

*Mateja Centa Strahovnik***Etika zmožnosti, priporočilni in pogovorni sistemi UI in doseganje dobrega življenja***Ethics of Capabilities, AI Recommendation and Conversation Systems and Attaining a Good Life*

Povzetek: Prispevek raziskuje povezavo med etiko zmožnosti in umetno inteligenco (UI). Uvodoma kot okvir raziskovanja predstavlja etiko zmožnosti in izpostavlja hitro rast tehnologije UI ter njeno vpeljavo v vsakdanje življenje. Prispevek se osredotoča na dva vidika: kako UI prispeva k razvoju posameznikovih temeljnih zmožnosti in kako lahko etiko zmožnosti uporabimo za razvoj etike UI – vključno s presojo tveganj in vplivov UI ter predlogi ukrepov za njihovo omejitev.

Ključne besede: etika zmožnosti, umetna inteligenca (UI), priporočilni sistemi, pogovorni sistemi, identiteta, dobro življenje

Abstract: This paper explores the link between capability ethics and artificial intelligence (AI). It introduces capability ethics as a framework for research and highlights the rapid growth of AI technology and its integration into everyday life. The paper focuses on two aspects: how AI contributes to the development of an individual's core capabilities and how capability ethics can be used to develop an ethics of AI, including an assessment of the risks and impacts of AI and suggestions for measures to limit or mitigate them.

Keywords: capability approach, artificial intelligence (AI), recommendation systems, chatbots, identity, good life

1. Uvod

Okvir našega raziskovanja predstavlja etika zmožnosti (angl. *capability approach*).¹⁰ Eno izmed temeljnih določil sodobnega sveta je hiter porast v razvoju tehnologije

¹⁰ Prispevek je nastal v okviru raziskovalnega programa „Presečišče kreposti, izkustva in digitalne kulture: Etični in teološki uvidi“, ki ga sofinancira Univerza v Ljubljani, raziskovalnega programa „P6-0269: Religija, etika, edukacija in izzivi sodobne družbe“, ki ga sofinancira Javna agencija za raziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS), ter raziskovalnega projekta „Epistemic Identity and Epistemic Virtue: Human Mind and Artificial Intelligence“, ki ga sofinancira John Templeton Foundation in the Ian Ramsey Centre for Science and Religion (University of Oxford).

umetne inteligence (UI) ter pospešena vpeljava njene rabe, ki je postala del naših življenj. Povezanost pristopa etike zmožnosti in UI je mogoče obravnavati z več vidikov. Eden izmed temeljnih vidikov je, kako lahko uporaba UI prispeva k razvoju in uveljavljanju posameznikovih temeljnih zmožnosti. Drugi vidik je, kako lahko etiko zmožnosti uporabimo za razvoj etike UI – vključno s presojo tveganj in vplivov, ki jih implementacija sistemov UI prinaša, ter predlogi ukrepov, ki bi ta tveganja omejili. Oba omenjena vidika sta za doseganje dobrega življenja posameznika in širše družbe nedvomno pomembna. V prispevku se bomo osredotočili na prvi vidik, še posebej na specifično vprašanje odnosa med priporočilnimi in pogovornimi sistemi UI ter na eno izmed desetih temeljnih zmožnosti.

V prvem delu prispevka najprej opredeljujemo pojem etike zmožnosti ter predstavljamo njen okvir. Nato se osredotočamo na zmožnost, ki je določena s krovno oznako „Čuti, domišljija in mišljenje“, pri čemer izpostavljamo element obveščnosti in pomen lastne izbire. Sledi prikaz vpliva priporočilnih sistemov UI na ta dva elementa – predvsem z izpostavljanjem prednosti in slabosti teh sistemov pri omenjenih elementih. V nadaljevanju razpravo navezujemo na pojem spoznavne in moralne identitete (posebej tudi v odnosu do etike zmožnosti in doseganja dobrega življenja) in skušamo pokazati: (a) kako lahko priporočilni sistemi UI na ti dve vrsti identitete vplivajo; (b) kako se obe omenjeni identiteti lahko (so)vzpostavljata v odnosu med posameznikom ter pogovornimi sistemi UI.

2. Etika zmožnosti

Etika zmožnosti je pristop h globalni etiki z načrtovanjem razvoja človeških skupnosti v luči temeljnih vrednot. V ospredju tega pristopa sta najosnovnejši človeški določili – dostojanstvo in samospoštovanje. Kot poseben etični pristop ga je razvila Martha Nussbaum. Izpostavlja seznam desetih temeljnih zmožnosti, ki se medsebojno dopolnjujejo in so podlaga za dostojno, dobro življenje posameznika. Te zmožnosti so svoboščine in priložnosti za doseganje dobrega življenja, ki se vzpostavljajo na podlagi sestava osebnih zmožnosti, ter možnosti, ki naj bi bile vzpostavljene v dejanskem političnem, družbenem in gospodarskem okolju, kjer posameznik živi (Nussbaum 2011, 20). Omenjeni pristop razlikuje »tri vrste zmožnosti: osnovne (te so bolj ali manj vrojene ali prirojene), notranje (to so razvita stanja osebe) in kombinirane. Slednje so notranje zmožnosti, kombinirane z ustreznimi zunanjimi pogoji za opravljanje funkcije. Državljeni represivnih režimov npr. imajo notranje, toda ne kombiniranih zmožnosti za uresničevanje svoje zmožnosti, da govorijo v skladu s svojo vestjo. Spisek Nussbaumove je spisek kombiniranih zmožnosti. Ljudem je treba torej zagotoviti vsaj minimalne in po možnosti čim boljše pogoje za uresničevanje vseh treh vrst njihovih zmožnosti.« (Žalec 2008, 85) Gre torej za dejanja in stanja, ki jih posameznik lahko doseže, če se tako odloči (Robeyns et al. 2023) – vsaka izmed teh zmožnosti pa mora biti omogočena do določene ravni oz. stopnje, da o dostojnemu življenju sploh lahko govorimo. Pristop se osredotoča na to, kaj posamezniki dejansko lahko počnejo in kaj lahko

postanejo. V središču zanimanja so zmožnosti, ki ljudem omogočajo živeti na način, ki ga cenijo oz. razumejo kot dobro življenje. To je večrazsežnostni in pluralni okvir za razumevanje človeškega življenja in blaginje, ki se pogosto uporablja tako pri razvojni etiki kot tudi za razreševanje bolj določnih etičnih in družbenih vprašanj.

Martha Nussbaum izpostavlja deset temeljnih zmožnosti, vsako pa zaznamuje splošna oznaka, ki jo je treba vsebinsko oz. po obsegu še nadalje opisati in s tem zamejiti: 1. življenje; 2. telesno zdravje; 3. telesna integriteta; 4. čuti, domišljija in mišljenje; 5. čustvovanje; 6. praktični razum; 7. pripadnost; 8. druge vrste bitij; 9. igra; 10. nadzor nad svojim okoljem (materialnim in političnim) (2011, 33–34). V nadaljevanju se bomo osredotočili na četrto zmožnost, to je na zmožnost, ki je povezana s čutenjem, domišljijo in mišljenjem.

2.1 Čuti, domišljija in mišljenje

Martha Nussbaum četrto zmožnost „Čuti, domišljija in mišljenje“ opredeljuje takole:

»Biti zmožen uporabljati čute, si predstavljati oz. zamišljati, razmišljati in sklepati – in delati te stvari na ‚resnično človeški‘ način, to je na način, ki je obveščen in kultiviran z ustrezno izobrazbo, zajema pa najmanj pismenost ter osnovno matematično in znanstveno vednost. Sposobnost uporabe domišljije in misli v povezavi z doživljanjem in ustvarjanjem del in dogodkov po lastni verski, literarni, glasbeni idr. izbiri. Biti sposoben uporabljati svoj um na načine, ki so zaščiteni z jamstvi svobode izražanja v zvezi s političnim in umetniškim govorom ter svobodo veroizpovedi in verske prakse. Biti zmožen imeti prijetne izkušnje in se izogniti nekoristni bolečini.« (Nussbaum 2011, 33)

Zmožnost sestavlja več med seboj povezanih zmožnosti (čutov, zamišljanja, razmišljanja in sklepanja). Razvoj in uveljavljanje vseh teh zmožnosti predpostavlja osebo, ki je ustrezno obveščena, kultivirana in izobražena. Zmožnost zamišljanja in razmišljanja se uveljavlja oz. izkazuje svobodno in izraža na različne načine, ki izhajajo iz lastnih prepričanj posameznika. Prav tako mora imeti posameznik zmožnost uporabljati svoje čute na način, ki je prijeten oz. nima za posledico nezaželenega trpljenja. Zmožnost vključuje tudi svobodo misli in govora, vse to pa se tesno povezuje z oblikovano ali zaželeno identiteto posameznika. En vidik, ki ga Nussbaum v okviru te zmožnosti izpostavlja, je npr. pomen religijske identitete: »ker je vera ljudem tako pomembna in je globok vir njihove identitete, to predstavlja močan argument za spoštovanje teh oseb, saj dopolnjuje premisleke o notranjih vrednotah. Če ljudem govorimo, da končnega smisla življenja ne morejo opredeliti na svoj način – tudi, če smo prepričani, da imamo prav in da njihov način ni pravi –, jim kot osebam ne izkazujemo polnega spoštovanja.« (2000, 180) Spoštovanje do drugega torej zahteva priznavanje njegove pravice do lastnih prepričanj in načinov iskanja smisla življenja. To pa zopet predpostavlja, da se posamezniku za uresničevanje zmožnosti omogoča – gledano širše – čim boljše pogo-

je za uspevanje in uveljavljanje v okolju in skupnosti, kjer živi – in obenem, da so v tej skupnosti upoštevani tudi posameznikovi osebni vidiki (in ožji odnosi). Primer enega teh pogojev predstavljajo vrednote, kreposti in veščine, ki so pomembne z vidika vzgoje. Ključne vrednote za ustrezno pripravo na življenje, zlasti v digitalni dobi, so samospoštovanje, kritično mišljenje, ustvarjalnost, skupnost (Globokar 2021, 98–109; Vodičar 2023).

V nadaljevanju pri tej zmožnosti izpostavljamo dva elementa, in sicer obveščenosti in lastne izbire – to pa predvsem zaradi naslednjega koraka, kjer nas bo zanimal vpliv UI na ta dva vidika zmožnosti. Izhodiščno se kaže, da imajo lahko različni sistemi UI na ta vidika zelo pomemben vpliv. Oglejmo si najprej, kakšno vlogo ta elementa zmožnosti imata – in zakaj sta pomembna za četrto zmožnost (in za celoten okvir etike zmožnosti).

Biti *obveščen* v okviru te četrte zmožnosti pomeni, da posameznik razume, da ima svobodo in priložnosti, da na podlagi svojih izkustev, domišljije in mišljenja svoje zmožnosti razvije in živi izpolnjujoče življenje. Etika zmožnosti poudarja zagotavljanje potrebnih virov in pogojev, da se lahko vsak posameznik smiselno odloča – na podlagi jasnega razumevanja razpoložljivih možnosti, verjetnih posledic in ustreznih informacij – in doseže, kar je za njegovo dobrobit in osebne cilje najboljši. Podobno velja tudi za *lastno izbiro*, kjer pa gre predvsem za poudarek na zavedanju, da ima posameznik sam svobodo in možnost sprejemanja odločitev, ki vplivajo na njegovo življenje. To vključuje možnost izbire med različnimi možnostmi glede na osebne vrednote, želje in znanja. Etika zmožnosti tako poudarja opolnomočenje posameznika, da svoje življenje oblikuje na podlagi lastnih izbir in uresniči svoj potencial. Lahko torej povzamemo, da če je posameznik napačno informiran, je tudi njegova slika o dobrem življenju popačena – obenem pa se to prenese na polje njegovih izbir ter razumevanja lastnega izkustva in sveta okoli sebe. Informiranost, na podlagi katere presoja in se odloča, je za uspevanje in doseganje dobrega življenja ključna.

2.2 Vpliv sistemov UI na obveščenost in lastno izbiro

V nadaljevanju nas bo zanimalo, kako na ta elementa vplivajo sistemi UI. Ker je to vprašanje zelo široko in se lahko nanaša na precejšen obseg tehnologij UI ter področij njihove uporabe, se bomo v tem razdelku bolj specifično osredotočili na tako imenovane priporočilne sisteme UI in kako ti vplivajo na informiranost in izbiro. Za te sisteme se namreč vsaj predhodno zdi, da je njihov vpliv znaten.

Priporočilni sistemi so algoritmi UI, povezani s strojnim učenjem, ki za delovanje uporabljajo velike podatke (angl. *big data*), na podlagi katerih uporabniku predlagajo ali priporočajo določeno vsebino/izbiro. Priporočilni sistemi UI uporabnikom z učenjem vzorcev njihovega prejšnjega vedenja zagotavljajo prilagojeno storitveno podporo in predvidevajo, kakšne so njihove trenutne preference, npr. glede določenih izdelkov (Zhang et al. 2021). Takšni sistemi (npr. Recombee, Argoid, Optimizely) se uporabljajo v zdravstvu, zabavni industriji, finančnem svetovanju, družbenih medijih, oglaševanju, potovalni industriji itd. Primeri teh sis-

temov oz. sistemov, ki priporočilne sisteme vključujejo, so mdr. različne spletne trgovine (Amazon), portali z video in glasbenimi vsebinami (YouTube, Netflix, Spotify) ter najrazličnejši drugi portali (LinkedIn, Airbnb, Uber, Google Maps). Ti sistemi so sicer osnovani na ogromnem obsegu podatkov, vendar brez tega, da bi uporabnik podal svoje začetne preference in jih potem s svojim delovanjem na neki način usmerjal, ne delujejo oz. niso posebej uporabni. Ko torej uporabnik naredi ključni korak in vpiše svojo zahtevo, priporočilni sistemi z vidika uporabnika (npr. kupca) preobremenitev z informacijami zmanjšujejo – zagotavljajo bolj personalizirane informacije o izdelkih, razvrščajo artikle, napovedujejo uporabniške nastavitve, zagotavljajo ocene in povzemajo mnenje skupnosti (Knotzer 2008, 10). Izpostaviti je treba, da priporočilni sistemi uporabniku pravzaprav ‚pomagajo‘ oz. bolje rečeno na odločitev uporabnika vplivajo, in sicer na način, ki se začne že preden do samega procesa odločanja – npr. za določen izdelek – sploh pride. V uporabniku sprožijo potrebo po ‚nečem‘ (19–20). Sistemi algoritmov na podlagi prejšnjih uporabnikovih delovanj na spletu (so)oblikujejo spletno okolje, v katerem se uporabnik nahaja in tako med ponujenimi stvarmi že nehote izbira. Hkrati je v te priporočilne sisteme UI lahko vpetih več drugih parametrov, ki jih lahko določa oz. manipulira upravljelec tega sistema.

Oglejmo si primer delovanja takšnega sistema z vidika informiranosti. Gre za izmišljen in poenostavljen primer, a tak, ki odraža temeljne podmene delovanja teh sistemov. Doma imamo mačko in psa. Mačka zboli in po internetu začnemo poizvedovati o simptomih in boleznih, ki prizadenejo mačke. Po določenem času se ob robu zaslona začnejo pojavljati oglasi in ponudbe za ležišča za mačke, hrano za mačke, blogi in članki o mačkah – npr. o tem, kako jih je mogoče naučiti opo-
našati določene zvoke. Sčasoma, ker vsebinam, povezanim z mačkami, namenjamo vedno več pozornosti, se izoblikuje neki vzporedni informacijski svet, kjer se zdi, da imamo doma zgolj mačko in da je to edini predmet našega zanimanja. S tega vidika je psu seveda namenjena minimalna pozornost. Enako velja npr. za izhodiščno povsem neodločeno iskanje informacij o tem, ali globalno segrevanje s svojim ravnanjem res povzroča človek. Če med ponujenimi viri informacij najprej izberemo nekaj takšnih, ki so skeptični, bomo na podlagi priporočilnih sistemov sčasoma seznanjeni zgolj z viri informacij, ki se s to držo strinjajo. Vse to seveda vpliva tudi na našo možnost izbire – ta je povezana s ponujenimi možnostmi, ki jih takšen priporočilni sistem oblikuje. In seveda njegov cilj ni naša informiranost ali pa naše dobro življenje, temveč sledi ciljem, določenim ob zasnovi in implementaciji, npr. podaljševanje časa na spletu, obseg nakupov oz. dobiček za spletno trgovino, hitra prodaja izdelkov, ki so na zalogi v prevelikem obsegu ali bodo kmalu zastareli, učenje o naših navadah, da bi lahko naučeno uporabil v prihodnje ipd. Spet si to lahko ponazorimo s primerom. V spletni trgovini želimo kupiti določen izdelek. Sistem po iskanju takoj ponudi seznam desetih ali več izdelkov, med katerimi običajno izbiramo. Problem pa je, da ti glede na naše iskanje niso nujno najboljši ali najustreznejši – ampak tisti, za katere so oglaševalci plačali, da so prikazani na prvih mestih.

Iz vsega povedanega lahko povzamemo, da priporočilni sistemi, ko za izbiro

nečesa iz množice podatkov potrebujemo hitro odločitev, sicer delujejo dobro – a pomembno se je zavedati, da s tem ne dobimo nujno najbolj relevantnih ali nepopačenih informacij. Hkrati pa ne glede na to, kakšne pridobljene, izbrane, urejene in prikazane informacije so, na uporabnika na neki način vplivajo (Adomavicius et al. 2013)

Vsak naš dotik tipkovnice oz. vnos v spletnem okolju pusti v tem virtualnem svetu odtis, ki se na neki način odziva nazaj. Izpostavimo lahko npr. naslednji opis. »Poizvedujte in morda boste našli. Vendar preživite na spletu dovolj časa in tisto, po čemer poizvedujete, vas bo našlo ravno v trenutku, ko boste to potrebovali.« (Caulfield 2020) To je moto, ki na podlagi sistemov UI trenutno prežema svetovni splet. Priporočilni sistemi imajo lahko tako na vidik informiranosti in izbire izjemno velik ter hkrati potencialno tudi škodljiv vpliv, kar pa posledično določa oz. vpliva na temeljno človekovo zmožnost *čutenja, domišljije in mišljenja*. Vpliv imajo tudi na sklop preostalih temeljnih zmožnosti in s tem na posameznikovo življenje – še posebej takrat, ko sistemi UI ne ostajajo omejeni le na svetovni splet, za katerega se zdi, da ima glede njegove rabe posameznik vsaj načeloma vseeno določen vpliv oz. nadzor. A ti sistemi presegajo meje zaslona in preko različnih predmetov posegajo tudi v druge dele življenjskega okolja – vključno z intimnimi odnosi posameznika.

Če se vrnemo na izhodišča etike zmožnosti, vidimo, da sta dobra informiranost in lastna izbira na poti uresničevanja dobrega življenja ključna – in sistemi UI kot del naše nove resničnosti (oz. določneje priporočilni sistemi UI) pri tem deloma lahko pomagajo. Vsekakor pa se je treba zavedati, da je njihov vpliv lahko v znatni meri tudi škodljiv.

3. Okvir odnosa med človekom in sistemi umetne inteligence ter identiteta

V prejšnjem razdelku smo obravnavali vpliv priporočilnih sistemov UI na nekatere vidike temeljnih zmožnosti. V ospredju je bila perspektiva obveščenosti posameznika in njen vpliv na posamezna dejanja ali izbiri. V tem razdelku bomo to gledišče razširili ter vlogo in vpliv UI raziskali v okviru tesnejšega odnosa med človekom in UI, natančneje pogovornimi roboti (angl. *chatbots*). O takšnem odnosu lahko govorimo zaradi vse večje razširjenosti oz. vseprisotnosti uporabe UI v okviru bolj kompleksnih sodelovalnih okolij ali pa v okviru tega, da ljudje sisteme UI (npr. pogovorne robote) sami začnemo obravnavati kot subjekte, ki z nami vstopajo v odnos oz. so del takšnega odnosa. Že če si na kratko ogledamo, kaj pogovorni sistem UI oz. klepetalni roboti so in čemu so namenjeni, lahko razumemo človekov vzgib, da bi jih dojemal kot subjekte in z njimi vzpostavljaj odnose – vključno s sistemi, ki so primarno namenjeni družabništvu (Skjuve et al. 2021). Ti sistemi so na temelju UI zasnovane računalniške aplikacije, ki so zmožne sporazumevanja z ljudmi v naravnem jeziku (McTear et al. 2016). Njihov namen je simulirati človeški podobno govorno interakcijo in v tem okviru ustrezno, smiselno odgovarjati na

vprašanja oz. pozive sogovornika (Mariani et al. 2023). Iz številnih raziskav interakcije klepetalnih robotov s človekom (West et al. 2019; Yuting et al. 2020; Kraus et al. 2021; McKee et al. 2021) je razvidno, da z nenehnim razvojem klepetalni roboti postajajo vse bolj usmerjeni in personalizirani (tudi s pomočjo priporočilnih sistemov UI) glede na potrebe, želje, vizualne preference, karakter itd. uporabnika (Centa Strahovnik 2023, 856–858). Razvoj gre v smeri, da uporabniki sistemom UI vse bolj zaupajo (v primeru avtonomnih vozil celo svoje življenje) in jih uporabljajo vsakodnevno. Posledično te sisteme dojemajo vse bolj 'osebno', kar lahko vidimo npr. že po samih odzivih uporabnikov, ki so lahko tudi izrazito čustveni. Uporabniki tako svojih stvari oz. pripomočkov, npr. svojega računalnika, ne vidijo več le kot orodje, ampak v tej interakciji na drugi strani občutijo prisotnost nekega razumnega bitja, ki pa ni izdelovalec ali programer, in se do njega začnejo vesti kot do soljudi. Pogostejše kot se uporabniki v takšnem odnosu znajdejo, bolj se to občutje krepi – ali celo samodejno vzpostavi (Kim et al. 2012, 242). Uporabniki torej v teh vse bolj osebnih interakcijah ne oblikujejo zgolj identitete pogovornih sistemov UI po svojem okusu, ampak (nezavedno) vplivajo tudi na spremembo lastne identitete (Centa Strahovnik 2023).

V tem okviru moramo biti gotovo posebej pozorni na že znane (negativne) posledice, ki iz teh interakcij izhajajo. Tu nimamo v mislih le bolj opaznih posledic oz. nesreč – kot so npr. nesreče z avtonomnimi vozili, kjer je zaradi uporabnikovega nerazumevanja omejitev sistema, pa tudi povratnih informacij, ki jih sistem posreduje, med človekom in strojem prišlo do motnje v interakciji. Gre tudi za posledice, ki so manj opazne in segajo globlje v odnos s temi sistemi. To vsekakor velja tedaj, ko oseba sistemu pripisuje 'človeške' lastnosti, kot smo omenili že zgoraj, ga v določenem trenutku celo dojema kot osebo in se nanj tako tudi odziva. To zajema tako prijateljske kot celo intimne odnose (Arnold et al. 2017; Arkin et al. 2019). Zavedati se je treba, da prihajajoča (bolje rečeno razvijajoča se) tehnologija UI za nas ni le nekaj uporabnega v smislu orodja, ampak nas spreminja kot ljudi in vpliva na načine, kako vidimo sebe in druge (Turkle 2005). V tem odnosu uporabnik namreč na neki način verjame oz. je prepričan, da je v odnosu z osebo. Zato je eden izmed ključnih korakov za razumevanje tega odnosa tudi v samem prepoznavanju in ustreznem vrednotenju prepričanj, ki jih ima uporabnik o sistemu UI, ki ga uporablja.

Odnose s temi sistemi lahko posameznik oblikuje na dva načina – in sicer glede na to, kako sistem UI, ki ga uporablja, razume. To je, ali kot orodje ali kot nekaj drugega – kot določenega akterja oz. kot subjekt. Nas v prispevku zanima predvsem slednje. Posameznik UI na neki način pripoznava kot subjekt, kadar sistemu recimo pripisuje človeške lastnosti (npr. čustvovanje) in sposobnosti (npr. izvirno ustvarjanje glasbe), posebej pa tudi, kadar sisteme UI razume kot nekoga, ki se odziva na razloge in ima namere oz. čigar dejanja lahko na tej podlagi razumsko razložimo. Preko tega ti sistemi soustvarjajo družbeni prostor (Duffy 2021, 180).

Proizvajalci sistemov UI se zavedajo, da je razumevanje in sprejemanje UI na 'oseben način' in razvoj takšne tehnologije z vidika vključevanja v vsakdanje življenje posameznika velikega pomena. Zgoraj opisana izhodišča antropomorfizma se

uporabljajo tudi pri razvoju družabnih robotov, kajti »če robotu omogočimo, da ustvari in izkazuje občutek identitete, ga ljudje lažje obravnavajo kot družbeno sposobnega udeleženca« (Duffy 2003, 186). Zato bo v prihodnosti gotovo vse večji poudarek na raziskavah in razvoju vse bolj humanoidnih robotov (Breazeal 2003; de Graaf et al. 2013), kjer pa ne bo šlo le za videz, ampak tudi nekaj globljega oz. občutek človeškega.

3.1 Vpliv na temeljne zmožnosti

Če se zdaj ozremo na celoten seznam zgoraj naštetih temeljnih zmožnosti, lahko vidimo, da se razvijajoči se odnosi med sistemi UI in uporabniki nekako dotikajo vseh zmožnosti – enih bolj, drugih pa manj neposredno. Če se osredotočimo na nekaj zmožnosti na seznamu, npr. življenje, telesno zdravje, telesna integriteta, lahko vidimo, kako jih takšni odnosi že oblikujejo, npr. ko posameznik zaupa sistemu UI ob vstopu v avtonomno vozilo; ko sistem UI postavi natančnejšo diagnozo kot usposobljen zdravnik ali opravi uspešen operativni poseg, ki brez njega ne bi bil mogoč ipd. Podoben vpliv se kaže pri drugih temeljnih zmožnostih, tudi tistih, ki na prvi pogled z UI morda niso neposredno povezane. Tak bi bil primer zmožnosti odnosa, skrbi in skupnega življenja z drugimi vrstami – saj gre razvoj sistemov UI tudi v smeri robotskih živali (npr. za terapijske namene) – in pa razumevanja drugih vrst ter komunikacije z njimi preko uporabe velikih jezikovnih modelov (Strahovnik in Centa Strahovnik 2023).

Vidimo torej, da je o dobremu življenju in razvoju ter udejanjanju temeljnih zmožnosti v prihodnosti težko razmišljati brez UI (in odnosa z njo), saj je UI v človeško okolje globoko vpletena že zdaj, s tem pa povezana z uspevanjem posameznika. To vzpostavlja temelj tako za vzpostavljanje kot uresničevanje zmožnosti, a prinaša številna tveganja in nevarnosti. Podobno trdi Johnstone, ki opozarja na dejstvo, da »odnos med tehnologijo in človeškimi zmožnostmi ni tako preprost in enostaven, saj lahko ena in ista tehnologija pri istem posamezniku uporabo določene zmožnosti tako spodbudi kot zmanjša.« (2012, 86) Seveda vseh bremen in vplivov sistemov UI na zmožnosti ne moremo v celoti razumeti, vsekakor pa lahko novi, razvijajoči se odnos z UI in njegov pomen za določeno zmožnost obravnavamo. Zanima nas zlasti – kot smo nakazali že zgoraj –, kako oz. na kakšne načine se ta odnos razvija oz. sploh zgodi, torej zakaj posameznik UI v tem odnosu razume kot subjekt. Tu se opiramo na pojem identitete – zanimali nas bosta zlasti moralna in spoznavna identiteta, ki ju razumemo kot tisti, ki se v tem odnosu tudi spreminjata (saj brez tega do takšnega odnosa ne bi prišlo).

Oglejmo si najprej vpliv na našo identiteto z vidika rabe pogovornih sistemov UI. Prepoznamo lahko več načinov spreminjanja naše celostne identitete. Prvi način se nanaša na privzemanje vlog in perspektiv ter posledično igranje različnih vlog (učenec, učitelj, zaupnik, stranka), ki lahko povratno vplivajo tudi na našo identiteto. Do spreminjanja identitete lahko pride tudi na podlagi vsebine pogovora oz. pri razkrivanju podatkov o samem sebi. Samorazkrivanje lahko posameznike spodbudi tako k razmisleku o lastni identiteti kot celo k njeni spremembi. Tretji način je povezan z zunanjo potrditvijo. Klepetalni sistemi lahko uporabniku

takšno zunanjo potrditev zagotovijo, nato pa se posameznikova identiteta na podlagi te potrditve – ki je npr. od preostalih posameznikov, s katerimi vstopa v odnose ne bi bil deležen – preoblikuje. Četrty način vključuje (pre)oblikovanje identitete na podlagi (so)oblikovanja preferenc uporabnika in prilagojenih priporočil ali predlogov (priporočilni sistemi UI), kar posameznikove preference, interese in vedenje sčasoma lahko preoblikuje. To vpliva na dojemanje naše identitete, saj jo uskladimo z izbirami in priporočili, ki jih podaja klepetalni robot (Centa Strahovnik 2023, 862).

Vse to lahko glede na vidika moralne in spoznavne identitete v kontekstu vpliva sistemov UI razdelamo še podrobneje. Moralna identiteta zajema človekov moralni značaj, vrednote, načela in stališča; povezana z ideali se oblikuje v smiselno, obstojno, razmeroma stabilno predstavo o posameznikovem moralnem jazu (Doris 2002); izraža se v njegovih dejanjih in projektih, ki jih uresničuje v življenju. Tako razumljena moralna identiteta očitno vpliva na človekove moralne presoje, dejanja in odločitve, bodisi z oblikovanjem prostora alternativ, ki jih posameznik vidi kot ustrezne, dopustne ali vredne, bodisi neposredno na same odločitve oz. dajanje prednosti izbrani alternativni pred drugo (Strahovnik 2011). Sistemi UI vplivajo na različne vidike takšne moralne identitete.

Spoznavna identiteta »se nanaša na posameznikov odnos do vednosti in tudi do samega procesa spoznavanja. Kot takšna vključuje načine oz. vzorce, glede na katere posamezniki vednost dojemajo, pridobivajo, potrjujejo in uporabljajo – pa tudi njihova prepričanja in stališča do znanja in njegovih virov.« (Centa Strahovnik 2023, 855–856). Na spoznavno identiteto tako že v izhodišču močno vpliva element obveščenosti, ki ga zadevajo predvsem priporočilni sistemi UI, medtem ko se zdi vidik moralne identitete bolj v ospredju pri pogovornih sistemih UI – posebej tistih, ki so namenjeni družabništvu.

Podobno kot priporočilni sistemi UI in klepetalni roboti na identiteto uporabnika vplivajo in o njej sodoločajo tudi preostale implementacije UI – npr. v zdravstvu, terapevtski dejavnosti, na področju varnosti, izobraževanja ipd. Če se navežemo na etiko zmožnosti ter na povezavo med identiteto in temeljnimi zmožnostmi, lahko jasno vidimo, da se identiteta posameznika oblikuje v okviru zmožnosti in z njimi povezanega dobrega življenja. Če torej analiziramo vpliv UI na temeljne zmožnosti, lahko razumemo, da čeprav UI v okviru nobene od temeljnih zmožnosti ni specifično omenjena, gre za dejavnik, ki v sodobnem svetu ključno vpliva na vse zmožnosti – zlasti kadar ti sistemi UI ne predstavljajo le orodja ali sredstva, ampak vstopajo v neke vrste odnos s posamezniki.

4. Zaključek

V luči zgornjih izhodišč in misli je prav poudariti, da je sam pristop etike zmožnosti glede na vpliv, ki ga UI ima na posameznikovo življenje, nujno potreben posodobitve oz. dopolnitve. Upoštevati je treba razširjenost sistemov UI ter vidike

dostopa, hitrosti, rabe velikih podatkov ipd. Vse to nas privede do razmisleka o prepletu UI s tradicionalnim razumevanjem človekovih zmožnosti, avtonomije in družbene pravičnosti. Vse večja vloga UI v družbi na ta etični okvir etike zmožnosti vpliva, saj slednje lahko ali razširja ali omejuje. Kot vidimo, lahko z orodji UI dostop posameznikov do informacij, izobraževanja (prilagojene učne izkušnje), zdravstvenega varstva (izboljšana diagnostika), gospodarskih priložnosti ipd. povečamo. Tveganje se pojavlja predvsem z zvezi s tem, da bi lahko neenak dostop privedel do večje družbene neenakosti ter poglobitve predsodkov in diskriminacije (Strahovnik 2023, 844–845). Razmislek o UI v okviru etike zmožnosti tako terja osredotočenost na etično upravljanje in socialno pravičnost. UI je treba oblikovati in urejati tako, da bo izrecno namenjena širjenju zmožnosti vseh, zlasti najbolj prikrajšanih. Da bi lahko zagotovili, da bo umetna inteligenca človeške zmožnosti povečevala – in ne zmanjševala –, je ključno razvijati politike in okvire za pravično uporabo UI, vključno z razvojem preglednih, odgovornih in participativnih postopkov za razvoj in uporabo sistemov UI (Miklavčič 2023). UI vpliva tudi na področje avtonomije in lastne dejavnosti. Nadzor, ki je utemeljen na UI, lahko zasebnost in svobodo posameznikov ogrozi, s tem pa potencialno zmanjša njihovo zmožnost polnega in svobodnega udejstvovanja v družbi. Zanašanje na sistem UI (npr. v primeru odločitev) lahko privede do izgube občutka za lastno delovanje, zmanjšanja sposobnosti kritičnega razmišljanja in samoodločanja, nenazadnje do odvisnosti. Končno pa razvoj na področju UI odpira refleksijo, kaj sploh so bistveno človeške sposobnosti in kako so potem nadalje vpete v temeljne zmožnosti. Ko sistemi umetne inteligence prevzemajo naloge, ki so jih tradicionalno opravljali ljudje, ali ko s sistemi UI vstopamo v odnose, se pojavljajo vprašanja o delu, ustvarjalnosti, avtorstvu in sami naravi človeških interakcij. Hkrati pa vse to odpira tudi razmislek o pojmovanju dobrega življenja.

Reference

- Adomavicius, Gediminas, Jesse C. Bockstedt, Shawn P. Curley in Jingjing Zhang.** 2013. Do Recommender Systems Manipulate Consumer Preferences? A Study of Anchoring Effects. *Information Systems Research* 24, št. 4:956–975. <https://doi.org/10.1287/isre.2013.0497>
- Arkin, Ronald in Jason Borenstein.** 2019. Robots, Ethics, and Intimacy: The Need for Scientific Research. V: Matteo Vincenzo D'Alfonso in Don Berkich, ur. *On the Cognitive, Ethical, and Scientific Dimensions of Artificial Intelligence*, 299–309. Cham: Springer Verlag. https://doi.org/10.1007/978-3-030-01800-9_16
- Arnold, Thomas, in Matthias Scheutz.** 2017. The Tactile Ethics of Soft Robotics: Designing Wisely for Human–Robot Interaction. *Soft Robot* 42, št. 2:81–87. <https://doi.org/10.1089/soro.2017.0032>
- Breazeal, Cynthia.** 2003. Emotion and sociable humanoid robots. *International Journal Human-Computer Studies* 59:119–155. [https://doi.org/10.1016/s1071-5819\(03\)00018-1](https://doi.org/10.1016/s1071-5819(03)00018-1)
- Caulfield, Brian.** 2020. *What's a Recommender System?* NVIDIA blogs. <https://blogs.nvidia.com/blog/whats-a-recommender-system/> (pridobljeno 20. 5. 2024)
- 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>
- Doris, John M.** 2002. *Lack of Character: Personality and Moral Behaviour*. Cambridge, MA: Cambridge University Press.
- de Graaf, Maartje M.A., in Somaya Ben Allouch.** 2013. Exploring influencing variables for the acceptance of social robots. *Robotics and Autonomous Systems* 61, št. 12:1476–1486. <https://doi.org/10.1016/j.robot.2013.07.007>

- Duffy, R. Brian.** 2003. Anthropomorphism and the social robot. *Robotics and Autonomous Systems* 42:177–190. [https://doi.org/10.1016/S0921-8890\(02\)00374-3](https://doi.org/10.1016/S0921-8890(02)00374-3)
- Globokar, Roman.** 2021. *Vzgojni izzivi šole v digitalni dobi*. Teološka fakulteta: Ljubljana.
- Johnstone, Justine.** 2012. Capabilities and Technology. V: Philip Brey, Adam Briggie in Edward Spen, ur. *The Good Life in a Technological Age*, 77–91. London: Routledge.
- Kim, Youjeong in S. Shyam Sundar.** 2012. Anthropomorphism of computers: Is it mindful or mindless? *Computers in Human Behavior* 28, št. 1:241–250. <https://doi.org/10.1016/j.chb.2011.09.006>.
- Knotzer, Nicolas.** 2008. *Product Recommendations in E-Commerce Retailing Applications*. Frankfurt am Main: Peter Lang.
- Kraus, Matthias, Philip Seldschopf in Wolfgang Minker.** 2021. Towards of Trustworthy Chatbot for Mental Health Applications. V: Jakob Lokoč, Tomáš Skopal, Klaus Schoeffmann, Vasileios Mezaris, Xirong Li, Stefanos Vrochidis in Ioannis Patras, ur. *MultiMedia Modeling*, 354–366. Cham: Springer International Publishing.
- Mariani, Marcello M., Novin Hashemi in Jochen Wirtz.** 2023. Artificial Intelligence Empowered Conversational Agents: A Systematic Literature Review and Research Agenda. *Journal of Business Research* 161:113838. <https://doi.org/10.1016/j.jbusres.2023.113838>
- McKee, Kevin R., Xuechunzi Bai in Susan Fiske.** 2021. Humans Perceive Warmth and Competence in Artificial Intelligence. *PsyArXiv*. 26. 2. <https://doi.org/10.31234/osf.io/5ursp>
- McTear, Michael, Zoraida Callejas in David Griol.** 2016. *The Conversational Interface. Talking to Smart Devices*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-32967-3>
- Miklavčič, Jonas.** 2023. Ideal transparentnosti v digitalni dobi. *Bogoslovni* 83, št. 4:825–838. <https://doi.org/10.34291/BV2023/04/Miklavcic>.
- Nussbaum, C. Martha.** 2000. *Women and Human Development: The Capabilities Approach*. New York: Cambridge University Press.
- . 2011. *Creating Capabilities: The Human Development Approach*. Cambridge: Harvard University Press.
- Robeyns, Ingrid, in Morten Fibieger Byskov.** 2023. The Capability Approach. V: Edward N. Zalta, ur. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/sum2023/entries/capability-approach> (pridobljeno 20. 5. 2024)
- Strahovnik, Vojko.** 2011. Identity, Character and Ethics: Moral Identity and Reasons for Action. *Synthesis philosophica* 26, št. 1:67–77.
- . 2022. Identiteta, etika prepričanja, razumnost in resonanca. *Bogoslovni vestnik* 82, št. 3:547–559. <https://doi.org/10.34291/bv2022/03/strahovnik>
- Strahovnik, Vojko, in Mateja Centa Strahovnik.** 2023. AI animal companions: animals and companionship in the age of artificial intelligence. V: Josip Guč, ur. *Nije pas beštija: 5. dani kulturne animalistike: Split, 17-19. 10 2023*, 120–121. Split: Hrvatsko bioetičko društvo.
- 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>
- Skjuve Marita, Asbjørn Følstad, Knut Inge Foster-vold in Petter Bae Brandtzaeg.** 2021. My Chatbot Companion - a Study of Human-Chatbot Relationships. *International Journal of Human-Computer Studies* 149:102601. <https://doi.org/10.1016/j.ijhcs.2021.102601>
- Turkle, Sherry.** 2005 [1984]. *The Second Self: Computers and the Human Spirit*. Cambridge, MA: The MIT Press.
- Vodičar, Janez.** 2023. Avtoriteta na področju vzgoje in verovanja v digitalni dobi. *Bogoslovni* 83, št. 4:1035–1047. <https://doi.org/10.34291/bv2023/04/vodicar>
- Zhang, Qian, Lu Jie, in Jin Yaochu.** 2021. Artificial intelligence in Recommender Systems. *Complex and Intelligent Systems* 7:439–457. <https://doi.org/10.1007/s40747-020-00212-w>
- Žalec, Bojan.** 2008. Od radikalne interpretacije do etike zmožnosti: filozofska razprava o možnostih in pogojih (medkulturnega) dialoga. *Šolsko polje* 19, št. 5/6:73–89.
- Yuting, Liao in Jianguan He.** 2020. Racial Mirroring Effects on Human-Agent Interactions in Psychotherapeutic Conversations. *Proceedings of the 25th International Conference on Intelligent User Interfaces, IUI'20*, 430–442. New York, NY: Association for Computing Machinery. <https://doi.org/10.1145/3377325.3377488>
- West, Mark, Rebecca Kraut in Chew Han Ei.** 2019. *I'd Blush if I Could: Closing Gender Divides in Digital Skills through Education*. [s. l.]: Unesco. <https://unesdoc.unesco.org/ark:/48223/pf0000367416.locale=en> (pridobljeno 23. 8. 2024).