

Pregledni znanstveni članek/Article (1.02)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 931—944

Besedilo prejeto/Received:08/2024; sprejeto/Accepted:08/2024

UDK/UDC: 004.89: 316.614

DOI: 10.34291/BV2024/04/Kraner

© 2024 Kraner, CC BY 4.0

David Kraner

Vloga moči in uporaba UI v mrežni družbi

The Role of Power and the Use of AI in the Network Society

Povzetek: Komunikacijska orodja so pomemben vir porazdelitve moči. Današnja mrežna družba, ki ima na razpolago mnoge komunikacijske tehnologije, doživlja nenehne komunikacijske zaplete. Eskalacija vplivanja in izkazovanja moči zaradi uporabe umetne inteligence (UI) bi se morala gibati znotraj nadzorovanega območja. To pomeni, da se mora razvoj UI in njena uporabnost podrediti spoštovanju človekovega dostojanstva. Dosedanje pozitivne izkušnje uporabe strojev z UI pri izvajanju opravil v pomoč človeku obetajo pozitivno prihodnost. Vendar pa ne manjka tistih, ki bi UI lahko izkoristili človeku v škodo. Zato je (in bo) potrebno neprestano etično spremljanje razvoja in vključevanja tovrstne tehnologije v družbo ter vzpostavljanje vzvodov za zagotavljanje varnosti, pravičnosti in kontrole.

Ključne besede: oblast, moč, umetna inteligenca, mrežna družba, komunikacija

Abstract: Means of communication are an important source of power distribution. In today's network society, which has most communication technologies at its disposal, there are constant complications in communication. The escalation of influence and the demonstration of power at the expense of the use of artificial intelligence should remain within a controlled zone. This means that the development of artificial intelligence and its use must be subject to respect for human dignity. The positive experiences to date with the use of AI machines in human-supported work promise a positive future. However, there is no shortage of those who can exploit this to the detriment of humans. It is, therefore, necessary and will continue to be necessary to ethically monitor the development and integration of this type of technology into society and to set levers to ensure safety, fairness and control.

Keywords: power, authority, force, artificial intelligence, network society, communication

1. Uvod

Tehnološke, izobraževalne, politične in druge institucije, ki se danes ponašajo z eksponentnim razvojem tehnologij, še posebej umetne inteligence (UI) (Dunnell et al. 2024; Castells 2006), nikakor ne morejo biti ponosne na medosebno komunikacijo na družbeni ravni (Kysar-Moon 2023; Tusalem 2024; Quinn et al. 2024).¹

Zdi se, da »mrežna družba« (Castells 2009), v kateri imamo na razpolago več komunikacijskih pripomočkov kot kdajkoli v zgodovini (Bahovec 2015; Gartner in Krašna 2023; Pasqualetti 2024), ne zna, ne zmore in/ali ne želi kvalitetno komunicirati. Komunikacija je sestavni del moči za doseganje političnih, gospodarskih ali osebnih interesov. Tudi delovanje strojev, ki jih poganja umetna inteligenca, danes predstavlja svojo komunikacijsko vrednost in moč. Etične, antropološke ter teološke izzive uporabe UI lahko razdelimo v tri sklope: a) umetna inteligenca in naši odnosi z drugimi in svetom; b) umetna inteligenca in naš odnos z Bogom oz. presežnim; c) umetna inteligenca in odnos do nas samih (Strahovnik 2023, 849). V tem prispevku se bomo osredotočili le na prvega.

Znanstveniki se danes v večini strinjajo, da človeka ne bi želeli zamenjati s stroji, ki jih poganja UI in da mora biti UI v službi človeka – torej mu mora biti podrejena in v pomoč (Tlili et al. 2023; Jung Youn 2023; Milberg 2024; Floridi 2024; Platovnjak in Svetelj 2023). Vendar ne manjka tistih, ki nevednost ljudi (zlasti mladih) že izkoriščajo za svoje partikularne interese (Pasqualetti 2024; Benanti 2019, 2022, 2020; Floridi 2022). Zato danes mnogi opozarjajo, da se je treba osredotočiti na človeka ter njegovo rabo tehnologij in UI (Pagliani 2004; Frančišek 2024). Predvsem se je treba izogibati lažnimi informacijam (Globokar 2021), religijskemu ekskluzivizmu (Kraner 2021; Jeglič 2023) in izkoriščanju uporabnikov digitalnih tehnologij – predvsem mladih (Vodičar 2020; Breton 2004; Kovač et al. 2020; Spitzer 2021). Pri tem imajo pomembno vlogo vse ustanove, ki se ukvarjajo z etično vzgojo: od religijskih do vzgojno-izobraževalnih (Nežič Glavica 2023, 1054; Šegula 2018; Stegu 2022; Platovnjak 2024).

V našem prispevku sledimo raziskovalnemu vprašanju: Ali je dovolj vedeti, kako digitalno tehnologijo uporabljati, ali pa je treba tudi razumeti, kakšen vpliv ima uporaba tehnologije na dinamiko moči? Kako zaščititi dostojanstvo človeka pred škodljivimi izkoriščevalskimi silami dobičarskih interesov? Zagovarjamo tezo, da v mrežni družbi razvoj in uporabo UI usmerjajo zlasti velike moči (tehnološke in finančne korporacije ter globalna politika). Posamezne manjše ustanove (raziskovalne, izobraževalne, religijske in druge) možnost vplivanja na UI imajo – vendar le z aktivnim spremljanjem njenega razvoja, proučevanjem etične, kazenske in politične odgovornosti razvijalcev tehnologij ter z vzgojno-izobraževalnimi programi.

¹ Prispevek je nastal v okviru raziskovalnega programa P6-0269 »Religija, etika, izobraževanje in izzivi sodobne družbe« in raziskovalnega projekta P6-24684 »Teologija, digitalna kultura in izzivi na človeka osredotočene umetne inteligence«, ki ju sofinancira Agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS), ter v okviru raziskovalnega projekta »Epistemic Identity and Epistemic Virtue: Human Mind and Artificial Intelligence« s podporo fundacije John Templeton (krovni projekt »New Horizons for Science and Religion in Central and Eastern Europe«).

2. Spreminjanje moči in avtoritete v mrežni družbi

Mrežna družba je koncept, ki ga je zasnoval španski sociolog Manuel Castells, ko je analiziral sodobno družbo na družbeni, politični, tehnični in gospodarski ravni. Prepričan je, da je vpliv digitalnih medijev – predvsem interneta – na družbo največji v zadnjih tridesetih letih. Internet je produkt procesa in razvoja človeštva. ‚Mrežno družbo‘ tvorijo digitalna omrežja in družbena struktura (tj. odnosi proizvodnje, potrošnje, reprodukcije, izkušenj in moči), ki tudi temelji na digitalnih omrežjih.

»Mrežna družba je družba, katere družbena struktura se vrti okoli omrežij, ki jih omogočajo digitalne, mikroelektronsko podprte informacijske in komunikacijske tehnologije. Družbene strukture razumem kot organizacijske ureditve ljudi, ki vstopajo v odnose proizvodnje, potrošnje, reprodukcije, izkušenj in moči, ki se v procesih komunikacije izražajo s pomenom, zakodiranim v kulturo.« (Castells 2009, 19) Mrežna družba – tako kot druge družbe – temelji na produkciji in distribuciji vrednosti, pri čemer je z vrednostjo mišljeno to, za kar se odločijo dominantna omrežja. Castells trdi, da je vrednost vedno izraz moči, mrežna družba pa v tem pogledu večjih sprememb ne prinaša (2009).

Z vsako spremembo družbene strukture – na primer zaradi pojava novih komunikacijskih tehnologij – se na novo opredelita tudi dve temeljni spremenljivki vsake družbe: čas in prostor. V mrežni družbi sta ti dve spremenljivki pomembni, saj predstavljata pretočnost prostorov in ne-časovnost časa. S pretočnostjo prostorov imamo v mislih možnost, da s pomočjo digitalnih tehnologij nekaj izvajamo sočasno, ne da bi se dva dogodka v prostoru in času dejansko stikala. Nečasovnost je značilnost digitalnih tehnologij, za katere se zdi, da zaradi zmogljivosti omrežja in hitrosti zanikajo čas in s tem zaporednost dogodkov, kar v praksi imenujemo večopravnost (Lévy 2008).

Piero Pratesi, ki proučuje odnos med komunikacijo in močjo, razlikuje dva načina njune korelacije. Prvi način je povezan z vlogo komuniciranja v okviru oblastne moči (politične, gospodarske, cerkvene in poklicne), drugi pa s komuniciranjem kot močjo. To lahko nadalje delimo na moč, ki izhaja iz izključne posesti informiranja, in moč, ki izhaja iz razpoložljivosti komunikacijskih orodij (Pratesi 2002, 895).

V mrežni družbi je odnos med komunikacijo in močjo treba ponovno proučiti. Velja poudariti, da je ta odnos v mnogih pogledih ambivalenten – z ene strani gre za učinkovito decentralizacijo moči, z druge pa za omrežje, ki ustvarja zelo centralizirane in vseprisotne sisteme nadzora (Castells 2009).

Preden se lotimo proučevanja sprememb moči v ‚mrežnih družbah‘, je treba razlikovati med pojmom ‚moč‘ in ‚oblast‘.

Ameriški sociolog, ki proučuje komunikacije in družbo, Joshua Meyrowitz, razliko med avtoriteto in oblastjo opredeljuje zelo jasno.

»Oblast se nanaša na sposobnost posameznika, da druge (s fizičnimi, ekonomskimi ali drugimi sredstvi) prisili, da ubogajo njegove zahteve. Posameznik ima lahko sredstva oblasti: fizično silo, orožje in denar. Oblast mora

vedno nekdo zastopati. Avtoriteta pa temelji na zmožnosti pridobiti zaupanje in pripravljenost za poslušnost drugih. Če oblast temelji na ustrahovanju, avtoriteta preživi z inspiracijo.« (Meyrowitz 1995, 102)

Ni naključje, da se danes učiteljska in cerkvena avtoriteta po večini zavrača, saj temelji na zaupanju in ne ustrahovanju. Kjer je nekoč avtoriteta temeljila na ustrahovanju, tam se ji danes izogibajo. »Svoboda, ki si jo vsi tako želimo, je zaradi možnosti ponudbe na digitalnem trgu vse bližja – zaradi osvobajanja od starih avtoritet, v duhu demokratizacije pa tudi izvedljiva.« (Vodičar 2023)

Meyrowitz meni, da si mora avtoriteta zaupanje ljudi zaslužiti, vendar ne s fizično silo, orožjem in denarjem, temveč s smiselnim odnosom do te mere, da so ljudje pripravljeni ubogati ali se vesti na določen način, ki je v skladu z navodili odgovorne avtoritete. Da bi dosegli poslušnost, morajo oblasti uporabiti natančne strategije. Ena od teh je nadzor nad informacijami – ni dovolj imeti komunikacijska sredstva, treba je imeti tudi sposobnost razumeti kode določenega medija (Meyrowitz 1995, 263).

V človeški zgodovini je eden izmed bolj opaznih vidikov moč pisanja. V antiki so imeli tisti, ki so znali brati in pisati, poseben položaj moči. Tudi s pojavom tiskarskega stroja, ki je sicer prek knjig in časopisov povzročil večjo distribucijo informacij, se je vprašanje pismenosti množic pojavilo znova. S prihodom elektronskih medijev (radio in televizije) je glede dostopa do informacij nastal nov položaj. Po Meyrowitzu elektronski mediji namreč ne zahtevajo velikih spretnosti kodiranja in dekodiranja. Zato se pričakovanja preusmerijo na en sam organ, ki mu ni treba pretirano blesteti s svojimi pisnimi in argumentacijskimi sposobnostmi, pač pa s sposobnostjo ustvarjanja podobe, ki je za javnost privlačna. Ta sprememba komunikacijske strategije se je začela v šestdesetih letih 20. stoletja – v okviru mešanice družbenih sprememb, ki so jih na eni strani sestavljali protesti proti tradicionalnim institucijam in avtoritetam, kot so družina, šola in država, na drugi strani pa mediji, ki so dali glas navadnemu državljanu, človeku z ulice, za katerega se je pogosto zdelo, da ima za konkretnost in pristnost več občutka kot pa same avtoritete ali celo specialisti. Elektronski mediji torej spreminjajo tok komunikacije in informacij ter tako vplivajo na hierarhijo oblasti. Meyrowitz trdi, da če v kakem podjetju organi vodenja, ki v okviru svoje moči običajno nadzorujejo pomembne informacije, začnejo del teh informacij prejemati od spodaj, se hierarhična struktura sčasoma sesuje in razpade (265–267).

3. Iluzornost demokratične moči interneta

Pri razvoju vsake komunikacijske tehnologije so obstajali skriti interesi za pridobivanje moči in prednosti pred drugimi (Ong 2011). Okrog interneta je bil v začetku zgrajen mit, da gre za tehnologijo, ki bo odprla um in vsakemu od nas pomagala, da bomo bolj informirani in bolj izobraženi. Čeprav večjega dostopa do znanja ne moremo zanikati, velja opozoriti, da prehod od možnosti 'biti informiran' do dejanja 'informirati se' ni samodejen.

Nevroznanstvenik Manfred Spitzer je še posebej zadržan – kritizira digitalne tehnologije in omrežja, ker potencialno vodijo v nazadovanje naših možganov in našega telesa. Digitalne tehnologije v otroški dobi učenja več prinašajo tveganj kot rešitev, saj otroku odvzamejo širšo kognitivno izkušnjo. Tudi za odraslega preobilica informacij predstavlja težave pri izbiri, kritičnem branju virov, ocenjevanju kaj je prav, kaj ima vrednosti in kaj ne (Spitzer 2017).

Tudi internet je orodje, ki je v rokah tistih, ki imajo družbeno moč. To dokazuje njegov razvoj in razcvet tehnologij, ki prinašajo bogastvo le nekaterim – kot bomo videli v nadaljevanju. Vse se je začelo s projektom ARPANET (1969), ki naj bi zagotovil ameriško vojaško prevlado v času hladne vojne. Na srečo je vojska zasnovo zaupala inženirjem, tehnikom in učiteljem, ki so bili bolj navdušeni nad znanstvenimi ideali kot nad vojaškimi strategijami. Projekt se je tako razvil zahvaljujoč sinergiji štirih družbenih akterjev: tehnokrat, hekerske skupnosti, virtualnih skupnosti in poslovnega sveta (Castells 2006, 45).

Z vstopom na področje poslovnega sveta je bilo uporabo mreže po eni strani mogoče razširiti zunaj elitnih krogov ameriških univerzitetnih kampusov in virtualnih skupnosti, po drugi strani pa je prišlo do tega, da je nadzor nad prihodnostjo omrežja prevzelo podjetje. Rodilo se je novo gospodarstvo, ki mu je zakon mreže naklonjen. Kakor trdi Castells, je »uresničitev skritega potenciala ob zmožnosti pretvorbe moči uma v proizvodnjo denarja postala temelj poslovne kulture Silicijeve doline in internetne industrije na splošno« (63).

V devetdesetih letih je bil komunizem močno poražen, kar je kot simbolni dogodek zaznamoval padec berlinskega zidu. Kapitalizem je dejansko zmagal, a tudi razkril svoj pravi obraz sistema, ki deluje ne glede na vrsto režima, ki ga sprejme – saj glavni cilj ni skupno dobro ljudi, temveč dobiček ekonomskih elit. Podjetja iz Silicijeve doline in denarni tokovi z Wall Streeta so postali referenčne točke za podjetniške in finančne inovacije, internet pa je deloval kot krmilo za osvajanje trgov po vsem svetu (Parisier 2012).

Web 2.0 (digitalna socialna omrežja) se je najprej odzval na poslovno usmerjene potrebe. Začela se je hitra in navdušujoča revolucija – Web 2.0 je namreč resnično omogočil ustvarjanje konvergentne kulture. Njegove zmožnosti interakcije in sodelovanja z uporabo vseh oblik sporočanja (zvok, slika, video, besedilo, grafika itd.) so v potencialnega tvorca informacij in kulture spremenili katerega koli uporabnika. Leto 2002 je zaznamovala eksplozija blogov. V javnosti je zaokrožila misel, da je prišel čas novega obdobja demokratizacije svobode misli in izražanja. Toda ob vsej svežini je bilo za informacije značilno tudi veliko zatohlosti, ki jo je ustvarilo kopičenje objav, katerih cilj je bil bolj razplamteti jezo kot pa spodbujati resne študije ali razprave. Blogosfera je tako takoj razkrila svojo ambivalentnost: naklonjenost alternativnim informacijam in povzročanje veliko nekoristnega hrupa. S pojavom družabnih omrežij, zlasti s Facebookom, se zgodovina zopet ponovi. Vedno se začne z nečim družbeno koristnim (tkanje odnosov, iskanje starih sošolcev itd.), potem pa zmogljiva aplikacija kmalu preide v družbeno negativno uporabo. Mreža sicer res daje prednost politični aktivnosti državljanov, družabno-

sti, prijateljstvu, solidarnosti, dobroteljnemu združenju ipd., vendar tudi poslovnim in brezvestnim zadevam, populizmu, rasizmu, fundamentalizmu, prostituciji, terorizmu itd. Pravzaprav je mreža skoraj v trenutku ustvarila nov digitalni svet z vsemi njegovimi krepostmi in slabostmi (Lanier 2010).

Iz začetnega mita o tehnologiji, ki naj bi pomagala vsakemu od nas, da bomo bolj informirani in izobraženi, je prišlo do prevlade elit. Ne le podrejanje tehnologije gospodarskim interesom – takšen trend lahko vodi v uničenje človeštva. O tem je prepričan tudi Pasqualetti, ki med drugim trdi, da je pohlep po moči največja nevarnost. Ogroža človek, ki tehnologijo zlorabi (Pasqualetti 2018, 63–64).

Naša resničnost je v zadnjih petdesetih letih vse bolj digitalna resničnost, sestavljena iz ničel in enic. Italijanski filozof – profesor na Univerzi v Oxfordu – Luciano Floridi spremembo človekovega odnosa do tehnologije zelo nazorno predstavlja na treh nivojih. Zanj je digitalna tehnologija tretji red. Ni tehnologija, ki bi bila postavljena neposredno med človeka in naravo – kot npr. sekira (prvi red). Tudi ni tehnologija, ki bi bila postavljena med človeka in drugo tehnologijo kot npr. motor (drugi red). Digitalna tehnologija stoji med eno in drugo tehnologijo – kot npr. računalniški sistem, ki nadzoruje robota pri barvanju avtomobila (tretji red). Floridi UI pripisuje izredno moč in opozarja na človekovo nesposobnost obvladovanja tehnologij: »Zaradi avtonomne računalniške moči digitalnega sistema nad procesom morda niti nimamo nadzora /.../. Po drugi strani pa digitalno ni le nekaj, kar neko resničnost izboljša ali poveča, temveč nekaj, kar jo radikalno preoblikuje, saj ustvarja nova okolja, ki jih naseljujemo, in nove oblike delovanja, s katerimi smo v interakciji.« (Floridi 2022, 31)

Danes smo uporabniki UI vsi. Od tega, da nosimo pri sebi pametni telefon, do uporabe vseh vrst pametnih naprav in aplikacij. Vendar – kdo se resno zaveda, katere gospodarske sile in skupine moči so v ozadju, ko priklopljen na svetovni splet prenaša podatke o sebi in drugih?

Tehnologija živi tam, kjer je moč in bogastvo. Neenakost med bogatimi in revnimi je jasen kazalec, da se o blaginji za vse in pravični družbi ne razmišlja. Agencija Oxford Committee for Famine Relief (OXFAM) že leta obsoja postopno koncentracijo bogastva v rokah peščice do te mere, »da ima samo osem ljudi toliko premoženja kot najrevnejša polovica človeštva. Od gospodarske rasti imajo koristi najbogatejši, medtem ko preostali del družbe, zlasti revni, trpi. Do tega skrajnega, nevzdržnega in nepravilnega položaja nas je pripeljala sama narava naših gospodarstev in temeljna načela naših gospodarskih sistemov.« (OXFAM 2017)

Danes gospodarske moči nimajo na voljo le določenega jezika, temveč vse jezike, ki jih je človek izumil skozi stoletja. Multinacionalke imajo moč, da svoje stranke dosežejo s svetlobno hitrostjo tako, da vlagajo v trženje, personalizirano oglaševanje, privlačno zunanjo podobo in embalažo izdelkov. Njihova moč je v vplivanju in prepričevanju (Pasqualetti 2018).

Učinkovitost prepričevanja in vplivanja sta v družbi dobro vidna: bogatenje zgolj peščice ljudi, eksponentna rast tehnoloških podjetij, raztresenost oz. problem koncentracije, zasvojenost s spletom itd. Obstaja torej resnična potreba, da bi se

človek iz te zasvojenosti prebudil – in svojo avtonomno zavest osvobodil. To danes ugotavlja tudi strokovnjak za interakcijo mladih z elektronskimi mediji, David Buckingham. Ne poudarja le digitalnega opismenjevanja mladih in starih z vključevanjem medijev v pedagoški proces (Buckingham 2003), temveč tudi potrebo po aktivni vključenosti ljudi v družbene in politične spremembe (2020). Primer takega vključevanja je brazilski pedagog Paulo Freire, ki je znan po pedagogiki zatiranih. Takole pravi: »Pedagogika zatiranih kot humanistična in osvobajajoča pedagogika bo imela dva različna momenta. Prvi, v katerem se zatirani sveta zatiranja zavedo in se v praksi zavežejo, da ga bodo preoblikovali; drugi, v katerem po preoblikovanju zatiralske stvarnosti pedagogika ni več pedagogika zatiranih – postane pedagogika ljudi, ki so v procesu trajne osvoboditve.« (Freire 2002, 40)

Čeprav Freire delovanja velikih omrežij in oblastnih moči, ki smo jih omenili zgoraj (političnih, gospodarskih, poklicnih, cerkvenih), ni doživel in analiziral, bi njegovo kritiko lahko mirno parafrazirali in usmerili na današnji čas. Tudi danes gre za zatiralsko stvarnost – za človeka, ki je ali na strani ,zatiranega' ali pa ,zatira'.

4. Moč UI in vloga človeka

Nenehen pretok podatkov med fizično in virtualno resničnostjo, ki ga omogoča svetovni splet z računalniki in pametnimi napravami, polni bazen zgodovinskih (avtentičnih), sintetičnih in hibridnih podatkov. Za Floridija so zgodovinski podatki tisti iz resničnega življenja ali vhodni podatki. Podatke, ki jih v celoti ustvari UI, imenuje sintetični (Floridi 2022, 70). Hibridni podatki so medtem produkt zgodovinskih in sintetičnih podatkov. Ravno zaradi anonimizacije so zgodovinski podatki najbolj ,osiromašeni' (72). Eden glavnih virov moči UI so podatki. Središčno zanimanje je, kako za reševanje zapletenih problemov ustvariti hibridne ali sintetične podatke. Bolj kot bo to mogoče, uspešnejša bo UI – je prepričan Floridi (84).

Ker je danes vse v povezavi z UI pod vplivom digitalne revolucije, se je tudi znanstvena misel o robotih preoblikovala, spremenila oz. prilagodila in izkrivila. O tem so razmišljali Reis in soavtorji, ki ugotavljajo, da stroji zaradi svoje mehanске in analitične narave pri opravljanju standardiziranih nalog ljudi praviloma prekašajo. Takoj nato dodajajo, da vendarle ne dosejajo takšne tehnološke usposobljenosti, da bi lahko človeka kar nadomestili (Reis et al. 2020).

Tudi vsakodnevne teme, o katerih pišejo mediji in so v pogovorih med ljudmi na neki način središčna tema, se dotikajo UI. Righetti in Carradore sta analizirala skoraj 3000 novic spletnih medijev v Italiji v letih 2014–2018. Ugotavljata, da se je v tem času število novic, povezanih z roboti, podvojilo in da so najpogostejše teme ,zaposlovanje' robotov v družbenih sektorjih (npr. negovalci v bolnišnicah, glasbeni vodje na festivalih in razstavah), fizične grožnje robotov ljudem, strah pred izgubo delovnih mest, vojna z roboti, sožitje med ljudmi in roboti (Righetti in Carradore 2019).

Kako uporabiti moč UI v korist človeku? V čem je potrebna regulacija moči in interesov, da se bo UI razvijala v dobro človeštva?

4.1 Vidna moč UI in njene reprezentacije

Preden preidemo na koristi UI, nas zanima, kaj imamo ljudje, ko govorimo UI, v mislih. Da bi lahko obravnavali najbolj točne miselne sisteme, ki imajo posebno logiko in jezik, strukturo razvoja (bodisi vrednot bodisi mišljenja), način posredovanja, ki jim je lasten, se bomo omejili na socialne reprezentacije. V njih ne vidimo zgolj ,mnenja o', ,slike od' in ,drže pred', ampak ,teorije', ,vede' *sui generis*, zasnovane za odkrivanje realnosti – in tudi njeno obvladovanje (Moscovici 2005). Močan vpliv, ki ga ima UI na človeka, je povzročil nastanek novih socialnih reprezentacij (SR) v povezavi s stroji (Brondi et al. 2021; Dunnell et al. 2024).

Brondi s soavtorji, ki je raziskoval, kakšne SR imajo mladi Italijani o strojih (vzorec obsega 422 mladih med 18 in 31 letom), ugotavlja, da gre za štiri glavne SR: 1. fizično utelešenje, 2. področje uporabnosti, 3. mesto v družbi, 4. idejna ali fantazijska vrednost robotov (Brondi et al. 2021).

V povezavi z ,utelešenjem' se pojavlja filozofsko vprašanje, kako razlikovati med roboti in človekom. Pogosto namreč robote vidimo preoblečene v lutke, podobne človeku. Poleg tega določene naloge opravijo bistveno bolje in učinkoviteje kot človek. Filozof Žalec je prepričan, da je lastnost samostojnega reševanja zapletenih problemov tista, ki je roboti niso sposobni. Ravno ta lastnost je tista, ki dela človeka kvalitativno drugačnega od stroja. »Ta kvalitativna razlika je vsekakor drugačna od tiste, ki jo označujejo pojmi, kot so psihično, fenomenalna zavest, zavedanje, živo, splošnost, intuicija itd.« (Žalec 2024, 823) Avtorica Centa Strahovnik, ki je analizirala oblikovanje in dožemanje identitete pogovornih sistemov umetne inteligence, je prepričana, da se »pri interakciji s temi sistemi lahko spremeni naša lastna identiteta. /.../ Sistemi umetne inteligence predstavljajo novo resničnost, s katero se človek srečuje /.../. Poleg tega, kako dojemamo njihovo zunanjo ,podo-bo' (spol, glas, barva kože ipd.), bo vse večji pomen dobivalo tudi vprašanje, kako dojemamo druge vidike njihove identitete.« (Centa Strahovnik 2023, 862–863)

Pri uporabnosti robotov se polje odpira v neskončnost, saj zaobjema vse kotičke človekovega življenja. Uporabniki trdijo, da mora biti UI človeku zgolj v pomoč, ne sme pa ga zamenjati. Izpostavimo le primer odprtokodnega sistema, ki je s pomočjo UI sposoben ustvarjati govor in sliko: spletna platforma „OpenOpenAI“, ki oblikuje besedila, govor in ustvarja videoposnetke za prikaz izvedbe govora. Dunnell s soavtorji ugotavlja, da želijo uporabniki UI za pomoč pri svojem delu uporabiti zlasti tam, kjer je boljša od človeške inteligence (Dunnell et al. 2024).

UI v družbi povzroča velike spremembe. Spremenil se je način učenja, igranja, dela, ljubezni, sovraštva, izbire, odločanja, proizvodnje, prodaje, nakupov, potrošnje, oglaševanja, uživanja, skrbi in nege, druženja, komuniciranja itd. S pametnimi napravami in pomočjo svetovnega spleta svoje življenje ovijamo v digitalna okolja. „Infosfera“ je okolje, ki omogoča uspešno delovanje robotov (Floridi 2022, 56). Vse več ljudi je v omrežja povezanih kar naprej – tudi takrat, ko ne brskajo po spletu. Zato Floridi pravi, da »vedno več ljudi živi *onlife*, tako na spletu kot brez njega, in *intolife*, tako digitalno kot analogno« (25).

UI je mogočen vir idej in fantazij. Nastala je množica mitov: UI bo rešila in premagala vse; povzročila bo usodno katastrofo; uničila bo milijone delovnih mest in nadomestila odvetnike, zdravnike, novinarje, raziskovalce, voznike tovornjakov, taksiste ter na koncu ljudi nadvladala (85–86).

Vendar so roboti kljub temu vsakdanjim praksam večine državljanov še vedno tuji. Zato so Brondi in soavtorji prepričani, da si posamezniki in družba srečanja z roboti predstavljajo, konstruirajo in pripovedujejo bolj na idejni ravni kot iz dejanskih izkušenj. S tem so povezani tudi njihovi strahovi in pričakovanja (Brondi et al. 2021).

Večina ljudi ima – kljub vsenavzočnosti razprav in diskusij o UI – zavestnih izkušenj z uporabo UI vendarle zelo malo ali nič. Hkrati pa strokovnjaki opozarjajo, da je moč, ki jo UI izkorišča, vse bolj problematična, saj se dotika različnih področij vsakdanjega življenja (zdravje, delo, izobraževanje, vzgoja itd.). Brondi s soavtorji opozarja, da so prakse in dejanske interakcije z roboti nepovezane (vsak v svojem okolju oz. kontekstualne) (2021).

Končno so ljudje prisiljeni, da zavzamejo stališče za ali proti, čeprav so pogosto zelo slabo informirani. Škodljivih tveganj, povezanih UI, ne smemo spregledati.

4.2 Etična načela, nevidna moč človeka

Zavedanje, da tista, ki bi vzbujala strah zaradi svoje potencialne moči, ni tehnologija, temveč človek, ki jo uporablja, je osrednjega pomena. Weidinger in soavtorji so z analizo velikih jezikovnih modelov izdelali sistematični pregled tveganj in jih razporedili v šest področij: 1) diskriminacija, sovražni govor in izključevanje; 2) informacijska tveganja oz. tveganja zlorabe informacij; 3) tveganja napačnih ali lažnih informacij; 4) zlonamerna uporaba sistemov; 5) tveganje za nastanek škodljivih posledic na podlagi interakcije med človekom in računalnikom; 6) tveganje za okoljsko in socialno-ekonomsko škodo (Weidinger et al. 2022).

Tveganja veljajo za vse stroje, ki jih poganja UI, in jih bomo (strnjena v dve točki) predstavili. Omejili se bomo na področji, kjer je današnji človek najbolj ranljiv: diskriminacija in zloraba informacij (Kraner 2023; 2024).

4.2.1 Problem diskriminacije

Nikoli v zgodovini noben medij (od knjige do elektronskih medijev) odprtega dostopa do vseh svojih privilegijev moči ni dopuščal vsem, ampak je vedno prihajalo do regulacije (Meyrowitz 1995). Lahko bi rekli, da je do diskriminacije pri dostopnosti medijev in uporabi njihove moči v določeni meri prihajalo vedno.² Zato je tudi danes utopično zahtevati, da bi UI bila dostopna vsem in z enakimi pooblastili. Prva ovira je finančna – aplikacije UI imajo svojo ceno. Z nastopom interneta in digitalizacije dobrin (dokumentov, izdelkov, odnosov itd.), je zrasla želja po zastojnosti ,vsega', dostopnosti ,povsod', vplivanja na vse. Pojavila se je tudi utopična ideja o kiberdemokraciji (Lévy 2008) – a kmalu se je izkazalo, da je to iluzija (Lanier 2010).

² Nekateri religiologi menijo, da mediji povzročajo diskriminacijo določenih verskih skupin, ko objavljajo zgolj pristranske novice – in da so glavni raznašalci islamofobije (Jeglič 2024, 302).

Pri tem Floridi opozarja na problem rasne, spolne in verske diskriminacije. Predlaga zahtevo po algoritemski pravičnosti na štirih področjih: področju antiklasifikacije (zaščita kategorij rase, spola), enakega razvrščanja (skupna merila za napovedovanje uspešnosti), kalibriranja algoritma (algoritem deluje do vseh zaščitnih skupin enako) in statistične enakosti (enak način izračuna srednje verjetnosti za vse člane zaščitnih skupin) (Floridi 2022, 163–164).

4.2.2 Problem zlorabe informacij in škodovanje

Ne želimo opozarjati zgolj na problem lažnivih informacij, ki so v digitalni dobi vsakodnevni pojav (Lah 2018), ampak tudi na netočnosti zaradi poenostavljanja. Na to opozarja Benanti s primerom zemljevida, ki bi bil natančna kopija resničnosti. Če bi bil teoretično izvedljiv, bi bil dejansko neuporaben, saj bi vseboval preveč podatkov. Podobno UI ne more upoštevati vseh podatkov, ker jih je preveč. Kot se odločimo za pot na osnovi zemljevida, tako se z uporabo UI odločimo za izvedbo operacije na osnovi nekaterih podatkov. In prav zato Benanti trdi, da etično vprašanje za UI ni prehodno vprašanje – ki bi ga na začetku upoštevali, potem pa ne več. Etično vprašanje je za UI konstitutivno vprašanje. Kot so zemljevidi redukcija resničnosti in zato sprejeti s poenostavitvami, tako tudi UI s pomočjo podatkov iz senzorjev, ki opravljajo podobne redukcije, po svoji naravi stvari prikaže zmotno. Benanti si zastavlja etično vprašanje: kako zagotoviti, da ta zmotljivost ne bo škodovala vrednotam, ki veljajo za ključne in temeljne? (Benanti 2022, 124)

Benanti poudarja, da mora etika v svojem minimalnem razumevanju zagotavljati sistem, ki je čim bolj varen pred napakami. To pomeni, da morajo biti vse nevarnosti preverljive na

»obeh polih sintetičnega spoznavnega razmerja: človek in stroj. Na človeški strani to pomeni, da algoritemska merila, na katerih stroj in njegovo delovanje temelji, postanejo pregledna in človeškemu spoznavanju dostopna. Na strani stroja pa to pomeni razvoj algoritmov, ki po eni strani ne povzročajo situacij, v katerih poseg sistema človeškega odločevalca zavaja, po drugi pa znajo človeško etično instanco ,prebrati' ali ,razumeti' tako, da jo prevedejo in izvedejo v ,strojnem jeziku'.« (127)

Za Benantija je najpomembnejše človeško dostojanstvo in podrejanje stroja človeku. Pri tem izpostavlja štiri komunikacijske lastnosti, ki da bi jih pametni stroji morali imeti – intuicija, iznajdljivost, prilagodljivost in podrejenost. Stroji so zanj sprejemljivi le, če bodo znali z ljudmi sodelovati v skladu z določenimi smernicami. To pomeni »ne le, da človeku ne bodo škodovali, temveč bodo znali tudi zaščititi njegovo dostojanstvo in iznajdljivost, ne da bi pri tem omrtvili njegovo notranjo vrednost« (128).

Umetni stroji bi morali pridobiti nekakšno ,umetno ponižnost', ki stroju omogoča, da ugotovi, kaj je pri njegovem delu prioriteta. Če gre za sesalec prahu, bi ta moral razumeti, ali naj poseša čim več prahu ali pa naj v tem trenutku poseša le malo prahu in se k tej funkciji vrne pozneje, ker so se pri ljudeh, ki so v tem tre-

nutku v sobi, pojavile druge prioritete. »Operativna prioriteta ni v algoritmu, temveč v osebi – tam je sedež dostojanstva. V mešanem okolju [med stroji in ljudmi] je oseba in njena edinstvena vrednost tista, ki vzpostavlja in določa prednostne naloge: robot je tisti, ki sodeluje s človekom, ne pa da bi človek pomagal stroju.« (131)

Spoštovanje človekovega dostojanstva želi človeka v temelju zaščititi pred manipulacijo, zavarovati njegovo zasebnost, ga pozvati k soglašanju, zagotoviti uporabniku pravičnost in z njim semantično prijazno sodelovati (Floridi 2022, 226; Globokar 2023, 866–869).

Za Benantija je odnosni vidik med človekom in strojem bistvenega pomena. »Ta odnos ne more potekati le v okviru racionalnega odnosa, temveč mora prevzeti – in biti sposoben obdelati – tudi celotno čustveno vprašanje, katerega nosilec je človeška stran.« (Benanti 2022, 126)

Za morebitno škodo, ki bi jo stroj človeku lahko povzročil, morajo biti predvidene sankcije. Za strojem je človek, ki mora za svoja dejanja odgovarjati. Floridi poudarja, da »je umetni agent za kaznivo dejanje lahko odgovoren vzročno, moralno pa je lahko odgovoren le človeški agent« (Floridi 2022, 185).

5. Zaključek

Pri predstavitvi kompleksne dinamike med močjo, UI in komunikacijskimi tehnologijami v sodobni mrežni družbi smo izpostavili več ključnih ugotovitev in odprtih vprašanj, ki zahtevajo nadaljnje razmisleke in ukrepe. Ključne ugotovitve so, da tehnologije danes oblikujejo načine, kako se moč in avtoriteta v družbi izražata. UI ne preoblikuje le komunikacije, temveč ostaja v službi močnih korporacij, kar vodi v centralizacijo moči. Ustvarjanje kumulativnih podatkov ter sintetičnih in hibridnih oblik informacij posega v temeljne vidike človekove identitete in vrednot. UI se pogosto opredeljuje kot podpora človeku – a obstaja nevarnost, da postane sredstvo za manipulacijo. V prispevku zato izpostavljamo potrebo po etičnem nadzoru nad razvojem UI in varovanju človekovega dostojanstva. Radikalna preobrazba, ki jo digitalizacija omogoča, kliče po skupni odgovornosti tako tehnoloških kot družbenih akterjev. UI lahko vodi do diskriminacije in širjenja lažnih informacij. Brez ustrezne regulacije obstaja nevarnost, da bo UI neenakosti v družbi še okrepila in škodovala ranljivim skupinam.

Odprta ostajajo naslednja vprašanja: Kako uravnovežiti moč UI in njeno etično odgovornost? Kako je mogoče regulirati razvoj UI, da bomo varni pred njenimi negativnimi učinki, ki vključujejo diskriminacijo in manipulacijo? Kdo nosi odgovornost? Kakšna je pravna in etična odgovornost posameznikov, podjetij in vlad pri uporabi in razvoju UI? Kako zaščititi ranljive skupine; kako lahko zagotovimo, da UI ne bo škodovala tistim, ki so na robu družbe? Kako se boriti proti širjenju dezinformacij? In kako lahko v svetu, kjer prevladujejo UI in digitalne tehnologije, ponovno vzpostavimo zaupanje med ljudmi in institucijami?

Pomen proučevanja etičnih, socialnih in pravnih vidikov tehnologije odraža nujnost, da se ljudje in institucije aktivno vključijo v oblikovanje prihodnosti, ki bo spodbujala pravičnost, odgovornost in človeško dostojanstvo. Ta odprta vprašanja ostajajo za oblikovanje trajnostne, vključujoče mrežne družbe ključna. Če parafraziramo McLuhanovo izjavo, lahko poudarimo, da tehnologije niso niti pozitivne niti negativne – pa tudi nevtralne niso (McLuhan in Quentin 1967). Tudi, če se strinjamo z Bregantom, ki trdi, da način uporabe UI daje vtis, da je meja med človeško in umetno inteligenco izbrisana (Bregant 2020, 17), moramo vendarle priznati, da je za vse, kar počnejo stroji, odgovoren človek. Skratka, sistemi UI nas ne smejo nadomestiti – čeprav nam lahko na mnogih področjih pomagajo (Dorobantu idr. 2022).

Kratika

UI – Umetna inteligenca.

Reference

- Bahovec, Igor.** 2015. Christianity in confrontation with individualism and crisis of Western culture: person, community, dialog, reflexivity, and relationship ethics. *Bogoslovni vestnik* 75, št. 2:335–346.
- Benanti, Paolo.** 2019. *Le Macchine sapienti: Intelligenze artificiali e decisioni umane*. Bologna: Marietti 1820.
- — —. 2020. *Digital age: teoria del cambio d'epoca: persona, famiglia e società*. Rim: San Paolo.
- — —. 2022. *Human in the loop: Decisioni umane e intelligenze artificiali*. Firenze: Mondadori Università.
- Bregant, Janez.** 2020. Umetna inteligenca v praksi. *Analiza* 24:5–20.
- Breton, Philippe.** 2004. *Elogio della parola: Il potere della parola contro la parola del potere*. Milano: Elèuthera.
- Brondi, Sonia, Monica Pivetti, Silvia Di Battista in Mauro Sarrica.** 2021. What do we expect from robots? Social representations, attitudes and evaluations of robots in daily life. *Technology in Society* 66:101663. <https://doi.org/10.1016/j.techsoc.2021.101663>
- Buckingham, David.** 2003. *Media education: Alfabetizzazione, apprendimento e cultura contemporanea*. Trento: Ericson.
- — —. 2020. *Un manifesto per la media education*. Milano: Mondadori Università.
- Castells, Manuel.** 2006. *Galassia Internet*. Milano: Feltrinelli.
- — —. 2009. *Comunicazione e potere*. Milano: Università Bocconi Editore.
- Centa Strahovnik, Mateja.** 2023. Identiteta in pogovorni sistemi umetne inteligence. *Bogoslovni vestnik* 83, št. 4:853–864. <https://doi.org/10.34291/bv2023/04/centa>
- Dunnell, Kevin, Gauri Agarwal, Pat Pataranuta-porn in Andrew Lippman.** 2024. AI-Generated Media for Exploring Alternate Realities. Prispevek predstavljen na znanstvenem zborovanju Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems. Honolulu, Association for Computing Machinery. Razmnoženina.
- Floridi, Luciano.** 2022. *Etica dell'intelligenza artificiale: Sviluppo, opportunità, sfide*. Milano: Raffaello Cortina Editore.
- — —. 2024. The Hardware Turn in the Digital Discourse: An Analysis, Explanation, and Potential Risk. *Philosophy & Technology* 37, št. 39:37–39.
- Frančiček.** 2024. Umetna inteligenca in mir. Poslanica papeža Frančiška za 57. svetovni dan miru. Katoliška cerkev. 1. 1. <https://katoliska-cerkev.si/poslanica-svetega-oceta-franciska-za-57-svetovni-dan-miru-umetna-inteligenca-in-mir>
- Freire, Paulo.** 2002. *La pedagogia degli oppressi: Edizione italiana a cura di Linda Bimbi*. Torino: EGA Editore.
- Gartner, Smiljana, in Marjan Krašna.** 2023. Etika umetne inteligence v izobraževanju. *Revija za elementarno izobraževanje* 16:221–237.

<https://doi.org/10.18690/rei.16.2.2846>

Globokar, Roman. 2021. *Vzgojni izzivi šole v digitalni dobi*. Ljubljana: Teološka fakulteta.

---. 2023. Razprava o dostojanstveni smrti. *Bogoslovni vestnik* 83, št. 4:865–882. <https://doi.org/10.34291/bv2023/04/globokar>

Jeglič, Urška. 2023. Širjenje salafizma na Balkanu s pomočjo digitalne tehnologije. *Bogoslovni vestnik* 83, št. 4:945–955. <https://doi.org/10.34291/bv2023/04/jeglic>

---. 2024. An Analysis of the Belief in Religious Truths Between Muslims and Roman Catholics in Slovenia and the Question of Identity. *Nova Prisutnost* 22, št. 2:301–315. <https://doi.org/10.31192/np.22.2.4>

Jung Youn, Soo. 2023. Test design and validity evidence of interactive speaking assessment in the era of emerging technologies. *Language Testing* 40, št. 1:54–60. <https://doi.org/10.1177/02655322221126606>

Kovač, Miha, Mojca Kovač Šebart, Tadej Vidmar in Jasna Mažgon. 2020. Bralne navade v Sloveniji in bralne navade študentov pedagoških programov. *Tekstleec* 71, št. 2:40–57.

Kraner, David. 2021. Intolerance in the media and representations of the Catholic Church. *Crkva u svijetu* 56, št. 4:731–766. <https://doi.org/10.34075/cs.56.4.8>

Kraner, David. 2023. Conflict as a Challenge for Personal Growth in the Educational Team. *Seminare* 44, št. 1:69–81. <https://doi.org/10.21852/sem.2023.1.07>

---. 2024. Negative News about the Catholic Church as a Challenge for a More Active and Critical Reading of the News. *Seminare* 45, št. 1:19–31. <https://doi.org/10.21852/sem.1822>

Kysar-Moon, Ashleigh. 2023. The power of family? Family social capital and the risk of externalising behaviours among youth with multiple childhood adversities. *Sociology of Health & Illness* 45, št. 4:810–836. <https://doi.org/10.1111/1467-9566.13621>

Lah, Peter. 2018. Laži, preoblečne v novice, da ljudje dobimo novice. Academia. 15. 2. www.academia.edu/35927042/Fake_news_Laži_preoblečene_v_novice_ki_povzročijo_da_ljudje_dobimo_novice (pridobljeno 28. 4. 2024).

Lanier, Jaron. 2010. *Tu non sei un gadget*. Milano: Mondadori.

Lévy, Pierre. 2008. *Cyberdemocrazia: Saggio di filosofia politica*. Milano: Mimesis.

McLuhan, Marshall, in Fiore Quentin. 1967. *The medium is the message*. New York: Random House.

Meyrowitz, Joshua. 1995. *Oltre il senso del luogo: L'impatto dei media elettronici sul comportamento sociale*. Bologna: Baskerville.

Milberg, Tanya. 2024. The future of learning: How AI is revolutionizing education 4.0. We forum. 28. 4. <https://www.weforum.org/agenda/2024/04/future-learning-ai-revolutionizing-education-4-0/> (pridobljeno 28. 4. 2024).

Moscovici, Serge. 2005. *Le rappresentazioni sociali*. Bologna: Il Mulino.

Nežič Glavica, Iva. 2023. 'Pedagogika 3.0' ali oris edukacije v digitalni dobi. *Bogoslovni vestnik* 83, št. 4:1049–1060. <https://doi.org/10.34291/bv2023/04/nezic>

Ong, Walter J. 2011. *Oralità e scrittura: Le tecnologie della parola*. Bologna: Il Mulino.

OXFAM. 2017. Un'economia per il 99%: È giunto il momento di costruire un'economia umana a vantaggio di tutti, non solo di pochi privilegiati. https://www.oxfamitalia.org/wp-content/uploads/2017/01/Rapporto-Uneconomia-per-il-99-percento_gennaio-2017.pdf (pridobljeno 1. 2. 2017).

Pagliani, Gabriella. 2004. *Il mestiere della guerra: Dai mecenari ai manager della sicurezza*. Milano: Franco Angeli.

Parisier, Eli. 2012. *Il filtro*. Milano: Il Saggiatore.

Pasqualetti, Fabio. 2018. Giovani, fede e vocazione cristiana nell'era della tecnologia digitale. V: Vito Orlando, ur. *I giovani, la fede e il discernimento vocazionale. Contributi di riflessione sulla realtà dei giovani di oggi*, 231–256. Rim: LAS.

---. 2024. *Ecologia, digitale, spiritualità: Un rapporto complesso e problematico*. Rim: Castelvecchi.

Platovnjak, Ivan. 2024. Human Vulnerability, Spiritual Abuse, and Its Prevention. *Obnovljeni život* 79:199–212. <https://doi.org/10.31337/oz.79.2.6>

Platovnjak, Ivan, in Tone Svetelj. 2023. Technology as the elixir of immortality: resurgent philosophical and spiritual enigma of human imprisonment. *Bogoslovni vestnik* 83, št. 4:973–984. <https://doi.org/10.34291/bv2023/04/platovnjak>

Pratesi, Piero. 2002. Potere e comunicazione. V: Franco Lever, Pier Cesare Rivoltella in Adriano Zancchi, ur. *La comunicazione: Il dizionario di scienze e tecniche*. Rim: LAS.

Quinn, Jason, T. David Mason, Mustafa Kirisci in Sally Sharif. 2024. Proto-Insurgency, Repression-Driven Contagion, and Civil War Onset. *Defence & Peace Economics* 35, št. 5:601–621. <https://doi.org/10.1080/10242694.2023.2183346>

Reis, João, Nuno Melão, Juliana Salvadorinho, Barbara Soares in Ana Rosete. 2020. Service robots in the hospitality industry: The case of Henn-na hotel, Japan. *Technology in Society* 63:101423.

- Righetti, Nicola, in Marco Carradore.** 2019. From Robots to Social Robots: Trends, Representation and Facebook Engagement of Robot-Related News Stories Published by Italian Online News Media. *Italian Sociological Review* 9, št. 3:433–454.
- Spitzer, Manfred.** 2017. *Digitalna demenca: Kako spravljamo sebe in svoje otroke ob pamet.* Celovec: Mohorjeva.
- —. 2021. *Epidemija pametnih telefonov: Nevarnosti za zdravje, izobraževanje in družbo.* Vetrinj: Mohorjeva založba; Mladinska knjiga.
- Stegu, Tadej.** 2022. Kerygmatic catechesis and resonance [Kerigmatična kateheza in resonanca]. V: Bojan Žalec in Stjepan Štivič, ur. *Program in povzetki*, 37. [s.l.]: [s.n.].
- Strahovnik, Vojko.** 2023. Etični in teološki izzivi velikih jezikovnih modelov. *Bogoslovni vestnik* 83, št. 4:839–852. <https://doi.org/10.34291/bv2023/04/strahovnik>
- Šegula, Andrej.** 2018. Evangelizacija kot ena od poti do enega Boga v sodobnih pastoralnih smernicah. *Bogoslovni vestnik* 78, št. 2:451–459.
- Tlili, Ahmed, Boulus Shehata, Michael Agyemang Adarkwah in Aras Bozkurt.** 2023. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments* 10, št. 1:15.
- Tusalem, Rollin F.** 2024. Citizen Perceptions of Crime and Their Effect on Support for Illiberal Democratic Rule: Evidence from the Philippines. *Asian Survey* 64, št. 3:514–543. <https://doi.org/10.1525/as.2023.2114587>
- Vodičar, Janez.** 2020. Kritika sekularizacije: tradicija kot pot do trdožive prihodnosti. *Bogoslovni vestnik* 80, št. 2:253–266. <https://doi.org/10.34291/bv2020/02/vodicar>
- —. 2023. Avtoriteta na področju vzgoje in učenja v digitalni dobi. *Bogoslovni vestnik* 83, št. 4:1035–1047. <https://doi.org/10.34291/bv2023/04/vodicar>
- Weidinger, Laura, Jonathan Uesato, Maribeth Rauh, Conor Griffin, Po-sen Huang, John Mellor, Amelia Glaese, Myra Cheng, Borja Balle, Atoosa Kasirzadeh, Courtney Biles, Sasha Brown, Zac Kenton, Will Hawkins, Tom Stepleton, Abeba Birhane, Lisa Anne Hendricks, Laura Rimell, William Isaac, Julia Haas, Sean Legassick, Geoffrey Irving in Iason Gabriel.** 2022. Taxonomy of Risks posed by Language Models. Prispevek predstavljen na znanstvenem zborovanju Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency. Seoul, Association for Computing Machinery.
- Žalec, Bojan.** 2024. Ali je umetna inteligenca inteligenca v pravem pomenu besede? *Bogoslovni vestnik* 83, št. 4:813–823. <https://doi.org/10.34291/bv2023/04/zalec>