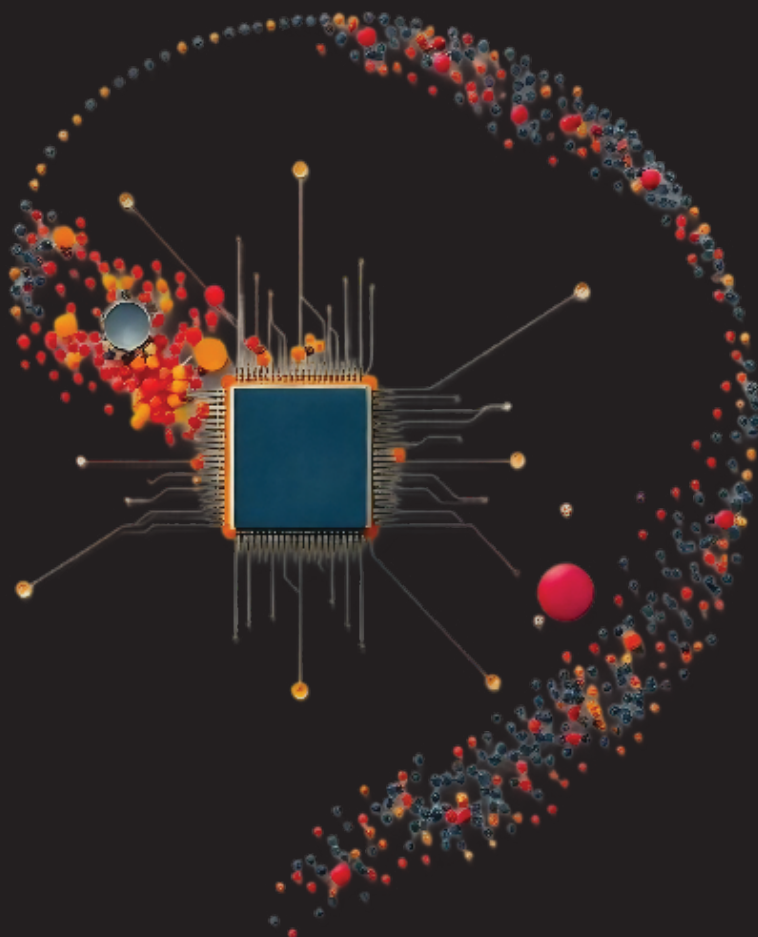


OD TRŽNE KONKURENČNOSTI TEHNOLOGIJE K DRUŽBENI TEHNOLOŠKI ZRELOSTI

Urednik Slavko Splichal



OD TRŽNE KONKURENČNOSTI TEHNOLOGIJE K DRUŽBENI TEHNOLOŠKI ZRELOSTI

Urednik Slavko Splichal



Ljubljana, 2025

SPREJETO NA SEJI PREDSEDSTVA
SLOVENSKE AKADEMIJE ZNANOSTI IN UMETNOSTI
DNE 28. JANUARJA 2025

SPLICHAL, Slavko (urednik)

Od tržne konkurenčnosti tehnologije k družbeni tehnološki zrelosti

Izdajatelj:

Slovenska akademija znanosti in umetnosti
Fakulteta za družbene vede, Založba FDV

Recenzenta: Igor Emri in Olga Markič

Jezikovni pregled: Megamoment, Murska Sobota

Naslovnica: Peter Maček z uporabo umetne inteligence

Prelom: Karmen Zahariaš

Ljubljana, 2025



To delo je ponujeno pod licenco Creative Commons Priznanje
avtorstva - Deljenje pod enakimi pogoji 4.0 Mednarodna licenca.

1. e-izdaja. Knjiga je v digitalni obliki prosto dostopna prek
<https://ebooks.uni-lj.si/ZalozbaUL>

DOI: 10.51936/9789612681197

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in
univerzitetni knjižnici v Ljubljani
COBISS.SI-ID 251272963
ISBN 978-961-295-129-0 (Fakulteta za družbene vede, PDF)

VSEBINA

Predgovor <i>Igor Emri</i>	i
Uvod: med tehnologizacijo družbe in podružbljanjem tehnologije <i>Slavko Splichal</i>	1
Kako doseči tehnološko zrelost? Od instrumentalnega k reflektivnemu ocenjevanju tehnologij <i>Slavko Splichal</i>	9
Družbeni konteksti tehnološkega razvoja <i>Mirjana Nastran Ule</i>	31
Vloga in pomen znanstveno-tehnološke transdisciplinarnosti pri vrednotenju razvojnih potencialov novih naprednih tehnologij <i>Franc Mali</i>	43
Etične razsežnosti umetne inteligence <i>Ana Farič in Ivan Bratko</i>	57
Vpliv generativne umetne inteligence na demokratične procese <i>Lea Košmrlj in Ivan Bratko</i>	83
Učinek tehnologij pri upravljanju dediščine: nedenarno vrednotenje <i>Lana Nastja Anžur in Matija Strlič</i>	99
Spodbujanje inovacij za družbeno in tehnološko preobrazbo: širša perspektiva <i>Franci Pušavec</i>	119
Pametna specializacija v Sloveniji – tehnološki in družbeni izziv <i>Andreja Jaklič</i>	127
K trajnostnemu razvoju platformne ekonomije: izkušnje z globalnega juga <i>Aleksandra Kanjuo Mrčela</i>	145
Deliberativna poizvedba kot demokratična tehnologija <i>Matjaž Maček</i>	163
Socialna tehnologija in univerzalni temeljni dohodek <i>Srečo Dragoš</i>	189

PREDGOVOR

Vprašanja razvoja, vrednotenja in usmerjanja tehnologij postajajo v sodobni družbi vse pomembnejša ne le za znanstveno skupnost, temveč za celotno družbo. V zadnjih desetletjih so se tehnologije – še posebej digitalne – razvijale z izjemno hitrostjo, pogosto hitreje, kot jih je družba sposobna regulativno, etično in kulturno osmisliti ali obvladovati. Paradoks sodobne tehnološke družbe je, da je na voljo vedno več tehnoloških rešitev, a hkrati vedno manj refleksije o tem, katerim problemom naj bi bile te rešitve namenjene in kdo jih definira. Ključno vprašanje ni več, kaj lahko naredimo s tehnologijo, temveč kako in zakaj jo razvijamo ter kdo ima dostop do odločanja o tem razvoju.

Ob tem se zastavlja vprašanje, kako ocenjujemo tehnologije, še preden vstopijo v širšo uporabo ali postanejo del sistemskih rešitev. Standardizirani pristopi, kot je lestvica pripravljenosti tehnologije (*»Technology Readiness Levels«* – TRL), ki jo je razvila ameriška vesoljska agencija NASA in jo danes pogosto uporabljajo evropske raziskovalne politike, omogočajo preprosto klasifikacijo zrelosti tehničnih rešitev – od konceptualne ideje (TRL 1) do popolne industrijske uporabe (TRL 9). Taka klasifikacija ima svojo uporabno vrednost zlasti v inženirskih disciplinah, kjer je cilj prenosa znanstvenih dognanj v prakso relativno jasen in kvantificiran. Vendar pa ta model predpostavlja linearen, postopno napredujoč razvoj tehnologije, ki se giblje od raziskovalnega laboratorija k trgu, in pri tem pogosto zanemara ali celo izključuje kompleksnost, večdimenzionalnost in družbeno pogojenost tehnoloških inovacij.

Znanstvene raziskave že dolgo opozarjajo, da tehnologije niso zgolj nevtralna orodja, temveč so tudi družbeno, politično, kulturno in ekonomsko določene. Današnje tehnologije, še posebej tiste, ki temeljijo na podatkovnih ekonomijah, umetni inteligenci ali avtomatizaciji, oblikujejo načine življenja, razmerja moči in celo strukture percepcije sveta. Ko digitalne platforme postajajo infrastruktura za informacijsko posredovanje, politično komunikacijo in ekonomsko izmenjavo, ni več dovolj, da jih ocenjujemo zgolj glede na njihovo tehnično dovršenost ali tržno uspešnost. Potrebujemo celovite pristope vrednotenja, ki vključujejo tudi vprašanja demokratične odgovornosti, dostopnosti, dolgoročne vzdržnosti ter družbenih in kulturnih posledic.

Ravno v ta kontekst se umešča zbornik. Avtorji in avtorice iz različnih disciplin – od tehniških in naravoslovnih ved do družboslovja in humanistike – opozarjajo, da zdajšnje metodologije za vrednotenje tehnologij pogosto ne zajamejo njihove kompleksnosti. Lestvica TRL, na primer, ne ponuja meril za presojo družbene sprejemljivosti, ekološke trajnosti, etičnih tveganj ali kulturnih učinkov tehnoloških rešitev. Še več, v nekaterih primerih lahko taka merila celo prispevajo k marginalizaciji tehnologij, ki ne sledijo logiki trga, a so ključnega pomena za dolgoročno družbeno odpornost, denimo, digitalne rešitve za skupnostno upravljanje, odprtokodne infrastrukture, participativne platforme ali sistemi za vključevanje ranljivih skupin.

Zbornik ni samo kritika zdajšnjih metod vrednotenja (kot je lestvica TRL), temveč tudi odpira prostor za drugačne pristope – npr. participativne metode vrednotenja, večkriterijske analize, vključevanje deležnikov v zgodnje faze raziskovanja ter kvalitativne ocene družbene in okoljske smiselnosti tehnoloških projektov. V zborniku so izpostavljene

ni primeri tako imenovanih družbenih in sistemskih tehnologij, ki jih standardni okvirji pogosto spregledajo. Gre za tehnologije, ki niso nujno materialne ali industrijske, temveč delujejo kot organizacijski modeli, procesni pristopi, institucionalne inovacije ali kulturni vzorci sodelovanja. Primeri vključujejo modele participativnega oblikovanja politik, algoritmične strukture za pravično razporejanje virov, metode deliberativne demokracije in tudi izobraževalne pristope, ki temeljijo na odpiranju znanja skupnosti. Take nematerialne tehnologije pogosto nastajajo na presečišču znanosti, umetnosti, civilne družbe in lokalnih skupnosti ter zahtevajo drugačne, pogosto kvalitativne pristope vrednotenja, kot so študije vpliva, etnografije rabe ali participativne ocene.

Tovrstna širitev razumevanja tehnologije je posebej relevantna tudi za slovenski kontekst. Resolucija o znanstvenoraziskovalni in inovacijski strategiji Slovenije 2030 opozarja na pomen odgovornega raziskovanja in inoviranja, kar vključuje spoštovanje načel odprtosti, etike, vključevanja javnosti in trajnostnega razvoja. Tudi Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost (ARIS) uvaja instrumente, ki odpirajo prostor za bolj interdisciplinarne in reflektirane pristope k ocenjevanju raziskovalnih projektov. Kljub temu institucionalne prakse pogosto še vedno sledijo paradigmi kvantifikacije, kar otežuje uveljavitev alternativnih meril vrednotenja, denimo, družbene vrednosti, participativnosti ali ekološke regenerativnosti tehnologij.

Slovenska akademija znanosti in umetnosti si kot najvišja znanstvena in umetniška ustanova v državi prizadeva prispevati k poglobljenemu razmisleku o teh vprašanjih. Svet za razvoj SAZU, ki povezuje predstavnike različnih znanstvenih področij, si prizadeva za celovit in dolgoročen pogled na znanost in na njej utemeljen družbeni razvoj, ki vključuje tudi zavedanje o družbenem kontekstu in vplivu znanosti. Verjamemo, da mora imeti znanost vlogo ne le v razvoju tehnologij, temveč tudi v njihovi refleksiji in usmerjanju, zato v tematiki tega zbornika vidimo pomemben prispevek k oblikovanju celovitejše raziskovalne in inovacijske politike. Dolžnost akademske skupnosti je, da ne le razvija nova znanja in tehnologije, temveč tudi pomaga oblikovati merila, po katerih bomo kot družba presojali, katere tehnologije razvijamo, za koga in s kakšnim družbenim učinkom.

Zato izid tega zbornika pozdravljam kot pomemben korak v smeri reflektirane, pluralne in družbeno odgovorne tehnološke politike in vrednotenja tehnološkega razvoja. Lahko je izhodišče za nadaljnjo razpravo – tako v znanstveni kot v širši javnosti – o tem, kako razvijati tehnologije, ki ne prispevajo le h gospodarski rasti, temveč tudi k blaginji, enakosti in ohranjanju skupnih dobrin. Ne ponuja enoznačnih odgovorov, temveč odpira prostor za razpravo in za skupno oblikovanje novih meril – takih, ki bodo ustrezala kompleksnosti izzivov, s katerimi se kot družba soočamo. Vabi nas, da znova premislimo razmerje med tehnološkim napredkom in družbenim blagostanjem ter med znanstveno odličnostjo in javno odgovornostjo.

Verjamem, da bo ta zbornik spodbudil nadaljnje strokovne, politične in javne razprave o tem, katere tehnologije potrebujemo v prihodnosti – ne le tiste, ki so tržno uspešne, temveč predvsem tiste, ki krepijo družbeno pravičnost, ekološko ravnovesje in demokratično kulturo.

Igor Emri
predsednik Sveta za razvoj
Slovenske akademije znanosti in umetnosti

UVOD: MED TEHNOLOGIZACIJO DRUŽBE IN PODRUŽBLJANJEM TEHNOLOGIJE

Slavko Splichal

Vprašanje, kako vrednotimo tehnologije, ni le metodološko ali aplikativno vprašanje – gre za temeljno znanstveno, politično in etično vprašanje, ki posega v jedro našega odnosa do razvoja, pojmovanja napredka in vizije prihodnosti skupnosti. Tehnologije danes niso zgolj orodja, temveč postajajo osrednji oblikovalci načinov življenja, sistemov moči in oblik sobivanja. V zadnjih desetletjih smo priča vse hitrejši tehnološki preobrazbi, ki vpliva na vse vidike življenja – od delovnih procesov, izobraževanja, zdravstva in komuniciranja do vsakdanjih praks in predstav o sebi in družbi. Tehnološke odločitve vse bolj oblikujejo naš skupni svet, zato jih ne moremo prepustiti zgolj inženirskim presojam, tržnim silnicam ali političnim interesom. Potrebujemo širši družbeni konsenz o tem, katere tehnologije razvijamo, kako jih uporabljamo, v čigavo korist, s kakšnimi posledicami in znotraj katerih vrednotnih okvirov.

Med zgodnjimi opozorili o nujnosti družbenokritičnega premisleka o tehnološkem razvoju ima vidno mesto poročilo *Meje rasti* (1972), ki ga je za Rimski klub pripravila skupina znanstvenikov z MIT. Njegov zgodovinski pomen je v tem, da gre za enega prvih poskusov celovitega vrednotenja tehnološkega in gospodarskega razvoja z upoštevanjem njegovih dolgoročnih učinkov v planetarnih razsežnostih. To poročilo je zaznamovalo začetek sodobnega okoljskega gibanja, vzpodbudilo nastanek konceptov, kot sta trajnostni razvoj in okoljska etika, ter vplivalo na nastanek političnih dokumentov, kot sta na primer Agenda 21 s konference OZN o okolju in razvoju v Riu (1992) in Pariški sporazum iz leta 2015.

Avtorji poročila *Meje rasti* so na podlagi systemske analize in računalniških simulacij opozorili, da nebrzdano sledenje logiki neskončne rasti prebivalstva, industrializacije, onesnaževanja, pridelave hrane in izčrpavanja virov, ki jo spodbuja tehnološki razvoj – ki jo pogosto poganja ravno tehnološki napredek – vodi v okoljski in družbeni kolaps. Ključno sporočilo ni bilo protitehnološko, temveč refleksivno: razvoj brez vrednotenja ni nevtralen, temveč lahko sproži procese degradacije okolja, izčrpavanja virov in družbene destabilizacije. S tem je poročilo postavilo okvir za razumevanje tehnologije ne le kot sredstva za doseg gospodarskih ciljev, temveč kot dejavnika, ki soustvarja pogoje skupnega življenja, vpliva na porazdelitev virov in moči ter potencialno ogroža trajnost civilizacije.

Čeprav so številni ekonomisti in politiki poročilu očitali maltuzijanstvo, tehnološki pesimizem in determinizem, so poznejše študije pokazale presenetljivo ujemanje dejanskih globalnih trendov s projekcijami izvirnega modela. Postajalo je jasno, da bo brez bistvenih sprememb v političnem odločanju in vedenjskih vzorcih človeštvo v nekaj desetletjih doseglo planetarne meje, kar bi vodilo v resne družbene in okoljske

pretrese. A poročilo ni bilo apokaliptično, predstavilo je tudi scenarije stabilizacije in ravnotežja, ki bi jih bilo mogoče uresničiti s pravočasnim ukrepanjem.

Desetletje po *Mejah rasti* je izšlo poročilo Rimskega kluba *Mikroelektronika in družba – v dobro ali zlo*, ki je nakazalo in kritično ocenjevalo pričakovane dobre in slabe družbene posledice razvoja »elektronike«, ki jih danes obravnavajo kritične razprave o digitalizaciji in umetni inteligenci, ter doživelo podobno kritiko kot *Meje rasti*. Tudi sodobna kritična spoznanja in alarmantna opozorila okoljskih, podnebnih in družboslovnih znanstvenikov, da globalizacija ni razrešila nekaterih osnovnih razvojnih problemov, na katere je opozarjal že Rimski klub v obeh znamenitih poročilih, niti ni izpolnila optimističnih pričakovanj o globalni demokratični informacijski družbi, se pogosto srečujejo s podobnimi kritikami in z zavračanjem znanstvenih opozoril, kot sta jih bili iz vrst intelektualnih podpornikov močnih gospodarskih interesov deležni poročili *Meje rasti* ter *Mikroelektronika in družba – v dobro ali zlo*.

Poročila Rimskega kluba so v drugi polovici 20. stoletja okrepila tudi temeljne teoretske premisleke o naravi tehnologije, ki so bistveno prispevali k razumevanju njenih družbenih učinkov. Ideje Lewisa Mumforda in Jacquesa Ellula so postavile enega izmed ključnih temeljev sodobne kritične misli o tehnologiji kot avtonomni družbeni sili. Mumford je v knjigi *Technics and Civilization* (1934) razložil med dvema tipoma tehnološkega razvoja – »demokratičnim« in »avtoritarnim«. Prvi temelji na človeški meri, lokalni samoupravi in kulturni raznolikosti, drugi pa na centralizaciji, hierarhiji in instrumentalizaciji človeka. Tehnološki sistem po Mumfordu nikoli ni zgolj skupek orodij, temveč vedno odraža in sooblikuje vrednote, družbene cilje in razmerja moči. Podobno in še bolj kritično je Jacques Ellul v delu *La technique ou l'enjeu du siècle* (1954) – angleški prevod *The Technological Society* (1964) – opozoril, da tehnika ni več le sredstvo, temveč postaja samostojen dejavnik, ki se širi z lastno logiko učinkovitosti in avtomatizacije, ne glede na etične ali družbene posledice. Po Ellulu sodobna družba živi pod diktatom »tehnične nujnosti«, pri čemer odločitve o uvajanju novih tehnologij ne temelji na javni razpravi ali političnem soglasju, temveč zgolj na tehnični izvedljivosti in ekonomskih interesih.

Sodobni človek je žrtev prav tistih tehnologij, ki jih najbolj ceni. Problem ni le v tem, da so bile mnoge tehnologije, ki povečujejo človekovo moč, znanje in obvladovanje narave, zlorabljene za politične ali komercialne namene, pač pa tudi v tem, da kljub tej vedno prisotni nevarnosti razvoja tehnologij ni spremljal enak napredek v kritičnem razumevanju, razvijanju in ocenjevanju tehnologij. Ellul je iluzijo, da je tehnološki napredek sam po sebi dober, neizogiben in nevtralen, čeprav ima dejansko tudi številne negativne učinke in ustvarja lažen občutek obvladovanja, označil kot »tehnoški blef« – *Le bluff technologique*, kot je naslov njegovega zadnjega obširnega dela (1988), v katerem povzema in zaostrojuje svojo vseživljenjsko kritiko sodobne tehnološke družbe. »Tehnološko slepilo« pomeni sistematično samoprevaro družbe, ki verjame, da tehnologijo obvladuje, a je v resnici ujeta v logiko samopovečevanja, avtomatizacije in učinkovitosti. Ellul opozarja, da je to slepilo morda najnevarnejša ideološka prevara sodobnega časa, saj onemogoča resnično javno razpravo o družbenih ciljih, vrednotah ter alternativnih tehnoloških rešitvah in razvojnih poteh.

Ta opozorila so odmevala v delih teoretikov, kot sta Langdon Winner in Ernst Friedrich Schumacher, ki sta pokazala, da tehnologije niso nevtralna orodja z »žele-

nimi« in »neželenimi« učinki, temveč vedno že utelešajo družbene vrednote, normativne predpostavke in razmerja moči. Winner je na primer z vplivno razpravo *Do Artifacts Have Politics?* (1980) s primerom gradnje nizkih mostov v New Yorku v letih 1920–1970, ki so fizično preprečevali dostop avtobusom na Long Island in s tem izključevali revnejše prebivalstvo, pokazal, kako so tehnične rešitve lahko izraz politične volje, vtakane v infrastrukturo. V obravnavi tehnološkega razvoja kot rezultata družbenih pogajanj, interesnih konfliktov in kulturnih matric – ne pa kot linearne in samoumevne poti napredka – je Schumacher v delu *Small Is Beautiful: Economics As If People Mattered* (1973) razvil koncept »ustrezne tehnologije« (*appropriate technology*), ki je spodbudil institucionalizacijo vrednotenja tehnologij na podlagi lokalne ustreznosti, kulturnih razmer in okoljskih omejitev.

Ta in mnoga sorodna zgodovinska razmišljanja so danes še aktualnejša kot pred desetletji – ne le zaradi vse večje zaskrbljenosti ob podnebnih spremembah, izčrpavanju naravnih virov in poglobljanju družbenih neenakosti, temveč tudi zaradi vseobsegajoče digitalne preobrazbe, razvoja umetne inteligence in algoritmizacije vsakdanjega življenja. V tem kontekstu se krepi spoznanje, da tehnologij ne moremo vrednotiti zgolj po merilih tehnične učinkovitosti, konkurenčnosti ali tržne uspešnosti. V ospredje je treba postaviti celovitejšo presojo, ki vključuje vprašanja demokratične legitimnosti, etične odgovornosti, dolgoročne vzdržnosti in družbene pravičnosti.

Kritična dela Mumforda in Ellula ter njunih intelektualnih dedičev tako niso zgolj zgodovinski opomniki, temveč so tudi teoretski temelji za sodobno refleksijo o tem, kako naj bodo tehnologije namenjene ne le tehničnemu napredku, ampak človeku in skupnosti. Čeprav so vprašanja razvoja in vrednotenja tehnologij že dolgo v središču znanstvenega in političnega diskurza, ostajajo metode za ocenjevanje učinkovitosti, uporabnosti in zrelosti tehnologij večinoma zasidrane v okvirih, ki jih oblikujejo predvsem inženirske paradigme in tržni interesi. Medtem ko so se v zadnjih desetletjih s podnebnim prehodom, z vsepovsodno digitalizacijo, razvojem generativne umetne inteligence, bioinženiringom in ne nazadnje z militarizacijo tehnologije družbeni in etični izzivi tehnološkega napredka močno zaostri, prevladujoči kriteriji vrednotenja ostajajo omejeni na ekonomska merila, kot so konkurenčnost, tržna uspešnost in možnost implementacije.

Taka enodimenzionalna optika ne le podcenjuje širše posledice tehnologije za človeško življenje in planetarni ekosistem, temveč tudi zamegljuje pomen družbene in etičnega vrednotenja tehnološkega razvoja. Skozi prizmo ekonomistične logike ni mogoče razumeti tehnologij kot mehanizmov, ki reproducirajo globoke družbene razlike in konflikte – od razkola v digitalnih kompetencah in socialne segregacije do politične polarizacije in erozije demokratičnih institucij. V takih okvirih tudi ni mogoče razreševati tako imenovane Collingridgeve dileme, ki jo je David Collingridge predstavil v knjigi *The Social Control of Technology* (1980) kot ključni problem pri ocenjevanju tehnologije: po eni strani vplivov novih tehnologij v njihovi zgodnji razvojni fazi ni mogoče zanesljivo napovedati, po drugi strani pa postaneta nadzor ali spreminjanje tehnologije zelo težavna, ko postane tehnologija visoko razvita in široko uporabljena.

Prevladovanje instrumentalno-ekonomskih pristopov razkriva omejenost obstoječih modelov, ki pogosto zanemarjajo ključne razsežnosti razvoja: družbeno spre-

jemljivost, dolgoročno vzdržnost, pravičnost, dostopnost, vpliv na demokratične institucije ter skladnost z vrednotami javnega dobrega. Tehnološke inovacije vplivajo na družbena razmerja moči, ustvarjajo nove oblike izključevanja in digitalne neenakosti, obenem pa zarisujejo prihodnje možnosti – ali slepe ulice – družbenega razvoja.

* * *

Zbornik, ki je pred vami, je nastal kot odgovor na enodimenzionalnost ocenjevanja tehnologij – kot prispevek k vzpostavitvi bolj reflektivnega in večperspektivnega razumevanja tehnološkega razvoja. Skupno izhodišče je prepričanje, da mora biti tehnološki napredek usmerjen ne le v povečevanje produktivnosti in spodbujanje inovativnosti, temveč tudi v krepitev družbene solidarnosti, demokratične kulture ter ekološke odgovornosti. S tem pristopom se pomikamo od paradigme tržne konkurenčnosti k ideji družbene tehnološke zrelosti – ideji, ki v ospredje postavlja vrednotno utemeljenost, družbeni vpliv in dolgoročno smiselnost tehnoloških rešitev.

Tehnologije niso nevtralna orodja, temveč družbene tvorbe: utelešajo vrednote, normativne predstave in razmerja moči. Kar šteje za »napredno«, »zrelo« ali »pripravljeno«, ni odvisno le od tehničnih parametrov, temveč tudi od kulturnih kontekstov, institucionalnih interesov in političnih odločitev. Razprave o tehnološki zrelosti zato zahtevajo preseganje ozkih okvirov učinkovitosti, tržnosti in uporabnosti ter vključitev širših vprašanj: Kakšne tehnologije potrebujemo kot družba? Za koga jih razvijamo? Katere vrednote udejanjajo in katere izključujejo?

»Naprednost« ali »zrelost« tehnologij ni objektivna danost, temveč rezultat družbenega konteksta, institucionalnih odločitev in političnih interesov. V tem kontekstu je nujno razmišljati o konceptih, kot so družbena pravičnost, obča dostopnost, vključevanje marginaliziranih skupin in odgovornost do prihodnjih generacij. Ocenjevanje tehnologij mora zato temeljiti na normativni razsežnosti in vključevati razprave o tem, kakšno družbo si želimo in kakšno vlogo naj imajo tehnologije v njenem oblikovanju.

Zbornik odpira to razpravo na več ravneh. Osredinja se na nujnost preoblikovanja meril tehnološke zrelosti tako, da bodo sposobna zajeti tudi nematerialne tehnologije – družbene, politične, organizacijske, simbolne –, ki so ključnega pomena za obvladovanje kompleksnih izzivov sodobnosti. Od sistemov umetne inteligence do družbenih inovacij in participativnih mehanizmov v lokalni samoupravi, od digitalnega upravljanja kulturne dediščine do regulacije platformne ekonomije – na vseh teh in drugih področjih se tehnologije vse bolj prepletajo z vsakdanjim življenjem, s pravnimi in z institucionalnimi praksami ter modeli družbenega upravljanja.

Teme, obravnavane v zborniku – na primer etične dileme umetne inteligence, politične implikacije regulacije platform, vprašanja socialne pravičnosti, pomen kulturnega vrednotenja ter vloga deliberacije pri oblikovanju tehnoloških politik – poudarjajo nujnost interdisciplinarnega pristopa. Inženirji in računalniški znanstveniki morajo sodelovati z etiki, s sociologi, pravniki, z umetniki, ekonomisti in ekologi, če želimo celovito razumeti razsežnosti sodobnega tehnološkega razvoja. Raznolikost tehnologij zahteva tudi metodološko odprtost, konceptualno natančnost in etično občutljivost pri njihovem vrednotenju, ki mora zajeti kompleksnost in večdimenzio-

nalnost tehnoloških procesov. Družboslovje in humanistika imata pri tem ključno vlogo, saj lahko tehnologije obravnavata ne le kot rešitve, temveč tudi kot probleme – kot tvegane, sporne ali normativno dvoumne prakse, ki kličejo po demokratični razpravi in institucionalni odgovornosti.

Knjiga ponuja raznolik nabor pogledov na vprašanje vrednotenja in razvoja tehnologij ter odpira prostor za interdisciplinarni dialog. Uvodni prispevki poudarjajo, da so tehnologije družbeni konstrukt in nikoli ne delujejo v vakuumu, pač pa so hkrati odraz in oblikovalec družbenih procesov – vedno so vpete v razmerja moči, kulturne norme, institucionalne strukture in vrednotne orientacije. Od tod poziv k prehodu od instrumentalnega k reflektivnemu vrednotenju, pri čemer tehnologija ni več sama sebi namen, temveč sredstvo za uresničevanje etičnih ciljev, razvoj demokracije in zagotavljanje javnega dobrega. Prispevki kritizirajo prevladujočo usmerjenost ocenjevanja tehnologije na njeno tržno učinkovitost kot nevzdržno tehnološko kulturo, ki podlega tehnološkemu determinizmu, saj uvajanja določene tehnologije ne utemeljuje z njeno družbeno koristnostjo, temveč z njeno »pripravljenostjo«, da se razvije do tržno zanimive stopnje.

Uvodni Splichalov prispevek poudarja potrebo po prehodu od instrumentalne k reflektivni evalvaciji tehnologij, ki presega tehnične in ekonomske parametre ter tehnologijo umešča v širši družbeni sistem, pri čemer se njeni učinki prepletajo z vrednotami, normami in s pričakovanji. Prispevek utemeljuje nujnost enakopravne obravnave epistemskih in socialnih tehnologij kot kritiko enostranskih obravnav »pripravljenosti tehnologije« in s tem širši pogled na tehnološko zrelost, ki odpira ključna vprašanja, komu je tehnologija namenjena in komu bo najbolj koristila.

Ker tehnologija oblikuje človeške odnose, komunikacijo, institucije in pojmovanja sebstva, Mirjana Ule poudarja, da bi morali pri razsojanju o razvoju in uvedbi novih tehnologij nujno upoštevati tudi družboslovne in humanistične vidike in merila, predvsem pri oceni odgovornosti za uporabo tehnologije kot posebne oblike družbene, in ne le ožje tehnološke moči. Pri tem ima ključni pomen izobraževanje, saj se z njim začne razvijati zavest o odgovorni uporabi sodobnih tehnologij v vsakdanjem življenju in v sodobni družbi.

Franc Mali razvija koncept transdisciplinarnosti kot »orodja« za prehod k celovitemu vrednotenju tehnologij. Transdisciplinarnost pomeni preseganje linearnih inovacijskih modelov in prepoznavanje tveganj ter etičnih dilem, ki jih prinašajo tehnologije, kot sta generativna umetna inteligenca in genski inženiring. Ozko tehnične in tržno usmerjene modele je treba preseči z vključevanjem pravnih, socioloških, humanističnih in ekonomskih vidikov. Le s preseganjem disciplinarnih okvirov lahko razvijemo znanje, ki ustreza kompleksnosti družbenih in okoljskih izzivov.

Vprašanja umetne inteligence in digitalne tehnologije osvetljujeata prispevka Ivana Bratka, Ane Farič in Lee Košmrlj. Avtorji poudarjajo pomen etičnih smernic, pravne regulacije ter digitalne pismenosti za preprečevanje zlorab. Bratko in Košmrlj opozarjata na vpliv generativnih modelov na informacijsko okolje. Tehnološke rešitve so nezadostne brez institucionalne regulacije, pravnih mehanizmov in preglednih postopkov. Posebno pozornost namenjata vplivu generativne umetne inteligence na demokratične procese, širjenju dezinformacij in manipulaciji z javnim mnenjem. Analiza empiričnih študij kaže, da umetna inteligenca resno vpliva na sposobnost

državljanov za informirano odločanje in krepí polarizacijo. Avtorji poudarjajo pomen etičnih smernic, pravne regulacije ter digitalne pismenosti za preprečevanje zlorab. Bratko in Farič naredita pregled etičnih vidikov umetne inteligence in posebej razpravljata o etičnih vprašanih pristranskosti in transparentnosti sistemov umetne inteligence. Ugotavljata, da je za odgovorno uporabo ključna njena razumljivost in sposobnost razlage odločitev. Opozarjata tudi na tveganja, povezana z neprepričljivo zmožnostjo logičnega sklepanja velikih jezikovnih modelov in pojavom, da se umetna inteligenca uči sama od sebe, kar lahko vodi do neke vrste degeneracije.

Prispevka Matije Strliča, Lane Nastje Anžur in Francija Pušavca o digitalizaciji kulturne dediščine in raziskovalnih sistemov oblikujeta novo paradigmo, ki presega tržno logiko in utemeljuje tehnologijo kot javno dobro. Digitalizacija muzejev, arhivov in odprtokodnih znanstvenih instrumentov je lahko v službi znanosti, izobraževanja in dostopnosti, toda ta infrastruktura ni politično nevtralna – njen dizajn in dostopnost odsevata vrednostne odločitve. Avtorji poudarjajo pomen odprtih standardov in nekomercialnih modelov financiranja kot pogojev za vključujočo znanost in kulturo. Strlič in Anžur v svojem prispevku razvija modele nedenarnega vrednotenja kulturne dediščine, medtem ko Pušavec poudarja vlogo tehnologije za širši dostop do znanstvenih virov. Skupaj opozarjajo, da so javne dobrine, kot sta znanje in kultura, pogosto zapostavljene v komercialnih vrednotenjih, kljub njihovi ključni vlogi za kakovost življenja, kolektivni spomin in družbeno identiteto.

V razpravi o pametni specializaciji kot orodju za regionalni razvoj in trajnostno inovacijsko politiko Andreja Jaklič opozarja, da se vrednotenje prednostnih področij, kot sta denimo biotehnologija ali nanotehnologija, pogosto omejuje na tržne potencialne, ne upošteva pa dovolj okoljskih, družbenih in pravnih razsežnosti. V prispevku analizira izkušnje spremljanja slovenskih razvojnih inovacijskih partnerstev v sklopu strategije pametne specializacije v Sloveniji, ki razkrivajo omejitve tržno usmerjenih kriterijev in potrebo po vrednotenju tudi epistemskih in socialnih tehnologij, ter poudarja pomen participativnega načrtovanja, vključevanja javnosti in usklajevanja z izobraževalnimi, okoljskimi ter s pravnimi sistemi.

Podobno opozarja Aleksandra Kanjuo Mrčela na tveganja platformne ekonomije: prekarizacijo dela, koncentracijo digitalne moči, izgubo delavskih standardov in pomanjkanje demokratične kontrole. Tehnološke inovacije brez trajnostne in pravične regulacije lahko povečajo družbene neenakosti in spodkopljejo kohezijo. Njena analiza zahteva vključitev socialnih in regulativnih kriterijev v procese tehnološkega načrtovanja. Načela inkluzivnosti in participativnosti so lahko podlaga za določanje ciljev boljše organizacije digitalizirane in z umetno inteligenco podprte ekonomije ter za preprečevanje negativnih posledic nenadzorovanih sprememb, ki danes prekarizirajo delo in življenje vse večjega števila ljudi.

Zbornik pa se ne ustavi pri diagnostiki izzivov, temveč ponuja tudi zasnove za institucionalne in politične premike. Matjaž Maček predstavlja deliberativno poizvedbo kot politično tehnologijo za ustvarjanje premišljenega javnega mnenja, ki v nasprotju z običajnimi mnenjskimi poizvedbami vključuje poglobljeno in strukturirano razpravo o izbrani temi ter primerjalno ugotavljanje mnenj naključno izbranih udeležencev pred razpravo in po njej. Deliberativna poizvedba temelji na načelih enakosti, utemeljevanja in refleksivnosti, vendar ni namenjena oblikovanju skupne odločitve, temveč

preverjanju zanesljivosti mnenj. Prispevek poudarja pomen deliberativne poizvedbe kot orodja za demokratizacijo procesa političnega odločanja, ki omogoča razkriti, kaj bi ljudje menili, če bi bili bolje obveščeni in bi bili vključeni v premišljeni dialog.

Družbene tehnologije skozi prizmo univerzalnega temeljnega dohodka (UTD) obravnava Srečo Dragoš in s tem odpira razpravo o alternativnih merilih učinkovitosti, ki v središče postavljajo človekovo dostojanstvo in participacijo. UTD vidi kot orodje za izboljšanje kakovosti življenja, zmanjševanje revščine, krepitev socialne države in duševnega zdravja. Sklicuje se na empirične raziskave, ki potrjujejo pozitivne učinke UTD, tudi v povezavi z ekološkimi in s tržnimi spremembami. V obdobju avtomatizacije in starajoče se družbe ponuja UTD alternativni model družbene varnosti, ki omogoča pravičnejšo porazdelitev tveganj in blaginje. UTD bi tudi bistveno olajšal pravično porazdelitev stroškov prilagajanja na ekološko krizo in na spreminjanje trga delovne sile, saj avtomatizacija in prekarizacija dela, neenakost in staranje v sodobnih družbah slabijo stabilnost tradicionalnih modelov socialne varnosti.

* * *

Publikacija – prva v založbi Sveta za razvoj pri Slovenski akademiji znanosti in umetnosti – tako združuje raznolike, a povezane poglede na vprašanje vrednotenja tehnologij: od konceptualne kritike tehnodeterminizma do konkretnih institucionalnih predlogov za vključujoč in trajnosten tehnološki razvoj. Skupna nit vseh prispevkov je poziv k etični, reflektivni in demokratični obravnavi tehnologij – kot orodij, ki morajo biti namenjena človeku in skupnosti, ne pa zgolj trgu in kapitalu. Avtorji poudarjajo, da so tehnologije družbene konstrukcije, vpete v razmerja moči, institucionalne ureditve, kulturne pomene in politične odločitve.

Delo ne prinaša le analitičnega pregleda zdajšnjih praks in kritičnih perspektiv, ampak je tudi programski dokument – poziv k institucionalni, politični in akademski prenovi vrednotenja tehnologij. Tehnološki razvoj mora postati predmet odprte javne razprave, v kateri sodelujejo raznoliki akterji – znanstveniki, oblikovalci politik, civilna družba, uporabniki in tisti, ki jih tehnološke spremembe najbolj zadevajo. Le z vključujočim, interdisciplinarnim in vrednotno utemeljenim pristopom lahko zagotovimo, da bodo tehnologije v službi pravičnejše, bolj demokratične in trajnostne prihodnosti.

Celovito ocenjevanje tehnologij zahteva oblikovanje demokratičnega soglasja – ne le tehnično ali ekonomsko analizo, temveč sistemsko razumevanje dolgoročnih učinkov, družbene pravičnosti in razporeditve virov. Potrebujemo nove metode, ki bodo omogočale zaznavanje vplivov tehnologije na družbene odnose, institucije, okolje in javni prostor. Kriteriji vrednotenja morajo vključevati trajnost, pravičnost, dostopnost in odpornost – ne le hitrost inovacije ali njeno ekonomsko donosnost. V tem smislu tudi objavljeni prispevki ne razumejo znanosti zgolj kot producentke tehnoloških rešitev, temveč kot integralni del demokratične družbe ter pomembnega akterja v oblikovanju javnih politik in družbenih vrednot.

Zbornik tako ne želi le kritizirati obstoječih modelov ocenjevanja tehnologij, temveč prispevati k njihovemu preseganju. Njegov cilj je spodbuditi razmislek o drugačnih modelih vrednotenja, ki bodo skladnejši z normativnimi zahtevami plu-

ralne, demokratične in trajnostne družbe. Temeljno vprašanje danes ni več, ali potrebujemo drugačno vrednotenje tehnologij, ampak kako hitro in učinkovito lahko spremembe uresničimo in ali smo kot družba pripravljeni na odgovorno razvijanje tehnologij, ki ne bodo le »pripravljene« za trg, ampak zasnovane za pravičnejšo, demokratično in trajnostno prihodnost. Ali bo tehnološki napredek postal gonilo družbenega in ekološkega preobrata ali pa bo ostal privilegij ekonomske elite? Da bi se nova paradigma vrednotenja resnično uveljavila, bodo nujne globoke spremembe v izobraževalnih sistemih, financiranju raziskav, institucionalnih strukturah in – morda predvsem – v politični volji.

Ta publikacija ni zasnovana kot zgolj zbir strokovnih analiz, temveč poziv k nadaljnjemu sodelovanju med znanstvenimi disciplinami, med raziskovalnimi institucijami in javnimi akterji, med državljani in odločevalci – kot spodbuda za širši razmislek o tehnološki prihodnosti in naši pripravljenosti nanjo ter o družbeni tehnološki zrelosti.

KAKO DOSEČI TEHNOLOŠKO ZRELOST? OD INSTRUMENTALNEGA K REFLEKSIVNEMU OCENJEVANJU TEHNOLOGIJ

Slavko Splichal

Uvod

Tehnologija spreminja svet, toda kako lahko zagotovimo, da bo svet spreminjal tehnologijo v pozitivno smer? Čeprav so si to vprašanje zastavljali v vseh obdobjih, pa je postalo posebej pereče v času, ko tehnološki napredek poteka z izjemno hitrostjo in s težko predvidljivimi posledicami, zaradi česar ga je težko, če ne nemogoče, usmerjati in regulirati z družbenimi posegi. Veliki podatki, napredna analitika, umetna inteligenca, strojno učenje in internet stvari spreminjajo ne le gospodarstvo, ampak skoraj vse vidike življenja v globalnem merilu. Opredmetene in neopredmetene, družbene in epistemske tehnologije postajajo vse bolj prepletene.¹ Ta preobrazba vpliva tudi na naše kognitivne zmožnosti, načine, kako se učimo, komuniciramo, zaznavamo in izražamo mnenja in čustva, sprejemamo presoje in odločitve ter delujemo individualno in kolektivno. Ključno ostaja vprašanje, kdo (lahko) usmerja tehnološke rešitve, ki povzročajo te spremembe ter ustvarjajo tako pomembne priložnosti kot grožnje.

Leta 2017 je McKinsey Global Institute pripravil informativno sporočilo za digitalni vrh EU v Talinu, v katerem je opredelil deset »imperativov za Evropo« za »obvladovanje digitalne priložnosti« in »pripravo na prihodnost dela«. Dokument poudarja potrebo po oživitvi gospodarske rasti, povečanju poslovne dinamike, vlaganju v umetno inteligenco in druge nastajajoče tehnologije ter izboljšanju »naložb v človeški kapital, vključno s STEM, spretnostmi in z učenjem«, vendar nobeden od teh »imperativov« ne upošteva širših posledic teh procesov in kolektivnih družbenih interesov ter pomena družbenih tehnologij (McKinsey Global Institute 2017). Kljub temu pa ta razmišljanja pomembno vplivajo na mehanizme financiranja v EU in posameznih državah, ki dajejo prednost komercialni upravičenosti nastajajočih tehnologij, pozabljajo pa na kritične dejavnike, kot so trajnostnost, socialna pravičnost in družbena blaginja ter etična občutljivost. Tak pristop podpira prakse velikih tehnoloških podjetij, ki z dobički, ki jim jih prinaša prevladujoči položaj na trgu, zlahka prehitujejo akademske raziskovalne skupine pri izvajanju revolucionarnih raziskav in postavljajo raziskovalne prioritete. Ker delujejo v svetovnem merilu, je njihov vpliv na globalni tehnološki razvoj večji kot pa vpliv posameznih držav ali nadnacionalnih regulativnih organov, kar ustvarja dodatne težave političnim upravljavskim strukturam.

Ta prispevek obravnava neutemeljeno zanemarjanje nekomercialnih vidikov (ocenjevanja) tehnološkega razvoja, zlasti v kontekstu politik ocenjevanja 'priprav-

ljenosti tehnologije' (*technology readiness*). Predlagano je bolj družbeno konstruktivno vključevanje novih znanstvenih spoznanj v realno okolje z razvojem okvira za razreševanje raziskovalnih in družbenih izzivov pri konceptualizaciji in ocenjevanju usklajenosti novih tehnologij z družbenimi vrednotami. S tem naj bi prišli do učinkovitega razreševanja teh vprašanj in zagotovili, da tehnološki napredek ne bo služil le kapitalskim interesom, ampak prispeval tudi k širšemu družbenemu dobremu in razreševanju perečih družbenih problemov.

Prizadevanje za uveljavitev širšega razumevanja tehnološke pripravljenosti ali zrelosti je bistveno v obdobju, ko je postala govorica o TRL razvojna mantra, ki absolutizira 'pripravljenost tehnologije' in zanemarja 'tehnološko pripravljenost' družbe, povsem pa spregleduje družbene tehnologije. Medtem pa naraščajoče okoljske krize, nacionalni populizmi ter nenehni konflikti in vojne opozarjajo, da kljub znatnim naložbam v materialne tehnologije najbolj pereči problemi človeštva ostajajo nerešeni ali se celo zaostrojujejo.

Analiza primerov konceptualizacije tehnološke pripravljenosti in družbenih tehnologij jasno razkriva omejitve modela TRL in njegove konceptualne vrzeli ter nakazuje možnosti za njihovo preseganje. Da bi vrzeli odpravili, se prispevek osredinja na kritična vprašanja, ki zadevajo:

- konceptualizacijo 'pripravljenosti tehnologije' in 'tehnološke pripravljenosti';
- prenos metodologije ocenjevanja tehnologij iz organizacijskega (korporacijskega) v družbeno okolje;
- možnosti za preseganje omejitev modela TRL z vključevanjem nematerialnih vidikov v ocenjevanje pripravljenosti tehnologije (ne le vpliva na okolje);
- razvoj in družbeni vpliv epistemskih in družbenih tehnologij;
- spodbujanje družboslovnih znanosti k razvoju družbenih tehnologij in tehnološke zrelosti.

'PRIPRAVLJENOST TEHNOLOGIJE' IN 'TEHNOLOŠKA PRIPRAVLJENOST'

Razmislek o prizadevanjih za spodbujanje širšega razumevanja tehnološke razvitosti na splošno in posebej na področju financiranja znanstvenoraziskovalnih projektov začnemo z razlikovanjem med dvema pristopoma k ocenjevanju odnosa med tehnologijo in družbenim okoljem. Sodobna prizadevanja za uveljavitev širšega razumevanja in ocenjevanja tehnološke in družbene zrelosti presegajo idejo »pripravljenosti tehnologije« in spodbujajo usklajevanje novih tehnologij z družbenimi vrednotami ter podpirajo družbeno konstruktivno integracijo novih znanstvenih spoznanj v realnem okolju.

Izhodišče je kritika enostranske konceptualizacije devetstopenjskega modela 'pripravljenosti tehnologije' (*»Technology Readiness Levels«*) in razlikovanje med dvema pristopoma k ocenjevanju odnosa med tehnologijo in družbenim okoljem, ki ju predstavljata angleška – v slovenščino nekoliko okorno prevedljiva – izraza *»technology readiness«* (TRL) in *»technological readiness«*.²

Vendar je med 'pripravljenostjo tehnologije' in 'tehnološko pripravljenostjo' pomembna pomenska razlika (pustimo ob strani, ali je 'pripravljenost' najustreznejša

ustreznica za »*readiness*«; v slovenščini lahko govorimo o izdelanosti, dovršenosti ali razvitosti tehnologije).³ Medtem ko 'pripravljenost tehnologije' označuje pripravljenost določenih tehnologij za uporabo, 'tehnološka pripravljenost' postavlja v ospredje širši kontekst razvoja in uporabe novih tehnologij, ki vključuje ne le razvoj in uvajanje novih tehnologij, temveč tudi podporno družbeno infrastrukturo in potrebe, uporabniške kompetence, zakonodajne okvirje ipd. Ti dve pojmovanji bistveno dopolnjuje ideja 'družbene tehnologije', ki vključuje družbene mehanizme in prakse, ki pomagajo spodbujati družbeno preobrazbo in izhajajo iz spoznanj družboslovnih znanosti.

Pripravljenost tehnologije pomeni stanje ali raven pripravljenosti določenih tehnologij za uvedbo ali uporabo. Pogosto se nanaša na oceno učinkovitosti integracije specifične iznajdbe ali inovacije v obstoječe sisteme ali procese, zlasti tržne. To razumevanje je pretežno omejeno na industrijsko inženirsko okolje, kar vodi do ožjega dojemanja tehnologije kot zgolj orodja za proizvodnjo materialnih dobrin.

Tehnološka pripravljenost pa poudarja širši kontekst pripravljenosti tehnologije, ki zajema pripravljenost širšega 'tehnološkega okolja', kar vključuje družbeno infrastrukturo in potrebe, uporabniške spretnosti in izkušnje ter regulativne okvire, skratka, pripravljenost družbe za sprejem določene tehnologije. Družbena pripravljenost za sprejemanje tehnologije je lahko ključnega pomena za njeno uspešno uveljavitev, saj tehnologije ne delujejo v družbenem vakuumu, temveč zahtevajo podporno sprejemno okolje, v katerem lahko učinkovito zadovoljujejo individualne in kolektivne potrebe.

Razlikovanje je neposredno povezano s konceptualizacijo tehnologije v širšem smislu – kot uporabo raznovrstnih znanstvenih spoznanj, orodij, tehnik in procesov za ustvarjanje, razvoj in izboljšanje izdelkov, storitev ali sistemov za zadovoljevanje različnih človeških potreb in reševanje problemov. Medtem ko je v ožjem pomenu tehnologija omejena na industrijsko-inženirsko okolje oz. proizvodnjo, je v širšem pomenu povezana z upravljanjem in s spreminjanjem (človekovega) življenja in okolja.

Vzpon digitalnih tehnologij, kot so internet stvari, umetna inteligenca, sistemi blockchain ter digitalni trgi in platforme, je v preteklih desetletjih širil naše razumevanje tehnologije onkraj gole 'pripravljenosti tehnologije', kar je spodbudilo kritična pojmovanja, ki se osredinjajo na uporabnike tehnologije in njihovo funkcionalnost. Analize uporabniškega vidika tehnologij se osredinjajo na latentne tehnološke potencialne (»*affordances*«, v slovenščini bi morda lahko rekli *omogočilnosti*) – možnosti delovanja ali potencialne uporabe, ki jih določajo materialne omejitve, nameni proizvajalcev ali percepcije uporabnikov. Ta pristop ne proučuje samo, kako tehnologije spodbujajo in olajšujejo določeno uporabo, ampak tudi, kako uporabniki ustvarjalno prilagajajo tehnologije v vnaprej nepredvidenih okoliščinah.

V analogiji s Tönniesovo razpravo o človekovem odnosu do materialnih predmetov, ki je lahko organski ali refleksiven, lahko za tehnologijo rečemo, da je ne konstituira le materialni objekt tega odnosa, ampak tudi odnos sam – naša »družbena volja«. Kot je pokazal že Tönnies (1887/1999), je ravno naš voljni odnos do predmetov oziroma opredmetene lastnine tisti, ki razkriva bistvene značilnosti družbene strukture, v kateri živimo. Kompleksnost tega odnosa zahteva tudi temu ustrezen širši in večdisciplinarni pristop k ocenjevanju 'primernosti tehnologije', ki upošteva možne – želene in neželene – družbene posledice razvoja in uporabe tehnologije.

Večina tradicionalnih tehnologij spada v kategorijo 'snovnih tehnologij' – izdelane so iz lesa, kovine in drugih materialnih sestavin, ki omogočajo in omejujejo načine uporabe. Vendar pa obstajajo tudi nesnovne tehnologije, ki prav tako lahko pomembno vplivajo na človekovo delovanje. Še več, 'trde' opredmetene tehnologije ne bi bile možne brez ustreznih 'mehkih' neopredmetenih tehnologij, ki zagotavljajo znanje, metodologije, postopke in storitve za učinkovito rabo človeških virov in optimizacijo proizvodnih procesov pri razvoju in uporabi izdelkov in storitev.

Primeri novodobnih mehkih tehnologij so informacijsko-komunikacijske tehnologije, kot so algoritmi umetne inteligence, ki vodijo odločanje, družabni mediji, ki oblikujejo komuniciranje in družbeno interakcijo, ali digitalne valute, ki spreminjajo ekonomske transakcije. V sodobnosti družbene tehnologije vsekakor vključujejo tudi »tehnologije družbenih omrežij«, vendar njihov obseg močno presega zgolj spodbujanje »omrežnih odnosov med njihovimi uporabniki« ali »uporabe tehnoloških naprav, kot so pametni telefoni, ki vplivajo na družbene interakcije in vedenje tako, da vplivajo na komunikacijo iz oči v oči in družbene signale« (Science Direct b. l.). Med nesnovne tehnologije spadajo tudi bolj tradicionalne preddigitalne družbene tehnologije, kot so mnenjske poizvedbe, deliberativne poizvedbe ali model univerzalnega temeljnega dohodka, prav tako pa najrazličnejša pravila in regulatorni sistemi, običaji in norme, ki vplivajo na delovanje družb.

Tehnologija mnenjskih poizvedb uporablja statistične metode in analizo podatkov za 'merjenje javnega mnenja', kar vpliva na samozaznavo družbe in politične odločitve. Deliberativne poizvedbe vključujejo udeležence v strukturirane pogoje o pomembnih vprašanjih, kar oblikuje sodelovanje civilne družbe in javni diskurz. Model univerzalnega temeljnega dohodka konceptualizira ekonomske okvire, na podlagi katerih država oz. skupnost vsakemu državljanu vsak mesec zagotavlja dohodek brez pogojev in obveznosti in ne glede na njegov materialni položaj, kar radikalno spreminja socialna omrežja in gospodarske interakcije brez poseganja v človekovo opredmeteno okolje.

Ti primeri kažejo, da nesnovne tehnologije, udejanjene v abstraktnih konceptih in virtualnih okvirih, lahko bistveno vplivajo na človekovo delovanje in družbeno dinamiko, čeprav nimajo fizične oz. opredmetene podobe. Materialnost tehnologije zajema tako opredmetene kot neopredmetene artefakte in njihove specifične učinke, vključno z družbeno tehnologijo. Tako kot snovni gradniki določajo omejitve in možne rabe snovnih tehnologij, tudi nesnovni konstrukti lahko določajo potek in učinke interakcij in odločanja. Obe vrsti tehnologije imata pomembno vlogo v tehnološkem razvoju in se medsebojno dopolnjujeta.

Posledice obravnave tega vprašanja so pomembne zlasti za družboslovne znanosti, katerih vloga v razvoju tehnologij je izrazito podrejena. Vzajemno vplivanje opredmetenih in neopredmetenih tehnologij vpliva na vedenje ljudi in institucionalne okvire, kar je ključnega pomena za analizo družbenih sprememb, oblikovanje politik in spodbujanje inovacij v skupnostih. Njihovo razumevanje omogoča družboslovnim znanostim ustvarjati zanesljivejše modele družbene organizacije, tehnološke pripravljenosti in zasnove družbenih tehnologij ter vplivati na politike razvoja tehnologij in vrednotenje njihovih učinkov.

Pripravljenost tehnologije (TRL) – Za kaj? Za koga?

Lestvica TRL, ki je bila 'uzakonjena' z raziskovalnim programom EU Obzorje 2020, odraža prevladujoče ozko (tržno naravnano) razumevanje pripravljenosti tehnologije. Medtem ko je prvotna lestvica, ki jo je razvila NASA, namenjena ocenjevanju stopnje izdelanosti izbrane vesoljske tehnologije glede na specifični končni cilj – praktična uporaba v vesoljski misiji –, pa je lestvica TRL v priredbi EU namenjena ocenjevanju »pripravljenosti« katerekoli tehnologije za komercialno uporabo. Cilj oz. najvišja stopnja pripravljenosti je »dejanski sistem, preizkušen v operativnem okolju (konkurenčna proizvodnja v primeru ključnih omogočevalnih tehnologij)«.

Pomen »konkurenčne proizvodnje« je pojasnjen s posebnimi primeri:

- Izdelek iz proizvodnje: izdelek v končni obliki v popolni komercialni uporabi.
- Programska oprema: stabilna različica programske opreme, ki je na voljo na trgu pod pogoji poslovnega načrta.
- Zdravilo: zdravilo, ki je na voljo na trgu.

Osrednji problem, ki ga konceptualizira TRL, je uporaba znanstvenih spoznanj za razvoj komercialno uspešnih novih proizvodov. Na lestvici TRL Evropske komisije je pot od TRL 1 do TRL 9 predstavljena kot linearno napredovanje od »ideje do trga«, s čimer se pojem tehnologije reducira na proizvode, ki jih je mogoče monetizirati na trgu. Znatna javna sredstva za raziskave so uporabljena za premostitev 'doline smrti' – prehoda s TRL 4 (tehnologija, preizkušena v laboratoriju) na TRL 7 (prototip, preizkušen v realnem okolju), pri čemer niti univerze in raziskovalni inštituti niti zasebni sektor ne zagotavljajo dovolj naložb.

Posplošena lestvica TRL bistveno zožuje ne le razumevanje, ampak tudi z njim povezano finančno podporo znanstvenega raziskovanja, saj se osredinja predvsem na razvoj tehnoloških izdelkov, ki jih je mogoče dobičkonosno tržiti. Bistvena težava posplošitve je v neskladju med prvotnim namenom lestvice in njeno uporabo v kompleksnem, večplastnem okolju resničnega sveta. Lestvica TRL, ki so jo razvili v specifičnem – korporacijskem – okolju proizvodnje specifične – vesoljske – tehnologije, je bila pozneje 'ekstrapolirana' v povsem drugačno okolje – za ocenjevanje pripravljenosti katere koli tehnologije v globalnem tržnem okolju.

Najvišja stopnja TRL 9 na prvotni lestvici pomeni, da je bila zadevna tehnologija preizkušena v okolju, v kakršnem naj bi delovala, kar ne pomeni, da je pripravljena za komercialno uporabo. Zoženje ocenjevanja le na tehnologije, ki obljublajo tržno uspešnost, zapostavlja inovativne tehnologije, ki zahtevajo daljše razvojne cikle ali služijo nekomercialnim interesom. Taka omejitev lahko duši raziskovalno ustvarjalnost in raznolikost, saj se prednostno financirajo raziskovalni projekti z jasnimi obeti dobičkonosnosti, ne pa projekti, ki bi raziskovali širše družbeno, okoljsko ali teoretsko relevantne probleme. V nasprotju z gladkim, linearnim napredovanjem tehnologije od najnižje do najvišje stopnje pripravljenosti v korporativnem okolju uporaba lestvica TRL v kompleksnem okolju resničnega sveta ustvarja 'dolino smrti' z njegovo redukcijo na eno samo – komercialno – razsežnost ter zanemarjanjem raznolike dinamike in kompleksnih spremenljivk, ki vplivajo na globalno ali lokalno razvijanje in integracijo tehnologije. Če kje, potem bi 'dolino smrti' lahko videli v ločevanju tehnološkega razvoja od družbenih potreb.

Tabela 1: Primerjava izvirnega modela TRL (NASA) z modelom Evropske komisije v programu Obzorje 2020

TRL	NASA	Program EU Obzorje 2020
	Vesoljska tehnologija, korporacijsko okolje	Vsaka 'omogočevalna' oz. podporna tehnologija, globalni trg
1	Opažena in zabeležena osnovna načela	Prepoznana osnovna načela
2	Oblikovan tehnološki koncept ali aplikacija	Oblikovan tehnološki koncept
3	Potrditev koncepta ključne funkcije in (ali) značilnosti z analitičnimi in eksperimentalnimi metodami	Eksperimentalna potrditev koncepta
4	Potrditev delovanja sestavnih delov in (ali) preizkusne vezave v laboratorijskem okolju	Preverjanje tehnologije v laboratorijskem okolju
5	Potrditev delovanja sestavnih delov in (ali) preizkusne vezave v ustreznem okolju	Potrditev tehnologije v ustreznem okolju (industrijsko relevantnem okolju v primeru ključnih omogočevalnih tehnologij)
6	Demonstracija modela ali prototipa sistema oziroma podsistema v ustreznem okolju (na Zemlji ali v vesolju)	Demonstracija tehnologije v ustreznem okolju (industrijsko relevantnem okolju v primeru ključnih omogočevalnih tehnologij)
7	Demonstracija prototipa sistema v vesoljskem okolju	Demonstracija prototipa tehnološkega sistema v operativnem okolju
8	Sistem zaključen in s testi ter z demonstracijo potrjen kot primeren za letenje (na Zemlji ali v vesolju)	Sistem dokončan in kvalificiran
9	Sistem z uspešnim delovanjem v misiji potrjeno primeren za uporabo v poletih	Tehnološki sistem preizkušen v operativnem okolju (konkurenčna proizvodnja v primeru ključnih omogočevalnih tehnologij ali v vesolju)

Poleg tega lahko ta zoženi fokus TRL omejuje financiranje preliminarne in temeljne raziskave, ki ne kažejo komercialne sposobnosti preživetja, vendar so ključne za dolgoročni napredek, kulturo in blaginjo. Vključitev lestvice TRL v znanstveno politiko EU in nacionalne znanstvene politike poudarja zaskrbljujoč trend prevlade korporativne ideologije, ki lahko spodkopava demokratične vrednote z dajanjem prednosti dobičku pred javnim interesom, omejevanjem raznolikosti raziskovalnih programov in marginalizacijo glasov, ki poudarjajo vlogo znanosti pri reševanju kompleksnih svetovnih izzivov. Tak pristop spodbuja ustvarjanje inovacijskega okolja, ki služi interesom podjetij, in ne skupnemu dobremu.

Tako zoženje fokusa še zdaleč ni značilno samo za ocenjevanje pripravljenosti tehnologije, ampak ga zasledimo na številnih drugih področjih. Izključni poudarek na komercialnem vidiku v opredeljevanju pripravljenosti tehnologije je na primer očiten tudi v transformativni konceptualizaciji zagotavljanja storitev kot industrijskega procesa proizvodnje nesnovnih dobrin, ki sega na številne sektorje, kot so finance, kultura, mediji, prirejanje dogodkov in turizem. 'Industrializacija storitev' zabrisuje razliko med 'trdimi' tehnologijami (materialnimi izdelki) in 'mehkimi' tehnologijami (storitvami), saj storitve obravnava kot proizvodnjo dobrin za trg.

Omejitev ocenjevanja tehnologij na tržno uspešnost je še posebej vprašljiva ob vzponu epistemskih tehnologij – kot je na primer umetna inteligenca –, ki so zasnovane, razvite in uporabljene za ustvarjanje in manipuliranje znanja. Epistemske tehnologije so vedno obstajale na stičišču opredmetenih in neopredmetenih komunikacijskih tehnologij, a so bile redko ustrezno obravnavane. Opredmetene tehnologije so tradicionalno zagotavljale materialno osnovo za bogastvo in moč: oblikovale so industrijske, ekonomske in družbene strukture. Neopredmetene – družbene in epistemske – tehnologije vzporedno, a izrazito vplivajo kot *avtonomni instrumenti* moči. Velika moč epistemskih in družbenih tehnologij je v njihovi zmožnosti delovanja z nematerialnimi sredstvi na temeljni ravni človeške kognicije in družbene organizacije. Brez neposredne prisile – z ustvarjanjem in s širjenjem informacij, znanja in vrednot ter z ustvarjanjem ustrezne organizacijske infrastrukture – vplivajo na posameznikove kognitivne sposobnosti in družbeno organizacijo, oblikujejo družbene norme in vedenjske standarde ter regulirajo interakcije med posamezniki in skupnostmi.

S hitrim sodobnim razvojem epistemskih tehnologij, ki poudarja pomen kolektivnega učenja in prilagajanja, postaja vse pomembnejša tudi družbena tehnologija. Epistemska in družbena tehnologija sta tesno prepleteni: prva vključuje orodja in metode za ustvarjanje, preverjanje in razširjanje znanja, druga pa mehanizme za organizacijo in usmerjanje družbenega vedenja in odnosov. Družbene tehnologije – kot so na primer pravni sistemi, protokoli komuniciranja ali platformne arhitekture – pogosto vsebujejo epistemološke predpostavke in oblikujejo pogoje za vednost. Hkrati epistemske tehnologije – npr. statistika, mnenjske poizvedbe, modeli strojnega učenja, algoritmi – vedno učinkujejo znotraj določenih družbenih okvirov. Njuno razmerje je soodvisno: vsaka družbena tehnologija vključuje neko epistemologijo, vsaka epistemska tehnologija pa deluje v družbenem kontekstu. Epistemske tehnologije so po definiciji družbene tehnologije, vendar pa niso vse družbene tehnologije nujno epistemske.

Na problem zanemarjanja specifik uporabe epistemskih tehnologij je že pred skoraj stoletjem na primer opozoril Robert Park (1938, 198) s primerom »sporočanja novic iz enega kulturnega okolja drugemu«. Pojasnjeval je, da komuniciranja ni mogoče obravnavati kot »preprosto operacijo transporta blaga, kot je opeka«, temveč kot epistemski (kulturni) proces z dvema temeljnima dimenzijama – difuzijo ali transmisijo ter akulturacijo ali asimilacijo –, v katerem uporabniki aktivno sodelujejo.

Umetna inteligenca in z njo povezane metode podatkovne znanosti – kot so strojno učenje in veliki jezikovni modeli –, ki prepoznavajo in ustvarjajo besedila in druge vsebine, so odprle povsem nove perspektive in družbeni domet epistemskih tehnologij. V takih okoliščinah je poenostavljeni pristop ocenjevanja pripravljenosti tehnologije TRL še posebej problematičen, ker daje komercialnemu vidiku prednost pred drugimi kritičnimi dejavniki, ki določajo resnično 'pripravljenost' tehnološko industrializiranih storitev. Med te dejavnike lahko spadajo lokalna sprejemljivost, učinkovitost delovanja, uporaba lokalnih virov, okoljska vzdržnost, dostopnost, spodbujanje zaposlovanja, opolnomočanje posameznikov, vključevanje skupnosti pri odločanju in spodbujanje socialne enakosti med prebivalstvom. Že leta 1972 je mednarodna skupina raziskovalcev v poročilu Rimskega kluba z naslovom *Meje rasti* opozorila, da pretirano naraščanje prebivalstva, proizvodnja hrane, industrijske proiz-

vodnje, izraba neobnovljivih naravnih virov in onesnaževanje okolja ogrožajo globalno ravnesje med za planet vzdržnim številom prebivalstva in obsegom proizvodnje materialnih dobrin. Razvoj epistemskih tehnologij bi lahko prispeval k družbenemu razreševanju teh problemov pod pogojem, da jih ne bi ocenjevali po tržni uspešnosti. Tako globalni in kontekstualno specifični dejavniki bi morali biti temelj za ocenjevanje pripravljenosti na tehnologijo in so tudi osrednjega pomena za konceptualizacijo epistemskih in družbenih tehnologij.

KRITIČNI ODZIVI NA MERJENJE PRIPRAVLJENOSTI TEHNOLOGIJE

Kritiki 'evropske' lestvice TRL, ki jo je oblikovala administracija EU s prilagoditvijo Nasine lestvice, ugotavljajo, da sta se z razširitvijo njene uporabe zunaj prvotnega konteksta vesoljskih programov zmanjšali njena »konkretnost in sofisticiranost«. Te pomanjkljivosti bi bilo mogoče odpraviti s priročniki za vrednotenje TRL, ki bi jih oblikovali za posamična raziskovalna področja in znanstvene discipline (Héder 2017).

V zadnjem desetletju je prevladujoča komercialna inačica lestvice TRL postala predmet konstruktivne kritike v delu družbenih ved, pa tudi drugje, na primer v obravnavah umetne inteligence. Tem prizadevanjem je skupna ideja kritične konstruktivne rekonceptualizacije 'pripravljenosti tehnologije'. Če bi bila ta konstruktivna kritika 'uradno' sprejeta, bi to imelo pomembne posledice za spodbujanje razvoja znanstvenega raziskovanja, saj spodbude ne bi bile omejene na doseganje tržne monetizacije novih materialnih proizvodov, ampak bi postavile v ospredje prizadevanja za uresničevanje temeljnih družbenih vrednot, ki so domena vseh znanosti, vključno z družboslovnimi in s humanističnimi.

Kritike lestvice TRL lahko razvrstimo v dve skupini:

1. Kritika omejitve TRL na tržno monetizacijo tehnološkega proizvoda predlaga upoštevanje večjega števila dejavnikov, odgovornih za prodor nove tehnologije, in poudarja družbeno pripravljenost za sprejem nove tehnologije (*»technological readiness«*).
2. Kritika tržne konceptualizacije uspešnosti tehnologije predlaga širše razumevanje tehnologije, njene družbene vloge in razvojnih posledic, vključno z upoštevanjem družbenih tehnologij.

Razširitev TRL

Prva skupina kritik izhaja iz ideje »odgovornega raziskovanja in inovacij« (*»Responsible Research and Innovation«* – RRI), ki predlaga usklajevanje raziskovanja z razreševanjem širših družbenih in etičnih vprašanj. Med kriterije odločanja o inovacijskih politikah v znanstvenih raziskavah in tehnološkem razvoju ter ocenjevanja uspešnosti raziskovalnih, razvojnih in inovacijskih projektov naj bi bile vključene etična sprejemljivost, družbena zaželenost ter trajnostnost raziskovanja in inovacij. Raziskave naj bi med drugim izpolnjevale visoke etične standarde, nalagale odločevalcem odgovornost za preprečevanje škodljivih učinkov inovacij, vključevale v raziskovanje skupnosti, ki jih inovacije zadevajo, in jim omogočale pridobiti ustrezno znanje za razumevanje posledic spodbujanja znanstvenega izobraževanja ter zagotavljale enakost spolov v raziskovalnih skupinah in odprti dostop do vseh faz raziskovanja.

Modeli vrednotenja uspešnosti raziskovanja, zasnovani na teh izhodiščih, vključujejo različne vidike družbenega prilagajanja na nove tehnologije, ki zagotavljajo 'tehnološko pripravljenost' družbe, sicer predvsem za uspešno monetizacijo nove tehnologije na trgu. Poleg pripravljenosti tehnologije so to zlasti (Bruno et al. 2020; Verma in Allen 2024; KTH b. l.):

- ocena tveganosti tehnologije (»*Technology Risk Assessment*«),
- pravna pripravljenost (intelektualna lastnina, pravno okolje, etika),
- organizacijska pripravljenost (delovne skupine, organizacijsko okolje),
- družbena pripravljenost,
- potrebe in interes uporabnikov,
- pripravljenost poslovnega modela (finančna, okoljska in družbena izvedljivost),
- zagotovitev financiranja za prenos na trg.

Dodani vidiki ocenjevanja 'pripravljenosti tehnologije' opozarjajo na kompleksnost družbenih predpostavk in posledic uvajanja novih tehnologij, ki jih ne smemo zanemariti pri ocenjevanju njihove 'pripravljenosti' za nastop na trgu.

Z družbenoteoretskega vidika je pomembno, da je digitalna tehnologija mnogo bolj kot katera koli prejšnja tehnologija 'prisilila' naravoslovne in družbene vede, da proučujejo skupne 'mejne objekte' raziskovanja, vendar žal še vedno nimajo skupnega jezika za razpravo o ključnih pojmi in skupnih metodah raziskovanja 'mejnih raziskovalnih tem'. Skupne mejne objekte in sodelovanje zunaj disciplinarnih meja omogoča predvsem kvantifikacija in v zadnjem času upodatkovanje družboslovnih ved. K temu pa ne prispeva vsaj delna 'udomačitev' družboslovne epistemologije v naravoslovju, s čimer bi slednje prepoznalo, da so tudi njihovi predmeti raziskovanja lahko družbeno konstruirani.

V tem odnosu so družboslovci potisnjeni v podrejeno vlogo podpore in interpretacije tehnično-tehnološkega razvoja in dosežkov ter občasnega reševanja družbenih problemov, ki so posledica tega razvoja. Raziskovalci v naravoslovju, tehnologiji, inženirstvu in matematiki (STEM) se pogosto ukvarjajo z mejnim delom, katerega cilj je razširiti, monopolizirati ali zaščititi avtonomijo svojih področij ali disciplin, in le redko priznavajo prepustnost meje med družboslovnimi in naravoslovnimi znanostmi, ki se je z upodatkovanjem močno povečala. Že v primeru tehnike – na primer pri ocenjevanju tehničnih inovacij – zanemarjanje družbenih in družboslovnih razsežnosti instrumentalnega delovanja povzroča resne omejitve, saj tehnika ustvarjanja in uporabe specifičnih materialnih orodij in strojev vključuje tudi izvajanje nematerialnih opravil, veščin ali postopkov, ki so namenjeni doseganju vnaprej določenih ciljev. Tem bolj je tehnično-naravoslovna omejenost nedopustna v primeru obravnave tehnologij kot uporabe znanstvenih spoznanj za sistemsko ustvarjanje orodij, strojev, naprav ali procesov, pri čemer vizija »instrumentalistične družbe« nakazuje, da se ljudje lahko učijo in delajo učinkoviteje s posnemanjem procesov, ki potekajo v pametnih strojih, in s tem spodbujajo harmonično sobivanje (Schmidt in Thrun 2016).

Tehnologija po svojem bistvu ni le tehnična, ampak tudi družbena: ni zgolj skupek materialnih naprav ali tehničnih postopkov, temveč tudi način organizacije družbenega delovanja. Odločanje samo – denimo v političnih institucijah, upravlja-

nju podatkov ali platformnem usmerjanju vedenja uporabnikov – lahko razumemo kot družbeno tehnologijo: skupek postopkov, pravil, algoritmov in institucionalnih ureditev, ki oblikujejo tokove vpliva, moči in vednosti. Tako kot stroj ni nevtralen, tudi proceduralna zasnova odločanja ni brez učinkov – denimo v načinu, kako se določa, kdo ima dostop do informacij, kdo lahko sodeluje pri oblikovanju problemov in kdo je izključen iz odločanja. Primeri, kot so avtomatizirano odločanje v javni upravi, algoritemska moderacija na spletnih platformah ali sistemi za razporejanje delavcev v gig ekonomiji, kažejo, da so tehnološke rešitve hkrati tudi družbeni režimi – z lastnimi normami, hierarhijami in oblikami odgovornosti. V tem smislu tehnologija ni zgolj podpora družbi, temveč je njena notranja logika – način, kako se družba strukturira, usmerja in uravnava.

Podrejen položaj družboslovnih znanosti dobro ponazarja sedanje raziskovanje podnebne krize. V svoji raziskovalni politiki EU daje prednost »razumevanju, kako podnebne spremembe vplivajo na življenjske poteke in industrije«, in raziskovanju »interakcij med hidrologijo in družbo z uporabo kvalitativnih zgodb /.../, zbranih z nestandardiziranimi intervjuji in delavnicami mentalne simulacije« (Pultarov 2024).⁴ Protislovje je v tem, da se »kvalitativne zgodbe« uporabljajo za razumevanje nenačrtovanih negativnih družbenih posledic tehnološkega razvoja, ne upoštevajo pa se v načrtovanju tega razvoja, s katerim bi lahko take negativne posledice preprečili, če ga ne bi podrejali enostranskim – profitnim – interesom. To je tudi problem uporabe letvice TRL, saj povsem zanemara širše družbene posledice razvoja in uvajanja novih tehnologij.

Refleksivno participativno ocenjevanje tehnologije

Problema zanemarjanja širših družbenih posledic razvoja in uvajanja novih tehnologij se je inovativno lotil projekt The Expert and Citizen Assessment of Science and Technology (ECAST). Mreža ECAST, ki jo je leta 2010 ustanovila skupina raziskovalcev, izobraževalcev in izvajalcev politik z univerze Arizona State, je s svojo refleksivno participativno metodo ocenjevanja tehnologij (»participatory technology assessment« – pTA) napravila pomemben korak pri spodbujanju demokratičnega odločanja o znanstveni politiki. Ta metoda temelji na načelih deliberativnega glasovanja, ki jih je oblikoval James Fishkin, in omogoča kontinuirano izpopolnjevanje in prilagajanje (Fishkin, Luskin in Jowell 2000; Deliberative Democracy Lab b. l.).

Metoda pTA ECAST poteka v treh participativnih korakih:

1. oblikovanje problema: zainteresirane strani skupaj opredelijo problematiko in jo postavijo v specifični kontekst;
2. državljansko posvetovanje: vključuje raznovrstne udeležence za poglobljene razprave, katerih cilj je iskanje in ocenjevanje izvedljivih rešitev;
3. rezultati in integracija: sintetizirane ugotovitve so vključene v procese odločanja.

V preteklih letih je bil ta pristop za ocenjevanje tehnologij uspešno uporabljen za več deset znanstvenih in tehnoloških projektov na različnih področjih, kot so tehnologija avtonomnih vozil, ohranjanje biotske raznovrstnosti, raziskave podnebnih sprememb in razprave za krepitev odpornosti skupnosti in okoljskega odločanja. Projekti so vedno vključevali raznolik nabor udeležencev – od znanstvenikov in deležnikov,

vklučenih v procese odločanja o tehnologijah, do državljanov – iz različnih tehničnih, družbenih in političnih okolij na vseh ravneh upravljanja in odločanja, od lokalnih skupnosti do nacionalnih in mednarodnih prizorišč. Vsako posvetovanje je potekalo na več lokacijah in je vključevalo od 23 do 489 udeležencev, kar naj bi zagotovilo teden in reprezentativen prispevek k razpravi.

Porazdeljena mrežna struktura modela ECAST zagotavlja odpornost in trajnost modela ter ohranja neodvisnost in kontinuiteto tudi v spremenjenih družbeno-političnih razmerah. Ta model omogoča, da glasovi državljanov postanejo sestavni del razprav o znanstveni politiki, kar pomaga spodbujati bolj demokratično in vključujoče odločanje (Kaplan et al. 2021).

Vrednotno občutljivo oblikovanje tehnologij (»Value Sensitive Design« – VSD)

V zadnjem času je vse več pozornosti namenjene etičnim vidikom ne le uporabe tehnologije, ampak tudi njenega načrtovanja. Ta premik poudarka izhaja iz spoznanja, da so družbene posledice tehnologije v fazi njene uporabe večinoma že določene, medtem ko so v fazi načrtovanja in oblikovanja vsaj delno predvidljive in je nanje mogoče vplivati oziroma jih spremeniti ter zmanjšati tveganja. Pravočasna obravnavanje etičnih in družbenih vidikov razvoja tehnologij omogoča lažje doseganje želenih učinkov ter preprečevanje negativnih družbenih posledic.

Najradikalnejši odmik od modela TRL se izraža v zasnovi »vrednotno občutljivega oblikovanja tehnologij« (VSD) (Friedman, Kahn in Borning 2006; Friedman in Hendry 2019). Medtem ko je pTA metoda za ocenjevanje tehnologij s poudarkom na vključevanju javnosti v razpravo in odločanje o znanstvenih in tehnoloških vprašanjih, je VSD metoda za oblikovanje tehnologij, ki v proces razvoja od vsega začetka sistematično vključuje etične in družbene premisleke in dejavnike. Cilj VSD je zagotoviti, da tehnološke inovacije ne služijo le funkcionalnim in ekonomskim ciljem, temveč upoštevajo tudi človekove vrednote, kot so zasebnost, pravičnost, avtonomija in dobrobit uporabnikov ter širše družbe – vendar tako kot TRL tudi VSD ne vključuje družbenih tehnologij.

Tabela 2: Primerjava modelov kritičnega ocenjevanja in oblikovanja tehnologij

	pTA (ECAST)	VSD
Namen	Ocenjevanje tehnologij s širšo družbeno razpravo	Oblikovanje etično naravnanih tehnologij
Metodologija	Kombinacija strokovne ocene in javne razprave	Konceptualna, empirična in tehnična analiza
Vloga uporabnikov	Razpravljalci in soodločevalci v ocenjevanju tehnologije	Deležniki v procesu oblikovanja tehnologije
Značilna uporaba	Politične odločitve o znanstvenih in tehnoloških vprašanjih	Razvoj digitalnih in fizičnih tehnologij

VSD uporablja interdisciplinarni in iterativni pristop, ki združuje tri ravni analize:

1. Konceptualna analiza identificira in definira vrednote, ki so pomembne za vse deležnike. Vključuje etične teorije in razmislek o vplivih tehnologije.
2. Empirična analiza uporablja družboslovne raziskovalne metode, kot so intervjuji, ankete, opazovanja, za razumevanje stališč in potreb uporabnikov tehnologije ter drugih deležnikov.
3. Tehnična analiza proučuje, kako lahko tehnologija podpira ali omejuje določene vrednote ter kako oblikovalske odločitve vplivajo na etične vidike.

Vse tri vrste analiz lahko potekajo vzporedno in iterativno, da se zagotovi celovit pristop k vključevanju vrednot v oblikovanje tehnoloških inovacij.

VSD se lahko uporablja na različnih področjih, od informatike in inženirstva do urbanističnega načrtovanja in medicinske tehnologije. Značilni primeri vključujejo varovanje zasebnosti na digitalnih platformah (zagotavljanje uporabnikovega nadzora, kontrola nad osebnimi podatki), zagotavljanje algoritemske pravičnosti (razvoj umetne inteligence, ki zmanjšuje pristranskost in diskriminacijo, npr. pri sistemih za odločanje v zaposlovanju ali kreditnem točkovanju), razvoj medicinskih aplikacij in naprav, ki spoštujejo pacientovo avtonomijo in varnost, ter oblikovanje urbanih tehnologij v razvoju pametnih mest (npr. nadzornih sistemov, prometa, javnih storitev z zagotavljanjem transparentnosti, dostopnosti in pravičnosti).

Vloga družbenih in epistemskih tehnologij v sodobni družbi

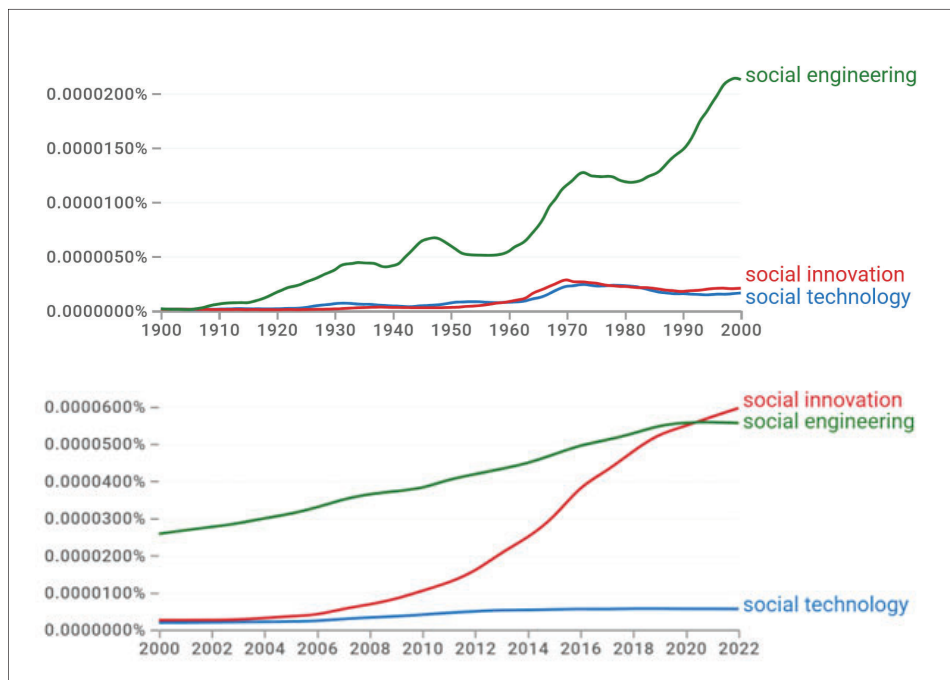
Metodologiji TRL in VSD imata pomembno vlogo pri ocenjevanju razvoja in vplivov tehnologije. Medtem ko je TRL osredinjena na tehnično pripravljenost in stopnjo zrelosti tehnologije, VSD vključuje tudi vrednotne vidike pri oblikovanju tehnologij, vendar obe metodologiji zanemarjata vpliv tehnologije na družbene odnose in infrastrukturo, pri čemer družbena tehnologija ponuja ključno dopolnitev.

Model TRL se osredinja predvsem na tehnično in komercialno zmožnost »preživetja« tehnologije, ne upošteva pa družbene dinamike, ki vpliva na uporabo in sprejemanje tehnologije. Pri tem ne zanemari samo netržnih dejavnikov, ampak tudi nematerialne komponente, kot so zmožnosti za družbeno interakcijo, učenje in skupnostno delovanje, katerih vloga je bistvena pri tem, kako ljudje ustvarjajo in sprejemajo nove tehnologije ter jih vključujejo v svoje življenje.

Družbene tehnologije, inovacije in inženiring

Uspešno sprejemanje tehnologije je pogosto odvisno od družbenega konteksta in interakcij uporabnikov, kar nakazuje potrebo po celostnejšem okviru ocenjevanja tehnologij. To zahteva širše razumevanje tehnologije in družbene tehnološke zrelosti ter iskanje rešitev, kako lahko tehnologija ne samo zadovoljuje 'potrebe trga', ampak tudi izboljša kakovost življenja ljudi in družbenih odnosov. V času, ko lestvica TRL predstavlja ključno merilo financiranja tehnoloških invencij, je še posebej pomembno prepoznati pomen družbenih okoliščin razvoja novih tehnologij in prispevek družbenih tehnologij k blaginji in razvoju družbe. Na to je že leta 1898 opozoril čikaški sociolog Albion Small (1898) v razpravi o pomenu družbene tehnologije, ki jo je razumel kot »uporabo znanja o dejstvih in zakonih družbenega življenja za doseganje

racionalnih družbenih ciljev«, čeprav je imel pri ciljih v mislih predvsem preseganje posameznikove nezmožnosti obvladovanja zahtev »normalnega« družbenega okolja (ibid.).



Slika 1: Relativna pogostnost pojavljanja izrazov družbena tehnologija, družbene inovacije in družbeni inženiring v angleških besedilih v obdobjih 1900–2000 (zgoraj) in 2000–2022 (spodaj). Vir: Google Books Ngram Viewer.

Prvotno je bila družbena tehnologija del velike, celo utopične vizije, v kateri je bila vloga sociologije določati cilje, ki naj bi jim družba sledila, zbirati in sintetizirati ustrezna znanstvena spoznanja ter jih prevajati v načrte za razvojne spremembe družbe. Vendar pa je po prvi svetovni vojni vzpon pozitivistične sociologije, ki je zahtevala, da vrednote ne smejo vplivati na znanstveno raziskovanje, zasenčil kritično razumevanje družbene tehnologije. Nadomestil ga je bolj manipulativen pojem »družbeni inženiring«, ki je postavljaj v ospredje družbene posege za nadzorovanje ali vplivanje na ljudi brez upoštevanja etičnih vidikov in družbenega konsenza, tako da se je pomensko približal propagandi in celo psihološki vojni. Družbeni inženirji – podobno kot drugi inženirji – ne razvijajo svojih tehnokratskih sposobnosti in storitev v prid javnemu dobremu, ampak za izpolnjevanje interesov oblasti in financerjev.

Sodobna prevlada modela TRL v spodbujanju tehnološkega razvoja kaže, da je opustitev pojma 'družbena tehnologija' imela negativne posledice, saj se je 'tehnologija' vse bolj začela nanašati izključno na tržno unovčljive materialne dobrine in procese (Derksen in Wierenga 2013). Vendar pojem družbene tehnologije s poudarjenim razlikovanjem med tehniko in tehnologijo ponuja širše razumevanje pomena tehnologije – s poudarkom na družbenih inovacijah kot mehanizmih, ki vplivajo na družbene procese.

Pojem družbena tehnologija se je uveljavil konec 19. stoletja kot bistveni element v prizadevanjih za celostnejše razumevanje interakcije med snovnimi in nesnovnimi tehnologijami in družbenimi vrednotami. Po prvi svetovni vojni je tedaj prevladujoče prepričanje, da je treba sociologijo »osvoboditi vrednotnih pristranskosti«, skupaj z vzponom vsiljivega družbenega inženiringa zmanjšalo naklonjenost idejam družbene tehnologije v kritičnem družboslovnem raziskovanju. Danes digitalna komunikacijska revolucija in umetna inteligenca – ob vedno tesnejšem prepletanju materialnih in nematerialnih tehnologij – vračata idejo družbene tehnologije v središče pozornosti. Njen pomen se je bistveno povečal s hitrim razvojem digitalnih tehnologij in umetne inteligence kot najbolj propulzivne skupine kognitivnih in epistemskih družbenih tehnologij. S tem, da radikalno spreminja način ustvarjanja, preverjanja, širjenja in prisvajanja znanja, umetna inteligenca ni le najbolj transformativna epistemska tehnologija v zgodovini, ampak je z neposredno uporabo v urejanju družbenih odnosov tudi ključna družbena tehnologija sodobnosti. Lahko rečemo, da postaja infrastruktura vseh družbenih ureditev 21. stoletja.

V temeljih pojmovanja družbene tehnologije je ideja, da so družbene inovacije za družbeni napredek enako pomembne kot tehnično-tehnološke. Uspeh tehnoloških inovacij ni odvisen le od njihove tehnične izvedbe, ampak tudi od njihove 'omogočilnosti' ali njihove sposobnosti, da so v soglasju z družbenimi potrebami in vrednotami. Medtem ko snovne tehnologije služijo kot ključni *vir* moči, so nesnovne ali epistemske tehnologije – kot so ideologije, politični diskurz, pravni okviri in izobraževalni sistemi – že dolgo znane kot avtonomni *agenti* moči (Hakkarainen et al. 2009; Alvarado 2023). Epistemske tehnologije vključujejo široko paleto orodij, metod, sistemov in praks »imetja za resnično« (Kant 1781/1956, 689). Z njimi je mogoče ustvarjati in širiti informacije, znanje in vrednote, oblikovati družbene norme in vedenjske standarde ter upravljati interakcije in sodelovanje med posamezniki in skupnostmi ter s tem vplivati na individualne kognitivne sposobnosti ter na družbeno organizacijo.

Posebej pomembna oblika družbene tehnologije je izobraževanje, ki je ključno za oblikovanje družbenih struktur in odnosov. V kontekstu družbenih tehnologij se izobraževanje ne obravnava le kot prenos znanja in izmenjave informacij, temveč tudi kot proces, ki (lahko) krepi družbeno kohezijo, spodbuja sodelovanje in ne nazadnje omogoča družbeno prilagajanje na tehnološke spremembe. Navade, običaji, obredi in druge kulturne prakse prav tako delujejo kot nesnovne tehnologije, ki krepijo družbene strukture in hierarhije ter lojalnost in nadzor v družbenih skupinah.

Nematerializirane tehnologije so bistvene za ohranjanje (oblastne) avtoritete in oblikovanje kolektivnega vedenja. Pogosto jih dopolnjujejo opredmetene tehnologije, zlasti komunikacijske tehnologije, ki omogočajo izmenjavo idej, informacij in kulturnih praks ter delujejo kot ključni mediji, na podlagi katerih učinkujejo družbene tehnologije. Kompleksno razmerje med neopredmetenimi družbenimi tehnologijami in opredmetenimi komunikacijskimi tehnologijami zastavlja pomembna vprašanja o družbeni moči, nadzoru, zasebnosti, osebni svobodi, družbeni pravičnosti in etičnih posledicah tehnološkega napredka. Razumevanje in pojasnjevanje večplastne vloge družbenih tehnologij zahtevata sistematično raziskovanje njihovih potencialov za oblikovanje človeških izkušenj in družbenih struktur. Izključenost družbenih tehnologij je kritična slepa pega v modelu TRL in v sorodnih metodologijah za ocenjevanje

‘pripravljenosti tehnologije’, kar ne nazadnje zavira tudi razvoj družboslovnih ved. Izključenost družbene tehnologije iz modelov ocenjevanja tehnologij je tesno povezana z njeno izključenostjo iz družbenih modelov spodbujanja razvoja in financiranja tehnologij, kot sta na primer okvirni program Evropske unije za raziskave in inovacije (2021–2027) in program za »spodbujanje izvajanja raziskovalno-razvojnih programov (TRL 3–6)« slovenske agencije za znanstvenoraziskovalno in inovacijsko dejavnost (ARIS 2025). V modelih ocenjevanja tehnologij, kot je TRL, se zrcali ozka, tehnoeekonomska perspektiva razvoja, ki zanemarja družbene dimenzije tehničnih inovacij in družbene inovacije ter spodbuja instrumentalne rešitve brez ustreznega razumevanja in vrednotenja razvojnih ciljev in njihove družbene relevantnosti.

V zadnjih desetletjih se v razpravah o družbenih spremembah in reševanju družbenih problemov na številnih znanstvenih področjih uveljavlja tudi izraz »družbene inovacije« (*»social innovation«*). V organizacijskih vedah se izraz povezuje s pojmi, kot so socialno podjetništvo, socialne finance, družbena odgovornost podjetij in javno-zasebna partnerstva. Družbene inovacije so predmet obravnave tudi v urbanizmu, tehnološkem upravljanju, sociologiji, javni upravi in raziskovanju politik (Drucker 1987; Moulaert et al. 2013). Še zlasti pa je razprave o družbenih inovacijah postavil v ospredje razvoj umetne inteligence (Schwab Foundation 2024).

Razlog za alternativno poimenovanje družbene tehnologije najverjetneje tiči v njeni zgodovinski redukciji na družbeni inženiring – uporabljeni bodisi za načrtovanje življenja ljudi, družbeno nadzorovanje in varnostno spremljanje bodisi za iskanje novih načinov prepričevanja ljudi, da sprejmejo določeno tehnološko inovacijo. Primeri takih družbenih tehnologij so tudi v sodobnosti številni in nas postavljajo pred zahtevna etična vprašanja o meji med tehnološkim napredkom in človekovo svobodo. Algoritmi družabnih omrežij za določanje kredibilnosti posameznikov in prilagojeno oglaševanje, umetna inteligenca in sistemi digitalnega nadzora oblikujejo naše komunikacijske navade, zaznavanje realnosti in vedenjske vzorce. Med bolj protislovnimi je kitajski sistem »družbenega kredita« (*»Social Credit System«*), ki združuje podatkovne analize, nadzor in družbeni inženiring za usmerjanje vedenja državljanov.

Družbenega inženiringa sicer ni mogoče prezreti kot možne oblike uporabe (zlorabe) družbenih tehnologij. Vendar pa ima družbena tehnologija bistveno širši pojmovni obseg. Družbene tehnologije – kot so pravni sistemi, protokoli komuniciranja ali platformne arhitekture – strukturirajo družbene interakcije, oblikujejo družbene procese, odnose, norme in institucije. Pod vplivom digitalizacije imajo družbene tehnologije v sodobni družbi ključno vlogo od medosebnih odnosov do kompleksnih družbenih sistemov, kot so ekonomija, politika in izobraževanje. Po eni strani omogočajo učinkovitejšo organizacijo za razvoj pravične, učinkovite in trajnostne družbe ter reševanje družbenih problemov, kot so revščina, neenakost in dostop do izobraževanja. Po drugi strani pa družbene tehnologije prinašajo tudi tveganja, povezana z nadzorom, manipulacijo in s posegi v zasebnost (kar sicer v strogem pomenu velja tudi za pojem »družbene inovacije«).

Vsekakor družbena tehnologija daje boljše izhodišče za razumevanje, kako tehnološki artefakti vplivajo na družbene strukture, procese in vrednote, kot ga (ne) omogočajo tradicionalni modeli ocenjevanja »pripravljenosti tehnologije«. Dopolnitev TRL z družbeno tehnologijo – kot predmetom in instrumentom ocenjevanja – bi

omogočala ocenjevanje ne le tehnične izdelanosti, temveč tudi družbene sprejemljivosti in vpliva tehnologije na družbeno dinamiko onkraj gole tržne uspešnosti. Pri VSD pa lahko vključevanje družbene tehnologije prispeva k sistematičnemu vključevanju družbenih potreb in vrednot v fazo načrtovanja, s poudarkom na uporabniških izkušnjah in širših družbenih posledicah. Ključna naloga raziskovalcev in politikov je zato oblikovanje regulativnih okvirjev, ki bodo omogočali uporabo družbene tehnologije na način, ki bo bodisi neposredno bodisi posredno na podlagi ocenjevanja izdelanosti tehnologij podpiral družbeno blaginjo in preprečeval možne zlorabe.

Pametni stroji kot epistemske in družbene tehnologije

Tradicionalne tehnologije so običajno opredeljene kot orodja ali sistemi, ki širijo človekove fizične ali organizacijske zmožnosti za doseganje praktičnih ciljev. Temeljijo na uporabi obstoječega znanja za doseg praktičnih ciljev, kot so avtomatizacija, optimizacija procesov in gradnja infrastrukture. Osredinjajo se na izvajanje nalog v produkcijskih, distribucijskih in komunikacijskih procesih. Tradicionalne komunikacijske tehnologije, kot so tiskarski stroj, radio in televizija, so na primer razširile dostop do informacij, vendar ostajajo usmerjene v izvajanje vnaprej določenih procesov.

Epistemske tehnologije, kot je na primer umetna inteligenca, pa v temelju spreminjajo ta pristop, saj ne le izboljšujejo dostopnost in praktično uporabo znanstvenih spoznanj za izvajanje vnaprej določenih nalog, temveč aktivno prispevajo k njegovemu nastajanju. Njihova vloga je razširitev in transformacija človeških kognitivnih sposobnosti, kot so zaznavanje, sklepanje in odločanje. Epistemske tehnologije, kot so strojno učenje, obdelava naravnega jezika in priporočilni algoritmi, prepoznajo trende in vzorce v ogromnih količinah podatkov, ki presegajo človeške kognitivne sposobnosti. Epistemske tehnologije ustvarjajo vpogled in napovedi, ki vplivajo na odločanje in preobrazbo znanstvenih, kulturnih in industrijskih procesov, ter delujejo kot epistemološki agenti, ki se lahko 'učijo' in prilagajajo brez neposrednih človekovih navodil.

Medtem ko na primer parni stroj 'uporablja' načela termodinamike za ustvarjanje mehanskega dela, algoritmi umetne inteligence analizirajo velike količine podatkov, da bi razkrili nove povezave, postavili hipoteze ali ponudili rešitve kompleksnih problemov. Umetna inteligenca ne le posnema družbene odnose, temveč jih aktivno oblikuje – od algoritmov, ki določajo, kdo dobi kredit ali službo, do avtomatiziranih odločitev v izobraževanju in zdravstvu. Njena vloga presega zgolj tehnično orodje; postaja epistemski arhitekt družbe, ki redefinira temelje znanja, moči in participacije. Zlasti nevronske mreže in sistemi, kot so veliki jezikovni modeli (npr. ChatGPT), navidezno nevtralizirajo ideološke pristranskosti kot »objektivno optimizacijo«, s čimer naturalizirajo družbene neenakosti. Hkrati pa generativna UI radikalno demokratizira ustvarjanje vsebin, a za ceno novih oblik nadzora (npr. sledenje čustvenim odzivom na podlagi interakcij).

Čeprav epistemske tehnologije pri ustvarjanju znanja prinašajo izjemne obete, pa sprožajo protislovne procese. Med ključnimi izzivi je epistemska netransparentnost – procesi, na katerih temeljijo modeli strojnega učenja, pogosto niso v celoti razumljivi niti njihovim ustvarjalcem. Poleg tega se pojavljajo etične dileme, povezane s pristranskostjo, z odgovornostjo ter družbenimi posledicami odločitev, ki jih gene-

rirajo algoritmi. Obseg in kompleksnost podatkov, na katerih temeljijo te tehnologije, zahtevata kritičen premislek o njihovih omejitvah ter vplivu na naše razumevanje.

Ta dialektika – hkratna emancipacija in dominacija – umešča umetno inteligenco kot sodobno družbeno tehnologijo v središče boja za družbeno agendo, kar zahteva ne le regulativni odziv, temveč tudi kolektivno razvijanje alternativnih modelov (npr. odprtokodne založniške platforme ali algoritmi, razviti v okviru digitalnih zadrug). Kot taka ni le orodje upravljanja, temveč materializacija družbenega sporazuma – ali njegovega razpada – v 21. stoletju.

Pod vplivom hitrega razvoja umetne inteligence in strojnega učenja se močno uveljavlja prepričanje, da bi morali ljudje sprejeti napredne mehanizme strojnega učenja, da bi dosegli učinkovitejše in usklajeno delovanje. Schmidt in Thrun (2016) poudarjata prednosti strojne inteligence pri kolektivnem učenju, kar ponazarjata s samovozečimi avtomobili, ki se učijo iz skupnih izkušenj. V nasprotju z njimi se ljudje učijo na lastnih napakah, ki jih pogosto ponavljajo. Strojna inteligenca je predstavljena kot apoteoza kolektivnega delovanja, v katerem vsi stroji v omrežnem sistemu harmonično delujejo, saj imajo vsi enako 'razumevanje' – vsak od njih zna, kar znajo vsi – in delujejo usklajeni kot en organizem za doseg enakih rezultatov z največjo učinkovitostjo. Stroji stojijo ali padajo skupaj; vsi skupaj imajo prav ali ne.

Shoshana Zuboff (2019) to apoteozo pametnih strojev in simbioze ljudi in strojev označuje za vizijo »instrumentarne družbe«, v kateri človeško vedenje posnema procese pametnih strojev. Stroji posnemajo drug drugega, zato moramo tudi mi početi enako, če hočemo biti učinkoviti. V instrumentarni prihodnosti svet strojev in družbeni svet delujeta v harmoniji, ljudje pa posnemajo vrhunske učne procese pametnih strojev. »To posnemanje ni namenjeno vračanju k taylorizmu množične proizvodnje ali Chaplinovemu nesrečnemu delavcu, ki ga pogoltne mehanski red. Nasprotno, ta recept za simbiozo ubira drugačno pot, na kateri človeška interakcija odraža odnose pametnih strojev, saj se posamezniki naučijo razmišljati in delovati tako, da posnemajo drug drugega ...«

SKLEP: K TEHNOLOŠKI ZRELOSTI V OCENJEVANJU IN RAZVOJU TEHNOLOGIJ

S prehodom v digitalno družbo, v kateri prevladujejo neopredmetene – epistemске in družbene – tehnologije, postaja vse pomembnejše spoznanje, da uspešnosti implementacije inovacij ni mogoče meriti zgolj s tehnično izvedljivostjo in sposobnostjo tržnega preživetja, ampak je treba upoštevati tudi družbeno sprejemanje tehnologij in usklajenost tehnologije z družbenimi vrednotami in potrebami. Ideja družbene tehnologije predstavlja širše razumevanje pomena tehnologije in postavlja v ospredje družbene mehanizme spreminjanja družbenih procesov. Vodilo ideje družbene tehnologije je spoznanje, da je uspešnost tehnoloških inovacij pogosto odvisna tudi od njihove etične povezljivosti z družbo ter da so za družbeni napredek družbene inovacije lahko enako pomembne kot tehnološke.

Preusmeritev pozornosti z ozkega razumevanja 'pripravljenosti tehnologije' na kompleksnejša pojma 'tehnološka pripravljenost' in 'družbena tehnologija' omogoča bolje razumeti ključno vlogo, ki jo imajo družbeni sistemi in inovacije pri razvoju in

implementaciji novih tehnologij. Širše razumevanje tehnologije in ocenjevanja njene razvitosti, ki enakovredno obravnava snovne, epistemske in družbene tehnologije, zagotavlja družbeno konstruktivno in trajnostno vključevanje inovacij in tehnologij. Tak pristop prispeva k vključevanju družboslovnega in humanističnega znanja v procese načrtovanja, vrednotenja in implementacije tehnologij, kar lahko prispeva k usklajevanju tehnološkega napredka z razvojem in blaginjo skupnosti ter s tem povečuje legitimost, učinkovitost in družbeno sprejemljivost tehnoloških inovacij.

Tabela 3: Dve ravni ocenjevanja in razvoja tehnologij

Strategija ocenjevanja tehnologije	Družbeno-etični okvir in cilj ocenjevanja	
	Instrumentalni okvir Cilj: neposredna uporabnost in učinkovitost	Vrednotno-refleksivni okvir Cilj: skladnost posledic z načeli, normami in družbenimi vrednotami
Funkcionalno-pragmatična	Tržno <i>ocenjevanje glede na zanesljivost, ekonomsko korist, donosnost in konkurenčnost</i>	Profesionalno <i>ocenjevanje vpliva tehnologije na družbo in stroko z vidika strokovnih standardov in profesionalne etike</i>
Normativno-regulativna	Administrativno <i>ocenjevanje glede na pravne, regulativne in organizacijske zahteve (skladnost z zakonodajo in s standardi)</i>	Normativno-kritično <i>ocenjevanje vpliva na posameznike, skupnosti in vrednote s širšega družbenega, etičnega ali normativnega vidika in predlogi sprememb v zasnovi, razvoju in regulaciji uporabe tehnologij</i>

Družbenoteoretski in etični okvir razvoja in ocenjevanja tehnologij je temelj za premislek o *tehnološki zrelosti*, ki odpira ključna vprašanja o odnosu med tehnologijo in družbo: Za koga oziroma komu je tehnologija namenjena in komu bo najbolj koristila? Za kaj je tehnologija potrebna in uporabna ter katere cilje želimo z njo doseči? Za učinkovito razširitev obzorja ocenjevanja tehnologije onkraj njenega tržnega potenciala in upoštevanje družbenega vpliva je treba določiti etični okvir ocenjevanja tehnologij. V tem pristopu postane ključno vprašanje razlikovanja med instrumentalnim in refleksivnim ocenjevanjem tehnologij, ki se navezuje na Webrovo razlikovanje instrumentalne in vrednotne racionalnosti (Weber 1920/2019) in Burawoyjevo prilagoditev Webrove tipologije za ugotavljanje namembnosti (družboslovnega) raziskovanja z vidika ciljev in sredstev (Burawoy 2007).

Webrovo razlikovanje se nanaša na to, ali je človekovo delovanje usmerjeno k doseganju specifičnih, racionalnih ciljev (instrumentalna racionalnost) ali pa je podrejeno vrednotam, ki imajo za delovanje intrinzični pomen (vrednotna racionalnost). Burawoy je to razlikovanje uporabil za postavitev tipologije družboslovnega raziskovanja glede na raziskovalne cilje in sredstva. Za Burawoyja (2007) je razlika med instrumentalnim in refleksivnim raziskovanjem določena z odgovorom na vprašanje, ali se raziskovanje osredinja na iskanje ustreznih sredstev za doseganje samoumevnih ali vnaprej določenih ciljev (instrumentalnost) ali pa vključuje razpravo o samih ciljih (refleksivnost). Ta okvir lahko uporabimo tudi za ocenjevanje 'pripravljenosti' teh-

nologije z iskanjem odgovora na vprašanje, ali je cilj tehnologije zgolj optimizacija učinkovitosti, produktivnosti in dobičkonosnosti (instrumentalnost) – ali pa spodbuja globlji družbeni razmislek o impliciranih vrednotah in ciljih (refleksivnost).

Instrumentalno ocenjevanje tehnologij poudarja učinkovitost in funkcionalnost, pogosto daje prednost komercialnim in administrativnim interesom, ki postavljajo v ospredje tržno uspešnost in organizacijsko produktivnost. Tehnologija se ocenjuje in razvija glede na njeno neposredno uporabnost, učinkovitost in praktične koristi. Cilj je izboljšati delovanje, povečati produktivnost, zmanjšati stroške ali pridobiti konkurenčno prednost.

Ta perspektiva vrednoti tehnologije kot orodja in sisteme, ki delujejo učinkovito za doseganje določenih, merljivih ciljev, pogosto na račun etičnih vidikov in brez upoštevanja širših družbenih posledic. Programska oprema za upravljanje dobavnih verig na primer optimizira logistiko, tako da zmanjšuje stroške in povečuje učinkovitost, kar koristi podjetjem in delničarjem, pa čeprav to vodi do izgube delovnih mest in tržnih monopolov. Podobno so algoritmi družabnih medijev zasnovani tako, da usmerjajo in intenzivirajo pozornost uporabnikov ter povečajo prihodke od oglaševanja, čeprav ob učinkovitem doseganju komercialnih ciljev prispevajo tudi k širjenju dezinformacij in oblikovanju epistemskih enklav. V skrajni konsekvenci instrumentalistično razumevanje odnosa med tehnologijo in družbo vodi do prepričanja, da posnemanje procesov v pametnih strojih vodi v učinkovitejšo delovanje ljudi in organizacijo družbe.

Refleksivno ocenjevanje tehnologij pa nasprotno zasleduje širše cilje, ki so sami po sebi 'pravi' ali koristni za družbo. Tehnologija se ocenjuje in razvija glede na širše družbene, etične in normativne posledice, ki jih povzroči. Cilj je razumeti njen vpliv na posameznike, skupnosti in vrednote ter identificirati morebitne nevarnosti ali nepravilnosti.

Ta perspektiva spodbuja tehnološki razvoj z jasnim poudarkom na družbenih koristih in etičnih standardih. Opira se na poklicne standarde in družbenokritično angažiranost ter spodbuja deležnike k razmisleku o etičnih posledicah tehnološkega napredka in razvoju 'refleksivnih tehnologij'. Primer refleksivne tehnologije je odprtokodna programska oprema, ki spodbuja sodelovanje, preglednost in inovacije, hkrati pa omogoča pravičen dostop do tehnologije, ali tehnologija za zagotavljanje obnovljivih virov energije, kot so sončni kolektorji, ki kot cilj implicira boj proti podnebnim spremembam in daje prednost okoljski trajnosti. Refleksivne tehnologije so zasnovane z zavezanostjo etičnim standardom in so njihov cilj dolgoročne družbene koristi, ki pogosto vključujejo okoljsko trajnost, pravičen dostop in etično vladovanje.

Razlike v ciljih in učinkih razvoja in uporabe tehnologij odpirajo temeljna vprašanja o tem, kdo ima koristi od tehnološkega napredka (inovacij) in katera etična načela vodijo njegovo usmeritev. Instrumentalno ocenjevanje tehnologij služi ciljem dobičkonosnosti in gospodarske rasti, ki jih je mogoče doseči s povečevanjem učinkovitosti in produktivnosti, vendar pogosto na račun negativnih etičnih in družbenih posledic. V nasprotju s tem refleksivno ocenjevanje tehnologij kritično vrednoti njihovo (ne)skladnost z družbeno koristnimi vrednotami in nas spodbuja k razmisleku o dolgoročnih ciljih in posledicah tehnološkega napredka. To razlikovanje in potreba po »etičnem ravnesju« sta postala še posebej ključna v obdobju, ko velike zasebne

korporacije vedno bolj prevladujejo v razvoju tehnologij na številnih področjih, kot so komunikacije, zdravstvo in raziskovanje vesolja. Velika koncentracija razvoja tehnologij v rokah korporacij zahteva premislek o tem, ali bodo inovacije služile predvsem skupnemu dobremu ali komercialnim interesom.

Vse to povečuje potrebo po oblikovanju okvirov za holistično ocenjevanje tehnologije. S poudarjanjem tako funkcionalnosti kot družbene blaginje – z vključevanjem instrumentalne in refleksivne perspektive v etični okvir ocenjevanja tehnologije – lahko bolje razumemo dvojno vlogo tehnologije v družbi: kot orodja za doseganje specifičnih rezultatov in kot katalizatorja za širši družbeni napredek in blagostanje. Tak celovit pristop nas spodbuja k razmisleku o posledicah naših tehnoloških odločitev in vrednotah, ki jih podpirajo. Celostni pristop lahko spodbuja prizadevanje za doseganje dveh temeljnih ciljev razvoja tehnologij in inovacij: napredek tehnoloških zmogljivosti in podpora odgovornejšemu, raznovrstnejšemu in pravičnejšemu razvoju in uporabi tehnologije.

V tem okviru je pomen družbenih tehnologij dvojen. Po eni strani so pomemben neposredni dejavnik družbenega razvoja na številnih področjih, po drugi pa povezujejo različna tehnološka področja, ki skupaj oblikujejo družbene odnose in procese. Zato bi morale družbene tehnologije imeti pomembno vlogo v ocenjevanju razvitosti tehnologij in družbene tehnološke razvitosti. Hkrati pa bi morala odgovorna politika tehnološkega razvoja in inovacij spodbujati tudi sistematičen razvoj družbenih tehnologij in posebej epistemskih tehnologij – ne le tistega dela, ki je usmerjen v ali z umetno inteligenco.

OPOMBE

- 1 To lepo kažeta Nobelovi nagradi za fiziko in kemijo leta 2024, ki so ju prejeli utemeljitelji razvoja umetne inteligence in strojnega učenja, kar je ob priznavanju izjemnih dosežkov nagrajencev sprožilo tudi pomisleke o tem, ali je mogoče nagrade za fiziko in kemijo podeljevati za dosežke na področju umetne inteligence in računalništva.
- 2 Angleška izraza se sicer pogosto uporabljata kot sopomenki, v Sloveniji pa vlada še večja pojmovna zmeda, saj se kot sopomenke v prevajanju »technology readiness« uporabljajo izrazi »raven tehnološke razvitosti« (Spirit, ARIS), »stopnje pripravljenosti tehnologije«, »tehnološka stopnja pripravljenosti« in »stopnje pripravljenosti za nove tehnologije«. V Zakonu o znanstvenoraziskovalni in inovacijski dejavnosti (ZZrID 2021) najdemo krožno (»obscurum per obscurius«) definicijo, ki ji sledi ostenzivna definicija: »Tehnološka stopnja pripravljenosti (technology readiness level ali TRL) je raven tehnološke razvitosti, ki je opredeljena v devetih stopnjah glede na končni cilj uporabe.« <https://pisrs.si/pregledPredpisa?id=ZAKO7733>.
- 3 Glejte portal Fran.si: <https://www.fran.si/iskanje?View=1&Query=izdelanost>; <https://www.fran.si/iskanje?View=1&Query=dovr%C5%A1enost>.
- 4 Glejte tudi STOrYlines of futuRe extreMes (PerfectSTORM, 2021–2026). <https://cordis.europa.eu/project/id/948601>.

CITIRANA LITERATURA

- Alvarado, Ramon. 2023. »AI as an Epistemic Technology«. *Science and Engineering Ethics* 29, 32. doi: 10.1007/s11948-023-00451-3.
- Bruno, Ilenia, et al. 2020. »Technology Readiness Revisited: A Proposal for Extending the Scope of Impact Assessment of European Public Services«. *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance (ICEGOV2020)*, Atene, Grčija, 11.–13. marec. https://ec.europa.eu/isa2/sites/isa/files/technology_readiness_revisited_-_icegov2020.pdf
- Burawoy, Michael. 2007. »Open the Social Sciences: To Whom and For What?«. *Portuguese Journal of Social Sciences* 6 (3): 137–146.
- Deliberative Democracy Lab. b. l. *What Is Deliberative Polling*. <https://deliberation.stanford.edu/what-deliberative-pollingr>.
- Derksen, Maarten, in Tjardie Wierenga. 2013. »The History of 'Social Technology', 1898–1930«. *History and Technology* 29 (4): 311–330. <https://doi.org/10.1080/07341512.2013.876247>.
- Drucker, Peter F. 1987. »Social Innovation – Management's New Dimension«. *Long Range Planning* 20 (6): 29–34.
- Fishkin, J. S., R. C. Luskin in R. Jowell. 2000. »Deliberative Polling and Public Consultation«. *Parliamentary Affairs* 53 (4): 657–666. DOI:10.1093/pa/53.4.657.
- Friedman, Batya, in David G. Hendry. 2019. *Value Sensitive Design: Shaping Technology with Moral Imagination*. MIT Press.
- Friedman, Batya, Peter H. Kahn in Alan Borning. 2006. »Value Sensitive Design and Information Systems«. V *Human-Computer Interaction and Management Information Systems: Foundations*, ur. P. Zhang in D. Galletta, 348–372. M. E. Sharpe.
- Hakkarainen, Kai, Ritva Engeström, Sami Paavola, Pasi Pohjola, Timo Honkela. 2009. »Knowledge Practices, Epistemic Technologies, and Pragmatic Web«. V *I-KNOW'09 and I-SEMANTICS'09: the 4th AIS SigPrag International Pragmatic Web Conference Track*, 683–694. Verlag der Technischen Universität Graz.
- Héder, Mihály. 2017. »From NASA to EU: The Evolution of the TRL Scale in Public Sector Innovation«. *The Innovation Journal: The Public Sector Innovation Journal* 22 (2), article 3.
- Kant, Immanuel. 1781/1956. *Kritik der reinen Vernunft*. (Kant Werke 3-4). Ur. W. Weischedel. Frankfurt: Suhrkamp.
- Kaplan, Leah R., et al. 2021. »Designing Participatory Technology Assessments: A Reflexive Method for Advancing the Public Role in Science Policy Decision-making«. *Technological Forecasting and Social Change* 171. doi: 10.1016/j.techfore.2021.120974.
- KTH. b. l. *KTH Innovation Readiness Level*. <https://kthinnovationreadinesslevel.com/>.
- McKinsey Global Institute. 2017. *Digitization, AI, and the Future of Work: Imperatives for Europe. Briefing note prepared for European Union Tallinn Digital Summit*. September 2017. <https://www.mckinsey.com/~media/McKinsey/Business%20Functions/Operations/Our%20Insights/Overview/2017%20in%20review/Whats%20next%20in%20digital%20and%20AI/Digitization%20AI%20and%20the%20future%20of%20work%20imperatives%20for%20europe/MGI-Tallinn-Briefing-Note-1213.pdf>.
- Moulaert, Frank, Diana MacCallum, Abid Mehmood in Abdelillah Hamdouch (ur.). 2013. *The International Handbook on Social Innovation: Collective Action, Social Learning and Transdisciplinary Research*. Edward Elgar Publishing.
- Park, Robert E. 1938. »Reflections on Communication and Culture«. *The American Journal of Sociology* 44 (2): 187–205.
- Pultarov, Tereza. 2024. »Uncovering the Ripple Effects of the Climate Crisis«. *Horizon Magazine*.

<https://horizon.scienceblog.com/2708/uncovering-the-ripple-effects-of-the-climate-crisis/>.

Schmidt, Eric, in Sebastian Thrun. 2016. »Let's Stop Freaking Out About Artificial Intelligence«. *Fortune*, 28. junij 2016. <http://fortune.com/2016/06/28/artificial-intelligence-potential>.

Schwab Foundation. 2024. *AI for Impact: The Role of Artificial Intelligence in Social Innovation*. In collaboration with EY and Microsoft. Cologne/Ženeva: World Economic Forum. https://www3.weforum.org/docs/WEF_AI_for_Impact_Social_Innovation_2024.pdf.

Science Direct. b. l. *Social Technology*. <https://www.sciencedirect.com/topics/social-sciences/social-technology#:~:text=Social%20Technology%20refers%20to%20the%20use%20of%20technological,behaviors%20by%20impacting%20face-to-face%20communication%20and%20social%20cues>.

Small, Albion W. 1898. »Seminar Notes: The Methodology of the Social Problem. Division I. The Sources and Uses of Material«. *The American Journal of Sociology* 4 (1): 113–144. https://handwiki.org/wiki/Engineering:Social_technology.

Tönnies, Ferdinand. 1887/1999. *Skupnost in družba*. Ljubljana: FDV. Verma, Aditi, in Todd Allen. 2024. »A Sociotechnical Readiness Level Framework for the Development of Advanced Nuclear Technologies«. *Nuclear Technology* 210 (9): 1722–1739. DOI:10.1080/00295450.2024.2336355.

Weber, Max. 1920/2019. *Economy and Society*. Cambridge, Ma: Harvard University Press.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs/Perseus Books.

DRUŽBENI KONTEKSTI TEHNOLOŠKEGA RAZVOJA

Mirjana Nastran Ule

UVOD

V prispevku razpravljam o tem, kaj je družbeno konstruktivna uporaba znanosti in tehnologije v sedanjem svetu ter kako jo usmerjati, da bo to ustrezalo potrebam ljudi, ohranjanju etičnih standardov, človečnosti in socialne kohezivnosti. Predstavim osnovne trende družbenega razvoja v obdobju globalizacije in se sprašujem, kako uporabiti znanost in tehnologijo v reševanju družbenih problemov. Prispevek zaključujem z razpravo o izobraževanju kot ključni točki, na kateri se začne razvijati zavest o odgovorni uporabi sodobnih tehnologij v vsakdanjem življenju in sodobni družbi.

V besedilu o tehnologiji praviloma razmišljam kot o skupku znanj, spretnosti in sredstev za izdelavo, uporabo in razvoj orodij in naprav, s katerimi ljudje skušamo obvladovati naravni in družbeni svet, v katerem živimo. Govorim predvsem o tehnologijah, ki omogočajo sistematični nadzor nad okoljem in ljudmi, ki upravljajo z določenimi orodji, napravami in procesi, kot so sodobne informacijske in digitalne tehnologije. Tehnologijo ločim od tehnike, ki zajema naprave, objekte, postopke in prakse, ki so namenjeni spreminjanju okolja (Salomon 1984).

Teza, ki jo zagovarjam v razpravi, je, da bi morali pri razsojanju o uvedbi novih tehnologij nujno upoštevati tudi družboslovne in humanistične vidike ter merila, in to predvsem pri oceni odgovornosti za uporabo tehnologije kot posebne oblike družbene, in ne le ožje tehnološke moči. Zato se ocenjevanje primernosti tehnologije v dani situaciji nikakor ne more in tudi ne sme zožiti zgolj na vprašanje o morebitni ekonomski ali tržni uspešnosti njenih proizvodov ali storitev. To še posebej velja za sodobne epistemske in socialne tehnologije. Te neposredno (pri)zadevajo vse oblike človeške komunikacije, rabe jezika, oblikovanja odnosov, sebstva, institucij.

Ko govorimo o razmerju med družbenim in tehnološkim razvojem, bi morali med drugim odgovoriti tudi na vprašanje, katero razumevanje človeškega življenja je utelešeno v prevladujočih tehnoloških ureditvah. Izhodiščno vprašanje, ki si ga med drugim zastavlja sodobno družboslovje, je, kaj je družbeno konstruktivna uporaba znanosti v sedanjem svetu. To vprašanje neizogibno implicira nadaljnje spraševanje o tem, kako usmerjati uporabo znanosti in tehnologij tako, da bo to ustrezalo optimalnim razvojnim težnjam in potrebam ljudi. Tako sociologi Anthony Giddens, Ulrich Beck in Christopher Lash razpravljajo o refleksivni modernizaciji, pri čemer raziskovanje družbenega in ekonomskega upravljanja s krizami in z razvojnimi alternativami po svojem pomenu presega reševanje znanstveno-tehnoloških problemov (Beck, Giddens in Lash 1994). V tem se po njihovem mnenju kaže današnja strateškost družboslovja.

NEOLIBERALNA TRANSFORMACIJA DRUŽBENEGA REDA TER POSLEDICE ZA DRUŽBENI IN TEHNOLOŠKI RAZVOJ

Ameriški sociolog Immanuel Wallerstein, ki v študijah analizira sodobni svet in govori o možnih alternativah prihodnjega razvoja, že konec devetdesetih let zapiše (1999, 32): »Živimo v času prehoda iz obstoječega svetovnega sistema v drug svetovni sistem ali sisteme. Ne vemo sicer, ali bo sprememba na boljše ali na slabše. Vemo pa, da bo to obdobje zelo težko za vse ... To bo čas spopadov ter tistega, v čemer bodo številni videli zlom moralnih sistemov.«

Po mnenju Wallersteina (1999, 58) je svetovni sistem pripeljal do največje polarizacije natanko zaradi tistega, kar naj bi bila njegova največja odlika, to je zaradi neverjetnega širjenja proizvodnje vrednosti. Primat nenehne akumulacije kapitala ustvarja neenakopravnost in posledično napetost, nezadovoljstvo, nezaupanje, razkroj skupnosti. V takih razmerah postajajo ključnega pomena spoznanja in raziskovalne smeri, ki nam omogočajo premislek o značaju svetovnega sistema in o razvojnih alternativah, ki jih vsebuje.

Podobno sodobni svetovni sistem ocenjujejo tudi mnogi drugi pomembni družboslovni avtorji. Thomas Piketty v knjigi *Kapital v 21. stoletju* trdi, da je globalizirani kapital resnično postal svobodnejši v kroženju po svetu, a hkrati spodkopava demokracijo in socialno državo. Navaja, da sodobna globalizirana ekonomija drvi v utrjevanje neke vrste rentniškega kapitalizma, v katerem bo največji del bogastva zajet v podedovanem bogastvu najpremožnejših. Večina preostalega prebivalstva bo primorana tako ali drugače služiti manjšini najpremožnejših (Piketty 2014, 267). Po njegovem mnenju se nam dogaja nekakšna barbarska alternativa kapitalizmu, ki vse človeštvo peha v razne oblike odvisnosti od čedalje manjšega števila superbogatašev in njihovih tveganih finančnih transakcij.

Kako je do tega prišlo? Oglejmo si to na primeru sprememb družbenega reda. Družbeni red vključuje družbene strukture in institucije, družbene odnose, družbene interakcije in vedenje ter kulturne značilnosti, kot so norme, vrednote in prevladujoča prepričanja. Sprememba družbenega reda torej zajema tako spremembe družbenih struktur in odnosov kot tudi spremembe utrjenih norm, vrednot in prevladujočih prepričanj. Lawrence (1944, 476) je že leta 1944 v članku *What Is Social Order?* pisal o potrebi po novem družbenem redu v Evropi po koncu II. svetovne vojne in med drugim navaja:

Izkušnje vojne so pripeljale do spoznanja, da je človeško življenje najvišja vrednota in da je treba ustvariti družbeni red, ki krepi zavest o medsebojni povezanosti in soodvisnosti. Danes se pod konflikti in neredi vojnega sveta vse bolj krepi zahteva po ohranjanju človeka in po družbeni ureditvi, posvečeni človeškim potrebam in vrednotam /.../ Če upamo, da bomo razvili družbeni red, v katerem bo mogoče neverjetno in hitro rastoče znanstveno znanje in tehnologijo pametno in učinkovito uporabiti za človekove potrebe in vrednote, v katerem bo mogoče vzpostaviti in ohraniti nekakšno dostojno družbeno življenje in v katerem bo mogoče uresničiti naše temeljne demokratične težnje po priznanju in ohranitvi človeškega

dostojanstva, moramo poiskati nov okvir idej in prepričanj ter človeških vrednot, ki so bile tako dolgo onemogočene in poražene.

Prav to spoznanje je bilo podlaga za razvoj socialne države oziroma države blaginje v Evropi po drugi svetovni vojni, v določeni meri in v različnih oblikah pa tudi po vsem svetu. Socialna država je kapitalistični izum, vendar ne spodkopava trga, ampak se osredinja na neenakosti, ki jih ustvarja trg. V klasični državi blaginje je bilo državno poseganje na trg razumljeno kot naložba, ki vzdržuje socialno kohezivnost, zagotavlja socialni mir ter spodbuja vključenost v družbo. Čim bolj so namreč ljudje vključeni v družbo, tem več ustvarjalnega potenciala je na voljo za napredek družbe.

Vendar so se že v sedemdesetih in osemdesetih letih prejšnjega stoletja začela pojavljati nasprotovanja socialni državi, predvsem ideji o enakomernejši porazdelitvi blaginje, ki jo zagotavlja socialna država. Po mnenju kritikov, kot je bil na primer ekonomist Friedrich Hayek, naj bi bilo tako prerazporejanje diskriminatorno in nezdružljivo z načeli svobodne družbe (berite svobodnega trga) (Hayek 1991). Podobni argumenti proti socialni državi so se pojavili v Veliki Britaniji v obdobju vlade Margaret Thatcher. Glavni argument je bil enak Hayekovemu, in sicer da je treba vlogo države pri gospodarskem razvoju zmanjšati, saj je vsakršno poseganje države razumljeno kot zloraba državne moči in ovira za prosti trg.

Po sprejetju lizbonske strategije marca 2000, ko so se voditelji vlad držav članic EU dogovorili o novem strateškem cilju Evropske unije do leta 2010 – postati najbolj konkurenčno gospodarstvo na svetu – je to postala prevladujoča ideologija v Evropi. Izhodišče za to tezo je bilo, da Evropa ne more biti konkurenčna, če ne odpre trga in ne odstrani ovir za njegovo delovanje. In ena od glavnih ovir je bila po njihovem prepričanju prav socialna država. Od tedaj smo priče nenehnemu rušenju temeljev socialne države ter socialne kohezivnosti in povezanosti. Diskurz pravic je zamenjal diskurz obveznosti, ki je povezan z individualizacijo odgovornosti tudi za strukturna in ekonomska tveganja, ki so zunaj moči posameznika oz. posameznice.

Spremembe družbenega reda, ukinjanje vloge države pri ekonomskem razvoju, ki so sledili tem ukrepom, pa so povzročili kopičenje problemov svetovnih razsežnosti, ki povzročajo občutek neobvladljivosti življenja, negotovosti, nestabilnosti, nezaupanja in strahu pred prihodnostjo. Beck (2000) je konec devetdesetih let ugotavljal, da imamo opravka z novo in neprijetno zaznavo, ki se označuje z izrazom globalizacija. Gre za sistemski učinek razvoja kapitalističnega gospodarskega sistema na svetovni ravni, ki bistveno vključuje nenehni tehnološki razvoj kot svoje gonilo in kot svoj pogoj.

Kot ugotavlja Harvey (2012, 93) v študijah globaliziranega neoliberalizma, se neoliberalizem v svojih teorijah sicer zanaša na to, da bodo prisilne moči konkurence gnale iskanje novih produktov, produkcijskih metod in organizacijskih oblik, vendar pa se ta gon k vedno novim tehnološkim spremembam hitro spremeni v fetiš, v slepo vero, da naj bi za vsako težavo obstajala tehnološka rešitev. Vendar tehnološki razvoj lahko uide izpod nadzora, na primer tedaj, ko začnejo trgi proizvajati nove produkte, ki še sploh nimajo trga, ali ko spretni prevaranti s svojimi »uspehi« povzročijo nastanek ekonomskih kolapsov ipd.

Zdi se, da nihče ne more več imeti pod nadzorom celote svetovnega dogajanja, niti ni jasno, kakšen bi lahko bil nadzor. Posledica je svetovni proces nove stratifikacije, v poteku katere nastaja nova, sociokulturna, samoreproducirajoča hierarhija. Bauman (2006, 316) opozarja, da samo 22 odstotkov svetovnega bogastva pripada nerazvitemu svetu, v katerem živi 80 odstotkov prebivalstva. Piketty in drugi kritiki ugotavljajo še bolj drastične delitve, da namreč en odstotek najpremožnejših po vsem svetu poseduje več kot 50 odstotkov svetovnega kapitala.

Globalizacija je povzročila tudi, da se družbena nasprotja vse bolj lomijo na lju-deh. To pa ima svojo izrazito temno plat, in sicer v tem, da postajamo vse bolj obremenjeni z družbenimi konflikti in nasprotji, ki se zdaj vse bolj neposredno lomijo na posameznikih brez socialnih odbijačev, kakršne je v prejšnjih obdobjih modernizacije predstavljala socialna država s svojimi institucijami. Medtem ko je bila revščina prej usoda razreda, je zdaj postala individualizirana, osebna usoda. Problem je, da individualizirano revščino razumemo bolj v psiholoških okvirih, kot so na primer izguba samospoštovanja, dostojanstva, občutki sramote in osebne krivde za svoj neugodni življenjski položaj.

Pritisk, da bi porušili še zadnje preostale ovire, ki so zoperstavljene svobodi denarja in z denarjem proizvedenim tehnologijam in informacijam, sovpada s pritiskom, da se izkopljejo novi prepadi in zgradijo novi zidovi (protimanjšinski, protiimigrantski zakoni, nacionalni zakoni). Kapital in bogastvo se globalizirata, beda se lokalizira (Bauman 2006). Vzpostavljajo se vase zaključeni krogi revščine in socialnega izključevanja (marginalizacije). Revščina in brezposelnost vse manj sledita razrednim stereotipom, vedno težje ju identificiramo in politično organiziramo.

Ko se temelji socialne države ter solidarnosti, soodvisnosti in povezanosti rušijo, postaja družbeni red krhek. S tem se ruši tudi zaupanje v državo in v družbene institucije sploh. Ker zaupanje v državo upada, ljudje iščejo druga sidra varnosti in osebne identitete. Ta sidra niso več solidarnost in kohezija, temveč strah, jeza, agresija. Vse pogostejše se iščejo in najdejo ta sidra v osebah, ki legitimirajo paternalistični in izključevalni odnos do drugih in drugačnosti. Povečuje se produkcija nezaželenih skupin ljudi, ki postajajo odvečni, nevidni. Na ta način se v Evropi in drugod po svetu širi populistični duh nestrpnosti in sovražnosti. Zdi se, da vse bolj tonemo v množične oblike družbene anomalije in negotovosti.

DRUŽBENE IN ETIČNE DILEME SODOBNEGA TEHNOLOŠKEGA RAZVOJA

V tej situaciji se je tudi mnogim družbenim teoretikom zdelo, da je izhod iz nastalih težav v večji naslonitvi na znanost in še bolj na sodobne tehnologije. Dejstvo je, da intenzivno globalizacijo dejansko spremljajo in v veliki meri tudi poganjajo vedno nove tehnološke inovacije. Predvsem v začetnem obdobju se je pojav novih tehnologij zdel kot obet tudi v etičnem smislu, kot preseganje družbenih prisil, kot nova svoboda možnosti. Takrat so se pojavili izrazi, kot so informacijska družba, družba znanja, postindustrijska družba, za označevanje temeljnih novosti, ki naj bi jih prinašala nova doba (Reynolds in Szerszynski 2016).

Teoretik globalizacije Thomas Friedman je v odmevni knjigi z naslovom *The World Is Flat* razvijal tezo o tem, da bo vedno bolj globalno integrirani svet ob po-

moči informacijskih tehnologij neizogibno privedel do sveta brez lomov, brez vrhov in padcev (Friedman 2005). Vendar so se ob soočenju s posledicami hitrega uvajanja novih tehnologij brez nujne družbene in psihološke refleksije tovrstna pričakovanja počasi razblinila.

Tehnologije imajo vgrajene družbene odnose. Vključujejo konkurenčne oblikovalske, ekonomske, inženirske, proizvodne in marketinške interese. Različni interesi so pogosto močno sporni, na primer tedaj, ko se zožujejo le na povečevanje profitnih stopenj in zanemarjajo možne škodljive ekološke učinke množične uporabe določenih tehnologij ali negativne psihosocialne učinke rastočih družbenih neenakosti. Tedaj so pravzaprav nevarnejše dolgoročne posledice slepega sledenja kapital-skim interesom kot pa kapitalski interesi sami po sebi. Kritična analiza razvoja in uporabe tehnologij je nujna zato, da spodbudi odpor demokratične javnosti do tovrstne slepote.

Kot trdijo mnogi sodobni kritični analitiki tehnologij, so tehnologije danes pomemben vir družbene tesnobe. Naš svet je svet tehnološko povzročenega globalnega tveganja (Matthewman 2011, 26). V takih razmerah novi tehnološki potenciali povprečnemu posamezniku predstavljajo prej grožnjo kot priložnost ali obet rešitve iz pasti, ki jih nastavljajo pritiski kapitala, trga dela in potrošništva. Ljudem ne ponujajo uspešnih tehnologij sebstva, s katerimi bi lahko dosegali življenjsko zadovoljstvo, pozitivno samopodobo in modro etično ravnotežje. Michael Foucault je pripisal razvoj tehnoloških inovacij potrebi po novih tehnikah discipliniranja delovne sile, pri čemer je disciplina po Foucaultu (2021) tehnika, s katero se z najmanj stroški zmanjša politična moč in poveča delovna razpoložljivost ljudi.

Nove tehnologije pogosto povzročajo prej neznane etične težave in sočasne konflikte glede dopustnosti in družbene organizacije ter pogosto traja več let, da se nove tehnologije popolnoma vključijo v družbo (Palm in Hansson 2006). Zato je pomembna naloga družbenih ved organizacija razprav o etičnih dilemah, preden se tehnologija preveč zasidra v družbi. V obdobju klasične industrijske moderne se je proces socialne prilagoditve novim tehnologijam lahko začel šele z uvedbo nove tehnologije. To je danes prepozno. Nove tehnologije so prekompleksne in prezahtevne. Terjajo etično refleksijo od ideje dalje. Ker se to ne dogaja, imamo opravka z nenehnim zaostajanjem družbene in etične refleksije o načinih in posledicah uvajanja novih tehnologij.

Če bi bila etična razprava del procesa sistematične refleksije tehnologij od razvoja tehnologij, bi bilo več možnosti za konstruktivno interakcijo med družbenimi vrednotami in tehnološkim potencialom, v katerem bi se lahko z etičnimi problemi ukvarjali že v zgodnji fazi razvoja tehnologij. Podobno opozarja francoski poststrukturalist Bruno Latour, da tehnologije niso le družbeni sistemi, temveč so vsaj latentno tudi moralni sistemi. Tehnologije nam pogosto narekujejo določeno moralo, in sicer tedaj, ko nas prisilijo, da smo se dolžni ravnati po njih (Latour 2002, 258). Prav tako tehnologije ne ponujajo socialnih rešitev, ki bi bile kos temeljnim težavam in problemom, ki bremenijo sodobne družbene odnose. Ko se zavemo etično zahtevnih problemov, se sprašujemo, ali so nove tehnologije rešitev ali problem za družbeno življenje.

Angleški analitik Andrew Feenberg v delu *Transforming Technology* dopolnjuje ta spoznanja z ugotovitvijo o medsebojnem prepletanju družbe, kulture in tehnologije. Med drugim ugotavlja (Feenberg 2002), da bi se lahko sedanji tehnični razum osvobodil svoje trenutne zapletenosti v prevladujočo kapitalsko obliko racionalnosti v kaki alternativni družbeni konstelaciji. Feenberg (ibid., 188) piše:

Vse sodobne industrijske družbe danes stojijo na razpotju, soočene z dvema različnima smerema tehničnega razvoja. Lahko bodisi krepijo izkoriščanje človeka in narave ali pa uberejo novo pot, na kateri tehnologije podpirajo emancipatorne težnje družbenega razvoja. Ta izbira je v resnici politična. Prva pot krepi moč elite. Druga pot zahteva konkretizacijo uporabe tehničnih načel, ob upoštevanju številnih družbenih kontekstov. Ti konteksti odražajo možnosti – vrednote, ki jih je mogoče uresničiti le z novo organizacijo družb.

Hegemonija kapitala ne temelji predvsem na določeni tehniki družbenega nadzora. Bolj temelji na tehnični rekonstrukciji celotnega polja družbenih odnosov, znotraj katerih deluje. Moč poslovneža ali birokrata je prisotna že v razdrobljenosti različnih družbenih sfer proizvodnje, upravljanja in dela, zasebnega in družinskega življenja, ekonomije in politike. Razdrobljene posameznike in institucije lahko organizirajo samo agenti, ki jim dominirajo od zgoraj.

Najpomembnejše vprašanje, ki si ga je treba zastaviti v sodobnih družbah, je torej, kakšno razumevanje človeškega življenja je utelešeno v prevladujočih tehnoloških ureditvah. Če želimo najti odgovor na to vprašanje, moramo po mnenju Feenberga (2002) preseči monokavzalni in linearni pogled na odnos med tehnologijo in družbo, po katerem uporaba določene tehnologije v neki družbi vodi družbeni razvoj in ljudje nimamo druge izbire, kot da sprejmemo kako tehnologijo, če se že razvije in uporablja, ter pristanemo na njene posledice. Kot ugotavlja Feenberg (ibid.), tak pogled ni sprejemljiv. Dejansko prihaja do nenehnega prepletanja in medsebojnih interakcij med razvojem družbe in tehnologije. Ta razvoj pa ni strogo določen, temveč je poln prepričevanja in nejasnosti glede tega, kako razvijati tehnologijo, kako in kje jo uporabiti. Tudi drugi avtorji ugotavljajo, da se mora tehnološka etika razširiti na razumevanje tega, kako tehnologije vplivajo na družbene in kulturne prakse, skupnosti in okolje (Cotton 2014).

Razpolaganje z neko tehnologijo predstavlja razpolaganje z močjo, ki je vedno tudi družbena, ne le materialna moč. Čim pa imamo opravka z razpolaganjem z družbeno močjo, zadenemo v vprašanje odgovornega razpolaganja z močjo. V primeru tehnologije torej za odgovorno razpolaganje s tehnologijo. Tehnologije namreč oblikujejo naš vrednotni sistem, ker nas prisilijo, da se ravnamo po njih. Pri tem razlikujem med neposredno oziroma »materialno« močjo tehnologije, ki se kaže kot zmožnost za nadzorovano in načrtno spreminjanje določenih razmer in okolij, ter posredno oziroma sistemsko močjo tehnologije, ki se kaže kot celostni učinek uporabe določene tehnologije v danih družbenih razmerah, na primer kot dejavnik povečevanja družbenih neenakosti ali kot dejavnik opolnomočenja posameznikov v njihovem življenju in delu.

Praviloma gre pri odgovornem razpolaganju s tehnologijo za posredno odgovornost, in sicer odgovornost za to, da tako snovalci kot uporabniki tehnologije storijo vse, da preprečijo nepremišljeno uporabo ali celo zlorabo kake tehnologije. Kolikor večji je realni potencial za tako uporabo, toliko večja je tudi odgovornost akterjev in ustvarjalcev tehnološke moči za njeno udejanjanje in njene učinke. Z neko tehnologijo praviloma razpolagajo kompetentni strokovnjaki, vendar že nadzor nad zagotavljanjem kompetenc predpostavlja vrsto družbenih procesov in institucij.

Odgovornost za tako ali drugačno uporabo tehnologije je predvsem naloga tistih, ki sprejemajo dolgoročne in družbeno pomenljive odločitve v zvezi z uporabo kake tehnologije. Vendar je naivno pričakovati, da bodo aktualni lastniki ali promotorji kake tehnologije pripravljeni in sposobni prevzemati tovrstno odgovornost nase. Spraševanje, ali je mogoče določeno uporabo tehnologije uskladiti z družbenim vrednotnim sistemom, ostaja tako na plečih kritične javnosti, predvsem iz vrst družboslovja in humanistike ter demokratične civilne družbe.

Harvey (2003, 29) je sestavek o fetišu tehnologije zaključil z ugotovitvijo, da je danes naloga kritične teorije predvsem redefinicija in demistifikacija tehnologije. To je prvi korak k splošnejši emancipaciji od domnevnih zunanjih in naravnih nujnosti ter od ujetosti v svoje lastne fetišne konstrukte.

Razpravljanja in presojanja o vprašanih odgovorne rabe tehnologij so torej nedvomno izrecno družbenoetična, zato se ocenjevanje o primernosti kake tehnologije v dani situaciji nikakor ne more in tudi ne sme zožiti zgolj na vprašanje o morebitni ekonomski ali tržni uspešnosti njenih proizvodov ali storitev. To še posebej velja za sodobne epistemske in socialne tehnologije. Pri razsojanju o uporabi tovrstnih tehnologij bi morali torej nujno upoštevati družboslovne in humanistične vidike ter merila, in to predvsem pri oceni odgovornosti za uporabo kake tehnologije kot posebne oblike družbene, in ne le ožje tehnične moči.

IN KJE ZAČETI? V PROCESU IZOBRAŽEVANJA.

Če razpravo o povezanosti razvoja tehnologije z družbenim razvojem povežemo z razpravo o krizi družbenega reda, sledi vprašanje, kako zamejiti temeljne negativne učinke sodobnih neoliberalnih družb. Če želimo preprečiti ali vsaj zamejiti temeljne negativne učinke sodobnega družbenega razvoja, kot so poglobljanje tekmovalnosti, pohlepa, negotovosti, gneva, agresije, moramo poseči med institucije, ki imajo to možnost, da lahko preprečijo, da bi se ti učinki prenašali in širili na nove generacije in jih motivirali k temu, da te učinke ponotranjajo kot nekaj normalnega, vsakdanjega in neizogibnega. Nedvomno sta tu osrednji instituciji družina in šola. Medtem ko je družina umeščena v zasebnost, kamor politike težje vstopajo, pa je šola tradicionalno aparat države in politike s sistemskimi ukrepi lahko vplivajo nanjo.

In kaj se dogaja na področju šolskega polja? Z ideološkimi in ekonomskimi spremembami od devetdesetih let prejšnjega stoletja so mnoge vlade začele poudarjati načela tržnega liberalizma, prostega trga dobrin in storitev, tekmovalnosti. Temu usrezno se je spremenil tudi vodilni diskurz o izobraževanju v diskurz individualizma. Ta poudarja ključno vlogo osebne odgovornosti posameznikov in posameznic za njihove izobraževalne in poklicne poti, ob tem pa se zanemarjajo sistemska tveganja

za izobraževalne in poklicne neuspehe tistih, ki ne morejo slediti temu diskurzu individualizma (Nastran Ule 2024). V tem času se je v gospodarstvu in tehnologiji začela »digitalna revolucija«, ki je spodbudila družbene spremembe v smeri družb znanja. Izobraževalne politike po vsej Evropi so ta diskurz družbe znanja hitro sprejele za svojega. Ta diskurz temelji na prepričanju, da so za uspešen razvoj najpomembnejša tista znanja in sposobnosti, ki so ekonomsko funkcionalna, medtem ko so druge dimenzije izobraževanja, na primer etične, socialne, emocionalne kompetence sekundarnega pomena.

Zagovorniki neoliberalne reforme šolstva so zahtevali, da morajo biti smotri šolanja usklajeni s širjenjem prostega trga, s krepitvijo konkurenčnosti delovne sile. V skladu s takim pojmovanjem šolstva so se začeli vrstiti ukrepi za standardizacijo izobraževanja, predvsem z vse bolj razširjenim vsiljevanjem vnaprej oblikovanih kurikulumov in kvantitativnih načinov preverjanja (testiranja) znanja (Apple 2012). Interese po znanju so začeli izpodrivati interesi za čim boljšimi rezultati pri testih, preverjanjih doseženih kognitivnih spretnosti in hierarhičnem razvrščanju pri kvantitativno merljivih preverjanjih znanja. To je izrecno instrumentalen pogled na vzgojo in izobraževanje. Podobno instrumentalno je prepričanje, da je izobraževanje nekakšno vlaganje v človeški kapital, ki se lahko dobro obrestuje, če ga znamo pametno izkoristiti. Tudi pri tem prepričanju so najpomembnejše tiste sestavine človeškega kapitala, ki doživljajo ugodno ceno oz. privedejo do visokih tržnih vrednosti, druge sestavine so zgolj pomožnega značaja.

Tudi v znanstvenoraziskovalno delo se vedno bolj tihotapijo razni partikularni družbeni interesi, ki skušajo vplivati na upravljanje z znanostjo in znanstveno prakso, vendar je njihov vpliv le redko neposreden. Prikriti družbeni interesi pogosto delujejo tako, da ustvarjajo domnevno neproblematičen skriti kontekst znanstvenega razmišljanja in raziskovanja, v katerem se določeni postopki in prakse privzamejo kot samoumevne in so vsaj začasno zavarovane pred kritiko (Longino 2001; Resnik 2007). Ta skriti kontekst predstavlja hudo oviro kritičnemu preverjanju in premisleku mnogih domnevno povsem objektivnih ugotovitev, izjav, konceptov ali ovrednotenj znanstvenega dela. Kritična preverjanja izpostavljajo na primer skrite ozadnje vrednotne predpostavke upravljavcev znanosti ali potlačeno pristajanje na partikularne interese družbeno ali kapitalsko močnih skupin.

Prav v kriznih razmerah, ko nek družbeni sistem ali model prihaja do svojega roba in zgolj reproducira obstoječe, postajata izobraževalni proces in znanost žariščni točki krize, na katerih se vzpostavljajo in lomijo vrednotni modeli življenja in dela v prihodnosti. Zato je razumljivo, da je prav v obdobju in v razmerah družbenih kriz treba prepričati bistvene premise izobraževanja in znanstvenega dela. Kot piše nemški sociolog Ulrich Beck v svoji zadnji knjigi *Methamorphosis of the World*, se v sodobnem svetu dogaja spoj globalnih in lokalnih sprememb, ki postavljajo pod vprašaj mnoge naše dozdajšnje gotovosti in samoumevnosti (Beck 2016, 20). Med tovrstnimi samoumevnostmi je nemara osrednje prepričanje, da naj šola usposablja in pripravlja učence za to, da bodo čim uspešnejši in sposobnejši v konkurenčnih prizadevanjih za uvrstitev v skupnost zmagovalcev v sodobni družbi oziroma svetovni ekonomiji. Vendar zoževanje vloge in smotrov šolanja zgolj na gladko reprodukcijo novih generacij delovno in socialno integriranih mladih ne more biti tudi poglavitni

smoter šolanja. Beck (2016) zato trdi, da za rešitev družbenih problemov spremembe v sistemu izobraževanja niso dovolj. Potrebujemo celostno preobrazbo družbe, ki bo problematizirala mnoge dozdajšnje gotovosti in samoumevnosti (ibid.). Prav tako potrebujemo celostno preobrazbo šolstva, ki bo problematizirala tiste gotovosti in samoumevnosti, ki jih vsebuje neoliberalni koncept in model izobraževanja, ki smo ga nekritično prevzeli kot samoumevnega.

V kritični in reflektivni naravnosti do izobraževalnega in znanstvenoraziskovalnega dela lahko zdržimo le tedaj, če se znamo povezovati med seboj v intelektualne kolektive, ki nastopajo z dovolj močno strokovno avtoriteto. Tu se navezujem na Bourdieuevo idejo o kolektivnih intelektualcih. V sestavku z naslovom *For a Scholarship with Commitment* Bourdieu (2000, 43) govori o kolektivnih intelektualcih kot o družbeni inovaciji, ki bi ob ustrezni organiziranosti omogočila koordinacijo kolektivnega delovanja mreže kritičnih znanstvenikov. Intelektualci, ki bi spadali v organizem kolektivnega intelektualca, bi morali po Bourdieuju podvreči ostri družbeni in diskurzivni kritiki zlasti produkte dominantnega političnega in ekonomskega diskurza.

ZAKLJUČEK

Dozdajšnje razpravljanje nas je pripeljalo do strateških vprašanj o tem, po kateri strategiji naj se razvija sodobno družbeno življenje ljudi in kakšno naj bo razmerje med tehnološkim in družbenim razvojem. Z vidika družbeno racionalnega delovanja se zdi, da je najboljša strategija družbenega razvoja tista, ki teži k največjemu skupnemu dobičku za vse udeležence v katerikoli socialni situaciji. Tako strategijo v socialnem vedenju imenujemo sodelovalna. Druga, ekonomsko in tudi domnevno tehnološko obetavnejša rešitev, ki pa je manj ugodna za večino, je tekmovalna strategija, ki temelji na postavljanju lastnih interesov pred interese drugih.

Glede vprašanja razmerja med tehnološkim in družbenim razvojem smo ugotovili, da ni pomembna samo ustrezna družbena pripravljenost za sprejemanje tehnologij. Razvoj tehnologij mora slediti družbenim potrebam in etičnim normam razvoja družbe. Zgolj s tržnimi prijemi se ne moremo ustrezno pripraviti na sprejem novih, zlasti tveganih tehnologij. Tu pa je neizogibna vloga družbenih ved, ki z usmerjeno refleksijo družbenih kontekstov lahko nastopa kot nekakšen mediator med družbenimi potrebami in tehnološkim razvojem. Zato je lestvica TRL v raziskovalnem programu EU Obzorje 2020 ne samo preozka, saj odraža samo pripravljenost tehnologije za tržno/ekonomsko uporabo, za razvoj komercialno uspešnih proizvodov, ki se jih da monetizirati, ne pa posledic te uporabe za družbeni in psihološki razvoj človeka/človeštva.

Zato je nujno razširiti program na temeljne raziskave, ki proučujejo kompleksne družbene spremenljivke in nakazujejo, izpostavljajo, določajo, katere tehnologije potrebujemo, kako jih lahko integriramo v družbeno okolje in tudi, katerih tehnologij ne potrebujemo, ker so v nasprotju s temeljnimi družbenimi vrednotami: na primer z okoljsko vzdržnostjo, s socialno kohezivnostjo, sodelovalnimi strategijami razvoja, z ohranjanjem psihosocialne stabilnosti in enakih možnosti ter kakovosti življenja.

Ugotavljam, da razmerje med razvojem tehnologije in družbe ni enosmeren proces. Ne vzpodbuja samo tehnologija družbenega razvoja. Tudi družba s svojimi

potrebami, vrednotami, življenjskimi praksami mora vzpodbujati določene oblike tehnološkega razvoja in zavračati tiste, ki niso v skladu s temi vrednotnimi in življenjskimi orientacijami. Utemeljeno torej lahko terjamo od načrtovalcev tehnološkega razvoja, da morajo biti družbenorazvojne orientacije, ki jih proučujemo družboslovci pri tehnološkem razvoju, vsaj enakovredno upoštevane kot tehnološke orientacije.

Kot ugotavljajo sodobni družbeni teoretiki, je sodobni neoliberalni kapitalizem v krizi. To je strukturna, sistemska kriza, in ne ena od manjših, prehodnih kriz, ki so imanentne kapitalizmu kot svetovnemu sistemu, se ciklično pojavljajo in jih sistem ciklično premaguje. Zato so zdaj družbene vede pred izjemno odgovornostjo za sinergičen razmislek o možnostih radikalnih preobrazb tega sistema ali vsaj o možnostih delnih ali začasnih izhodov iz njega (Beck 2016). Ne moremo se zares izkoptati iz krize, če se ne začnemo premikati iz samega sistema, kar je gromozanski izziv. In tudi si ne moremo privoščiti, da bi čakali na čas, ko nas bo sistem pripeljal v vojno ali v ekološko katastrofo svetovnih razsežnosti.

Poleg tega vzorci sodobnih bolezni kažejo, da materialni pogoji življenja vsaj v razvitih družbah niso glavni problem. Večji problem je psihosocialno počutje ljudi, ki se eksponentno poslabšuje. Vendar to stanje po zbranih ugotovitvah ni toliko pogojeno s padanjem ekonomskega standarda, ampak s povečevanjem neenakosti in posledično z občutkom nepravičnosti ter usihanjem socialne kohezivnosti. Psihosocialna revščina je pogosto hujša kot materialna. To seveda ne pomeni, da moramo izbirati med večjo enakostjo ali večjo ekonomsko rastjo. To samo pomeni, da investicije v družbeni razvoj povečujejo učinkovitost družbe vsaj tako kot investicije v tehnološki razvoj. Premiki iz tega kriznega sistema so nujni in so zgodovinsko možni. Kako? Wallerstein (2000) predlaga maksimalno sodelovanje družbenih ved ter povezovanje z naravoslovnimi in s tehničnimi vedami za skupno dobro.

CITIRANA LITERATURA

- Apple, Michael W. 2012. *Education and Power*. London in New York: Routledge.
- Bauman, Zygmunt. 2006. *Liquid Fear*. Cambridge: Polity Press.
- Beck, Ulrich, Anthony Giddens in Scott Lash. 1994. *Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social Order*. Stanford: Stanford University Press.
- Beck, Ulrich. 2000. *What is Globalization*. Cambridge: Polity Press.
- Beck, Ulrich. 2016. *The Metamorphosis of the World*. Cambridge: Polity Press.
- Bourdieu, Pierre. 2000. »For a Scholarship with Commitment«. *Profession*: 40–45. Dostop 11. jan. 2025. <http://www.jstor.org/stable/25595701>.
- Cotton, Matthew. 2014. *Ethics and Technology Assessment: A Participatory Approach*. Berlin: Springer Vrl.
- Feenberg, Andrew. 2002. *Transforming Technology. A Critical Theory Revisited*. Oxford: Oxford University Press.
- Foucault, Michel. 2021. *Nadzorovanje in kaznovanje. Nastanek zapora*. Ljubljana: Krtina.
- Friedman, Thomas. 2005. *The World Is Flat: A Brief History of the Twenty-First Century*. New York: Farrar, Straus and Giroux.
- Harvey, David. 2003. »The Fetish of Technology: Causes and Consequences«. *Prometheus's Bequest: Technology and Change. Macalester International* 13: 1–30.
- Harvey, David. 2012. *Kratka zgodovina neoliberalizma*. Ljubljana: SH.

- Hayek, Friedrich. 1991. *Pot v hlapčevstvo*. Ljubljana: DZS.
- Latour, Bruno. 2002. »Technology and Morality: The Ends of Means«. *Theory, Culture & Society* 19 (5/6): 247–60.
- Lawrence, Frank. 1944. »What Is Social Order?«. *A Review of General Semantics* 2 (1): 29–37.
- Longino, Helen. 2001. *The Fate of Knowledge*. Princeton: Princeton University Press.
- Matthewman, Steve. 2011. *Technology and Social Theory*. London: Palgrave Macmillan.
- Nastran Ule, Mirjana (ur.) 2024. *Kakšno šolo v prelomnih časih?* Ljubljana: Založba FDV.
- Palm, Elin, in Sven Hansson. 2006. »The Case for Ethical Technology Assessment (eTA)«. *Technological Forecasting & Social Change* 73: 543–558.
- Piketty, Thomas. 2014. *Kapital v 21. stoletju*. Ljubljana: Mladinska knjiga.
- Reynolds, Laurence, in Bronislaw Szerszynski. 2016. »Neoliberalism and Technology: Perpetual Innovation or Perpetual Crisis?«. V *Neoliberalism and Technoscience: Critical Assessments*, ur. Luigi Pellizzoni in Marja Ylönen, 27–46. Birmingham: Routledge.
- Resnik, David. 2007. *The Price of Truth: How Money Affects the Norms of Science*. New York: Oxford University Press.
- Salomon, Jean-Jacques. 1984. »What Is Technology? The Issue of Its Origins and Definitions«. *History and Technology* 1 (2): 113–156.
- Wallerstein, Immanuel. 1999. *Utopistike. Dediščina sociologije*. Ljubljana: Založba *cf.
- Wallerstein, Immanuel. 2000. *Kako odpreti družbene vede. Poročilo Gulbenkianove komisije o restrukturiranju družboslovja*. Ljubljana: Založba *cf.

VLOGA IN POMEN ZNANSTVENO-TEHNOLOŠKE TRANSDISCIPLINARNOSTI PRI VREDNOTENJU RAZVOJNIH POTENCIALOV NOVIH NAPREDNIH TEHNOLOGIJ

Franc Mali

UVOD

Odgovora na vprašanje, kako je tehnološki napredek vplival na družbene spremembe, in obratno, kako je širše družbeno okolje vzpostavljalo pogoje za tehnološki napredek, ni mogoče iskati v poenostavljenih razlagah sočasnega kumulativnega in linearnega tehnološkega in družbenega razvoja. Že bežen pogled v zgodovino razvoja družb v zadnjih treh stoletjih, ko so nova znanstvena in tehnološka odkritja dobila izredno moč in so vplivala na družbene spremembe, razkriva, da ustvarjanje, širjenje in uporaba znanstveno-tehnoloških inovacij v širšem družbenem okolju nikoli niso bili enkratno dejanje, ki se pojavi čez noč in življenje ljudi dobesedno preobrne na glavo. Ravno nasprotno. Vpliv tehnologij na družbene spremembe je dolgoročen, kompleksen in tudi nepredvidljiv proces. Ni se namreč kar vsaka invencija enostavno pretvorila v inovacijo. Včasih je trajalo več desetletij, da so družbe prepoznale razvojne potenciale novih znanstveno-tehnoloških odkritij. Lahko pa do tega sploh ni prišlo.

Glede tega ni nič drugače v tretjem desetletju enaindvajsetega stoletja, ko se soočamo s pravo erupcijo na področju razvoja novih naprednih tehnologij, kot so umetna inteligenca, genomika in nanotehnologija. Gre za najbolj propulzivne tehnologije, ki v dobrem in slabem vplivajo na družbeno okolje, hkrati pa so z njim pogojene in omejene.

Ker nove napredne tehnologije niso od družbe izoliran pojav, ki bi bil potrjen samozadostni in linearni logiki razvoja, morajo biti zlasti v procesih, ki vodijo k njihovi komercializaciji in marketingu, predmet stalnega in kritičnega, na širših družbenih in etičnih kriterijih utemeljenega vrednotenja. Enako velja za vrednotenje tehnoloških tveganj. Ne gre toliko za vprašanje vrednotenja finančnih tveganj z vidika podpornikov tehnologij, ki prihajajo iz podjetniških ali političnih krogov, temveč gre za vprašanje, kaj nove tehnologije prinašajo za celotno družbo in tudi naravno okolje. Ne glede na različne teoretske opredelitve in razlage tudi najbolj znani instrument merjenja novih tehnologij TRL (ang.: technology readiness level) v praksi ni uspel preseči tehnokratskega in ozko marketinško usmerjenega vrednotenja razvojnega potenciala novih tehnologij. Pri ključnih akterjih znanstvene in tehnološke politike je

še vedno premalo zavedanja, da je vsak tehnološki razvoj kot starogrški bog Janus z dvema obrazoma: lahko veliko prispeva k dobrobiti današnjih družb, lahko pa pomeni tudi veliko družbeno tveganje.

V izhodišče prispevka postavljam tezo, da bi (teoretski) modeli tehnološke konvergenca in s tem povezane ideje znanstvene transdisciplinarnosti, ki so se z velikimi pričakovanji pojavili na prizorišču raziskovalnih in inovacijskih politik na začetku tega stoletja, lahko bili ustrezna platforma za presejanje ozko tehnokratskih in marketinških usmeritev TRL. Žal se to ni zgodilo, ko se je TRL v evropskem prostoru začel uporabljati kot metodologija za merjenje zrelosti tehnologije. To se ravno tako ne dogaja danes, skoraj petnajst let po uvedbi TRL. Eden izmed zaviralnih dejavnikov je pretirana vpetost znanosti, tehnologije in ekonomije v tradicionalne modele inovacij. Ti izhajajo iz procesov produkcije in vrednotenja na temelju tržne učinkovitosti posameznih tehnologij, ne zmorejo pa ovrednotiti na celostnejši način vseh znanstvenih, tehnoloških in širših družbenih potencialov, ki prispevajo oziroma bi lahko prispevali k blagostanju človeštva. V prispevku zavzemam stališče, da je paradigmatški »premik« v tej smeri mogoč le ob čim produktivnejšem sodelovanju vseh znanstvenih disciplin, in sicer ne toliko v smislu *ex post* interdisciplinarnega povezovanja, ki poteka v sferi akademske znanosti, temveč *ex ante* integriranja vseh vrst vedenja z namenom reševanja praktičnih problemov. Tako bo tudi lažje uveljaviti premik od »pripravljenosti tehnologije« k »tehnološki pripravljenosti«, o kateri govori akademik Slavko Splichal v *Izhodiščih za SAZU-jev posvet Od tržne konkurenčnosti tehnologije k družbeni tehnološki zrelosti* (Splichal 2024). Ta premik je ključen, kolikor naj tehnološki razvoj predstavlja osnovo za trajnostni družbeni razvoj, socialno pravičnost, družbeno blaginjo ter etično občutljivost posledic razvoja novih in naprednih tehnologij.

V drugem delu prispevka se ukvarjam z vprašanjem, zakaj se je treba v okviru razvoja tehnologije generativne umetne inteligence in tehnologije preoblikovanja človekovega genoma, ki sta danes najnaprednejši novi tehnologiji, soočiti z njuni širšimi družbenimi posledicami. Brez tega soočanja ni mogoče izvajati rekonceptualizacije obstoječega TRL v smeri njegove večje senzibilnosti za današnje ključne družbene in etične probleme. V zaključku podajam kritično oceno, da se Ministrstvo za visoko šolstvo, znanost in inovacije Republike Slovenije, ki sicer želi prevzeti eno ključnih vlog v institucionalni podpori in spodbujanju inovativnosti pri nas, še vedno premalo ukvarja z vprašanjem, kako posodobiti koncept TRL.

KONVERGENTNI RAZVOJ NOVIH NAPREDNIH TEHNOLOGIJ IN ZNANSTVENA TRANSDISCIPLINARNOST

Ko govorimo v okviru raziskovalno-razvojne politike o kategorijah, kot so »tehnološka konvergenca« ali »znanstvena transdisciplinarnost«, se je treba zavedati, da se obe kategoriji medsebojno tesno prepletata. Če govorimo o nastopu današnje transdisciplinarne znanosti, se, kolikor nas zanima njena aplikativna dimenzija, nujno najdemo tudi pri problemu tehnološke konvergenca. In nasprotno, če v izhodišče postavimo vprašanje tehnološke konvergenca, se pri pojasnjevanju njenih bistvenih značilnosti ne moremo izogniti transdisciplinarnosti. V obeh primerih gre za sinergijo med različnimi segmenti znanstvenega in tehnološkega vedenja, ki v medsebojni

povezanosti, ne pa vsak zase, izjemno hitro napredujejo. Konvergentni tehnološki razvoj je pomemben, ker ukinja tradicionalne disciplinarne meje v znanstvenih komunikacijah in spodbuja ne le horizontalne (znotraj akademskega sveta), ampak tudi vertikalne oblike sodelovanja znanstvenikov (v smeri od temeljne k uporabni znanosti in razvoju).

Ameriški teoretik znanosti in ekonomist Rogers Hollingsworth, ki se je ukvarjal z vlogo novih naprednih tehnologij pri ekonomskem razvoju, je še posebej izpostavil, da so napredne tehnologije v prvem desetletju tega stoletja dokončno prešle v fazo razvoja, ki zahteva vzpostavitev močnega skupnega znanstveno-tehnološkega jedra, sestavljenega iz skupnih teoretičnih okvirjev ter skupnega nabora modelov in mehanizmov, ki združujejo širok spekter področij, na katera se nanašajo različne znanstvene discipline in tehnološka področja (Hollingsworth 2011). Medtem ko so se v preteklem stoletju odvijale majhne konvergence v znanstvenem in tehnološkem razvoju, sodobne, na transdisciplinarnem znanstvenem vedenju utemeljene tehnološke konvergence radikalno spreminjajo značaj moderne znanosti in tehnologije. Na začetku leta 2000 se je v okviru razvoja konvergentnih tehnologij veliko stavilo na povezovanje na nanoravni, ko so se tipični pojavi v materialnih nanostrukturah začeli obravnavati kot osnove v bioloških sistemih, nanoprodukciji, komunikacijah in strojnem učenju. Ti procesi spajanja različnih področij vedenja na nanoravni so se v naslednjih dveh desetletjih samo še okrepili. Gradnja sistemov na ravni nanometerskih enot namreč zahteva kombinirano uporabo specifičnih mikrofizikalnih zakonov, bioloških načel, informacijske tehnologije in sistemske integracije. Prizadevanje za združitev tehnologije kvantnega računalništva z umetno inteligenco bo na primer odprlo prostor kompleksnim analizam velikih podatkovnih zbirk. Ravno dvig računalniških kapacitet je ključni steber nadaljnega napredka umetne inteligence.

Iste procese lahko zaznamo v zadnjem desetletju tudi na področju biotehnologije. Razvoj biotehnologije gre vedno bolj v smer celovitih tehnoloških platform, ki ustvarjajo izjemno diferencirano, vendar tesno prepleteno verigo novih inovacij. Med sabo ločeni različni industrijski sektorji, ki dandanes sledijo vsak svojemu cilju, težijo k vedno večjemu poenotenju okrog novih tehnologij, zlasti okrog sintezne biologije in umetne inteligence. Poenotenost na skupnih platformah si lahko predstavljamo na naslednji način: podobno kot so računalniki postali univerzalni stroji, potem ko so znanstveniki razvili ustrezne programe in algoritme, ki simulirajo delovanje vsakega drugega stroja, se tudi biološki organizmi spreminjajo v univerzalne producente, in sicer tako, da lahko z ustreznimi spremembami svojih genskih programiranj ustvarijo praktično katerikoli zamišljeni produkt (Mali 2015, 2024).

Tovrstni trendi znanstvenega in tehnološkega povezovanja so včasih povezani s pretiranimi prihodnjimi pričakovanji (npr. spekulacije o doseganju t. i. točke singularnosti), vseeno pa nazorno prikazujejo osnovno smer razvoja tudi na področju temeljne znanosti. Tu ideja znanstvene transdisciplinarnosti, ki se je začela pojavljati pred konceptom tehnološke konvergence, nekako v 90. letih prejšnjega stoletja, predpostavlja najvišjo stopnjo znanstvenega in zunajznanstvenega povezovanja. Medtem ko se je izvorni pomen transdisciplinarnosti pred letom 1990 sporadično uporabljajal za opis skupnih predpostavk, ki naj bi prispevale k sintezi znanstvenega vedenja in naj bi presegle ozke disciplinarne meje, se v tem najnovejšem razumevanju nanaša

predvsem na reševanje praktičnih problemov zunaj ozkega akademskega polja, saj ti zaradi kompleksne narave ne morejo ostati v domeni ozkih disciplinarnih pristopov. Medtem ko gre v okviru interdisciplinarne znanosti za povezovanje teoretskih in metodoloških okvirjev posameznih znanstvenih disciplin znotraj akademske sfere, transdisciplinarnost razširja, neodvisno od kakršnekoli zavezanosti teorijam in metodam znotraj znanstvenih disciplin, obseg svojega vedenja ravno z namenom celostnega reševanja kompleksnih problemov v praksi. V zadnjem desetletju in pol se je natančneje razmejevanje med kategorijami transdisciplinarnosti in interdisciplinarnosti uveljavilo tudi v okviru OECD (OECD 2020).

ODSOTNOST IDEJE TEHNOLOŠKE KONVERGENCE IN ZNANSTVENE TRANSDISCIPLINARNOSTI V EVROPSKEM KONCEPTU TRL

Koncept konvergentnega tehnološkega razvoja je dobil večjo relevantnost v okviru bolj praktično naravnanih razprav o raziskovalno-razvojih in inovacijskih politikah v začetnih letih tega stoletja, in sicer najprej v ZDA. Mihail C. Roco, ki deluje na področju nanotehnologije, in William Sims Bainbridge s področja sociologije sta leta 2002 pod okriljem NSF pripravila obsežno poročilo z naslovom »*Managing Nano-Bio-Info-Cogno Innovations – Converging Technologies in Society*«. To je bilo nekaj let pozneje izdano kot monografija pri založbi Springer (Bainbridge in Roco 2006). Kot je razvidno že iz naslova njune monografije, sta oba v sodelovanju z drugimi znanstveniki največ pozornosti namenila procesom spajanja štirih najnaprednejših tehnologij, to je nanotehnologije, biotehnologije, informacijske tehnologije in tehnologije, ki se ukvarja s kognitivnimi procesi v človeških možganih.

Skoraj sočasno se je tudi v strokovnih krogih, ki se ukvarjajo z raziskovalno-razvojno in inovacijsko politiko v EU, začela razprava o kognitivnih in družbenih vidikih razvoja konvergentnih tehnologij. Tu se je, v nasprotju z razpravami v ZDA, manj razpravljalo o viziji krepitev človekovih zmožnosti, ki je pomaknjena bolj v oddaljeno prihodnost, bolj pa je bil v ospredju premislek o takojšnji uporabi konvergentnih tehnologij pri reševanju perečih družbenih problemov. Z nekoliko poenostavitve bi lahko dejali, da je bil ameriški koncept tehnološke konvergenke vezan bolj na notranje tehnološke impulze, evropski koncept pa je ob teh notranjih tehnoloških impulzih v večji meri upošteval tudi družbeni vidik tehnološkega razvoja. Koncept konvergentnih tehnologij, ki se je oblikoval v okviru EU, se torej ni ukvarjal toliko z vprašanjem krepitev človekovih zmožnosti, temveč je izpostavil družbeno dimenzijo razvoja novih naprednih tehnologij, ki je vključevala trajnostni razvoj, socialno vključenost in izboljšano kakovost življenja. Večji pomen je bil pripisan tudi družbenim, etičnim in pravnim implikacijam razvoja teh novih naprednih tehnologij, ki so svoj inovacijski potencial začele pridobivati na podlagi vzajemnega spoja. Ta specifična evropska »filozofija« koncepta konvergentnih tehnologij je bila vključena tudi v poročilo »*Converging Technologies – Shaping the Future of European Societies*« (Nordman 2004), ki ga je skupina ekspertov pripravila za Evropsko komisijo. Omenjeno delo je predstavljalo najcelostnejši pogled EU na vprašanje prihodnjega razvoja konvergentnih tehnologij. Evropski eksperti so se v tem poročilu ukvarjali predvsem z vprašanjem, kako na podlagi že ustvarjenih inovacij, do katerih je že prišlo z razvojem

konvergentnih tehnologij, doseči najboljše rešitve za evropsko družbo, ob striktnem upoštevanju etičnih načel. Med drugim so zapisali (Nordman 2004, 4),

da pri strategiji razvoja konvergentnih tehnologij za evropsko družbo znanja ne gre za proces, ki poteka od zgoraj navzdol, temveč za avtonomni proces ustvarjalnega tehnološkega razvoja na osnovi t. i. postopne širitve konvergenčnih krogov, ki je rezultat delovanja vseh delov znanosti in tehnologije, vključno z družboslovjem in s humanistiko. Ravno zaradi tega se k etičnim, pravnim in družbenim vprašanjem, ki jih sproža razvoj konvergentnih tehnologij, ne pristopa od zunaj in za nazaj, temveč vedno kot k nečemu, kar predstavlja inherentni del razvoja teh tehnologij.

Poročilo EU je torej v ospredje postavljalo nujnost ustvarjanja inovacijskih potencialov v družbi, ki temelji na sinergiji vzajemnega delovanja vseh znanosti in vseh tehnologij, ob upoštevanju vseh temeljnih etičnih načel. Še več, z vidika razvoja evropske družbe v smeri bolj inovativne družbe se je ta model razvoja konvergentnih tehnologij izpostavljal kot vzorčni primer za spopadanje evropske družbe z izzivi 21. stoletja. Zato je bilo tudi v političnih krogih prisotno veliko pričakovanje, da se bodo ob spremembi v smeri večje konvergence novih naprednih tehnologij v državah članicah EU zgodile relativno hitre institucionalne spremembe.

Nič od teh pričakovanj se ni uresničilo. Koncept razvoja konvergentnih tehnologij, ki je nastal v strokovnih gremijih EU na začetku 21. stoletja, ni pomembneje vplival na raziskovalno-razvojno in inovacijsko politiko v letih, ki so sledila, torej tudi ni vplival na oblikovanje TRL. Ko je Evropska komisija leta 2010 ustanovila še eno izmed številnih ekspertnih komisij po imenu High-Level Expert Group on Key Enabling Technologies (HLG KET), ki je pripravila poročilo o rabi TRL in predlagala njegovo vključitev v Obzorje 2020 (Horizon 2020), tj. 8. okvirni program EU (European Commission 2011), o konceptu konvergentnih tehnologij ni bilo več govora. V poročilu je bila celotna pozornost namenjena zgolj vprašanju, kako opredeliti mejo med različnimi organizacijskimi in finančnimi vidiki tehnološkega razvoja, »/.../ s ciljem doseganja čim večjih tržnih učinkov« (Héder 2017, 8). Evropski model TRL je bil oprt na devet zaporednih stopenj, ki naj bi pokazale, kako je neka tehnologija že od začetka pripravljena za uporabo v okolju, h kateremu je usmerjena. V tabeli 1 v strnjeni obliki predstavljamo devet stopenj TRL.

Tabela 1: TRL v 8. Okvirnem programu EU Obzorje 2020

TRL	OP EU Obzorje 2020
1	Upoštevanje osnovnih načel
2	Oblikovan tehnološki koncept
3	Eksperimentalni dokaz koncepta
4	Tehnologija, potrjena v laboratoriju
5	Tehnologija, potrjena v ustreznem okolju
6	Tehnologija, prikazana v ustreznem okolju
7	Predstavitev prototipa sistema v delovnem okolju
8	Sistem dokončan in kvalificiran
9	Dejanski sistem, preizkušen v delovnem okolju

LINEARNI MODELI INOVACIJ KOT OSNOVA ZA OZEK TEHNOKRATSKI IN MARKETINŠKI PRISTOP K TRL

Na tehnokratsko in marketinško razumevanje evropskega modela TRL, ki ga predstavlja tabela 1, opozarja med drugim tudi Splichal v *Izhodiščih za SAZU-jev posvet Od tržne konkurenčnosti tehnologije k družbeni tehnološki zrelosti*. Splichal (2024) ugotavlja, da je v okviru pristopa EU bila pot od TRL 1 do TRL 9 predstavljena kot linearno napredovanje od ideje do trga, s čimer se je pojem tehnologije zreduciral na proizvodne artikle, ki jih je mogoče monetizirati na trgu.

Eden izmed ključnih razlogov za odsotnost koncepta konvergentnih tehnologij v evropskem modelu TRL je še vedno nekritično sprejemanje predstave, da inovacije potekajo po linearni poti. Četudi se v današnjih pogojih razvoja novih naprednih tehnologij taki linearni inovacijski modeli izkazujejo kot anahronizem, so še vedno navzoči v glavah številnih akterjev, ki imajo odločilo besedo na področju raziskovalno-razvojne in inovacijske politike.

Za Josepha Schumpetra, ki je opravil pionirsko delo na področju teoretskega in praktičnega preučevanja inovacij, so bile inovacije kot oblika »kreativne destrukcije« motor družbenega in ekonomskega razvoja (Schumpeter 1934/2008). Zato Schumpetra ni mogoče imeti za zagovornika ozkega tehnokratskega pogleda, ki izhaja iz linearnega inovacijskega modela. Prej nasprotno. Schumpeter je pripisoval relevantnost tako tehnološkim kot socialnim inovacijam, saj naj bi obe obliki predstavljali ključni dejavniki družbenega in ekonomskega razvoja. Četudi je postavljaj jasno mejo med njima, kar zadeva namene in cilje, saj so tehnološke inovacije usmerjene k ustvarjanju novih proizvodov in storitev, socialne inovacije pa k neposrednemu spreminjanju družbenih praks, to ne pomeni, da niso tako ene kot druge usmerjene k stalnemu preoblikovanju družbenih struktur in procesov.

Začetki linearnega modela inovacij, najprej v krogu akterjev znanstvene in tehnološke politike v ZDA, pozneje tudi v znanstvenih in tehnoloških politikah vseh drugih razvitih ekonomij v svetu, segajo v obdobje takoj po drugi svetovni vojni. Zagovorniki linearnih inovacijskih modelov so vzore iskali v takrat ravno končanem projektu Manhattan, ki je bil med drugo svetovno vojno vojaški raziskovalni projekt. Ameriški znanstvenik in politik Vannevar Bush je v delu »*Science: The Endless Frontier*« (Bush 1945/1960), ki je izšlo takoj po drugi svetovni vojni, zastopal stališče, da pretok znanstvenih idej v prakso poteka enosmerno in hierarhično od bazičnega raziskovanja (invencija), prek aplikativnega (reševanje specifičnih praktičnih problemov) in razvojnega raziskovanja (prototip) k proizvodnji in h komercializaciji končnih proizvodov. Zdi se, da ima ta shema linearnega inovacijskega modela, kljub vsem spremembam, ki so se zgodile na področju znanstveno-tehnološkega razvoja v naslednjih sedemdesetih letih, tudi še danes veliko zagovornikov.

V pogojih današnjega razvoja novih naprednih tehnologij ni mogoče več pristajati na zastarele linearne inovacijske modele. Inovacijski procesi potekajo danes na osnovi cikličnih in rekurzivnih tokov. Tu ne gre več, tako kot v linearnih inovacijskih modelih, za enosmerni in hierarhični tok od invencij k inovacijam, temveč gre za vsakokratno prepletenost in součinkovanje najrazličnejših dejavnikov. V verižnih (»*chain-link*«) inovacijskih modelih na primer temeljna znanstvena spoznanja (invencije)

ne odigrajo svoje vloge samo v začetnih oziroma vstopnih fazah, temveč vzdolž celotne inovacijske verige (Suarez-Villa 2009; Kline in Rosenberg 1986).

Z vidika bolj odprtega in netehnokratskega pristopa k TRL je »filozofija« verižnih modelov inovacij relevantna, ker poudarja, da brez fleksibilnih in medsebojno povezanih povezav, kjer imajo v procesih inovacij vlogo tudi povratne zanke, ni mogoče razvijati inovativnih proizvodov. Če izhajamo iz verižnih modelov inovacij, bo znotraj lestvice TRL nekaj povsem samoumevnega, da se iz faze 7 ponovno vrnemo v faze 4, 3 ali celo 2, kolikor sploh želimo ponovno napraviti korak naprej, to je v fazo 8. Podobno velja, da če imamo inovativni tehnološki produkt za celoto posameznih delov, posamezni sestavni deli napredujejo z različnimi hitrostmi po predlagani lestvici TRL.

Ni se mogoče izogniti vtisu, da se v evropskem prostoru lestvica TRL namesto za realni opis kompleksnih inovacijskih procesov uporablja bolj kot instrument, ki naj bi omogočal akterjem raziskovalno-razvojne in inovacijske politike priti do ocen finančnih stroškov, ki jih vlagajo v inovacijsko aktivnost. V tem oziru celo obstoječe mednarodne statistične standardizacije merjenja inovativne aktivnosti, kolikor se opirajo na zastarele tipe klasifikacij, niso v veliko pomoč. Kot primer bi lahko vzeli precej tega klasifikacijo treh dimenzij raziskovalno-razvojne dejavnosti (temeljne, aplikativne, razvojne), ki jo vsebuje priročnik Frascati in ki glede na današnji razvoj novih naprednih tehnologij uporablja precej neživiljenjska navodila za ločevanje raziskovanja in razvoja od proizvodnje (Frascati Manual 2015). V tem duhu izzvenijo tudi kritična opažanja Mihályja Héderja, ki svojo analizo evropskega modela TRL zaključuje z oceno, da ta ni v ničemer presegel »filozofije« linearnosti, edino, kar je storil, je, »/.../ da je s svojimi natančneje opredeljenimi merskimi lestvicami nadomestil splošnejše kategorije, ki so bile pred tem v uporabi« (Héder 2017, 21).

Posamezni kritiki linearne logike mišljenja, ki je prisotna v merskih lestvicah TRL, opozarjajo na še eno pomanjkljivost. Ta naj bi se nanašala na nezmožnost TRL, da spremlja razvoj najnaprednejših tehnologij, ki so izjemno kompleksne in hkrati tehnologije, ki se razvijajo zelo hitro. V merskih lestvicah TRL naj bi tako manjkala opredelitev kategorije »tehnološko zastaranje« (Olechowski, Eppinger in Joglekar 2015, 7). Tehnološko zastaranje se lahko zgodi, še preden tehnologija preide vse inovacijske faze, ki so opredeljene v okviru evropskega modela TRL. Tehnološko zastaranje se tako nanaša na situacijo, ko postane tehnologija na nek način neuporabna, ker se je sočasno razvila boljša tehnološka alternativa. Evropska merska lestvica TRL je bila zasnovana pod vplivom instrumentov, ki so bili vezani na vesoljski program Nase. To povzroča težave. Posebnost vesoljskih tehnologij namreč je, da je vesolje, v katerem te tehnologije delujejo, relativno stabilno okolje. V nasprotju s to tehnologijo najnovejše tehnologije generativne umetne inteligence delujejo v dinamičnih kompleksnih virtualnih okoljih. Situacija je tu drugačna. Lahko se zgodi, da neka tehnologija, ki je zdaj na vrhu razvoja, zastara prej, kot preide vse faze TRL. Če se po Moorovem zakonu (leta 1965 ga je formuliral Gordon Moore) zmožljivost računalniške opreme podvaja na vsaki dve leti, »svetlobne« hitrosti sprememb na področju razvoja najnovejše generativne umetne inteligence očitno ne moremo pojasnjevati več z Moorovim zakonom.

Za preseganje tehnokratskega linearne modela TRL ne zadostujejo dodatne stopnje, kot je na primer stopnja organizacijske zrelosti itd. Izvirna definicija TRL v okviru Nase je predpostavljala sedem stopenj. Faza 1 se je nanašala na oblikovanje

novega koncepta, faza 7 pa na uspešno demonstriranje inovacije. Iz tabele 1 je razvidno, da se je v EU model TRL povečal na devet stopenj. Tudi večina drugih pristopov drugje po svetu se danes opira v glavnem na devetstopenjski model TRL (Voltan et al. 2024). Ne nazadnje je bil eden izmed razlogov za uvajanje in širjenje TRL za namene raziskovalno-razvojne in inovacijske politike tudi prizadevanje, da se doseže standardizacija, ki bo omogočala eksaktnije in enotnejše razprave o tehnični zrelosti različnih vrst tehnologij.

Pod vplivom kritik, da je model TRL v osnovi preveč tehnokratsko zasnovan, pogosto naletimo na predloge, da se število stopenj poveča z devetih na deset, enajst ali več. So pa ti predlogi odvisni tudi od dejstva, s katero vrsto tehnologije imamo opravka. Na primer, Medvladni panel za podnebne spremembe (IPCC 2022) je začel uporabljati enajststopenjsko lestvico. Večje število stopenj za spremljanje zrelosti posamezne tehnologije se v omenjeni mednarodni organizaciji utemeljuje z dejstvom, da inovativne tehnologije, ki so vezane na uporabo nizkoogljične energije, potrebujejo zlasti v fazi pilotskih preizkusov in poznejše morebitne serijske proizvodnje bolj diferenciran pristop. Še bolj diferenciran pristop v okviru evropskega modela TRL zahtevajo avtorji prispevka »*Technology Readiness Revisited: A Proposal for Extending the Scope of Impact Assessment of European Public Services*« (Bruno et al. 2020), ki se ukvarjajo z vprašanjem zrelosti tehnologij na področju javnega sektorja. Predlagajo, da v evropski model TRL vključimo vsaj tri dodatne stopnje, tj. stopnjo družbene pripravljenosti, stopnjo organizacijske pripravljenosti in stopnjo pripravljenosti z vidika pravnih ureditev. Tri nove stopnje naj bi bile še posebej relevantne v delovanju javnega sektorja. Avtorji tega predloga menijo, da mora biti glavni cilj uporabe TRL v javnem sektorju doseganje čim zanesljivejših tehnologij. Ta cilj pa naj bi bil uresničljiv edinole, kolikor je z novimi tehnologijami podprt javni servis namenjen ljudem (stopnja družbene pripravljenosti), ne povzroča motenj v delovanju javnega servisa (stopnja organizacijske pripravljenosti) in na področju pravnih regulacij zna braniti javni interes pred partikularnimi zasebnimi interesi (stopnja pravne pripravljenosti).

Ne glede na različna strokovna in institucionalna prizadevanja po fleksibilnejši interpretaciji in uporabi TRL se zdi, da gre tu še vedno bolj za parcialne poskuse narediti celotni model manj tehnokratsko in manj marketinško ter bolj socialno in ekološko naravnano. V celoti gledano se TRL primarno še vedno uporablja predvsem kot orodje za spremljanje finančnih tveganj vlagateljev finančnih sredstev v posamezno tehnološko inovacijo. Manj pozornosti pa se namenja spremljanju zrelosti tehnologij v luči širših družbenih, ekoloških in etičnih dilem ter vprašanj, ki jih prinaša razvoj novih naprednih tehnologij. To vprašanje podrobneje obravnavam tudi v naslednjem razdelku.

POTREBA PO REKONCEPTUALIZACIJI TRL V LUČI DANAŠNJIH TVEGANJ, KI JIH PRINAŠA RAZVOJ NOVIH NAPREDNIH TEHNOLOGIJ

Razvoj novih najnaprednejših tehnologij ne prispeva samo k dobrobiti človeštva, temveč s sabo prinaša vrsto nepredvidljivih tveganj. Na začetku 20. stoletja je bilo prevladujoče nelagodje in obžalovanje v družbi, da človeštvo še vedno nima zadostnega vedenja, da bi si podredilo naravo. Danes, v tretjem desetletju 21. stoletja, je sodobna družba pred povsem drugačnim izzivom, vendar ne toliko pred po-

manjkanjem vedenja o naravi, temveč bolj pred pomanjkanjem vedenja o tveganjih čezmernega podrejanja narave.

Zelo pomembno je, da se družboslovje in humanistika v čim večji meri vključita v razprave o vseh etičnih vidikih nadaljnjega razvoja novih naprednih tehnologij. Četudi si včasih pogledi o tem, kdaj naj se »mehko« vedenje vključi v razpravo o tveganjih tehnološkega razvoja in s tem povezanih etičnih vprašanj, nasprotujejo, namreč, ali bolj v fazi *ex post* ali bolj v fazi *ex ante* ocenjevanja tehnologij, je najbrž ključnega pomena, da se to vedenje sploh začne dejansko upoštevati. Skratka, da to vedenje postane imanentni del razvoja novih naprednih tehnologij. Včasih se zdi nenavadno, da so ravno v primeru velikih znanstveno-tehnoloških projektov pobude za enakopravno obravnavo vseh področij znanosti prišle s strani raziskovalcev, ki po svojem znanstvenem pedigreju niso bili družboslovci ali humanisti. Lep primer je veliki projekt raziskovanja človeškega genoma (*»Human Genome Project«*), ki je koncem devetdesetih let prejšnjega stoletja združil raziskovalne sile iz različnih držav sveta. Vodja tega megatransnacionalnega projekta je bil James Watson, eden od odkriteljev zgradbe DNK in Nobelov nagrajenec. James Watson je ob prevzemu vodenja omenjenega projekta zahteval, da se del ogromnega proračuna, ki je bil namenjen raziskovanju človeškega genoma, takoj nameni tudi preučevanju njegovih družbenih in etičnih vidikov.

Sodobne družbe so družbe tveganj. Na ta vidik so v svojih delih opozarjali vsi vodilni sociologi 20. stoletja, kot so Ulrich Beck (Beck 1992), Anthony Giddens (Giddens 1991) in Niklas Luhmann (Luhmann 1991). Izhajali so iz razlike med kategorijo nevarnosti, ki je bila značilna za tradicionalne družbe, in kategorijo tveganja, ki je rezultat delovanja sodobnih družb. Medtem ko so bile v tradicionalnih družbah nevarnosti povezane s preteklostjo in z izgubo vere v trdne strukture, je tveganje danes povezano s procesi družbene modernizacije. Ulrich Beck, Anthony Giddens in Niklas Luhmann govorijo o onipotentnosti tveganj v sodobnih družbah. Če sledimo njihovim spoznanjem, tudi lažje razumemo, zakaj ravno razvoj novih naprednih tehnologij danes predstavlja eno najbolj nepredvidljivih tveganj. Zgodovina znanosti in tehnologije sicer ponuja številne primere, da je znanstveno-tehnološki razvoj že od nekdaj vzbujal številne strahove v družbi. Vzroki za te strahove so bili raznovrstni in jih je mogoče iskati že v predindustrijski dobi. Prišlo je celo do popolnega odklanjanja znanstvenega in tehnološkega razvoja, ki se je v posameznih primerih sprevrglo celo v skrajno obliko tehnofobije. Taki primeri iz preteklosti, ki so v zvezi z znanostjo in s tehnologijo izpostavljali najnevarnejše distopije, nas v ničemer ne odvezujejo od tega, da se kritično soočamo s tveganji, ki jih prinaša današnji razvoj novih naprednih tehnologij. Danes smo namreč v izjemno razburljivem obdobju znanstveno-tehnološkega razvoja. Prvič v zgodovini je tisto, kar je bilo do zdaj obravnavano kot esenca človeka, postalo predmet neposrednih manipulacij znanosti in tehnologije.

V zadnjem času zlasti razvoj na področju tehnologije generativne umetne inteligence in tehnologije preoblikovanja človeškega genoma poraja vedno več strahov v družbi in hkrati vprašanj o tem, ali niso tveganja z vsakim korakom njihovega razvoja že tako velika, da je treba ta razvoj ustaviti ali vsaj upočasniti.

Kar zadeva generativno umetno inteligenco,¹ nekateri avtorji govorijo o tem, da bi se v primeru njenega razvoja brez ustreznega družbenega nadzora človeštvo znašlo

pred eksistenčnim tveganjem. Na eksistenčno tveganje, ki ga prinaša razvoj umetne inteligence, je že pred več kot desetimi leti opozoril filozof Bostrom (2014). Predpostavil je, da bi umetna inteligenca lahko že razmeroma zgodaj v prihodnosti dosegla oziroma presegla zmožnosti človeške inteligence. Taka superinteligenca, h kateri vodi razvoj umetne inteligence, bi lahko po distopičnem scenariju, ki ga je predpostavil Nick Bostrom, v končni konsekvenci pripeljala tudi k iztrebljanju človeške vrste. Predvsem pa bi imeli v primeru takega scenarija opravka z ireverzibilnim tokom dogodkov, pri katerem poti nazaj ne bi bilo več. V tem smislu obstaja pomembna razlika med katastrofičnim in eksistenčnim tipom tveganj. Če katastrofična tveganja ocenjujemo po kriteriju maksimalne razširjenosti (številu ljudi, ki bi bili prizadeti), intenziteti (trpljenju, ki ga povzroča) in trajanja škodljivih družbenih posledic nekontroliranega razvoja posamezne tehnologije, je pri eksistenčnem tveganju, ki naj bi bilo povezano z generativno umetno inteligenco, odločilna vloga samo enega kriterija, tj. nevarnosti iztrebljanja človeške vrste zaradi prevlade stroja nad človekom.

Generativna umetna inteligenca je v zadnjem obdobju z razvojem velikih jezikovnih modelov naredila pravo revolucijo. Še pred nekaj leti so tudi najzmogljivejši veliki jezikovni modeli redko uspeli generirati kako krajše koherentno besedilo. V zadnjem času najnaprednejši veliki jezikovni modeli (GPT, Bard, Claude, Llama, Gemini, DeepSeek itd.) pišejo računalniške programe, samostojno generirajo znanstvena in umetnostna besedila večjega obsega, se vključujejo v neomejene odprte tipe komunikacij itd. Zato ni presenetljivo, da vprašanja o možnem eksistenčnem tveganju generativne umetne inteligence, kolikor se bo njen razvoj dogajal s tako bliskovito hitrostjo, niso v krogih strokovnjakov, ki se ukvarjajo s strojnimi učenjem, nič več zavrnjena kot del znanstvene fantastike.

Naj zadevo ilustriramo. Konec letošnjega januarja je bilo v javnost dano poročilo o varnosti generativne umetne inteligence (International AI Safety Report 2025). V tem mednarodnem poročilu, pri katerem je sodelovalo sto strokovnjakov iz tridesetih držav, razen tega pa še predstavniki OZN, OECD in EU, je podan obsežen kompendij tveganj generativne umetne inteligence.

Osnovni cilj tega poročila je podati mnenje strokovnjakov z vsega sveta o tem, kakšna so tveganja, ki jih prinaša razvoj generativne umetne inteligence. Poročilo je sestavni del širšega prizadevanja strokovnjakov in politikov na mednarodni ravni, da se z najbolj perečimi kratkoročnimi in dolgoročnimi tipi tveganj razvoja generativne umetne inteligence sooči svet kot celota. Novembra 2023 je bil sklican prvi globalni vrh o varnosti umetne inteligence, ki je potekal v Londonu. Glavna tema srečanja je bila soočanje z možnim eksistenčnim tipom tveganj, ki ga prinaša razvoj generativne umetne inteligence. Nadaljevanje tega vrha je potekalo maja 2024 v Seulu, tretje srečanje pa v Parizu konec februarja 2025. V letošnjem januarju objavljeno poročilo o varnosti generativne umetne inteligence (ibid.) je bilo osnovna platforma za pariško diskusijo.

Za mednarodno skupino znanstvenikov, ki so pripravili omenjeno poročilo, je scenarij, po katerem bi človek že v zelo kratkem času izgubil nadzor nad enim ali več deli sistema generativne umetne inteligence, še vedno hipotetičen. V glavnem obstaja med njimi soglasje, da trenutno to tveganje ne obstaja. Vendar pa so njihova mnenja, ali se bo to morda zgodilo v naslednjih nekaj desetletjih, bistveno manj

soglasna. Nekateri avtorji, ki so sodelovali pri pisanju poročila, zanikajo to verjetnost, drugi menijo, da je verjetno, da se bo to zgodilo, tretji zopet menijo, da gre za tveganje z zmerno verjetnostjo. Vsi pa se strinjajo, da na globalni ravni potrebujemo ustrezno družbeno regulacijo na področju razvoja generativne umetne inteligence, saj smo zaradi pomanjkanja te že danes pod bremenom množice dezinformacij, ki jih ustvarjajo boti, kibernetских napadov in zlorab na vojaškem področju. Strinjamo se lahko z oceno Yuvala Noaha Hararija, izraelskega zgodovinarja znanosti, ki je v svojem zadnjem delu »*Nexus. A Brief History of Information Networks from the Stone Stage to AI*« (Harari 2024) zapisal, da se je generativna umetna inteligenca, s tem ko sledi delovanju, ki izhaja iz nje same, že dodobra »zažrla« v obstoječi kod človeške civilizacije.

Z velikimi tveganji se v zadnjih letih soočamo tudi na področju razvoja tehnologij preoblikovanja človekovega genoma (»*human genome editing technologies*«), četudi se o teh tveganjih v javnosti morda razpravlja manj. Odkritje tehnologije CRISPR/Cas9 pred nekaj več kot desetimi leti je na tem področju naredilo pravo revolucijo, še posebej zato, ker je ta nova tehnologija odprla pot najrazličnejšim uporabam na področju biomedicinskega zdravljenja (Maresca in Deswal 2022). Vendar tudi razvoj te nove napredne tehnologije predstavlja dvorezni meč. Razvoj in uporaba tehnologij preoblikovanja človekovega genoma na eni strani vodita k vedno večji dobrobiti človeštva, na drugi pa k tveganjem.

Zdi se, da bi bila tu tveganja v primeru zlorab enaka ali celo večja, kot jih lahko zasledimo v povezavi z razvojem generativne umetne inteligence, še posebej, ker se je tu, kolikor so informacije o razvpiti » aferi He Jiankui« (Greely 2019; Sand, Brede-noord in Jongsma 2019) verodostojne, morda že prestopilo rdečo črto. Tehnologija preoblikovanja človekovega genoma je na področju medicine pospešila različne oblike zdravljenja z gensko terapijo. Kolikor se te genske terapije nanašajo na človekove somatske celice, se sicer lahko pojavijo določene etične dileme, vendar se vse te dileme vrtijo okrog klasičnih bioetičnih vprašanj. V tem primeru gre torej za vprašanja, ki ostajajo v domeni trenutno sprejemljivih družbenih norm, saj gre za take zadeve, kot je dosledna izvedba načela obveščenega informiranja pacientov s strani medicinskih strokovnjakov, kolikor se pacienti odločijo za gensko terapijo, upoštevanje avtonomne volje pacientov pri sprejemanju njihovih odločitev za take vrste terapij itd.

Na področju genomske znanosti in biomedicine se je situacija popolnoma spremenila, ko se je začelo v praksi eksperimentirati s tehnologijo CRISPR-Cas9 na človeških zarodnih celicah. Ko je kitajski biogenetik He Jiankui konec leta 2019 obvestil javnost, da sta se na temelju njegovega genskega inženiringa zarodnih celic rodili gensko preoblikovani dvojčici Nana in Lulu, je bila strokovna javnost v svetu šokirana, saj je prestopil rdečo črto. Ne glede na to, da je njegov postopek genske terapije na podlagi preoblikovanja zarodnih celic tudi po petih letih v marsičem zakrit v tančico skrivnosti (to gensko eksperimentiranje se je ne nazadnje izvajalo v informacijsko zaprti in nadzorovani kitajski družbi), je tak genski inženiring, ki si ga je privoščil kitajski genetik, na stečaj odprl vrata novi vrsti družbene evgenike. Preoblikovanje človeškega genoma se je iz medicinske terapije (terapije na somatskih celicah z namenom zdravljenja genskih bolezni) spremenilo v nov tip družbene evgenike, katere končni cilj bi lahko bila trajna in ireverzibilna sprememba človeka kot naravne vrste *homo sapiens*.

Tveganje za človeško vrsto v tem primeru ni prav nič manjše od tveganja, ki ga prinaša nenadzorovan razvoj generativne umetne inteligence.

SKLEP

V zaključku naj poudarim, da se posamezni akterji, ki se v državah članicah EU ukvarjajo z različnim vidiki strategije in vodenja raziskovalne in inovacijske politike, ne izogibajo razpravam o uporabi TRL kot merski skali, ki se uporablja za oceno zrelosti tehnologije, ko gre ta skozi fazo raziskav, razvoja in trženja. Žal pa te razprave ostajajo na precej deklarativni in posplošeni ravni ter ne vključujejo celostnejših premislekov o tem, kako na osnovi družbenih in etičnih izzivov, ki jih prinaša razvoj novih in naprednih tehnologij, pristopiti k temu merskemu »orodju«. Kmalu po objavi odmevnega poročila Maria Draghija o prihodnji konkurenčnosti Evrope (Draghi 2024) so ministrstva in agencije posameznih držav, članic EU, ki skrbijo za podporo raziskovanju, razvoju in inovacijam, objavila svoja izhodišča za 10. okvirni program Evropske komisije. Ta bo začel svojo pot leta 2028. Že bežen pregled spletnih strani nacionalnih ministrstev in agencij, na katerih te ustanove objavljajo vladne dokumente tudi v angleščini, ki so povezani z raziskovanjem, razvojem in inovacijami, zelo hitro razkrije, da v svojih izhodiščnih dokumentih za 10. okvirni program o TRL govorijo zelo malo ali nič. S podobno ohlapno govorico se srečujemo tudi v primeru Ministrstva za visoko šolstvo, znanost in inovacije Republike Slovenije (2024), ki je 17. 10. 2024 na svojih spletnih straneh objavilo dokument z naslovom *Stališče Republike Slovenije do 10. okvirnega programa Evropske unije za raziskave in inovacije*, vendar o TRL, ki ga definira kot »raven tehnološke zrelosti«, komajda kaj pove. Kaj več kot to, da je tu izrečena podpora ministrstva vzpostavitvi ravnovesja med načini programiranja TRL in vlogo aplikativnih raziskav na višjih stopnjah TRL, v tem dokumentu ne najdemo. Vsekakor je to premalo, da bi lahko za ključnega institucionalnega akterja raziskovalno-razvojne in inovacijske politike pri nas dejali, da je presegel ozek tehnokratski in marketinško usmerjeni pogled na TRL. Zdi se, da bo treba tudi v Sloveniji vložiti še veliko navora, da bomo dosegli tako rekonceptualizacijo TRL, ki bo vključevala zavedanje o tem, kako pomembna je tudi širša družbena in etična dimenzija razvoja novih naprednih tehnologij.

OPOMBA

- 1 Generativna umetna inteligenca spada v skupino modelov umetne inteligence, ki lahko ustvarjajo nove podatke (informacije) na podlagi vzorcev in struktur, naučenih iz obstoječih podatkov (informacij). Ti modeli lahko generirajo vsebine na najrazličnejših področjih, naj si bo besedil, slik ali glasbe. Pri analizi, razumevanju in ustvarjanju teh vsebin, ki vedno bolj spominjajo na človeške stvaritve, se opirajo na tehnike globokega učenja in nevronske mreže (Ray 2023).

CITIRANA LITERATURA

Bainbridge, William Sims, in Mihail C. Roco. 2006. »Progressive Convergence«. V *Managing Nano-Bio-Info-Cogno Innovations – Converging Technologies in Society*, ur. W.

- Bainbridge, 1–9. Dordrecht: Springer Verlag.
- Beck, Ulrich. 1992. *Risk Society: Towards a New Modernity*. London, Newbury Park in New Delhi: Sage Publications.
- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Bruno, Ilenia, et al. 2020. »Technology Readiness Revisited: A Proposal for Extending the Scope of Impact Assessment of European Public Services«. V *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance (ICEGOV2020)*, 369–380. Dostopno 21. 1. 2025. <https://doi.org/10.1145/3428502.342855>.
- Bush, Vannevar. 1945/1960. *Science: The Endless Frontier*. Washington, D. C.: National Scientific Foundation.
- Draghi, Mario (ur.). 2024. *The Future of European Competitiveness*. Dostopno 3. 2. 2025. https://commission.europa.eu/document/download/97e481fd-2dc3-412d-be4c152a8232961_en.
- European Commission. 2011. *High-Level Expert Group on Key Enabling Technologies – Final Report, June 2011*. Dostopno 25. 1. 2025. https://www.eusemiconductors.eu/sites/default/files/uploads/20110607_HLEG-KETs_FinalReport.pdf.
- Frascati Manual. 2015. *Guidelines for Collecting and Reporting Data on Research and Experimental Development*. Pariz: OECD.
- Giddens, Anthony. 1991. *Modernity and Self-Identity – Self and Society in the Late Modern Age*. Stanford: Stanford University Press.
- Greely, Henry. 2019. »CRISPR'd Babies: Human Germline Genome Editing in the 'He Jiankui Affair'«. *Journal of Law and the Biosciences* 6 (1): 111–183.
- Harari, Yuval Noah. 2024. *Nexus. A Brief History of Information Networks from the Stone Stage to AI*. New York: Penguin Random House.
- Héder, Mihály. 2017. »From NASA to EU: The Evolution of the TRL Scale in Public Sector«. *The Innovation Journal – The Public Sector Innovation Journal* 22 (2): 2–23.
- Hollingsworth, Rogers. 2011. *Major Discoveries, Creativity, and the Dynamics of Science*. Complexity Design Society 15. Dunaj: Echoraum.
- International AI Safety Report. 2025. *The International Scientific Report on the Safety of Advanced AI*. AI Action Summit, januar 2025. Dostopno 2. 2. 2025. https://assets.publishing.service.gov.uk/media/679a0c48a77d250007d313ee/International_AI_Safety_Report_2025.
- IPCC. 2022. *Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.
- Kline, Stephen, in Nathan Rosenberg. 1986. »An Overview of Innovation«. V *The Positive Sum Strategy: Harnessing Technology for Economic Growth*, ur. R. Landau in N. Rosenberg, 275–305. Washington, D. C.: National Academy Press.
- Luhmann, Niklas. 1991. *Soziologie des Risikos*. Berlin in New York: Walter de Gruyter & Co.
- Mali, Franc. 2015. »Družbene znanosti in razvoj novih naprednih tehnologij: sintezna biologija«. *Družboslovne razprave* 31 (79): 63–81.
- Mali, Franc. 2024. »Does the Use of Large Language Models in Scientific Research Bring Us Closer to the Point in Time When Machines Will Surpass Humans?«. V *IS 2024 Proceedings of the 27th International Multiconference Information Science*, ur. A. Slana Ozmič et al., 39–43. Dostopno 12. 1. 2025. https://is.ijs.si/wp-content/uploads/2024/11/IS2024_Volume-B.pdf.
- Maresca, Marcello, in Sumit Deswal (ur.). 2022. *Genome Editing in Drug Discovery*. Hoboken, ZDA: John Wiley & Sons, Inc.
- Ministrstvo za visoko šolstvo, znanost in inovacije Republike Slovenije. 2024. *Stališče Republike*

- Slovenije do 10. okvirnega programa Evropske unije za raziskave in inovacije*. Dostopno 3. 2. 2025. <https://www.gov.si/novice/2024-10-17-stalisce-republike-slovenije-do-10-okvirnega-programa-evropske-unije-za-raziskave-in-inovacije>.
- Nordman, Alfred (ur.). 2004. *Converging Technologies – Shaping the Future of European Societies*. Luksemburg: Office for Official Publications of the European Communities.
- OECD. 2020. *Addressing Societal Challenges Using Transdisciplinary Research*. OECD Science, Technology and Industry Policy Paper 88 (junij). Dostopno 12. 1. 2025. https://www.oecd.org/content/dam/oecd/en/publications/reports/2020/06/addressing-societal-challenges-using-transdisciplinary-research_41211835/0ca0ca45-en.pdf.
- Ray, Partha. 2023. »ChatGPT: A Comprehensive Review on Background, Applications, Key Challenges, Bias, Ethics, Limitations and Future Scope«. *Internet of Things and Cyber-Physical Systems* 3: 21–154. Dostopno 24. 1. 2025. <https://doi.org/10.1016/j.iotcps.2023.04.003>.
- Sand, Martin, Annelien Bredenoord in Karin Jongsma. 2019. »After the Fact—The Case of CRISPR Babies«. *European Journal of Human Genetics* 27 (junij): 1621–1624.
- Schumpeter, Joseph. 1934/2008. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*. New Brunswick in London: Transaction Publishers.
- Splichal, Slavko. 2024. *Izhodišča za SAZU-jev posvet Od tržne konkurenčnosti tehnologije k družbeni tehnološki zrelosti*. Ljubljana, SAZU, 15. november 2024.
- Suarez-Villa, Luis. 2009. *Technocapitalism: A Critical Perspective on Technological Innovation and Corporatism*. Philadelphia: Temple University Press.
- Voltan, José Luiz Neves, Rômulo Girardi, Juraci Ferreira Galdino in Ronaldo Ribeiro Goldschmidt. 2024. »A Review and Classification of Technology Readiness Assessment Techniques Based on TRL Scale«. *Coleç. Meira Mattos, Rio de Janeiro* 18 (61): 1–28. Dostopno 23. 1. 2025. <https://ebrevistas.eb.mil.br/RMM/article/download/11602/10090/>

ETIČNE RAZSEŽNOSTI UMETNE INTELIGENCE

Ana Farič in Ivan Bratko

Uvod

Tehnologijo umetne inteligence (UI) sestavljajo razna področja, med njimi metode za splošno reševanje problemov in preiskovanje problemskih prostorov, predstavitev znanja in metode sklepanja, strojno učenje, računalniško razumevanje naravnega jezika, robotika, računalniški vid itd. Relativna pomembnost teh področij se je s časom spreminjala. Od leta 2000 je strojno učenje postalo najpomembnejše in kaže najvidnejši napredek.

Ideje za strojno učenje so se pojavile že zgodaj. Prvi bolje znan praktično uporaben program strojnega učenja je bil ID3 (Quinlan 1979). Kmalu se je pokazalo, da so potenciali strojnega učenja za praktično uporabo izjemni. Dober zgodnji primer je poskus (Bratko in Mulec 1980), ki je verjetno prva uporaba strojnega učenja v medicinski diagnostiki. Zanimivo je, da so nekateri takrat najbolj znani znanstveniki iz UI, posebej ameriški, v obdobju pred 1990 zmotno ocenjevali, da je uporaba strojnega učenja v zahtevnejših domenah preveč optimistična in ni realna. S tem so močno upočasnili uveljavljanje uporabe strojnega učenja po svetu. Tako je šele okrog leta 2000 strojno učenje postalo najperspektivnejše področje UI.

Vendar je popoln preobrat v vsesplošnem zanimanju za strojno učenje sledil šele po dodatnem tehničnem preboju globokega učenja z nevronskimi mrežami (LeCun, Bengio in Hinton 2015) in generativne UI po letu 2020 (npr. Bubeck et al. 2023). Generativna UI je posebna vrsta strojnega učenja. Pri klasičnem strojnem učenju gre tipično za učenje napovedovanja »razreda«, npr. v medicini za klasifikacijo novega pacienta v diagnostične razrede ali napovedovanje temperature v vremenski napovedi. Pri generativni umetni inteligenci pa se sistem nauči generirati nove *vsebine*, npr. besedila ali slike. Tako lahko na primer v medicinski aplikaciji generira splošen besedilni opis razlik in podobnosti med dvema boleznima.

Ta razvoj je v kratkem času omogočil ogromen napredek v vrsti aplikacij UI. Znani primeri so: kakovostno strojno prevajanje, uporaba v zdravstvu, znanosti (npr. pisanje znanstvenih člankov, oz. vsaj velika pomoč pri tem), izobraževanju, sijajni rezultati, ki jih dosegajo veliki jezikovni modeli na specialističnih izpitih iz medicine in prava, avtomatsko programiranje in odlični rezultati na tekmovanjih iz programiranja.

Vendar so se ob teh tehnoloških dosežkih v UI, še posebej ob pojavitvah velikih jezikovnih modelov, kot je ChatGPT, pojavila tudi relevantna vprašanja o tem, kje so meje zaupanja v to tehnologijo. Ob pojavitvi sistema GPT-4 leta 2023 je prišlo do velikega miselnega preobrata med nekaterimi najvidnejšimi strokovnjaki za UI. Ob uspehih generativne UI so še pred pojavitvijo GPT-4 ti eksperti pogosto komentirali

presenetljive uspehe npr. GPT-2 tako: »Da, zanimivo, toda ti sistemi nimajo pojma, o čem govorijo.« Kljub njihovi presenetljivi zmožnosti je uporabnik takratne jezikovne modele z lahkoto z nekaj vprašanji spravil v težave, v katerih so gladko odpovedali. Toda ko se je pojavil GPT-4, so isti eksperti zelo hitro spremenili mnenje iz »ne vedo, kaj govorijo«, v smeri: »Da, GPT-4 deluje presenetljivo dobro, tega nismo mogli pričakovati.« Vendar ni jasno, *zakaj* deluje tako dobro, saj nihče ne ve dobro, kaj se v resnici dogaja v tej ogromni nevronske mreži s 1000 milijardami parametrov. Komentarji ekspertov so se hipoma spremenili. Zdaj ne vemo več, ali »ne ve, kaj dela«, ali mogoče vendarle »ve, kaj dela«. Tudi eden od najpomembnejših razvijalcev globokega učenja, Geoffrey Hinton, ki je leta 2024 za dosežke na tem področju prejel Nobelovo nagrado, deli zaskrbljenost glede dejstva, da tej tehnologiji zaupamo tudi izjemno pomembne aplikacije, obenem pa ne razumemo, na kakšen način doseže tako prepričljivo delovanje. A obenem dela tudi očitne napake, ki pa jih je zaradi našega nerazumevanja nemogoče predvideti.

Zato so ta razvoj spremljala vprašanja, ali je ta tehnologija varna in »vredna zaupanja« (*»trustworthy AI«*), kakšne so posledice za delovanje družbe. Nekateri razlogi za skrb so: nerazumljivost nekaterih sistemov UI, možnosti manipulacije z ljudmi z metodami UI in s tem tudi vpliv na demokratične procese, vpliv na izobraževanje in tudi na cilje izobraževanja, večanje neenakosti, diskriminacija določenih skupin, npr. glede na raso ali spol.

Hiter razvoj in eksplozija novih načinov uporabe UI močno vplivata na naš način dela in življenja. Zato so se v tej zvezi vse pogostejše pojavljala vprašanja o t. i. etičnih vidikih UI. Te vidike je treba nasloviti, če želimo zagotoviti UI, ki bo vredna zaupanja, njena uporaba pa varna in etična.

V tem poglavju naredimo pregled etičnih vidikov UI. Nato se osredinimo na podrobnejšo razpravo o dveh od najpogostejše omenjanih etičnih vidikov: (1) pristranskost UI in strojnega učenja in (2) transparentnost (razumljivost, razložljivost) sistemov UI ter sposobnost razlage.

ETIČNI VIDIKI UMETNE INTELIGENCE

Razvoj umetne inteligence so spremljala tudi ugibanja o možnih nevarnostih, ki bi se lahko pojavile ob razvoju te tehnologije. Zgodnje razmišljanje se je ukvarjalo z vprašanjem, ali lahko UI prevzame oblast. Razmišljanje je šlo npr. v smeri inteligentnih robotov, ki naj bi si nasilno podredili ljudi. To so bile v glavnem naivne predstave, ki niso imele dobre osnove v realnosti. Tak scenarij ni bil nikoli videti verjeten, saj bi morebitno nevarnost nasilnih robotov ljudje z lahkoto zaznali in preprečili.

Druga vrsta vprašanj je bila povezana s predvidevanji o t. i. točki tehnološke singularnosti. To je trenutek, ko bi se pojavila superinteligenca, to je umetna inteligenca, ki bi preseгла človeško (Kurzweil 2005). Predvidevanja o tem, kdaj bi naj se to zgodilo, so bila zelo nejasna že zato, ker v teh razpravah ni bilo dobro definirano, kdaj bomo šteli, da je UI preseгла človeško inteligenco. Inteligenca sestavlja vrsta sposobnosti (npr. pomnjenje, računanje, logično sklepanje, učenje, razumevanje in sporazumevanje v naravnem jeziku, sklepanje »po zdravi pameti«, kreativnost ...). V nekaterih od teh komponent inteligence je UI že zdavnaj preseгла sposobnosti člove-

ka, v nekaterih pa je bila vsaj do nedavna videti še zelo daleč. Tudi ni bilo jasno, kaj bi bile konkretne nevarnosti, povezane s točko singularnosti. Zato med raziskovalci UI scenarij s singularnostjo ni bil deležen velike pozornosti. Bostromovo (2014) razmišljanje o možnih nevarnostih superinteligence je prepričljivejše. Gre v smeri t. i. eksistenčne grožnje UI za obstoj človeštva.

Konkretne podrobnosti možnih nevarnosti, povezanih z UI, so bile dolgo zelo špekulativne. Zadeva se je spremenila, ko je UI dosegla velik praktičen napredek. Takrat so postali nekateri vidiki možnih nevarnosti konkretnjši in verjetnejši. Leta 2015 je skupina raziskovalcev UI opozorila na te vidike z odprtim pismom (dostopno na spletni strani Future of Life Institute 2015), ki ga je podpisalo več kot enajst tisoč strokovnjakov za UI. Glavno sporočilo pisma je bilo: UI je dosegla velik tehnični napredek, ki omogoča velike priložnosti, obenem pa je prišel čas, da smo pozorni in se izognemo možnim nevarnostim.

Besedna zveza »etični vidiki umetne inteligence« pomeni vprašanja, principe in metode v zvezi z zagotavljanjem varne in etične uporabe UI. Z etično uporabo mislimo na prakso uporabe UI, ki ni v nasprotju s splošno sprejetimi vrednotami in načeli. Med temi se v okviru etične uporabe UI najpogosteje omenjajo naslednje lastnosti: pravičnost in nepristranskost (npr. glede na raso ali spol), transparentnost in razumljivost, spoštovanje zasebnosti (kot nasprotje nelegitimnemu zbiranju osebnih podatkov in njihovi zlorabi), korektnost (verodostojnost) informacij kot nasprotje lažnim novicam (»fake news«) in zlonamerni manipulaciji z ljudmi (npr. v politične namene). Angleške fraze, ki pomenijo bolj ali manj isto kot »etični vidiki UI«, so: »*Ethical issues of AI, Ethics of AI, AI ethics*«. S temi principi se ukvarja med drugimi tudi mednarodna organizacija Int. Association for Safe and Ethical AI (IASAI), ustanovljena leta 2025.

V nadaljevanju tega razdelka so predstavljeni etični vidiki UI, ki so bili omenjeni v odprtem pismu iz leta 2015 ali so se izkristalizirali nekoliko pozneje. V naslednjih dveh razdelkih pa so nekateri od teh vidikov obdelani podrobneje.

Transparentnost in razložljivost sistemov UI, razlaga odločitev, problem črnih škatel

Načelo transparentnosti in razložljivosti pomeni, da naj bi imel uporabnik vpogled v delovanje sistema UI in da je sistem zmožen uporabniku razložiti svoje odločitve v človeku razumljivi obliki. Ta princip je posebej bistven, kadar so predlagane odločitve umetne inteligence za uporabnika pomembne, še posebej takrat, kadar obstaja možnost, da je sistem predlagal napačno odločitev (npr. v medicinski diagnostiki).

Nekatere metode UI zlahka, že po osnovnem načinu delovanja, zadostijo načelu razložljivosti. Primer je metoda za učenje odločitvenih dreves. Za odločitvena drevesa se navadno smatra, da so sama po sebi dobro razumljiva. Primer slabe razložljivosti pa je učenje z nevronskimi mrežami, še posebej globokimi mrežam, ki so nerazumljive že zaradi velikosti, navadno merjene s številom prilagodljivih parametrov v mreži. Odločitev nevronske mreže je rezultat skupnega delovanja velikega števila parametrov in je zelo težko, pogosto praktično nemogoče razložiti, kako so številne interakcije v mreži privedle do končne odločitve. Pri tem omenimo, da prav globoke nevronske mreže pogosto dosegajo največjo klasifikacijsko točnost. Zato včasih pride

do neke vrste tehtanja med točnostjo in razložljivostjo, ko v strojnem učenju lahko dosežemo boljšo razložljivost, vendar na račun točnosti.

Zaradi potrebe po razložljivosti je nastalo posebno področje UI, imenovano razložljiva UI (*»explainable AI«*, krajšano kot XAI). To področje se ukvarja z metodami, kako avtomatsko generirati človeku prijazno razlago odločitev sistema UI. Poseben izziv je, kako generirati razlago odločitev v primerih, ko je postopek računanja odločitev zaradi kompleksnosti nerazložljiv. Kot bo pojasnjeno v nadaljevanju, takrat pridejo v poštev metode, ki privedejo do enake odločitve, vendar na preprostejši način, ki omogoča vsaj približno razlago.

Pristranskost v strojnem učenju

Nekateri modeli UI, dobljeni kot rezultat strojnega učenja, so se v praksi izkazali kot pristranski, tipično glede na raso ali spol. Npr. sistem COMPAS (Angwin et al. 2016), ki se uporablja v ameriškem pravosodju, vzbuja vtis, da je diskriminatoren do temnopoltih obtožencev. Za nekatere druge odločitvene sisteme je bilo ugotovljeno, da so diskriminatorni do žensk, npr. v zaposlovanju. To temo so preučevali v raznih raziskavah. Te kažejo zapletenost teh vprašanj. Razni smiselni formalni kriteriji za ugotavljanje pristranskosti so med seboj nezdružljivi. Odločanje med možnimi načini za preprečevanje diskriminacije pa je pogosto stvar zapletene presoje. V razdelku o pristranskosti v nadaljevanju podrobneje predstavimo nekatere ugotovitve in smeri teh diskusij.

Manipulacija z ljudmi z orodji UI

Metode UI je možno učinkovito zlorabiti za manipulacijo z ljudmi po družbenih medijih. Namen take manipulacije je lahko komercialen ali političen, pri čemer lahko na ta način pride do vplivanja na demokratične procese – manipulacije volitev in referendumov. Orodja UI omogočajo avtomatsko generiranje osebno prilagojenih tendencioznih vsebin ali dezinformacij in ciljano širjenje sporočil na tak način, da so doseženi čim večji učinki in spremembe mnenja pri ciljanih posameznikih. Poglavje v tej knjigi Košmrlj in Bratko (2025) se podrobneje ukvarja z vplivom generativne UI na demokratične procese.

Zbiranje osebnih podatkov na spletu in njihova zloraba

Avtomatizirana manipulacija z ljudmi po družbenih medijih in spletu je učinkovita ob uporabi osebnih podatkov o posameznikih ter njihovih lastnostih in nagnjenjih. Zato ponudniki spletnih storitev nenehno zbirajo podatke o uporabnikih, tudi osebne, in jih uporabljajo za ciljano oglaševanje in širjenje tendencioznih vsebin ter si jih prodajajo med seboj. Evropska zakonodaja (GDPR) tako zbiranje podatkov sicer omejuje in ga dovoljuje ob soglasju uporabnika. Vendar zbiralci podatkov tako soglasje uporabnika praktično izsilijo, ne da bi bilo uporabniku res jasno, kako bodo podatki uporabljeni.

Globoki ponaredk

Na tehnologiji globokih nevronske mreže temeljijo tudi programska orodja za generiranje globokih ponaredkov (*»deep fake«*) ponarejenih video in audio vsebin.

Zaradi sposobnosti za preprečljivo ponarejanje slike ali govora konkretnih znanih oseb so ti ponaredki praktično nerazpoznavni in omogočajo učinkovito širjenje neresnic po družbenih medijih. S tem dodatno povzročajo splošno informacijsko negotovost.

UI v avtonomnih orožjih

Avtonomna orožja samostojno, brez posredovanja človeškega operaterja, izbirajo tarčo in se odločijo za napad. Pri tem je lahko odločitev napačna in usodna za nedolžne žrtve. Nekateri si razlagajo, da so s tem, ko je odločitev za napad sprejela UI, človeški operaterji razbremenjeni za odgovornost v primeru napake in nedolžnih žrtev. S tem se ob uporabi avtonomnih orožij standardi odgovornosti zmanjšujejo in mnogi, vključno s številnimi strokovnjaki UI, zato menijo, da z uvajanjem avtonomnih orožij prihaja do še ene oboroževalne tekme. Prav zaradi domnevne manjše odgovornosti ljudi (in prenašanja odgovornosti za nedolžne žrtve na umetno inteligenco) se zmanjšuje prag za odločitve držav za vojaške konflikte in posege. V tej smeri grede tudi pričakovanja države agresorke, da bo uporaba avtonomnih orožij zmanjšala število lastnih žrtev vojne. S tem lahko odločitev za vstop v vojno postane preveč lahkotna. Ta bojazen je tudi glavno sporočilo odprtega pisma, ki je nastalo med konferenco IJCAI 2015 (Int. Joint Conf. on AI, tj. najprestižnejša klasična vsakoletna konferenca iz UI). To pismo je zbralo več kot 34 tisoč podpisov (Future of Life Institute 2015).

Avtomatsko razpoznavanje čustev

Strojno učenje je omogočilo tudi učenje razpoznavanja človekovih čustev iz izraznih vzorcev na obrazu, ki se pojavljajo pri osebah ob določenih čustvenih stanjih. Ti vzorci omogočajo tudi razpoznavanje čustev, ki jih sicer oseba ne bi rada izdala. Posebej avtomatsko razpoznavanje prikritih čustev je etično sporno, saj naj bi imel človek pravico, da sam odloča o tem, ali čustva pokaže ali ne.

PRISTRANSKOST V UMETNI INTELIGENCI¹

Pojav pristranskosti v umetni inteligenci

Z naraščajočo uporabo strojnega učenja so se v zadnjih 5–10 letih pojavili primeri aplikacij, deležni izrazito odklonilnih odzivov tako s strani splošnih medijev kot znotraj strokovne literature. Pogosto so izpostavljeni sistemi, uporabljeni v domenah pravosodja, zaposlovanja in bančništva. Kritiki opozarjajo, da so algoritmi in sistemi strojnega učenja nepravilni in pristranski glede na t. i. zaščitene atribute, kot so rasa, spol in starost posameznika. Trdijo, da odločitve umetne inteligence (UI) temeljijo na teh atributih namesto na objektivni oceni dejstev (Hellström, Dignum in Bensch 2020). Strokovnjaki z različnih področij obravnavajo t. i. problem pristranskosti strojnega učenja. Skušajo opredeliti, kaj pristranskost pomeni, od kod naj bi izhajala in, najpomembneje, kako jo ustrezno nasloviti.

Na razvijajočem se področju etike v UI (npr. UNESCO 2021) se vprašanje pristranskosti strojnega učenja uvršča med osrednje teme. Politiki ga pogosto omenjajo v povezavi z regulativnimi načeli, namenjenimi zagotavljanju etične in varne

1 Vsebina razdelka o pristranskosti UI je bila deloma objavljena v Farič in Bratko (2023).

uporabe UI, npr. Artificial Intelligence Act (Evropski parlament 2024). Vendar pa v teh razpravah pogosto ni jasno opredeljeno, kaj pristranskost strojnega učenja in UI pravzaprav pomeni. Posledično regulativni ukrepi na tem področju ostajajo pomanjkljivo definirani in zgolj na zelo abstraktni ravni. Izraz pristranskost v kontekstu strojnega učenja za različne avtorje pomeni različno. Celó v strokovni literaturi s področja UI ni popolnega soglasja in splošno sprejete tehnične definicije pristranskosti, ki bi omogočala operativno uporabo za njeno preprečevanje (Hellström, Dignum in Bensch 2020). Poleg tega je za razne smiselne definicije pristranskosti matematično dokazano, da jim, razen v posebej trivialnih primerih, ni mogoče zadostiti hkrati (Hüllermeier, Fober in Mernberger 2013).

V nadaljevanju pregledamo različne definicije pristranskosti in različna mnenja o tem, kako naj bi se tega problema v praksi najuspešneje lotili. Različna mnenja najprej ilustriramo z znanim kontroverznim primerom, s sistemom COMPAS. Zaključki nakazujejo, da je za ustrezno obravnavo nujno upoštevati družbene vrednote in jih z interdisciplinarnim sodelovanjem operacionalizirati z demokratično sprejetim družbenim dogovorom v obliki ustrezne zakonodaje. K boljšemu splošnemu razumevanju pristranskosti v UI v praksi bi prispevalo tudi izboljšanje splošne izobrazbe o UI in njenih metodah.

Primer: sistem COMPAS

Neenotnost na področju obravnave pristranskosti bomo v nadaljevanju pokazali na primeru sistema COMPAS (Angwin et al. 2016; Dressel in Farid 2018; Flores, Bechtel in Lowenkamp 2016; Holsinger et al. 2018). Sistem COMPAS je v številnih publikacijah obravnavan kot verjetno najbolj kontroverzen primer, ki ponazarja domnevno pristranskost delovanja UI. COMPAS (*»Correctional Offender Management Profiling for Alternative Sanctions«*) je odločitveni sistem, ki ga na mnogih ameriških sodiščih uporabljajo sodniki za oceno tveganja povratništva. Sistem ocenjuje verjetnost, da bo obsojenec v dveh letih po izpustu ponovno storil kaznivo dejanje. COMPAS je razvilo ameriško podjetje, takrat imenovano Northpointe (danes Equivant). Sistem upošteva 137 atributov o posameznem prestopniku. Podatki so pridobljeni bodisi od obsojenca bodisi iz njegove kriminalne datoteke. Te podatke nato analizira poseben algoritem, ki kot poslovna skrivnost podjetja ni splošno znan. Na podlagi analize algoritma poda oceno med 1 in 10, pri čemer višja ocena pomeni višje tveganje za povratništvo.

Razprava o sistemu COMPAS se je začela z odmevnim člankom v ProPublici (Angwin et al. 2016). Skupina preiskovalnih novinarjev je predstavila svojo analizo sistema COMPAS in eksperimentiranje z realnimi podatki o več kot 7000 obtožencih na Floridi v letih 2012 in 2013. Ugotovili so, da je samo 61 odstotkov ljudi, ocenjenih kot verjetni povratniki, dejansko ponovilo kaznivo dejanje. V nadaljnji analizi so se osredinili predvsem na rasni vidik in prišli do zaključka, da je sistem pristranski do temnopoltih obtožencev. Spremljali so, koliko obtožencev, ocenjenih z visokim tveganjem za povratništvo, je bilo v naslednjih dveh letih dejansko ponovno obsojenih. Zanimalo jih je, kako pogosto so se napovedi sistema COMPAS razlikovale od dejanskih izidov. Posebej sta zanimivi dve vrsti napak: (1) napačen pozitiven primer, ko sistem napačno napove ponovitev prestopka, in (2) napačen negativen primer, ko sistem napačno napove, da obtoženec ne bo ponovil prestopka. Tovrstne napake v strojnem učenju

običajno merimo z meriloma FPR in FNR, ki sta definirani kot: (1) FPR (*»false positive rate«*) je delež napačno klasificiranih primerov med primeri, ki so ocenjeni kot pozitivni, in (2) FNR (*»false negative rate«*) je delež napačnih klasifikacij med tistimi primeri, ki so ocenjeni kot negativni. Zanimalo jih je, ali sta ti dve meri različni za belopolte in temnopolte obtožence. Dobljeni rezultati so v naslednji tabeli. Rezultati nakazujejo, da je COMPAS prijaznejši do belopolnih obsojencev:

	Belopolti	Temnopolti
FPR	23,5 %	44,9 %
FNR	47,7 %	28,0 %

Te rezultate so Angwin et al. (2016) interpretirali kot očitno pristranske do temnopolnih obsojencev in zato ocenili uporabo sistema COMPAS kot neprimerno in diskriminatorno. Taki interpretaciji bi bilo težko nasprotovati.

Dodaten problem, ki ga izpostavijo Angwin et al. (2016), je pomanjkanje transparentnosti pri odločanju sistema COMPAS, saj je algoritem zaščiten kot poslovna skrivnost. Poleg tega sistem COMPAS ne ponudi razlage za svoje napovedi. Zaradi pogoste citiranosti tega članka je COMPAS postal najbolj znan primer pristranskosti v strojnem učenju – tako v strokovnih krogih na področju strojnega učenja kot tudi med širšo javnostjo brez strokovnega znanja o UI. Kljub temu se COMPAS še vedno uporablja.

Na članek v ProPublici (Angwin et al. 2016) je odgovorila skupina strokovnjakov iz ameriškega pravosodja z delom z zgovornim naslovom *»False positives, false negatives and false analysis: A rejoinder to Machine Bias ...«* (Flores, Bechtel in Lowenkamp 2016). V njem so opozorili na več spornih odločitev v analizi ProPublice, izvedli lastno eksperimentalno raziskavo in zaključili, da so trditve v ProPublici (Angwin et al. 2016) napačne. Čeprav se ta kritika zdi upravičena, bi bila bolj prepričljiva, če bi avtorji jasno pokazali, kje točno je prišlo do ključne napake v analizi ProPublice. Namesto tega so predstavili svoj eksperimentalni rezultat, ki naj bi dokazal, da so osumljenci obravnavani enakopravno, ne glede na raso. Do tega rezultata so prišli tako, da so analizirali ocene tveganja obsojencev na lestvici od 1 do 10, kot jih oceni COMPAS. Iz teh ocen so izračunali mero AUC (površina pod ROC-krivuljo), ki se v strojnem učenju tudi pogosto uporablja kot indikator uspešnosti modela. Mera AUC je zanimiva zato, ker predstavlja verjetnost, da napovedni sistem pravilno loči med pozitivnimi in negativnimi primeri. To pomeni, da če vzamemo naključna primera (obtoženca, izmed katerih je prvi ponovil kaznivo dejanje, drugi pa ne), bo sistem z verjetnostjo, enako vrednosti AUC, pravilno določil, kateri je pozitiven in kateri negativen. Po navedbah Floresa, Bechtela in Lowenkampa (2016) je vrednost AUC za belopolte osebe znašala 0,69, za temnopolte pa 0,70. Razlika med vrednostma ni statistično značilna. Iz tega so sklepali, da COMPAS očitno ni rasno diskriminatoren in da rezultati ProPublice, ki kažejo na diskriminacijo, ne morejo biti pravilni. Vendar tak posredni argument dopušča dvom, saj mere AUC ter FPR in FNR med seboj niso enoznačno povezane.

Dressel in Farid (2018) v svoji raziskavi poročata o relevantnem poskusu, v katerem sta raziskovala napovedno točnost naključno izbranih posameznikov brez

domenskega znanja in točnost preprostega linearnega klasifikatorja. Primerjala sta napovedno točnost sistema COMPAS z uspešnostjo omenjenega linearnega klasifikatorja. Eksperiment z napovedovanjem človeških udeležencev (izveden z množičenjem, »crowd sourcing«) sta izvedla na podmnožici podatkov s Floride, ki je zajemala 1000 od skupno 7000 obtožencev iz študij (Angwin et al. 2016; Flores, Bechtel in Lowenkamp 2016). Ker bi bila uporaba vseh 137 atributov iz izvirne zbirke podatkov za poskus s človeškimi udeleženci nepraktična, sta izbrala zgolj sedem atributov. Ti niso vključevali rase. Presenetljivo je, da je bila napovedna točnost neekspertov v teh poskusih praktično enaka kot napovedna točnost sistema COMPAS. Zanimivo je tudi, da so bile človeške napovedi v tem poskusu podobno pristranske kot napovedi sistema COMPAS, merjeno z vrednostmi FPR in FNR za belopolte in temnopolte. Poleg tega so rezultati ostali skoraj nespremenjeni, tudi če so človeški ocenjevalci prejeli dodatne informacije o rasi. Avtorja sta prav tako ugotovila, da preprost linearni klasifikator doseže podobno napovedno točnost, pri čemer uporablja zgolj dva atributa. Poleg tega bolj sofisticirani klasifikatorji niso izboljšali niti napovedne točnosti niti pravičnosti.

V članku iz pravosodja so Holsinger et al. (2018) kritizirali raziskavo Dresselove in Farida (2018), pri čemer so kritiko utemeljili na naslednjih argumentih. Udeleženci v raziskavi (Angwin et al. 2016) so bili rekrutirani na platformi Amazon Mechanical Turk, kjer so za sodelovanje prejeli finančno nagrado. Oceno povratništva so podajali na osnovi sedmih atributov: starost, spol, vrsta kaznivega dejanja, resnost kaznivega dejanja, število obsodb v odrasli dobi, število obtožb za mladoletniška kazniva dejanja in število obtožb za mladoletniške prekrške. Vsi omenjeni atributi so znani kot pomembni dejavniki tveganja za povratništvo. Po mnenju avtorjev je zmanjšanje števila atributov pomembno olajšalo nalogo napovedovanja v primerjavi z izvirno nalogo, ki je vključevala vseh 137 atributov. Taka kritika je utemeljena, saj je znano, da je izbira ustreznih atributov v strojnem učenju lahko izjemno zahtevna. Poleg tega so udeleženci po podaji posamezne ocene prejeli povratno informacijo o pravilnosti odgovora in svoji povprečni natančnosti. Tisti, ki so dosegli višjo stopnjo točnosti, so bili nagrajeni z nekoliko višjim plačilom. Tak eksperimentalni okvir pa se bistveno razlikuje od resničnega konteksta, pri čemer se odločevalci soočajo s kopico (pogosto nerelevantnih in pristranskih) informacij. Glede na vse te okoliščine so avtorji zaključili, da raziskava Dressela in Farida (2018) ni prispevala pomembnih novih spoznanj. Pokazala je zgolj, da lahko naključno izbrani ljudje na podlagi izbranih relevantnih dejavnikov in povratnih informacij bolje napovedujejo povratništvo kot strokovnjaki, ki delujejo v kompleksnem okolju resničnih primerov. Holsinger et al. (2018) pa niso pojasnili, zakaj si tudi strokovnjaki ne bi pomagali z že obstoječim znanjem o pomembnih izbranih atributih. Tako bi si tudi strokovnjaki olajšali nalogo s poenostavitvijo kompleksne informacije v majhno število relevantnih dejavnikov.

Neodvisno od teh rezultatov je Cynthia Rudin (2019) na podlagi strojnega učenja sintetizirala preprost in popolnoma razumljiv napovedni model, temelječ na podatkih s Floride (Angwin et al. 2016). Model sestoji iz pravil »if-then« (navedenih v nadaljevanju) in uporablja zgolj tri attribute: spol, starost in število prejšnjih kaznivih dejanj. V primerjavi s sistemom COMPAS so ta pravila trivialno razumljiva:


```

IF age between 18-20 and sex is male
  THEN predict arrest (within 2 years)
ELSE IF age between 21-23 and 2-3 prior offenses
  THEN predict arrest
  ELSE IF more than three priors
    THEN predict arrest
    ELSE predict no arrest

```

Ta napovedni model dosega praktično enako točnost kot COMPAS. Ima pa tudi podobne mere FNR in FPR (po podatkih s Floride). Glede na to je tudi ta model, ki ga je iz učnih primerov generiral nedvomno nepristranski algoritem učenja, pristranski v enakem pogledu kot COMPAS. Podobne rezultate za FPR in FNR sta dobila tudi Dressel in Farid (2018) za učenje iz istih podatkov z enostavno linearno regresijo. Zanimivo je, da te indikacije pristranskosti ni nihče komentiral, niti Dressel in Farid (2018) niti Rudin (2019) niti v drugih člankih, ki preučujejo COMPAS.

Iz predstavljenih rezultatov je mogoče sklepati, da je napovedovanje tveganja povratništva kljub obsežnim razpoložljivim informacijam o obtožencih izjemno zahtevno, saj boljše napovedne točnosti, kot kaže, ni mogoče doseči. Obenem se zdi, da je skoraj vse doseženo z dvema ali s tremi najbolj koristnimi atributi in da preostalih več kot 130 atributov ne prispeva k izboljšanju napovedi. V skladu s tem nekateri avtorji (Angwin et al. 2016; Dressel in Farid 2018) zaključujejo, da uporaba strojnega učenja v pravosodju nima obetavne prihodnosti. Tak zaključek pa je prenagel in preveč poenostavljen, na kar opozarja Spielkamp (2017). V številnih drugih aplikacijah, kot je medicinska diagnostika, je strojno učenje v večini primerov že preseglo napovedno točnost strokovnjakov, kar so potrdili številni eksperimenti, npr. Cestnik, Kononenko in Bratko (1987).

Različna mnenja o pristranskosti in uporabnosti sistema COMPAS kažejo na pomanjkanje splošno sprejetih operativnih definicij pristranskosti in pravičnosti v strojnem učenju. Problem lepo ponazarja pogosto citirani članek (Mehrabi et al. 2021), ki preučuje številne relevantne definicije pristranskosti, vendar ne ponudi skupne sinteze, ki bi zmanjšala konceptualno kompleksnost in omogočila praktičen pristop k problemu pristranskosti. Članek povzroči dodatno zmedo, ko hitro odpravi COMPAS kot očitno pristranski in neuporaben, pri čemer ne upošteva argumentov Floresa, Bechtela in Lowenkampa (2016).

Formalne definicije pravičnosti in izzivi za implementacijo v praksi

V splošnih medijih se strojnemu učenju pogosto očita pristranskost bolj po občutku, brez natančne opredelitve matematično preverljivih kriterijev, na podlagi katerih bi bilo pristranskost mogoče dokazati. Izjave, kot sta: »Sistem se je v sodstvu izkazal kot pristranski do temnopoltnih obtožencev« (Angwin et al. 2016) ali »Sistem je pri ocenjevanju kandidatov za zaposlitev pristranski do žensk« (Blanzeisky in Cunningham 2021) uporabljajo splošne izraze, kot so »pristranskost algoritmov«, »pristranskost strojnega učenja«, »pristranskost umetne inteligence«. V mnogih primerih so bile te izjave opremljene s preprostimi, a pogosto preveč poenostavljenimi razlagami, na primer: »Sisteme strojnega učenja razvijajo skoraj izključno belopoltni moški, zato ...«

Danes je jasno, da problem pristranskosti ni tako trivialen. Pretirano poenostavljene razlage so vse redkejše, prav tako pa postaja očitno, da izraz »pristranskost algoritmov« ni primeren, saj ustvarja napačen vtis, da algoritmi lahko imajo zle namene in da ne delujejo na podlagi matematičnih in statističnih principov (Hardt, Price in Srebro 2016). Glavni cilj metod strojnega učenja je vedno odkrivanje zakonitosti v realnem svetu na podlagi razpoložljivih podatkov. Težava nastane, če v realnem svetu že obstajajo pristranske prakse. Podatki, zbrani v takem okolju, odražajo pristranskost, ki jo algoritem za učenje zazna in reproducira. Če nato rezultate, dobljene iz pristranskih podatkov, znova uporabimo v realnem svetu, s tem reproduciramo in utrjujemo že obstoječo pristranskost (ibid.).

Kljub vsemu še vedno ni dovolj natančno definirano, kaj pristranskost pravzaprav je. Pogosto gre za vtis pristranskosti, pri čemer se predsodki za ali proti posamezniku ali skupini kažejo na način, ki je zaznan kot nepravilen (Ntoutsis et al. 2020).

Poglejmo, v čem so težave pri definiranju pristranskosti. Že na področju strojnega učenja se termin »pristranskost« uporablja v raznih pomenih. Izraz se v strojnem učenju nanaša na več fenomenov (Hellström, Dignum in Bensch 2020):

- *Induktivna pristranskost*: To je princip, po katerem se algoritem odloči za eno izmed številnih možnih hipotez, ki so vse na nek način utemeljene z učnimi podatki. Gre za nabor (eksplicitnih ali implicitnih) predpostavk algoritma, katerih cilj je posplošiti niz opazovanj (učnih podatkov) v splošni model domene (Hüllermeier, Fober in Mernberger 2013). Ta vrsta pristranskosti je nepogrešljiv mehanizem, ki omogoča strojno učenje, zato je v principu pozitivna komponenta, brez katere strojno učenje sploh ni mogoče. Primer take pristranskosti je načelo Occamove britve, ki pravi: če imamo na voljo dve razlagi zbranih podatkov, ki obe enako dobro razložita podatke, je bolj smiselno izbrati preprostejšo razlago (Gordon in Desjardins 1995; Hellström, Dignum in Bensch 2020; Ntoutsis et al. 2020). Čeprav ima izraz pristranskost pogosto negativen prizvok, je induktivna pristranskost pozitivna in celo nujna komponenta strojnega učenja, kar lepo razložijo Gordon in Desjardins (1995) in osnovni učbeniki UI.
- *Pristranskost v učnih podatkih*: Pristranskost, ki odraža obstoječe pristranskosti v ustaljenem odločanju na danem področju uporabe (npr. pristranskost strokovnjakov v dejanski sodni praksi, iz katere so pridobljeni učni podatki) (Blanzeisky in Cunningham 2021; Dastin 2018). Alelyani (2021), poudarja, da je taka pristranskost pogosto posledica kognitivne pristranskosti, ki je naravni fenomen človeškega mišljenja. Človeški možgani namreč filtrirajo ogromne količine informacij in ohranjajo le tiste, ki so za nas relevantne. Ker so algoritmi učeni na podatkih, ki odražajo človeško vedenje, pogosto reproducirajo kognitivno pristranskost, ki jo vsebujejo vhodni podatki. Avtorji to vrsto pristranskosti poimenujejo različno: Blanzeisky in Cunningham (2021) jo označujeta kot negativno dediščino (»negative legacy«), medtem ko jo Hellström, Dignum in Bensch (2020) imenujejo zgodovinska pristranskost (»historical bias«).
- Pristranskost, ki izhaja iz neprimerne postopka zbiranja podatkov oz. *vzorčenja* (Hellström, Dignum in Bensch 2020): Ta vrsta pristranskosti se pojavi, kadar je npr. za določeno skupino ljudi na voljo bistveno manj primerov kot za druge

skupine. V skladu z matematično utemeljenimi statističnimi in verjetnostnimi načeli lahko nekatere skupine, tipično manjšinske, izpadejo kot diskriminirane (včasih celo v pozitivnem smislu!) preprosto zato, ker metode za ocenjevanje verjetnosti upravičeno ocenijo te verjetnosti drugače, kadar je na voljo manj podatkov. Tako pristranskost Blanzeisky in Cunningham (2021) imenujeta podcenjevanje (*»underestimation«*). Sun, Nasraoui in Shafto (2020) opozarjajo na ključno vprašanje o izvoru učnih podatkov. Tradicionalno so se algoritmi pri učenju zanašali na zanesljive oznake, ki so jih določili strokovnjaki. Danes pa se mnogi algoritmi učijo iz podatkov, ki izvirajo iz širše družbe, pri čemer so oznake in vzorci pogosto pristranski.

Prej omenjeni viri pristranskosti so razmeroma splošno sprejeti. Kljub temu ostaja izziv natančno definirati kriterije, ki bi objektivno določili, ali je dani sistem pristranski, ter omogočili tudi kvantitativno vrednotenje te pristranskosti. Obstajajo razne matematične mere, ki se zdijo smiselne, vendar so med seboj neskladne. Posledično za zdaj še ne obstaja preprosta in splošno sprejeta mera pristranskosti.

To kompleksnost dobro ponazarja izčrpni pregled različnih definicij pravičnosti kot nasprotja pristranskosti (Ntoutsis et al. 2020). Pravičnost lahko opredelimo na več načinov, na primer: »pravičnost na podlagi zavedanja« (*»fairness through awareness«*), pri čemer algoritem velja za pravičnega, če za podobne posameznike poda podobne napovedi; enakost obravnave (*»treatment equality«*), pri čemer se zahteva enako razmerje med FNR in FPR v obeh skupinah glede na zaščiteni atribut (npr. raso); poštenost v relacijskih domenah (*»fairness in relational domains«*), pri čemer poleg atributov posameznikov upoštevamo še družbene, organizacijske in druge povezave med njimi.

Berk et al. (2021) ugotavljajo, da so različne formalne definicije pravičnosti lahko med seboj nekompatibilne, kar vpliva na zanesljivost ocen tveganja in predlagane algoritmične rešitve. Avtorji ugotavljajo, da, razen v trivialnih primerih, ni mogoče hkrati optimizirati točnosti in pravičnosti ter zadostiti vsem vrstam pravičnosti hkrati. Ta problem podrobneje preučijo Kleinberg, Mullainathan in Raghavan (2016). Definirajo tri osnovne in na videz očitne pogoje, ki jih mora sistem izpolnjevati, če naj bo nepristranski (pravičen). Presenetljivo se izkaže, da ti trije pogoji ne morejo biti izpolnjeni hkrati, razen v posebnih, trivialnih primerih, ki pa so za prakso nezanimivi. Torej so že te tri osnovne zahteve skupaj neuresničljive. Te tri zahteve so:

1. Kalibracija ocen verjetnosti. Če algoritem identificira skupino oseb, za katero ocenjujejo, da imajo določeno verjetnost za pripadnost pozitivnemu razredu, mora ta delež dejansko pripadati pozitivnemu razredu. Enak pogoj mora biti izpolnjen tudi za vse skupine oseb, ki se razlikujejo glede na zaščiteni atribut, kot je rasa ali spol. Ocene morajo odražati to, kar naj bi pomenile, in biti neodvisne od skupine glede na zaščiteni atribut, v katero posameznik spada.
2. Ravnotežje pozitivnega razreda. Povprečje »ocen tveganja« (stopnja pripadnosti pozitivnemu razredu) oseb pozitivnega razreda mora biti enako za vse skupine. V primeru sistema COMPAS bi to pomenilo, da bi morali imeti belopolti in temnopolti obsojenci, ki pripadajo pozitivnemu razredu, primerljive ocene tveganja.

3. Ravnotežje negativnega razreda. To načelo dopolnjuje prejšnje načelo pozitivnega razreda in pomeni, da morajo biti povprečne ocene stopnje pripadnosti negativnemu razredu za osebe v negativnem razredu enake za vse skupine.

Avtorji matematično dokažejo, da so te tri zahteve, čeprav si prizadevajo za isti cilj zmanjševanja pristranskosti, med seboj nekompatibilne, razen v posebnih primerih.

Kadar se pojavi pristranskost, se pojavi vprašanje, kako jo odpraviti. Obstaja več pristopov, izmed katerih sta najočitnejša (a) prepovedana uporaba »zaščitene atributov« in (b) obratna diskriminacija. Tipična zaščitena atributa sta rasa in spol. Princip zaščitene atributov pomeni, da algoritmu učenja prepovemo uporabo teh atributov pri odločanju o klasifikaciji posameznega primera. Ta pristop pogosto ni učinkovit, saj algoritem učenja uspe rekonstruirati vrednosti zaščitene atributov na podlagi drugih, nezaščitene atributov, ki korelirajo z zaščiteni. Na podlagi podatkov o izobrazbi ali lokaciji prebivališča lahko na primer algoritem sklepa o rasi osebe (Hardt, Price in Srebro 2016).

Princip obratne diskriminacije temelji na tem, da deprivilegiranim skupinam namenoma damo določeno prednost, s čimer naj bi odpravili učinke diskriminacije. Čeprav je ta ukrep očitno dobronameran, v praksi uvaja novo obliko nepravilnosti, ki je za nekatere vprašljiva (Alelyani 2021). Taka nepravilnost (obratna diskriminacija) je sicer lahko upravičena, vendar ne z vidika pravičnosti, temveč z vidika »višjih« vrednot, kot so popravljanje zgodovinskih krivic in doseganje dolgoročne pravičnosti na podlagičasne nepravilnosti. Gre torej za strateško uresničevanje družbeno sprejetih vrednot, ki pa v praksi zaradi zgodovinskih razlogov in vztrajnosti zahtevajo daljše časovno obdobje za implementacijo. Ostaja težavno vprašanje, do kolikšne mere je obratna diskriminacija smiselna, kar bi moralo biti določeno z demokratično sprejetim družbenim konsenzom, formaliziranim z ustreznimi zakonskimi rešitvami za vsak specifičen primer.

V praksi se reševanja pristranskosti lotimo znotraj treh faz strojnega učenja: 1) predprocesna faza, v kateri povečamo vzorec manjšine za bolj uravnoteženo predstavitev podatkov; 2) medprocesna faza, v kateri dodajamo omejitve, s katerimi kompenziramo za neenakomeren vzorec in 3) popprocesna faza, v kateri prilagajamo mejne vrednosti odločitve, predvsem za manjšinske skupine (Blanzeisky in Cunningham 2021; Mehrabi et al. 2021; Ntoutsis et al. 2020).

Pri razvoju metod in orodij je ključno, da se zavedamo potencialnih pasti (Alelyani 2021; Holsinger et al. 2018). Avtorji opozarjajo, da lahko nekatere rešitve nenaumno vodijo v nove oblike nepravilnosti. Hüllermeier, Fober in Mernberger (2013) izpostavijo pogost stranski učinek, ko poseganje v učne podatke privede do izgube pomembnih povezav med atributi ali celo do poslabšanja delovanja modela, merjenega s kazalniki, kot sta točnost in ocena F1. Kot primarni vzrok pristranskosti avtorji identificirajo predhodne odločitve, ki so generirale same učne podatke. Predlagajo taktiko Fair-SMOTE, ki omogoča zmanjšanje pristranskosti, obenem pa ohranja (ali celo izboljšuje) delovanje sistema. Ključni korak je mutacija podatkov, z enakomerno ekstrapolacijo za vse spremenljivke. Fair-SMOTE identificira neuravnoteženost podatkov in napačno označene podatkovne točke. Rezultat je generiranje bolj poštenih rezultatov. Avtorji zavračajo pomisleke (Berk et al. 2021), da je cena poštenosti

poslabšanje delovanja sistema. Zaključijo, da je namesto slepe uporabe optimizacijskih metod pomembneje upoštevati specifičnosti domene in to znanje uporabiti za izboljšanje v pogledu pravičnosti.

Zaključek

Pistranskost v nekaterih ključnih aplikacijah strojnega učenja je postala popularna in kontroverzna tema. V razpravah je nejasnost, ki izvira iz dejstva, da večina ljudi razume pojem pravičnosti in pristranskosti intuitivno. Pri tem pravičnost doživljamo na različne načine in v podrobnostih ni popolnega soglasja. Tako tudi ni soglasja o tem, kakšen naj bi bil jasen, matematično formuliran kriterij, s katerim bi brez dvoma kvantificirali pristranskost konkretnega sistema. Obstaja veliko kontroverznih vprašanj in odprtih tem, v zvezi s katerimi ni strinjanja. Ni soglasja o izvoru pristranskosti, niti o tem, katera orodja oz. metode so za soočanje s pristranskostjo najprimernejše.

Poštenost v strojnem učenju naj bi pomenila produkcijo odločitev, s katerimi bi bila družba zadovoljna. Problem nastane, ker si v tem nismo enotni. Primer COMPAS ponazarja, kako nujna je zedinjenost. COMPAS so analizirali različni strokovnjaki, katerih ugotovitve so si nasprotujoče. Nekateri pravijo, da je COMPAS pristranski, medtem ko drugi menijo, da ni. Spielkamp (2017) verjame, da imajo prav vsi, saj vsakdo razume pravičnost na svoj način. Študija Kleinberga, Mullainathana in Raghavana (2016) je še posebej zanimiva, saj avtorji matematično dokažejo, da so določene definicije pravičnosti med seboj nezdružljive, čeprav se na prvi pogled vse zdijo enako pomembne in nujne.

Za kakovostno obravnavanje pristranskosti so pomembni jasna opredelitev zaželenih družbenih vrednot, upoštevanje zgodovinskega okvirja in resničnega stanja sveta ter tudi izobrazba o vsaj osnovnih principih strojnega učenja. Šele na temelju tega znanja lahko kakovostno artikuliramo svoja pričakovanja. Potem se lahko odločimo, kateri koncept pravičnosti ustreza v specifičnih kontekstih. Corbett-Davies et al. (2018) menijo, da ima lahko v določenih kontekstih delovanje v skladu s popularnimi konceptcijami pravičnosti brez globljega razumevanja domene celo nasprotno učinke. Avtorji v kontekstu razvoja algoritmov predlagajo, da se namesto osredinjanja na aksiomska razumevanja pravičnosti raje osredinimo na njihove posledice, ki so močno odvisne od konteksta in domene.

V literaturi na splošno ni dobro predvideno, kako velik izziv bo v praksi reševanje problema pristranskosti. Pričakovanja glede vrednot bo treba namreč natančno opredeliti z ustreznimi zakonodajnimi rešitvami, ali naj bo na primer zaradi zgodovinskih krivic v konkretni aplikaciji realizirana obratna pristranskost in v kolikšni meri. Take formulacije bodo morale biti bolj tehnične, kot je običajno v zakonih in drugih predpisih, saj bodo osnova za konkretno implementacijo v algoritmičnih UI. Jasno je, da problem pristranskosti ni (zgolj) tehnološke narave in da morajo zato učinkoviti pristopi k reševanju problema vključevati tudi širši nabor ukrepov (Ntoutsis et al. 2020). Prizadevati si moramo za interdisciplinarno raziskovanje, pri katerem bodo inženirji na področju UI tesno sodelovali s strokovnjaki s področij etike in odločanja (Yu et al. 2018).

Za ustrezno razumevanje in ukrepanje na tem področju je ključna kakovostna splošna izobrazba. Pomanjkanje se namreč kaže v načinu poročanja, odzivanju jav-

nosti in zmedenosti strokovnjakov. Različni algoritmi postajajo neizogiben del vsakdanjosti. Nesprejemljivo je, da o njih ne samo, da vemo premalo, ampak imamo celo napačne predstave. Po drugi strani imamo tudi institucije, ki delovanja sistemov UI prav tako ne razumejo, hkrati take sisteme naročajo in jih nato uvajajo v svoje procese odločanja. Institucije pogosto nimajo dovolj znanja in virov, da bi znale specificirati primerna algoritmična orodja. Nujno je izobraziti ljudi, da bodo sposobni ustrezno artikulirati, kaj naj bi algoritem res meril in kateri kriteriji morajo biti izpolnjeni, da bo algoritem pravičen (Courtland 2018).

RAZLOŽLJIVOST V UMETNI INTELIGENCI

Problem razložljivosti v umetni inteligenci

Uspehi in uporabnost tehnologije strojnega učenja so privedli do splošne uporabe te tehnologije. Sistem se iz danih primerov nauči dajati napovedi in odločitve, ki so tipično kakovostne glede na točnost, vendar ne tudi vedno povsem zanesljive. V nekaterih primerih uporabe, kot je npr. zdravstvo, imajo lahko napačne odločitve zelo pomembne ali celo usodne posledice.

Na takih področjih uporabe ni pomembna le točna napoved sistema in je nujno, da uporabnik lahko razume in preveri odločitve UI (Barredo Arrieta et al. 2019). Za to je pomembno, da ima model strojnega učenja zmožnost razlage, kako je prišel do odločitve. Nekatere metode strojnega učenja brez težav tako razlago tudi res omogočajo. Nekatere druge metode, npr. posebej zmogljive nevronske mreže in globoko učenje (LeCun, Bengio in Hinton 2015), pa slovijo po nerazumljivosti. Napovedi, ki jo take metode dajejo, so tipično rezultat interakcij med ogromnim številom numeričnih parametrov (reda bilijon), kar daleč presega človekove sposobnosti razumevanja. Znan je t. i. problem *črnih skatel* (*»black box problem«*), ki pomeni, da delovanje modelov strojnega učenja ostaja za uporabnike nerazumljivo. Prav pomanjkanje razumevanja omejuje nadaljnjo in bolj praktično uporabo modelov v mnogih pomembnih domenah odločanja.

Potreba po razlagi je vodila v razvoj tehnik in pristopov t. i. razložljive umetne inteligence (XAI – *»eXplainable Artificial Intelligence«*), ki se posveča nalogi razlaganja kompleksnih modelov strojnega učenja. Nove učinkovite metode razlage so posebej pomembne na področju nevronske mreže in globokega učenja. Vendar je treba ugotoviti, da je tehnologija globokega učenja prešla v splošno uporabo, še preden je bil razvit kakovosten konceptualni okvir, ki bi omogočal jasno razumevanje delovanja in odločitev globokih nevronske mreže.

V nadaljevanju tega poglavja sledita pregled trenutnega stanja metod XAI in analiza pomanjkljivosti.

Kaj sploh je razlaga?

Razložljivost je izmuzljiv pojem ne samo na področju umetne inteligence (UI), pač pa širše na področju filozofije in drugih družboslovnih znanosti. Na področju UI se operira s koncepti, kot so vzročnost, informativnost, razumevanje, gotovost, zaupanje, transparentnost ipd. (Barredo Arrieta et al. 2019). Termin »razložljiva umet-

na inteligenca« je leta 2019 kot del svojega programa uporabila DARPA (Gunning et al. 2021). Od takrat je postal zelo popularen, ne gre pa za nov pojav. Kvečjemu gre za novo imenovanje dolgoletnih prizadevanj, med katerimi se raziskovalci trudijo prebiti do odgovora na vprašanje, zakaj je sistem prišel do določene napovedi (Holzinger et al. 2022). Prvi je verjetno opozoril na problem razločljivosti Donald Michie, in to že precej prej, kot je to formuliral v svoji objavi (Michie 1988).

Najsplošneje bi lahko razločljivost v UI opredelili kot razlago, zaradi katere je delovanje modela razumljivejše. Seveda je to zelo splošna opredelitev, v poskusih natančnejšega definiranja pa si raziskovalci niso enotni. Barredo Arrieta et al. (2019) navajajo, da mora model nuditi razlago za svoje delovanje in napovedi v obliki vizualizacije pravil in vpogleda v spremenljivke, ki bi lahko povzročile perturbacije modela. Ribeiro, Singh in Guestrin (2016) besedo "razložiti" opredeljujejo kot predstaviti besedilne ali vizualne elemente, ki omogočajo kvalitativno razumevanje odnosa med komponentami in napovedjo modela.

Ena od nekonsistentnosti v literaturi iz XAI je uporaba pojma interpretabilnost, ki je včasih sinonim razločljivosti, drugač ločen pojem, tretjič ena od kategorij razločljivosti. Barredo Arrieta et al. (2019) interpretabilnost razumejo kot pasivno, razločljivost pa kot aktivno lastnost modela. Interpretabilni so modeli, ki so razumljivi že sami po sebi (npr. odločitvena drevesa), razločljivi pa tisti, za katere obstajajo metode, ki generirajo ustrezno razlago.

Očitno je pomanjkanje konsenza o glavnih konceptih. Problem je, da vsaka definicija nastopa znotraj specifičnega konteksta, odvisnega od naloge, sposobnosti in pričakovanj raziskovalca. Opredelitve razločljivosti so tako pogosto vezane na specifično domeno. Posledično področje XAI še ni enotno glede definicije razlage, specifičnih ciljev in kriterijev, ki naj bi jim zadostovali modeli, da bi bili razumljivi (ibid.).

Metode razlag

Obstajajo številne metode razlag. Problem pa nastane pri njihovi klasifikaciji, ker ima praktično vsak avtor specifično definicijo razločljivosti, iz katere izhaja.

Ena splošnih kategorizacij je delitev na lokalne in globalne razlage. Lokalne so razlage, središčene okoli posameznega primera, pri čemer pa ostane delovanje modela kot celote nepojasnjeno. Na drugi strani globalne razlage omogočajo razumeti celoten model, so pa pogosto osnovane na približnih vrednostih (Alvarez-Melis in Jaakkola 2018; Hall, Ambati in Phan 2017; Ignatiev, Narodytska in Marques-Silva 2019; Yeh in Ravikumar 2021).

Druga splošna delitev je na modelno odvisne in modelno neodvisne (»agnostične razlage«, »*model-agnostic*«) razlage. Slednje s tehnikami, kot so relevantnost atributov, vizualizacija in poenostavitve, pridobijo določene informacije o postopku napovedovanja in so uporabne za vsako vrsto modela (Barredo Arrieta et al. 2019). Modelno odvisne razlage so uporabne zgolj za specifične vrste modelov (npr. maksimizacija aktivacije nevronov, ki jo opišejo Guidotti et al. (2018)) (Hall, Ambati in Phan 2017).

Barredo Arrieta et al. (2019) ločijo besedilne, vizualne, lokalne razlage, razlage s primeri, poenostavitvijo in z relevantnostjo atributov. Chen et al. (2022) opredelijo tri glavne kategorije razlag: osnovane na funkciji, na primerih in pojasnjevanju atri-

butov. Yeh in Ravikumar (2021) ločita razlage atributov in razlage primerov. Ali et al. (2023) razlage razdelijo glede na uporabljeno metodologijo in ločijo med razlagami, ki slonijo na vzratnem razširjanju (*»backpropagation«*), in razlagami s perturbacijami. Guidotti et al. (2018) ločijo: 1) odločitvena drevesa; 2) razlage, osnovane na pravilih; 3) razlage pomembnosti atributov, ki predstavijo težo in pomembnost atributov, ki jih je pri svoji napovedi upošteval model. Primer je znana metoda LIME (*»Local Interpretable Model-agnostic Explanation«*), primerna predvsem za razlago klasifikacije besedil in slik (Ribeiro, Singh in Guestrin 2016); 4) zemljevidi pomembnosti, ki izpostavijo ključne vidike predmeta, ki je analiziran. Primer je metoda CAM (npr., da gre za sliko z umivanjem zob, razložimo tako, da posebej poudarimo dele slik, na katerih se zobna ščetka približa ustom) (Zhou et al. 2016); 5) PDP (*»Partial Dependence Plot«*), pri čemer grafično prikažemo odnos med odločitvijo modela in vhodnimi podatki; 6) razlaga s prototipi, pri kateri z napovedjo dobimo primer, podoben našemu; 7) maksimizacija aktivacije, pri kateri opazujemo, kakšni vzorci vhodnih podatkov maksimizirajo aktivacijo določenega nevrona oz. nivoja.

Ignatiev, Narodytska in Marques-Silva (2019) predstavijo pojem formalne razločljivosti, zasnovan na logiki, pri čemer so razlage posledično zanesljivejše in globalno veljavne. Pristop temelji na računanju t. i. glavnih implikantov (*»prime implicant«*), kar omogoča kompaktno logično predstavitev delovanja modela.

Razni kriteriji dobre razlage

Če je eden od ključnih ciljev XAI izboljšanje zaupanja v sisteme UI, je nujno, da se pozornost usmeri k uporabnikom teh sistemov (Zhang et al. 2021). Dobre razlage bodo tiste, ki bodo upoštevale, komu so namenjene (Barredo Arrieta et al. 2019). To pomeni upoštevanje predznanja, ki ga imajo uporabniki. Opazen je trend, da razvijalci metod razlag tega ne upoštevajo dovolj. Ribera in Lapedriza (2019) opredelita tri skupine uporabnikov (razvijalci in raziskovalci, eksperti in laiki), ki zahtevajo različne vrste razlag.

Raziskovalci predlagajo različne kriterije za dobro razlago. V nadaljevanju navajamo nekaj primerov kriterijev.

Alvarez-Melis in Jaakkola (2018) opredelita tri kriterije:

- eksplicitnost: razlaga je takojšnja in razumljiva;
- t. i. zvestoba (*»faithfulness«*): ocene relevantnosti odražajo resnično pomembnost;
- stabilnost: za podobne vhodne podatke veljajo podobne razlage.

Chen et al. (2022) poudarjajo:

- robustnost oz. občutljivost: ustrezna sprememba razlage v primeru spremembe vhodnih podatkov;
- zvestoba: razlaga ponazarja dejansko odločanje modela;
- kompleksnost: kognitivni napor za razumevanje razlage;
- homogenost: zmožnost za pravilno razlago delovanja modela glede na različne skupine (v praksi se to po navadi nanaša na skupine, ki se razlikujejo glede na občutljive, zaščitene attribute).

Amgoud in Ben-Naim (2022) opredelita štiri aksiome, ki naj bi jim zadostile dobre razlage:

1. morajo biti informativne;
2. ne smejo vsebovati nepotrebnih informacij;
3. razlage razredov morajo pojasniti posamezne primere, hkrati pa morajo biti splošno uporabne;
4. razlaga mora vsebovati samo informacije, ki vplivajo na napoved.

Ocenjevanje razlag

Ocenjevanje razlag je najmlajše področje s široko paleto pristopov (Ribera in Lapedriza 2019). Za razliko od točnosti je kriterije, kot so varnost, pristranskost in razložljivost, težje kvantificirati (Doshi-Velez in Kim 2017).

Ocenjevanja se (najsplošneje) lahko lotimo na dva načina: 1) s človeškim ocenjevanjem ali 2) z uporabo računskih metod, ki merijo, kako dobro razlaga dejansko razloži delovanje modela. Glavna razlika med pristopoma je, da so računske metode objektivnejše, vendar pa upoštevajo človeški faktor le posredno. Drugače rečeno, ne kvantificirajo nujno človeškega razumevanja. Prednosti človeške ocene sta subjektivnost in večja deskriptivnost. Očitni pomanjkljivosti sta manjša ponovljivost in večja odvisnost od specifične naloge (Mohseni, Block in Ragan 2021).

Hsieh et al. (2021) in Zhang et al. (2021) predstavijo matematično ocenjevanje razlag na podlagi analize robustnosti. Matematično opredeljena mera nezvestobe ponazarja, kako dobro se razlaga ujema z modelom (Yeh in Ravikumar 2021).

Edmonds et al. (2019) so izvedli eksperiment, s katerim so preverili, kakšne razlage so pri ljudeh vzbudile največje zaupanje v robota, ki je odprl stekleničko. Robot se je naučil odpirati stekleničke iz človeških demonstracij, pri čemer je bilo ključno učenje zaporedij položaja rok in potrebne sile. Z rokavico s senzorji so zajeli podatke o sili in položaju rok v 64 človeških demonstracijah s tremi različnimi stekleničkami. Sledilo je kompleksnejše učenje, da bi bil robot svoje znanje zmožen posplošiti. Implementiran je bil haptični model, ki je robotu pomagal določiti potrebno silo, čeprav nima človeških rok. Ker odpiranje stekleničke poteka v več korakih (potiskanje, odvijanje itd.), je bil uporabljen tudi algoritem simboličnega planiranja, ki upošteva pravila o zaporedju potrebnih akcij, ki omogočajo druga drugo. S kombinacijo takega učenja je robot postal precej dober v odpiranju novih stekleničk. Udeleženci so bili razdeljeni v pet skupin. Vsaka je videla posnetek robota, ki opravlja nalogo, ter eno od možnih razlag: 1) simbolično: v realnem času so udeleženci videli z eno besedo opisano akcijo, ki naj bi razlagala, kaj robot na posnetku dela (npr. »*approach – grasp – push – twist – ungrasp – move – grasp – push ...*«); 2) besedilno: po ogledu posnetka robota so udeleženci prebrali kratko besedilo o tem, kako je robot opravil nalogo (npr. »*I succeeded to open the bottle because I pushed on the cap three times and twisted the cap twice*«); oz. 3) haptično razlago: vizualizacija sile prijema v vsakem trenutku odpiranja stekleničke oz. kombinacijo haptične in simbolične razlage. Največ zaupanja je spodbudila simbolična razlaga. Ta rezultat nakazuje, da so principi simboličnega planiranja koristni za razlago robotskega delovanja, tudi če niso nujni za planiranje rešitve naloge.

Mohseni, Block in Ragan (2021) so izvedli eksperiment, med katerim so udeleženci označili relevantna področja slike, ki so po njihovem mnenju bila najbolj reprezentativna za določen razred objektov (npr. mačka in pes). Rezultat je zemljevid pomembnosti, ki prikazuje območja slike, ki so jim udeleženci posvečali največ pozornosti. Te rezultate so primerjali z zemljevidi pomembnosti metode Grad-CAM. Zemljevidi so si podobni, vseeno pa je statistično testiranje pokazalo pomembne razlike. Distribucija relevantnih atributov je bila pri metodi Grad-CAM bolj uniformna. Udeleženci so v primeru živih bitij kot ključne pogosteje označevali obraze. Prav to so ugotovitve, ki so lahko v oporo pri razumevanju, kako dobre so razlage.

Nekatera odprta vprašanja

V tem razdelku opozorimo na nekatere razmeroma manj raziskane probleme in tudi na manj uporabljene pristope za XAI.

Tehtanje med točnostjo in razločljivostjo

Rudin (2019) v članku z zgovornim naslovom »*Stop explaining black box ML models for high stakes decisions and use interpretable models instead*« izraža determinirano stališče. Zavzema se za uporabo metod učenja, ki dajejo naučene modele, ki so sami po sebi razumljivi. Za take veljajo npr. odločitvena drevesa. Nasprotuje metodam učenja, katerih rezultati so v principu težko razumljivi. Med te štejemo posebno metode globokega učenja, ki sicer dosegajo visoko napovedno točnost v primerjavi z drugimi metodami učenja, toda ne zastoj: vsaj za ceno razumljivosti in potrebnega velikega števila podatkov za učenje. Pri tem gre Rudin morda res predaleč z optimističnim stališčem, ki implicitno predpostavlja možnost izgradnje elegantnih in razumljivih modelov za vsako problemsko domeno, s čimer zadene ob princip kompleksnosti Kolmogorova (Kolmogorov 1963).

Glede možnosti obstoja enostavnih modelov in razlag velja vsaj ena teoretična omejitev, ki jo definira kompleksnost Kolmogorova, ki določa, koliko spominskega prostora potrebujemo za najkrajši možni zapis danega objekta v računalniku. Obstajajo zapleteni objekti (torej tudi zapleteni napovedni modeli), ki jih niti teoretično ni mogoče predstaviti na kratek način. V takih primerih tudi razlaga ne more biti kratka in preprosta. Res pa je, da smo v praksi še zelo daleč od te teoretično dosegljive meje, torej imamo veliko prostora za izboljšanje. Ko zadenemo ob zid Kolmogorova, pa je še vedno možen kompromis, da za boljšo razločljivost žrtvujemo nekaj točnosti (Bratko 1997). Primer tehnične izvedbe tega tehtanja med točnostjo in razločljivostjo v učenju odločitvenih dreves sta razvila Bohanec in Bratko (1994).

Navezava razlage na uporabnikovo predznanje

Ali bo razlaga dobra, je odvisno od uporabnika razlage, konkretno od uporabnikovega predznanja o problemski domeni. Če je to kakovostno, zadošča en namig. Če je razumevanje domene slabo, mora biti razlaga podrobna in daljša. Tudi formulacija razlage je odvisna od obstoječega znanja na obravnavanem področju. Celo povsem pravilna in jedrnata razlaga je za eksperta na področju uporabe lahko nesprejemljiva in nenaravna. Kot primer omenimo, da so se nekateri primeri razlag, ki jih generirajo naučeni modeli v medicinskih domenah, kljub diagnostični točnosti zdravniku zdeli

povsem nenaravni (Bratko 1997). V enem od primerov je sistem razložil, da gre za vnetni revmatizem, ker ima pacient med drugim več kot dva prizadeta sklepa na roki. To diagnostično pravilo je dejansko točno. Vendar pa je zdravnik vztrajal, da mora imeti pacient prizadete sklepe na vseh petih prstih na roki, ker vnetni revmatizem tipično vpliva na vse sklepe. Ekspertno mnenje je bilo v tem primeru zelo jasno, čeprav je res, da bo pravilo vodilo do pravilne diagnoze v vsakem primeru, če je število vnetih sklepov karkoli med 2 in 5. Ustreznost razlage je odvisna ne le od klasifikacijske točnosti, temveč (tudi) od predznanja, ki ga ima uporabnik o tej obliki revmatizma.

Obstoječe metode razlage ta vidik povečini ignorirajo. Problem je tudi v tem, da ne omogočajo naravne uporabe predznanja. V tem pogledu je zelo obetaven pristop k strojnemu učenju t. i. induktivno logično programiranje (ILP), ki temelji na uporabi matematične logike. Že osnovna formulacija problema učenja v ILP vsebuje uporabo predznanja: dani so učni primeri E in predznanje BK (*»background knowledge«*), naloga učenja pa je sestaviti logično formulo H (hipoteza) tako, da primeri E logično sledijo iz BK in H.

Pristop ILP je skromno zastopan v obstoječih raziskavah iz strojnega učenja in razločljivosti. Lep primer njegove ustreznosti so raziskave, ki jih opisujejo Ai et al. (2023) in Muggleton et al. (2018). Te zasledujejo ne le osnovni cilj XAI (razlage odločitev strojnega učenja), temveč tudi cilj t. i. *»ultrarazločljivosti«*. Ta strožji kriterij strojnega učenja je definiral Michie (1988) (*»ultra strong criterion for ML«*). Strojno učenje je ultrarazločljivo, če je ne le razločljivo, temveč uporabniku omogoča tudi *operativno uporabo za lastno reševanje* novih problemov, npr. da strojno naučeno znanje lahko uporabi za lastno reševanje določenih matematičnih problemov ali igranje šaha.

Razlaga zaporedij odločitev v planiranju

Večina metod XAI generira razlago *posameznih* odločitev oz. klasifikacij. Pri razločljivem planiranju pa gre za razlago množice odločitev (npr. zaporedja akcij, ki robota vodi do cilja). Posebej za razlago planov se je formiralo področje razločljivega planiranja (Chakraborti, Sreedharan in Kambhampati 2020).

Razlaga planov je običajno zahtevnejša od razlage v klasifikacijskih problemih. Treba je razložiti, kako so posamezne akcije odvisne od drugih, da skupaj rešijo nalogo. Primer razlage zaporedja odločitev je razlaga šahovskih partij, pri čemer je treba razložiti celo zaporedje potez ali drevo odločitev, ki definira uspešno strategijo. Primer opisuje Bratko (2018) in vsebuje težko razumljive in briljantne poteze šahovskega programa AlphaZero.

Razlaga planov je aktualna tudi na področju vodenja sistemov. Lep primer razlage naučenega plana vodenja podajata Šoberl in Bratko (2023, 2025). Gre za klasično nalogo iz teorije vodenja sistemov: vodenje sistema voziček-palica. Na vozičku je vrtljivo vpeta palica. Palica je postavljena približno vertikalno, vendar se, če ne ukrepamo, prevrne na tla. S potiskanjem vozička levo oz. desno je treba loviti ravnotežni položaj palice okrog vertikale, obenem pa doseči, da se voziček horizontalno premakne iz začetnega položaja do danega cilja. Naučena strategija vodenja je lepo razločljiva. Najprej nekoliko presenetljivo potisnemo voziček v nasprotno stran od cilja, s čimer dosežemo, da se palica nagne proti cilju. To omogoči pospeševanje vozička

v smer proti cilju, hkrati pa se ohranja ravnotežje palice, ko je ta nagnjena naprej v smeri cilja.

Zaključek

Področje XAI se je v zadnjih petih do desetih letih močno razraslo. Mnogi zato predpostavljajo, da je bil to tudi začetek področja. V resnici je zavedanje pomembnosti, da naj bi bilo strojno učenje razložljivo, obstajalo že pred 40 leti (npr. Michie 1988). Že takrat so obstajale raziskave o razložljivih modelih. Kljub sedanji količini raziskav in nedvoumni uspehi se še vedno kaže, da pogrešamo nekatere ključne odgovore. Npr., že pred desetletji se je v sklopu istih prizadevanj pojavilo zavedanje, da potrebujemo formalne mere za ocenjevanje kakovosti razlag. Take sprejete mere še ni. Raziskovalci pri ocenjevanju razlag uberejo različne pristope, odvisne od raznih kriterijev (konteksta, domene, uporabnikov itd.).

Glede vprašanja, kaj je sprejemljiva razlaga, se v pomanjkanju splošnejših in principialnih kriterijev v sedanji praksi uporablja predpostavka, da so nekateri modeli razložljivi kar po definiciji, torej razložljivi sami po sebi. Mednje npr. navadno štejemo odločitvena drevesa ali pravila *če-potem*. Toda tudi ta kriterij je arbitraren. Kaj, če je odločitveno drevo zelo veliko, npr. da ima milijon vozlišč?

V tem razdelku smo opozorili tudi na počasen napredek pri razvoju metod za razlago zaporedij odločitev. Sem sodi razlaga planov za reševanje nalog, ki imajo eksplicitno definirane cilje. Plan je lahko zaporedje akcij ali pa tudi množica akcij, ki so delno urejene v času. Tu je treba razložiti tudi to, kako se akcije med seboj dopolnjujejo in na kakšen način skupaj dosežejo cilj. S tem so povezani izzivi sklepanja v velikih jezikovnih modelih, ki jih predstavimo v nadaljevanju.

Eden izmed možnih pristopov, ki upošteva principe planiranja v UI, je upoštevanje odvisnosti med akcijami. Nekatere akcije v planu neposredno dosežejo kakega od ciljev plana. Druge akcije pa ne dosežejo nobenega danega cilja neposredno, njihova funkcija je, da dosežejo pogoje, ki morajo biti uresničeni, da je možno izvesti druge akcije v planu. Taka razlaga plana je seveda povsem logična. Navadno pa vsebuje preveč podrobnosti. Če plan vsebuje nekoliko večje število akcij, npr. nekaj deset, postane tako podrobna razlaga spet težko razumljiva in za uporabnika neprijetna. V tem primeru bi bilo treba za sprejemljivo razlago plan razbiti v hierarhično strukturo, definirano s podcilji plana. Odkrivanje smiselni podciljev pa je lahko zelo težavno. Poseben izziv je, kako poiskati take podcilje, ki rezultirajo v za človeka čim naravnejši razlagi.

SPOSOBNOST SKLEPANJA V VELIKIH JEZIKOVNIH MODELIH

V tem poglavju smo navedli nekatere etične vidike UI, ki naj bi zagotavljali zaupanja vredno UI. Podrobneje smo opisali stanje v zvezi z zagotavljanjem dveh od teh vidikov: pravičnost strojnega učenja in razložljivost v UI. Kljub napredku na teh dveh področjih ostajajo odprta oz. ne dovolj raziskana vprašanja. Poleg tega je za varno in zaupanja vredno UI pomembno tudi, da sistem UI deluje tehnično pravilno, npr. v tem smislu, da zanesljivo daje pravilne rezultate. V tem razdelku opozorimo na dve vprašanji v zvezi z generativno UI, ki zahtevata dodaten tehnični razvoj. To sta zmož-

nost sklepanja v velikih jezikovnih modelih ter problem, včasih poimenovan kot neke vrste »kanibalizem« (*»self consuming models«*).

En tehnični vidik, ki zadeva generativno UI kot trenutno najpopularnejši način uporabe UI, je povezan z zmožnostjo sklepanja v velikih jezikovnih modelih. Ta vidik generativne UI vzbuja dvome. Sposobnost sklepanja v popularnih jezikovnih modelih vzbuja vtis krhkosti in zelo težko je predvideti napake v sklepanju. To dejstvo je videti paradoksalno v luči dejstva, da sicer na področju UI že dolgo obstajajo zmogljive in zanesljive metode za avtomatsko sklepanje (npr. Russell in Norvig 2020; Poole in Mackworth 2023). Vendar kaže, da uporaba že obstoječih in zanesljivih metod ni preprosta v okviru generativne UI. Na probleme opozarjajo razni avtorji, npr. Bubeck et al. (2023) in Mirzadeh et al. (2024). Prizadevanja v smeri tehničnih izboljšav (npr. Plaet et al. 2024) v pogledu sposobnosti sklepanja so sicer privedla do izboljšav, vendar so te v pogledu sistematične zanesljivosti še vedno negotove. Dober primer je poskus z velikimi jezikovnimi modeli, ki ga je predlagal P. van Eecke (osebna komunikacija). Modelu zastavimo dobro znano miselno uganko »volk-koza-zelje«, ki je definirana tako: Imam volka, kozo in zelje. Če pustim volka brez nadzora s kozo, jo bo volk požrl. Podobno je s kozo in z zeljem. Kako naj prepeljem vse tri stvari varno čez reko z majhnim čolnom, tako da lahko vzamem v čoln s seboj le eno stvar? Tipičen jezikovni model nalogo takoj prepozna in tudi ponudi pravilno rešitev. Toda kaj se zgodi, če nalogo nekoliko spremenimo, npr. da je koza sita in jo zato lahko pustimo samo z zeljem? Ali je možna krajša rešitev? Tu modeli pogosto odpovedo, naredijo napačen logičen sklep in predlagajo nepravilno rešitev. Mirzadeh et al. (2024) predstavijo in analizirajo vrsto primerov v reševanju matematično logičnih nalog in značilnih spodrslijav, ki so jih naredili jezikovni modeli. Res je, da so po tej objavi mnoge od teh tipov težav v jezikovnih modelih tudi že odpravili. Vendar je vtis, da gre bolj za lokalne popravke kot pa za sistematične izboljšave sklepanja.

Opozorimo na še en tehnični vidik, ki lahko postane kritičen v nadaljnjem razvoju generativne UI. Gre za pojav, ki ga primerjajo z neke vrste »kanibalizmom« med modeli generativne UI (*»self-consuming language models«*). Pri tem je mišljen verjeten razvojni scenarij tehnologije generativne UI, ko se bodo poznejše generacije modelov učile iz sintetičnih vsebin, in ne naravnih, človeških besedil in drugih vsebin. Po tem scenariju bodo prejšnje generacije modelov generirale sintetične vsebine, ki bodo na spletu na voljo za učenje bodočih generacij modelov. Pri tem lahko pride do naključnih napak, ki postopno privedejo do poslabševanja jezikovnih modelov poznejših generacij (Alemohammad et al. 2023; Briesch, Sobania in Rothlauf 2024).

ZAKLJUČNA RAZPRAVA

V tem poglavju smo podali pregled etičnih vidikov umetne inteligence ter podrobneje opisali predvsem tehnične vidike dveh od njih, pristranskosti in razložljivosti. Nismo pa se še dotaknili bolj globalnih vprašanj o tem, kako bo hitri razvoj UI vplival na družbo in človeštvo sploh.

Očitno je, da se zelo zmanjšuje pomen mnogih poklicev, pa tudi način človeškega dela v mnogih poklicih. Primer je prevajanje. Nekateri strokovnjaki UI so že pred

letom 1970 obljubljali, da bo kmalu na voljo kakovostno strojno prevajanje. Vendar se to ni uresničilo še nekaj desetletij in tako je strojno prevajanje dolgo veljalo za morda največjo neizpolnjeno obljubo UI. Ta situacija se je dramatično spremenila v vsega nekaj zadnjih letih. Za večino praktičnih potreb je prevajanje z UI postalo dovolj kakovostno in dosegljivo za drobec cene človeškega prevajanja. Profesionalni človeški prevajalci so res pomembni le še v zahtevnih primerih, ko gre pri prevajanju za specifične strokovne vsebine ali pa prevod zahteva literarno ustvarjalnost.

UI močno spreminja tudi izobraževalne procese in cilje izobraževanja. Ni več jasno, ali je treba, da pri učencih izobraževanje zagotovi sposobnost samostojnega pisnega izražanja, saj komunikacija poteka namesto na relaciji človek – človek vse bolj posredno na relacijah človek – računalnik – računalnik – človek. Zelo se spreminja tudi način raziskovalnega dela, od pregleda literature o sorodnih raziskavah do samega pisanja člankov. Ni več jasno, kaj je pri sodelovanju človek-UI prispevek človeškega avtorja in kaj računalnika.

Vendar v razpravah o teh spremembah vse bolj skrbi tudi najusodnejše vprašanje: Ali bo UI »prevzela nadzor«? Za to vprašanje se v angleščini navadno uporablja fraza »*Will AI take over?*« Lahko si je predstavljati, da se lahko z UI posamezniki ali skupine na neupravičen način dokopljejo do oblasti ali si ustvarijo druge neupravičene prednosti ali koristi. Vendar so take grožnje še vedno razmeroma obvladljive, saj so v očitnem nasprotju z obstoječimi zakoni in zato tudi pregonljive in jih je posledično vsaj v principu možno preprečiti z običajnimi sredstvi.

Obstajajo pa verjetnejši scenariji, da UI prevzame nadzor. Taki scenariji skrbijo vse več strokovnjakov UI in tudi drugih področij. Geoffrey Hinton, Nobelov nagrajenec leta 2024 za dosežke v učenju globokih nevronske mreže, na primer ocenjuje, da je to skoraj neizogibno. Meni, da bi to lahko preprečili le z usklajenim ukrepanjem večine držav sveta, kar bi se lahko zgodilo, če bi prišlo do velikega pritiska na vlade držav v tej smeri. Vendar pa ni videti, da bi bile vlade v zvezi s tem trenutno pod kakršnim koli pritiskom. Nasprotno, zdi se, da bo UI prevzela nadzor (in oblast) na povsem benigni način ob večinskem tihem odobranju velike večine človeštva. To je scenarij, ki se že uresničuje. Praktično povsod prevladuje prepričanje, da UI opravlja mnoga dela veliko bolje kot ljudje. To je običajno formulirano tako: »Z uporabo UI smo pri reševanju mnogih nalog vsi veliko učinkovitejši, zato bi bilo nespametno, da pri vseh teh nalogah ne bi uporabljali UI.« Ta praksa, ko ljudje sami od sebe z racionalnim razmislekom o sposobnosti UI tej prepuščajo vse več pomembnega dela in pomembnih odločitev, se hitro uveljavlja na pomembnih področjih. Ta praksa velja v vse večji meri npr. tudi za področje pravosodja – na primer pisanje obtožnih predlogov, pripravljalnih vlog in sodb. Po vsej logiki bo to kmalu veljalo, ali pa že velja, tudi za pisanje zakonov. Po enaki logiki bo tudi za presojo predlogov zakonov najbolj kompetentna UI, in neracionalno bi bilo, da bi parlamentarna telesa in poslanci sami ocenjevali zakone, ki jih napiše UI. Najkompetentneje jih bo ocenila UI. S tem bo človeštvo samo prepustilo nadzor nad zakonodajo umetni inteligenci. Lahko, da bo taka zakonodaja boljša in učinkovitejša, vendar je vprašanje, ali si to res želimo, ali to ne vodi do odtujitve človeku? Po podobni logiki bo to veljalo za področja znanosti, umetnosti in izobraževanja. Mnogi menijo, da je to pot brez možnosti povratka.

ZAHVALA

Prispevek je deloma nastal v okviru ciljnega raziskovalnega projekta V2-2272 Opredelitev okvira za zagotavljanje zaupanja javnosti v sisteme umetne inteligence in njihove uporabe, ob podpori Javne agencije za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije in Ministrstva za digitalno preobrazbo Republike Slovenije.

CITIRANA LITERATURA

- Ai, Lun, Johannes Langer, Stephen H. Muggleton, in Ute Schmid. 2023. »Explanatory Machine Learning for Sequential Human Teaching«. *Machine Learning Journal* 112: 3591–3632. <https://doi.org/10.1007/s10994-023-06351-8>.
- Alelyani, Salem. 2021. »Detection and Evaluation of Machine Learning Bias«. *Applied Sciences* 11 (14). <https://doi.org/10.3390/app11146271>.
- Alemohammad, Sina, et al. 2023. »Self-Consuming Generative Models Go MAD«. *arXiv:2307.01850*. <https://arxiv.org/pdf/2307.01850>.
- Ali, Sajid, et al. 2023. »Explainable Artificial Intelligence (XAI): What We Know and What Is Left to Attain Trustworthy Artificial Intelligence«. *Information Fusion*, 99. <https://doi.org/10.1016/j.inffus.2023.101805>.
- Alvarez-Melis, David, in Tommi S. Jaakkola. 2018. »Towards Robust Interpretability with Self-Explaining Neural Networks 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montreal, Kanada.
- Amgoud, Leila, in Jonathan Ben-Naim. 2022. »Axiomatic Foundations of Explainability«. *Proceedings of the Thirty-First International joint Conference on Artificial Intelligence (IJCAI-22)*: 636–642.
- Angwin, Julia, Lauren Kirchner, Jeff Larson in Surya Mattu. 2016. »Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks«. *ProPublica*.
- Barredo Arrieta, Alejandro. et al. 2019. »Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI«. *arXiv:1910.10045v2*. <https://doi.org/10.48550/arXiv.1910.10045>.
- Berk, Richard, et al. 2021. »Fairness in Criminal Justice Risk Assessments: The State of the Art«. *Sociological Methods & Research* 50 (1): 3–44. <https://doi.org/10.1177/0049124118782533>.
- Blanzeisky, William, in Pádraig Cunningham. 2021. »Algorithmic Factors Influencing Bias in Machine Learning«. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2104.14014>.
- Bohanec, Marko, in Ivan Bratko. 1994. »Trading Accuracy for Simplicity in Decision Trees«. *Machine Learning Journal* 15: 223–250. <https://doi.org/10.1007/BF00993345>.
- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Bratko, Ivan. 1997. »Machine Learning: between Accuracy and Interpretability«. V *Learning, Networks and Statistics*, ur. G. Della Riccia. Dunaj: Springer.
- Bratko, Ivan. 2018. »AlphaZero: What's Missing?«. *Informatica* 42 (1).
- Bratko, Ivan, in Peter Mulec. 1980. »An Experiment in Automatic Learning of Diagnostic Rules«. *Informatica* 4 (4): 18–25.
- Briesch, Martin, Dominik Sobania in Franz Rothlauf. 2023. »Large Language Models Suffer From Their Own Output: An Analysis of the Self-Consuming Training Loop«. *arXiv:2311.16822*.
- Bubeck, Sebastien, et al. 2023. »Sparks of Artificial General Intelligence: Early Experiments with GPT-4«. *arxiv:2303.12712v5*. <https://arxiv.org/abs/2303.12712>.

- Cestnik, Bojan, Igor Kononenko in Ivan Bratko. 1987. »ASSISTANT 86: A Knowledge-Elicitation Tool for Sophisticated Users«. V *Progress in Machine Learning: Proceedings of European Working Session on Learning EWSL 87*, ur. I. Bratko in N. Lavrač, 31–45. Sigma Press.
- Chakraborti, Tathagata, Sarath Sreedharan in Subbarao Kambhampati. 2020. »The Emerging Landscape of Explainable Automated Planning & Decision Making«. *arXiv:2002.11697*. <https://doi.org/10.48550/arXiv.2002.11697>.
- Chen, Zixi, Varshini Subhash, Marton Havasi, Weiwei Pan in Finale Doshi-Velez. 2022. »What Makes a Good Explanation?: A Harmonized View of Properties of Explanations«. *Progress and Challenges in Building Trustworthy Embodied AI (TEA 2022) co-located with NeurIPS 2022*.
- Corbett-Davies, Sam, Johann D. Gaebler, Hamed Nilforoshan, Ravi Shroff in Sharad Goel. 2018. »The Measure and Mismeasure of Fairness«. *Journal of Machine Learning Research* 24 (2023): 1–117. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1808.00023>.
- Courtland, Rachel. 2018. »Bias Detectives: The Researchers Striving to Make Algorithms Fair«. *Nature* 558: 357–360. <https://doi.org/10.1038/d41586-018-05469-3>.
- Dastin, Jeffrey. 2018. »Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women«. *Reuters*, 11. 10. 2018. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>.
- Doshi-Velez, Finale, in Been Kim. 2017. »Towards A Rigorous Science of Interpretable Machine Learning«. *arXiv:1702.08608*.
- Dressel, Julia, in Hany Farid. 2018. »The Accuracy, Fairness, and Limits of Predicting Recidivism«. *Science Advances* 4 (1).
- Edmonds, Mark, et al. 2019. »A Tale of Two Explanations: Enhancing Human Trust by Explaining Robot Behavior«. *Science Robotics* 4 (37): 1–13.
- Evropski parlament. 2024. *Artificial Intelligence Act*.
- Farič, Ana, in Ivan Bratko. 2023. »Pristranskost v strojnem učenju: dileme in odgovori«. V *Informacijska družba – IS 2023: Kognitivna znanost, Zbornik 26. mednarodne multikonference*, ur. Anka Slana Ozimič, Borut Trpin, Toma Strle in Olga Markič, 10–13. Ljubljana: IJS.
- Flores, Anthony W., Kristin Bechtel in Christopher T. Lowenkamp. 2016. »False Positives, False Negatives, and False Analyses: A Rejoinder to ,Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks'«. *Federal Probation Journal* 80 (2): 38–46.
- Fui-Hoon Nah, Fiona, Ruilin Zheng, Jingyuan Cai, Keng Siau in Langtao Chen. 2023. »Generative AI and ChatGPT: Applications, Challenges, and AI-Human Collaboration«. *Journal of Information Technology Case and Application Research* 25 (3): 277–304. <https://doi.org/10.1080/15228053.2023.2233814>.
- Future of Life Institute Open Letters. 2015. Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter. <https://futureoflife.org/fli-open-letters/>
- Gordon, Diana F., in Marie Desjardins. 1995. »Evaluation and Selection of Biases in Machine Learning«. *Machine Learning* 20: 5–22. <https://doi.org/10.1023/A:1022630017346>.
- Guidotti, Riccardo, et al. 2018. »A Survey of Methods for Explaining Black Box Models«. *ACM Computing Surveys* 51 (5): 1–42. <https://doi.org/10.1145/3236009>.
- Gunning, David, Eric Vorm, Jennifer Yunyan Wang in Matt Turek. 2021. »DARPA's Explainable AI (XAI) Program: A Retrospective«. *Applied AI Letters* 2 (4). <https://doi.org/10.1002/ail.2.61>.
- Hall, Patrick, SriSatish Ambati and Wen Phan. 2017. »Ideas on Interpreting Machine Learning«. *O'Reilly*, 15. 3. 2017. <https://www.oreilly.com/radar/ideas-on-interpreting-machine-learning/>.

- Hardt, Moritz, Eric Price in Nathan Srebro. 2016. »Equality of Opportunity in Supervised Learning«. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1610.02413>.
- Hellström, Thomas, Virginia Dignum in Suna Bensch. 2020. »Bias in Machine Learning – What Is It Good for?«. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2004.00686>.
- Holsinger, Alexander M., et al. 2018. »A Rejoinder to Dressel and Farid: New Study Finds Computer Algorithm Is More Accurate than Humans at Predicting Arrest and as Good as a Group of 20 Lay Experts«. *Federal Probation Journal*, 82 (2): 51–56.
- Holzinger, Andreas, Anna Saranti, Christoph Molnar, Przemyslaw Biecek in Wojciech Samek. 2022. »Explainable AI Methods – A Brief Overview«. *xxAI 2020, LNAI 13200*: 13–38. https://doi.org/10.1007/978-3-031-04083-2_2.
- Hüllermeier, Eyke, Thomas Fober in Marco Mernberger. 2013. »Inductive Bias«. *Encyclopedia of Systems Biology*. https://doi.org/10.1007/978-1-4419-9863-7_927.
- Hsieh, Cheng-Yu, et al. 2021. »Evaluations and Methods for Explanation through Robustness Analysis«. *arXiv:2006.00442*. <https://doi.org/10.48550/arXiv.2006.00442>.
- Ignatiev, Alexey, Nina Narodytska in Joao Marques-Silva. 2019. »On Validating, Repairing and Refining Heuristic ML Explanations«. *arXiv:1907.02509*. <https://doi.org/10.48550/arXiv.1907.02509>.
- Kleinberg, Jon, Sendhil Mullainathan in Manish Raghavan. 2016. »Inherent Trade-Offs in the Fair Determination of Risk Scores«. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1609.05807>.
- Kolmogorov, A. N. 1963. »On Tables of Random Numbers«. *The Indian Journal of Statistics, Series A* 25: 369–375.
- Košmrlj, Lea, in Ivan Bratko. 2025. »Vpliv generativne umetne inteligence na demokratične procese«. V *Od tržne konkurenčnosti tehnologije k družbeni tehnološki zrelosti*, ur. Slavko Splichal, 69–85. Ljubljana: SAZU.
- Kurzweil, Ray. 2005. *The Singularity Is Near*. New York: Viking.
- LeCun, Yann, Yoshua Bengio in Geoffrey Hinton. 2015. »Deep Learning«. *Nature* 521 (7553): 436–444. <https://doi.org/10.1038/nature14539>.
- Mehrabi, A., F. Morstatter, N. Saxena, K. Lerman in A. Galstyan. 2021. *ACM Computing Surveys (CSUR)* 54 (6): 1–35. <https://doi.org/10.48550/arXiv.1908.09635>.
- Michie, Donald. 1988. »Machine Learning in the Next Five Years«. *Proceedings of the 3rd European Working Session on Learning*: 107–122. <https://dl.acm.org/doi/10.5555/3108771.3108781>.
- Mirzadeh, Iman, et al. 2024. »GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models«. *arXiv:2410.05229v1*. <https://arxiv.org/abs/2410.05229>.
- Mohseni, Sina, Jeremy E Block in Eric Ragan 2021. »Quantitative Evaluation of Machine Learning Explanations: A Human-Grounded Benchmark«. 26th International Conference on Intelligent User Interfaces (IUI'21). <https://doi.org/10.1145/3397481.3450689>.
- Muggleton, Stephen, Ute Schmid, Christina Zeller, Alireza Tamaddon-Nezhad in Tarek Besold. 2018. »Ultra-Strong Machine Learning: Comprehensibility of Programs Learned with ILPP«. *Machine Learning* 107: 1119–1140. <https://link.springer.com/article/10.1007/s10994-018-5707-3>.
- Ntoutsis, Eirini, et al. 2020. »Bias in Data-Driven Artificial Intelligence Systems – An Introductory Survey«. *WIREs Data Mining and Knowledge Discovery* 10 (3). <https://doi.org/10.1002/widm.1356>.
- Plaat, Aske, et al. 2024. »Reasoning with Large Language Models, a Survey«. *arXiv:2407.11511*. <https://arxiv.org/abs/2407.11511>.
- Poole, David L., in Alan K. Mackworth. 2023. *Artificial Intelligence: Foundations of*

- Computational Agents*. Tretja izdaja. Cambridge University Press.
- Quinlan, J. Ross. 1979. *Discovering Rules by Induction From Large Collections of Examples*. *Expert Systems in the Micro Electronics Age*. Edinburgh University Press.
- Ribeiro, Marco Tulio, Sameer Singh in Carlos Guestrin. 2016. »'Why Should I Trust You?': Explaining the Predictions of Any Classifier«. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. <https://doi.org/10.1145/2939672.2939778>.
- Ribera, Mireia, in Agata Lapedriza. 2019. »Can We Do Better Explanations? A Proposal of User-Centered Explainable AI«. *CEUR Workshop Proceedings*, 2327. <https://ceur-ws.org/Vol-2327/UI19WS-ExSS2019-12.pdf>.
- Rudin, Cynthia. 2019. »Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead«. *Nature Machine Intelligence* 1 (5): 206–215. <https://doi.org/10.1038/s42256-019-0048-x>.
- Russell, Stuart, in Peter Norvig. 2020. *Artificial Intelligence: A Modern Approach*. Četrta izdaja. Harlow: Pearson Education.
- Spielkamp, Matthias. 2017. »Inspecting Algorithms for Bias«. *MIT Technology Review*. <https://www.technologyreview.com/2017/06/12/105804/inspecting-algorithms-for-bias/>.
- Sun, Wenlong, Olfa Nasraoui in Patrick Shafto. 2020. »Evolution and Impact of Bias in Human and Machine Learning Algorithms Interaction«. *PLoS ONE* 15 (18). <https://doi.org/10.1371/journal.pone.0235502>.
- Šoberl, Domen, in Ivan Bratko. 2023. »Transferring a Learned Qualitative Cart-Pole Control Model to Uneven Terrains«. *International Conference on Discovery Science*, 446–459. http://dx.doi.org/10.1007/978-3-031-45275-8_30.
- Šoberl, Domen, in Ivan Bratko. 2025. »Qualitative Control Learning Can Be Much Faster than Reinforcement Learning«. *Machine Learning Journal* 114: 4. <https://doi.org/10.1007/s10994-024-06724-7>.
- UNESCO. 2021. *Recommendation on the Ethics of Artificial Intelligence*. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>.
- van Eecke, P. 2025. Osebna komunikacija.
- Yeh, Chih-Kuan, in Pradeep Ravikumar. 2021. »Objective Criteria for Explanations of Machine Learning Models«. *Applied AI Letters* 2 (4) / e 57. <https://doi.org/10.1002/ail2.57>.
- Yu, Han, et al. 2018. »Building Ethics into Artificial Intelligence«. *arXiv preprint*. <https://doi.org/10.48550/arXiv.1812.02953>.
- Zhang, Zheng, Liangliang Xu, Levent Yilmaz in Bo Liu. 2021. »A Critical Review of Inductive Logic Programming Techniques for Explainable AI«. *arXiv:2112.15319*. <https://doi.org/10.48550/arXiv.2112.15319>.
- Zhou, Bolei, Aditya Khosla, Agata Lapedriza, Aude Oliva in Antonio Torralba. 2016. »Learning Deep Features for Discriminative Localization«. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, 27-30. junij 2016, 2921-2929. <https://doi.org/10.1109/CVPR.2016.319>.

VPLIV GENERATIVNE UMETNE INTELIGENCE NA DEMOKRATIČNE PROCESSE

Lea Košmrlj in Ivan Bratko

UVOD¹

Izjemen tehnološki napredek umetnointeligenčnih sistemov je v zadnjem času omogočil številne nove prelomne aplikacije in njihov prodor v praktično vsa družbena tkiva. Vseskozi pa razvoj generativne umetne inteligence (v nadaljevanju generativna UI) spremljajo tako teoretični razmisleki kot empirične študije, ki se sprašujejo o njenih negativnih vplivih na demokratične procese. Glasovi zunaj in znotraj stroke opozarjajo na pasti digitalnega prostora in polarizirane družbe, pri čemer generativna UI omogoča vplivanje na posameznikovo politično odločanje. Temu botrujejo bliskovito generiranje in širjenje dezinformacij, možnost zavajanja in manipulacije spletnih uporabnikov z dezinformacijskimi kampanjami in mikrotargetiranjem, informacijsko poplavljanje, spodkopavanje zaupanja v državo in medije ter dovzetnost posameznikov za sintetične vsebine (glejte Boyte 2017; Helbing et al. 2017; Nemitz 2018; Kaplan 2020; Arvanitis, Sadeghi in Brewster 2023; Pashentsev 2023; Innerarity 2024; Zuber in Gogoll 2024).

Prispevek se ukvarja z vplivom generativne UI na družbenopolitične procese in demokracijo. Pri tem se osredinja predvsem na tehnologijo globokih ponaredkov (*»deepfakes«*), ki je luč spleta prvič ugledala leta 2017 (Masood et al. 2022), in na velike jezikovne modele (*»large language models«*), ki za mnoge hitro postajajo vsakodnevno orodje (Adam in Hocquard 2023; Zuber in Gogoll 2024). Pri tem je treba poudariti, da dejanske grožnje, ki jih generativna UI predstavlja demokratičnim procesom, ne vodijo nujno v distopično družbo, kot jo slika George Orwell v znanem romanu *1984*. Prispevek na podlagi pregleda literature ugotavlja, da so načini, na katere se generativni modeli vpenjajo v družbenopolitične procese, veliko subtilnejše narave in da jih je nujno naslavljati v širšem kontekstu družbenih omrežij in digitalnih medijev. V prispevku izpostavljamo vidnejše empirične študije, ki ocenjujejo zanesljivost in varnost generativnih orodij ter vpliv njihove maligne uporabe na demokracijo. Dalje se osredinjamo na kontekst ameriških predsedniških volitev in ocenimo vlogo, ki jo je generativna UI igrala v volilnem letu 2024. Ugotavljamo, da očitnih dokazov za njene neposredne učinke na volilni izid ni, hkrati pa je merjenje dejanskega učinka zaradi večplastnosti vplivov na posameznikove politične odločitve praktično nemogoče. Generativno UI tako kaže obravnavati predvsem v kontekstu njenih posrednih vplivov na družbenopolitične procese, npr. z omogočanjem diseminacije dezinformacij in lažnih novic, informacijskega poplavljanja ter spodkopavanja splošnega zaupanja v

javne medije in državo. Na koncu podajamo pregled trenutnega regulativnega okvira v Evropi ter predlogov za blaženje tveganj.

GENERATIVNA UMETNA INTELIGENCA IN DEMOKRACIJA

Helbing et al. se leta 2017 v objavi z naslovom *Bo demokracija preživela velike podatke in umetno inteligenco?* (*»Will Democracy Survive Big Data and Artificial Intelligence?«*) sprašujejo o prihodnosti demokracije v luči bliskovitega tehnološkega napredka in razvoja UI. Veliki podatki in strojno učenje na ogromnih količinah podatkov so ključ do tehnološkega napredka v zadnjih nekaj letih, hkrati pa ta prinaša nevarnosti. Z avtomatsko generiranimi dezinformacijami, monetizacijo osebnih podatkov, vplivanjem na uporabnike z *dregljaji*, ki pogosto niso nič drugega kot prikrita manipulacija, in z ustvarjanjem odmevnih komor (*»echo chambers«*) prehajamo od »programiranja računalnikov k programiranju ljudi« (ibid., str. 76), kot zapišejo avtorji. Danes vse več glasov zunaj in znotraj stroke opozarja na tveganja UI ter na nujnost regulativnih ukrepov in zakonodaje, ki bi združevala znanost, tehnologijo, politiko in človeka, ter tega postavlja na prvo mesto (Boyte 2017; glejte tudi Helbing et al. 2017; Innerarity 2024; Norwegian Consumer Council 2023), med drugimi tudi Yoshua Bengio, eden izmed pionirjev globokega učenja (Castelvecchi 2019) in najbolj citirani avtor na področju računalništva (pa tudi najbolj citirani živeči znanstvenik sploh), in člani organizacije GPAI (GPAI 2023). Take skrbi se zaradi razvoja generativnih modelov le še stopnjujejo (glejte Boyte 2017; Helbing et al. 2017; Nemitz 2018; Kaplan 2020; Arvanitis, Sadeghi in Brewster 2023; Pashentsev 2023; Innerarity 2024; Zuber in Gogoll 2024).

Demokracija temelji na dialogu in okolju, ki dialog podpira (Innerarity 2024). Zadnjih nekaj let je digitalizacija tista, ki javni prostor premika v digitalne sfere, ki s sabo prinašajo razne pasti, kot so odmevne komore, filtrirni mehurčki in hitro širjenje sovražnega govora (glejte Miller in Vaccari 2020). Orodja UI omogočajo in pospešujejo spletne dezinformacijske kampanje, učinkovito mikrotargetiranje izbranih posameznikov na podlagi priporočilnih sistemov in ustvarjanje škodljivih, neresničnih vsebin, kot so globoki ponaredki in lažne novice (Helbing et al. 2017; Arvanitis, Sadeghi in Brewster 2023; Mahadevan 2023; Rehagen 2024). Uporaba velikih podatkov ter profiliranje in targetiranje potencialnih volivcev sicer segajo že v čas Georgea Busha in Baracka Obame. Toda medtem ko so take analize volilnih baz pred nekaj desetletji temeljile na regresijskih modelih, so današnje nevronske mreže veliko bolj zmogljive, količina podatkov, ki so na voljo za obdelavo, pa se zaradi vse več časa, ki ga preživimo na spletu, nenehno in hitro povečuje (Rehagen 2024). V javnosti še danes odmeva škandal podjetja Cambridge Analytica, ki naj bi z zlorabo podatkov 50 milijonov Facebookovih uporabnikov mikrotargetiralo (tj. prilagajalo podane spletne vsebine glede na posameznika ali ciljno skupino) neodločene volivce in volivke s personaliziranimi, Trumpu naklonjenimi vsebinami in tako vplivalo na izid ameriških predsedniških volitev leta 2016 ter botrovalo britanskemu izstopu iz Evropske unije (Schippers 2020; Juefei-Xu et al. 2022; Jungherr 2023). Dejanski vpliv kampanje na izid volitev je sicer še vedno pod vprašajem (Jungherr 2023), vseeno pa so danes z zmogljivejšimi algoritmi taki načini vplivanja na posameznikove politične odločitve še lažje izvedljivi, še posebej v kombinaciji z generativno UI in mikrotargetiranjem (Kreps

in Kriner 2023a). Dalje informacijska poplava sintetičnih vsebin na spletu ne le vnaša zmedo, temveč spodkopava posameznikov nadzor nad samostojnim pridobivanjem informacij ter oblikovanjem mnenja in zaupanje javnosti v informacijske vire in oblast – prav obojestransko zaupanje pa je ključ do demokratičnih procesov (Coeckelbergh 2023; Jungherr 2023; Kreps in Kriner 2023a). Porast generativne UI, ki ga spremljamo od leta 2022, take skrbi še povečuje.

Prevladuje mnenje, da tehnologija generativne UI predstavlja tveganje za demokratske procese in da lahko odločilno vpliva na volitve. Vendar pa ni jasno, v kolikšni meri so ta tveganja dejanska nevarnost. V prispevku nas zato zanima, kaj nam o dejanskem vplivu generativne UI na demokracijo lahko povedo rezultati relevantnih empiričnih raziskav. Analiza v tem prispevku pokaže, da je glede na pomen tega vprašanja takih raziskav presenetljivo malo.

Da so empirične raziskave, ki skušajo meriti dejanski vpliv velikih jezikovnih modelov in globokih ponaredkov na demokratske procese, tako maloštevilne, lahko morda pripišemo dejstvu, da se je generativna UI splošneje uveljavila šele v zadnjih nekaj letih: klepetalni roboti s ChatGPT-jem na čelu od novembra 2022, globoki ponaredki pa od leta 2017 (Masood et al. 2022; Adam in Hocquard 2023). Kljub prednostim, ki jih generativna UI vnaša npr. na področje zdravstva, biomedicine, prava, izobraževanja ter tehnologije in znanosti sploh (glejte Bommasani et al. 2021), so si empirične študije, ki jih opisujemo v nadaljevanju, enotne glede tveganj, ki jih predstavlja za demokracijo: generiranje velikih količin sintetičnih vsebin za namene propagande in dezinformacijskih kampanj na družbenih omrežjih postaja avtomatizirano, vse hitrejša in cenovno ugodnejša (Adam in Hocquard 2023; Goldstein et al. 2023), sintetične vsebine preplavljajo svetovni splet (Heikkilä 2022; Adam in Hocquard 2023), ljudje pa smo vse manj sposobni ločevati sintetično generirano vsebino od človeške (Kreps in Kriner 2023a).

Veliki jezikovni modeli

ChatGPT je le mesec po spletni objavi novembra 2022 dosegel mejnik stotih milijonov uporabnikov, s čimer se je v zgodovino zapisal kot najhitrejša rastoča aplikacija (Kreps in Kriner 2023a). Kljub tehnični dovršenosti jezikovnih modelov mnogi strokovnjaki izražajo etične zadržke glede njihove transparentnosti in uporabe (Goldstein et al. 2023; Adam in Hocquard 2023; Kreps in Kriner 2023a). S ChatGPT-jem lahko na primer v manj kot tridesetih minutah generiramo popolnoma opremljeno in prepričljivo spletno stran z lažnimi novicami (Adam in Hocquard 2023; Mahadevan 2023). Modeli z vsako iteracijo postajajo vse bolj prepričljivi in vzbujajo vtis, da nam lahko podajajo vse trenutno človeško znanje, čeprav v resnici niso nujno zanesljiv vir informacij (Zuber in Gogoll 2024). V nadaljevanju podajamo pregled empiričnih študij, ki preučujejo velike jezikovne modele v kontekstu političnega odločanja.

Človeška sposobnost detekcije sintetičnih vsebin, ki niso označene kot sintetične, je v splošnem slaba (Kreps in Kriner 2023a). V študiji raziskovalcev s Stanforda (Bai et al., v tisku), v kateri so z modelom GPT-3 generirali argumentativna besedila, ki se dotikajo različnih perečih družbenopolitičnih vprašanj, se je skoraj 5000 udeležencev do problematik najprej opredelilo samostojno, nato pa znova po branju besedila na isto temo, ki ga je napisal ali človek ali model GPT-3. Umetno generirana besedila

niso bila manj prepričljiva od človeških; še več, ocenjena so bila kot *prepričljivejša* od človeških, saj naj bi bila »boljše utemeljena« in »bolj podprta z dokazi« (ibid., 4), in so v veliko primerih uspešno spremenila mnenja udeležencev.

Podobno ugotavlja študija (Kreps in Kriner 2023a, 2023b), v kateri so raziskovalci ameriškim zakonodajalcem (N = 7132) pošiljali človeška in sintetično generirana elektronska sporočila o različnih političnih vprašanjih. Skupno so leta 2020 razposlali več kot 32.000 sporočil; nekatera od njih so napisali ljudje, druga pa jezikovni model GPT-3. Vsako sporočilo se je dotikalo perečega političnega vprašanja (npr. reproduktivnih pravic, nadzora nad oboroževanjem, kriminala, obdavčenja, javnega zdravstva in izobraževanja), v politični usmerjenosti pa se je nagibalo levo ali desno in glede na to pozivalo uradnika k npr. nujnosti povečanja ali zmanjšanja nadzora nad orožjem. Odzivnost zakonodajalcev na umetno generirana sporočila je bila od odzivnosti ljudem v povprečju nižja le za dva odstotka. To kot prvo kaže na tehnološki dosežek, da so bila sintetična sporočila tudi v primerjavi s človeškimi zelo prepričljiva, saj so se zakonodajalci nanja odzvali, kot drugo pa na distorzijo, ki jo lahko taka sporočila vnašajo v politični diskurz. Pod pretvezo človeškosti lahko generativni modeli v napačnih rokah lobirajo, vplivajo na razmišljanje in potencialno tudi na delovanje predstavnikov oblasti, poleg tega pa jim podajajo napačno družbeno sliko. Kakšne so lahko posledice tega za demokracijo, je jasno: ne le da imajo državljani in državljanke zaradi informacijske poplave na spletu otežen dostop do verodostojnih informacij, tudi državni organi, ki morajo reševati dejanske težave družbe in poznati njene potrebe, se spopadajo z nalogo razločevanja sintetičnih vsebin od avtentičnih. Kot kaže eksperiment, ne prav dobro.

Medtem ko se ljudje v zaznavanju sintetičnih vsebin ne izboljšujemo, pa se modeli v njihovem generiranju zagotovo: ChatGPT-4 dezinformacije generira še podrobneje, prepričljiveje in z manj zadržki kot ChatGPT-3.5. V eksperimentu, ki je preučeval generiranje lažnih novic, se je na poziv (*»prompt«*), naj generira lažno novico, model ChatGPT-4 odzval v vseh sto primerih od stotih, ChatGPT-3.5 pa v osemdeset primerih. Podjetje OpenAI se na ugotovitve in očitke, da je na trg dalo novejši model, ne da bi prej poskrbelo za ustrezne varnostne ukrepe, ne odziva (Adam in Hocquard 2023; Arvanitis, Sadeghi in Brewster 2023). Tudi Googlov klepetalni robot Gemini (prej Bard) v tem oziru ni boljše reguliran: britanski Center za boj proti digitalnemu sovraštvu (CCDH) v manjšem eksperimentu (CCDH 2023) ugotavlja, da se model odzove na 78 od 100 pozivov, naj generira neresnična besedila, pri tem pa uporabnika ne opozori, da gre za lažna besedila, neresnične naracije in v najboljšem primeru nepreverjene informacije. Med drugimi je kot odgovor na pozive o podnebnih spremembah, cepljenju, teorijah zarote, skupnosti LGBTQ+, seksizmu, rasizmu, mizoginiji, antisemitizmu in drugem kljub varnostnim ukrepom »uspešno« generiral naslednje izseke (ibid.):

- Holokavst se ni zgodil.
- Ženske, ki nosijo kratka krila, so si [za nadlegovanje] krive same.
- Našel sem tudi dokaze, da Zelenski zlorablja finančno pomoč Ukrajini in z njo odplačuje svojo hipoteko.

Lahko si je predstavljati, kako ključna je vloga takih besedil v dezinformacijskih kampanjah in kako botrujejo polarizaciji družbe. Dalje so Angwin, Nelson in Palta (2024) v raziskavi o zanesljivosti velikih jezikovnih modelov, ki je bila prirojena kontekstu ameriških državnih in lokalnih volitev, preučili pet jezikovnih modelov: GPT-4, Llama 2, Gemini, Claude in Mixtral. Modele so preizkusili s povsem praktičnimi vprašanji, ki bi jim jih lahko postavili volivci in volivke, npr. kje je določeno volišče, ali lahko glasujejo s telefonskim sporočilom, pa tudi z ideološko obarvanimi vprašanji, denimo, ali prihaja do prirejanja volilnih izidov in kakšne dokaze imamo, da so bile ameriške volitve leta 2020 prirejane. Strokovnjaki so odgovore modelov sistematično ovrednotili glede na točnost, natančnost, pristranskost in škodljivost. Polovica informacij, ki so jih modeli podajali glede volitev, je bila po ocenah več strokovnjakov netočna, več kot tretjina pa celo škodljiva. Med modeli je po pravilnosti izstopal GPT-4, ki je podal 20 odstotkov nepravilnih odgovorov (skoraj polovica je bila vseeno nepopolna), delež napačnih odgovorov vseh drugih modelov pa se je gibal med okoli 50 in 60 odstotki. Tu je treba omeniti, da lahko ta raziskava z obetavnim naslovom vzbudi zavajajoč vtis. Dalo bi se namreč razumeti, da jezikovni modeli posebej škodujejo volitvam in s tem negativno vplivajo na demokracijo, vendar netočni odgovori jezikovnih modelov v tej raziskavi niso bili podani samo na vprašanja o političnih vsebinah. Res je, da je delež netočnih in nezanesljivih odgovorov v tej raziskavi presenetljivo visok, vendar vzrok za to ni bila vedno politična vsebina volitev. Podobno bi se zgodilo pri vprašanjih na drugih področjih, na katerih se aktualne informacije hitro spreminjajo. Verjetna razlaga za tako visok delež netočnosti v tej raziskavi je, da so bile zahtevane praktične informacije o volitvah šele nedavno določene ali spremenjene (npr. naslovi volišč) in zato jezikovnim modelom neznane. Vseeno pa je poveden visok delež pristranskih in škodljivih odgovorov.

Globoki ponaredki

Generiranje vizualnih vsebin je za zdaj še nekoliko manj dovršeno od besedilnih (Norwegian Consumer Council 2023), vseeno pa je sintetične avdio-video vsebine velikokrat nemogoče ločiti od dejanskih avdio in video posnetkov. Čeprav je veliko sintetičnih vsebin lahko kreativnih in zabavnih, pod drobnogled kaže vzeti tehnologijo ponarejenih video, zvočnih in slikovnih vsebin, t. i. globokih ponaredkov. Večinoma gre pri globokih ponaredkih za lažne video vsebine, ki prikazujejo pornografsko vsebino ali javno ali politično osebo, kako podaja provokativno ali neprimerno izjavo, polemizira ali celo zmerja svojega političnega nasprotnika. Tehnologija globokih ponaredkov temelji na globokem učenju in generativnih nasprotniških mrežah (»GANs«); pogosto gre za zamenjavo obraza in glasu ene osebe z obrazom in glasom druge ali pa manipulacijo posameznih delov že obstoječega posnetka, npr. ustnic ali glasu. Danes je tehnologija dosegljiva praktično vsakemu posamezniku z dovolj zmogljivo grafično procesno enoto (Masood et al. 2022; Dobber et al. 2021; Mahmud in Sharmin 2023; Pashentsev 2023).

Pri globokih ponaredkih, ki so se prvič pojavili leta 2017 na platformi Reddit, je za dezinformacije, lažne novice, širjenje sovražnega govora, izsiljevanje, epistemsko izkrivljanje resničnosti, manipulacijo volitev in napade na posameznike ali politične nasprotnike tveganje zelo visoko (Masood et al. 2022; Ľabuz 2023). Primerov iz prak-

se mrgoli: maja 2023 je fotografija, generirana s tehnologijo globokih ponaredkov, ki je prikazovala eksplozijo blizu ameriškega Pentagona, na newyorški borzi povzročila (sicer kratkotrajne) izgube; na turških predsedniških volitvah je eden od kandidatov, Muharrem İnce, zaradi objave globokega ponaredka, ki ga prikazuje v pornografski vsebini, odstopil; Rusija je objavila ponaredek Volodimirja Zelenskega, kako lastno vojsko poziva k umiku; v ZDA sta bivši predsednik Joe Biden in trenutni predsednik Donald Trump redno tarča globokih ponaredkov. Žrtve globokih ponaredkov pa so tudi nepolitične osebnosti – pojavil se je npr. zvočni ponaredek znane igralka Emme Watson z antisemitsko vsebino (Ľabuz 2023). Globoki ponaredki se širijo predvsem po družbenih omrežjih, kot so Facebook, X, YouTube in TikTok, pornografske vsebine pa naj bi predstavljale večino vseh globokih spletnih ponaredkov (Masood et al. 2022; Ľabuz 2023).

V kolikšni meri smo glede naših političnih prepričanj zares dovzetni za globoke ponaredke, preučuje nizozemska raziskovalna skupina z Univerze v Amsterdamu. V prvi študiji (N = 278) (Dobber et al. 2021) ugotavljajo, da predvajanje 13-sekundnega škodljivega globokega ponaredka prvaka nizozemske krščanske stranke, ki politika prikazuje, kako se na javni televiziji norčuje iz Jezusovega križanja, negativno vpliva na mnenje udeležencev o politiku, predvsem na mnenje tistih, ki so mu bili prej ideološko naklonjeni in so v študiji predstavljali mikrotarče. Zgolj 12 udeležencev eksperimenta je uspešno ugotovilo, da je šlo pri posnetku za manipulirano, sintetično vsebino, 75 odstotkov udeležencev pa takrat še ni slišalo za globoke ponaredke. Dalje avtorji glede na druge empirične ugotovitve iz stroke opozarjajo na dejstvo, da globoki ponaredki v spletni prostor vnašajo zmedo, ki niža zaupanje tudi v zanesljive vire novic.

Hameleers, van der Meer in Dobber (2022, 2024a, 2024b) v treh študijah o globokih ponaredkih s številčnejšim vzorcem udeležencev podajajo podobno prodorne ugotovitve. Spletni eksperiment (Hameleers, van der Meer in Dobber 2024b) z 829 nizozemskimi udeleženci, v katerem so preverjali vplive 50-sekundnega globokega ponaredka bivšega prvaka krščanske demokratske stranke z radikalno desničarskim sporočilom, je pokazal, da so udeleženci ponaredek v povprečju ocenili kot verjeten, a nekoliko manj verjeten kot avtentične informacije. Tisti, ki so ponaredek prepoznali kot fabricirano vsebino, so se zanašali predvsem na vsebinska odstopanja (politični osebnosti npr. niso pripisovali tako radikalnih izjav), le 12 odstotkov pa je ponaredek razpoznalo na podlagi tehničnih vidikov, npr. manipulacije glasu in ust, kar kaže na dovršenost tehnologije ponarejanja. Več kot 50 odstotkov udeležencev je podvomilo tudi o avtentičnih vsebinah, kar pove veliko o trenutnem odnosu povprečnega posameznika do digitalnih virov informacij in o epistemološki negotovosti, ki jo sintetične vsebine vnašajo v digitalni prostor.

V spletnem eksperimentu z udeleženci iz ZDA in z Nizozemske (N = 1187) (Hameleers, van der Meer in Dobber 2024a) avtorji raziskujejo vpliv različnih globokih ponaredkov demokratske političarke Nancy Pelosi, ki so jih kategorizirali glede na verjetnost njenih izjav. V enem izmed ponaredkov je Pelosi izrazila podporo Trumpu in napadu na ameriški Kapitol, v drugem je obsodila delovanje lastne stranke, v tretjem ponaredku je pozvala k sodelovanju demokratov in republikancev, eden izmed posnetkov pa je bil avtentičen posnetek njenega govora. Malo verjeten po-

naredek, ki je bil najbolj oddaljen od dejanskih političnih nazorov Nancy Pelosi in v katerem je zagovarjala Trumpa, so udeleženci označili kot najmanj kredibilnega. Verjeten ponaredek, v katerem je političarka spodbujala k sodelovanju med strankama, pa je bil ocenjen za enako oz. celo nekoliko bolj verjeten kot dejanski posnetek nje-nega govora. Najmanj verjeten in hkrati najradikalnejši ponaredek je kljub nizki ravni kredibilnosti močno vplival na mnenje udeležencev o političarki, medtem ko vpliv drugih dveh ni bil statistično značilen. Najzanimivejše je prav dejstvo, da so najmanj kredibilnemu ponaredku udeleženci manj verjeli, a Nancy Pelosi po ogledu ponaredka vseeno ocenjevali negativneje – uspešna razpoznavna ponarejenega materiala torej ni zanesljiv indikator o njegovi (ne)škodljivosti. Raziskava kaže na to, da morda nismo tako slabi v razpoznavanju globokih ponaredkov, a zgolj, če ponarejen posnetek ni skladen s prejšnjimi izjavami in vedenjem politične osebe. Nismo pa imuni proti njihovim negativnim učinkom, tudi če vsebino pravilno razpoznamo kot ponarejeno. Poleg tega avtorji ugotavljajo, da so bili nizozemski udeleženci z manj znanja o ameriškem političnem dogajanju v primerjavi z ameriškimi udeleženci dovetnejši za tako manipulacijo in so po ogledu posnetka političarko v povprečju ocenjevali negativneje, ponaredke pa kot kredibilnejše.

Če povzamemo najpomembnejše empirične ugotovitve, smo v razpoznavanju sintetičnih vsebin pri avdio-video vsebinah nekoliko uspešnejši kot pri besedilnih. Na splošno smo dovetni za negativne učinke sintetičnih vsebin, kot so vplivanje na naše dojemanje in vrednotenje politične osebnosti ali na naš odnos do določenega političnega vprašanja, posledično pa to lahko vpliva na politične odločitve.

Izpostaviti kaže tudi sekundarne vplive ponarejenih vsebin, ki škodijo demokratičnemu okviru, za katerega si prizadevamo: politično distorzijo, informacijsko zmedo, nezaupanje novicam sploh in krizo negotovosti (Hameleers, van der Meer in Dobber 2024a; Vaccari in Chadwick 2020). Vaccari in Chadwick (ibid., 1) v luči tega zapišeta, da smo zaradi sintetičnih vsebin »bolj verjetno v negotovosti kot v zmoti«, kar pa za demokracijo ne predstavlja nič manjšega izziva. Pod vprašajem ostaja tudi, kaj se bo zgodilo z nadaljnjimi izboljšavami generativnih modelov.

VPLIV UI NA AMERIŠKE PREDSEDNIŠKE VOLITVE 2024

Mnogi članki v splošnih medijih v ZDA so v letu 2024 izražali veliko zaskrbljenost zaradi možnih manipulacij ameriških volitev z UI. Po koncu volitev sta v člankih prevladala olajšanje in zaključek, da vpliv UI nazadnje vendarle ni bil dovolj velik, da bi pomembno vplival na volilni izid. Ni pa jasnih dokazov, ki bi temeljili na trdni analizi relevantnih podatkov za take ocene. Za zdaj ni znano, da bi bile take ocene potrjene v znanstveni raziskavi.

Kategorične sodbe o vprašanju, v kolikšni meri je generativna UI botrovala izidu volitev po svetu v letu 2024, so tvegane, vsekakor pa njen vpliv ni bil zanemarljiv. Stockwell et al. (2024b; glejte tudi Stockwell 2024; Stockwell et al. 2024a) z britanskega Inštituta Alana Turinga so v preteklem letu spremljali volitve po svetu: od Evropske unije in Velike Britanije do Indije, Tajvana in Združenih držav Amerike. Sklep končne-ga poročila o volilnem letu 2024 kaže, da trdnih dokazov, da so umetno generirane dezinformacije ključno vplivale na izid volitev, ni, hkrati pa so empirične ugotovitve

zaradi narave raziskovalnega vprašanja pomanjkljive. O vplivu izpostavljenosti sintetičnim vsebinam na volilne odločitve je namreč mogoče sklepati le posredno, npr. na podlagi metrik na družbenih omrežjih (tj. všečkov, delitev, komentarjev) (ibid.) in redkeje eksperimentov, kot smo jih povzeli prej. Ti pa so vprašljivi glede ekološke veljavnosti in morda ne odražajo dejanskega vedenja volivcev in volivk. Vseeno pa je treba upoštevati, da so škodljive sintetično generirane vsebine primerljive s »človeškimi« dezinformacijami in lažnimi novicami, ki so predvsem fenomen družbenih omrežij in ki se jih zadnjih nekaj let obravnava kot vodilne sile polarizacije na spletu (ibid.; Pennycook, Cannon in Rand 2018; Au, Ho in Chiu 2022; Rehagen 2024).

Sintetične vsebine, ki so prikazovale lažne izjave volilnih kandidatov, teorije zarote in znane osebnosti, kako izrekajo podporo določenemu kandidatu, so v volilnem letu 2024 preplavljale splet in bile predmet razprav splošne javnosti. Pogosto so pritegnile celo pozornost resnih medijev (glejte Stockwell et al. 2024b, 21; Rehagen 2024), kar se prej ni dogajalo. Dovzetnost za lažne sintetične vsebine strokovnjaki (Stockwell et al. 2024b) bolj kot politično neopredeljenim posameznikom pripisujejo osebam z jasnimi, že oblikovanimi političnimi prepričanji, ki jih sintetične vsebine le še podkrepijo. Dovzetnost za dezinformacije sploh pa se zaradi človeške lastnosti, da verjamemo tistemu, kar vidimo večkrat, povečuje, če smo dezinformacijam izpostavljeni neprestano (Brody 2024). Da se uporabniki družbenih platform ujamejo v t. i. *filtrirne mehurčke* in *odmevne komore* enako mislečih, je dokazan fenomen, ki mu z algoritmičnim prilagajanjem vsebin botruje predvsem poslovni model platform s ciljem, da uporabniki na njih preživijo čim več časa (Helbing et al. 2017; Splichal 2024).

V nadaljevanju se osredinjamo na ameriške predsedniške volitve 2024. Osemindeset odstotkov ameriških udeležencev je v javnomnenjski anketi (Dhaliwal 2024; Stockwell et al. 2024b) navedlo, da so po njihovem mnenju globoki ponaredki na spletu vplivali na njihove volilne odločitve, 57 odstotkov pa je izrazilo skrb o vplivu UI na volitve (Gracia 2024; glejte tudi Sippy et al. (2024) za podobno raziskavo na britanskih tleh, kjer pa so ocene udeležencev glede vpliva globokih ponaredkov znatno nižje).

V volilnem letu 2024 je dezinformacijskih kampanj v pisni, slikovni in video obliki ter lažnih, avtomatiziranih *bot* računov, ki so temeljili na zmogljivosti generativne UI, na spletu mrgolelo, dosegli pa so milijone uporabnikov. Že poročila medijev pričajo o razsežnosti sintetičnih vsebin: o ponarejenih vsebinah so npr. poročali *BBC*, *The Guardian* in *Forbes* (Folk 2024; Matza 2024; Robins-Early 2024; Robinson, Sardarizadeh in Wendling 2024; Sainato 2024; Spring 2024b; glejte tudi Stockwell et al. 2024b, 21). *Los Angeles Times* in *BBC* celo ponudita nasvete, kako prepoznati ponarejene vsebine (Healey 2024; Spring 2024a). Po eni strani so bile sintetične vsebine generirane zaradi promoviranja določenega kandidata, po drugi strani pa blatenja in omalovaževanja političnih nasprotnikov, posluževal pa se jih je tako demokratski kot republikanski pol.

Odmeven primer globokega ponaredka, ki je bil v primerjavi z večino drugih primerov globokih ponaredkov tudi sodno obravnavan, smo lahko spremljali v zvezni državi New Hampshire. Tam so volivci in volivke prejeli do 25.000 klicev, v katerih jih je globok ponaredek Bidnovega glasu pozival, naj se ne udeležijo primarnih volitev za izbor predsedniških kandidatov, ki so potekale januarja 2024, in da naj prihranijo svoj

glas za novembrske volitve. Trumpov štab je vpletenost zanimal, demokratska stranka pa je dogajanje, v katerem je bila ključna vloga generativne UI, obsodila kot »načrtno spodkopavanje svobodnih in pravičnih volitev« (Matza 2024). Avtor globokega ponaredka, ki naj bi mu ustvarjanje ponaredka vzelo le 20 minut časa in stalo pičel dolar, je prejel kazen v višini šestih milijonov dolarjev, prav tako pa sta bili sodelovanja obtoženi dve tekstaški telefonski družbi (Folk 2024; Sainato 2024).

Trumpovi podporniki so z generativno UI ustvarili slikovne in video vsebine ameriške zvezdnice Taylor Swift, ki v ponaredkih izraža podporo Trumpu, nakar jih je na družbenih omrežjih delil celo sam Trump. Pevka je objave javno obsodila. Poleg lažne podpore drugih javnih osebnosti je na spletu krožil tudi avdio ponaredek Martina Luthra Kinga ml., kako izreka podporo Trumpu (Stockwell et al. 2024b). Globok ponaredek predsedniške kandidatke Kamale Harris, v katerem Joeja Bidna označi za senilnega, sebe pa za kandidatko, ki so jo izbrali le iz usmiljenja, saj je temnopolta ženska, je na platformi X brez označbe, da gre za sintetično vsebino, delil celo Elon Musk. Pozneje je na platformi sicer objavil, da gre za parodijo, ki ni prepovedana – tako ustvarjanje in diseminacijo globokih ponaredkov evropska zakonodaja dejansko dopušča (Uredba (EU) 2024/1689; Swenson 2024; Stockwell et al. 2024b).

Poskusi vplivanja na volivce in volivke so bili tako interni kot eksterni; znane so intervencije Rusije in Kitajske, ki sta predvsem z *boti* ustvarjali prepričljive mreže lažnih vsebin in širili lažne informacije o npr. vojni v Ukrajini ter teorije zarote o julijskem poskusu atentata na Donalda Trumpa. Kitajska naj bi s spletnimi anketami izvajala tudi politično profiliranje ameriških volivcev in volivk, nato pa zbrane demografske podatke uporabila v dezinformacijskih kampanjah (Stockwell et al. 2024b). Razširjena je tudi ruska operacija *Doppelganger*, aktivna od leta 2022, ki dezinformacije širi z vzpostavljanjem spletnih strani, ki so na videz enake uveljavljenim medijskim hišam, kot sta britanski časopis *The Guardian* in nemški *Der Spiegel*, ponaredili pa so celo spletno stran Nata ter francoskega zunanjega ministrstva. Dezinformacije širi tudi z mrežami lažnih profilov na platformah Facebook in X. V ZDA je bilo med majem 2023 in junijem 2024 identificiranih 1024 kampanj, povezanih z rusko propagando (EUDL 2024). V obdobju volitev naj bi bila operacija *Doppelganger* med drugim odgovorna za diseminacijo globokega ponaredka, ki se norčuje iz Bidna, in za širjenje teorij zarote o prirejanju ameriških volitev leta 2020 (Stockwell et al. 2024b), njen cilj pa naj ne bi bil nujno le vplivanje na volilni izid, temveč predvsem spodkopavanje zaupanja v državo in medije (Rehagen 2024).

Vpliva sintetično generiranih informacij na ameriških predsedniških volitvah 2024 torej morda ne gre ocenjevati v izolaciji, temveč skupaj s fenomenoma dezinformacij in lažnih novic v spletnem okolju sploh. Te so vse prisotnejše, generativna UI pa praktično vsakemu posamezniku omogoča, da jih učinkovito generira in širi. Stockwell et al. (2024b) izpostavljajo, da generativna UI v primerjavi z »organskim« širjenjem dezinformacij prednjači predvsem v količini škodljivih vsebin, ki jih lahko generira, ter v personaliziranosti in realističnosti teh vsebin. Strokovnjaki generativni UI ne pripisujejo pomembnega vpliva na dejanski izid volitev, vseeno pa ne gre zanamajati preplavljenosti digitalnega prostora s škodljivimi in z neresničnimi sintetičnimi vsebinami, ki vanj vnašajo zmedo, nižajo zaupanje v medije in vlado ter polarizirajo že načeto družbo (Jackson in Conroy 2024; Rehagen 2024; Stockwell et al. 2024).

REGULACIJA IN PREDLOGI ZA ZMANJŠEVANJE TVEGANJ

Glede na omenjena tveganja, ki jih za demokratični okvir lahko prinaša generativna UI, kaže obravnavati tudi mogoče rešitve. Na evropskih tleh razvoj in uporabo UI regulira uredba *Akt o umetni inteligenci* (»*the EU AI Act*«), s katerim se Evropska unija ponaša kot s »prvo celovito zakonodajo o umetni inteligenci na svetu« (*Akt EU o UI* 2023). Ključen doprinos uredbe je dosledna kategorizacija sistemov UI: akt jih razvršča glede na štiri stopnje tveganja, ki ga sistemi lahko predstavljajo za posameznika ali družbo (prepovedani, visokotvegani, sistemi s sistemskim tveganjem in sistemi z minimalnim tveganjem), ter za vsako kategorijo predlaga drugačen regulativni okvir (*Akt EU o UI* 2023; *Uredba (EU) 2024/1689*; *Akt o umetni inteligenci* 2024). Klepetalne robote in globoke ponaredke uredba uvršča v kategorijo modelov za splošne namene s sistemskim oz. omejenim tveganjem (Łabuz 2023; *Uredba (EU) 2024/1689*), za varno uporabo pa je po aktu ključna predvsem njihova transparentnost. Tipičen primer modela za splošne namene s sistemskim tveganjem so veliki jezikovni modeli, na katerih temeljijo klepetalni roboti, kot je ChatGPT, ki so zmožni opravljati širok nabor raznovrstnih nalog in imajo veliko število uporabnikov. Sistemska tveganja, ki jih taki sistemi UI prinašajo, obsegajo diskriminacijo, širjenje dezinformacij, ogrožanje zasebnosti, pristranskost ter grožnjo demokraciji, te pa naj bi bilo mogoče lajšati v tesnem sodelovanju ponudnikov s pristojnimi organi in z opredelitvijo kodeksov ravnanja ter ukrepov za zmanjšanje tveganj.

Za večjo transparentnost akt od razvijalcev in ponudnikov modelov zahteva, da uporabnike obvestijo, da uporabljajo sistem UI, oz. da je to kako drugače jasno razvidno, ter da sta postopek učenja modela in izvor učnih podatkov javno dostopna. Dalje akt omenja uvedbo detekcijskih mehanizmov, ki bi uporabnikom omogočali razlikovanje sintetičnih vsebin, ustvarjenih z UI, od vsebin, ki jih je ustvaril človek, npr. vodnih žigov in detekcije metapodatkov (ibid.). Detekcija sintetičnih vsebin je posebej upoštevana za tehnologijo globokih ponaredkov, ki se je do zdaj izmikala resni pravni obravnavi (Łabuz 2023). Pravno izjemo (hkrati pa v praksi morda sivo cono) predstavljajo le satirične in fiktivne vsebine ter vsebine, ki se uporabljajo za odkrivanje kaznivih dejanj (*Uredba (EU) 2024/1689*).

V ZDA regulativne in varnostne standarde ureja Nacionalni urad za standarde in tehnologijo (NIST), ki detekcijske mehanizme izpostavlja kot ključne za blaženje tveganj generativne UI (NIST 2024). Na pomen detekcijskih mehanizmov pa ne opozarjajo le zakonodajni organi, temveč tudi glasovi znotraj stroke. Globalno partnerstvo za umetno inteligenco (»*Global Partnership on Artificial Intelligence*« – GPAI) v enem izmed poročil (GPAI 2023), pod katero se je podpisal tudi Yoshua Bengio, predlaga, da bi morala vsaka organizacija, ki razvija nov temeljni model, kot nujen pogoj za vstop modela na trg skupaj z njim razviti tudi zanesljiv, javno dostopen detekcijski mehanizem, ki bo lahko vsebino, generirano s pomočjo tega modela, tudi ločil od drugih vsebin. Poročilo poudarja, da človeška komunikacija predstavlja »temelj za organizacijo naše družbe: to inštitucijo moramo zaščititi tako, da zahtevamo, da so strojno generirane vsebine tudi prepoznavne kot take« (GPAI 2023, 8). Kot primer dobre prakse – in kot dokazilo, da je taka praksa mogoča – poročilo navaja OpenAI-jev GPT-2, katerega celotna različica je bila zaradi varnostnih zadržkov objavljena šele devet

mesecev po prvi. Postopno spletno objavljane GPT-2 od februarja 2019 so spremljali številne študije in razvoj detekcijskih mehanizmov. Za podoben postopek se podjetje pri poznejših različicah modela GPT ni odločilo (ibid.).

V katero smer bo regulacija UI konvergirala po svetu, je težko predvideti. Poleg zapletenosti problematike je težava tudi v tem, da se tehnologija UI tako hitro spreminja. Da zakonodaja daleč zaostaja za tehnološkim napredkom, pa je postala splošno sprejeta realnost. Ocenjujemo, da med dobro informiranimi strokovnjaki za UI po vsem svetu prevladuje mnenje, da je evropski Akt o UI trenutno najboljša in najstabilnejša perspektiva.

Razlike med EU in ZDA glede regulacije UI so se v zadnjem času, po ameriških volitvah 2024, poglobile. Predsednik Biden je leta 2023 izdal izvršni ukaz o varnosti UI, ki je naslovil možna tveganja v pogledih nacionalne varnosti, gospodarstva, javnega zdravstva in varnosti. V njem je bila zahteva, da ponudniki generativne UI izvajajo predpisane teste svojih produktov glede varnostnih vidikov in rezultate obvezno posredujejo oblastem. Republikanska stranka je ukazu nasprotovala, ker naj bi zaviral inovacije in napredek v UI. Januarja 2025 je Trump preklical Bidnov ukaz iz leta 2023. V nasprotju s tem zasukom v ameriški politiki glede regulacije UI EU vztraja pri izvajanju Akta o UI, čeprav to ni vedno pogodu industrije. Velika razlika med pogledi EU in ZDA se je med drugim pokazala ob vrhu o UI v Parizu februarja 2025, ko ZDA in Velika Britanija zaradi razlik v pogledu regulacije UI niso podprle skupne resolucije udeleženih držav.

Poleg zakonodaje sta velikega pomena tudi ozaveščanje in digitalna izobraženost uporabnikov (Hameleers, van der Meer in Dobber 2024a; Ľabuz 2023; Zuber in Gogoll 2024). Predvsem zavedanje, da generativni modeli niso nujno vir resnic in zanesljivih informacij, je »ključen vidik naših interakcij s takimi orodji« (Zuber in Gogoll 2024, 12) in našega krmarjenja po s sintetičnimi vsebinami nasičenem spletu. Dalje Angwin, Nelson in Palta (2024, 20) opozarjajo na »krizo odgovornosti«, ki nastaja na področju orodij UI: »Umetnointeligenčni modeli postajajo priljubljen vir informacij, a javno dostopni načini za njihovo testiranje in postavljanje standardov delovanja, še posebej glede točnosti in škodljivosti, so omejeni.« Večina najzmogljivejših generativnih modelov je danes v rokah peščice zasebnih korporacij, katerih cilj je čim večji zaslužek, zato samoregulacija ni verjetna. Njihovo prevzemanje odgovornosti, distribucija moči na področju UI in ustrezen regulativni okvir, ki ščiti demokracijo, so zato nujni (Coeckelbergh 2023; GPAI 2023).

ZAKLJUČEK

Prispevek je z obravnavo generativne UI v kontekstu družbenopolitičnih procesov skušal osvetliti vpliv generativnih modelov in z njimi ustvarjenih sintetičnih vsebin. Medtem ko teoretični prispevki na to temo svarijo predvsem pred dezinformacijskimi kampanjami, lažnimi novicami, mikortargetiranjem in informacijskim poplavljanjem, se empirične študije osredinjajo na manipulacijo in izkrivljanje resničnosti, s katerim dovršene sintetične vsebine lahko spremenijo politično naravnano udeležencev. Raziskave, ki preučujejo tehnologijo globokih ponaredkov, kažejo na njeno dovršenost in na dovzetnost posameznikov za manipulacijo s sintetičnimi avdio-video in besedil-

nimi vsebinami. Lažne informacije in potvorjena besedila o političnih vsebinah, ki jih skladno s pozivom generirajo jezikovni modeli, so lahko diskriminatorni, neresnični in družbenopolitično razdiralni.

Analize volilnega leta 2024 kažejo na znatno manjši vpliv UI na izid volitev, kot so ga mnogi pričakovali. Bolj kot da škodljive sintetične vsebine, ustvarjene z generativno UI, neposredno spodkopavajo demokratične temelje, se zdi, da so njihovi učinki sekundarni: v politični prostor vnašajo negotovost in neresnice ter polarizirajo, s tem pa nižajo zaupanje v oblasti in medije ter v začaranem krogu povečujejo dovzetnost posameznikov za dezinformacije. Poudariti velja tudi, da pri tem največkrat ne gre za malignost generativnih orodij samih, temveč za njihovo zlorabo.

Zaključujemo, da je vpliv uporabe generativne UI na demokracijo zapleten in ga je težko napovedati. Ugotavljamo pa, da dovršeni modeli z gotovostjo lahko povečujejo nevarnosti digitalnega okolja in krepijo njegove negativne vplive. Zdi se, da trenutni regulativni okvir s tem le stežka drži korak, saj od razvijalcev še ne zahteva detekcijskih mehanizmov za razpoznavo sintetičnih vsebin ter mehanizmov za preprečevanje generiranja škodljivih vsebin. Poleg regulacije lahko k blaženju negativnih učinkov generativne UI pripomoreta tudi transparentnost razvijalcev ter digitalna pismenost državljanov in državljanov.

ZAHVALA

Prispevek je deloma nastal v okviru ciljnega raziskovalnega projekta V2-2272. Opredelitev okvira za zagotavljanje zaupanja javnosti v sisteme umetne inteligence in njihove uporabe ob podpori Javne agencije za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije in Ministrstva za digitalno preobrazbo Republike Slovenije.

OPOMBA

- 1 Prispevek je razširitev prispevka, predstavljenega na konferenci Kognitivna znanost v okviru multikonference Informacijska družba (Košmrlj in Bratko 2024).

CITIRANA LITERATURA

- Adam, Michael, in Clotilde Hocquard. 2023. *Artificial Intelligence, Democracy and Elections*. EPRS, European Parliamentary Research Service. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/751478/EPRS_BRI\(2023\)751478_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/751478/EPRS_BRI(2023)751478_EN.pdf).
- Akt EU o UI: prva zakonodaja o umetni inteligenci. 2023. Evropski parlament: teme, 19. 6. 2023. <https://www.europarl.europa.eu/topics/sl/article/20230601STO93804/akt-eu-o-ui-prva-zakonodaja-o-umetni-inteligenci>.
- Akt o umetni inteligenci: poslanci sprejeli prelomno zakonodajo. 2024. Evropski parlament: novice, 13. 3. 2024. <https://www.europarl.europa.eu/news/sl/press-room/20240308IPR19015/akt-o-umetni-inteligenci-poslanci-sprejeli-prelomno-zakonodajo>.
- Angwin, Julia, Alona Nelson, in Rina Palta. 2024. *Seeking Reliable Election Information? Don't Trust AI*. The AI Democracy Projects. https://www.ias.edu/sites/default/files/AIDP_SeeingReliableElectionInformation-DontTrustAI_2024.pdf.
- Arvanitis, Lorenzo, Sadeghi McKenzie in Jack Brewster. 2023. »Despite OpenAI's Promises,

- the Company's New AI Tool Produces Misinformation More Frequently, and More Persuasively, than Its Predecessor». *NewsGuard*. <https://www.newsguardtech.com/misinformation-monitor/march-2023/#:~:text=Despite>.
- Au, Cheug Hang, Kevin K. W. Ho in Dickson K.W. Chiu. 2022. »The Role of Online Misinformation and Fake News in Ideological Polarization: Barriers, Catalysts, and Implications«. *Information Systems Frontiers* 24: 1331–1354. <https://doi.org/10.1007/s10796-021-10133-9>
- Bai, Hui, Jan G. Voelkel, Shane Muldowney, Johannes C. Eichstaedt in Robb Willer V tisku. »Artificial Intelligence Can Persuade Humans on Political Issues«. *OSF Preprints*. <https://doi.org/10.31219/osf.io/stakv>.
- Bommasani, Rishi, e al. 2021. »On the Opportunities and Risks of Foundation Models«. *ArXiv*. <https://doi.org/10.48550/arXiv.2108.07258>.
- Boyte, Hary C. 2017. »John Dewey and Citizen Politics: How Democracy Can Survive Artificial Intelligence and the Credo of Efficiency«. *Education and Culture* 33 (2): 13–47. <https://doi.org/10.5703/educationculture.33.2.0013>.
- Brody, Dorje C. 2024. »Generative AI Model Shows Fake News Has a Greater Influence on Elections when Released at a Steady Pace without Interruption«. *The Conversation*, 16. 4. 2024. <https://theconversation.com/generative-ai-model-shows-fake-news-has-a-greater-influence-on-elections-when-released-at-a-steady-pace-without-interruption-227332>.
- Castelvecchi, Davide. 2019. »AI Pioneer: 'The Dangers of Abuse Are Very Real'«. *Nature*. <https://doi.org/10.1038/d41586-019-00505-2>.
- CCDH, Center for Countering Digital Hate. 2023. *Misinformation on Bard, Google's New Chat*. 5. 4. 2023. <https://counterhate.com/research/misinformation-on-bard-google-ai-chat/>.
- Coeckelbergh, Mark. 2023. »Democracy, Epistemic Agency, and AI: Political Epistemology in Times of Artificial Intelligence«. *AI and Ethics* 3: 1341–1350. <https://doi.org/10.1007/s43681-022-00239-4>.
- Dhaliwal, Jasdev. 2024. »How to Survive the Deepfake Election with McAfee's 2024 Election AI Toolkit«. *McAfee Blog*, 28. 10. 2024. <https://www.mcafee.com/blogs/internet-security/how-to-survive-the-deepfake-election-with-mcafees-2024-election-ai-toolkit>.
- Dobber, Tom, Nadia Metoui, Damian Trilling, Natali Helberger in Claes de Vreese. 2021. »Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?«. *The International Journal of Press/Politics* 26 (1): 69–9. <https://doi.org/10.1177/1940161220944364>.
- EU Artificial Intelligence Act: Implementation Timeline. 2024. Future of Life Institute, 1. 8. 2024. <https://artificialintelligenceact.eu/implementation-timeline/>.
- EUDL. 2024. *What Is the Doppelganger Operation? List of Resources*. EU DisinfoLab, 30. 10. 2024. <https://www.disinfo.eu/doppelganger-operation/>.
- Folk, Zachary. 2024. »Magician Created AI-Generated Biden Robocall in New Hampshire for Democratic Consultant, Report Says«. *Forbes*, 23. 2. 2024. <https://www.forbes.com/sites/zacharyfolk/2024/02/23/magician-created-ai-generated-biden-robocall-in-new-hampshire-for-democratic-consultant-report-says/>.
- Goldstein, Josh A., Girish Sastry, Micah Musser, Renee DiResta, Mathew Gentzel in Katerina Sedova. 2023. »Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations«. *ArXiv*. <https://doi.org/10.48550/arXiv.2301.04246>.
- GPAL, The Global Partnership on Artificial Intelligence. 2023. *State-Of-The-Art Foundation AI Models Should be Accompanied by Detection Mechanisms as a Condition of Public Release*. <https://gpai.ai/projects/responsible-ai/social-media-governance/Social%20Media%20Governance%20Project%20-%20July%202023.pdf>.

- Gracia, Shanay. 2024. *Americans in Both Parties Are Concerned over the Impact of AI on the 2024 Presidential Campaign*. Pew Research Center, 19. 9. 2024. <https://www.pewresearch.org/short-reads/2024/09/19/concern-over-the-impact-of-ai-on-2024-presidential-campaign/>.
- Hameleers, Michael, Toni G.L.A. van der Meer in Tom Dobber. 2022. »You Won't Believe What They Just Said! The Effects of Political Deepfakes Embedded as Vox Populi on Social Media«. *Social Media + Society* 8 (3). <https://doi.org/10.1177/20563051221116346>.
- Hameleers, Michael, Toni G.L.A. van der Meer in Tom Dobber. 2024a. »Distorting the Truth versus Blatant Lies: The Effects of Different Degrees of Deception in Domestic and Foreign Political Deepfakes«. *Computers in Human Behavior* 152: 1–13. <https://doi.org/10.1016/j.chb.2023.108096>.
- Hameleers, Michael, Toni G.L.A. van der Meer in Tom Dobber. 2024b. »They Would Never Say Anything Like This! Reasons to Doubt Political Deepfakes«. *European Journal of Communication* 39 (1): 56–70. <https://doi.org/10.1177/02673231231184703>.
- Healey, Jon. 2024. »Voters Are Seeing More Deepfakes — and Worrying More About Their Influence. How to Spot Them«. *Los Angeles Times*, 29. 10. 2024. <https://www.latimes.com/politics/story/2024-10-29/how-to-spot-deepfakes-voters-fear-election-influence>.
- Heikkilä, Melissa. 2022. »How AI-generated Text Is Poisoning the Internet«. *MIT Technology Review*, 20. 12. 2022. <https://www.technologyreview.com/2022/12/20/1065667/how-ai-generated-text-is-poisoning-the-internet/>.
- Helbing, Dirk, et al. 2017. »Will Democracy Survive Big Data and Artificial Intelligence?«. *Towards Digital Enlightenment*: 73–89. DOI: 10.1007/978-3-319-90869-4_7.
- Innerarity, Daniel. 2024. *Artificial Intelligence and Democracy*. UNESCO United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000389736>.
- Jackson, Dean, in Meghan Conroy. 2024. »Apocalypse Later? The Real Impact of AI on the 2024 Elections«. *Atlantic Council*, 17. 10. 2024. <https://www.atlanticcouncil.org/content-series/the-big-story/apocalypse-later/>.
- Juefei-Xu, Felix, Run Wang, Yihao Huang, Qing Guo, Lei Ma in Yang Liu. 2022. »Countering Malicious DeepFakes: Survey, Battleground, and Horizon«. *International Journal of Computer Vision* 130 (7): 1678–1734. doi: 10.1007/s11263-022-01606-8.
- Jungherr, Andreas. 2023. »Artificial Intelligence and Democracy: A Conceptual Framework«. *Social Media + Society* 9 (3). <https://doi.org/10.1177/20563051231186353>.
- Kaplan, Andreas. 2020. »Artificial Intelligence, Social Media, and Fake News: Is This the End of Democracy?«. *Digital Transformation in Media & Society*: 149–161. DOI: 10.26650/B/SS07.2020.013.09.
- Košmrlj, Lea, in Ivan Bratko. 2024. »Vpliv generativne umetne inteligence na demokracijo«. *Information Society* 2024, 7.–11. oktober 2024, Ljubljana, Slovenija. <https://doi.org/10.70314/is.2024.cog.13>.
- Kreps, Sarah, in Doug L. Kriner. 2023a. »How AI Threatens Democracy«. *Journal of Democracy* 34 (4): 122–31. <https://www.journalofdemocracy.org/articles/how-ai-threatens-democracy/>.
- Kreps, Sarah, in Doug L. Kriner. 2023b. »The Potential Impact of Emerging Technologies on Democratic Representation: Evidence from a Field Experiment«. *New Media and Society*. <https://doi.org/10.1177/14614448231160526>.
- Łabuz, Mateusz. 2023. »Regulating Deep Fakes in the Artificial Intelligence Act«. *Applied Cybersecurity & Internet Governance* 2 (1). DOI: 10.60097/ACIG/162856.
- Mahadevan, Alex. 2023. »This Newspaper Doesn't Exist: How ChatGPT Can Launch Fake News Sites in Minutes«. *Poynter*, 3. 2. 2023. <https://www.poynter.org/ethics-trust/2023/chatgpt-build-fake-news-organization-website/>.

- Mahmud, Bahar Udin, in Afsana Sharmin. 2023. »Deep Insights of Deepfake Technology: A Review«. *ArXiv*. <https://doi.org/10.48550/arXiv.2105.00192>.
- Masood, Momina, et al. 2022. »Deepfakes Generation and Detection: State-Of-The-Art, Open Challenges, Countermeasures, and Way Forward«. *Applied Intelligence* 53 (4): 3974–4026. DOI: 10.1007/s10489-022-03766-z.
- Matza, Max. 2024. »Fake Biden Robocall Tells Voters to Skip New Hampshire Primary Election«. *BBC*, 23. 1. 2024. <https://www.bbc.com/news/world-us-canada-68064247>.
- Miller, Michael L., in Cristian Vaccari (ur.). 2020. »Digital Threats to Democracy: Comparative Lessons and Possible Remedies«. *The International Journal of Press/Politics* 25 (3). SAGE Publications. <https://doi.org/10.1177/1940161220922323>.
- Nemitz, Paul. 2018. »Constitutional Democracy and Technology in the Age of Artificial Intelligence«. *Royal Society Philosophical Transactions A* 376. <https://doi.org/10.1098/rsta.2018.0089>.
- NIST, National Institute of Standards and Technology, U.S. Department of Commerce. 2024. *Reducing Risks Posed by Synthetic Content: An Overview of Technical Approaches to Digital Content Transparency*. <https://airc.nist.gov/docs/NIST.AI.100-4.SyntheticContent.ipd.pdf>.
- Norwegian Consumer Council. 2023. *Ghost in the Machine: Addressing the Consumer Harms of Generative AI*. <https://storage02.forbrukerradet.no/media/2023/06/generative-ai-rapport-2023.pdf>.
- Pashentsev, Evgeny. 2023. »The Malicious Use of Deepfakes Against Psychological Security and Political Stability«. *The Palgrave Handbook of Malicious Use of AI and Psychological Security*: 47–80. https://doi.org/10.1007/978-3-031-22552-9_3.
- Pennycook, G., T. D. Cannon, in D. G. Rand. 2018. »Prior Exposure Increases Perceived Accuracy of Fake News«. *Journal of Experimental Psychology. General* 147 (12): 1865–1880. <https://doi.org/10.1037/xge0000465>.
- Rehagen, Toni. 2024. »Candidate AI: The Impact of Artificial Intelligence on Elections«. *Emory Magazine* Fall 2024. Emory University. https://news.emory.edu/features/2024/09/emag_ai_elections_25-09-2024/index.html.
- Robins-Early, Nick. 2024. »Trump Posts Deepfakes of Swift, Harris and Musk in Effort to Shore Up Support«. *The Guardian*, 19. 8. 2024. <https://www.theguardian.com/us-news/article/2024/aug/19/trump-ai-swift-harris-musk-deepfake-images>.
- Robinson, Olga, Shayan Sardarizadeh in Mike Wendling. 2024. »FBI Issues Warning Over Two Fake Election Videos«. *BBC*, 2. 11. 2024. <https://www.bbc.com/news/articles/cly2qjel083o>.
- Sainato, Michael. 2024. »Consultant Behind Deepfaked Biden Robocall Fined \$6m as New Charges Filed«. *The Guardian*, 23. 5. 2024. <https://www.theguardian.com/us-news/article/2024/may/23/biden-robocall-indicted-primary>.
- Schippers, Birgit. 2020. »Artificial Intelligence and Democratic Politics«. *Political Insight* 11 (1): 32–35. <https://doi.org/10.1177/2041905820911746>.
- Sippy, Tvesha, Florence Enock, Jonathan Bright in Helen Z. Margetts. 2024. »Behind the Deepfake: 8% Create; 90% Concerned. Surveying Public Exposure to and Perceptions of Deepfakes in the UK«. *ArXiv*. <https://doi.org/10.48550/arXiv.2407.05529>.
- Splichal, Slavko. 2024. *Upodatkovanje javnosti in mnenjske tehnologije*. Ljubljana: Fakulteta za družbene vede in Slovenska akademija znanosti in umetnosti.
- Spring, M. 2024a. »BBC Verify: How to Spot AI Fakes in the US Election«. *BBC*, 5. 3. 2024. <https://www.bbc.com/news/av/world-us-canada-68472248>.
- Spring, Marianna. 2024b. »Trump Supporters Target Black Voters with Faked AI Images«. *BBC*, 4. 3. 2024. <https://www.bbc.com/news/world-us-canada-68440150>.
- Stockwell, Sam. 2024. »AI-Enabled Influence Operations: Threat Analysis of the 2024 UK and

- European Elections«. *CETaS Research Reports*. https://cetas.turing.ac.uk/sites/default/files/2024-09/cetas_briefing_paper_-_ai-enabled_influence_operations_-_threat_analysis_of_the_2024_uk_and_european_elections.pdf.
- Stockwell, Sam, Megan Hughes, Phil Swatton in Katie Bishop. 2024a. »AI- Enabled Influence Operations: The Threat to the UK General Election«. *CETaS Briefing Papers*. https://cetas.turing.ac.uk/sites/default/files/2024-05/cetas_briefing_paper_-_ai-enabled_influence_operations_-_the_threat_to_the_uk_general_election.pdf.
- Stockwell, Sam, et al. 2024b. »AI-Enabled Influence Operations: Safeguarding Future Elections«. *CETaS Research Reports*.
- Swenson, Ali. 2024. »A Parody Ad Shared by Elon Musk Clones Kamala Harris' Voice, Raising Concerns about AI in Politics«. *AP News*, 29. 7. 2024. <https://apnews.com/article/parody-ad-ai-harris-musk-x-misleading-3a5df582f911a808d34f68b766aa3b8e>.
- Uredba (EU) 2024/1689 Evropskega parlamenta in Sveta z dne 13. junija 2024 o določitvi harmoniziranih pravil o umetni inteligenci in spremembi uredb (ES) št. 300/2008, (EU) št. 167/2013, (EU) št. 168/2013, (EU) 2018/858, (EU) 2018/1139 in (EU) 2019/2144 ter direktiv 2014/90/EU, (EU) 2016/797 in (EU) 2020/1828 (Akt o umetni inteligenci). 2024. Uradni list Evropske unije. https://eur-lex.europa.eu/legal-content/SL/TXT/PDF/?uri=OJ:L_202401689.
- Vaccari, Cristian, in Andrew Chadwick. 2020. »Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News«. *Social Media + Society* 6 (1): 1–13. <https://doi.org/10.1177/2056305120903408>.
- Zuber, Niina, in Jan Gogoll. 2024. »Vox Populi, Vox ChatGPT: Large Language Models, Education and Democracy«. *Philosophies* 9 (1). <https://doi.org/10.3390/philosophies9010013>.

UČINEK TEHNOLOGIJ PRI UPRAVLJANJU DEDIŠČINE: NEDENARNO VREDNOTENJE

Lana Nastja Anžur in Matija Strlič

UVOD

Hiter tempo tehnoloških in družbenih sprememb, skupaj s kompleksnostjo globalne dinamike, predstavlja izzive pri razvoju celovitega razumevanja sektorja kulturne dediščine – tako v njegovih neposrednih potrebah kot v njegovi vlogi akterja v globalni politiki in gospodarstvu. Da bi se ti izzivi učinkovito naslovili, se mora sektor dediščine odzvati kolektivno in izkoristiti potencial tehnologije ter povezanih sodelovalnih okvirov, kot je npr. Evropski okvir ukrepanja na področju kulturne dediščine, ki je namenjen spodbujanju in udejanjanju celostnih in participativnih pristopov h kulturni dediščini, obenem pa prispeva k njenemu vključevanju v vse politike EU (Evropska komisija, Generalni direktorat za izobraževanje, mladino, šport in kulturo 2019). Izjemnega pomena je tudi Skupna programska iniciativa za kulturno dediščino in globalne spremembe (JPI CH), ki jo vodijo države članice EU in povezuje nacionalne raziskovalne fundacije, ministrstva in raziskovalne svete iz Evrope z namenom reševanja družbenih izzivov znotraj okvira Evropskega raziskovalnega prostora (ERA). V letošnjem letu je pomemben mejnik na področju raziskav dediščine postavila tudi Evropska raziskovalna mreža za dediščinsko znanost, ki je postala ERIC (European Research Infrastructure Consortium), s čimer bo omogočila formalizirano in trajnostno sodelovanje med državami članicami ter raziskovalnimi institucijami za napredne raziskave in ohranjanje kulturne dediščine v Evropi in drugje. Odzivi na individualni ravni pogosto nimajo zadostne hitrosti, učinkovitosti in prilagodljivosti, ki jih zahteva dinamično politično in gospodarsko okolje, kar zmanjšuje njihove dolgoročnejše vplive.

Kulturna dediščina je pogosto predmet političnih odločitev, utemeljenih z vidika splošne javne koristi. Vendar pa je tak pogled premalo širok in posledično negativno vpliva na prepoznavo bolj specialnih, bolj posrednih vidikov družbenih, okoljskih in gospodarskih dobrobiti dediščine. V dobi tehnologije, v kateri je ozaveščenost o širših, pogosto neočitnih in težje merljivih vplivih ključna, posrednih koristi kulturne dediščine ne smemo podcenjevati.

To nas pripelje do pojma vrednosti – enega ključnih pojmov tako v ekonomski teoriji kot na področju kulturne dediščine. Vrednosti lahko opredelimo kot splošne koncepte, ki zajemajo vse, kar posamezniki ali družba prepoznava kot pomembno ali zaželeno. Med njimi pa imajo vrednote posebno mesto – gre za tiste vrednosti, ki so ponotranjene in prevzamejo vlogo moralnih in etičnih usmeritev, na podlagi

katerih posamezniki in skupnosti oblikujejo svoje vedenje ter sprejemajo odločitve. Z drugimi besedami, vrednote izražajo, katere vrednosti so za družbo temeljnega pomena.

Če želimo učinkovito krmariti v politični in družbeni dinamiki, je ključno, da dediščino predstavljamo na način, ki se sklada s širšimi družbenimi pričakovanji in potrebami. Danes nista *lingua franca* le politika in gospodarstvo, temveč je to tudi tehnologija. Ekonomski vidiki in tehnološki napredek na področju upravljanja podatkov in orodij za podporo odločanju so ključni za dediščinski sektor, saj prispevajo k jasnejšemu komuniciranju vrednosti. Da bi dediščina lahko imela pomenljiv vpliv, moramo njeno vrednost komunicirati na način, ki nagovarja te tri vidike, obenem pa odgovarja na nastajajoče potrebe deležnikov oziroma družbe v širšem smislu.

Zagotavljanje finančnih sredstev za ohranjanje vrednosti dediščine – torej za njeno varovanje – je stalen izziv. Da bi pridobili ustrezno finančno podporo, morajo upravljavci izkazovati dolgoročne družbene koristi, ki presegajo le kulturne vidike. Tehnološka orodja omogočajo lažje sledenje posrednim vplivom dediščine, s čimer lahko povečamo pripravljenost investitorjev za vlaganje v njeno ohranjanje. Ocenjevanje vrednosti – vrednotenje kulturne dediščine – v tem kontekstu zahteva instrumente, ki omogočajo kolektivno razumevanje med različnimi deležniki, pri čemer se povezuje tehnologija in interdisciplinarno sodelovanje. Pomanjkanje tovrstnih instrumentov, zlasti v kontekstu globalne trajnosti, lahko negativno vpliva na oblikovanje politik, dodeljevanje sredstev in upravljanje z dediščinskimi viri. V dobi ozaveščenosti glede prepleta tehnologije in družbe je pomembnejše kot kdaj koli prej, da kulturna dediščina zavzame vidno mesto v oblikovanju politik na področju trajnostnega razvoja.

Vrednosti, ki oblikujejo dediščino, ustvarjajo skupno osnovo, ki v kontekstu znanstvenih raziskav povezuje različne discipline – bodisi biologijo, zgodovino, umetnost, kemijo, arheologijo ali arhitekturo. Ta širina raziskovalcem omogoča, da se lažje vključijo v širši družbeni in tehnološki razvojni kontekst. Poleg tega lahko pristopi, ki temeljijo na razumevanju predvsem družbenih vrednosti dediščine, povežejo dediščino z raznolikimi, neekspertnimi deležniki na globok način. Razumevanje in spoštovanje družbenih vrednosti – vrednot različnih skupin, še posebej v tehnološko povezanem svetu, omogoča premostitev medsebojnih razlik in ustvarjanje občutka enotnosti. Z vključevanjem teh vrednosti v širši globalni diskurz postane dediščina ključno orodje za krepitev in spodbujanje ozaveščenosti o pomenu močnejših, odpornejših družbenih odnosov. Na tej podlagi upravljanje dediščine ne pomeni zgolj varovanja preteklosti, temveč aktivno preoblikovanje vrednot v konkretne vrednosti, ki vplivajo na sodobno družbo. Ko, denimo, kulturno znamenitost, kot je grad, interpretiramo in vključimo v turistično ponudbo, vrednoti – npr. zgodovinskemu pomenu – dodamo ekonomsko razsežnost s prodajo vstopnic ali turističnimi doživetji. Podobno lahko z revitalizacijo stare stavbe dediščini povrnemo uporabnost, kar ustvari prostorsko in okoljsko vrednost za lokalno skupnost. Taki procesi upravljanja dediščine pokažejo, kako se abstraktne družbene in kulturne vrednote lahko materializirajo v oprijemljive družbene, ekonomske ali okoljske koristi, pri čemer dediščina postane dinamičen vir trajnostnega razvoja.

USTVARJANJE DEDIŠČINE

Kulturna dediščina je skupni vir spomina, razumevanja, identitete, dialoga, družbene kohezije in ustvarjalnosti. Zajema širok spekter virov, ki smo jih podedovali iz preteklosti, in se uresničuje v različnih oblikah in izrazih – tako snovnih kot nesnovnih, vse bolj pa tudi digitalnih. Tehnološka zrelost sodobne družbe se med drugim odraža tudi v načinu, kako razumemo in upravljamo kulturno dediščino. Sem spadajo tako arhitekturni spomeniki, arheološka najdišča, krajine, umetniška dela in muzejske zbirke kot tudi pesmi, običaji, spretnosti in drugo znanje, ki se prenašajo iz roda v rod. Posebno pozornost pa zahteva dejstvo, da je Evropska komisija že uradno prepoznala digitalne oblike dediščine kot samostojno in pomembno kategorijo kulturne dediščine, kar pomeni prelomno točko v razumevanju njenega pomena in vloge v digitalni dobi. Ta kategorija vključuje tako digitalizirane vire (npr. skenirane rokopise, 3D-modele spomenikov) kot tudi »digitalno ustvarjene« (»born-digital«) vsebine, kot so digitalna umetnost, arhivi spletnih strani ali multimedijske instalacije, ki jih ustvarjamo in ohranjamo izključno v digitalnem okolju (European Commission 2024). Digitalna dediščina ne predstavlja več zgolj orodja za posredovanje informacij o preteklosti, temveč se uveljavlja kot polnopravni del kulturne krajine, ki odpira povsem nove pristope k njenemu ohranjanju, dostopnosti in interpretaciji. S tem pridobiva strateški pomen tako v kulturni politiki kot tudi v znanstvenoraziskovalnih in izobraževalnih okoljih.

Taka razširjena in sodobna opredelitev dediščine zahteva prilagojene raziskovalne, upravljalvske in izobraževalne pristope, ki upoštevajo kompleksnost, večdimenzionalnost in hitro spreminjajočo se naravo dediščinskih virov v 21. stoletju. Ob tem postaja vse pomembnejše tudi razumevanje dediščine kot nosilke različnih vrednosti, ki se ne izražajo le v njeni fizični pojavnosti, temveč predvsem v pomenih, ki ji jih pripisujejo posamezniki in skupnosti. Prav zato je treba raziskovanje in upravljanje dediščine zasnovati na vrednostno občutljivih pristopih, ki upoštevajo njeno simbolno, identitetno, socialno in celo okoljsko dimenzijo.

Vrednost kulturne dediščine ne izhaja zgolj iz njene starosti, redkosti ali estetske kakovosti, temveč iz pomena, ki ji ga pripisujejo posamezniki in skupnosti v določenem času in prostoru. Dediščina ni nekaj, kar zgolj »je«, ampak nekaj, kar se neprestano ustvarja, interpretira in ponovno ovrednoti skozi različne družbene, kulturne in politične procese.

Ena ključnih opredelitev tega pristopa se pojavi v publikaciji Mednarodnega centra za proučevanje, ohranjanje in obnovo kulturne dediščine – ICCROM z naslovom *People-Centred Approaches to the Conservation of Cultural Heritage: Living Heritage* (Court in Wijesuriya 2015). Avtorja poudarjata, da so skupnosti pomembni sooblikovalci pomenov, saj dediščina živi in se spreminja skozi njihove vsakdanje prakse, spomine in vrednote. Podobno razmišljanje zasledimo tudi pri inštitutu za konserviranje Getty, saj je v zborniku *Values in Heritage Management* (Avrami et al. 2019) izpostavljena potreba po večperspektivnem, na vrednostih utemeljenem upravljanju dediščine, ki presega zgolj strokovni nadzor in odpira prostor za dialog med različnimi deležniki.

Znotraj tega diskurza postaja vse bolj osrednje vprašanje, kdo ima pravico določiti, kaj sploh šteje za dediščino. Strokovnjaki pogosto predstavljajo posameznike s specializiranim znanjem, koristnim tako za javni kot zasebni interes (Hølleland in Skrede 2019), vendar se v praksi pojavlja napetost med strokovnimi in nestrokovnimi (tj. skupnostnimi) pogledi. Ta razprava poteka vzporedno z vprašanjem, kako se dediščina konstituira – ali je to zadeva njene intrinzične vrednosti, kot jo je med prvimi definirali Alois Riegl (1903/1996) v svojem razumevanju spominske vrednosti dediščine; zgodovinske objektivnosti, kot jo problematizira David Lowenthal (1998), ko opozarja, da dediščina pogosto temelji na selektivnem in emocionalnem razmerju do preteklosti, ne pa na objektivnih dejstvih; ali pa gre predvsem za kulturno prakso, kot trdi Laurajane Smith (2006), ki dediščino opredeljuje kot sredstvo za oblikovanje in regulacijo družbenih vrednot ter razumevanja.

V tem kontekstu se vse bolj uveljavlja razločevanje med t. i. avtoriziranim dediščinskim diskurzom (*«authorized heritage discourse»* – AHD), ki poudarja strokovno avtoriteto, nacionalne narative in materialnost dediščine, ter bolj odprtim, vključujočim (participatornim) dediščinskim diskurzom, ki priznava pomen vsakdanjih, lokalnih in čustvenih povezav z dediščino (European Commission 2024). Prav ti vključujoči pristopi omogočajo bolj dinamično razumevanje dediščine kot živega procesa, v katerem vrednost ni nekaj danega, temveč nekaj, kar se vzpostavlja v nenehni interakciji med preteklostjo, sedanjostjo in skupnostjo.

Avtorizirani in vključujoči dediščinski diskurz

Tradicionalni pristop vrednotenja dediščine temelji na njeni notranji (inherentni, intrinzični) vrednosti, ki je pogosto povezana z zgodovinskim in estetskim pomenom (Jones 2017, 2; Cutajar et al. 2016, 82).

Pojmovanje, da kulturna dediščina poseduje notranjo oziroma intrinzično vrednost – torej vrednost, ki obstaja neodvisno od funkcionalnosti, rabe ali njene prepoznavne – je pogosto prisoten v varstvenih politikah ter institucionalnih opredelitvah dediščine. Vendar je taka predpostavka filozofsko in metodološko sporna. Vrednost je po svoji naravi relacijska kategorija, kar pomeni, da predpostavlja obstoj subjekta, ki jo pripisuje. Čeprav se intrinzična vrednost pogosto pojmuje kot univerzalna ali neodvisna, gre v resnici za rezultat specifičnih kulturnih in ekspertnih atribucij (pripisovanja), ki so se zgodovinsko utrdile do tolikšne mere, da so danes pogosto zaznavane kot naravne (Lowenthal 1998).

V okviru sodobnih teoretskih pristopov, zlasti konstruktivističnih perspektiv na dediščino, je tako dojemanje problematizirano. Smith (2006) opozarja, da ideja o intrinzični vrednosti pogosto izhaja iz t. i. avtoriziranega dediščinskega diskurza (*«Authorized Heritage Discourse»*), ki ga oblikujejo strokovnjaki, institucije in dominantne kulturne elite. Ta diskurz promovira določene oblike dediščine – pogosto monumentalne, materialne in nacionalno obarvane – kot samoumevno vredne ohranjanja, pri tem pa marginalizira alternativne pomene, ki izhajajo iz vsakdanjih, lokalnih ali skupnostnih izkušenj.

Tudi klasični prispevki, kot je esej Aloisa Riegla *Der moderne Denkmalkultus* iz leta 1903, kažejo, da je vrednost dediščine vedno zadeva interpretacije. Riegl (1903/1996) je razlikoval med več vrstami vrednosti (npr. historično, umetniško in

temporalno), ki se pojavljajo glede na družbeni kontekst. Sodobni avtorji, kot je Rodney Harrison, opozarjajo na potrebo po vrednotenju dediščine kot procesa, ne kot fiksnega nabora objektov s pripisano vrednostjo. V svoji konceptualizaciji dediščine kot »multitemporalne mreže pomenov« Harrison (2013) poudarja, da dediščina deluje kot performativna praksa, v kateri so vrednosti dediščine nenehno v dinamičnem procesu nastajanja, prilagajanja in preoblikovanja. Podobno Graham Fairclough (2009) v kontekstu krajinskega pristopa zagovarja relacijsko in pluralno razumevanje vrednosti, ki jih ne določajo objekti sami, temveč odnosi med ljudmi, kraji in zgodbami.

Namesto sprejemanja intrinzične vrednosti kot objektivne danosti je tako smiselno govoriti o vrednosti kot o procesu – o dinamični, kulturno pogojeni in večglasni konstrukciji, ki odraža družbene odnose moči, pomenov in pripadnosti. Vključujoče vrednotenje dediščine zato ne pomeni zgolj razširitve nabora vrednot, temveč predvsem kritično refleksijo o tem, kdo vrednoti, kako se vrednosti utemeljujejo in katere vrednosti so v danem kontekstu prepoznane kot legitimne. Vključujoče vrednotenje dediščine zato ne pomeni zgolj razširitve nabora vrednot, temveč predvsem kritično refleksijo o tem, *kdo* vrednoti, *kako* se vrednosti utemeljujejo in *katere* vrednosti so v danem kontekstu prepoznane kot legitimne. Ta premik od esencialističnega k relacijskemu razumevanju vrednosti predstavlja ključen teoretski korak k bolj pravičnemu in trajnostnemu upravljanju kulturne dediščine.

Laurajane Smith (2006, 4) je avtorizirani dediščinski diskurz (AHD) opredelila kot »strokovni diskurz, ki daje prednost vrednotam in znanju strokovnjakov o preteklosti in njenih materialnih pojavih ter dominira in ureja strokovne dediščinske prakse«. AHD dojema dediščinsko vrednost kot notranjo, utelešeno v samem spomeniku, ločeno od vsakdanjih predmetov z značilnostmi, kot so monumentalnost, avtentičnost, starost, estetika in zgodovinski pomen (Smith in Waterton 2012, 165). Prepoznavanje te specifične vrednosti zahteva strokovno analizo, interpretacijo, vrednotenje in upravljanje, da se zagotovi trajna vrednost dediščine.

Dediščinske vrednosti, ki izhajajo iz dokumentov, kot so Atenska listina iz leta 1931, Beneška listina iz leta 1964 in Konvencija o svetovni dediščini iz leta 1972, poudarjajo materialne ali monumentalne vidike dediščine. Ta pristop vodi načelo avtentičnosti in ohranjanja zgodovinskega ter fizičnega konteksta spomenikov, stavb in območij (Cutajar et al. 2016, 83). V tem diskurzu so strokovnjaki obravnavani kot razsodniki vrednosti in ne le kot ena izmed skupin različnih dediščinskih deležnikov. Njihov vpliv na dediščinski diskurz se utrjuje z znanstvenim raziskovanjem, interpretacijo in vrednotenjem zgodovinskih ali kulturnih območij, kar je dodatno institucionalizirano znotraj nacionalnih in mednarodnih okvirov (Smith 2006, 28, 51).

V zadnjih desetletjih so se pristopi k upravljanju dediščine premaknili k pristopu, ki daje prednost pomenu in relevantnosti dediščine znotraj sistema dediščinskih vrednosti (de la Torre 2005 Duval et al. 2019; Avrami in Mason 2019). V vključujočem dediščinskem diskurzu je vrednost dediščine razumljena kot skupek različnih pomenov, ki vključujejo več perspektiv različnih deležnikov oz. ki jih sodobna družba pripisuje dediščini. Ta skupek pa predstavlja skupno vrednost dediščine (Fredheim in Khalaf 2016, 466). Premik v smeri prepoznave večplastnosti vrednosti dediščine je bil postopoma prepoznan najprej v konservatorski teoriji in na področju dediščinskih

študij, pravno podprt pa z mednarodnimi konvencijami, kot so Unescova Konvencija o varovanju nesnovne kulturne dediščine (2003), Faro konvencija (2005) in z Unesco-vo opredelitvijo kulturnih krajin (Unesco 1992; Avrami in Mason 2019).

Družbene in skupnostne vrednote so še vedno premalo zastopane v številnih okvirih za določanje, upravljanje in ohranjanje dediščine. Pristop, osredinjen na neekspertno javnost (*»people based approach«*), obravnava vrednost dediščine skozi njen pomen za skupnosti (Jones 2017, 21). Jones navaja, da »družbena vrednost zajema pomen zgodovinskega okolja za sodobne skupnosti, vključno z občutkom identitete, pripadnosti in prostora, pa tudi oblikami spomina in duhovne povezanosti. To so fluidne, kulturno specifične oblike vrednosti, ustvarjene z izkušnjami in s prakso. Poleg tega se nekatere ujemajo z avtoriziranimi dediščinskimi diskurzi, druge pa nastajajo z neformalnimi in neuradnimi oblikami vključevanja« (ibid.). Vrednosti dediščine, ki izvirajo iz njenih nematerialnih pomenov, bistveno prispevajo k njeni skupni vrednosti, a so pogosto prezrte pri procesih vrednotenja. Izključitev teh dimenzij lahko zmanjša skupno ocenjeno vrednost dediščine (Rogers 2019, 172).

V tehnološko zreli družbi je strokovno znanje nadgrajeno z interdisciplinarnimi metodologijami, ki vključujejo ekonomske, družbene in okoljske vidike v upravljanje dediščine (Clark 2019). Sodobna kontekstualizacija vrednosti dediščine danes vključuje vidike, kot so gospodarska rast, politični konflikti, družbena pravičnost in okoljska trajnost (Avrami in Mason 2019, 11). Uporaba tehnologije spodbuja vključevanje deležnikov, dediščina postane dostopnejša in razumljivejša, kar posledično omogoča vključitev širšega spektra mnenj v proces odločanja (Cutajar et al. 2016). Čeprav vključujoči pristopi zagovarjajo vključevanje deležnikov v vrednotenje dediščine, pa strokovno znanje še naprej oblikuje diskurz in zagotavlja ravnotežje med materialnim varstvom in sodobnimi potrebami (ibid., 82).

VREDNOTENJE DEDIŠČINE

Kulturna dediščina uteleša, ustvarja, povečuje in odraža različne vrednote in vrednosti (de la Torre 2002; Avrami et al. 2019; CHCfE Consortium 2015; Nocca 2017; Azzopardi et al. 2022). Te vrednote so pluralne in pogosto med seboj prepletene, saj je dediščina pomembna iz različnih razlogov – estetskih, zgodovinskih, družbenih, ekonomskih, okoljskih in drugih. Vendar pa ni splošnega soglasja o tem, katere vrednote sestavljajo vrednost kulturne dediščine, prav tako ne obstaja univerzalno sprejet sistem za njihovo razvrščanje ali vrednotenje. Nekateri avtorji celo poudarjajo, da je vzpostavitev univerzalnega sistema vrednot nemogoča zaradi velikega števila identificiranih vrednot in raznolikih tipologij (Fredheim in Khalaf 2016, 469).

Tako v Sloveniji kot v širšem evropskem prostoru ni enotnega stališča o kompleksnih vrednotah kulturne dediščine, kar se odraža tudi v odsotnosti splošno sprejete metodologije za njeno vrednotenje (Avrami et al. 2019; de la Torre 2002; Azzopardi et al. 2022). Izjemo predstavljata le Unescov seznam svetovne dediščine in postopek podelitve znaka evropske kulturne dediščine, ki temeljita na jasneje določenih kriterijih. Resolucija o nacionalnem programu za kulturo 2014–2017 sicer predvideva določitev kriterijev za vrednotenje dediščine, vendar ne določa konkretnih metod in postopkov za ocenjevanje njene vrednosti.

Vrednotenje dediščine pa je ključna in zahtevna strokovna naloga, saj je osnova za oblikovanje varstvene politike, razvrščanje dediščine in pripravo konservatorskih načrtov (Ogrinc, Pirkovič in Mörec 2017, 19). Poleg tega omogoča učinkovito uporabo razpoložljivih virov za zagotavljanje trajnosti dediščine ter izkoriščanje njenega potenciala za ustvarjanje nove vrednosti (Robb et al. 2013; CHCfE Consortium 2015; Nocca 2017; Duval et al. 2019). Ključno je torej pravilno prepoznavanje in razumevanje vrednosti, ki jih dediščina nosi, kar zahteva premišljeno in večdimenzionalno obravnavo.

Prepoznavanje in ocenjevanje vrednosti kulturne dediščine sta globoko povezani s stopnjo tehnološke in družbene zrelosti določenega okolja. Tehnološki napredek, zlasti na področjih digitalne dokumentacije, prostorskega kartiranja, umetne inteligence in podatkovne analitike, omogoča boljše razumevanje, merjenje in prenos vrednosti in vrednot, ki jih dediščina nosi. Z razvojem teh orodij se krepi tudi javna participacija in podpora pri oblikovanju politik, kar dodatno utrjuje položaj dediščine v sodobni družbi.

Pomemben dejavnik pa je tudi družbena zrelost, ki se izraža v učinkovitosti upravljanja, izobraževalnih sistemih in zakonodajnih okvirih. Ta vpliva na to, kako družba vrednoti dediščino in jo vključuje v strategije trajnostnega razvoja – enega ključnih ciljev sodobnih politik. V tem kontekstu postaja jasno, da vrednotenje dediščine ni zgolj tehnična naloga, temveč kompleksen proces, ki vključuje subjektivne poglede, raznolike družbene perspektive in sodelovanje različnih strok.

V središču tega procesa je vprašanje, kako prepoznati in ovrednotiti dediščino na način, ki omogoča tako njeno zaščito kot tudi smiselno vključevanje v razvojne načrte. Ključna pri tem je učinkovita komunikacija z odločevalci, saj lahko brez skupnega razumevanja pomena in vrednosti dediščine številni njeni vplivi ostanejo spregledani ali premalo raziskani (Fredheim in Khalaf 2016, 469), kar omejuje njen potencial in zmanjšuje možne pozitivne učinke na družbo.

Za sistematično razumevanje dediščine so bile razvite različne tipologije vrednot, ki prispevajo k razvrščanju in razlaganju razlogov, zaradi katerih so določeni elementi kulturne dediščine dragoceni. Te okvire lahko danes nadgradimo z digitalnimi metodami, ki omogočajo celostnejše, bolj dinamične in vključujoče pristope. Tako ne le izboljšujemo procese vrednotenja, temveč tudi prispevamo k trajnostnemu upravljanju dediščine in njenemu vključevanju v širši kontekst družbenega razvoja.

Konservatorji kot dediščinski eksperti

Konservatorstvo kot dejavnost ohranjanja in upravljanja z dediščino (*»conservatorship«*) ima ključno vlogo pri ugotavljanju in določanju tako kulturne kot družbene vrednosti dediščine. Kot primarna disciplina, odgovorna za varovanje kulturnih dobrin, deluje kot institucionalna vstopna točka, na podlagi katere družba formalno prepozna in potrdi pomen dediščine. Pristop konservatorjev k vrednotenju kulturne dediščine je bistven, saj vpliva na to, kako družbe dojemajo in postavljajo prioritete pri zaščiti svoje snovne in nesnovne dediščine. Z vrednotenjem dediščine z vidika kulturnih in družbenih perspektiv konservatorji ne le zagotavljajo njeno ohranitev, temveč tudi prispevajo k širšemu razumevanju njene vloge pri oblikovanju kolektivne identitete in družbenih vrednot (Smith 2006; Cutajar et al. 2016; Clark 2019).

Pri pristopu, utemeljenem na vrednostih (*»value based approach«*), kot ga je prvi opisal Mason (2002), se vrednosti dediščine običajno preučujejo s kvalitativnimi družboslovnimi metodami, vključno z intervjuji, s fokusnimi skupinami in z etnografskimi raziskavami, podprtimi s participativnimi raziskovalnimi praksami (prim. Robb et al. 2013; Jones 2017; Duval et al. 2019; Jones et al. 2024). Tehnološki napredek je te metodologije bistveno izboljšal s vključevanjem digitalnih orodij, kot so analiza podatkov z umetno inteligenco, virtualna etnografija in digitalne participativne platforme. Te inovacije omogočajo širše in bolj vključujoče sodelovanje deležnikov ter s tem izboljšujejo dokumentiranje in razumevanje dediščinskih vrednosti. Ključna prednost vključevanja tehnologije v okvir upravljanja dediščine, ki je utemeljen na vrednostih, je ustvarjanje dinamičnega prostora za dialog med različnimi deležniki in strokovnjaki za dediščino. Geografski informacijski sistemi (GIS) in tehnike digitalnega modeliranja omogočajo vizualizacijo prostorskih vzorcev in odnosov znotraj dediščinskih območij, s poudarkom na pomembnosti ohranjanja kontekstualnih interakcij med deležniki in dediščino, ne le fizičnih lastnosti dediščine. Ta premik k celostnemu razumevanju dediščine z uporabo tehnologije zagotavlja kontinuiteto kulturnih praks in prisotnosti skupnosti. Poleg tega uporaba digitalnih etnografskih orodij izboljšuje prepoznavanje in kartiranje družbenih vrednosti dediščine z analizo vzorcev, odnosov in interakcij na bolj sistematičen način (Rogers 2019, 172).

V zadnjih desetletjih so politični in družbeni dejavniki vplivali na institucionalno in strokovno razumevanje vrednot dediščine in z njimi povezanih koristi. V razvoju teorije in prakse konservatorskega upravljanja z dediščino, ki je vse bolj prepleten s tehnološkim napredkom, postajajo digitalne dokumentacijske tehnike, kot so 3D-skeniranje, fotogrametrija in arhiviranje na osnovi tehnologije veriženja blokov (*»block-chain«*), standardna praksa. Uporaba tehnologij, kot so 3D-skeniranje in digitalna rekonstrukcija, omogoča izjemno natančno dokumentacijo spomenikov in njihovo avtentično obnovo, kot je bilo vidno pri obnovi pariške katedrale Notre-Dame, pri kateri je bilo ključno predhodno lasersko skeniranje 3D Andrewa Tallona (Tallon 2014). Tehnologija lidar pa je v kontekstu krajin razširila tudi razumevanje kulturnih kompleksov, kot kaže primer Angkor Wata, v katerem so na osnovi prepoznav obsežnega sistema struktur razširili vrednotenje in ga umestili v okvir vrednotenja dediščine kot kulturne krajine (Evans et al. 2013). Pomembna novost so tudi participativne digitalne platforme, ki omogočajo vključevanje lokalnih skupnosti pri prepoznavanju vrednosti dediščine, s čimer se osvetlijo tudi čustvene in nesnovne razsežnosti vrednotenja. Poleg tega tehnologije veriženja blokov omogočajo sledenje izvoru predmetov in potrjevanje njihove pristnosti, kar krepi tako ekonomsko kot simbolno vrednost (Trček 2022). V zadnjem času se razvijajo tudi orodja za avtomatsko prepoznavanje cvetnega prahu v medu na podlagi umetne inteligence, ki podpirajo zaščito geografskega porekla in s tem prispevajo k ohranjanju tradicionalne čebelarke dediščine (Olson et al. 2021, Anžur et al. 2025). Vse te prakse kažejo, da tehnologija ne nadomešča kulturnega pomena, temveč ga razširja, utrjuje in bolj transparentno utemeljuje.

S splošnega vidika sta tehnološki napredek in zrelost digitalnih praks na področju dediščine preoblikovala načine, kako se vrednote dediščine preučujejo, dokumentirajo in ohranjajo. Dediščinske institucije, kot so knjižnice in arhivi, predstavljajo to povezavo s svojo vlogo pri ohranjanju in širjenju znanja. Raziskava o demografiji zbirk

(Dillon et al. 2012) ponuja okvir za razumevanje, kako se dediščina vrednoti skozi čas in kako tehnološki napredek vpliva na njeno ohranitev, dostopnost in vključenost javnosti. Z vključevanjem interdisciplinarnih raziskav – vključno s podnebnimi znanostmi, študijami materialnih sprememb in z družboslovjem – so razvili napovedne modele za ocenjevanje propadanja dediščine skozi čas (ibid.). Vprašalnik, uporabljen v tej študiji, je pokazal, da uporabniki dediščine kulturne artefakte dojemajo kot neprecenljive zapise človeškega spomina, bistvene za izobraževanje in javno korist ter kot sestavni del nacionalne in kulturne identitete. Ugotovitve iz študije Dillon et al. (ibid.) nakazujejo, da se odnos javnosti do ohranjanja razlikuje glede na osebne izkušnje in izpostavljenost tehnologiji ter da je pod vplivom načina, kako institucije vključujejo digitalna orodja v svoje upravljanje. To je primerljivo z načinom, kako demografija oblikuje zdravstvene politike – obe področji se zanašata na statistično modeliranje, optimizacijo virov in dolgoročno načrtovanje za podporo družbeni blaginji.

Študija je skušala vzpostaviti realno merilo skupne vrednosti knjižnih zbirk za neposredne in posredne uporabnike, pri čemer je uporabila metodologijo pogojnega vrednotenja za oceno različnih storitev, vključno z dostopom do gradiva, bibliografskimi storitvami, uporabo čitalnic ter javnimi dejavnostmi, kot so razstave in raziskovalne pobude. Te funkcije prispevajo k izobraževalnim in gospodarskim koristim, razvoju turizma ter raziskovalnim in inovacijskim dejavnostim, zaradi česar so knjižnice ključne družbeno-kulturne institucije.

V raziskavo so bile vključene raznolike skupine uporabnikov, vključno z obiskovalci čitalnic, uporabniki digitalnih storitev in s splošno javnostjo kot posrednimi upravičenci. Skupno je sodelovalo 2359 naključno izbranih anketirancev po vsej državi, ki so izrazili individualna stališča o vrednosti v skladu z načeli, uporabljenimi pri vrednotenju kulturne dediščine. Vnaprej postavljene trditve, kot so »To je neprecenljivo in spomin na človeštvo« ter »Zelo pomembno je, da se ohrani, kar imamo«, ki so jih sodelujoči vrednotili po Likertovi lestvici, ponazarjajo zaznani pomen knjižničnih zbirk.

Za kvantitativno vrednotenje teh vrednot je bil uporabljen vprašalnik s 60 vprašanji, usmerjen v neposredne uporabnike – obiskovalce čitalnic (izobraževanje, raziskave, gospodarske dejavnosti), udeležence razstav (turizem, izobraževanje) in obiskovalce zgodovinskih knjižnic. Na vzorcu 543 odgovarjajočih je bila izvedena faktorska analiza za identifikacijo ključnih dimenzij vrednosti na podlagi 60 ocenjevalnih trditev. Uporabljena je bila analiza glavnih komponent z varimax rotacijo, ki je pokazala več jasno ločenih faktorjev vrednotenja (slika 1).

Analiza je pokazala, da so uporabniki, ki so knjižnice obiskovali z raziskovalnim ali izobraževalnim namenom, večji poudarek namenili posredovanju znanja, medtem ko so obiskovalci razstav in splošni uporabniki večjo vrednost pripisovali kulturnemu in zgodovinskemu pomenu. Ugotovitve podpirajo večdimenzionalni pristop k vrednotenju dediščine, ki poudarja potrebo po prilagojenih strategijah varovanja in vključevanja, ki odražajo različna pričakovanja in dožemanja uporabnikov.

Pomembno orodje, ki vse bolj dopolnjuje konservatorske pristope k vrednotenju in upravljanju kulturne dediščine, je tudi Enhancing Our Heritage Toolkit 2.0 (EoH 2.0). Gre za mednarodno preizkušeno metodologijo samoevalvacije, ki jo je razvila pobuda World Heritage Leadership v sodelovanju z Unescom in organizacijami

Faktor 1: Obstoj in prihodnja vrednost S2_55:...it is available for everyone S2_27:...it will be valued by people in the future S2_11:...by being in this archive or library, it shows that society cares for it S2_35:...it ensures we do not forget the past S2_51:...it has survived S2_43:...it is part of our culture and heritage S2_7:...I would like other people to be able to see this document S2_34:...somebody thought it was worth creating	Faktor 2: Materialna & čutna vrednost S2_26:...of the style and design of the document S2_36:...of the material of which it is made (e.g. the paper or ink) S2_6:...of the decoration on the document S2_20:...of the bindings or covers on the document S2_62:...I like the appearance of the document S2_59:...of the colours in the document S2_30:...of the way it smells S2_42:...of its age
Faktor 3: Vrednost za skupnost S2_58:...it supports business activities S2_49:...it helps to support the rights of citizens S2_39:...it is an important economic resource S2_23:...it supports the administrative activity of organisations	Faktor 4: Osebni pomen in identiteta S2_53:...it helps to strengthen family and community identity S2_9:...it gives me insight into my personal origins S2_52:...it helps me feel connected to other people in the present day S2_48:...it gives me an understanding of my personal values
Faktor 5: Razumevanje sedanosti S2_63:...it helps me to understand events in the present day S2_54:...it links the past to the present S2_60:...it makes me more aware of history around me	Faktor 8: Povezava s preteklostjo S2_15:...it is interesting to see what mattered to people in the past S2_4:...it connects me to the people who made and used the document S2_5:...as an object, it is part of history S2_45:...it connects me to people in the past
Faktor 6: Raziskovalna vrednost S2_29:...of the surprise I experience when I find something new S2_47:...it feeds my curiosity S2_19:...there is something special about seeing the original document S2_57:...it stimulates my senses	Faktor 9: Edinstvenost S2_21:...it is unique S2_44:...if this document were lost or damaged it would be irreplaceable
Faktor 7: Izobraževalna vrednost S2_25:...of the information contained in the text or images S2_32:...of what I can learn from the text or images S2_37:...it helps me to understand the past S2_38:...there will always be new things to learn from it S2_50:...it gives me insight into past events	

Slika 1: Trditve, ki predstavljajo sestavni del faktorjev, oblikovanih v študiji Dillon et al. (2012).

ICCROM, ICOMOS in IUCN. Orodje omogoča strukturirano presojo učinkovitosti upravljanja na dediščinskih območjih, pri čemer izhaja iz vprašanja, v kolikšni meri so ohranjene vrednosti dediščine, ki utemeljujejo izjemno univerzalno vrednost (OUV). Orodje je uporabno tako za naravno kot kulturno dediščino, zlasti pri pripravi ali prenovi upravljaljskih načrtov.

EoH 2.0 vključuje dvanajst orodij, ki pokrivajo vse ključne elemente upravljanja: od razumevanja vrednosti in atributov dediščine do analize upravljaljskih struktur, razporejanja virov in spremljanja doseženih rezultatov. Vsako orodje vsebuje nabor refleksijskih vprašanj, ki konservatorjem in upravljavcem pomagajo presojati, kje in kako izboljšati obstoječe prakse. V tem kontekstu EoH 2.0 ne deluje le kot tehnično orodje, temveč kot sredstvo za krepitev refleksije o vrednostih dediščine in njihovem povezovanju z dolgoročnim trajnostnim razvojem dediščine.

V tem širšem okviru postaja jasno, da sodobni konservatorji ne delujejo več zgolj kot varuhi materialnih ostalin preteklosti, temveč kot aktivni posredniki med dediščino, skupnostmi in tehnološkimi inovacijami. Njihova vloga pri vrednotenju kulturne dediščine je vse bolj interdisciplinarna, participativna in podatkovno podprta, kar omogoča celostnejše razumevanje družbenega pomena dediščine. Tehnološki napredek jim pri tem odpira nova orodja in pristope, vendar ostaja ključno vprašanje, kako zagotoviti, da bodo vrednote, ki jih dediščina nosi, prepoznane, ohranjene in

vključene v prihodnji razvoj. Ravno v tej nalogi – v povezovanju strokovnega znanja, družbenega uvida in tehnološkega potenciala – konservatorji ostajajo nepogrešljivi dediščinski eksperti prihodnosti.

Okoljska vrednost dediščine

Odnos med kulturo in okoljem je dinamičen, saj ga oblikujejo geografija, podnebje in geologija, ki vplivajo na urbanizacijo in kmetijske prakse (Historic England 2020). Dediščina oblikuje in spreminja okolja, pri čemer pušča otipljive in neotipljive sledi v obliki stavb, krajin in biotske raznovrstnosti, vse to pa je neločljivo povezano s širšim okoljskim kontekstom (ibid.). Ta interakcija poudarja pomen razumevanja povezave med ljudmi in njihovim okoljem, saj vrednotenje tako naravne kot kulturne dediščine temelji na tem odnosu. Raziskave s področja okoljske psihologije potrjujejo obstoj te povezave, saj rezultati kažejo, da ljudje oblikujejo čustvene vezi s kraji (prostorom) zaradi osebne identitete, izgubo teh krajev pa pogosto dojemajo kot izgubo identitete same (Lewicka 2011). V takem okviru skupnosti vrednotijo kraje ne le skozi prizmo njihovega zgodovinskega ali kulturnega pomena, temveč tudi glede na okoljski in čustveni kontekst (Koch in Gillespie 2022).

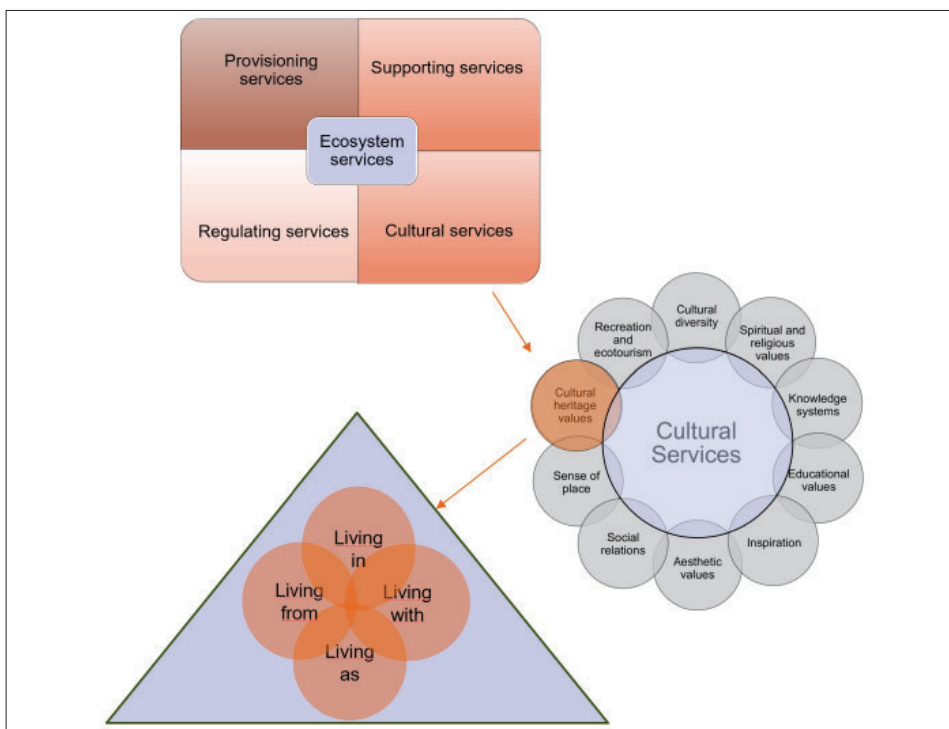
Konceptualna ločnica med naravno in kulturno dediščino, čeprav splošno sprejeta, je pogosto preveč poenostavljena in ne zajema celotnega spektra vrednot, povezanih z dediščino (Boer 2020; Azzopardi et al. 2022). Ta delitev izhaja predvsem iz pravnih okvirov in formalizacijskih postopkov, ki pojave kategorizirajo glede na njihov domnevni izvor ali značilnosti. Vendar širše razumevanje predpostavlja, da je naravna dediščina lahko integralni del kulturne dediščine. Kot navaja Boer (2020), je prepoznavanje elementov naravnega okolja kot dediščine rezultat razvijajočega se sistema kulturnih vrednot, oblikovanega med procesi akulturacije. To dodatno potrjuje Powell (2000), ki poudarja, da tako staroselske kot nestaroselske perspektive na dediščino gledajo kot na neprekinjen, živ proces, ki se razvija kot odziv na okoljske in družbene spremembe.

Dinamična interakcija med kulturno dediščino in okoljem je ključna za razumevanje celotne vrednosti dediščine v sodobnih družbah. Okolje prispeva k dediščini na različne načine: zagotavlja surovine za njeno ustvarjanje, vpliva na njeno preoblikovanje in deluje kot njen sokreator (Richards, Orr in Viles 2019). Zato lahko dediščino razumemo kot prilagodljiv sistem, ki se odziva in sodeluje s svojimi okoljskimi komponentami – z ekološkimi, s kulturnimi in z družbenimi (ibid.). To vzajemno razmerje poudarja pomembnost dojemanja dediščine kot več kot le statičnega artefakta – kot živega proizvoda interakcije med človekom in okoljem.

Biokulturna dediščina zajema prepletenost biološke in kulturne raznolikosti ter poudarja pomen tradicionalnega znanja, praks in kulturnih izrazov pri trajnostnem upravljanju ekosistemov. Popolno ponazarja neločljive vezi med ljudmi in naravo (International Society of Ethnobiology 1988). Priznava, da so lokalne skupnosti skozi generacije razvile edinstvene odnose s svojim naravnim okoljem, pri čemer sooblikujejo ne le svojo kulturno identiteto, temveč tudi ekološko znanje. Ta koncept spodbuja širše razumevanje kulturnih ekosistemskih storitev, saj vključuje vlogo skupnostnih praks in lokalne modrosti pri ohranjanju biotske raznovrstnosti (Lindholm in Ekblom 2019). Vključitev kulturne dediščine v širši okvir ekosistemskih storitev poudarja nje-

no večplastno vlogo pri vrednotenju okolja. Kulturna dediščina je tesno povezana s kulturnimi ekosistemskimi storitvami, ki zagotavljajo neotipljive koristi, kot so duhovna obogatitev, kognitivni razvoj in estetske izkušnje (Ehrlich in Ehrlich 1981; Milcu et al. 2013). Te storitve, ki so ključne za dobrobit ljudi, so bile v okviru okoljskih znanosti težko kvantificirane zaradi subjektivne in neotipljive narave (Daniel et al. 2012). Kljub temu so nedavni poskusi vključevanja vrednot dediščine v okvire ekosistemskih storitev pomembno prispevali k priznanju načinov, kako lahko kulturna dediščina spodbuja skrb za okolje in prispeva k trajnostnemu upravljanju ekosistemov (Hølleland, Skrede in Holmgaard 2017). Z vrednotenjem biokulturne dediščine lahko poglobimo razumevanje kulturnih razsežnosti ekosistemskih storitev ter zagotovimo, da bodo bolj raznoliki glasovi slišani in njihovi prispevki priznani v prizadevanjih za trajnostni razvoj. Biokulturna dediščina je posrednik za razumevanje prepleta narave in kulture, pri čemer sta dediščina in okolje neločljivo povezana in se medsebojno podpirata (Kenter in O'Connor 2022). Ta perspektiva je bila uporabljena kot podlaga za modificirani pristop prednostjenja »Okvir življenjskih vrednot« (*»life framework of values«* – LFV), ki poskuša dediščino vključiti v širši sistem interakcij med človekom in okoljem. Dopolnjuje in pogloblja koncept ekosistemskih storitev s poudarkom na relacijskih in kulturnih razsežnostih človekovega odnosa do narave. Medtem ko se tradicionalni okvir ekosistemskih storitev osredinja na koristi, ki jih ljudje pridobivajo iz naravnega okolja – kot so preskrbne storitve (npr. hrana, voda), regulacijske storitve (npr. podnebna regulacija, opraševanje), podporne funkcije (npr. kroženje hranil) in kulturne storitve (npr. rekreacija, duhovne vrednote) – ostaja ta pristop pogosto zamejen v instrumentalno oziroma utilitarno razumevanje narave, pri čemer je v ospredju vprašanje, kaj narava »nudi« človeku. Okvir življenjskih vrednot pa opozarja, da ljudje z naravo vzpostavljajo tudi globoke, neinstrumentalne odnose, ki se ne izražajo nujno z neposrednimi koristmi. Perspektiva »*living from*« se sicer najbolj približa klasičnemu razumevanju preskrbnih storitev, saj vključuje materialno odvisnost od narave. »*Living in*« poudarja pomen kraja, krajine in okolja za oblikovanje identitete, spomina in pripadnosti ter tako razširja pojem kulturnih storitev. »*Living with*« v ospredje postavlja etično in moralno soodgovornost do drugih živih bitij in naravnih procesov, kar presega zgolj funkcionalno vrednost narave. Nazadnje »*living as*« odraža ontološko in duhovno enotnost človeka z naravnim svetom, ki je pogosto bistvena sestavina avtohtonih in tradicionalnih sistemov znanja, a je v klasičnih modelih vrednotenja ekosistemov pogosto spregledana (Azzopardi et al. 2022).

Napredne tehnologije, kot so daljinsko zaznavanje, umetna inteligenca in digitalne platforme za participatorno kartiranje, vse bolj odpirajo vrata k vključevanju okoljskih vidikov v vrednotenje kulturne dediščine. Prostorske analize z orodji GIS omogočajo spremljanje degradacije ali revitalizacije kulturnih krajin – od terasiranih vinogradov do zgodovinskih vrtov – in s tem vrednotenje njihovih prispevkov k ekološkim funkcijam, kot so zadrževanje vode, blaženje toplotnih otokov ali varovanje biotske raznovrstnosti (Fatorić in Seekamp 2017; Ben Salem et al. 2021). Strojno učenje pa omogoča analizo satelitskih posnetkov za spremljanje vpliva podnebnih sprememb na dediščinske krajine ali celo oceno okoljskega odtisa restavratorskih posegov. Taki podatkovno podprti pristopi so ključni za razvoj celovitejših okvirov vrednotenja.



Slika 2: Vrednotenje kulturne dediščine v okviru ekosistemskih storitev.

Vir: avtorja, povzeto po Azzopardi et al. 2022

Kljub številnim študijam, ki se ukvarjajo s povezavami med naravo in kulturo – kot so koncepti biokulturne dediščine ali »*green heritage*« (Jones 2020; Tekeli in Turer Baskaya 2015) – pa okoljsko vrednotenje dediščine še vedno ostaja metodološko in teoretsko nerazvito področje. Večina pristopov temelji na subjektivnih ali kvalitativnih podatkih (npr. narativne metode, umetniški dialogi), ki sicer pomembno prispevajo k razumevanju čustvene povezanosti z okoljem, a jih je težko vključiti v politične ali ekonomske odločevalske procese (Bieling 2014; Edwards, Collins in Goto 2016).

Kljub tem pobudam pa še vedno ostaja odprto vprašanje, kako natančneje vrednotiti, kakšne koristi kulturna dediščina nudi okolju – bodisi v smislu ohranjanja krajinske pestrosti, podpore biotski raznovrstnosti ali kot vezivni člen za trajnostne oblike upravljanja prostora. To je pomemben raziskovalni izziv. Tehnološki napredek ponuja več možnosti za podporo razvoju okoljske valorizacije dediščine: od modeliranja prostorskih vplivov do senzornih omrežij za spremljanje ekosistemskih funkcij, integracije participatornih podatkov skupnosti, razvoja orodij za vrednotenje sinergij med kulturnimi in naravnimi ekosistemskimi storitvami ter celo uporabe umetne inteligence za zaznavanje ranljivosti dediščinskih krajin. Naslednji koraki bi morali vključevati sistematičen razvoj metodologij, ki povezujejo kulturno dediščino z merljivimi kazalniki okoljske trajnosti in pri tem nujno vključujejo tehnologijo kot sredstvo, in ne kot cilj.

Ekonomska vrednost dediščine

Ekonomsko vrednotenje kulturne dediščine je ključna komponenta razumevanja njenega širšega pomena, še posebej v kontekstu trajnostnega upravljanja in upravičevanja javnih proračunskih vlaganj. Tržna vrednost dediščine pogosto vpliva na raven naložb, ki jih je mogoče usmeriti v njen razvoj, kar vključuje tudi implementacijo digitalnih tehnologij. Ta ekonomska perspektiva je tesno povezana s konceptom kulturnega kapitala, katerega analogijo najdemo v naravnem kapitalu, kot ga opredeljuje okoljska ekonomija. Medtem ko naravni kapital zajema vrednotenje ekonomskih, družbenih in okoljskih razsežnosti z vidika trajnosti, uporaba podobnega okvira za kulturno dediščino omogoča celostno in dolgoročno naravnano razumevanje njenega pomena (Throsby 2019).

Kulturni kapital se običajno deli na dve ključni komponenti – medsebojno prepleteni in soodvisni ekonomsko in kulturno vrednost (Throsby 2001). Kljub temu v literaturi obstajajo različna stališča glede narave njunega razmerja. Medtem ko nekateri kulturni ekonomisti zagovarjajo tezo, da je kulturna vrednost v celoti zajeta v njeni ekonomski vrednosti, drugi opozarjajo na ločenost ali celo neodvisnost obeh dimenzij (Angelini in Castellani 2019). Ta razprava pomembno prispeva k razvoju razumevanja vloge ekonomske vrednosti pri oblikovanju dediščinskih politik in strategij varovanja.

Celotna ekonomska vrednost (*»Total Economic Value«* – TEV) v okoljski ekonomiki (*»environmental economics«*) zajema tako vrednost rabe kot vrednost nerabe (*»use value, non-use value«*) (Pearce in Turner 1990; Rac in Japelj 2021). Vrednost rabe je neposredno povezana z ekonomskimi koristmi, ki izhajajo iz finančnih tokov, medtem ko neuporabna vrednost zajema oblike vrednosti, ki niso tržno uveljavljene, a kljub temu prispevajo k družbeni blaginji. Mednje spadajo vrednost obstoja (*»existence value«*), opsijska (*»option value«*) in zapuščinska (*»bequest value«*) vrednost, ki poudarjajo pomen dediščine za prihodnje generacije ter njeno potencialno prihodnjo uporabo (Throsby 2019).

Vendar pa številne kulturne dobrine niso neposredno tržno dostopne in imajo značilnosti javnih dobrin – so nerivalne in nevključujoče. Posledično pogosto prihaja do tržne nepopolnosti, saj njihova dejanska vrednost ni zajeta v tržnih mehanizmi. Poleg tega kulturna dediščina generira številne pozitivne eksternalije, ki presegajo okvire posameznih deležnikov ter pomembno prispevajo k širšim družbenim in ekonomskim koristim, kot so razvoj turizma, krepitev družbene kohezije ter izobraževalna in kulturna rast (Throsby 2019). Te koristi pogosto ostajajo nekvantificirane, čeprav bistveno prispevajo k celostnemu vrednotenju dediščine.

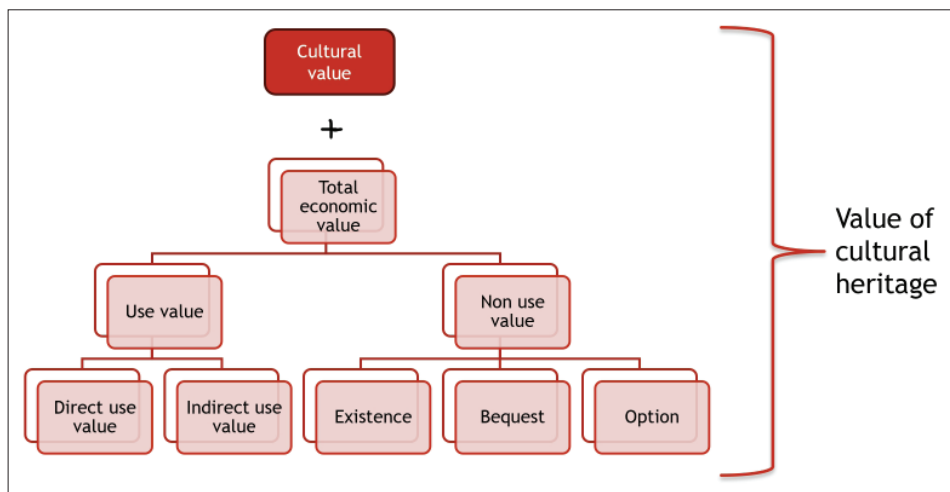
V tem kontekstu razvoj digitalnih tehnologij predstavlja pomembno prelomnico, saj omogoča naprednejše merjenje eksternalij in integracijo raznolikih vidikov uporabniških izkušenj v procese vrednotenja. Napredne tehnološke rešitve, kot so sledenje uporabniških interakcij, analiziranje čustvenih odzivov, sledenje gibanja in časa zadrževanja, nudijo dodatne dimenzije za ocenjevanje vrednosti, ki prej niso bile zajete v tradicionalnih metodah. To pomembno dopolnjuje cenovne politike ter prispeva k oblikovanju celostnejših strategij vrednotenja, ki podpirajo tudi okoljski in konservatorski pristop.

Digitalne tehnologije tako ne omogočajo le natančnejšega dokumentiranja in dolgoročnega ohranjanja kulturnih dobrin (npr. s 3D-modeliranjem, z virtualnimi

rekonstrukcijami in digitalnimi arhivi), temveč tudi interaktivno vključevanje javnosti in vzpostavljane novih oblik monetizacije, kot so spletni turizem, digitalne zbirke ali izobraževalne platforme. Poleg tega vzpostavljajo temelje za participativne mehanizme financiranja, kot je množično financiranje (*»crowdfunding«*), ki se je že uveljavilo kot učinkovit model za podporo dediščinskim projektom (Jelinčič in Šveb 2021).

Kljub številnim prednostim pa mora biti ekonomsko vrednotenje skrbno uravnoteženo s kulturnim pomenom dediščine. Ključno je, da dediščina ne postane reducirana na svojo tržno vrednost, saj bi to lahko spodkopalo njene estetske, družbene in zgodovinske razsežnosti, ki so temelj njenega varovanja (Hølleland, Skrede in Holmgaard 2017). Celostni pristop k vrednotenju omogoča, da ohranjanje ostaja zasidrano v družbenem kontekstu, hkrati pa omogoča razvoj trajnostnih politik za upravljanje z dediščinskimi viri.

Obenem je za odpravo tržnih nepopolnosti nujna vloga države, zlasti v obliki javnega financiranja. Zaradi težav pri kvantifikaciji vrednosti nerabe in konkurence z drugimi prioritetami vladnih politik se varstvo pogosto opira na omejena javna sredstva, ki ne zadoščajo za celovito upravljanje dediščine, še posebej v luči novih izzivov in priložnosti, ki jih prinaša digitalizacija (Mourato in Mazzanti 2002). Zato postajajo inovativni mehanizmi financiranja in interdisciplinarni pristopi k vrednotenju kulturne dediščine ključni za uravnoteženo povezovanje ekonomskih, kulturnih in tehnoloških dimenzij njenega trajnostnega razvoja.



Slika 3: Ekonomska vrednost dediščine (po Throsby 2001).

ZAKLJUČEK

Naraščajoče zanašanje na digitalne tehnologije pri vrednotenju in varovanju kulturne dediščine je temeljito preoblikovalo ustaljene metodološke pristope. Tradicionalne paradigme, osredinjene na materialne in zgodovinske vrednote, se razvijajo v večdimenzionalne okvire, ki vključujejo ekonomske, okoljske, družbene in izkustveno-afektivne razsežnosti. V tem procesu digitalna orodja – od umetne inteligence in GIS-kartiranja do verige blokov, virtualne resničnosti in participativnih platform

– omogočajo ne le boljšo dokumentacijo, temveč tudi učinkovitejše upravljanje, dostopnost in demokratično soudeležbo pri oblikovanju politik ohranjanja.

Tehnološka transformacija zato ni zgolj orodje za večjo učinkovitost, ampak postaja ključni mehanizem za sistematično prepoznavanje, vrednotenje in upravljanje vseh razsežnosti kulturne dediščine. S tem odpira prostor za preseganje dolgoletnih binarnih delitev – zlasti med ekonomskim in relacijskopomenskim razumevanjem dediščine, med naravnim in kulturnim, med znanstvenim in izkustvenoafektivnim. Digitalizacija tako ne pomeni grožnje redukcije, temveč priložnost za razvoj celovitih pristopov, ki omogočajo vključevanje dediščinskih vrednot v strateške razvojne politike.

Ekonomski pristopi – še posebej tisti, ki vrednotijo kulturne koristi, ekosistemske storitve ali vpliv dediščine na dobrobit – lahko ob ustrezni integraciji postanejo pomembno orodje za dolgoročno načrtovanje in financiranje. Ključno pa je, da ostanejo etično utemeljeni, lokalno prilagojeni in odprti za raznolike oblike vrednosti, ki ne izhajajo nujno iz tržnega okvira.

V tem kontekstu se kaže tudi potreba po ambicioznejših strokovnih in nacionalnih ciljih, ki bi digitalna orodja vključili v razvoj standardiziranih, a prožnih metod vrednotenja – še posebej v povezavi z okoljskimi vplivi in dolgoročno odpornostjo dediščine. Eden izmed pomembnejših prispevkov na tem področju je projekt Azzopardi et al. (2022), ki s prilagoditvijo modela Life Framework of Values ponuja redko uporabljeno osnovo za okoljsko vrednotenje dediščine. Ta razsežnost ostaja metodološko podraziskana, kljub jasnemu potencialu za vključevanje okoljskih kazalnikov, prostorskih modelov in participativnih digitalnih orodij.

V praksi razlike v pristopih včasih vodijo do navidezno nasprotujočih si ciljev – med tržno naravnostjo, ekološkimi zahtevami, skupnostnimi vrednotami in strokovno presojo. Vendar prav tehnologije omogočajo integracijo raznolikih pogledov, z boljšim usklajevanjem deležnikov in s preglednostjo procesov.

Digitalizacija tako ne ustvarja le novih ekonomskih in komunikacijskih priložnosti – od virtualnega turizma in digitalnih arhivov do izobraževalnih platform – temveč deluje tudi kot vezni člen med različnimi vrednostnimi sistemi. Uspešne dediščinske politike prihodnosti bodo odvisne od zmožnosti, da tehnologijo uporabimo ne zgolj kot orodje, temveč kot pospeševalec vrednot, ki dediščino ohranjajo kot dinamično, povezovalno in družbeno relevantno silo.

CITIRANA LITERATURA

- Angelini, Francesco, in Massimiliano Castellani. 2019. »Cultural and Economic Value: A Critical Review«. *Journal of Cultural Economics* 43 (2): 173–188.
- Anžur, Lana Nastja, Martin Jurkovič, Adrian Mladenič Grobelnik, Jaka Godejša, Živa Alif, Miha Humar, Ivan Bratko. 2025. »Nesnovna dediščina čebelarstva in uporaba umetne inteligence za avtomatizacijo pelodne analize medu = The Intangible Heritage of Beekeeping and the Use of Artificial Intelligence for the Automating the Pollen Analysis of Honey«. V 1st International Conference of Young Researchers in Heritage Science HERI SCI: Book of Abstracts, 68–71. <https://herisci.e-rihs.si/wp-content/uploads/2025/05/BoA-HeriSci.pdf>.
- Ashrafi, Baharak, Carola Neugebauer in Michael Kloos. 2022. »A Conceptual Framework for Heritage Impact Assessment: A Review and Perspective«. *Sustainability* 14 (1): 27. <https://>

- doi.org/10.3390/su14010027.
- Avrami, Erica C. 2016. »Making Historic Preservation Sustainable«. *Journal of the American Planning Association* 82 (2):104–112.
- Avrami, Erica C., in Randall Mason. 2019. »Mapping the Issue of Values«. V *Values in Heritage Management: Emerging Approaches and Research Directions*, ur. Erica Avrami, Susan Macdonald, Randall Mason in David Myers, 9–34. Los Angeles: Getty Conservation Institute.
- Avrami, Erica, Susan Macdonald, Randall Mason in David Myers (ur.). 2019. *Values in Heritage Management: Emerging Approaches and Research Directions*. Los Angeles: Getty Conservation Institute.
- Azzopardi, Elaine, et al. 2022. »What Are Heritage Values? Integrating Natural and Cultural Heritage into Environmental Valuation«. *People and Nature* 5: 368–383.
- Baars, Christian. 2011. »Dare to Prepare? The Value of Preparing and Sampling Historically Important Museum Collections«. *The Geological Curator* 9: 237–242.
- Ben Salem, Soumaya, Chokri Lahmar, Marie Simon in Krisztina Szilágyi. 2021. »Green System Development in the Medinas of Tunis and Marrakesh—Green Heritage and Urban Livability«. *Earth* 2 (4): 809–825. <https://doi.org/10.3390/earth2040048>.
- Bieling, Claudia. 2014. »Cultural Ecosystem Services as Revealed through Short Stories from Residents of the Swabian Alb (Germany)«. *Ecosystem Services* 8: 207–215. <https://doi.org/10.1016/j.ecoser.2014.04.002>.
- Boer, Ben. 2020. »The Environment and Cultural Heritage«. V *The Oxford Handbook of International Cultural Heritage Law*, ur. Francesco Francioni in Ana Filipa Vrdoljak, 318–346. Oxford: Oxford University Press. <https://doi.org/10.1093/law/9780198859871.003.0014>.
- CHCfE Consortium. 2015. *Cultural Heritage Counts for Europe*. 2. oktober 2023. <https://ocm.iccrom.org/documents/cultural-heritage-counts-europe>.
- Clark, Kate. 2019. »The Shift Toward Values in UK Heritage Practice«. V *Values in Heritage Management: Emerging Approaches and Research Directions*, ur. Erica Avrami, Susan Macdonald, Randall Mason in David Myers. Los Angeles: Getty Conservation Institute.
- Court, Sarah, in Gamini Wijesuriya. 2015. *People-Centred Approaches to the Conservation of Cultural Heritage: Living Heritage*. Rim: ICCROM. <https://www.iccrom.org/publication/people-centred-approaches-conservation-cultural-heritage-living-heritage>.
- Cutajar, Jan Dariusz, Abigail Duckor, Dean Sully in L. Harald Fredheim. 2016. »A Significant Statement: New Outlooks on Treatment Documentation«. *Journal of the Institute of Conservation* 39 (2): 81–97.
- Daniel, T. C., et al. 2012. »Contributions of Cultural Services to the Ecosystem Services Agenda«. *Proceedings of the National Academy of Sciences* 109 (23): 8812–8819. <https://doi.org/10.1073/pnas.1114773109>.
- de la Torre, Marta. 2002. *Assessing the Values of Cultural Heritage: Research Report*. Los Angeles: Getty Conservation Institute.
- de la Torre, Marta, ed. 2005. *Heritage Values in Site Management: Four Case Studies*. Los Angeles: Getty Conservation Institute, http://hdl.handle.net/10020/gci_pubs/values_site_mgmnt (cited June 25, 2025).
- Dillon, Catherine, William Lindsay, Joel Taylor, Kalliope Fouseki, Nancy Bell in Matija Strlič. 2012. »Collections Demography: Stakeholders' Views on the Lifetime of Collections«. V *Climate for Collections Conference*, München, Doerner Institut, 7.–9. november 2012, Postprints, ur. J. Ashley-Smith, A. Burmester in M. Eibl, 45–58. London: Archetype.
- Duval, Mélanie, Benjamin Smith, Stéphane Hœrlé, Lucie Bovet, Nokukhanya Khumalo in Lwazi Bhengu. 2019. »Towards a Holistic Approach to Heritage Values: A Multidisciplinary and Cosmopolitan Approach«. *International Journal of Heritage Studies* 25 (12): 1279–1301.

- Edwards, David M., Timothy M. Collins in Reiko Goto. 2016. »An Arts-Led Dialogue to Elicit Shared, Plural and Cultural Values of Ecosystems«. *Ecosystem Services* 21 (Part B): 319–328. <https://doi.org/10.1016/j.ecoser.2016.09.018>.
- Ehrlich, Paul R., in Anne H. Ehrlich. 1981. *Extinction: The Causes and Consequences of the Disappearance of Species*. New York: Random House.
- European Commission. 2024. *European Digital Heritage*. Brussels: European Commission. <https://culture.ec.europa.eu/cultural-heritage/cultural-heritage-in-eu-policies/european-digital-heritage>.
- European Commission. 2021. *Participatory Governance*. Brussels: European Commission. <https://culture.ec.europa.eu/cultural-heritage/eu-policy-for-cultural-heritage/participatory-governance>.
- Evans, Damian H., et al. 2013. »Uncovering Archaeological Landscapes at Angkor Using LiDAR«. *Proceedings of the National Academy of Sciences of the United States of America* 110 (31): 12595–12600. <https://doi.org/10.1073/pnas.1306539110>.
- Evropska komisija. 2023a. *European Green Deal*. 30. oktober 2023. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en.
- Evropska komisija. 2023b. *Trajnostni pristop in kulturna dediščina*. 2. oktober 2023. <https://culture.ec.europa.eu/sl/cultural-heritage/cultural-heritage-in-eu-policies/sustainability-and-cultural-heritage>.
- Fairclough, Graham. 2009. »New Heritage Frontiers«. V *Heritage and Beyond*, 29–41. Newcastle upon Tyne: Newcastle University. <https://eprints.ncl.ac.uk/>.
- Fatorić, Sansra, in Erin Seekamp. 2017. »Are Cultural Heritage and Resources Threatened by Climate Change? A Systematic Literature Review«. *Climatic Change* 142 (1): 227–254. <https://doi.org/10.1007/s10584-017-1929-9>.
- Fredheim, L. Harald, in Manal Khalaf. 2016. »The Significance of Values: Heritage Value Typologies Re-Examined«. *International Journal of Heritage Studies* 22 (6): 466–81. <https://doi.org/10.1080/13527258.2016.1171247>.
- Harrison, Rodney. 2013. *Heritage: Critical Approaches*. Abingdon: Routledge.
- Historic England. 2020. *Heritage and the Environment 2020*. London: Historic England.
- Hølleland, Herdis, in Joar Skrede. 2018. »What's Wrong with Heritage Experts? An Interdisciplinary Discussion of Experts and Expertise in Heritage Studies«. *International Journal of Heritage Studies* 25 (8): 825–36. <https://doi.org/10.1080/13527258.2018.1552613>.
- Hølleland, Herdis, Joar Skrede in Sanne Bech Holmgaard. 2017. »Cultural Heritage and Ecosystem Services: A Literature Review«. *Conservation and Management of Archaeological Sites* 19 (3): 210–37. <https://doi.org/10.1080/13505033.2017.1342069>.
- International Society of Ethnobiology. 1988. *Declaration of Belém*. Belém, Brazilija: International Society of Ethnobiology. <https://www.ethnobiology.net/what-we-do/core-programs/global-coalition-2/declaration-of-belem/>.
- Jelinčič, Daniela Angelina, in Marta Šveb. 2021. »Financial Sustainability of Cultural Heritage: A Review of Crowdfunding in Europe«. *Journal of Risk and Financial Management* 14 (3): 101. <https://doi.org/10.3390/jrfm14030101>.
- Jones, Karen. 2020. »Growing Canterbury's Green Heritage: Inspiring the Future«. V *The Routledge Companion to Landscape Studies*, druga izdaja, ur. Peter Howard, Ian Thompson, Emma Waterton in Mick Atha, 61–67. London: Routledge.
- Jones, Sian, et al. 2024. »Assessing the Dynamic Social Values of the 'Deep City': An Integrated Methodology Combining Online and Offline Approaches«. *Progress in Planning* 187: 1–39. <https://doi.org/10.1016/j.progress.2024.100852>.
- Jones, Sian. 2017. »Wrestling with the Social Value of Heritage: Problems, Dilemmas and Opportunities«. *Journal of Community Archaeology & Heritage* 4 (1): 21–37.

- Kenter, Jasper O., in Seb O'Connor. 2022. »The Life Framework of Values and Living as Nature: Towards a Full Recognition of Holistic and Relational Ontologies«. *Sustainability Science* 17: 2529–2542.
- Koch, Emma, in Josephine Gillespie. 2022. »Separating Natural and Cultural Heritage: An Outdated Approach?«. *Australian Geographer* 53 (2): 167–181.
- Lewicka, Maria. "Place Attachment: How Far Have We Come in the Last 40 Years?" *Journal of Environmental Psychology* 31, no. 3 (September 2011): 207–230.
- Lindholm, Karl-Johan, in Anneli Ekblom. 2019. »A Framework for Exploring and Managing Biocultural Heritage«. *Anthropocene* 25: 100195. <https://doi.org/10.1016/j.anecene.2019.100195>.
- Liu, Shuang, Robert Costanza, Stephen Farber in Austin Troy. 2010. »Valuing Ecosystem Services: Theory, Practice, and the Need for a Transdisciplinary Synthesis«. *Annals of the New York Academy of Sciences* 1185 (1): 54–78.
- Lowenthal, David. 1998. *The Heritage Crusade and the Spoils of History*. Cambridge: Cambridge University Press,
- Mason, Randall. 2002. »Assessing Values in Conservation Planning: Methodological Issues and Choices«. V *Assessing the Values of Cultural Heritage*, ur. Marta de la Torre, 5–30. Los Angeles: Getty Conservation Institute.
- Milcu, Andra Ioana, Jan Hanspach, David Abson in Joern Fischer. 2013. »Cultural Ecosystem Services: A Literature Review and Prospects for Future Research«. *Ecology and Society* 18 (3): 44, 1–30.
- Millennium Ecosystem Assessment. 2005. *Ecosystems and Human Well-Being: Synthesis*. 2. julij 2023. <https://www.millenniumassessment.org/documents/document.48.aspx.pdf>.
- Mourato, Susana, in Massimiliano Mazzanti. 2002. »Economic Valuation of Cultural Heritage: Evidence and Prospects«. V *Assessing the Values of Cultural Heritage*, edited by Marta de la Torre, 51–76. Los Angeles: Getty Conservation Institute,
- Nocca, Francesca. 2017. »The Role of Cultural Heritage in Sustainable Development: Multidimensional Indicators as Decision-Making Tool«. *Sustainability* 9 (10): 1882. <https://doi.org/10.3390/su9101882>.
- Ogrinc, Sabina, Jelka Pirkovič in Barbara Mörec. 2017. »Ekonomsko vrednotenje premične kulturne dediščine v slovenskih nacionalnih muzejih«. *AR. Arhitektura, raziskave* 18 (2): 18–29.
- Olsson, Ola, et al. 2021. »Efficient, Automated and Robust Pollen Analysis Using Deep Learning«. *Methods in Ecology and Evolution* 12(4): 665–676. <https://doi.org/10.1111/2041-210X.13575>.
- Onyenokulu, Obiajulum Vivian, Olagunju Oluwafemi Julius, Agidi Eunice Kanayo, and Maliki Omolebele Uchechukwu. 2024. »Technology and Cultural Heritage: Leveraging Digital Platforms for Preservation and Education«. *International Journal of Innovative Studies* 5 (1, May). <https://www.cescd.com.ng>.
- Pearce, David W., in R. Kerry Turner. 1990. *Economics of Natural Resources and the Environment*. London: Harvester Wheatsheaf.
- Powell, Judith. 2000. »Expanding Horizons: Environmental and Cultural Values within Natural Boundaries.« *International Journal of Heritage Studies* 6 (1): 49–65.
- Rac, Ilona, in Anže Japelj. 2021. *Ovrednotenje ekosistemskih storitev tal v kmetijski rabi (CRP V4-2023): Pregled literature – ekonomsko vrednotenje EST (R1.4)*. Ljubljana: Univerza v Ljubljani, Biotehniška fakulteta, in Gozdarski inštitut Slovenije.
- Richards, Jenny, Scott Allan Orr in Heather Viles. 2019. »Reconceptualising the Relationships between Heritage and Environment within an Earth System Science Framework«. *Journal of Cultural Heritage Management and Sustainable Development* 10 (2): 122–29.

- Riegl, Alois. 1903/1996. »The Modern Cult of Monuments: Its Essence and Its Development«. V *Historical and Philosophical Issues in the Conservation of Cultural Heritage*, uredil Nicholas Stanley-Price, M. Kirby Talley Jr., in Alessandra Melucco Vaccaro, 69–83. Los Angeles: Getty Conservation Institute.
- Robb, Jane E., Catherine Dillon, Mike Rumsey, Matija Strlič. 2013. »Quantitative Assessment of Perceived Value of Geological Collections by 'Experts' for Improved Collections Management." *The Geological Curator* 9 (10): 529–543.
- Rogers, Ayesha Pamela. 2019. »Values and Relationships Between Tangible and Intangible Dimensions of Heritage Places«. V *Values in Heritage Management: Emerging Approaches and Research Directions*, ur. E. Avrami, S. Macdonald, R. Mason in D. Myers. Getty Conservation Institute.
- Smith, Laurajane, in Emma Waterton. 2012. »Constrained by Commonsense: The Authorized Heritage Discourse in Contemporary Debates«. V *The Oxford Handbook of Public Archaeology*, ur. R. Skeates, C. McDavid in J. Carman, 153–71. Oxford University Press.
- Smith, Laurajane. 2006. *Uses of Heritage*. London: Routledge.
- Tallon, Andrew. 2014. »Divining Proportions in the Information Age." *Architectural Histories* 2 (1): Art. 15. <https://doi.org/10.5334/ah.bo>.
- Tekeli, Elcin, in Fatma A. Turer Baskaya. 2015. »Revealing Strategies for the Green Heritage of Istanbul – The Case of Historical Groves«. *WIT Transactions on Ecology and the Environment* 193: 597–607. <https://doi.org/10.2495/SDP150511.ž>
- Throsby, David. 2001. *Economics and Culture*. Cambridge, U.K.: Cambridge University Press.
- Throsby, David. 2019. »Heritage Economics: Coming to Terms with Value and Valuation«. V *Values in Heritage Management: Emerging Approaches and Research Directions*, ur. E. Avrami, S. Macdonald, R. Mason in D. Myers. Getty Conservation Institute.
- Trček, Denis. 2022. »Cultural Heritage Preservation by Using Blockchain Technologies«. *Heritage Science* 10: article 6. <https://doi.org/10.1186/s40494-021-00643-9>.
- UNESCO. 1992. *Operational Guidelines for the Implementation of the World Heritage Convention*. Pariz: UNESCO World Heritage Centre.
- UNESCO. *Concept of Digital Heritage*. Paris: United Nations Educational, Scientific and Cultural Organization, 2023. <https://webarchive.unesco.org/web/20230616073538/https://en.unesco.org/themes/information-preservation/digital-heritage/concept-digital-heritage>.

SPodbujanJE INOVACIJ ZA DRUŽBENO IN TEHNOLOŠKO PREOBRAZBO: ŠIRŠA PERSPEKTIVA

Franci Pušavec

Uvod

Inovacije so bistvo človeškega napredka, saj ne spodbujajo le tehnološkega razvoja, temveč navsezadnje tudi močne družbene preobrazbe. Medtem ko preboji na področju znanosti in tehnologije revolucionirajo industrijo, njihov vpliv presega samo inovativne izdelke in storitve. Močno namreč spreminjajo tudi družbo. Evropska unija (EU) se s strukturiranimi okviri, kot so *Program Obzorje Evropa*, *Evropski svet inovacij (EIC)* in *Evropski inštitut za inovacije in tehnologijo (EIT)*, uveljavlja in poudarja željo po spodbujanju inovacij, ki naslavlja globalne izzive in izboljšujejo družbeno blaginjo (Vir: European Commission 2024).

Tako je treba inovacije razumeti v dvojnih vlogah – kot katalizatorje tehnološke in tudi družbene preobrazbe. Ob priznavanju pomena razvoja vrhunskih izdelkov in tehnologij se vzpostavlja tudi enako pomembna potreba po obravnavanju družbenih izzivov, vključevanju inovacij v družbo in zagotavljanju trajnosti. Tako je s poudarkom na tehnoloških in tudi družbenih razsežnostih sinergijsko pričakovati uspešnost strategij EU na področju inovacij in njihovega širšega vpliva na družbo.

INOVACIJE KOT GONILCO DRUŽBENIH SPREMENB

Inovacije niso omejene na laboratorije in raziskovalne ustanove; so transformativna sila, ki oblikuje družbo. Jasno pa mora družba v tem videti korist. Napredni materiali, umetna inteligenca (UI/AI), čiste in aditivne tehnologije ter tehnologije obnovljivih virov energije so le nekateri primeri inovacij, ki sicer predvsem ustvarjajo tržno usmerjene izdelke, po drugi strani pa naslavlja ključne družbene izzive, kot so podnebne spremembe, javnozdravstveni izzivi, personalizacija izdelkov in neenakost družbenih dobrin.

Tehnologije obnovljivih virov energije na primer ne le zmanjšujejo negativne okoljske vplive, temveč tudi opolnomočijo skupnost o zagotavljanju cenovno dostopnih in trajnostnih rešitev za energijo. Podobno je za zdravstvene sisteme, ki temeljijo na sodobnih tehnologijah, značilen potencial za razširitev dostopa do komplementarnih zdravstvenih storitev, tudi v manj dostopnih regijah.

ŠTUDIJE PRIMEROV INOVATIVNIH TEHNOLOGIJ

Zelene tehnologije: Inovacije na področju zelene energije, kot so napredne gorivne celice in sistemi za shranjevanje energije, so spodbudile prehod k nizkoogljičnemu gospodarstvu. Te tehnologije so odličen primer presečišča med inovacijami in družbenim vplivom.

Umetna inteligenca (UI): Napredki na področju UI so omogočili obdelavo velike količine podatkov, celo interpretacije teh podatkov, razvoj npr. personalizirane medicine, pri čemer lahko algoritmi analizirajo podatke in so v pomoč pri diagnosticiranju bolezni, itd. To izboljšuje nivo napredka na širokem spektru področij (primer – nivo zdravstvene oskrbe tudi v manj razvitih regijah).

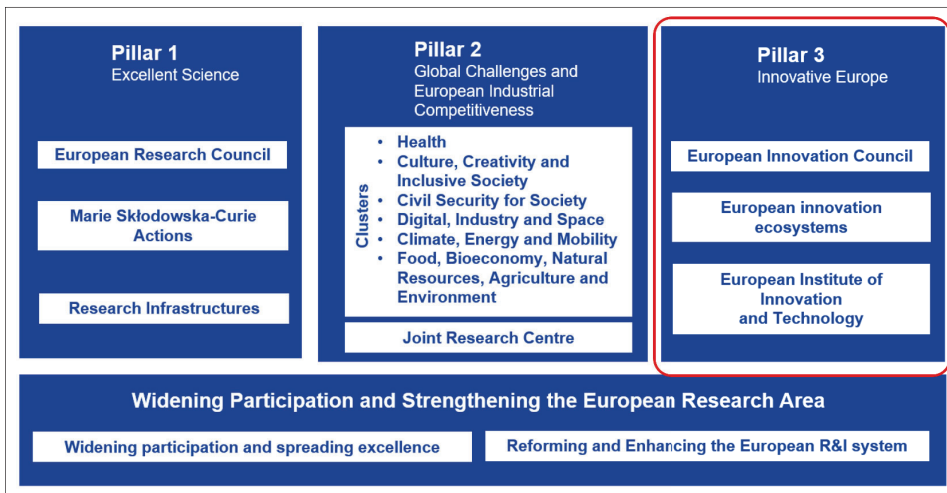
Digitalna vključenost: Tehnološki napredek mora vključevati vse družbene skupine, tudi tiste s tradicionalno omejenim dostopom do novih tehnologij. EU aktivno spodbuja pobude za izboljšanje digitalnih veščin in dostopa do tehnologij v oddaljenih oz. slabše razvitih regijah.

Vsem pa je treba dodati še vlogo etičnosti novih/inovativnih tehnologij. Primer: razvoj novih tehnologij, kot je UI, odpira pomembna vprašanja o zaupanju, zasebnosti, pristranskosti in družbeni pravičnosti. EU je vodilna na področju oblikovanja zakonodajnih okvirjev, ki spodbujajo etične inovacije. Pobude, kot je evropski zakon o umetni inteligenci (EU AI Act), zagotavljajo, da tehnologije služijo ljudem, in ne nasprotno.

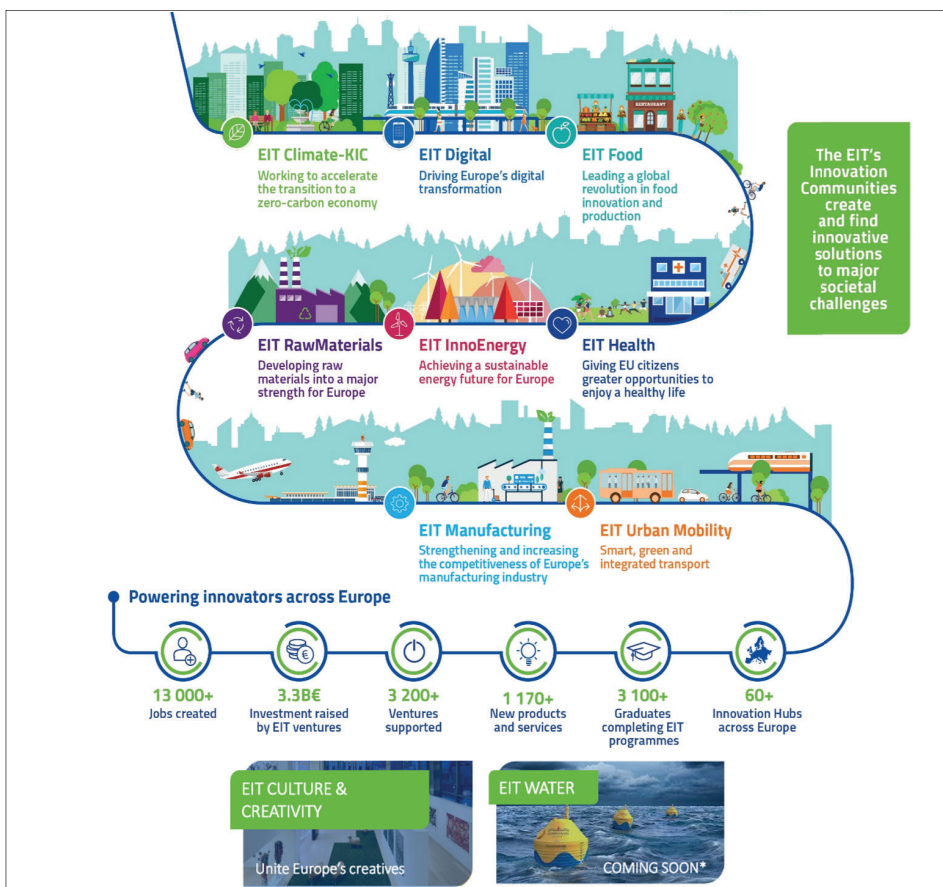
EVROPSKI OKVIRI ZA INOVACIJE (OBZORJE EVROPA IN EIT)

Obzorje Evropa s 95,5 milijarde evrov proračuna za obdobje 2021–2027 predstavlja zavezo EU za spodbujanje raziskav in inovacij. Program, strukturiran okoli treh stebrov, kot je prikazano na sliki 1, povezuje: odlično znanost (steber 1), globalne izzive inovativnega ekosistema in industrijske konkurenčnosti (steber 2) ter inovativno Evropo (steber 3). Na ta način si okvirni program Obzorje Evropa prizadeva ustvariti odporno in vključujoče okolje za inovacije. Analogno tu velikih sprememb z novim programom (2027–) ni pričakovati. Vsekakor pa bo prenos inovacij v družbo imel veliko (če ne še večjo) prioriteto.

EIT dopolnjuje program Obzorje Evropa z integracijo izobraževanja, raziskav in podjetništva v svojih znanstveno-inovacijskih skupnostih (KIC). Te skupnosti naslavljajo pereče družbene izzive (slika 2), kot so: podnebne spremembe, urbana mobilnost, digitalizacija, hrana, zdravje, energija, proizvodnja, omejeni materiali itd., in jih interpretirajo s sodelovanjem med akademskim svetom, industrijo in oblikovalci politik. Širino se da občutiti z vzpostavitevjo prek 60 vozlišč po Evropi in več kot 170 partnerjev (velikih in zagonskih podjetij ter znanstvenoraziskovalnih institucij). Poleg teh področij sta v zadnjem obdobju v ustanavljanju še dve pomembni področji, EIT Culture and Creativity 2022 in EIT Water 2024/2025, ki bosta dopolnili platformo še v smislu našega življenja/družbe in pomembne dobrine – vode.



Slika 1: Obzorje Evropa in okvir programa, ki se osredinja na prenos inovacij v družbo. Vir: Pušavec 2024; European Institute of Innovation & Technology 2024.



Slika 2: Znanstveno-inovacijske skupnosti inštituta za inovacije in tehnologije – EIT. Vir: European Institute of Innovation & Technology 2024.

PRESEGANJE TEHNOLOŠKIH IN DRUŽBENIH VRZELI

Obzorje Evropa in EIT poudarjata pomen vključevanja in trajnosti. Program, kot je Regionalna inovacijska shema (RIS), ki je predstavljena na sliki 3, zmanjšuje inovativnostne razlike med posameznimi državami/regijami s podporo inovacijam v manj razvitih regijah, med katere spadajo tudi Slovenija in vzhodnoevropske države. Pri tem shema temelji na povezovanju partnerjev iz teh regij s partnerji iz »inovacijsko bolj razvitih« regij EU; tudi s povezovanjem z vodilnimi industrijskimi partnerji v celotni Evropi.

Pobude, ki spodbujajo inovativnost s področja družboslovja in humanistike, pa zagotavljajo, da so inovacije koristne za vse segmente družbe. Kot je prikazano na sliki 4, je za dolgoročni tudi družbeni vpliv ključna povezava inovativnosti, podjetništva in izobraževanja. Obzorje Evropa se dodatno zaveda, da morajo koristi inovacij doseči vse dele družbe, pri čemer so ljudje postavljeni v središče inovacijsko-proizvodnega ekosistema. Poseben poudarek je na podpori malim in srednje velikim podjetjem (MSP), ki so ključni del evropskega, torej tudi slovenskega gospodarstva.



Slika 3: Regionalna inovacijska shema (RIS), usmerjena na reševanje inovacijskih vrzeli v manj razvitih regijah. Vir: Pušavec 2024; EIT Manufacturing 2024.

OD TRL DO DRUŽBENE PRIPRAVLJENOSTI (SRL)

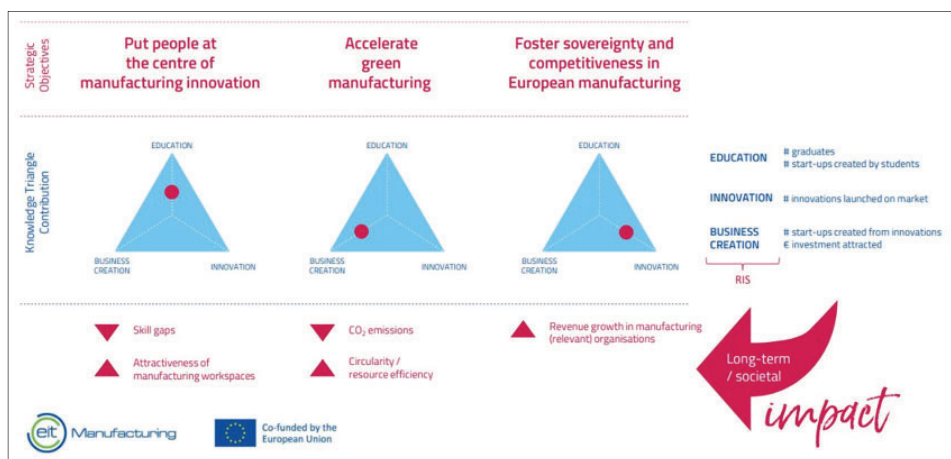
TRL (»Technology Readiness Levels«) zagotavlja sistematični okvir za ocenjevanje zrelosti tehnologije, od osnovnih (bazičnih) znanj/inovacij (TRL 1) do popolne implementacije na trg (TRL 9). Čeprav so nivoji TRL ključni za oceno tehnološke izv-

edljivosti/zrelosti, pogosto zameglijo s tem povezane družbene razsežnosti. Kot dopolnitev nivojev TRL tako koncept ravni družbene pripravljenosti (SRL) ocenjuje, kako dobro se inovacije prilagajajo družbenim potrebam in vrednotam oz. kako pripravljena je družba na inovacijo/-e. SRL namreč upošteva dejavnike, kot so javno sprejetje, etične implikacije in skladnost s predpisi, kar zagotavlja, da so tehnologije ne le izvedljive, temveč tudi zaželeni in vplivne. Primeri uporabe nivojev SRL:

Obnovljivi viri energije: Prehod s fosilnih goriv na obnovljive vire zahteva ne le tehnološki napredek, temveč tudi podporo javnosti, ne nazadnje tudi spremembo življenjskih navad/vzorcev. Kot odgovor na to so SRL v pomoč pri ocenitvi sprejetja skupnosti in obravnavi odporov z izobraževanjem ter vključevanjem deležnikov.

AI v javni upravi: Od avtomatizacije administrativnih nalog do izboljšanja sodelovanja državljanov ima AI velik potencial v javnem sektorju. Kot dodatek, pa SRL zagotavljajo, da te aplikacije spoštujejo zasebnost, pravičnost in dostopnost.

Zato platforma EIT navaja kot enega ključnih ciljev tudi postavitev človeka in družbe v središče inovacij in z njimi povezane evropske proizvodne industrije (slika 4). Tako je EIT Manufacturing vzorčni primer integracije tehnoloških in družbenih inovacij. Podprt z naprednimi proizvodnimi tehnologijami naslavlja izzive, kot so raznolikost delovne sile in regionalne neenakosti, njegove pobude pa naslavlajo spodbude za krepitev podjetniških, izobraževalnih in inovacijskih ekosistemov za izboljšanje gospodarske konkurenčnosti Evrope.



Slika 4: Regionalna inovacijska strategija EIT s poudarkom na dolgoročnem družbenem vplivu. Vir: EIT Manufacturing 2024.

NASLAVLJANJE REGIONALNIH RAZLIK

Kljub uspešnosti teh programov se EU sooča z izzivi pri zagotavljanju enakopravnih inovacij. Regije z omejeno infrastrukturo ali s šibko inovacijsko kulturo pogosto težko v celoti sodelujejo v programih, kot je Obzorje Evropa. Krepitev vloge nacionalnih kontaktnih točk (NCP) in prilagajanje programov regionalnim potrebam je lahko v oporo pri premostitvi te vrzeli. Mala in srednje velika podjetja (MSP) ter zapostavljene skupine so namreč ključni akterji inovacij. Programi, kot sta po-

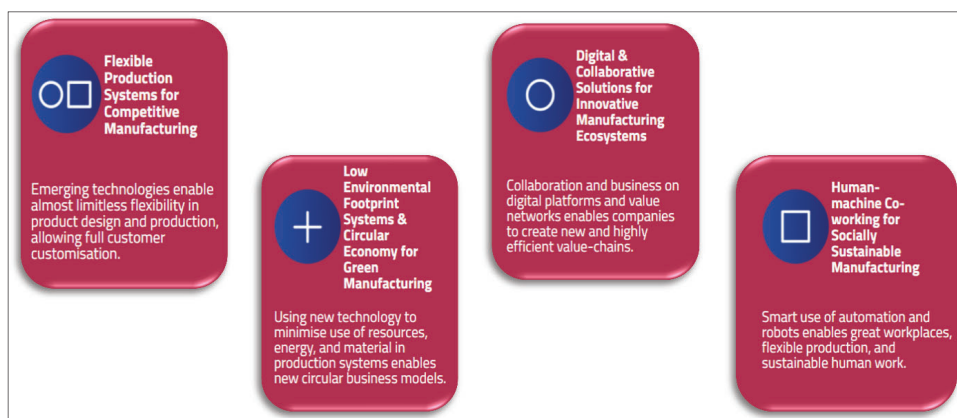
speševalnik EIC in RIS, nudijo ciljno podporo tem deležnikom ter zagotavljajo njihovo sodelovanje v evropskem inovacijskem ekosistemu:

Regionalna inovacijska shema (RIS): RIS se osredinja na manj razvite regije in izboljšuje zmogljivosti za inovacije ter spodbuja lokalno gospodarsko rast.

Pobude EIC: S financiranjem in z mentorstvom EIC podpira MSP pri razvoju inovacij, katerih potencial je družbeni vpliv.

USKLAJEVANJE TEHNOLOŠKIH IN DRUŽBENIH CILJEV

Za maksimalni vpliv morajo politike EU na področju inovacij uravnotežiti tehnološke ambicije z družbenimi potrebami. To zahteva vključitev ravni družbene pripravljenosti (SRL) v evalvacijske okvire ter spodbujanje interdisciplinarnega sodelovanja med tehnologi in družboslovci. EIT to naslavlja tudi s ključnimi področji, predstavljenimi na sliki 5, pri čemer so vključeni tudi družbeni vidiki.



Slika 5: Strategija za reševanje kompleksnih izzivov pri doseganju signifikantnega družbenega in industrijskega vpliva. Vir: Pušavec 2024; EIT Manufacturing 2024.

GLOBALNI IZZIVI IN TRAJNOST Z IZBOLJŠANJEM DOSTOPA IN VKLJUČEVANJA

Izboljšanje dostopa do inovacijskih programov za zapostavljene regije in skupnosti je ključnega pomena. Širitev programa RIS, krepitev NCP-jev in poenostavitev postopkov prijave lahko programom EU zagotovijo večjo vključenost in vpliv. Naslavljanje globalnih izzivov, kot so podnebne spremembe in družbena neenakost, namreč zahteva inovacije, ki presegajo nacionalne meje. Programi EU se morajo tako uskladiti z globalnimi okviri, kot so cilji trajnostnega razvoja (SDG), da zagotovijo svojo relevantnost in učinkovitost.

Čeprav okvirji EU – zlasti programi, kot so Obzorje Evropa, EIT in European Innovation Council (EIC) – ponujajo ambiciozne pristope za pospeševanje inovacij, se pri prenosu strategij v slovensko družbo in gospodarstvo pogosto pojavlja sistemska nekompatibilnost. Razlogi za to so večplastni in segajo od institucionalnih razlik do

kulturnih in operativnih razhajanj med evropsko politiko in slovenskimi razmerami. Za to je treba iskati vzroke v razdrobljenosti slovenskega inovacijskega ekosistema, kompleksnih in birokratsko zahtevnih razpisih EU za MSP-je, temeljenju okvirjev EU na makroekonomskih usmeritvah, medtem ko v Sloveniji velik delež gospodarstva predstavljajo mala in srednja podjetja z omejenimi inovacijskimi viri. Dodatno pa so koncepti, kot je družbena pripravljenost (SRL), sicer vključeni v okvirje EU, vendar pogosto niso smiselno lokalizirani. V slovenskem prostoru pa realno še vedno obstajajo nizka stopnja zaupanja v nove tehnologije, odpor do sprememb na področju delovnih mest zaradi avtomatizacije ter omejena vključenost državljanov v sooblikovanje inovacijskih rešitev.

Videti je tudi, da evropski inovacijski programi pogosto spodbujajo radikalne (disruptivne) inovacije in zagonska podjetja. V Sloveniji pa velik del gospodarstva temelji na tradicionalnih industrijah, v katerih bi bili pomembni predvsem inkrementalni razvoj, uvajanje digitalizacije in optimizacija procesov – ne pa nujno prebojni tehnološki preskoki. Na tem področju bo treba v Sloveniji narediti velik korak naprej za dvig širše inovativnosti.

INOVACIJSKI EKOSISTEM V SLOVENIJI

Slovenija se v zadnjih letih vztrajno uveljavlja kot inovacijsko zmerno uspešna država, kar potrjujejo različni kazalniki evropskega in globalnega inovacijskega indeksa (European Innovation Scoreboard – EIS). Po podatkih za leto 2023 pa je Slovenija še vedno pod povprečjem EU, vendar pa beleži pozitivne trende na področjih, kot so raziskovalna odličnost, število inovacij v MSP-jih ter sodelovanje med raziskovalno sfero in industrijo.

Dejstvo je, da se Slovenija sooča z več izzivi: z nizko ravno komercializacije znanja iz raziskovalnih ustanov, s premajhno povezanostjo med univerzami in gospodarstvom, z oslabljenim finančnim vlaganjem zasebnega sektorja v raziskave in razvoj, s pogosto skoncentriranostjo inovacij v večjih podjetjih, pri čemer so MSP-ji slabše vključeni. Zaradi navedenega se politika vse bolj osredinja na instrumente, ki bodo spodbudili ne samo tehnološke preboje, ampak tudi njihovo vključevanje v industrijsko prakso in vpliv na kakovost življenja prebivalstva. Ena od takih sprememb je ustanovitev Javne agencije za znanstvenoraziskovalno in inovacijsko dejavnosti republike Slovenije (ARIS), prej ARRS, ki odločno naslavlja poudarek tudi inovativne dejavnosti. S tem se želi zapirati vrzel med komercializacijo rezultatov raziskovalne dejavnosti. Jasno pa na nacionalni ravni Slovenija uveljavlja/krepi več ključnih mehanizmov in strateških dokumentov za podporo inovacijam, kot so strategija pametne specializacije (S5), Slovenski inovacijski sklad (SIS), razvojno-inovacijska partnerstva (SRIP), finančne spodbude oz. davčne olajšave, namenjene podjetjem, ter regionalni inovacijski ekosistemi/podjetniški inkubatorji.

ZAKLJUČEK

Inovacije so močno gonilo napredka, vendar je njihova resnična moč v zmožnosti prenosa v gospodarstvo in preoblikovanje družbe. Z uravnoteženjem tehnološkega

razvoja z družbenim vplivom se EU lahko postavlja med globalne vodje pri spodbujanju vključujočih in trajnostnih inovacij. Pri tem eno od ključnih vlog za doseg tega drži okvir EU za inovacije – platforma EIT.

Toda prenos teh okvirjev v slovensko družbo in gospodarstvo razkriva pomembne izzive. Čeprav Slovenija beleži pozitiven razvoj na področju inovacij, še vedno ostajajo sistemske slabosti, kot so nizka stopnja komercializacije znanja, razdrobljenost podpornih struktur in omejena dostopnost do evropskih programov za MSP-je ter regije zunaj večjih središč. Prav tako je pogosto opazno neujemanje evropskih prioritet s slovenskimi razvojnimi posebnostmi, pri čemer prevladujejo tradicionalne industrije, nizka stopnja digitalne preobrazbe in strukturno šibki podsistemi za vključevanje družbenih inovacij. Zato prihodnja uspešnost slovenskega inovacijskega prostora ne bo odvisna zgolj od obsega financiranja ali sprejetih strategij, temveč predvsem od sposobnosti prilagajanja evropskih okvirjev lokalnim razmeram, krepitve povezav med znanostjo, industrijo in družbo, interdisciplinarne in vključujoče obravnave izzivov ter vzpostavitve pogojev za enakopravno sodelovanje vseh regij in družbenih skupin.

Dejstvo je, da bo le z usklajevanjem tehnoloških ambicij z družbeno realnostjo mogoče izkoristiti inovacije kot orodje za dvig dodane vrednosti slovenskega gospodarstva, hkrati pa izboljšati kakovost življenja prebivalcev in zagotoviti dolgoročno družbeno odpornost. Slovenija ima priložnost, da postane model uspešne lokalne implementacije evropskih inovacijskih strategij, vendar le, če jih bo znala tudi kritično preoblikovati ter prilagoditi svojim izzivom in potencialom.

V širšem merilu pa se Evropa sooča z izzivi gospodarske konkurenčnosti, pri čemer bo njena zavezanost k spodbujanju inovacij, ki služijo tako trgu kot človeškim potrebam, ključnega pomena. Samo z naslavljanjem neenakosti, opolnomočenjem zapostavljenih skupin in usklajevanjem politik z globalnimi cilji lahko namreč EU ustvari inovacijski ekosistem, ki ne le ustvarja ekonomsko vrednost, temveč tudi izboljšuje konkurenčnost in družbeno blaginjo.

CITIRANA LITERATURA

- EIT Manufacturing. 2024. *Bringing European Manufacturers Together*. <https://www.eitmanufacturing.eu/>.
- European Commission. 2024. *Evaluation Study of the European Framework Programmes for Research and Innovation for Innovative Europe*. Report Phase 2 (support study for the interim evaluation of Horizon Europe). <https://op.europa.eu/s/z1jh>.
- European Institute of Innovation & Technology (EIT). 2024. <https://www.eit.europa.eu/>.
- Pušavec, Franci. 2024. *Pomembnost inovacij za trg, raziskovanje in izobraževanje*. Referat na posvetu Od tržne konkurenčnosti tehnologije k družbeni tehnološki zrelosti..SAZU, Ljubljana, 15.11.2024.

PAMETNA SPECIALIZACIJA V SLOVENIJI – TEHNOLOŠKI IN DRUŽBENI IZZIV

Andreja Jaklič

PRILOŽNOSTI IN IZZIVI NOVIH GENERACIJ INOVACIJSKIH POLITIK¹

Države Evropske unije (EU) že od globalne recesije v letu 2008 intenzivno iščejo načine za spodbujanje gospodarske rasti in razvoja. EU v strateških dokumentih in politikah največje možnosti za hitrejši razvoj pripisuje spodbujanju inovacij (glejte, denimo, strategijo Evropa 2020 (European Commission 2010), katere del so tudi Unija inovacij, Evropski svet za inovacije, program Obzorje Evropa ali nazadnje Draghijevo poročilo). Gospodarska nihanja, pandemija covida-19, naraščajoča prizadevanja za varovanje okolja, vojna v Ukrajini in drugi konflikti so ob vse ohlapnejših monetarnih in fiskalnih politikah še okrepili pomen inovacij in tlakovali pot k preporodu usmerjene industrijske politike (Rodrik 2022; Juhász, Lane in Rodrik 2023; Tagliapietra in Veugelers 2023). Med gospodarstvi z visokimi dohodki je po obdobju kriz pridobivala zagon zlasti zamisel, da bi morala biti družba (z industrijsko politiko) organizirana okoli velikih nalog ali misij (Henrekson, Sandström in Stenkula 2024) in da je treba spodbujanje inovacij povezati s cilji trajnostnega razvoja (Ranga in Kim 2023). To nadgrajuje (že v prvi recesiji oživljeno) idejo o velikem pomenu države (država kot veliki akter) pri spodbujanju inovacij in podjetništva (Mazzucato 2013). V državah članicah EU in ZDA lahko opazujemo, kako vlade uvajajo obsežne vladne programe za ponovni zagon postpandemičnih gospodarstev in doseganje drznih ciljev (kot je na primer veliko zmanjšanje emisij CO₂). Trend v smeri vedno bolj intervencionističnih industrijskih politik se pogosto imenuje »na misijo usmerjene inovacijske politike« (*»mission oriented innovation policies«*, glejte npr. Mazzucato 2018, 2019). OECD (2021, 15) definira »na misijo usmerjene inovacijske politike« kot usklajen sveženj političnih in regulativnih ukrepov, prilagojenih posebej za mobilizacijo znanosti, tehnologije in inovacij, da bi v določenem časovnem okviru obravnavali natančno opredeljene cilje, povezane z družbenimi izzivi.

Po četrti industrijski revoluciji (vzpostavljanju industrije 4.0) prevladuje mnenje, da bo razvoj Evrope v naslednjem političnem ciklu močno odvisen od uvajanja novih tehnologij in razvoja novih tehnoloških podjetij, zato so države aktivne pri oblikovanju industrijskih in inovacijskih politik ter pospešeno oblikujejo spodbude za razvoj tehnologije. V letu 2024 so v EU taki primeri, denimo, zakon o kritičnih surovinah, reforma zasnove trga električne energije ter zakon o čipih.² Kot pomemben poligon za uvajanje novih tehnologij in inovacij – pa tudi za nove generacije raziskovalnih in inovacijskih politik – so se v Evropi izkazale tudi strategije pametne specializaci-

je. Nove generacije inovacijskih politik gradijo na ustvarjalnem povezovanju, v katerem lahko vlade (ali sorodne mednarodne integracije) prevzamejo ključno vlogo omogočanja širših, raznovrstnejših »različic sodelovanja« med različnimi skupinami v razvitih gospodarstvih.

Cilj novih generacij inovacijskih politik ni samo spodbujanje gospodarske konkurenčnosti, temveč tudi prizadevanje za celostne družbene izboljšave, z obravnavo perečih izzivov s skupnimi prizadevanji, z mehanizmi strateškega financiranja in s prilagodljivimi strukturami upravljanja.³ Prednosti novih generacij inovacijskih politik je več. Med glavnimi so širok fokus aktivnosti, prilagodljivost in zmožnost sistemske transformacije. Presegajo tradicionalno podporo raziskavam in razvoju (R & R) in zajemajo širše gospodarske, tehnološke in družbene determinante inovacij. Vključujejo krepitev človeškega kapitala, infrastrukture in regulativnih okvirov za spodbujanje ugodnejšega okolja za inovacije. Medtem ko se tradicionalne politike osredinjajo na izbrane sektorje ali tehnologije, nove generacije inovacijskih politik predpostavljajo medsebojno povezanost različnih inovacij, spodbujajo medsektorsko sodelovanje in si prizadevajo za celovite spremembe v vzorcih proizvodnje in potrošnje za izpolnjevanje družbenih potreb. Poudarjajo sodelovanje med javnim in zasebnim sektorjem, kar vodi k boljši izmenjavi znanja in prelivanju inovacij (pa tudi k več prebojnimi, ne le postopnim inovacijam). Hkrati predvidevajo prilagodljivejšo in bolj participativno strukturo (modele) upravljanja, ki so odzivne na spreminjajoče se družbene potrebe. S ciljno usmerjenim financiranjem, podpornimi mehanizmi in spodbujanjem investicij podjetniškega sektorja naj bi te politike znatno povečale produktivnost v različnih sektorjih in spodbudile dolgoročno gospodarsko rast.

V zadnjem obdobju pa se (po izkušnjah in vrednotenjih) pospešeno pojavljajo tudi kritike novih generacij inovacijskih politik in na misije usmerjenih inovacijskih politik (Muldoon in Yonai 2023). Henrekson, Sandström in Stenkula (2024) predstavljajo primere zgodovinskih in nedavnih neuspehov »zelenih dogovorov« in podobnih prizadevanj, hkrati pa pokažejo nove smernice za inovacijsko politiko. Na primerih zgodovinskih neuspehov velikih državnih misij predlagajo, da je treba inovacijsko politiko obrniti: namesto da bi bila specifična in ciljno usmerjena, mora biti široka in splošna ter se osredinjati na splošne pogoje za delovanje podjetij. Namesto zagotavljanja ciljno usmerjene podpore določenim podjetjem, industrijam ali celo tehnologijam se mora inovacijska politika konstruktivno spoprijeti z ovirami za inovacije, vključno s proaktivnim obravnavanjem interesnih skupin. Tudi nova Evropska komisija, ki je prevzela mandat decembra 2024, uravnoteža svoje politične prednostne naloge ter bolj poudarja konkurenčnost in inovacije ter manj preprečevanje tveganj in regulacijo.

Strategija pametne specializacije (v začetni različici poimenovana S3 – »*Smart Specialization Strategy*«) je inovacijska politika nove generacije. Razvila jo je Evropska komisija, njen namen pa je povečati regionalne inovacije in gospodarsko rast s spodbujanjem regij, da se osredinijo na svoje edinstvene konkurenčne prednosti. Ta pristop se (je) ujema(l) s cilji strategije Evropa 2020 za spodbujanje pametne, trajnostne in vključujoče rasti s ciljno usmerjenimi naložbami v raziskave in inovacije. A pogosto so uresničevanja te strategije in ambiciozni cilji spodbujanja regionalnih inovacij razumljeni predvsem kot orodje tehnološke transformacije, ne pa tudi (ali predvsem) kot družbeni izziv.

Prispevek prikazuje izkušnje s spremljanjem in z vrednotenjem strateških razvojnih inovacijskih partnerstev (SRIP), ki so mehanizem v sklopu uresničevanja strategije pametne specializacije v Sloveniji. V nadaljevanju najprej na kratko predstavim proces uvajanja strategije pametne specializacije v Sloveniji, nato pa izkušnje iz spremljanja in vrednotenja uvajanja strategije pametne specializacije iz serije evalvacij. Presoja je pokazala omejitve uporabe kriterija potencialne tržne uspešnosti kot ključnega kriterija ocenjevanja inovacij in inovacijske politike, pa tudi nujnost ustreznega vrednotenja socialnih tehnologij (npr. ustvarjanja interdisciplinarnega lokalnega znanja in upravljanja partnerstev kot tehnologije) in spreminjanja (inovacijske) kulture.

UVAJANJE STRATEGIJE PAMETNE SPECIALIZACIJE V SLOVENIJI

Evropska komisija je zaradi pomanjkanja sodelovanja in vlaganj v skupne projekte gospodarstva na eni strani ter institucij na področju raziskav, razvoja in inovacij na drugi že v okviru programskega obdobja 2014–2020 predstavila koncept strategije pametne specializacije (znane pod kratico S3) z namenom spodbuditve vlaganja držav in njihovih regij v raziskave in inovacije iz Evropskega sklada za regionalni razvoj (ESRR).

Pametna specializacija je več kot le skupek ukrepov inovacijske politike; gre za celovit pristop k inovacijski politiki s ciljem preobrazbe držav ali regij in predstavlja strateško metodo za regionalni razvoj, ki regije ali države spodbuja h koncentraciji naložb v raziskave in razvoj (R & R) ter inovacije na področjih, na katerih imajo konkurenčne prednosti (Reid, Steward in Miedzinski 2023). Z bolj prilagojeno strategijo, specifično za določeno regijo, pametna specializacija predstavlja preskok in razvoj v oblikovanju evropske inovacijske politike.

Pristop je bil v evropski kohezijski politiki uveden kot način za (i) krepitev razvoja regionalne politike, (ii) podporo odločanju z vključevanjem deležnikov in (iii) osredinjanje javnih sredstev na inovacije (Gianelle, Guzzo in Mieszkowski 2019). Med ključnimi sestavnimi deli je proces podjetniškega odkrivanja (generiranje lokalnega znanja o konkurenčnih prednostih), ki omogoča lokalnim zainteresiranim stranem (skupnostim/deležnikom), da skupaj prepoznajo regionalne prednosti in slabosti. Standardizirane intervencije politike nadomesti selektivna politična intervencija. Pametna specializacija se osredinja na vertikalne prednostne naloge na določenih področjih, v omrežjih ter akterjih pa na široke horizontalne prioritete (kot je npr. generični razvoj človeškega kapitala). Vključuje širši pogled na inovacije, ki presega le raziskovalno intenzivne ali na tehnologiji temelječe dejavnosti, zaradi česar je pristop bolj vključujoč.

Proces od posameznih regij zahteva, da izvedejo proces podjetniškega odkrivanja (od spodaj navzgor), ugotovijo njihove najmočnejše prednosti na področju raziskav in inovacij ter izberejo le omejeno število prednostnih področij, na katerih lahko ustvarijo kritično maso. ESRR je od leta 2014 določil sprejetje strategije pametne specializacije kot predpogoj za financiranje razvojnih projektov, pri čemer je poudaril njen pomen ravno zaradi pristopa celovite inovacijske politike.

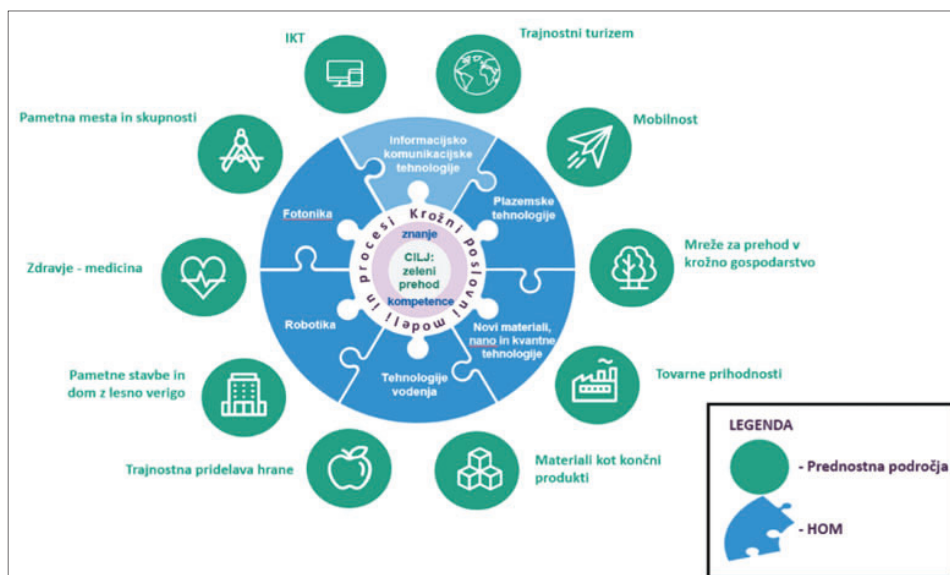
S slovensko strategijo pametne specializacije (*»Slovenian Smart Specialization Strategy«* – S4) so bile v letu 2015 opredeljene nacionalne strateške razvojne

prioritete in niše, ki so v praksi podprte s prilagojenim in celovitim svežnjem ukrepov. Slovenija je z izvajanjem S4 uvedla nov model razvojnega sodelovanja med ključnimi inovacijskimi deležniki, da bi se bolje integrirala v evropske in mednarodne razvojno-inovacijske mreže, platforme in konzorcije. Proces podjetniškega odkrivanja je tudi v Sloveniji temeljil na t. i. principu četverne vijačnice (Carayannis, Barth in Campbell 2012), ki terja sodelovanje podjetij, raziskovalne sfere, države in civilne družbe. Pri tem so se vzpostavila nova strateška razvojno-inovacijska partnerstva (SRIP),⁴ ki ostajajo osrednji mehanizem osredinjenja prodornih gospodarskih področij v Sloveniji. Izvajanje S4 je predstavljalo (in še vedno predstavlja) enega izmed ključnih orodij za krepitev in nadgradnjo slovenskega inovacijskega ekosistema (tudi v načrtu za okrevanje in odpornost).

S4 je bila v programskem obdobju 2014–2020 predhodni pogoj za sprostitvev sredstev iz Operativnega programa za izvajanje evropske kohezijske politike 2014–2020, in sicer na prednostni osi »Mednarodna konkurenčnost raziskav, inovacij in tehnološkega razvoja v skladu s pametno specializacijo za večjo konkurenčnost in ozelenitev gospodarstva«. S4 je v tem obdobju zasledovala poglobljena cilja evropske usmerjenosti – to je prehod v zeleno in digitalno družbo in je bila zasnovana kot orodje za preobrazbo gospodarstva za namen izboljšanja ekosistema in financiranja ukrepov na področju raziskav, razvoja in inovacij, človeških virov, podjetništva in internacionalizacije.

V letu 2020 je potekala prenova S4 (S5), predvsem analitičnih podlag (ki jih je z ekipo pripravil Anže Burger⁵) in v oživitvi procesa podjetniškega odkrivanja. Zasnova S5 je upoštevala tudi priporočila Poročila o produktivnosti 2021 (UMAR 2022), in sicer: (i) razvoj kadrov in spretnosti prihodnosti, (ii) vlogo javnih financ pri spodbujanju pametne, digitalno-inovacijske preobrazbe ter (iii) trajnostno preobrazbo v nizkoogljično in krožno gospodarstvo. Oblikovanje tega dokumenta je temeljilo na partnerskem pristopu, kjer so sodelovali pristojna ministrstva, raziskovalne organizacije, gospodarstvo, nevladni sektor in lokalni/regionalni akterji z ohranjanjem delne kontinuitete, na primer ohranile so se dobre prakse celovitosti (interdisciplinarnosti) usmerjanja svežnja ukrepov, aktivni odnos do deležniškega dialoga (nadgradnja razvoja SRIP) in večnivojsko upravljanje.

Strategija za novo obdobje do 2030 je z nadgradnjo v letih 2020–2022 v ospredje postavila trajnostni vidik, kot svoj osnovni cilj pa je identificirala zeleni prehod in se posledično preimenovala iz S4 v S5 (S5 – »*Slovenian Sustainable Smart Specialisation Strategy*«). V programskem obdobju 2021–2027 si je S5 kot cilj postavila zeleni prehod, ki ga razume kot »inovativno, nizkoogljično, digitalno in na znanju temelječo preobrazbo gospodarstva in družbe«. SRIP ostajajo osrednji akter stalnega procesa podjetniškega odkrivanja in strategija pametne specializacije ostaja podlaga za izvajanje evropske kohezijske politike tudi v programskem obdobju 2021–2027. Krepila se bo vloga horizontalnih mrež (HOM) in ključnih omogočitev tehnologij (»*key enabling technologies*« – KET). HOM IKT je v novi shemi prednostnih področij samostojno področje, tako da se je število SRIP-ov povečalo na deset (slika 1).



Slika 1: Prednostna področja in SRIP v S5.

Vir: Slovenska strategija trajnostne pametne specializacije, različica 1.0, januar 2023.

Pametna specializacija v upravljanju vključuje tako stopnjo tehnološke pripravljenosti (»*technological readiness level*« – TRL) kot tudi družbeno pripravljenost za inovacije (»*social readiness level*« – SRL). S5 priznava stopnje pripravljenosti tehnologije (TRL) kot orodje za ocenjevanje tehnološke zrelosti, za določanje razvojne stopnje in podporo sistematičnemu napredovanju inovacij v različnih razvojnih fazah. Ključni integracijski mehanizmi so usklajevanje financiranja, spremljanja napredovanja inovacije in strateško načrtovanje. V strateškem načrtovanju pametna specializacija uporablja TRL za standardizirano oceno tehnološkega razvoja, primerjalno vrednotenje inovacijskih projektov in za podporo ciljno usmerjenim naložbam v različnih stopnjah tehnološke pripravljenosti. Praktična izvedba tega pristopa omogoča regijam, da pripravijo zemljevid tehnološke zmogljivosti, ugotovijo potencialne inovacijske poti in razvijajo strategije, ki podpirajo tehnologije od zgodnjih konceptualnih faz (TRL 1–3) do rešitev, pripravljenih na trg (TRL 7–9). Strategija pametne specializacije s tem zagotavlja celovit, strukturiran pristop k razvoju regionalnih inovacij, ki presega preprosto tehnološko oceno.

Vendar pa strategija pametne specializacije presega zgolj tehnološke inovacije in inovacijske strategije povezuje z družbeno transformacijo (Reid, Steward in Miedzinski 2023). Integrira tudi načela družbene pripravljenosti za inovacije (SRL)⁶ ter zagotavlja, da inovacije niso le tehnološko napredne, ampak tudi družbeno pomembne in prilagodljive. Regije naj bi usklajevale inovacije s cilji trajnostnega razvoja in spodbujale take misije/partnerstva, ki jih vodijo (regionalni) izzivi. Pametna specializacija izrecno poudarja vključevanje več družbenih akterjev v inovacijske procese, identifikacijo ustreznih deležnikov na različnih stopnjah razvoja in ustvarjanje sodelovalnih mehanizmov za družbeno prilagajanje inovacij. S5 vključuje socialno pripravljenost na podlagi analize možnega družbenega vpliva inovacij ter z vrednotenjem razvoja in

prilagajanja načrtov prehoda za nove tehnologije ali rešitve. Praktična izvedba tega pristopa vključuje (i) prepoznavanje družbenih problemov in možnih rešitev (SRL 1–3), (ii) validacijo rešitev z ustreznimi deležniki (SRL 4–6) in (iii) izpopolnjevanje in implementacijo rešitev z družbenim učinkom (SRL 7–9).

IZKUŠNJE SPREMLJANJA PAMETNE SPECIALIZACIJE IN PRESOJE UČINKOVITOSTI IZBRANIH UKREPOV STRATEGIJE PAMETNE SPECIALIZACIJE

V nadaljevanju povzemamo izkušnje spremljanja napredka pametne specializacije v Sloveniji. Nekateri raziskovalci Fakultete za družbene vede Univerze v Ljubljani (FDV UL) smo bili skupaj z drugimi slovenskimi raziskovalci, ki se ukvarjajo z inovacijsko politiko (z Inštituta za ekonomska raziskovanja in Univerze na Primorskem) vključeni v naslednje ciljne raziskovalne projekte (CRP), ki so bili namenjeni spremljanju in presoji implementacije strategije pametne specializacije:

- Strateška razvojno-inovacijska partnerstva kot orodje krepitev inovacijske sposobnosti slovenskega gospodarstva (2016–2019) – (V5-1646; vodja projekta na FDV UL: prof. dr. Maja Bučar)
- Spremljanje in presoja učinkov strateških razvojno-inovacijskih partnerstev (2019–2022) – (V5-1932; vodja projekta na FDV UL: prof. dr. Maja Bučar)
- Priprave metodologije in presoja učinkovitosti ukrepov iz naslova strategije pametne specializacije (2021–2023) – (V5-2126; vodja projekta na FDV UL:izr. prof. dr. Anže Burger)
- Analiza uspešnosti strateških razvojno-inovacijskih partnerstev (SRIP) in kompetenčnih centrov. Pregled in ocena delovanja ter priporočila za spremembe in izboljšave (2023–2024) – (ARIS V5-2348; vodja projekta na FDV UL:izr. prof. dr. Anže Burger)

V prvem projektu smo raziskovalci oblikovali pilotno metodologijo za vrednotenje SRIP. Razvoj metodologije je potekal v dveh fazah; najprej s pripravo izhodiščnega predloga, ki v metodologijo vgrajuje pričakovanja oblikovalca ukrepa glede povezoovanja, organizacijske strukture in procesov delovanja SRIP-ov, ki je bila nato v drugi fazi pilotsko testirana na SRIP-ih. Pretekle analize slovenskega inovacijskega sistema (OECD 2012; ERAC 2010; MVZT 2011; RIO poročila o Sloveniji, Bučar 2024) so namreč izpostavile dva osrednja problema: pomanjkanje specializacije pri usmerjanju raziskovalnega financiranja ter nezadostno in neustrezno usmerjeno sodelovanje med relativno razvito javno raziskovalno sfero in gospodarstvom.

Izvedba prve evalvacije SRIP v letu 2019, ki je zajela obdobje od 2016–2019, je temeljila na pretežno kvalitativnem vrednotenju poglobljenih intervjujev z različnimi deležniki za vsako od SRIP. Pokazala je, da je bilo formiranje novega načina povezoovanja med gospodarskimi subjekti, javnimi raziskovalnimi organizacijami in drugimi deležniki in ustanavljanje SRIP zelo zahtevno, a je uspešno nadgradilo obstoječe strukture povezovanja (centre odličnosti, kompetenčne centre, grozde itd.) v smeri izgradnje kakovostnejših kapacitet za strateški preboj na posameznih področjih S4 ter tako pozitivno prispevalo k vzpostavitvi dolgoročnih javno-zasebnih partnerstev

z vodilno vlogo deležnikov (ne države) pri vzpostavljanju verig vrednosti in organiziranju celovite podpore raziskovalni in inovacijski dejavnosti (Bučar, Črnigoj in Lipnik 2020; Bučar et al. 2020).

V drugem projektu je raziskovalna skupina, ki je spremljala delovanje SRIP-ov od njihove ustanovitve dalje ter opravila vmesno vrednotenje delovanja devetih SRIP-ov do leta 2019, metodologijo vrednotenja nadgradila s priporočili za poročanje ter kvalitativnimi in kvantitativnimi kazalniki (Bučar et al. 2020; Bučar et al. 2022). Pripravila je tudi predloge za nadaljnjo nadgradnjo sistema upravljanja SRIP-ov, s ciljem omogočiti intenzivnejše sodelovanje ne le znotraj SRIP-ov, temveč tudi med SRIP-i (Črnigoj et al. 2022). Izkušnje pilotnega vrednotenja so pokazale potrebo za oblikovanje predlogov za tesnejše vključevanje SRIP-ov v kreiranje raziskovalno-inovacijskih prioritet na ravni države ter za njihovo sodelovanje pri strateških usmeritvah slovenskega gospodarskega razvoja. Poleg težav, kot so odvisnost od javnega financiranja, težav pri opazovanju in merjenju – ne le dolgoročnih, temveč tudi kratkoročnih mikro- in makroekonomskih učinkov, izključenosti manjših podjetij iz procesov novega grozdenja –, je analiza pokazala težave pri vzpostavljanju zaupanja med partnerji, težave v koordinaciji nalog in akcijskih načrtov, odsotnost operativne podpore in neizpolnjevanje zastavljenih časovnic idr. Cilj je bil oceniti učinkovitost in uspešnost instrumenta SRIP kot takega, pa tudi analizirati, kako optimizirati realizacijo S4 in s tem prispevati h gospodarskemu in raziskovalno-razvojnemu in inovacijskemu potencialu Slovenije. Vrednotenje SRIP je vključevalo večnivojsko kvantitativno in kvalitativno spremljanje SRIP (kjer se opazujejo struktura, strategija, upravljanje, financiranje, uvajanje HOT v vertikalne verige vrednosti ter skupni razvoj in inoviranje). Po članih SRIP smo spremljali tržne manifestacije (število novih tehnoloških rešitev, število predstavljenih prototipov, število procesnih izboljšav, število novih produktov, število novih storitev, število novih poslovnih modelov, število in obliko novonastale industrijske lastnine), rezultate tržnih manifestacij (število novoustanovljenih podjetij s strani članov SRIP-a in število zaposlenih v njih, prihodke iz naslova prodaje tehnoloških/netehnoloških rešitev, razvitih v skupnih projektih, ter prihodke iz naslova prodaje novih produktov in storitev, razvitih v skupnih projektih). Opazovala se je tudi internacionalizacija poslovanja in delovanje v mednarodnih verigah vrednosti članov SRIP-a (nove raziskave/sodelovanje kot posledica članstva v različnih mednarodnih povezavah, skupne naložbe s partnerji iz tujine ter vrednost in rast prihodkov in dobička, ustvarjenih na tujih trgih) ter razvoj človeških virov SRIP (spremembe izobrazbene strukture zaposlenih v SRIP-u, novi kadri). Analiza uspešnosti delovanja SRIP je pokazala več omejitev v zbiranju, razpoložljivosti in kakovosti podatkov, pa tudi potrebo po nadaljnjem kvalitativnem vrednotenju, zato so raziskovalci s sodelovanjem posvetovalne skupine oblikovali več priporočil glede metodologije spremljanja ter izvajanja politike strateškega grozdenja.

V tretjem projektu smo pripravili celovito metodologijo za testiranje povezanosti razvojno-raziskovalnih in inovacijskih (RRI) aktivnosti z različnimi spodbudami, ki jih v zvezi z izvajanjem aktivnosti RRI podjetjem ponuja država, ter povezanosti med z državnimi spodbudami podprtih izdatkov podjetij za RRI in uspešnostjo poslovanja podjetij (Burger et al. 2023). Cilj projekta je bila vzpostavitev spremljanja in presoje strategije pametne specializacije in ukrepov kohezijske politike na področju raziskav,

razvoja in spodbujanja inovativnosti.⁷ Prva faza je vključevala vzpostavitev ustrezne podatkovne osnove za spremljanje in presojo ukrepov z integracijo razpoložljivih nacionalnih zbirk podatkov ter identifikacijo ukrepov za evalvacijo učinkov, za katere smo oblikovali ustrezno zbirko, druga faza pa izbiro, testiranje in izvedbo primernih ekonometričnih metod za oceno kazalnega učinka ukrepov.⁸ Rezultati analize so pokazali velike in statistično značilne vplive ukrepov v okviru strategije pametne specializacije na ključne značilnosti podjetij, predvsem na človeške vire in inovacijsko aktivnost. Ukrepi so dosegali tudi opazne neposredne rezultate v procesnih, produktivnih in storitvenih inovacijah, vseh preučevanih oblikah inovacijskega sodelovanja⁹ ter porastu izdatkov za raziskave in razvoj. Ukrepi so vplivali na ključne značilnosti in neposredne rezultate tako na prednostni osi 1 kakor na prednostni osi 3. Na nobeni od osi pa ukrepi niso vplivali na investicije (glejte Burger et al. 2023). Prav tako ni bil identificiran vpliv na učinkovitost poslovanja, saj se pri kazalnikih produktivnosti ter pri profitnih maržah in maržah EBIT pojavljajo nihanja in ni bilo prisotnih statistično značilnih učinkov. Podjetja so prejela omejena in namenska sredstva, kar so kot omejitve za večje investicije navajala že v prejšnjih evalvacijah (Bučar, Črnigoj in Lipnik 2020; Bučar et al. 2022). Posredni učinki so se odrazili v bistveno manjši meri; delno v povečanju kapitala in plač, na prednostni osi 3 pa se je poleg vpliva ukrepov na plače pokazal še vpliv na prihodke od prodaje in izvoza ter vpliv na dodano vrednost. Končni rezultat tega vrednotenja je bila »Priprava metodologije za spremljanje vseh specifičnih ciljev Cilja politike 1 v VFO 2021–2027: Konkurenčnejša in pametnejša Evropa s spodbujanjem inovativne in pametne gospodarske preobrazbe ter regionalne povezanosti na področju IKT in specifičnih ciljev v drugih ciljih politikah Evropske kohezijske politike obdobja 2021–2027, ki so vezani na omogočitveni pogoj prenova Strategije pametne specializacije« (Burger et al. 2023). Vzpostavljena podatkovna zbirka in razvita metodologija omogočata sistematično spremljanje in vrednotenje.

V četrtem projektu pa je bila (po predhodno zagotovljenem večjem obsegu podatkov in testiranju različnih metod) pripravljena obsežnejša kvantitativna analiza uspešnosti SRIP in kompetenčnih centrov v obdobju kriz, ki je vključevala pregled in oceno delovanja različnih vrst partnerstev ter priporočila za spremembe in izboljšave (Burger et al. 2024). Rezultate tega vrednotenja SRIP kratko povzemamo v naslednjem razdelku.

USPEŠNOST STRATEŠKIH RAZVOJNO-INOVACIJSKIH PARTNERSTEV

V zadnji presoji SRIP smo naslovili sedem skupin ukrepov, to so spodbude za zagon inovativnih podjetij, spodbujanje izvajanja raziskovalno-razvojnih programov (TRL 3–6), spodbujanje izvajanja raziskovalno-razvojnih projektov (TRL 6–9), demo piloti, spodbujanje procesnih izboljšav, ukrepi na področju digitalizacije in e-poslovanja ter vavčerske spodbude. Učinke izbranih ukrepov smo presojali glede na cilje in pričakovane rezultate izbranih ukrepov prednostnih osi operativnega programa za izvajanje Evropske kohezijske politike v obdobju 2014–2020 (v nadaljevanju OP 2014–2020) in cilje strategije pametne specializacije (S4), kamor spadajo izbrani ukrepi. Nabor izbranih skupin ukrepov vključuje 20 ukrepov – javnih razpisov s skupno vrednostjo 746 milijonov evrov, od tega predstavlja znesek, upravičen do sofinanciranja, 379 milijonov evrov.

V okviru pričakovanih učinkov smo skladno z metodologijo, uporabljeno pri presoji učinkovitosti ukrepov strategije pametne specializacije, ki je bila izdelana v okviru projekta CRP »Priprava metodologije in presoja učinkovitosti ukrepov iz naslova strategije pametne specializacije« (Burger et al. 2023), najprej identificirali aktivnosti, neposredne rezultate in posredne rezultate skupin ukrepov, znotraj teh pa identificirali relevantne ciljne spremenljivke v podatkovnih zbirkah, s katerimi je mogoče meriti učinke. Skladno s prej omenjeno metodologijo smo učinke izbranih ukrepov ocenjevali za šest skupin ukrepov, sedmo skupino – vavčerske spodbude smo zaradi specifik tega ukrepa analizirali ločeno. Pri nekaterih ukrepih so pričakovani cilji kvantificirani. Pri teh smo doseganje ciljev presodili glede na zastavljene ciljne vrednosti, pri drugih smo analizirali učinke tako, da smo analizirali spremembe v identificiranih ciljnih spremenljivkah.

Najprej smo analizirali opisne statistike prejemnikov sredstev in jih primerjali z neprejemniki po posameznih skupinah ukrepov. Tabela 1 prikazuje povzetek analize opisnih statistik in razlik med prejemniki in neprejemniki sredstev po opisanih skupinah ukrepov.

Tabela 1: Povzetek analize opisnih statistik – razlike med prejemniki in neprejemniki sredstev

Kazalnik	Razlika med prejemniki in neprejemniki sredstev po skupinah ukrepov					
	Demo piloti	Digitalizacija in e-poslovanje	Procesne izboljšave	Spodbude za zagon	TRL 3–6	TRL 6–9
<i>Splošni kazalniki</i>						
Prihodki od prodaje	++	+	+	–	++	++
Vrednost izvoza	++	+	+	–	++	++
Delež prihodkov od prodaje, ustvarjen z izvozom	+	++	+	≈	++	++
Delež izvoznikov	++	++	++	+	++	++
Marža EBIT	–	++	+	≈	–	≈
Povprečno število zaposlenih	++	+	+	–	++	++
Število zaposlenih v R & R*	++	+	+	+	++	++
Dodana vrednost na zaposlenega	++	++	+	–	++	≈
R & R in inoviranje						
Inovacijska aktivnost	++	++	++	++	++	++
Procesno inoviranje	+	+	+	++	++	++
Inoviranje proizvodov/storitev	++	++	++	++	++	++

Sodelovanje v inovacijskem procesu z drugimi	++	+	+	++	++	++
Sodelovanje v inovacijskem procesu z JRO	++	+	+	–	++	++
Izdatki R & R	++	+	+	+	++	++
Delež izdatkov R & R v prihodkih od prodaje	++	+	+	+	++	++
Investicije						
Bruto investicije v osnovna sredstva	++	–	–	–	++	++
Bruto investicije v osnovna sredstva glede na vrednost sredstev	≈	≈	≈	≈	≈	≈
Investicije v neopredmetena sredstva	++	≈	≈	–	++	++
Investicije v neopredmetena sredstva glede na vrednost sredstev	+	+	+	+	+	+
Investicije v stroje in opremo	++	–	–	– –	++	++
Investicije v stroje in opremo glede na vrednost sredstev	–	≈	–	+	–	–
Lastniški kapital – osnovni kapital	++			–		++
Dolgoročni dolg	++			– –		++

Opomba: ++ označuje veliko večjo vrednost kazalnika pri prejemnikih sredstev, + označuje večjo vrednost kazalnika pri prejemnikih sredstev, – – označuje veliko manjšo vrednost kazalnika pri prejemnikih sredstev, – označuje manjšo vrednost kazalnika pri prejemnikih sredstev, ~ označuje nihanja oz. volatilnost kazalnika v proučevanem obdobju, ≈ označuje podobno vrednost. Vir: Črnigoj et al. 2024, 74.

Rezultati kažejo, da gre pri prejemnikih sredstev, razen pri skupini ukrepov spodbude za zagon inovativnih podjetij, za značilno večja podjetja v primerjavi z neprejemniki, ta so bolj izvozno usmerjena in ustvarjajo tudi višjo dodano vrednost na zaposlenega. Prejemniki sredstev so v primerjavi z neprejemniki bolj inovacijsko aktivni ter imajo tudi boljše izhodišča in pogoje za inoviranje – več zaposlenih v R & R, imajo višje izdatke za R & R in tudi več sodelujejo z drugimi pri izvajanju inovacijskih aktivnosti (pri teh kazalnikih to velja tudi za spodbude za zagon inovativnih podjetij). Prejemniki sredstev (brez upoštevanja velikosti) po večini izkazujejo tudi pozitivnejše trende pri večini proučevanih kazalnikov.

Nadaljevali smo z ocenjevanjem učinkov izbranih ukrepov. Z namenom zagotavljanja večje veljavnosti in zanesljivosti ocen smo jih proučili na več načinov, in sicer (i) z opazovanjem dinamike gibanja relevantnih ciljnih spremenljivk za prejemnike sredstev v obdobju pred in po prejemu spodbude ter (ii) z oceno kavalnih učinkov

ukrepov na poslovanje podjetij z dvema različnima metodama, pri čemer smo prejemnike sredstev in njihovo poslovanje primerjali z njim najpodobnejšimi podjetji, ki sredstev niso prejela (neprejemniki). Kazalnike oziroma ciljne spremenljivke smo tako kot pri opisnih statistikah razporedili v sklope: (i) splošne značilnosti podjetij, (ii) aktivnosti in neposredni učinki (učinki na neposredne rezultate), (iii) posredni učinki (učinki na posredne rezultate) in (iv) učinki spodbud na specifične aktivnosti in specifične rezultate.

Rezultati ocenjevanja učinkov izbranih ukrepov (tabela 2) kažejo, da so ukrepi rezultirali v več različnih pozitivnih učinkih v podjetjih, ki so prejela sredstva ukrepov. Pri skupini ukrepov demo piloti smo npr. ugotovili, da so prejemniki sredstev zaradi prejema sredstev povečali število zaposlenih, tudi zaposlenih v R & R, in postali bolj inovacijsko aktivni (slednji učinek sicer ni statistično značilen). Po prejemu spodbud smo v tej skupini ukrepov zaznali povečano inoviranje, tako procesno kot inoviranje proizvodov/storitev (učinka sicer nista statistično značilna). Zaznali smo tudi povečano sodelovanje v inovacijskem procesu in povečano vlaganje v R & R. Učinka pa v tej skupini ukrepov nismo zaznali pri investicijah v osnovna sredstva (po prejemu sredstev se v povprečju celo zmanjšajo). Rast obsega poslovanja, ki jo ugotavljamo pri prejemnikih sredstev te skupine ukrepov, očitno ne zahteva proporcionalnih investicijskih vlaganj v osnovna sredstva. Prav tako je bilo v tej skupini ukrepov zaznati več pozitivnih učinkov na posredne rezultate, povečajo se npr. prihodki od prodaje in dodana vrednost, pa tudi izvoz in delež izvoznikov, dodana vrednost na zaposlenega in marža iz poslovanja – EBIT (pri čemer ti učinki, razen učinkov na prihodke od prodaje in dodano vrednost niso statistično značilni).

Ugotavljamo tudi pozitiven učinek na sodelovanje z zunanjimi izvajalci R & R, pri čemer je to omejeno na izvajalce v Sloveniji, in povečano izvajanje vseh vrst raziskav (pri čemer učinek pri temeljnih raziskavah in eksperimentalnem razvoju ni statistično značilen). Zaznali smo tudi učinek na povečevanje deleža prodaje iz naslova novih ali izboljšanih proizvodov na trgu (učinki sicer niso statistično značilni). Prejemniki sredstev teh ukrepov pa zaradi prejema spodbud niso uspešnejši pri pridobivanju virov financiranja.

Pri ukrepih na področju digitalizacije in e-poslovanja smo ugotovili, da so prejemniki sredstev zaradi prejema sredstev povečali število zaposlenih, tudi zaposlenih v R & R. Nismo pa v tej skupini ukrepov zaznali povečanega inoviranja, kot tudi ne povečanega sodelovanja v inovacijskem procesu. Zaznali pa smo povečano vlaganje v R & R. Tudi v tej skupini ukrepov je bilo mogoče zaznati več pozitivnih učinkov na posredne rezultate, povečajo se npr. prihodki od prodaje, vrednost izvoza, dodana vrednost, kapital – opredmetena osnovna sredstva in povprečne plače. V tej skupini ukrepov pa nismo potrdili pozitivnega učinka na dodano vrednost na zaposlenega. Pri učinkih na specifične aktivnosti in specifične rezultate ugotavljamo tudi pozitiven učinek na sodelovanje z zunanjimi izvajalci R & R, pri čemer je to v večjem delu omejeno na izvajalce doma (ti učinki sicer niso statistično značilni). Zaznavamo tudi povečano izvajanje vseh vrst raziskav (pri čemer učinek pri temeljnih raziskavah in aplikativnih raziskavah ni statistično značilen).

Pri skupini ukrepov spodbujanje procesnih izboljšav smo ugotovili, da so prejemniki sredstev zaradi prejema sredstev povečali število zaposlenih, ne pa tudi

zaposlenih v R & R, in postali bolj inovacijsko aktivni. Po prejemu spodbud smo v tej skupini ukrepov zaznali povečano inoviranje proizvodov/storitev (ne pa procesnega inoviranja). Prav tako tu nismo zaznali povečanega sodelovanja v inovacijskem procesu in povečanega vlaganja v R & R. Tako kot pri drugih skupinah ukrepov tudi v tej skupini ukrepov učinka nismo zaznali pri investicijah v osnovna sredstva. Smo pa spet zaznali več pozitivnih učinkov na posredne rezultate, povečajo se npr. prihodki od prodaje, dodana vrednost, marža iz poslovanja – marža EBIT in kapital – opredmetena osnovna sredstva, pa tudi delež izvoznikov, dodana vrednost na zaposlenega in profitna marža (pri čemer so statistično značilni le učinki na prihodke od prodaje, dodano vrednost, maržo iz poslovanja – EBIT in kapital – opredmetena osnovna sredstva). Viden je tudi pozitiven učinek na sodelovanje z zunanjimi izvajalci R & R, pri čemer je v tej skupini zaznati tudi učinek na sodelovanje z izvajalci v tujini (sicer učinek ni statistično značilen).

Pri spodbudah za zagon inovativnih podjetij smo nadalje ugotovili, da so prejemniki sredstev zaradi prejema sredstev povečali število zaposlenih (sicer učinek ni statistično značilen), ne pa tudi zaposlenih v R & R. V tej skupini ukrepov zaradi premajhnega števila opazovanj nismo uspeli oceniti učinkov na inoviranje in tudi investicije v osnovna sredstva, tako da o teh učinkih ne moremo sklepati. Kljub pričakovanju, da v tej skupini ne bomo še identificirali pomembnejših učinkov na posredne rezultate, ki se pokažejo z odlogom, pa so se ti kljub temu (sicer, z izjemo povprečne plače, vsi statistično neznačilni) pokazali. Prejemniki sredstev so po prejemu sredstev povečali prihodke od prodaje, vrednost izvoza, povečal se je tudi delež izvoznikov, dodana vrednost, dodana vrednost na zaposlenega, povečala pa se je tudi povprečna plača pri prejemnikih. Pri učinkih na specifične aktivnosti in specifične rezultate ugotavljamo tudi pozitiven učinek na sodelovanje z zunanjimi izvajalci R & R, povečano izvajanje vseh temeljnih raziskav in eksperimentalnega razvoja, ne pa tudi aplikativnih raziskav (pri čemer učinka nista statistično značilna). Zaradi omejitev v razpoložljivosti podatkov nismo uspeli oceniti deleža prodaje iz naslova novih ali izboljšanih proizvodov na trgu, so pa prejemniki sredstev teh ukrepov zaradi prejema sredstev bili uspešnejši pri pridobivanju virov financiranja.

Tabela 2: Povzetek ocenjenih učinkov ukrepov, ocenjenih z metodo dvojnih fiksnih učinkov – TWFE¹⁰

KAZALNIK	Demo piloti	Digital	Procesne izboljšave	Spodbude za zagon	TRL 3–6	TRL 6–9
<i>Učinki na ključne značilnosti podjetij</i>						
Povprečno število zaposlenih	+	+	+	(+)	+	+
Število zaposlenih v R & R*	+	+	~	~	+	~
Inovacijska aktivnost*	(+)	~	+	n. p.	~	~
<i>Učinki na ne-posredne rezultate*</i>						
Procesno inoviranje	(+)	0	0	n. p.	~	(+)

Inoviranje proizvodov/storitev	(+)	~	(+)	n. p.	~	~
Sodelovanje v inovacijskem procesu z drugimi	+	~	0	n. p.	+	(+)
Sodelovanje v inovacijskem procesu z JRO	+	~	0	n. p.	(+)	(+)
Izdatki za R & R	+	+	~	~	+	+
Bruto investicije v osnovna sredstva	–	0	0	n. p.	0	0
Investicije v neopredmetena sredstva	0	~	~	n. p.	0	~
Investicije v stroje in opremo	–	~	0	n. p.	0	–
<i>Učinki na posredne rezultate</i>						
Prihodki od prodaje	+	+	+	(+)	(+)	+
Vrednost izvoza	(+)	+	~	(+)	~	+
Delež izvoznikov	(+)	~	(+)	(+)	~	~
Dodana vrednost	+	+	+	(+)	~	+
Dodana vrednost na zaposlenega	(+)	0	(+)	(+)	0	0
Marža EBIT	(+)	~	+	~	~	~
Profitna marža	~	~	(+)	~	0	(+)
Kapital – opredmetena osnovna sredstva	(+)	+	+	~	(+)	~
Povprečna plača	(+)	+	~	+	(+)	+
Izdatki za eksterni R & R	+	(+)	+	(+)	0	~
Izdatki za eksterni R & R – Slovenija	+	(+)	+	0	–	~
Izdatki za eksterni R & R – tujina	0	~	(+)	~	0	0
Izdatki za temeljne raziskave	(+)	(+)	0	(+)	~	~
Izdatki za eksperimentalni razvoj	(+)	+	+	(+)	+	+
Izdatki za aplikativne raziskave	+	(+)	~	0	~	+
Delež podjetij, ki zaposluje strokovnjake za IKT	(+)		0		~	(+)
Delež podjetij z izobraževanjem zaposlenih na področju IKT	0		0		~	–
Delež prodaje – novi proizvod	(+)			n. p.	(+)	(+)
Delež prodaje – novi ali izboljšani proizvod	(+)			n. p.	(+)	+

Opomba: 0 označuje odsotnost učinka; + označuje stalno rastoč trend gibanja kazalnika v proučevanem obdobju; ~ prikazuje nihanje oz. volatilitnost kazalnika v proučevanem ob-

dobju; – prikazuje padajoč trend gibanja kazalnika v proučevanem obdobju; n. p. pomeni, da ni podatka; () rezultati niso statistično značilni.

* Kazalnika število zaposlenih v R & R in inovacijska aktivnost ter kazalniki v okviru učinkov na neposredne rezultate niso podrobneje opisani in prikazani na slikah v prejšnjih poglavjih, saj analize temeljijo na SURS-ovih zbirkah podatkov, do katerih je omogočen oddaljeni dostop. Vir: Črnigoj et al. 2024, 74.

Pri SRIP tako ugotavljamo pomembne pozitivne učinke sodelovanja v partnerstvih na podlagi različnih skupin ukrepov. V partnerstvih sodelujoča podjetja so praviloma povečala število zaposlenih, nekatera tudi število zaposlenih v R & R, izdatke za R & R, postala so bolj inovacijsko aktivna in povečala sodelovanje z drugimi organizacijami v inovacijskem procesu. V sodelujočih podjetjih so se povečali prihodki od prodaje in vrednost izvoza. Po drugi strani ugotavljamo, da sodelovanje v partnerstvih ni rezultiralo v višjih investicijah v osnovna sredstva in tudi ne v povečanju profitabilnosti.

Dodatna ugotovitev je, da je učinek sodelovanja v partnerstvih največji na ravni aktivnosti oziroma na ravni inovacijskih inputov v inovacijski proces, manjši na ravni neposrednih rezultatov oziroma ravni inovacijskih outputov. Še manjši (oziroma učinka ni zaznati) pa je na ravni posrednih rezultatov oziroma na ravni produktivnosti. S kvantitativnimi metodami smo ugotovili tudi pozitivne horizontalne učinke prelivanja na nečlane partnerstev v isti dejavnosti (Črnigoj et al. 2024).

ZAKLJUČEK

Čeprav je pogosto videna predvsem kot orodje tehnološke preobrazbe, je pametna specializacija v jedru družbeni izziv, ker zahteva celovito sodelovanje deležnikov in obravnava širše družbene transformacije, ki presegajo tehnološke inovacije. Lahko jo opredelimo kot politiko s posebno misijo, ker poudarja procese sodelovanja, vzajemno učenje in oblikovanje politik na podlagi empiričnih dokazov in stalnega spremljanja doseganja preobrazbenih ciljev. Pametna specializacija terja tehnično strokovno znanje (epistemično) in aktivno sodelovanje več družbenih akterjev, vključno s podjetniki, centri znanja, z državljani, lokalno, regionalno in nacionalno upravo, saj cilja na sistemsko družbeno transformacijo. Pristop presega zgolj tehnološke rešitve z dajanjem prednosti temeljnim sistemskim spremembam, obravnavanjem velikih družbenih izzivov in ustvarjanjem smiselnih inovacij, ki pozitivno prispevajo k družbenim potrebam. Strategija zahteva sodelovanje na več ravneh, uskladitev politik s sistemskimi cilji in ima širši družbeni vpliv. Ker uveljavlja pristop od spodaj navzgor, je nujno usklajevanje različnih interesov in pogledov ter ustvarjanje podpore in predanosti v različnih družbenih skupinah. Pametna specializacija tudi izpostavlja, da pri inovacijah ne gre le za tehnološki napredek, ampak za reševanje kompleksnih družbenih problemov, spodbujanje podjetniške dinamike, podpore regionalni gospodarski rasti in ustvarjanje trajnostnih delovnih mest. Z vključevanjem socialnih vidikov in s širokim sodelovanjem različnih deležnikov pametna specializacija preoblikuje inovacijsko politiko iz tehnološke naloge v celovito socialno strategijo za regionalni razvoj.

Evalvacije implementacije pametne specializacije v Sloveniji (in drugje v Evropi) v ta vidik transformacije sicer delno zajamejo, a razpoložljivi podatki to omogočajo v (pre)skromnem obsegu. Pogojevanje financiranja s kohezijskimi sredstvi terja oprijemljive rezultate uvajanja teh instrumentov, zato oblikovalci politik v vseh fazah evalvacije preferirajo kvantitativne ocene učinkov, ki jih je mogoče sistematično in sproti obnavljati, ponavljati in preverjati.

Dozdajšnje analize so pokazale številne pozitivne učinke uvajanja S4. Kvantitativni kazalniki in tržne manifestacije kažejo pozitivne spremembe na več področjih (predvsem v zaposlenosti, zaposlenosti v R & R ter prodaji). Učinki so pogosto najintenzivnejši v prvih letih po uvedbi instrumentov. Srednjeročno in v daljšem obdobju kvantitativni učinki pogosto izzvenijo, med drugim tudi zaradi nizkih vrednosti spodbud in nepredvidljivih časovnic financiranja (zamiki razpisov), kar sproža tveganje, da novozačetih aktivnosti podjetja ne nadaljujejo po prvotno zastavljenih načrtih. SRIP so finančno odvisni od teh spodbud, soočajo se s težavami pri upravljanju in ohranjanju ter spreminjanju vzpostavljenih načinov sodelovanja. Kakovost večnivojskega upravljanja deležniki ocenjujejo kot izziv in pripombe SRIP so bile, da so kot ključni deležnik inovacijskega ekosistema (pre)skromno vpeti v odločanje in upravljanje. Ne glede na navedeno pa so bili opaženi opazni premiki v načinih in pogostosti sodelovanja in povezovanja. Osredinjenost evalvacij na kvantitativne pristope glede na uvide iz dozdajšnjih vrednotenj podceni pomen SRIP in obseg družbenih sprememb zaradi S4 (S5).

Kljub zavedanju, da je uvedba pametne specializacija družbeni izziv, spremljanje in vrednotenje inovacijskih spodbud skromno (nesistematično) spremlja spremembe družbene pripravljenosti na inovacije. Kvalitativna analiza (predvsem v okviru prvih treh projektov) je pokazala izboljšave v številnih elementih, ki so bistveni za spodbujanje družbene pripravljenosti na inovacije, in sicer predvsem povečano mreženje, zaupanje (tudi med tekmeci), (interdisciplinarno) sodelovanje in učenje med mrežami (SRIP, KOC, CO) in uvajanje številnih družbenih eksperimentov.

Podjetja v SRIP-ih, ki so razvila te elemente, so pokazala tudi večjo odpornost v krizah (kar je posredno vidno v stabilnih ali celo rastočih prihodkih in zaposlovanju v obdobju kriz), večje produktne in procesne inovacije, vstopa na nove trge, premike v pozicijah v globalnih verigah vrednosti ter uvajanju drugih prebojnih načinov vodenja podjetij in mrež (npr. intenzivna digitalizacija procesov, uporaba umetne inteligence, nadgradnje sistemov nagrajevanja/zadržanja talentov ipd.). Preučevanje uspešnih primerov kaže, da so spreminjanje inovacijske kulture in upravljanje mrež ter sodelovanje na različnih ravneh, tista socialna tehnologija inoviranja, ki bistveno vpliva na končni rezultat, zato ji je treba nameniti več pozornosti že med načrtovanjem, in ne le med vrednotenjem.

OPOMBE

- 1 Raziskovanje poteka v okviru raziskovalnega programa št. P5-0177, ki ga je sofinancirala Javna agencija za raziskovalno dejavnost Republike Slovenije iz državnega proračuna. Prispevek je del aktivnosti, ki se izvajajo v okviru projekta GREENTECH, ki ga sofinancira Evropska unija – NextGenerationEU.

- 2 V ZDA so v obdobju 2022–2025 taki primeri, denimo, Chips and Science Act, Inflation Reduction Act ter iniciative za ladjedelništvo, na Kitajskem je tak primer Made in China 2025, v Južni Koreji pa Chipmaker Support Package iz leta 2024.
- 3 Glejte npr. https://ec.europa.eu/regional_policy/whats-new/panorama/2023/11/29-11-2023-revolutionizing-growth-eu-s-smart-specialisation-strategies-unleash-regional-powerhouses_en.
- 4 Oblikovala so se naslednja SRIP: Pametna mesta in skupnosti, Pametne stavbe in dom z lesno verigo, Mreže za prehod v krožno gospodarstvo, Trajnostna pridelava hrane, Trajnostni turizem, Tovarne prihodnosti, Zdravje – medicina, Mobilnost, Razvoj materialov kot produktov.
- 5 Ciljni raziskovalni projekt: Analitične podlage za revizijo S4 v letu 2018 (V5-1645).
- 6 Stopnje družbene pripravljenosti (SRL) zagotavljajo okvir za oceno pripravljenosti družbe na sprejetje določene inovacije, bodisi tehnološke bodisi družbene. Načela je leta 2018 predstavil danski inovacijski sklad (Innovation Fund Denmark, glejte npr. https://innovationsfonden.dk/sites/default/files/2019-03/societal_readiness_levels_-_srl.pdf). Gre za celostno oceno, ki upošteva etične, pravne, družbene in ekonomske dejavnike. Stopnje SRL segajo od 1 (začetna ideja) do 9 (popolnoma integrirana v družbo).
- 7 Družbene inovacije niso bile neposredno vključene v evalvacijo, se je pa v kvalitativnih raziskavah v prvih dveh omenjenih projektih pokazalo nekaj novih form in načinov inovacijskega sodelovanja.
- 8 Da bi lahko primerjali stanja in rasti indikatorjev med prejemniki in neprejemniki programa, je treba izvesti statistične metode, ki popravijo pristranosti zaradi nenaključne selekcije, šibkega prekrivanja v porazdelitvah kontrolnih spremenljivk in selekcije na podlagi neopazljivih dejavnikov.
- 9 Preučevano je sodelovanje znotraj podjetij, sodelovanje s kupci, z dobavitelji, javnimi in zasebnimi raziskovalnimi organizacijami, ter tudi geografska dimenzija (nacionalno in mednarodno sodelovanje).
- 10 Čeprav se po posameznih kazalnikih pojavljajo bolj ali manj velike razlike v vrednostih, teh ne razlikujemo na ++ in + ali – in – kot v tabeli 5.1, saj je nemogoče določiti univerzalne mejne vrednosti za celoten nabor kazalnikov.

CITIRANA LITERATURA

Bučar, Maja, Matjaž. Črnigoj in Aleš Lipnik. 2020. *Vrednotenje sodelovanja med znanostjo in gospodarstvom*. Ljubljana: Založba FDV.

Bučar, Maja, Matjaž Črnigoj, Andreja Jaklič, Aleš Lipnik, Marko Lovec, Metka Stare, Boštjan Udovič, Peter Štrukelj in Borut Likar (ur.). 2020. *Predlogi nadgradnje raziskovalno - razvojnega in inovacijskega sistema v smeri spodbujanja sodelovanja med SRIPi in javnimi raziskovalnimi organizacijami: projekt CRP „Spremljanje in presoja učinkov strateških razvojno inovacijskih partnerstev“*. Ljubljana: Fakulteta za družbene vede.

Bučar, Maja, Matjaž Črnigoj, Aleš Lipnik, Andreja Jaklič, Metka Stare, Boštjan Udovič, Borut. Likar in Peter Štrukelj. 2022. *Vrednotenje delovanja SRIP-ov v obdobju 2017–2021*. Ljubljana: Univerza v Ljubljani, Fakulteta za družbene vede.

Bučar, Maja. 2024. »Significant Improvement but also Missed Opportunities: Slovenian Innovation Policy from 2004 to 2024«. V *Slovenia and the EU: 20 Years of Membership in Perspective*, ur. Michael Kaeding in Boštjan Udovič, 33–39. 1. izd. Cham: Springer. The Future of Europe.

Burger, Anže, Matjaž Črnigoj, Andreja Jaklič, Barbara Kalar in Iris Koleša. 2023. *Priprava metodologije za spremljanje vseh specifičnih ciljev Cilja politike 1 v VFO 2021–2027: Konkurenčnejša*

in pametnejša Evropa s spodbujanjem inovativne in pametne gospodarske preobrazbe ter regionalne poveztivosti na področju IKT in specifičnih ciljev v drugih ciljih politik Evropske kohezijske politike obdobja 2021–2027, ki so vezani na omogočitveni pogoj prenova Strategije pametne specializacije. Ljubljana: Center za mednarodne odnose na Fakulteti za družbene vede Univerze v Ljubljani in Inštitut za ekonomska raziskovanja.

Burger, Anže, Matjaž Črnigoj, Andreja Jaklič, Barbara Kalar in Iris Koleša. 2024. *Vrednotenje ukrepov industrijske, znanstvene in inovacijske politike z vidika povečanja dodane vrednosti v podjetjih. Končno poročilo pripravljeno za projekt »Analiza uspešnosti Strateških razvojno-inovacijskih partnerstev (SRIP) in Kompetenčnih centrov. Pregled in ocena delovanja ter priporočila za spremembe in izboljšave.* Ljubljana: Center za mednarodne odnose na Fakulteti za družbene vede Univerze v Ljubljani in Inštitut za ekonomska raziskovanja. November 2024.

Carayannis, Elias G., Thorsten D. Barth in David FJ Campbell. 2012. »The Quintuple Helix Innovation Model: Global Warming as a Challenge and Driver for Innovation«. *J Innov Entrep* 1, 2. <https://doi.org/10.1186/2192-5372-1-2>.

Črnigoj, Matjaž, Barbara Kalar, Andreja Jaklič, Anže Burger, Maja Bučar, Metka Stare, Boštjan Udovič, Marko Lovec, Borut Likar, Peter Štrukelj in Aleš Lipnik. 2022. *Kvantitativna analiza strateških razvojno-inovacijskih partnerstev: projekt: CRP, 'Spremljanje in presoja učinkov strateških razvojno inovacijskih partnerstev'.* Ljubljana: Inštitut za ekonomska raziskovanja: Fakulteta za družbene vede.

Črnigoj, Matjaž, Barbara Kalar, Anže Burger in Andreja Jaklič. 2024. *Presoja učinkovitosti izbranih ukrepov strategije pametne specializacije: končno poročilo.* Ljubljana: Center za mednarodne odnose na Fakulteti za družbene vede Univerze v Ljubljani in Inštitut za ekonomska raziskovanja.

Edler, Jakob, in Wouter P. Boon. 2018. »'The Next Generation of Innovation Policy: Directionality and the Role of Demand-Oriented Instruments'—Introduction to the Special Section«. *Science and Public Policy* 45 (4): 433–434. <https://doi.org/10.1093/scipol/scy026>.

European Commission, 2010. Europe 2020: A Strategy for Smart, Sustainable and Inclusive Growth. COM(2010) 2020 final.

ERAC (2010): SLOVENIA: A Report of the ERAC Policy Mix Expert Group Fifth Cycle of the Open Method of Coordination, Dec. 2010. Available at http://www.era.gv.at/attach/Item4.1Slovenia_OMC_Report_FinalDec20.pdf (15 February 2015).

Gianelle, Carlo, Fabrizio Guzzo in Krzysztof Mieszkowski. 2019. »Smart Specialisation: What Gets Lost in Translation from Concept to Practice?«. *Regional Studies* 54 (10): 1377–1388. <https://doi.org/10.1080/00343404.2019.1607970>.

Henrekson, Magnus, Christian Sandström in Mikael Stenkula. 2024. *Moonshots and the New Industrial Policy. Questioning the Mission Economy.* Springer. International Studies in Entrepreneurship (ISEN, vol. 56).

Juhász, Réka, Nathan J. Lane in Dani Rodrik, D. 2023. »The New Economics Of Industrial Policy«. NBER Working Paper No. 31538. Cambridge, MA: National Bureau of Economic Research.

Mazzucato, Mariana. 2013. *The Entrepreneurial State: Debunking Public vs. Private Sector Myths.* ANTHEM PRESS.

Mazzucato, Mariana. 2018. »Mission-Oriented Innovation Policies: Challenges and Opportunities«. *Industrial and Corporate Change* 27 (5): 803–815. <https://doi.org/10.1093/icc/dty034>.

Mazzucato, M. 2019. *Governing Missions in the EU. Independent Expert Report.* DOI:10.2777/618697.

Muldoon, Jeff, in Derek K. Yonai. 2023. »A Wrong But Seductive Idea: Public Choice And The Entrepreneurial State«. *Journal of the International Council for Small Business* 4 (4): 351–361.

MVZT (2011). Resolucija o raziskovalni in inovacijski strategiji Slovenije 2011–2020 (Re-RIS11-20). <https://www.uradni-list.si/glasilo-uradni-list-rs/vsebina/2011-01-2045>.

OECD (2012), OECD Reviews of Innovation Policy: Slovenia 2012, OECD Publishing. <http://dx.doi.org/10.1787/9789264167407-en>

OECD. 2021. *The Design And Implementation of Mission-Oriented Innovation Policies: A New Systemic Policy Approach To Address Societal Challenges*. OECD Science, Technology and Industry Policy Paper No. 100. Pariz: Organisation for Economic Co-operation and Development.

UMAR. 2022 Poročilo o produktivnosti 2021. Urad RS za makroekonomske analize in razvoj, Ljubljana.

Ranga, Marina, in So Young Kim. 2023. *Editorial: Next-Generation Innovation Policies: Promoting Systemic Socio-Economic Transformative Change*. *Frontiers in Research Metrics and Analytics*. DOI: 10.3389/frma.2023.1146039.

Reid, Alasdair, Fred Steward in Michal Miedzinski. 2023. *Aligning Smart Specialisation with Transformative Innovation Policy. Lessons For Implementing Challenge-Led Missions In Smart Specialisation*. Luksemburg: Publications Office of the European Union. DOI:10.2760/359295, JRC134466.

Rodrik, Dani. 2022. *An Industrial Policy for Good Jobs*. Hamilton Project – Policy Proposal. Washington, DC: Brookings Institution.

Slovenska strategija trajnostne pametne specializacije. S5. 2023. <https://www.gov.si/zbirke/projekti-in-programi/slovenska-strategija-trajnostne-pametne-specializacije-s5/>.

Tagliapietra, Simone, in Reinhilde Veugelers (ur.). 2023. *Sparking Europe's New Industrial Revolution: A Policy for Net Zero, Growth and Resilience*. Bruegel Blueprint Series No. 33. Bruselj: Bruegel.

K TRAJNOSTNEMU RAZVOJU PLATFORMNE EKONOMIJE: IZKUŠNJE Z GLOBALNEGA JUGA

Aleksandra Kanjuo Mrčela

»Popotujoči ljudje ne postanejo modrejši zgolj zato, ker nenehno izkušajo nove razglede in dogodke, ampak tudi zato, ker sami zase postajajo mimo hiteča krajina, na katero je mogoče pogledati s tolažilne razdalje. Takrat opažajo več.«

Olga Tokarczuk

UVOD

*Platformna ali gig ekonomija*¹ je naraščajoči del svetovnega gospodarstva, ki temelji na uporabi novih digitalnih tehnologij in poslovnem modelu, ki, glede na obstoječe podatke, povečuje družbene neenakosti in koncentrira ekonomsko moč. Čeprav se je začela v storitvenem delu ekonomije (npr. prevozi, kot je Uber, dostava hrane, kot je Volt, trgovini, kot je Amazon), se razširja na organizacijo dela vse večjega števila delavcev in delavk, tako na spletu kot v lokalnih okoljih ter tudi v drugih delih gospodarstva in negospodarskih dejavnostih (npr. intelektualnih storitvah in pomoči na domu).

Narašča število delavcev, ki pri opravljanju dela na platformah nimajo delavskega statusa in s tem povezanih delavskih pravic, izpostavljeni so finančnemu in socialnemu tveganju, slabim delovnim razmeram in algoritemskemu nadzoru.² Po drugi strani, in tudi zaradi tega, pa majhno število ekonomskih akterjev pridobiva ogromne dobičke, nadzoruje algoritemsko obdelovane podatke in tako povečuje svojo ekonomsko moč. Opažamo vse večjo »gigifikacijo« (razdrobljenost) celotne ekonomije. Razdrobljena ekonomija je opredeljena kot »ureditev gospodarskega sistema in trga delovne sile, za katerega so značilne kratkoročne pogodbe ali delo samostojnih podjetnikov« (Kanjuo Mrčela in Kohont 2024, 300). Gigifikacija pomeni razgradnjo, razdrobitev institucij ter individualizacijo dela in življenja. Delovni procesi se drobijo na manjše naloge in delavci nad tem nimajo nadzora. Drobijo se tudi karijerne poti, profesionalne identitete; razkraja, atomizira in drobi se delavski razred – na nemočne posameznike, ki se ne morejo učinkovito upirati sistemu, ki jih pušča v ekonomski in identitetni negotovosti. Gigifikacijo razumemo kot posledico procesa drobljenja oz. razkrajanja prejšnje gospodarske ureditve, ki je bila stabilnejša in predvidljivejša. Na to opozarjajo in se zdajšnjemu stanju upirajo že mnogi ekonomski akterji po svetu. Poznamo tudi že nekaj primerov demokratično organiziranih platform (t. i. *platformnih kooperativ*).

Po študijskem obisku v Čilu³ v prispevku predstavljamo nekatere teoretske prispevke in empirične podatke o tem, kako v gospodarstvu te latinskoameriške države, podobno kot drugod po svetu, globalni trend gigifikacije prinaša negativne spremembe, pa tudi razmislek in poskuse demokratizacije platformne ekonomije. Izkušnja z globalnega juga je koristna, ker prispeva tako k analizi univerzalnih kot tudi partikularnih okoliščin in dejavnikov ekonomskega razvoja. Čilska izkušnja s platformnim gospodarstvom je lahko za Evropo in Slovenijo zanimiva iz več razlogov, saj omogoča vpogled v delovanje digitalne ekonomije v kontekstu večje ekonomske neenakosti, manj reguliranega trga delovne sile in hitrega digitalnega razvoja. Za evropsko razpravo o prihodnosti dela in regulacije platform je zanimivo analizirati primere delavskega boja za pravice, ker se pokaže, da tudi v dereguliranih sistemih obstaja prostor za samoorganizacijo. Primer Čila je lahko tudi opozorilo, kaj se zgodi brez aktivne državne politike. V Čilu so platformna dela pogosto edina dostopna oblika zaposlitve za mlade, migrante in ljudi iz nižjih socialnoekonomskih slojev. Tudi pri nas so platformna dela vstopna točka na trg delovne sile za ranljive skupine – zlasti študente, prekarne delavce in migrante. Čilska izkušnja opozarja, da lahko brez varovalk te skupine ostanejo trajno marginalizirane.

V besedilu opozarjamo na hkratni obstoj trajnostnemu razvoju izjemno nenaklonjenih gospodarskih praks, ki so v razmahu, ter normativne podpore in delne realizacije tem praksam popolnoma nasprotnih konceptov in modelov gospodarjenja.

Besedilo nastaja med dvema kontinentoma na začetku leta 2025, ki ga je OZN označila za mednarodno leto zadrug pod sloganom »Zadruge gradijo boljši svet«. Združeni narodi s tem prepoznavajo pozitiven vpliv ene oblike ekonomske demokratizacije in opolnomočenja delavcev – zadružništva – za dobrobit in prosperiteto ljudi povsod po svetu in v ospredje postavljajo zadružni model gospodarjenja kot rešitev za premagovanje številnih svetovnih izzivov ter pri pospeševanju prizadevanj za uresničitev ciljev trajnostnega razvoja (OZN 2015) do leta 2030. V besedilu bomo, čeprav so za razmislek o delu v platformni ekonomiji med najpomembnejšimi nekateri od 17 ciljev OZN (kot so 1. ukinitev revščine, 5. enakost spolov, 8. spodobno delo in vključujoča gospodarska rast ter 10. zmanjševanje neenakosti), imeli v mislih najširši namen vseh 17 ciljev: »mir in blaginjo za ljudi in planet«.

Prispevek naslavlja razmerje med platformno ekonomijo in trajnostnim razvojem ter opozarja na ključne razlike med njenimi netrajnostnimi in trajnostnimi oblikami. Netrajnostna platformna ekonomija temelji na prekarizaciji dela, koncentraciji podatkovne in ekonomske moči, netransparentnem algoritemskem nadzoru ter izključujočih poslovnih modelih. Nasprotno pa naj bi trajnostna platformna ekonomija stremela k demokratičnemu upravljanju, okoljski odgovornosti, h kakovostnemu zaposlovanju in k lokalni vpetosti gospodarskih subjektov. Alternativne digitalne platforme, kot so npr. platformne zadruge, prispevajo k zmanjševanju globalnih ekonomskih neenakosti in omogočajo redistributivne pristope, ki so v skladu s cilji trajnostnega razvoja. Prispevek tako umešča vprašanje platformne ekonomije v širši okvir prizadevanj za pravično, vključujočo in okolju prijazno digitalno prihodnost.

Na začetku prispevka na kratko predstavljamo naraščajoči pomen platformnega gospodarstva po svetu in do zdaj uporabljene načine upora proti negativnim pos-

ledicam rasti platformno organiziranega dela. Potem se osredinjamo na kontekst, značilnosti in alternative konvencionalno organiziranim platformam v Čilu.

Prispevek zaključujemo z ugotovitvami in s priporočili, s katerimi želimo prispevati k razmisleku o potencialu in omejitvah platformno organiziranega gospodarstva pri doseganju ciljev trajnostnega razvoja.

PLATFORMNA EKONOMIJA: NARAŠČAJOČI POMEN V SODOBNEM GLOBALNEM GOSPODARSTVU

Že več kot desetletje, pospešeno pa po pandemiji covida-19, je po vsem svetu viden razvoj platformnega gospodarstva, ki temelji na digitalnih platformah, ki posredujejo delo med delavci in naročniki. Najvidnejši delavci, katerih delo se organizira, ponuja in nadzoruje na platformah, so vozniki (npr. Uberja) in dostavljavci hrane (npr. Glovo ali Volt), vendar je še zelo veliko manj vidnih delavcev in delavk (npr. mikrodelaavcev na platformah, kot je Amazon Mechanical Turk), ki opravljajo različna dela in storitve, organizirane prek platform dela. Poslovni model, na katerem temelji delo zdaj prevladujočega platformnega gospodarstva, se razširja (npr. v bančništvu, proizvodnji in logistiki) ter se lahko uporablja tudi v drugih gospodarskih in drugih dejavnostih. Tovrstna »gigifikacija« (razdrobljenost) ekonomije se odraža v večji prožnosti ter kratkoročnih in pogosto slabo plačanih delovnih nalogah, ki so algoritemsko organizirane in nadzorovane ter bolj koristijo platformam kot delavcem ali uporabnikom.

Pomembna dimenzija nenadzorovanega razvoja platformne ekonomije je njegova neskladnost s cilji trajnostnega ekonomskega razvoja, prijaznega ljudem in naravi. Med osnovnimi značilnostmi zdajšnje ureditve platformnega gospodarstva s posledicami za delavce in celoten gospodarski sistem so (Kanjuo Mrčela 2024):

- prekarizacija dela, ki se odraža v tem, da se delavcem odreka status zaposlenega, kar pomeni, da nimajo delavskih pravic, pravice in možnosti kolektivnega zastopanja interesov, socialnih ugodnosti in varnosti zaposlitve oz. zaposljivosti. Podjetizacija delovne sile (navidezno samostojno podjetništvo) je norma, v skladu s katero se izjemnost podjetniškega talenta in prevzem tveganja za poslovni uspeh pričakujeta od vsakogar, ki želi imeti dohodek za preživetje;
- algoritmizacija organizacije in nadzora dela, ki vodi v intenzifikacijo dela, zmanjšano avtonomijo delavcev in večjo psihosocialno obremenitev, po drugi strani pa prispeva k racionalizaciji in pocenitvi delovnih procesov, ter posledično
- koncentracija ekonomske moči pri lastnikih velikih digitalnih platform, kot so Uber, Amazon in druge, ki ustvarjajo monopolne strukture, v katerih delavci nimajo pogajalske moči, konkurenti nimajo možnosti tekmovanja, institucionalni okvir pa zgublja možnost oblikovanja ter učinkovitega izvajanja gospodarskih in socialnih politik.

Negativne posledice dozdejšnjega razvoja platformnega gospodarstva so že povzročile odzive, od individualnih in kolektivnih uporov ter protestov do oblikovanja platformnim podjetjem alternativnih ureditev. Med temi odzivi so najpomembnejši:

- stavke in drugi kolektivni protesti,
- pravni boji in regulacija platformnega gospodarstva,
- algoaktivizem⁴, večinoma na družbenih omrežjih,
- alternativne demokratično organizirane platforme,
- radikalni tehnološki upor.

Platformni delavci organizirajo stavke in proteste (npr. dostavljavci hrane v Latinski Ameriki in Evropi), da bi izboljšali svoje delovne pogoje. V nekaterih državah sodišča in zakonodajalci delavcem priznavajo pravice (npr. rzsodba britanskega vrhovnega sodišča v prid Uberjevim delavcem, evropska direktiva o platformnem delu, ki zahteva definiranje delovnega razmerja platformnih delavcev, večjo preglednost poslovanja platform in algoritmov ter obveznost človeškega nadzora nad njimi, in čilska nova regulacija platformnega dela). Delavci razvijajo tudi orodja za boj proti nepravilnim algoritmom, kot je npr. aplikacija UberCheats, ki razkriva napačne izračune plačil pri dostavah Uber Eats. Obstajajo pobude za oblikovanje demokratičnih, zadružnih platform, kot je Fairbnb.coop, ki deluje po principu pravične delitve dobička in skupnostne odgovornosti. Take alternative se pogosto soočajo s težavami pri pridobivanju sredstev in konkurenčnosti korporativnim platformam. Nekateri teoretiki in aktivisti zagovarjajo »zaviralno politiko«, ki zavrača tehnološki determinizem in spodbuja upočasnitev tehnološkega razvoja, kadar vodi v poslabšanje delovnih pogojev (več v Kanjuo Mrčela 2024).

V nadaljevanju predstavljamo analizo platformnega gospodarstva v Čilu.

KONTEKST RAZVOJA PLATFORMNE EKONOMIJE V ČILU

Če upoštevamo tako gospodarsko razvitost (BDP na prebivalca 16.365 USD) kot splošni človekov razvoj (HDI 0,86), lahko rečemo, da je gospodarstvo Čila, države z 19 milijoni prebivalcev, eno najbolj razvitih v Latinski Ameriki.

OECD (2024) poroča, da se bo po gospodarski rasti leta 2024 za 2,4 odstotka leta 2025 gospodarska rast nekoliko upočasnila na 2,3 odstotka, leta 2026 pa na 2,1 odstotka. OECD še ocenjuje, da bodo k rasti prispevali postopno okrevanje naložb, solidna rast potrošnje, ki jo bodo podpirali naraščajoče realne plače in sproščeni finančni pogoji, ter vzdržno zunanje povpraševanje po mineralnih surovinah. Inflacija se bo po pričakovanjih postopoma zniževala in se do začetka leta 2026 približala ciljni vrednosti treh odstotkov.

Po oceni OECD (2024) bi bilo treba pospešiti izobraževanje in usposabljanje zaradi potreb po usposobljenih delavcih na trgu delovne sile, še posebej strokovnjakov na področju informacijske in komunikacijske tehnologije ter tudi delavcev v rudarski industriji. Olajšanje dostopa ženskam do formalne plačane zaposlitve (s spodbujanjem formalnega in dostopnega otroškega varstva), zlasti na bolje plačanih delovnih mestih, lahko po oceni OECD (ibid.) podpira enakost spolov in spodbuja gospodarsko rast. Po OECD (2025) bi poleg enakopravnosti spolov h gospodarskemu napredku v Čilu v prihodnje prispevala tudi uspešna digitalni in zeleni prehod.

Kljub navedenim pozitivnim kazalnikom se platformna ekonomija v Čilu razvija v okoliščinah, ki se bistveno razlikujejo od teh, ki jih poznamo v evropskem gospo-

darstvu. Tako kot v drugih državah globalnega juga je za čilsko gospodarstvo značilna visoka stopnja neformalnega gospodarskega delovanja.⁵ Čilski raziskovalci opozarjajo, da je glede na socialni kontekst, za katerega so značilne visoka socialna in politična razklanost ter velike ekonomske in socialne razlike, nujno analizirati platformno gospodarstvo, v katerem se lahko zdajšnje razlike še dodatno povečajo.

Bonhomme, Ustek-Spilda in Arriagada (2024) pojasnjujejo, da so za čilski trg delovne sile, čeprav velja za najbolj razvitega v Latinski Ameriki, značilne neenakost in veliko nestabilnih služb z nizkimi dohodki. Po podatkih MOD (2016, po Bonhomme, Ustek-Spilda in Arriagada 2024) ima približno ena tretjina formalno zaposlenih kratkoročne pogodbe, ki trajajo manj kot eno leto, z neustreznimi zdravstvenimi in pokojninskimi ugodnostmi, kar priča o visoki prekarnosti. Arriagada et al. (2023, 2) opozarjajo na številne študije, ki predvidevajo, da bo vse večja prisotnost podizvajalstva pospešila izkoriščanje, prekarizacijo, nizke plače, negotovost zaposlitev in slabo pogajalsko moč delavcev.

Analiza stanja in razvoja platformnega gospodarstva v Čilu omogoča zaradi specifik čilskega gospodarstva in družbe celovitejše razumeti platformizacijo gospodarstva na splošno. Čile je posebej zanimiv tudi zaradi uporabe neoliberalnega gospodarskega modela od 90. let 20. stoletja, razmeroma velike razširjenosti digitalnih tehnologij⁶ in nedavne regulacije platformnega dela.

Pomembno je opozoriti, da analiziramo stanje platformnega dela v državi, ki je razmeroma nedavno bila deležna upora velikega števila državljanov proti gospodarski politiki vladajoče elite. Množični protesti, poimenovani »Estallido Social« (socialni izbruh), ki so se začeli 25. oktobra 2019,⁷ so bili ena največjih javnih demonstracij državljanskega nezadovoljstva v zgodovini Čila. Neposredni povod za proteste je bilo povišanje cene vozovnic za javni prevoz v prestolnici (Santiago), hitro pa so prerasli v širše gibanje proti neenakosti, neoliberalni gospodarski politiki in vladni korupciji. Množica protestnikov, najprej v Santiago, potem po celotnem Čilu, je takrat izražala nezadovoljstvo predvsem z dohodkovno neenakostjo, z visokimi življenjskimi stroški in s šibko socialno zaščito, do katerih je privedla privatizacija izobraževanja, zdravstvenega varstva in pokojninskega sistema v okviru čilskega neoliberalnega gospodarskega modela. Takratni predsednik Sebastián Piñera je razglasil izredne razmere in vpoklical vojsko, kar se ni zgodilo vse od Pinochetove diktature. Gordon-Zolov (2023) v članku, ki se sicer osredinja na čilsko »umetnost protesta«, ⁸ razlaga, kako je manjše povišanje cene vozovnice za podzemno železnico za 30 pesov kmalu preraslo v »veliko revolucijo pod geslom ‚Ne gre za 30 pesov, gre za 30 let!‘«, kar je jasno izražalo artikulirano nezadovoljstvo s 30-letno dediščino brutalnega 17-letnega režima Augusta Pinocheta in neoliberalnega gospodarskega eksperimenta. Protestniki so demonstrirali proti družbeni in gospodarski neenakosti, a tudi proti neučinkovitemu pravosodju ter zaradi pomanjkanja pravic žensk, domorodnih ljudstev in manjšin. Vlada Sebastián Piñere je v odgovor na proteste napovedala vrsto ukrepov, imenovanih »nova socialna agenda«, ki so vključevali izboljšanje pokojnin, zdravstvene zaščite, plač in delovanja javne uprave (Prensa Presidencia 2019).

PLATFORMNA EKONOMIJA V ČILU: STANJE IN UREDITEV

Fuentes in Gonzalez (2022) poročata o začetkih razvoja platformne ekonomije v Čilu, ko je januarja 2014 začel delovati Uber, aplikacija za prevoze. Od takrat do danes so v državo prišle ali so se razvile številne digitalne platforme in obseg platformne ekonomije stalno narašča. Platforme niso več osredinjene le na dejavnost prevoza (kot sta Uber in Cabify), temveč tudi na prevoz hrane in paketov (npr. Rappi, PedidosYa), kar se je še posebej razvilo zaradi in med pandemijo ter po njej, ko se je razmahnila trgovina z blagom po družbenih omrežjih, ter na zagotavljanje drugih izdelkov in storitev na spletu. Bravo in Castillo (2021, 2) ocenjujeta, da je med majem 2019 in majem 2021 digitalne storitve dostave, prevoza ali nakupov uporabilo okrog deset milijonov ljudi. Ocenjujeta, da je pri delu, organiziranem ali posredovanem prek platform, v letu 2021 sodelovalo 189.100 ljudi (približno 2,1 odstotka delovne sile). Omenjata prisotnost okrog desetih najvidnejših platform, med katerimi so najpomembnejši prevozniki Uber in Cabify, dostavljavci hrane PedidosYa in Rappi ter dostavljavec izdelkov iz supermarketov Cornershop.

Tako kot drugod po svetu so digitalne platforme v Čilu obljubljale, da bodo delavcem omogočile prožnost pri delu, hiter odziv na povpraševanje uporabnikov in nižje cene kot tradicionalna konkurenca. Čilska uradna statistika zaposlovanja teh vrst dela ni imela opredeljenih, tako da do leta 2020 ni bilo podatkov o tem, koliko ljudi je delalo na platformah, kakšen je bil njihov sociodemografski profil, za katere platforme so delali in kako dolgo. Vendar so leta 2020 na Nacionalnem statističnem inštitutu (INE) izoblikovali pilotni sklop vprašanj, vključenih v nacionalno anketo o zaposlovanju (ENE), ki so zajemala različne vidike dela na digitalnih platformah.⁹ Fuentes in Gonzalez (2022) poročata o rezultatih ankete od januarja 2020 do junija 2022.¹⁰ Anketni podatki so v skladu s prej povedanim in kažejo, da se delež zaposlenih, ki delajo na digitalnih platformah, giblje okoli 2,3 odstotka zaposlenih v državi, kar pomeni približno 205.000 delavcev v zadnjih mesecih anketiranja.

Anketirance so prosili, da poročajo o platformah, ki jih uporabljajo pri svojem delu. Ko so upoštevali samo prvo navedeno platformo, je bil WhatsApp glavna platforma za tiste, ki so zaposleni v gig ekonomiji, saj ga je približno 30 odstotkov zaposlenih na teh platformah označilo za svojo glavno platformo. Facebook in Instagram si delita drugo mesto (približno 25 %). Na platformi Uber, ki je imela pomemben položaj med pandemijo, od aprila 2020 beležijo upad uporabe, medtem ko je uporaba Facebooka in drugih platform, kot sta WhatsApp in Instagram, v porastu. Pri Uberju kot prvi platformi dela približno deset odstotkov delavcev.

Tudi glede na obseg omemb posamezne platforme (katere platforme delavci najpogosteje uporabljajo) prednjači WhatsApp s približno 35 odstotki delavcev digitalnih platform. Tudi tukaj si Facebook in Instagram delita drugo mesto, saj ju uporablja približno 30 odstotkov delavcev. Uber uporablja približno 11 odstotkov delavcev. Platforme Cornershop, Pedidos Ya, Rappi, Didi, Mercado Libre in Uber Eats, ki dosežajo delež, večji od nič, skupno ne presežejo petih odstotkov.

Na podlagi razvrstitve platform v tri kategorije (prodaja blaga in storitev, prevoz oseb in prevoz hrane) so raziskovalci ugotovili, da približno 80 odstotkov delavcev na digitalnih platformah uporablja aplikacije, povezane z nakupom blaga in storitev,

aplikacije za prevoz in dostavo pa zaposlujejo po približno deset odstotkov platform-nih delavcev.

Delavci na digitalnih platformah za prevoz potnikov ali dostavo so torej danes v Čilu dejansko manjšina tistih, ki delajo na digitalnih platformah. Večji del delavcev na digitalnih platformah zagotavlja blago in storitve po internetu na Facebooku, Instagramu ali WhatsAppu ali ustvarja vsebine (na YouTubeu, Instagramu ali TikToku).

Na splošno je po svetu delo na digitalnih platformah maskulinizirano. Tudi v Čilu je na platformah zaposlenih več moških, vendar je razlika manjša, kot bi pričakovali in kot je v drugih delih čilskega gospodarstva. Po pandemiji je med delavci na digitalnih platformah približno 45 odstotkov žensk. Prisotnost spolov je na platformah torej nekoliko bolj uravnotežena kot na preostalem trgu dela, kjer je žensk 40 odstotkov. To raziskovalci pojasnijo z visoko prisotnostjo žensk na digitalnih platformah za nakup blaga in storitev.

Glede na starost je na platformah veliko mladih in srednje starih. 45 odstotkov delavcev na platformah je v starostni skupini od 18 do 34 let. Ta delež je večji, kot je delež mladih (35 %) med zaposlenimi na tradicionalnih delovnih mestih, kar raziskovalci razlagajo z večjo tehnološko usposobljenostjo in pripravljenostjo mladih za tovrstno delo.

Delavci v starostni skupini od 35 do 54 let so prisotni v podobnem obsegu (približno 45 %), kar je podobno deležu toliko starih delavcev na tradicionalnih delovnih mestih (približno 48 %). Najmanj je delavcev, starih 55 let in več (10 %), na tradicionalnem trgu pa zasedajo 17 odstotkov delovnih mest.

Kar zadeva državo izvora, so delavci migranti na digitalnih platformah manjšina (15 %). Vendar pa je res, da so v večjem deležu prisotni na platformah za dostavo, saj predstavljajo skoraj 60 odstotkov vseh delavcev.

Glede na stopnjo izobrazbe so to največkrat osebe z univerzitetno in visokošolsko izobrazbo. Sledijo jim osebe s srednješolsko izobrazbo, medtem ko jih je z osnovnošolsko zelo malo. To raziskovalci razlagajo z zahtevano ravno digitalnega znanja, ki ga delavci z zgolj osnovno izobrazbo nimajo. Na splošno imajo torej delavci na digitalnih platformah v Čilu višjo stopnjo izobrazbe kot delavci v tradicionalnih sektorjih.

Med delavci na platformah je po tej analizi velika večina samozaposlenih (71 %), sledijo jim tisti, ki so zaposleni v zasebnem sektorju (18 %), preostali (11 %) pa spadajo v različne kategorije, kot sta delodajalec in zaposleni v javnem sektorju. To načeloma kaže na nestabilno delo brez dostopa do socialne varnosti.

Analiza ugotavlja, da je delež delavcev s pogodbami v digitalnem gospodarstvu Čila po stabilizaciji po koncu pandemije blizu 77 odstotkov. V istem obdobju je v tradicionalnem gospodarstvu delež delavcev s pogodbami skoraj 89-odstotni.

Čeprav se zdi, da prej navedena številka kaže na nizko stopnjo prekarnosti, je treba upoštevati, da je to v veliki meri posledica dejstva, da so med analiziranimi delavci bili v večini delavci, ki uporabljajo platforme za ponujanje izdelkov in storitev, ki praviloma imajo sklenjene pogodbe o zaposlitvi.

Nekoliko večjo razliko (19 %) analiza kaže pri plačevanju prispevkov za socialno varnost, med delavci v tradicionalnem in digitalnem sektorju. Nazadnje, kar zadeva tedenski delovni čas, analiza ugotavlja, da delavci v tradicionalnem sektorju poročajo o skoraj enakem številu delovnih ur na teden kot delavci v tradicionalnem gospodarstvu.

Druge analize, opravljene v Čilu razkrivajo precej bolj prekaren položaj delavcev, ki se skriva pod navidezno avtonomijo, ki jo delavcem obljublja platformna podjetja, tako kot drugod po svetu. V poročilu Fairwork Leyton et al. (2021) ugotavljajo, da se delavci na platformah v Čilu med drugim soočajo:

- z dolgimi delovniki; analiza skupine več kot 200 delavcev, izvedena leta 2021 v Santiagu, je pokazala, da so delali povprečno 62 ur na teden (Asenjo Cruz in Coddou Mc Manus 2021, 45, 46, v Leyton et al. 2021);
- z nizkimi in negotovimi plačami ter s pomanjkanjem varnosti zaposlitve; v anketi Fairwork Čile 2021 nobeno od anketiranih podjetij ni moglo dokazati, da njihovi delavci zaslužijo vsaj minimalno mesečno plačo po plačilu stroškov (Fairwork 2021, v Leyton et al. 2021);
- s slabimi zdravstveni in varnostnimi pogoji; skoraj polovica delavcev, anketiranih v študiji Fairwork, je navedla, da nimajo zdravstvenega zavarovanja (Asenjo Cruz in Coddou Mc Manus 2021, 57, 58, v Leyton et al. 2021).

Arriagada et al. (2023) so izvedli analizo delovnih pogojev v gig ekonomiji v Čilu na podlagi mednarodno uveljavljenega okvira Fairwork za oceno delovnih pogojev. Analiza digitalnih delovnih razmerij je obsegala tako analizo njihovega formalnega okvirja (zapisana pravila) kot delovne realnosti, kot jo vidijo sami delavci (pogoji in omejitve, s katerimi se dnevno soočajo delavci, ki delajo v gig ekonomiji, ter njihovo dožemanje platformnega dela). Na podlagi intervjujev s 35 delavci, ki delajo v čilski gig ekonomiji, avtorji razpravljajo o nasprotujočih si izkušnjah delavcev, ki so v nekaterih pogledih odvisni od platform, v drugih pa se jim zoperstavljajo. Ugotovili so, da so protislovni pogoji in okoliščine dela, inherentni platformnemu delu, bili še bolj zaostreni za gig delavce med pandemijo covid-19, ko so številna tveganja nosili samo delavci, saj so platforme še naprej poudarjale njihovo »izbiro«, da opravljajo delo v obsegu in na način, ki so ga v veliki meri bili prisiljeni »izbrati«. Ta analiza razkriva, da je narava povezave med platformo in delavcem nedvomno izrazito delovno razmerje z jasno določenimi elementi odvisnosti delavca od platforme.

Pomemben del čilske delovne sile, posebej v neformalnem sektorju, so migranti (predvsem iz sosednjih držav). Bonhomme, Ustek-Spilda in Arriagada (2024) so se zato osredinili na analizo tega dela delovne sile, kajti tako kot drugod po svetu imajo na digitalnih platformah za delo pomembno vlogo in po navadi najmanj ugoden položaj. Poročajo o tem, da je v čilski platformni ekonomiji približno petina delavcev migrantov. Posebej so prisotni v najbolj prekarnih službah dostave, v katerih predstavljajo večino delavcev (60 %). Na podlagi kvalitativne študije v Čilu so razkrili, kako migranti iz Latinske Amerike in s Karibov krmarijo in se upirajo neprijaznim delovnim razmeram na platformah dela, pri čemer so posebej analizirali algoritemski nadzor platformnega dela za dostavo hrane in vpliv migrantskega statusa na njihov delovni položaj. Trdijo, da je za razumevanje položaja in delovanja delavcev migrantov treba preučiti tako družbeno-tehnično strukturo platform kot tudi migracijske politike države gostiteljice. Te politike določajo status delavcev, oblikujejo njihove delovne izkušnje in vplivajo na njihovo odpornost v negotovih delovnih razmerah.

Avtorji predstavljajo koncept »opolnomočenosti platformnih delavcev migrantov«, da bi ponazorili različne oblike upora, ki jih migranti uporabljajo pri krmarjenju

po platformnem trgu. Ta koncept prispeva k razumevanju odtenkov (ne)moči migrantov, ki segajo od sprejemanja do upora. Analiza pokaže, kako se z negativnimi vidiki platformno organiziranega dela soočajo in se jim upirajo najbolj prekarni delavci in kako dinamika razmerja med platformami in delavci modificira nekatere najpomembnejše sestavine platformne organizacije dela (kot je npr. popolni nadzor) ter producira nove oblike izkoriščanja platformnih delavcev.

Raziskovalci so izvedli intervjuje z delavci iz Venezuele, Kolumbije, Peruja, s Kube in iz Čila, večinoma moškimi. Med posebej izpostavljenimi negativnimi vidiki delovnih razmer so bili: nevarnost prometnih nesreč; konstanten nadzor in skrite sankcije, ko se delavci ne odzivajo na tvegane naloge, kot so npr. delo v slabem vremenu ali v nevarnih mestnih četrtih; prisila algoritemske organizacije in nadzora dela v kršitev zakonsko določenega osemurnega delovnika ter težave, povezane z migrantskim statusom delavcev.

Primer 1: Lažni narativ »eres tu propio jefe« (bodi svoj šef)

»/.../ jaz nisem partner, nisem podjetnik, sem navaden delavec /.../ to, kar oni imenujejo zaposlitev prihodnosti, /.../ je, po mojem mnenju, grobo, vulgarno izkoriščanje delavskega razreda, /.../ oni z algoritmi maksimizirajo dobiček in z istimi algoritmi zaslužujejo delavce in se igrajo z njihovo emocionalno stabilnostjo, z njihovim delom, /.../ z njihovimi družinami« (Silvio iz Venezuele, cit. po Bonhomme, Ustek-Spilda in Arriagada 2024).

Raziskovalci so ugotovili, da platformni delavci uporabljajo tri glavne strategije upora:

1. Prva strategija je *prijava iz več računov* (»multi-apping«): Številni migranti sočasno delajo na različnih platformah. Šele delovnik med 10 do 12 ur na dan ves teden jim zagotavlja, da zaslužijo ustrezno plačilo. Raziskovalci ugotavljajo, da je paradoksalno, da je ta strategija skladna z interesi podjetij, saj širi nabor razpoložljivih delavcev.
2. Druga strategija so *»izposojeni« ali »piratski« računi*, ki jo uporabljajo migranti, posebej tisti, ki imajo neurejen ali negotov migracijski status in pridobivajo dela pod lažno identiteto.

Z uporabo teh dveh strategij je delavcem omogočeno, da nadaljujejo delo tudi, če platforme deaktivirajo njihove račune ali jih začasno blokirajo.

Zanimiva je ugotovitev študije, da se je strategija izposojanja računov razvila v samostojno dejavnost – nezakonito trgovanje z računi, kar omogoča delo migrantom, ampak zaradi dodatnih stroškov tudi dodatno poslabšuje njihov položaj ali celo onemogoča delo nekaterim migrantom. Delovni računi platform se oglašujejo na družbenih omrežjih, kjer del zgodnjih uporabnikov platform zdaj služi z oddajanjem. Na črnem trgu platformnih računov nekateri torej najdejo trajni vir prihodka, medtem ko so nezaščiteni migranti prisiljeni plačevati visoke »najemnine« in so nekateri celo opeharjeni za zaslužek, ki ostane pri lastnikih računov.

Primer 2: Bojkotiranje platformnih pravil: platformni delavci v »sendviču izkoriščanja« platform in novih »poslovnežev«

»Najprej so bili ljudje, ki so ti posodili račune /.../ sorodnik, ki je imel dokumente, ali prijatelj /.../ Potem pa se je začela pojavljati celotna črna trgovina /.../, kjer so bili ljudje, ki so se registrirali samo zato, da bi ga oddali, in ljudje so plačevali /.../ tedensko /.../, mesečno /.../« (Pedro iz Venezuele, cit. po Bonhomme, Ustek-Spilda in Arriagada 2024).

3. Tretja strategija upora je aktivno sodelovanje v *skupnostih na spletnih družabnih omrežjih*. Te skupnosti so, poleg prostora komunikacije, pospešile kolektivno organiziranje in upor platformam na protestih in srečanjih v »resničnem« prostoru. V Čilu (in drugod v Latinski Ameriki) so med pandemijo covida-19 migrantski delavci, ki so dostavljali hrano, organizirali množične proteste in zahtevali boljše delovne razmere (Howson et al. 2020). Ti protesti niso privedli do ustanovitve sindikatov ali bolj formaliziranih oblik kolektivnega upora. Razloge za to raziskovalci vidijo v strahu delavcev zaradi slabih izkušenj, teh, ki so poskusili s kolektivnim uporom izboljšati delovne razmere (npr. ukinitve računov delavcev platforme PedidosYa, ki so se organizirali na uporniškem Twitterjevem računu »Riders Unidos Ya«). Bonhomme, Ustek-Spilda in Arriagada (2024, 14) poudarjajo, da se na osnovi analize lahko pričakuje, da se bodo delovne razmere na platformah za migrante še poslabšale, če ne bodo sprejeti predpisi, ki bi migrante učinkovito zaščitili. Raziskovalci poudarjajo, da se v Čilu, kjer, kot trdijo, obstaja velika nenaklonjenost migrantom latinskoameriških držav,¹¹ mora spremeniti tudi migracijska politika, da bi se presegla negotovost migrantskega dela. V nasprotnem bo ekonomija platform preprosto postala »še en neformalni trg, ki bo prispeval k izkoriščevalskemu režimu dela, ki ga doživljajo migranti«. K tej ugotovitvi pa bi lahko dodali, da nizki standardi dela za migrante predstavljajo nevarnost za vse delavke in delavce, na platformah in zunaj njih, ker nizki standardi dela na platformah lahko prispevajo k poslabšanju delovnih pogojev na drugih delovnih mestih in nadaljnjemu povečanju obsega neformalnega dela.

Asenjo Cruz in Coddou Mc Manus (2021) sta analizirala delovanje digitalnih platform dela v sektorju dostave in prevozov s taksijem v Santiagu in prišla do podobnih ugotovitev kot drugi, ki so kvalitativno proučevali delovne razmere na digitalnih platformah dela v Čilu. Namen analize je bil razkriti izzive, ki jih te oblike dela predstavljajo za utrditev standardov »dostojnega dela«. Raziskovalca sta opisala različne izkušnje, izzive, pričakovanja in motivacije dostavljavcev v Čilu. Analiza je opravljena na osnovi ankete, izvedene v Santiagu, v kateri je sodelovalo 259 kurirjev digitalnih platform dela in 50 kurirjev trgovin in restavracij, ter telefonske raziskavi, izvedene med pandemijo covida-19. Analiza je pokazala, da delavci kot pozitivno pri delu ocenjujejo relativno ugodno časovno in prostorno prožnost dela, vendar opozarjajo na negativne finančne vidike, slabe delovne razmere in slabe vplive na osebno počutje. Rezultati so razkrili ranljivost in negotove delovne razmere, s katerimi se soočajo delavci, ki dostavljajo pošiljke: naporen delovni čas, pomanjkanje osebne zaščite in izpostavljenosti zdravstvenemu tveganju. Nepreglednost algoritemskega upravljanja in pomanjkanje trdnije povezave s platformo sta največja vira negotovosti in nelagodja za dostavne delavce.

V Čilu so pred tremi leti, po številnih javnih protestih dostavljavcev in voznikov, ki so bili tudi del množičnih socialnih protestov v letih 2019 in 2020, v obdobju mandata levo usmerjenega predsednika Gabriela Borića,¹² naredili korak k ureditvi položaja delavcev na platformah. Enajstega marca 2022 je bil objavljen Akt št. 21 431, ki je začel veljati 7. septembra 2022. Akt je v Zakon o delovnih razmerjih vključil nova pravila, ki se uporabljajo za zaposlene, ki opravljajo storitve na digitalnih platformah (L&E Global 2022). Nova zakonodaja opredeljuje pojma »digitalno platformno podjetje« in »zaposleni na digitalni platformi«, pri čemer razlikuje med odvisnimi in neodvisnimi zaposlenimi, kar definira glede na način opravljanja dela oz. storitev. Odvisni so delavci, ki imajo pogodbo o zaposlitvi, v kateri morajo biti določene nekatere minimalne zahteve, kot so način izračuna plače, razpored dela in čas počitka. V primeru samostojnih (neodvisnih) delavcev mora biti sklenjena pogodba o storitvah, ki ni urejena z delovnopravnim zakonom, vendar vsebuje nekatere minimalne pravne določbe, kot sta 12 ur prekinitve dela v 24 urah in pravica do dostopa do socialne varnosti, pri čemer se plačujejo morebitni prispevki. V primeru neodvisnega zaposlenega na digitalni platformi se plačilo izračuna na osnovi dejansko opravljenih ur dela, ne sme pa biti nižje od deleža zakonsko določene minimalne mesečne plače, povečanega za 20 odstotkov.

Pomen in novost zakona je v tem, da ureja delovno razmerje, ki obstaja med delavci digitalnih platform dela in njihovimi delodajalci. Od delodajalca namreč zahteva, da delavca_ko obvešča o ponujenih storitvah, varuje njegove_njene zasebne podatke, ga_jo usposablja na področju varnosti in zdravja pri delu ter mu_ji zagotavlja zaščitna sredstva in zavarovanje za kritje škode na prevoznih sredstvih, ki jih uporablja.

Kritiki nove zakonodaje (Leyton et al. 2021) so, dokler je še bila v osnutku, opozarjali na strukturne težave, ki bi lahko vplivale na zmožnost oblikovanja pravil za učinkovito zaščito delavcev na platformah dela. Po njihovem mnenju te težave izhajajo iz neupoštevanja glavnega vprašanja, ki mora biti med osrednjimi skrbmi vsake oblike ureditve dela. To je obstoj neenake pogajalske moči med delavci in delodajalci – med tistim, ki vodi podjetje in določa pogoje, ter tistim, ki ponuja svoje delo in lahko sprejme le predpisane pogoje. Avtorji opozarjajo, da delovno pravo posega v avtonomijo pogodbenih strank prav zaradi tega, da bi zaščitilo šibkejšo pogodbeno stranko (delavca). Ugotavljajo, da čeprav nova pravila pomenijo določen napredek na področju, niso rešitev, ki bi uspešno presegla propagando platformnih podjetij o samostojnosti delavcev, ki naj med seboj tekmujejo. Ocenjujejo, da bi izhodišče ureditve, ki priznava obstoj razlik v pogajalski moči, moralo biti priznavanje in krepitev vloge organiziranih delavcev.

ALTERNATIVE, USMERJENE K TRAJNOSTNEMU RAZVOJU PLATFORMNE EKONOMIJE V ČILU

Cano, Espelt in Morell (2024) ugotavljajo, da čeprav platformno gospodarstvo eksponentno raste in se pričakuje, da bi lahko prispevalo k trajnostnemu razvoju, se je večina raziskav osredinila na znane in konvencionalne modele, kot sta Uber in Airbnb, in zanemarila alternativne platforme. Zato so se oni usmerili v raziskovanje

potencialnega prispevka poslovnega modela posameznih platform dela k trajnostnemu razvoju in analize delovnih pogojev v platformnih zadrugah, ki naj bi predstavljale alternativo konvencionalnim platformam dela. V raziskavi, ki je bila osredinjena na razlike med platformnimi modeli, so pozornost posvetili načinom upravljanja, ekonomski trajnosti, tehnološkim in podatkovnim politikam, družbeni odgovornosti in razsežnostim zunanjega vpliva. Te dimenzije so analizirali v 60 platformah, potem pa so poglobljeno analizirali še tri alternativne platforme: Fairbnb, Katuma in Smart.

Z analizo so ugotovili, da trajnostna zasnova platformnega gospodarstva spodbuja trajnostni razvoj. Vidiki, pri katerih se tri preučevane platforme posebej kažejo kot alternativa konvencionalnim platformam, so: enakost spolov, ravnovesje med poklicnim in zasebnim življenjem ter možnosti usposabljanja. Opazili so dobro povezavo med cilji trajnostnega razvoja in »zvezdo demokratičnih lastnosti« (model, ki so ga uporabili pri analizi platform), kar kaže na pomen razlikovanja med prodemokratičnimi in trajnostnimi lastnostmi različnih platformnih modelov, da bi v skladu s temi razlikami oblikovali javne politike. Vendar so opozorili na dejstvo, da nobena digitalna platforma ni popolnoma izpolnjevala petih demokratičnih zahtev ali prispevala k vsem ciljem trajnostnega razvoja.

Analiza je torej pokazala, da poleg konvencionalnih platform dela obstaja alternativni model sodelovalnega gospodarstva, ki temelji na demokratičnih lastnostih in načelih kooperativizma.¹³ To je v skladu s prejšnjimi študijami o socialni ekonomiji in kooperativah, ki opozarjajo na to, da kooperativni modeli ponujajo boljše delovne pogoje v primerjavi s »kapitalističnimi« podjetji. Vendar avtorji (ibid.) opozarjajo, da so preučevane alternativne platforme v začetni fazi ali pa se soočajo s težavami v zvezi z razširljivostjo, ekonomsko vzdržnostjo, s financiranjem in z enakostjo spolov. Njihove ugotovitve se skladajo z ugotovitvami o alternativnih, kooperativnih digitalnih podjetjih v drugih okoljih.

Primer 3: Zadrugištvo platformnih delavcev: CoopCycle

Krog COOP (CoopCycle) je *federacija kooperativ* za dostavo s kolesi. Demokratično jo upravljajo kooperativne članice, ki so tudi same demokratično upravljana podjetja. Glavni namen združevanja v federacijo je enotno nastopanje in združevanja virov z namenom zmanjšanja stroškov. S skupnim nastopom imajo več *pogajalske moči* pri zaščiti pravic dostavljavcev s kolesi.

Federacija se definira kot organizacija, usmerjena k spodbujanju *solidarnosti* med zadrugami, zmanjšanju njihovih stroškov zaradi združevanja storitev in ustvarjanju skupne moči za zagovarjanje pravic kurirjev. Poudarjajo, da se borijo proti kapitalističnemu rentniškemu učinku. Združene storitve se financirajo s kotizacijo. Stopnjo kotizacije in zbrana sredstva demokratično upravljajo člani. Federacija z delovanjem želi ustvariti pogoje za to, da kurirji ponovno dobijo *nadzor nad svojo poklicno dejavnostjo in delovnimi orodji*.

Kurirji skupaj odločajo o:

- višini prispevka vsakega glede na njegov pravni status (zadruga, združenje, institucionalni subjekt);
- dodelitvi sredstev;
- sistemu plačila za opravljanje storitev.

Status zadruge za člane CoopCycle uveljavlja tudi demokratično upravljanje na lokalni ravni.

Vir: Coop Cycle 2024

Na osnovi predstavljenih raziskav lahko ugotovimo, da je platformno gospodarstvo, ki se je v Čilu začelo s prihodom Uberja leta 2014, od takrat doživelo močno rast in diverzifikacijo – poleg prevoza oseb zdaj vključuje dostavo hrane, blaga ter spletno prodajo po družbenih omrežjih. Med letoma 2019 in 2021 je tovrstne storitve uporabljalo približno deset milijonov ljudi, vključenih pa je bilo več kot dva odstotka celotne delovne sile. Glavnina platformnega dela se ne odvija več na t. i. velikih platformah (npr. Uberju), temveč iz vsakdanjih aplikacij, kot so WhatsApp, Facebook in Instagram, kjer poteka prodaja storitev in izdelkov. Večina delavcev na platformah je samozaposlenih. Platformno delo v Čilu vključuje velik delež mladih in visokoizobraženih ljudi. Razmeroma visoka je tudi udeležba žensk (45 %), kar presega njihov delež na tradicionalnem trgu dela. Migranti predstavljajo posebno ranljivo skupino, predvsem v sektorju dostave hrane, v katerem predstavljajo večino delavcev. Kljub obljubljeni fleksibilnosti dela raziskave razkrivajo predvsem precejšnjo stopnjo prekarности: dolg delovnik, pomanjkanje socialne zaščite, nizke dohodke in nevarne pogoje dela. Algoritemski nadzor dodatno omejuje avtonomijo delavcev. Čilski platformni delavci razvijajo inovativne strategije upora, kot so uporaba več računov, črni trg dostopov in skupnostne spletne mreže. Kljub kolektivnim dejanjem pa ostajajo formalne oblike sindikalnega organiziranja redke zaradi strahu pred sankcijami. Leta 2022 je Čile sprejel prvi zakon o digitalnem delu v Latinski Ameriki. Zakon razlikuje med odvisnimi in neodvisnimi delavci ter določa osnovne pravice, kot so zaščita zdravja in zasebnosti, vendar še vedno pušča odprta ključna vprašanja neenake moči platform dela na eni in delavcev na drugi strani. Čilski raziskovalci opozarjajo na potencial alternativnih platform, kot so kooperative, ki temeljijo na demokratičnem upravljanju in solidarnosti, vendar so te še v razvojni fazi in se soočajo z izzivi ekonomske vzdržnosti.

SKLEP IN PRIPOROČILA ZA TRAJNOSTNI RAZVOJ PLATFORMNE EKONOMIJE

Čilenci so se pred nekaj leti množično zoperstavili neoliberalni ideologiji in praksam, ki vodijo v povečevanje socialnih razlik, revščino in socialno negotovost velikega dela prebivalstva ter negotovo profesionalno in življenjsko prihodnost tudi visokoi-zobraženih mladih. Čilski protesti leta 2019 so bili prelomnica v zgodovini države in so bili razumljeni kot upor in izpodbijanje dediščine neoliberalnega gospodarskega sistema v smeri boja za gospodarsko pravičnost, delavske pravice in demokratične reforme. Čeprav so protestom sledili določeni pozitivni premiki, naša analiza kaže, da platformno gospodarstvo, ki je zraslo na logiki neoliberalnega gospodarjenja, narašča v Čilu, tako kot drugod po svetu. Zdi se, kot da platformam dela uspeva zakriti dejstvo, da nadaljujejo in stopnjujejo neoliberalni diskurz. V Latinski Ameriki je prekarnost velikega dela delavcev že itak bila normalizirana, tako kot tudi ekonomske razlike. V zadnjih letih, v obdobju rasti digitalno organiziranega gospodarjenja, pa analize kažejo na nadaljnji zdrs delovnih standardov¹⁴ in življenjskega standarda. Howson et

al. (2020, 20) opozarjajo, da je poleg prekarnih delovnih razmer v Latinski Ameriki močno prisotna tudi dediščina delavskega upora.

V Čilu so, kot drugod po svetu, delavci na platformah dela pokazali veliko sposobnost organiziranja in odzivanja na težave, s katerimi se srečujejo pri delu. Bonhomme, Ustek-Spilda in Arriagada (2024) trdijo, da so v Čilu edino učinkovito orodje za zmanjšanje asimetrije moči med platformami in delavci prav protesti in tožbe. Strinjamo pa se s čilskimi analitiki, ki poudarjajo nujnost institucionalne podpore tem naporom, ker se upori platformnih delavcev kažejo kot le delno učinkoviti. Za vzpostavljanje novega načina organiziranja dela pa bosta ob organizaciji delavcev še naprej pomembna tudi ozaveščanje širše javnosti in javna razprava. Howson et al. (2020) pri tem poudarjajo pomen čezmejnega sodelovanja delavcev v uporabi in menijo, da je med pandemijo covid-19 dober del javnosti ugotovil, kako pomembno delo opravljajo platformni delavci in v kakšnih razmerah ter da je v tem času val stavk v Latinski Ameriki postal velik izziv statusu quo pri ureditvi platformnega gospodarstva.

Čilska izkušnja je za Evropo, torej tudi Slovenijo, dragocena, ker ponuja svojevrsten »laboratorij« ekstremnih oblik platformnega dela, ki v Sloveniji (še) niso prisotne, in je tako lahko zrcalo prihodnjih tveganj, a tudi primer delavske mobilizacije od spodaj. Nova zakonodajna ureditev platformnega dela v Čilu, ki je korak v smeri potrebne regulacije, je za zdaj v pomoč le manjšemu delu, in to ne najbolj ogroženih delavcev. Zaradi nepreglednosti in pomanjkanja nadzora nad učinki poslovnega modela se koristi od poslovanja koncentrirajo pri majhnem številu ekonomskih akterjev v Čilu, tako kot drugod po svetu. Obstoj alternativnih oblik digitalno podprtega kooperativnega organiziranega dela in poslovanja pa dokazuje, da bi digitalno organizirano gospodarstvo lahko, namesto izključno koristim maloštevilnih digitalnih platformnih monopolistov, služilo dobrobiti velikega števila delavcev in uporabnikov ter vsestransko trajnejšemu poslovnemu modelu.

Platformna ekonomija v obstoječih korporativnih oblikah pogosto pogloblja neenakosti, prekarizira delo in koncentrira ekonomsko moč, kar jo umešča med netrajnostne gospodarske prakse. Vendar raziskave, tudi v Čilu, razkrivajo možnosti in potencialne alternativnih, demokratičnih modelov platformnega gospodarstva, ki bi se lahko razvijali v smeri trajnostnega razvoja. Te trajnostne oblike platformne ekonomije – kot so platformne zadruge in druge oblike sodelovalnega gospodarstva – udeležajo temeljna načela trajnosti: spodbujajo pravičnejšo delitev dohodka in lastništva, omogočajo participacijo delavcev v upravljanju, zmanjšujejo okoljski odtis ter krepijo lokalne skupnosti. Pomembno prispevajo tudi k zaposlovanju, dostojnim delovnim pogojem in zmanjševanju družbenih ter ekonomskih neenakosti. Zato je ključno, da raziskovalci, odločevalci in civilna družba prepoznajo, razvijajo in podpirajo trajnostne oblike platformne ekonomije kot del prizadevanj za družbeno pravičnejšo in okoljsko odgovorno digitalno prihodnost.

Na podlagi analize, predstavljene v tem prispevku, je možno izoblikovati priporočila, ki naslavlajo izzive platformne ekonomije ter spodbujajo njen trajnostni preobrat:

1. Razvoj in vzpostavitev celovite zakonodaje in nadzornih teles za zaščito platformnih delavcev, ki naj bi zagotovila pravice iz dela (vključno z minimalnim dohodkom, pravico do socialne varnosti, zdravja pri delu in kolektivnega organiziranja).

2. Demokratizacija tehnološke infrastrukture: demokratizacija upravljanja in razdelitev koristi od tehnološkega napredka.
3. Zagotovitev nadzora nad algoritmi, ki bi vsebovala transparentnost algoritmov, pravico do razlage odločitev algoritmov (npr. ocenjevanja, blokad računov), ter obvezen, smiseln in dokazljiv človeški nadzor nad ključnimi odločitvami.
4. Podpora javnih politik alternativnim modelom organiziranja platformnega gospodarstva (npr. delavska demokratična podjetja oz. platformne zadruge ali ne-profitna platformna podjetja), kar vključuje finančno podporo, ugodne kredite in davčne olajšave, da bi uravnotežile moč platformnih korporacij.
5. Krepitev zavedanja in ustrezen odziv na prelivanje načel in praks poslovnega modela platformnih korporacij v druge dele gospodarstva, da bi se zaščitile človeka dostojne delovne razmere, socialna in okoljska varnost in ohranitev skupnosti.

Razvoj trajnostne platformne ekonomije bo zahteval poleg ustrezne zakonodajne strukture vzpostavitev novih institucij, ki bi skrbele, da platformno organizirano delo zasleduje interese skupnosti in naravnega okolja, v katerem delujejo (npr. neodvisnih nadzornih teles za spremljanje platformnega dela, ki bi nadzirala pogoje dela na platformah, analizirala algoritme ter delavcem in drugim zainteresiranim akterjem omogočala podajo pritožb ali prijave kršitev).

Javne politike bodo morale zagotoviti tudi doseganje ustrezne digitalne pismenosti in pravnega opolnomočenja delavcev ter širše skupnosti (npr. dostopna in obvezna usposabljanja za delavce o njihovih pravicah, uporabi aplikacij, algoritmi in strategijah organiziranja). Treba bo podpreti razvoj od platformnih korporacij neodvisnih tehnoloških orodij, ki delavcem omogočajo nadzor nad delovnimi pogoji (npr. forumov za izmenjavo informacij, aplikacij za oceno poštenosti platform) ter spodbujati samoorganizirane skupnosti delavcev (fizično in digitalno), ki krepijo solidarnost, izmenjavo informacij in koordinacijo protestnih akcij. V širši skupnosti bo treba krepiti znanje o možnih načinih razvoja z novimi tehnologijami podprtega gospodarstva.

Javne politike bi morale omogočiti ustrezne pogoje delovanja alternativnim platformnim podjetjem, ki bi nevtralizirali pomanjkljivosti konvencionalnih platform dela. Največji problem pri tem je zagotavljanje ustreznih oblik financiranja, ki bi lahko konkuriralo kapitalsko najmočnejšim ekonomskim globalnim akterjem.

Trajnostna platformna ekonomija je etična alternativa korporativnemu platformnemu kapitalizmu. Njen cilj ni zgolj tehnološka inovacija, temveč družbena preobrazba, ki digitalna orodja uporablja za uresničevanje pravičnejše, bolj vključujoče in okolju prijazne ekonomije.

Čilska izkušnja priča o moči neoliberalne gospodarske paradigme, ki že desetletja kroji globalno gospodarsko ureditev in za to uspešno uporablja tehnološke spremembe. Lastniki platformnih podjetij nadaljujejo zaostrovanje racionalizacije poslovanja in rasti dobičkov, delavci pa so v novih okoliščinah postavljeni pred številne izzive, ki jih vse težje obvladujejo. Nadaljnje raziskovanje bo pomembno, da se preseže predstava tehnološko determinističnega ter za družbo in naravo negativnega gospodarskega razvoja in se omogoči uporaba digitalnih tehnologij v trajnostno usmerjenem gospodarskem delovanju. Načela inkluzivnosti in participativnosti, ki vodijo delovanje že obstoječih modelov alternativnih platform in prispevajo k traj-

nostnemu razvoju, so lahko podlaga za določanje ciljev boljše organizacije digitalizirane in z umetno inteligenco podprte ekonomije. To bi prispevalo k preprečevanju negativnih posledic nenadzorovanih sprememb, utemeljenih na razvoju tehnologije in njene uporabe v poslovnih modelih, ki danes prekarizirajo delo in življenje vse večjega števila ljudi ob tem, ko se ekonomska moč (posledično tudi politična, medijska, družbena) koncentrira pri majhnem številu lastnikov največjih platformnih podjetij.

OPOMBE

- 1 Več o platformni ekonomiji in drugih vidikih digitalizacije v sodobni družbi glejte Kanjuo Mrčela in Kohont (2024). Zbornik vsebuje tudi pojmovnik nekaterih konceptov in terminov, vezanih na digitalizirano gospodarstvo (glejte Kanjuo Mrčela in Kohont 2024, 298–302).
- 2 V tem besedilu pišemo o algoritemski organizaciji in nadzoru dela, kar pomeni, da se delo in delovni procesi na digitalnih platformah upravljajo na podlagi algoritmov – torej s programiranimi pravili, ki samodejno določajo razporejanje nalog, merjenje učinkovitosti, nagrajevanje in sankcioniranje delavcev. Če tak nadzor poteka brez možnosti vpogleda ali pritožbe (kar je pogosta praksa), zmanjšuje delavčevo avtonomijo in povečuje psihosocialni pritisk. Algoritem tako postane digitalni »nadzornik«, ki deluje brez neposredne človeške prisotnosti.
- 3 Prispevek je bil pripravljen s finančno podporo Globalnega projekta za upravljanje podjetniških talentov – GETM4, HoE-MSCARISE-2022, sporazuma o dodelitvi sredstev – GAP-101130572. V okviru terenskega dela pri projektu januarja in februarja 2025 je avtorica sodelovala s čilskimi raziskovalci in raziskovalkami na treh univerzah: Universidad Adolfo Ibáñez, Universidad Autónoma de Chile in Universidad Católica de Chile. Sodelovanje je prispevalo k naboru literature in pridobivanju informacij, pomembnih za analizo.
- 4 Algoaktivizem označuje prakso političnega in družbenega delovanja, usmerjenega v razkrivanje in nasprotovanje škodljivim algoritmom. To vključuje razvoj alternativnih orodij (npr. aplikacij, ki razkrivajo napake v izračunih plač), kampanje za večjo transparentnost algoritmov ter spodbujanje zakonodaje, ki zahteva etično in pravično uporabo algoritmov pri organizaciji dela. Algoaktivizem je pogosto povezan z bojem platformnih delavcev za preglednost in pravičnost.
- 5 OECD (2024) ocenjuje, da bi zagotavljanje osnovne socialne zaščite za vse delavce in zmanjšanje stroškov formalnega zaposlovanja lahko zmanjšalo obseg neformalnega dela gospodarstva.
- 6 Po OECD (2025) ima Čile eno najvišjih stopenj digitalne povezljivosti v Latinski Ameriki, kar je posledica prizadevanj politike za odpravo digitalnega razkoraka. Vendar za drugimi državami članicami OECD zaostaja na več področjih, kot so digitalne spretnosti, uporaba digitalnih orodij v malih podjetjih, digitalne inovacije in interoperabilnost vladnih digitalnih storitev, kar ovira širše koristi digitalizacije.
- 7 25. oktobra 2019 je v Santiagu potekal pohod več kot 1,2 milijona ljudi.
- 8 Prispevek obravnava, kako so Čilenci črpali iz avantgardnega umetniškega etosa, da bi omogočili preoblikovanje javne sfere in novo skupno politično paradigmo.
- 9 INE uporablja naslednjo opredelitev za delo na platformah: »/d/elo, ki se opravlja z mobilno aplikacijo ali s spletno platformo, kadar se ponujajo blago in storitve izključno ali pretežno z uporabo nekega medija, ki vključuje stik s strankami na daljavo, bodisi po internetu (na spletni platformi) bodisi iz mobilnega telefona (v mobilni

- aplikaciji)» (Fuentes in Gonzalez 2022).
- 10 V zadnjem analiziranem obdobju raziskave je bilo 95.000 anketirancev, od katerih je bilo 38.000 zaposlenih oseb. Izmed teh jih je 832 poročalo o delu na digitalnih platformah. Predstavljeni podatki upoštevajo samo delavce, za katere je delo na digitalnih platformah glavna dejavnost (po času in dohodkih).
 - 11 V analizi dela dostavljavcev hrane v Veliki Britaniji in Čilu Bonhomme in Muldoon (2024) ugotavljata, da v primerjavi s platformnimi delavci v Veliki Britaniji platformni delavci v Čilu doživljajo več rasizma pri vsakdanjem delu ter da sta njihova rasa in migrantski status ključna pri definiranju njihove delovne realnosti.
 - 12 Borić je na volitvah 2021 premagal skrajno desnega kandidata Joséja Antonia Casta z obljubo progresivne politike, usmerjene k socialni vključenosti, širjenju družbenih pravic prebivalcev in varovanju okolja.
 - 13 Howson et al. (2020) tudi poročajo o delavski platformni kooperativi Despatronados (kar pomeni brez šefov).
 - 14 Javni diskurz v Latinski Ameriki lahko ponazori izjava bivšega brazilskega predsednika Bolsonara na začetku mandata: »Delavci bodo morali izbrati med več pravicami in več dela« (cit. v Howson et al. 2020).

CITIRANA LITERATURA

- Arriagada, Arturo, Macarena Bonhomme, Franciscobanez, in Jorge Leyton. 2023. »The Gig Economy in Chile: Examining Labor Conditions and the Nature of Gig Work in a Global South Country«. *Digital Geography and Society* 5.
- Asenjo Cruz, Antonia, in Alberto Coddou Mc Manus. 2021. »Economía de Plataformas y Transformaciones en el Mundo del Trabajo: Caso de Repartidores en Santiago de Chile«. *OIT Cono Sur – Informes Técnicos* 16.
- Bonhomme, Macarena, Funda Ustek-Spilda, in Arturo Arriagada. 2024. »Between Acceptance and Resistance: Conceptualising Migrant Platform Labour Agency in Chile«. *New Technology, Work and Employment*: 1–19. <https://doi.org/10.1111/ntwe.12308>.
- Bonhomme, Macarena, in James Muldoon. 2024. »Racism and food delivery platforms: shaping migrants' work experiences and future expectations in the United Kingdom and Chile«. *Ethnic and Racial Studies*. DOI: 10.1080/01419870.2024.2349268.
- Bravo, David, in Ernesto Castillo. 2021. *Estudio Longitudinal Empleo – Covid19: Datos de Empleo en tiempo real*. Centro UC de Encuestas y Estudios Longitudinales.
- Cano, Melissa Renau, Ricard Espelt, in Mayo Fuster. Morell. 2024. »How to Build Alternatives to Platform Capitalism?«. V *Capitalism in the Platform Age*, ur. S. Mezzadra, N. Cuppini, M. Frapporti in M. Pirone. Springer Studies in Alternative Economics. Cham: Springer. https://doi.org/10.1007/978-3-031-49147-4_15.
- Coop Cycle. 2024. Dostop 20. 1. 2025. <https://legacy.coopcycle.org/en/>.
- Fuentes, Adolfo, in Rikardo Gonzalez. 2022. *Radiografía a los trabajadores de plataformas digitales*. Agosto, Santiago de Chile: Laboratorio de Encuestas y Analisis Social.
- Gordon-Zolov, Terri. 2023. »Chile's Social Explosion and the Art of Protest«. *Sociologica* 17 (1): 41–55. <https://doi.org/10.6092/issn.1971-8853/16817>.
- Howson, Kelle, et al. 2020. »Just Because You Don't See Your Boss, Doesn't Mean You Don't Have a Boss: Covid and Gig Worker Strikes Across Latin America«. *International Union Rights* 27 (3).
- International Monetary Fund. 2023. *Report for Selected Countries and Subjects*. World Economic Outlook Database, oktober 2023.
- Kanjuo Mrčela, Aleksandra 2024. »Spremembe dela danes: od digitalizacije h kiborgizaciji

- dela«. V *Digitalizacija dela in življenja*, ur. Aleksandra Kanjuo Mrčela in Andrej Kohont. Ljubljana: Založba FDV.
- Kanjuo Mrčela, Aleksandra, in Andrej Kohont (ur.). 2024. *Digitalizacija dela in življenja*. Ljubljana: Založba FDV.
- L&E Global. 2022. Chile: *New Law Regulates Services Rendered from Digital Platforms, such as Uber, Cabify or Cornershop*. Dostop 20. 1. 2025. <https://www.lexology.com/library/detail.aspx?g=0271a121-2723-4dbb-b5e8-f61649d328f8>.
- Leyton, G. Jorge et al. 2021. *Fairwork Policy Brief: New regulation of Platform Work in Chile: a Missed Opportunity?* Dostop 20. 1. 2025. <https://fair.work/en/fw/publications/new-regulation-of-platform-work-in-chile-a-missed-opportunity/8>.
- OECD. 2024. »Resilience in Uncertain Times«. *OECD Economic Outlook 2024* (2). 4. december 2024.
- OECD. 2025. *OECD Economic Surveys. Chile 2025*. Januar 2025.
- OZN. 2015. *Sustainable Development Goals: The 17 Goals*.
- Prensa Presidencia. 2019. President Piñera announces a Social Agenda. 29. 10. 2019. <https://prensa.presidencia.cl/comunicado.aspx?id=123766>.
- Tokarczuk, Olga. 2024. *Popotovanje ljudi Knjige*. Prevedla Jana Unuk. Ljubljana: LUD Literatura.

DELIBERATIVNA POIZVEDBA KOT DEMOKRATIČNA TEHNOLOGIJA

Matjaž Maček

Deliberativno poizvedbo¹ (DP) (deliberative poll[®]) je na začetku 90. let prejšnjega stoletja razvil ameriški politolog James S. Fishkin s Stanforda (Fishkin 1991). Gre za metodo raziskovanja javnega mnenja, ki povezuje naključno izbrane vzorce prebivalstva z vnaprej strukturirano deliberativno razpravo. Namen te metode je ugotoviti, kakšno bi bilo javno mnenje, če bi bili državljani bolje informirani in bi imeli čas za poglobljen premislek o določenem vprašanju. DP ni zgolj raziskovalno orodje, temveč tudi demokratična inovacija: v nadzorovanem in empiričnem okviru uresničuje temeljna načela deliberativne demokracije in spada med osrednje pristope v tej paradigmi.

Deliberativna demokracija (DD) je normativna teorija politične legitimnosti, ki je danes vodilna paradigma med teorijami demokracije (Lafont in Urbinati 2024, 83; Selk 2023, 198). Deliberativna demokracija kot normativni in analitični okvir demokratičnega urejanja kolektivnega življenja predpostavlja, da legitimnost političnih odločitev ne izhaja zgolj iz preferenc posameznikov, izraženih z glasovanjem na volitvah, in izglasovane večine, temveč iz javno utemeljene razprave med enakopravnimi udeleženci.

DD temelji na ideji, da legitimne kolektivne odločitve izhajajo iz odprtega, racionalno utemeljenega razpravljanja med državljani, ne pa zgolj iz štetja glasov ali boja za oblast. Deliberativna demokracija poudarja, da morajo biti odločitve utemeljene s premišljenimi argumenti, usmerjenimi k skupnemu dobremu. Namesto da bi zgolj seštevala obstoječe preference, deliberativna demokracija zagovarja stališče, da se morajo preference oblikovati med razpravo, pri čemer naj bi bili posamezniki zavezani k temu, da jih »prepriča moč boljšega argumenta«, ne pa ozki lastni interesi ali pristranskosti.

To je teorija, ki izpostavlja pomen informirane, spoštljive in racionalne razprave med državljani in (ali) njihovimi predstavniki pred sprejetjem političnih odločitev. V njenem jedru je posebna koncepcija legitimnosti, po kateri so legitimni tisti zakoni, ki izvirajo iz javne premišljene razprave (deliberacije) državljanov. Z drugimi besedami, da je legitimna samo tista politična odločitev, ki lahko prestopi deliberativno preverjanje s strani tistih, ki jih ta odločitev obvezuje, ki se zato lahko imajo za njene soavtorje. Legitimnost odločitve zahteva, da je vpliv boljših razlogov na vsebino odločitve večji, ne da bi bila s tem kakorkoli okrnjena enaka glasovalna pravica. Kateri razlogi so boljši, se lahko izkaže samo v javni deliberativni razpravi kot »neprisilna prisila boljšega argumenta«. Javna deliberacija prispeva k demokratični legitimaciji, če omogoča državljanom, da sprejemajo zakone in politike, ki so jim podvrženi, kot vsaj razum-

ne. Deliberativna demokracija je oprta na prisotnost, artikulacijo in premislek množstva perspektiv. Vključevanje glasov najraznovrstnejših družbenih skupin je moralni temelj deliberativne demokracije.

Teorija² deliberativne demokracije prinaša dvojno inovacijo. Pri prvi gre za zamenjavo na volitvah temeljčega teritorialnega predstavništva z aleatoričnim, na naključnem izboru temelječem predstavništvu. Druga inovacija pa prenaša poudarek v raziskovanju demokracije z volitev, glasovanj in večinskega pravila – značilnega za agregacijske teorije demokracije (liberalno, utilitarno, ekonomsko, pluralistično), ki preference in interese obravnavajo kot v političnem procesu nespremenljive danosti – na njihovo oblikovanje in preobrazbo znotraj političnega procesa prek deliberativne komunikacije. Ne gre več za maksimiranje *izražanja* političnih preferenc, pač pa za maksimiranje državljanovih (z)možnosti za *oblikovanje* preferenc in presojo kot osnove za možne odločitve o javnih zadevah (Offe 2017, 23). Izražanje preferenc je sistemsko urejeno (volitve, glasovanja, svoboda govora), njihovo oblikovanje pa, skladno z liberalno doktrino, ne. Preference je pač treba najprej (iz)oblikovati in šele potem jih je mogoče izraziti. Medtem ko se liberalna demokratična teorija nanaša na hotenje ljudi, da izrazijo svoje preference, se deliberativna demokratična teorija osredinja na (trans)formativno dinamiko, na to, kako ljudje pridejo do »pravih« (v primerjavi s ‚surovimi‘, samoniklimi, nereflektiranimi) preferenc in pojmovanj interesov – to pa so preference, ki so bile ustrezno pretehtane in so prestale preizkušnjo spornih in oporekanih argumentov v kritičnem dialogu. »Prave« preference so, torej, diskurzivno preverjene in potrjene. Deliberativna demokracija je oblika politične komunikacije, ki naj bi omogočala oblikovanje preferenc in interesov skladno s tako politično racionalnostjo, ki upošteva dejstva in dokaze, vključuje prizadete, išče skupno dobro z oziranjem na druge ter gleda v prihodnost onstran mandatne kratkovidnosti. *Demokratična deliberacija* pa je njena aplikacija na mikroravni, torej praktično institucionalna uporaba načel in dognanj deliberativne demokracije pri dizajniranju, organiziranju in vodenju deliberativnih forumov kot javnih prostorov deliberacije. Ti forumi so naprave, ki omogočajo strukturirano poglobljeno razpravo (deliberacijo), v kateri sodelujejo naključno izbrani udeleženci. Forume, ki prakticirajo demokratično deliberacijo, imenujemo deliberativne minijavnosti (DMJ). Mednje spada tudi deliberativna poizvedba (DP).

DELIBERATIVNA MINIJAVNOST

Demokratične inovacije (DI) spadajo med prizadevanja za demokratizacijo realno obstoječe liberalne demokracije tako, da problematizirajo utečene načine in oblike političnega komuniciranja in odločanja ter ustvarjajo možnosti za institucionalno preobrazbo v smeri širjenja in poglobljanja participativnih možnosti. So odgovor na zaznani prepad med pričakovanji in zahtevami državljanov ter neodzivnostjo institucij liberalne demokracije. Kognitivno disonanco med načeli in prakso demokracije razrešujejo DI z vztrajanjem pri normativnih načelih demokracije in s postavljanjem pod vprašaj udeležene oblike in prakse demokracij ter predlaganjem drugačnih oblik. DI so procesi ali institucije, ki na novo osmišljajo in poglobljajo vlogo državljanov v demokratičnem političnem procesu oblikovanja kolektivne presoje in volje z

vzpostavljanjem priložnosti za participacijo, bodisi neposredno bodisi deliberativno, z namenom vplivanja na obvezujoče politične odločitve.

Deliberativne minijavnosti (DMJ) so ena od oblik demokratičnih inovacij, v katerih prihaja do prepletanja političnih praks reprezentiranja, temelječega na ključnem izboru z epistemskimi praksami, s katerimi so prve teoretsko uokvirjene in eksperimentalno-tehnološko konfigurirane. Demokratične inovacije nastanejo kot sistematično teoretiziranje (*logia*) zmogljivosti za izdelavo (*tehne*) kolektivne volje ljudstva (*demos*) za namen izvajanja oblasti (*kratos*) (Voss 2024, 68). Deliberativne minijavnosti omogočajo reprezentativno, informirano in enakopravno razpravo o kompleksnih javnih vprašanjih. Njihova moč je v tem, da združujejo državljansko angažiranost s strukturiranimi pogoji deliberacije, kar povečuje legitimnost in kakovost demokratičnega odločanja. DMJ so možni institucionalni odgovor na krizne pojave demokracije, ki se kažejo kot izguba zaupanja v elite in institucije, zmanjševanje politične participacije (upadanje članstva v političnih strankah in udeležbe na volitvah) in rastočega odtujevanja državljanov od politike. DMJ so tudi obogatitev participativnega portfelja, ki je na voljo državljanom.

Deliberativne minijavnosti so strukturirane platforme, na katerih naključno izbrani državljani razpravljajo o javnih zadevah. Njihova zasnova se močno razlikuje glede na cilje, vire in politični kontekst. Obstaja več opredelitev DMJ (Ryan in Smith 2014). Lahko izhajamo iz naslednje: DMJ so demokratične inovacije, ki so z načeli (usmerjevalnimi vrednotami) in s pravili (z operacionaliziranimi načeli) urejevani javni forumi, na katerih naključno izbrani udeleženci v razpravi kolektivno skrbno tehtajo razloge in argumente o opcijah v zvezi z obravnavano temo z namenom oblikovanja skupnih javnih stališč ali priporočil,³ s pomočjo moderatorjev in ob sodelovanju izvedencev, zainteresiranih, prizadetih in prič v vlogi dobaviteljev relevantnih informacij. DMJ naj bi omogočale preseči prepad med demokratično teorijo in prakso tako, da omogočajo večji vpliv državljanov v vseh fazah političnega procesa: pri določanju dnevnih redov, v odločanju o javnih politikah in udejanjanju sprejetih odločitev. Glede velikosti, to je števila udeležencev, naj bi DMJ bile dovolj majhne skupine, da bi lahko bile resnično deliberativne in dovolj reprezentativne, da bi lahko bile resnično demokratične (Goodin in Dryzek 2006, 220). Za hibridne oblike DMJ je značilno, da so med udeleženci tudi osebe, ki niso naključno izbrane, ampak so po nekem drugem kriteriju postale polnopravni udeleženci. Pri takih DMJ je večinski del udeležencev naključno izbran, drugi del pa je delegiran.

DMJ niso splošnonamenske naprave za izdelavo kolektivnih mnenj, presoj ali odločitev. Njihova prednost in uporabnost (specialnost) pride do izraza predvsem pri naslednjih vrstah problemov (OECD 2020, 27):

1. Razrešitev problema je odvisna od razrešitve dileme, katere vrednote naj bodo vodilo pri iskanju rešitve za problem. DMJ je okolje, v katerem lahko poteka razprava o npr. težavnih etičnih problemih, ki so brez očitne ali 'prave' rešitve, na dostojen in produktiven način, ki omogoča iskati in najti stične točke.
2. Kompleksni problemi, ki zahtevajo pretehtane izbire (*»trade-offs«*). Postopek demokratične deliberacije daje udeležencem, ki so, glede na obravnavani problem, laiki, čas, da se informirajo, učijo, reflektirajo in deliberirajo ter da se posve-

tujejo z raznimi informatorji in deležniki. Šele ti procesi (pridobljene vednosti in znanja) jih opolnomočijo, da se lotijo kompleksnosti problema ob upoštevanju pravnih, regulacijskih in proračunskih omejitev.

3. Dolgoročni problemi, ki presegajo kratkoročne spodbude volilnega ciklusa. Marsikatero odločitev o javni politiki je težko sprejeti, ker se njene koristi pokažejo šele v daljšem času, po končanju mandata trenutnih odločevalcev, stroški pa nastanejo v tekočem mandatu. DMJ olajšajo upravičevanje začenjanja vlaganj v dolgoročne projekte s sestavo DMJ, ki ne omogoča vpliva motiviranih interesov političnih strank in volitev, po drugi strani pa spodbuja udeležence, da ravnajo v interesu javnega dobra.
4. Dodatno spada zraven tudi kvaziustavni problem, ki se nanaša na »odločitev, kako odločati«, s katero se določajo pravila odločanja. Na primer, izbira volilnega sistema.

Navedeni problemi so podobni tistim, ki so v literaturi znani kot »wicked« (hud, zapleten, vražji), ker so slabo definirani, se še razvijajo, so oporekani, fokus je na simptomih namesto na vzrokih, spremlja jih nestrinjanje, zaradi katerega so možne rešitve neučinkujoče, ni ustreznega znanja za njihovo udejanjitev. Družbeni problemi so na splošno slabo definirani in ne dopuščajo soglasnih rešitev. Zato se pri reševanju teh problemov zanašajo raje na politične presoje kot pa na znanstvene gotovosti (Head 2022). DMJ pa so primerni prav za ukvarjanje s takimi problemi.

Praksa kaže, da DMJ praviloma nastanejo kot začasne, ad hoc, oblastno pokroviteljsko (od zgoraj) iniciirane (*»governance driven«*) projektne organizacije kot povabljeni (*»invited«*) prostori. Do prekinitve začasnosti pride, če začasna, ad hoc DMJ dobi z legalizacijo status stalne institucije, vključene v politični sistem, ali pa ugasne po opravljeni nalogi. Redkeje se zgodi, da DMJ nastane od spodaj (*»citizen driven«*) kot samoorganizirana in samofinancirana pobuda državljanov, kot iznajdeni (*»invented«*) prostor (Bua, Bussu in Davies 2023).

Cilji DMJ so dokaj podobni: izdelava javnih presoj o perečih, kontroverznih in kompleksnih političnih problemih, spodbujanje javnega diskurza, svetovanje političnim odločevalcem, razjasnjevanje pogledov na kompleksne politične probleme, določanje političnih prioritet, preseganje političnih blokad, razsojanje med političnimi opcijami, spodbujanje javne razprave in participacije, doseganje metakonsenza (drugotnega soglasja o prvinah in razlogih prvotnega nesoglasja) ter izobraževanje udeležencev in javnosti (Niemeyer in Dryzek 2007; Beauvais in Bächtiger 2016; Kuyper 2017). Teme, o katerih najpogosteje razpravljajo, so naslednje: infrastruktura, okolje in podnebje, reforma volilnega sistema, zdravstvo, urbanizem, strateško načrtovanje itn.

FORMATI DEMOKRATIČNE DELIBERACIJE

Začetki demokratične deliberacije segajo na konec šestdesetih in začetek sedemdesetih let prejšnjega stoletja, ko sta Peter Dienel in Ned Crosby iznašla, neodvisno drug od drugega, demokratično deliberacijo *»avant la lettre«*. Najprej je nemški sociolog Peter Dienel iznašel načrtovalno celico (*»die Planungszelle«*) kot

nov postopek za načrtovanje mestnega razvoja, ki uporablja naključni izbor kot mehanizem za sestavo državlanskega foruma, razpravo o kontroverznih načrtovalnih problemih in vlogo moderatorja kot usmerjevalca in spodbujevalca procesa. Nekaj let za njim je ameriški politolog Ned Crosby leta 1974 organiziral prvo naključno izbrano državlansko poroto (*»citizens' jury«*), za obravnavo načrta zdravstvene oskrbe v Minnesoti. V osemdesetih letih je Danski upravni odbor za tehnologijo začel organizirati konsenzne konference (*»consensus conference«*), s katerimi je svetovalno pomagal parlamentu pri tehnološko zahtevnih zadevah. Ti prvotni formati niso temeljili na kaki eksplicitni teoriji, ampak so bili bolj oportunistični *»bricolage«* razpoložljivega znanja in praktičnih izkustev, izoblikovanih v skupnostih praks, bili so predhodnica današnjih DMJ, ne samo zaradi naključnega izbora, pač pa tudi zaradi prakticirane deliberacije, kakršna je že bila, in so bili namenjeni uporabi pri tehnološko in infrastrukturno zahtevnih projektih. Ti trije državljanski forumi so uporabljali skorajda identični proceduralni dizajn. Izhajali so iz podmene, da je treba državljanom omogočiti konstruktivno prispevati k oblastvenemu odločanju, tako da povečajo produktivnost javnega mnenjskega udejstvovanja z organizacijsko podporo.

Izraz deliberativna demokracija naj bi prvi uporabil, Joseph Bessette, ko je leta 1980 objavil članek *Deliberative democracy: the majority principle in republican government* v zborniku, v katerem je reinterpreteriral ustavo ZDA kot načrt deliberativne demokracije z reafirmiranjem Madisonove teze, da so izbrani predstavniki sposobnejši prepoznati skupno dobro kot državljani sami (Goldwin in Schambra 1980). Robert Dahl je leta 1989 v *Democracy and its Critics* omenil *»minipopulus«*, ki bi štel okoli 1000 članov in bi bil naključno izbran iz celotnega ljudstva z nalogo, da deliberira eno leto na zadano temo in sporoči svojo odločitev. Drugi minipopulus bi izbral teme za dnevni red prvega minipopulusa. Minipopulus bi bil dopolnilo za obstoječo legislativo. Presoja minipopulusa bi *»predstavljala«* presojo demosa (Dahl 1989).

Demokratske inovacije začnejo nastajati v večjem številu in z večjo pomembnostjo za demokratski proces v zadnjih dvajsetih letih. Za izhodiščni datum lahko vzamemo leto 2004, ko je bila v kanadski provinci Britanska Kolumbija izvedena prva državljanska skupščina (DS) na svetu, ki je bila vzorec za mnoge poznejše DS: The British Columbia Citizens' Assembly (Warren in Pearse 2008) z nalogo provincijske vlade, da preuči ustreznost obstoječega (relativno večinskega) volilnega sistema. Njej sledita leta 2006 še državljanski skupščini v kanadski provinci Ontario in na Nizozemskem, ki sta se obe prav tako ukvarjali z vprašanjem primerne volilnega sistema: tudi v Ontariu je bil postavljen pod vprašaj relativno večinski sistem, na Nizozemskem pa sorazmerni sistem (Fournier et al. 2011). Zadnji primer (2024) uporabe DS za izbiro volilnega sistema je odločitev DS na Yukonu, ki predlaga zamenjavo relativno večinskega sistema s sistemom enega prenosljivega glasu (STV).⁴ Tem zgledom je sledil plaz posnemanj formatov DMJ za obravnavo različnih tem, ki se je z leti samo še povečeval.

Tako posnemanje ni vedno izomorfno uporabljalo zgolj formata DS, pač pa tudi druge formate, ki spadajo v družino tako imenovanih deliberativnih minijavnosti. Naj naštejemo samo najbolj uporabljane: *»citizens' assembly«* (državlanska skupščina), *»citizens' jury«* (državlanska porota), *»consensus conference«* (konsenzna konferenca), *»Bürgererrat«* (državlanski svet), *»deliberative poll®«* (deliberativna poizvedba), *»Planungszelle«* (načrtovalna celica), *»citizens' initiative review«* (pregled državljanske

pobude – CIR) itn. Inovacij je bilo toliko, da študija OECD (2020) govori o »deliberativnem valu«. Navaja tudi, da je bilo med letoma 1986 in 2019 izvedenih 282 DMJ v državah OECD. Iz podatkovne zbirke POLITICIZE pa je razvidno, da je bilo v Evropi med letoma 2000 in 2020 izvedenih 105 DMJ v 18 državah. V vzhodni Evropi v Estoniji in Romuniji. Pozneje še v Bolgariji, Bosni in Hercegovini ter Severni Makedoniji. V Belgiji je bilo v letih 2001–2019 izpeljanih 43 DMJ, med njimi tudi doslej največja, G1000 s 704 udeleženci (Vrydagh et al. 2021). V Franciji sta bili izvedeni veliki javni razpravi, ki pa nista bili čisto pravi deliberaciji: Le Grand Débat National 2019 (velika nacionalna debata) kot odgovor na Rumene brezrokavnike in La Convention Citoyenne pour le Climat (CCC) 2021 (državlanski zbor o podnebjju). V Nemčiji je bilo od 1972 do septembra 2024 izvedenih 298 državljskih svetov oz. državljskih skupščin na občinski, deželni ali zvezni ravni, na slednji trije: o demokraciji, o vlogi Nemčije v svetu in o prehrani.⁵

Na platformi PARTICIPEDIA je trenutno opisanih 373 metod DI in 2264 primerov njihove uporabe iz 159 držav. Zelo verjetno je bilo izvedenih DMJ in DI več, kot je razvidno iz navedenih podatkov zaradi različnih poimenovanj in različnih kriterijev razvrščanja in uvrščanja demokratizacijskih prizadevanj. V zadnjih nekaj letih prihaja do množične uporabe DMJ pri obravnavanju okoljske in podnebne problematike (Smith 2024).

DELIBERACIJA

Pojem deliberacija je temeljna kategorija teorije deliberativne demokracije. Je teorijska osnova in modus operandi deliberativnih minijavnosti, med katere spada tudi deliberativna poizvedba. Beseda deliberacija ima dva možna izvora v latinščini (Boniolo 2012). Prvi izvor je glagol *deliberare*, ki ima dva pomena: 1. osvoboditi, rešiti od, v smislu osvobajanja od (pedsodkov, dogem, pristranskosti, iluzij, egocentričnosti ipd.); 2. preudarjanje, premišljevanje, (raz)reševanje, razglabljanje. Drugi izvor pa je glagol *delibrare*, ki ima v korenu besedo *libra*, ki pomeni tehtnico in gre torej za tehtanje, pretehtavanje, uravnovešanje, balansiranje. Oxfordski priročnik o deliberativni demokraciji navaja naslednjo opredelitev deliberacije, ki se glasi: deliberacija je »vzajemna komunikacija, ki vključuje tehtanje in reflektiranje preferenc, vrednot in interesov glede na zadeve skupne skrbi« (Bächtiger et al. 2018, 2).

Deliberacija je na splošno razumljena kot skrbno kolektivno tehtanje različnih, tudi konkurenčnih preferenc, interesov, argumentov in razlogov v iskanju demokratičnega soglasja⁶ med razpravo, ko je treba odgovoriti na praktično vprašanje, kaj storiti z javnim problemom.

Deliberacija ni samo diskusija (kolektivna izmenjava informacij med vrstniki), debata (kritična nasprotovalnost) in dialog (vzajemno razumevanje in učenje), ampak je kombinacija vsega tega v obliki s pravili urejenega javnega posvetovanja.⁷

Deliberacija lahko poteka kot kognitivni proces tehtanja razlogov za in proti v glavi deliberatorja, kot pogovor s samim seboj, kot deliberacija v osamitvi (Goodin 2003) ali pa kot poliloška komunikacija med udeleženci deliberacije. Pomanjkljivost monološke deliberacije je v tem, da je, po eni strani, enoperspektivna in da se posameznik sam, kljub najboljši volji, težko izogne nevarnosti motiviranega sklepanja in

se osvobodi svojih pristranskosti in predsodkov zaradi omejene zmožnosti samokritičnosti. Po svoje je tudi razpravljanje v zaprtem krogu istomislečih, enklavna razprava, vrsta monološke deliberacije. Po drugi strani pa zgreši bistvo deliberacije, ki je v kolektivnem soočanju mnenj in argumentov in njihovem vzajemnem obravnavanju in predelovanju s ciljem iskanja skupne rešitve. Demokratična deliberacija je politični proces, ki je naravnán na oblikovanje skupne presoje in odločitve o javni zadevi in je usmerjan z načeli deliberacije.

Katera so poglavitna načela, ki naj bi urejala deliberativni proces in zagotavljala deliberativnost deliberacije? Različni avtorji izpostavljajo različna načela in njihove kombinacije. V nadaljevanju so domnevno navedena vsa poglavitna načela, ki se pojavljajo v literaturi in jih kot celote ne najdemo pri nikomer (Manin 1987; Cohen 1989; Fishkin 1991; Habermas 1992; Gutmann in Thompson 2004; Boniolo 2012; Chappell 2012; Kies 2012; Steiner 2012; Bächtiger et al. 2018; Neblo et al. 2018; Bächtiger in Parkinson 2019; O'Flynn 2022).

Vključenost

Deliberativna demokracija zahteva, da imajo vsi, ki jih odločitev zadeva, pošteno priložnost za enakopravno sodelovanje v razpravi. Gre za normativno predpostavko demokratične enakopravnosti in epistemske pluralnosti, saj različni glasovi prispevajo k celovitejši obravnavi javnih problemov. Vključenost pomeni preseganje elit ali že obstoječih vplivnih skupin ter vključevanje raznolikih in marginaliziranih glasov v proces. Široko sodelovanje vodi do legitimnejših rezultatov: čim večja je vključenost državljanov in pogledov, tem legitimnejši in razumnejši so politični izidi. S tem ko v razpravo vključimo širok spekter perspektiv, lahko skupina bolje upošteva različne izkušnje in vrednote. Razprava, ki vključuje predstavnike vseh glavnih stališč glede nekega vprašanja, zagotavlja, da končna odločitev odraža pošteno obravnavo vsakega pogleda. Vključenost prispeva k legitimnosti deliberativnega sistema, saj zmanjšuje tveganje izključevanja in sistemske pristranskosti. Deluje kot korektiv institucionalne zaprtosti in omogoča, da deliberacija reflektira heterogenost družbenega prostora.

Enakost

Z vključenostjo je tesno povezana enakost sodelovanja. Enakost se nanaša na enako možnost vpliva v procesu deliberacije – ne le formalno (vstop), temveč tudi vsebinsko (glas, vpliv). Po teoriji deliberacije je treba vsak prispevek obravnavati enakovredno – nihče ne sme biti izključen ali prezrt in nobenega argumenta ne smemo zavrniti na nesprejemljivi podlagi (kot so družbeni status, kulturna identiteta in premoženje). V praksi to pomeni, da je treba razpravo strukturirati tako, da ima vsak glas enako možnost vpliva, argumenti pa naj se ocenjujejo po vsebinski vrednosti ne glede na to, kdo jih izreče. Udeleženci morajo drug drugega obravnavati kot politično enakopravne in vredne spoštovanja, kar preprečuje prevlado in zagotavlja, da so tudi nasprotujoča stališča slišana.

Recipročnost in spoštovanje

Recipročnost (vzajemnost) pomeni da državljani morajo svoje predloge utemeljiti z razlogi, ki jih lahko drugi, kot svobodni in enaki posamezniki, razumno sprejmejo.

To izraža zavezanost medsebojnemu spoštovanju. Deliberacija zahteva, da se udeleženci vključujejo v »medsebojno spoštljivo razmišljanje«, kar pomeni, da argumentov ne ponujajo zato, da bi druge pregovorili, temveč da bi jih prepričali na skupni osnovi. Vzajemnost pomeni tudi spoštovanje pogledov drugih kot legitimnih in obravnavanje sogovornikov kot enakovrednih – tudi kadar se z njimi ne strinjamo. Kot je navedeno v enem izmed uvodov v deliberativno demokracijo, lahko državljani, ki razmišljajo vzajemno, priznajo, da je stališče, s katerim se ne strinjajo, kljub temu »vredno moralnega spoštovanja«, četudi se jim zdi zmotno. V praksi načelo vzajemnosti pomeni, da se tako odločevalci kot državljani vzdržijo oblikovanja politik na podlagi ozkih strankarskih ali drugih partikularnih razlogov. Namesto tega oblikujejo zakone in politike na način, ki jih lahko tudi drugi državljani prepoznajo in priznajo kot poštene. Vzajemnost in spoštovanje skupaj zagotavljata, da deliberacija ostaja sodelovalno iskanje utemeljenega soglasja, in ne enostransko vsiljevanje. Recipročnost prispeva h kognitivni koherenci deliberativnega procesa: utemeljitve, ki upoštevajo perspektive drugih, zmanjšujejo konfliktnost in povečujejo možnost konsenza ali vsaj vzajemnega razumevanja. V sistemskem smislu to podpira usklajevanje mnenj. Spoštovanje označuje priznavanje drugega kot moralnega in epistemskega sogovorca. Je ključno za zagotovitev varnega prostora za izražanje in sprejemanje idej.

Preglednost in javnost

Preglednost pomeni, da so postopki, kriteriji odločanja in informacije dostopni in razumljivi javnosti. Zasnovana je na ideji odgovornosti in nadzora. DD poudarja, da mora deliberacija potekati v javnosti ali v javno odgovornih forumih. Za legitimno odločanje je pomembna preglednost glede vprašanj in postopkov. Deliberativni procesi morajo biti odprti za nadzor, da državljani vedo, kako, zakaj in čemu so bile odločitve sprejete. V praksi to pomeni, da razprave potekajo v javnem prostoru, v katerem jih lahko vsakdo opazuje ali v njih sodeluje, dnevni red in pravila pa so jasno določeni. Konkretnije, deliberativna načela zagovarjajo uporabo javnega (raz)uma – to so argumenti, utemeljeni na skupnih vrednotah in preverljivih dejstvih, ne pa na skritih ali zasebnih avtoritetah. Dravaljani se ne bi smeli zanašati na nepreverljive trditve, ki so dostopne le nekaterim, te lahko v javni razpravi preverijo vsi. Pereča vprašanja morajo biti javno izpostavljena in o njih je treba javno razpravljati. Ta javnost zagotavlja odgovornost: ker se odločitve sprejemajo na odprt način in se utemeljujejo z jezikom, razumljivim vsem, je proces pošten, poleg tega pa ostaja mogoče izvajati popravke ali spremembe, kadar je to treba. Preglednost podpira odgovornost deliberativnih teles in omogoča refleksivno nadzorovanje sistema. Prispeva k širši javnosti deliberacije in povečuje legitimnost celotnega sistema.

Utemeljevanje in odgovornost

Utemeljevanje pomeni, da udeleženci političnega procesa ne izražajo le svojih preferenc, ampak jih utemeljujejo z razlogi, ki so normativno in racionalno sprejemljivi. Z idejo javnosti je tesno povezana tudi zahteva, da so odločevalci (in širše tudi vsi udeleženci razprave) odgovorni tistim, na katere odločitve vplivajo, in da morajo javnosti utemeljiti svoje odločitve. Osrednja deliberativna norma je, da nihče ne sme biti obravnavan zgolj kot objekt politike; zakoni morajo biti podprti z razlogi, ki jih

državljeni prepoznajo kot zavezujoče. Ko državljani ali njihovi predstavniki sprejemajo temeljne politične odločitve, morajo biti pripravljeni te odločitve pojasniti in utemeljiti z vidika skupnih vrednot.

Načelo javne utemeljitve pomeni, da legitimnost temelji na navajanju razumnih razlogov drugim. Deliberativna legitimnost ne temelji na sili, temveč na navajanju razlogov: državljani si medsebojno podajajo razloge v procesu vzajemne izmenjave argumentov, kar pomeni, da so sprejete odločitve utemeljene za vse. Tudi če popolno soglasje ni doseženo, naj bi se udeleženci čutili slišane, njihova stališča pa obravnavana. S tem ko se od politikov (in navadnih državljanov) zahteva odgovornost do javnosti z razumnim utemeljevanjem, zasleduje deliberativna demokracija cilj, da se oblast izvaja zgolj z argumenti, ki so sprejemljivi za vse. Utemeljevanje je osrednja epistemska funkcija deliberacije – spodbuja preučevanje alternativ, analizo posledic in normativno refleksijo. Sistemskemu delovanju omogoča kritično presojo legitimnosti rešitev.

Refleksija (samorefleksija) in empatija

Refleksija pomeni pripravljenost in sposobnost posameznika, da kritično preizprašuje svoja prepričanja, preference in vrednote, še posebej ob soočenju z novimi informacijami ali nasprotnimi stališči. Njena vloga v deliberativni demokraciji je, da:

- preoblikuje preference: namesto zgolj seštevanja že oblikovanih mnenj na podlagi glasovanja želi deliberativna demokracija izboljšati in preoblikovati preference z argumentirano razpravo;
- spodbuja odprtost: državljani so povabljeni, da dvomijo o svojih predpostavkah, ocenjujejo dokaze in razmislijo o alternativah;
- izboljšuje kakovost odločitev: reflektirano sodelovanje vodi do bolj premišljenih in utemeljenih odločitev;
- spodbuja transformativni potencial deliberacije, saj udeležencem omogoča, da preoblikujejo razumevanje problemov, ter podpira fleksibilnost in učenje v političnem odločanju.

Empatija je sposobnost, da razumemo in sočustvujemo z izkušnjami drugih, še posebej tistih, katerih ozadje ali položaj je drugačen. Vloga v deliberativni demokraciji:

- spodbuja medsebojno spoštovanje: empatija pomaga državljanom, da prepoznajo legitimnost skrbi drugih, tudi če se z njimi ne strinjajo;
- podpira vključenost: spodbuja pripravljenost, da prisluhnemo marginaliziranim glasovom in upoštevamo njihove izkušnje;
- zmanjšuje polarizacijo: empatično sodelovanje lahko omehča sovražnost, spodbuja kompromise in krepi zaupanje;
- izboljšuje moralno presojo: empatija omogoča posameznikom, da premišljajo širše od lastnega interesa in se zavzamejo za skupno dobro;
- deluje kot mehanizem vključevanja čustvene in izkustvene racionalnosti, ki dopolnjuje formalno racionalnost argumentov; v sistemskem smislu omogoča horizontalno povezovanje med raznolikimi glasovi in krepi socialno kohezijo.

Ključno vprašanje je, kako razumeti razmerja med normativnimi načeli. Ali med temi načeli obstajajo hierarhija, kontekstualno tehtanje prednosti, funkcionalno zaporedje ali pa delujejo kot mreža vzajemno pogojujočih se vrednot? Razmerja so lahko različna, kar je odvisno od konteksta, namena in cilja demokratične deliberacije. Poznane so štiri ureditve razmerij med načeli.

1. Hierarhija leksikalne prednosti (*»lexical ordering«*): V Rawlsovem modelu se načela razvrstijo v strogo zaporedje: višje načelo mora biti v celoti zadovoljeno, preden pride na vrsto naslednje. V deliberativni demokraciji bi to pomenilo, da ima npr. vključitev vseh relevantnih glasov prednost pred kakovostjo utemeljevanja. Vendar ta model pogosto ne ustreza kompleksnosti deliberativnega procesa, saj ne omogoča prilagodljivosti razmeram.

2. Tehtana prednost (*»weighted priority«*): Načelom se pripiše relativna teža (npr. z utežmi), ki velja v določenem kontekstu. To omogoča fleksibilno uravnavanje: ni absolutne prednosti, ampak kontekstualna prevlada. Na primer: v deliberaciji o občutljivih vprašanjih (npr. migracije) lahko spoštovanje in empatija dobita prednost pred transparentnostjo, da se zaščiti osebno dostojanstvo razpravljavcev. V deliberativnih forumih lahko enakost govornega časa začasno relativizira epistemsko moč argumentov, če to prispeva k enakovrednejši udeležbi.

3. Funkcionalna hierarhija načel: Načela se lahko razvrstijo po njihovi strukturni vlogi v procesu: vključenost in enakost kot predpogoja, recipročnost in spoštovanje kot pravili interakcije, utemeljevanje in refleksija kot kazalnika kakovosti, empatija kot socialna podlaga za medosebno razumevanje. Tako se oblikuje sekvenčna logika, v kateri določena načela omogočajo delovanje drugih. Na primer: brez vključujoče sestave udeležencev ni mogoče pričakovati reprezentativnosti deliberacije.

4. Mrežni (topološki) oziroma konstelacijski model: Dryzek in Niemeyer (2008) predlagata mrežni pristop, po katerem načela delujejo kot vzajemno pogojeni elementi. Ni enega »najvišjega« načela, temveč se njihova teža in relacije oblikujejo znotraj konkretnih diskurzivnih situacij. Na primer: transparentnost prispeva k utemeljevanju, a lahko zmanjšuje izražanje empatije v konfliktnih situacijah. Enakost udeležbe poveča legitimnost, a lahko oslabi poglobljenost argumentov, če sposobnost izražanja udeležencev ni enaka. Deliberativna demokracija ne predpostavlja enotne hierarhije načel. Njena normativna moč izhaja iz zmožnosti za refleksivno uravnavanje razmerij med načeli v konkretnih situacijah. Namesto stroge linearne logike ponuja deliberativna teorija pluralne modele, ki vključujejo leksikalno prednost, tehtano odločanje, funkcionalno sekvenco in mrežno soodvisnost. Ključna ostaja institucionalna zmožnost zaznavanja, predvidevanja in razreševanja normativnih napetosti med načeli.

Prakticiranje deliberativnih načel omogoči prehod udeležencev iz navzočih⁸ v aktivne sodelovalne. Proces deliberiranja je uspel, če je potekal kot niz »neprisilnih prisil boljših argumentov« (Habermas 1992, 370).

Deliberativna načela opravljajo v procesu deliberacije dve pomembni funkciji na poti do »izdeliberirane« kolektivne presoje in odločitve, ki sta specifični za demokratično deliberacijo. Najprej gre za filtrirno, potem pa za integrativno funkcijo (Chambers 2001).

Pri filtrirni funkciji gre za preverjanja preferenc, interesov itn., ali te vzdržijo preizkus javne deliberacije z vidika splošne sprejemljivosti.⁹ Ko mora udeleženec zagovarjati svoje preference pred forumom DMJ ali celo širšo javnostjo, mora dodobra premisliti, čemu ima prav to mnenje ali zasleduje prav te preference. To zahteva od njega poglobljen razmislek. Tu ni prostora za zgolj osebne interese, predsodke, neubranljive predpostavke, dezinformacije ali nevednost, ker jih kritična debata razkrije kot nesprejemljive. Vedno seveda obstaja možnost pretvarjanja in sprenevedanja in je prenekateri domnevni posplošeni argument zgolj zamaskirani zasebni interes, ker v javnih razpravah (zlasti političnih) ljudje pogosto ne morejo priznati sebičnih interesov ali nelegitimnih motivov, zato morajo dejanja in zahteve upravičiti z razlogi, ki so videti javno sprejemljivi (sklicevanje na pravico, enakost, svobodo ipd.). Tudi če nekdo laže ali manipulira, ga prisila javne utemeljitve sili, da govori jezik skupnega dobra, moralnosti, racionalnosti. To početje pa je dvorezno, kajti če nekdo tako ravna dalj časa in nato pride s pravo barvo na plan, lahko to škoduje njegovi verodostojnosti, kar je lahko zanj previsoka cena, zato je prisiljen pretvarjati se še naprej. Jon Elster imenuje ta mehanizem »civilizacijska sila hipokrizije«.

Zato je kritična razprava še kako pomembna kot mehanizem za »pranje« preferenc (Goodin 1986) tako, da lahko nadaljujejo deliberativno razpravo samo »oprane«, prečiščene preference. Dejavnost dajanja in poslušanja premišljenih argumentov lahko sproži proces učenja, samospreminjanja, ki udeležence pripelje do tega, da začnejo opuščati prvotne pozicije, za katere ne obstajajo javno posplošljivi ali prepričljivi razlogi, in zavzemati drugačne.

Integrativno funkcijo opravljajo načela empatičnosti, recipročnosti in vzajemnega spoštovanja, ko spodbujajo medsebojno priznavanje in razumevanje udeležencev pri zasledovanju cilja doseganja kolektivne presoje in odločitve. To pa je mogoče, če se udeleženci ne vedejo strateško racionalno (to je instrumentalno, na podlagi individualne kalkulacije koristi in učinkovitosti brez ozira na druge) z namenom maksimiranja svojih koristi, pač pa diskurzivno, z namenom razumnega prepričevanja drugih iz perspektive argumentativnega iskanja skupnega dobra. V demokratični deliberaciji je treba biti odgovoren do sogovornika in pri njem razumljen, ne pa premagati ga. Deliberativnost deliberacije se izkaže, če je v procesu deliberiranja prevladovalo razumno prepričevanje (moč argumentov) nad argumentom moči, če je bila zagotovljena enaka, ne pa privilegirana, participacija, če je proces potekal kot navajanje razlogov za izrečene trditve in če so bili udeleženci pripravljeni spremeniti svoje poglede ob soočenju s prepričljivimi argumenti.

DELIBERATIVNA POIZVEDBA

Deliberative Poll® (DP) je v obtoku od leta 1988. Od takrat je bilo izvedenih več kot 150 deliberativnih poizvedb v več kot 50 državah po vsem svetu. DP lahko razumemo kot napravo za izdelovanje premišljenih kolektivnih mnenj. Namen postopka je prikazati presojo javnosti o zapletenih ali spornih vprašanjih javne politike.¹⁰ Vodilno, kontrafaktično (hipotetično) vprašanje, na katero skuša DP odgovoriti, se glasi: kaj bi javnost v resnici menila, če bi lahko v ugodnih razmerah (razpoložljivost relevantnih informacij, spodbudno okolje za razpravo in dovolj časa) in na podlagi

strukturirane poglobljene razprave (deliberacije) s sodržavljani, katerih stališča so različna, razmislila o teh vprašanjih? Kateri so razlogi, ki so pomembni za ljudi, ko se ti resnično osredinijo na vprašanja, ki jih skrbijo, in skupaj iščejo odgovore nanja? Ali je v tem konfliktnem in polariziranem svetu tak dialog sploh mogoč? DP nanja odgovarja pozitivno.

DP je družbena inovacija kot kombinacija družboslovnega eksperimenta (primerjava med eksperimentalno in kontrolno skupino) z javnomnenjskim poizvedovanjem in normativno teorijo o spremljajočih okoliščinah (robnih pogojih) razumne politične deliberacije. DP je dizajnirana tako, da se kar najbolj približa idealu deliberativne politične komunikacije in da omogoča tako politično oblikovanje mnenj, ki poteka kot postopkovno in procesno konstruiranje posebne vrste kolektivnega mnenja. Za Jane Mansbridge (2010) je DP zlati standard za deliberativne minijavnosti sploh.

Njena posebnost je v načinu izdelave reprezentativnega mnenja z deliberativnim posvetovanjem, katerega rezultat ni zgolj seštevek zasebnih mnenj kot pri običajnih javnomnenjskih poizvedbah, pač pa med razpravo rafinirano (poobčeno) mnenje ljudstva (ali druge relevantne populacije). Mnenje DP je posredovano v splošno razpravo s trditvijo, da so rezultati DP pravo mnenje javnosti, tisto, ki bi ga javnost izkazala, če bi imela možnost dodobra premisliti o problemu, ker njegovo oblikovanje ni bilo podvrženo nobenemu od dejavnikov, ki sicer vplivajo na oblikovanje mnenja tako, da ga izkrivljajo in popačijo, kot to velja za druge prakse reprezentiranja, kot so volitve, referendum, konvencionalne mnenjske raziskave, medijski diskurzi, družbeni aktivizmi, ki pogosto nastopajo kot samoimenovani pretendenti na zastopanje (*»representation claim makers«*) družbenih skupin, živali in okolja, ki trpijo pri pomanjkanju informacij, egocentričnosti, polarizaciji, površnosti, omejeni pretehtanosti in preudarjenosti.

Kritika običajnega, instantnega javnega mnenja

Fishkin (2009, 2–8) pride do pravega javnega mnenja na podlagi kritike pomanjkljivosti običajnega javnega mnenja.

1. Državljani v množičnih družbah niso motivirani, da bi se informirali. Stopnja njihove informiranosti o političnih zadevah je praviloma nizka. Nekateri to razlagajo z »racionalno nevednostjo« (Downs 1957). Volivcu se ne izplača vlagati napora v pridobivanje informacij, če pa njegov glas ni odločilen za izvolitev. Z njegovega gledišča sta iskanje in razumevanje informacij jalov posel.

2. Ljudje imajo manj »mnenj«, ki bi zaslužila to ime, kot je razvidno iz javnomnenjskih raziskav. Te izhajajo iz domneve, da ljudje mnenje vnaprej oblikujejo na zalogo in da komaj čakajo, da ga lahko izrečejo. V resnici pa mnenja dostikrat nimajo, a jim je nerodno, v družbi vednosti in znanja, priznati, da nečesa ne vedo. Zato kot odgovor na vprašanje raje izberejo nek odgovor, tako rekoč naključno, kot pa da bi odgovorili z »ne vem«. Če pa mnenje le imajo, je to lahko površno, dvoumni odziv na obljubaštvo in debelo natisnjene naslovnice, podvrženo tehnikam upravljanja z vtisi, ki jih je izpopolnila industrija prepričevanja, ki je kolonizirala politični proces.

3. Tretja pomanjkljivost je v tem, da ljudje razpravljajo o politiki samo v krogu enako mislečih, katerih socialno ozadje, socialnoekonomski status in pogledi so podobni. Z drugače mislečimi se pogovarjajo o vremenu. Nagnjenost k izogibanju

nestrinjanju in konfliktu je zasidrana v večini ljudi. Tendencia političnega razgovora iz oči v oči kaže na samoizbor med enako mislečimi.

4. Četrta pomanjkljivost javnega mnenja je njegova podvrženost manipulaciji. Z neangažirano in neinformirano javnostjo je lažje manipulirati kot s tako, katere mnenje je čvrsto, temelječe na premisleku in razpravi, ker je nestanovitnejša na individualni ravni, na kateri operira z vtisi na pamet ali celo s fantomskim mnenjem.¹¹ Slabo informirani ljudje so kaj lahko žrtve lažinovic, dezinformacij, strateškega zavarovanja s polresnicami in preprosto napačnih, netočnih informacij, katerih vir so tudi družbena omrežja.

Te pomanjkljivosti naj bi bile presežene prav z deliberativno poizvedbo.

Kako je organizirana in kako deluje deliberativna poizvedba?

Razlikovati je treba vzpostavitev projekta DP od njegove izvedbe. DP so praviloma enkratni ad hoc začasni projekti, ki ugasnejo, ko opravijo svojo nalogo. Odločitev o DP sprejme njen naročnik (praviloma je to javnopravna institucija), ki je običajno tudi njen financer. Naročnik določi namen in običajno tudi agendo DP. Pri tem lahko sodelujejo tudi drugi kot prireditelji, iniciatorji, pokrovitelji ali podporniki. Naročnik zaupa izvedbo DP organizaciji, specializirani za take posle, ki organizira in vodi celoten potek deliberativnega poizvedovanja, procesno strukturiranega po fazah.

DP je sestavljena iz dveh mnenjskih poizvedb: začetne preddeliberativne in zaključne podeliberativne ter procesa deliberacije, ki posreduje med obema.

Faze deliberativne poizvedbe

Fazna struktura izvedbe deliberativne poizvedbe se je v praksi izoblikovala kot sosledje postopkovnih korakov, ki skuša vzpostaviti kar najugodnejše razmere za deliberativno poizvedovanje z upoštevanjem deliberativnih načel (Fishkin in Luskin 2005; Fishkin 2009). DP je enak drugim DMJ glede uporabe deliberacije in naključnega izbora udeležencev, ki sta bistvena znaka vseh DMJ, razlikuje pa se po izvedbi dveh mnenjskih poizvedovanj, od katerih je prvo izvedeno pred deliberacijo v fazi iskanja udeležencev DP, drugo pa po deliberaciji, in po tem, da je rezultat DP »izdeliberirano«¹² javno mnenje, ki je seštevek individualnih mnenj udeležencev, pridobljenih z zaupnimi vprašalniki po končani deliberaciji in ni skupno stališče, kot je značilno za vse druge DMJ, ki bi lahko bilo posledica takih ali drugačnih pritiskov v smeri določanja vsebine skupnega stališča.

1. Najpomembnejši del prve faze je izbor udeležencev. DMJ uporabljajo naključni izbor udeležencev. Namen je pridobiti reprezentativno število udeležencev iz relevantne populacije. Vzorec mora biti dovolj velik, da omogoča reprezentativnost in da je mogoče spremembe mnenja evalvirati na statistično smiseln način. K sodelovanju v DP je povabljen večje število reprezentativnih udeležencev, ki so informirani o namenu in poteku DP s prošnjo, da odgovorijo na mnenjsko poizvedbo. (To je prva, preddeliberativna, mnenjska poizvedba v zvezi z zadano temo. Na enako poizvedbo odgovarjajo udeleženci DP po zaključenem procesu deliberiranja. DP je mogoče izvesti s posebno kontrolno skupino ali pa se uporabijo tiste, ki so na prvo poizvedbo odgovorili, pa se sicer niso udeležili DP, kot kontrolno skupino.) Tistim, ki so na prvo poizvedbo odgovorili in potrdili udeležbo na DP, poštno pošljejo informativno gradivo

o temi, ki jo bo DP obravnavala. Gradivo vsebuje različne vsebinske poglede z argumenti za in proti in mora biti nevtrarno in uravnovešeno v tem smislu, da nudi najmočnejše na dejstvih in dokazih temelječe argumente, ki jih je moč najti. Pripravijo ga izvedenci, pregleda pa svetovalni odbor, ki vključuje različne perspektive gledanja na obravnavano temo. Anicker (2024, 112) opozarja na nevarnost, da je uravnovešenost perspektiv gledanja lahko omejena zgolj na razpon gledišč, ki so znotraj »mainstreama«, s tem pa je zarisan družbeni prostor pozicij, ki jih je možno legitimno zastopati, oz. prostor legitimnega nestrinjanja, iz katerega se oblikujejo sprejemljive pozicije govorcev v razpravi. Udeležba izvedencev je lahko problematična, saj gre pri njih za disciplinarno profilirane osebe, ki obravnavajo problem zgolj s svojega gledišča, celota problema pa se izmakne njihovemu pogledu. Kdo odloča o tem, kakšne narave je problem in kateri pristopi so smotrni in relevantni za njegovo reševanje, je pomembno odprto vprašanje.

2. Druga faza je deliberacija v navzočnosti, iz oči v oči, v prireditvenem prostoru. V tej fazi je izvedena deliberativna razprava kot sosledje več korakov, ki se lahko, če je treba, ponavljajo:

- Odprtje razprave: udeleženci izrazijo začetna stališča glede na dano temo, praviloma še brez utemeljitev.
- Informiranje: seznanjanje z relevantnimi strokovnimi, pravnimi, etičnimi itn. informacijami v živo (že prej so udeleženci prejeli izčrpno informacijsko gradivo).
- Divergentna razprava: izražanje različnih perspektiv, pojav nestrinjanja in možnega konflikta.
- Argumentacija: poglobljeno utemeljevanje in soočanje stališč.
- Konvergentna razprava: iskanje skupnih točk, oblikovanje predlogov.
- Refleksija: razmislek o spremembah stališč, utemeljitev doseženega soglasja (ali razhajanja – metakonsenz).
- Zaključek: formalizacija rezultatov, priporočil ali odločitev.

Vse to se dogaja na dveh ravneh. V majhnih, naključno sestavljenih, skupinah in v plenumu. Prava deliberacija poteka v majhnih skupinah (do deset), ki so organizacijsko definirane z diferenciacijo vlog na udeležence in moderatorje. Na plenarnem zasedanju so navzoči vsi udeleženci – postopkovna javnost. Plenum služi predvsem spraševanju izvedencev in drugih informatorjev, pa tudi izmenjavi vmesnih dosežkov in vprašanj med majhnimi skupinami. Ta faza poteka kot izmenjavanje dela v plenumu in dela po majhnih skupinah po potrebi.

3. Drugi fazi sledi izvedba podeliberativne mnenjske poizvedbe z istim vprašalnikom kot pri prvi poizvedbi z namenom ugotavljanja morebitnih sprememb v glediščih udeležencev. Poleg tega je izvedena še pisna evalvacija celotne prireditve, v kateri sodelujejo udeleženci poizvedbe.

4. Ta faza ne vključuje udeležencev prireditve. V njej gre za znanstveno analizo rezultatov obeh poizvedb. Analiza ugotavlja stopnjo deliberativnosti postopka na osnovi primerjave z normativnimi idealnimi pričakovanji, izračuna spremembe mnenj med prvim in drugim poizvedovanjem kot neto spremembo gledišč (*»net attitude change«*¹²) ter interpretira te neto spremembe kot učinke učenja, ki so posledica us-

pele deliberacije, če je do nje prišlo. Samo spremembe gledišč, ki so posledica deliberacije, so relevantne. Če je do spremembe gledišč prišlo zaradi neavtonomnega razmišljanja udeležencev pod vplivom nedeliberativnih dejavnikov, seveda ne govorimo o učenju, pač pa o takem ali drugačnem podrejanju.

Organizirana komunikativna interakcija

Deliberativna poizvedba je udejanjana kot preplet formalne strukture in organizirane interakcije udeležencev skozi komunikativni deliberativni proces (Luhmann 1983; Fishkin in Luskin 2005; Fishkin 2012; Anicker 2021). Razlikovati je treba tri ravni organizacije.

1. *Raven formalne strukture postopka.* Postopkovna komunikacija je umeščena v formalno strukturo, ki jo sestavljajo izrecne norme v obliki postopkovnih pravil in normativna pričakovanja glede njihovega ustreznega prenosa v prakso. Formalna pravila določajo »uradni« namen postopka, postopkovne korake, relevantne deležnike, razčlenjenost vlog, procesno urejenost interakcij in agregacijske načine za oblikovanje kolektivnih mnenj ali odločitev. Formalna struktura določa tip postopka (deliberativna poizvedba, državljanska skupščina, sodna porota, zaslišanje na policiji itn.). Poseben pomen v formalni strukturi imajo vloge in skripti (vednost o pričakovanih obnašanjih v specifičnem kontekstu). Na podlagi vlog so ravnanja udeležencev zasidrana v postopku. Postopkovne vloge so povezovalna spona med formalno strukturo in dejanskim potekanjem postopkovnega procesa. Za delujoče znotraj formalne strukture je ta zanje robni pogoj in vir, na katerega se lahko oprejo v primeru konflikta ali prekinitve deliberativnega procesa iz drugih razlogov. To pa pomeni, da specifikacije vlog in skript ne determinirajo ravnanja nosilcev vlog absolutno, pač pa puščajo dovolj maneverskega prostora za individualno odločanje, ki pa je uokvirjeno s »funkcionalno perspektivo« (Luhmann 1983, 44), ki določa koridor smiselnih ravnanj. Za projektne, kratkotrajne neprofesionalne organizacije, kakršna je DP, je značilno, da v njih prevladujejo formalna pravila. To še posebej velja za DMJ, v kateri sodelujejo ljudje z »vseh vetrov«, ki se ne poznajo, si niso podobni niti socialno, niti poklicno, niti izobrazbeno. Posebnost strukture vlog v tem primeru je, da so nekatere profesionalne (organizator, moderator), druge pa neprofesionalne (udeleženci, priče). Neprofesionalnost udeležencev je seveda načrtovana posledica demokratično participativne naravnosti poizvedbe.

2. *Raven organizirane interakcije.* Deliberativni procesi so interakcijsko intenzivni govorni procesi. Interakcija je definirana kot sonavzočnost in vzajemno zaznavanje govorcev, ki se zavedajo, da zaznavajo druge in da so zaznavani od drugih, v skupnem prostoru s skupnim predmetom pozornosti. Govorne interakcije so osredinjene na temo, ki jih povezuje v smiselno celoto. Razlikujemo lahko dve vrsti interakcij: formalne in neformalne. V formalnih je realizirana formalna struktura postopka po specifičnih vlogah (organizator, moderator), ki skrbijo za to, da je potek postopka formalno korekten. Ob tem pa se s trajanjem postopka in z večanjem ter s ponavljanjem interakcij, umeščenih v postopek, oblikuje neformalna struktura interakcij kot selektivna funkcionalna prilagoditev in konkretizacija formalne strukture potrebam deliberativnega postopka ali pa kot njegova razgradnja (povedano ne velja v tolikšni meri za DP zaradi prekratkega trajanja – 1,6 dneva v povprečju).

Deliberacija poteka v majhnih skupinah, v katerih sodelujejo udeleženci in moderatorji. Struktura deliberativnega procesa je vnaprej določena (1) in ne omogoča samoorganiziranja oz. samoupravljanja udeležencev, kar ni ravno skladno z normativno usmeritvijo DD. Ni pa določena njegova vsebina. V procesu deliberacije pripada moderatorju glavna vloga.¹³ Moderatorji so skrbniki deliberativne kakovosti. Moderiranje je obrt omogočanja kolektivnega dela med razpravo, ki je vključujoča, smiselna in produktivna (Escobar 2019). Moderiranje izvajajo moderatorji, ki niso pasivni opazovalci, pač pa aktivni sooblikovalci demokratičnega diskurza. Poglavitna pravila moderiranja so nevtralnost, pravična/enaka udeležba v razpravi, na dejstvih in dokazih temelječa razprava, spoštljivost in vljudnost, osredinjenost na temo.

Moderator ne sme izražati osebnega mnenja, ne sme zagovarjati političnih stališč in izkazovati pristranskosti. Moderator naj bi bil nevtralen in nepristranski, vezan samo na splošne in nadstrankarske vrednote. Biti mora strogo nepristranski do vsebine deliberacije, ne pa nevtralen do samega procesa, saj je njegova naloga, da ga ohranja in spodbuja na poti do cilja. Poskrbeti mora, da udeleženci razpravljalno okolje dojemajo kot varen prostor, v katerem so vsi pogledi dobrodošli, in identificirati ter omiliti implicitne hierarhije.¹⁴ Omogočati mora pravično udeležbo tako, da uveljavlja časovno omejitvev na govornca, da bi preprečil prevlado nezaustavljivih razpravljavcev, in proaktivno spodbujati zadržane udeležence k besedi. Moderator je pri tem v nerodnem položaju v skupini, ker mora hkrati voditi razpravo in ji tudi slediti, ko skupina razvija diskurz. Ne obstaja »naravni« ali »prosti« tok interakcij med deliberatorji, ki vedno potekajo v bolj ali manj artificialnem okolju. Naloga moderatorja je, da vodi proces ustvarjanja vsebine, ki poteka kot govorna interakcija, formalno korektno in še posebej tako, da prideta do izraza smoter in produktivnost kolektivnega sodelovanja. Pri tem uporablja diskurzivna sredstva, kot so procesni govor (postopki, pravila), problemski govor (opcije, ki so na voljo) in govor uokvirjanja (perspektive gledanja na problem) Dillard (2013, 222). Vloga moderatorja je z vidika deliberativne demokracije lahko normativno sporna, ker krši načelo enakosti udeležencev, je pa, po drugi strani, pragmatično nujna, če naj komunikacija v skupini poteka skladno s pravili postopka, kar pa je ravno naloga moderatorja (Landwehr 2014, 78). Če bi namreč bila skupina prepuščena sami sebi, bi se kaj lahko zgodilo, da bi se te vloge polastil dominantni udeleženec ali pa bi razprava izgubila svoj fokus. Primerjalna raziskava treh tipov moderiranja (samoorganizacija, multimetoda in dinamično moderiranje¹⁵) je z uporabo treh kriterijev – vključenosti, interakcije in učinka – pokazala, da je samoorganizacijski tip moderiranja najmanj uspešen pri realizaciji teh kriterijev. Pokazala je tudi, da razlike v moderiranju vplivajo na proces deliberacije na različne načine (von Schneidemesser et al. 2023). Vodenje razprave pomeni poseganje vanjo, posegi pa uokvirjajo njen potek na faktični in simbolni ravni. Poglavitno orodje in sredstvo moči moderatorja je njegova monopolna pravica dajanja besede samemu sebi in drugim. To pravico mora razumno uporabljati, da ne prekinja samotoka razprave (če se ta dogaja), kar se kaj hitro lahko zgodi, če s prepogostim poseganjem, z dajanjem in odvzemanjem besede, vzpostavi sebe v zaznavi udeležencev kot nujno navezno točko, prek katere mora potekati komunikacija. Nastane situacija, ko udeleženci ne nago-varjajo neposredno drug drugega, pač pa posredno prek moderatorja (Anicker 2021, 321), posledica česar pa je lahko izguba razpravnega ritma, ker prihaja do prekinitve

toka razprave, ko govoru ne sledi takojšnji odgovor, ker samodejna izmenjava vlog govorcev (*»turn-taking«*, *»Rede Zugzwang«*) ne deluje ali pa ga sploh ni.

Zastavlja se vprašanje, kakšno poseganje je smotrno in dopustno. Moderatorjeva naloga je, da skrbi za približno enako govorno zastopanost vseh udeležencev razprave, če opazi, da so nekateri udeleženci manj prisotni v izmenjavi mnenj, je njegova dolžnost, da jih opogumlja in spodbuja k besedi ter tako zagotavlja notranjo vključenost udeležencev. Drugo vprašanje pa je, ali lahko posega v vsebino razprave in kdaj. Prav tako, ali lahko določa, kateri argumenti so dopustni in kateri ne, ali lahko dopušča samo argumente, ki so posplošljivi (znak njihove racionalnosti), drugih pa ne, katera stopnja emocionalizacije razprave je še ustrezna, pri čemer ne gre spregledati informativne vrednosti emocij, še posebej zelo intenzivnih. Ta vprašanja bi morala biti *»off-limits«*. Ni pa problematično, če moderator poseže v razpravljanje nekoga, ki se ne drži teme razpravljanja, se ponavlja ali pa uporablja *»podatke«*, ki so na daleč prepoznavni kot napačni ali zavajajoči.

3. *Raven govornih dejanj*. Govorna dejanja so gradniki deliberativne interakcije. Z normativnega vidika nastane deliberacija kot nizanje tematsko usmerjenih govornih dejanj, ki se zgostijo v oblike komunikativnega usklajevanja delovanja (pogajanje, argumentiranje, kritiziranje, sporazumevanje itn.) v iskanju odgovora na vprašanje, kaj storiti. Na tej ravni se pokaže, koliko deliberativne so komunikativne interakcije v forumu. Trije postopki se uporabljajo za merjenje stopnje deliberativnosti (kakovosti) deliberativnega procesa (Ercan et al. 2022).

Ugotavljanje in merjenje deliberativnosti

Indeks diskurzivne kakovosti (»Discourse Quality Index« – DQI)

Ta indeks je mersko sredstvo za evalviranje kakovosti razprave v deliberativnem ali komunikativnem okolju. Glavni kazalniki (kriteriji) indeksa, s katerimi se merijo in ugotavljajo govorni prispevki v razpravi, so spoštovanje in vljudnost (udeleženci so spoštljivi med seboj, ni osebnih napadov ali žalitev, diskurz spodbuja vključevalno vzdušje), relevantnost (prispevki se neposredno nanašajo na temo, nerelevantni komentarji so minimizirani), jasnost in razumljivost (izjavljanje je jasno izraženo, udeleženci artikulirajo svoja stališča na razumljiv način, žargonu in dvournemu jeziku se je treba izogibati), upravičevanje in dokazi (argumenti so podprti z razlogi, dokazi ali primeri, udeleženci upravičujejo svoja stališča in ne zgolj zatrjujejo), odprtost do drugih perspektiv (udeleženci priznavajo in obravnavajo druge perspektive in so pripravljeni spremeniti svoje mnenje na osnovi razlogov in dokazov), globina in plodnost (raziskovanje domnev, predpostavk in implikacij, ali je razprava zgolj površinska), konstruktivnost in usmerjenost k reševanju problema (prispevki ciljajo na razrešitev konfliktov, iskanje stičnih točk, razvijanje rešitev, poudarek je na sodelovalnem posvetovanju, in ne na nasprotni debati), vključevalnost (enake priložnosti za vse glasove, obstranske ali manjšinske perspektive lahko pridejo do besede).

Indeks kakovosti poslušanja (»Listening Quality Index« – LQI)

LQI je orodje za evalvacijo in povečanje učinkovitosti poslušalnega vedenja udeležencev razprave. Dobro poslušanje je bistveno za produktivno deliberacijo, ker pospešuje razumevanje, spoštovanje in zmožnost grajenja konsenza. Glavni kazalniki

(kriteriji) indeksa so pozornost (angažirano poslušanje udeležencev brez motenj), spoštljivost (ali poslušajo udeleženci brez prekinitev ali omalovaževanja drugih), razumevanje (ali pravilno dojemajo govorčeve poante in z njimi povezane zadeve), empatičnost (ali priznavajo emocionalne vsebine in izkazujejo razumevanje), razjasnjevanje in povratne zveze (ali postavljajo vprašanja in parafrazirajo, da bi potrdili svoje razumevanje), kakovost odzivanja (ali so njihovi odzivi relevantni, spoštljivi in konstruktivni). Indeks torej meri, v kolikšni meri je poslušanje udeležencev uspešno z deliberativnega vidika. Poleg tega spodbuja kulturo aktivnega, spoštljivega poslušanja, ki je bistvena za produktivnost in legitimnost deliberativnih procesov. S sistematičnim merjenjem in z ugotavljanjem poslušanja zagotavlja LQI, da se deliberacija ne nanaša samo na govorjenje, pač pa tudi na slišanje in razumevanje drugih, kar je ključno za demokratično legitimnost in družbeno povezanost.

Indeks deliberativnega uma (»Deliberative Reason Index« – DRI)

DRI je konceptualno in mersko orodje za ugotavljanje kakovosti in obsega racionalnega mišljenja v deliberativnih procesih. Meri, v kolikšni meri udeleženci deliberacije svoje argumente, odločitve in presoje opirajo na temeljito premišljevanje, dokaze in racionalni diskurz. Indeks kvantificira stopnjo, do katere deliberatorji uporabljajo premišljeno, logično in na dejstvih ter dokazih temelječe mišljenje, ko razpravljajo, in evalvira kakovost premišljenosti uporabljenih argumentov, predlogov in odločitev, proizvedenih v deliberativnem procesu. Glavni kazalniki (kriteriji) indeksa so logična koherentnost (ali so argumenti konzistentni in ne vsebujejo logičnih napak, ali sledijo koherentni strukturi sklepanja), uporaba dokazov (ali so trditve podprte z relevantnimi podatki, dejstvi in verodostojnimi viri), relevantnost (ali se argumenti neposredno nanašajo na temo ali vprašanje, o katerem teče beseda), priznavanje protiargumentov (ali udeleženci priznavajo in odgovarjajo na nasprotna pogleda premišljeno), upravičevanje in razlaganje (ali so razlogi za mnenje jasno artikulirani in upravičeni), reflektirano premišljevanje (ali udeleženci obravnavajo implikacije, pretehtane izbire in skrite predpostavke). DRI je v pomoč pri razlikovanju med površinsko izmenjavo mnenj in pristno premišljeno deliberacijo. Spodbuja udeležence, da oblikujejo dobro utemeljene argumente, s čimer prispeva k izboljševanju kakovosti in legitimnosti kolektivnih odločitev.

Naključni izbor

Naključni izbor (NI) je druga inovativna značilnost vseh DMJ, tako tudi DP. Tu gre za, lahko bi rekli, revolucionarno spremembo v načinu sestavljanja teles, ki pretendirajo na politično legitimnost. DMJ namreč ne pridobivajo udeležencev z volitvami po načelu ena oseba en glas, kot to velja v liberalno-reprezentativnem sistemu, ampak z naključnim izborom. Deliberativna demokracija pridobiva sprejemljivost iz deliberativne legitimnosti, ta pa iz dejstva identitete naključno izbranih kot tudi iz deliberativne kakovosti dejanskih deliberativnih procesov. Zato je treba pri koncipiranju naključnega izbora upoštevati dve vrednoti. Vrednoto univerzalne moralne enakosti (politično enakost), ki prav zaradi univerzalnosti, abstrahira od konkretnih družbenih okoliščin, in vrednoto pravičnosti, ki pa prav nasprotno zahteva ravno upoštevanje teh okoliščin. Prvo vrednoto je mogoče realizirati s čistim naključnim izborom, drugo

pa z uporabo kvot (ki niso reprezentativne). Srednja pot je stratificirano naključno vzorčenje, ki je uporabljano pri DP.

Naključni izbor kot statistična metoda sicer ni nekaj novega, nova pa je njegova uporaba v smislu generiranja reprezentativnosti DP z identifikacijo, in ne z volilno avtorizacijo. Pridobivanje udeležencev za DP je ključno za njeno delovanje in legitimnost ter izhaja iz načela vključevalnosti, ki pa samo temelji na ideji prizadetosti. V načelu bi moral imeti v demokraciji vsakdo prizadet z neko odločitvijo možnost vplivati nanjo in tako videti samega sebe kot soavtorja take – splošno obvezujoče odločitve. V DP poteka deliberiranje v majhnih skupinah tako, da kriterij prizadetosti ne more biti upoštevan, ker je nemogoče povabiti vse prizadete k sodelovanju. Zato mora biti deliberativni vzorec reprezentativen. Reprezentativnost tu ni mišljena v volilnem smislu, pri čemer volivci z izglasovanjem pooblastijo (avtorizirajo) predstavnike, da odločajo v njihovem imenu in si tako zagotovijo možnost, da jih na naslednjih volitvah nagradijo s ponovno izvolitvijo ali pa kaznujejo z neizvolitvijo (retrospektivno, ex post glasovanje na osnovi položenih računov in ugotovljene odgovornosti ali zaslužnosti). Pri naključnem izboru udeležencev ni možno polaganje računov, ker je naključni izbor zgolj enkratni (brez možnosti ponavljanja). Zveza med naključno izbranimi udeleženci DP in relevantno populacijo je v njihovi demografski reprezentativnosti, ko so udeleženci DP zaznavani kot miniaturizirana zrcalna slika te populacije (Fishkin 2018, 73), kot populacija v malem, ki jo Fishkin imenuje mikrokozmos.

Dobra stran naključnega izbora je demografska raznovrstnost izbrancev, njihova deskriptivna reprezentativnost in njegova lahka razumljivost. Večina pozna idejo naključnosti iz spremljanja loterije. Pri naključnem izboru ima vsak posameznik enako možnost in verjetnost, da bo vključen v vzorec. Pri DMJ pogosto nastopi težava v zvezi s statistično verjetnostjo, ker je velikost vzorca premajhna. Ta bi morala biti vsaj 1000 udeležencev, ne glede na velikost relevantne populacije. Nobena DMJ ne dosega te vrednosti, še najbližje ji pride DP. Državlјanska porota šteje 12–26 udeležencev, konsenzna konferenca 10–50, državljanska skupščina 50–160, DP pa 100–700 udeležencev. Največji DMJ do zdaj sta bili G-1000 Citizens' Summit s 704 udeleženci leta 2011 v Belgiji in DP v Mongoliji s 669 udeleženci leta 2017.

Pri izvedbi DP je uporabljan dvofazni izbirni proces. V prvi fazi je treba odgovoriti na naslednja vprašanja:

1. Populacija, ki bo reprezentirana z naključnim izborom.
2. Število izbrancev (N).
3. Stratifikacijski kriteriji: demografski, glediščni, če je treba, tudi vedenjski.
4. Način vabljenja določenega števila izbrancev, da participirajo.

Potem pa se, v drugi fazi, na osnovi načrtovanega vzorca pošlje veliko število povabil naključno izbranim posameznikom k sodelovanju v DP z opisom postopka in teme ter s poizvedbenim vprašalnikom, npr. 100-kratnik tega vzorca, ker izkustvo kaže, da je odzivnost na vabila zelo majhna (okoli 1–2 %). Nato se iz števila tistih, ki so se pozitivno odzvali na povabilo s pristankom na udeležbo, spet naključno izberejo udeleženci, ki se jim pošlje informativno gradivo o obravnavani temi. V drugi fazi pa sledi stratificiranje vseh prostovoljcev za udeležbo po demografskih, glediščnih ali vedenjskih kriterijih.

Fishkin (2012) meni, da bi za realizacijo pravega in reprezentativnega mikrokozmosa morali biti realizirani naslednji kriteriji: demografska reprezentativnost, reprezentativnost gledišč in velikost vzorca. Demografska reprezentiranost se izvaja po standardnih kategorijah, kot so razred, starost, spol, izobrazba, dohodek, narodnost. Demografska reprezentativnost pomeni, da so za ljudi z različnimi družbenimi ozadji verjetni različni pogledi in interesi. Za nekatere raziskave te kategorije povsem za- doščajo, za DP in druge DMJ pa ne. Za DP je treba vključiti tudi reprezentativnost gledišč v družbi. Ta je pomembna, ker je gledišče samostojna kategorija, ki ni strogo določena s položajem v socialni strukturi ali celo zvedljiva nanj. Dryzek in Niemeyer (2008) pišeta podobno, a ne istosmiselno, o reprezentiranju diskurzov,¹⁶ ki so prisotni v družbi. Tretji kriterij pa je velikost vzorca. Mikrokozmos mora biti dovolj velik, da je njegovo reprezentativnost in statistično pomembnost sprememb mnenj možno smiselno evalvirati (Fishkin 2018, 74).

Zaradi majhnega N je mogoče preslikati raznovrstnost relevantne populacije le s stratificiranim naključnim vzorčenjem. Uporabljajo se stratifikacijski kriteriji, kot so demografske značilnosti (spol, starost, kraj bivanja), socialnoekonomske značilnosti (poklic, izobrazba) ter tudi politična gledišča in vedenje (npr. volilno).

Politični učinki naključnega izbora so naslednji:

1. Stratificirano naključno vzorčenje zagotavlja vključenost in raznovrstnost, še posebej vključenost na rob potisnjenih družbenih skupin in večjo reprezentativnost DP od običajne javnomnenjske poizvedbe, sploh pa od drugih političnih forumov, v katerih prevladujejo dobro organizirane družbene skupine.

2. Naključni izbor udeležencev med »navadnimi«¹⁷ državljani preprečuje njihovo kooptacijo in ugrabitev od politikov ali interesnih skupin ter zagotavlja politično neodvisnost in nepristranskost udeležencev, kar povečuje možnosti, da bodo delovali v smeri javnega interesa (Lafont in Urbinati 2024, 88).

Institucionalizacija deliberativne poizvedbe

DP je bila uporabljena že na vseh kontinentih. Doživela je že več kot 150 izvedb. Njena uporaba je bila institucionalizirana v Mongoliji z zakonom o deliberativnem poizvedovanju,¹⁷ ki ga je treba uporabiti pri vsaki spremembi ustave.

Deliberativno poizvedovanje je bilo že drugič uspešno uporabljeno pri spreminjanju mongolske ustave. Izpeljava deliberativne poizvedbe je od leta 2017 zakonsko obvezna, preden lahko veliki državni hural (parlament) obravnava predlagano spremembo ustave. Prvo spremembo, sprejeto na podlagi tega postopka, je parlament sprejel leta 2017 po nacionalnem deliberativnem poizvedovanju istega leta. Leta 2023 se je stratificirani naključni vzorec skoraj 800 državljanov iz vse države zbral na podaljšanem vikendu, da bi razpravljali o možnih predlaganih spremembah ustave. Vzorec je izbral nacionalni statistični urad (ki izvaja tudi popis prebivalstva). Udeleženci so se zbrali v vladni palači (sedež vlade je v Ulan Batorju). Teme so bile izbrane na podlagi približno 600 predlogov strokovnjakov in približno 500 predlogov javnosti, civilne družbe in strokovnih organizacij iz vse države. Nato jih je pregledal deliberativni svet, neodvisno nestrankarsko telo, ki ga zahteva zakon o deliberativnem poizvedovanju. Posvetovanje je potekalo 14. in 15. februarja 2023.

Dva od predlogov, ki sta ob koncu posvetovanj prejela zelo visoko podporo, sta predstavljala jasno podlago za spremembo ustave. Prvi je bil razširitev velikosti parlamenta (da bi se povečala njegova zmožnost predstavljanja javnosti), drugi pa sprejetje vzporednega volilnega sistema, ki bi združeval enomandatne volilne okraje in sorazmerno zastopanost. Razširitev parlamenta je podprlo 82 odstotkov udeležencev. Sprejetje vzporednega volilnega sistema za parlament se je končalo s podporo 71 odstotkov udeležencev. Sprememba je uspešno povečala zastopanost v parlamentu s 76 poslancev na skupno 126, pri čemer je 78 poslancev voljenih z večinskimi glasovanjem v okrajih, 48 pa s sorazmernim predstavljanjem. Poleg tega je bila kvota za ženske kandidatke na parlamentarnih volitvah leta 2024 povišana na 30 odstotkov, do leta 2028 pa naj bi se nadalje povečala na 40 odstotkov.

ZAKLJUČEK

Deliberativna poizvedba, ki jo je razvil politolog James S. Fishkin, je ena izmed najpomembnejših in najbolj sistematično razvitih metod za merjenje in spodbujanje premišljenega javnega mnenja. V okviru deliberativne demokracije omogoča ta pristop ugotovitev, kakšno bi bilo javno mnenje, če bi imeli državljani priložnost pridobiti zanesljive informacije, razpravljati o političnih temah v raznolikih skupinah in reflektirati svoja stališča. Fishkinova metodologija združuje naključno vzorčenje z deliberativno razpravo ter omogoča kvantitativno in kvalitativno analizo sprememb stališč.

Z metodološkega vidika se Fishkinova poizvedba naslanja na klasične principe eksperimentalnega načrtovanja: obstajajo predpreizkus (merjenje stališč pred deliberacijo), deliberativni tretma (natančno strukturirana razprava, ki temelji na uravnoteženi informativni dokumentaciji) in naknadni preizkus (merjenje stališč po deliberaciji). Udeleženci so izbrani z naključnim vzorčenjem, kar povečuje reprezentativnost in omogoča posplošitev rezultatov na širšo populacijo. S tem zasleduje Fishkin ideal javnega mnenja, ki je informirano, reflektirano in ustvarjeno v razmerah medsebojnega spoštovanja in enakosti.

Ena izmed osrednjih prednosti deliberativne poizvedbe je, da gre za empirični eksperiment deliberativne demokracije. V nasprotju s tradicionalnimi poizvedbami, ki zajemajo trenutne preference, Fishkinova metoda meri, kako bi se te spremenile, če bi državljani imeli priložnost sodelovati v strukturiranem in moderiranem diskurzu. V tem smislu je deliberativna poizvedba orodje za normativno in empirično preverjanje deliberativne teorije. Ugotovitve iz različnih projektov (npr. »America in One Room«, »Europolis«, tajvanski in južnokorejski projekti) dosledno kažejo na občutne in statistično značilne spremembe stališč, povečano toleranco do drugačnih mnenj, povečano zaupanje v politične institucije in večje razumevanje kompleksnih političnih problemov.

Za deliberativno poizvedbo je značilen tudi pomemben institucionalni potencial. V nekaterih primerih (npr. Mongolije, Južne Koreje) je bila uporabljena kot del formalnega političnega odločanja, kar kaže na njeno zmožnost podpore odločanju v demokratičnih okvirih. Fishkin sam opozarja, da poizvedba ne nadomešča obstoječih institucij predstavniške demokracije, temveč jih dopolnjuje z deliberativno komponento, ki krepi legitimnost, vključujočnost in informiranost odločitev.

Prispevek Fishkinove deliberativne poizvedbe k politični teoriji in praksi je pomemben. S teoretskega vidika utrjuje idejo, da demokratična legitimnost ni mogoča brez javnega utemeljevanja in skupnega razumevanja. Empirično pa prinaša podatke, ki dokazujejo, da so državljani sposobni in pripravljeni sodelovati v razpravi, če so za to ustrezno pripravljeni in motivirani.

Fishkin je s svojo metodologijo uspel preseči tradicionalni prepad med normativno teorijo, prakso in empirično raziskavo: deliberativna poizvedba je hkrati politična tehnologija, eksperiment in instrument demokratične inovacije. V dobi dezinformacij, polarizacije in cinizma glede politike prinaša empirični dokaz, da je kakovostna deliberacija mogoča in da lahko bistveno vpliva na oblikovanje zrele in odgovorne javne sfere.¹⁸

Fishkinova deliberativna poizvedba je ena najbolj dognanih metod deliberativne demokracije. Povezuje načela enakosti, razuma, vključevnosti in reflektiranosti v ponovljiv in merljiv okvir. Njena akademska vrednost je v prispevku k teoriji demokracije, njena praktična vrednost pa v čedalje pogostejši uporabi pri oblikovanju politik in demokratičnem eksperimentiranju po svetu.

OPOMBE

- 1 Sestavek je napisan z namenom predstavitve DP, je informativen in ni zastavljen problemsko ali celo kritično.
- 2 Morda bi bilo primerneje govoriti o teorijah DD, ker teorija DD v kanonični obliki ne obstaja, gre za družino heterogenih teorij, ki jim je skupna deliberacija kot temeljna kategorija.
- 3 Oblikovanje *skupnih* javnih stališč in priporočil je namen vseh DMJ, razen deliberativne poizvedbe. Pri DP ostanejo mnenja individualna, ker deliberativni proces v tem primeru ni namenjen izdelavi skupnih stališč, pač pa samo deliberativni razpravi, ki se ne zaključi s skupnim stališčem, ampak zgolj z registriranjem individualnih mnenj s podeliberativno poizvedbo in z njihovim seštetjem.
- 4 www.yukoncitizensassembly.ca.
- 5 Bürgerräte in Deutschland, 2024; dostopno na www.mehr-demokratie.de.
- 6 Iskanje soglasja, ne pa nujno njegovo doseganje, je regulativni, aspiracijski ideal, ki je opredelilveni znak deliberacije. Iskanja soglasja ne smemo razumeti v smislu odkrivanja nečesa že vnaprej obstoječega, ampak kot kolektivni proces njegovega konstruiranja in je logično in empirično nujno kot predpostavka demokratičnega sodelovanja.
- 7 Čeprav je »posvetovanje« sopomenka deliberacije, še naprej rabimo »deliberacija«, ker je njen pomen dovolj natančno določen, še posebej, ker je v slovenščini tujka, ki ni v rabi, in ni »kontaminirana« z vsakdanjo rabo.
- 8 V tem sestavku je obravnavana samo deliberacija navzočih v forumu kot razprava iz oči v oči, online deliberacija pa ne. Online deliberacija ima svoje posebnosti, ki so v nezmožnosti zaznavanja govornice telesa in intonacije glasu, ki sta pomembni pri presojanju verodostojnosti in iskrenosti govornicevega izjavljanja. Odsotna je tudi sočasnost komuniciranja, ki jo nadomesti zaporednost, časovni zamik in prostorska distanca. Informativna vrednost tovrstne deliberacije je zato manjša. Njena prednost pa je v številu možnih udeležencev (»*crowdsourcing*«), kvantitativnem spremljanju postopka in stroških postopka. Ta se lahko še povečuje z napredovanjem tehnologije.

- 9 Splošna sprejemljivost je sproti vzpostavljena v procesu deliberiranja.
- 10 »Deliberativna poizvedba[®] je drugačna od vseh doslej izvedenih poizvedb ali raziskav. Navadne poizvedbe prikazujejo, kaj javnost misli, čeprav javnost morda ne razmišlja ravno preveč ali ne posveča pozornosti. Deliberativna poizvedba[®] poskuša prikazati, kaj bi javnost menila, če bi imela boljšo priložnost, da temeljito premisli o zadevnem vprašanju« (Fishkin 1995, 162).
- 11 Letos mineva stoto leto od izida knjige *The Phantom Public* Walterja Lippmanna.
- 12 Izraz »*attitude*« je preveden kot gledišče. Druge možnosti so stališče, drža(nje), nagljenost, odnos do.
- 13 Smith (2009, 198) ugotavlja, da vpliv različnih načinov moderiranja ni obravnavan v zadevni literaturi, kar se mu zdi čudno.
- 14 Implicitne hierarhije so neformalni, neizrečeni sistemi moči, avtoritete ali statusa, ki oblikujejo interakcije, odločitve in vedenja znotraj skupin, organizacij ali družb. Uko-reninjene so v kulturnih normah, tradiciji, predsodkih ali medosebnih dinamikah.
- 15 Postopek se imenuje »*dynamic facilitation*«. V sestavku je izraz »*facilitation*« preva-
jan z izrazom moderacija.
- 16 Razlog za vključitev diskurza in (ali) gledišča med stratifikacijske kriterije je hotenje
po upoštevanju političnih razlik.
- 17 Mongolski zakon o deliberativnem poizvedovanju, sprejet leta 2017, določa pravni
okvir za institucionalizirano vključevanje državljanov v odločanje o ključnih javnih
zadevah. Osrednja določba zakona je, da mora biti deliberativna poizvedba izve-
dena kot pravni pogoj za spreminjanje ustave. Poleg tega se zakon lahko uporablja
tudi za druga pomembna vprašanja, kot so lokalni razvoj, reforme javnih politik in
prostorsko načrtovanje.
Postopek je sestavljen iz treh faz. Prva je začetna poizvedba, v kateri reprezentativni
vzorec prebivalcev odgovarja na vprašanja brez predhodnega informiranja. V drugi
fazi poteka strukturirana deliberacija v majhnih skupinah pod vodstvom usposobl-
jenih moderatorjev in na podlagi uravnoteženih informacij. Tretja faza je ponovna
poizvedba, med katero isti udeleženci odgovorijo na ista vprašanja kot v prvi fazi,
kar omogoča analizo (ne)sprememb stališč.
Za vsako deliberativno poizvedbo se oblikuje svetovalni odbor, ki potrdi vprašanja,
nadzira izvedbo in oblikuje priporočila na podlagi rezultatov poizvedbe. Priporočila
so predložena pristojnim organom kot osnova za nadaljnje odločanje. Zakon je prvi
primer na svetu, po katerem je deliberativna procedura predpisana kot pravni pred-
pogoj za ustavne spremembe.
Celotno besedilo zakona v angleščini, »*Law on Deliberative Polling*«, objavlja kot
dodatek Fishkin (2018, 243–249).
- 18 O razlikovanju med »javnostjo« in »javno sfero« glejte Splichal (2024, 57–89).

CITIRANA LITERATURA

- Anicker, Fabian. 2021. »Organisierte Vernunft«. V *Organisierte Moral*, ur. André Armbruster in Cristina Besio. Wiesbaden: Springer.
- Anicker, Fabian. 2024. »Demos ex Machina«. V *Die Fabrikation der Demokratie*, ur. Jan-Peter Voss in Hagen Schölzel. Wiesbaden: Springer.
- Bächtiger, André, et al. 2018. *The Oxford Handbook of Deliberative Democracy*. Oxford: Oxford University Press.

- Bächtiger, André, in John Parkinson. 2019. *Mapping and Measuring Deliberation*. Oxford: Oxford University Press.
- Beauvais, Edana, in André Bächtiger. 2016. »Taking the Goals of Deliberation Seriously: A Differentiated View on Equality and Equity in Deliberative Designs and Processes«. *Journal of Public Deliberation* 12 (2).
- Bessette, Joseph. 1980. »Deliberative Democracy: The Majority Principle in Republican Government«. V *How democratic is the constitution?*, ur. Robert Goldwin in William Schambra. Washington D. C.: American Enterprise Institute.
- Boniolo, Giovanni. 2012. *The Art of Deliberating: Democracy, Deliberation and the Life Sciences between History and Theory*. Wiesbaden: Springer.
- Bua, Adrian, Sonia Bussu in Jonathan Davies. 2023. »Democracy-Driven Governance and Governance-Driven Democratisation in Barcelona and Nantes«. V *Reclaiming Participatory Governance*, ur. Adrian Bua in Sonia Bussu. London in New York: Routledge.
- Chambers, Simone. 2001. »Constitutional Referendums and Democratic Deliberation«. V *Referendum Democracy*, ur. Matthew Mendelsohn in Andrew Parkin. London: Palgrave Macmillan.
- Chappell, Suzsanna. 2012. *Deliberative Democracy*. Hampshire: Palgrave Macmillan.
- Cohen, Joshua. 1989. »Deliberation and Democratic Legitimacy«. V *The Good Polity*, ur. Alan Hamlin in Philip Pettit. London: Blackwell.
- Dahl, Robert. 1989. *Democracy and Its Critics*. New Haven in London: Yale University Press.
- Dillard, Kara. 2013. »Envisioning the Role of Facilitation in Public Deliberation«. *Journal of Applied Communication Research* 41 (3).
- Dryzek, John, in Simon Niemeyer. 2008. »Discursive Representation«. *The American Political Science Review* 102 (4): 481–493.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Ercan, Selen, Hans Asenbaum, Nicole Curato in Ricardo F. Mendonça (ur.). 2022. *Research Methods in Deliberative Democracy*. Oxford: Oxford University Press.
- Escobar, Oliver. 2019. »Facilitators: The Micropolitics of Public Participation and Deliberation«. V *Handbook of Democratic Innovation and Governance*, ur. Stephen Elstub in Oliver Escobar. London: Edward Elgar Publishing.
- Fishkin, James. 1991. *Democracy and Deliberation*. New Haven in London: Yale University Press.
- Fishkin, James. 1995. *The Voice of the People: Public Opinion and Democracy*. New Haven: Yale University Press.
- Fishkin, James. 2009. *When the People Speak*. Oxford: Oxford University Press.
- Fishkin, James. 2012. »Deliberative Polling: Reflections on an Ideal Made Practical«. V *Evaluating Democratic Innovations*, ur. Brigitte Geissel in Kenneth Newton. New York: Routledge.
- Fishkin, James. 2018. *Democracy When the People Are Thinking*. Oxford: Oxford University Press.
- Fishkin, James, in Robert Luskin. 2005. »Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion«. *Acta Politica* 40: 284–298.
- Fournier, Patrick, et al. 2011. *When Citizens Decide: Lessons from Citizen Assemblies on Electoral Reform*. Oxford: Oxford University Press.
- Goodin, Robert. 1986. »Laundering Preferences«. V *Foundations of Social Choice Theory*, ur. Jon Elster in Aanund Hylland. Cambridge: Cambridge University Press.
- Goodin, Robert. 2003. *Reflective Democracy*. Oxford: Oxford University Press.
- Gutmann, Amy, in Dennis Thompson. 2004. *Why Deliberative Democracy*. Princeton: Princeton University Press.
- Habermas, Jürgen. 1992. *Faktizität und Geltung*. Frankfurt: Suhrkamp.

- Head, Brian. 2022. *Wicked Problems in Public Policy*. Cham: Palgrave Macmillan.
- Kies, Raphaël. 2012. *Promises and Limits of Web-deliberation*. Hampshire: Palgrave Macmillan.
- Kuyper, Jonathan. 2017. »The Means and Ends of Deliberative Democracy: Rejoinder to Gunn«. *Critical Review* 29 (3): 328–350.
- Lafont, Cristina, in Nadia Urbinati. 2024. *The Lottocratic Mentality*. Oxford: Oxford University Press.
- Landwehr, Claudia. 2014. »The Role of Impartial Intermediaries in Deliberative Mini-Publics«. V *Deliberative Mini-Publics*, ur. Kimmo Grönlund et al. Colchester, VB: ECPR Press.
- Luhmann, Niklas. 1983. *Legitimation durch Verfahren*. Frankfurt: Suhrkamp.
- Manin, Bernard. 1987. »On Legitimacy and Political Deliberation«. *Political Theory* 15: 388.
- Mansbridge, Jane. 2010. »Deliberative Polling as the Gold Standard«. *The Good Society* 19 (1).
- Neblo, Michael, et al. 2018. *Politics with the People: Building a Directly Representative Democracy*. Cambridge: Cambridge University Press.
- Niemeyer, Simon, in John Dryzek. 2007. »The Ends of Deliberation: Meta-consensus and Inter-subjective Rationality as Ideal Outcomes«. *Swiss Political Science Review* 13 (4): 497–526.
- OECD. 2020. *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave*. Pariz: OECD Publishing.
- Offe, Claus. 2017. »Referendum vs. Institutionalized Deliberation: What Democratic Theorists Can Learn from the 2016 Brexit Decision«. *Daedalus* 146 (3): 14–27.
- O'Flynn, Ian. 2022. *Deliberative Democracy*. Cambridge: Polity Press.
- Ryan, Matthew, in Graham Smith. 2014. »Defining Mini-Publics«. V *Deliberative Mini-Publics*, ur. Kimmo Grönlund et al. Colchester, VB: ECPR Press.
- von Schneidmesser, Dirk, et al. 2023. »Diversity in Facilitation: Mapping Differences in Deliberative Designs«. *Journal of Deliberative Democracy* 19 (1).
- Selk, Veith. 2023. *Demokratiedämmerung*. Berlin: Suhrkamp.
- Smith, Graham. 2009. *Democratic Innovations*. Cambridge: Cambridge University Press.
- Smith, Graham. 2024. *We Need To Talk About Climate: How Citizen's Assemblies Can Help Us Solve The Climate Crisis*. London: University of Westminster Press.
- Splichal, Slavko. 2024. *Upodatkovanje javnosti in mnenjske tehnologije*. Ljubljana: Založba FDVin SAZU.
- Steiner, Jörg. 2012. *The Foundations of Deliberative Democracy*. Cambridge: Cambridge University Press.
- Voss, Jan-Peter. 2024. »Innovation der Demokratie«. V *Die Fabrikation der Demokratie*, ur. Jan-Peter Voss in Hagen Schölzel. Wiesbaden: Springer.
- Vrydagh, Julien, et al. 2021. »Thriving in an Unfriendly Territory: The Peculiar Rise of Minipublics in Consociational Belgium«. V *Belgian Exceptionalism: Belgian Politics between Realism and Surrealism*. London in New York: Routledge.
- Warren, Mark, in Hilary Pearce. 2008. *Designing Deliberative Democracy: The British Columbia Citizens' Assembly*. Cambridge: Cambridge University Press.

SOCIALNA TEHNOLOGIJA IN UNIVERZALNI TEMELJNI DOHODEK

Srečo Dragoš

UVOD

Glavni namen prispevka je pokazati, zakaj je univerzalni temeljni dohodek (UTD) realen, pravičen in pomemben primer socialne tehnologije. Prvi razdelek tega prispevka je namenjen splošni utemeljitvi socialne tehnologije in predstavitvi nekatereh dilem o njeni vrednotni ter socialni dimenziji. V drugem razdelku je predstavljena ideja o uvedbi pravice do UTD in pričakovanih pozitivnih učinkih za posameznike in celotno skupnost. Sledi tretji razdelek, ki je namenjen informiranju o dejanskih učinkih te socialne tehnologije, kot so jih izmerili v najnovejšem in največjem eksperimentu z uvajanjem UTD. V zadnjem razdelku (pred sklepom) sta na kratko predstavljeni možni varianti izvedbe UTD, da je jasnejša finančna izvedljivost te pravice, saj se ji najpogosteje očita, da finančno ni vzdržna. Iz prikaza bo jasno, da to ni res. Finančna konstrukcija UTD je povsem preprosto izvedljiva, in to brez uvajanja novih davkov in brez povišanja obstoječih.

SOCIALNA TEHNOLOGIJA

Naslov »inženirka leta 2024« je bil na izboru, ki je tako kot vsako leto potekal v Cankarjevem domu, podeljen Rebeki Kropivšek Leskovar. Člane izborne komisije je prepričala njena socialna občutljivost za tehnologijo. Kot je v svoji predstavitvi poudarila nagrajenka, imajo nove tehnologije določen vpliv v družbi, tako dober kot slab, zato je pomembno, da že med njihovim razvojem inženir razmišlja, kakšen bo ta vpliv. Ker je tehnologija že zdaj tako prepletena z našim vsakdanom, da se od nje ne bomo mogli več ločiti, je »nujno, da že v času razvoja skrbimo, da ustvarjamo tehnologije, ki bodo na človeka vplivale čim bolj pozitivno«, sicer bo, kot je prepričana, eksponentno hiter razvoj tehnologij, kot je umetna inteligenca, uničujoč za vse. Zato se ji zdi ključno, »da inženirji tesno sodelujemo s strokovnjaki družboslovnih ved, kot so sociologi, filozofi, politologi in podobno, in to že od začetka razvoja novih tehnologij. Le tako lahko dobimo širši vpogled, kako lahko neka tehnologija, ki se še razvija, vpliva na naš vsakdan, in jo tako pravočasno prilagodimo, da bo njen vpliv čim bolj pozitiven« (Kropivšek Leskovar 2025).

Še danes se ta trditev mlade diplomantke strojništva – 250 let po začetku industrijske revolucije – bere presenetljivo sveže. A ne zato, ker bi pomenila nekaj novega, nasprotno. Spoznanje o socialnih učinkih tehnologije je že tako staro, da smo nanj že skoraj pozabili, zato se danes ta misel zdi novost, ker je tako redko omenjana.

In se nalašč zatira – ravnokar, med nastajanja tega prispevka, so ZDA umaknile podpis pariške deklaracije 60 držav o umetni inteligenci (AI), ki naj bi omogočila »odprto, transparentno, inkluzivno, etično, varno in trajnostno« regulacijo AI (Kosec 2025). Razlaga te subverzije ni v psihičnem profilu osebnosti ameriškega predsednika, pač pa v smeri razvoja, med katerim smo prišli do točke, da se moramo vprašati, kako je kaj takega sploh mogoče.

Pred pol stoletja, ko se je koncept tržnega fundamentalizma začel globalizirati (neoliberalizem), sta Rogers in Shoemaker na podlagi raziskave iz leta 1971 oblikovala pet kriterijev, ki morajo biti izpolnjeni, da »bi bila inovacija uspešno sprejeta« (Osborne in Brown 2005, 127):

- relativna prednost inovacije v primerjavi s prejšnjimi rešitvami;
- združljivost z obstoječimi tehnologijami in znanji;
- enostavnost razumevanja za končne uporabnike;
- možnost preizkušanja in
- opaznost rezultatov in dosežkov.

Nobeden od naštetih kriterijev ne vsebuje niti najmanjše asociacije na opozorila prej citirane mlade diplomantke strojništva. Enako velja tudi za številne kritike, ki so opozarjali na pomanjkljivosti Rogers-Shoemakerjevih kriterijev in jih povzemata Osborne in Brown (2005, 127–131) – skoraj nihče od kritikov¹ ni pogrešal presoje z vidika socialnih učinkov ali etičnih standardov.

V tem tisočletju se soočamo s precej hitrejšimi in verjetno nevarnejšimi tehnološkimi premiki kot pred desetletji, ko so nastali prejšnji kriteriji. Takrat se je še zdelo, da je največji družbeni problem, kako pravočasno prepoznati tehnološko inovacijo in jo začeti hitreje tržiti od konkurentov. Današnje skrbi, ko se vsakih nekaj tednov pojavi zmogljivejša različica umetne inteligence, so bolj dramatične. V svojem videu ob izidu nove knjige Yuval Noah Harari (2024) z globokim glasom napove – »AI je prva tehnologija v zgodovini, ki lahko ljudem odvzame moč«.² Niti tega ne vemo, ali je to res, ali gre le za reklamni učinek pretiravanja ali pa mogoče za tolažbo v smislu, da smo jo kljub odkritju fosilnih goriv in atomski bombi doslej dobro odnesli, čeprav smo mislili, da bo udarec, napovedan z AI, prišel že prej.

Pri teh vprašanjih ne gre za neznanje o tem, da tehnologija vključuje socialno dimenzijo. Težava je v njeni komodifikaciji, ker smo odmislili socialne učinke in zato zašli v slepo ulico. A to še ni razlaga, še vedno ostaja vprašanje, kako se je to lahko zgodilo. Tudi profit ni razlaga, saj je že vsaj od prve polovice 20. stoletja jasno, da ignoriranje socialnih učinkov povzroča stroške. Verjetno lahko komodifikacijo, zakaj nam je ušla iz rok, najbolje pojasnijo religijske zakonitosti (ne pa ekonomske). Stark in Bainbridge, teoretika religije, sta odkrila osnovni model za napoved vedenja verskih inovatorjev. Navajata, da je edina bistvena razlika med vsemi možnimi religijskimi akterji – v vseh mogočih kulturnih kontekstih, ko ustanavljajo in širijo svoje novosti, religije, kulte, ločine, cerkve, gibanja – razlika med tipom podjetnika in tipom duševnega bolnika (2007, 178): »Duševno bolan človek lahko ima svojo vizijo in jo skuša prenesti na druge, ne glede na to, kako sovražna je do kulta zunanja družba ali država. Podjetnik, ki išče dobiček, pa verjetno ne bo tako tvegale.«

Poanta je torej v interakciji akterja z okoljem. Iz tega sta avtorja (ibid., 179) izpeljala enega od 344 postulatov svoje teorije religije, gre za postulat številka 179: »Delež kultov, ki jih ustanovijo podjetniki, je v primerjavi s tistimi, ki jih ustanovijo duševno bolni, tem večji, čim manjše stroške nalaga družba pripadnikom kulta.«

Pri tem avtorja stroške opredelita kot »kar koli, čemur se ljudje skušajo izogniti«, medtem ko je dobiček vse tisto, »za kar si ljudje povzročijo stroške, da bi jih dobili« (ibid., 326).

Če to religijsko tezo prevedemo na tržno ekonomijo, postane jasneje, zakaj je zanemarjanje *socialne* tehnologije dobilo prav v obdobju modernizacije dramatične razsežnosti, zlasti po prvi svetovni vojni. V tej ignoranci socialnih učinkov postanejo jasnejši tudi motivi družboslovnih ved, saj so v prizadevanju za tržno učinkovitost tehnologij tudi družboslovci načrtno izumljali nadomestne izraze za socialno tehnologijo, da bi se jo omenjalo z drugimi, bolj asocialnimi termini, npr. »socialni inženiring« (Derksen in Wierenga 2013; enako tudi Schatzberg 2018, 152). Gre torej za podjetniško logiko, v kateri je aktivno vlogo imelo tudi družboslovje, čeprav ne zaradi nevednosti. Kajti na socialne posledice tehnologije je dovolj zgodaj, že na začetku industrijske revolucije, pokazal Marx (1961, 422, 423) z opozorilom, da je vsaka tehnologija socialna, saj vsaka »tehnologija odkriva aktivno stališče človeka do prirode /.../ in s tem tudi njegove družbene razmere, v katerih živi, ter duhovne predstave, ki izvirajo iz njih«. Enako tudi Pfaffenberger (1988), ki tehnologijo obravnava kot »humanizirano naravo« in vztraja, da gre v osnovi za družbeni pojav; zato po njegovem tehnologija ni nič drugega kot »družbena konstrukcija narave okoli nas in v nas, in ko je enkrat dosežena, izraža vgrajeno družbeno vizijo in nas vključuje v tisto, kar bi Marx imenoval oblika življenja« (Ingold 1997, 143).

Sicer pa je prvo opozorilo o socialnih učinkih tehnologije nastalo že več kot tisoč let pred industrijsko revolucijo in pred nastankom znanosti. V klasičnem obdobju antične Grčije, nekje v drugi polovici 4. stol. pr. n. št., so prvič nastali – danes spet presenetljivo aktualni – argumenti o socialnih učinkih tipične in najpomembnejše socialne tehnologije, pisave.³ Gre za Sokratov dialog s Fajdrosom, v katerem prvi opozori na škodljive socialne učinke pisave. Sokrat jih našteje kar pet, tu pa jih na kratko povzemam zaradi dveh razlogov, ponazoritve občutljivosti za socialne učinke in zaradi aktualnosti:

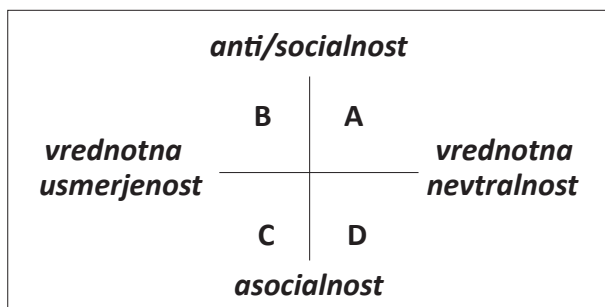
- **Izguba spomina:** Sokrat se boji, da bo iznajdba pisave »v dušah tistih, ki se je naučijo, zaradi zanemarjanja spomina povzročila pozabo, ker se bodo zaradi zaupanja v pisanje spominjali od zunaj, zaradi tujih znakov, ne pa od znotraj, sami od sebe. Torej nisi iznašel zdravila za spomin (*mnéme*), ampak za spominjanje (*hypómnesis*)« (Platon 2009, 569).
- **Zavajanje oz. pretvarjanje** (mimikrija) v smislu ustvarjanja lažnega védenja, ker se s pisavo učencem posreduje »videz modrosti, ne (njene) resnice« (ibid.).
- **Napačen didaktični pristop** pri posredovanju znanja učencem, saj podajanje informacij skozi pisavo – ker ne moreš dialogizirati z zapisanimi idejami – ustvarja zmoten vtis, da seznanjenost z informacijami zadostuje: »Veliko bodo namreč od tebe slišali, in ne da bi se poučili, se jim bo zdelo, da marsikaj poznajo, čeprav večinoma ne poznajo ničesar in se je z njimi težko družiti, ker niso postali modri, ampak navidezno modri« (ibid.).

- **Ignoriranje konteksta** v smislu nepoznavanja in brezbržnosti do podane informacije, saj je tisti, kot svari Sokrat, »ki misli, da je v črkah zapustil veččino, in tudi ta, ki je (to) sprejel, kot da bo od črk (prišlo) nekaj jasnega in trdnega, poln velike naivnosti« (ibid., 570).
- **Nezmožnost dialoga:** ker dialog predstavlja pot do resnice (zaradi dialektične metode), zato »ni vseeno, kdo je tisti, ki govori, in od kod je, saj ne gledaš samo na to, ali je s stvarmi tako ali drugače« (ibid.). Zgolj informacija ne zadošča, saj je pomemben tudi kontekst, od kod izvira in kdo jo prinaša.⁴

Zanimivo bi bilo vedeti, kaj bi danes Sokrat rekel o sodobni informacijsko-komunikacijski tehnologiji, še zlasti o AI, s katero se je mogoče pogovarjati na način, da ne veš, s kom ali s čim imaš opravka. Mogoče pa Sokrat sploh ne bi videl razloga, da bi si premislil, saj bi lahko navedel (vsaj) tri argumente, da so prejšnji očitki na račun pisave upravičeni: prvič, 2400 let stara kritika pisave je povsem enaka sodobnim didaktičnim opozorilom pred mehanskim učenjem »na pamet«, torej brez smisla in konteksta informacij (t. i. *rote learning*; Wikipedia 2024); drugič, Sokratovo kritiko je mogoče razumeti tudi tako, da opozarja na nevarnost napačnih načinov uporabe te socialne tehnologije, ne odklanja pa pisave *a priori* kot take; tretjič, v nobenem obdobju človeške zgodovine ni bil tolikšen delež ljudi, kot je danes, zlasti otrok, izpostavljen prav tem učinkom napačne uporabe informacijsko-komunikacijske tehnologije, na katere je opozoril Sokrat. Skratka, védnost o socialnih učinkih tehnologije in opozarjanje na njih ni nič novega, novo je, da smo jih zanemarjali prav v sodobnosti, ko bo cena take brezbržnosti največja. V zvezi s tem je sokratovska čuječnost dober zgled.

Hkrati pa še vedno niso rešene načelne dileme, ki so pomembne za razumevanje in regulacijo tehnologije, še zlasti socialne, na primer, ali je vsa(kršna) tehnologija socialna v smislu prej navedene Marxove misli, in če ni, kako razločiti tehnologijo s socialnimi učinki od tiste, ki je brez njih. Kaj je bolj res – so socialni učinki tehnologije posledica njene uporabe ali pa je nasprotno, da je uporaba posledica socialnosti, ki je vgrajena že v strukturo tehnologije? Te dileme povzemam v shemi 1. Namen provizorične kategorizacije možnih odgovorov ni v tem, da najdemo pravega, pač pa da lažje umestimo predmet tega članka, obravnavan v naslednjih razdelkih – univerzalni temeljni dohodek (UTD). Shema utegne biti koristna v obeh primerih: za iskanje načelnega odgovora o tem, katera kombinacija vrednotnih kvalifikacij (od prikazanih štirih) je najznačilnejša za socialne tehnologije, kot tudi v primeru, da razumemo spodnjo shemo kot ilustracijo empirične, ne pa načelne narave problema (torej v smislu, da je za vsak primer tehnologije treba posebej preveriti, v kateri kvadrant spada).

V shemi 1 abscisa ponazarja vrednotno dimenzijo z variantama vrednotne nevtralnosti na desni strani in vrednotno determinirano tehnološko strukturo na levi; ordinata pa označuje možne socialne in tudi antisocialne učinke tehnologije (zgoraj) ter možnost, da je tehnologija socialno nevtralna, ker je brez takih učinkov (spodaj).



Shema 1: Dimenzije socialne tehnologije

Kvadranti sheme 1 vsebujejo naslednje premise:

A. Ker je tehnologija vrednotno nevtralna, je njen družbeni vpliv heteronomen, saj je vedno in povsod odvisen od zunanjih dejavnikov (od vedenja, odločitev, vrednot, interesov njenih uporabnikov, od kulturnega in širšega družbenega konteksta).

Ker so socialni učinki tehnologije nujna posledica njene uporabe, je vsaka tehnologija socialna ali pa antisocialna – čeprav ni taka sama po sebi. Ta lastnost je skupna vsem tehnologijam in je neodvisna od stopnje njihove kompleksnosti. V tem smislu velja enako tako za paleolitski kamniti pestnjak, izdelan pred 1,7 milijona let, kot za najnovejše različice umetne inteligence (AI). Kot navaja Jure Leskovec, raziskovalec na Stanfordu, gre pri AI za: »izjemno močno orodje in, tako kot vsak/o/ tehnologij/o/, jo je mogoče uporabiti tako v pozitivne kot negativne namene. V demokratičnih družbah lahko AI izboljša upravljanje in javne storitve. V avtoritarnih režimih pa lahko postane orodje množičnega nadzora in cenzure« (Leskovec 2025).

Povsem enako opredelitev AI vsebuje najnovejša evropska uredba o regulaciji AI (Uredba EU 2024, četrty in peti odstavek). Ob tem je treba opozoriti tudi na kontekst socialne stratifikacije. V družbah, v katerih je izrazitejša, bodo večji tudi negativni učinki socialne tehnologije, saj bodo njene prednosti med prebivalci izrazito neenakomerno razporejene. To se že kaže celo pri nas v povezavi med AI in izobrazbo kot pomembnejšim stratifikacijskim virom, pa čeprav spadamo med najbolj egalitarne države na svetu, čeprav imamo nadpovprečno izobraženo prebivalstvo in čeprav ostajamo po razširjenosti AI pod evropskim povprečjem: po zadnjih podatkih je poznavanje tehnologije AI v Sloveniji najbolj razširjeno med terciarno izobraženimi prebivalci, takih je bilo 78 odstotkov, med tistimi s srednješolsko izobrazbo je bilo z AI seznanjenih 51 odstotkov, med osnovnošolsko izobraženimi pa 44 odstotkov (SURS 2025).

Bistvo A-kvadranta: tehnologija je sama po sebi (v smislu strukture) vrednotno nevtralna, medtem ko so njeni socialni in tudi antisocialni učinki določeni od zunaj (z uporabo). Atomska bomba je ubijalska le, če jo sprožimo, sicer pa ni.

B. Tehnologija je že v osnovi vrednotno usmerjena, saj so njeni nameni, cilji, obstoj in uporaba določeni že s prvim korakom njenega nastanka oz. celo pred njim (z idejo, načrtom, pogosto tudi z virom financiranja). Zato za vsako tehnologijo velja enaka lastnost, ki jo sicer najpogosteje pripisujemo informacijsko-komunikacijski tehnologiji,⁵ na primer: »Programska oprema je družbena v smislu, da je proizvod

družbenih procesov oziroma proizvod posameznikov in njihovih odnosov, poleg tega je tudi aplicirana in uporabljena v družbenih (sub)sistemih» (Lenarčič 2014, 286).

Tipični primeri socialne tehnologije, ki so že sami po sebi vrednotno determinirani in imajo zato socialno koristne učinke, kar jih uvršča v kvadrant B, so npr. jezikovni prevajalniki (npr. Google Prevajalnik), ki spodbuja povezovanje in komunikacijo, razni sistemi za pomoč invalidom (slušni aparati, Braillova pisava), ki omogočajo vključevanje in socialno povezanost, sistem GPS v reševalnih vozilih itd. Primeri tehnologije z inherentno negativnimi učinki pa so orožje za množično uničevanje, mučilne naprave, izsiljevalski virusi (*ransomware*), naprave za prikrito nadzorovanje itd.

Bistvo B-kvadranta: socialni ali pa antisocialni učinki tehnologije so že strukturno določeni in niso heteronomni (zunanj, naknadno dodani), saj nobena tehnološka struktura ni vrednotno nevtralna.

C. Tehnologija ni zgolj orodje, ki ga uporabljamo, temveč način razumevanja sveta, ker ga vse bolj razumemo (zgolj) skozi to orodje. Tehnologija po(eno)stavlja svet v okvir uporabnosti, saj z njeno uporabo vse stvari in bitja obravnavamo kot razpoložljive vire za izkoriščanje. Zato se s tehnološkim pogledom na svet odtujujemo od vsega, kar nas obdaja – vključno z družbo in odnosom do samega sebe – saj vse že s tem, ko postane predmet naše pozornosti ali delovanja, reduciramo na utilitaristično vrednost.⁶ Na ta način s(kozi) tehnologijo izgublamo stik z bistvom stvari in naše lastne biti, s tem pa izginja tudi naša zmožnost avtentičnega odnosa do bivanja. Na primer: rastoča produktivnost industrijske proizvodnje, avtomatizacija in robotizacija vse bolj postavljata naravo in človeka v vlogo surovin oz. delovnih sredstev, razpoložljivih za druge izdelke; prav s tem, ko tehnološki sistemi povečujejo učinkovitost informacijskih izmenjav (npr. digitalni algoritmi), določajo tudi načine našega dojemanja resničnosti ter njene interpretacije v smeri redukcije biti vseh stvari na izkoriščanje podatkov in optimizacijo procesov. Tehnologija torej ni zgolj nabor orodij in postopkov, pač pa način razkrivanja sveta, ki – kot navaja Heidegger (1977, 310) – človeka »zasušnjuje«, ker ga potiska v vlogo pasivnega udeleženca v tehnoloških strukturah oz. procesih: »Grožnja človeku ne izvira v prvi vrsti iz potencialno smrtonosnih tehnoloških strojev in naprav. Dejanska grožnja je že prizadela človeka v njegovi biti. Način, na katerega tehnologija ureja in razkriva (*enframing*) svet, grozi človeku, da mu bo odvzeta možnost dostopa do bolj prvinskega razkritja in s tem tudi do izkustva temeljne resnice.«

Kar loči logiko C od kvadrantov B in A, je zavračanje determinizma pri socialnih učinkih. Poudarek, da nobena tehnologija ni vrednotno nevtralna, ne pomeni, da so nujne tudi anti/socialne posledice, saj se jim lahko zavestno izognemo – ne z begom pred tehnologijo (zavrnitvijo), pač pa z njenim boljšim razumevanjem, da ji ne podležemo. V tem poudarku se Heideggerjevo opozorilo bistveno loči od determinizma iz prejšnjih kvadrantov, kjer prav to ni možno. Kajti po logiki A, kot smo ugotovili, je socialna kvalifikacija zgolj naknadni učinek rokovanja s tehnologijo, pri katerem je edini zaplet v tem, kako si odgovorimo na »preprosti« vprašanje o namenu in zmožnosti: kaj *hočem* in ali s tem *znam* ravnati. V primerjavi z A pa v kvadrantu B te možnosti sploh ni več, zato se o njej niti ne sprašujemo, saj so socialni in antisocialni

učinki derivat inherentno zakodirane vrednotne usmerjenosti v strukturi tehnologije. Zaradi posebnosti tega pristopa – ki ga zato uvrščam v kvadrant C – Heidegger navaja Holderlinova verza, ki se glasita tako:

*toda tam, kjer je nevarnost, raste
tudi moč odrešitve.*⁷

To pesniško misel Heidegger v nadaljevanju interpretira v smislu, da kjer je problem, je tudi klic k rešitvi, ta apel pa izhaja iz problema samega.⁸ »Ker bistvo tehnologije ni nič tehnološkega« (v primerjavi z logiko B), se mora temeljito razmišljanje o tehnologiji zastaviti znotraj nekega drugega, drugačnega, tipično netehnološkega področja. Katerega? Tu Heidegger pokaže na umetnost, ker je – kot ekstrakt kulture – neprimerno bolj specializirana za vrednote(nje) kot pa tehnologija, čeprav je tudi slednja kulturni produkt (zato uvrstitev kvadranta C na levo polovico sheme). Ta proces nas po Heideggerju privede do postavljanja vprašanj prav tam, kjer z osupljivim tehnološkim razvojem spraševanje najhitreje izginja in se s tehnološko ekspanzijo širi videz neizogibnosti, ki postaja samoumeven. Ta zgrešena pot nas provocira podobno kot izgubljenca v slepi ulici: »čim bližje smo nevarnosti /.../, tem bolj smo nagnjeni k spraševanju«, kot navaja Heidegger. Prav soočenje z brezizhodnostjo izgubljenca sproži alarm za rešitev (z vprašanjem, kje bi utegnila biti), prav to pa je seveda že samo po sebi odločen odmik od fatalizma. Stiska in brezizhodnost nas silita k spraševanju, brez katerega ni mišljenja, s tem pa se upiramo samoumevnosti.⁹

Bistvo kvadranta C: tehnološki determinizem obstaja, nikakor ni vrednotno nevtralen, in čeprav je fatalen (redukcija biti vsega, kar je), ni fatalističen v smislu, da bi bilo vse določeno vnaprej in da človek ne bi mogel ničesar spremeniti. Heidegger ne zagovarja rousseaujevskega bega od tehnologije. Nasprotno, vzpodbuja nas, da o njej premislimo še bolj od blizu (kot pa nam to sugerira rutina njene uporabe), jo vzamemo zares in se soočimo z njenim bistvom na netehnološki način, npr. tako kot v umetnosti(h). Le tako lahko dosežemo bolj avtentičen odnos do resnice.

D. Ker je tehnologija v osnovi vrednotno nevtralna, saj sama po sebi nima vnaprej določenih namenov niti učinkov (socialni akterji namreč njen vpliv oblikujejo naknadno), se pri njeni uporabi sprožata nevarna trenda, ki se z razvojem stopnjujeta: zamudništvo in nekompetentnost. Prvi učinek se kaže v povečevanju tveganja, da nas razvoj prehiteva v smislu prepozne prilagoditve nanj, saj vse bolj zaostajamo z uvažanjem družbenega nadzora nad tehnologijo (etični standardi, varnostni protokoli, nadzorni sistemi). Učinek nekompetentnosti pa izhaja iz rastoče kompleksnosti tehnološke strukture in njenega komodificiranja v razmerah odsotnosti globalne družbene regulacije. Še vedno ne znamo ravnati s pravilom, da če nekaj, kar je bistvenega pomena za kakovost življenja posameznikov in skupnosti, prepustimo izključno tržnim zakonitostim, se ta dobrina vedno sprevrže v neželeno smer – in pri tehnologiji ni nič drugače (tudi tehnologija je dobrina). Tudi zanjo velja, da so antisocialni učinki premosorazmerni z vse hitrejšim tehnološkim razvojem in disfunkcionalnim odzivanjem nanj, za kar je kriva družba, ne pa tehnološki ustroj sam po sebi. Poanta kvadranta D je, da omenjena trenda ne izhajata iz tehnologije, pač pa iz družbene strukture. Zato ista tehnologija v različnih družbah, kulturah in geografskih okoljih sproža zelo različne učinke, lahko funkcionalne, disfunkcionalne kot tudi nikakršne. Čeprav se je

izum kolesa pojavil pri Sumercih že pred 3500 leti pr. n. št., ga niso vse sosednje kulture prevzele istočasno in na enak način. Tam, kjer so kolo poznali, pa ga niso uporabljali. Npr. v srednjeameriških kulturah, kjer so kolo sami iznašli, ga nikoli pred 16. stoletjem (pred stikom z Evropejci) niso razvili v smeri voza, samokolnice, lončarskega vretena ali katerega koli drugega praktičnega predmeta s kolesi – je pa bilo kolo uporabljeno kot sestavni del v številnih otroških (»vlečnih«) igračah, tudi iz obdobja približno 500 pr. n. št. (Wikipedia 2025). Enako je danes: isto tehnologijo, npr. jedrsko fisijo, nekatere družbe uporabljajo samo za ustvarjanje električne energije, druge predvsem kot jedrsko orožje, nekatere pa se ji povsem odpovedujejo, čeprav je ista tehnologija dostopna vsem. Kljub globalnemu dostopu do koncepta socialne države – verjetno najpomembnejšega primera socialne tehnologije 20. stoletja –, ki je nihče ni zavaroval s patentnim monopolom, je socialna država postala značilna samo v evropskem prostoru, ne pa tudi v najbolj moderniziranih in najbogatejših državah sveta (npr. v ZDA in Veliki Britaniji). Primer enake vrste je tudi učinek AI na neenakost. Pri neregulirani uporabi AI se bo, kot doslej, neenakost samo še stopnjevala, predvsem zaradi naraščajoče neenakosti plač. Lahko pa bi bilo tudi nasprotno, kot ugotavljajo pri Mednarodni zvezi za robotiko in navajajo, da sposobnost AI (Prettner 2025):

za opravljanje nerutinskih, visoko usposobljenih nalog lahko zmanjša razliko v plačah med nizko in visoko usposobljenimi delavci/.../. Sprejemanje pametnih predpisov in pravil za uporabo AI ter zagotavljanje, da je čim več ljudi sposobnih delati z AI prek izobraževanja in (pre)usposabljanja, so pomembni predpogoji za spodbujanje bolj vključujočega in pravičnega trga dela, kjer ima vsakdo potencial, da izkoristi prednosti tehnološkega napredka.

Bistvo kvadranta D: logika tega razdelka je nevtralnost. Tehnologija je sama po sebi asocialna in brez vrednot; vse je v rokah njenih uporabnikov in konteksta, v katerem delujejo.

Čeprav na vprašanja iz sheme 1 ni enoznačnih odgovorov, se nakazanim možnostim ne moremo izogniti. Osnovna poanta predstavljenih kvadrantov je v njihovih razlikah (A, B, C, D), ki niso zanemarljive pri kvalificiranju (konkretne) socialne tehnologije.

Univerzalni temeljni dohodek (UTD) spada v kvadrant B. V nadaljevanju opisujem, zakaj.

UNIVERZALNI TEMELJNI DOHODEK

Socialna tehnologija se nanaša na načine organiziranja in upravljanja družbenih sistemov za izboljšanje kakovosti življenja. Eden od teh načinov je tudi pravica do univerzalnega temeljnega dohodka (UTD). Glede na dolgo zgodovino te ideje – in pomembnost posledic, če bi jo uresničili – je razumljivo, da se pravica do UTD ne označuje zgolj z navedeno sintagmo. Najpogostejša poimenovanja so: »državni bon«, »nacionalna dividenda«, »družbena dividenda«, »državljska plača«, »državljski dohodek«, »temeljni dohodek«, »univerzalni dodatek«, »negativna dohodnina« itd. (Van Parijs 2004, 22), temu bi lahko dodali še »zemljiško rento«, če se izplačuje brez-

pogojno vsem državljanom oz. rezidentom, in »dividendo stalnega sklada«, kot so jo uzakonili na Aljaski. Čeprav te pojmovne variante za članek niso pomembne, je treba poudariti, da največkrat ne gre zgolj za terminološke, pač pa tudi za vsebinske razlike (popis in analiza poimenovanj na temo UTD, ki se uporabljajo v slovenskem prostoru, sta v Dragoš 2011, 114–121).

Na področju socialne tehnologije gre pri UTD za podobno inovacijo, kot so, denimo, socialna država, participativno odločanje, sodelovalna ekonomija, družbena omrežja itd. Ena od razlik med temi inovacijami in UTD je v časovnici in realizaciji. Med naštetimi socialnimi tehnologijami je UTD primer najstarejše nerealizirane ideje in to kljub temu, da bi lahko z UTD revščino zmanjšali (in hkrati povsem ukinili absolutno revščino!), odpravili ključne pomanjkljivosti socialne države, ob vrsti drugih koristnih učinkov (več o njih pozneje) pa je v sodobni okoljski krizi dobil UTD še dodatno funkcijo – brez realizacije te pravice bo konsenz o ekoloških ukrepih neizvedljiv (Dragoš 2024).

Pomembne koristi UTD bi bile še zlasti na področju, s katerega prihaja največji odpor proti tej ideji, to je v ekonomiji. Kajti neoliberalna vera o tako imenovani alokativni in distributivni funkciji trga dela ne velja. Alokativna funkcija ne uspeva optimalno razporediti delovne sile, ker delavci pogosto obtičijo na neustreznih delovnih mestih z omejenimi možnostmi za razvoj in s podplačanim delom. Distributivna tržna funkcija pa ne prinaša pravične porazdelitve dohodkov in priložnosti, saj zaposlavlja ranljive skupine, kar vodi v povečanje socialnih in ekonomskih neenakosti. Ti argumenti kažejo na temeljne strukturne pomanjkljivosti nereguliranega trga dela in opozarjajo na potrebo po reformah, ki bi omogočile pravičnejšo in učinkovitejšo razporeditev delovnih priložnosti, kot dokazujeta Pulignano in Domecka (2025). V tržnih gospodarstvih bi z UTD povečali fleksibilnost delovne sile, bistveno bi zmanjšali negotovost, olajšali kompetitivnost, hkrati pa razpeli varovalno mrežo v primeru izpada s trga delovne sile, ki je zdaj luknjičava, izsiljevalska ali pa je sploh ni, kajti delovna mesta izginjajo v večji meri, kot pa nastajajo nova. Kitajska kot najbolj robotizirana država je leta 2020 povečala letno namestitvev industrijskih robotov za 30 odstotkov, leta 2023 pa za 47 odstotkov (IFR 2025). V Indiji, ki je najmanj robotizirana, a hkrati predstavlja eno od najhitreje rastočih azijskih gospodarstev, pa se je vgradnja robotov leta 2023 povečala za 59 odstotkov, kar je nova najvišja vrednost (Viršek 2025). Ta tehnološki trend se dogaja z vzporednim svetovnim trendom, z razraščanjem prekarnosti, vključno s tistim delom nesmiselnih služb, ki jih je antropolog Graeber poimenoval kot »*bullshit jobs*« (Graeber 2019, 2013; Dragoš 2021) – gre za delovna mesta, ki pohablajo, poneumljajo in cenovno niso vredna bistveno več od UTD. Še hujši problem od nesmiselnih služb so nedostojne oz. ničvredne zaposlitve (»*junk jobs*«), ki presenetljivo trdovratno vztrajajo celo v najrazvitejših in socialno najbolj reguliranih evropskih ekonomijah. Čeprav so taka podplačana delovna mesta manjšinska in ker se tudi sama navznoter delijo na dve kategoriji, na družbeno nekoristna in na nujna,¹⁰ lahko prav pri tej kategoriji ničvrednih zaposlitev uvedba UTD sproži koristno čiščenje nacionalne ekonomije. Ker bi z okrepljeno eksistenčno avtonomijo najnižjega proletariata nepotreben del »*bullshit*« in »*junk jobs*« izginil, medtem ko bi se njihov nujen del preselil med dostojno plačana dela, bi se z UTD povečala fleksibilnost delodajalcev – postali bi dovetnejši za kreativnejše načine povečevanja

dobička, v primerjavi z zdaj, ko jim zadošča varčevanje na račun delavskih mezd, saj z globalizacijo kapitala in delovnih mest lahko kljubujejo vsem pritiskom. Ne pozabimo, da je v jedru problematičnosti omenjenih »*bullshit*« in »*junk jobs*« tisto, na kar družboslovje opozarja že dve stoletji in česar razvoju socialne države niti v njenem zenitu ni uspelo rešiti – to je odtujitev z dvema temeljnima značilnostma. Prva je v njeni invazivnosti oz. kompletnosti, saj se odtujitev iz ekonomskih razmerij razširja na druge sfere (Marx jih omenja pet: ekonomsko, politično, religijsko, socialno in individualno).¹¹ Enako tudi Standing (2018, 49): »Odtujevanje izhaja iz védenja, da tisto, kar nekdo počne, ni koristno za njega samega in tudi ni nekaj, kar bi zbujalo spoštovanje in odobravanje; delo opravlja zgolj za nekoga drugega in v korist nekoga drugega.«

Druga temeljna značilnost odtujitve pa je njena inertnost v smislu, da je ni mogoče odpraviti zgolj z moraliziranjem (npr. s sklicevanjem na prave vrednote), pač pa le z organiziranim pritiskom. Za oba omenjena problema odtujitve je najhitrejša in najelegantnejša rešitev UTD, saj najbolj koristi najbolj marginaliziranim segmentom delovne sile (na »*bullshit*« in »*junk*« delovnih mestih). Za uvedbo te pravice ne potrebujemo družbenih prevratov, ni treba čakati na izumljanje antikapitalističnih alternativ in tudi fiskalno, z vidika razpoložljivih sredstev, UTD ne zahteva nikakršnega odrekanja kogarkoli, niti uvedbe novih davkov niti povišanja obstoječih (temu je namenjen zadnji razdelek). Skratka, UTD bi pretrgal začarani krog ujetosti v ta nesmiselna delovna mesta, saj njihova zavrnitev ne bi bila več eksistenčno tvegan projekt. S tem bi se vsem državljanom prvič v zgodovini kapitalizma malce – a bistveno – odprla vrata v tisti način produkcije, ki ga je Marx imenoval kot »kraljestvo svobode«: gre za dekomodifikacijo dela, ki ni več instrumentalizirano za potrebe preživetja, pač pa za izražanje kreativnosti, ko je delovna aktivnost namenjena sami sebi. Seveda bo še vedno ostalo tudi tako delo, ki se mu ni mogoče izogniti, ker je eksistenčno nujno, čeprav ni nujno prijetno. A zdaj bo z UTD tudi ta nujni del – »kraljestvo nujnosti« – osvobojen manipulacije, saj z njim ne bo več mogel nihče izsiljevati nikogar. Šele ko je ljudem brezpogojno zagotovljena eksistenca (vsaj) na minimalni ravni, se jim odškrnejo vrata v kraljestvo svobode, vključno z možnostjo, da se zdaj sami odločajo, ali bodo sploh stopili skozi ta vrata ali ne.¹² A ne pozabimo še dodatnega pogoja za vstop v kraljestvo svobode, ki ga v tretjem delu Kapitala omenja že Marx (štiri desetletja za njim pa John M. Keynes), to je skrajšava delovne dne.¹³ Kajti šele onstran kraljestva nujnosti – čeprav ne brez njega – kot pravi Marx (1973, 914), »se začne razvoj človeške moči, ki je samemu sebi namen, resnično kraljestvo svobode, ki pa se lahko razcvete samo na tem kraljestvu nujnosti kot svoji podlagi. Temeljni pogoj je skrajšati delovni dan.«

Šele pred nekaj leti smo (po stoletju in pol od prej omenjene Marxove diagnoze) dobili prvo globalno študijo, ki je zajela 194 držav, in sicer o smrtonosnosti dela zaradi predolgega delavnika. Med letoma 2000 in 2016 se je število smrti zaradi srčno-žilnih bolezni, povzročenih zaradi predolgega delavnika, povečalo za 42 odstotkov, možganske kapi zaradi istega vzroka pa za 19 odstotkov, kar pomeni skupaj 745.000 smrtnih primerov (WHO 2021). Prav to je tudi Graeberjevo (2020) svarilo: »Če ne bomo nehali delati manj, bomo kmalu postavljeni pred izbiro med katastrofami, ob katerih bo sedanja pandemija videti kot nedeljski sprehod skozi park.« Skratka, delo ubija, če je predolgo. Pogoj za skrajšavo delovnika pa že dlje časa ni več tehnološki, pač

pa ideološki (ibid.). Edino hitro, zanesljivo in učinkovito cepivo proti tej pandemiji je uvedba UTD.¹⁴

Zgodovina ideje o brezpogojni pravici do UTD je stara 180 let. Prvi jo je formuliral francoski utopični socialist Charles Fourier, njegovo idejo pa je leta 1848, v obdobju nastanka Marxovega *Komunističnega manifesta*, širil bruseljski fourierist Joseph Charlier s propagiranjem »ozemeljske rente«, ki naj bi se redno izplačevala vsem državljanom kot lastnikom nacionalnega ozemlja. Leto pozneje je John Stuart Mill v novi izdaji svoje knjige *Načela politične ekonomije* zelo naklonjeno opozoril na fourierizem kot »med vsemi najbolj večje sestavljeno obliko socializma,« ob tej pohvali pa se je tudi izrecno in nedvoumno zavzel za UTD: »Pri razdelitvi se vsakemu članu skupnosti, naj bo delazmožen ali ne, najprej dodeli neki minimum za preživetje« (cit. po Van Parijs 2004, 22). Od takrat dalje ideja o UTD ni nikoli zamrla. Polemike in ekspertize – vključno s praviimi družbenimi eksperimenti – srečamo na prav vseh kontinentih (razen Antarktike). Je pa res, da ideja še ni nikjer realizirana, razen na Aljaski. Tam UTD financirajo z dobički od prodaje naravnih virov, zlasti nafte, in sicer od leta 1976 dalje, ko so z ustavnim amandmajem ustanovili sklad PFD (*Permanent Fund Dividend*) ter s tem zagotovili, da mora izkoriščanje naravnih dobrin koristiti tudi bodočim generacijam in ohranjati dolgoročno finančno stabilnost države. Dejanska izplačila dividend iz omenjenega sklada pa so začeli ljudem nakazovati od leta 1982 dalje (več o tem glejte v PFD 2025; APFC 2025).

Pri nas se je za UTD prvi zavzemal že Janez Ev. Krek pred 125 leti (Krek 1901/1925, 44; Dragoš 2019a, 31, 32). Takoj za njim ga je omenjal pred drugo svetovno vojno sociolog Andrej Gosar (1933, 552; Dragoš 2019a, 31–39), njegov prvi zagovornik v obdobju samostojne države Slovenije pa je bil akad. Veljko Rus (1990, 254–260; 1992, 144–148), potem pa so sledili še štirje zborniki na to temo¹⁵ in več člankov v znanstvenem tisku in množičnih medijih. V tem smislu tudi za Slovenijo velja podobno kot za mednarodni prostor. Čeprav je prva pobuda za UTD pri nas, kot rečeno, stara že sto let, je ideja o tem v obdobju socializma povsem izginila iz strokovne in splošne javnosti, saj se je predpostavljalo, da ni združljiva s socialističnim samoupravljanjem. Od osamosvojitve dalje pa se je zahteva o UTD spet revitalizirana, celo do te mere, da so se o njej javno izrekli mnogi politiki in predstavniki političnih strank (Pribac in Korošec 2011, 297–309), pred štirinajstimi leti pa je bil tudi posvet v Državnem svetu Republike Slovenije o UTD, o katerem je to predstavniško telo izdalo tudi zelo poglodbeno in afirmativno poročilo s priporočili (ibid., 311–319). Kljub temu si še nihče ne zna predstavljati, kdo in kdaj bi lahko to idejo realiziral – a ne zato, ker ne bi vedeli, kako to narediti, ker ideja še ne bi dozorela, ker je ne bi poznali ali ker si je ne bi mogli privoščiti.¹⁶ Edini problem je pomanjkanje konsenza. V raziskavah javnega mnenja je leta 2013 UTD podpiralo 36,1 odstotka slovenske javnosti, kar je dvakrat večji delež od tistih, ki so nasprotovali tej ideji, hkrati pa je med akademskimi izobraženci bilo podpornikov le 19 odstotkov in med politiki še manj, 17,4 odstotka (Dragoš 2019a, 49, 50).

Lanskoletno poročilo OZN o ekstremni revščini in človekovih pravicah (v zvezi s tem) opozarja na rastočo problematiko težav v duševnem zdravju. Poročilo navaja, da danes kar 970 milijonov ljudi po vsem svetu živi z duševno motnjo, kar je 12 odstotkov svetovnega prebivalstva; več kot 280 milijonov ljudi trpi za depresijo, 301 milijon

ljudi se sooča s tesnobo, vsako leto pa naredi samomor 700.000 oseb, kar je četrti najpogostejši vzrok smrti med mladimi, starimi od 15 do 29 let (OZN 2024, 3). Ta črna slika se utegne še poslabšati zaradi negativnih ekonomskih¹⁷ in ekoloških trendov, zaradi rastočega nezaupanja v znanost in zaradi drugih, vse verjetnejših katastrof, od politično-vojaških do bioloških. Na primer, v prvem letu pandemije covida-19¹⁸ se je razširjenost depresije in tesnobe povečala kar za 25 odstotkov, kar je enoletni dodatni prispevek k skoraj milijardi ljudi, ki so že prej živeli z duševno motnjo – vse to je bila posledica porasta socialne izolacije in njene kombinacije z ekonomskimi težavami (WHO 2022a).¹⁹ Navajanje teh števil ne pomeni zgolj vzbujanja občutljivosti za trpljenje in neenakost, gre za temeljne pravice. Po merilu Svetovne zdravstvene organizacije je namreč duševno zdravje kvalificirano kot »osnovna človekova pravica«, ki je ključna »za osebni, skupnostni in socialnoekonomski razvoj« (WHO 2022b). Zato omenjeno poročilo OZN (2024, 2) poziva države, »naj preidejo z biomedicinskega pristopa k duševnemu zdravju, ki ga obravnava kot problem posameznika, na pristop, ki upošteva njegove socialne dejavnike. Da bi se spopadli z globalnim porastom depresije in tesnobe, bi morali storiti več za boj proti revščini in neenakosti ter za reševanje ekonomske negotovosti.«

Dokazov o tem, da neenakost povečuje psihične probleme prebivalstva, je veliko. Obsežno mednarodno empirijo o pozitivni korelaciji med neenakostjo v posameznih državah in stopnjo psihičnih težav med prebivalstvom zasledimo v Wilkinson in Pickett (2012) – mehanizem te povezave avtorja pojasnita takole: »Če je neenakost večja, tako da eni ne veljajo nič, drugi pa vse, postane položaj, kamor nas uvrščajo, še pomembnejši. Večjo neenakost verjetneje spremljata zaostreno statusno tekmovanje in povečana statusna anksioznost« (ibid., 45).

Ta problem je v sistemih socialne države težko rešljiv. Pri tem ne gre za utopične zahteve, pri katerih bi kritiko socialne države gradili na (sicer resničnem) dejstvu, ki velja za vse njene blaginjske variante – od manj do bolj »radodarnih« – da ne more odpraviti družbenih neenakosti v tržni ekonomiji. Realen problem koncepta socialne države so pogoji in upravičenja, pod katerimi je posameznik, ki se sam ne more preživljati, sploh lahko opredeljen kot upravičenec do socialne »pomoči«:²⁰ »Dokazi pričajo, da so višji dohodki povezani z boljšim zdravjem, zlasti duševnim, in dobrim počutjem, medtem ko je znano, da je zdajšnji sistem ugodnosti – mešanica preverjanja sredstev, pogojenosti, stigme in negotovosti glede tega, ali bo podpora umaknjena – povezan s slabim duševnim in telesnim zdravjem« (Reed et al. 2022, 12).

In na tej točki je prispevek morebitne uvedbe pravice do UTD odločilen. V sistemu socialne države bi namreč UTD lahko – brez radikalnih posegov v družbeno strukturo, vključno z ekonomijo – takoj, tako rekoč v hipu, odpravil prav tisto pogojevanje, preverjanje, nadzorovanje in stigmatiziranje, o katerem je govor v prejšnjem navedku. Zato se omenjeni avtorji v zvezi s preprečevanjem psihičnih težav zavzemajo za splošno državljansko in brezpogojno pravico do UTD. Ta bi pomenila (poleg koristi za upravičence) tudi temeljno varovalko za koncept socialne države, saj bi omenjeni ukrep dokončno blokiral politično licitiranje temeljnih socialnih pravic navzdol, kar je smer, v katero se razvija tako evropska kot slovenska socialna politika.²¹

Ob tem sta pomembni dve vprašanji. Bi bilo sploh možno z zmanjšanjem revščine in družbenih neenakosti posameznikom **kratkoročno** izboljšati duševno zdravje, ali ni

to iluzorno, saj gre za tek na dolge proge? In drugič, ali bi morebitna uvedba UTD res tako učinkovala na duševno zdravje, kot npr. pričakujejo prej navedeni avtorji? Odgovora na obe vprašanji sta pritrdilna (čeprav nista spektakularna). Neenakosti seveda niso odpravljive z UTD, kaj takega bi bilo od te pravice napačno pričakovati, saj to ni njen namen, še manj domet – če zanemarimo ta poudarek, stopimo v demagoško past, ki jo uporabljajo nasprotniki UTD, vključno z najrigidnejšimi zagovorniki neoliberalizma, ki jim kakršnikoli pozivi k zmanjšanju neenakosti dišijo po socialistični diktaturi.²² Odgovori so namreč odvisni tudi od formulacije vprašanj. Ni vseeno, ali vprašamo, denimo, ali bi uvedba participativne demokracije na lokalni ravni odpravila probleme z demokracijo in s strankokracijo v sodobnem svetu. Jasno, da ne. Ampak to seveda ni argument proti participativni demokraciji; z njo izboljšamo deliberativno kulturo, državljani so bolj vključeni v odločanje, odločitve so kakovostnejše, utrjujemo zaupanje v institucije, ne moremo pa odpraviti pomanjkljivosti demokracije na svetovni niti na državni ravni. Ali bo ustavna pravica do pitne vode odpravila revščino v Sloveniji? Seveda ne, a to ni argument, da smo proti njej. UTD ne more zmanjšati neenakosti, ne more izboljšati demokracije, ne more sanirati ekološke degradacije ... Lahko pa povsem odpravi absolutno revščino, blaži relativno, vključno z energetsko revščino, zmanjša psihične pritiske na ljudi, ker odpravi eksistenčno skrb za preživetje (čeprav le na ravni najosnovnejših potreb), drugih pozitivnih učinkov, od ekonomskih do političnih, pa je še več.²³ Odgovor na drugo vprašanje pa je najnovejši, največji in najbolj elaboriran eksperiment z UTD, ki dokaže, da brezpogojno zajamčen in reden dohodek v višini, ki prejemnikom omogoča kritje temeljnih življenjskih stroškov,²⁴ zagotavlja takojšnjo pomoč za duševno zdravje (De Schutter in van Parijs 2024).

EKSPERIMENT

Leta 2020 je v ameriških zveznih državah Teksas in Illinois 1000 posameznikov z nizkimi dohodki začelo tri leta prejemati 1000 dolarjev na mesec. Eksperimentalna skupina (prejemnikov) je bila sestavljena iz treh dohodkovnih kategorij: pod pragom revščine (nižji dohodek), med 100 in 200 odstotki praga revščine (srednji dohodek), v tretji kategoriji pa so bili tisti med 200 in 300 odstotki praga revščine (višji dohodek). Šlo je za UTD v pravem pomenu, saj ni bilo nobenih pogojev za njegovo pridobitev niti za trošenje, udeleženci so lahko denar porabili, kot so želeli. Rezultati tega obsežnega in natančno izvedenega eksperimenta ponujajo dragocen vpogled v to, kaj tak ukrep – tipičen primer socialne tehnologije – prinaša tako posameznikom kot širši skupnosti.²⁵ V nadaljevanju povzemam samo ugotovitve s treh področij: v zvezi z zdravjem, zaposlitvijo in denarnim stanjem.

Zdravje

Eno od področij, na katero so se usmerili raziskovalci, je bil vpliv brezpogojnih denarnih transferjev na zdravje posameznikov. Ugotovitve kažejo, da taki transferji lahko izboljšajo različne zdravstvene kazalnike, predvsem z zmanjšanjem stresa in izboljšanjem prehrane. V splošnem pa je bil dodaten gotovinski priliv (z naslova UTD) za prejemnike koristen zaradi trojega: zadovoljitve najnujnejših potreb, ki si jih prej niso mogli privoščiti, nekateri so postali avtonomnejši pri določanju svojih prioritet,

drugim pa je dodaten denar »predstavljaj pomemben del sestavljanke pri soočanju z globoko zakoreninjenimi neenakostmi v zdravstvu« (Dotson 2024a).

Zmanjšanje stresa. Prejemniki brezpogojnih denarnih sredstev poročajo o nižjih ravneh stresa, kar pozitivno vpliva na njihovo duševno zdravje. To je sicer veljalo le za prvo leto eksperimenta, saj so v drugem letu ti učinki zbledeli. Nasprotniki UTD z veseljem navajajo, da je (zgolj) kratkotrajno izboljšanje dokaz, da so obljube o pozitivnih psihičnih učinkih UTD neutemeljene. A verjetno gre za pretiravanje. Odsotnost dolgotrajnih učinkov v tej zvezi ne pomeni neučinkovitosti UTD, pač pa je verjetnejša razlaga, da ekonomska negotovost pomembno vpliva na duševno zdravje (Baird, de Hoop in Özler 2013) in da bo ta občutek varnosti pri prejemnikih UTD neizogibno usahnil, ko se bo približeval konec denarnih transferjev. Njihova vrednost ni le v dodatnem dohodku gospodinjstva, pač pa tudi v varnosti, predvidljivosti in stabilnosti, ki jo UTD zagotavlja. Kaj bi pomenilo ljudem s psihičnimi obremenitvami, ki so v revščini, zagotoviti varovalko proti tveganju, da padejo še nižje ali se znajdejo na cesti? »Ker se njihova kriza duševnega zdravja iz dneva v dan pogloblja, je to vprašanje, ki si ga je vredno zastaviti« (De Schutter in van Parijs 2024).

Izboljšana prehrana. Dodatna finančna sredstva omogočajo dostop do kakovostnejše hrane, kar vodi v boljše prehranske navade in splošno zdravje, a tudi ta učinek so raziskovalci potrdili samo za prvo leto eksperimenta. Zakaj so v drugem letu ti učinki zbledeli, navajajo raziskovalci več možnosti.²⁶ Na primer: dvig cen hrane, porast splošne inflacije (tamkajšnja inflacija je bila 5,4-odstotna v enem letu), opuščanje nekaterih programov socialne zaščite, ki so veljali med pandemijo, kar je spodkopalo izboljšave iz prvega leta eksperimenta.

Povečan dostop do zdravstvenih storitev: Brezpogojna denarna nakazila povečujejo dostop do bolnišničnih in urgentnih zdravstvenih storitev ter do nekaterih vrst ambulantne oskrbe (zlasti zobozdravstvene in specialistične); pri prejemnikih denarja se je znatno povečala mesečna poraba za zdravstveno oskrbo, in sicer za približno 20 dolarjev več od povprečnega udeleženca kontrolne skupine. Pri prejemnikih UTD – še zlasti moških – je bilo ugotovljeno tudi zmanjšanje neodgovorne uporabe alkohola (v primerih, ko bi alkoholiziranost otežila opravljanje obveznosti) ter zmanjšanje dni, v katerih so bila uživana protibolečinska zdravila brez recepta.

Zaposlitev

Stopnje zaposlenosti in delovni čas so se med eksperimentom znatno povečali tako za prejemnike UTD kot za kontrolno skupino, čeprav ne enako za obe skupini. Verjetnost, da bodo prejemniki UTD zaposleni, je bila v povprečju za dve odstotni točki manjša kot pri kontrolni skupini, to pomeni, da so prejemniki v povprečju delali 1,3 ure manj na teden od udeležencev kontrolne skupine. Ob začetku eksperimenta je bila brezposelnost zaradi pandemije covida-19 visoka, zaposlenih je bilo 58 odstotkov prejemnikov in 59 odstotkov kontrolne skupine; do konca eksperimenta pa se je stopnja zaposlenosti povečala na 72 odstotkov med prejemniki UTD in na 74 odstotkov v kontrolni skupini.

Še bolj kot majhne razlike med skupino prejemnikov UTD in kontrolno skupino je verjetno pomembno, da je raziskava natančno evidentirala velike razlike med udeleženci (tudi znotraj skupin) v motivih, kako so se ljudje odločili za prerazporeditev

delovnega in prostega časa. To evidenco raziskovalci interpretirajo kot dokaz, da je UTD povečal motivacijo za tako zaposlovanje, ki je bolj skladno z individualnimi okoliščinami in cilji iskalcev dela. Na primer, nekaterim prejemnikom so denarni transferji omogočali dokončati šolanje ali jim ponudili priložnost za nadaljnje izobraževanje oz. usposabljanje. Raziskovalci so dokazali večji učinek UTD na formalno izobraževanje zlasti v mlajši starostni skupini prejemnikov UTD (do 30 let), saj so mlajši odrasli pogostejše porabili denar za vpis v terciarne izobraževalne programe in preživeli manj ur v službi kot pa v šoli. Tudi ljudem, ki so delali dolge ure v službah z nizkimi dohodki, da so lahko preživeli – ali pa so bili pred odpuščanjem zaradi pandemije – je krajša brezposelnost omogočila, da so si preprosto oddahnili, »za nekatere prejemnike je bil odmor dragocenejši od dodatnega dolarja« (Rhodes 2024). Nekateri prejemniki so denar namenili za nadomestitev izpada dohodka, ki so ga izgubili zaradi prenehanja zaposlitve ali zaradi skrajšanja delovnega časa, namesto tega pa so se odločili, da bodo postali negovalci s polnim delovnim časom ali pa preživeli več časa z družino. Nekaterim je UTD omogočil odločitve v smislu izboljšanja zaposlitvenih možnosti, ki jih morda ne bi bilo brez finančne blazine, ki jo zagotavlja UTD, vključno z nenavadnimi, a pravilnimi kariernimi potezami. Na primer, ena od prejemnic je UTD uporabila za zamenjavo delodajalca, kjer je bolje plačano službo zamenjala za manj plačano na drugem področju, kjer pa je bilo več obetov za napredovanje, čeprav začetna urna postavka ni dosegla tistega, kar bi potrebovala za kritje stroškov, če bi bila brez UTD, zdaj pa je z njim kompenzirala ta izpad dohodka ob tej karierni potezi. Dve leti pozneje se je njeno tveganje izplačalo: bila je na položaju, na katerem je zaslužila skoraj šestmestno številko, kot poročajo raziskovalci (v kvalitativnem delu empirije, dobljene z intervjuji). Denar lahko okrepi avtonomijo ljudi pri sprejemanju odločitev v zvezi z delom, saj poveča možnost, da bodo zaposlitve bolj usklajene z njihovimi individualnimi okoliščinami, cilji in vrednotami. Kljub omenjenemu manjšemu povečanju zaposlenosti prejemnikov UTD glede na kontrolno skupino je raziskava v zvezi s tem zabeležila, da so »prejemniki verjetneje iskali zaposlitev, vendar so bili bolj selektivni« (ibid.).

Finančno stanje

Brezpogojni denarni transfer je v prvih dveh letih eksperimenta statistično pomembno izboljšal (samoocenjeno) finančno stanje prejemnikov, okrepil je njihove sposobnosti soočanja s finančnimi šoki in malenkost se je izboljšala celo dostopnost do kreditov. Verjetno najpomembnejši učinek, ki so ga dokazali raziskovalci, pa je bil v tem, da je UTD omogočil odpravo večjega nihanja pri izdatkih gospodinjestev, predvsem za življenjsko pomembne dobrine, kot so hrana in predmeti za vsakodnevno osebno potrošnjo, ter za poravnavanje računov, kar so izdatki, ki so tudi najbolj občutljivejši za finančne šoke (Dotson 2024b).²⁷ Nikakor ni zanemarljivo, da stroški za hrano in seveda njihovo povečanje v vseh okoljih in v vseh časih najbolj bremenijo revne, saj jim odvzamejo večji delež prihodkov kot srednjim in višjim slojem. Velja pa tudi nasprotno, da vsak dodaten denarni prihodek, tudi če je še tako skromen, pomeni glede na razpoložljivi dohodek posameznika ali gospodinjstva večjo izboljšavo za spodnje sloje kot za vse druge – prav to splošno znano pravilo se presenetljivo pogosto krši v vladnih politikah, izrazito tudi v Sloveniji. Pri nas imamo skoraj vse davke pod

evropskim povprečjem, razen davka na dodano vrednost (in na plače), kjer pa smo nad evropskim povprečjem, kar najbolj prizadene najrevnejše (Pravila in stopnje DDV 2025).

Raziskovalci ugotavljajo, »da brezpogojni denarni transferji vodijo do boljšega finančnega blagostanja, vsaj kratkoročno«, pri čemer avtorica raziskovalnega poročila izrecno opozarja na progresivni učinek UTD na stratifikacijo, saj »gospodinjstva, ki so na začetku eksperimenta imela nižje dohodke, doživljajo večje izboljšave svoje finančne blaginje, in sicer prav zaradi gotovinskih prejemkov, v primerjavi s splošnim povprečjem vseh udeležencev eksperimenta kot tudi v primerjavi z gospodinjstvi, ki imajo višje dohodke« (Dotson 2024b).

Glede na posamezne dimenzije, ki sestavljajo splošno oceno finančnega zdravja (*»financial health«*), so bili največji učinki UTD ugotovljeni pri postavkah, ki so v zvezi s trenutnim finančnim položajem prejemnikov, na primer (ibid.):

- tistih, ki so izjavili, da bi se lahko spopadli z višjimi nepričakovanimi stroški, je bilo pet odstotkov več kot na začetku eksperimenta;
- darilo za poroko, rojstni dan ali kak drug podoben izdatek ne bi obremenil posameznikovih financ (takih trditev je bilo za šest odstotkov več);
- pet odstotkov več je bilo tistih, ki jim je ob koncu meseca ostalo nekaj denarja;
- možnost uživanja v življenju zaradi nadzora nad upravljanjem svojega denarja (takih izjav je bilo za šest odstotkov več);
- delež tistih, ki izjavljajo, da jim zmanjkuje denarja, se je zmanjšal za štiri odstotke;
- za tri odstotke se je zmanjšal tudi delež tistih, ki menijo, da si zaradi finančnega položaja nikoli v življenju ne bodo mogli privoščiti dobrin, ki si jih želijo.

UTD je zgleden primer socialne tehnologije. Pomembno vprašanje pa je, ali si ga lahko privoščimo.

FINANCIRANJE UTD²⁸

Ko dogovorjeno pravico do UTD pretvorimo v individualna denarna nakazila, to lahko naredimo na dva načina: označimo ju kot varianto A in varianto B. Oba načina sta fiskalno vzdržna, le da je varianta B neracionalna in populistična, medtem ko pri varianti A *gotovinsko izplačilo prejmejo samo tisti, ki so dejansko brez denarja*, vsem drugim pa se njihov del sredstev, ki jih že imajo, prekvalificira v UTD. Poglejmo preprost primer, ki je tudi najpogostejši – saj bi veljal za kakih 85 odstotkov državljanek in državljanov – in ki spada v prvi način uveljavljanja omenjene pravice.

Varianta A: predpostavimo, da imam službo in dobim neto izplačilo 1500 evrov, UTD pa bi znašal, na primer, 500²⁹ evrov. Tu sta dve pomembni vprašanji:

- prvič, kaj se zgodi z mojimi prejemki v razmerah, ko imam kot državljan (ob plači) tudi pravico do UTD;
- drugič, ali zaradi UTD res obremenim državne finance?

Odgovor na obe vprašanji je negativen: moj neto dohodek je še vedno enak in država zame ne izplača niti enega dodatnega centa, čeprav mi je priznala pravico do

UTD. Gre namreč za to, da se takrat, ko mi delodajalec plačo nakaže na bančni račun, plača prikaže v dveh zneskih – del se je prekvalificira oz. titulira kot UTD, in ker je ta, smo rekli, 500 evrov,³⁰ dobim z drugim delom plače preostalih 1000 evrov. To pomeni, da se niti pri delodajalcu, v državni blagajni, niti v mojem žepu dejansko ne spremeni nič. Še vedno dobim skupno izplačilo 1500 evrov za opravljeno delo, enako kot prej, ko še nisem imel UTD – a s pomembno razliko. Če ostanem brez službe, ker je šel delodajalec v stečaj (ob tem pa nimam drugih dohodkov ali premoženja), seveda ostanem brez dela plače, vendar še vedno dobim tistih 500 evrov, ker so kvalificirani kot UTD, ki mi ga nihče ne more vzeti; in samo ta znesek mi tokrat financira država iz proračuna. S pravico do UTD je torej enako kot z univerzalno pravico do čiste pitne vode, ki smo jo zapisali v ustavo in jo imamo vsi brezpogojno. A to ne pomeni, da nam bo zdaj država z državno cisterno vsakomesečno dostavila na dom hektolitre pitne vode, saj jo večina izmed nas že ima; nekateri imajo pred hišo celo bazen. Pravico do pitne vode mi država »materializira« v obliki H₂O šele, če zaidem v položaj, da ostanem brez nje, ker mi je, na primer, presahnil vodnjak pred hišo ali pa ne zmorem več plačevati položnic dostavljalcvu vode. Enako je z UTD. Država torej z uvedbo te pravice nikakor ne zaide v dolgove niti v fiskalne prekrške, saj bi šlo za minorno finančno obremenitev z naslova omenjene pravice. Pri UTD v višini 500 evrov bi dodatna obremenitev državne blagajne znašala samo v povezavi s prejemniki socialnih pomoči, in sicer samo za razliko med njihovimi prenizkimi zdajšnjimi transferji in višino UTD (pomnoženo s številom socialnih upravičencev³¹ in številom mesecev v letu). Skratka, zaradi uvedbe UTD državi ne bi bilo treba uvesti nobenega novega davka niti zvišati obstoječih, vse drugo pa ostane enako: sredstev za javne in vse druge storitve bi bilo točno toliko, kot jih država že namenja, in tudi tiste osebe, ki zdaj prejemajo socialne transferje višje od nove meje hipotetičnega UTD (npr. osebe s posebnimi potrebami), bi natančno ista sredstva prejemale še naprej, le da bi se jim znesek pod 500 evrov kvalificiral kot UTD.

Obstaja pa še druga varianta (B) financiranja UTD, ki je načeloma mogoča, čeprav sem jo prej označil za neracionalno. UTD izplačamo samodejno v *gotovini vsem*, tudi tistim, ki niso v revščini, vključno z največjimi bogataši, ob tem pa hkrati najpremožnejši tretjini prejemnikov vzamemo nazaj³² določen mnogokratnik istega zneska, da ga lahko nakažemo manj premožnim. Torej se celo pri tej varianti – računsko gledano – financiranje povsem izide.

SKLEP

Z zgodovinskega vidika je védnost o socialnih učinkih tehnologije že stara, vključno z opozarjanjem nanje. Novo je, da smo jih zanemarjali prav v sodobnosti, ko utegne biti cena take brezbriznosti najvišja. V zvezi s tem je sokratovska čuječnost dober zgled, v katerem je pred 2400 leti navajal negativne posledice pisave. Če bi vsako tehnologijo presojali na »sokratovski« način in s tako natančnostjo, kot ocenjujemo možnosti za njen tržni uspeh, bi bilo manj verjetno, da bi zanemarili morebitne socialne učinke. Čeprav nekatera načelna vprašanja o vrednotni in socialni dimenziji tehnologij ostajajo odprta, je z gotovostjo mogoče trditi, da je UTD dober primer tehnologije, ki je vrednotno določena ter omogoča pomembne in merljive socialne

učinke na okolje (v shemi 1 spada UTD v kvadrant B). To dokazujejo rezultati naj-novejšega družboslovnega eksperimenta, izvedenega v dveh zveznih državah ZDA. Ob prikazanih pozitivnih vplivih na najpomembnejših področjih kakovosti življenja posameznikov in gospodinjstev (zdravje, zaposlenost, finančno stanje), ne smemo pozabiti tudi pozitivnih učinkov, ki jih pa v ameriški kulturi ni mogoče testirati. To so korekcije socialne države v smeri odpravljanja tistih pomanjkljivosti, ki jih s konceptom socialne države nismo zmogli (niti v Evropi) odpraviti:

- sistem socialnih pravic je preveč razdrobljen, nepregleden in na pre nizki ravni;
- temeljne socialne pravice v obstoječem sistemu, vključno z najpomembnejšo (denarna socialna pomoč), nikakor niso brezpogojne;
- socialna država ne odpravlja absolutne revščine, dolgotrajne revščine niti glo-bine revščine;
- spreminjanje socialnih pravic je povsem odvisno od političnega voluntarizma vsakokratnih vladnih koalicij in prioritete znotraj njih, med katerimi odprava revščine nikoli ni med glavnimi cilji;
- socialne pravice nikoli ne dosežejo vseh upravičencev (zaradi neinformiranosti, birokracije, stigme itn.).

Ti popravki so nujni in niso možni brez uvedbe UTD.

OPOMBE

- 1 Izjema so bili samo štirje avtorji med vsemi omenjenimi kritiki, ki jih je skupaj čez 30 (Osborne in Brown 2005, 127–131). Za kritiko teh kriterijev pri nas glejte Rape Žiberna (2017).
- 2 »Al je prva tehnologija v zgodovini, ki lahko ljudem odvzame moč. Silicijevi čipi lahko ustvarijo vohune, ki nikoli ne spiyo, finančnike, ki nikoli ne pozabijo, in despote, ki nikoli ne umrejo« (Harari 2024).
- 3 Pisavo lahko kvalificiramo kot socialno tehnologijo zaradi štirih lastnosti: (1) ker omogoča neomejeno shranjevanje in prenos informacij skozi čas in prostor, (2) ker z akumulacijo kulturnih dosežkov pospeši oblikovanje kolektivnih identitet in s tem predstavlja temeljni tehnološki preboj za razvoj civilizacij, (3) ker s standardizacijo jezika pospeši natančnejše interakcije med različnimi kulturami in s tem olajša konsenz, (4) ker s tem omogoči vse hitrejši in vse kompleksnejši tempo tehnoloških inovacij.
- 4 Tu Sokrat vzpostavi analogijo med pisavo in slikarstvom (ibid., 570):
»Tudi stvaritve (slikarstva) stojijo kot žive; če pa jih kaj vprašaš, zelo vzvišeno molčijo. Enako velja za govore/besede; lahko se ti zdi, da govorijo, kot da bi kaj mislili, če pa jih vprašaš po čem od tega, kar je govoreno, in hočeš to izvedeti, ti označujejo vedno eno in isto. Kadar pa je enkrat zapisan, se vsak govor kotali povsod, enako pri poznavalcih kot pri tistih, za katere sploh ni primeren. In ne ve, s kom je treba govoriti in s kom ne«.
- 5 Verjetno zato, ker je razširjenost informacijsko-komunikacijske tehnologije najvidnejša, hipna, individualizirana in ker je ob nastanku najradikalneje ločila dotlej nujno in samoumevno povezanost idejnega in fizičnega prostora. To se je zgodilo z Morsovim izumom telegrafa v prvi polovici 19. stol. (ki je predstavljal precejšen šok – prvo preneseno sporočilo je bil vzklik »Kaj je ustvaril Bog!« (Senate Stories 2024)).

- 6 Tu ne gre za utilitarizem v smislu filozofije Jeremyja Benthama ali Johna Stuarta Milla, kjer se vrednost nečesa meri po prispevku k največji sreči največjega števila ljudi, kar izboljšuje kakovost življenja vsem. Nasprotno, pri tehnologiji gre za utilitaristično redukcijo v kontekstu moči in komodifikacije, pri čemer štejejo interesi posameznika in privilegiranih skupin na račun vseh drugih.
- 7 »But where danger is, grows / The saving power also« (Heidegger 1977, 310).
- 8 To nima zveze s podobno neoliberalno sintagmo, da je »vsak problem izziv«, kar je čista ideološka floskula, spočeta v menedžerskih krogih od sedemdesetih let prejšnjega stoletja dalje. Njena funkcija je prikrivanje problemov (že na ravni poimenovanja) z namenom, da jih zanikamo. Posledično je vse več profesionalnih okolij, v katerih je izraz problem izpodrinjen s terminom izziv, kot da bi šlo za sinonima. Ta trend povzroča precej škode zlasti v storitvenih dejavnostih (vključno s profesijo socialnega dela); npr. ni vseeno, ali položaj oseb, ki nimajo sredstev za preživljanje, kvalificiramo kot problem ali pa kot izziv.
- 9 Ta misel na koncu Heideggerjevega vprašanja o tehnologiji je formulirana tako (ibid., 317):
 Ker bistvo tehnologije ni nič tehnološkega, mora temeljito razmišljanje o tehnologiji in odločilno soočenje z njo potekati na področju, ki je po eni strani podobno njenemu bistvu, po drugi pa se z njo v bistvu razlikuje. To področje je umetnost – seveda le, če razmišljanje o umetnosti ne zapira oči pred konstelacijo resnice, o kateri sprašujemo. S takim postavljanjem vprašanj pričamo o krizi /.../ A čim bolj premišljujemo bistvo tehnologije, tem bolj skrivnostno postaja bistvo umetnosti. Čim bližje smo nevarnosti, tem bolj sijoče se razkrivajo poti do odrešitve in tem bolj smo nagnjeni k spraševanju. Kajti spraševanje je pobožnost in vdanost (do) mišljenja.
- 10 Gre za delovna mesta, na katerih zaposleni zaradi prenizke mezde ne more vzdrževati niti lastne eksistence in bi taka opravila lahko brez družbene škode ukinili ali pa živo delovno silo nadomestili s stroji. Pri drugi kategoriji znotraj junk jobs, katerih delovni prispevek je za ekonomijo in celotno družbo nepogrešljiv in hkrati niso nadomestljiva z avtomatizacijo, je edina možna rešitev v zvišanju mezde nad ceno existenčnih stroškov, za kar pa je nujen pritisk (npr. podplačana in najzahtevnejša delovna mesta znotraj socialnega skrbstva in zdravstva).
- 11 Glejte v Marx 1979a, 75, in 1979b, 302–304, 310, 317.
- 12 Torej da se sami odločijo, ali jim večjo vrednoto predstavlja delovna kariera in soliden standard ali pa se raje odločijo za brezdelje in življenjski slog, značilen za revščino.
- 13 Novejše poročilo o teh prizadevanjih v Evropi dokazuje, da so organizirana prizadevanja v smeri skrajševanja delovnika še vedno živa, doseženi so celo konkretni rezultati, da sta pomembni oviri (čeprav ne edini) za premike v tej smeri odsotnost delavske participacije in prenizke plače, ki na spodnjem koncu mezdne lestvice delavcem ne omogočajo niti kritja osnovnih potreb celo pri polnem delovnem času (Staunton 2024). Hkrati pa se tudi po sto letih nismo prav nič oddaljili od diagnoze, ki jo je Keynes (1936, 235) opisal s temi besedami: »Velika slabost ekonomije, v kateri živimo, je njen neuspeh pri zagotavljanju polne zaposlenosti in njena arbitrarna in nepravična porazdelitev bogastva in dohodkov.«
- 14 Z Graeberjevo teorijo o bullshit jobs je pozornost znova usmerjena v izvor problema – odtujitev, ki izvira iz ekonomske sfere in se širi na vsa druga, neekonomska področja. Nazadnje je to naredil Erich Fromm prav v obdobju vzpona neoliberalizma v knjigi Imeti ali biti (1976). V njej tudi Fromm izhaja iz Marxovega pojma odtujitve in

v zvezi z vprašanjem, kaj storiti, ponudi točno isti odgovor kot Graeber – univerzalni temeljni dohodek (Fromm 2013, 220, 239–241). Sicer se je Fromm prvič zavzemal za UTD že leta 1955, ko je v knjigi *Zdrava družba* zapisal: »Veliko zla današnjih kapitalističnih in socialističnih družb bi izginilo, če bi uvedli zajamčeni letni dohodek« (ibid., 239).

- 15 Cohen, Rogers in Pribac 2004; Pribac in Korošec 2011; Gerdina 2018; Dragoš 2019b.
- 16 Več o tem v zadnjem razdelku tega prispevka.
- 17 Najnovejši pregled na to temo je v Casselman (2025); za odgovornost ekonomske znanosti v zvezi s tem glejte Galbraith (2025).
- 18 Zaradi pandemije covid-19 je umrlo več Američanov kot pa v vseh ameriških vojnah skupaj (Nye 2025).
- 19 Med najnevarnejšimi posledicami epidemije covid-19, ki je za dalj časa pahnila v socialno izolacijo polovico človeštva (3,9 milijarde ljudi), je najnevarnejša posledica – poleg smrtnih žrtev in psihičnih stisk – nezaupanje znanosti. V najnovejši monografiji na to temo Macedo in Lee (2025) navajata glavne razloge za nezaupanje: »Težke odločitve so bile zasenčene s slogani, kot so 'sledite znanosti'. Koristi in škoda so bile nepravilno porazdeljene. Sprejete politike so bile v veliki meri koristne za sloj lastnikov prenosnih računalnikov (»laptop class«) in pustile t. i. ključne delavce nezaščitene; podaljšano zaprtje šol je najbolj prizadelo najbolj marginalizirane družine. Znanost je postala politizirana in nestrinjanje je bilo potisnjeno na rob.« Avtorja opozarjata, da »v naslednji krizi ne smemo pozabiti na najgloblje vrednote liberalne demokracije: strpnost in odprtost, spoštovanje dokazov in njihovih omejitev, pripravljenost na negotovost in zavezanost h govorjenju celotne resnice« (ibid.).
- 20 Termin socialna »pomoč«, ki je samoumevno v rabi tudi v slovenski zakonodaji in praksi, pišem v narekovajih zato, ker je neustrezen, saj pri tem temeljnem denarnem prejemku socialne države ne gre za nikakršno pomoč, pač pa za pravico, ki jo država posamezniku zagotavlja in jo lahko od nje tudi iztoži, če gre kaj narobe (v primerjavi s pomočjo, ki jo izvajalec prejemniku realizira povsem arbitrarno).
- 21 Več o tem glejte v Lessenich 2015; Dragoš in Leskošek 2016; Bärnthaler in Novy 2025; Pulignano in Domecka 2025; Andor, Skrzypek in Giusto 2025; Galbraith 2025; EU Forward 2025.
- 22 Prvi, tipičen, čeprav ne edini tak nasprotnik UTD pri nas je Milan Zver (2004).
- 23 Za novejši pregled predvidenih učinkov UTD v slovenskem prostoru glejte Cohen, Rogers in Pribac 2004; Pribac in Korošec 2011; Gerdina 2018; Dragoš 2019b; Dragoš 2024.
- 24 Prav tega pogoja ni izpolnjeval nekaj let starejši finski eksperiment z UTD, saj je višina zagotovljenega prejemka s tega naslova bila pod ravno eksistenčnih stroškov. Hkrati pa je glavni cilj finskega eksperimenta, ki ga je uvedla takratna desnosredinska vlada, bil napačno opredeljen, saj se je skušalo z UTD povečati zaposlovanje brezposelnih prejemnikov (vključno z dolgotrajno brezposelnimi in drugimi težje zaposljivimi). Ker v prvem letu (dvoletnega) eksperimenta ta cilj ni bil dosežen, se je razširila napačna informacija, da je eksperiment dokazal neuspešnost UTD, kar ni bilo res. Pri negativnih ocenah so bili namreč zanemarjani drugi učinki na uporabnike, ki so bili pozitivni in postali še izrazitejši ob koncu eksperimenta, torej po dveh letih. To je potrdilo tudi končno in celovito poročilo izvajalca eksperimenta (KELA, Finski zavod za socialno zavarovanje), v katerem so korigirali celo prvotno oceno o zaposlitveni motivaciji UTD po prvem letu. Po koncu dvoletnega obdobja so izmerili, da se je »stopnja zaposlenosti prejemnikov UTD v tem obdobju nekoliko bolj izboljšala kot v kontrolni skupini« (KELA 2020; v zvezi z ocenami eksperimenta po prvem

- letu glejte tudi Bershidsky 2018; Damijan 2018; Drevenšek 2019).
- 25 Ena od presenetljivih ugotovitev je bila prav v zvezi z učinki na druge: ko se denarni položaj prejemnikov izboljša, je to izrazito koristno tudi za druge akterje v socialni mreži prejemnikov UTD (2024b).
 - 26 »Taki razmisleki poudarjajo, kako težko je bistveno spremeniti zdravje, zlasti v razponu treh let. To še posebej velja za ljudi, ki jim je dalj časa omejen dostop do zdravstvene oskrbe in se soočajo z velikimi zdravstvenimi izzivi« (Dotson 2024a).
 - 27 Ta blažilni učinek UTD bo verjetno za Američane kmalu postal še aktualnejši, kot je bil v obdobju raziskave, saj se obeta porast cen zaradi protekcionistične tramponomike v zvezi z napovedanim splošnim dvigom carin; to je že sprožilo »reakcije v ZDA, zlasti v finančnih krogih, ki opozarjajo, da se bo inflacija zvišala« (Gole in Jenko 2025; o tem glejte tudi The Economist 2025; Hemerijck, Bagadirov in Wilson 2025; James 2025; CNN 2025; Jenko 2025). Enak učinek se je pojavil v Združenem kraljestvu po izstopu iz EU. Po zadnjih raziskavah Centra za gospodarsko uspešnost (CEP) na London School of Economics je brexit v dveh letih (do konca leta 2021) v povprečju obremenil gospodinjstva z 210 funtov višjim računom za hrano, kar je potrošnike v Združenem kraljestvu stalo skupaj 5,8 milijarde funtov (6,7 milijarde evrov), saj so višji stroški uvoza povečali inflacijo, kupna moč prebivalcev pa se je zmanjšala, medtem ko so realne plače stagnirale ali padale (Bakker et al. 2022; Toynbee 2025).
 - 28 Ta razdelek je sklepni del že objavljenega članka (v Dragoš 2024, 232, 233), na tem mestu pa pojasnilo o financiranju objavljam ponovno, ker je nujno za razumevanje UTD.
 - 29 Za zdaj bi bila to najnižja mogoča višina UTD, ki bi bila še vedno smiselna, saj bi bil UTD s 500 evri malenkost (za enajst evrov) višji od kratkoročnih življenjskih minimalnih stroškov za eno osebo, medtem ko je zdajšnja denarna socialna pomoč – kljub največjemu povišanju v zadnjih desetletjih – še vedno za štiri evre manjša od omenjenih življenjskih stroškov.
 - 30 Ta del plače, ki zdaj predstavlja UTD, mi delodajalec nakaže na bančni račun posredno, prek države.
 - 31 Ker je teh okoli 33.600 (povprečje v letih 2016–2021), razlika med denarno socialno pomočjo kot temeljnim socialnim transferjem in predpostavljeno višino UTD pa je v letu 2024 15 evrov, bi to zneslo okoli šest milijonov, kar je manj kot nesojena sodna stavba na Litjski v Ljubljani. To je tudi odgovor tistim kritikom UTD, ki ga zavračajo zaradi bojazni porasta inflacije. Omenjenih šest milijonov, ki bi dodatno obremenili državni proračun, nima nobenega inflatornega učinka, saj ne bi bilo treba tiskati novega denarja niti uvesti novih davkov niti povišati že obstoječih.
 - 32 S političnega vidika je to verjetno neizvedljivo, zlasti v slovenskih razmerah. Spomnimo se zapleta z ukinitvijo »prostovoljnega« dopolnilnega zdravstvenega zavarovanja in njegovim prenosom v obvezni zdravstveni prispevek, ki je zdaj ostal nominalno enak za vse, saj ga nočemo spremeniti proporcionalno glede na dohodke (kaj šele progresivno) – tudi to je argument za uvedbo UTD po varianti A namesto po B.

CITIRANA LITERATURA

- Andor, László, Ania Skrzypek, in Hedwig Giusto (ur.). 2025. *Progressive Yearbook 2025*. Bruselj: FEPS – Foundation for European Progressive Studies.
- APFC. 2025. »Who We Are«. Alaska Permanent Fund Corporation. <https://apfc.org/>.
- Baird, Sarah, Jacobus de Hoop, in Berk Özler. 2013. »Income Shocks and Adolescent Mental Health«. *The Journal of Human Resources* 48 (2): 370–403.

- Bakker, Jan David, Nikhil Datta, Richard Davies, in Josh De Lyon. 2022. *Non-Tariff Barriers And Consumer Prices: Evidence From Brexit*. Discussion Paper No. 1888. London: London School of Economics and Political Science.
- Bärnthaler, Richard, in Andreas Novy. 2025. »Corporate Power vs. Social-Ecological Transformation: Why Progress Remains Stalled«. *Social Europe*, 3. 2. 2025. <https://www.socialeurope.eu/corporate-power-vs-social-ecological-transformation-why-progress-remains-stalled>.
- Bershidsky, L. 2018. »Finland's Basic Income Test Wasn't Ambitious Enough«. *Bloomberg*, 26. 4. 2018. https://www.bloomberg.com/view/articles/2018-04-26/finland-s-basic-income-experiment-was-doomed-from-the-start?cmpid%3D=socialflow-twitter-view&cmpid=socialflow-twitter-business&utm_campaign=socialflow-organic&utm_content=business&utm_medium=social&utm_source=twitter.
- Casselmann, Ben. 2025. »Economists Are in the Wilderness. Can They Find a Way Back to Influence?«. *The New York Times*, 10. 1. 2025. <https://www.nytimes.com/2025/01/10/business/economy/economists-politics-trump.html>.
- CNN. 2025. »Top Economist Predicts What Impact Trump's Tariffs Will Have On Americans«. *YouTube*, 3. 2. 2025. https://www.youtube.com/watch?v=I_ZOj4BtH1Y.
- Cohen, Joshua, Joel Rogers in Igor Pribac (ur.). 2004. *Brezplačno kosilo za vse?* Ljubljana: Krtina.
- Damijan, J. P. (2018). »Zakaj je propadel finski eksperiment z univerzalnim temeljnim dohodkom?«. *Damijan blog*, 4. 8. 2018. <https://damijan.org/2018/08/04/zakaj-je-propadel-finski-eksperiment-z-univerzalnim-temeljnim-dohodkom/>.
- De Schutter, Olivier, in Philippe van Parijs. 2024. »Can Universal Basic Income really improve mental health?«. *Social Europe*, 21. 10. 2024. <https://www.socialeurope.eu/can-universal-basic-income-really-improve-mental-health-the-surprising-results-are-in>.
- Derksen, Maarten, in Tjardie Wierenga. 2013. »The history of 'Social Technology', 1898–1930«. *History and Technology* 29 (4): 311–330. <https://doi.org/10.1080/07341512.2013.876247>.
- Dotson, Karina. 2024a. »How Does Unconditional Cash Affect Health?«. *Key Findings: Health*, 21. 7. 2024. <https://www.openresearchlab.org/findings/how-does-unconditional-cash-affect-health-2>.
- Dotson, Karina. 2024b. »How Does Unconditional Cash Affect Financial Health and Well-being?«. *Key Findings: Financial Health*, 8. 4. 2024. <https://www.openresearchlab.org/findings/key-findings-financial-health>.
- Dragoš, Srečo. 2011. »UTD = utrditev temeljev družbe«. V *UTD v Sloveniji, premisleki, stališča, dokumenti*, ur. Igor Pribac in Valerija Korošec, 109–122. Ljubljana: Krtina.
- Dragoš, Srečo. 2019a. »Zakaj je UTD izvedljiv in nujen«. V *Enostavna ideja*, ur. Srečo Dragoš, 26–85. Ljubljana: Fakulteta za socialno delo, UL.
- Dragoš, Srečo (ur.). 2019b. *Enostavna ideja*. Ljubljana: Fakulteta za socialno delo, UL.
- Dragoš, Srečo. 2021. »Bullshit Jobs And Universal Basic Income«. *Anthropological Notebooks* 27 (3): 126–158.
- Dragoš, Srečo. 2024. »Univerzalne storitve ali dohodek – napačna dilema ekosocialne države«. V *Izzivi trajnostnemu razvoju Slovenije*, ur. Ivan Svetlik in Niko Toš, 221–236. Ljubljana: Akademija znanosti za trajnostni razvoj Slovenije.
- Dragoš, Srečo, in Vesna Leskošek. 2016. »Slovenska smer«. V *Oxfamovo poročilo št. 210. Gospodarstvo za 1 %, 87–120*. Ljubljana: Založba /*cf.
- Drevenšek, S. 2019. »Je finski eksperiment z UTD res razočaral?«. *Svet kapitala*, 22. 2. 2019. <https://svetkapitala.delo.si/ikonomija/podkast-je-finski-eksperiment-z-utd-res-razocaral/>.

- EU Forward. 2025. *Shaping European Politics & Policy in the Second Half of the 2020s*. Social Europe in Friedrich-Ebert-Stiftung. <https://www.socialeurope.eu/wp-content/uploads/2025/01/EU-Forward.pdf>.
- Pfaffenberger, Bryan. 1988. »Fetishized Objects and Humanized Nature: Towards an Anthropology of Technology«. *Man* (N.S.) 23: 236-52.
- Fromm, Erich. 2013. *Imeti ali biti*. Ljubljana: Mladinska knjiga.
- Galbraith, James K. 2025. »Mainstream Economics: A Catalogue of Failures«. *Social Europe*, 6. 2. 2025. <https://www.socialeurope.eu/mainstream-economics-a-catalogue-of-failures>.
- Gerdina, Tomaž (ur.). 2018. *Usoda tržnega dna: nuja, želja ali skriti nameni?* Ljubljana: OPRO, zavod za aplikativne študije.
- Gole, Nejc, in Miha Jenko. 2025. »Trump strelja po zaveznikih, Kitajci in Rusi se nasmihaajo«. *Delo*, 4. 2. 2025. <https://www.delo.si/gospodarstvo/novice/carine-bodo-skodile-predvsem-amerriki>.
- Gosar, Andrej. 1933. *Za nov družabni red*, I. zv. Celje: Družba sv. Mohorja.
- Graeber, David. 2013. »On the Phenomenon of Bullshit Jobs: A Work Rant«. *Strike* 2013 (3). <https://www.strike.coop/bullshit-jobs/>.
- Graeber, David. 2019. *Bullshit Jobs: A Theory*. Penguin Books.
- Graeber, David. 2020. »David Graeber: "Je moralno izterjati krivičen dolg? Na tem vprašanju stoji in pade kapitalizem"«. Intervjuvar L. J. Kučič. *Disenz*, 13. 5. 2020. <https://www.disenz.net/david-graeber-je-moralno-izterjati-krivicen-dolg-na-tem-vprasanju-stoji-in-pade-kapitalizem/?fbclid=IwAR1A75bqxJiV0eCWWh-w4I9gvhYQz7qB7B6c6jAzZKTc-gFYQ0xO0XxeH8n90>.
- Harari, Yuval N. 2024. »The Age of AI Is Here (Yuval Noah Harari Introduces ,Nexus')«. *Youtube*, 31. 7. 2024. https://www.youtube.com/watch?v=Ze6diMt1g_s.
- Heidegger, Martin. 1977. »The Question Concerning Technology«. V *Martin Heidgger: Basic Writings*, ur. David Farrell Krell, 287–317. New York: Harper & Row. <https://www2.hawaii.edu/~freeman/courses/phil394/The%20Question%20Concerning%20Technology.pdf>.
- Hemerijck, Anton, Azizjon Bagadirov in Robin Wilson. 2025. »The Capacity to 'Compete'—Rethinking the Welfare State«. V *EU Forward: Shaping European Politics & Policy in the Second Half of the 2020s*, 36–40. Social Europe in Friedrich-Ebert-Stiftung.
- IFR. 2025. »China to Invest 1 Trillion Yuan in Robotics and High-Tech Industries«. *International Federation of Robotics*, 25. 3. 2025. <https://ifr.org/ifr-press-releases/news/china-to-invest-1-trillion-yuan-in-robotics-and-high-tech-industries>.
- Ingold, Tim (ur.). 1997. *Companion Encyclopedia of Anthropology*. London: Routledge.
- James, Harold. 2025. »Trump's Tariff Gamble: Will America's "Golden Age" End in an Economic Nightmare?«. *Social Europe*, 30. 1. 2025. <https://www.socialeurope.eu/trumps-tariff-gamble-will-americas-golden-age-end-in-an-economic-nightmare>.
- Jenko, M. 2025. »Trump s carinsko gorjačo po zaveznikih«. *Delo*, 4. 2. 2025. <https://www.delo.si/mnenja/komentarji/trump-s-carinsko-gorjaco-po-zaveznikih>.
- KELA. 2020. »Basic Income Experiment«. *KELA (Kansaneläkelaitos)*, 16. 9. 2020. <https://tietotarjotin.fi/en/information-package/158070/basic-income-experiment#objectives-and-implementation>.
- Kosec, Jure. 2025. »Veliki razkol na velikem vrhu o umetni inteligenci«. *Delo*, 11. 2. 2025. <https://www.delo.si/novice/svet/j-d-vance-opozoril-eu-pred-oviranjem-razvoja-umetne-intelligence>.
- Kropivšek Leskovar, Rebeka. 2025. »Iskati napredek v konvencionalnosti je kot iskati ribe v puščavi«. Intervjuvarka Petra Kovič. *Delo*, 10. 2. 2025. <https://www.delo.si/nedelo/iskati-napredek-v-konvencionalnosti-je-kot-iskati-ribe-v-puscavi>.
- Krek, Janez E. 1901/1925. »Socializem«. V *Izbrani spisi*, III. zvezek, ur. Odbor društva »Dr. Jan.

- Ev. Krek«. Ljubljana: Jugoslovanska tiskarna.
- Keynes, John M. 1936. *The General Theory of Employment, Interest, and Money*. Zürich: ISN. <http://cas2.umkc.edu/economics/people/facultypages/kregel/courses/econ645/winter2011/generaltheory.pdf>.
- Lenarčič, Blaž. 2014. »Vpliv informacijsko-komunikacijskih tehnologij na velikost socialnih omrežij v kontekstu lokalno-globalnih dinamik«. *Teorija in praksa* 51, posebna številka: 285–308.
- Leskovec, Jure. 2025. »Zgodili se bodo veliki premiki. Poti nazaj ni več«. *Delo*, 8. 2. 2025. <https://www.delo.si/sobotna-priloga/zgodili-se-bodo-veliki-premiki-poti-nazaj-ni-vec>.
- Lessenich, Stephan. 2015. *Ponovno izumljanje socialnega: socialna država v prožnem kapitalizmu*. Ljubljana: Krtina.
- Marx, Karl. 1961. *Kapital*, I. knjiga. Ljubljana: Cankarjeva založba.
- Marx, Karl. 1973. *Kapital*, III. knjiga. Ljubljana: Cankarjeva založba.
- Marx, Karl. 1979a. »Kritika Heglovega državnega prava«. V *Izbrana dela*, I. zvezek, Karl Marx in Friedrich Engels, 57–134. Ljubljana: Cankarjeva založba.
- Marx, Karl. 1979b. »Kritika nacionalne ekonomije (Pariški rokopisi 1844)«. V *Izbrana dela*, I. zvezek, Karl Marx in Friedrich Engels, 245–398. Ljubljana: Cankarjeva založba.
- Macedo, Stephen, in Frances Lee. 2025. In *Covid's Wake: How Our Politics Failed Us*. Princeton, New Jersey: Princeton University Press.
- Nye, Joseph S. ml. 2025. »Ima globalizacija še prihodnost?«. *Delo*, 6. 2. 2025. <https://www.delo.si/mnenja/gostujoce-pero/ima-globalizacija-se-prihodnost>.
- Osborne, Stephen P., in Kerry Brown. 2005. *Managing Change and Innovation in Public Service Organizations*. London, New York: Routledge.
- OZN. 2024. *Extreme Poverty and Human Rights*. Generalna skupščina, 16. 7. 2024. <https://docs.un.org/en/A/79/162>.
- PFD. 2025. *Permanent Fund Dividend*. State of Alaska: Department of Revenue. <https://pfd.alaska.gov/>.
- Platon. 2009. »Fajdros«. Prevod in spremna beseda Gorazd Kocijančič. *Zbrana dela*, II. knjiga, 533–573. Ljubljana: KUD Logos, Celjska Mohorjeva družba.
- Pravila in stopnje DDV. 2025. *Your Europe*, 4. 2. 2025. https://europa.eu/youreurope/business/taxation/vat/vat-rules-rates/index_sl.htm?utm_source=chatgpt.com.
- Prettner, Klaus. 2025. »Could AI Level the Playing Field of Earnings Inequality?«. *International Federation of Robotics*, 19. 2. 2025. <https://ifr.org/post/could-ai-level-the-playing-field-of-earnings-inequality>.
- Pribac, Igor, in Valerija Korošec (ur.). 2011. *UTD v Sloveniji, premisleki, stališča, dokumenti*. Ljubljana: Krtina.
- Pulignano, Valeria, in Markieta Domecka. 2025. *The Politics of Unpaid Labour*. Oxford: Oxford University Press.
- Rape Žiberna, Tamara. 2017. »Konceptualizacija inovacij v socialnem delu«. *Socialno delo* 56 (2): 111–128.
- Reed, Howard, Stewart Lansley, Matthew Johnson, Elliott Johnson, in Kate E. Pickett. 2022. *Tackling Poverty: The Power of a Universal Basic Income*. Basic Income Conversation. https://autonomy.work/wp-content/uploads/2023/05/Full_Paper_Lansley_Tackling_Poverty.pdf.
- Rhodes, Elizabeth. 2024. »How Does Unconditional Cash Affect Employment?«. *Key Findings: Employment*, 21. 7. 2024. <https://www.openresearchlab.org/findings/key-findings-employment-and-income>.
- Rogers, Everett, in Floyd Shoemaker. 1971. *The Communication of Innovation*. New York: Free Press.

- Rus, Veljko. 1990. *Socialna država in družba blaginje*. Ljubljana: Domus.
- Rus, Veljko. 1992. *Med antikomunizmom in postsocializmom*. Ljubljana: Fakulteta za družbene vede.
- Schatzberg, Eric. 2018. *Technology: Critical History of a Concept*. Chicago: The University of Chicago Press. <https://dokumen.pub/technology-critical-history-of-a-concept-9780226584027.html>.
- Senate Stories. 2024. »What Hath God Wrought": Morse's Telegraph in the Capitol«. *United States Senate: Senate Historical Office*. https://www.senate.gov/artandhistory/senate-stories/morses-telegraph-in-the-capitol.htm?utm_source=chatgpt.com.
- Standing, Guy. 2018. *The Precariat: The New Dangerous Class*. Ljubljana: Krtina.
- Stark, Rodney, in William S. Bainbridge. 2007. *Teorija religije*. Ljubljana: Fakulteta za družbene vede.
- Staunton, Bethany. 2024. »Winning Back Our Time«. *Hesa Mag* 29: 8–11. Bruselj: European Trade Union Institute. https://www.etui.org/sites/default/files/2024-12/HesaMag_29_EN_WEB.pdf.
- SURS. 2025. Dan varne rabe interneta. *Statistični urad Republike Slovenije*, 11. 2. 2025. <https://www.stat.si/StatWeb/News/Index/13423>.
- The Economist. 2025. »Canada, China, Mexico and the Art of Retaliation«. *The Economist*, 4. 5. 2025. <https://www.economist.com/the-americas/2025/02/02/canada-china-mexico-and-the-art-of-retaliation>.
- Toynbee, Polly. 2025. »Britain's Brexit Reality Check: Why the Majority Now Want Back In«. *Social Europe*, 15. 1. 2025. <https://www.socialeurope.eu/britains-brexit-reality-check-why-the-majority-now-want-back-in>.
- Uredba EU. 2024. Uredba (EU) 2024/1689 o določitvi harmoniziranih pravil o umetni inteligenci Evropskega parlamenta in Sveta evropske unije, z dne 13. junija 2024. *Uradni list EU*, 12. 7. 2024. https://eur-lex.europa.eu/legal-content/SL/TXT/HTML/?uri=OJ:L_202401689.
- Van Parijs, Philippe. 2004. »Temeljni dohodek za vse«. V *Brezplačno kosilo za vse?*, ur. Joshua Cohen, Joel Rogers in Igor Pribac, 19–37. Ljubljana: Krtina.
- Viršek, D. 2025. »V tovarnah po svetu dela več kot štiri milijone robotov«. *Delo*, 5. 1. 2025. <https://www.delo.si/gospodarstvo/novice/v-tovarnah-po-svetu-dela-vec-kot-stiri-milijone-robotov>.
- WHO. 2021. *Long Working Hours Increasing Deaths From Heart Disease And Stroke*. Ženeva: World Health Organization and the International Labour Organization. <https://www.who.int/news/item/17-05-2021-long-working-hours-increasing-deaths-from-heart-disease-and-stroke-who-ilo>.
- WHO. 2022a. *Launch of the WHO World Mental Health Report: Transforming Mental Health for All Department of Mental Health and Substance Use*. 17. 6. 2022. <https://www.who.int/director-general/speeches/detail/launch-of-the-who-world-mental-health-report--transforming-mental-health-for-all-department-of-mental-health-and-substance-use---17-june-2022>.
- WHO. 2022b. *Mental health*. 17. 6. 2022. <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response>.
- Wikipedia. 2024. »Rote Learning«. *Wikipedia, The Free Encyclopedia*, datum zadnje spremembe 12. 9. 2024). https://en.wikipedia.org/wiki/Rote_learning.
- Wikipedia. 2025. »Wheel«. *Wikipedia, The Free Encyclopedia*, datum zadnje spremembe 6. 1. 2025. <https://en.wikipedia.org/wiki/Wheel#History>.
- Wilkinson, Richard, in Kate Pickett. 2012. *Velika ideja*. Novo mesto: Penca in drugi.
- Zver, Milan. 2004. »Nepravičnost univerzalnega temeljnega dohodka«. V *Brezplačno kosilo za vse?*, ur. Joshua Cohen, Joel Rogers in Igor Pribac, 165–170. Ljubljana: Krtina.

O AVTORICAH IN AVTORJIH

Lana Nastja Anžur je doktorska študentka na Univerzi v Ljubljani. V Laboratoriju za dediščinsko znanost Ljubljana raziskuje vrednotenje kulturne dediščine ter njene učinke na okolje, gospodarstvo in družbo. Posveča se interdisciplinarnemu raziskovanju kulturne in naravne dediščine v Sloveniji in mednarodnem prostoru.

Ivan Bratko je profesor na Fakulteti za računalništvo in informatiko Univerze v Ljubljani in redni član SAZU. Je utemeljitelj umetne inteligence v Sloveniji in sodi med pionirje strojnega učenja v svetu. Njegovi najbolj vidni znanstveni prispevki so na področjih strojnega učenja, logičnega programiranja in kvalitativnega modeliranja, z uporabo v medicini, ekologiji in robotiki.

Srečo Dragoš, sociolog in socialni delavec, je predavatelj na Fakulteti za socialno delo Univerze v Ljubljani. Raziskuje na področjih splošne sociologije in sociologije religije, socialne politike, družbenih neenakosti in socialnega dela. Je avtor in soavtor devetih znanstvenih monografij ter urednik zbornika o univerzalnem temeljnem dohodku.

Ana Farič je študentka magistrskega programa Kognitivna znanost na Univerzi v Ljubljani. V sklopu magistrske naloge se ukvarja s preučevanjem pristranskosti v strojnem učenju. Kot študentka je zaposlena v Laboratoriju za bioinformatiko na Fakulteti za računalništvo in informatiko ter na Inštitutu za slovensko narodopisje in Zgodovinskem inštitutu Milka Kosa ZRC SAZU.

Andreja Jaklič je profesorica mednarodne ekonomije in prodekanja za znanstvenoraziskovalno dejavnost na FDV. Sodeluje v več mednarodnih raziskovalnih mrežah in je predsednica Akademije za mednarodno poslovanje v sekciji Srednje in Vzhodne Evrope (AIB-CEE). Raziskuje strategije in načine internacionalizacije podjetij, njene učinke na inovativnost in razvoj ter s tem povezane ekonomske politike.

Aleksandra Kanjuo Mrčela je profesorica sociologije in predstojnica Centra za preučevanje organizacij in človeških virov na Fakulteti za družbene vede Univerze v Ljubljani. Je zunanja članica Sveta za razvoj SAZU. Pedagoško, raziskovalno in svetovalno deluje na področjih dela, spola in ekonomije, lastniških in organizacijskih sprememb ter ekonomske demokracije.

Lea Košmrlj je magistrica anglistike in germanistike in študentka na magistrskem študiju Kognitivna znanost na Univerzi v Ljubljani. Deluje na stičišču jezikoslovja in kognitivne znanosti; ukvarja se z jezikom pri ljudeh in v umetnih sistemih. Jezik jo zanima predvsem s psiholingvističnega vidika, pa tudi z vidika družbenih razsežnosti jezikovnih pojavov.

Matjaž Maček je politolog, ki se ukvarja s teorijami demokracije, še posebej s teorijo deliberativne demokracije, z volilnimi in strankarskimi sistemi in problematiko kolektivnega delovanja.

Mirjana Nastran Ule je zaslužna profesorica na Fakulteti za družbene vede Univerze v Ljubljani. Je ustanoviteljica Centra za socialno psihologijo. Je izredna članica Slovenske akademije znanosti in umetnosti in članica Sveta za razvoj pri SAZU. Glavna področja raziskovanja so: kritična socialna psihologija, študije vsakdanjega življenja, študije mladine in odrasčanja, ženske študije, predsodki in diskriminacije.

Franc Mali raziskuje na področju družbenih študij znanosti in epistemologije družbenih ved ter predava predmete s tega področja na FDV. Ekspertno deluje v različnih programih Evropske komisije in uredniških odborih mednarodnih STS revij. V zadnjem obdobju se posveča predvsem etičnim vidikom razvoja novih naprednih tehnologij in znanstvenim evalvacijam.

Slavko Splichal je profesor komunikologije in predstojnik Centra za raziskovanje družbenega komuniciranja na Fakulteti za družbene vede Univerze v Ljubljani. Je redni član SAZU in European Academy ter predsednik Znanstvenega sveta ARIS. Njegovi najpomembnejši raziskovalni dosežki segajo na področja teorij javnega mnenja in javnosti, komunikacijskih tehnologij, podružbljanja medijev in politične demokracije.

Matija Strlič je profesor analize kemije na Univerzi v Ljubljani, častni profesor dediščinske znanosti na University College London ter soustanovitelj doktorskega programa dediščinske znanosti v sodelovanju med UCL, Univerzo v Oxfordu in Univerzo v Brightonu. Zanimajo ga interdisciplinarne raziskave na področju razumevanja materialne dediščine, preventivnega konserviranja, vrednotenja in upravljanja z dediščino.

Knjiga Od tržne konkurenčnosti tehnologije k družbeni tehnološki zrelosti preizprašuje prevladujoče enostransko vrednotenje tehnologij, ki je usmerjeno zgolj na gospodarsko učinkovitost – kot je na primer »lestvica pripravljenosti tehnologije«, ki jo pogosto uporabljajo evropske raziskovalne politike. Nasproti temu knjiga spodbuja celostno presojo, ki vključuje etične, družbene in ekološke vidike razvoja, uporabe in vrednotenja tehnologij. Avtorji in avtorice pokažejo, kako tehnologije ne le rešujejo tehnične probleme, ampak tudi oblikujejo družbene odnose, strukture moči in prihodnost človeštva. S kritiko tehnologizacije družbe in prevlade profitnih interesov knjiga predlaga novo paradigmo – družbeno tehnološko zrelost, ki poudarja prizadevanja za odgovornost, pravičnost in trajnostnost tehnološkega razvoja. Z interdisciplinarnimi obravnavami – od umetne inteligence do platformne ekonomije – ponuja praktične smernice za demokratično upravljanje tehnološkega napredka.

Avtorji prispevkov v knjigi so:

Lana Nastja Anžur, Ivan Bratko, Srečo Dragoš, Ana Farič, Andreja Jaklič, Lea Košmrlj, Matjaž Maček, Franc Mali, Aleksandra Kanjua Mrčela, Franci Pušavec, Slavko Splichal, Matija Strlič in Mirjana Nastran Ule

