajajo pomisleke manjših laboratorijev, večinoma iz manj razvitih držav, da bodo deljenje podatkov vseprek izkoristili tisti z zmožnostmi analiz velikega obsega, z zmogljivejšimi orodji in računalniško infrastrukturo ter z znanjem, prehiteli druge in uporabili podatke brez upoštevanja prispevka dajalcev podatkov. Regulacijo in pogojevanje dostopa predstavniki omenjenega GISAID utemeljujejo s tem, da ščitijo manjše in šibkejšje dajalce podatkov, saj ti sicer svojih zbranih podatkov sploh ne bi delili. Podobni pomisleki so prisotni na drugih področjih. Obstaja bojazen, da bodo podatke o kliničnih testiranjih na bolnikih izkoristile farmacevtske tovarne za produkte in zaslužek, kmetijske pa za prodajo semen in zaščitnih sredstev, ter da bodo bogatejše države in inštituti izkoristili manj razvite, manjše. Teme neenakosti in izkoriščanja na račun dostopa do podatkov se v pričujoči številki posebej zavzeto lotevajo prispevki S. Masten in S. Bezjak, ki govorita o podatkovnem kolonializmu in podatkovljenju, M. Nardoni z izpostavljanjem algoritemskega izžemanja prosto dostopnih podatkov s strani industrije spletnih platform in S. Madackega o pogosto občutljivih podatkih, ki so jih zbrali in izkoristili tuji gostujoči raziskovalci v okoliščinah vojne in nasilja v Bosni in Hercegovini, za njimi pa se je izgubila vsaka sled. Če popolna odprtost škodi tistim, katerih podatki so zbrani, pa naj so to družbene manjšine, bolniki, žrtve nasilja ali podatki uporabnikov spletnih platform, je nujno regulirati njihovo razpoložljivost.

Kljub različnim pomislekom je poziv k čim večji odprtosti, razen izjemoma, še vedno osnovno vodilo, zlasti ko je posredi javni interes.

4 Glej <https://creativecommons.org/share-your-work/public-domain/cc0/>.

V pozivu v *Nature* kritizirajo tudi, da so v GISAID na voljo samo že delno predelani podatki. Nepredelani podatki, včasih jim rečemo tudi »su- rovi«, drugim uporabnikom omogočajo, da neodvisno od prvotnega raziskovalca sledijo in preverjajo postopek priprave podatkov za nadaljnje analize. S tem se odpre možnost odkrivanja in odpravljanja napak, tako pa prihajajo do veljavnejših rezultatov. Problematizirajo tudi različne pogoje za dostop do podatkov. Vsi uporabniki ne dobijo dovoljenja za prenos celotnega obsega podatkov, ampak so omejeni na manjše izseke. Tu je v odtenkih odprtosti prisoten še en vidik podatkovne pravičnosti, tj. diskriminacija različnih uporabnikov podatkov. Na podobne dileme naletijo raziskovalci, ki želijo uporabljati podatke družbenih medijev. Tudi če korporacije načeloma dovolijo uporabo za raziskovalne namene, so podrobnosti pogojev uporabe nejasne in vodijo v negotovost. Kakšen obseg, katere vsebine lahko uporabljamo, ali lahko podatke predelujemo in delimo dalje? Za etično skladno rabo je pomembno razčistiti tudi odsotnost soglasja za uporabo podatkov s strani posameznikov, ki objavljajo vsebine (Littman, 2019). Prav tako pa je vprašanje, tako kot pri očitkih GISAID, komu in pod kakšnimi pogoji ter za kakšen namen so podatki dostopni. Nekateri imajo prednost pri dostopu do podatkov družbenih medijev s posebnimi dogovori z lastnikom, to so privilegirani raziskovalci z dostopom do zlate jame podatkov (Scola, 2018), medtem ko so običajni raziskovalci omejeni na izsek podatkov (glej na primer Bishop, 2019). Drugo področje, a skoraj povsem enake dileme glede odprtosti podatkov za raziskovanje.

Ali ni znanost sama po sebi odprta?

Odprtost znanosti je v okviru etosa znanosti nekaj, česar ni treba posebej poudarjati; je samoumevna v predstavi znanosti kot občestva, ki sledi ciljem kodificiranja veljavnega znanja, v kontrastu do epidemije antiintelektualizma, lažnih novic in teorij zarote. Začetnik sociologije znanosti Robert Merton je enega od štirih temeljev tega etosa poimenoval *komunizem* (Merton, 1942), polje občega človeškega, kjer so dosežki, nastali s sodelovanjem več generacij, last vseh. Izraz so zaradi politično-ideološkega predznaka pozneje raje prevajali kot *komunitarizem*. Izhajajoč iz občesti kot oznake, ki nasprotuje posedovanju in lastnini (pogosto prisotnemu razumevanju zbranih

raziskovalnih podatkov kot »mojih«), lahko pritegnemo argumentaciji Z. Mlinarja v prispevku v pričujoči številki in govorimo o *občanskosti*, ki nosi tako občestveno skupnostno sodelovalno komponento pri ustvarjanju produktov znanosti kot obče človeško javno dobro pri posedovanju in uporabi teh produktov.

V čem je potem problem, ob privzeti samoumevnosti odprtosti, zakaj o tem razpravljati? Tu smo, razumljivo, že na meji etosa idealnotipske podobe znanosti ter realnosti opazovanja in razumevanja znanstvenega delovanja v različnih kontekstih, tako omejitev in zahtev vsakodnevnega dela raziskovalcev v njihovih mikrookoljih kot v odnosu do družbe. Govorimo lahko o različnih vidikih odprtosti, o odprtosti navzven, v javnost, v družbo, v naravo, ter o odprtosti in demokratizaciji znanosti navznoter, v deljenju podatkov s kolegi, solidarnosti in transparentnosti raziskovanja. Realnost so med drugim različne ovire ter pomanjkanje spodbud in motivacije za zares odprto delovanje, ko se raziskovalci sprašujejo o smiselnosti odpiranja: razumljivo je, da ne vsega in ne za vsako ceno (Borgman, 2012). V tekmovalnem okolju znanosti je pomembno nadomestiti izgubo konkurenčne prednosti, ki bi lahko sledila iz pripravljenosti deliti ustvarjene raziskovalne podatke, katerih potencial za lastno analizo in objavljanje publikacij ni izčrpan.

Da samo sledenje etosu ni dovolj in da so potrebne korekcije, ugotavlja vrsta pobud. Govorijo o odprtih publikacijah (znanstvenih člankih in monografijah), odprtem recenziranju, odprtih raziskovalnih podatkih, odprti programski kodi, odprtem in alternativnem vrednotenju znanosti ter o občanski znanosti kot delu odprte znanosti.⁵ Vsako od področij zaznamuje svojstvena zavzetost za spremembe, ki se odraža v različnih deklaracijah in manifestih v prid reševanju zagat odprtosti. Najbolj znani so Plan S, deklaracija DORA ter Leidenski manifest s področja publikacij in vrednotenja raziskovalnih dosežkov. Z odprtimi raziskovalnimi podatki se posebej ukvarjajo priporočila uglednih svetovnih konglomeracij, kot sta OECD in UNESCO.⁶

5 Prispevkom Z. Mlinarja, M. Pušnika ter A. Slavec in A. Srakarja je skupno, da se ukvarjajo z občansko znanostjo, pri tem pa se je med avtorji razvila terminološka razprava o slovenjenju *citizen science*, ki ji lahko sledite predvsem v besedilu Z. Mlinarja. Omenjeni avtorji so v svojih prispevkih uporabljali dve različici, in sicer občanska znanost in skupnostna znanost, odločitev smo prepustili njim. V uvodniku uporabljamo termin občanska znanost.

6 Izšla je tudi posodobljena različica priporočil OECD z naslovom *Recommendation of the Council Concerning Access to Research Data from Public Funding* (glej OECD Legal Instruments, 2021). UNESCO trenutno razpravlja o osnutku priporočil o odprti znanosti (glej UNESCO, 2021).

Za podlago posebne številke smo izhajali iz delovanja Research Data Alliance (RDA), tako tudi pričujoče razmišljanje o odprti znanosti pretežno preizprašuje odprtost raziskovalnih podatkov.

V različnih pobudah se je izoblikoval načrt odprte znanosti, ki ga v svojem prispevku podrobneje predstavlja M. Pušnik. Mednarodne pobude, deklaracije ter različne nacionalne, institucionalne in publicistične politike vzpostavljajo okvir za premagovanje ovir in pripravo spodbud za odprtost raziskovalnih podatkov in drugih vidikov odprte znanosti, tako da vzpostavljajo okvir pravičnosti, kjer veljajo enake obveznosti za vse sodelujoče: če boš ti delil svoje podatke, jih bom tudi jaz, in če so odprti podatki obvezni za vse, odpade *free-riding*.

Med področji Odprte znanosti (z veliko začetnico) pa so tudi prehodi in medsebojna prekrivanja. Podatki pomenijo prednost pri objavi (če dostop do njih omejimo), lahko pa prispevajo k citiranju (če jih delimo). S. Bezjak iz Arhiva družboslovnih podatkov predstavi koncept podatkovne objave, ki lahko kot nekakšen hibrid med besedilnim člankom in raziskovalnimi podatki omili protislovja in krivičnosti sistema vrednotenja v znanosti. Uporabljene podatke tako kot članke citiramo v seznamu literature. Podatkovna objava je ovrednotena kot prispevek znanosti po analogiji z znanstvenim člankom. Njenemu prispevku ter prispevku M. Petek, K. Gruden in Š. Baebler z Nacionalnega inštituta za biologijo je skupno omenjanje še enega hibrida v svetu znanstvenega publiciranja, podatkovnih revij in podatkovnih člankov.

Vrednotenje in točkovanje znanstvenih dosežkov

Vrsta prispevkov v pričujoči številki se dotika problematičnih vidikov vrednotenja znanosti ter do skrajnosti prignane tekmovalnosti v točkovanju objav v revijah s faktorji vpliva in lovljenju citatov, ki je na sistemski ravni ena od ovir na poti večjemu deljenju produktov znanosti. G. Majdič in S. Dolenc, vsak s svojega vidika, opažata hiperprodukcijo z namenom nabiranja točk znanstvenih objav ter uporabo različnih pollegalnih ali nelegalnih prijemov, ki pa so v ekscesni obliki samo simptom pritiska, ki ga občutijo vsi raziskovalci – preživetvene nuje za dosežkarstvo. Revije skupaj z raziskovalci logiki nabiranja odmevnosti sledijo tudi na škodo napredovanja spoznanj, na primer s tem ko dajo prednost objavam, ki prikazujejo potrditve preizkusov, s čimer se objavljeni rezultati različnih preizkusov pristrano

nagibajo v pozitivno smer, saj negativni izidi ne pridejo do objav. M. Pušnik poziva k bolj sodelovalni, skupnostni oziroma občanski znatnosti za odmik od arbitrarnosti kvantitativnih metod vrednotenja.

Vloga politike, države in financerjev je, da sledeč mednarodnim pobudam nehotene posledice točkovanja blaži in uravnotežuje, deluje preventivno in vzpostavi okvir pravičnosti, za vse enakih zahtev glede odprtosti, tako da ni razloga za občutek izgube konkurenčnosti. Slovenski Akcijski načrt in strategija za uveljavljanje načel odprte znatnosti (glej Vlada Republike Slovenije, 2015; 2017) sta bila zastavljena po vzoru nekaterih najrazvitejših držav in v skladu s priporočilom EU (Sveinsdottir idr., 2020).⁷ Osrednji zastavek je bil določiti, da za vse, ki prejemajo javna sredstva za raziskovanje, veljajo enake obveznosti glede odprtega dostopa do publikacij in raziskovalnih podatkov, ki izhajajo iz projektov z javnim financiranjem. Dovoljene so izjeme, ki pa ne smejo biti poljubne, kakor so nekateri razumeli določilo t. i. *opt-out*,⁸ ki je bilo povzeto po pravilnikih pogodb za raziskovalne projekte razpisa Obzorje 2020, temveč morajo tisti, ki jih želijo uveljaviti, dokazati, da imajo legitimne razloge in izpolnjujejo pogoje, pod katerimi so izjeme dovoljene – na primer zaščita osebnih podatkov in drugi taksativno naštetih veljavni razlogi za izjeme odprtega dostopa.

Odprto, kolikor je mogoče, in zaprto, kolikor je nujno

Prav na področju vprašanj izjem, povezanih z varstvom osebnih podatkov – če ga izpostavimo kot primer za nadaljnjo ponazoritev odtenkov odprtosti –, se pojavljajo različne interpretacije, ki lahko imajo interesno ozadje in se iztečejo v pogosto nerazumnem omejevanju dostopa do podatkov. Za podatke v terminologiji odprte znatnosti velja, da so odprti, če so »odprti, kolikor je mogoče, in zaprti, kolikor je nujno«. Pri interpretaciji možnosti odpiranja podatkov na plano pridejo stališča in prepričanja, ki jih je na področju odpiranja podatkov javnega sektorja z uporabo teorije racionalne izbire sistematično razdelal F. Ritchie (2016). Lastniki podatkov, v njegovem

7 Analiza slovensko strategijo na več mestih izpostavlja kot zgled.

8 Možnost, da se posamezen raziskovalni projekt izvame iz obveznosti odprtega dostopa do raziskovalnih podatkov. Glej OpenAIRE, n. d.

primeru uradniki na različnih agencijah, imajo do podatkov pretirano zaščitniško držo. Namesto privzete odprtosti (*Open Access By-Default*; glej Deklaracija EOSC, 2017) je njihova udobna pozicija skrajno omejevanje dostopa do podatkov ne glede na koristi za raziskovalce, javne koristi in navsezadnje koristi za njihove lastne ustanove, če vzpostavijo zaupanje z zunanjimi uporabniki podatkov – ti prispevajo k metodološkemu izboljšanju podatkov in novim rezultatom analiz, ki so uporabni za cilje teh ustanov. Drža javnih ustanov so torej privzeto zaprti podatki, ki jo utemeljujejo z iskanjem vseh mogočih razlogov, zakaj podatkov ne bi delili. Med razlogi so priročni takšni, ki dajejo videz skrbnega ravnanja s podatki pri varovanju zasebnosti. Popularni priročniki za zakrivanje podatkov zaradi poenostavitve matematičnega postopka statističnega nadzora pred razkritjem izhajajo iz predpostavke skrajnih scenarijev zlorabe podatkov nekoga s slabimi nameni. Z uporabo takšnih vodil se na podlagi možnosti razkritja posamezne enote v podatkih, ki bolj ali manj ostaja le v teoriji, dostop do za analizo primernih podrobnih izvornih podatkov do skrajnosti omeji in, denimo, v odprti dostop ponudi povsem okleščeno in za resno analizo neuporabno različico agregiranih podatkov.⁹ Nasprotno lahko ob seznanjanju javnosti s koristmi odprtih podatkov in z regulacijo dostopa, ki upošteva realne okvire tveganj ter vzpostavlja zaupanje med uporabnikom in lastnikom podatkov, poskrbimo, da je veliko več podatkov bolj dostopnih, tj. dostopnih, kolikor je mogoče, na primer po modelu »petih ukrepov za varno rabo podatkov« (Desai idr., 2016). Podobno držo *privzeto zaprtega* imajo včasih raziskovalci, ki ustvarjajo podatke v akademskem okolju, z naštevanjem vseh mogočih razlogov, zakaj podatkov ne bi delili, pa naj bo grožnja, da bo podatke kdo zlorabil in izkoristil njim v škodo, utemeljena ali ne.

F. Ritchie (ibid.) na vprašanje, ali lahko zagotovimo učinkovito uporabo administrativnih (in drugih javnih) podatkov, odgovarja optimistično: lahko, če resnično hočemo.¹⁰ V pričujoči številki je kar nekaj prispevkov, ki se dotikajo tovrstnih dilem ter vsebujejo tako argumente koristi od večje dostopnosti in uporabnosti podatkov kot tudi predlagajo rešitve, kako narediti podatke bolj dostopne brez pretiranih tveganj. A. Slavec in A. Srakar o tem, kaj je mogoče storiti, pišeta na primeru zdaj že vsem znanega in zaupanja vrednega Sledilnika

9 Primerjaj mnenje Informacijskega pooblaščenca z dne 18. decembra 2017 (Informacijski pooblaščenec, 2017).

10 Glej na primer praktični vodič za odločanje Elliot, Mackey in O'Hara, 2020.

COVID-19. Sledilnik je odprl vrata sodelovanju bolj kot vse deklaracije, pozivi in strategije; v projekt so vključene različne ustanove (sodelujeta tudi NIJZ in IJS), ki posredujejo in si izmenjujejo podatke. In vendar so, kot pokažeta avtorja, še vedno prisotne bolj ali manj razumne omejitve, ki bi jih bilo treba zmanjšati. V skupnosti Sledilnika in tudi širše zanje ponujajo različne rešitve. Avtorja navajata primer enkratno hitrega odziva podatkovnih specialistov z različnih področij, od genomike in epidemiologije do družboslovja, skupnostne in občanske znanosti ter pravic staroselcev, v okviru združenja Research Data Alliance (RDA), ki so že maja pripravili vodiče za pripravo podatkov v zvezi s COVID-19 za najširšo rabo (glej RDA COVID-19 Working Group, 2020). Na prakso in delovanje RDA se naslonita tudi M. Kerstein in M. Lovec, ki v uvodu prispevka na temo uporabnosti administrativnih podatkov s področja kmetijstva za načrtovanje politik predstavitva prizadevanja za poenotenje pristopov v mednarodnem okolju, ki bi povečalo preglednost, povezljivost, dostopnost in uporabnost podatkov. Priporočila skupine RDA za agrisemantiko, etična priporočila za pripravo podatkov kmetijstva in prehrane združenja GODAN (Global Open Data for Agriculture and Nutrition) in druga priporočila pomagajo pri presoji *stanja zrelosti* organizacij pri odpiranju svojih podatkov. V sklepu avtorja ponudita predloge za izboljšanje stanja in dvig kulture deljenja podatkov, in sicer v realističnem prepoznanju, da vsega ni mogoče doseči naenkrat, da pa imajo že majhni koraki in poskusi oblikovanja dostopnejših podatkov demonstracijske učinke za podporo uvajanju obsežnejših izboljšav. Ali po priljubljenem reflu prof. Zupana, s katerim smo sodelovali pri projektu, katerega rezultate prikazujeta Kerstein in Lovec: »Najprej se lotimo *low hanging fruit*.« Naj pri tem omenimo, da smo v razpisu za prispevke te številke izpostavili željo, da bi prispevki poleg načel tematizirali tudi njihov praktični vidik: za pot naprej potrebujemo eksperimentiranje pri implementaciji in s preizkušanjem rešitev postopno izboljšujemo stanje. Take vrste je prikaz dobre prakse za načrtovanje in ravnanje s podatki biotehnoloških raziskav; vpeljuje ga prispevek M. Petka, K. Gruden in Š. Baebler, ki je lahko zgled tudi za druga področja.

FAIR, pa tudi CARE

Pri ocenjevanju dostopnosti raziskovalnih podatkov so se v zadnjem času uveljavila načela FAIR (*Findable, Accessible, Interoperable, Reusable* ali najdljivi, dostopni, povezljivi in uporabni za druge), v vseh odtenkih prepoznavanja, da FAIR ni nujno povsem odprto. Smiselno je izpostaviti različne zahteve glede doseganja posameznih načel, odvisno od tradicij in posebnosti področij. Nekje bo cilj že doseganje F, kakor Kerstein in Lovec ugotavljata pri kmetijskih podatkih, pri katerih so dosežek že identifikacija obstoja, opis lastnosti in popis vsebin. Pomen načel pa je tudi v težnji po povečanju uporabnosti podatkov med disciplinami, odprtost v smislu dostopnosti za čim širšo drugo rabo, tudi s pomočjo razumljivosti vsebin in načinov dela s podatki za širši krog uporabnikov. Načela FAIR, ki jih v svojem prispevku predstavlja M. Dolinar in za presojo uporabnosti podatkov uporablja vrsta prispevkov v nadaljevanju, dopolnjujejo druga načela, predstavljena v besedilih S. Madackega ter S. Masten in S. Bezjak. To so načela CARE, s katerimi presojamo prispevek podatkov k občemu dobremu (*Collective benefit*); določajo, da morajo o podatkih odločati tisti, ki jih zadevajo (*Authority to control*), ter kličejo k odgovorni rabi (*Responsibility*) in upoštevanju etike (*Ethics*). Govorimo o podatkovni pravičnosti.

V prispevkih različnih avtorjev se vlečejo vzporednice pri opozarjanju tako na problem dostopnosti rezultatov znanosti kot na vlogo znanosti in znanja v relacijah moči in dominacije. A. Inkret omenja paradoks nedostopnosti znanja ob vseh tehnoloških možnostih digitalnega okolja. Z. Mlinar in D. Zadnikar opažata pristranost etabriranega znanstvenega znanja v svojem položaju: ali odvisnosti od naročnikov ali igranja orodja politik in oblasti, evropocentričnega univerzalizma. Izboljšanje bi dosegli z vključevanjem vednosti udeležencev, ki jih tovrstno znanje zadeva. S svojim bližnjim in izčrpnim razumevanjem okoliščin in posebnosti sveta, v katerem živijo, pomenijo korekcijo etabriranega znanja, pa naj bo to občanski prispevek pri znanstvenem projektu ali lastni subverzivni alternativni imaginarij znanja in vednosti od spodaj, ki oporeka univerzalnemu, s tem pa prispeva k pluralnosti razumevanja, ki raste iz skupnostnih družbenih gibanj in avtonomnih kulturnih tradicij, na primer predkolonialnih. Zaradi nasprotja interesov v razmerjih neenake moči se lahko zaupanje v znanost vzpostavlja v njeni odprtosti do različnih

»epistemoloških socialnih izhodišč«, pravi Zadnikar v svojem prispevku. Ni naključje, da sta oba avtorja družboslovca; delujeta torej na področju, na katerem so heterodoksije in multiparadigmatski pristopi uveljavljeni že zaradi prepletenosti subjekta in objekta raziskovanja, občana znanstvenika in znanstvenika občana. Vendar iz primerov, ki jih navajata, tudi sledi, da lahko določena mera odprtosti za avtentične in nezamenljive občanske izkušnje in poglede prispeva tudi k boljšim rezultatom drugih raziskovalnih področij, tako pri kakovosti kot pri pravičnosti rezultatov – komu koristijo in predvsem komu naj ne škodijo.

Politike spodbujanja odprtosti podatkov

Tisti, ki delajo¹¹ znanost, so načeloma pripravljene deliti podatke; navsezadnje želijo uporabljati podatke drugih za izboljšanje svojih rezultatov. Sistem tekmovalnosti in želja po prestižu, ki je eno od gonil uveljavljanja v akademskem okolju, jih silita v iskanje konkurenčnih prednosti oziroma njihovo ohranjanje, tudi z monopolizacijo podatkov. Status, ki si ga pridobijo raziskovalci ali raziskovalne skupine, jim omogoča, da znanstveno delujejo na zanje pomembnih področjih, da torej dobijo nove projekte. Pri razpisih, na katere se prijavljajo, pa je pomembno prikazati pretekle dosežke, katerih ne preveč posrečeno merilo so točkovane objave v prestižnih revijah. Kakor hitro se določeno merilo začne uporabljati za namene ocenjevanja znanstvenih dosežkov, se bodo raziskovalci temu prilagodili, merilo pa bo postalo neuporabno.¹²

Z vprašanji, povezanimi z odprtostjo, se sooča tudi glavni financer znanosti v Sloveniji, tj. Javna agencija za raziskovalno dejavnost RS (ARRS). Raziskovalcem bo treba pri obveznostih odprtih podatkov zaupati glede identifikacije podatkov in samopresoje njihovega pomena za drugo rabo, saj podatke s svojega raziskovalnega področja sami najbolje razumejo in vrednotijo. Zaupanje bo potrebno tudi pri utemeljevanju izjem od popolne odprtosti in temu prilagojenih režimov ravnanja s podatki. Skupni okvir pa naj bodo načela odprtosti

11 J. Adema v svojem prispevku govori o *doing of scholarship* v povezavi s kontekstom znanstvene produkcije, načinov, pogojev in posledic.

12 Kar je ponazoritev Goodhartovega zakona na področju znanstvenih politik. Glej Biagioli, 2016.

podatkov, ki temeljijo na kulturi odprtosti: na podpori načelom, kot so FAIR, in sicer na ravni stališč, na njihovem razumevanju v okviru posebnosti področij, njihovem sprejetju in ponotranjenju njihove smiselnosti, ter na znanju, kako jih izpeljati, ki se izteče v večjo razširjenost navad deljenja podatkov.

Konkretno vprašanje financerja – naj bo to ARRS ali kdo drug, podobna vprašanja si zastavljajo povsod po svetu – je, kako zagotoviti izvajanje, nadzor in vrednotenje rezultatov projektov ter preverjanje, ali so izpolnjene obveznosti, in če so, nagrajevanje. Za financerja, ki si prizadeva za uvajanje obveznosti odprtih podatkov, je ključno, da podpre pripravo in izvajanje načrtov ravnanja s podatki (NRRP),¹³ kar je še ena od inovacij na področju odprte znanosti. Za raziskovalce, če tega ne ponotranjijo, je NRRP videti kot še en obrazec, ki ga morajo izpolniti in ki jim zaradi birokracije krade čas, ko bi se lahko ukvarjali s pomembnimi stvarmi, tj. z raziskovanjem. Ob ustrezni podpori raziskovalne ustanove, financerja in podatkovnih repozitorijev sta načrtovanje in sprotna priprava podatkov za odprti dostop jamstvo, da bodo podatki bolj kakovostni, pregledni in uporabni že znotraj projekta.

Skupna izhodišča za NRRP, ki postaja obvezen izroček raziskovalnega projekta, njegova prva različica pa se v novem okvirnem programu razpisa Horizont Evrope upošteva že pri ocenjevanju znanstvene odličnosti predloga projekta, je leta 2017 pripravilo elitno združenje evropskih raziskovalnih agencij Science Europe. Njegov letos posodobljen praktični vodič (glej Science Europe, 2021) za uporabo pri raziskovalnih agencijah vsebuje novo poglavje o tem, kako ocenjevati NRRP. Tu je nakazana pot tudi za financerje in ni brez izzivov, vendar ni nepremostljivih ovir za izvedbo projektov.

Delovanje RDA vozlišča Slovenije

V projektu vzpostavitve RDA vozlišča Slovenije, ki je tudi v ozadju priprave pričujoče številke ČKZ,¹⁴ smo del aktivnosti posvetili prav

¹³ Uporabljamo slovenjenje za v tuji literaturi uveljavljen izraz Data Management Plan (DMP).

¹⁴ Posebno številko ČKZ je podprl projekt RDA Europe 4.0, ki mu je sredstva dodelil Okvirni program EU 2020 za raziskave in inovacije po pogodbi št. 777388. V okviru tega projekta so potekale aktivnosti vzpostavitve nacionalne skupnosti Združenja za raziskovalne podatke, RDA (Research Data Alliance) vozlišča Slovenije. Na tem mestu se za sodelovanje in prispevek zahvaljujemo kolegom iz projekta RDA Europe 4.0 ter drugim iz svetovne RDA, posebej vsem

vzpostavljanju podpore načrtovanju in izvajanju ravnanja s podatki v raziskovalnih ustanovah, ki je v sozvočju z omenjenimi pričakovanimi zahtevami financerjev znanosti. Logika pri tem je, da če financer v svojih pravilnikih zahteva odprte podatke in pripravo NRRP, mora poskrbeti tudi za to, da se raziskovalne ustanove pripravijo in opolnomočijo za izvajanje. Tako raziskovalce razbremenimo negotovosti in dela naporov ter povečamo pripravljenost za sodelovanje pri odprti znanosti. Glavna rezultata RDA vozlišča s tega področja sta delavnica s predstavitvami področnih posebnosti, orodij in primerov priprave NRRP¹⁵ ter prilagoditev kažipota 23 točk za raziskovalne knjižničarje in knjižničarke,¹⁶ ki lahko v ustanovah pomagajo pri izvajanju odprtega dostopa.

Druga hrbtencična dejavnost vozlišča je bila pobuda slovenskim znanstvenim revijam in založbam, da pripravijo pravila za avtorje za deljenje raziskovalnih podatkov v povezavi s člankom. Rezultat dela so osnutki navodil za deljenje podatkov pri nizu slovenskih revij na podlagi mednarodno usklajenih smernic (glej Štebe, Bezjak in Dolinar, 2020). Pomembnost teh smernic presega goli okvir uporabe pri revijah in je lahko osnova za ustrezno prilagojena navodila raziskovalcem v ustanovah in za uskladitev z zahtevami financerja. S pripravo pravil za raziskovalne projekte po zgledu smernic lahko ARRS zagotovi mednarodno in nacionalno usklajenost pričakovanj in zahtev za raziskovalce.¹⁷ Smernice se že sicer sklicujejo na Akcijski načrt, katerega posodobitev, predvsem pa kontinuiteto pri korakih razvoja ekosistema odprte znanosti pričakujemo tudi v prihodnjem obdobju.¹⁸

Tretja smer delovanja RDA vozlišča je delovna skupina za usklajevanje razvoja raziskovalne infrastrukture za podatke. Skupaj z Mlado

članom in podpornikom RDA vozlišča Slovenije. Vozlišče je odprto za nove člane, ki se lahko pridružijo prek portala <https://www.rd-alliance.org/groups/rda-slovenia>.

15 Predstavitve in prispevki z delavnice Ravnanje z raziskovalnimi podatki so dostopni na: <https://www.adp.fdv.uni-lj.si/dogodki/delavnica-rda-nrrp-2020/program/>. Za prihodnjo usmeritev predlagamo sledenje priporočilom interesne skupine RDA *Active Data Management Plans, Machine Actionable DMP Specification*, in izkušnjam delovne skupine *Exposing Data Management Plans*.

16 Originalno delo Interesne skupine RDA *Libraries for Research Data Interest Group* je v slovenski priredbi dostopno na: https://zenodo.org/record/4436033#.YEDt5W_OnIU.

17 O povezavi med politikami revij in financerjev razpravljamo v Štebe, Dolinar, Bezjak in Inkret, 2020.

18 Po naših informacijah na MIZŠ posebna delovna skupina posodablja poglavje o odprti znanosti v Razvojni in inovacijski strategiji Slovenije in akcijski načrt odprte znanosti za prihodnje obdobje.

akademijo in projektom Ni4OS je pripravila konferenco s pregledom infrastrukture odprtih podatkov v Sloveniji. Zaključki konference se opirajo na izkušnje z uvajanjem CoreTrustSeal za ocenjevanje delovanja podatkovnih storitev in načel FAIR¹⁹ za ocenjevanje dostopnosti in uporabnosti raziskovalnih podatkov za druge uporabnike. Nadaljevanje dela te skupine je *Pregled slovenskih repozitorijev raziskovalnih podatkov*, ki sta ga pripravila K. Meden in T. Erjavec iz CLARIN.SI.

Številki na pot

V razpisu za tematsko številko smo zapisali, da pričakujemo prikaze zgledov in predlogov, ki naslavlajo reševanje kulturnih, družbenih, organizacijskih in tehničnih zagat in problemov v povezavi z raziskovanjem, podatki in odprtim dostopom. Prikazi naj vsebujejo predloge in primere možnosti implementacije.

Odziv na razpis po številu in raznovrstnosti prispevkov kaže, da zastavek misliti odprto znanost po eni strani zahteva bolj temeljni konceptualni razmislek o tem, kakšna so sploh epistemološka izhodišča sodobne znanosti (o čemer v svojem prispevku kritično razpravlja D. Zadnikar), o položaju znanosti in njeni družbeni vlogi, ter kako naj odprto znanost razumemo kot etično prakso (prispevek J. Adema), po drugi strani pa je razmislek tesno povezan s praktičnimi vidiki, posebnimi po področjih znanosti in uporabe, ter konkretnimi razmisleki in rešitvami v situacijah, v katerih se problemi porajajo.

Prispevki prinašajo tudi načelne razprave o terminologiji ter nakazujejo različna področja uvajanja koncepta odprtosti: odprtost za sodelovanje v različnih fazah raziskovanja, ne samo s pomočjo izkoriščanja množic pri zbiranju podatkov (*crowdsourcing*), ampak pri vseh korakih, od izbire raziskovalnih problemov in iskanja razlag do analize in predstavitve rezultatov ter razmišljanja o njihovih posledicah. Možnosti sodelovanja v znanosti, ki jih prinaša večja odprtost, z vključenimi prispevki niso izčrpane, nekatere so zgolj omenjene, kot na primer pristopi odprtega recenziranja in poskusi ponovitve (ang. *replication*) v člankih objavljenih rezultatov s ponovno analizo

¹⁹ Zaključki konference »Odprti raziskovalni podatki v Sloveniji« so dostopni na: https://www.adp.fdv.uni-lj.si/media/img/datoteke/Porocilo_Odprta_Znanost_Maribor_2019_v1.0.pdf. O CTS glej: <https://www.rd-alliance.org/groups/repository-audit-and-certification-dsa%E2%80%93partnership-wg.html>. O skupini FAIR pri RDA glej: <https://www.rd-alliance.org/groups/fair-data-maturity-model-wg>. Več o CTS in FAIR v prispevku M. Dolinar.

dostopnih podatkov. Vse to prinaša večjo preglednost nastajanja znanstvenih rezultatov ter prispeva k razumevanju in zaupanju javnosti. Želimo si tudi, da bi se z večjo dostopnostjo razširil krog vključenih v razprave o odprti znanosti. In ko govorimo o vrednotenju, si želimo, da bi bil priznan prispevek tako pri ustvarjanju novega znanja kot pri skrbi za podatke in druge rezultate znanosti, za njihovo dostopnost drugim, za katero pogosto poskrbijo spregledani izvajalci storitev za podporo odprti znanosti.

Uredništvu ČKZ se zahvaljujemo za priložnost, da smo lahko začeli v Sloveniji o odprti znanosti in podatkih širšo razpravo. Njen namen je po prevratu, ki ga povzroča v ustaljenih razmerjih *delanja znanosti*, odprto znanost udomačiti, da postane del običajne znanosti. Posebna zahvala gre sourednici posebne številke in odgovorni urednici ČKZ Danijeli Tamše ter kolegici Ani Inkret iz Arhiva družboslovnih podatkov za zavzeto in hkrati sproščeno sodelovanje pri pridobivanju in pregledovanju besedil ter pri vseh ostalih opravilih, ki so pripomogla pri uresničevanju zamisli. Zahvaljujemo se tudi številnim anonimnim recenzentom za hiter odziv in tehtne pripombe.

Literatura

- Biagioli, Mario (2016): Watch Out for Cheats in Citation Game. *Nature* 535(7611). Dostopno na DOI: 10.1038/535201a.
- Bishop, Libby (2019): New Data Types in Social Science Research and Data Archives. Predstavitev na delavnici *CESSDA Strengthening and Widening of the European Infrastructure of Social Science Data Archives*, 5. november. Dostopno na: https://www.cessda.eu/skopje2019/presentations/5.BISHOP_NewData_Skopje_05Nov2019_v2.pdf (3. marec 2021).
- Borgman, Christine L. (2012): The Conundrum of Sharing Research Data. *Journal of the American Society for Information Science and Technology* 63(6): 1059–1078. Dostopno na DOI: 10.1002/asi.22634.
- Deklaracija EOSC* (2017). Dostopno na: https://eosc-portal.eu/sites/default/files/eosc_declaration.pdf (4. marec 2021).
- Desai, Tanvi, Felix Ritchie in Richard Welpton (2016): The Five Safes: Designing Data Access for Research. *Working Papers in Economics no. 1601*, University of the West of England, Bristol. Dostopno na: <https://www2.uwe.ac.uk/faculties/bbs/Documents/1601.pdf> (4. marec 2021).
- Elliot, Mark, Elaine Mackey in Kieron O'Hara (2020): *The Anonymisation Decision-Making Framework: European Practitioners' Guide*. Dostopno na: <https://msrb-cel.files.wordpress.com/2020/11/adf-2nd-edition-1.pdf> (4. marec 2021).

- Informacijski pooblaščenec (2017): Anonimizacija, psevdonimizacija in uporaba osebnih podatkov v znanstvene in raziskovalne namene. *Mnenje – VOP*, 18. december. Dostopno na: https://www.ip-rs.si/varstvo-osebni-podatkov/iskalnik-po-odlocbah-in-mnenjih/odlocbe-in-mnenja-vop/?tx_jzvop-decisions_pi1%5BshowUid%5D=3027&tx_jzvop-decisions_pi1%5BhighlightWord%5D=psevdonimizacija (4. marec 2021).
- Littman, Justin (2019): Twitter's Developer Policies for Researchers, Archivists, and Librarians. *On Archivy*, 7. januar. Dostopno na: <https://medium.com/on-archivy/twitters-developer-policies-for-researchers-archivists-and-librarians-63e9ba0433b2> (4. marec 2021).
- Merton, Robert K. (1942): *Science and Technology in a Democratic Order*. (Prenaslovljeno v *The Normative Structure of Science*.) Dostopno na: <https://www.panarchy.org/merton/science.html> (10. februar 2021).
- OECD Legal Instruments (2021): *Recommendation of the Council concerning Access to Research Data from Public Funding*. Dostopno na: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347> (4. marec 2021).
- OpenAIRE (n. d.): *Can I Opt Out?* Dostopno na: <https://www.openaire.eu/can-i-opt-out> (6. februar 2021).
- RDA COVID-19 Working Group (2020): Recommendations and Guidelines on Data Sharing. *Research Data Alliance*, 30. junij. Dostopno na DOI: 10.15497/rda00052.
- Ritchie, Felix (2016): *Can a Change in Attitudes Improve Effective Access to Administrative Data for Research?* Dostopno na: <http://www2.uwe.ac.uk/faculties/BBS/Documents/1601.pdf> (4. marec 2021).
- Science Europe (2021): *Practical Guide to the International Alignment of Research Data Management*. Dostopno na: https://www.scienceurope.org/media/4br-kxxe5/se_rdm_practical_guide_extended_final.pdf (4. marec 2021).
- Scola, Nancy (2018): Facebook's Next Project: American Inequality. *Politico*, 19. februar. Dostopno na: <https://www.politico.com/story/2018/02/19/facebook-inequality-standford-417093> (4. marec 2021).
- Sveinsdottir, Thordis, Vanessa Proudman in Joy Davidson (2020): *An Analysis of Open Science Policies in Europe, v6 (Version 6)*. Dostopno na DOI: 10.5281/zenodo.4005612.
- Štebe, Janez, Sonja Bezjak in Maja Dolinar (2020): *Smernice za oblikovanje politik znanstvenih založb glede navajanja raziskovalnih podatkov v znanstvenih publikacijah in zagotavljanja dostopa do primarnih podatkov, uporabljenih v člankih*. Dostopno na DOI: 10.5281/zenodo.37572875.
- Štebe, Janez, Maja Dolinar, Sonja Bezjak in Ana Inkret (2020): Implementing the RDA Research Data Policy Framework in Slovenian Scientific Journals. *Data Science Journal* 19: XX: 1–15. Dostopno na DOI: 10.5334/dsj-2020-052.

- UNESCO (2021): *UNESCO Recommendation on Open Science*. Dostopno na: <https://en.unesco.org/science-sustainable-future/open-science/recommendation> (4. marec 2021).
- Van Noorden, Richard (2021): Scientists Call for Fully Open Sharing of Coronavirus Genome Data. *Nature* 590: 195–196. Dostopno na DOI: 10.1038/d41586-021-00305-7.
- Vlada Republike Slovenije (2015): *Nacionalna strategija odprtega dostopa do znanstvenih objav in raziskovalnih podatkov v Sloveniji 2015–2020*. Dostopno na: https://www.gov.si/assets/ministrstva/MIZS/Dokumenti/ZNANOST/Strategije/Nacionalna_strategija_odprtega_dostopa.pdf (4. marec 2021).
- Vlada Republike Slovenije (2017): *Akcijski načrt izvedbe Nacionalne strategije odprtega dostopa do znanstvenih objav in raziskovalnih podatkov v Sloveniji 2015–2020*. Dostopno na: http://mizs.arhiv-spletisc.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Odprti_dostop/Akcijski_nacrt_-_POTRJENA_VERZIJJA.pdf (4. marec 2021).