

To je recenzirani rokopis članka, objavljenega v Časopisu za kritiko znanosti, 2021, letnik 49, št. 282, str. 101–119.



Sonja Bezjak

Znanstveno objavljanje raziskovalnih podatkov v odprti znanosti

Abstract

Publishing Research Data in Open Science

Data is often one of the main research results underpinning scientific knowledge. Researchers report their results in publications, thereby indirectly building their reputation and career. Even though research data is a fundamental element of a research process, the scientific community pays little attention to data as an independent scientific result. In order to shed light on the importance of this question, the article presents the concept of data publication, the importance of acknowledging authorship and citing data, and the role of the scientific evaluation of databases as independent scientific results. We outline the landscape of stakeholders who (should) participate in the quality publication of research data in the domestic and international space, and especially the role of scientific journals in publishing data and data-based scientific publications. The purpose of the article is to encourage actors in the Slovenian research community to be more committed to the transparent and timely publication of research data, which is a key building block of open science.

Keywords: data publication, data article, data citation, data services

Sonja Bezjak holds a PhD in Sociology and is currently working at the Slovenian Social Science Data Archives, where she is primarily focused on the subject of open science principles. This article combines her professional and personal interests, as she is actively engaged in social and environmental justice initiatives (sonja.bezjak@fdv.uni-lj.si).

Povzetek

Podatki so pogosto eden temeljnih raziskovalnih rezultatov, ki je podlaga za znanstvena spoznanja. Raziskovalci o njih poročajo v publikacijah, s tem pa posredno gradijo svoj ugled in kariero. Četudi gre za temeljne prvine v raziskovalnem procesu, znanstvena skupnost podatkom kot samostojnemu znanstvenemu rezultatu posveča le malo pozornosti. Da bi osvetlili pomen tega vprašanja, v preglednem prispevku predstavimo koncept podatkovne publikacije, pomen priznanja avtorstva in citiranja podatkov ter vlogo znanstvenega vrednotenja podatkovnih zbirk kot samostojnih znanstvenih rezultatov. Orišemo krajino deležnikov, ki (naj) sodelujejo pri kakovostnem objavljanju raziskovalnih podatkov v domačem in mednarodnem prostoru, ter še posebej vlogo znanstvenih revij pri objavljanju podatkov in znanstvenih publikacij, temelječih na podatkih. Namen članka je akterje slovenske raziskovalne skupnosti spodbuditi k večji zavzetosti za transparentno in pravočasno objavljanje raziskovalnih podatkov, ki so pomemben gradnik odprte znanosti.

Ključne besede: podatkovna objava, podatkovni članek, citiranje podatkov, podatkovne storitve

Sonja Bezjak je doktorica sociologije, zaposlena v Arhivu družboslovnih podatkov, kjer se še posebej zanima za načela odprte znanosti. V tem članku se je njeno poklicno zanimanje združilo s pristočnimi aktivnostmi, pri katerih si angažirano prizadeva za uresničevanje socialne in okoljske pravičnosti (sonja.bezjak@fdv.uni-lj.si).

Uvod

Razvoj informacijskih tehnologij, vzpon svetovnega spleta in velike količine podatkov prinašajo spremembe ne le v vsakdanjem življenju, o čemer pišemo v Bežjak in Masten (2021), pač pa tudi na področju znanosti. Znanstveniki z različnih področij vse več spoznanj utemeljujejo na velikih količinah podatkov, nekateri govorijo o nastanku »podatkovno intenzivne« znanstvene paradigme (Gray, 2009: xxx; Elger idr., 2015: 119). To pa s seboj prinaša številne tehnološke, pravne, etične in druge strokovne izzive glede možnosti izmenjave podatkov med raziskovalci; ti individualno ali skupinsko delajo na podatkih, ki so jih zbrali sami ali drugi raziskovalci.¹ Pri tem velja razjasniti, na kaj mislimo, ko govorimo o podatkih, saj se ti med seboj razlikujejo po metodi, s katero so bili zbrani, formatu, v katerem so predstavljeni, vrsti, nedinamičnosti in drugih značilnostih, kar med deležnike v življenjskem krogu podatkov nemalokrat vnaša zmedo.

Koncept »odprti podatki« spodbuja izmenjavo oziroma deljenje podatkov tako, da lahko podatke načeloma kdor koli uporabi za kar koli (Open Knowledge International). V arhivih podatkov zastopamo stališče, da je treba podatke kolikor mogoče hitro narediti dostopne za čim več zainteresiranih skupnosti. Pri tem moramo upoštevati načelo, da naj bodo podatki odprti, kolikor je mogoče, in zaprti, kolikor je nujno. Ni nujno oziroma obvezno, da so podatki prosto dostopni ali povsem brez pravnih omejitev,² toda kadar je dostop omejen, morajo biti pogoji jasni in nediskriminatorni. Ni sprejemljivo, da bi bilo pisanje prošnje avtorjem raziskave za dostop do podatkov del običajnega postopka, prav tako morajo biti jasno urejene pravice intelektualne lastnine (Kratz in Strasser, 2014: 4; Shin idr., 2019: 703). Za tem se lahko namreč skriva neetična praksa trgovanja s podatki, na primer ko za dostop do podatkov kdo zahteva soavtorstvo pri morebitnem članku. Obenem mora biti zagotovljena zahteva po priznanju avtorstva, to je citiranje podatkov kot povračilo ustvarjalcu podatkov za vložen trud. S promocijo kulture odprtega dostopa igramo arhivi podatkov aktivno in angažirano vlogo pri spreminjanju praks izmenjave podatkov ter odpravljanju kulturnih in tehničnih ovir, ki stojijo na poti do odprtih podatkov.

Kultura deljenja podatkov se med področji močno razlikuje. Na področju proučevanja biodiverzitete je že dlje časa jasno, da je za napredek v spoznanjih nujen dostop do velikih

¹ Nekatera področja, ki delajo na primer s prostorskimi, astronomskimi in meteorološkimi podatki, tudi družboslovni podatki z arhivi, katerih začetki segajo v 60. leta 20. stoletja, imajo torej daljšo tradicijo in so se z navedenimi izzivi soočali že prej.

² Zaprtost podatkov se nanaša na zaščito pred razkritjem identitete udeležencev raziskav, dragocenega najdišča ali državne in poslovne skrivnosti.

količin podatkov iz različnih virov, saj je le tako mogoče razumeti kompleksnost ekoloških sistemov na veliki prostorski in časovni skali. Posledično so razvili orodja in platforme za množično deljenje in rabo podatkov, s tem pa omogočili sodelovanje širokega nabora deležnikov pri izmenjavi in uporabi podatkov (Kelling idr., 2009: 613–619). Prakse deljenja pa niso povsod in zmeraj samoumevne. V zadnjem desetletju je bilo v tujini opravljenih več raziskav o tem, zakaj raziskovalci podatkov ne delijo (Tenopir idr., 2011, 2015; Borgman, 2013; Arzberger idr., 2004; Piwowar, 2011), o razlogih pa smo v okviru projekta Odprti podatki spraševali tudi raziskovalce v Sloveniji (Štebe idr., 2013). Raziskovalci so pogosto zadržani, ker se jim ne povrne napor, ki so ga vložili v zbiranje, obdelavo in opis podatkov. Ovira jih tudi strah, da bi lahko drugi podatke napačno razumeli, jih zlorabili ali z njimi povzročili škodo (Elger idr., 2015: 119). Razmere, v katerih deljenje podatkov še ni rutinsko in samoumevno, so hkrati priložnost za neetične prakse, kot na primer manipulacije pri dostopu in nove oblike podatkovnega kolonializma (Chase idr., 2020: 58).³ Standardizirane prakse in sprejeta načela⁴ deljenja podatkov so lahko ključnega pomena pri premoščanju ovir, kot je potreba po priznanju avtorstva, to je povračilo avtorju podatkov za vloženi napor, zaščita pred zlorabami oziroma goljufijami, kot na primer ponarejanje podatkov, nenamernimi napakami, na primer zaradi nepoznavanja izvornih metod zbiranja podatkov, in pri omejevanju dostopa do podatkov določenim zainteresiranim skupinam ali posameznikom (Shin idr., 2019: 701–702; Kratz in Strasser, 2014: 4).

V Arhivu družboslovnih podatkov (v nadaljevanju ADP), kjer vsakodnevno sodelujemo z raziskovalci pri predaji oziroma prevzemu raziskav s podatki, ugotavljamo, da raziskovalci slabo poznajo podatkovne storitve, ki so na voljo v domačem in mednarodnem okolju. Trije raziskovalci so kot razlog, zakaj podatkov na temo Covid-19 ne bodo objavili v ADP, navedli, da bodo podatke raje objavili pri reviji Open Access Data Journal, saj želijo večjo mednarodno prepoznavnost.⁵ Pri tem so spregledali, da revija objavi samo opis podatkov, ne pa samih podatkov.⁶ Tako se vedno znova potrjuje, da je osveščanje, izobraževanje in usposabljanje o podatkovnih produktih in storitvah nujen del arhivskega vsakdanjika, saj se z novimi storitvami na mednarodni ravni tudi doma širi skupnost uporabnikov. Še posebej tiste, ki se z arhiviranjem in objavljanjem podatkov srečujejo prvič, je treba o obstoju in uporabi storitev v nacionalnem okolju poučiti in jih za to delo usposobiti. Namen članka je tako prispevati k večji ozaveščenosti

³ Več o tem v tej številki ČKZ v članku Sonje Bezjak in Sergeje Masten.

⁴ Načela FAIR (Wilkinson idr., 2016), načela CARE (skupina RDA, 2019), načela TRUST (Lin idr., 2020).

⁵ Vir: osebna korespondenca avtorice z raziskovalci, september 2020.

⁶ Glej Navodila za avtorje (*Instructions for Authors*), dostopna na: <https://www.mdpi.com/journal/data> (10. november 2020).

o podpori, ki je raziskovalcem na voljo v krajini podatkovne infrastrukture in podatkovnih storitev. S pregledom aktualne literature in na podlagi potreb, identificiranih v okviru sodelovanja z raziskovalci v fazi odločitve za arhiviranje in predajo podatkov v ADP, želi avtorica opredeliti temeljne pojme in koncepte – vse s ciljem akterje slovenske raziskovalne skupnosti, tako raziskovalce kot založbe znanstvenih revij, spodbuditi k večji zavzetosti za transparentno in pravočasno objavljanje raziskovalnih podatkov, ki so pomemben gradnik odprte znanosti in podpornik hitrejšega napredka v znanstvenih spoznanjih. Naš cilj je promovirati nenehen zajem, arhiviranje in objavljanje podatkov ter njihovo vrednotenje (Shin idr., 2019: 702). V tem članku bomo uporabili definicijo, ki jo uporabljamo v ADP: »Raziskovalni podatki predstavljajo osnovno podlago za znanstveno raziskovanje in z analizo omogočajo izpeljavo teoretično ali uporabno naravnanih zaključkov« (Štebe idr., 2015: 1).

Od deljenja podatkov k znanstveni objavi podatkov

V zadnjih letih se namesto zelo splošnega termina deljenje podatkov, ki vključuje tudi izmenjavo oziroma posredovanje podatkov med raziskovalci brez vnaprej določenih pravil in v skupnosti sprejetih načel, na primer po elektronski pošti, na USB ključu ali v oblaku, vse bolj uveljavlja pojem objava podatkov. K temu jezikovnemu premiku je prispevala ideja, da je treba podatke oziroma podatkovne zbirke prepoznati kot samostojen znanstveni rezultat, podoben znanstvenemu članku, ki dobi priznanje šele z objavo. Kdaj lahko rečemo, da smo podatke objavili? V času spletnih tehnologij lahko podatke in povezano gradivo mimogrede objavimo kjerkoli na spletu. Ker smo podatke ponudili javnosti, lahko rečemo, da smo jih objavili (Kratz in Strasser, 2014: 4). Ob tem pa je treba razjasniti slabosti takšne objave, ki bi se zgodila »kjerkoli že«. Problematičen je predvsem vidik dolgotrajne hrambe, ko ni garancije, da bodo podatki tam tudi ostali, na primer raziskovalec se preseli na drugo ustanovo, povezava do podatkov izgine ali projektna stran ugasne, ker ni več plačnika domene. Vprašljivi pa so tudi kakovost dokumentacije, omejen vpliv takšne objave, odsotnost ocene kakovosti raziskave in priznanja znanstvene vrednosti. V znanstveni skupnosti tako obstaja konsenz o treh temeljnih značilnostih, ki opredeljujejo podatkovno objavo. Prvič, objavljeni podatki so javno dostopni zdaj in za nedoločeno prihodnost, avtor pa ima zagotovilo, da bodo tam tudi ostali.⁷ Četudi je za dostop morebiti treba plačati ali podpisati pogodbo, ta ne sme biti odvisen od individualnih kapric avtorjev. Drugič, objavljeni podatki morajo biti ustrezno dokumentirani, da jih je

⁷ Kadar iz upravičenih razlogov podatkov ni mogoče objaviti, lahko avtorji objavijo samo metapodatke in s tem zadostijo značilnostim podatkovne objave. Pri tem je pomembno, da predstavijo razloge za omejitve dostopa in morebitne posebne pogoje dostopa.

mogoče ponovno ustvariti ali ponovno uporabiti. Tretjič, podobno kot velja za knjigo ali članek, mora biti na voljo dovolj bibliografskih informacij, da lahko podatkovno objavo citiramo, s čimer avtorjem priznamo vložen napor (Kratz in Strasser, 2014: 3–4). Da bi podatke sploh lahko uporabljali zunaj lastne raziskovalne skupine, od raziskovalcev zahteva nekaj napora, saj je treba zagotoviti opisne informacije (to so metapodatki), ki podatke opišejo, razložijo in omogočajo njihovo preverbo (Royal Society London, 2012) (Elger idr., 2015: 121). Da bi podatki oziroma podatkovna zbirka dobila status znanstvene objave, je treba speljati podoben proces kot za znanstveni članek (Lawrence idr., 2011: 7; Callaghan idr., 2012: 108–109).⁸ Kako torej opisati podatke in kje jih objaviti, da bi dobili status samostojne znanstvene objave?

Krajina deležnikov oziroma kako opisati podatke in kje jih objaviti

Podatke je mogoče opisati in povezati z znanstvenimi deli na različne načine. Gre za različne *formate* oziroma oblike podatkovnih objav. Opisi podatkov so lahko bolj ali manj strukturirani. Podatki so lahko opisani kar znotraj znanstvenega članka, v katerem so analizirani. Avtorji jih lahko opišejo v t. i. podatkovnem članku, ki ga objavijo v t. i. podatkovni reviji. Podatkovni repozitoriji nudijo opis v obliki standardnega metapodatkovnega zapisa, lahko pa so opisani tudi v obliki t. i. podatkovnega poročila (Kratz in Strasser, 2014: 4; Elger idr., 2015: 121). Različni formati podatkovnih objav imajo različne prednosti in slabosti, o katerih velja razmisliti, preden se raziskovalec odloči za objavo opisa podatkov in mesta objave samih podatkov.

Podatki kot spremno gradivo znanstvenemu članku

Po raziskavi Wiley (2014) so podatki najpogosteje objavljeni kot spremno gradivo znanstvenemu članku.⁹ Taki podatki so običajno omejenega obsega, namen njihove objave pa je omogočiti ponovitev analize oziroma ugotovitev iz članka. Poleg tega, da avtorji ne objavijo

⁸ V tem članku se osredotočamo na koncept znanstvene objave, velja pa omeniti, da so v znanstveni skupnosti uveljavljeni tudi drugi koncepti oziroma metafore, kakor jih imenujeta Parsons in Fox, na primer »Big Iron«, ki se pogosteje pojavi v inženirskem okolju v povezavi z velikimi ustvarjalci podatkov, kot je NASA, in »Science Support«, ki se nanaša na znanstveno podporo v obliki operacijske podporne strukture, ki jo tipično najdemo v laboratorijih in raziskovalnih postajah, vendar se lahko v različnih okoljih zelo razlikuje, »Map Making«, ki se uporablja na področju prostorskih podatkovnih infrastruktur in povezanih geografskih informacijskih sistemov, kjer je večji poudarek na geografski vizualizaciji in manj na ohranjanju podatkov, ter »Linked Data«, kjer je v ospredju mreža podatkov, ki so med seboj interoperabilni in omogočajo interdisciplinarnost (Parsons in Fox, 2013).

⁹ Raziskava o globalnih trendih deljenja podatkov iz leta 2014 je pokazala, da je 67 % sodelujočih raziskovalcev podatke delilo v obliki spremnega gradiva, 37 % na osebni, institucionalni ali projektni spletni strani, 26 % v institucionalnem repozitoriju, 19 % v področnem repozitoriju, 6 % v splošnem repozitoriju in 5 % drugje (Wiley, 2014).

vseh podatkov, ki so jih zbrali v raziskavi, tudi dokumentacija ni nujno pripravljena tako, da bi bilo mogoče podatke uporabiti neodvisno od vsebine članka (Kratz in Strasser, 2014: 5). Če znanstveni članek ni objavljen v odprtodostopni reviji, tudi podatkov, ki so njegovo spremno gradivo, ne moremo razumeti in ponovno uporabiti, zato priporočamo objavo podatkov zunaj revije, v odprtodostopnih podatkovnih repozitorijih (Elger idr., 2015: 120). S povečevanjem števila podatkovnih repozitorijev, ki imajo na voljo strokovnjake za dolgotrajno hrambo in dostop, revije te naloge vedno pogosteje prepuščajo njim.¹⁰ Raziskave kažejo, da tudi danes le okoli 25 % raziskovalcev podatke objavi v repozitorijih (Hrynaszkiewicz, 2019: 6).

Podatkovni članek, objavljen v podatkovni reviji

Poudariti velja, da podatkovne revije največkrat sploh ne objavljajo podatkov, ampak se v njih objavijo samo opisi podatkovnih zbirk. Podatkovne revije objavljajo članke o izvornih raziskovalnih podatkih, v katerih avtorji natančno opišejo logiko podatkov oziroma podatkovne zbirke in metode zbiranja. V opisu je predstavljena celotna podatkovna zbirka, vključno z nestandardnimi podatki, odkloni in redkimi dogodki, ki postanejo zanimivi šele, če imamo na voljo obsežne zbirke podatkov. Sami podatki pa so objavljeni v kakšnem repozitoriju. Prednost podatkovnega članka je v zagotavljanju bogate dokumentacije, ki je še posebej primerna za množico manjših in heterogenih (*long-tail*) raziskovalnih podatkov, za katere ni mogoče zagotoviti ozko standardiziranega formata opisa podatkov.¹¹ Podatkovne revije se lahko razlikujejo po dolžini in strukturi podatkovnega članka, večinoma pa težijo h kratkim, strnjanim formatom. Na področju družboslovja in humanistike se uveljavlja podatkovna revija *Research Data Journal for the Humanities and Social Sciences*, ki so jo zasnovali v nizozemskem arhivu podatkov.¹² Sprejemajo podatkovne članke v obsegu največ 2.500 besed, ki naj vsebujejo opis podatkov, raziskovalni kontekst in prispevek k akademski razpravi. Od avtorjev zahtevajo, da pred objavo članka podatke shranijo v zaupanja vrednem digitalnem arhivu oziroma repozitoriju, ki zagotovi stalni identifikator (kot na primer DOI, HANDLE ali URN).¹³ Primer podatkovne revije s področja biologije je *Scientific Data*, ki jo izdajajo pri *Nature*; od avtorjev med drugim zahtevajo, da podatke shranijo v uveljavljenem področnem repozitoriju, če ta ni na

¹⁰ Pri reviji *The Journal of Neuroscience* so prenehali objavljati podatke kot spremno gradivo leta 2010 ter to nalogo zaupali repozitorijem, in sicer z argumentom, da gre za ustrežnejši način diseminacije podatkov (Kratz in Strasser, 2014: 5).

¹¹ Pojem mali podatki v tem članku razumemo kot majhne sklope podatkov z velikega števila različnih raziskovalnih področij.

¹² *DANS – Data Archiving and Networked Services*. Dostopno na: <https://dans.knaw.nl/en> (10. november 2020).

¹³ Glej <https://brill.com/view/journals/rdj/rdj-overview.xml?rskey=Y9aoVu&result=1&contents=toc-34057> (10. november 2020).

voljo, pa v splošnem repozitoriju.¹⁴ Podatkovnim revijam je skupno, da zahtevajo povzetek ter opis metode zbiranja in podatkovne zbirke; nekatere predlagajo, naj opišejo tudi možno uporabo podatkov, druge imajo dodatne področno specifične zahteve, vendar pa ne vključujejo analize in raziskovalnih ugotovitev. Podatki, predstavljeni v podatkovnem članku, so običajno objavljeni v podatkovnem repozitoriju in so recenzirani skupaj z znanstvenim člankom. Podatkovne revije se hitro razvijajo na vseh znanstvenih področjih, nekatere so že indeksirane v Web of Science (na primer Earth System Science Data – ESSD) (Kratz in Strasser, 2014: 5; Elger idr., 2015: 122; Shin idr., 2019: 701).

Objava podatkov v podatkovnem repozitoriju

Podatkovni repozitorij je digitalni arhiv, ki zbira, ohranja in objavlja podatkovne zbirke, povezano dokumentacijo in metapodatke. [...] Da bi bili podatki dostopni na dolgi rok, sta potrebna znanje in storitve konvertiranja podatkov v druge formate in dodajanje drugih funkcionalnosti, s katerimi podatki dobijo dodano vrednost (OpenAIRE, 2017: 7).

Objava podatkov je v tem primeru dokumentirana pri arhivu oziroma repozitoriju kot neodvisno oziroma samostojno gradivo in opremljena z opisom v obliki strukturnih metapodatkov. V takih primerih je repozitorij oziroma arhiv v vlogi založnika.

Obdelava podatkov v arhivih podatkov poteka fazno po načelih standardnega modela OAIS.¹⁵ Arhiv oziroma repozitorij prevzame podatke od avtorja (SIP), jih arhivsko obdela (AIP) ter preda v uporabo skupnosti (DIP).¹⁶ Ko raziskovalci podatke arhivirajo, so gotovi, da bodo do njih dostopali na točno določenem mestu in da bodo lahko tudi v prihodnje »brali« vsebino podatkovne zbirke. Pri tem je ključno, da arhiviranje ni enkratno dejanje, pač pa kontinuiran proces, ki obsega skrb za dolgoročno ohranjanje raziskovalnih podatkov. Arhivirani podatki, ki so varno hranjeni, v primernem formatu in ustrezno dokumentirani, so lahko na zahtevo oziroma po potrebi dostopni za verifikacijo oziroma preverbo, na primer recenzentom. Da so podatki ponovno uporabni še za druge namene s strani tretjih, ki niso sodelovali niti pri zbiranju niti pri analizi, jih je treba objaviti. Z objavo v arhivu ali repozitoriju raziskovalci podatke naredijo najdljive, dostopne, razumljive ter ponovno uporabne zase, za svojo raziskovalno

¹⁴ Glej Recommended Data Repositories, dostopno na: <https://www.nature.com/sdata/policies/repositories> (10. november 2020).

¹⁵ Open Archival Information System, glej <http://www.oais.info/>.

¹⁶ SIP je *Submission Information Package*, AIP je *Archival Information Package*, DIP je *Dissemination Information Package*. Glej <https://public.ccsds.org/pubs/650x0m2.pdf> (10. november 2020).

skupino in druge zainteresirane skupnosti (CESSDA Training Team, 2017–2020).¹⁷ Da so opisi podatkov primerljivi in povezljivi s podatki v drugih arhivih oziroma repozitorijih, se uporabljajo uveljavljeni *metapodatkovni standardi*; v ADP in drugih arhivih Konzorcija arhivov družboslovnih podatkov je v uporabi t. i. standard DDI.¹⁸

Podatkovno poročilo

Elger idr. opozorijo na še en tip podatkovne objave, to je *podatkovno poročilo*. Podatkovno poročilo je uporabno, kadar je na voljo premalo metapodatkovnih polj, da bi podatke zadostno opisali. Poročilo nadomešča ponekod že uveljavljeni »read me« dokument, ki ga je treba običajno pretočiti na računalnik, da bi ga lahko preverili. Podatkovno poročilo je še posebej primerno za podatke, ki niso tako obsežni oziroma kompleksni, da bi jih opisali v podatkovnem članku. Na primer: German Research Centre for Geosciences (GFZ) objavlja t. i. znanstvena tehnična poročila, ang. *Scientific Technical Reports – STR*). V njem navadno najdemo opise podatkov, zbrane v okviru doktorskih disertacij, ali projektna poročila. V enem poročilu je mogoče opisati več podatkovnih zbirk in vse relevantne vidike podatkovne zbirke (Elger idr., 2015: 123).

Podatkovni repozitoriji specialisti za objavo podatkov

Krajina repozitorijev je raznolika, razlikujejo se po tem, kateri skupnosti služijo in katere vrste podatkov sprejemajo ter po tipu ekspertize. Obstajajo institucionalni, področni, multidisciplinarni, projektno usmerjeni in drugi repozitoriji. Njihove prednosti in slabosti je dobro poznati, da bi se osveščeno odločili, katere podatke kje objaviti.

Področni podatkovni repozitoriji imajo strokovnjake, specializirane za določene vrste podatkov, določene teme ali določena področja.¹⁹ Poleg tega imajo specializirane kataloge in so povezani z drugimi repozitoriji. Splošno pravilo je, da sprejemajo podatke višje kakovosti, ki imajo potencial za nadaljnjo in čim širšo rabo. V repozitorijih opravijo tehnični pregled gradiva, s katerim zagotovijo, da so podatki celoviti, da je opis celovit in da sta opisa med seboj skladna. V znanstvenem pregledu, ki je neke vrste vsebinska recenzija, ocenijo metode zbiranja podatkov, verodostojnost podatkov in potencial ponovne rabe. Ker je znanstveni pregled težje

¹⁷ Za več glej poglavje *Archive & Publish*. Dostopno na: <https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/6.-Archive-Publish/Towards-archiving-publication> (10. november 2020).

¹⁸ Glej Data Documentation Initiative, dostopno na <https://ddialliance.org/>.

¹⁹ Na voljo so tudi zelo specializirane baze, na primer *WormBase in FlyBase*, ter baze, ki zbirajo podatke določenega tipa, na primer *Gene Expression Omnibus (GEO)* in seizmološki podatki v *SeismicPortal* (Kratz in Strasser, 2014: 6).

zagotoviti, ga ne omogočajo vsi področni repozitoriji (Callaghan idr., 2012; Kratz in Strasser, 2014: 7). Področni repozitoriji so običajno dobro poznani v raziskovalni skupnosti, ki so ji namenjeni, so standardizirani in prijazni za rabo. Slabost je, kadar so zaradi specifik repozitorijev podatki enega projekta razpršeni med več repozitorijev.²⁰ Za raziskovalce pa je pomembna tudi informacija, ali je repozitorij pridobil certifikat zaupanja vrednega. Tak certifikat je na primer CoreTrustSeal,²¹ ki zagotavlja, da je ustanova usposobljena in opremljena za dolgotrajno arhiviranje podatkov in zagotavljanje dostopa do njih. V Sloveniji sta certifikat pridobila Arhiv družboslovnih podatkov²² in CLARIN Slovenija.²³

Institucionalni repozitoriji služijo skupnostim svoje ustanove, na primer raziskovalcem določene univerze.²⁴ Univerza v Oxfordu je ustanovila Research Data Oxford, ki med drugim ponuja arhivsko storitev ORA-Data za podatke vseh disciplin. Njen namen ni hraniti vseh podatkov, ki jih zberejo raziskovalci z Oxforda, ampak predvsem tiste, ki jih je treba hraniti zaradi politik financiranja ali založnika, na primer so podlaga znanstvenemu članku, vendar zanje ni ustreznega področnega arhiva.²⁵

Splošne repozitorije priporočamo, kadar objava podatkov ni možna niti v področnem niti v institucionalnem repozitoriju. Splošni repozitoriji, na primer Zenodo,²⁶ Mendeley Data,²⁷ FigShare²⁸ in Harvard Data Verse,²⁹ objavljajo podatke katerega koli raziskovalca s katerega koli področja. Njihova storitev je prilagojena heterogenim in malim (ang. *long-tail*) podatkom, ki ne sodijo v katerega od področnih repozitorijev. Storitve obsega hrambo, deljenje in registracijo podatkov, vendar ni nujno, da je zagotovljena dolgoročna hramba. Prav tako ne zagotavljajo tehničnega in znanstvenega pregleda podatkov in dokumentacije.

Pri odločitvi o tem, kje objaviti podatke, raziskovalci upoštevajo politike financiranja in založnikov. Nekatera vidnejša znanstvena združenja in financiranja priporočajo, naj bodo objave

²⁰ Na primer na Irskem kvalitativne podatke s področja družboslovja raziskovalci objavijo v Irskem arhivu kvalitativnih podatkov (IQDA – The Irish Qualitative Data Archive, dostopno na <https://www.maynoothuniversity.ie/iqda>). Če so v sklopu raziskave nastali tudi kvantitativni podatki, jih objavijo v Irskem arhivu družboslovnih podatkov (ISSDA) (<https://www.ucd.ie/issda/>). Slovenski Arhiv družboslovnih podatkov sprejema tako kvantitativne kot kvalitativne podatke.

²¹ Dostopno na: <https://www.coretrustseal.org/why-certification/> (10. november 2020).

²² Dostopno na: <https://www.adp.fdv.uni-lj.si/> (10. november 2020).

²³ Dostopno na: <https://www.clarin.si/info/o-repozitoriju/> (10. november 2020).

²⁴ Pri nas je v tej funkciji repozitorij Univerze v Ljubljani (dostopen na <https://repozitorij.uni-lj.si/>), ki pa shranjuje v glavnem nepodatkovne publikacije zaposlenih in študentov univerze ter je zatočišče za originalne podatke le v primerih, ko nimajo primernejšega mesta hrambe v področnem ali tematskem repozitoriju.

²⁵ Glej Research Data Oxford, dostopno na: <https://researchdata.ox.ac.uk/preserving-your-data/archives-and-other-options/> (10. november 2020).

²⁶ Dostopno na: <https://zenodo.org/> (10. november 2020).

²⁷ Dostopno na: <https://data.mendeley.com/> (10. november 2020).

²⁸ Dostopno na: <https://figshare.com/> (10. november 2020).

²⁹ Dostopno na: <https://dataverse.harvard.edu/> (10. november 2020).

v področnih repozitorijih za raziskovalce prva izbira. Evropska komisija v novem finančnem sklicu Obzorje Evropa od raziskovalcev zahteva, da čim prej in skladno z roki, ki so si jih raziskovalci postavili v načrtu ravnanja z raziskovalnimi podatki, shranijo podatke v zaupanja vrednem repozitoriju, upoštevajoč načelo: »odprti, kolikor je mogoče, in zaprti, kolikor je nujno«. Prav tako morajo v repozitoriju objaviti vsa gradiva, potrebna za presojo uporabnosti in razumevanje podatkov (Dechamp, 2020: 10).³⁰

Vloga znanstvenih revij pri zagotavljanju dostopa do podatkov in vrednotenju podatkov

Čeprav znanstvene revije na določenih področjih že več kot dve desetletji od avtorjev zahtevajo, da objavijo svoje podatke, šele v zadnjem obdobju narašča število revij in založnikov z različnih področij, ki uvajajo zahtevo, da morajo avtorji člankov zagotoviti informacijo o tem, kje je mogoče dostopati do podatkov, uporabljenih v člankih (Colavizza idr., 2020: 2). Da lahko revije z uredniškimi politikami pri objavi podatkov pomembno dopolnijo zahteve financerjev in prispevajo k spremembi kulturnih praks (Borgman idr., 2018), priča raziskava o razpoložljivosti podatkov, na katerih temeljijo članki. Od 516 člankov, objavljenih v obdobju 1991–2011, so lahko do podatkov dostopali le v 101 članku. Skupno so dostopali do podatkov iz 19,5 % člankov, za članke, objavljene pred letom 2000, pa le iz 11 % člankov. Ugotovili so, da verjetnost, da bo mogoče dostopati do podatkov, letno pade za 17 %. Glavni razlog, da podatki iz starejših člankov niso dostopni, je, da so jih avtorji izgubili ali da ne morejo več dostopati do nosilca, na katerem jih hranijo. Zaskrbljujoča je tudi ugotovitev, da verjetnost, da bodo e-naslovi prvega, zadnjega ali odgovornega avtorja delovali, letno pade za 7 %. Sklepna ugotovitev poročila je, da so avtorji dolgoročno slabi skrbniki raziskav (Vines idr., 2014). Colavizza idr. (2020: 7) so proučili revije založnikov PLOS in BMC.³¹ V njihovih založniških politikah so identificirali štiri kategorije,³² med katerimi avtorji izberejo eno in jo navedejo v izjavi o razpoložljivosti podatkov, ki obvezno dopolnjuje članek.³³ Primerjava za leti 2017 in 2018 je pokazala bistvene razlike med založnikoma v deležu kategorij. V opazovanem obdobju

³⁰ OpenAire (katerega članica je poleg drugih 65 evropskih univerz tudi UL) priporoča, naj, če je na voljo, izberemo področno specifičen podatkovni repozitorij, ki sledi podatkovnim standardom z našega področja. Glej <https://www.openaire.eu/opendatapilot-repository-guide> (10. november 2020).

Založniki, združeni v Coalition on Publishing Data in the Earth and Space Sciences, pa so v skupni izjavi *Statement of Commitment* priporočili, naj bodo zbirke raziskovalnih podatkov prednostno hranjene v ustreznih področnih repozitorijih (COPDESS, 2015).

³¹ Za seznam revij glej https://www.biomedcentral.com/p/the-bmc-series-journals?gclid=Cj0KCQiA5bz-BRD-ARIsABjT4njoWVOV7JKnxajpB_rVwY2HIBJICRE6wZKqD1VhdBCEPh_xmymdrJEaAu2DEALw_wcB za revije BMC in <https://plos.org/> za revije PLOS (10. november 2020).

³² Gre za naslednje možnosti: 0 – podatki niso na voljo; 1 – podatki so na voljo na zahtevo; 2 – podatki so na voljo v članku ali kot spremno gradivo; 3 – podatki so na voljo v repozitoriju.

³³ V angleščini *Data Availability Statement* (DAS).

se je pri založbi BMC 60 % člankov uvrstilo v kategorijo 1 – podatke je mogoče dobiti na zahtevo, 19,2 % v kategorijo 2 – podatki so na voljo v članku ali kot spremno gradivo – in 12,2 % v kategorijo 3 – podatki so objavljeni v repozitoriju. V istem obdobju se je 65,2 % člankov pri založbi PLOS uvrstilo v kategorijo 2 – podatki so na voljo v članku ali kot spremno gradivo, 14 % v kategorijo 1 – podatki so na voljo na zahtevo – in 20,8 % v kategorijo 3 – podatki so objavljeni v repozitoriju. Ali raziskovalci podatke delijo oziroma kje jih objavijo, je povezano s tem, ali v politikah zgolj priporočajo ali zahtevajo deljenje podatkov in med katerimi možnostmi lahko raziskovalci izbirajo.³⁴ Številne revije s področja družboslovja že imajo politike rokovanja s podatki in navodila za avtorje (78 % na področju sociologije in 44 % na področju političnih ved), pa vendar so avtorji le za okoli polovico člankov zagotavljali, da so podatki na voljo, le v 37 % primerih pa je dejansko bilo mogoče dostopati do podatkov (Zenk-Möltgen idr., 2018: 1053).

Med motivi za objavo podatkov je tudi povračilo za vložen trud. Leta 2014 je mreža organizacij na pobudo skupnosti znanstvenikov, knjižničarjev, arhivistov, založnikov in financerjev, združenih v FORCE11, sprejela *Skupno deklaracijo načel citiranja podatkov* (orig. *Joint Declaration of Data Citation Principles*, 2014), ki v prvem načelu pravi: »Citiranju podatkov je treba priznati enak pomen v znanstvenih objavah kot drugim raziskovalnim objektom.« V praksi bi to pomenilo, da kadar objava temelji na podatkih oziroma podatkovni zbirki, je treba podatke uradno navesti v seznamu virov tako kot drugo uporabljeno gradivo. Vendar praksa zaostaja za konsenzom, številni založniki namreč ne zahtevajo citiranja podatkov ali ne zagotavljajo navodil za citiranje, ki bi bila v pomoč pri uresničevanju dogovora (Kratz in Strasser, 2014: 6). V *Coalition on Publishing Data in the Earth and Space Sciences* so leta 2015 pripravili *Izjavo o zavezanosti* (orig. *Statement of Commitment*), s katero so se založbe podpisnice zavezale, da bodo zahtevale ustrezno citiranje znanstvenih podatkov, kadar so ti uporabljeni v znanstvenih člankih. Ta korak je pomemben premik v smeri prepoznavanja in priznanja znanstvenega napora, ki prispeva k izboljšanju dostopa do znanstvenih podatkov, saj je treba jasno navesti, kje natančno so podatki in kako je mogoče do njih dostopati – priporočili so rabo stalnih identifikatorjev. Podatke, opremljene z enoličnim stalnim identifikatorjem (na primer DOI – Digital Object Identifier), je mogoče citirati in so trajno dostopni, spremlja pa jih vsaj minimalen nabor metapodatkov, ki omogočajo najdljivost in izboljšajo uporabnost objavljenih podatkovnih zbirk (Elger idr., 2015: 120).

³⁴ Na primer v 37,3 % vseh člankov v revijah PLOS, ki vključujejo DAS, je navedeno naslednje besedilo: »Vsi relevantni podatki so v članku in spremnem gradivu« (Colavizza idr., 2020: 7).

Število citatov ostaja pomembno merilo vpliva in ponovne rabe raziskav ter sredstvo vrednotenja raziskovalnih del. Analiza revij založnikov PLOS in BMC za leti 2017 in 2018 je pokazala, da avtorji, ki so priložili izjavo o razpoložljivosti podatkov, le v manjšem deležu podatke objavijo v repozitoriju – pri PLOS 12,2 % in pri BMC 20,8 %. Analiza citiranosti člankov pa je pokazala, da so članki, v katerih je za dostop do podatkov navedena povezava do repozitorija, za 25,36 % ($\pm 1,07\%$) večkrat citirani v primerjavi s tistimi, pri katerih so podatki pri avtorju na zahtevo ali kot spremno gradivo članku (Colavizza idr., 2020: 2, 14). Da so znanstveni članki, s katerimi so objavljeni tudi podatki, pogosteje citirani, so ugotavljali tudi za raziskave s področja družboslovja (Pienta idr., 2010), genetike (Piwowar in Vision, 2013; Botstein, 2010), astronomije (Henneken in Accomazzi, 2011; Dorch 2012) in oceanografije (Sears, 2011; Belter, 2014; CESSDA Training Team, 2017–2020). Z že preverjenimi ukrepi in dobrimi praksami lahko znanstvene revije pomagajo raziskovalni skupnosti pri večji preglednosti, najdljivosti in vrednotenju znanstvenih podatkov.

Kljub temu, da so arhivi družboslovnih podatkov po svetu že kaj kmalu promovirali objavljanje in citiranje podatkov (Gherghina in Katsanidou, 2013; Zenk-Möltgen idr., 2018), čemur sledimo tudi v ADP, so v slovenskih družboslovnih revijah veljali različni, nestandardizirani načini navajanja uporabljenih podatkov. Hiter pregled znanstvenih revij, ki jih financira ARRS, je leta 2013 pokazal, da njihove uredniške politike ne zahtevajo deljenja in citiranja podatkov.³⁵ Ta ugotovitev je bila spodbuda, da smo v ADP v naslednjih letih izvedli več aktivnosti z založniki. V okviru projekta RDA vozlišče Slovenija³⁶ smo za revije pripravili *Smernice za oblikovanje politik znanstvenih založb glede navajanja raziskovalnih podatkov v znanstvenih publikacijah in zagotavljanja dostopa do primarnih podatkov, uporabljenih v člankih* (Štebe idr., 2020), ki med drugim vključujejo navodila za avtorje glede citiranja podatkov v znanstvenih člankih – vse s ciljem, da bi založbe prispevale k uveljavljanju transparentnosti in širjenju standardiziranih praks objave raziskovalnih podatkov. V pilotnih aktivnostih, namenjenih preoblikovanju uredniških politik, so sodelovale štiri znanstvene revije: *Slovenščina 2.0: empirične, aplikativne in interdisciplinarne raziskave*, *Socialno delo*, *Prispevki za novejšo zgodovino* in *Documenta Praehistorica*.³⁷ V pilotu so skladno s smernicami pripravile predlog sprememb uredniških politik ter se v uredništvih dogovorile za

³⁵ Pregled smo opravili pred izvedbo delavnice z založbami znanstvenih revij in financerjem (ARRS), ki je potekala v okviru projekta Odprti podatki (2010–2013). Glej <https://www.adp.fdv.uni-lj.si/spoznaj/projekti/pretekli/>.

³⁶ RDA Node Slovenija (2019–2020). Glej <https://www.adp.fdv.uni-lj.si/spoznaj/projekti/>.

³⁷ Spletne strani revij so: <https://revije.ff.uni-lj.si/slovenscina2> (Slovenščina 2.0), <https://www.revija-socialnodelo.si/> (Socialno delo), <https://ojs.inz.si/pnz> (Prispevki za novejšo zgodovino) in <https://revije.ff.uni-lj.si/DocumentaPraehistorica> (Documenta Praehistorica).

bolj ali manj postopno uvajanje. To je eden od načinov, kako znanstvene revije v sodelovanju s področnimi repozitoriji prispevajo k uresničevanju odprte znanosti in odprtih podatkov.³⁸

Sklep

Skladno s svojim poslanstvom ADP kot področno podatkovno središče objavlja »raziskovalne podatke, zanimive za družboslovne analize, s poudarkom na problemih, povezanih s slovensko družbo ali sicer pomembnih za slovensko družbo in družboslovje ne glede na geografske meje« (glej Poslanstvo ADP). Gradiva, objavljena v ADP, so vpisana v katalog COBISS ter avtomatsko povezana v skupen katalog Konzorcija CESSDA. Opis raziskave je zagotovljen tudi v angleškem jeziku, s čimer so raziskave najdljive v širši mednarodni raziskovalni skupnosti. Vsem uporabnikom so prosto dostopna nepodatkovna gradiva, kot so opis raziskave ter spremljajoči dokumenti za razumevanje vsebine raziskave in dokumenti za razumevanje vsebine podatkov. Raziskave, ki so pod režimom omejenega dostopa zaradi zaščite podatkov, so pod posebnimi pogoji dostopne le raziskovalcem, bodisi prek varne povezave ali v varni sobi. Dostop do podatkovnih gradiv za izobraževalne in raziskovalne namene je omogočen uporabnikom, ki se predhodno registrirajo. Uporabniki se v postopku registracije zavežejo, da bodo podatke pri uporabi citirali in tako priznali avtorstvo raziskovalcem, ki so podatke zbrali. Z objavo v ADP avtorji raziskave pridobijo vse potrebne bibliografske podatke, s katerimi je mogoče njihovo raziskavo citirati. Navodila za citiranje so na voljo ob vsaki raziskavi. Tako ADP nastopa tudi v vlogi založnika objav podatkov.

Kadar izbirajo med založniki, naj bodo raziskovalci pozorni na razlike v povezavi z zagotavljanjem kakovosti podatkovne objave. ADP prejeto podatkovno gradivo pregleda z vidika celovitosti ter opravi tehnični pregled in oceno znanstvene vrednosti raziskave. V postopku, ki ga je mogoče primerjati z recenzijo, ocenjujemo naslednje:

znanstvena pomembnost se kaže v uporabnosti za raziskovanje širokega nabora aplikativnih ali teoretsko osmišljenih problemov. Zbirka podatkov mora biti rezultat zaključene raziskave in ustrezati visokim merilom kakovosti, ki se jo ocenjuje na podlagi izčrpne spremljajoče dokumentacije. Zbirka podatkov mora biti javno dostopna v nacionalnem ali mednarodnem znanstvenem podatkovnem arhivu. Zbirka podatkov mora biti dokumentirana in dostopna v takšni

³⁸ Za več o sodelovanju med ADP in znanstvenimi revijami glej Štebe idr., 2020.

obliki, da omogoča ponovitev objavljenih znanstvenih ugotovitev, izvedenih na njeni podlagi (Bibliografska merila znanstvene in strokovne uspešnosti).³⁹

Postopek pregleda in znanstvenega vrednotenja podatkov ADP uvršča na seznam agencij, na katerega so uvrščene znanstvene revije, pri katerih raziskovalci z objavo pridobijo znanstvene točke. Po merilih ARRS lahko zaključena znanstvena zbirka podatkov ali korpus (2.20) s seznama agencije (BIBLIO-D), na katerega je uvrščen ADP, dobi 30 znanstvenih točk. ADP v pregledu ugotovi celovitost podatkovne objave. Pri vrednotenju je v ospredju, ali je zbirka podatkov rezultat zaključene raziskave in je razumljivo predstavljena v vsej celovitosti ter jo je mogoče predstaviti in razumeti kot samostojen znanstveni rezultat. Če gre za zbirko, ki je zgolj v funkciji spremnega gradiva znanstvenemu članku in je njen obseg omejen zgolj na razumevanje rezultatov, predstavljenih v članku, je s tem okrnjena uporabnost za raziskovanje širokega nabora aplikativnih ali teoretsko osmišljenih problemov. V tem primeru ugotovimo, da ne gre za podatkovno objavo, ki bi ustrezala merilom za zagotovitev znanstvenih točk. Če se navežemo še na druga prej omenjena formata podatkovnih objav, to sta podatkovni članek, objavljen v podatkovni reviji, in podatkovno poročilo, v katerih gre zgolj za opisa podatkovne zbirke, kot takšna nista zaključeni podatkovni zbirki.

V množici podatkovnih storitev in raznoliki podatkovni krajini so raziskovalci, ki še nimajo veliko izkušenj s podatkovnimi objavami, upravičeno pred dilemo, kje objaviti podatke. Namen tega članka je bil pomagati pri razumevanju raznolikosti in orientaciji, ko se odločamo o objavi podatkov ali opisov podatkov. Če povzamemo: kadar želijo raziskovalci objaviti specializirane celovite zbirke podatkov, za katere doma nimajo niti infrastrukture niti podatkovnih storitev, bodo te iskali v mednarodnem okolju. Kadar pa imajo podporo domačih strokovnjakov ter so jim na voljo razvite storitve in infrastruktura, naj bo nacionalni področni podatkovni repozitorij prva izbira, ki jim bo v oporo pri pripravi podatkov, tako da bodo objavili še en samostojen znanstveni rezultat ter ga varno in trajno delili s svetovno raziskovalno skupnostjo. ADP je za področje družboslovja primer take domače infrastrukture oziroma podatkovnih storitev in je v funkciji založnika podatkov, ki raziskovalcem zagotovi povračilo vloženega truda v pripravo samostojnega znanstvenega rezultata v obliki podatkovne zbirke – to so pogoji za citiranje take objave –, v primeru znanstvene odličnosti pa tudi znanstvene točke.

³⁹ Ker je ADP na seznamu agencij ARRS, se zaključene znanstvene zbirke, objavljene v ADP, upoštevajo pri kategorizaciji znanstvenih publikacij. Objava na primer v repozitoriju Univerze v Ljubljani ne zahteva izčrpane spremljajoče dokumentacije in ocene, ali zaključena zbirka ustreza visokim merilom kakovosti, zato ni na seznamu agencij ARRS in znanstvenih točk ne prinese.

Literatura

- Arzberger, Peter, Peter Schroeder, Anne Beaulieu, Geoffrey C. Bowker, Kathleen Casey, Leif Laaksonen, David Moorman, Paul Uhlir in Paul Wouters (2004): Promoting Access to Public Research Data for Scientific, Economic, and Social Development. *Data Science Journal* 3(29): 135–152. Dostopno na DOI: [10.2481/dsj.3.135](https://doi.org/10.2481/dsj.3.135).
- Bezjak, Sonja in Sergeja Masten (2021): Od podatkovnega kolonializma do podatkovne pravičnosti: primer obravnave manjšin v dobi podatkov. *Časopis za kritiko znanosti* 282: 203–218.
- Bibliografska merila znanstvene in strokovne uspešnosti* (2014). Dostopno na: <http://www.arrs.si/sl/akti/inc/2015/Bibl-merila-Priloga1.pdf> (10. november 2020).
- Borgman, Christine L. (2013): Why You Should Care About Open Data: Open Access Week Thoughts on Why Research Data Rarely Are Reused. *Open Access Week at UCLA*. Dostopno na: <http://works.bepress.com/borgman/274/> (10. november 2020).
- Borgman, Christine L., Andrea Scharnhorst in Milena S. Golshan (2018): Digital Data Archives as Knowledge Infrastructures: Mediating Data Sharing and Reuse. *Journal of the Association for Information Science and Technology*: 1–31. Dostopno na: <https://arxiv.org/ftp/arxiv/papers/1802/1802.02689.pdf> (10. november 2020).
- Callaghan, Sarah, Steve Donegan, Sam Pepler, Mark Thorley, Nathan Cunningham, Peter Kirsch, Linda Ault, Patrick Bell, Rod Bowie, Adam Leadbetter, idr. (2012): Making Data a First Class Scientific Output: Data Citation and Publication by Nerc's Environmental Data Centres. *International Journal of Digital Curation* (7/1): 107–113. Dostopno na: <http://www.ijdc.net/article/view/208> (10. november 2020).
- CESSDA Training Team (2017–2020). *CESSDA Data Management Expert Guide*. Bergen, Norveška: CESSDA ERIC. Dostopno na: <https://www.cessda.eu/DMGuide> (10. november 2020).
- Chase, Adrian, Diane Chase in Arlen Chase (2020): Ethics, New Colonialism, and Lidar Data: A Decade of Lidar in Maya Archaeology. *Journal of Computer Applications in Archaeology* (3): 51–62.
- Colavizza, Giovanni, Iain Hrynaszkiewicz, Isla Staden, Kirstie Whitaker in Barbara McGillivray (2020): The Citation Advantage of Linking Publications to Research Data. *PLoS ONE* (15/4): 1–18, e0230416. Dostopno na DOI: [10.1371/journal.pone.0230416](https://doi.org/10.1371/journal.pone.0230416).
- Data Citation Synthesis Group (2014): *Joint Declaration of Data Citation Principles*. San Diego, Kalifornija: FORCE11. Dostopno na DOI: [10.25490/a97f-egykh](https://doi.org/10.25490/a97f-egykh).
- Dechamp, Jean-François (2020): The European Commission and Open Science: Recent Developments. *UNICA & EURLO Scholarly Communication Working Groups Online Workshop*, 27. november.
- Elger, Kirsten, Boris Biskaborn, Heinz Pampel in Hugues Lantuit (2016): Open Research Data, Data Portals and Data Publication – An Introduction to the Data Curation Landscape. *Polarforschung* (85/2): 119–133. Dostopno na DOI: [10.2312/polfor.2016.009](https://doi.org/10.2312/polfor.2016.009).
- Gherghina, Sergiu in Alexsia Katsanidou (2013): Data Availability Policies in Political Science Journals. *Eur Polit Sci* (12): 333–349. Dostopno na DOI: [10.1057/eps.2013.8](https://doi.org/10.1057/eps.2013.8).
- Gray, Jim (2009): Jim Gray on eScience: A Transformed Scientific Method. V *The Fourth Paradigm: Dataintensive Scientific Discovery*, T. Hey, S. Tansley in K. Tolle (ur.): xvii–xxxii. ZDA: Microsoft Research.
- Hrynaszkiewicz, Iain (2019): Building Trust to Break Down Barriers. V *The State of Open Data Report 2019. Digital Science Report. A Selection of Analyses and Articles about Open Data, Curated by Figshare*, 6–7. Dostopno na DOI: [10.6084/m9.figshare.9980783.v2](https://doi.org/10.6084/m9.figshare.9980783.v2).
- Kelling, Steve, Wesley M. Hochachka, Daniel Fink, Mirek Riedewald, Rich Caruana, Grant Ballard in Giles Hooker (2009): Data-Intensive Science: A New Paradigm for Biodiversity Studies. *BioScience* (59/7): 613–620. Dostopno na DOI: [10.1525/bio.2009.59.7.12](https://doi.org/10.1525/bio.2009.59.7.12).

Kratz, John in Carly Strasser (2014): Data Publication Consensus and Controversies *F1000Research* (3:94): 1–19. Dostopno na DOI: [10.12688/f1000research.3979.3](https://doi.org/10.12688/f1000research.3979.3).

Lawrence, Bryan, Catherine Jones, Brian Matthews, Sam Pepler in Sarah Callaghan (2011): Citation and Peer Review of Data: Moving Towards Formal Data Publication. *International Journal of Digital Curation* (6/2): 4–37. Dostopno na DOI: [10.2218/ijdc.v6i2.205](https://doi.org/10.2218/ijdc.v6i2.205).

Lin, Dawei, Jonathan Crabtree, Ingrid Dillo, Robert R. Downs, Rorie Edmunds, David Giaretta, Marisa de Giusti, Hervé L'Hours, Wim Hugo, Reyna Jenkyns, idr. (2020): The TRUST Principles for Digital Repositories. *Scientific Data* (7): 1–5. Dostopno na DOI: [10.1038/s41597-020-0486-7](https://doi.org/10.1038/s41597-020-0486-7).
Open Knowledge International. Dostopno na: okfn.org.

OpenAIRE (2017): *Briefing Paper Research Data Management*. Dostopno na: <https://www.openaire.eu/briefpaper-rdm-infonoads> (10. november 2020).

Parsons, Mark in Peter Arthur Fox (2013): Is Data Publication the Right Metaphor? *Data Science Journal* (12): 32–46. Dostopno na DOI: [10.2481/dsj.WDS-042](https://doi.org/10.2481/dsj.WDS-042).

Piwovar, Heather (2011): Who Shares? Who Doesn't? Factors Associated with Openly Archiving Raw Research Data. *PLoS One* (6/7): 1–13. Dostopno na DOI: [10.1371/journal.pone.0018657](https://doi.org/10.1371/journal.pone.0018657).

Poslanstvo ADP. Dostopno na: <https://www.adp.fdv.uni-lj.si/spoznaj/adp/poslanstvo/> (10. november 2020).

RDA International Indigenous Data Sovereignty Interest Group (2019): *CARE Principles for Indigenous Data Governance*. Dostopno na: <https://www.gida-global.org/care> (15. oktober 2020).

Shin, Nagai, Hideaki Shibata, Takeshi Osawa, Takehisa Yamakita, Masahiro Nakamura in Tanaka Kenta (2020): Toward More Data Publication of Long-Term Ecological Observations. *Ecological Research* (35/5): 700–707. Dostopno na DOI: [10.1111/1440-1703.12115](https://doi.org/10.1111/1440-1703.12115).

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.4040643](https://doi.org/10.5281/zenodo.4040643).

Štebe, Janez, Sonja Bezjak in Sanja Lužar (2013): *Odprti podatki: načrt za vzpostavitev sistema odprtega dostopa do raziskovalnih podatkov v Sloveniji*. Ljubljana: Fakulteta za družbene vede. Dostopno na: http://www.adp.fdv.uni-lj.si/media/publikacije/2013/Celotno_porocilo_odpp10_zakljucno.pdf (10. november 2020).

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-2020-049](https://doi.org/10.5334/dsj-2020-049).

Štebe, Janez, Sonja Bezjak in Irena Vipavc Brvar (2015): Priprava raziskovalnih podatkov za odprti dostop: priročnik za raziskovalce. Dostopno na: <http://www.dlib.si/stream/URN:NBN:SI:DOC-06SLBVXX/507094f4-625d-4539-9722-7da5b8f433d0/PDF> (10. november 2020).

Tenopir, Carol, Elizabeth D. Dalton, Suzie Allard, Mike Frame, Ivanka Pjesivac, Ben Birch, Danielle Pollock in Kristina Dorsett (2015): Changes in Data Sharing and Data Reuse Practices and Perceptions amongst Scientists Worldwide. *PLoS ONE* 10(8): 1–24. Dostopno na: DOI: [10.1371/journal.pone.0134826](https://doi.org/10.1371/journal.pone.0134826).

Tenopir, Carol, Suzie Allard, Kimberly Douglass, Arsev Umur Aydinoglu, Lei Wu, Eleanor Read, Maribeth Manoff in Mike Frame (2011): Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* (6/6): 1–21. Dostopno na DOI: [10.1371/journal.pone.0021101](https://doi.org/10.1371/journal.pone.0021101).

Vines, Timothy H., Arianne Y. K. Albert, Rose L. Andrew, Florence De'Barre, Dan G. Bock, Michelle T. Franklin, Kimberly J. Gilbert, Jean-Sebastien Moore, Sebastien Renaut in Diana J. Rennison (2014): The Availability of Research Data Declines Rapidly with Article Age. *Current Biology* (24): 94–97. Dostopno na DOI: [10.1016/j.cub.2013.11.014](https://doi.org/10.1016/j.cub.2013.11.014).

Wiley (2014): *Researcher Data Sharing Insights*. Dostopno na: <https://3spxpi1radr22mzge33bla91-wpengine.netdna-ssl.com/wp-content/uploads/2014/11/researcher-data-insights-infographic-final.pdf> (3. marec 2021).

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, idr. (2016): The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data* (3/160018): 1–9. Dostopno na DOI: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).

Zenk-Möltgen, Wolfgang, Esra Akdeniz, Alexia Katsanidou, Verena Naßhoven in Ebru Balaban (2018): Factors Influencing the Data Sharing Behavior of Researchers in Sociology and Political Science. *Journal of Documentation* (74/5): 1053–1073. Dostopno na DOI: [10.1108/JD-09-2017-0126](https://doi.org/10.1108/JD-09-2017-0126).