

Od podatkovnega kolonializma do podatkovne pravičnosti: primer obravnavе manjšin v dobi podatkov

Abstract

From Data Colonialism to Data Justice: The Treatment of Minorities in the Data Age

In the age of data, it is important to rethink established theoretical concepts that define power relations and consequently the positions, roles and agency of individuals, communities, organizations and even countries. The article highlights data colonialism and the related issue of data ownership. Data is often referred to as the new gold or the new oil, and the authors address the questions of who owns data and the ways in which it can be appropriated. In addition, the article also tackles questions such as who can access data and under what conditions, and outlines how data relations are defined in an age where virtually all aspects of individual and social life are commodified. The article also explores the possible avenues for resistance and whether historical experiences can help us identify them. Special attention is given to minorities, whose position in the data age can be rethought through concepts such as data inequality, data inclusion, data justice and data activism. In this context, open data can be viewed as an important building block of open science, allowing the research community to achieve more visibility and greater representation of all minorities. In order to achieve data justice, science also needs to (re)consider various approaches to data.

Keywords: data colonialism, data justice, open data, open science

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).

Sergeja Masten holds a Master's degree in Sociology and is currently working at the Slovenian Social Science Data Archives, where her primary role is the digital curation of

social science research data, as well as the training of data providers and users. She is an active proponent of open science. Her professional interests include topics such as gender and economic inequalities, as well as ecological/environmental issues (sergeja.masten@fdv.uni-lj.si).

Povzetek

V dobi podatkov velja na novo premisliti stare, uveljavljene teoretske koncepte, s katerimi so opredeljena razmerja moči, s tem pa tudi določeni položaji, vloge in možnosti delovanja posameznikov, skupnosti, organizacij in držav. Med njimi bomo posebej izpostavili podatkovni kolonializem ter z njim povezano vprašanje lastništva podatkov. Kdo in kako si prilašča podatke, ki jih v sodobnem jeziku pogosto imenujemo kar nova nafta ali novo zlato, ter kdo in pod kakšnimi pogoji ima do njih dostop in posledično koristi? Kako so opredeljeni podatkovni odnosi v dobi, ko so praktično vsi vidiki posameznikovega in družbenega življenja poblagovljeni? Kakšne so možnosti za odpor in kako nam lahko pri tem pomagajo zgodovinske izkušnje? Posebno pozornost bomo namenili manjšinam, katerih položaj v dobi podatkov bomo premislili s koncepti, kot so podatkovna neenakost, podatkovna vključenost, podatkovna pravičnost in podatkovni aktivizem. Odprte podatke vidimo kot enega od pomembnih gradnikov odprte znanosti, s katerim lahko raziskovalna skupnost prispeva k večji zastopanosti vseh manjšin. Da bi dosegli podatkovno pravičnost, je tudi v znanosti potreben premislek o podatkovnih prijemih.

Ključne besede: podatkovni kolonializem, podatkovna pravičnost, odprti podatki, odprta znanost

Sonja Bežjak 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 prostočasnimi aktivnostmi, pri katerih si angažirano prizadeva za uresničevanje socialne in okoljske pravičnosti (sonja.bezjak@fdv.uni-lj.si).

Sergeja Masten je magistrica sociologije, zaposlena v Arhivu družboslovnih podatkov (ADP), kjer je njena primarna vloga digitalno arhiviranje družboslovnih raziskav. Posveča se še usposabljanju dajalcev in uporabnikov podatkov. Je aktivna promotorka odprte znanosti. Poklicno jo zanimajo teme, kot so spolna in ekonomska neenakost ter okoljska problematika (sergeja.masten@fdv.uni-lj.si).

Uvod

V članku predstavlja osnovne pojme, s katerimi je mogoče kritično preištevati *dobro podatkov* v luči *prilaščanja podatkov* in *podatkovnega nadzora* s strani podjetij in državnih organov ter podatkovnega aktivizma in podatkovne pravičnosti, ki si jo bo treba še izboriti. V prvem delu govoriva o podatkovnem kolonializmu, pri čemer opozoriva na *podatkizem*, ideologijo, ki ga preprečuje uzreti v njegovi izkoriščevalski naravi in ga kaže kot naravno stanje, neodvisno od družbenih razmer in razmerij. Nujnost preiščevanja utemeljujeva z *upodatkovljenjem* družbenega življenja, ko so vsi segmen-

ti človekovega družbenega delovanja kvantificirani in preoblikovani v podatkovne zbirke, s katerimi upravljajo podjetja in države, ne da bi pri tem upoštevali temeljne elemente podatkovne pravičnosti, kot so nediskriminacija, pravica do nevidnosti, nevključenosti v tehnologijo oziroma neuporabe tehnologije, kar vključuje pravico do zasebnosti in anonimnosti. V dobi podatkov se vsi uporabniki digitalnih tehnologij srečujemo z diskriminacijo, neosveščenostjo in nemočjo, pri čemer so manjšine, revni, begunci in lokalne skupnosti izpostavljeni dvojni diskriminaciji, ko se zoper njih uvajajo dodatni podatkovni prijemi, kot so izdatno izžemanje podatkov o beguncih ob prestopu meje, podatkovni nadzor prejemnikov socialne pomoči, onemogočanje dostopa do podatkov lokalnim skupnostim v degradiranih okoljih in podobno. Ta dvojna diskriminacija je bila razlog, da sva ob strani pustili večino in v središče razprave postavili tiste, ki so v dobi podatkov znova soočeni z reprodukcijo obstoječih družbenih neenakosti in asimetrijo družbene moči.

Prizadevali sva si iskati ustrezne prevode pojmov, ki so izvirno uporabljeni v angleškem jeziku in se nekateri tukaj prvič pojavijo v slovenskem jeziku. Raziskovalno javnost vabiva, da prevode kritično premisli in po potrebi predlaga popravke.

Podatkovni kolonializem

Razvoj in porast medijskih tehnologij sta na eni strani poenostavila in pospešila prenos informacij med ljudmi ter med podjetji, državnimi organi in ustanovami, po drugi strani pa sta prispevala k nastanku ne le novega informacijskega okolja, pač pa tudi podatkovnega okolja. V tem članku naju zanimajo informacije kot podatki, ki jih nekdo zbira, organizira v podatkovne zbirke in z njimi upravlja. Te so lahko o ljudeh, o njihovem življenjskem okolju ali o drugih virih in vrednostih, ki določajo razmerja moči med posamezniki, organizacijami in državami. Številni domači in tuji spletni portali pišejo, da so podatki nova nafta in podjetniška priložnost. Couldry in Mejias (2019a: 3) menita, da podatki niso *kot* nafta, ampak jih moramo videti kot družbeni konstrukt, ki deluje v danem zgodovinskem trenutku ter ga usmerjajo ekonomske in družbene sile (Newman, 2014: 11). Van Dijck opozarja, da ljudje niso osveščeni o tem, na kaj pristanejo z uporabo digitalnih tehnologij. Američani so šele z razkritjem Edwarda Snowdena leta 2013 ugotovili, da v zameno za uporabo spletnih platform ponudnikom teh storitev odstopijo del svoje zasebnosti. Uporabniki so postopno sprejeli in se navadili na to,

da *delijo* informacije o sebi, na primer o svojem zakonskem stanu, bolezni, prehranjevalnih navadah in najljubši glasbi, kar je postalo nova norma. Na osnovi te nove norme lastniki spletnih platform agregirane metapodatke rutinsko delijo s tretjimi strankami, in sicer z namenom, da bodo trženje svojih produktov oziroma storitev prilagodili uporabnikom (van Dijck, 2014: 197).

Proces zajema in procesiranja podatkov o posameznikih in družbi vključuje tudi prvino prilaščanja podatkov. V procesu prilaščanja podatkov ne gre za nikakršno »naravno konverzijo«, pač pa za podatkovna razmerja, ki vzpostavljajo nov družbeni red in nove oblike kolonializma (Mann in Daly, 2019: 381). Podatkovni kolonializem deluje s pomočjo poblagovljenja podatkov in akumulacije kapitala z razlastninjenjem (Thatcher idr., 2016: 991). Da bi razumeli sodobne razmere in procese, moramo upoštevati zgodovinske okoliščine. Couldry in Mejias ugotavljata, da je zgodovinski kolonializem izkoriščal ljudi ter jim vladal s prilaščanjem ozemelj in virov, danes pa se kolonializem kaže v izkoriščanju ljudi s pomočjo podatkov (Couldry in Mejias, 2019: 336). Kolonialni razvoj lahko opazujemo v kvantifikaciji socialnih reprezentacij, to je pretvarjanju in prikazovanju človekovega delovanja v številskih tabelah, kar poteka z ekstrakcijo podatkov. Sočasno poteka še en pomemben proces, to je razlastninjenje lastnikov podatkov in prisvajanje oziroma prilaščanje podatkov s strani lastnikov orodij, platform za ekstrakcijo podatkov.¹

Uporabniki imajo lažen občutek, da v zameno za osebne podatke dobijo nekaj storitev od platform, kot sta Google in Facebook, v resnici pa so ta podjetja odvisna od brezplačnega dela posameznikov, ki ažurno objavljajo na Facebooku, pregledujejo Amazon, urejajo gesla na Wikipediji in pišejo bloga, s čimer te storitve šele dobijo svojo vrednost (Newman, 2014: 16).²

Čeprav so se spremenili načini, intenzivnost, razsežnosti in okoliščine razlastninjenja, podstat sodobnega podatkovnega kolonializma ostaja enaka: osvojiti ozemlje in vire, iz katerih bo mogoče pridobiti ekonomsko vrednost. Da bi razumeli družbene procese v dobi podatkov, je treba preučevati poslovni model, utemeljen na ekstrakciji in menedžmentu podat-

1 Platforme razumeva kot spletne strani družbenih omrežij (na primer Facebook), spletne iskalnike (na primer Google), pametne telefone in operacijske sisteme, kot na primer Android, ki kot digitalni posredniki močno vplivajo na človekovo vsakdanje življenje (Jin, 2013: 145).

2 Velja dodati, da je med vodilnimi spletnimi platformami Wikipedija ena redkih, ki je organizirana po neprofitnem modelu, zato je v političnoekonomskem smislu ne gre umeščati med podjetja, ki z delom uporabnikov platform akumulirajo kapital.

kov, ter njegovo samoumevnost; z njim pa tudi temeljna vidika preobrazbe družbenega življenja,³ ki bi ju sicer pustili ob strani zgolj kot stranska učinka, to sta *podreditev* človeških bitij, ki je nujna za prilastitev virov, in utemeljitev te preobrazbe s splošno racionalnostjo, ki se svetu vsiljuje kot edinstvena vizija in vednost o t. i. velikih podatkih (Couldry in Mejias, 2019a: 2–3). Opazovati velja tudi razporejanje moči na globalni ravni. V 21. stoletju imamo na eni strani peščico zahodnih držav, ki so razvile digitalne platforme in z njimi upravljajo, med njimi izstopajo ZDA, na drugi strani pa veliko večino nezahodnih držav, ki takih naprednih platform nimajo v lasti in jih ne upravljajo. ZDA so s platformami obnovile imperializem oziroma njegovo različico, platformni imperializem pa podpirajo politika, ekonomija, vojska in kultura (Jin, 2013: 167).

Podatkizem in upodatkovljenje

»Kako pojasniti toleranco do velikega brata in velikega biznisa, ki rutinsko dostopata do osebnih informacij, znanih pod imenom veliki podatki?« sprašuje van Dijck in kot odgovor ponudi ang. *dataism* oziroma sl. podat-kizem, ideologijo, s katero tehnološka in oglaševalska podjetja pridobivajo podatke o uporabnikovem delovanju (van Dijck, 2014: 198). Ta se kaže v razširjenosti *verjetja* v objektivnost kvantifikacije ter v potencial sledenja vseh vrst človekovega vedenja in delovanja s pomočjo spletnih medijskih tehnologij. Ideologija temelji na *zaupanju* uporabnikov do tistih, ki zbira-jo, interpretirajo in delijo (meta)podatke, pridobljene z digitalnih platform. Pritrjujejo ji tudi številni raziskovalci, ki tako pridobljene podatke vidijo kot dokaze o posameznikovem dejanskem vedenju ali razpoloženju, medtem ko platforme razumejo kot nevtralne posrednike oziroma zbiralce podatkov (van Dijck, 2014: 198–199). Mosco tak pristop definira kot digitalni po-zitivizem. Opiše ga kot mit, po katerem velja, da znanje ne prihaja od ljudi, pač pa iz čistih podatkov, shranjenih v oblaku (Mosco, 2014: 193).

Zbiranje, kvantificiranje in razvrščanje podatkov na podlagi družbe-nih aktivnosti uporabnikov digitalnih tehnologij omogoča, da jim platfor-me sledijo in jih nadzorujejo v realnem času ter pripravljajo napovedne analize. Mayer-Schoenberger in Cukier (2013 v van Dijck, 2014: 198) po-pis družbenega delovanja v obliki spletnih številskih podatkov imenujeta

3 Nekaj primerov: preobrazba družbenega življenja je mnogo več od spremembe v panogi socialnih medijev, so tudi nove metode nadzora na delovnem mestu, vzpon logistike, pojav novih platform za generiranje profita v prometu in turizmu (Couldry in Mejias, 2019a: 2).

datafikacija oziroma sl. upodatkovljenje.⁴ Govorimo lahko o *družbenem upodatkovljenju*, kjer poteka pretvorba informacij o posameznikovem življenju v podatkovne zbirke (nize). To omogočajo digitalne tehnologije in naprave, s pomočjo katerih posamezniki vsakodnevno vstopajo v podatkovne odnose. Podatkovne odnose je mogoče nadzorovati ter posameznike spremeniti v podatkovne subjekte, podatke o njihovem delovanju pa v ekonomsko vrednost s pomočjo na osebe usmerjenih praks in napovedne analitike (Calzati, 2020: 2). Veliko podatkov, ki jih zbirajo digitalne tehnologije, se sicer uporabi za klasično oglaševanje, vendar poleg tega podjetja na podlagi delovanja svojih uporabnikov zbirajo njihove osebne podatke in oblikujejo uporabniške profile, ki pa omogočajo druge možnosti izkoriščanja.⁵ Podjetja uporabniško vedenje na spletu povezujejo z dostopnimi podatki (plačila, demografija, lokacija) in na podlagi tega oblikujejo profilne skupine, ki so posledično tarča prilagojenih oglasov in ponudb. Industrija podatkovnih posrednikov pozna celo pojem »seznam naivnežev« za revne, stare in manj izobražene skupine uporabnikov, ki jih prodajajo oglaševalcem, ti pa se poslužujejo neetičnih tržnih praks.⁶ Obravnavati velja še en vidik, in sicer povečevanje ekonomske neenakosti, h kateri prispevajo komercialne platforme. Ko s podatkovnim rudarjenjem in oblikovanjem uporabniških profilov zbirajo velike podatke, prispevajo k informacijski asimetriji na potrošniških trgih (Newman, 2014: 11, 14, 21).

Podatkovni nadzor

Zbiranje podatkov s strani podjetij in držav poleg akumulacije kapitala prinaša še več mehanizmov nadzora in dominacije. Digitalna infrastruktura z nadzornimi programi, ki upravljajo z velikimi podatki, ustvarja nesorazmerja moči med tistimi, ki nadzorujejo, ter nadzorovanimi, ki imajo manj

4 Tehnološka podjetja zbirajo podatke o posameznikovih prijateljstvih in preferencah (Facebook), »sledilcih« in »zasledovanih« (Twitter), iskalcih in ponudnikih zaposlitev (LinkedIn), izmenjevalcih avdiovizualnih vsebin, številna pa tudi o povsem vsakdanjih pogovorih in zanimanjih (van Dijck, 2014: 198–199).

5 Med njimi tudi takšne, ki niso nujno v korist uporabnikov. Na primer: podjetje jim na podlagi uporabniškega profila ponudi višjo ceno produkta kot drugim, ne da bi to vedeli, uporabnikom iz gospodinjstev z nižjimi dohodki pa ponudi slabše posojilne pogoje. Sploh ne gre za to, da bi se podjetja ukvarjala s posamezniki, saj oblikujejo ponudbo za uporabniške kategorije, problematičen je vidik, da uporabniki s tem niso seznanjeni (Newman, 2014: 11).

6 Lahko pa se tudi zgodi, da se posameznik zaradi napačnih izračunov spletnih algoritmov znajde v »napačni« skupini. Ne da bi vedel zakaj, mu banka zniža limit na kreditni kartici (Newman, 2014: 14).

moči in sredstev, da bi se nadzoru izognili. Ameriški žvižgač Edward Snowden, ki je leta 2013 novinarjem razkril množico tajnih dokumentov iz ameriške Nacionalne varnostne agencije ter odprl vprašanje nacionalne varnosti in posameznikove zasebnosti, je pokazal, da so vladne agencije v ZDA in Veliki Britaniji aktivno nadzirale politične skupine, med drugim mednarodne organizacije, kot na primer UNICEF, Amnesty International in Human Rights Watch, pa tudi ugledne posameznike (na primer vodjo pisarne Al-Jazeera Pakistan) in državljane, ki so obiskali spletno stran Wikileaks⁷ (Dencik idr., 2016: 3). Študije iz ZDA kažejo, da so podatkovne prakse škodljive za nebelce, saj so izpostavljeni dvojnemu podatkovnemu kolonializmu. Podobno Mann in Daly ugotavljata za Avstralijo, kjer so leta 2007 v oddaljenih staroselskih skupnostih s projektom *The Intervention* skušali zajezi zlorabe in nasilje. V okoli 70 staroselskih skupnostih so na javne računalnike namestili program za nadzor nad rabo spleta in program za preprečevanje ogleda strani z odraslo pornografijo, ki je v Avstraliji zunaj teh skupnosti sicer veljala za legalno. Prav tako so ponudniki interneta na javnih računalnikih v teh skupnostih morali hraniti podatke o rabi računalnika in uporabnikih. Projekt je bil deležen kritik zaradi kršitev mednarodnih standardov človekovih pravic, označili so ga za »izjemen primer institucionalnega nadzora marginaliziranih in ranljivih skupin«. Še en primer je uvedba t. i. *BasicsCard*, s katero nadzorujejo porabo denarne socialne pomoči v staroselskih skupnostih. S kartico na primer ni mogoče kupiti cigaret in alkohola ter dvigniti gotovine (Mann in Daly, 2019: 384–385). Digitalni identifikacijski sistemi naredijo revne ali priseljence vidnejše, kar je lahko koristno, hkrati pa jih izpostavijo možnim manipulacijam tistih, katerih interes je, da jih nadzirajo ali vplivajo na njihovo vedenje. »Ljudje morajo imeti digitalno identiteto, ki jo lahko preberejo najrazličnejši akterji za potrebe kategorizacije, klasifikacije, pa tudi potencialno za potrebe varnosti, preverbe in nadzora nad pravico do virov na mednarodni ravni« (Martin in Taylor, 2020: 2). Zbiranje podatkov s strani držav in podjetij deluje znotraj širših nadzornih sistemov kapitalizma in zagotavlja nadvlado, zato aktiviste s področja podatkovne pravičnosti skrbi obsežen in rasističen nadzor (Vera idr., 2019: 1017).

Podatki so lahko *nasilni*, kadar so učinkovito sredstvo za dodatno izključevanje in marginaliziranje s pomočjo različnih metod, na primer prognostične policijske ocene,⁸ prostorsko zamejevanje manjšin (ang. *redlining*),⁹

7 Spletna stran multinacionalne medijske organizacije, ki jo je leta 2006 ustanovil Julian Assange (<https://wikileaks.org/>).

8 Ang. *predictive policing*.

9 Termin »redlining« je leta 1960 skoval John McKnight. Z njim je opisal delovanje bank, ki prebivalcem z določenega območja niso dale posojil. Kasneje so termin uporabljali za različne

sistem finančnega točkovanja (ang. *financial scoring*) in nadzor.¹⁰ Staroselski strokovnjaki opozarjajo, da je treba zbiranje podatkov preusmeriti stran od tistih, ki so že tako ali tako nadzorovani in kot taki že označeni za »problematične«. Uporabiti jih je treba za emancipacijo in krepitev moči. Kadar prispevajo k boljšemu razumevanju kompleksnih pojavov, kot je podnebna kriza, ali k izboljšanju življenjskih razmer v skupnosti temnopoltih, so podatki izjemno koristni, pravijo Vera idr. (2019: 1016–1017).

Zbiranje podatkov o koloniziranih ljudeh in njihova uporaba ter kategorizacija kot sredstvo kolonialnega nadzora pri kolonialnih silah niso nekaj novega (Mann in Daly, 2019: 381). Novo je, da digitalne tehnologije omogočajo množično zbiranje informacij o ljudeh ter njihovo upodatkovljenje, deljenje in analiziranje.

Podatkovna pravičnost

Da bi se lahko ustrezno ukvarjali z vprašanjem demokratičnosti in odpravljanjem družbenih neenakosti, v razpravo vpeljujemo pojem podatkovne pravičnosti. Podatki igrajo pomembno vlogo pri (samo)odločanju in uveljavljanju individualnih ali skupinskih pravic, vendar imata procesa pozitiven vpliv na vse udeležence le, če sta upravljanje s podatki in dostop do podatkov urejena demokratično. Razvoj novih, še kompleksnejših digitalnih tehnologij tega velikokrat ne omogoča, bodisi ker to ni v interesu kapitala bodisi ker je podatkov toliko, da je v praksi težko premišljeno spremljati vse vidike njihovega nastanka in uporabe. Čeprav dobo podatkov lahko razumemo kot pomembno stopnjo v razvoju človeštva, prinaša nove izzive ter s tem potrebo po novih premislekih in zagotovilih, da tehnologija in z njo podatki ne bodo diskriminatorni in izključujoči (Taylor, 2017). O podatkih je treba premišljevati tudi z vidika podatkovnega nasilja in si prizadevati za popis podatkovne škode (Vera idr., 2019: 1015).

Zaradi kompleksnosti položaja ne zadošča več izhajanje iz okvira temeljnih človekovih pravic, marveč moramo iskati nov pristop, ki bo upošteval raznovrstnost udeležencev in možnosti v dobi podatkov. Demokratičen pristop mora upoštevati novosti in kompleksnost načinov, na katere lahko

diskriminatorne prakse, ki pomenijo izključevanje iz določenih storitev na podlagi ocene določene skupnosti.

10 Nasilne prakse je mogoče prepoznati tudi na straneh platform, ki zbirajo velike podatke o uporabnikih in jih prodajajo tretjim, na primer oglaševalcem, in sicer tako, da uporabniki na tej osnovi izgubijo nadzor nad svojimi podatki, ki ne služijo več njim, pač pa oglaševalcem (Newman, 2014: 18).

podatki diskriminirajo, disciplinirajo in nadzirajo. Upoštevati mora pozitivne in negativne potenciale novih podatkovnih tehnologij – njihov prispevek k razvoju človeštva in tveganje, da razvoj zavrejo. Prav tako mora razvijati dvojne standarde, upoštevajoč razliko med državami z višjim oziroma nižjim dohodkom glede na zasebnost in vidnost (Taylor, 2019: 8–10).

Dencik idr. (2016) v središče pojma podatkovne pravičnosti postavijo razumevanje, da kapitalizem, ki nadzoruje, omejuje aktivizem in državljanstvo. Podatkovno pravičnost uporabljajo za opisovanje upora proti (ne) vladnemu nadzoru. Za uresničevanje podatkovne pravičnosti je potrebno sodelovanje med aktivisti, ki so proti nadzoru, in tistimi za socialno pravičnost; v svoji raziskavi ugotavljajo, da je tega vse premalo in da med aktivisti ni stika: prvi zagovarjajo pravice in svobodo ter se ne ukvarjajo z nadzorom, drugi sledijo tehnološkim razsežnostim nadzora in odpora, vendar se v tem boju ne povežejo z drugimi aktivisti (Dencik idr., 2016: 10).

Taylor (2019: 8–10) je podatkovno pravičnost opredelil s tremi stebri: (1) vidnost, (2) digitalna (ne)vklučenost in (3) nasprotovanje digitalni diskriminaciji. Vidnost se nanaša na pravici biti zastopan in imeti informacijsko zasebnost. Pri tem je treba upoštevati zaščito zasebnosti na individualni ravni ter skupinsko zasebnost, ki je lahko ogrožena s tehnikami oblikovanja skupinskih profilov. Digitalna vključenost se nanaša na pravico uživanja prednosti, ki izhajajo iz podatkov, digitalna nevklučenost pa zagotavlja avtonomijo pri tehnološki izbiri. Avtonomija pomeni, da posamezniki niso prisiljeni v rabo točno določenih tehnologij, če želijo dostopati do določenih storitev ali produktov. Del avtonomije je tudi pravica do »biti izbrisan«. Tretji steber je nediskriminacija, ki se nanaša na zmožnost identifikacije in spopada s pristranostmi pri obravnavi in rabi podatkov ter na preprečevanje diskriminacije. Kroll idr. (2017: 636) opozarjajo, da se z razvojem vse bolj kompleksnih tehnologij, temelječih na strojnem učenju (ang. *machine learning*), zmanjšujejo možnosti za prepoznavo pristranosti, ki izhajajo iz velikih podatkovnih sistemov in algoritmov, zato je soočenje s tem problemom nujno in iskanje pravnih rešitev v interesu javnosti. Johnson (2014) opozarja še na naslednje vidike, ki so pomembni za uresničevanje podatkovne pravičnosti: (1) opredeliti je treba načine uporabe podatkov, ki podpirajo dominacijo in neenakosti v družbi; (2) osredotočiti se na tehnologije, ki omogočijo večjo distributivno pravico s tem, ko revne naredijo vidne; (3) proučevati vpliv podatkovnega nadzora na delo organizacij s področja socialne pravičnosti.

Caroll idr. (2019: 3–6) poudarjajo pomen podatkovne neodvisnosti (ang. *data sovereignty*) kot pravico manjšin, da upravljajo z lastnimi podatki, jih sami uporabljajo in so njihovi lastniki. Za staroselce v ZDA je podatkovna

neodvisnost pomembna za uresničevanje podatkovne pravičnosti. Ostanke kolonializma iz preteklosti se v dobi podatkov kažejo v kolonialnem prilaščanju podatkov o skupinah, ki so v zgodovinskih procesih izgubile neodvisnost zaradi aktivnosti kolonialistov. Podatkovno neodvisnost dosežemo le, ko posameznik in skupina neodvisno upravljata s svojimi podatki (Caroll idr., 2019: 3–6; Taylor, 2017).

Vera idr. (2019: 1017) poudarjajo, da zbiranje podatkov s strani držav in podjetij deluje znotraj širših nadzornih sistemov kapitalizma, ki zagotavljajo dominacijo, zato aktiviste s področja podatkovne pravičnosti skrbi obsežen in rasističen nadzor. Nasprotno si aktivisti s področja okoljske pravičnosti prizadevajo za še več podatkov v imenu »pravice vedeti«, saj podatke običajno vidijo kot »vrednostno nevtralne objekte«.

Podatkovni aktivizem in možnosti za spremembe

Ena najdlje trajajočih oblik aktivizma je tekma za zemljevide in druge kartografske vizualizacije. Spomnimo se konfliktov imperialnih sil v odnosu do staroselcev – zemljevide in druge podatkovne baze so izdelovalci pripravljali za imperialne sile z namenom izkoriščanja surovin zase, ne glede na škodo, ki so jo povzročili staroselcem. Prve zemljevide so osvajalci pridobili od staroselcev, ki so zemljo poznali skladno s svojimi potrebami, znanji, prepričanju in verovanji. Zemlja ni bila lastniška, ampak so si jo delili. Osvajalci so kasneje ustvarili svoje zemljevide, s katerih so izbrisali staroselska znanja in ustvarili nikogaršnjo zemljo, ki so jo pretvorili v lastniški prostor za izkoriščanje naravnih virov. V nadaljevanju so staroselce z nikogaršnje zemlje preselili v rezervate (Kidd, 2019: 957–958). Poveden primer je iz Kanade, kjer z zakoni in davčnimi olajšavami na škodo staroselcev podpirajo ekstrakcijske industrije, ključen gradnik neoliberalnega kapitalizma. Izčrpavanje neobnovljivih virov spremlja odpor staroselskih, okoljskih in drugih gibanj. Nancy Peluso (leta 1995) opredeli posebno vrsto podatkovnega aktivizma (ang. *counter mapping*), s katerim se staroselci upirajo uradnim zemljevidom, ki ne upoštevajo njihove zgodovine in interesov. Ta oblika podatkovnega aktivizma ustvarja znanje o svetu, ki nasprotuje prevladujočim reprezentacijam, osvetluje neenakosti in nepravičnost ter ponuja alternativne poglede (Kidd, 2019: 955).

Praksa kaže, da se vprašanje podatkovne pravičnosti pogosto pojavi v primerih, ko obravnavamo okoljsko pravičnost. Marginalizirane skupine, ki živijo na (družbenih) obrobjih in so podvržene okoljski degradaciji ali slabim

okoljskim razmeram, pogosto nimajo dostopa do velikih količin podatkov, s katerimi bi lahko dokumentirale okoljsko neenakost in okoljski razizem.¹¹ Vera idr. opozarjajo na t. i. helikopterske znanosti; s tem pojmom označujejo znanstvenike in strokovnjake, ki v imenu znanstvenega raziskovanja vzamejo podatke iz okolja, jih analizirajo v svojih laboratorijih in rezultate uporabijo za znanstvene članke, ne pa tudi za dobrobit skupnosti. Zato so prepričani, da je treba take načine zbiranja podatkov nadomestiti z drugimi, pri čemer pripadniki skupnosti, v kateri raziskava poteka, v njej aktivno sodelujejo, in to kot raziskovalni kolegi v vseh fazah raziskovanja – zasnova raziskave, zbiranje podatkov, analiza in priprava rezultatov. Da bi premostili primanjkljaj, je treba oblikovati spletne skupnosti, v katerih aktivisti skupaj s tehnologi in raziskovalci razvijajo orodja, s katerimi lahko skupnosti same zbirajo okoljske podatke na cenovno dostopne načine (Vera idr., 2019: 1015).¹²

Znanost, podprta s pravičnimi podatki

V dobi podatkov se tudi raziskovalna področja med seboj razlikujejo po tem, kako dobro, hitro in spretno znajo ravnati s podatki, kakšne količine podatkov analizirajo, od kod jih pridobijo, kje jih naredijo dostopne za druge – če sploh – in za kakšne namene jih uporabijo.

Analiziranje velikih podatkov je videti še posebej atraktivno. Van Dijck ugotavlja, da velike podatke, izžete z družbenih omrežij, v znanosti prikazujejo kot pridobljene z zelo natančno in izčrpno metodo za merjenje vsakdanjih interakcij. Veljajo celo za več vredne od podatkov, pridobljenih na vzorcih, in zanesljivejše od podatkov iz intervjujev in anket. Veliki sklopi »neurejenih« podatkov nadomeščajo manjše sklope vzorčenih podatkov, in kot pravijo zagovorniki velikih podatkov, velikost podatkovnih nizov kompenzira njihovo neurejenost (van Dijck, 2014: 199). Danes za nosilce novega znanja veljajo podatkovni znanstveniki, specializirani za kvantitativne pristope. Fuchs opozarja, da pri tovrstnih raziskovalnih pristopih manjka poznavanje uporabnikove subjektivnosti, njegovih izkušenj, norm, vrednot in interpretacij ter umeščenost družbenih medijev v strukture moči in družbene boje. Poudarja, da potrebujemo paradigmski premik iz digitalnega pozitivizma v smeri kritičnega raziskovanja družbenih medijev (Fuchs, 2017: 40).

¹¹ Vprašanje dostopa do podatkov je sestavni del okoljskega boja tudi na slovenskih tleh, o tem glej na primer Weiss, 2020; Bezjak in Zaviršek, 2020.

¹² Spletna skupnost Public Lab ponuja orodja in načine za zbiranje okoljskih podatkov s področja kmetijstva, kakovosti zraka, podnebni sprememb, rabe zemlje, mikroplastike, hrupa, prometa, odpadkov, mokrišč, kakovosti vode in drugo (<https://publiclab.org>).

Drug poveden primer je s področja arheologije, na katerem je lidarska metoda zbiranja podatkov prinesla nove uvide. Ker gre za drago metodo, se veliko razpravlja o vprašanju dostopa do lidarskih podatkov – raziskovalci, regije ali države dostop do podatkov omejujejo z argumentom zaščite najdišč pred izropanjem. S tem pa odpirajo prostor za manipulacije in nove oblike kolonializma. Da bi se izognili novemu kolonializmu, morajo arheologi sodelovati z lokalnimi skupnostmi, vladnimi agencijami in ljudmi, ki živijo ali delajo na najdiščih; to še posebej velja za tiste, ki prihajajo iz drugih držav. Poleg tega morajo izoblikovati smernice, ki bodo preprečevale manipulacije pri dostopu do podatkov (Chase idr., 2020: 58).

Številni avtorji (Taylor, 2017; Carroll idr., 2019; Chase idr., 2020) pri opisu dobrih praks v smeri podatkovne pravičnosti in neodvisnosti izpostavljajo *akcijsko raziskovanje*. Gre za emancipatorno obliko raziskovanja, kjer s stroko sodelujejo tudi udeleženci raziskave, ki so hkrati partnerji raziskave in ne le pasivni udeleženci. To pomeni, da z raziskovanjem hkrati že krepimo moč vključenih, saj je raziskovalna metoda odprta za dialog in komunikacijo med vsemi sodelujočimi. Takšno raziskovanje, ki služi reševanju praktičnih problemov (Mesec, 1994), izhaja iz konkretnih potreb in je zato postavljeno v praktične okoliščine, neobremenjene s teoretskimi okviri. Aktivno vključevanje udeležencev raziskave je pot k podatkovni pravičnosti. Raziskovalec je povezan z udeleženci in skupnostjo, člani skupnosti sodelujejo kot raziskovalni kolegi v vseh raziskovalnih fazah, od zasnove raziskave in zbiranja podatkov do analiziranja s poudarkom na součenju. V dobi podatkov lahko upodatkovljenje uporabimo za emancipacijo posameznikov in skupnosti in ne le za akumulacijo kapitala. Tak primer je pobuda Data 4 Black Lives (d4bl, n. d.), v okviru katere skuša skupina aktivistov s pomočjo podatkovne znanosti oblikovati konkretne merljive spremembe za izboljšanje življenja temnopoltih ljudi. Menijo, da lahko podatkovni sistemi, ki so bili zgodovinsko uporabljeni kot orodje zatiranja in diskriminacije, opolnomočijo temnopolte ljudi (Vera idr., 2019: 1016).

Obstaja vse več razlogov, ki kličejo h kritičnemu premisleku o rabi podatkovnih prijemov v znanosti. Goldkind idr. poudarjajo, da so *majhni podatki* priložnost za veliko pravičnost; po njihovi oceni je v središču majhnih podatkov človek. Še več prednosti vidijo v *heterogenosti*¹³ majhnih podatkov, ki izhajajo iz *konkretnih okoliščin*,¹⁴ ko poznamo *poreklo podatkov*¹⁵ in kadar njihova uporaba služi izboljšanju življenja posameznikov (Goldkind idr., 2019: 175).

13 Pridobljeni so iz različnih virov, so različni, imajo različne pomeni.

14 Poznamo sociokulturni kontekst, v katerem so nastali.

15 Vemo, od kod prihajajo in kakšno je lastništvo v njihovem celotnem življenjskem krogu.

Gibanje v smeri *odprte znanosti* in *odprtih podatkov* lahko tlakuje pot demokratske rabe podatkov, vendar le, če v znanstveni skupnosti premišljujemo o podatkovnih prijemih ter njihovem pomenu za posameznike in širšo skupnost. Interesna skupina pri Research Data Alliance (skupina RDA) opozarja, da trenutni trendi k odprti znanosti in odprtim podatkom niso usklajeni s pravicami in interesi družbenih manjšin, saj se načela FAIR¹⁶, po katerih deluje gibanje za odprte podatke, osredotočajo na (predvsem tehnične) lastnosti podatkov, hkrati pa ignorirajo oziroma ne upoštevajo razlik v moči in (zgodovinskih) okoliščin nastanka teh podatkov. Deljenje podatkov ustvarja napetosti, saj si manjšine prizadevajo, da bi pridobile nadzor nad podatki, ki so osnovani na podlagi njihovih prepričanj oziroma njihovega pogleda in vsebujejo informacije o njih (RDA International Indigenous Data Sovereignty Interest Group, 2019).

Da bi zagotovili podatkovno pravičnost, skupina RDA (2019) predlaga pripravo podatkov za odprti dostop po t. i. načelih CARE, ki so naravnana k ljudem, saj upoštevajo ključno vlogo podatkov pri izboljšanju položaja manjšin. Skladno s temi načeli morajo podatki služiti skupnemu dobremu (ang. *Collective benefit*), manjšine morajo nad njimi imeti pooblaščen nadzor (ang. *Authority to control*), z njimi pa morajo vsi udeleženci ravnati odgovorno (ang. *Responsibility*) in etično (ang. *Ethics*). Podatkovni sistemi morajo biti zasnovani tako, da bodo koristili vključevanju manjšin in podpirali njihove vrednote, obogatili (na primer okoljska) načrtovanja ter služili njihovim težnjam po boljših življenjskih razmerah. Interesi in pravice manjšin glede podatkov morajo biti priznani s strani drugih, prav tako morajo manjšine imeti nadzor nad podatki o sebi oziroma dostop do podatkov, ki so pomembni za njihov razvoj. Vsi, ki delajo s podatki o manjšinah, so odgovorni, da s podatki in podatkovnimi prijemi vplivajo na izboljšanje življenjskih razmer manjšin. Prav tako morajo podatki vključevati etično komponento, kar pomeni, da ne smejo stigmatizirati ali biti neresnični. T. i. načela CARE tako dopolnjujejo načela FAIR in bodo, ko bodo postala norma in raziskovalni standard, v podporo podatkovni pravičnosti.

Vlogo v procesu uresničevanja podatkovne pravičnosti imamo tudi arhivi podatkov in druge podporne storitve v raziskovalnem okolju. IASSIST, ameriško združenje informacijskih storitev in tehnologije s področja družboslovja, je v času velikega upora »*Black lives matter*«, ki ga je leta 2020 sprožil uboj Georgea Floydja v Minneapolisu, ZDA, pripravilo uradno stališče, v katerem sporoča:

16 F – Findable, A – Accessible, I – Interoperable, R – Reusable, kar v slovenščino prevajamo kot F – najdljivi, A – dostopni, I – interoperabilni oziroma združljivi z drugimi podatki, R – ponovno uporabni.

Razumemo, da zbiranje, opredeljevanje in uporaba podatkov v veliki meri povzroča ravno nasprotno od pravičnosti, enakosti in vključevanja, s čimer se nadaljuje utrjevanje systemskega izključevanja manjšin. Na drugi strani lahko podatki služijo kot izjemno orožje pri rušenju trenutnega stanja neravnovesja ter vodijo v smer podatkovne in družbene pravičnosti. To naj bo izziv za sedanje in prihodnje generacije.¹⁷

Izjava vključuje tudi zavezo, da bodo dejavno nastopili proti rasizmu in postali protirasistična organizacija. K razmisleku o obravnavi podatkov, povezanih z manjšinami in drugimi marginaliziranimi skupinami, pozivava tudi slovensko raziskovalno skupnost.

Literatura

- Aaron, Martin in Linnet Taylor (2020): Exclusion and Inclusion in Identification: Regulation, Displacement and Data Justice. *Information Technology for Development* 27(1): 50–66. Dostopno na DOI: 10.1080/02681102.2020.1811943.
- Bezjak, Sonja in Darja Zaviršek (2020): Ženske so reko varovale podnevi, moški ponoči: okoljska pravičnost na področju hidroenergije. *Časopis za kritiko znanosti* 279: 216–240.
- Calzati, Stefano (2020): Decolonising “Data Colonialism”. Propositions for Investigating the Realpolitik of Today’s Networked Ecology. *Television & New Media*. Dostopno na DOI: 10.1177/1527476420957267.
- Carroll, Stephanie, Desi Rodriguez-Lonebear in Andrew Martinez (2019): Indigenous Data Governance: Strategies from United States Native Nations. *Data Science Journal* 18(31): 1–15. Dostopno na: <https://datascience.codata.org/articles/10.5334/dsj-2019-031/> (27. januar 2021).
- 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. Dostopno na: <https://journal.caa-international.org/articles/10.5334/jcaa.43/> (27. januar 2021).
- Couldry, Nick in Ulises A. Mejias (2019): Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject. *Television & New Media* 20(4): 336–349. Dostopno na DOI: 10.1177/1527476418796632.
- Couldry, Nick in Ulises A. Mejias (2019a): Making Data Colonialism Liveable: How Might Data’s Social Order Be Regulated? *Internet Policy Review* 8(2): 1–16. Dostopno na DOI: 10.14763/2019.2.1411.

¹⁷ Izjavo smo prejeli kot naročniki novic združenja IASSIST, posredovana je bila na discuss@lists.iassistdata.org, 22. 9. 2020, z naslovom: Racism, Data, and IASSIST.

- Data for Black Lives*. Dostopno na: <https://d4bl.org/programs.html> (15. oktober 2020).
- Dencik, Lina, Arne Hintz in Jonathan Cable (2016): Towards Data Justice? The Ambiguity of Anti-Surveillance Resistance in Political Activism. *Big Data & Society* 3(2): 1–12. Dostopno na DOI: 10.1177/2053951716679678.
- Fuchs, Christian (2017): From Digital Positivism and Administrative Big Data Analytics Towards Critical Digital and Social Media Research! *European Journal of Communication* 32(1): 37–49. Dostopno na DOI: 10.1177/0267323116682804 (10. januar 2021).
- Goldkind, Lauri, Mamello Thinyane in Moon Choi (2018): Small Data, Big Justice: The Intersection of Data Science, Social Good, and Social Services. *Journal of Technology in Human Services* 36(4): 175–178. Dostopno na DOI: 10.1080/15228835.2018.1539369.
- International Association for Social Science Information Service & Technology (2020): *Racism, Data and IASSIST*. Dostopno na: <https://iassistdata.org/blog/2020/09/13/racism-data-and-iassist/> (13. september 2020).
- Jin, Dal Yong (2013): The Construction of Platform Imperialism in the Globalization Era. *TripleC* 11(1): 145–172. Dostopno na DOI: 10.31269/triplec.v11i1.458.
- Kidd, Dorothy (2019): Extra-Activism: Counter-Mapping and Data Justice. *Information, Communication & Society* 22(7): 954–970. Dostopno na DOI: 10.1080/1369118X.2019.1581243.
- Kroll, Joshua A., Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson in Harlan Yu (2017): Accountable Algorithms. *University of Pennsylvania Law Review* 165: 633–705. Dostopno na: https://scholarship.law.upenn.edu/penn_law_review/vol165/iss3/3 (27. januar 2021).
- Mesec, Blaž (1994): Model akcijskega raziskovanja. *Socialno delo* 33(1): 3–16.
- Mosco, Vincent (2014): *To the Cloud: Big Data in a Turbulent World*. Boulder, CO: Paradigm.
- Nathan, Newman (2014): *How Big Data Enables Economic Harm to Consumers, Especially to Low-Income and Other Vulnerable Sectors of The Population*. Dostopno na: <https://perma.cc/VB4Y-53SR> (1. april 2017).
- 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).
- Taylor, Linnet (2017): What Is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally. *Big Data & Society*: 1–14. Dostopno na DOI: 10.1177/2053951717736335.
- Thatcher, Jim, David O'Sullivan in Dillon Mahmoudi (2016). Data Colonialism through Accumulation by Dispossession: New Metaphors for Daily Data. *Environment and Planning D: Society and Space* 34(6): 990–1006. Dostopno na DOI: 10.1177/0263775816633195.

- Van Dijck, José (2014): Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology. *Surveillance & Society* 12(2): 197–208. Dostopno na DOI: 10.24908/ss.v12i2.4776.
- Vera, Lourdes A., Dawn Walker, Michelle Murphy, Becky Mansfield, Ladan Mohamed Siad, Jessica Ogden in EDGI (2019): When Data Justice and Environmental Justice Meet: Formulating a Response To Extractive Logic Through Environmental Data Justice. *Information, Communication, Society* 22(7): 1012–1028. Dostopno na DOI: 10.1080/1369118X.2019.1596293. <https://www.tandfonline-com.nukweb.nuk.uni-lj.si/doi/full/10.1080/1369118x.2019.1596293>
- Weiss, Monika (2020): Šibitev okoljskega boja v Sloveniji – oris dveh primerov. *Časopis za kritiko znanosti* 279: 241–261.