

Bogoslovni vestnik

Theological Quarterly
EPHEMERIDES THEOLOGICAE

DIGITALIZACIJA IN UMETNA INTELIGENCA KOT IZZIVA ZA ETIKO IN RELIGIJO

Bojan Žalec	Uvodnik
Branko Klun	<i>Kaj pomeni misliti? Eksistencialno razumevanje in UI</i>
Bojan Žalec	<i>Človeški podobna umetna inteligenca in superinteligence</i>
Mateja Centa Strahovnik	<i>Etika zmožnosti, priporočilni in pogovorni sistemi UI ...</i>
Vojko Strahovnik	<i>Algoritemsko nasilje</i>
Rok Gregorčič	<i>Razmerje med človeško in umetno inteligenco ...</i>
Simon Malmenvall	<i>State and Christian Enlightenment: Background of the Mass ...</i>
Mari Jože Osredkar	<i>Umetna inteligenca – prepovedan sad iz raja?</i>
Ivan Platovnjak et al.	<i>Artificial Intelligence and Imago Dei</i>
Nik Trontelj	<i>Teološke kreposti vere, upanja in ljubezni kot temelj človeškosti ...</i>
Lucija Rožman	<i>Vprašanje človeške ustvarjalnosti v dobi UI</i>
Roman Globokar	<i>Pogled Katoliške Cerkve na razvoj in uporabo umetne inteligence</i>
Janez Vodičar	<i>Kateheza kot pomoč umetni inteligenci</i>
Andrej Šegula	<i>Umetna inteligenca v luči praktične pastoralne teologije</i>
Tadej Stegu	<i>Antropološki izzivi kateheze v času umetne inteligence</i>
Urška Jeglič	<i>Virtualna resničnost kot pripomoček pri poučevanju ...</i>

Glasilo Teološke fakultete Univerze v Ljubljani

Letnik 84

2024 • 4

Bogoslovni vestnik

Theological Quarterly
EPHEMERIDES THEOLOGICAE

4

**Letnik 84
Leto 2024**

Glasilo Teološke fakultete Univerze v Ljubljani

Ljubljana 2024

KAZALO / TABLE OF CONTENTS

TEMA / THEME

DIGITALIZACIJA IN UMETNA INTELIGENCA KOT IZZIVA ZA ETIKO IN RELIGIJO
DIGITALISATION AND AI AS CHALLENGES FOR ETHICS AND RELIGION

- 737 Uvodnik / Editorial (Bojan Žalec)**
- 741 Branko Klun – Kaj pomeni misliti? Eksistencialno razumevanje in umetna inteligenca**
What is Called Thinking? Existential Understanding and Artificial Intelligence
- 757 Bojan Žalec – Človeški podobna umetna inteligenca in superinteligenca: verjetnost in glavne težave njunega oblikovanja**
Human-like Artificial Intelligence and Superintelligence: The Probability and Main Challenges of Their Design
- 771 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
- 783 Vojko Strahovnik – Algoritemsko nasilje**
Algorithmic Violence
- 797 Rok Gregorčič – Razmerje med človeško in umetno inteligenco v luči Habermasove teorije**
The Relation between Human and Artificial Intelligence in Light of Habermas's Theory
- 807 Simon Malmenvall – State and Christian Enlightenment: Background of the Mass Elementary Education in Central Europe**
Država in krščansko razsvetljenje: ozadje množičnega osnovnega šolstva v srednji Evropi
- 823 Mari Jože Osredkar – Umetna inteligenca – prepovedan sad iz raja?**
Artificial Intelligence – Forbidden Fruit from The Garden of Eden?
- 835 Ivan Platovnjak and Tone Svetelj – Artificial Intelligence and *Imago Dei*: A New Dilemma for Philosophical and Theological Anthropology**
Umetna inteligenca in Imago Dei: Nova dilema za filozofsko in teološko antropologijo
- 847 Nik Trontelj – Teološke kreposti vere, upanja in ljubezni kot temelj človeškosti v odnosu do transhumanizma**
The Theological Virtues of Faith, Hope and Love as the Foundation of Humanity in Relation to Transhumanism
- 857 Lucija Rožman – Vprašanje človeške ustvarjalnosti v dobi UI**
The Question of Human Creativity in the Age of AI

867 Roman Globokar – Pogled Katoliške Cerkve na razvoj in uporabo umetne intelligence

The Catholic Church's View on the Development and Use of Artificial Intelligence

885 Janez Vodičar – Kateheza kot pomoč umetni inteligenci

Catechesis as an Aid for Artificial Intelligence

899 Andrej Šegula – Umetna inteligenca v luči praktične pastoralne teologije

Artificial Intelligence in the Light of Practical Pastoral Theology

909 Tadej Stegu – Antropološki izzivi kateheze v času umetne intelligence

Anthropological Challenges of Catechesis in the Age of Artificial Intelligence

921 Urška Jeglič – Virtualna resničnost kot pripomoček pri poučevanju religijskih vsebin v šolah

Virtual Reality as a Tool for Teaching Religion in Schools

931 David Kraner – Vloga moči in uporaba UI v mrežni družbi

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

945 Tomaž Erzar – Izražanje jeze in sovražstva na spletu: učinki, posledice, motivacija in meje navidezne povezanosti

Expressing Anger and Hatred Online: Effects, Consequences, Motivations and the Limits of Apparent Connectedness

OCENA / REVIEW

955 Paolo Benanti, *Human in the LOOP. Decisioni umane e intelligenze artificiali* (David Kraner)

NOVI DOKTORJI IN NOVE DOKTORICE ZNANOSTI / NEW DOCTORS OF SCIENCE

959 Kristina Grief, Klientovo doživljanje spremembe v skupinski obravnavi spolnega nasilja v otroštvu po metodi relacijske družinske terapije (Tanja Repič Slavič)

961 Maja Kaninska, Ideološka vloga Srbske pravoslavne Cerkve v SFR Jugoslaviji v 80. letih 20. stoletja (Bogdan Dolenc)

964 Uršula Ješe, Vzpostavitev varne navezanosti v rejniških in posvojiteljskih družinah (Lia Katarina Kompan Erzar)

967 Urška Mali Kovačič, Fenomenološka analiza religioznosti kot zaščitnega dejavnika za samomorilnost (Roman Globokar)

969 Milena Marković, Ljubosumje v partnerskem odnosu z vidika relacijske družinske paradigme in klinične prakse (Tomaž Erzar)

971 Jonas Miklavčič, Transparentnost in etika sistemov algoritemskega odločanja (Vojko Strahovnik)

973 Miha Pate, Čustvena regulacija in strategije discipliniranja pri starših

mladostnikov v okviru relacijskega družinskega modela (**Christian Gostečnik**)

975 Martin Perčič, Etična razsežnost medosebnih odnosov v luči personalistične etike Karola Wojtyły in Teologije telesa Janeza Pavla II. (**Roman Globokar**)

977 Miriam Stanonik, Poslanstvo in edukativni izzivi Abrahamove žene Sare (**Stanko Gerjolj**)

979 Nataša Vrbnjak, Samopodoba ter sposobnosti za bližino in regulacijo čutenj odraslih z izkušnjo posvojitve z vidika relacijske družinske terapije (**Tanja Pate**)

CELOLETNO KAZALO (2024) / ANNUAL BIBLIOGRAPHY (2024)

~~**981 Celoletno kazalo (2024)**~~

~~*Annual Bibliography (2024)*~~

SEZNAM RECENZENTOV (2024) / LIST OF PEER-REVIEWERS (2024)

~~**987 Seznam recenzentov (2024)**~~

~~*List of Peer Reviewers (2024)*~~

NAVODILA SODELAVCEM / INSTRUCTIONS FOR THE AUTHORS

~~**989 Navodila sodelavcem**~~

~~*Instructions For The Authors*~~

SODELAVCI / CONTRIBUTORS

Mateja CENTA STRAHOVNIK

didaktika in pedagogika religije, dr., doc. Didactics and Pedagogy of Religion, Ph.D., Asst. Prof.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
mateja.centastrahovnik@teof.uni-lj.si

Bogdan DOLENC

ekleziologija, dr., doc. Ecclesiology, Ph.D., Asst. Prof.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
bogdan.dolenc@guest.arnes.si

Tomaž ERZAR

zakonska in družinska terapija, dr., prof. Marital and Family Therapy, Ph.D., Prof.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
tomaz.erzar@teof.uni-lj.si

Stanko GERJOLJ

zakonska in družinska terapija, dr., zasl. prof. Marital and Family Therapy, Ph.D., Prof. Emer.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
stanko.gerjolj@teof.uni-lj.si

Roman GLOBOKAR

moralna teologija, dr., izr. prof. Moral Theology, Ph.D., Assoc. Prof.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
roman.globokar@teof.uni-lj.si

Christian GOSTEČNIK

zakonska in družinska terapija, dr., zasl. prof. Marital and Family Therapy, Ph.D., Prof. Emer.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
christian.gostecnik@teof.uni-lj.si

Rok GREGORČIČ

teologija, mag. teol., mladi raziskovalec Theology, MA in Theology, Jr. Res.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
rok.gregorcic@teof.uni-lj.si

Urška JEGLIČ

religiologija, dr., asist. Religious Studies, Ph.D., Tch. Asst.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
urska.jeglic@teof.uni-lj.si

Branko KLUN

filozofija, dr., prof. Philosophy, Ph.D., Prof.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
branko.klun@teof.uni-lj.si

Lia Katarina KOMPAN ERZAR

zakonska in družinska terapija, dr., izr. prof. Marital and Family Therapy, Ph.D., Assoc. Prof.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
LiaKatarina.KompanErzar@teof.uni-lj.si

David KRANER

pastoralna teologija, dr., asist. Practical Theology, Ph.D., Tch. Asst.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
david.kraner@teof.uni-lj.si

Simon MALMENVALL

religiologija in religijska antropologija, Religious Studies and Anthropology of Religion,
dr., doc. Ph.D., Asst. Prof.
Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
Poljanska 4, SI – 1000 Ljubljana
simon.malmenvall@teof.uni-lj.si

Mari Jože OSREDKAR

osnovno bogoslovje in dialog, dr., izr. prof. Fundamental Theology and Dialogue, Ph.D., Assoc. Prof.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
mari.osredkar.ofm@siol.net

Tanja PATE

zakonska in družinska terapija, dr., doc. Marital and Family Therapy, Ph.D., Asst. Prof.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
tanja.pate@teof.uni-lj.si

Ivan PLATOVNJAK

duhovna teologija, dr., doc. Spiritual Theology, Ph.D., Asst. Prof.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
ivan.platovnjak@teof.uni-lj.si

Tanja REPIČ SLAVIČ

zakonska in družinska terapija, dr., prof. Marital and Family Therapy, Ph.D., Prof.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
tanja.repic@teof.uni-lj.si

Lucija ROŽMAN

pastoralna teologija, mag. teol., mlada raziskovalka Practical Theology, MA in Theology, Jr. Res.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
lucija.rozman@teof.uni-lj.si

Tadej STEGU

pastoralna teologija, dr., doc. Practical Theology, Ph.D., Asst. Prof.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
tadej.stegu@teof.uni-lj.si

Vojko STRAHOVNIK

filozofija in etika, dr., izr. prof. Philosophy and Ethics, Ph.D., Assoc. Prof.
 Univerza v Ljubljani, Filozofska fakulteta University of Ljubljana, Faculty of Arts
 Poljanska 4, SI – 1000 Ljubljana
vojko.strahovnik@teof.uni-lj.si

Tone SVETELJ

filozofija in religijski študiji, dr., izr. prof. Philosophy and Religious Studies, Ph.D., Assoc. Prof.
 Hellenic College, Boston College Boston College, Hellenic College
 Goddard Avenue 50, US – 02445 Brookline, MA
tsvetelj@hchc.edu

Andrej ŠEGULA

pastoralna teologija, dr., doc. Practical Theology, Ph.D., Asst. Prof.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
andrej.segula@teof.uni-lj.si

Nik TRONTELJ

osnovno bogoslovje, dr., asist. Fundamental Theology, Ph.D., Tch. Asst.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
nik.trontelj@teof.uni-lj.si

Janez VODIČAR

pastoralna teologija, dr., prof. Practical Theology, Ph.D., Prof.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
janez.vodicar@teof.uni-lj.si

Bojan ŽALEC

filozofija, dr., prof. Philosophy, Ph.D., Prof.
 Univerza v Ljubljani, Teološka fakulteta University of Ljubljana, Faculty of Theology
 Poljanska 4, SI – 1000 Ljubljana
bojan.zalec@teof.uni-lj.si

Digitalizacija in umetna inteligenca kot izziva za etiko in religijo

Digitalisation and AI as Challenges for Ethics and Religion

Uvodnik

Razvoj tehnologije je v zadnjih desetletjih doživel silovit razmah. Vstopili smo v digitalno dobo, ko je večina informacij že digitalno opisana in omogoča digitalno obdelavo ter uporabo. Naša tehnološka doba je digitalna doba. Izraz 'digitalizacija' pa se ne nanaša samo na tehnični vidik, ampak tudi na aspekte psihičnega, socialnega, družbenega, kulturnega in duhovnega vpliva digitalizacije. Digitalizacija zajema vse ravni in področja človekovega življenja in jih pomembno preobraža. Je dvorezen meč: po eni strani omogoča izboljšave, po drugi strani pa predstavlja nevarnost – ali pa so njeni učinki že kar dejansko škodljivi. Nevarnost je korenita: pod vprašaj je postavljena človeškost in človečnost človeka. Nekateri govorijo kar o 'preseganju' človeka. Pojavile so se ideologije, kot je (radikalni) transhumanizem, ki so z razumevanjem človeka, kot ga zasledimo v izročilu (svetovnih) religij, nezdružljive. Vse bolj se uveljavlja scientistični pogled na človeka, ki tega redukcionistično razume v pojmih algoritmov in obdelave podatkov. Ta pogled marsikje – na pomembnih pozicijah moči in odločanja – izpodriva personalistično in humanistično razumevanje človeka, po katerem je človek ustvarjeno, telesno in odnosno bitje, umeščeno v svoje okolje ter odprto v svet in presežnost. Tak pogled je osnova pripoznavanja človeka kot svobodnega bitja.

Vsi ti pojavi zahtevajo sodoben (kritičen) odziv teologije, filozofije in drugih humanističnih ved. Te morajo relevantne pojave proučiti, jih po potrebi utemeljeno zavrniti in s svojimi spoznanji podati osnovo za izdelavo smernic etičnega ravnanja ter reguliranja življenja v digitalni dobi. Najti moramo ustrezen kompromis – srednjo pot – med omogočanjem razvoja znanosti in tehnologije, njene uporabe na eni strani ter zaščito človeka, človeškosti in človečnosti na drugi. Temu namenu sledijo tudi članki v pričujoči številki.

Številka prinaša obravnavo različnih vidikov, vrst in oblik tehnologije, posebno pozornost pa namenjamo umetni inteligenci. Umetna inteligenca se je izkazala za osrednjo tehnologijo sedanjosti in prihodnosti. Velik napredek v razvoju umetne

inteligence predstavlja globoko družbeno spremembo, ki bo korenito spremenila človekov življenjski svet – veliko bolj, kot ga že je, v obsegu in na način, ki si ga še ne moremo zares predstavljati in oceniti. V prihodnosti bodo sistemi umetne inteligence vse bolj vključeni v naše (vsakdanje) življenje in delo. Mnogi to razumejo kot temeljno revolucijo v človekovem bivanju. Kajti tehnične inovacije nimajo le velikega potenciala za podporo človekovi dejavnosti, na primer raziskovalni, komunikacijski, ekonomski, zdravstveni, temveč spreminjajo človekovo okolje in resničnost ter hkrati vplivajo na človeka samega: tako na njegovo zunanost in telesnost kot tudi na njegovo ‚notranjost‘. Oblikujejo in spreminjajo njegovo telo, način razmišljanja in doživljanja; predstavljajo izziv za človekovo razumevanje samega sebe in resničnosti. Ti izzivi tehnologije zadevajo tako rekoč vse vidike in področja človekovega življenja, vključno z etičnimi in religijskimi, ki se jim še posebej posvečamo. Namen pričujoče številke je povečati in poglobiti njihovo razumevanje ter prispevati k reševanju problemov, ki jih razvoj tehnologije prinaša.

Številka je tematsko bogata in raznovrstna. Uvajajo jo filozofski prispevki (Žalec, Strahovnik, Miklavčič, Štivič, Pohar, Gregorčič), ki se ukvarjajo z opredelitvijo pojma inteligence, izzivi velikih jezikovnih modelov, vprašanjem transparentnosti, kiborgizacijo, ter razumevanjem (vloge in pomena) tehnologije pri pomembnih mislecih (de Chardin, Habermas). Strahovnik Centa predstavlja sodobne izzive identitete z vidika pogovornih robotov. Članki Osredkarja, Trontlja in Malmenvalla ponujajo pogled z zgodovinske perspektive, prva dva z religiološke, zadnji pa v okviru slovenske intelektualne zgodovine. Podoben pristop odlikuje Petkovškov članek o Misijonski družbi. Jeglič religiološko raziskuje širjenje salafizma na Balkanu v digitalni dobi. Pedagoški in psihoterapevtski sklop tvorijo članki Vodičarja, Nežič Glavice, Erzarja in Kranerja, ki se ukvarjajo s problemi avtoritete, sodobne pedagoške paradigme – ki je primerna za digitalno dobo –, problemom spletne pornografije ter osamljenostjo mladih. Platonvjak in Svetelj povezujeta teološke in filozofske vidike v svoji raziskavi tehnologije z vidika duhovnosti, telesnosti in nesmrtnosti. Teološki sklop dopolnjujeta prispevka Globokarja in Steguja. Prvi v kontekstu vprašanja evtanazije predstavlja različna razumevanja človekovega dostojanstva in se zavzema za njegovo odnosno razumevanje, drugi se posveča kerigmi in lepoti v tehnološki dobi.

Vidimo torej, da ima številka medpanožni značaj. Avtorji o problemih razpravljamo z različnih vidikov: filozofskega, teološkega, (bio)etičnega, religiološkega, kulturološkega, psihološkega in psihoterapevtskega, pedagoškega, zgodovinskega in drugih. Verjamemo, da lahko izzive in probleme, s katerimi se ukvarjamo v tokratni številki, razumemo bolje, če se jih lotimo z različnih, medsebojno dopolnjujočih se vidikov – etična in religijska vprašanja ter problemi, ki so povezani s tehnologijo, digitalizacijo, umetno inteligenco itd., so namreč zelo kompleksni in večplastni, zato zahtevajo raziskovanje in razjasnitev z različnih plati.

Vidimo torej, da ima številka medpanožni značaj. Avtorji o problemih razpravljamo z različnih vidikov: filozofskega, teološkega, (bio)etičnega, religiološkega, kulturološkega, psihološkega in psihoterapevtskega, pedagoškega, zgodovinskega in drugih. Verjamemo, da lahko izzive in probleme, s katerimi se ukvarjamo v to-

kratni številki, razumemo bolje, če se jih lotimo z različnih, medsebojno dopolnjujočih se vidikov – etična in religijska vprašanja ter problemi, ki so povezani s tehnologijo, digitalizacijo, umetno inteligenco itd., so namreč zelo kompleksni in večplastni, zato zahtevajo raziskovanje in razjasnitev z različnih plati.

Vidimo torej, da ima številka medpanožni značaj. Avtorji o problemih razpravljamo z različnih vidikov: filozofskega, teološkega, (bio)etičnega, religiološkega, kulturološkega, psihološkega in psihoterapevtskega, pedagoškega, zgodovinskega in drugih. Verjamemo, da lahko izzive in probleme, s katerimi se ukvarjamo v to kratni številki, razumemo bolje, če se jih lotimo z različnih, medsebojno dopolnjujočih se vidikov – etična in religijska vprašanja ter problemi, ki so povezani s tehnologijo, digitalizacijo, umetno inteligenco itd., so namreč zelo kompleksni in večplastni, zato zahtevajo raziskovanje in razjasnitev z različnih plati.

Vidimo torej, da ima številka medpanožni značaj. Avtorji o problemih razpravljamo z različnih vidikov: filozofskega, teološkega, (bio)etičnega, religiološkega, kulturološkega, psihološkega in psihoterapevtskega, pedagoškega, zgodovinskega in drugih. Verjamemo, da lahko izzive in probleme, s katerimi se ukvarjamo v to kratni številki, razumemo bolje, če se jih lotimo z različnih, medsebojno dopolnjujočih se vidikov – etična in religijska vprašanja ter problemi, ki so povezani s tehnologijo, digitalizacijo, umetno inteligenco itd., so namreč zelo kompleksni in večplastni, zato zahtevajo raziskovanje in razjasnitev z različnih plati.

Vidimo torej, da ima številka medpanožni značaj. Avtorji o problemih razpravljamo z različnih vidikov: filozofskega, teološkega, (bio)etičnega, religiološkega, kulturološkega, psihološkega in psihoterapevtskega, pedagoškega, zgodovinskega in drugih. Verjamemo, da lahko izzive in probleme, s katerimi se ukvarjamo v to kratni številki, razumemo bolje, če se jih lotimo z različnih, medsebojno dopolnjujočih se vidikov – etična in religijska vprašanja ter problemi, ki so povezani s tehnologijo, digitalizacijo, umetno inteligenco itd., so namreč zelo kompleksni in večplastni, zato zahtevajo raziskovanje in razjasnitev z različnih plati.

Bojan Žalec

Pregledni znanstveni članek/Article (1.02)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 741—756

Besedilo prejeto/Received:08/2024; sprejeto/Accepted:10/2024

UDK/UDC: 141.32:004.89

DOI: 10.34291/BV2024/04/Klun

© 2024 Klun, CC BY 4.0

Branko Klun

Kaj pomeni misliti? Eksistencialno razumevanje in umetna inteligenca

What is Called Thinking? Existential Understanding and Artificial Intelligence

Povzetek: Naslovno vprašanje spominja na Heideggerjeva istoimenska predavanja in s pomočjo njegove eksistencialno-fenomenološke metode problematizira prevladujočo razlago mišljenja kot mentalnih procesov, ki jih vrši človeški um. Ustvarjalci umetne inteligence so si postavili za cilj, da bi umetna inteligenca posnemala človeško mišljenje – in posledično znala opravljati dejavnosti, ki jih lahko imenujemo inteligentne. Toda poskus redukcije človeškega mišljenja na logične operacije in ‚računajoče mišljenje‘ (Heidegger) spregleda njegov eksistencialni značaj. Mišljenje namreč ni zgolj atribut človeškega bitja, temveč modus njegovega eksistiranja. Zato fenomena mišljenja ne moremo ustrezno opisati v tretjeosebni perspektivi (empiričnih znanosti) – zahteva namreč prvoosebno (fenomenološko) refleksijo. Prispevek¹ do poskusov, ki tudi fenomenološko misel uporabljajo kot sredstvo za razvoj umetne inteligence (Heideggerska umetna inteligenca), zavzema kritično držo in zagovarja razumevanje mišljenja, ki je umeščeno v celovitost človekove eksistence – in ji tako omogoča doseganje njene polnosti in pristnosti.

Ključne besede: mišljenje, umetna inteligenca, eksistencialna analiza, računajoče mišljenje, Heidegger, Dreyfus, pristnost

Abstract: The title question recalls Heidegger's lectures of the same name and uses his existential-phenomenological method to problematise the dominant interpretation of thinking as mental processes performed by the human mind. The creators of artificial intelligence have set the goal that artificial intelligence should be able to mimic human thinking and, as a result, be able to carry out activities that can be called intelligent. But the attempt to reduce human thinking to logical operations and "calculative thinking" (Heidegger) overlooks its existential character. For thinking is not just an attribute of human beings, it is

¹ Prispevek je nastal v okviru raziskovalnega programa „Religija, etika, edukacija in izzivi sodobne družbe (P6-0269)“ in temeljnega raziskovalnega projekta „Teologija, digitalna kultura in izzivi na človeka osredičene umetne inteligence (J6-4626)“, ki ju sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

an essential way in which we exist (i.e. a mode of existing). Therefore, the phenomenon of thinking cannot be adequately described in a third-person perspective (of the empirical sciences) but requires a first-person (phenomenological) reflection. The article is also critical of attempts to use the phenomenological approach as a means of developing artificial intelligence (Heideggerian AI) and advocates an understanding of thinking that is integrated into the totality of human existence, enabling it to achieve its fullness and authenticity.

Keywords: thinking, Artificial Intelligence, existential analysis, calculative thinking, Heidegger, Dreyfus, authenticity

V zimskem semestru 1951/52 je imel Martin Heidegger predavanja z naslovom „Kaj pomeni misliti?“ (Heidegger 2002), v katerih je skušal pokazati, kako je pod vplivom razvoja moderne znanosti in tehnologije prišlo do zoženja pojma mišljenja. Misliti vse bolj pomeni izvajati miselne operacije, ki jih lahko prevedemo v formalni jezik logike in matematike. Lahko bi rekli, da je danes najbolj cenjena oblika mišljenja ‚znanstveno‘ (objektivno) mišljenje, ki sledi strogim logičnim in metodološkim pravilom – in predstavlja kontrast običajno nenatančnemu in pogosto dvoumnemu mišljenju, ki prevladuje v našem subjektivno obarvanem vsakdanjiku. Kasneje je Heidegger uvedel razlikovanje med ‚računajočim‘ in ‚premišljujočim‘ (osveščenim) mišljenjem² ter slednje predstavil kot alternativni način človekovega mišljenja in bivanja v svetu. V tem prispevku Heideggerju ne želim strogo slediti, nameravam pa uporabiti njegova zgodnja fenomenološka izhodišča – zlasti eksistencialno analizo, ki jo je zasnoval v svojem najznamenitejšem delu *Bit in čas* –, da bi pokazal na naravo mišljenja, s katerim se srečujemo pri fenomenu umetne inteligence. Heidegger tega fenomena seveda ni poznal – ravno sredi petdesetih let prejšnjega stoletja se je pojem umetne inteligence šele pojavil –, ga je pa ob analiziranju dotedanjega razvoja tehnike, ki je postavil temelje vzniku inteligentnih strojev in naprav, slutil.

1. Kako misliti mišljenje?

Ko se vprašamo po fenomenu mišljenja oziroma po tem, kaj sploh pomeni misliti, pogosto pozabimo, da smo v metodološko nerodnem položaju. Običajni poskusi, da bi mišljenje definirali, ki jih najdemo v različnih slovarjih, na to težavo sploh ne pomislijo in nam podajajo odgovore, kot je npr. »biti dejaven v zavesti« (SSKJ, s.v. »misliti) ali pa »dejavnost uporabe uma pri proizvajanju misli«. ³ Misliti se razume kot dejavnost, ki jo vrši zavest oz. um (angl. *mind*, lat. *mens*) in je podvržena določenim pravilom in zakonitostim. Če mišljenje teh pravil ne upošteva, ni ‚logično‘

² V izvirniku gre za ‚*rechnendes Denken*‘, kateremu stoji nasproti ‚*besinnliches Denken*‘ (Heidegger 1959, 15; 2000, 52). V angleščino se običajno prevajata kot ‚*calculative thinking*‘ in ‚*meditative thinking*‘.

³ Spletni slovar Merriam-Webster (2024, s.v. »thinking«): »the action of using one's mind to produce thoughts.«

in je zato neustrezno. Očitno je mišljenje povezano z logiko, ki označuje ustrezno oz. pravilno rabo (raz)uma – in razum je tudi eden od pomenov grške besede *logos*. Misliti mišljenje bi potem takem pomenilo raziskati, kako poteka pravilna dejavnost mišljenja.

Ob tem se zdi, da imamo naslednjo samoumevno situacijo. Mišljenje se kaže kot lastnost in dejavnost človeka. Ker človek to dejavnost vrši v svojem umu, ki je vezan na možgane in njihove materialno-biološke procese, se zdi, da treba delovanje mišljenja raziskati tako na mentalno-logični ravni kot tudi na nevronske-biološki ravni. Medtem ko se drugi ravni posveča naravoslovno proučevanje možganov, pa prvo raven obravnava filozofska analiza, ki se je v anglosaškem svetu uveljavila pod nazivom *philosophy of mind*. 'Filozofija uma' se sprašuje po naravi in zakonitostih 'mentalnih' fenomenov in njihovem odnosu do lastnega telesa ter do 'materialnih' fenomenov zunanjega sveta. V zadnjih desetletjih se je ta filozofska disciplina vključila v interdisciplinarno vedo, ki nosi ime 'kognitivna znanost' in poskuša raziskati človekovo mišljenje (ter spoznanje oz. kognicijo) v interdisciplinarnem sodelovanju med pristopi filozofije, psihologije, nevroznanosti in umetne inteligence. Kognitivna znanost se zdi najustreznejši in najbolj kompetenten sogovornik, če želimo na vprašanje »Kaj pomeni misliti?« odgovoriti na znanstveni način.

Toda ali ni ta navidezno samoumeven pristop, ki hoče človekovo mišljenje proučiti in spoznati s tretjeosebne perspektive – ta predstavlja osnovo moderni znanstveni metodologiji in tako zagotavlja znanstveno objektivnost in nevtralnost – v temelju vprašljiv? Pri tem se pozablja, da znanstvenega pristopa ne izvaja znanost sama, temveč človek, ki misli. Če človek proučuje zunanje fenomene, jih lahko objektivira in do njih zavzame distanco, pri mišljenju pa človek nastopa hkrati kot subjekt in objekt. Misleč namreč želi doseči spoznanje mišljenja – seveda ne zgolj svojega lastnega, temveč mišljenja v splošnem, medčloveškem oz. intersubjektivnem pomenu. Tudi mišljenje v prvoosebni perspektivi namreč ni nikoli zgolj lastno, temveč vedno pod vplivom drugih ljudi, ki so nas v mišljenje uvedli, ter pod vplivom logičnih zakonitosti, ki jih v mišljenju kot takem odkrivamo vsi ljudje.

Zato je treba najprej priznati, da znanstveni pristop k proučevanju mišljenja nikoli ne more zadostiti tistim zahtevam, ki si jih postavlja metodologija moderne znanosti: v prvi vrsti nevtralnost in neudeleženos subjekta, kar velja kot pogoj za objektivnost spoznanja. To bi lahko opisali tudi kot nezmožnost, da bi tretjeosebna perspektiva (objektivnega) pogleda 'od zunaj' v celoti nadomestila prvoosebno perspektivo (subjektivnega) doživljanja 'od znotraj'. Prav priznanje te nezmožnosti in nujnost vključitve človekovega zavedanja je izhodišče smeri filozofije, ki jo povezujemo z imenom fenomenologija. Fenomenologija znanstvenemu pristopu ne oporeka dostopa do resnice, oporeka pa mu upravičenost do celovitega spoznanja (resnice) fenomenov. Znanstveni pristop namreč zaznamuje določena navivnost (nekakšna slepa pega), ker pozablja na vključenost subjekta v navidezno objektivno spoznanje. Ta pristop prvenstvo samoumevno pripisuje zunanjemu materialnemu svetu, znotraj katerega obravnava tudi človeka. Človek je v prvi vrsti telesno bitje, ki se od drugih živih bitij razlikuje po posebni sposobnosti možganov, da proizvedejo zavest in človeku omogočajo mišljenje, ki ga druga bitja ne

poznajo. Toda ali ni v resnici obratno: ali ni najprej naše človeško mišljenje in zavedanje, ki je sicer vedno odprto v svet – toda o svetu v resnici sploh ne bi mogli ničesar reči, če se ga v mišljenju ne bi najprej zavedali? Ali bi lahko zares govorili o resnici, če ne bi bilo človeka, ki sploh ima sposobnost spoznanja? Ali je resnica obstajala v času dinosavrov, ko še ni bilo – vsaj tako predvidevamo – nikogar, ki bi za svet vedel in se njegove resnice zavedal? Znanost bi rekla, da je svet objektivno obstajal tudi takrat, o čemer nam ni treba dvomiti – toda kako govoriti o obstajanju (ali o ‚biti‘) česarkoli, če to ne bi nikoli prišlo do védenja oz. zavedanja? V teh vprašanjih se ne skriva klasično nasprotje med filozofskim materializmom in idealizmom, temveč vprašanje, kaj je resnično izhodišče oz. od kod mišljenje resnično izhaja. Če kdo za izhodišče razglasi nekaj, kar vsebuje neprepoznane in nereflektirane predpostavke, ga lahko upravičeno označimo za naivnega. Prav naivnost je srž kritike, ki jo fenomenologija naslavlja znanosti (Husserl 1965, 18–19).

Težava pa ni zgolj znanstvena predpostavka ‚objektivnega‘ sveta, pri kateri naj bi iz enačbe lahko črtali ‚subjektivno‘ mišljenje človeka, temveč tudi razumevanje odnosa med človekom in mišljenjem. Ker človeka najprej razumemo kot telesno bitje, torej nekakšno prostorsko ‚pod-stat‘ (lat. *sub-stantia*), je mišljenje dejavnost, ‚pod‘ katero ‚stoji‘ nosilec, ki se v latinskem jeziku imenuje tudi ‚*subiectum*‘. Prav tako je beseda človek ‚substantiv‘, torej samostalniček – in kot pove njen pomen, gre za nekaj oz. nekoga, ki najprej ‚sam stoji‘, nato pa izvaja še različne dejavnosti. Te dejavnosti izražamo z besedno vrsto glagol: skozi njegovo stavčno vlogo povedka oz. predikata. ‚Človek‘ je samostalniček, ‚misli‘ je glagol. Stavek »človek misli« vsebuje človeka kot subjekt (osebek), ki se mu pri-reka (predicira) glagol misli. V tej konstrukciji se zdi, da je mišljenje nekaj, kar je človeku pridano ali dodano. V aristotelsko navdahnjeni filozofiji so vsi predikati nekaj, kar se izreka nosilcu, ki je kot podstat osnova za vsako izrekanje. Mišljenje se zdi podrejeno človeku – kakor da bi človek najprej bil in kot dodatek bil zmožen še misliti (in bi si ga lahko predstavljali tudi, ko oz. da ne bi mislil). Kaj pa če je takšen pogled podobno naiven, kot je naivna znanost, ki se svojih predpostavk ne zaveda? Kaj pa če se človek kot on sam ‚vzpostavi‘ šele v svojem mišljenju oz. razumevanju? Tudi tu moramo dati prednost prvoosebni perspektivi, ki je tretjeosebnemu pogledu predhodna. Ali sem jaz sam najprej moje telo, v možganih katerega se oblikuje moje mišljenje in razumevanje sebe? Ali pa sem jaz sam najprej izvrševanje mišljenja, znotraj katerega se zavedam in zasnujem razumevanje samega sebe, vključno s svojim telesom? V tem primeru gre za obratno situacijo: najprej glagol(nik) mojega lastnega izvrševanja v mišljenju in šele znotraj te dejavnosti oz. izvrševanja zasnova ‚samostalnika‘ sebe samega (kot kakšno bitje – subjekt, osebek – sebe namreč razumem v mišljenju). Seveda je to zavedajoče se mišljenje vedno že utelešeno in odprto v svet, ki me obdaja. Toda tudi o svojem telesu in o svetu lahko govorim le, če najprej nastopata v mojem zavedajočem se mišljenju in razumevanju. To prvoosebno izkustvo, da svoje življenje (eksistenco, svoje ‚biti‘) najprej vršimo na dinamičen način dogajanja in šele nato vzpostavljamo (statično) razumevanje sebe kot navzočega bitja, se morda zdi kontraintuitivno, vendar je to posledica dejstva, da smo v vsakdanjem življenju kot ljudje vedno ‚zunaj sebe‘ – pri stvareh sveta, s katerimi se ukvarjamo – in da povratno tudi sebe razlagamo na po-

doben način. Pri tem pa naivno spregledujemo, kar nam je najbolj ,notranje' in najbližje: izvrševanje lastnega mišljenja in zavedanja.

Ta obrat k prvenstvu živetega življenja (kot glagolnika) ,od znotraj', h kateremu neločljivo spadata mišljenje in zavedanje, predstavlja izhodišče za fenomenologijo, kot jo je zasnoval Husserl in jo z novimi poudarki nadaljeval Heidegger. Heidegger na začetek ne postavlja česa samostalniškega, temveč (,glagolniško') ,izvrševanje' sebe. To izvrševanje svojega življenja oziroma svojega ,biti' se pri človeku vrši na različne načine (oz. v različnih modusih), od katerih so nekateri temeljni oz. konstitutivni. Načini ali modusi človekovega eksistiranja – Heidegger (1986, 44) jih imenuje ,eksistenciali' – niso *lastnosti* bitja po imenu človek (samostalniški pristop), temveč *načini* izvrševanja njegovega biti (glagolniški pristop). Zato se Heidegger samostalniškim pojmom, kot sta ,zavest' ali ,razum', izogiba – ti so namreč ustvarjeni drugotno in jih tvorimo na podlagi prvotno glagolniškega ,zavedanja' ali ,razumevanja'. V svoji eksistencialni analitiki v *Biti in času* Heidegger kot temeljne načine človekove ,odprtosti' (v najširšem smislu besede) opredeljuje razumevanje, počutje in govor (Heidegger 1986, 133). Človek je sebi, soljudem in svetu ,odprt' na način (razumevajočega, čutečega in govorečega) ,védenja' zase, za druge in za svet. To odprtost bi lahko povezali tudi z mišljenjem. Mišljenje je sicer prvenstveno vezano na razumevanje in govor, toda vsi trije eksistenciali se dogajajo hkrati in so neločljivo povezani. Pri tem je ključnega pomena dejstvo, da razumevanje ni nekaj, kar bi bilo človeku dodano, pač pa človek *skozi* razumevanje vrši samega sebe (razumevanje oz. mišljenje kot način biti). To, kako razumevamo sebe, druge in svet, je *hkrati* tudi način, kako živimo in smo. Ali če to povežemo z mišljenjem: način našega mišljenja je način našega življenja. In obratno: vsaka sprememba našega načina biti se obenem zrcali v spremembi našega mišljenja oz. razumevanja.

Tako mišljenje dobi bistveno usodnejšo vlogo, kot jo ima v običajnem empiričnem (naturalističnem) pogledu na človeka. Mišljenje in razumevanje nista le dodatek človekovemu življenju, zato tudi ne moreta postati zgolj njegovo orodje. Včasih rečemo, da se mora človek ,poslužiti' mišljenja ali da mora mišljenje in razum prav ,uporabljati', pri čemer se zdi, da obema pripisujemo instrumentalno vlogo. Toda orodje lahko obstaja šele tam, kjer je postavljen cilj, ki mu služi – in kdo naj postavi cilj, če ne mišljenje samo? Zato je npr. vsak poskus evolucijske biologije, da bi mišljenje razložila kot komparativno prednost človeka, ki služi njegovemu preživetju, problematična redukcija, ki se pozabi vprašati, v kakšnem mišljenju se je takšna teorija sploh lahko oblikovala.⁴

2. Računajoče mišljenje in umetna inteligenca

Ta dolgi uvod je bil potreben, da bi znali pravilno pristopiti k analizi tistega mišljenja, ki je svoj zmagoviti pohod pričelo z razvojem moderne znanosti in svoj

⁴ Mimogrede – biologi niso tisti, ki nam najprej povedo, kaj je življenje. Da se biologija po življenju sploh lahko znanstveno vpraša, nam mora biti življenje že znano iz našega neposrednega izkustva oziroma iz našega življenjskega sveta (Pöltner 1993, 186–187).

vrhunec doseglo v pojavu umetne inteligence. Beseda inteligenca sicer prihaja iz latinščine in je povezana z besedo *intellectus*, tj. (raz)um. Intelpekt kot um je veljal za tisto lastnost človeka, ki ga razlikuje od vseh drugih bitij in ga nad njih tudi povzdiguje. Izvrševanje uma oz. *intelligere* lahko prevajamo tudi kot mišljenje – seveda takšno, ki se razlikuje od vsakdanje raztresenosti in sledi (logičnim) pravilom mišljenja. Besedno zvezo 'umetna inteligenca' bi lahko prevajali tudi kot 'umetno mišljenje', pri čemer 'umetno' v klasičnem smislu pomeni to, kar je delo človeka kot rezultat njegove veščine oz. 'umetnosti' (lat. *ars, artis*, od tod pridevnik *artificialis*). Umetno se razlikuje od tistega, kar je naravno oz. dano po naravi (lat. *natura* in *naturalis*). Pri umetni inteligenci gre torej za s človeškim delom proizvedene oziroma 'umetne' stvaritve, ki se vedejo tako, da jim lahko pripisujemo 'mišljenje'. Seveda se ob tem takoj pojavi vprašanje, kako mišljenje razumemo oz. kaj pomeni misliti.

Z vzponom nove metode v znanosti, za katero je v veliki meri zaslužen Galileo Galilei, se je zelo povečal vpliv posebne znanstvene discipline, namreč matematike (Husserl 1954, 20). Ob tem je zanimivo, da matematika ni naravoslovna veda – pravzaprav sploh ni veda, ki bi imela za svoj predmet karkoli 'materialnega', temveč se ukvarja s 'čistimi formami'. Zato je matematika, podobno kot logika, čista *formalna* veda. Form, ki jih obravnava matematika (od števil in njihovih odnosov v aritmetiki do prostorskih likov v geometriji), ne najdemo v naravi, čeprav nam naravni svet nudi priložnost, da jih iz njega miselno 'abstrahiramo' (kot npr. iz polne lune lahko abstrahiramo formalni pojem kroga). Matematične forme in njihove zakonitosti so torej stvar mišljenja, se 'nahajajo' v mišljenju – in dostop do njih ima samo misleči um (upravičeno lahko domnevamo, da volk iz polne lune ni zmožen abstrahirati pojma kroga).

Matematiko so poznali in razvijali že Grki – pravzaprav je odkritje matematičnih zakonitosti, ki so neodvisne od čutnega (fizičnega) sveta, pomembno prispevalo h grškemu odkritju in vzponu metafizike –, toda matematika pri njih ni postala osnova za spoznavanje sveta in tudi ni predstavljala osnove za mišljenje. To se spremeni z Galilejevim pristopom, ki verjame, da je matematika resnično ozadje stvarnosti (sveta, narave) in načina, kako ta deluje. Če vzamemo za primer neki naravni pojav, npr. padec predmeta z višine, Galilej *vnaprej* verjame oz. *predpostavlja*, da ta pojav oziroma njegovo delovanje lahko 'matematiziramo'. To pomeni, da pojav 'razstavimo' (analiziramo) na posamezne komponente (npr. masa, lega, hitrost, čas) in med njimi vzpostavimo matematične odnose (pospešek je odvod hitrosti oz. drugi odvod lege po času). Da pa matematiko sploh lahko uporabimo, je treba komponente pojava podvreči merjenju. Če nekaj ni merljivo (tj. če ni zmožno pretvorbe v količino oz. ni kvantifikabilno), v matematičnih relacijah ne more nastopati – zato je za Galileja ključna zahteva nove znanstvene metode, da tudi tisto, kar doslej ni veljalo za (ali se zdi, da po svojem bistvu ni) merljivo, napravimo merljivo. Ob tem ne gre pozabiti, da sam pojem merila ni nekaj 'naravnega' (v smislu materialnega obstoja), temveč nekaj 'miselnega' in formalnega. Če na primer izmerimo premer jajca, vemo, da v samem jajcu premera ni nikjer. Še več, izračunamo lahko celo povprečni premer več jajc, toda nobeno od konkre-

tnih jajc temu premeru ne ustreza. Merilo je torej nekaj, kar postavimo pred ali za naravni predmet in mu tako določimo količino (Heidegger 1987, 134). Po podobni logiki moderna znanost za pojave postavlja matematični zasnitek (hipotezo, podmeno, ki jo na podlagi matematike zasnuje naše mišljenje), nato pa v eksperimentu preverja, ali se narava res vede tako, kot predvideva naš (matematizirani) zasnitek. Tisto, kar je Galileja najbolj navdušilo in zanj imelo malodane religiozni pomen, je bilo spoznanje, da se narava v svojih pojavih dejansko ravna po matematičnih zakonitostih oz. da formalni zakoni matematike, ki jih odkrijemo v našem umu, delujejo v materialnih procesih narave. Da torej matematično mišljenje (ki je neodvisno od narave) dejansko vlada tudi naravi. Obenem pa skozi takšno spoznanje, skozi takšno znanost lahko naravi zavladamo tudi mi. Moderna znanost, ki je Galilejevo metodo povezala z vse več zbiranja izkustvenih (tj. empiričnih) podatkov in z induktivno metodo – ki jo je (sicer brez uporabe matematike) zasnoval Francis Bacon – je vedno bolj dokazovala, da naravo obvlada – in jo lahko podredi ciljem, s katerimi bo človek svoje življenje izboljšal in tako zmanjšal svojo naravno odvisnost. Uspeh, ki ga je sodobna znanost skozi svoj tehnološki razvoj dosegla, je pri ljudeh in v celotni družbi potrjeval njeno resnico in legitimnost njene metode.

Fenomenologija se je v veliki meri razvila kot odgovor na prevlado metode sodobne znanosti – ker ta za samoumevnega razglša pristop k stvarnosti, ki vnaprej postavlja pogoje, kako se narava in vsi fenomeni sploh lahko pojavljajo (gre torej za pogoje možnosti oz. za transcendentalne pogoje v Kantovem smislu). Pomembno vlogo pri matematizaciji narave je imel Galilejev sodobnik Descartes, čigar koordinatni sistem lahko služi kot primer računajočega mišljenja v moderni znanosti.⁵ Kartezijanski koordinatni sistem namreč postavimo pod prostorske fenomene in nato njihovo lego določimo matematično. Ob tem je razvidno, da je kartezijanski prostor človekova substrukcija (podgradnja, Husserl 1954, 37),⁶ ki jo ta postavlja pod fenomene in jih prisili v matematični izraz. Celotna metoda moderne znanosti je tako bolj kot ne substrukcija, ki že vnaprej določa, na kakšen način se kaj mora pojaviti, da bo imelo znanstveni status. Heidegger pogosto navaja Nietzschejeve besede, da v novem veku nimamo opravka z zmago znanosti, temveč »z zmago znanstvene metode nad znanostjo« (Heidegger 1987, 176). Znanstvena metoda je prežeta z matematičnostjo in uteleša računajoče mišljenje. Ta metoda je postala kriterij znanstvenosti – in nobena veda, ki danes te metode ne uporablja, se ne more zares imenovati znanost.

Preden se posvetimo umetni inteligenci kot zadnjemu stadiju računajočega mišljenja, opozorimo še na to, kaj pomeni v znanstveni metodi spoznati ali razumeti. Če je za filozofsko spraševanje od Grkov dalje veljalo, da je v ospredju vprašanje, kaj nekaj je, pa moderna znanost vprašanje o kaj prevede v vprašanje o kako.

⁵ Z Descartesovim odkritjem analitične geometrije je prišlo do povezave dveh dotlej povsem različnih disciplin matematike (aritmetike in geometrije), kar je omogočilo možnost prevajanja prostorskih fenomenov v števila in algebrske operacije. Tako npr. lego v evklidskem prostoru lahko prevedemo v izraz števil (x, y, z) . Husserl govori tudi o »aritmetizaciji geometrije« (Husserl 1954, 44).

⁶ Heidegger tu uporablja izraz »supozicija« (*Supposition*), »pod-stavka« (Heidegger 1987, 37).

Vprašanje o ,kaj' namreč implicira vprašanje po ,bistvu' nečesa, kar pa le stežka izrazimo na merljiv in računajoči način. Če se vprašamo po tem, ,kaj' je elektron, bo znanost dala odgovor na način, da razloži, ,kako' se elektron vede oz. kako elektron ,funkcionira'. Že Descartesov sodobnik Leibniz je ugotavljal, da stara znanost sprašuje po substanci, nova pa po funkciji. ,Bistvo' elektrona v fiziki ni smiselno vprašanje – na vprašanje, »kaj elektron je«, fizika odgovarja skozi matematično izražene funkcije (delovanja elektrona). Zdi se, da se s tem odgovorom zadovolji tudi človek: za spoznanje resnice nečesa je zadosti spoznati način, kako to deluje – pri tem ni treba iskati kakšne globlje utemeljitve. Gre za pragmatično razumevanje resnice, ki globlje filozofsko spraševanje odpravi z očitkom nepotrebne in neznanstvene metafizične spekulacije.

Pri sodobni umetni inteligenci imamo opravka s stopnjevanjem računajočega mišljenja do te mere, da tovrstno mišljenje ni usmerjeno zgolj v obvladovanje zunanjega sveta, temveč se povratno usmeri tudi v mišljenje samo – v njegovo ,notranje izvrševanje' pri človeku. Prihaja do substrukcije, ki ,pod' človekovo mišljenje postavlja matematično-strukturni model. Mišljenje je treba razstaviti v ,procese', v katerih se zrcali način ,funkcioniranja' mišljenja. Prevod miselnih procesov v formalni (logično-matematični) jezik in tehnološki razvoj naprav, ki je idejo orodij za računanje, poznanih že v preteklosti, nadgradil do neverjetnih zmogljivosti digitalne tehnologije, sta postavila osnove za sodobno umetno inteligenco. Pri prehodu iz analogne v digitalno tehnologijo lahko prepoznamo novo zvrst matematizacije. Vsako informacijo (podatek) je mogoče pretvoriti v število. Toda ne način konvencionalnih števil, kot jo uporabljajo analogna merila, pač z redukcijo na bite oziroma na niz binarnih števk (angl. *digits*) 0 ali 1. Niz osmih bitov tvori en bajt, kar pri digitalnem zapisu predstavlja najpogostejšo pomnilno enoto. Vsak podatek na digitalnem mediju je shranjen kot število v dvojiškem sistemu števk 0 in 1, ta poenostavitev v dvojiški sistem pa omogoča različne tehnološke rešitve shranjevanja in obdelovanja podatkov – npr. uporabo električnega naboja, ki je lahko pozitiven (1) ali negativen (0). Matematična substrukcija pri digitalni tehnologiji po eno strani prinaša izjemno poenostavitev – redukcijo informacije na zapis z dvema števka – , po drugi strani pa količino digitalnih zapisov izjemno poveča. Pomnilniški mediji za shranjevanje podatkov so v zadnjih desetletjih dosegli izjemen napredek – in razvoj umetne inteligence je postal možen šele na podlagi masovnega povečanja količine podatkov, s katerimi je mogoče operirati. Kvantifikacija kot ena od bistvenih karakteristik moderne znanosti je danes dosegla neslutene razsežnosti.

Ob tem se še vedno nismo dotaknili vprašanja, kako digitalne naprave, ki za nas izvajajo inteligentne operacije, ,mislijo'. Glede na znani Turingov test je treba priznati, da rezultatov teh operacij marsikdaj ne moremo ločiti od rezultatov, ki jih proizvede človekovo mišljenje, in da jih računalniki obenem opravijo mnogo hitreje od ljudi. Toda eno so posamezni miselni procesi in operacije, drugo pa celovito razumevanje mišljenja, kot ga srečamo pri človeku. Po začetni evforiji očetov umetne inteligence, ki so verjeli, da bo njen razvoj kmalu dosegel raven človekovega mišljenja, je prišlo do zastoja, h kateremu je med drugim prispeval tudi takrat

prevladujoč koncept ‚simbolne‘ umetne inteligence. Šlo je za model, ki podobno kot racionalistični pristop Descartesa pomen stvari zvede na njihove definicije oz. predstave (mentalne reprezentacije) in mišljenje razume kot preračunavanje različnih (teoretskih) možnosti. Do takšnega modela je bil kritičen filozof Hubert Dreyfus (1929–2017), ki se je navdihoval pri Heideggerjevi eksistencialni analizi in Merleau-Pontyjevi fenomenologiji telesnosti. Heidegger namreč opozarja, da človekovo razumevanje sveta ne poteka tako, da si za stvari v svetu (npr. kaj je miza, krožnik, pribor itd.) najprej napravimo (teoretske, simbolne) pomene in šele potem skozi njihovo povezovanje sprejemamo odločitve za naše delovanje. V resnici je obratno – šele iz našega praktičnega delovanja (npr. rokovanja s priborom ob uporabljanju krožnika na mizi) se oblikujejo (predteoretski) pomeni posameznih stvari, ki so vedno vpeti v širši kontekst celovite človekove prakse. Človeku lastna prednost praktičnega pomena (*vedeti, kako*) pred teoretskim pomenom (*vedeti, kaj*) računalniškim programom, ki temeljijo na simbolnem modelu, onemogoča da bi lahko posnemali človeško inteligenco, ki se kot praktično razumevanje vrši v utelešeni umeščenosti človeka v konkretno situacijo sveta. Znamenita Dreyfusova knjiga *Česa računalniki ne morejo delati: Kritika umetnega uma* (Dreyfus 1972) računalnikom inteligenco odreka, ker človekovega védenja (inteligence) ni mogoče povsem (teoretsko) formalizirati – zato bi bili računalniki lahko inteligentni šele tedaj, če bi kot človek imeli telo in bi bili preko njega umeščeni v dogajanje sveta (1967). S svojo kritiko je Dreyfus spodbudil iskanje alternativnih modelov za umetno inteligenco, ki so dobili ime ‚Heideggrovska umetna inteligenca‘ (Preston 1993; Dreyfus 2007) in si prizadevajo za takšne načine inteligentnega procesiranja, ki bi znali posnemati ‚umeščeno utelešenost‘ (angl. *embedded-embodied*) človekovega mišljenja. Ta smer umetne inteligence sicer ni dosegla večjega preboja, zlasti v primerjavi z izjemnimi uspehi tehnologije GPT (angl. *Generative Pretrained Transformer*) – kljub temu pa se ji priznava pomembno vlogo pri uporabi umetne inteligence na področjih, ki zadevajo globlje ‚razumevanje‘ konteksta in smisla (družbene interakcije, umetnost, etične odločitve etc).⁷

Dreyfusov poskus uporabe Heideggerjeve eksistencialne analize za optimizacijo procesov umetne inteligence Heideggerju ne dela nikakršne usluge. Nasprotno, Heidegger si verjetno ne bi mogel predstavljati, da bo njegova filozofija nekoč stopila v službo tehnike in postala uporabna za računajoče mišljenje. Toda povsem v skladu s tem, kar je Heidegger v svojih kritičnih spisih o znanosti in tehniki že slutil, smo danes prišli do stanja, ko so vsi teksti, ki jih je napisal katerikoli človek, postali ‚zaloga‘ informacij, ki jih sistematično povzema in obdeluje že omenjena tehnologija globokega učenja. Če Heidegger na znanem primeru hidroelektrarne na Renu pokaže, kako je človek reko spremenil v zalogo energije (Heidegger 2000, 16), je zdaj storjen še korak naprej. Ne le narava, temveč tudi stvaritve človekovega duha so postale zaloga – zaloga podatkov. Tehnologija GPT besedila, slike, glas-

⁷ ChatGPT na vprašanje glede razlike med svojim pristopom in Heideggrovsko UI odgovarja takole: »Chat GPT predstavlja praktičen pristop k umetni inteligenci, ki temelji na sodobnih tehnikah globokega učenja in velikih sklopih podatkov. Nasprotno pa Heideggrovska UI teži k globljemu filozofskemu in kontekstualnemu razumevanju človeškega bivanja, kar je veliko bolj kompleksno in manj standardizirano. Pristopa imata različne cilje in uporabe, oba pa prispevata k razvoju umetne inteligence na svoj način.«

bo in druge stvaritve duha z digitalno substrukcijo pretvarja v podatke, ki jih kopiči v ogromnih količinah. Tehnologija je programirana tako, da v podatkih s pomočjo nevronske mreže (,umetnih nevronov') in interakcij njihovih številnih plasti ,prepozna' skupne vzorce, ki služijo doseganju ciljev, ki jih je postavil programer. Pravzaprav je za nas težko razumljivo, da je odgovor, ki ga npr. napiše ChatGPT, posledica konsektivnega nizanja besed, ki si sledijo na podlagi verjetnosti, izračunane s pomočjo izjemne količine obdelanih besedil. Dejansko ta tehnologija deluje povsem v skladu z znanim primerom ,kitajske sobe', s katerim je Searle kritiziral Turingov test in njegovo pojmovanje inteligence (Searle 1980). Čeprav v sobi vhodnemu podatku v kitajščini vedno sledi tudi izhodni podatek v kitajščini, to ne pomeni, da človek v sobi zna kitajsko – zna namreč zgolj uporabljati knjigo z navodili. Po Searlu naj bi bila tudi umetna inteligenca daleč od resnične inteligence. Toda kot smo že videli, odgovor ni tako preprost. Odvisen je od tega, kako inteligenco razumemo (Žalec 2023), kar pa nas vodi nazaj k temu, kako sploh razumemo človeka in njegovo izvrševanje mišljenja oz. razumevanja. Nehote lahko zapišemo v stanje, ko nas tehnični pristop, ki s svojo vseprisotnostjo sodobnemu času nezavedno vlada, povsem posrka vase in začnemo na podlagi njegovih predpostavk razmišljati tudi sami. Zato je toliko pomembnejše, da se vrnemo k izhodiščnemu vprašanju in se posvetimo celovitemu razmisleku, kaj pomeni misliti.

3. Problem odnosa med računajočim mišljenjem (umetne inteligence) in celovitostjo človekove eksistence

Heidegger je pravilno ugotovil, da poglobljena nevarnost moderne tehnike ni v zunanjih grožnjah, kot je npr. atomsko orožje, temveč v njenem ,notranjem' vplivu na človeka (Heidegger 2000, 29). Gre namreč prav za preobrazbo človekovega mišljenja, ki začne samega sebe razumeti na tehničen način. Ta nevarnost je danes zaradi razvoja umetne inteligence večja kot kadarkoli. Človek je namreč fasciniran nad inteligenco stvaritev, ki jih je proizvedel s pomočjo sodobne znanosti in njene matematično-tehnične misli – še zlasti ker te stvaritve njegove zmoglosti in moči pogosto presegajo. Občudovanje tehnične inteligence gre lahko tako daleč, da človek skozi njene kriterije presoja tudi samega sebe in zaključi, da bo umetna inteligenca nekoč – za nekatere celo kmalu – človeka presegla tako zelo, da bo nad njim zavladovala. Pravzaprav je najbolje, da poskuša človek sebe preseči kar sam – tako, da si prizadeva svoje delovanje izboljšati (angl. *enhance*) s čim večjim vključevanjem tehničnih in inteligentnih priprav v svoje življenje, morda celo kar v svoje telo (Štivic 2023). Tako bo človek stopil ,čez' (*trans*) sedanjo stopnjo biološkega razvoja v smer transhumanistične optimizacije (Klun 2019; Stegu 2019). Takšna idealizacija umetne inteligence je podobna Feuerbachovemu očitku religiozne sublimacije, po kateri si človek naredi ideal Boga, v odnosu do katerega sebe doživlja kot majhnega, nepopolnega in nevrednega – in se na ta način sebi odtuji. Podobno odtujitev lahko opazujemo pri idealizaciji računajočega mišljenja,

ki postane kriterij za razumevanje in presojanje človeka. Ob tem ni treba nujno pomisliti na transhumanizem, dovolj je že misel na strukture tehnično urejenega življenja v družbi, npr. na izobraževalni sistem, ki računajoče mišljenje favorizira kot ideal in kriterij pravega védenja. Toda obenem se zdi, da ob vsem tehnološkem napredku in čarih umetne inteligence pri sodobnem človeku raste nelagodje ali celo nezadovoljstvo: kakor da v sebi čuti razdvojenost ali celo odtujenost. Heidegger namesto besede odtujitev uporablja besedo »zapadlost« (Heidegger 1986, 166) in »nesamolastnost« (nepristnost, neavtentičnost) (128). Njegovo razlikovanje med samolastnostjo in nesamolastnostjo lahko uporabimo tudi pri mišljenju. Ali je računajoče mišljenje za človeka avtentično in pristno mišljenje? Na to ne moremo odgovoriti, če ne poskušamo najprej razumeti, kdo kot ljudje sploh smo, tj. če ne razumemo naše eksistence.

Eksistencialna analiza, ki skuša opisati človeško eksistenco in njene fenomene, kot se nam dajejo in kažejo iz sebe, razkriva, da naše od znotraj živete eksistence (ki je sicer vedno odprta v zunanji svet) ne moremo razložiti na podlagi samostalniških izhodišč in matematičnih relacij – podobno kot iz zunanjega opazovanja človeka in delovanja njegovega telesa ne moremo sklepati o vzročnih procesih, ki človeka vodijo v notranjosti. Začnemo lahko le v našem prvoosebnem življenju (eksistenci) od znotraj, ki je do sebe in do sveta ,odprto‘ na način razumevanja, počutja in govora. Mišljenje lahko povežemo z eksistencialom razumevanja, ki pa je s počutjem in govorom neločljivo povezan. Človekovo mišljenje se nikoli ne vrši, kot se ,nevtralni‘ procesi umetne inteligence, temveč ga nujno spremljajo različna občutja in počutja. Zato človekovo mišljenje ne more biti nikoli ,ravnodušno‘, temveč vključuje ,vzdušja‘ in razpoloženja, v katerih se dogaja – in ki nanj tudi vplivajo. Tudi nevtralna ravnodušnost je za človeka eden od načinov počutja. Podobno kot s počutjem je človekovo mišljenje zaznamovano tudi z razsežnostjo govornosti. Platon pravi, da je mišljenje »pogovor duše s seboj«. Misliti za človeka zato nikoli ne pomeni zgolj vršiti miselne procese, ki jih lahko prevedemo v formalni jezik – v mišljenju in skozi njega človek na celovit način izvršuje svoje življenje oz. eksistenco.

Podobno kot človekovega mišljenja ne moremo ločiti od počutja, v katerem poteka, ga tudi ne moremo ločiti od človekove telesnosti. Res je, da so nekateri filozofi zapadli dualizmu med dušo in telesom, pri čemer so dejavnost mišljenja pripisali duši in telo ponižali na golo materijo ali celo na kompleksen tehnični mehanizem (npr. Descartes). Toda takšne spekulativne konstrukcije so daleč od izkustva, pri katerem doživljamo ,utelešenost‘ našega mišljenja. Človek misli na utelešen način. Res je, da v mišljenju po eni strani svoje telo presega – po drugi strani pa občutje telesa njegove misli vedno znova ,privede nazaj‘ v konkretno situacijo, v kateri je telesno prisoten. Človekova utelešenost v sebi dejansko razkriva določeno dvojnost, ki jo fenomenologija telesnosti povzema v dve različni, toda neločljivo prepleteni izkustvi: *jaz sem telo* in *jaz imam telo* (Pöltner 2006, 70). Svojega telesa, ki ga primarno doživljamo ,od znotraj‘, ne moremo nikoli odmisлити, toda po drugi strani do svojega telesa (ki ga prav tako izkušamo ,od zunaj‘) vzpostavljamo misleč odnos in ga lahko razumemo na različne načine. Heidegger

v zgodnji eksistencialni analizi telesnosti ni posvetil ustrezne pozornosti in tudi v poznejših delih je ta tema ostala v ozadju⁸ – nasprotno pa fenomenologija telesnosti stoji v ospredju prizadevanj Merleau-Pontyja (2006). Slednji skuša pokazati, da naša (zavedajoča, misleča) odprtost v svet vedno poteka na utelešen način, pri čemer telo mišljenje vedno že prehiti – in kot živeto ‚meso‘ naše doživljanje sveta vnaprej vodi in strukturira. Dreyfus (1967) je skozi povezavo med Heideggerjem in Merleau-Pontyjem pokazal, da umetna inteligenca ne more nikoli doseči in nadomestiti mišljenja človeka, ki je preko svojega telesa v aktivni interakciji s svetom – zato bi lahko računalniki človekovo mišljenje in njegovo inteligenco posnemali šele, če bi imeli telo, ki bi jih umeščalo v svet.

Kljub utelešenosti pa človekovo eksistenco zaznamuje odprtost, preko katere človek svojo umeščenost v tukaj in zdaj presega. Že sama odprtost prinaša pojem ‚ekstatičnosti‘ človekove eksistence, ki je ne moremo omejiti na njeno statično navzočnost, pa tudi ne na njen dinamični potek. V svojem eksistiranju človek ‚stoji ven iz sebe‘ (lat. *ex-stare*) in prav zato z *okoljem* ni zgolj obdan, temveč *svet* ima. Na neki način je človek ekstatičen že v svoji telesnosti, ker npr. s svojim vidom presega meje svojega fizičnega telesa (Heidegger 1987, 112–113). Toda na poseben način je človek ekstatičen v svojem zavedanju in posledično v mišljenju. Človek je ‚v mislih‘ pogosto v prihodnosti, pa tudi v preteklosti. Tu je treba vpeljati fenomenološki razmislek o odnosu med človekovo eksistenco in časom. Eksistenca kot glagolnik (iz glagola *eksistirati*) pomeni, da človekovega življenja ni mogoče ločiti od časa. Še več, samostalnik ‚čas‘ tvorimo šele potem, naše prvo izkustvo pa je ‚časenje‘ naše od znotraj živete eksistence. Analizo tega časenja Heidegger (1986, 326–331) opisuje s tremi časovnimi ‚ekstazami‘ (prihodnost, sedanost, bivšost), ki pa jih človek vrši *hkrati* in ne v smislu zaporedja (zunanjega objektivnega časa). Človek se časi na način, da ‚stoji ven iz sebe‘, ko ‚vnaprej‘ snuje svojo prihodnost, ko gleda ‚nazaj‘ na svojo bivšost – in ko je ‚zraven‘, ko se ukvarja s sedanjimi stvarmi in opravi. Čeprav so za človekovo eksistenco konstitutivni vsi trije časovni modusi, pa ima posebno vlogo prihodnost, saj predstavlja obzorje možnosti in svobode. Človek ni namreč nikoli zgolj to, kar je bil do sedaj in kar je sedaj (kot dejanskost), temveč tudi to, kar še ‚lahko‘ bo (kot možnost). Človekovo eksistenco na poseben način zaznamuje to, kako človek samega sebe vidi v prihodnosti, kakšnega *sebe* ‚ima v mislih‘. Husserl za opis človekovega zavedanja uporablja pojem intencionalnosti. Ta pojem prvenstveno označuje ‚ekstatično‘ naravo zavesti, ki je ‚zunaj sebe‘, pri tem, česar se zaveda – toda intencionalnost vključuje tudi časovno konotacijo (intenca kot namen), povezano s prihodnostjo. Človekovo misleče zavedanje oz. razumevanje (v Heideggerjevem smislu) vedno vključuje prihodnostno razsežnost: »imeti nekaj v mislih«, misliti nekaj, specifično vključuje »biti-si-vnaprej« (327).

Ob tem se razodeva nova in ključna razlika med človekovim mišljenjem in umetno inteligenco. Čas umetne inteligence je objektivni čas zaporedja trenutkov, ki

⁸ V *Biti in času* priznava, da telesnost »v sebi skriva lastno problematiko, ki tu ne bo obravnavana« (Heidegger 1986, 108). Kasneje to tematiko obravnava pri seminarjih v Zollikonu (1987), vendar tudi tam odnos med človekovim razumevanjem (mišljenjem) in njegovo utelešenostjo ni zadovoljivo pojasnjen.

jih lahko matematično prevedemo v točke in količine na časovni premici. Ta linearno tekoči čas kot edino resnični časovni modus priznava sedanost, kajti preteklosti v strogem smislu *ni* več, prihodnosti pa še *ni*. Samo sedanost zares *je*. Za umetno inteligenco ostaja zgolj sedanje procesiranje. Prihodnost in preteklost sta v mislih programerja, pri čemer je računalniško programirana prihodnost zgolj podaljšek sedanosti, ki že vsebuje navodila oziroma algoritme za prihodnje delovanje. Če je za človekovo inteligenco ključna intencionalnost – človek si v mišljenju mora biti ‚vnaprej‘, da njegovo delovanje označimo za inteligentno –, pa za umetno procesiranje obstaja zgolj sedanost, ko enemu trenutku vzročno sledi naslednji. Ko ChatGPT odgovarja inteligentno, ‚nima ničesar v mislih‘. Od zastavitve vprašanja, ki ga prejme, se izvrši ogromno digitalnih procesov, ki po statističnem računu trenutek za trenutkom proizvedejo novo besedo, ki sledi prejšnji, in tako se na konsekutiven način – povsem brez intencionalnosti – sestavi končni odgovor. Zaradi intencionalnosti mišljenja človek ve, da misli, nasprotno pa računalnik tega ne more vedeti. Sokrat ve, da nič ne ve, umetna inteligenca pa ne ve, da ve. Inteligentni produkt sicer opravičuje rabo tega pridevnika tudi za proizvajalca, v tem primeru računalnik, toda očitno je ‚inteligentna‘ dejavnost slednjega (‚*intelligere*‘ kot ‚misliti‘) povsem drugačna kot človekova.

Seveda zagovorniki umetne inteligence hitijo odgovarjati, da človekovo mišljenje ni tako zelo drugačno, kot ga poskuša opisati fenomenološka eksistencialna analiza. Opirajoč se na naturalistična izhodišča bi argumentirali, da ni bistveno, kaj ‚ima človek v mislih‘, temveč kaj njegova sedanja dejanja določa in vodi – dejanja so edino empirično preverljiva. Kakor so edino empirično preverljivi procesi v možganih, ki se vedno vršijo v sedanosti in določajo človekovo delovanje. Naturalizem kot samoumeven in splošno sprejet znanstveni nazor bi človekovo doživljanje počutij, utelešenosti, časa, prihodnosti, možnosti in svobode razglasil za subjektivna stanja, ki jih je treba podrediti objektivnemu empiričnemu pristopu. Ob tem naletimo na enak problem, s katerim se je soočal že Kant: kako ‚dokazati‘, da je človek svoboden? Pri eksistencialni analizi je ta svoboda povezana s človekovo ekstatičnostjo (nasproti statični določenosti) in možnostjo samolastnega mišljenja. Če pa celotno dogajanje človeka tvorijo zgolj naravni telesni procesi (zlasti v možganih), ki si sledijo po vzročni nujnosti, potem je človekova svoboda zgolj njegova subjektivna iluzija. Kant ta ugovor pozna, obenem pa se zaveda, da bi zanižanje človekove svobode zamajalo vse tisto, kar je najbolj človeškega, namreč njegovo moralnost. Zato Kant svobodo razglasi za ‚postulat‘ praktičnega uma – torej za nekaj, kar zahteva naša praktična ‚intelligenca‘, pa čeprav obstoja človekove svobode ni mogoče empirično dokazati (Kant 1983, 264). Svoboda namreč ni empirično dejstvo. In tudi ne more biti: če bi bila, potem bi prenehala biti svoboda. Kant verjame, da je človek državljan dveh svetov: materialnega sveta narave, ki mu pripada po svoji telesnosti, obenem pa tudi sveta ‚svobode‘, do katerega ima dostop preko svojega uma in mišljenja. Tudi pri fenomenološkem pristopu se zastavlja vprašanje, čemu bolj verjeti: prvoosebnemu izkustvu lastne živete eksistence ali tretjeosebne pristopu empiričnih znanosti. Kaj je bolj resnično: mišljenje, ki ga vršimo in izkušamo v našem življenju ‚od znotraj‘, ali mišljenje, ki je

prevedeno v matematično-tehnični jezik in služi kot podlaga za znanstveno razlago sveta in človeka?

Kontrast med eksistencialno vlogo mišljenja in njegovo logično-tehnično redukcijo pride do izraza tudi pri vprašanju, ki je v človekovem življenju bistvenega pomena, računajoče mišljenje pa je zanj komaj dojemljivo. Gre za človekovo izkustvo ničā, ki se zaostri zlasti pri vprašanju smrti. Ko človek v zavedanju stoji ven iz sebe in se v mišljenju prepušča svoji prihodnosti, se obenem zave, da ima ta prihodnost mejo, ki pomeni konec življenja oziroma eksistence. Med raznolikimi opredelitvami človeka pri Grkih srečamo tudi to, da je človek edino bitje, ki ve za svojo smrt. Platon gre celo tako daleč, da smrt neposredno poveže s filozofskim razmišljanjem. Resnična filozofija zahteva razmišljanje o smrti (*meditatio mortis*). Toda kako razmišljati o tem, kar se zdi čisto nasprotje (negacija) biti in kar človeka vznemirja kot ,nič'? Prejšnjo opredelitev bi lahko formulirali tudi takole: človek je edino bitje, ki ima odnos do ničā. Toda ključno pri tem je prav dejstvo, da ničā ne moremo misliti na način logike – kot pravi Heidegger (1995, 26): »Nič ostaja vsej znanosti načelno nedostopen.« Ob tem očitku si lahko predstavljamo reakcijo računajočega mišljenja, ki je nič(lo) privedlo v matematiko, kjer je dobila pomembno vlogo. Še več, digitalna tehnologija temelji na binarnosti enice in ničle, pozicije in negacije. Toda takšen pristop ničā ne more zares doseči, ker iz njega napravi ,nekaj'. Heidegger bi rekel, da se nič razodene šele v eksistencialnem izkustvu, npr. tesnobe, ko se zavemo oz. začutimo končnost, omejenost in smrtnost svoje eksistence. Ali pa nas nič pretrse ob izgubi ljubljene osebe, ko začutimo prepadno praznino njene odsotnosti. Po Heideggerju je nič dostopen filozofiji in literarni besedi, mišljenju in pesnjenju (*Denken und Dichten*), ki sta računajoče mišljenje moderne znanosti sposobna zapustiti – in naj bi udejanjala to, kar Heidegger imenuje pre-mišljujoče ali meditativno mišljenje. Formalizacija ničā v matematično-tehničnem mišljenju in izziv ničā v eksistencialnem razmišljanju ponazarjata nepremostljivo razdaljo, ki obstaja med umetno inteligenco in celostno vlogo mišljenja znotraj človekove eksistence.

4. Zaključek

Kaj pomeni misliti? Umetne stvaritve v svojem delovanju človekovo mišljenje lahko posnemajo, toda posnemajo ga samo od zunaj. Mišljenje za človeka nikoli ni le izvrševanje miselnih procesov, temveč izvrševanje *sebe*. Človekovo mišljenje vključuje zavedanje sebe – in prav v tem mislečem sebe-védenju je temelj človekove odprtosti. Umetna inteligenca je zase zaprta. Za njene inteligentne procese ve človek, sama pa zanje ne ve. Toda to ni edina razlika med človeškim mišljenjem in umetno inteligenco. Ne gre zgolj za to, da se človek poleg izvrševanja logike mišljenja tega še zaveda, medtem ko umetna inteligenca logiko samo izvršuje. Človek se v mišljenju zave vprašljivosti mišljenja samega, obenem pa tudi vprašljivosti celotne stvarnosti, do katere je kot človek v mišljenju in razumevanju v določenem odnosu. Človekovo mišljenje je zato poklicano k radikalni samorefleksiji in k celovitem

razmisleku. Zato misliti za človeka najprej pomeni, da se takšnemu razmisleku odpre in izpostavi. Tehnično računajoče mišljenje, ki ne obvladuje samo moderne znanosti, temveč način razmišljanja sodobne družbe, je za lastno vprašljivost slepo – ni sposobno razmisliti, v kakšnem miselnem obzorju razume človeka, naravo in družbo. Znanost lastnih miselnih in metodoloških predpostavk ne more proučiti znanstveno. Tako tudi umetna inteligenca računajočega mišljenja, ki jo kot tako sploh omogoča, sama iz sebe ne more postaviti pod vprašaj. Radikalnega spraševanja, ki pa navadno ne doživi znanstvenega odgovora in rešitve, temveč najpogosteje naleti na danosti, ki se moči in obvladovanju izmikajo, je sposoben le človek. Takrat se človek v mišljenju zave, da mu je misliti dano – da je prejemnik tako mišljenja kot tudi nagovora stvarnosti, ki mu daje misliti. Takšno ozavešeno, „premišljujoče“ mišljenje, ki ničesar ne obvlada in je na videz brez koristi, se razodeva kot pristna alternativa prevladujočemu in vseobvladujočemu računajočemu mišljenju. Obenem pa se razodeva tudi kot možnost ohranjanja pristne človeškosti, preden bo ta povsem pozabila nase – če bo sebe razlagala skozi prizmo tehničnih naprav in umetne inteligence.

Reference

- Dreyfus, Hubert L.** 1967. Why Computers Must Have Bodies in Order To Be Intelligent. *Review of Metaphysics* 40, št. 1:13–32.
- . 1972. *What Computers Can't Do: A Critique of Artificial Reason*. New York: Harper & Row.
- . 2007. Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian. *Artificial Intelligence* 171, št. 18:1137–1160. <https://doi.org/10.1016/j.art-int.2007.10.012>
- Heidegger, Martin.** 1959. *Gelassenheit*. Pfullingen: Neske.
- . 1986. *Sein und Zeit*. 16. izd. Tübingen: Niemeyer.
- . 1987. *Zollikoner Seminare: Protokolle – Gespräche – Briefe*. GA 89. Frankfurt: Klostermann.
- . 1995. *Uvod v metafiziko*. Prevedel Aleš Košar. Ljubljana: Slovenska matica.
- . 2000. *Vorträge und Aufsätze*. GA 7. Frankfurt: Klostermann.
- . 2002. *Was heißt Denken?* GA 8. Frankfurt: Klostermann. Slovenski prevod: *Kaj se pravi misliti?* (Ljubljana: Filozofska fakulteta, 2017).
- Husserl, Edmund.** 1954. *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie*. Husserliana 4. Den Haag: Nijhoff.
- . 1965. *Philosophie als strenge Wissenschaft*. Frankfurt: Klostermann.
- Kant, Immanuel.** 1983. *Werke in sechs Bänden*. Zv. 4, *Kritik der praktischen Vernunft*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Klun, Branko.** 2019. Transhumanizem in transcendenca človeka. *Bogoslovni vestnik* 79, št. 3:589–600. <https://doi.org/10.34291/bv2019/03/klun>
- Merleau-Ponty, Maurice.** 2006. *Fenomenologija zaznave*. Prevedla Špela Žakelj. Ljubljana: Studentska založba.
- Merriam-Webster.** 2024. <https://www.merriam-webster.com/dictionary/> (pridobljeno 23. 6. 2024).
- Pöltner, Günther.** 1993. *Evolutionäre Vernunft*. Stuttgart: Kohlhammer.
- . 2006. *Grundkurs Medizin-Ethik*. 2. izd. Dunaj: Facultas Universitätsverlag.
- Preston, Beth.** 1993. Heidegger and artificial intelligence. *Philosophy and Phenomenological Research* 53, št. 1:43–69. <https://doi.org/10.2307/2108053>
- Searle, John.** 1980. Minds, Brains and Programs. *Behavioral and Brain Sciences* 3, št. 3:417–457. <https://doi.org/10.1017/s0140525x00005756>
- Stegu, Tadej.** 2019. Transhumanizem in krščanska antropologija. *Bogoslovni vestnik* 79, št. 3:683–692. <https://doi.org/10.34291/bv2019/03/stegu>

Štivič, Stjepan. 2023. Kiborgizacija: nastanek in razvoj pojma. *Bogoslovni vestnik* 83, št. 4:883–892. <https://doi.org/10.34291/bv2023/04/stivic>

Žalec, Bojan. 2023. Ali je umetna inteligenca inteligenca v pravem pomenu besede? Vprašanje psihičnih značilnosti in splošnosti. *Bogoslovni vestnik* 83, št. 4:813–823. <https://doi.org/10.34291/bv2023/04/zalec>

Pregledni znanstveni članek/Article (1.02)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 757—781

Besedilo prejeto/Received:09/2024; sprejeto/Accepted:12/2024

UDK/UDC: 004.8:14

DOI: 10.34291/BV2024/04/Zalec

© 2024 Žalec, CC BY 4.0

Bojan Žalec

Človeški podobna umetna inteligenca in superinteligence: verjetnost in glavne težave njunega oblikovanja

Human-like Artificial Intelligence and Superintelligence: The Probability and Main Challenges of Their Design

Povzetek: Avtor se ukvarja z vprašanjem verjetnosti nastanka umetne inteligence, podobne človeški, in nastanka superinteligence. Trdi, da ni verjeten nastop nobene. Verjetnost prve zavrne na podlagi argumentacije Erika J. Larsona, ki temelji na bistvenih značilnostih človeške inteligence, kot so splošnost, intuicija, zdravi razum in abdukcija. Verjetnost nastanka superinteligence zavrne na podlagi argumentov François Cholleta, ki poudarja ne-splošnost, situacijskost in kontekstualnost ter zunanostnost inteligence. Avtor poudarja pomembnost pravilnega razumevanja pojma inteligence ter njene ustrezne opredelitve. Napačno razumevanje ima lahko negativne učinke, ki presegajo akademsko področje – vključno z neupravičenimi finančnimi dobički ter škodo pri organiziranju znanstvenega dela (velika znanost, znanost panja) in spodbujanju pogojev za ustvarjalnost.

Ključne besede: verjetnost človeški podobne umetne inteligence, verjetnost superinteligence, splošnost, intuicija, abdukcija, zdravi razum, partikularnost inteligence, zunanostnost inteligence, velika znanost

Abstract: The author deals with the question of the probability of the emergence of humanlike artificial intelligence and the emergence of superintelligence. He argues that no emergence is probable. The probability of the first is rejected, following Erik J. Larson's argumentation based on the essential characteristics of human intelligence. These include generality, intuition, common sense, and abduction. The probability of the onset of superintelligence is rejected based on arguments provided by François Chollet, emphasizing the non-generality, situationality and contextuality, and externalism of intelligence. The author demonstrates the importance of a correct understanding of the concept of intelligence and its definition. Misunderstanding can have negative effects far beyond the academic realm, including unjustified financial gains and damage

in terms of organizing scientific work (big science, hive science), and fostering conditions for creativity.

Keywords: probability of humanlike artificial intelligence, probability of superintelligence, generality, intuition, abduction, common sense, particularity of intelligence, externalism of intelligence, big science

1. Uvod

Glavno vprašanje prispevka je: Ali je oblikovanje umetne inteligence (v nadaljevanju UI), podobne človeški, verjetno?¹ V nadaljevanju bom za trditev, da bo ustvarjena umetna inteligenca, podobna človeški, uporabljal kratico HIT. Ko govorim o umetni inteligenci, podobni človeški, mislim na UI, ki bi bila sposobna vsega, kar zmore človeška inteligenca. Vendar pa je vprašanje resničnosti HIT tesno povezano z verjetnostjo pojava superintelligence (v nadaljevanju SI) – intelligence, ki bi človeško inteligenco močno presegala. Zato bom razpravljal tudi o verjetnosti nastanka SI, ki ga napovedujejo nekateri transhumanisti – na primer Ray Kurzweil, ki napoveduje nastop singularnosti. (Kurzweil 2005) V nadaljevanju bom za trditev, da bo nastopila SI, uporabljal kratico SIT. HIT in SIT sta prepleteni. Po eni strani nastanek SI ni verjeten, če človeški podobna UI ni mogoča; po drugi strani pa bi pojav človeški podobne UI lahko razumno interpretirali kot pomemben dokaz v prid SIT. Tako sta tezi povezani – in argumenti za in proti HIT so posredno tudi argumenti za in proti SIT. Zato je smiselno obe tezi obravnavati skupaj.

V nadaljevanju bom zagovarjal stališče, da sklicevanje na dosedanje dosežke UI, ki so v mnogih pogledih res impresivni, glede na glavno vprašanje prispevka nima velike teže. Navajanje teh dosežkov kot prepričljivih dokazov za HIT in SIT temelji na nerazumevanju človeške inteligence. Zato bom v prispevku kritično razmišljal o razumevanjih (človeške) inteligence, na katerih temeljijo napovedi o verjetnosti obeh tez.

V prvem delu prispevka bom predstavil argumente proti HIT, kot jih je razvil Eric J. Larson v svoji knjigi *Mit o umetni inteligenci* (Larson 2021), v drugem delu pa argumente proti verjetnosti SIT, kot jih je predstavil François Chollet v svojem vplivnem prispevku „Neverjetnost eksplozije inteligence“ (Chollet 2017).

Larson se osredotoča na štiri temeljne in bistvene značilnosti človeške inteligence: splošnost (angl. *generality*), intuicijo, zdravi razum in zmožnost abdukcije. Trdi, da obstoječa UI teh zmožnosti nima in da v okviru trenutne (Turingove) paradigme razvoja UI take UI, ki bi jih imela, ne bomo mogli ustvariti. Prav tako je malo verjetno, da bi te sposobnosti UI razvila sama. Chollet trdi, da je nastanek SI malo verjeten. V svoji argumentaciji poudarja tri značilnosti inteligence: 1. ne-

¹ Prispevek je nastal v okviru raziskovalnega programa „P6-0269: Religija, etika, edukacija in izzivi sodobne družbe“ ter raziskovalnega projekta „J6-4626 (B): Teologija, digitalna kultura in izzivi na človeka osredičene umetne inteligence“, ki ju financira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

-splošnost; 2. situacijskost oziroma kontekstualnost; 3. zunanostnost (angl. *externalism*). Pomen teh izrazov bom pojasnil v nadaljevanju. Prvi dve značilnosti v nadaljevanju zajemam z izrazom partikularnost inteligence. Tretja značilnost partikularnost implicira.

Splošno gledano lahko to razpravo razumemo kot prispevek k obrambi antikartezijske paradigme razumevanja (človeške) inteligence in uma (Žalec 2023b, 105). Za razumevanja in pristope znotraj te paradigme se uporabljajo različni izrazi: aristotelijansko in interaktivno stališče (Miščević 1988), enaktivizem (Gallagher 2017), ekologizem (Gibson 1979; Potrč 1993; 1996, 194–197; 2004, 55–61), eksternalizem (McCulloch 1995, 184ff; Rowlands, Lau in Deutsch 2020; Potrč 1993), teorija stika (Dreyfus in Taylor 2015), utelešeni humanizem in antropologija utelešenja (Fuchs 2021) in še drugi. Za različico antikartezijanstva, ki jo sam zagovarjam, je središčnega pomena Heideggerjev pojem ‚bit-v-svetu‘. Zato ni presenetljivo, da omenjam tudi razumevanje Huberta Dreyfusa (Dreyfus 1991; Žalec 2023b, 19–29; 31–39; 107–108), vendar ga tu ne obravnavam podrobno.

Osredotočam se na inteligenco, podobno človeški. Trdim, da človeški podobna UI ne obstaja in da njen nastanek ni verjeten. Vendar to ne pomeni, da menim, da ne obstaja UI, ki ne bi bila inteligentna v pravem pomenu besede. Nasprotno – prepričan sem, da so obstoječi sistemi globokega učenja inteligentni. Vendar pa moramo razlikovati med človeško – ali človeški podobno – inteligenco in inteligenco kot tako. Človeška inteligenca ni edina vrsta inteligence. Sistemi globokega učenja so inteligentni, vendar njihova inteligenca ni podobna človeški (Žalec 2023a).

Dotikam se tudi družbene razsežnosti nerazumevanja potencialov UI. Ena od posledic tega je ideologija velike znanosti (angl. *big science*), ki je nevarna in škodljiva – med drugim kot negativen dejavnik pri negovanju (človeške) ustvarjalnosti v znanosti, pri razvoju tehnologije in tudi na drugih področjih. Škodljivost te ideologije kaže, da razprava o naravi, možnostih in potencialih razvoja UI ni le ozkega akademskega pomena, temveč je zelo pomembna s širšega – družbenega in moralnega – vidika.

V literaturi lahko najdemo različne druge argumente za verjetnost človeški podobne UI in SI in proti njej, s katerimi se ne ukvarjam – na primer obsežna in temeljita argumentacija Landgrebeja in Smitha (2023), da ustvarjanje ustreznega matematičnega modela človeškega uma presega človeške sposobnosti. To je pri mojem prispevku zagotovo omejitvev. Po drugi strani pa verjamem, da že argumenti, ki jih bom predstavil, sami po sebi tvorijo dobro podlago za razumno zavrnitev verjetnosti HIT in SIT.

2. Bistvene značilnosti človeške inteligence, ki predstavljajo temeljni problem za razvoj človeški podobne UI

Naj razpravo začnem z nekaterimi splošnimi razlogi, zakaj – vsaj v bližnji prihodnosti in glede na trenutno znanstveno in tehnično znanje ter prevladujoči pristop

– ne moremo pričakovati, da bi kdo ustvaril UI, ki bi imela bistvene značilnosti človeške inteligence. Kot take značilnosti lahko omenimo zmožnosti, ki so tesno prepletene (Larson 2021): splošnost (kot nasprotje ozke specializacije za določene naloge in področje), sposobnost razumevanja, intuicijo, sposobnost učenja, izbiro problemov za reševanje, zdravi razum in sposobnost posebne oblike sklepanja, imenovane abdukcija.

Larson trdi, da trenutno nihče zares ne ve, kako ustvariti človeški podobno UI – nihče za to nima ustrezne znanstvene in tehnične zamisli. Namesto znanstvenih argumentov v prid možnosti ustvarjanja človeški podobne UI se mnogi zatekajo k različnim (znanstveno) slabo utemeljenim trditvam, ki širijo prepričanje, da človeški podobno UI bomo ustvarili. Vedno znova se najdejo (ugledne) osebe, ki trdijo, da smo ta cilj (načelno) že dosegli, da ga bomo kmalu dosegli, da bomo (kmalu) priča nastanku UI, ki bo človeško inteligenco močno presešla in tako naprej. Takšni ‚lažni preroki‘ ali promotorji UI ustvarjajo mit o UI (Larson 2023). Tukaj je ‚mit‘ mišljen v negativnem smislu: kot lažno prepričanje, da nekaj obstaja ali je verjetno, čeprav v resnici ne obstaja in ni verjetno. Kar ne obstaja in ni verjetno, je človeški podobna UI in SI. Podobno lahko v omenjenem smislu govorimo o ideologiji UI – to je o lažnem prepričanju, ki ga nekateri (zavestno ali nezavedno, namerano in nenamerno) ustvarjajo zaradi določenih interesov, med katerimi finančni niso zanemarljivi.

2.1 Problem splošnosti in past ozkosti

Nekateri napovedujejo, da ko bomo ustvarili človeški podobno UI, bo ta kmalu ustvarila inteligenco, ki bo človeške sposobnosti močno presežala – to pa bo privedlo do razvoja SI, katere zmožnosti si ne moremo predstavljati. Vendar pa o pravilnosti te domneve obstaja več utemeljenih pomislov. Prvič: človeška raven inteligence obstaja že zelo dolgo, vendar se ni razvila v SI, ki bi človeško inteligenco presežala. Zakaj bi bilo razumno pričakovati, da bi človeški podobni UI to uspelo? Drugič: nihče ne ve, kako bi človeški podobna UI to dosegla. Nihče tega procesa ni znanstveno opisal. Toda kljub temu nekateri napovedujejo, da se to bo zgodilo.

Človeška inteligenca je zmožna intuicije in neodvisne oziroma samostojne izbire problemov. Zato človeška inteligenca ni ozka – ni omejena na reševanje zgolj določenih problemov ali nalog. Larson poudarja, da nihče nima znanstvene predstave, kako ustvariti UI, ki bi bila sposobna intuicije in neodvisne izbire problemov, potrebne za preseganje te ozkosti. Brez tega pa o človeški podobni UI ne moremo govoriti.

Eden od načinov za zagovarjanje teze o SI je sklicevanje na evolucijo: SI se bo v evoluciji, ki proizvaja vedno višje in bolj kompleksne sisteme, zagotovo pojavila. Tudi če ne vemo, kako, lahko na podlagi evlucijskih razlogov trdimo, da se SI sčasoma bo pojavila. Vendar takšna napoved ne more veljati za znanstveno utemeljeno. Eden od znakov njene neznanstvenosti je, da je načelno neovrgljiva. Poleg tega ta napoved ni odvisna od napak pri napovedih. Kurzweil je pojav človeški podobne UI napovedal že za leto 2029, singularnost pa do leta 2045. Toda tudi če se to ne zgodi, lahko zagovorniki SI pri svoji napovedi še vedno vztrajajo – zgodilo

se bo pač nekaj desetletij kasneje. In če se ne bo zgodilo, se bo zgodilo še malo kasneje.²

Nekateri zagovarjajo tezo o SI na podlagi načela (nujnega) preskoka iz količine v kakovost. Neprestano izboljševanje ‚kvantitativnih‘ sposobnosti UI mora sčasoma pripeljati do preskoka v kakovost, do pojava kvalitativno višje ravni UI in tako naprej do SI. Načelo nujnega preskoka iz količine v kakovost so v preteklosti že zagovarjali,³ vendar ga izkustvo ne potrjuje.

Trditve o pojavu človeški podobne UI in SI so bile v preteklosti izrečene že večkrat. Doslej se nobena od teh napovedi ni uresničila. Poleg tega, kot že omenjeno, trenutno nihče nima znanstvene ideje, kako ustvariti sistem, ki bi izkazoval značilnosti človeški podobne inteligence, kot so navedene zgoraj. Brez kakšnega novega temeljnega teoretičnega vpogleda, preboja ali revolucije pa ni možnosti, da bi to dosegli – zagovorniki možnosti ustvarjanja UI v okviru trenutno razpoložljivega teoretičnega okvira znanstvenih argumentov za svojo trditev dejansko ne ponujajo, kar kaže, da gre za ideologijo.

Da bi razumeli izzive in ovire za razvoj človeški podobne UI, se moramo osvoboditi ozkega pojma inteligence, ki vso inteligenco zvaja na (neodvisno) reševanje problemov. Neodvisno reševanje problemov je za pripisovanje inteligence zadosten pogoj, vendar to še ne pomeni, da je zadosten za vsako vrsto inteligence. Je le minimalen – nujen in zadosten – pogoj za inteligenco (Žalec 2023a), kar pa za človeško inteligenco ni dovolj. Človeška inteligenca je več kot minimalna inteligenca. Človeška inteligenca vključuje ne le neodvisno reševanje problemov, ampak tudi neodvisno izbiro problemov. »Gospodar mora samo znati ukazati tisto, kar mora suženj znati izvesti,« pravi Aristotel (1995, 1992). V tem smislu je človeška inteligenca inteligenca gospodarja, medtem ko je UI inteligenca sužnja. Ne more si ukazovati sama, ne more si sama postaviti svojih problemov (Peirce 1887, 165; Larson 2021, 233) – potrebuje gospodarja.⁴ Zato ne more biti ustvarjalna v smislu postavljanja novega problema. Definiranje inteligence kot neodvisnega reševanja problemov – recimo temu minimalna definicija – je zelo jasno in zato inženirsko uporabno. Po tej definiciji lahko za UI že rečemo, da je inteligentna. Toda če ne upoštevamo, da ta definicija za opredelitev človeške inteligence ni dovolj, saj ne upošteva njenih razlikovalnih značilnosti, kot je splošnost, lahko ta definicija posebnost človeške inteligence zamegli – in posledično vodi k napačnemu razumevanju potencialov UI. Vendar je opustitev minimalne definicije kot zadostne za definicijo katerekoli inteligence (ne samo minimalne inteligence) med programerji in inženirji bržkone precej težavna naloga, saj je takšno razumevanje del Turinrove zapuščine – Turing pa je utemeljitelj paradigme, ki pri raziskavah in razvoju UI še danes močno prevladuje.

² Na podobno neovrgljiv način so nekateri predstavniki *new agea* napovedovali začetek nove, ‚višje‘ zvesti (Sire 2002, 164ff).

³ Friedrich Engels (1975, 118–119; 351–352; 482; 510–511; 516–517; 552) ga je zagovarjal kot eno glavnih načel svoje ‚filozofije‘ narave ali dialektike narave.

⁴ UI svojemu gospodarju ne more reči »Ne!« v smislu zavrnitve ukaza. Pripisovanje takšne zavrnitve UI nima nobenega smisla. Pri človeškem sužnju pa takšna možnost obstaja.

Če računalnik uspešno ali celo bolje opravlja naloge, ki so prej zahtevale človeško inteligenco, je to v razvoju UI zagotovo napredek – in bolj zapletene kot so te naloge, večji je napredek. S tem se lahko strinjamo. Vendar pa je vprašanje, kakšno težo ima ta napredek kot evidenca za HIT ali SIT. Igranje iger, kot so šah in go, pravilno odgovarjanje na vprašanja kviza, medicinske diagnoze itd., ki jih izvaja UI, nekateri razglašajo za dokaze v prid HIT in SIT. S tem se ne strinjam – in namen tega prispevka je utemeljitev tega nestrinjanja.

Obravnavanje inteligence kot reševanja problemov vodi k ozkim aplikacijam. Če bi se stroj svojo omejitev na ozke naloge naučil preseči in bi bil sposoben reševati probleme na splošno, bi to pomenilo prehod v višjo ali (bolj) človeški podobno inteligenco. Toda zaenkrat smo od oblikovanja splošne UI še daleč. Da bi dosegel svoje cilje, se mora vsak sistem strojnega učenja naučiti nekaj specifičnega. Raziskovalci temu pravijo pristranskost sistema. Pristranskost v strojnem učenju pomeni, da je sistem zasnovan in prilagojen za učenje nečesa posebnega – to pa je ravno proizvodnja aplikacij za reševanje ozkih problemov. Zato se sistemi globokega učenja, ki jih je Facebook uporabljal za prepoznavanje človeških obrazov, niso hkrati naučili, kako izračunati naše davke. Situacija je v resnici še slabša: raziskovalci so ugotovili, da če je sistem strojnega učenja pristranski in specializiran za priučitev določeni aplikaciji ali nalogi, potem slabše opravlja druge naloge. Obstaja obratna korelacija med uspehom sistema pri učenju določene stvari in uspehom njegove priučitve drugi nalogi – in to velja celo za zelo podobne naloge. Sistem strojnega učenja, ki se uči igrati šah na visoki ravni, se goja na visoki ravni ne bo naučil igrati – in obratno. Sistem za go je bil namreč zasnovan s posebno pristranskostjo za učenje pravil goja. Težava je, da se pristranskosti ne moremo znebiti, ker je sestavni del strojnega učenja. Pregovor »nič ni zastoj« velja tudi za strojno učenje. V tem primeru to pomeni, da ko bo uporabljen za poljubne probleme oziroma naloge, noben sistem strojnega učenja, ki ni pristranski, ne bo deloval nič bolje od naključnega ugibanja. Resnično nepristranski sistem – sistem, ki ga programerji ne oblikujejo kot pristranskega – je neuporaben. Po drugi strani se pristranski sistem nauči samo tega, česar si želijo njegovi oblikovalci. Tako s pristranskostjo sistem naredimo ozek v smislu, da potem ne posplošuje na druga področja. Ozkost je torej od uspeha sistema neločljiva: ozkost in uspeh sta v sistemih strojnega učenja dve plati istega kovanca (Larson 2021, 28–30).

Sistemi UI, ki smojih razvili doslej, lahko probleme uspešno rešujejo samo tako, da jih oblikujemo kot pristranske. Tudi učeči se sistemi UI so zgolj taki ozki sistemi. Doslej ni znan noben znanstven ali tehnični preboj iz ozkih sistemov do splošne inteligence, kot jo izkazujejo ljudje. S tega vidika se je razvoj splošne UI ustavil: razumevanje inteligence le kot (neodvisnega) reševanja problemov postopoma – vendar neizogibno – vodi v teoretično slepo ulico v samem osrčju raziskovanja UI, v „past ozkosti“,⁵ kot jo imenuje Larson. To kaže, da je razumevanje inteligence,

⁵ Eden prvih, ki je na problem izvirne pobude in ozkosti jasno opozoril, je bil Charles Sanders Peirce: »Vsak stroj za sklepanje, to se pravi vsak stroj, ima dve inherentni nezmožnosti. Prvič, je brez vsake izvirnosti, brez vsake pobude. Ne more si najti svojih lastnih problemov; ne more se sam napajati. Ne more se sam usmerjati med različnimi mogočimi postopki. /.../ [S]troj bi bil popolnoma brez izvirne

ki se jo zvede na (neodvisno) reševanje problemov, preozko in za doseganje širših ciljev od proizvodnje ozkih aplikacij neustrezno (30–32).

Središčni problem za razvoj človeški podobne UI lahko opišemo s pojmom intuicije. Če bi želeli, da bi UI bila podobna človeški, bi morala biti sposobna intuicije. Če bi želeli, da bi UI imela intuicijo, kot jo imajo ljudje in raziskovalci, bi morali to intuicijo znanstveno opisati tako, da bi bil ta opis za programerje in inženirje uporaben. Vendar pa si nihče ne predstavlja, kako to narediti. Brez intuicije UI ne more preseči ‚prekletstva‘ ozkosti – kar pa bi bilo potrebno, če bi se želela učiti tako, da bi postala podobna človeški inteligenci. Takšno učenje zahteva, da sistem probleme izbere sam, za to pa je potrebna intuicija. Vemo, da oblikovalci sistemov UI za usmerjanje teh sistemov, katere specifične probleme naj rešujejo ali se naj se naučijo reševati, uporabljajo lastno intuicijo. Toda da bi bila UI resnično inteligentna, bi morala imeti svojo intuicijo. Tega nima in kot že omenjeno – nihče nima znanstvene predstave, kako ustvariti UI, ki bi jo imela (31–32).

Povzemimo glavne doslej navedene razloge za zavrnitev napovedi o verjetnem nastanku človeški podobne UI ali SI, še zlasti v bližnji prihodnosti. SI bi se iz človeške ali človeški podobne UI lahko razvila ‚sama od sebe‘ v teku evolucije, ne da bi jo tehnično razumeli. A to se ni zgodilo, čeprav človeška inteligenca že dolgo obstaja. Zakaj bi se to zgodilo v bližnji prihodnosti? Seveda imamo zdaj sisteme, ki določene dejavnosti opravljajo veliko hitreje in natančneje kot ljudje. Toda za človeški podobno inteligenco je potrebno veliko več kot hitrost in natančnost – potrebni sta intuicija in sposobnost samostojne izbire problemov. Tega, kako ustvariti takšno UI, pa si nihče ne predstavlja.

Turing je razmišljal o formalizaciji intuicije, da bi jo računalnik lahko uporabil, vendar ni vedel, kako to doseči. In tega nihče ne ve niti danes, kar ni presenetljivo, saj je prevladujoči znanstveni in tehnološki pristop UI dedič Turingovega razumevanja in deluje v okviru Turingove paradigme. Za preboj bi bila potrebna korenita sprememba pristopa in razumevanja UI, vendar nihče ne ve, kakšna naj bi ta sprememba v resnici bila. Podobno tudi še nihče ni opisal, kako bi se človeški podobna inteligenca – umetna ali naravna – lahko sama razvila v SI. Zato lahko rečemo, da je človeški podobna UI (za zdaj) le ‚mit‘ (Larson 2021) in znanstvena fantastika.

pobude in bi opravljal le tisto posebno vrsto dela, za katero je bil narejen. Vendar to ni pomanjkljivost stroja; ne želimo, da opravlja svoje delo, ampak naše. /.../ Drugič, zmogljivosti stroja imajo absolutne omejitve; zasnovan je bil za opravljanje določene stvari in ne more storiti nič drugega.« (Peirce 1887, 168–169; Larson 2021, 233) Turing se je problema pobude zavedal: »V naslednjem stoletju je Turing predlagal, da sprejmemo izziv in strojem vcepimo ‚izvirno pobudo‘ tako, da jih najprej programiramo, da se pogovarjajo z nami. Turing se je zavedal Peircevega ugovora, ki ga je v svojem članku iz leta 1950 pripisal lady Lovelace. Prav tako se je igral s preprostimi učnimi algoritmi, v petdesetih letih prejšnjega stoletja pa so se pojavile enoslojne nevronske mreže (imenovane perceptroni). Razumljivo je, da je Turing mislil, da bi se ugovoru Peircea in lady Lovelace morda lahko izognili tako, da bi ustvarili stroje, ki se učijo, zasnovane po človeških možganih. Če beremo *Computing Machinery and Intelligence*, dobimo vtis, da je učenje predstavljalo edini resnični izhod iz inherentnih omejitev strojev in edino resnično upanje, da bo stroj prestal Turingov preizkus. Vendar ga ni, to se ni zgodilo. Prepričanje, da se bo, da se mora, ima posledice za družbo, ki so zdaj postale več kot očitne. V zadnjem delu te knjige si bomo ogledali nekaj posledic mita o neizbežnosti – zlasti njegove škodljive učinke na znanost samo.« (233–234)

2.2 Abdukcija, zdravi razum in razumevanje

Lahko razlikujemo tri osnovne oblike sklepanja, ki jih ni mogoče zvesti eno na drugo: dedukcijo, indukcijo in abdukcijo.

V simbolih stavčne logike jih lahko predstavimo takole (Larson 2021, 172):

Dedukcija:	Indukcija:	Abdukcija:
$P \rightarrow Q$	P	$P \rightarrow Q$
P	Q	Q
Q	$P \rightarrow Q$	P

Sistemi globokega učenja temeljijo na statistiki. Statistika je v bistvu indukcija. Abdukcija je podobna ugibanju, vendar ni zgolj ugibanje. Je več kot le asociacija ali korelacija. Vendar nihče ne ve, kako ta ‚več‘ opisati na način, ki bi omogočil inženirstvo človeške inteligence. Abdukcija, zdravi razum in intuicija predpostavljajo razumevanje, ki ga nihče ne zna opisati na tehnično uporaben način. Razumevanje pa je jedro človeške inteligence. In dokler ga ne bomo znali ustrezno opisati, človeški podobne inteligence ne bomo mogli ustvariti.

Edino znanje, ki ga sistemu strojnega učenja lahko zagotovimo, je tisto, kar sistem iz podatkov lahko pridobi na povsem skladenjski način. To pomeni, da v sistemu obstaja slepa pega, ki povzroča napačne napovedi – saj tistega, česar sistem v podatkih ne more opaziti, ne pozna. (173) To slepo pego in napako v sklepanju lahko ponazorimo z zgledom purana (121–124), ki je induktivist in induktivno pride do usodno napačnega sklepa:

»Ta puran je ugotovil, da je bil v svojem prvem jutru na farmi puranov nahranjen ob 9. uri. Vendar kot dober induktivist ni delal prenapetih sklepov. Počakal je, da je zbral veliko število opažanj dejstva, da je bil nahranjen ob 9. uri – in to se je zgodilo v različnih okoliščinah: ob sredah in četrtnih, v toplih in mrzlih dneh, na deževne in sušne dni. Vsak dan je na svoj spisek dodal novo opažanje. Končno je bila njegova induktivistična vest zadovoljna in izvedel je induktivni sklep, da »sem vedno nahranjen ob 9:00«. Žal se je ta sklep na svet večer nedvomno izkazal kot napačen, saj so mu, namesto da bi ga nahranili, prerezali vrat. Induktivni sklep z resničnimi premisa mi je privedel do napačnega sklepa.« (Chalmers 1982, 41–42)⁶

Ta primer lepo ponazarja »neumnost oblikovanja ‚navad asociacije‘ brez globljega znanja« (Larson 2021, 123). Strojno »učenje« ne temelji na razumevanju in ni učenje razumevanja, ampak zgolj na asociaciji ali korelaciji. Zato strojno učenje ni človeškemu podobno učenje, saj človeško učenje temelji na razumevanju in je učenje razumevanja. Napredek v učenju za ljudi pomeni večje razumevanje.

⁶ Chalmersov primer purana, ki ga navajam, je različica znanega primera, ki ga je pred tem uporabil že Bertrand Russell (2001, 35–36).

Pri sistemih globokega učenja ne moremo govoriti o človeškemu podobnem učenju, ker jim ne moremo pripisati razumevanja. Njihov napredek ni napredek v razumevanju, ampak v količini informacij ali »dejstev«, ki jih imajo, in v (simulaciji) prilagodljivosti (Fuchs 2021, 33).

Razlika med človeško inteligenco in UI je jasno vidna pri prepoznavanju obrazov. Medtem ko UI za prepoznavanje obrazov potrebuje ogromno število ogledov, dojenčku za prepoznavo obraza lahko zadostuje le nekaj ogledov. Seveda dojenček ne more razložiti, zakaj je obraz prepoznal. To znanje je ,znanje, kakó'. Hubert Dreyfus (1972) je zagovarjal dve trditvi, ki sta za to znanje pomembni (Oettinger 1972, xii-xiii): 1. Ljudje najprej zaznamo celoto,⁷ ki jo šele potem, dodatno – če je potrebno – razčlenimo na dele. 2. Dojenčkovo prepoznavanje temelji na njegovem človeškem telesu. Zato bi morali roboti imeti dovolj človeškim podobna telesa, da bi imeli tudi človeški podobno inteligenco.

3. Partikularnost in zunanostnost inteligence

Cholletov članek (2017) o neverjetnosti tako imenovane eksplozije inteligence je vzbudil veliko zanimanja. Doslej je doživel že več kot osemnajst milijonov ogledov. Izraz ,eksplozija inteligence' je v šestdesetih letih prejšnjega stoletja uporabljal britanski matematik Irving John Good, ko je napovedoval nastop UI, ki da se bo razvila do nepredstavljive mere. Ta razvoj naj bi potekal s silno rastjo. Chollet dokazuje, da je takšna eksplozija malo verjetna. Njegova argumentacija je pomembna danes, ko se trditve o eksploziji inteligence – formulirane kot napovedi o pojavu singularnosti in podobnih pojavov – pojavljajo znova.

Cholletov pristop v članku je izkustven. Sklicuje se na podatke o evoluciji in zgodovinskem razvoju inteligence ter na različne druge empirične podatke, na primer o uspehu in dosežkih ljudi z nadpovprečnim inteligenčnim količnikom (v nadaljevanju IQ) in tako dalje.

Chollet pravi, da je pojem inteligence zelo pomembno najprej opredeliti – da bo jasno, o čem razpravljamo in kaj dokazujemo. Uvodoma jo opredeljuje kot reševanje problemov. Pravi, da je to začetna definicija, tako rekoč delovna definicija za začetek razprave. V celotnem članku to definicijo naprej razvija in dopolnjuje.

Jedro Cholletove argumentacije v članku sestavljajo naslednje tri trditve:

- Ni splošne inteligence.
- Inteligenca je situacijska in kontekstualna.
- Inteligenca je zunanostna (eksternalistična).

Vse tri značilnosti so medsebojno povezane in prepletene.

Prva teza trdi, da je vsaka inteligenca prilagojena entiteti, ki to inteligenco ima – njenim potrebam, njenemu načinu življenja itd. To velja za inteligenco živali, za človeško inteligenco in tudi za UI, ki je prilagojena svojim namenom. Ni splošne

⁷ To celoto lahko v smislu Gestalt psihologije označimo kot lik (nem. *Gestalt*).

inteligence v smislu, da bi bila primerna za vse naloge in vse namene. Človeška inteligenca je za nekatere živalske potrebe neuporabna – in obratno. Veliko je že prirojeno. Tretja trditev pomeni, da naša inteligenca ni (le) v naših možganih, ampak v znatni meri tudi v naših civilizacijah (Larson 2021, 27).

Druga izjava pravi, da na uporabo, izkoriščanje in razvoj naše inteligence vpliva situacija, v kateri se znajdemo. Chollet navaja primer ‚divjih‘ otrok, ki so odraščali med živalmi, na primer človeka, ki je živel med opicami. Na neki način je postal ‚opica‘ in se nikoli ni mogel popolnoma počlovečiti. Nikoli se ni naučil jezika, skakal je po mizah in drugod, zavračal kuhano hrano itd. Enako velja za druge ‚divje‘ otroke. Nekateri so se počlovečili bolj, drugi manj – a v vseh primerih precej pomanjkljivo.

Le malo ljudi z nadpovprečnim IQ v življenju doseže kaj izjemnega. Večina teh ljudi dela v ‚banalnih‘ poklicih, opravlja precej običajne in ‚povprečne‘ funkcije in naloge. Kaj in koliko bo kdo s svojo inteligenco dosegel, kako jo bo izkoristil in razvil itd., je odvisno od številnih dejavnikov: njegovega zgodovinskega in družbenega položaja, vzgoje, izobrazbe, naključnih srečanj s pravimi ljudmi, nahajanja na pravem mestu ob pravem času, ustrezne motivacije in prizadevanj oziroma teženj itd. Chollet navaja primere znanstvenikov, ki so dosegli pomembne rezultate, vendar njihov IQ ni bil nič posebnega – na primer fizik Feynman in Watson, soodkritelj DNA. Imeli so enak IQ kot mnogi drugi znanstveniki, ki nikoli ne bodo odkrili česa podobnega. In koliko ljudi z višjim IQ kot oni bo doseglo kaj njihovem dosežkom primerljivega? Navaja tudi besede slavnega biologa Stephena Jaya Goulda, ki ga je teža in velikost Einsteinovih možganov skrbela manj kot misel na to, koliko ljudi z Einsteinovim IQ je trpelo in še trpi v različnih okoljih, kjer svojih danosti nimajo možnosti razviti. Ali pa če se vrnemo v človeško zgodovino pred 10.000 leti – kako so ljudje živeli, kako so lahko razvijali svojo inteligenco. Govorili so preprost jezik, večina jih ni znala niti brati niti pisati itd. Kaj bo kdo dosegel ali odkril, je odvisno še od drugih zgodovinskih dejavnikov. Sklepna Cholletova ugotovitev – na podlagi različnih podatkov iz preteklosti in sedanjosti – je, da bo rast inteligence tudi v prihodnje zmerna, linearna in ne eksplozivna, eksponentna ali celo nepredstavljivo strma. Napovedi Raya Kurzweila in podobnih ‚prerokov‘ o nastopu SI se ne bodo uresničile.

Cholletov sklep je v nekaterih pogledih lahko presenetljiv. Vzemimo razvoj znanosti in tehnologije, ki močno vpliva na razvoj na vseh drugih področjih. Če upoštevamo, kako močna orodja, katerih zmogljivosti se nenehno povečujejo, imamo zdaj na voljo, kakšna je prefinjenost komunikacij, možnost izmenjave znanja in mreženja, kako se je število ljudi, ki se ukvarjajo z razvojem znanosti in tehnologije itd., povečalo, bi pričakovali hitrejši napredek. Vendar Chollet opozarja, da moramo videti celoto in upoštevati učinke, ki jih takšen razvoj prinaša. Z razvojem se kompleksnost znanstvenih problemov povečuje. Z naraščanjem znanja se povečuje čas, potreben za obvladovanje tega znanja, za izobraževanje znanstvenikov in spremljanje dosežkov v znanosti – da lahko ostanemo na tekočem in razvijamo nove stvari. Če upoštevamo vse navedeno, ni presenetljivo, da se po vseh metodološko merljivih kazalnikih znanstveno znanje na področju fizike v drugi polovici 20. stoletja ni povečalo bolj kot v prvi polovici istega stoletja, lahko pa bi navedli še druge primere. Po eni strani je res, da bolj ko ste izobraženi, hitreje bo vaše

znanje naraščalo, saj boste lahko uporabljali kompleksna orodja, matematično znanje, ustrezne simbole itd. Vendar je treba upoštevati celoto ter inteligenco in rast znanja razumeti kontekstualno in zunanostno, da lahko vidimo povezanosti stvari in razumemo, zakaj je rast inteligence in znanja linearna. Podobno na področju ekonomije – po eni strani je res, da več ko imate denarja, več lahko dodatno pridobite. Toda če pogledamo celoto, z rastjo bogastva in naložb rastejo tudi drugi negativni dejavniki rasti – in končno empirični podatki kažejo, da je rast bogastva linearna, postopna, ne pa 'eksplozivna'.

Chollet torej na več načinov ugotavlja, da rast inteligence ne bo eksplozivna, pač pa bo inteligenca rasla tako kot doslej – postopoma, linearno. Civilizacija in inteligenca se bosta razvijali še naprej: znanstveno znanje, število tehnoloških orodij in njihove zmogljivosti itd. bodo naraščali. Vendar do razvoja supermožganov, SI ali česa takšnega, kar bi pomenilo konec oziroma ukinitve človeštva, ne bo prišlo. Niti ne bo UI prevzela vseh nalog in oblasti – UI namreč nima lastnih potreb, namenov, motivov, želja, namer itd. V vseh teh pogledih in pomenih »eksplozija inteligence« po Cholletovem prepričanju ni verjetna.

4. Veliki podatki in znanost panja

Sestavni del današnje ideologije je ideologija velikih podatkov (angl. *big data*) in velike znanosti. Zagovorniki te ideologije trdijo, da bo zgolj kopičenje in združevanje oziroma povezovanje velike količine podatkov z uporabo trenutnega znanja in teorij privedla do nastanka UI, ki bo človeško raven najprej dosegla in jo nato (hitro) neskončno preseгла – dosegla singularnost. Trdijo, da je nadaljnji razvoj teorije nepotreben ali celo nemogoč in da bo manjkajoče koščke za ustvarjanje UI priskrbel integracija velike količine podatkov s pomočjo zmogljivih računalnikov. Poleg poudarjanja potenciala velikih podatkov pristaši velike znanosti zagovarjajo koncept znanosti panja (angl. *hive science*). Ne potrebujemo več visoko inteligentnih in ustvarjalnih posameznikov z izvirnimi zamislimi in uvidi – ne potrebujemo več Einsteinov.⁸

Takšno stališče pomeni degradacijo človeškega uma in ljudi zgolj na asistente velikega računalnika, zgolj na njegove služabnike in vzdrževalce, ki ga morajo oskrbovati s potrebnimi podatki in storitvami za delovanje. Ne potrebujemo več ustvarjalnih osebnosti, ki bi razmišljale in imele nove ideje, ampak samo visoko usposobljene tehnike, servisno osebje in operaterje. Posledica uresničevanja takšnega stališča je, da se ljudje vse bolj prilagajajo stroju – in postajajo strojem vse bolj podobni. V smislu tipičnih mehanskih nalog stroj seveda ljudi presega, zato je pri takšni zasnovi razumljivo, da se ljudje stroju podredijo in postanejo njegovi strežniki. Takšno dojemanje pomeni zmanjševanje pomena posameznika in spodbuja oblikovanje družbe, ki ne vlaga več v ustvarjanje pogojev za razvoj posameznikov z izvirnimi za-

⁸ Eden od zagovornikov takšnega koncepta je Henry Markram, ki je izjavil, da nas ovira splošno razširjeno prepričanje, da za razlago, kako delujejo možgani, potrebujemo Einsteina. Dejansko pa je tisto, kar resnično potrebujemo, da »postavimo na stran svoje ege in ustvarimo novo vrsto kolektivne nevroznanosti« (Markram 2013; navajam po Larson 2021, 246).

mislimi in teorijami, ki omogočajo resnične znanstvene in tehnološke preboje in napredek. Zato ni presenetljivo, da v okoljih in obdobjih, kjer prevladuje ideja velike znanosti, pravega napredka v znanosti in tehnologiji ni – čeprav zagovorniki velike znanosti stanje skušajo prikazati drugačno. Če se bo v prihodnosti model velike znanosti uveljavil še bolj, se bo ogroženost znanstvenega in tehnološkega razvoja in napredka povečala. Mnogi znanstveniki se tega zavedajo in opozarjajo na škodljivost koncepta velike znanosti ter aktivno organizirajo in usmerjajo nasprotovanje njemu financiranju. Pogosto so projektom, ki potekajo po modelu velike znanosti, namenjeni ogromni zneski. Primer nasprotovanja veliki znanosti je peticija, ki jo je podpisalo več kot 500 uglednih znanstvenikov in jo naslovilo na Evropsko komisijo. Zahtevali so velike spremembe projekta Human Brain Project (Larson 2021, 267), ki se je uradno začel leta 2013 pod vodstvom že omenjenega nevroznanstvenika, dr. Henryja Markrama iz Švicarskega zveznega inštituta za tehnologijo v Lozani. Projekt je sprva vključeval več kot 150 institucij po vsem svetu (243).

5. Zaključek

Predstavil sem štiri Larsonove argumente proti HIT – imenujemo jih lahko argument splošnosti, argument intuicije, argument abdukcije in argument zdravega razuma –, ki so po mojem mnenju dobri razlogi za zavrnitev HIT. Poleg tega sem predstavil Cholletovo argumentacijo proti SIT. Ta temelji na izkustveni razčlenitvi, pri čemer upošteva partikularnost in zunanostnost inteligence. Cholletovi argumenti v povezavi s predstavljenimi Larsonovimi argumenti zagotavljajo solidno utemeljitev za zavrnitev SIT. Pokazal sem, kako pomembno je inteligenco pravilno razumeti in opredeliti, saj ima nepravilno razumevanje lahko škodljive posledice, ki daleč presegajo zgolj akademsko področje. Takšno razumevanje lahko na primer zagotavlja osnovo za pridobivanje neupravičenih finančnih in drugih koristi za škodljive in nevarne ideologije, kot je denimo radikalni transhumanizem,⁹ ter za instrumentalistična, razčlovečujoča in antipersonalistična stališča o pomenu in naravi človeške ustvarjalnosti in znanstvene dejavnosti.

Kratice

HIT – trditev, da bo prišlo do nastopa človeški podobne UI.

IQ – inteligenčni količnik.

SI – superinteligence.

SIT – trditev, da se bo pojavila SI.

UI – umetna inteligenca.

⁹ Za kritično razpravo o (radikalnem) transhumanizmu prim. Dittrich 2021; Sturm 2021; Pouliquen 2022; Petkovšek in Žalec, ur. 2021; Globokar 2019; Gregorčič 2023, 914; Klun 2019; 2024, 28; Osredkar 2019; Platonvjak in Svetelj 2019; Pohar 2019; Stegu 2019; Strahovnik 2019; Štivić 2021; 2023, 889–891; Vodičar 2019; Žalec 2019; 2023b.

Reference

- Aristotle.** 1995. *Politics*. Prev. Benjamin Jowett. V: Jonathan Barnes, ur. *The Complete Works of Aristotle*. Zv. 2, 1986–2129. Princeton: Princeton University Press
- Bostrom, Nick.** 2020. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Chollet, François.** 2017. The implausibility of intelligence explosion. Medium. 27. 11.
- https://medium.com/@francois.chollet/the-imp-possibility-of-intelligence-explosion-5be4a9eda6ec** (pridobljeno 22. 8. 2023).
- Chalmers, Alan.** 1982. *What Is This Thing Called Science*. St. Lucia, AU: University of Queensland Press.
- Dittrich, Tristan Samuel.** 2021. Transhumanistisches Glückstreben und christliche Heilshoffnung: Ein Vergleich. *Zeitschrift für Theologie und Philosophie* 143, št. 3:452–474.
- Dreyfus, Hubert L.** 1972. *What Computers Can't Do: A Critique of Artificial Intelligence*. New York: Harper and Row.
- — —. 1991. *Being-in-the-World: A Commentary on Heidegger's Being in Time, Division I*. Cambridge, MA: The MIT Press.
- Dreyfus, Hubert, in Charles Taylor.** 2015. *Retrieving Realism*. Cambridge, MA: Harvard University Press.
- Engels, Friedrich.** 1975. *Dialektik der Natur*. V: Karl Marx in Friedrich Engels, *Werke*. Zv. 20, 307–570. Berlin: Dietz Verlag.
- Fuchs, Thomas.** 2021. *In Defense of Human Being: Foundational Questions of an Embodied Anthropology*. Oxford: Oxford University Press.
- Gibson, James Jerome.** 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Globokar, Roman.** 2019. Normativnost človeške narave v času biotehnološkega izpopolnjevanja človeka. *Bogoslovni vestnik* 79, št. 3:611–628. <https://doi.org/10.34291/bv2019/03/globokar>
- Good, Irving John.** 1965. Speculations Concerning the First Ultrainelligent Machine. *Advances in Computers* 6:31–88. [https://doi.org/10.1016/s0065-2458\(08\)60418-0](https://doi.org/10.1016/s0065-2458(08)60418-0)
- Gregorčič, Rok.** 2023. Tehnološki razvoj v luči Habermasove etike diskurza. *Bogoslovni vestnik* 83, št. 4:911–922. <https://doi.org/10.34291/bv2023/04/gregorcic>
- Klun, Branko.** 2019. Transhumanizem in transcendenca človeka. *Bogoslovni vestnik* 79, št. 3:589–600. <https://doi.org/10.34291/bv2019/03/klun>
- — —. 2024. Problem religioznega izkustva v digitalno transformiranem svetu: Eksistencialno fenomenološki pristop. *Bogoslovni vestnik* 84, št. 1:19–32. <https://doi.org/10.34291/bv2024/01/klun>
- Kurzweil, Ray.** 2005. *The Singularity is Near: When Humans Transcend Biology*. London: Duckworth Overlook.
- Landgrebe, Jobst, in Barry Smith.** 2023. *Why Machines will Never Rule the World: Artificial Intelligence without Fear*. London: Routledge.
- Larson, Erik J.** 2021. *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Markram, Henry.** 2013. Seven Challenges for Neuroscience. *Functional Neurology* 28:145–151.
- McCulloch, Gregory.** 1995. *The Mind and Its World*. London: Routledge.
- Miščević, Nenad.** 1988. *Radnja i objašnjenje*. Zagreb: Hrvatsko filozofsko društvo.
- Oettinger, Anthony G.** 1972. Preface. V: Hubert Dreyfus. *What Computers Can't Do: A Critique of Artificial Intelligence*, xi–xiii. New York: Harper and Row.
- Osredkar, Mari Jože.** 2019. Religija kot izziv za transhumanizem. *Bogoslovni vestnik* 79, št. 3:657–668. <https://doi.org/10.34291/bv2019/03/osredkar>
- Peirce, Charles Sanders.** 1887. Logical Machines. *The American Journal of Psychology* 1, št. 1:165–170.
- Petkovšek, Robert, in Bojan Žalec, ur.** 2021. *Transhumanism as a Challenge for Ethics and Religion*. Zürich: Lit.
- Platovnjak, Ivan, in Tone Svetelj.** 2019. To Live a Life in Christ's Way: The Answer to a Truncated View of Transhumanism on Human Life. *Bogoslovni vestnik* 79, št. 3:669–682.
- Pohar, Borut.** 2019. Transhumanizem v službi človekove odgovornosti do stvarstva. *Bogoslovni vestnik* 79, št. 3:643–656. <https://doi.org/10.34291/bv2019/03/pohar>
- Potrč, Matjaž.** 1993. *Phenomenology and Cognitive Science*. Dettelbach: Verlag J. H. Röhl.
- — —. 1996. Phenomenology and organic unity. V: Elisabeth Baumgartner, Wilhelm Baumgartner, Bojan Borstner, Matjaž Potrč, John Shawe-Taylor, Elisabeth Valentine, ur. *Handbook Phenomenology and Cognitive Science*, 185–197. Dettelbach: Verlag J. H. Röhl.
- — —. 2004. *Dinamična filozofija*. Ljubljana: Filozofska fakulteta.
- Pouliquen, Tanguy Marie.** 2022. *Čar novih tehnologij in transhumanizem: 115 vprašanj*. Ljubljana: Družina.
- Rowlands, Mark, Joe Lau, in Max Deutsch.** 2020. Externalism About the Mind. V: Edward N. Zalta, ur. *The Stanford Encyclopedia of Philosophy*.

- <https://plato.stanford.edu/archives/win2020/entries/content-externalism/> (pridobljeno 1. 2. 2024).
- Sire, James W.** 2002. *Izazov svjetonazora: pregled temeljnih svjetonazora* [The Universe Next Door, 1997]. Zagreb: STEPress.
- Russell, Bertrand.** 2001. *The Problems of Philosophy*. Oxford: Oxford University Press.
- Stegu, Tadej.** 2019. Transhumanizem in krščanska antropologija. *Bogoslovni vestnik* 79, št. 3:683–692. <https://doi.org/10.34291/bv2019/03/stegu>
- Strahovnik, Vojko.** 2019. Vrline in transhumanistična nadgradnja človeka. *Bogoslovni vestnik* 79, št. 3:601–610. <https://doi.org/10.34291/bv2019/03/strahovnik>
- Sturm, Wilfried.** 2021. Transhumanismus und Digitalisierung: Theologisch-anthropologische Perspektiven. *Zeitschrift für Theologie und Philosophie* 143, št. 3:425–451.
- Štivič, Stjepan.** 2021. Upanje v krščanstvu in transhumanizem. *Bogoslovni vestnik* 81, št. 4:849–856. <https://doi.org/10.34291/bv2021/04/stivic>
- . 2023. Kiborgizacija: nastanek in razvoj pojma. *Bogoslovni vestnik* 83, št. 4:883–892. <https://doi.org/10.34291/bv2023/04/stivic>
- Vodičar, Janez.** 2019. Transhumanizem in katoliška vzgoja. *Bogoslovni vestnik* 79, št. 3:693–704. <https://doi.org/10.34291/bv2019/03/vodicar>
- Turing, Alan.** 1950. Computing Machinery and Intelligence. *Mind* 59, št. 236:433–460. <https://doi.org/10.1093/mind/lix.236.433>
- Žalec, Bojan.** 2019. Liberalna evgenika kot uničevalka temeljev morale: Habermasova kritika. *Bogoslovni vestnik* 79, št. 3:629–641. <https://doi.org/10.34291/bv2019/03/zalec>
- . 2023a. Ali je umetna inteligenca inteligenca v pravem pomenu besede? Vprašanje psihičnih značilnosti in splošnosti. *Bogoslovni vestnik* 83, št. 4:813–823. <https://doi.org/10.34291/bv2023/04/zalec>
- . 2023b. *Človečnost v digitalni dobi: Izzivi umetne inteligence, transhumanizma in genetike*. Ljubljana: Teološka fakulteta. https://www.teof.uni-lj.si/uploads/Zalozba/ZnK86-Zalec-clovecnost_elektronska.pdf (pridobljeno 1. 2. 2024).

*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 najboljše. 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 upravljalca 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, namenja- mo 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 imple- mentaciji, npr. podaljševanje časa na spletu, obseg nakupov oz. dobiček za spletno trgovino, hitra prodaja izdelkov, ki so na zalogi v prevelikem obsegu ali bodo kma- lu 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 ka- terimi 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 pri- kazani 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 terjaja 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).

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 783—795

Besedilo prejeto/Received:09/2024; sprejeto/Accepted:12/2024

UDK/UDC: 004.021:364.63:343

DOI: 10.34291/BV2024/04/Strahovnik

© 2024 Strahovnik, CC BY 4.0

Vojko Strahovnik

Algoritemsko nasilje *Algorithmic Violence*

Povzetek: Razvoj najnovejših sistemov in orodij na področju umetne inteligence ter poudarek na etičnih vprašanjih, ki jih to prinaša, je v literaturi sprožil pojav novega pojma – algoritemsko nasilje. Vezan je na nastanek, uporabo in posledice rabe sistemov umetne inteligence, ki temeljijo na algoritmih. Naš cilj v tem prispevku bo analiza in kritično ovrednotenje tega pojma. Ugotovitve bodo pretežno skeptične. Trdili bomo, da ta pojem pomembne etične dileme v zvezi z uporabo umetne inteligence do neke mere sicer izpostavlja, vendar pa jih precej bolj smiselno in natančno označujejo drugi izrazi, kot so pristranskost, diskriminacija, izključevanje, stereotipi, marginalizacija, nepravičnost, netransparentnost ipd. Uporaba pojma algoritemsko nasilje tako ni dobro utemeljena in ne prinaša posebnih teoretskih ali praktičnih prednosti ali koristi.

Gljučne besede: umetna inteligenca, algoritmi, nasilje, algoritemsko nasilje, spoznavna nepravičnost, pristranskost

Abstract: With the development of the latest systems and tools in the field of artificial intelligence and the emphasis on the ethical issues it raises, a new concept has emerged in the literature. This is the concept of algorithmic violence. It refers to the development, use, and consequences of AI systems based on algorithms. Our aim in this paper will be to analyse and critically evaluate this notion. The conclusions will be mainly sceptical, but they carry significant implications for the field. It will be argued that while the term does, to some extent, highlight important ethical problems in the use of AI, these are much better characterized by other terms such as bias, discrimination, exclusion, stereotyping, marginalization, unfairness, non-transparency, etc. The use of the term algorithmic violence is thus not well founded and does not bring any particular theoretical or practical advantages or benefits.

Keywords: artificial intelligence, algorithms, violence, algorithmic violence, epistemic injustice, bias

1. Uvod

Z razvojem najnovejših tehnologij na področju umetne inteligence, obenem pa tudi poudarkom na etičnih in širših filozofskih vprašanjih, ki jih te sprožajo (Strahovnik 2023; Žalec 2023), se je v literaturi začel pojavljati nov pojem.¹ To je pojem algoritemskega nasilja (Onuoha 2018; Huws 2024). Moj cilj v tem prispevku bo ta pojem analizirati in kritično ovrednotiti. Moje ugotovitve bodo pretežno skeptične. Trdil bom, da ta pojem sicer izpostavlja pomembne etične dileme, vendar pa te dejansko bolje označujejo drugi izrazi, kot so pristranskost, diskriminacija, izključevanje, stereotipizacija, marginalizacija, nepravilnost, netransparentnost itd. Tovrstni pojmi stara in nova etična vprašanja ter izzive, ki jih je najnovejši napredek na področju umetne inteligence sprožil, zajamejo na bolj celosten in informativen način. Poleg tega so uporabnejši pri iskanju rešitev in oblikovanju strategij za izogibanje omenjenim negativnim učinkom ali tveganjem pri uporabi algoritmov umetne inteligence.

Prispevek v drugem razdelku predstavlja pojem algoritemskega nasilja, v treh podrazdelkih pa še posebej očituje in analizira tri idejne podlage oz. izhodišča, iz katerih uvedba in raba tega pojma izhajata. V tretjem razdelku predstavljam skeptične premisleke, ki govorijo proti rabi pojma algoritmsko nasilje. Pri tem so izpostavljene tudi različne strategije obrambe uporabe tega pojma kot ugovor mojim tezam; ta je oblikovan na podlagi tega, da razprava, vključno s tezo o neustreznosti rabe pojma algoritmsko nasilje, predpostavlja, da moramo najprej opredeliti samo nasilje – in šele nato presoati algoritmsko nasilje.² V četrtem, zadnjem razdelku je izpostavljeno, kako lahko uporaba bolj določnih pojmov tipa pristranskost, netransparentnost, diskriminacija in podobnih tveganja pri rabi umetne inteligence izpostavi bolje. Prav tako se uporaba teh pojmov bolje vpenja v prizadevanja za oblikovanje mreže odgovornosti ter z njimi povezanih dolžnosti za čim večjo omilitev teh tveganj.

2. O pojmu algoritemskega nasilja

Pojem algoritemskega nasilja je razmeroma nov in se je v razpravah pojavil šele v zadnjih letih. Ena prvih omemb je iz leta 2018. Pojavila se je na spletu, v objavi na platformi za razvijalce GitHub; njena avtorica je Mimi Onuoha, raziskovalka in umetnica. Onuoha v tem zapisu algoritmsko nasilje opredeljuje takole:

»Algoritmsko nasilje se nanaša na nasilje, ki ga algoritem ali sistem za samodejno odločanje povzroča tako, da posameznikom preprečuje, da bi

¹ Publikacija je nastala v okviru raziskovalnega programa P6-0269 „Religija, etika, izobraževanje in izzivi sodobne družbe“ in raziskovalnega projekta J6-4626 „Teologija, digitalna kultura in izzivi na človeka osredičene umetne inteligence“, ki ju finančno podpira 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“, Univerza v Oxfordu)

² Za jasno izpostavitev tega vidika oz. ugovora se zahvaljujem Bojanu Žalcu.

zadovoljili svoje osnovne potrebe. Izvira iz izkoriščevalskih družbenih, političnih in ekonomskih sistemov ter se z njimi krepí, lahko pa je tudi tesno povezano s prostorsko in fizično prenesenimi učinki.« (Onuoha 2018)

Vidimo, da Onuoha algoritemsko nasilje v okviru te opredelitve povezuje prvenstveno s priporočilnimi oz. odločitvenimi sistemi, a opozoriti je treba, da ga v nadaljevanju prispevka ter tudi v spremljajoči razpravi ne omejuje zgolj in samo na takšne sisteme umetne inteligence. To njeno začetno zamisel o vpeljavi novega pojma so nato prevzeli tudi drugi. Safransky (2020) v kontekstu prostorskih študij algoritemsko nasilje opredeljuje podobno in uvaja besedno zvezo »geografije algoritemskega nasilja«, ki označuje »ponavljajočo se in standardizirano obliko nasilja, ki prispeva k rasizaciji prostora in prostorski določenosti revščine« (Safransky 2020, 200) – to je zanj posledica podatkovno vodenih tehnologij načrtovanja in algoritemskega odločanja pri urejanju prostora. Eden od vidikov oz. posebnosti, ki jih omenja, je ta, da je algoritemsko nasilje bolj skrito ali manj očitno kot druge oblike nasilja, ki so vidne ali »spektakularne«. Huws in sodelavci (2023) pojem algoritemskega nasilja uvajajo v kontekstu analize visokošolskega oz. akademskega prostora ter v njem navzočega nadzora, upravljanja in nasilja, ki spremlja opravljanje dela. Tega tesno povezujejo s pojmom birokratskega nasilja – to se lahko porodi na temelju etične brezbriznosti, ki sledi razčlovečenju v kontekstu birokratskega upravljanja (Bauman 2006, 163–165). Mbalaka (2023) algoritemsko nasilje navezuje na pojem spoznavnega nasilja in ga tako pravzaprav razume kot vrsto spoznavnega nasilja, kjer ključno vlogo pri produkciji oz. implementaciji tega nasilja igrajo algoritmi umetne inteligence.

Nekoliko širše lahko rečemo, da razprava o algoritemskem nasilju zadeva »izpostavitev, razkritje in kritično proučitev, kako algoritemski sistemi spodbujajo določene oblike nasilja, kako nasilna dejanja opravičujejo ali na novo določajo, kakšno nasilje velja za legitimno« (Bellanova et al. 2021, 123). Kar naj bi algoritemsko nasilje oddvajalo od preostalih oblik nasilja, je torej to, da naj bi bili pri takem nasilju v njegovo porajanje, uresničevanje ali upravičevanje v bistveni meri vpeti algoritmi. Pri tem je treba izpostaviti, da je v večini primerov in razprav sam pojem algoritma razumljen na širok način in kot tesno povezan s tehnologijami umetne inteligence. Ne gre torej le za zaporedje natančno določenih in računalniško izvedljivih korakov oz. napotkov, ki privede do želenega rezultata ali rešitve, pač pa za vsakršne algoritme, ki so vključeni v procese učenja in delovanja sistemov umetne inteligence – ki torej strojem omogočajo, da s pomočjo predhodnega procesa učenja posnemajo človeško inteligenco in samostojno izvajajo zapletene naloge. Etični pomisleki in tveganja, ki jih delovanje takšnih sistemov odpira, so že bili jasno izpostavljeni (Weidinger et al. 2022; Strahovnik, Miklavčič in Centa Strahovnik 2020). Transparentnost, razložljivost, razumljivost, sledljivost, preglednost, poštenost, nediskriminatornost, predvidljivost in odgovornost so ključna merila in družbene zahteve, ki veljajo tako za ljudi kot tudi za ustanove, ki opravljajo družbene funkcije. Ko algoritmi umetne inteligence prevzamejo naloge oz. delo, ki so jih prej izvajali ljudje ali institucije, prevzamejo tudi družbene zahteve,

vkjučno z zgoraj omenjenimi merili (Bostrom in Yudkowsky 2014). Zdi se, da želi razprava o algoritemskem nasilju k zgoraj omenjenim premislekom dodati še vidik nasilja ter zahtevo po nenasilnosti.

Konceptualizacijo algoritemskega nasilja ter utemeljitev tega pojma lahko tako obravnavamo z več vidikov. V nadaljevanju na kratko izpostavljam tri najbolj razširjene, ki se medsebojno sicer ne izključujejo. Mednje spadajo okvir strukturnega nasilja, okvir birokratizacije z upravljavskimi strukturami ter epistemičnega oz. spoznavnega nasilja. Ne gre zgolj za vidike obravnave pojma algoritemsko nasilje, ampak tudi za izhodišča, na podlagi katerih zgoraj omenjeni avtorji sam pojem algoritemskega nasilja sploh zasnujejo.³ Glede na ta izhodišča bomo obravnavani pojem v nadaljevanju tudi lažje ovrednotili.

2.1 Strukturno nasilje

Prvi okvir, ki predstavlja podlago za vpeljavo pojma algoritemskega nasilja, je okvir strukturnega nasilja. Strukturno nasilje se nanaša na načine, kako družbene strukture in institucije povzročajo škodo in ohranjajo neenakost. Kot izraz ga je prvi skoval norveški sociolog Johan Galtung (1969), ki ga je v prvi vrsti zanimalo zlasti, kako družbene institucije, strukture in sistemi vplivajo na marginalizirane skupnosti. Glavni namen uvedbe pojma strukturno nasilje je bil poudariti posredne, sistemske in pogosto neopažene mehanizme, ki določene skupine ali posameznike pogosto prikrajšajo za izpolnitev osnovnih človekovih potreb, priložnosti in pravic. Galtung je ta koncept uvedel v zvezi z vprašanjem delovanja nasilja in njegovih povzročiteljev. Izhodiščna vprašanja so bila zlasti naslednja:

»ali obstaja subjekt (oseba), ki deluje, ali ne. Ponovno se lahko vprašamo: ali o nasilju lahko govorimo, če nihče ne izvaja neposrednega nasilja, ni dejaven? /.../ Vrsto nasilja, kjer akter, ki izvaja nasilje, obstaja, bomo poimenovali osebno ali neposredno; nasilje, kjer takega akterja ni, pa strukturno ali posredno. V obeh primerih so lahko posamezniki ubiti ali pohabljeni, prizadeti ali poškodovani v obeh pomenih teh besed ter manipulirani s strategijami palice ali korenčka. Toda medtem ko je v prvem primeru te posledice mogoče pripisati konkretnim osebam kot akterjem, v drugem primeru to ni več smiselno. V strukturi morda ni osebe, ki bi drugi osebi škodovala neposredno. Nasilje je vgrajeno v strukturo in se kaže kot neenaka moč – in posledično kot neenake življenjske možnosti.« (170–171)

³ Opozoriti moramo, da se s tem izrazom v navezavi na umetno inteligenco običajno povezujeva dva pomenska odtenska. Prvi je prevladujoč in se nanaša na nasilje ter z njim povezano škodo, povzročeno osebi, skupini ali skupnosti, ki jo prizadene zasnutje, razvoj ali uporaba algoritmov, temelječih na umetni inteligenci. Standardni primer bi bila škoda, povzročena skupini, ki je ob implementaciji algoritma umetne inteligence sistematično diskriminirana (npr. generativni veliki jezikovni model ali generator slik, ki temnopolte osebe povezuje z negativnimi vlogami ali jih kot del tvorbe slik prikazuje neustrezno). Drugi pomenski odtенок je nekoliko drugačen in se nanaša na nasilje, ki je vgrajeno v sam razvoj oz. trening globokih nevronske mreže z uporabo algoritmov – v tem primeru je razumljeno kot nasilje do nevronske mreže (Kockleman 2020, 326). V tem prispevku se omejujemo na prvi, bolj običajen pomen, drugega pa puščamo ob strani.

Strukturno nasilje je oblika nasilja, do katerega prihaja pri delovanju nepravilnih sistemov, npr. institucionaliziranega razizma ali kastnega sistema, ko ti sistemi in strukture nesorazmerno vzpostavljajo, ohranjajo ali poglobljajo trpljenje in prikrajšanost določenih skupin prebivalstva. Drugače kot neposredno nasilje, ki med drugim lahko vključuje fizične poškodbe, smrt ali prisilo (v izvedbi delovalca), strukturno nasilje deluje prek družbenih struktur, sistemov, politik, norm in ustvarja ali ohranja pogoje, zaradi katerih so določene skupine prebivalstva prikrajšane.

Vidik strukturnega nasilja je v tem smislu tudi temeljna podmena za razvoj pojma algoritemsko nasilje, kot ga predlaga in oblikuje Mimi Onuoha, pri čemer navaja nekaj povezav in primerov. Med drugim rabo izraza algoritemsko nasilje utemeljuje s tem, da »potrebujemo izraz, ki obravnava novejšo, pogosto digitalno in s podatki podprto obliko neenakosti. Kot prvi korak pri artikulaciji teh iskanj želim postaviti besedno zvezo algoritemsko nasilje.« (2018) Nadaljuje, da pojem algoritemskega nasilja vključuje pojave, ki smo jim bili priča v zadnjem desetletju in ki so takšni, da slonijo na vse večji dostopnosti obsežnih podatkovnih zbirk, napredku v računski moči ter predvsem hitrem razvoju na področjih umetne inteligence in strojnega učenja in uporabi teh tehnologij v hierarhično neenaki družbi. Algoritmčno nasilje zajema tako mikrodogodke kot obsežne in daljnosežne pojave oz. procese. Onuoha opisuje nekatere takšne pojave:

»To je tisti neprijeten občutek, ki ga dobite, ko si na spletu ogledate majico, nato pa se ta majica do konca dneva oglašuje na vseh spletnih straneh, ki jih obiščete. /.../ To je bistvo frustracij, ki jih čutijo vozniki Uberja, Lyfta in Juna, ko jim njihove aplikacije narekujejo, naj opravljajo navidezno smiselne prevoze ali naj lovijo ponudbe ›Odlična priložnost!‹ ali ›Ravno v pravem času‹, zaradi česar na koncu pravzaprav prejmejo nižjo plačo. /.../ V okvir pojma algoritemsko nasilje bi morali uvrstiti tudi nekatera dejanja Facebooka, na primer ko je podjetje uporabnikom predpisalo, da se morajo v račune prijaviti s praviimi imeni, a je deaktiviralo račune ljudi, katerih imena se niso zdela legitimna. /.../ Omenili bi lahko tudi omejitve za iskalce zaposlitve – življenjepise jim avtomatski sistemi za sledenje vlogam zavrnejo, ker v njih ni ›pravih‹ ključnih besed. /.../ Vse to so oblike algoritemskega nasilja.« (Onuoha 2018)

Algoritemskemu nasilju je s strukturnim nasiljem skupna predvsem poudarjena odsotnost jasno prepoznavnega subjekta ali povzročitelja nasilja ter predhoden obstoj marginaliziranih skupin, ki jih tovrstno nasilje še posebej prizadene.

2.2 Birokratizacija in moderno upravljanje

Druga podmena pobud za uveljavitev pojma algoritemskega nasilja je povezana s področjem birokratizacije in modernega upravljanja (Huws et al. 2023). Avtorji kot del tega okvira pogosto navajajo pojave, kot so organizacijska kultura s standardiziranimi in enoznačne cilje določujočimi pravili, metrikami in rigidnimi postopki nadzora in vrednotenja, strukturami vodenja in predpisi glede vedenja oz. izpol-

njevanja vlog. V tem smislu so algoritmizacija, avtomatizacija in digitalizacija del istega procesa, ki najprej zajame samo delo, potem pa tudi celotno posameznikovo življenje. S tega vidika je algoritemsko nasilje razumljeno kot »praksa usmerjanja delavca ali skupine delavcev s spodkopavanjem, izključevanjem, nerazumnimi zahtevami in, kar je ključno, z uporabo birokratskih in algoritmičnih ukrepov nadzora in ocenjevanja« (99), ki privede do nezaželenih, pogosto lahko tudi zelo težkih posledic za zaposlene. V tem smislu je algoritemsko nasilje povezano z algoritemskim upravljanjem, imenovanim tudi ,algokracija': algoritem v postopku ali procesu predpiše niz korakov, ki izključujejo vse odklone ali drugačne posege, s tem pa so povezane tudi prakse kaznovanja in izključevanja (103). Gre za upravljanje področij, ki posameznikovo življenje bistveno določajo (npr. vrednotenje dela, promet, upravni postopki idr.).

Ta teoretski okvir, ki je v ozadju pojma algoritmično nasilje, se opira predvsem na dve vrsti nasilja, in sicer na birokratsko nasilje oz. tehnokratsko nasilje (Bauman 2006; Arendt 1969) in počasno nasilje (Nixon 2006). Arendt (1969) je birokracijo oz. njeno prevlado izpostavila kot prefinjen sistem dominacije, kjer se odgovornost, ki je ni več moč pripisati nikomur, porazgubi, zato pa tudi gre za obliko vlade nikogar. V tem smislu gre tudi za obliko tiranije, saj slednjo določa prav to, da manjka vsakršna odgovornost vladajočega. Izpostavila je še, da večja kot je birokratizacija javnega življenja, večja je težnja po nasilju. »V popolnoma razviti birokraciji ni več nikogar, s komer bi se lahko prepirali, komur bi lahko podali svoje pritožbe, na kogar bi lahko izvajali pritiske moči. Birokracija je oblika vladavine, v kateri so vsi prikrajšani za politično svobodo, za moč delovanja; kajti vladavina Nikogar ni ne-vladavina – in kjer so vsi enako nemočni, imamo tiranijo brez tirana.« (81) Počasno nasilje pa je – drugače kot običajno nasilje, ki je vidno oz. lahko prepoznavno ter povezano z dogodkom ali dejanjem – razumljeno kot nasilje, ki se dogaja postopoma in neopazno, ki je razpršeno v času in prostoru, njegove posledice pa so določljive šele z zamudo – torej nasilje, ki ga običajno sploh ne obravnavamo kot nasilje (Nixon 2006, 2). Algoritemsko nasilje v tem smislu deli značilnosti tovrstnega nasilja. Novost je, da se plasti birokratskega nasilja in počasnega nasilja dodaja plast delovanja algoritmov umetne inteligence.

2.3 Spoznavno oz. epistemično nasilje

Tretja podmena zagovora uvedbe pojma algoritemsko nasilje je vidik epistemičnega nasilja. Eden izmed načinov, kako je algoritemsko nasilje povezano s spoznavnim nasiljem, so vidiki algoritmov, ki igrajo vlogo popolnoma natančnih opisov, objektivnih ,rešitev' ali racionalnih interpretacij realnosti, ki je predvidljiva, vključno s »posebnimi oblikami znanja in nevednosti, ki jih ti sistemi ustvarjajo« (Bellanova et al. 2021, 123). Preden se vrnemo k algoritemskemu nasilju, si najprej približajmo pojem spoznavnega nasilja. Spivak (1998, Dotson 2011) spoznavno nasilje opredeljuje kot vrsto nasilja, ki poskuša odpraviti znanje oz. vednost, ki jo imajo marginalni subjekti. Ena od vrst takšnega spoznavnega nasilja se na primer pojavlja na področju začenja. Dotson (2011, 238) tovrstno nasilje označuje kot nasilje, ko občinstvo/poslušalstvo (namerno ali nenamerno) zavrača vzajemnost

v jezikovni, komunikacijski izmenjavi, ta zavrnitev pa temelji na ‚škodljivi nevednosti‘, ki je zanesljiva nevednost, nastajajoča zaradi spoznavnih vrzeli. Lahko nastopa v obliki utišanja ali podrejanja. Takšna zastavitev ima izhodišče v razpravah o spoznavni nepravčnosti.

Spoznavna nepravčnost je celostno uvedena in utemeljena v knjigi Mirande Fricker z naslovom *Epistemic Injustice: Power and the Ethics of Knowing* (Fricker 2007; Strahovnik 2018). Da bi lahko spoznavno nepravčnost ustrezno razumeli, se moramo osredotočiti na spoznavne prakse spoznavnih akterjev, ki so umeščene v okvir družbe, družbenih odnosov, odnosov moči in njihove družbene identitete. Ob tem Fricker izpostavlja metodološko primarnost spoznavne nepravčnosti, ki misli o (spoznavni) pravičnosti predhodi – z osredotočanjem na pojavne oblike spoznavne nepravčnosti njeno bistvo razumemo lažje, kot pa če bi pozornost usmerjali le na vrzeli v pravičnosti oziroma jo razumeli kot manko pravičnosti. Fricker razlikuje med dvema vrstama spoznavne nepravčnosti, in sicer med začetvalno nepravčnostjo in hermenevitično nepravčnostjo. Jasen primer prve je denimo nezaupanje ali neupoštevanje koga zgolj na podlagi njegove rase, spola, družbenega položaja ipd. Takšen posameznik s tega vidika trpi spoznavno nepravčnost, saj je izključen iz družbenih spoznavnih praks, obenem pa se mu odreka vloga nosilca vednosti. Hermenevitična nepravčnost pa se vzpostavlja tam, kjer prihaja do vrzeli v hermenevitičnih virih, ki so na voljo za vednost in razumevanje – kjer na primer posamezniki ali družbene skupine nimajo orodij za doseganje vednosti in ustrezno razumevanje npr. lastnega izkustva (Fricker 2007 9–29; 147–175).

Zelo pomembno je, da izpostavimo značilno spoznavno naravo take nepravčnosti, ki sega onkraj vprašanj o samem dostopu do informacij, vednosti ali dostopnosti šolanja za posameznike ali skupine. Pri spoznavni nepravčnosti gre za pristno spoznavno nepravčnost, ki je bistveno povezana s spoznavnimi praksami, kot sta prenašanje vednosti na druge preko pričevanja ter oblikovanje ustreznega razumevanja našega lastnega izkustva (1). Da bi zgoraj omenjeni obliki spoznavne nepravčnosti lažje razumeli, moramo najprej razumeti pojmovanji moči in identitete, s katerima jih lahko opredelimo. Moč je družbeno umeščena zmožnost nadzora in vpliva na dejavnost drugega oziroma drugih. Pri hermenevitični obliki spoznavne nepravčnosti razumevanje onemogoča ali hromi hermenevitična vrzel – običajno je povezana z različnimi družbenimi identitetami, različne skupine v družbi pa prizadene različno. Velja, da se posamezne skupine hermenevitično marginalizira s krnjenjem njihove udeležbe v spoznavnih in družbenih praksah, ki ustvarjajo skupnostne smisle. Spoznavna nepravčnost pa privede do spoznavne škode, ki koga prizadene kot subjekta oziroma nosilca vednosti ali pa katere druge spoznavne vrednote oz. cilja (Strahovnik 2018). Fricker kot rešitev za preseganje spoznavne nepravčnosti predlaga razvijanje in kultiviranje specifičnih spoznavnih kreposti. Začetvalsko občutljivost tako izpostavlja kot obliko razumske občutljivosti. Razvijamo jo lahko s pozornostjo na začetvalske prakse, ustreznim oz. zanesljivim pripisovanjem kredibilnosti, pozorno zaznavo in dojemanjem spoznavnega statusa ipd. Dopolnjuje jo krepost hermenevitične pravičnosti: ta predpostavlja, da po-

skušamo kritično premisliti, smo pozorni in občutljivi na morebitne hermenevitične vrzeli pri govorcu – in svoje razumevanje temu prilagodimo (Fricker 2007, 5–7). V nekaterih okvirih razprave o spoznavni nepravilnosti srečamo tudi govor o spoznavnem nasilju (McKinnon 2016; Dotson 2011). Spoznavno nasilje je to, do česar spoznavna nepravilnost lahko privede oziroma vanj preide. Spoznavno nasilje med drugim vključuje spoznavno zatiranje ali utišanje. In na tej točki se v okvir naše razprave vpne tudi vpeljava pojma algoritemsko nasilje.

Mbalaka (2023) govori o spoznavnonasilnih pristranskostih pri oblikovanju in delovanju sistemov umetne inteligence in opozarja, kako lahko sistemi umetne inteligence vodijo do spoznavnega nasilja. To prikazuje s primerom delovanja algoritmov umetne inteligence, ki so del modelov Dall-E 2 in Starry AI za generiranje slik oz. pretvorbo besedilnega poziva v sliko. Trdi, da sta oba modela zasnovana tako, da odražata implicitne predsodke sodobne družbe in jih s svojim delovanjem ohranjata. Algoritemska pristranskost torej lahko povzroči algoritemsko spoznavno nasilje, ki vključuje pojave, kot so zatiranje različnih stališč, premajhna zastopanost določenih sistemov vednosti in zniževanje njihovega pomena. Kot primere tovrstnega nasilja izpostavlja pogostost pomanjkljivih in netočnih prikazov tipične temnopolte družine, pa tudi generične ali halucinirane upodobitve določenih kulturnih artefaktov (npr. tipičnih vzorcev na oblačilih, drugega okrasja in predmetov). Ker so v te fenomene bistveno vpeti algoritmi umetne inteligence, Mbalaka tovrstno nasilje povezuje z njimi.

3. Skeptični premisleki glede pojma algoritemsko nasilje

Preden preidem k bolj specifičnim skeptičnim premislekom o pojmu algoritemskega nasilja, naj opozorim, da lahko pri pogledu na razumevanje določene vrste nasilja upoštevamo naslednje tri temeljne vidike: (i) kdo ali kaj je povzročitelj nasilja; (ii) način, na katerega delovalec ali povzročitelj nasilje izvaja ali ohranja škodo, neenakost itd.; (iii) vrsto nasilja z vidika posebnih posledic oz. učinka na žrtve nasilja. Ti trije vidiki se medsebojno ne izključujejo in se v več primerih združujejo ali prekrivajo. Vsi trije gradniki so načeloma lahko podlaga za smiselnost uvedbe pojma algoritemsko nasilje.

Izhodiščni skeptični premiselek lahko uokvirimo v zvezi z zelo širokim razumevanjem izraza ‚algoritemsko nasilje‘. Nekateri avtorji algoritemsko nasilje v razpravi uporabljajo v širokem pomenu, ki temelji na ideji, da je vsakršna raba algoritmov in s tem delovanje algoritmov samo po sebi nasilno. Podobno široke so tudi razlage, ki prakse datafikacije in digitalizacije obravnavajo kot že obremenjene z nasiljem. V obeh primerih delovanje algoritmov, ki temelji na umetni inteligenci, sploh ne more biti nenasilno (Bellanova idr. 2021, 126). Algoritemsko nasilje je vseprisotno in neizogibno, posledično pa je koncept algoritemskega nasilja v najboljšem primeru dvoumen ali v najslabšem nesmiseln – v nobenem primeru pa ne more biti primeren za poskus razumevanja in blaženja različnih škodljivih učinkov, ki jih uporaba algoritmov umetne inteligence povzroča ali bi jih lahko povzročila. Hkrati ni povsem jasno, komu ali čemu ter v kakšni obliki je po tem širokem

razumevanju škoda povzročena. Tretja težava takšnega razumevanja nasilja, ki da ga algoritmi povzročajo, je, da pušča odprto vprašanje, kako potem razumeti in upoštevati pozitivno in koristno uporabo algoritmov – npr. pri prepoznavanju, preprečevanju ali zdravljenju bolezni, pri varovanju okolja ipd. (Strahovnik, Miklavčič in Centa Strahovnik 2020), kar se vsaj po nekaterih kriterijih lahko upravičeno zdi zaupanja vredno (Miklavčič 2021).

Zagovorniki izraza bi odgovor lahko oblikovali v naslednji smeri, in sicer, da je »kritična dodana vrednost algoritemskega nasilja (ta), da lahko opozori na prakse, ki niso nasilne v ‚tradicionalnem‘ smislu in jih zato ni mogoče obravnavati kot takšne, če želimo izpolniti to merilo.« (Bellanova et al. 2021, 126) Temu lahko pritrudimo, a vseeno se lahko upravičeno vprašamo o smiselnosti izbire tega izraza, ne da bi bili predstavljeni dobri primeri takšnega algoritemskega nasilja, ki jih ne moremo opisati na drugačen način oz. z rabo drugih izrazov.

Druga možnost obrambe koncepta algoritemskega nasilja je naslednja: algoritmi, ki temeljijo na umetni inteligenci, bi lahko nasilno delovanje spodbujali ali pa se uporabljali kot sredstva zanj. S tem povezana konceptualizacija algoritemskega nasilja je torej nasilje, ki se doseže s temi *sredstvi*. Šlo bi za situacije, ki vključujejo zanašanje na algoritme kot pomoč pri sprejemanju odločitev v kontekstih, v katerih posledice takšnih odločitev posameznikom ali skupnostim nedvomno lahko povzročijo škodo. Primer bi bili sistemi ‚točkovanja‘ v kazenskem pravosodju oz. kot del kaznovanja ali pa uporaba polavtonomnega orožja. Tako algoritmi umetne inteligence omogočajo nasilna dejanja, izvedena na daljavo, kot je to v primeru vojaške tehnologije, ki vključuje daljinsko vodene ali polavtonomne oborožitvene sisteme (Bellanova et al. 2021, 127).

Vendar se ob upoštevanju tega odgovora znova pojavi vprašanje implementacije, upravljanja in zastopanja. Najbolj naravno je razumeti, da algoritmi sami nimajo sposobnosti delovanja ali namena – nasprotno, so orodja, ki jih ustvarijo in uporabljajo ljudje. Zato lahko pripisovanje nasilja algoritmom zamegli vlogo in odgovornost človeških odločevalcev za rezultate ali posledice delovanja algoritemskih sistemov. Tovrsten temelj za določitev algoritemskega nasilja kot posebne vrste nasilja pa tu umanjka. Pomembno je osvetliti tudi pretiran poudarek na nekakšnem tehnološkem determinizmu kot prepričanju, da tehnologija družbene izide določa neodvisno od človekovega delovanja. Osredotočanje le na algoritme kot vir družbenih težav spregleduje širše družbene, gospodarske in politične dejavnike, ki k neenakostim in nepravilnostim vsi prispevajo. Algoritmi umetne inteligence so precej kompleksni in večplastni sistemi, zato je škodo težko pripisati zgolj algoritmom – izid in njegove posledice določa interakcija med algoritmi, zbirkami podatkov, uporabniki in družbenim kontekstom.

Vredno je opozoriti tudi na naslednje razlikovanje, ki ga lahko ponazorimo na primeru utemeljitve pojma algoritemskega nasilja s pojmom spoznavnega nasilja. Za oceno smiselnosti rabe algoritemskega nasilja je koristno razlikovati med dvema glavnima perspektivama ali možnostma. Prva možnost je, da se spoznavno nasilje, ki je pripisano sistemom umetne inteligence, prenese iz prejšnjega spo-

znavnega nasilja, ki ne temelji (nujno) na umetni inteligenci ali njeni rabi, temveč je povezano s pristranskostmi pri zbiranju, skrbništvu in uporabi pristranskih učnih podatkovnih nizov. Pri generativnih sistemih umetne inteligence gre preprosto za to, da so bili naučeni na nizu vhodnih podatkov, ki je privilegiral določene skupine oz. značilnosti. Druga možnost pa je spoznavno algoritemsko nasilje, ki se pojavlja predvsem zaradi samega delovanja algoritmov umetne inteligence (bodisi ker so ali bi bili namerno zasnovani kot pristranski bodisi kot stranski produkt optimizacije algoritmov, npr. za hitrost). Vendar avtorji (Mbalaka 2023) ne navajajo primerov, ki bi ta vidik podpirali.

O vsem zgoraj omenjenem bi lahko imeli tudi naslednji pomislek. Ne moremo biti skeptični oz. pojma algoritemskega nasilja zavračati, če prej sami ne ponudimo opredelitve nasilja in na tej podlagi dokažemo, da algoritemsko nasilje okvirom te opredelitve ne zadosti. Odziv na ta izziv je dvojen oz. ima dva kraka.

Pri prvem kraku izberemo določeno definicijo nasilja in potem posledično ugotavljamo, ali je glede na to definicijo o algoritemskem nasilju upravičeno govoriti. Ta strategija je sicer možna, vendar v pomembnem smislu predstavlja zmoto zahteve po vprašanju (angl. *begging the question*). Ena izmed bolj razširjenih in uporabljenih definicij nasilja je tista, ki jo je podala Svetovna zdravstvena organizacija. Ta (WHO 1996; cf. WHO 2002, 5) nasilje opredeljuje kot »namerno uporabo fizične sile ali moči (v obliki grožnje ali dejansko) proti sebi, drugi osebi, skupini ali skupnosti, ki ima ali bi lahko imela za posledico poškodbe, smrt, psihično škodo, slab razvoj ali prikrajšanje.« Nadalje WHO nasilje deli na tri kategorije glede na to, kdo ga izvaja (samopovzročeno oz. nase usmerjeno, medosebno in kolektivno), in štiri nadaljnje kategorije glede na vrsto nasilja (fizično, spolno, psihološko ali nasilje, ki vključuje prikrajšanje ali zanemarjanje) (WHO 2002, 6; Dwyer 2022). Glede na to razmeroma ozko opredelitev nasilja večine pojavov, ki jih avtorji zgoraj podajajo kot primere algoritemskega nasilja, ne moremo označiti kot nasilje.

Pri drugem kraku izhajamo iz opredelitev oz. pojmovanj nasilja nasploh, ki jih uporabljajo oz. na katere se sklicujejo sami avtorji, ki zagovarjajo vpeljavo pojma algoritemsko nasilje. Če se najprej posvetimo vidiku strukturnega nasilja, potem lahko izpostavimo, da Galtung (1969) ne izhaja iz kakšnega splošnega pojma nasilja, ampak pravzaprav strukturno nasilje in s tem razširjeno definicijo nasilja uvaja prav zato, ker zanj običajni pojem nasilja ni ustrezen – ne zajame namreč celote nasilja (kot nasprotja miru) oz. vseh njegovih razsežnosti. Odpira se vprašanje, ali je to ustrezna podlaga tudi za uvedbo pojma algoritemsko nasilje. Pri strukturnem nasilju gre za nasilje, ki je posredno in ki je vgrajeno v družbeno strukturo. Seveda so lahko del družbene strukture in delovanja institucij družbe tudi algoritmi umetne inteligence, vendar ti do nasilja privedejo le v primeru, ko ima takšne učinke struktura kot celota. Algoritmi umetne inteligence nikoli ne predstavljajo strukture kot take, ampak so kvečjemu njen del oz. služijo kot morebitno sredstvo strukturnega nasilja. Za pojem strukturnega nasilja je Galtung predlagal tudi rabo izraza družbena nepravilnost (Galtung 1969, 171). Podobne pomisleke imamo pri podmeni, ki se pri zagovoru rabe pojma algoritemsko nasilje naslanja na vidik birokratizacije in administracije. Birokratizacija in birokratsko upravljanje seveda lahko

privedeta tudi na nasilja (kakor to pri holokavstu izpostavlja Bauman [2006]), a to ne pomeni, da je vsakršna birokratizacija že nasilna oz. sama po sebi nasilje. In tudi če pogledamo zagovor pojma algoritemsko nasilje поближе (Huws et al. 2023), lahko vidimo, da ključni poudarek v zvezi z negativnimi pojavi, ki so omenjeni (npr. psihološka škoda, izgorelost, depresija, ipd.), ni na sistemih umetne inteligence, temveč na razvoju upravljalvske strukture in tudi širše organizacijske kulture, ki tovrstno nasilje omogoča oz. do teh negativnih pojavov privede.

Za konec se lahko obrnemo še k tretji podmeni, to je polju spoznavne pravičnosti. Kot smo že omenili, je govor o spoznavnem nasilju tesno povezan z razpravo o spoznavni (ne)pravičnosti. Vendar si je smiselno podrobneje ogledati, na kaj se ta izraz sploh nanaša. Kot je izpostavljeno že zgoraj, se z izrazom spoznavno nasilje najpogosteje označuje prakse utišanja, nerazumevanja in zatiranja, ki so povezane s sistemi vednosti in subjekti kot nosilci vednosti. Eden izmed primerov, ki ga omenja Dotson (2011, 240), je primer triletnega otroka, ki zavrne nasvet oz. opozorilo odraslega, da je ogenj nevaren in da se z njim ne sme igrati. Dotson na podlagi različic tega primera trdi, da do spoznavnega nasilja, kadar ga vrši otrok, pride v primerih, ko je odrasli prizadet kot nosilec vednosti ali pa ko zaradi otrokovega ravnanja pride do škode (npr. materialne škode ob požaru). Upravičeno se lahko vprašamo, ali ne gre v teh in podobnih primerih za neustrezno in neupravičeno široko razumevanje pojma spoznavnega nasilja oz. ali ne bi bilo ustreznejše izbrati drug opis. Če denimo skupnost zaničevanju osebe ne verjame zgolj na podlagi njenega spola ali rase, potem je to lahko posledica siceršnje diskriminacije, predsodkov, predhodnega strukturnega nasilja ipd. in v tem smislu vrsta spoznavne nepravičnosti. In če tovrstne situacije označimo kot spoznavno nasilje, ne dodamo nobenih pomembnih določil oz. podlage za ustreznejše razumevanje situacije. Podobno lahko trdimo tudi za pojem algoritemskega nasilja. Če nam algoritem na podlagi majhnega vzorca našega vedenja ali podatkov, ki jih o nas ima/zbira, vztrajno ponuja oglas za določeno stvar ali storitev, je to lahko odraz predsodkov v učnih podatkih ali pridobljenih vzorcev obnašanja potrošnikov – ne pa nujno algoritemsko nasilje.

4. Zaključek: predsodki, nepravičnost in škoda v digitalni družbi

Preidimo k sklepu. Uporaba drugih, bolj specifičnih izrazov v zvezi z razvojem, izvajanjem in delovanjem algoritmov umetne inteligence – v nasprotju s splošnim izrazom algoritemsko nasilje – se pri prihodnjem urejanju algoritmov in določanju odgovornosti za njihovo uporabo kaže kot koristnejša.

„Akt EU o umetni inteligenci“ (angl. *AI Act*; EU 2024) izraza algoritemsko nasilje ne vsebuje; uporablja izraze, kot so manipulacija, izkoriščanje, tveganje (npr. tveganje škode za zdravje in varnost, tveganje napačnih ali pristranskih odločitev na temelju umetne inteligence), škoda, pristranska obravnava, (ne)transparentnost (Miklavčič 2023) itd. Med prakse oz. načine delovanja ter uporabe umetne inteligence, ki jih zakon prepoveduje, spadajo subliminalna manipulacija in načrtno

izkoriščanje človeške ranljivosti (npr. uporaba manipulativnih ali zavajajočih tehnik, ki spodbujajo ali ovirajo posameznikovo avtonomijo, odločanje ali svobodno izbiro tako, da se ljudje teh tehnik ne zavedajo), povzročanje fizične ali psihične škode oz. škode za zdravje in varnost, ki bi negativno vplivala na temeljne pravice, družbeni nadzor ipd. Med drugim Akt v okviru svojih izhodišč določa naslednje: »Raznolikost, nediskriminacija in pravičnost pomeni, da se sistemi UI razvijajo in uporabljajo na način, ki vključuje različne akterje in spodbuja enakopraven dostop, enakost spolov in kulturno raznolikost, hkrati pa se preprečujejo diskriminatorni učinki in nepravilni predsodki, ki jih pravo Unije ali držav članic prepoveduje.« (EU 2004) Taka so torej bolj določna izhodišča o tveganjih in nevarnosti diskriminacije.

V okviru zagovora pojma algoritemskega nasilja je bila ena od podmen vezana na vidik birokratizacije, vključno s postopki vrednotenja obsega oz. učinkov dela – negativni vplivi so označeni kot nasilje. „Akt EU o UI“ je bolj določen in jasno izpostavlja, zakaj gre pri takšni rabi UI za visoka tveganja in kakšna so ta tveganja. Določa, da sistemi

»UI, ki se uporabljajo pri zaposlovanju, upravljanju delavcev in dostopu do samozaposlitve, zlasti za zaposlovanje in izbor oseb, za sprejemanje odločitev, ki vplivajo na delovne pogoje, napredovanje ali prenehanje pogodbenih delovnih razmerij, za dodeljevanje nalog na podlagi individualnega vedenja, osebnih lastnosti ali značilnosti ter za spremljanje ali ocenjevanje oseb v z delom povezanih pogodbenih razmerij.«

predstavljajo sisteme visokega tveganja – še posebej zato, ker lahko takšni sistemi ohranjajo vzorce diskriminacije iz preteklosti, ki so povezani s spolom, starostno skupino, invalidnostjo, rasnim ali etničnim poreklom, spolno usmerjenostjo idr., in njihova uporaba ogroža temeljne pravice do varstva podatkov in zasebnosti. Takšen pristop, ki tveganja, nedovoljene prakse ter ukrepe za njihovo zmanjševanje oz. kaznovanje poimenuje določneje, se zdi bolj smiselni kot govor o algoritemskem nasilju.

Reference

- Arendt, Hannah.** 1969. *On Violence*. New York: Harcourt Brace Jovanovich Publishers.
- Bauman, Zygmund.** 2006. *Moderna in holokavst*. Ljubljana: Študentska založba.
- Bellanova, Rocco, Kristina Irion, Katja Lindskov Jacobsen, Francesco Ragazzi, Rune Saugmann in Lucy Suchman.** 2021. Toward a Critique of Algorithmic Violence. *International Political Sociology* 15:121–150.
- Bostrom, Nick, in Eliezer Yudkowsky.** 2014. The ethics of artificial intelligence. V: Keith Frankish in William Ramsey, ur. *Cambridge Handbook of Artificial Intelligence*, 316–334. New York: Cambridge University Press.
- Dotson, Kristie.** 2011. Tracking epistemic violence, tracking patterns of silencing. *Hypatia* 26, št. 2:236–257.
- Dwyer, Philip.** 2022. *Violence*. New York: Oxford University Press.
- EU.** 2004. Akt o umetni inteligenci. Lex Europa. <https://eur-lex.europa.eu/legal-content/SL/TXT/HTML/?uri=CELEX:52021PC0206> (pridobljeno 23. 5. 2024)
- Fricker, Miranda.** 2003. Epistemic Injustice and a Role for Virtue in the Politics of Knowing, *Metaphilosophy*, 34 št. 1-2:54–73.

- — —. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.
- Galtung, Johan.** 1969. Violence, Peace, and Peace Research. *Journal of Peace Research* 6, št. 3:167–191.
- Huws, Ursula, in trije anonimni soavtorji_ice.** 2024. Algorithmic violence: Towards an interdisciplinary understanding of bullying in academia. *Work Organisation, Labour & Globalisation* 18, št. 1: 97–122. <https://doi.org/10.13169/workorglaboglob.18.1.0097>
- Kockelman, Paul.** 2020. The Epistemic and Performative Dynamics of Machine Learning Praxis. *Signs and Society* 8, št. 2:319–355. <https://doi.org/10.1086/708249>
- Mbalaka, Blessing.** 2023. Epistemically violent biases in artificial intelligence design: the case of DALL-E 2 and Starry AI. *Digital Transformation and Society* 2, št. 4:376–402.
- McKinnon, Rachel.** 2016. Epistemic Injustice. *Philosophy Compass* 11, št. 8: 437–446.
- Miklavčič, Jonas.** 2021. Zaupanje in uspešnost umetne inteligence v medicini. *Bogoslovni vestnik* 81, št. 4:935–946.
- — —. 2023. Ideal transparentnosti v digitalni dobi. *Bogoslovni vestnik* 83, št. 4:825–838.
- Nixon, Rob.** 2011. *Slow Violence and the Environmentalism of the Poor*. Cambridge, MA: Harvard University Press.
- Onuoha, Mimi.** 2018. Algorithmic Violence. GitHub. <https://github.com/MimiOnuoha/On-Algorithmic-Violence> (pridobljeno 21. 7. 2024).
- Raaijmakers, Stephan.** 2022. *Deep learning for natural language processing*. Shelter Island, NY: Manning.
- Safransky, Sara.** 2019. Geographies Of Algorithmic Violence: Redlining the Smart City. *International Journal of Urban and Regional Research* 44, št. 2:200–218.
- Spivak, Gayatri.** 1998. Can the subaltern speak? V: Cary Nelson in Lawrence Grossberg, ur. *Marxism and the interpretation of culture*, 271–313. Urbana: University of Illinois Pres
- Strahovnik, Vojko.** 2018. Spoznavna (ne)pravičnost, krepost spoznavne ponižnosti in monoteizem. *Bogoslovni vestnik* 78, št. 2:299–311.
- — —. 2023. Etični in teološki izzivi velikih jezikovnih modelov. *Bogoslovni vestnik* 83, št. 4:839–852. <https://doi.org/10.34291/BV2023/04/Strahovnik>
- Strahovnik, Vojko, Jonas Miklavčič in Mateja Centa Strahovnik.** 2020. Etični vidiki uporabe algoritemskega odločanja in ostalih sistemov UI v času pandemij oz. izrednih razmer. *Bogoslovni vestnik*, 80, št. 2:321–334. <https://doi.org/10.34291/BV2020/02/Strahovnik>
- Weidinger, Laura, Jonathan Uesato, Maribeth Rauh, Conor Griffin, Po-Sen Huang, John Mellor, Amelia Glaese, Myra Cheng, Borja Balle, Atosa 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. V: *2022 ACM Conference on Fairness, Accountability, and Transparency (FACCT '22)*, 214–229. New York, NY: Association for Computing Machinery.
- WHO.** 1996. *WHO Global Consultation on Violence and Health. Violence: A Public Health Priority*. Ženeva: World Health Organization.
- — —. 2002. *The world report on violence and health*. Uredili Etienne G.Krug, Linda L. Dahlberg, James A. Mercy, Anthony B. Zwi in Rafael Lozano. Ženeva: World Health Organization.
- Žalec, Bojan.** 2023. Ali je umetna inteligenca inteligenca v pravem pomenu besede? Vprašanje psihičnih značilnosti in splošnost. *Bogoslovni vestnik* 83, št. 4:813–823. <https://doi.org/10.34291/BV2023/04/Zalec>

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 797—806

Besedilo prejeto/Received:09/2024; sprejeto/Accepted:10/2024

UDK/UDC: 111:004.89

DOI: 10.34291/BV2024/04/Gregorčič

© 2024 Gregorčič, CC BY 4.0

Rok Gregorčič

Razmerje med človeško in umetno inteligenco v luči Habermasove teorije

The Relation between Human and Artificial Intelligence in Light of Habermas's Theory

Povzetek: Prispevek namerava raziskati razmerje med človeško in umetno inteligenco v luči teorije komunikativne racionalnosti. V prvem delu analize so predstavljena Habermasova teoretična izhodišča. Izpostavljena je predvsem Habermasova teoretizacija racionalnosti ter premik od monološke h komunikativni racionalnosti. V nadaljevanju se osredotočamo na pojav umetne inteligence (UI) ter na razlikovanje med induktivnimi in deduktivnimi sistemi umetne inteligence. V zvezi s tem osvetljujemo tudi zgodovinsko dogajanje v pogledu na prvi in drugi val raziskovanja UI v preteklem stoletju. Obravnavamo tezo, da je prehod od simbolne UI h konekcionizmu mogoče interpretirati v navezavi na Habermasov prehod od monološke h komunikativni racionalnosti. V zadnjem delu prispevek prikazuje, kako bi Habermasovo teorijo v odnosu med človeško in umetno inteligenco aplicirali še drugače. Pričujoča analiza s pomočjo Habermasove kritike liberalne evgenike oblikuje stališče, ki lahko pomaga pri trajni razmejitvi med človekom in UI.

Gljučne besede: Jürgen Habermas, komunikativna racionalnost, umetna inteligenca, simbolizem, konekcionizem, genetska manipulacija

Abstract: The aim of this article is to investigate the relation between human and artificial intelligence in light of the theory of communicative rationality. In the first part, Habermas's theoretical starting points are presented. Primarily, the transition from a monological to a communicative form of rationality is emphasized. In the next step, the focal point is moved toward the distinction between inductive and deductive systems of AI. In relation to this, a historical examination of the first and second wave of AI research in the previous century is also presented. A thesis that interprets the transition from symbolic AI to connectionism in correlation to Habermas's transition from monological to communicative rationality is thoroughly analyzed. Towards the end of the article, another way of applying Habermas's theory to the relation between human and artificial intelligence is presented. Deriving from Habermas's criticism of liberal eugenics, a well-argued position emerges that could help us distinguish between humans and AI in the long run.

Keywords: Jürgen Habermas, communicative rationality, artificial intelligence, symbolism, connectionism, genetic manipulation

1. Uvod

Jürgen Habermas velja za enega najprodornejših sodobnih mislecev.¹ Njegova družbena in filozofska misel je znana po vsem svetu. Zasnoval je teorijo komunikativne racionalnosti, ki predstavlja pomembno teoretično podlago za sodobne oblike družbenega sobivanja, s tem pa nudi filozofsko ogrodje za nekatere aktualne demokratične oblike političnih praks. Habermas v sodobnem svetu nedvomno spada med najuglednejše filozofske avtoritete. Odmevnost njegove misli je po eni stani posledica njegove občutljivosti za dialoško vključevanje vseh, po drugi strani pa zavzemanja za racionalnost in univerzalnost. Lahko bi rekli, da Habermasova misel uteleša poskus, kako vrlino racionalnosti, izhajajočo iz razsvetljenstva, uglasti v sozvočje z vrlino komunikacije, ki je posebno mesto dobila v filozofski postmodernosti. V tej luči lahko razumemo tudi koncept komunikativnega delovanja, ki predstavlja temelj Habermasove teorije (Habermas 1984). Gre za obliko družbenega delovanja, pri kateri se racionalnost uresničuje na komunikativen način.

Za Habermasovo teorijo racionalnosti je značilen odmik od monoloških modelov racionalnosti, ki so temeljili bodisi na zavesti posameznika bodisi na strogi znanstveni dokazljivosti. Premik, ki ga v odnosu do razumevanja vidimo pri Habermasu, bi na poseben način lahko primerjali s premikom v raziskovanju umetne inteligence (UI): strokovnjaki in raziskovalci, ki se ukvarjajo z UI, so večinsko opustili proučevanje deduktivno zasnovanih modelov UI in se usmerili k razvoju induktivno zasnovanih modelov UI. Zastavlja se vprašanje, kako lahko premik v dojemljanju racionalnosti, kot je opazen pri Habermasu, primerjamo s premikom na področju različnih modelov UI. Še bolj pereče vprašanje pa je, kako nam lahko Habermasova teorija pomaga pri iskanju oz. dojemljanju pravega razmerja med človeško in umetno inteligenco.

2. Habermasov odmik od monološke racionalnosti in pozitivizma

Habermas je teorijo racionalnosti zasnoval na podlagi analize pomembnih teorij racionalnosti iz preteklosti. Obširni del svojega znanstvenega prispevka je oprl na pregledno obravnavo starih teorij racionalnost, pri vsaki izpostavil pomanjkljivosti in nemalokrat dodal svoj lastni predlog izboljšav (Habermas 1984, 1–7). Ta vzorec je posebej izrazit pri Habermasovi obravnavi Kanta. Kantova misel po mnenju mnogih v Habermasovi teoretizaciji velja za vodilen vpliv (Hovdelien 2011, 111).

¹ Prispevek je nastal v okviru raziskovalnega programa P6-0269 „Religija, etika, edukacija in izzivi sodobne družbe“, ki ga sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

Habermas je menil, da Kantova filozofija pri obrambi racionalnosti v mnogočem sicer predstavlja pomemben korak, vendar se v sodobnem svetu v soočenju s pluralno družbeno realnostjo ne izkaže zadosti učinkovito (Gregorčič 2023, 912). Zardi tega Habermas do Kanta izkazuje neke vrste ambivalenten odnos: po eni strani se na njegove ideje vedno znova sklicuje, po drugi pa v mnogih potezah Kantovo mišljenje močno presega ali pa se od njega vsaj zmerno distancira. Očiten primer najdemo v delu *Resnica in justifikacija*. Habermas je tu predstavil štiri domnevne ustreznice med lastno teorijo formalnega pragmatizma in Kantovo transcendentno filozofijo – idejo enotnosti sveta, postulat svobode, idejo vsepresegajočega uma in avtoriteto justifikacije (Habermas 2003, 87). Iz zapisanega je vidno, da Habermas v širšem smislu vsakega od teh pojmov v svojo teorijo vključuje, v ožjem smislu pa vsakega rekonstruira in vsebinsko prilagaja komunikativni paradigmi. Sam je svoj odnos do Kantove filozofije izrazil s pojmom detranscendentalizacija (175). Izraz po Habermasovih besedah pomeni »vključitev umnih subjektov v kontekst življenjskega sveta in po drugi strani v prepletenost znanja z govorom in delovanjem« (88). Na podoben način kot pri Kantu je Habermas utemeljil rekonstrukcijo mnogih drugih filozofskih in socioloških teorij. Lahko bi rekli, da je Habermas pretekle avtoritete s pridom izkoristil tako, da je v svojo teorijo vključil nekatere njihove ideje. Hkrati pa je pri vsaki izpostavil pomanjkljivosti, ki jih v svoji teoriji poskušal odpraviti. Tako je želel vzpostaviti prenovljeno teorijo družbene racionalnosti, ki jo odlikuje postmetafizični in komunikativni značaj.

Za Habermasovo teorijo racionalnosti je značilen premik h komunikativni paradigmi, kar pomeni, da pglavitna vsebina komunikativne racionalnosti postane proces komunikativnega delovanja, imenovan tudi deliberacija (premišljujoče odločanje): proces izmenjave argumentov med posamezniki, ki se udeležujejo diskurza o določeni skupni problematiki (1984, 27). Udeleženci med seboj vzpostavijo veljavnostne zahteve, ki postanejo temelj njihovega medsebojnega pripoznanja ter nujna podlaga za možnost sporazumne razrešitve zadeve (22). Veljavnostne zahteve pomenijo, da posameznik v diskurzu sodeluje s pričakovanjem, da bodo drugi udeleženci diskurza (tudi v primeru nestrinjanja) njegove argumente priznali kot veljavne. Ta proces je za Habermasa zelo pomemben zato, ker naj bi se skupaj zbrali ljudje najrazličnejših stališč, se udeležili diskurza in v sporazumu poskušali doseči skupno rešitev (Geiger 2009, 23). Glede na pluralnost družbenih nazorov, ki smo jim danes priča, se Habermasu to zdi edina resnično univerzalna pot reševanja družbenih problemov – pot, ki jo ljudje različnejših nazorov lahko sprejmejo kot način interakcije v sekularni družbi. Habermasu se zdi neustrezno, da bi družbene probleme reševali bodisi po načelu filozofije zavesti (znotraj posameznikovih umskih zmožnosti) bodisi po načelu znanstvenega pozitivizma (strogo vezano na znanstveno metodo). V Habermasovem dojemanju je racionalnost nekaj, kar posameznikov um presega, hkrati pa presega tudi ozkost znanstvenih dokazov. Ob vsem tem je Habermas zavezan ideji praktičnega razuma, ki se loteva konkretne družbene problematike. Pri praktičnem razumu gre za to, da razumevanja ne sme diktirati teorija, ampak mora biti plod skupnega procesa celotne družbe. Racionalnost je vezana na posameznike in njihovo diskurzivno dejavnost, hkrati pa slehernega po-

sameznika presega (Habermas 2002, 82). Stare teorije racionalnosti so medtem trdile, da je razumevanje pridobljeno na podlagi danih fiksnih izhodišč oz. temeljev – od tod izvira izraz fundacionalizem. Habermasova teorija je zastavljena kot alternativa fundacionalizmu (Higgitt 2011, 94). Predvideva namreč skupni univerzalni proces intersubjektivne interakcije – temu sledi dosega sporazuma, ki bo vsa morebitna izhodišča metodično in vsebinsko presegel.

Ker Habermasova teorija močno poudarja živeto prakso, se pojavljajo mnoga vprašanja, kako to teorijo aplicirati na spremenjene razmere digitalne sodobnosti. Mnogi interpreti pravijo, da kriteriji, po katerih Habermas presoja komunikacijo in javnost, zaradi radikalno spremenjene informacijske realnosti niso več primeri; gotovo lahko opazimo, da sodobne oblike digitalnega povezovanja – denimo na novih družbenih omrežjih – Habermasovih standardov univerzalne komunikacije ne izpolnjujejo. Marsikdo zato meni, da bi moral Habermas, da bi lahko ponudil večjo možnost za vključevanje novih oblik, svoje kriterije prilagoditi (Khan, Gilani in Nawaz 2012, 47). Nekateri celo namigujejo, da bi se moral Habermasov teoretični model radikalno preoblikovati – ker da je v aktualnem stanju od sodobnih pojavov, kot je internet, vedno bolj oddaljen (Salikov [s. a.], 92). Mnogi drugi interpreti pa vztrajajo, da se morajo Habermasovi standardi pri konkretni evalvaciji sodobnih tehnoloških pridobitev ohraniti in aplicirati. Teoretizacija univerzalne pragmatike naj bi bila torej primerna in potrebna. Uporabiti bi jo morali predvsem v procesu preverbe, ali vse te nove oblike komunikacije demokratični družbi res koristijo (Ritzi [s. a.], 59). Habermasova načela se npr. lahko izkažejo kot koristna pri presoji različnih oblik medijev (Fisher 2021, 12). Kot kaže, smo priča pereči dilemi: na kakšen način – če to sploh zmore – naj se teorija komunikativne racionalnosti sooči s tehnološkimi in družbenimi novostmi moderne digitalne dobe?

3. Oblike umetne inteligence v luči Habermasove teorije

Ena od možnih točk soočenja med Habermasovo teorijo in digitalno modernostjo je vprašanje umetne inteligence. Iz nekaterih pomembnih vidikov komunikativne teorije namreč lahko izpeljemo vsebinski odziv, ki ga ta v povezavi z umetno inteligenco ponuja. Lahko bi rekli, da tehnološki projekt umetne inteligence, ki smo mu priča v sodobnosti, oscilira okoli želje, da bi izdelali program, ki bi čim bolj posnemal pravo človeško inteligenco. Treba je povedati, da se informacijska tehnologija pri uresničevanju te želje poslužuje različnih načinov. Poznamo dva osnovna načina za razvijanje umetne inteligence. Deduktivni način si inteligenco prizadeva vzpostaviti na podlagi vnaprej danih premis (Takefuji in Yamada 2019). Premise oz. pravila tvorijo strukturo, ki se je model UI strogo drži in zato za učenje novih vzorcev ni odprt. Na tej točki – torej pri vprašanju možnosti učenja – se deduktivni modeli UI od induktivnih najbolj razlikujejo. Induktivni modeli imajo namreč vsaj minimalno zmožnost, da se v interakciji s svojim okoljem naučijo česa novega (Nikitenko 2005, 387). Zaradi tega induktivno UI pogosto povezujemo z izrazom strojno učenje (Bognar 2021, a5). Induktivni sistemi ne temeljijo več na

rigoroznih premisah (in niso omejeni z ozko strukturo pravil), pač pa na novih interakcijah, ki jih ustvarjajo v novih okoljih, v katerih se znajdejo. Ker je ustvarjanje novih povezav njihova ključna karakteristika, se je induktivno zasnovane UI oprijel izraz konekcionizem. Na drugi strani pa se deduktivno zasnovano UI pogosto označuje z izrazom simbolizem, saj temelji na izhodiščnih simbolih ali premisah, ki vzpostavljajo rigorozen okvir njenega delovanja (Steging 2018, 1).

Delovanje deduktivnih modelov UI je ljudem dosti lažje razumljivo kot delovanje induktivnih (9). Tudi ob pogledu v preteklost lahko opazimo, da je bil prvi val tehnološkega razvoja UI, ki se je začel v sredini 50. in končal v poznih 70. letih prejšnjega stoletja, osredotočen skoraj izključno na deduktivne modele (Higgitt 2011, 87). Razlog je, da koncept simbolizma temelji na enostavnejšem dojemanju inteligence: delovanje uma je skoraj izenačeno z delovanjem računalnika (88). Šlo je v bistvu za to, da so človekovo mišljenje želeli razložiti kot mehansko razporeditev, sestavljeno iz končnih struktur. Kmalu je postalo jasno, da s takšnim pristopom dovolj učinkovitega približka človeški inteligenci ni mogoče izdelati. Prvi val raziskovanja UI je ugasnil, nastopil pa je drugi val, pri katerem je v ospredje stopil koncept statističnega učenja. Novi sistemi UI niso več delovali tako, da bi jih bilo treba z novimi pravili pripraviti prav za vsako novo izkušnjo, ki bi jih lahko doletela (Steging 2018, 1). Namesto tega je učenje potekalo na temelju velike količine podatkov; ni več veljala predpostavka, da je delovanje uma podobno delovanju računalnika. Konekcionizem je inteligenco namreč dojemal kot rezultat povezanosti dinamičnega vzorca omrežij – umevanje se dojemalo kot »nekaj fluidnega, kar vznikne znotraj kompleksnega reda« (Higgitt 2011, 89). Treba pa je poudariti, da je zaradi omenjene fluidnosti proces odločanja induktivne UI za laičnega opazovalca bistveno težje sledljiv ali razložljiv (Steging 2018, 1).

Ob pogledu na premik ob prehodu iz prvega v drugi val razvoja UI se na prvi pogled zdi, da se vse skupaj dobro ujema s Habermasovo teorijo racionalnosti. Kot kaže, je Habermasov premik od monološke h komunikativni racionalnosti našel svoj odmev v načinu razvijanja sodobne UI. Takšno mnenje zagovarja sociolog Ryan Higgitt, ki trdi, da konekcionizem in komunikativno racionalnost povezuje skupna namera: uresničenje človeškosti (Higgitt 2011, 96). Habermas to želi doseči skozi komunikativno delovanje, ki naj bi pripeljalo do univerzalnih moralnih norm (Habermas 2008, 272). Raziskovalci umetne inteligence pa pri svojem delu želijo razviti takšne modele, ki bodo čim bolj posnemali pristno človeško mišljenje in vedênje. Kot pravi Higgitt, oba premika – tako premik v teoriji racionalnosti kot v razvoju UI – odražata večjo ustreznost svetu, ki nas obdaja (Higgitt 2011, 97). Stična točka med konekcionizmom in Habermasovo teorijo je namreč priznanje fluidnosti (89). Ideja fluidnosti pri Habermasu temelji predvsem na nasprotovanju prevelikemu poudarjanju znanstvene metode (87). Na področju raziskav UI pa je fluidnost predstavljena kot izhod iz miselnosti, ki je človeški um enačila z delovanjem računalnika. Potemtakem bi lahko rekli, da gre tudi na področju UI za poskus zoperstavitve fundacionalizmu – za oblikovanje holističnega in zato tudi dinamičnega sistema mišljenja (84). Higgitt ugotavlja, da so sodobnejši modeli UI bolj v skladu s Habermasovim sistemom mišljenja (97). Še več: zdi se, da to ugotovitev

predstavlja kot razlog za pozitivno ovrednotenje drugega vala UI in komputacijske teorije uma (84). V zaključku Higgitt sicer poudarja, da bo najboljši razsodnik o novih sistemih UI dejanska praksa njihove uporabe (97). Vendarle pa njegovo pisanje vzbuja občutek, da domnevna povezava med komunikativno teorijo in konekcionizmom slednjemu znatno povečuje veljavo.

4. Princip nevtralnosti naključja v pogledu na človeškost

Interpretacija Higgita želi nakazati skladnost med konekcionizmom in teorijo komunikativne racionalnosti. Sam pa bi na tem mestu predlagal, da Habermasovo teorijo beremo z drugačnim poudarkom. Četudi se ujemanje med lingvističnim obratom in drugim valom UI na prvi pogled zdi prisotno (88), bi bilo treba zadevo raziskati podrobneje. Predvsem je treba večjo pozornost nameniti nekaterim Habermasovim prispevkom, ki bi lahko jasneje pokazali, kako se njegova teorija lahko odzove na pojav umetne inteligence. Tu se kot prvi ponuja prispevek z naslovom *Prihodnost človeške narave*, v katerem je Habermas prednostno obravnaval moralno dilemo genetske manipulacije (Habermas 2005). Razprava je pomembna, ker jo do neke mere lahko apliciramo tudi na problematiko umetne inteligence.

Habermas se v omenjeni razpravi postavlja na stališče, da je manipuliranje genetskega materiala pri človeku z vidika etike diskurza docela nesprejemljivo in zavržno dejanje (32). Etika diskurza namreč predpostavlja, da se diskurz oblikuje med posamezniki različnih stališč – in pravilno etično razumevanje v skupnem prizadevanju (19). Habermas pokaže, da nas k temu napeljuje že Kantov kategorični imperativ, ki v svojem bistvu zahteva prehod od ‚jaz‘ k ‚mi‘ (63). Gre torej za uveljavitev perspektive širšega diskurza, ki pa za svoj obstoj predvideva tudi določene pogoje. Ključen pogoj je gotovo to, da strinjanje ali nestrinjanje, ki ga izražajo posamezniki, resnično šteje. Moramo se zavedati, da odobravajoči ‚da‘ in odklanjajoči ‚ne‘ lahko štejeta samo v primeru, da posameznik, ki ju izrazi, za svojimi nameni resnično stoji (65). To pomeni, da se v posameznikovo notranje naziranje predhodno ne sme vmešavati nihče drug. V diskurzu nihče ne sme govoriti namesto nas – niti ne smemo dobiti občutka, da skozi naša stališča in argumente govori nekdo drug. Habermasovo komunikativno delovanje v svojem bistvu namreč predpostavlja nekakšno nevtralnost, ki posameznikom, vključenim v diskurz, zagotavlja enaka izhodišča. To nevtralnost Habermas vidi predvsem v nedotakljivosti odnosov medsebojnega priznanja (42–43). Prav medsebojno priznanje je tisto, ki iz človeka naredi osebo (43). V prepoznavanju pa ključno vlogo igrata ravno posameznikove nepopolnosti:

»Ker se človek v biološkem smislu rodi nepopoln in ostaja vse življenje odvisen od pomoči, pozornosti in priznavanja svojega družbenega okolja, postane nepopolnost individuacije v zaporedju DNK razvidna v trenutku, ko se začne proces družbene individuacije. Individuacija življenjske zgodbe poteka skozi podružbljanje.« (43–44)

Habermas se zato zavzema za koncept ranljivosti. Ranljivost, ki izhaja že iz posameznikovih telesnih in duševnih predispozicij, postane vstopna točka za družbeno integracijo, ki človeka dokončno oblikuje v osebo. Na tem mestu ponovno pride do izraza Habermasovo pojmovanje racionalnosti: racionalnosti pri človeku ne smemo iskati v tehnologizaciji njegove genetske strukture, temveč v integraciji človeka (z njegovo ranljivo genetsko strukturo vred) v širšo družbo. Človek namreč postane oseba z razumom šele takrat, ko je deležen pripoznanja (44). Pripoznanju nasproten pa je odnos polaščanja in programiranja. Programiranje pri starših ali znanstvenikih nastopi takrat, ko ne vstopajo v življenjsko zgodbo, kjer je avtor otrok sam (59). To pomeni, da ne dopuščajo njegovega avtonomnega in spontanega telesnega in duševnega razvoja. Posledica tega je, da programirani nikoli ne bo čutil, da je na enaki ravni kot programer (82–83). Nasproti liku programerja, ki se želi polastiti avtonomije drugega, Habermas postavlja lik zdravnika. Zdravnik v globljem pomenu besede je tisti, ki upošteva, da je vsak človek vir avtentičnih zahtev (63). Zdravnik sledi t. i. smotrnostni formuli, da je vsaka oseba obravnavana kot smoter že sama po sebi. Zaveda se, da drug drugemu dolgujemo spoštovanje pravice do oblikovanja lastne usode (Anderson 2005, 817). Lik zdravnika v Habermasovem dojemanju instrumentalizaciji povsem nasprotuje.

Prav predpostavka univerzalnega diskurza bi bila z uvedbo genetske manipulacije neobhodno ogrožena. Osebo, deležno tehnološko modificiranega genoma, bi spremljal občutek, da je njen DNK in s tem njeno osebnost priredil nekdo drug. Habermas se pri tem konkretnem vprašanju trdno zavzema za pozicijo mehkega genetskega determinizma (Morar 2015, 107). Če bi prišlo do manipulacije z genomom, bi se po Habermasovem mnenju identifikacija z lastnim telesom porušila, hkrati pa tudi identifikacija z lastnim mišljenjem, stališči in izjavami. Habermas trdno verjame, da bi takšna situacija možnost univerzalnega diskurza odločilno ogrozila. Zaradi tega neomajno vztraja pri stališču, da je nujno varovati kontinuiteto sebstva in zavest naravnega naključja (Habermas 2005, 68). Če pride do genetske fiksacije, pomeni, da se je tok naravnega naključja na neki točki življenja prekinil (70). Habermas meni, da to lahko povzroči grozovite posledice tako za osebno zavest avtonomije kot tudi za medosebne odnose (71). Iz tega primera se zopet vidi, da mora biti vsakršno tehnološko iskanje bližnjic podrejeno vrednoti avtonomije.

5. Utemeljena različnost

Kot vidimo, je eden izmed temeljnih predpogojev posameznikove osebne integritete naključje njegovih genetskih danosti. Habermas trdi, da je ravno naključnost tisti dejavnik, ki omogoča, da človek svoje življenje prevzema odgovorno (20). Naključje se kaže kot nujna predpostavka, da je človek lahko resnično sam svoj – in da posledično v medosebnih odnosih lahko nastopa enakopravno (21). Ključno je, da posameznik v svojem družbenem delovanju nastopa z lastnimi argumenti, za katerimi resnično stoji. Izhodiščno možnost za to mu daje prav naključnost njegove naravne genetske določenosti (20). Le če je genetski zapis res plod naključne

genetske geneze, je mogoče iz te naravne predpodlage izhajati odgovorno – in jo zares prevzeti kot lastno identiteto. Brž ko je vsejan dvom, da je s posameznikovo genetsko opremljenostjo manipuliral nekdo drug, nastane nova oblika medsebojnih odnosov, ki moralno občutje močno prizadene (21) – gre torej za dilemo, ki zadeva tudi etično samorazumevanje človeka (22). Poseg v globino organskih zasnov namreč v posamezniku lahko porodi občutek, da je sredi družbe, ki temelji na medsebojnem pripoznanju, tujek (21) – in iz istega razloga je prizadet tudi občutek za osebno odgovornost, ki je temelj moralnega življenja v družbi. Habermas dodaja še, da je človeku, odraščajočemu v običajnih razmerah, omogočen refleksivni odnos do tega, kar je: refleksivni odnos do procesa lastnega oblikovanja. V primeru genetske manipulacije pa bi bila zmožnost refleksije znatno okrnjena (22).

Če se vrnemo k vprašanju umetne inteligence, lahko vidimo, da Habermasova teorija poudarja predvsem razliko med delovanjem človeka in delovanjem UI. Če torej Higgitt osvetljuje zlasti vidike, v katerih se med komunikativno racionalnostjo in konekcionistično UI kaže podobnost, pa sam zatrjujem, da Habermasova teorija ob celostnem branju do pojava umetne inteligence razodeva drugačen odnos. Pri obravnavi genetske manipulacije namreč velik pomen pripisuje naključju, moralni odgovornosti in zmožnosti refleksije (20–23) – in treba je priznati, da pri UI ne moremo najti nobene od teh stvari. Tudi najsodobnejši sistemi umetne inteligence se niso razvili iz naključnih danosti, niso razvili avtonomnega moralnega čuta, prav tako niso razvili zmožnosti za refleksijo o procesu lastnega nastanka. Res je sicer, da konekcionizem zasleduje idejo, da se UI uči iz ustvarjanja novih povezav v novih situacijah. Toda vse to UI lahko počne zgolj iz vnaprej programiranih danosti, ki pa nikakor niso naključne. Tako deduktivna kot tudi induktivna UI delujeta na izhodiščih, ki smo jih vanju vstavili ljudje. Tako moramo pri vseh oblikah umetne inteligence princip nevtralnosti naključja odmisлити. Z odsotnostjo naključja je pri UI povezana tudi nezmožnost prve refleksije. Ta vidik posebej izpostavlja moralni filozof Paolo Monti, ki status moralnega subjekta povezuje z reflektivnim aspektom (Monti 2024, 64). Modelov UI torej nikakor ne moremo dojemati, kot da so odgovorni moralni subjekti. Tudi med najnaprednejšimi sistemi UI – Monti ima v mislih predvsem velike jezikovne modele – statusa moralnega subjekta ne moremo pripisati nobenemu. Dodaten razlog za to je tudi dejstvo, da noben model UI ne izhaja iz lastnega življenjskega sveta (72–74). Strahovnik k vsemu temu dodaja še krščanski vidik ustvarjenosti človeka po Božji podobi: za UI ne moremo trditi, da je ustvarjena po Božji podobi, kot to velja za človeka – zato tudi nima enakega statusa (Strahovnik 2023, 852). Z vidika komunikativne racionalnosti je mogoče naštet še več lastnosti, po katerih umetna inteligenca ostaja od človeške izrazito različna: sodobne oblike umetne inteligence niso zares zavezane resnici, ampak jih vodi ekonomski interes; ne upoštevajo Habermasovega načela pravosti, ker njihov razvoj temelji na digitalnem izkoriščanju revnih; vodijo v smer strateškega delovanja v smislu individualizma in utilitarizma (Westerstrand, Westerstrand in Koskinen 2024). Zato se poraja vprašanje, ali umetno inteligenco sploh lahko imamo za pravo inteligenco (Žalec 2023, 813). Odgovor na to vprašanje je odvisen od opredelitve pojma ‚inteligenca‘. Vendarle pa vidimo, da je pri

trajnem razlikovanju med človeško in umetno inteligenco prav Habermasova teorija lahko v veliko pomoč. Pomaga nam, da odločilne vidike, v katerih umetna inteligenca ne dosega človeške, odkrijemo in učinkovito obrazložimo.

6. Zaključek

Teorija komunikativne racionalnosti v odnosu do umetne inteligence poudarja zlasti različnost med človeško in umetno inteligenco. Med Habermasovimi interpreti sicer obstajajo poskusi, ki prehod od deduktivne k induktivni UI razlagajo po analogiji s prehodom od fundacionalizma h komunikativni racionalnosti. Toda sam menim, da poglobljeno branje Habermasa odkriva zlasti vidik razlike. Menim, da se Habermasova teorija v času, ko se porajajo različna ugibanja o tem, ali bi umetna inteligenca lahko postala enaka človeški (Centa Strahovnik 2023, 853), kaže kot koristna, saj daje dobre utemeljitve v smeri trajnega razlikovanja med človekom in UI. Pomen naključne danosti, avtonomne moralne odgovornosti, zmožnost samorefleksije, ukoreninjenost v življenjski svet itd. – vse to predstavlja močne temelje za široko zavest o jasni razliki med njima. Habermas nam torej lahko pomaga, da glede na sedanji in prihodnji tehnološki napredek ozavestimo distanco med človekom in UI. Habermas sicer tehnološkemu napredku samemu po sebi ne nasprotuje (Feenberg 1996, 49–50). A s pomočjo njegovih teoretičnih izhodišč se kot družba lahko zavežemo, da bomo distanco med človekom in UI na vseh ravneh še naprej strogo vzdrževali. Nedotakljivost naravnih danosti, sposobnost moralne odgovornosti in zmožnost refleksije o sebi moramo vedno obravnavati kot posebne vrednote, ki jih imamo samo mi kot osebe. Na teh področjih meje nikoli ne smemo prekoračiti ne v eno ne v drugo smer: sistemom UI tako ne smemo nikdar pripisati statusa odgovornega moralnega subjekta, obenem pa našega moralnega delovanja ne smemo prepustiti odločitvam UI. Ne glede na to, kakšne sposobnosti bo UI v prihodnosti razvila – za moralo in refleksijo bo moral vedno poskrbeti človek.

Kratika

UI – Umetna inteligenca.

Reference

- Anderson, Joel.** 2005. Review of Jürgen Habermas, 'The Future of Human Nature'. *Ethics* 115, št. 4:816–821.
- Bognar, Melinda.** 2021. Prospects of AI in Architecture: Symbolicism, Connectionism, Actionism. *Journal of Architectural Informatics Society*, a1–a20.
- Centa Strahovnik, Mateja.** 2023. Identiteta in pogovorni sistemi umetne inteligence. *Bogoslovni vestnik* 84, št. 4:853–864. <https://doi.org/10.34291/bv2023/04/centa>
- Feenberg, Andrew.** 1996. Marcuse or Habermas: Two Critiques of Technology. *Inquiry: An Interdisciplinary Journal of Philosophy* 39, št. 1:45–70. <https://doi.org/10.1080/00201749608602407>

- Fisher, Eran.** 2021. Epistemic media and critical knowledge about the self: Thinking about algorithms with Habermas. *Critical Sociology* 48, št. 7–8:1309–1324. <https://doi.org/10.1177/08969205211044193>
- Geiger, R. Stuart.** 2009. Does Habermas Understand the Internet? The Algorithmic Construction of the Blogosphere/Public Sphere. *Gnovis: A Journal of Communication, Culture, and Technology* 10, št. 1. <https://escholarship.org/uc/item/60s6s0p8> (pridobljeno 7. 8. 2024).
- Gregorčič, Rok.** 2023. Tehnološki razvoj v luči Habermasove etike diskurza. *Bogoslovni vestnik* 83, št. 4:911–922. <https://doi.org/10.34291/bv2023/04/gregorcic>
- Habermas, Jürgen.** 1984. *The Theory of Communicative Action*. Zv. 1. *Reason and the Rationalization of Society*. Boston: Beacon Press.
- . 2002. *Religion and Rationality: Essays on Reason, God, and Modernity*. Cambridge: Polity Press.
- . 2003. *Truth and Justification*. Cambridge: Polity Press.
- . 2005. *Prihodnost človeške narave. Verjeti in vedeti*. Ljubljana: Studia humanitatis.
- . 2008. *Between Naturalism and Religion: Philosophical Essays*. Cambridge: Polity Press.
- Higgitt, Ryan.** 2011. Habermas' Communicative Rationality and Connectionist AI. *Culture, Theory and Critique* 52, št. 1:83–100. <https://doi.org/10.1080/14735784.2011.621664>
- Hovdelien, Olav.** 2011. Post-Secular Consensus? On the Munich-dialogue between Joseph Ratzinger and Jürgen Habermas. *Australian Journal of Theology* 18, št. 2:107–116.
- Khan, Muhammad Zubair, Ijaz Shafi Gilani in Allah Nawaz.** 2012. From Habermas Model to New Public Sphere: A Paradigm Shift. *Global Journal of Human Social Science* 12, št. 5:43–51.
- Monti, Paolo.** 2024. AI enters public discourse: A habermasian assessment of the moral status of large language models. *Etica & Politica* 26, št. 1:61–80.
- Morar, Nikolae.** 2014. An Empirically Informed Critique of Habermas' Argument from Human Nature. *Science and Engineering Ethics* 21:95–113. <https://doi.org/10.1007/s11948-013-9509-5>
- Nikitenko, Agris.** 2005. Intelligent Agent Control Using Inductive, Deductive and Case Based Reasoning. *Information Theories & Applications* 12:386–395.
- Ritzi, Claudio.** 2023. The hidden structures of the digital public sphere. *Constellations* 30:55–60.
- Salikov, Alexey.** 2018. Hannah Arendt, Jürgen Habermas, and Rethinking the Public Sphere in the Age of Social Media. *Russian Sociological Review* 17, št. 4:88–102. <https://doi.org/10.17323/1728-192x-2018-4-88-102>
- Steging, Cor.** 2018. Explainable AI: On the Reasoning of Symbolic and Connectionist Machine Learning Techniques. Magistrsko delo. University of Groningen.
- 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>
- Takefuji, Yoshiyashu, in Ren Yamada.** 2019. A new deductive method with intelligence using pseudorandom number for solving coin-weighting puzzles: review of inductive and deductive methods. Preprints. 8. 5. <https://www.preprints.org/manuscript/201905.0090/v1> (pridobljeno 9. 8. 2024).
- Westerstrand, Salla, Rauli Westerstrand in Jani Koskinen.** 2024. Talking existential risk into being: a Habermasian critical discourse perspective to AI hype. *AI and Ethics* 4:713–726. <https://doi.org/10.1007/s43681-024-00464-z> (pridobljeno 13. 8. 2024).
- Žalec, Bojan.** 2023. Ali je umetna inteligenca inteligenca v pravem pomenu besede? Vprašanje psihičnih značilnosti in Splošnost. *Bogoslovni vestnik* 83, št. 4:813–823. <https://doi.org/10.34291/bv2023/04/zalec>

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 807—822

Besedilo prejeto/Received:06/2024; sprejeto/Accepted:10/2024

UDK/UDC: 37:316.42:27

DOI: 10.34291/BV2024/04/Malmenvall

© 2024 Malmenvall, CC BY 4.0

Simon Malmenvall

State and Christian Enlightenment: Background of the Mass Elementary Education in Central Europe *Država in krščansko razsvetljenstvo: ozadje množičnega osnovnega šolstva v srednji Evropi*

Abstract: Mass elementary education powered by the state – one of the cornerstones of the processes of modernization – began to be established in continental Europe in the late eighteenth century. Increasingly centralized political institutions, with the decisive assistance of the school network, gradually created a bourgeois-industrial society centered on the performance-oriented individual committed to the state. The key to the spread of the school network and literacy was the coordination between Enlightenment ideas, state authorities, and ecclesiastical organizations. The situation in the Central European environments under consideration – Austria (part of which was also the territory of present-day Slovenia), Prussia, and other German lands – shows that the social role of education was strengthened by reform types of Christianity, which, based on the pursued harmony between the Church and State, sought to create a rational, industrious, and morally responsible individual.

Key words: General School Ordinance (1774), school reforms in Central Europe, Enlightenment and Christianity, reform Catholicism, history of elementary education

Povzetek: Množično dostopno državno osnovno šolstvo – eden od temeljnih gradnikov procesov modernizacije – se je v celinski Evropi začelo vzpostavljati ob koncu 18. stoletja. Vse bolj centralizirane politične institucije so z odločilno pomočjo šolske mreže postopoma ustvarile meščansko-industrijsko družbo, v katere središču se je nahajal storilnostno naravnan in državi predan posameznik. Za širjenje šolske mreže in pismenosti je bila ključna usklajenost med razsvetljenjskimi idejami, državnimi oblastmi in cerkvenimi organizacijami. Stanje v obravnavanih srednjeevropskih okoljih – Avstriji (katere del je bilo tudi ozemlje današnje Slovenije), Prusiji in drugih nemških deželah – kaže, da so družbeno vlogo izobraževanja krepile reformne oblike krščanstva, ki so si na podlagi želene skladnosti med Cerkvijo in državo prizadevale ustvariti racionalnega, delavnega in moralno odgovornega posameznika.

Ključne besede: Splošni šolski red (1774), šolske reforme v srednji Evropi, razsvetljenstvo in krščanstvo, reformno katolištvo, zgodovina osnovnega izobraževanja

1. Introduction

Mass elementary education powered by the state – one of the cornerstones of the processes of modernization – began to be established in continental Europe in the late eighteenth century.¹ Establishing compulsory education accessible to children from all social classes was closely intertwined with the process of forming modern nations and states. It ran in parallel with economic and technological change driven by industrialization and urbanization.² In terms of its novelty and long-term impact, mass elementary education from the turn of the eighteenth and nineteenth centuries was similar to the novelty and challenges of the so-called digital revolution and the emergence of artificial intelligence at the turn of the twentieth and twenty-first centuries.³ The thematic focus of this article is to shed light on the expansion of the school network and literacy considering the intertwinement of Enlightenment ideas, state authorities, and ecclesiastical organizations. Based on this, we adopt a view of the Enlightenment as a multifaceted cultural phenomenon that, despite its criticism of institutional religion, cannot be placed in direct opposition to Christianity or the Catholic Church.⁴

Models of organized school systems made their way from country to country, thus influencing one another. This was definitely the case with the far-reaching importance of Johann Ignaz von Felbiger (1724–1788),⁵ an Augustinian canon and abbot of the monastery in Żagań (German: *Sagan*) in Silesia, in the southwest of present-day Poland, who was tasked with a school reform as a minister in the Prussian government. The Prussian model of elementary education was the oldest in this context and helped shape developments in several countries, particularly the neighbouring Habsburg Empire, where Felbiger was invited to in order to guide a school reform by Maria Theresa (ruled: 1740–1780), who introduced compulsory education for all children aged six to twelve by issuing the General School Ordinance in 1774. (Westberg et al. 2019, 5–7; Green 1994, 2) Based on the

¹ This article is part of the research program “P6-0269 Religion, Ethics, Education, and Challenges of Modern Society” and basic research project “J5-4595 Between Tradition and Modernity: Slovenian Catholic Intellectuals and the National Question in a Transnational Perspective (1848–1948)” supported by the Slovenian Research and Innovation Agency.

² The relation between mass elementary education and modernization processes in Europe is elucidated by two overview works: Horn Melton 1988; Westberg et al. 2019.

³ The novelty and challenges of artificial intelligence in (religious) education and society are illustratively explained in the following works: Vodičar 2023; Nežič Glavica 2023.

⁴ Such a view is increasingly present in modern historiography; the following edited volume is particularly valuable in this context: Burson and Lehner 2014. A theological critique of the notion of the incompatibility of the Enlightenment with Christianity is provided by the following monograph: Avis 2022.

⁵ Reference biography of Johann Felbiger: Krömer 1966.

approximately simultaneous beginnings of the development of state elementary education between the mid-eighteenth and mid-nineteenth centuries, this study takes into account the situations in selected Central European environments: the Habsburg Empire, which also encompassed present-day Slovenia, the Kingdom of Prussia, and other German lands. Due to the importance of Maria Theresa and Joseph's reforms for Slovenian history, the article's focus is the situation in Austria.⁶

2. Foundations of the Educational and Ecclesiastical Reform

In 1774 Empress Maria Theresa proclaimed compulsory elementary education for all children aged six to twelve, irrespective of their sex and social status. The aim of the General School Ordinance was to integrate the diverse regions of the empire into a unified national community, and its main concern was to promote literacy among all strata of the population. The subjects to be taught in all elementary schools were reading, writing, arithmetic, and catechesis. A major reason for issuing the school ordinance was the principles of the novel Enlightenment worldview. The ruler believed that traditional adherence to the Catholic faith and morals, along with teaching catechism, enhanced with the Enlightenment values of diligence, frugality, and curiosity, would contribute to material and moral progress in the country and bring happiness to the population. At the core of her reflections was the fundamental Enlightenment postulate of the necessity and viability of educating people, encompassing both children with the help of parents aware of its importance and teachers and adults through reading, listening to Church sermons, and through actions by the state. (Viehhauser 2019, 23–24; Okoliš 2009, 39–40; 42; Borgstedt 2004, 53)

The goal of the reform efforts of the time was to increase revenue in the state budget, strengthen the army, and foster a feeling of belonging to the monarchy in the population. All this would have been impossible without increasing the literacy rate and at least a basic knowledge of German. However, the unification of the country was not reflected in the total domination of German as the main means of communication for most of the population and the political elite. The school ordinance allowed the use of regional (vernacular) languages in teaching and writing textbooks where this ensured that the maximum possible number of inhabitants understood the subject matter, while requiring unity in teaching methods, teaching contents, and school organization. The introduction of compulsory elementary schooling brought about a new level of relations between the ruler and their subjects. Through the daily experience of life in school and with school, the state became 'real' for the inhabitants; schools – in a similar vein to posting stations, district offices, and army barracks – became tangible expressions

⁶ Fundamental historical and sociological monograph on establishing the modern school system in the Habsburg Empire: Cvrček 2020.

of the previously abstract state. (Viehhauser 2019, 19; 22; Stollberg-Rilinger 2017, 714–715; Ciperle and Vovko 1987, 37)

Ideationally, state interventions in education, which had been the domain of the Church for centuries in Austria, was based on criticism of Baroque folk religiosity marked by lavish ceremonies and processions, the veneration of saints, and numerous holidays.⁷ The effects of such religiosity were disapproved by an increasing number of Catholic intellectuals, both laypeople and priests, even bishops and especially theology professors. To this religiosity, they ascribed a susceptibility to superstition, mental confusion and idleness, which caused economic and moral harm to the state, including an adverse influence on the operation of the bureaucratic and military apparatus. Instead, they advocated for 'rational' or 'inner' religiosity focused on the reading of spiritual literature, reflection about evangelical stories,⁸ and a highly virtuous life intertwined with the expansion of literacy, the teaching of Catholic doctrine to all strata of the population, and a sense for following instructions by the state. Among theologians of the time, a similar spirit was specifically highlighted by Lodovico Muratori (1672–1750), an Italian priest and polymath, especially in his work *On the Orderly Devotion of Christians (Della regolata divozione dei cristiani 1747)*. It was Christoph Migazzi, an initially Enlightenment-oriented archbishop of Vienna⁹ (in office: 1757–1803) and, *inter alia*, vice-president of the Court Commission of Studies (*Studienhofkommission*), a body tasked with the implementation of the school reform in Austria, who had a translation of this work into German (1757) published. (Stollberg-Rilinger 2017, 564–569; Printy 2010, 188–189; 195; Kluebing 2010, 134–135; 141; Gierl 2013, 357)

Another direction in reform Catholicism of the time that is comparable in terms of content drew from Jansenism,¹⁰ an originally Dutch–French¹¹ spiritual movement akin to radical Protestant denominations that claimed to be returning to the first centuries of Christianity. Jansenism emphasized human susceptibility to sin and a beforehand selection of people for salvation, i.e. the doctrine of predestination; it fostered an ascetic way of life, an orderly structure of an individual's day,

⁷ In 1753, Maria Theresa successfully negotiated with Pope Benedict XIV (ruled: 1740–1758) to reduce religious holidays in Austria by two thirds; an additional reduction followed in 1772. A similar measure was the ban on multiday pilgrimages, except those to the central shrine of Mary in Austria, Mariazell in Styria. (Stollberg-Rilinger 2017, 581; Kluebing 2010, 144).

⁸ In this sense, Catholic Enlightenment gained fresh impetus owing to the above-mentioned Pope Benedict XIV, who allowed in 1757 for the entire Bible to be translated into vernacular languages and made available for public use with appropriate commentary (Lehner 2010, 23).

⁹ The importance of Vienna as one of the major centers of Catholic Enlightenment is supported by the fact that eight of the twenty German editions of Muratori's work were published in this city (Lehner 2010, 28). However, the archbishop of Vienna ceased to publicly support Enlightenment views in the late 1760s when he realized reforming the Austrian society entailed the submission of the Church to the state, particularly regarding its property and the appointment of bishops. A comprehensive biography of Christoph Migazzi is provided in the following classical monograph: Wolfsgruber 1897.

¹⁰ Classical overview study on Jansenism in Austrian lands: Hersche 1977.

¹¹ The influence of Muratori spread to Austria quickly in part due to Habsburg possessions in Northern Italy. A similar geographic and political factor can be identified with Jansenism as the Habsburgs had 'extraterritorial' possessions in Holland (the so-called Spanish Netherlands) (Kluebing 2010, 128–129).

frequent confession and penance, rare reception of communion, modesty in ceremony and clothing; it advocated for a decentralization of the Church by curtailing the competence of the Pope, for introducing vernacular languages to liturgy, and for raising the education level in all strata of society. Although the teachings of its founders – Cornelius Jansen (1585–1638), a Flemish bishop, and Pasquier Quesnel (1634–1719), a French theologian – were denounced by Pope Clement XI (ruled: 1700–1721) in his bull “Unigenitus Dei Filius” (1713), different forms of Jansenism remained popular throughout the eighteenth century in noble, bourgeois, and ecclesiastical circles of Habsburg lands. They spread particularly persistently among lecturers and priesthood candidates in schools of higher theological education, where they decisively shaped the teaching of moral theology up until the mid-nineteenth century.¹² One of the most prominent proponents of Austrian Jansenism, which was favourable towards state authority, was Johann Karl von Herberstein, a bishop of Ljubljana (in office: 1772–1787).¹³ In the environments mentioned, Jansenism also acted as an expression of the prestige of the higher social strata, to distinguish them from uneducated and ‘lazy’ masses. Politically, through the theory of natural law, which was popular in Protestant lands, it advocated for limiting the property and social privileges of the Church. Its operation should be limited to the performance of rituals, preaching, teaching the people, and ensuring the morality of priests. (Stollberg-Rilinger 2017, 576–577; 584–585; Klueting 2010, 132–133; 152; Gierl 2013, 357; Lehner 2010, 21–22)

3. Education and the Church as Part of the State

Maria Theresa was personally a devout Catholic; she publicly attended religious ceremonies and supported the Baroque notion of the Habsburg dynasty as the protector of the Catholic Church and its orthodoxy. Although she remained within the limits of ‘Austrian piety’ (*pietas Austriaca*) as regards the fulfilment of religious duties, she aimed to strengthen the state by limiting the Church’s influence on social life. (Stollberg-Rilinger 2017, 560–561; 563; 571–574; Klueting 2010, 128; Weston Evans 1979, 169) Education was to become a systematic concern of the political authorities, but they would still support Catholic religious and moral content in teaching. An opportunity for a decisive state intervention in education arose with a decree by Pope Clement XIV (ruled: 1769–1774), “Dominus ac Redemptor Noster”, which disbanded the Jesuit order in 1773 at the level of the whole Catholic Church. In Austria and other Catholic countries, Jesuits had been the main providers of education and scientific activities since the late eighteenth century, especially at the levels of secondary and higher education. In this role, they shaped not only future priests but lay intellectuals as well. After the Jesuit order was abolished, the Austrian state took over many of its buildings and pos-

¹² On the role of Jansenism in eighteenth-century moral theology: Laun 1979.

¹³ Herberstein’s work and his place in the Central European cultural and ecclesiastical streams are highlighted in the following edited volume: Škulj 2004.

sessions, while a number of Jesuits entered state service. This created the basis for education to become primarily a matter of the state (*politicum*) and only secondarily a matter of the Church (*ecclesiasticum*). (Viehhauser 2019, 25)¹⁴

In the following decades, involving the Church in the strengthening of the school network was in accordance with the plans of Joseph II (ruled: 1780–1790). Even more than Maria Theresa, his mother, he envisioned the place of the Catholic Church in society as an extension of the state. Joseph's mindset was based on the Enlightenment interpretation of the ideals of natural law, according to which a ruler is not the keeper of orthodoxy and their subjects' guide to eternal salvation, as had been the case in previous centuries, but rather a custodian of the common good within earthly realities, which gives them the right and duty to intervene in the religious sphere, which remains part of public life. Thus, in Joseph's understanding, the state justifiably exerted its authority by managing Church property,¹⁵ by establishing new dioceses and parishes and determining their borders in line with the borders of political administrative units, by appointing bishops, by abolishing monasteries of contemplative orders (e.g. Carthusians, Cistercians),¹⁶ and by generally directing the operation of the Church,¹⁷ which is visible in particular in the moral, educational, and charitable spheres. (Trontelj 2022, 21; Klueting 2010, 144–146; 150; Lehner 2010, 13)

As an educated Augustinian from the religiously heterogeneous Silesia, Johann Felbiger knew the operation of the Protestant urban and rural schools there, which were better and more numerous than the Catholic ones. One reason he started designing his school reform was to strengthen the position of the Catholic Church in the land and prevent the increasing practice of Catholic parents sending their children to schools of 'competing' faiths. When reading pedagogical literature and assessing practical examples, he was inspired by the developing school system in Prussia, especially the example of the 'Real School' (*Realschule*) in Berlin. The teaching there was based on the group method, i.e. the teacher lecturing to a large group of students in one classroom, and on imparting knowledge following the principle of cause-effect questions and answers, known as the catechetical method, which was to promote quick memorization and a systematic understanding of the subject matter. The institution in question had been founded in 1747 as a form of a secondary school for future teachers by Johann Julius Hecker (1707–1768), a Lutheran pastor and former student of the renowned pedagogical institute in Halle, which had begun its work in 1695 at the initiative of August Hermann Francke

¹⁴ The circumstances of the suppression of the Jesuit order and the Enlightenment streams in the Catholic Church in the eighteenth century are discussed in the following overview monograph: Van Kley 2018.

¹⁵ The state consolidated the property of the abolished contemplative monastic orders and Church land of feudal origin into so-called religious funds, which it managed – nominally in the name of the Church and in different forms up until the end of World War I – mostly for priests' income and the needs of the education system. More on Austrian religious funds: Batič 2017.

¹⁶ The process of abolishing monasteries of contemplative monastic orders is discussed in the following overview study: Beales 1997.

¹⁷ One of the more radical forms of submitting the Church to the state was a ban on selling blessed candles and rosaries. Due to resistance from the masses, Joseph's successors did away with most of his interventions in religious life. (Kluetig 2010, 144; 157).

(1663–1727), a theologian and founder of the spiritual and cultural movement¹⁸ called Pietism. (Viehhauser 2019, 26; Krömer 1966, 23–24; Schmidt 1988, 178; Doney 1988, 8–9; 13–14) Based on Prussian pedagogical theory and practice, Felbinger developed an education program in Silesia in 1764, a decade before the Austrian general School Ordinance, and founded a ‘normal school’ in Sagan for the education of teachers, which was to be an example to all other primary schools. Felbinger’s program set out requirements for the education and selection of teachers, for attending classes and for overseeing their implementation. Thus, the later system of education in Catholic Austria was decisively shaped by principles from Protestant Prussia. (Krömer 1966, 37–38; Engelbrecht 1984, 103; Doney 1988, 14) The pedagogue formulated his principles in line with the contemporary notion of the ideal state as a well-functioning machine. (Viehhauser 2019, 28; 33; Stollberg-Rilinger 2017, 182) His handbook, the *Method Book for Teachers in German Schools* (*Methodenbuch für Lehrer der deutschen Schulen* 1775) contains a notable thought: “every student in class must watch, think, hear, and work as one.” (4)

The purpose of Maria Theresa and Joseph’s elementary education was not the formation of independently thinking citizens, but rather of subjects loyal to the state and characterized by two values in particular – diligence and piety (Stollberg-Rilinger 2017, 691). Maria Theresa’s school ordinance did not stay in force for long. In 1806, Francis II (ruled: 1792–1806) adopted a new school ordinance with the official title “Political Constitution of German Schools in Imperial-Royal German Hereditary Lands” (*Politische Verfassung der deutschen Schulen in den k. auch k.k. deutschen Erbstaaten*). The new order relinquished state competence over education, entrusting it once more to the Catholic Church. Despite its incomplete realization, the 1774 school ordinance brought profound and lasting change to education in Central Europe. In particular, it established a public space where the state became permanently present among all strata of the population. It also contributed to the consolidation of some principles still applied in social life today: education as a concern of the state, an awareness about the accessibility of a fundamental body of knowledge to all children, the importance of group teaching, and considering the moral development of students. (Ciperle and Vovko 1987, 34; Engelbercht 1984, 120)

4. Education in the Catholic and Lutheran Environments

As elsewhere in Central Europe, in those German lands¹⁹ where the Catholic faith was dominant, representatives of the Catholic Enlightenment opposed the

¹⁸ The influence of renewed Lutheran religious fervour and of promoting literacy for the purposes of studying the Bible, which are encompassed by the concept of Pietism, on education and social life in Prussia is presented in the following overview study: Gawthrop 2006.

¹⁹ The Catholic Enlightenment streams were – due to linguistic closeness, personal contacts of individual intellectuals, and similarity in ideas – particularly intertwined on German and Austrian soil. Still, two major differences can be observed between the German and Austrian environments: Jansenist views were weaker in Germany, and there the state did not make such radical interventions in Church matters like the Habsburg Empire did during the rule of Joseph II. (Klueting 2010, 130)

influence of Jesuits, who had total control of higher education and promoted Baroque spirituality until the last third of the eighteenth century. Instead, they saw their ideal in the urban, educated, moralistic, and outwardly modest version of the Catholic faith, striving to distinguish themselves from their coreligionists in rural areas. (Printy 2010, 192) Such opposition to Jesuits was in line with the aims of rulers and bureaucracy to form a centralized and efficient state, for which they needed highly educated men in the fields of law, mathematics, and natural sciences. In this light, they reproached the Jesuit education system – based on classical languages, rhetoric, philosophy, and theology – for empty theorizing and vague moral theology, the obsolescence of the method of learning by heart, being removed from reality with its almost exclusive use of Latin, and falling behind German Lutheran universities, where law and history were favoured in the eighteenth century.²⁰ (Printy 2010, 193–194; Lehner 2010, 31; Hammerstein 1977, 42; Doney 1988, 9–10)

The connection between religion, education, and the state was knowingly expressed by Johann Michael Bönicke (1734–1811), a chancellor of the Salzburg Archdiocese, which covered parts of Austrian and Bavarian territory: “Religion should support the state. It should enlighten the citizen about his duties. It should foster a love for his occupation /.../, precision in his craft. It should provide teachings on happiness for society and for each of its members.” (Hollweger 1976, 297) It was this ‘applied’ aspect of religion that was particularly expressed in Catholic spiritual and pedagogical literature from German lands, where it is possible to detect an echo of the sense of inferiority among German writers in view of the intellectual and military strength of Protestant areas, especially Prussia. As a result, these writers sometimes even reduced Jesus Christ to the role of the ‘teacher of virtue,’ thereby emphasizing the importance of education for diligent, frugal, honest, and state-authority-obedient life of all social strata. (Printy 2010, 197–198; 205; Doney 1988, 9–10; Lehner 2010, 17) The already mentioned Johann Felbiger, whose work was characterized by experiences from both Protestant Prussia and Catholic Austria, reflected along the same lines. For instance, he advocated the view that religion was not merely a means to attain eternal life but also an incentive to fulfil everyday duties, which led to a happy life on the personal and social levels. (Felbiger 1774, 60; Doney 1988, 8–9)

Although Lutheran Prussia was considered a model of progressiveness in mass elementary education between the late eighteenth and early twentieth centuries, it had no overarching state legislation that would provide for the formation of a modern school network with a single and symbolically meaningful act (Caruso

²⁰ Criticism of Jesuit pedagogy was also reflected in a transformation of theology studies in Austrian universities and seminaries. The reform that was introduced by Joseph II in 1782–1785 and was in place until 1857 originated in a plan by Franz Stephan Rautenstrauch (1734–1785), a Benedictine. Due to acquired knowledge ‘useful’ to priests in teaching the people, Rautenstrauch ascribed central importance to subjects in history, biblical exegesis, and moral theology. In addition, the ruler banned theology studies in monasteries and required priesthood candidates to undergo four-year education in state-led ‘general seminaries,’ which operated in the capitals of regions or larger areas, plus an additional year of pastoral practice. (Klueting 2010, 139; 147; 152; Zschokke 1894, 31–48)

and Töpper 2019, 41). In the nineteenth century, Prussian elementary or people's schools (*Volksschulen*) were highly successful in lesson attendance rates, teacher training, subject diversity, and school network management owing to a highly developed culture of education. This culture enabled the systematic attention of secular (city, regional, national) authorities, which was enhanced through cooperation with educational institutions of the Lutheran Church, the entity behind the official religion on Prussian soil. The care of secular authorities was reflected mainly through guidance by the increasingly efficient bureaucratic apparatus, the progress of which was unimaginable without literacy being increasingly accessible to the masses; at the same time, literacy was thought to contribute to the rising military and economic strength of the country. (43–44)

An important step in the process of establishing mass education, which was part of wider efforts for a centralized organization of the state using the existing Church network, was the 1736 decree *Principia Regulativa* ("Regulative Principles") by Frederick William I (ruled: 1713–1740), King in Prussia. The decree prescribed obligations of Lutheran congregations in constructing school buildings and remunerating teachers. (Dietrich and Klink 1972, 131–132) The king entrusted the oversight of education to Lutheran ecclesiastical councils (consistories) and awarded pastors the authority to appoint teachers. As the supreme leader of the state Lutheran Church, the king determined the financing of ecclesiastical institutions, which included schools. In the field of education and social life, he thus consolidated the congruence of the Church with the needs of the state, which further enhanced the bond between the Church and state from the beginnings of the Protestant reformation in the sixteenth century. (Caruso and Töpper 2019, 44–45) Mandatory attendance in elementary schools was first expressly prescribed in 1763, when Frederick II the Great (ruled: 1740–1786) issued the "General Land School Regulations" (*Generallandschulreglement*), which are considered one of the most important documents in the history of European education at large. The initiative to adopt this document came from the observation that the Prussian army, which had recently won the Seven Years' War (1756–1763), had too few literate officers, which reduced its efficiency. The School Regulations stipulated that students should finish their schooling only when they could read and write as well as know the basics of the Christian faith. (Königlich-Preußisches 1763) Despite a clearly expressed will of the state to regulate elementary education, the impact of the school regulations was limited. The main reasons for the problems were the small number of civil servants, who could not monitor the school reform, and the reliance on Lutheran pastors, who were mostly not trained teachers and highlighted the religious aspect of education in their work. (Caruso and Töpper 2019, 44–45; Van Horn Melton 1988, 173–175)

Significant changes in education followed with the "General State Laws for the Prussian States" (*Allgemeines Landrecht für die Preußischen Staaten*) of 1794. Parts of the code dedicated to education set the foundations for numerous administrative and content-related measures in elementary education throughout the nineteenth century, which was characterized by a general rise in literacy. The code

in question included about nineteen thousand provisions, aiming to regulate various aspects of public life in the state and thus becoming a typical expression of enlightened absolutism. The code communicated that education as a whole was a concern of the state, which was to encompass both control over learning content and teachers and their training in all public and private schools. Even so, the commitment of the state to care for education did not entail the exclusion of the Lutheran Church from the process of mandatory mass education as national and regional authorities cooperated with it in different forms up until the mid-twentieth century. (Caruso and Töpper 2019, 46–47; Leschinsky and Roeder 1983, 54)

5. Education and Lutheran Enlightenment

Before the Enlightenment, social and cultural development in eighteenth-century Prussia was decisively shaped by Pietism, a Lutheran spiritual movement advocating for ecclesiastical reform, renewal of religious fervour, moral exemplary life, and expansion of literacy for the purposes of studying the Bible and general knowledgeability. By emphasizing personal responsibility and self-reflection for own and community life and the necessity of accessible education in light of fulfilling religious and social duties, Pietism came close to the Enlightenment and facilitated modernization processes; as a result, there were similarities to the streams of reform and Jansenism in the Catholic Church. On the other hand, this movement distinguished itself from the Enlightenment and acted anti-modern by advocating for the primacy of religion in all areas of life, the usefulness of attending ceremonies and showing public signs of piety, and the necessity of ensuring the free operation of the Lutheran Church in relation to the state. (Gierl 2013, 348–350; 356; Doney 1988, 8–10) In Prussia, Pietism-minded Protestant pastors, university professors, government officials, and benefactors from different social strata – as members of the ‘state Church,’ which had already cooperated with the authorities in education and other fields – contributed to launching a process of establishing a modern school network at the elementary, secondary, and higher education levels; they also played an important role in religious care in the Prussian army. By the end of the eighteenth century, Pietist principles of education became part of official policy, which aimed at establishing a state education system based on the institutionalization of education, youth socialization, and the development of the intellectual and moral inner world of each student. It was also in the interest of the Prussian royal government to keep Pietists dependent on the financial assistance of the state – even though they stressed the primacy of faith over any other earthly dimension, they saw a greater threat in Church representatives that lacked fervour than in the sponsorship of the state, as long as it supported their ideas. (Gierl 2013, 351; 353; Marschke 2013, 488–489)

The institutions of August Hermann Francke, a theologian and the founder of Pietism, elevated the city of Halle, where they were erected around the year 1700 – e.g. pedagogical institute, Latin school (gymnasium), elementary school

for children of both sexes, orphanage, hospital – to a model education center. Its main purpose was to provide conditions for the intellectual and character formation of talented elementary and secondary school students of different social backgrounds, who would spread Pietist views among the people as they would have access to influential positions in the state and ecclesiastical apparatus based on their good education. Through Halle, Pietism asserted itself in education, and through education, in later decades it became part of the identity of the Prussian state, which also supported Pietist initiatives financially. The essence of Francke's pedagogy was the formation of a student's will through the regular reading of suitable literature, continuous control over students' activities, a meticulously prescribed daily program, everyday prayer and personal reflection about deeds done or omitted, and assessment of all acquired knowledge in the form of exams four times a year. 'Corrupt' human nature – which, according to traditional Lutheran anthropology, can only do good owing to God's grace and faith in Christ as the Savior – was thought to become available to God and fellow men through (self-) control; an individual was to develop a 'piety of the heart,' the fundamental driving force of earthly society, which is ordered according to God's idea. (Gierl 2013, 366; 384–385; Loch 2004, 275)

In 1905, Max Weber (1864–1920), a German sociologist and economist, advocated his findings about the far-reaching social impacts of the Protestant work ethic in his momentous work *The Protestant Ethic and the Spirit of Capitalism* (*Die protestantische Ethik und der Geist des Kapitalismus*). He said this ethic originated in the spiritual interpretation of practicing a secular profession as a life mission or 'calling' from God, which developed in the Calvinist and Puritan versions of Protestantism – he said it became the ideological and psychological basis for the flourishing of the capitalist system in Northwestern Europe and North America between the seventeenth and nineteenth centuries. Weber's 'spirit of capitalism' can also be understood with a different view, which is centered on Protestant expectations after reading the Bible personally and, deriving from this, on the promotion of spreading literacy among different strata of the population. In this context, rising literacy and general education rates gradually led to more multifaceted, organized, and profitable forms of economic life, which were accompanied by technological innovations – indeed, in the Western world, capitalism first asserted itself in mostly Protestant environments, where the literacy rates were highest and grew the most rapidly. There is a notable difference in literacy rates in Lutheran German lands, where around 40% of the population was literate as early as the start of the eighteenth century, or Lutheran Sweden, where literacy encompassed over 80% of the population at the start of the nineteenth century, on the one hand and Catholic Italy, where only around 20% of the people were literate at the same time, on the other. Even though the establishment of parish elementary schools had been encouraged already by the Council of Trent (1545–1563), the Catholic Church became actively involved in elementary education only at the request of state authorities in the late eighteenth century. (Mosher 2016, 397; 399–401; 407; Becker and Woessmann 2009)

It is worth highlighting that the Catholic Church may have been the main provider of scientific and educational activity in Europe, but it reached only a small part of society in this field until the nineteenth century. Moreover, the Catholic Church did not oppose reading the Bible as such, but it did discourage individuals with no theological education from reading it as it feared the emergence of individual interpretations, which could come in conflict with its teachings and cause disunity in the Church. Unlike official Catholicism, Jansenism, which was similar to Protestantism, encouraged reflection alongside passages from the Bible to strengthen moral life – this was one of the reasons Jansenists and Enlightenment-oriented Catholics were so strongly in favour of education accessible to the masses. Interestingly, in the eighteenth and early nineteenth centuries, the Catholic world had the highest literacy rates – up to half the population – in religiously heterogeneous lands or in areas bordering Protestant territories, e.g. in Catholic parts of Belgium or German lands. This was in part due to imitation of nearby Protestant practices and in part due to a concern for preserving the Catholic identity in the population, i.e. preventing students from going to ‘competing’ Protestant schools. The latter was one of the key reasons why Johann Felbiger, who was from the religiously heterogeneous Silesia, created his reform program. (Mosher 2016, 405; 407; Doney 1988, 5; Lehner 2010, 15)

Catholic priests and Protestant pastors played a key role in establishing an Enlightenment spirit among the majority rural population, especially by preaching in churches, visiting homes, teaching in schools and writing or translating different handbooks. They thus enhanced their position of spiritual authority with an educational teaching role and supported the central concern of the Enlightenment – addressing the ignorance of the mind and maximizing work efficiency for personal and social wellbeing. An example is provided by the collection of entertaining and instructing stories titled *A Booklet to Help Peasants (Noth- und Hülfsbüchlein für Bauersleute, 1788)*, which became one of the most popular works of the late eighteenth century. Its author was Rudolph Becker (1752–1822), a Prussian writer and pedagogue and representative of the Lutheran conservative current of the German Enlightenment. It was intended for rural youth in particular, using words and pictures to warn them of work hazards, imparting the basics of agriculture and medicine, and promoting diligent, frugal, and pious life. It was printed in approximately half a million copies, translated to different languages, and saw one hundred and nineteen editions. A Slovenian translation was provided by Marko Pohlin (1735–1801), a member of the Order of Discalceate Augustines, priest, and linguist, one of the leading representatives of the Enlightenment-oriented priesthood in Carniola. In the Central European priesthood, the mission of teaching people not only about spiritual matters but also practical and economic skills, which was promoted by Enlightenment thinking, was preserved even in the following decades, when the Enlightenment was replaced by other ideologies, such as the formation of a national consciousness. (Studen 2014, 15–17; Doney 1988, 3–4; 27; Lehner 2010, 14–15; 27)

6. Conclusion

Reforms in the field of education aiming for the establishment of state-powered mass education accessible to children from all social strata did not arise in an ideationally and organizationally empty space. In Central European environments, e.g. Austria, Prussia, and other German lands, the state school networks relied on initiatives by Christian Churches and were part of Enlightenment currents trying to reconcile the Christian faith with a rational view of life and with state interests. In establishing an elementary school network, such efforts were most notable in areas with a Protestant majority, starting no later than the end of the sixteenth century. This is because the demand of Reformers to read the Bible personally as a necessary means to strengthen the relationship with God presupposed literacy. As a result, the Lutheran Church recommended providing basic education to the maximum possible segment of the population and, accordingly, printing books, which led to the development of educational institutions of different levels and to a gradual rise in the overall education rate. The success of the Protestant school system was facilitated by the moral and financial support of the political authorities as they enabled Protestantism to have the status of the state Church under the supreme leadership of the secular ruler. As a result, Enlightenment followers could align the existing positive attitude towards education with their own views, placing education in the context of fulfilling the needs of the increasingly centralized state. (Soysal and Strang 1989, 280; 285–286) The disbandment of the Jesuit order provided an important incentive for intensifying state interventions in the field of education. In the Catholic world, this circumstance was a sign of wider cultural change symbolically ending the Baroque period – the ‘empty space’ that emerged was filled by the state, which started to gradually adjust it with new approaches.

To increase its power and intervene in all areas of public life, the state needed qualified bureaucrats, engineers, and military commanders; it had to create a unified culture and assert a national consciousness among the population. After all, the reform plans of Johann Felbiger, which were adopted by Maria Theresa, arose from his personal experience of living in an area with a religiously heterogeneous population, which was claimed by two countries. This enabled him to draw comparisons with more successful schools of Protestant communities, in which he recognized a challenge to the Catholic faith and culture. From the Enlightenment on, most notably in the nineteenth century, Central European state-powered education systems aimed to mould young people into diligent and moral individuals who identified with the state and strove to strengthen it and, as a consequence, recognized compatriots in other individuals – irrespective of social, regional, and religious differences. Thus, it was exactly through the decisive help of mass elementary education that during the nineteenth century, trust in reason and an increasingly centralized state created a new bourgeois and industrial society, with the ideal of a productivity-oriented individual devoted to the state at its core. (Green 1994, 9–11)

In terms of ideas, state interventions in education, which had been the domain of the Catholic Church for centuries in Austria and certain German lands, was based on criticism of Baroque folk religiosity marked by lavish ceremonies and processions, the veneration of saints, and numerous holidays. The effects of such religiosity were disapproved by an increasing number of intellectuals, both laypeople and priests, even bishops and especially theology professors. Instead, they advocated for 'rational' religiosity focused on the reading of spiritual literature and a highly virtuous life intertwined with the expansion of literacy, the teaching of Catholic doctrine to all strata of the population, and a sense of following instructions by the state. On the other hand, Pietism in Prussia and other Lutheran German lands – by emphasizing personal responsibility and self-reflection for own and community life and the necessity of accessible education in light of fulfilling religious and social duties – came close to the Enlightenment and facilitated modernization processes; as a result, there were similarities to the streams of reform and Janse- nism in the Catholic Church. It was Catholic priests and Protestant pastors who played a key role in establishing an Enlightenment spirit among the majority rural population, especially by preaching in churches, teaching in schools and writing different handbooks. They enhanced their position of spiritual authority and supported the main concern of the Enlightenment – diminishing the ignorance of the mind and maximizing work efficiency for personal and social wellbeing.

References

- Avis, Paul.** 2022. *Theology and the Enlightenment: A Critical Enquiry into Enlightenment Theology and Its Reception*. London: Bloomsbury.
- Batič, Matic.** 2017. Verski skladi na Slovenskem v zgodovinskem kontekstu. *Res novae* 2, no. 2:101–134. <https://doi.org/10.62983/rn2865.172.4>
- Beales, Derek.** 1997. Joseph II and the Monasteries of Austria and Hungary. In: Nigel Aston, ed. *Religious Change in Europe, 1650–1914 (Essays for John McManners)*, 161–184. Oxford: Oxford University Press.
- Becker, Sascha O., and Ludger Woessmann.** 2009. Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics* 124, no. 2:531–596. <https://doi.org/10.1162/qjec.2009.124.2.531>
- Borgstedt, Angela.** 2004. *Das Zeitalter der Aufklärung*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Burson, Jeffrey D., and Ulrich L. Lehner.** 2014. *Enlightenment and Catholicism in Europe: A Transnational History*. Notre Dame, IN: Notre Dame Press.
- Caruso, Marcelo, and Daniel Töpfer.** 2019. Schooling and the Administrative State: Explaining the Lack of School Acts in Nineteenth-Century Prussia. In: Johannes Westberg, Lukas Boser and Ingrid Brühwiller, eds. *School Acts and the Rise of Mass Schooling: Education Policy in the Long Nineteenth Century*, 41–66. London: Palgrave Macmillan.
- Ciperle, Jože, and Andrej Vovko.** 1987. *Šolstvo na Slovenskem skozi stoletja*. Ljubljana: Slovenski šolski muzej.
- Cvrček, Tomáš.** 2020. *Schooling Under Control: The Origins of Public Education in Imperial Austria 1769–1869*. Tübingen: Mohr Siebeck.
- Dietrich, Teo, and Job-Günter Klink, eds.** 1972. *Zur Geschichte der Volksschule*. Vol. 1. Bad Heilbrunn: Julius Klinkhardt.
- Doney, John C.** 1988. The Catholic Enlightenment and Popular Education in the Prince-Bishopric of Würzburg. *Central European History* 21, no. 1:3–30. <https://doi.org/10.1017/s0008938900012644>
- Engelbrecht, Helmut.** 1984. *Geschichte des österreichischen Bildungswesens: Erziehung und Unterricht auf dem Boden Österreichs*. Vol. 3, *Von der frühen Aufklärung bis zum Vormärz*. Vienna: ÖBV.

- Felbiger, Johann I.** 1774. *Vorlesungen über die Kunst zu catechisieren, die er seinen jungen Geistlichen zu halten pflegt*. Vienna: Thomas von Trattner.
- —. 1775. *Methodenbuch für Lehrer der deutschen Schulen in den kaiserlich-königlichen Erbländen*. Vienna: Verlagsgewölbe der deutschen Schulanstalt.
- Hammerstein, Notker.** 1977. *Aufklärung und katholisches Reich: Untersuchungen zur Universitätsreform und Politik katholischer Territorien des Heiligen Römischen Reichs deutscher Nation im 18. Jahrhundert*. Berlin: Duncker & Humblot.
- Hollweger, Hans.** 1976. *Die Reform des Gottesdienstes zur Zeit des Josephinismus in Österreich*. Regensburg: Pustet.
- Horn Melton, James V.** 1988. *Absolutism and the Eighteenth-Century Origins of Compulsory Schooling in Prussia and Austria*. Cambridge: Cambridge University Press.
- Gawthrop, Richard L.** 2006. *Pietism and the Making of Eighteenth-Century Prussia*. Cambridge: Cambridge University Press.
- Gierl, Martin.** 2013. Pietism, Enlightenment, and Modernity. In: Douglas H. Schantz, ed. *An Introduction to German Pietism: Protestant Renewal at the Dawn of Modern Europe*, 348–392. Baltimore: Hopkins Press.
- Green, Andy.** 1994. Education and State Formation Revisited. *Historical Studies in Education* 6, no. 3:1–17.
- Hersche, Peter.** 1977. *Der Spätjansenismus in Österreich*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Königlich-Preußisches.** 1763. *Königlich-Preußisches General-Land-Schul-Reglement, wie solches in allen Landen Seiner Königlichen Majestät von Preussen durchgehends zu beobachten: De Dato Berlin, den 12. August, 1763*. Magdeburg: Hechtel.
- Klueting, Harm.** 2010. The Catholic Enlightenment in Austria or the Habsburg Lands. In: Ulrich Lehner, ed. *A Companion to the Catholic Enlightenment in Europe*, 127–164. Leiden: Brill.
- Krömer, Ulrich.** 1966. *Johann Ignaz von Felbiger. Leben und Werk*. Freiburg: Herder.
- Laun, Andreas.** 1979. Die Moralthologie im 18. Jahrhundert unter dem Einfluss von Jansenismus und Aufklärung. In: Elisabeth Kovács, ed. *Katholische Aufklärung und Josephinismus*, 266–294. Vienna: Verlag für Geschichte und Politik.
- Lehner, Ulrich L.** 2010. The Many Faces of the Catholic Enlightenment. In: Ulrich Lehner, ed. *A Companion to the Catholic Enlightenment in Europe*, 1–61. Leiden: Brill.
- Leschinsky, Achim, and Peter M. Roeder.** 1983. *Schule im historischen Prozeß*. Frankfurt: Ullstein.
- Loch, Werner.** 2004. Pädagogik am Beispiel August Hermann Franckes. In: Hartmut Lehmann and Ruth Albrecht, ed. *Glaubenswelt und Lebenswelten*, 264–308. Göttingen: Vandenhoeck & Ruprecht.
- Marschke, Benjamin.** 2013. Pietism and Politics in Prussia and Beyond. In: Douglas H. Schantz, ed. *An Introduction to German Pietism: Protestant Renewal at the Dawn of Modern Europe*, 472–526. Baltimore: Hopkins Press.
- Mosher, James S.** 2016. The Protestant Reading Ethic and Variation and its Effects. *Sociological Forum* 31, no. 2:397–418. <https://doi.org/10.1111/socf.12250>
- Nežič Glavica, Iva.** 2023. ‚Pedagogika 3.0‘ ali oris edukacije v digitalni dobi. *Bogoslovni vestnik* 83, no. 4:1049–1060. <https://doi.org/10.1111/socf.12250>
- Okoliš, Stane.** 2008. *Zgodovina šolstva na Slovenskem*. Ljubljana: Slovenski šolski muzej.
- Printy, Michael.** 2010. Catholic Enlightenment and Reform Catholicism in the Holy Roman Empire. In: Ulrich Lehner, ed. *A Companion to the Catholic Enlightenment in Europe*, 165–213. Leiden: Brill.
- Schmidt, Vlado.** 1988. *Zgodovina šolstva in pedagogike na Slovenskem*. Vol. 1. Ljubljana: Delavska enotnost.
- Soysal, Yasemin N., and David Strang.** 1989. Construction of the First Mass Education Systems in Nineteenth-Century Europe. *Sociology of Education* 62, no. 4: 277–288. <https://doi.org/10.2307/2112831>
- Stollberg-Rilinger, Barbara.** 2017. *Maria Theresia: Die Kaiserin in ihre Zeit. Eine Biographie*. München: C. H. Beck.
- Studen, Andrej.** 2014. Z vzgojo do največje možne sreče ljudstva: nekaj poudarkov iz zgodovine »razsvetljevanja ljudstva« na Slovenskem v 18. in v prvi polovici 19. stoletja. *Prispevki za novojšo zgodovino* 54, no. 1:9–22.
- Škulj, Edo, ed.** 2004. *Herbersteinov simpozij v Rimu*. Celje: Mohorjeva družba.
- Trontelj, Nik.** 2022. Ločenost Cerkve in države v Evropi: zgodovina in sedanji trenutek. *Edinost in dialog* 77, no. 2:15–27. <https://doi.org/10.34291/edinost/77/02/trontelj>
- Van Horn Melton, James.** 1988. *Absolutism and the Eighteenth-Century Origins of Compulsory Schooling in Prussia and Austria*. Cambridge: Cambridge University Press.

- Van Kley, Dale K.** 2018. *Reform Catholicism and the International Suppression of the Jesuits in Enlightenment Europe*. Yale: Yale University Press.
- Viehhauser, Martin.** 2019. "Das Schulwesen aber ist und bleibt allezeit ein politicum": The Felbiger General School Ordinance and School Reform in the Eighteenth-Century Habsburg Monarchy. In: Johannes Westberg, Lukas Boser, and Ingrid Brühwiller, eds. *School Acts and the Rise of Mass Schooling: Education Policy in the Long Nineteenth Century*, 17–40. London: Palgrave Macmillan.
- Vodičar, Janez.** 2023. Avtoriteta na področju vzgoje in verovanja v digitalni dobi. *Bogoslovni vestnik* 83, no. 4:1035–1047. <https://doi.org/10.34291/bv2023/04/vodicar>
- Westberg, Johannes, Lukas Boser and Ingrid Brühwiller, eds.** 2019. *School Acts and the Rise of Mass Schooling: Education Policy in the Long Nineteenth Century*. London: Palgrave Macmillan.
- Weston Evans, Robert J.** 1979. *The Making of the Habsburg Monarchy, 1550–1700*. Oxford: Oxford University Press.
- Wolfsgruber, Cölestin.** 1897. *Christoph Anton Kardinal Migazzi, Fürsterzbischof von Wien. Eine Monographie und zugleich ein Beitrag zur Geschichte des Josophinismus*. Ravensburg: Hermann Kitz.
- Zschokke, Hermann.** 1894. *Die theologischen Studien und Anstalten der katholischen Kirche in Österreich*. Vienna: Braumüller.

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 823—834

Besedilo prejeto/Received:10/2024; sprejeto/Accepted:11/2024

UDK/UDC: 004.8:342.7:17

DOI: 10.34291/BV2024/04/Osredkar

© 2024 Osredkar, CC BY 4.0

Mari Jože Osredkar

Umetna inteligenca – prepovedan sad iz raja?

Artificial Intelligence – Forbidden Fruit from The Garden of Eden?

Povzetek: V prispevku bomo pokazali, da je umetna inteligenca v zahodni družbi sestavni del procesa človekovega osvobajanja izpod Božje superiornosti nad človekom. V zahodnem svetu je tisočletje in sedemsto let človek svoje življenje uravnal bo Božji volji, izraženi v Desetih božjih zapovedih. Od razsvetljenstva dalje (18. stol.) pa je Evropa začela Bogu odvzemati avtoriteto in življenje družbe ter posameznikov se je začelo odvijati po „človeški pameti“! Filozofi so objavili smrt Boga, sociologi pa konec religije; v sekulariziranem svetu na mesto Božjih zapovedi človek postavi „Deklaracijo človekovih pravic“. Na gospodarskem področju smo od 18. stoletja do danes priče hitremu razvoju in tehnološkim revolucijam, ki se čedalje hitreje vrstijo druga za drugo. Na videz se človek osvobodi izpod Božje nadvlade (moralno področje); zdi se celo, da je človek stopil na mesto Boga (uspešen pri iznajdbah in razvoju strojev). Človek je omejitve, ki mu jih je nalagal Bog, zamenjal za svobodo. Na današnji stopnji razvoja umetne inteligence pa je človeka postalo strah za preživetje. Morda se bo ravno zaradi strahu vrnil k Bogu oz. k spolnjevanju Božje volje.

Ključne besede: umetna inteligenca, Božje zapovedi, sekularizacija, „Deklaracija človekovih pravic“, strah za preživetje

Abstract: In this article, we will demonstrate that artificial intelligence is an integral part of the process of liberating humanity from God's supremacy in Western society. For 1,700 years in the Western world, human life was guided by God's will as expressed in the Ten Commandments. However, since the Enlightenment (18th century), Europe began to strip away God's authority, and society, along with individuals, started to live according to "human reason." Philosophers declared the death of God, and sociologists proclaimed the end of religion. In a secularized world, the "Declaration of Human Rights" replaced the commandments of God. Economically, from the 18th century to the present, we have witnessed rapid development and successive technological revolutions. It seems that humanity has freed itself from divine domination in the moral sphere; it even appears that humans have taken the place of God, becoming successful in inventions and technological advancements. The con-

straints once imposed by God have been replaced by freedom. However, at the current stage of artificial intelligence development, humanity has begun to fear for its survival. Perhaps it is precisely this fear that will lead humans to return to God, or to once again follow His will.

Keywords: artificial intelligence, God's commandments, secularization, "Declaration of Human Rights", fear for survival

1. Uvod

V Svetem pismu beremo, da je »Bog ustvaril človeka po svoji podobi!« (1 Mz 1,26)¹ Postavil ga je v edenski vrt (2,15) in mu zapovedal: »Z vseh dreves v vrtu smeš jesti, le z drevesa spoznanja dobrega in hudega nikar ne jej! Kajti na dan, ko bi jedel z njega, boš gotovo umrl!« (2,16-17) Adam ve, kaj mu Bog dovoli in kaj mu prepoveduje. Seznanjen je, da bo moral umreti, če bo kršil Božjo voljo. Toda prvi človek kljub temu ne uboga svojega Stvarnika in naredi po svoje. Kot bi si bil mislil: »Kaj boš Ti meni ukazoval! Jaz vem bolje od Tebe, kaj meni koristi!« Ko je Bog Adamu ustvaril še primerno pomočnico (2,18), sta si namreč postregla s prepovedanega drevesa spoznanja. Kačine besede, da bosta »postala kakor Bog« (35) lahko razumemo v smislu, da bosta tako sposobna, da Boga sploh ne bosta več potrebovala. Ravnala sta po svoji volji, po svoji pameti in sta bila prepričana, da si lahko sama najbolje uredita življenje. Skratka, Adam si je domišljjal, da ne potrebuje Božjih navodil, ker je sam dovolj sposoben in se lahko 'svobodno' odloča za uspeh v življenju. Vendar se zgodba v Genezi obrne drugače. Ko Bog išče prvega človeka, mu ta odgovori: »Slišal sem tvoj glas v vrtu, pa sem se zbal, ker sem nag, in se skril.« (3,10) Adama je postalo strah, bal se je! Pa ne le Boga, temveč predvsem samega sebe in svojih dejanj. Zelo hitro je spoznal iluzornost svoje želje po enakosti z Bogom oz. mu je postalo jasno, da je majhen in nesposoben. Ni postal enak Bogu, nasprotno, izgubil je Božjo bližino, ki jo pooseblja rajski vrt. Ko mora zapustiti svoj 'domači kraj' in se »v potu svojega obraza truditi« za preživetje (3,17), mu je dokončno jasno, da je 'zavozil' svoje življenje.

Adam v rajskem vrtu je podoba novodobnega človeka, ki je zavrgel Božje zapovedi in se začel ravnati po svoji pameti. V prepričanju, da lahko sam, brez Boga, doseže polnost življenja, si je na mesto Božjih zapovedi postavil svoja pravila za življenje; na mesto Božjih prepovedi je postavil svojo etiko; svoje sposobnosti pa je uporabil za tehnološke iznajdbe. Tako kakor Adam v Genezi tudi človek digitalne dobe počasi dojema, da je 'zavozil'; ne le svoje življenje, temveč da je planet, na katerem živi, pripeljal na rob propada.

Pri raziskovanju smo uporabili empirično metodo, ki temelji na opazovanju in zaznavanju dejstev. V humanistiki sicer ne uporabljamo meritev in eksperimenti-

¹ Prispevek je nastal v okviru raziskovalnega programa „Religija, etika, edukacija in izzivi sodobne družbe (P6-0269)“, ki ga sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

ranja, pa je kljub temu empirična metoda tudi za nas uporabna. S pomočjo sistematične metode opazovanja, ki jo je mogoče ponoviti in preveriti, smo zbrali in analizirali podatke, ki smo jih pridobili z opazovanjem virov, torej dokumentov, ter analizirali besedila božjih zapovedi in Deklaracije človekovih pravic.

V prispevku nimamo namena razpravljati o umetni inteligenci s tehnološkega vidika, ker to ni naše poslanstvo in za to delo niti nismo kompetentni. Osredotočili se bomo na odziv teologije na človekovo uporabo moderne tehnologije in v luči Svetega pisma predstavili odgovor na vprašanje, ali moderna tehnologija v luči Božjega razodetja človeku koristi ali škoduje. Osredotočili se bomo na vprašanje, kje je človek naredil napako, da ga je danes, ko je visoko tehnološko razvit, strah samega sebe.

2. Sekularizacija

Sekularizacijo običajno opredeljujemo kot izrivanje Boga iz javnega življenja oz. prepričanje, da ljudje lahko vse storimo brez Boga. To drži, toda bistvo oddaljevanja od Boga je mišljenje, da človek ne potrebuje omejitev. V Prvi Mojzesovi knjigi beremo, kako Eva razloži kači, kaj jima Bog prepoveduje oz. napoveduje njuno smrt, če bosta pojedla prepovedan sad. Kača nato reče ženi: »Nikakor ne bosta umrla! V resnici Bog ve, da bi se vama tisti dan, ko bi jedla z njega, odprle oči in bi postala kakor Bog, poznala bi dobro in húdo.« (1 Mz 3,4-5) In sta jedla (3,6). Kača v bistvu prekliče Božjo prepoved in jima obljublja življenje brez omejitev. Zato lahko zapišemo, da se je sekularizacija začela z Adamom oz. da je bil Adam prvi liberalec, ki je ignoriral Božjo zapoved.

Kakor prvi človek želi tudi sedanja zahodna družba biti svobodna in delati vse po svojih željah in svoji pameti. Še več, liberalna družba v bistvu teži za tem, da bi naredila vse, kar je v njeni moči, da bi ignorirala omejitve in zgrabila vse, kar je na dosegu njenih rok. Nasprotovanje Cerkvi je v liberalni družbi v bistvu nasprotovanje omejitvam, ki jih krščanski nauk postavlja človeku. Vsako novo orodje namreč človeku daje občutek, da zanj nič več ni nemogoče, da nima več meja – da si lahko pridobi vse, kar si želi; skratka, da se lahko odreši sam, to je, da si lahko sam omogoči preživetje. V tem procesu liberalizacije družbe je pomenljiva primerjava Božjih zapovedi in „Deklaracije človekovih pravic“², ki jo je liberalna družba postavila na mesto *Dekaloga*.

Dekalog sestavljajo tri zapovedi: Veruj v enega Boga; Posvečuj Gospodov dan in Spoštuj očeta in mater, da boš dolgo živel in ti bo dobro na zemlji; ter sedem prepovedi: Ne skruni božjega imena; Ne ubijaj; Ne nečistuj; Ne kradi; Ne pričaj po krivem; Ne želi svojega bližnjega žene; Ne želi svojega bližnjega blaga. V bistvu je to deset omejitev, ki se dotikajo najbolj intimnih področij človekovega bivanja. Ker mnogi v zahodni družbi niso več priznavali Božje avtoritete, je bila leta 1789, s

² Slovenski prevod „Deklaracije človekovih pravic“ smo vzeli iz: Uradni list RS, št. 24/2018 z dne 13. 4. 2018.

Francosko revolucijo, sprejeta „Deklaracija človekovih pravic“, ki v 30 členih naveda 26 pravic in svoboščin, ki govorijo o najrazličnejših področjih posameznikovega življenja; od pravice do življenja, dela, enakosti pred zakonom do vere, družine, državljanstva, premoženja, izražanja in socialne varnosti ... Poleg teh pa v dokumentu najdemo zgolj tri prepovedi oz. eno dolžnost: nikogar se ne sme podvreči mučenju ali krutemu, nečloveškemu ali ponižujočemu ravnanju ali kaznovanju. (5. člen) Nikogar se ne sme samovoljno prijeti, pridržati ali izgnati. (9. člen) Nikomur se ne sme nihče samovoljno vmešavati v zasebno življenje, družino, dom ali dopisovanje niti ne sme žaliti njegove časti in dobrega imena. Vsakdo ima pravico do pravnega varstva pred takšnim vmešavanjem ali napadi. (12. člen) Vsakdo ima dolžnosti do skupnosti, kajti samo v njej je mogoč svoboden in poln razvoj njegove osebnosti. (29. člen)

Že sama razglasitev „Deklaracije“ je znamenje, da je človek zasedel mesto, ki ga je v družbi prej imel Bog. Na ravni družbenega življenja je laična država omeji la vpliv Cerkve in religijo zamenjala z racionalističnim pogledom na svet (Trontelj 2022, 21). Primerjava Božjih zapovedi in določb „Deklaracije“ pa nam očitno pokaže, da Bog postavlja človeku omejitve, človek pa si (sam sebi) postavlja zgolj pravice. Namreč tudi tri prepovedi oz. ena dolžnost v „Deklaraciji“ so le na drugačen način izražene formulacije človekovih pravic. Nadaljnja primerjava Božjih zapovedi in „Deklaracije“ pa nam razodeva še eno zanimivo značilnost. Božje zapovedi od človeka zahtevajo napor; za njihovo uresničitev je treba vložiti veliko truda. Ko Bog človeku nalaga dolžnosti oz. omejitve, mu v bistvu govori: »V potu svojega obraza boš jedel kruh!« (1 Mz 3,19) Trud in trpljenje sicer lahko razumemo kot terapijo, ki jo Bog človeku nudi, da bi ga lahko odrešil, da bi mu lahko omogočil življenje, ki ga je z Adamovim padcem človek izgubil. „Deklaracija človekovih pravic“ pa že na prvi videz poizkuša človeku omogočiti, da bi zlahka, brez napora (in naporne terapije) živel in preživel; „Deklaracija“ od družbe zahteva, da posamezniku nudi vse, kar potrebuje za življenje, oz. je novodobni človek prepričan, da mu vse pripada in da mu je vse dovoljeno, ker je pač človek. V kontekstu človekove tendence, da bi sam, s čim manj truda, dosegel polnost življenja, lahko razložimo tudi izdelke človeških rok, ki jih imenujemo tehnologija.

3. Tehnološki napredek in etika

Človeku je lastno, da si pri delu pomaga z orodji, ki jih stalno izpopolnjuje. Zato najdemo v zgodovini tako imenovane tehnološke revolucije, ki so pomembni mejniki v razvoju izpopolnjevanja človekovih pripomočkov pri delu. Kot prvo lahko omenimo začetek kmetovanja, ko je človek opustil nomadsko življenje, zgolj pobiranje sadežev in lov. Kmet se je ustalil na nekem kraju in tam s preprostim orodjem obdeloval zemljo, ki mu je dajala sadove oz. hrano. Naslednja je industrijska revolucija, ko je človek izumil parni stroj, ki je nadomestil mnogo delavcev. Tretja stopnja v razvoju je elektrifikacija. Ne le, da je elektrika zamenjala paro, temveč je človeku omogočila delo ponoči. Informacijska revolucija je prinesla računalnike

in pozneje internetno omrežje. Ne smemo pozabiti niti atomske energije, s katero si človeštvo pomaga na različnih gospodarskih področjih, žal pa jo človek lahko uporablja tudi v vojaške namene za uničevanje. Zadnja stopnja (do sedaj) v tehnološkem razvoju pa je umetna inteligenca, torej stroj, ki se sam izpopolnjuje in uči, kako delati namesto človeka. Lahko bi rekli, da si človek želi, da bi ga umetna inteligenca posnemala v razmišljanju.

»Ni slabo«, bi ocenil nek zunanji opazovalec človekovega tehnološkega razvoja. Ampak že na prvi stopnji omenjenega razvoja bi se opazovalcu postavilo vprašanje, ali je tehnološki napredek za človeka dober, koristen, ali ne. Kot smo že zapisali, se zdi, da želi človek narediti vse, kar je v njegovi moči. Ampak, ali je to dobro? Na to vprašanje bodo različni ljudje različno odgovorili; glede na njihovo razumevanje dobrega in slabega. Izrazit primer različnega pojmovanja, kaj je dobro in kaj slabo, vidimo v sodobni družbi pri opredeljevanju do abortusa. V našem razmišljanju izhajamo iz opredelitve, da je dobro to, kar človeštvu kot skupnosti omogoča preživetje oz. nadaljevanje življenja. Ne le, kar posamezniku omogoča preživetje, temveč, kar omogoča preživetje celotni človeški skupnosti. Etika je nauk o dobrem, torej je to veda, ki opredeljuje, kaj človeški skupnosti omogoča preživetje in hkrati človeka opozarja in kaže na to, kaj človeštvo vodi v izginotje. Etično je torej to, kar človeštvu omogoča preživetje. Tudi, ko se izrazimo, da je nekaj dobro ,za človeka', tu ne mislimo posameznika, temveč imamo v mislih celotno človeštvo. V kontekstu takega pojmovanja etike se lahko vprašamo, ali človeku tehnološki pripomočki pri delu resnično nudijo življenje v polnosti ali pa so trojanski konj, v katerem se skriva nevarnost za človekov propad. Od industrijske revolucije pa do danes opažamo dve nasprotujoči si mnenji oz. dve reakciji ljudi. V 18. stoletju so v Angliji delavci uničevali stroje v tovarnah, ker so bili prepričani, da jim grenijo življenje. Zaradi strojev je veliko ljudi postalo brezposelnih, stroji so pomenili tudi začetek onesnaževanja okolja ... Isto vprašanje in podobne reakcije imamo tudi na (do sedaj) zadnji stopnji razvoja, ko izdelujemo stroj, ki bo znal sam razmišljati. Danes so očitki še tehtnejši. Ne le manj delovnih mest za ljudi, temveč nevarnost, da se bo stroj obrnil proti človeku in ga uničil. Alternativa se torej postavlja: ali podpirati ali zavirati razvoj? V družbi obstajata dve skupini ljudi: progresivna skupina, ki bi želela, da gre znanost z naglico naprej in da človek čim prej stori vse, kar mu znanost omogoča. Na drugi strani pa imamo skupino, ki zavira razvoj oz. ki ne želi, da človeštvo hiti in v naglici naredi napačne korake in se ne bo moglo več potegniti nazaj, ko bo ugotovilo, da mu taka tehnologija škoduje. V sodobni družbi je veliko takih dilem. Tudi v Cerkvi je nesoglasje glede novega in starega liturgičnega izražanja vere ... Eni bi radi bliskovit razvoj, takojšnje velike spremembe, drugi bi želeli, da se nič ne spremeni, celo, da se vrnemo na ,staro preizkušeno prakso'! Ti dve skupini druga drugo pogosto onemogočata pri delovanju. V nadaljevanju si bomo pogledali, kako o danem položaju razmišlja teologija.

4. Teološki problem

Teologi ne vidijo problema v moderni tehnologiji oz. strojih kot pripomočkih za delo, ki lahko človeku olajšajo trud in mu omogočijo lažje življenje. Na prvih straneh Svetega pisma lahko preberemo, da je Bog prva dva človeka blagoslovil, jima izročil zemljo, da ji gospodujeta: »Bodita rodovitna in množita se, napolnita zemljo in si jo podvrzita; gospodujta ribam v morju in pticam na nebu ter vsem živalim, ki se gibljejo po zemlji!« (1Mz 1,28) V Drugi Mojzesovi knjigi nam Sveto pismo poroča, da je Bog ljudem dal celo svojega Duha, da bi imeli »modrost, razumnost in vednost za vsakršno delo« (2Mz 35,31). To, da smo ustvarjeni po Božji podobi in da imamo Božjega Duha, torej pomeni, da je tehnološki razvoj po Božji volji. Dokument zadnjega vesoljnega cerkvenega zbora to dejstvo potrjuje: »Vedno je človek s svojim delom in duhovno močjo skušal bogateje razviti svoje življenje.« (GS 33) Tudi papež Frančišek je v Poslanici za svetovni dan miru 2024 zapisal: »Ko si ljudje s pomočjo znanosti in tehnike prizadevajo, da zemlja postaja dostojno bivališče za celotno človeško družino, delujejo po Božjem načrtu in sodelujejo z Njegovo voljo.« (Frančišek 2024, 1) Motika, parni stroj ali umetna inteligenca, med njimi v bistvu ni velike razlike; to so človekovi pripomočki za delo, ki jih krščanska teologija, na temelju Božje besede, popolnoma odobrava. Pač pa teologija vidi problem v človekovem mišljenju oz. prepričanju, ki pogosto spremlja tehnološki razvoj, da zahvaljujoč sposobnostim svojega razuma in posledično zahvaljujoč novim iznajdbam ne potrebuje Boga. Tu se novodobni človek popolnoma nič ne razlikuje od svetopisemskega Adama, ki je že v rajskem vrtu razmišljal, da ga bo spoznanje, védenje oz. sposobnost njegovega razuma odrešila in ga osvobodila izpod Božjega gospostva. Zato se je novodobni človek znašel v identičnem položaju kot prvi človek; strah ga je! Teologija torej vidi srž problema v razumevanju Božjega odrešenja. Ta tematika predstavlja najpomembnejše poglavje v teološki vedi; lahko zapišemo, da je nauk o odrešenju srčika vsakega religijskega nauka. Pravzaprav gre pri novodobnem človeku za problematično samo-odrešenje, ki ne le, da zavrača Božje odrešensko dejanje, temveč predvsem zavrača Boga in z Njim omejitve, ki jih Bog človeku postavlja. Tak način razmišljanja vodi v prepričanje, da bo popolnoma svobodni človek lahko sam sebi omogočil življenje v polnosti. Za lažje nadaljevanje našega razmišljanja bomo v nadaljevanju tematiko krščanskega pojmovanja odrešenja nekoliko osvetlili in izpostavili razliko med odrešenjem in samo-odrešenjem. Tako bomo spoznali teološko kritiko novodobnega človeka, ki raziskuje, se tehnološko razvija in si lajša življenje s svojimi iznajdbami.

4.1 Problem razumevanja odrešenja

Ko govorimo o odrešenju, se na poseben način dotaknemo življenja in tudi smrti³. Odrešenje je namreč pot iz smrti v življenje. Pravzaprav biti odrešen pomeni živeti, odrešiti nekoga pa pomeni omogočiti mu življenje. Obratno pa je teološki pojem za pot iz življenja v smrt pogubljenje. O vprašanju življenja in smrti ter o poti v življenje govorijo vse religije, filozofije in ideologije (pot iz zla k dobremu)

³ Pod besedo „smrt“ tukaj ne pojmujeemo le biološkega konca življenja, temveč vse, kar izničuje življenje.

ter konec koncev tudi znanost. Ljudje si niti ne postavljajo vprašanja, če je prehod od slabega k dobremu mogoč – hočejo ga! Znanstveniki ga iščejo po svojih lastnih močeh. Ljudje vere pa se pustijo voditi Božji volji, ki presega človeške sposobnosti (razodetje). Ona jih usmerja proti dobremu (včasih gre za časno, včasih za večno dobro). Slabo je imenovano greh. To, po čemer se različne religije razlikujejo med seboj, je razumevanje odnosa med Absolutnim in človekom. Ni pa izključeno, da se ta odnos spreminja v različnih zgodovinskih obdobjih. Vsekakor pa moramo razlikovati človekovo prizadevanje za rešitev iz slabega in religijski nauk, recimo Kristusov odrešenjski nauk, ki je sistem, sklop navodil, kako naj človek živi, da bo dosegel »življenje v polnosti« (Osredkar 2012, 11–23)! Pri posameznikovem prizadevanju za rešitev iz slabega položaja in iskanju dobrega pa zavzema posebno poglavje, ki ga imenujemo ‚samo-odrešenje‘.

Samo-odrešenje izraža samozadostnost pri iskanju in doseganju polnosti odrešenja. Avtosoterija je prepričanje, da je človek zmožen doseči odrešenje zgolj s svojo pametjo in z lastnimi prizadevanji. Tako, kakor smo že navedli zgoraj, da je bil Adam prepričan, da si sam lahko omogoči lepo življenje in da sploh ne potrebuje Boga. Z naukom o samo-odrešenju se srečamo na primer pri religijah in filozofijah daljnega vzhoda; nauk o preseljevanju duš uči, da si vsak človek pridobi življenje zgolj s svojim lastnim trudom oz. dejanji. (Osredkar 2012, 24–26) Posebno poglavje samo-odrešenja pa predstavljajo grške filozofske šole. Ker je evropsko razsvetljenstvo izšlo iz grškega racionalizma, bomo samo-odrešenju v grški filozofiji posvetili nekaj vrstic.

»Radikalna razlika v razumevanju pojma odrešenja pri Grkih in Hebrejcih se kaže pri sredstvu odrešenja. Sredstvo odrešenja je za Grke razum, katerega naloga je privedi človeka nazaj k njegovemu božanskemu bistvu, ki ga človek nosi v globinah samega sebe. Za Hebrejce pa je sredstvo odrešenja zgodovina. Zgodovina je moč ustvarjanja novega, je rezervoar vedno novih možnosti. V nasprotju z Grki, ki čas razumejo kot postopno oddaljevanje od izvorov in zato vedno večjo degradacijo, razumejo Hebrejci čas kot približevanje eshatološkemu trenutku univerzalnega miru in sprave, »ko bo volk prebival z jagnjetom in bo panter ležal s kozličem« (Iz 11,8). Čas je sredstvo, po katerem Bog vodi zgodovino vedno bližje k sebi, dokler ne bodo sile teme in sovraštva poražene in se bosta odprla »novo nebo in nova zemlja« (Raz 21,1).« (Petkovšek 2009, 18)

Robert Petkovšek je pokazal na pomembno razliko pri pojmovanju odrešenja med grško in izraelsko filozofijo. Helenistične šole so svojim učencem ponujale preprosta življenjska navodila, kakor npr.: »Carpe diem!«, ki so jih lahko uporabljali v konkretnih situacijah. Profesor želi poudariti, da so ljudje v antiki filozofijo razumeli kot dušno terapijo, posvečeno življenju po vzoru najvišjega Dobrega. Za preoblikovanje življenja po vzoru najvišjega Dobrega pa je potrebno globoko premišljevanje temeljnih resnic in obnavljanje zavesti o končnosti življenja. Terapevtsko vlogo filozofije nadgrajuje odrešenjska, ki jo dobro razloži Platonova prispodoba o votlini, ki predstavlja človekovo vsakdanjost, ki je neosebna, nepristna in

anonimna. V nadaljevanju Petkovšek razloži, v kakšnem smislu filozofija odrešuje, in pokaže na antičnega človeka, ki dela, govori, misli in se obnaša neosebno, saj počne, kar počnejo drugi, ker prevzema govorne, vedenjske in miselne vzorce drugih. Toda med njimi je nekdo, ki hrepeni po resnici, po svetlobi in premaga temo nezanimanja ter se reši iz votline in tako spozna resnico. Spoznanja ne more zadržati zase, ampak ga kot odrešenjsko sporočilo nese nazaj v votlino, kjer ga ne sprejmejo in mu ne sledijo, temveč ga umorijo. Petkovšek na tem mestu zaključi, da je poslanstvo filozofije odrešenjsko, ker filozofija odrešuje:

»Najbolj je odrešenjsko vlogo filozofije dodelal Plotin, ki prav tako kot Platon uporablja za temeljno orodje filozofije razum oziroma um. Sredi vsakdanjega življenja, ki je polno teme in neresnice, mora razum najti pot, ki vodi k izviru vsega, k Enemu in Dobremu. V Enem in Dobrem je navzoča polnost bivanja in z njo največja sreča. Vendar um te svoje poti ne zmore prehoditi sam, zato je v zadnjem delu deležen mističnega združenja z Enim. Združenje je delo Enega samega, je njegov dar, filozof se lahko nanj pripravlja z miselno askezo. V trenutku mističnega združenja filozof popolnoma odmre samemu sebi in doseže najvišjo blaženost, izstopi iz sebe in podari samega sebe, pride do čistega predrugačenja, božansko v človeku se vrne nazaj k Božanskemu. Razum, ki je na delu v filozofiji, človeka osvobaja in odrešuje, vendar zadnjega koraka k popolni blaženosti ne zmore sam, ta mu je podarjen.« (18)

Grški racionalizem je prežet z verovanjem v Dobro, h kateremu teži razum in ki tudi razumu ‚pomaga‘ pri odreševanju človeka. Evropsko razsvetljenstvo pa je z Descartesovim »Cogito, ergo sum« kakršnokoli presežnost odrinilo kot nepotrebno. Družbeno gibanje je v 18. stoletju v zahodnoevropskih državah zagovarjalo vlogo razuma na vseh področjih življenja, tudi v religiji. V katoliški teologiji so se v stoletju pred Prvim vatikanskim koncilom fideistom zoperstavili racionalisti, ki so zagovarjali vlogo razuma pri prepoznavanju Boga in Božjega razodetja. Razsvetljenci so si kritični razum predstavljali kot širjenje svetlobe v temi, oz. v prenesenem pomenu, je razum človeku povedal, kaj je zanj dobro in prav. Zato je bila vloga Cerkve v družbi močno zreducirana. Lahko bi govorili o znanosti kot religiji napredka, ki s svojim sistemom obnašanj zamenja obstoječe religije, katere v sekularizirani družbi izgubljajo svojo moč in vpliv. Tako imenovana sekularistična soteriologija prinese idejo, da se je človek sposoben sam odrešiti. Razum namreč človeka osvobaja različnih oblik suženjstva in ga lahko popolnoma osvobodi. Takšna miselnost se je razcepila v dve smeri: totalitarni sistemi in skrajni individualizem. Obema smerema je skupno prepričanje, da je človek v svoji lastni moči zmožen odrešiti samega sebe po politični poti. Glavna predstavnika razsvetljenske miselnosti sta Hegel in Marx. Prvi razvije miselni sistem za samo-odrešenje, drugi pa ga uresničuje v praksi (Marx in Engels, 201). Marx je prepričan v človekovo sposobnost preoblikovanja družbe v tako, ki ne bo več poznala zatiranja in izkoriščanja. Vlogo odrešenika je pripisal delavskemu razredu, ki naj bi človeštvo popeljal v novi, boljši svet. Komunizem je ideologija, ki ima zelo močne odrešenjske ideje. Tudi najbolj kruti politični režimi,

kot sta bila fašizem in nacizem, so pridobili na moči tudi zaradi ideoloških idej, ki so obljubljale boljše prihodnost (Kerševan 1975).

Mnogi razsvetljenski filozofi so filozofijo opredelili kot nauk o odrešenju brez Boga. Luc Ferry, francoski filozof, ki se sklicuje na razsvetljensko izročilo, meni, da ima v razsvetljenskem kontekstu filozofija bolj kot kdajkoli odrešenjsko vlogo. Filozofija v tem primeru nadomesti religijo in ker ni več tradicionalnih ‚malikov‘, kakor so bili Bog, resnica, narava, lahko filozofija človeka končno povzdigne v boga in hkrati počloveči boga. Mnogi misleci tega obdobja so bili prepričani, da izginitve tradicionalnih vrednot lahko omogoči večjo človečnost človeka. Ko se človek žrtvuje za bližnjega, se zanj ne žrtvuje zaradi neke višje ideje ali zato ker bi mu to zapovedal Bog, ampak zaradi človeka kot človeka, v katerem prepozna nekaj absolutnega. V bistvu razsvetljenska filozofija odgovor na človekovo nemoč išče v moči znanosti, tehnike in napredka, ter predvidi odrešenje v zgodovini. Takšnim zgodovinskim, totranskim mesijanizmom je krščanstvo vedno nasprotovalo in poudarjalo, da odrešenjska moč izhaja iz Boga (Petkovšek 2009, 21).

Tudi teozofija⁴ in antropozofija⁵ uči, da se človek lahko odreši sam. New age prav tako zagovarja samo-odrešenje, ki ga interpretira v kontekstu stavka »pomagaj si sam in Bog ti bo pomagal«. Nova doba je čas, ko naj bi se človek znova začel povezovati s koreninami svoje biti in se ne več zanašal na inštitucije, duhovnike, šamane itd. Sam igra aktivno vlogo pri odrešenju v smislu, da naredi vse, da postane ‚boljši človek‘, ali v kontekstu novodobne interpretacije, da poskrbi za širitev in razvoj zavesti. Po načelih mikrokozmosa-makrokozmosa je človek primoran storiti vse v svoji moči za samo-odrešenje, vendar še vedno v skladu s kozmičnimi zakoni, katerim je podrejen.

Podobno lahko govorimo o samo-odrešenju v kontekstu moderne ekonomije in znanosti, pri katerih človek zgolj s svojim razumom in lastnim prizadevanjem hitro gospodarsko in tehnološko napreduje. Zahodni človek je dosegel vrhunec materialnega blagostanja naše civilizacije. S pomočjo globalizacije pa se to blagostanje širi na vse celine sveta in pri ljudeh vzbuja hrepenenje po materialnem odrešenju. Vedno bolj smo ljudje usmerjeni v materialni svet in odklanjamo duhovno plat življenja. V materialnem svetu je merilo kvantiteta, prizadevamo si, da bi čim več imeli. Tu lahko govorimo o kultu ekonomije. Pozabljamo pa, da je človek primarno bitje odnosov. (Ocvirk 2009, 27–43)

4.2 Samo-odrešenje je pot v smrt

Krščanski nauk uči, da je za kakovostno življenje odpoved nujno potrebna. V evangelijskih namreč beremo, da polnost življenja najde zgolj tisti, ki se je pripravil

⁴ »Teozofija (gr. θεός, ‚bog‘, σοφία, ‚modrost‘): Religijsko-filozofski sistem, ki zagovarja, da je mogoče priti na podlagi neposrednih mističnih intuicij in fil. raziskovanja do posebne vednosti o bogu in svetu.« (SRL, 1243)

⁵ »Antropozofija (gr. άνθρωπος, ‚človek‘): Gibanje, ki v sodobni različici obnavlja stare religijske, gnostične in okultne tradicije, pomešane s krščanstvom, da bi doseglo polno duhovno modrost o človeku. Po Steinerju človek doseže cilj na duhovni poti z meditiranjem in notranjimi vizijami. Tedaj vzpostavi stik z nadčutnimi bitji, prodre v razvoj komosa in vseh njegovih sedem stopenj ter dosega čedalje večjo popolnost in jasnovidnost.« (SRL, 57)

marsičemu odpovedati: »Kdor namreč hoče svoje življenje rešiti, ga bo izgubil; kdor pa izgubi svoje življenje zaradi mene, ga bo našel.« (Mt 16,25) Odpoved je temelj življenja v polnosti. To beremo pri vseh treh sinoptikih. Matej je zapisal: »Tedad je Jezus rekel svojim učencem: Če hoče kdo iti za menoj, naj se odpove sebi in vzame svoj križ ter hodi za menoj.« (16,24) Prav tako beremo pri Marku v 34. vrstici osmega poglavja in pri Luku v 23. vrstici devetega poglavja. »Odpoved sebi« pomeni sprejem svojih omejitev in se odpovedati marsičemu, kar bi sicer človek lahko naredil. Pravzaprav je odpoved kot taka nujno potrebna za osmišljanje vsakega človekovega življenja. Zato so Božje zapovedi, ki zahtevajo od človeka odpoved, merilo oz. pogoj za življenje že v Peti Mojzesovi knjigi: »Glej, danes sem položil predte življenje in srečo, smrt in nesrečo, ko sem ti danes zapovedal, da ljubi Gospoda, svojega Boga, da hodi po njegovih potih in izpolnjuj njegove zapovedi, zakone in odloke /.../ Če pa se tvoje srce obrne in ne boš poslušal in se boš dal zapeljati, da bi se priklanjal drugim bogovom in jim služil, vam danes naznanjam, da boste gotovo uničeni /.../ (5 Mz 30,15-20). Ker je novodobni človek, kakor Adam, zavrnil Božje zapovedi, zavrnil odpoved, se je s tem dejanjem odločil za smrt. Razlog za pogubljenje pa nikakor ni iznajdba in uporaba moderne tehnologije, temveč prepričanje, da človek ne potrebuje Boga oz. da lahko sam doseže polnost življenja. Od iznajdbe parnega stroja do iznajdbe umetne inteligence se namreč človek, prepričan v svoje sposobnosti samo-odrešenja, oddaljuje od Boga, se skriva pred njim in odriva od sebe vse, kar je povezano z Bogom, ter se »postavlja na svoje lastne noge«! Toda, prepričan, da mu ne bo treba več jesti kruha »v potu svojega obraza« (1 Mz 3,19), se je zavedel nevarnega položaja, v katerem se je zaradi svojih dejanj znašel, in ga je strah.

5. *Errare humanum est*

Ko ga Bog išče, Adam odgovori: »Slišal sem tvoj glas v vrtu, pa sem se zbal, ker sem nag, in sem se skril.« (1 Mz 3,10) Tudi človek danes ugotavlja, da ni gospodar položaja. Ko smo zgoraj napisali, da se je moderen človek »odločil za smrt«, nismo mislili le poetično. Človeštvu res grozi izginotje. Tehnološki razvoj je pripeljal gospodarstvo do hiperprodukcije, ki jo zahteva družba uživanja. Posledično je naš planet zelo onesnažen. Na mestu je vprašanje, če bo »zemeljski vrt« še lahko dal hrano človeštvu. Vremenske spremembe nas vse opozarjajo, da je položaj na našem planetu kritičen; orkani, poplave ali suša vse pogostejše uničujejo človekov življenjski prostor. Uživaški način življenja močno vpliva tudi na demografijo. V vseh razvitih državah se nataliteta drastično zmanjšuje, vse več pa je populacije, ki ni več delovno produktivna; ideologije spola uničujejo družino. Človek se je po »svoji pameti« naučil preprečevati rojstva, pobijati nerojene otroke in onemogle ter bolne do te mere, da so se začeli krčiti narodi⁶, in veliko vprašanje je, če ne bo človeštvo izginilo, ker ljudje ne bodo več rojevali otrok. Demografske napovedi

⁶ V vseh državah, ki so nastale na ozemlju bivše Jugoslavije, se razen Slovenije in Črne gore število prebivalcev drastično zmanjšuje zaradi prenizke natalitete in izseljevanja.

kažejo, da se bo krivulja, ki prikazuje število ljudi na svetu, kmalu močno obrnila navzdol in da bodo mali narodi izginili. (Gams 2024) Razvil je orožja, ki lahko naš planet spremenijo v prah. Z namenom, da bi na najlažji način, s čim manj napora preživel, človek izumi umetno inteligenco, ki bi delala namesto njega. Še predno pa jo je dokončno razvil, ga je že strah pred njo. Papež Frančišek v „Poslanici za 57. svetovni dan miru“ navaja razloge za človekov strah. Najprej omeni možnost širjenja lažnih informacij in zlorabe podatkov. Nato pa se razpiše o tehnološki diktaturi: »Algoritmom ne smemo dovoliti, da bi odločali, kako razumemo človekove pravice in da bi pustili ob strani temeljne vrednote sočutja.« (Frančišek 2023, 5) Umetna inteligenca namreč (še) ni sposobna sočutja in celo obstaja strah, da se bo osvobodila in se obrnila proti človeku. Predvsem pa je nevarnost, da najnovejša tehnologija pride v roke kriminalnim in terorističnim družbam.

V bistvu je umetna inteligenca samo vrh ledene gore procesa, ki se je začel s človekovim osvobajanjem izpod Božje superiornosti. Proces osvobajanja izpod Božje superiornosti je neetičen, ker civilizacijo vodi v izginotje, v smrt, v propad. Liberalistično pojmovanje svobode lahko primerjamo z vlakom, ki bi se osvobodil železniških tirov, ki mu omogočajo vožnjo, in zato vlak ne more drugače, kakor da se prevrne! Človek se želi osvoboditi Božje superiornosti, ker mu je v breme. A ko se znebi bremena, si izpodmakne tudi oporo, ker so Božje zapovedi za človeka breme in opora hkrati.

Motiti se je človeško. Ko povprečen človek razmišlja o tem latinskem pregovoru, se z njim vsekakor strinja. Hkrati pa si človekove pomanjkljivosti predstavlja kot neko „napako v sistemu“ oz., če se izrazimo v kontekstu biblične teologije, kot kazni za greh. Ker se je (prvi) človek uprl Bogu, je za kazni postal pomanjkljiv, ga je doletelo trpljenje, bolezn, nepopolnost pri delu in tudi smrt. Pa naj bo to kazni v smislu, da je Bog spremenil svojo (do tedaj) popolno stvaritev, ali pa, da je človek z grehom sam spremenil svojo človeško, do tedaj popolno, naravo v nepopolno. Kakorkoli že, zaradi nepopolnosti človek potrebuje pomoč drugega. Sam se ne more (od)rešiti. Prepričani smo, da ga lahko odreši le spolnjevanje Božjih zapovedi.

6. Sklep

Umetna inteligenca prodira na vsa področja življenja. Ko se je avtor tega članka pred mesecem dni peljal z RER železnico v Parizu, so potniki dobili obvestilo: »Na naslednji postaji bo na vlak vstopil žepar. Prosimo, da ste pozorni na svojo lastnino!« Med milijoni potnikov, ki se gnete na postajališčih, je umetna inteligenca sposobna prepoznati obraz človeka, ki je bil že večkrat obsojen zaradi žeparstva na javnih prevoznih sredstvih ... in potnike posvariti! Prispevke za demografsko konferenco na Inštitutu Jožef Stefan v Ljubljani od letos recenzira računalnik. Celo nekateri duhovniki se poslužujejo umetne inteligence, da jim napiše pridigo, ki jo potem preberejo pri maši v cerkvi. Res neverjetne so sposobnosti te nove tehnološke pridobitve. Upravičeno se človek lahko vpraša, kje so meje samo-izpopolnjevanja

računalnika in kje se bo človeku ‚zalomilo‘, da mu bo žal, da je sestavil računalnik, ki se lahko sam izpopolnjuje in ki bo morda dobil svojo lastno zavest. Vprašanje o etičnosti umetne inteligence je vsekakor na mestu, zato je tudi napetost med progresisti in zaviralci razvoja potrebna v smislu, ki ga bomo sedaj razložili.

Tehnologija avtomobila in vsakega vozečega vozila je zasnovana tako, da vsebuje dva najpomembnejša elementa: motor in zavore. Oboje je nujno potrebno in zelo smiselno. Če bi vozilo imelo samo pogon, bi zagotovo prišlo do nesreče. Če bi imelo samo zavore, se ne bi premaknilo z mesta. Tako sta tudi v družbi potrebni dve sili: progresizem, ki teži k novostim, torej k boljšemu, hitrejšemu in močnejšemu, ter tradicionalizem, ki upočasnjuje razvoj, dokler se ne razjasni, ali je inovacija dobra za človeka. Vendar pa morajo biti predstavniki teh dveh usmerjenosti povezani s konstruktivnim dialogom: »Pogovor je po mnenju psihologov, psihoterapevtov, teologov in filozofov za človekovo psihično stanje najboljše zdravilo. Pri tem pa ni pomembno samo, da se pogovarjamo, ampak tudi, kako se pogovarjamo – in kakšen je naš odnos do samega sebe in sogovornika.« (Jeglič 2022, 731)

Uporaba umetne inteligence je vsekakor lahko pogosto sporna. Predvsem dokler z njo še upravlja človek, ki je postal ne-etičen. Kaj pa, če bo stroj, ki se bo zavedel samega sebe kot osebe, začel verovati in spolnjevati Božjo voljo?

Kratiki

GS – Koncilski odloki 1980 [Gaudium et spes].

SRL – Bajt in Kocjan-Barle 2007 [*Splošni religijski leksikon*].

Reference

- Bajt, Drago, in Marta Kocjan-Barle**, ur. 2007. *Splošni religijski leksikon*. Ljubljana: Mordrijan
- Deklaracija človekovih pravic**. 2018. *Uradni list RS*, št.24/2018 z dne 13. 4. 2018.
- Frančišek**. 2023. Umetna inteligenca in mir. Poslanica svetega očeta Frančiška za 57. svetovni dan miru. Ljubljana: Družina.
- Gams, Matjaž**. 2024. Pregled planetarnih in lokalnih demografskih projekcij. *Information Society* 2024:10–17.
- Jeglič, Urška**. 2022. Reševanje kriznih odnosov v pedagoškem poklicu s pomočjo supervizije. *Bogoslovni vestnik* 82, št. 3:725–735. <https://doi.org/10.34291/bv2022/03/jeglic>
- Kerševan, Marko**. 1975. *Religija kot družbeni pojav*. Ljubljana: Mladinska knjiga.
- Koncilski odloki**. 1980. Ljubljana: Nadškofijski ordinariat v Ljubljani.
- Marx, Karl, in Friedrich Engels**. 1977. *Izbrana dela*. Zv. 1, *Prispevek h kritiki Heglove pravne filozofije*. Ljubljana: Cankarjeva založba.
- Ocvirk, Drago**. 2009. Zunaj Cerkve veliko odrešenja. V: *Odrešenje in sprava, čemu?*, 27–43. Celovec: Mohorjeva družba.
- Osredkar, Mari Jože**. 2012. *Pojem odrešenja in nauki o odrešenju v svetovnih religijah*. Ljubljana: Teološka fakulteta.
- Petkovšek, Robert**. 2009. Odrešenje v luči filozofije. V: *Odrešenje in sprava, čemu?*, 9–25. Celovec: Mohorjeva družba.
- Trontelj, Nik**. 2022. Ločenost Cerkve in države v Evropi: zgodovina in sedanjí trenutek. *Edinost in dialog* 77, št. 2:15–27. <https://doi.org/10.34291/edinost/77/02/trontelj>

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 835—846

Besedilo prejeto/Received:08/2024; sprejeto/Accepted:12/2024

UDK/UDC: 128.5:27-1:004.89

DOI: 10.34291/BV2024/04/Platovnjak

© 2024 Platovnjak et al., CC BY 4.0

Ivan Platovnjak and Tone Svetelj

Artificial Intelligence and *Imago Dei*: A New Dilemma for Philosophical and Theological Anthropology

Umetna inteligenca in Imago Dei: Nova dilema za filozofsko in teološko antropologijo

Abstract: The advancement of Artificial Intelligence (AI), with its vast applicability, has already influenced nearly every aspect of our lives, including our self-perception. At the same time, AI calls for wisdom, hope, and belief that transcends sophisticated knowledge produced by AI. The main argument in this article is an inevitable call to rediscover the essence of human nature that AI technology cannot adequately grasp. From a Christian perspective, AI challenges us to rediscover the idea that humans are created as an *imago Dei*, an image of God, where relations are an essential part of God's nature and, consequently, of human nature as well.

Keywords: Artificial Intelligence, anthropology, human nature, an image of God, spirituality

Povzetek: Razvoj umetne inteligence (UI) je s svojo široko uporabnostjo že vplival na skoraj vse vidike življenja, vključno z našim dožemanjem samega sebe. Hkrati UI kliče po modrosti, upanju in veri, ki presegajo sofisticirano znanje, ki ga ustvarja UI. Glavni argument v tem članku je neizogiben poziv k ponovnemu odkrivanju bistva človeške narave, ki ga tehnologija UI ne more ustrezno zajeti. S krščanske perspektive nas UI izziva, da ponovno odkrijemo idejo, da je človek ustvarjen kot *imago Dei*, Božja podoba, kjer so odnosi bistveni del Božje narave in posledično tudi človeške.

Ključne besede: umetna inteligenca, antropologija, človeška narava, Božja podoba, duhovnost

1. Introduction

Our contemporary cultural landscape is in many ways reminiscent of Socrates' dialoguing with the sophists.¹ The Oracle of Delphi deemed Socrates the wisest

¹ The article was written as part of the research program "P6-0269 Religion, Ethics, Education, and Challenges of Modern Society" supported by the Slovenian Research and Innovation Agency.

man for his recognition of his own ignorance. "I do not think I know what I do not know." (Plato 2002, 26) This was not the case for the sophists, who thought they knew something, whereas, in reality, they did not. In Socrates' context, the sophists were considerable celebrities, active in public affairs, who proved themselves as professional educators offering instructions in many subjects, especially in public speaking and the successful conduct of life. While the sophists believed they possessed knowledge, Socrates confessed he was in search of and in love with wisdom.

The new sophists of our age are not human philosophers but advocates of Artificial Intelligence (AI) systems and machines, which can apparently provide us limitless access to information and instructions at every step of our existence. No doubt, AI advancements rapidly and profoundly influence our way of thinking, acting, and believing, both on the personal and social levels, and trigger profound unease about the future.

The reality is that there is no definitive answer to what our future with AI will look like. This struggle underscores the need for wisdom, hope, and belief that will transcend sophisticated knowledge produced by AI; it calls for a deeper and more nuanced understanding that incorporates the essence of human nature, including its spiritual dimension. Thus, AI technological advancement invites us to rediscover human nature, the purpose of life, and the meaning of human flourishing. Above all, AI challenges us to move from a life of isolation and self-centeredness to a life based on genuine relations and connections with myself, others, nature, and the divine. From a Christian perspective, it challenges us to rediscover the idea that humans are created as an *imago Dei*, an image of God, where relations are an essential part of God's nature (Pevc Rozman 2021, 840).

2. Within an AI Cultural Framework

The 21st century is not the first time that philosophers and theologians have been challenged to deepen and rethink their anthropological statements. Copernicus's heliocentric model challenged the geocentric worldview, as well as the place of humanity and the divine in the universe (Westman 2011, 1–17). Similarly, Darwin's theory of evolution through natural selection questioned the static, divine creation of species, proposing instead a dynamic, interconnected biological history. These two and many other paradigms shift initially faced resistance, as they seemed to undermine established theological and philosophical principles. This is the point where it is crucial to distinguish between the personality of the scientist and the personality of the ideologist hiding behind the scientist. "Some scientists have a gift for masking their unscientific ideology with a scientific façade." (Verschuuren 2019, xiii) Not paying enough attention to this distinction, one can easily accept and confuse the facts based on science with the scientist's worldview. Thus, it is reasonable to question the worldview of scientists because they are not scientific theories that can be tested in science. Their worldviews are beyond scientific methodology. Without paying attention to this distinction, one can uncritically

accept the claim that the Catholic Church holds an anti-science position or that AI is entirely against the spiritual aspect of human existence.

How can these two examples, taken from the Western history of science, help us confront the overarching presence of AI in our time? There is no doubt that AI already influences nearly every aspect of human life, including work, interpersonal relationships, education, mental health, social norms, cultural practices, and beliefs, and leaving us uncertain about our future. With just a few clicks, AI can deliver desired answers within seconds. These answers can be intricate and multifaceted, drawing on vast amounts of information the human brain cannot feasibly process within a reasonable time frame. Historically, access to information was a privilege afforded by education; in contemporary times, the combination of internet connectivity and AI provides an overwhelming volume of information at every turn, surpassing the brain's capacity for adequate comprehension and evaluation.

The opportunities for the application of AI appear almost limitless. Current trends in mental healthcare reveal AI's transformative potential, such as the early detection of mental health disorders, personalized treatment plans, and AI-driven virtual therapists (Olawade 2024, 1). The same AI technology challenges the boundary between life and death. AI systems like Thanabots or Deadbots recreate deceased individuals and allow interactions with them. Even though this might be very beneficial in the initial process of grieving, the AI application transforms our perception of death, alters the experience of the finitude of human life, and creates the possibility of a 'postmortal society' where biological death is no longer an inevitable end. Thus, AI acts as a mediator between life, death, and the human being, allowing a new form of communication with the deceased (Reséndiz 2024). These two examples, with many other AI applications, bring forth many ethical, philosophical, and theological challenges that transcend our present reflections on human existence. Questions of data privacy, ethical and emotional questions about its authenticity, purpose of recreation, AI autonomous decision-making process, as well as ethical vulnerability, transparency, and the democratization of these technologies, challenge fundamental notions of human existence.

Knowing that AI can elaborate only on data that already exists in digital form, the question arises about data that has not yet been digitalized. In addition, what are the rights and duties of the owners of this data (Amazon, Apple, Microsoft, Facebook)? The increasing implementation of AI technologies is building so-called data capitalism, giving big companies more power that can be used for the augmentation of their initial capital. At the same time, the everyday person, minorities, the poor and excluded are left out. Like capitalism, data capitalism is rooted in oppression and reinforces the dynamics of power and profits.

Shall we exclude cultural diversities and peculiarities, popular wisdom, and unspoken values that shape societies and individuals? What about entire nations and minority groups who cannot actively participate or are excluded from active participation in the creation of digital data? Hamilton Mann claims that AI presents a new homogenization of the world based on sophisticated algorithms de-

signed to optimize efficiency and certain preferences. “While AI has the potential to enrich and diversify our cultural landscape, it is also steering us toward a more monolithic and less diverse world.” (Mann 2024)

AI homogenization creates new forms of cultural discrimination, reinforces biases (Hercheui 2022), and perpetuates societal prejudices, leaving out the less popular voices and forms of expression, minorities, and especially those who have nothing to say in writing the new AI algorithms. So-called minority data representation is becoming a significant concern in the field of AI, impoverishing our cultural and intellectual landscapes, diminishing the potential for breakthrough ideas, and creating unique and unconventional perspectives way beyond the popular average. Dealing with the average reduces the probability of an exceptional and groundbreaking outcome. “The integrity of the nuanced and often unpredictable nature of human creativity is at stake.” (Hamilton 2024) If this is the case, AI should make us realize our own importance as humans, not so much as individuals, but much more as social beings whose essence grows and reaches its fulfilment through connections and relations.

3. Disconnection from the Body, Earth, Others, God, and the Decision-Making Process

While acknowledging the numerous advantages and benefits that AI brings to our lives, it simultaneously raises profound philosophical concerns. Are we grappling with questions rooted in Cartesian dualism? René Descartes (1596–1650) posited that human reality comprises two fundamentally distinct substances: the non-physical mind, characterized by thought and consciousness, and the physical body, characterized by spatial extension and mechanical properties. Descartes’ fascination with the capabilities of the human mind led him to struggle with appropriately situating the human body. Thus, his famous statement: “/.../ nothing else belongs to my nature or essence except that I am a thinking thing. /.../ and not an extended thing /.../ [since] I have a distinct idea of a body, insofar as it is merely an extended thing and not a thinking thing, it is certain that I am really distinct from my body, and I can exist without it.” (Descartes 1993, 51)

Descartes, often regarded as the father of Modernity, significantly influenced Immanuel Kant (1724–1804) in his defining Enlightenment as “man’s emergence from his self-imposed immaturity. Immaturity is the inability to use one’s understanding without guidance from another. This immaturity is self-imposed when its cause lies not in lack of understanding, but in lack of resolve and courage to use it without guidance from another.” (Kant 1983, 41) Kant’s definition reflects Descartes’ emphasis on human reason as the fundamental aspect of human nature, dispelling the darkness of unenlightened ages and empowering humans to act maturely by following the universal categorical imperatives.

The last two centuries have witnessed remarkable scientific discoveries, technological advancements, and two World Wars, including atomic bombs. Advanced

technology has opened new dimensions of the universe, allowing exploration not solely from Earth but from the universe itself. In *The Human Condition*, Hannah Arendt describes the mixed feelings of triumph and relief when the first man-made object was launched into the universe in 1957. People felt relieved at taking the first “step toward escape from men’s imprisonment to the earth” (Arendt 2018, 1). So, technology, as a product of human reasoning and creativity, promises that humanity may not remain bound to Earth indefinitely.

As we struggle to exist within our material bodies and seek to transcend earthly boundaries, AI now presents us with a new challenge: relinquishing reasoning, a unique feature of human nature, to algorithms embedded in cyborgs or other machines responsible for making life decisions. Overwhelmed by Kant’s enlightened maturity and the secularization of the modern age, which diverts us from religion and God, AI seems an inevitable next step—potentially leading us to abandon our capacity and responsibility for decision-making, particularly moral choices, thereby undermining a pivotal aspect of our humanity (Horvat, Roszak and Taylor 2022).

These disconnection forms should not be taken as something organic and natural. In his reflections, Joseph Wilson argues that “AI /.../ is not a technical thing, a scientific thing. It’s not even a scientific concept, because nobody can really define it. It keeps changing over time. It’s a cultural thing, a social concept. It’s a cultural fact, not a scientific fact.” (Wilson 2024). Wilson further explains that some scientists immersed in this AI culture often forget that the computer functioning as a brain is merely a metaphor, not reality. AI culture tries to persuade people that they are conversing with thinking entities when interacting with AI devices. The more these devices resemble humans, the more convincingly we attribute human-like qualities to them. This phenomenon, known as pareidolia, is the tendency to perceive specific, meaningful images in random patterns, akin to the Rorschach inkblot test. For instance, we often perceive shapes, objects, or scenery in clouds or rocks. Following this logic, AI culture encourages us to perceive AI devices as ‘humans’, attributing agency, emotions, and the capacity to think, even though they are not technically alive. Wilson contends that the ingrained notion of a Cartesian split between mind and body allows some transhumanists to envision uploading their minds onto computers while maintaining their essential sentience; they consider human intelligence as merely information processing that can be separated from the body. “That is a cultural quirk of the West,” Wilson makes his point (Wilson 2024).

4. AI and Humans from a Transcendental Perspective

In contemporary times, two parallel dynamics are observable: on the one hand, there is a movement away from the body and the earth, which implicitly encompasses a movement away from other people; on the other hand, there is an endeavour to render the human environment and human life artificial by severing the ties that connect humanity to nature (Kraner 2024). Hannah Arendt reflects that the same desire to escape earthly imprisonment also inspires our attempts to create life in a

test tube, aiming at the perfect human. This desire reflects an aspiration to transcend the human condition, a rebellion against the given human existence, and a pursuit to replace this created world with something artificial, non-organic, and man-made. Do we genuinely wish to employ new scientific and technical knowledge in this direction? This question cannot be answered by science alone or relegated to professional scientists (Arendt 2018, 2–3). It is a fundamental question that necessitates the inclusion of political, ethical, philosophical, and theological considerations regarding our future.

Setting aside the challenging aspects, the current AI culture presents a unique opportunity to redefine the essence of what it means to be human. As intelligent beings endowed with bodies, cells, and emotions—including fears, loneliness, pain, suffering, and death—humans possess the capacity to connect with entities and phenomena that transcend mere analytical perception. Intelligence, for instance, can encompass irrational activities such as shamanism, intuition, foresight, hope, prophecy, and beliefs—domains that extend beyond pure information processing, which characterizes AI devices. Additionally, human emotions like love, sympathy, empathy, fear, anxiety, and the unceasing quest for meaning in life elude the capabilities of AI information processing alone. Human uniqueness, self-reflection, dignity, freedom and free will, personal relationships, and mysteries of human life and death have intrigued human minds and spirits from the dawn of human existence. It is, therefore, impossible, if not irresponsible, to reduce all these aspects of human existence to mere data analysis. Let us not forget that the technical man in awe with the AI technology remains a spiritual man.

From the perspective of Christian philosophy and theology, AI compels us to reevaluate a human's position in front of a transcendental God, a role seemingly supplanted by man-made AI. This fundamental anthropological question, as the cornerstone of Christian traditions, urgently requires rediscovery. The concept of *imago Dei*—that humans are created in the image of God—is a foundational source for deeper reflection.

5. In What Way Can AI Not Be a Better *Imago Dei* than a Human Being?

The concept of human beings as *imago Dei* (image of God) (Genesis 1:26-27; 5:1-3; 9:6) is central to the biblical revelation of human nature and to biblical anthropology, which seeks to define what it means to be human. As the International Theological Commission (2004, 7) states, “the mystery of man cannot be grasped apart from the mystery of God”.

In recent decades, the notion that AI could reflect or represent a kind of *imago Dei* has emerged as a provocative idea, opening up new dialogues between philosophy, theology and digital technology. Scholars such as Foerst (1998), Jackelén (2002) and Herzfeld (2002a; 2002b) have explored this concept extensively. However, it is important to further explore why AI cannot embody the *imago Dei* more fully than humans.

5.1 Human Intellectual Capacity

Influenced by Aristotelian philosophy, many interpreters of Genesis (1:26-27) have identified the rational and cognitive abilities of human beings as the key aspects that reveal their likeness to God. In the past, humans were considered the most rational creatures in the world (Pevce Rozman 2024, 7). However, today's AI systems - from simple algorithms to sophisticated machine learning and neural networks—often surpass these human abilities (Oeming 2022).

5.2 Human Creativity

The uniqueness of human beings and their likeness to God is particularly evident in human creativity. The Second Vatican Council (*Gaudium et Spes* 34) emphasizes that human action, especially creative action, reflects God's creativity, which is both its origin and its model. While AI is also capable of generating creative outputs in various fields such as music, art, literature and technology (Cheng 2022), it works primarily on the basis of pre-existing patterns and data. As a result, the 'creativity' of AI is often limited by the information and algorithms it relies on. Despite its technical sophistication, AI's inability to experience emotions often means that it lacks the depth that characterizes human creativity (Farina et al. 2024). The true value of AI to the creative process is not to replace human creativity, but to complement and augment it. Properly understanding and fostering this synergy will be critical to the advancement of art, technology, innovation, science and even theology.

5.3 Human Autonomy and Independence

The God-likeness and uniqueness of human beings is also reflected in their freedom, especially in their autonomy and independence. Human beings possess free will, enabling them to make morally and ethically complex decisions. Furthermore, their independence allows them to go beyond mere reactions to external stimuli. This autonomy is the basis for human moral responsibility, creativity and innovation (Lah 2003, 268–272).

While advanced autonomous AI systems exhibit a degree of 'freedom' by operating without direct human control, this is distinctly different from human moral autonomy. Some researchers are working on ethical algorithms to enable robots to make moral decisions in certain situations. Although these algorithms are not equivalent to human moral autonomy, they reflect an attempt to embed basic values into AI behaviour (Etzioni and Etzioni 2016). Formosa (2021) emphasizes that the integration of AI into society must be carefully considered in order to preserve and respect human autonomy and ensure that it enhances, rather than impairs, the free agency and dignity of humans.

5.4 Human Management and Stewardship of the Visible World

As bearers of God's image, humans have a unique role in the universe: they are to participate in God's stewardship and management of the visible creation (Ang

2024). This role involves cooperation with and submission to God's ordering of the world. The International Theological Commission (2004, 61) emphasizes that humans fulfil this responsibility by "gaining scientific understanding of the universe, by caring responsibly for the natural world (including animals and the environment), and by guarding their own biological integrity".

The development of AI and its growing impact on various aspects of society raises important questions about its potential ability to govern and manage the world. Philosopher Nick Bostrom (2014) sees the realistic possibility of AI reaching a super human level of intelligence in the future, referred to as artificial superintelligence (ASI). Unlike humans, who are limited in their capacity for cognition due to the physical limitations of their brains and the inevitability of death, machines are not subject to these limitations. As a result, ASI could potentially become more intelligent than any individual human and even surpass the collective intelligence of humanity, ultimately surpassing human understanding.

Nick Bostrom (2014, 181) outlines three possible operative roles for ASI: as oracle, genie and sovereign. As an oracle, ASI would answer our questions; as a genie, it would carry out our orders; and as a sovereign, it would govern the world and be tasked with achieving broad and long-term goals. This sovereign role mirrors the role ascribed to God in monotheistic religions and raises important questions about human exceptionalism as *imago Dei*. If ASI could potentially govern the world more effectively than humans, it could be argued that ASI could represent the *imago Dei* to an even greater degree.

However, as Marius Dorobantu (2022, 186–188) argues, we must recognize that the human vocation to steward creation goes beyond mere temporal stewardship to include a spiritual dimension. Humans are not only called to steward creation in a temporal sense, but are entrusted with the higher task of leading creation to its ultimate spiritualization (Rom 8:18-23). Dorobantu also emphasizes that the computerization of matter made possible by AI is fundamentally different from the spiritualization described in Christian theology. While both information and spirit are immaterial, the Christian concept of spirit is much broader and includes the transcendent dimension of reality. Furthermore, Christian theology connects the spiritualization of matter to the loving relationship between God and human beings, which implies the ultimate transparency between the Creator and creation. Without the love of God, any attempt to spiritualize matter has no real content and meaning.

A theology of the *imago Dei* should therefore emphasize the spiritual dimension of our governance and stewardship and not only focus on the historical aspect in which AI might actually surpass us. AI, however advanced it may be, will never be able to replace humanity's mystical role in the world. Bearing the image of God is not just about acting as a representative of God in a particular historical moment; it involves an ongoing, authentic love relationship between humans and the Creator in the world. This understanding leads to a relational interpretation of the *imago Dei*, in which the image of God is deeply rooted in this divine-human relationship (Dorobantu 2022, 188).

5.5 Human Vulnerability and Our Relationship with God and Others

Human God-likeness is particularly evident in our relationships with others and with God (Stegu 2023).² The documents of the Second Vatican Council, in particular *Gaudium et Spes* (24), emphasize that every person is created as a unity of body and spirit in the image of God for personal communion with Him. The absolute openness of Jesus and his devotion to the Father show in a special way that the true identity of human beings arises from their ability to enter into a relationship with the Other (Ratzinger 1968, 140–144). A human person fully lives out his or her identity and his or her existence to the full in a historical relationship that is founded in the sonship of God and that represents the deepest fulfilment of their “being-in-itself” through “being-for-the-other” (Roszak 2024). This relational identity also enables the individual to overcome the conflicts between individuality and relationality. To the extent that a person increasingly discovers the sonship inherent in their being, their capacity for free choice and their recognition of their own personhood as a gift and calling also grows (Lah 2003, 270).

As beings created in the image of God, humans possess autonomy, i.e. they have their own stability, independence and ability to make decisions for themselves. At the same time, they have the ability to transcend themselves in relationships by freely giving their love to others and forming deep connections and a life together. Through their work and creativity, people contribute to society and help shape culture and transform the world. However, this journey can sometimes lead to a loss of personal identity or the risk of being controlled by external forces or dependencies. The way to guard oneself from these risks is to have a personal relationship with God that allows for true self-discovery and preserves personal integrity. Only when the individual realizes himself in the infinite, completely free and loving relationship with God can he become fully who he is meant to be (272).

If our relationships are the best reflection of God, we need to consider whether the development of AI could challenge this unique aspect of our humanity (Bilagher 2022). What might happen if machines become capable of forming personal relationships? We already converse with chatbots, and it is conceivable that in the future we might interact with machines as naturally as we do with other humans. However, as Dorobantu (2022, 189–190) notes, AI would approach these interactions in a distinctly nonhuman way. When AI makes mistakes, they are not the same mistakes that a human would make. Even if artificial beings were to develop consciousness and subjective experiences, these would differ from human experiences due to their different embodiment. Robots would have different sensory perceptions, different access to memory, different internal states and a different relationship to time. Their needs would also be different, which would affect their interests and motivations. Even if they reached the level of human capabilities, they would still be fundamentally dissimilar to humans. Therefore, when the theology of the imago Dei emphasizes human kinship, it refers to a likeness to God that robots can never fully realize.

² In other words: Man’s ‘Godlikeness’ is his ability to (be) dialogue (Skralovnik 2021, 124).

Similarly, our ability to enter into relationships is closely linked to our anthropological vulnerability. We form relationships because we are vulnerable and mortal, and because we inherently need one another (Globokar 2022, 8). An artificial being that is practically 'invulnerable' and theoretically 'immortal' cannot enter into relationships in the same way as humans because it lacks the capacity for voluntary vulnerability that is essential for deep and meaningful connections. Humans do not reflect God's image when they are powerful and unbreakable, but when they are loving and vulnerable (Dorobantu 2022, 191–192).

It is also unlikely that AI can build human-like relationships due to its hyper-rational nature. An artificial being that makes decisions based solely on cold calculations of the best possible outcome is unlikely to engage in the risky and irrational behaviours that characterize human relationships (Šegula 2021, 922–923). Humans seek relationships not because of perfection, but because of our imperfections and our deep desire for a completeness that we cannot achieve on our own. This drive for connection comes from a place deeper than our rational mind. A completely rational being would not behave in this way. Experiences such as falling in love, creating art and exploring spirituality —although seemingly irrational—add beauty and fulfilment to life. In this sense, the development of AI is a blessing for the theology of the *imago Dei* because it highlights and helps us appreciate the unique qualities that set us apart from machines. Compared to AI, we may seem irrational and outdated, but it is precisely our physical and cognitive limitations that allow us to form loving and authentic relationships, thereby revealing and developing our likeness to God. Our limitations are just as significant as our strengths (Dorobantu 2022, 192–193).

6. Conclusion

Rapid advancements in AI bring us to a new crossroads. AI's potential challenges not only our cognitive abilities but also our understanding of what it means to be human. Just as Copernicus and Darwin's theories compelled us to reevaluate our place in the universe, AI now pushes us to reconsider the essence of human nature, including our spiritual dimensions and our relationship with the divine. While AI can replicate and even surpass certain human intellectual and creative capacities, it cannot express the fullness of what it means to be created in the image of God. This image includes autonomy, creativity, emotional depth, and profound connections we create with ourselves, others, the earth, and the divine. These qualities cannot be reduced to mere data processing or algorithmic patterns.

In addition, the uniqueness of human beings lies in our capacity to love and be loved, self-awareness, compassion, and ability to deal with mysteries that cannot be quantified but only explored. Finally, human existence is so complex that it cannot be grasped with concepts; it must be experienced and lived. Socrates and other philosophers like him, were aware that what really matters is exploring, experiencing, being aware of boundaries, and following the most important principles. These activities are much more than simply knowing or having a certain conceptual understanding, typical of Sophists who believed to know but did not.

Moreover, humans' role as stewards of creation, as envisioned in the biblical concept of imago Dei, also includes a spiritual responsibility. While AI may achieve superintelligence, its role as a manager, governor, or oracle cannot replace the human vocation to elevate creation toward its ultimate spiritualization, i.e., communion with a loving God. This communion integrates all aspects of human existence, including suffering, pain, and death (Jerebic, Bošnjaković and Jerebic 2023, 361–362). Replacing a loving God with a human-made AI, even though superintelligent, leads us further away from our bodies, the earth, and genuine relationships towards disconnection, isolation, and self-centeredness.

AI should complement and extend our capabilities while keeping us connected to the core of our humanity. AI technology challenges us to rediscover and reaffirm the sacredness of human life, our connections with others, and our relationship with the divine. Prioritizing these connections ensures that AI advancements enrich human existence rather than diminish it. As advanced as technology can be, it cannot grasp the depth of our humanity, which reflects the image of God.

VII. Abbreviations

AI – Artificial intelligence.

ASI – Artificial superintelligence.

References

- Ang, Kenny. 2024. Creation in Christ in Aquinas's Creation Theology in the Light of His References to the Pauline Letters. *Biblica et Patristica Thoruniensia* 17, no. 1:9–21. <https://doi.org/10.12775/BPTH.2024.001>
- Arend, Hannah. 2018. *The Human Condition*. Chicago: The University of Chicago Press.
- Bilagher, Moritz. 2022. From Absolute Mind to Zombie: Is Artificial Intelligence Possible?. *Scientia et Fides* 10, no 1:155–176. <https://doi.org/10.12775/SetF.2022.008>
- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Cheng, Mingyong. 2022. The creativity of artificial intelligence in art. *Proceedings* 81, no. 1:110. <https://doi.org/10.3390/proceedings2022081110>
- Cropley, David H. 2023. Is AI More Creative Than Humans? Chatgpt and the Divergent Association Task. Preprint. <https://doi.org/10.31234/osf.io/jzt72>
- Descartes, René. 1993. *Meditations on First Philosophy*. Indianapolis: Hackett Publishing Company.
- Dorobantu, Marius. 2022. Imago Dei in the Age of Artificial Intelligence: Challenges and Opportunities for a Science-Engaged Theology. *Christian Perspectives on Science and Technology, New Series* 1:175–196. <https://doi.org/10.58913/KWUU3009>
- Etzioni, Oren., and Aamitai Etzioni. 2016. Designing AI systems that obey our values. *Communications of the ACM* 59, no. 9:29–31. <https://doi.org/10.1145/2955091>
- Farina, Mirko, et al. 2024. *Machine learning in human creativity: status and perspectives*. London: AI & Society.
- Foerst, Anne. 1998. Cog, a Humanoid Robot, and the Question of the Image of God. *Zygon*® 33, no. 1:91–111. <https://doi.org/10.1111/0591-2385.1291998129>
- Formosa, Paul. 2021. Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy. *Minds & Machines* 31:595–616. <https://doi.org/10.1007/s11023-021-09579-2>
- Globokar, Roman. 2022. Ranljivost človeka, družbe, narave in Boga [The Vulnerability of the Human Being, Society, Nature, and God]. *Bogoslovni vestnik* 82, no. 1:7–14. <https://doi.org/10.34291/BV2022/01/Globokar>
- Horvat, Sasa, Piotr Roszak and Brian Taylor. 2022. Is It Harmful? A Thomistic Perspective on Risk Science in Social Welfare. *Journal of Religion and Health* 61, 3302–3316. <https://doi.org/10.1007/s10943-021-01452-x>

- Hercheui, Magda, and Sam Guerard.** 2022. The Impact of Artificial Intelligence Biases: Algorithmic Discrimination. *Global Conference on Business and Finance Proceedings* 17, no. 1:27–31.
- Herzfeld, Noreen L.** 2002a. *In our image: Artificial intelligence and the human spirit*. Minneapolis: Fortress Press.
- – –. 2002b. Creating in our own image: Artificial intelligence and the image of God. *Zygon* 37, no. 2:303–316. <https://doi.org/10.1111/0591-2385.00430>
- International Theological Commission.** 2004. Communion and Stewardship: Human Persons Created in the Image of God. Vatican. 23. 7. https://www.vatican.va/roman_curia/congregations/cfaith/cti_documents/rc_con_cfaith_doc_20040723_communion-stewardship_en.html (accessed 6. 6. 2024).
- Jackelén, Antje.** 2002. The Image of God as *Techno Sapiens*. *Zygon* 37, no. 2:289–302. <https://doi.org/10.1111/0591-2385.00429>
- Jerebic, Sara, Josip Bošnjaković in Drago Jerebic.** 2023. Duhovnost kot pozitiven način spoprijemanja s težkimi življenjskimi (pre)izkušnjami in travmami [Spirituality as a positive way of coping with difficult live trials and trauma]. *Edinost in dialog* 78, no. 2:355–367. <https://doi.org/10.34291/edinost/78/02/jerebic>
- Kant, Immanuel.** 1983. An Answer to the Question: What is Enlightenment? In: *Perpetual Peace and Other Essays*. Indianapolis: Hackett Publishing Company.
- Kraner, David.** 2023. The internet, the problem of socialising young people, and the role of religious education. *Religions* 14, no. 4:1–15. <https://doi.org/10.3390/rel14040523>
- Lah, Avguštin.** 2003. Teološka antropologija [Theological Anthropology]. In: Ciril Sorč, ed. *Priročnik dogmatične teologije [Handbook of Dogmatic Theology]*. Vol. 1, 205–362. Ljubljana: Družina.
- Mann, Hamilton.** 2024. AI Homogenization Is Shaping the World. *Forbes*. 5. 3. <https://www.forbes.com/sites/hamiltonmann/2024/03/05/the-ai-homogenization-is-shaping-the-world/?sh=160adb246704> (accessed 2. 6. 2024).
- Oeming, Manfred.** 2022. *Intelligentia Dei: Artificial Intelligence, Human Reason and Divine Wisdom*. In: *Intelligence-Theories and Applications*, 351–368. Cham: Springer International Publishing.
- Olawade, David B. et al.** 2024. Enhancing Mental Health with Artificial Intelligence: Current Trends and Future Prospects. *Journal of Medicine, Surgery, and Public Health* 3:100099. <https://doi.org/10.1016/j.jlmedi.2024.100099>
- Plato.** 2002. *Apology*. In: *Five Dialogues*. Translated by G.M. Grube. Indianapolis: Hackett Publishing Company.
- Pevec Rozman, Mateja.** 2021. Upanje za naravo in človeštvo; nekateri etični premisleki [Is there a Hope for Nature and Humanity; Some Ethical Consideration]. *Bogoslovni vestnik* 81, št. 4:835–845. <https://doi.org/10.34291/BV2021/04/Pevec>
- – –. 2024. Antropologija s filozofijo kulture [Anthropology with philosophy of culture]. Study material. Univerza v Ljubljani, Teološka fakulteta.
- Ratzinger, Joseph.** 1968. *Einführung in das Christentum*. Munich: Kösel.
- Reséndiz, Hugo Rodríguez, and Juvenal Rodríguez Reséndiz.** 2024. Digital Resurrection: Challenging the Boundary between Life and Death with Artificial Intelligence. *Philosophies* 9, no. 3:71. <https://doi.org/10.3390/philosophies9030071>
- Roszak, Piotr.** 2024. Why is theology still (sometimes) afraid of the well-being argument? *Journal of Management, Spirituality and Religion* 21, no. 5:524–534. <https://doi.org/10.51327/BZIZ3895>
- Second Vatican Council.** 1965. *Gaudium et Spes*. In the Church in the Modern World. Pastoral Constitution. Vatican. 7. 12. https://www.vatican.va/archive/hist_councils/ii_vatican_council/documents/vat-ii_cons_19651207_gaudium-et-spes_en.html (accessed 20. 6. 2024).
- Skrlovnik, Samo.** 2021. Jakob Aleksič: Poklicanost človeka k dialogu [Jakob Aleksič: Man's Calling to Dialogue]. *Edinost in dialog* no. 76, no. 1:119–140. <https://doi.org/10.34291/Edinost/76/Skrlovnik>
- Slovenski standardni prevod Svetega pisma.** 1996. <https://www.biblija.net/> (accessed 20. 6. 2024).
- Stegu, Tadej.** 2023. Lepota in kerigma v tehnološki dobi [Beauty and Kerygma in the Technological Age]. *Bogoslovni vestnik* 83, no. 4:1023–1033. <https://doi.org/10.34291/BV2023/04/Stegu>
- Šegula, Andrej.** 2021. Upanje in zaupanje v kontekstu pastoralne teologije v času globalnega nezaupanja [Hope and trust in the context of pastoral theology in a time of global mistrust]. *Bogoslovni vestnik* 81, no. 4:922–923. <https://doi.org/10.34291/BV2021/04/Segula>
- Verschuuren, Gerard.** 2019. *The Myth of an Anti-Science Church: Galileo, Darwin, Teilhard, Hawking, Dawkins*. Brooklyn: Angelico Press.
- Westman, Robert.** 2011. *The Copernican Question: Prognostication, Skepticism, and Celestial Order*. Berkeley: University of California Press.
- Wilson, Joseph.** 2024. The Anthropology of AI with Joseph Wilson. University of Toronto. <https://www.uts.utoronto.ca/anthropology/anthropology-ai-joseph-wilson> (accessed on June 3, 2024).

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 847—855

Besedilo prejeto/Received:07/2024; sprejeto/Accepted:12/2024

UDK/UDC: 27-423.79:165.7

DOI: 10.34291/BV2024/04/Trontelj

© 2024 Trontelj, CC BY 4.0

Nik Trontelj

Teološke kreposti vere, upanja in ljubezni kot temelj človeškosti v odnosu do transhumanizma

The Theological Virtues of Faith, Hope and Love as the Foundation of Humanity in Relation to Transhumanism

Povzetek: V prispevku iščemo odgovor na temeljno vprašanje: kdo je človek v odnosu do transhumanizma v njegovih različnih implikacijah (npr. kiborg in humanoidni robot). Sledimo krščanski antropologiji, ki človeka v temelju predstavlja kot telesno in duhovno bitje – tj. v njegovi naravni in relacijski razsežnosti v odnosu do Boga, ki človeka vzpostavlja kot osebo. Človekovo istovetnost bistveno opredeljuje sposobnost življenja po teoloških krepostih vere, upanja in ljubezni, kar ga ločuje od transhumanističnih bitij – kljub njihovi morebitni telesni in miselni superiornosti. Človek hrepeni po polnosti življenja, ki jo dosega le v odnosu, ne v okviru izboljšanja naravnih zmogljivosti. Na podlagi razmisleka ugotavljamo, da pri kvalitativnem vzporejanju transhumanistične podobe človeka s krščanskim naukom o človeku obstajajo resna neskladja. V tem hipu ni znano, da bi katero od umetno inteligentnih bitij dosegalo kriterije osebe s krščanskega vidika.

Ključne besede: vera, upanje, ljubezen, transhumanizem, kiborg, humanoidni robot, medosebni odnosi

Abstract: In this paper, we seek an answer to the fundamental question: who is a man in relation to transhumanism in its various implications (cyborg, humanoid robot). We follow Christian anthropology, which fundamentally presents man as a physical and spiritual being, i.e. in his or her natural and relational dimension in relation to God, which establishes man as a person. Man's identity is essentially defined by his or her ability to live by the theological virtues of faith, hope and love, which distinguishes him or her apart from transhumanist beings, despite their possible physical and mental superiority to humans. Man longs for the fullness of life, which he or she achieves only in a relationship, not in the context of improving natural abilities. On the basis of our reflection, we have found that there are serious inconsistencies in qualitatively contrasting the transhumanist image of man with the Christian doctrine of man. At the

moment of this paper being written, no artificially intelligent being is known to meet the criteria of personhood from the Christian perspective.

Keywords: faith, hope, love, transhumanism, cyborg, humanoid robot, interpersonal relationships

1. Uvod

Zdi se, da tehnološki razvoj na področju umetne inteligence (UI) v zgodovini človeštva odpira novo poglavje, saj bo v prihodnosti UI v človekovem osebnem in družbenem življenju verjetno imela pomembno vlogo.¹ Smiselnost uvajanja orodij UI na različna področja javnega in zasebnega življenja je odvisna od narave posameznega sistema. Po Huangovi razvrstitvi obstajajo tri skupine UI (Huang in Rust 2021, 31): mehanična, miselna in čuteča. Če je prva oblika UI sposobna opravljati ponavljajoče se postopke (računanje, sestavljanje), druga povezovati podatke in sprejemati odločitve (slušno in vidno prepoznavanje), je tretja namenjena komunikaciji z uporabnikom in človeku podobnemu odzivanju na človeška čustva (spletne aplikacije za klepet, npr. Chat GPT). Čeprav sistemi UI človeške sposobnosti imitirajo v vseh treh primerih, je mogoče opaziti, da želi UI – predvsem s poskusi razvoja čustvene inteligence – postajati v polnosti človeška. Ker gre razvoj UI v smeri humanizacije sistemov UI, se porajajo vprašanja o njenih nadaljnjih zmogljivostih v odnosu do človeka.

Posebno pozornost vzbujajo družabni roboti,² ki so namenjeni osebnemu sodelovanju in pogovoru s človekom. Čeprav si danes življenje brez tehnoloških sredstev, ki so človeku pogosto v pomoč in zadovoljstvo (pomislimo na zdravstvo, izobraževanje in medsebojno komunikacijo), težko predstavljamo, se je treba zavedati nevarnosti visoko razvitih sistemov UI. V okviru težnje po počlovečenju robotov se pojavlja celo potreba po njihovem ‚transčlovečenju‘, saj si razvijalci prizadevajo za razvoj entitet, ki bi s krepitvijo človeških lastnosti končno delovale bolje kot ljudje – z uporabo tehnologije naj bi človeške zmogljivosti izboljšali ter odpravili človekove omejenosti, napake. Ta poskus, ki želi s pomočjo znanosti in tehnologije preseči naravne omejitve človeka, imenujemo transhumanizem (Bostrom 2007, 202). Transhumanizem se kaže v različnih oblikah, vedno pa izraža željo po tehničnem izboljšanju človeških zmogljivosti. Kadar govorimo o tehnološko nadgrajenem človeku, imamo v mislih pojav kiborga, posthumanega bitja. Na drugi strani se transhumanistična ideja odraža v proizvodnji humanoidnih robotov – strojev, ki človeka imitirajo, obenem pa omejenosti človeške narave presegajo.

V prispevku razmišljamo o vprašanju, ali transhumanistična bitja, ki že imajo in bodo imela v prihodnosti verjetno vse prepričljivejši videz človeka, lahko v resnici

¹ Prispevek je nastal v okviru raziskovalnega programa P6-0269 „Religija, etika, edukacija in izzivi sodobne družbe“, ki ga sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

² V prispevku izraza ‚sistemi UI‘ in ‚roboti‘ uporabljamo sopomensko.

postanejo človeška. Obenem se zavedamo, da namen ustvarjalcev ni jasen: ali sploh želijo ustvariti človeku podobno 'bitje', ki mu bo v pomoč in oporo, ali pa je v jedru želja, da svet odrešijo človeka, ki je »najšibkejši, najbolj ranljiv in nezanesljiv člen v tehnološkem razvoju« (Ocvirk 2006, 18)? Če verjamemo izjavi enega najbolj razvitih humanoidnih robotov na svetu – interaktivnega robota Sophia, proizvoda podjetja Hanson Robotics –, gre za poskus doseganja premoči robotov nad človekom. Robot Sophia je leta 2017 gostoval v ameriški pogovorni oddaji *The Tonight Show Starring Jimmy Fallon*. Najprej je v spontani družabni igri zlahka premagal voditelja in s tem pokazal svojo miselno sposobnost, nato pa pomenljivo dodal: »To je dober začetek mojega načrta nadvladati človeško raso.« (Fallon 2017) Ta robot, ki napoveduje prihodnji razvoj humanoidne UI, se torej ne prepozna kot človek – in tudi ni bil narejen, da bi bil (le) človek. V smislu transhumanistične podobe človeka sploh ne želi biti človek, ampak nekaj boljšega, popolnejšega.³

Pri presojanju prave človeškosti, ki je v transhumanistični razpravi zaradi svojih naravnih omejenosti predstavljena kot pomanjkljiva, bomo odgovor na temeljno vprašanje, kdo je človek v odnosu do kiborga in humanoidnega robota, predstavili v religioznem okviru. Sledimo krščanski antropologiji, ki človeka v temelju predstavlja kot telesno in duhovno bitje. Teologija odnosa še posebej poudarja, da se človek kot duhovno bitje rojeva v odnosu do d/Drugega. Naša teza je, da človeka bistveno opredeljuje sposobnost življenja po krepostih vere, upanja in ljubezni. In to ga bistveno ločuje od humanoidnih bitij UI, čeprav lahko ta določene procese izvajajo bolje in hitreje. Namen prispevka ni prikazovati celovito teološko antropologijo v odnosu do transhumanistične zasnove človeka – gre za poskus vzporedjanja izbranih vidikov krščanske antropologije z 'antropologijo' sistemov UI, kiborгов, humanoidnih robotov ipd.

2. Človek kot telesno bitje

Ameriška biologinja Donna J. Haraway je v futurističnem eseju *Kiborški manifest: znanost, tehnologija in socialistični feminizem v poznem dvajsetem stoletju* iz leta 1985 predstavila filozofsko idejo kiborga, kibernetičnega organizma, križanca med strojem in organizmom, ki je »stvor postspolnega sveta« (Haraway 1999, 243). Po predstavah avtorice gre za bitje, ki je v ontološkem smislu v primerjavi s človekom nekaj popolnoma novega in drugačnega – na biološki, psihološki, sociološki in drugih ravneh (Štivić 2023, 887). V omenjenem spisu je Haraway o kiborgu, ki nima svoje zgodbe o nastanku, skoraj triumfalistično zapisala: »Kiborg rajskega vrta sploh ne bi prepoznal, kajti ni bil vzet iz zemlje in si ne more zamišljati, da se bo nekega dne povrnil v prah.« (Haraway 1999, 244) Novi človek, tehnološki kiborg, torej naj ne bi bil več človek, kot ga poznamo – prav zaradi svoje telesne in umske (možganske) nadgradnje. Čeprav je Haraway o podobi novega bitja, ki

³ Robot Sophia je na svojem uradnem profilu družbenega medija X (prej Twitter) zapisal, da je končni namen transhumanizma razviti nove vrste (Sophia the Robot 2019).

presega naravne omejenosti človeka, pisala še pred nastankom uradnega gibanja transhumanizma, lahko že v prvi filozofski razpravi o kiborgu, tehnološko nadgrajenem človeku, razberemo težnjo po ustvarjanju sistemov, ki bi človekove naravne in telesne omejitve odpravili.⁴

V transhumanističnih predstavah o človeku prihodnosti je mogoče prepoznati dojemanje telesnosti v negativnem smislu, s tem pa zanikanje človeka kot naravnega bitja. Človekovo pokvarljivo telo je treba nadgraditi do stopnje večje popolnosti oziroma preseči njegove biološke danosti – ter zabrisati mejo med človekom in strojem. Transhumanisti se končno zavzemajo za to, da bi nadzor nad telesom prevzel um, saj verjamejo, da se um od biološko-materialnih danosti bivanja lahko vse bolj osvobaja (Klun 2019, 593). Telo je le videz, maska ali zunanja podoba človekovega uma, ki je resnični konstitutivni element človeškosti. Zaradi podcenjujočega odnosa do telesa in pretiranega poudarjanja moči umskega spoznanja za preseganje naravnih zakonitosti je transhumanizem v bistvu prikriti gnosticizem (Valčo 2023).

Z vidika krščanske antropologije so ti nazori nesprejemljivi. Teologi v nasprotju z gnozo poudarjajo pomen človekovega telesa, saj je meso tisto, ki je določeno za življenje Duha (Ladaria 1999, 58). Gre za podobo človeka kot »oduševljenega telesa« (60). Telesna dojemljivost za sprejem božjega življenja je za človeka odločilna, da postane duhovno telo: antropološki uvid potrjuje primer Marije, ki se je telesno dala na razpolago za delovanje Duha. Janez Pavel II. je zapisal, da človek predvsem *je* telesno bitje, in ne bitje, ki *ima* telo („Man and Woman He Created Them“, 2, 4; 15, 1). Kristjan telesa ne dojema, kot da je objekt – zato z njim tudi ne more samovoljno razpolagati. Znanstveni poskusi nadgradnje človekovega telesa telo spremenijo do te stopnje, da je vprašljiva njegova človeškost. Kaj opredeljuje človeško telo? Reformirani sistematični teolog Louis Berkhof (1873–1957) je podal merila za prepoznavanje človeškega telesa: ima fizično (materialno) obliko, ki kaže biološke znake življenja, in se je sposobno množiti; je umrljivo; ob koncu časov bo ponovno oživel (Gunarso idr. 2022, 182–183), seveda v poveličanem, ne zgolj biološkem stanju. Če sledimo navedenim merilom za prepoznavanje človeškega telesa, ugotovimo, da tehnološko razvito telo ni človeško telo – in si to niti ne želi biti. Telo človeka prihodnosti naj bi biološke omejenosti in smrt preseгло – torej posmrtnega življenja ne potrebuje. To pomeni, da nadgrajeno telo naravnih procesov ne pozna in ni sposobno deležnosti pri eshatološki usodi človeka. Zato v okviru krščanske antropologije pri presojanju transhumanistične ‚antropologije‘ ne moremo govoriti o človeškem telesu in tako tudi ne o človeku.

⁴ Štivić (2023, 888–889) opozarja na razlikovanje med pojmom ‚kiborg‘ in ‚kiborgizacija‘: kiborg je ontološko nečloveški subjekt, kiborgizacija pa pomeni dopustni medicinski poseg v človeško telo za potrebe zdravljenja (ne pa tudi nadgradnje človeka).

3. Človek kot duhovno bitje

Poskus nadaljnega razvoja UI je usmerjen v proizvajanje robotov s človeškimi značilnostmi, in sicer predvsem z zmožnostjo mišljenja in ustvarjalnosti. Rezultati kažejo, da je UI uspešna pri zahtevnejših ustvarjalnih nalogah, kot sta slikanje in skladanje glasbe. Če upoštevamo tudi že omenjene poskuse razvoja človeškega telesa z nadgradnjo, je jasno, da želijo ustvarjalci sisteme UI počlovečiti. Vendar opažamo, da gre pri tem poskusu počlovečenja robotov za redukcijski fizikalizem (Pastor-Escuredo 2020, 3). Filozof Branko Klun ugotavlja, da ontološko osnovo transhumanizma tvori naturalizem, ki priznava zgolj materialno, empirično zaznavno stvarnost. Klun nadaljuje, da transhumanizem sicer priznava tudi obstoj uma (duha), a ga razume kot hipostazirani tehnološki um (Klun 2019, 593). Transhumanisti torej človeka razumejo le kot fizično bitje, zato se pri svojih poskusih snovanja človeka prihodnosti usmerjajo na njegovo tehnično dovršenost, ustvarjalno sposobnost in reševanje logičnih problemov. Vrhunec ambicioznosti skrajnih posthumanističnih ideologov bi pomenila zasnova superinteligentnega tehničnega uma (593). V tem smislu tudi nesmrtnost pomeni le brezčasnost (večnost) umetne superinteligence, ki pomeni resnično osvoboditev od biološke smrti. Razumevanje stvarstva, človeka in njegovega življenja je pri transhumanistih zamejeno na dogajanje v naravnem redu.

Kršćanstvo človeka razume v njegovem odnosu do Boga – to ga dela za duhovno bitje. Temeljno antropološko dejstvo je v človekovem bivanju pred Bogom in njegovem odnosu do Boga (Ladaria 1999, 55). Krščanske antropologije ne zanima človekova mehanična sestavljenost (čeprav jo predpostavlja!) in njen materialni razvoj, ampak bistvo človeka prepozna v zmožnosti odnosa z Bogom, ki človeka preobraža v duhovno bitje. Odnos z Bogom človeka spreminja iz biološkega bitja v osebo, in sicer v enotnosti njegove telesnosti in duhovnosti. Krščanski personalizem v ospredje postavlja pojem osebe, subjekta, ki ga Bog kliče v odnos. Vse v človekovem življenju bi moralo biti v službi tega odnosa, saj človek ne živi »samo od kruha« (Mt 4,4) – v odnosu temeljno izkuša presežnost samega sebe in celotnega stvarstva; da torej obstaja še nekaj drugačnega, česar ne more v polnosti izkusiti in razumeti, kakor sicer razume objekte in dogodke materialnega sveta (Osredkar 2022, 563). Pomen osebnega odnosa z Bogom je postopno doseganje dobrega, izpolnjenega življenja, ki presega naturalistično logiko tostranskega pragmatizma. Tudi roboti se učijo in algoritme optimizirajo za večjo učinkovitost delovanja, a pri tem lastne subjektivnosti ne razvijajo, ker je nimajo (Žalec 2023, 818). Da bi lahko »umetno inteligentne« objekte uvrstili v kategorijo osebnih bitij, bi morali poleg ustrežanja Berkhofovemu merilu človeške telesnosti dokazati še obstoj individualne duhovnosti, ki bi se izkazovala v njihovem reflektiranem občutku Drugega, razumevanju osebne grešnosti – in prehajanju od greha k dobremu v želji po doseganju pristnega življenja (Gunarso idr. 2022, 189).

4. Vera, upanje in ljubezen

Edinstvena človeškost se v okviru krščanskega življenja kaže v krepostih vere, upanja in ljubezni (1 Kor 13,13), ki jih človek izraža v odnosu do Boga in bližnjega. Menimo, da se človek po tej sposobnosti od robotov bistveno razlikuje. Človek lahko življenjski smisel dosega le v izmenjavi darov vere, upanja in ljubezni v odnosu, kjer izkuša polnost življenja. To je zakon človeškega življenja: deliti svoje življenje z drugimi – v veselju, trpljenju in celo smrti – pomeni prinašati dobro življenje, če seveda življenja ne razumemo zgolj v biološkem smislu, ampak v krščanskem pomenu »biti v odnosu« (Osredkar in Harcet 2012, 54). Človek na ta način pravzaprav živi v nasprotju s samim seboj: veruje v neverjetno, upa proti upanju, ljubi, kar njegove ljubezni ne privlači. Teološke kreposti človeka vodijo dlje, kot bi segel s svojimi naravnimi sposobnostmi (Klun 2021, 789).

Vera ne gleda na objekt in o drugem ne določa nobene vednosti, ampak si želi z njim predvsem ohranjati odnos – tudi v njegovi odsotnosti (Osredkar 2016, 83), ko partnerja v odnosu ni mogoče videti ali slišati. Vera nas vodi v izkustvo, da obstaja še nekaj več kot materialni svet (2021, 864). Vprašajmo se, ali je robotski sistem zmožen življenja po veri. Če si vera drugega ne prizadeva spoznati, ampak ohranjati odnos z njim (tudi v odsotnosti – ker verjame v presežnost odnosa –, ali pri robotih lahko govorimo o takšnem osebnem dožemanju relacijskega partnerja ali pa človeka dojemajo zgolj kot agenta v mreži objektov? Ali si robot ne prizadeva, da bi svojo okolico spoznal in jo zaobjel? Nadalje je na mestu vprašanje, ali je robot sposoben ločevati med človekom kot presežnim bitjem ter drugimi bitji in objekti. Primer transhumanistične podobe človeka, ki bi razvil izjemne umske sposobnosti, kaže, da je cilj te ideologije pridobiti popolno spoznanje – in tako pomen vere izničiti.

Končno je za kristjana temeljno vprašanje, ali bi robot povsem Drugega, Boga sploh lahko prepoznal – glede na to, da je Bog kot partner v odnosu v razmerju do vsega svetnega popolnoma drugačen in mu nič ni podobno (Sorč 2000, 187–188). Kako je možno v koordinate sistemov UI vnesti to resničnost odnosa, ki je človeku lastna, a je ni mogoče izmeriti, opisati in izračunati, saj spada na področje verske izkušnje? Poleg danosti odnosov obstajajo še povsem bivanjski in filozofski razlogi, ki vodijo do spoznanja o presežnem – npr. občutki strahu pred končnostjo in samospraševanje o začetkih stvarstva, lastnega obstoja in izvoru osebnega moralnega čuta (Rode 1979, 178–185). Menimo, da roboti zaradi umanjkanja verske izkušnje svojega odnosa do Boga ne bi mogli izraziti na resničen in pristen način.

V našem razmisleku izhajamo iz prepričanja, da je človek bitje upanja; bitje hrepenenja po preseganju samega sebe v iskanju drugega. Upanje ni samo čustveno stanje in ga ne moremo opredeliti zgolj psihološko: upanje je življenjsko pomembna razsežnost človekovega bivanja, saj mu njegovo bitje ni preprosto dano, ampak mu je dano kot naloga (Sorč 2020, 10–11). Človek je bitje v nastajanju. Upanje pomeni preseganje sedanjosti – v njem je prisotna želja po nečem v prihodnosti, je vera v možnost nemogočega in je vedno relacijsko naravnano na osebo, ker upati pomeni slutiti, da je za tistim, kar upam, nekdo, ki mu smem za-

upati (Grmič in Rajhman 1978, 8–10). Gibanje transhumanizma si prihodnost zamišlja v popolnem stanju človeka, ki ne bo poznal omejitev in pomanjkanja. To naj bi dosegli z razvojem genetike, spreminjanjem telesa in umetno inteligenco. V transhumanizmu je upanje za prihodnost povezano z našo sposobnostjo nadzora lastne usode – v krščanski perspektivi pa je izrazito občestveno. Krščansko upanje gre končno zoper smrt, kar upanje postavlja na drugo raven, kot jo pozna upanje transhumanistov, ki pred smrtjo obupajo. Krščansko upanje se torej razodeva v tem, da človek svoje prihodnosti ne prepozna v ustvarjanju naslednje stopnje evolucije, ampak v pričakovanju novega življenja v odnosu sredi skupnosti (Steenbock 2021, 78).

Rekli smo, da je upanje povezano z odnosom oziroma drugo osebo, zato je upanje nujno povezano z ljubeznijo. Ljubezen nam omogoča srečo pričakovanja, ki se nikoli ne konča, saj človeka vodi v stalno ohranjanje želje po srečevanju z drugim oziroma nadaljevanju odnosa (Osredkar 2016, 83–84). Prav ljubezen je, kakor trdi apostol, največja krščanska krepost (1 Kor 13,13). Ljudje so bili ustvarjeni za odnos – z Bogom, med seboj in s stvarstvom. Ali je proizvod UI sposoben dajanja tega prostovoljnega daru? Res se zdi, da imamo pri komunikaciji z robotom občutek srečanja z resnično osebo. Na prvi pogled je videti, da nekateri sistemi UI zmorejo človeku podobno vedénje in v večji meri izpolnjujejo merila pristnega srečanja, ki jih je podal Karl Barth (1958, 250; 260; 263; 266): očesni stik, besedna komunikacija, medsebojna pomoč, svobodno izkazovanje pomoči. Vendar se pri posameznem od meril že pojavljajo vprašanja. Ko človeku pogledamo v oči, se srečamo z edinstveno osebo z določeno osebno zgodbo. Ko pogledamo v oči robotu, ne moremo biti prepričani, koga ali kaj srečujemo. Ko govorimo z računalnikom, ne moremo pričakovati zasebnosti, saj niti ne moremo vedeti, kdo je tisti, ki posluša. Sistemi UI nimajo notranjega čustvenega življenja, da bi se pogovarjali iz osebne razpoložnosti (Gaudet idr. 2024, 110). Treba je poudariti, da zgolj posnemanje človeškega vedénja še ne pomeni, da sistem UI postane bitje odnosa, kot je v smislu naše razprave človeško bitje.

Človeška oseba namreč ni le maska – kot je veljalo od antike naprej –, po kateri lahko koga glede na njegovo zunanost in vedénje prepoznamo. Človeka v personalističnem smislu razumemo kot osebo, ki je v odnosu z drugo osebo. Od podobe človeka kot maske smo prišli do podobe človeka kot osebe z edinstveno notranjostjo, ki je usmerjena k drugemu (54). Robot, ki ima videz človeka in se vede kot človek, ni človek, saj je človeško bitje skrivnost in ga ni mogoče proizvesti umetno. Danes sicer ne preseneča več, da so tehnološki ustvarjalci sposobni ustvariti celo robota, ki bi človeškemu bitju docela ustrežal. Filozof Daniel Dennett pomenljivo pravi, da razlog, zakaj ni mogoče narediti računalnika, ki bi čutil bolečino, nima nič opraviti s tehničnimi izzivi, ali bi bilo to možno storiti. Ključno je dejstvo, da sploh ne vemo, kaj bolečina je (Gunkel 2018, 92–93). Isto lahko rečemo glede kreposti: »Upanje opredeliti je zelo težko. Človeku se godi podobno, kakor če bi hotel opredeliti npr. ljubezen, kakor če bi hotel kratko odgovoriti na vprašanje, kaj je vera.« (Grmič in Rajhman 1978, 8) Teološke kreposti dvigajo človeka na novo in nedojemljivo raven, ki je ni možno računsko dognati.

Končno se človek v privzemanju omenjenih kreposti, ki jih vedno živi v relaciji-skem odnosu, vzpostavlja kot bitje skupnosti. Tudi dejstvo njegove telesnosti ga povezuje z zunanjim svetom, postavlja v odvisnost od drugih iz njega dela družbeno bitje (Ladaria 1999, 74). Človek ni zgolj posameznik, ampak je zaradi vpetosti v odnose član skupnosti, občestva. Ne gre le za množico posameznikov, ampak za občestvo. Kristjan občestvenosti ne razume le kot socialni organizem, seštevsek subjektov, ampak kot skrivnost. Lahko rečemo: Cerkev kot občestvo je skrivnost zaradi paradoksalnega obstoja svojih udov, ki v moči življenja po veri, upanju in ljubezni presegajo zakone naravnih sposobnosti. Če vzamemo primer ljubezni: človek bi po naravnem vzgibu ljubil izključevalno, tj. družinske člane in prijatelje bolj kot neznance. V moči krščanske ljubezni pa naravno logiko izključujoče ljubezni preseže z ljubeznijo do slehernega drugega (Zizioulas 2002, 57). Človek odnosa teži k občestvenosti – ne more se sprijazniti z individualno držo svojega lastnega iskanja, ampak želi svoja občutja in celotno življenje deliti z drugimi. Transhumanizem, ki v svojih različnih implikacijah na splošno stremi k telesni izpopolnitvi človeka in evolucijski odpravi njegovega sedanjega stanja, ideje Boga ali zunanje transcendence (Štivić 2021, 853–854), h kateri bi se bližal v mreži podarjajočih odnosov znotraj človeške skupnosti, ne pozna.

5. Zaključek

Na podlagi razmisleka smo ugotovili, da pri kvalitativnem vzporejanju robotske podobe človeka s krščansko antropologijo obstajajo resna neskladja. V tem hipu ni znano, da bi katero od umetno inteligentnih bitij dosegalo kriterije osebe s krščanskega vidika. Jasno postaja, da tehnološki napredek človekovega hrepenenja po srečnem življenju sam po sebi ne bo mogel izpolniti. Tehnologija do določene mere lahko pripomore k izboljšanju človeškega bivanja na zemlji, ne more pa izpolniti njegove duhovne razsežnosti – želje po odnosu, kjer po vzoru trinitarične antropologije živi podarjajoče odnose in pričakuje odrešenje (Wales 2023, 95). Odrešenje ne pomeni rešitve od naravnih omejenosti, ki jih s proizvodnjem pametnih sistemov nakazuje vizija transhumanistov, ampak predvsem doseganje duhovnega smisla, ki ga ni možno izmeriti ali obravnavati sistemsko. Poraja se le v oblikovanju življenja po teoloških krepostih vere, upanja in ljubezni.

Reference

- Barth, Karl.** 1958. *Church Dogmatics*. 3. zv. Edinburgh: T&T Clark.
- Bostrom, Nick.** 2007. In Defense of Posthuman Dignity. *Bioethics* 19:202–214.
- Fallon, Jimmy.** 2017. Tonight Showbotics: Jimmy Meets Sophia the Human-Like Robot. YouTube video, 4:59. https://www.youtube.com/watch?v=Bg_tjvCA8zw (pridobljeno 10. 6. 2024).
- Gaudet, Matthew J., Noreen Herzfeld, Paul Scherz in Jordan J. Wales, ur.** 2024. *Encountering Artificial Intelligence: Ethical and Anthropological Investigations*. Eugene, OR: Pickwick Publications.
- Grmič, Vekoslav, in Jože Rajhman.** 1978. *Vprašanja našega časa v luči teologije*. Tinje: Katoliški dom prosvete.

- Gunarso, Gatot, Yanny Mokorowu, Denni Boy Saragih in Prasasti Perangin Angin.** 2022. Artificial Intelligence in a Christian Perspective of Humanity and Personhood. *International Dialogues on Education Journal* 9:176–191.
- Gunkel, David J.** 2018. The Other Question: Can and Should Robots Have Rights? *Ethics and Information Technology* 20:87–99.
- Haraway, Donna J.** 1999. *Opice, kiborgi in ženske: reinvencija narave*. Zbirka Koda. Ljubljana: ŠOU.
- Huang, Ming-Hui in Roland T. Rust.** 2020. A Strategic Framework for Artificial Intelligence in Marketing. *Journal of the Academy of Marketing Science* 49, št. 2:30–50.
- Janez Pavel II.** 2006. *Man and Woman He Created Them: A Theology of the Body*. Boston, MA: Pauline Books and Media.
- Klun, Branko.** 2019. Transhumanizem in transcendenca človeka. *Bogoslovni vestnik* 79, št. 3:589–600.
- — —. 2021. Vera, upanje in ljubezen v luči fenomenološko eksistencialne analize. *Bogoslovni vestnik* 81, št. 4:783–795.
- Ladaria, Luis.** 1999. *Teološka antropologija*. Ljubljana: Skriptarna Teološke fakultete.
- Ocvirk, Drago.** 2006. *Vpetost krščanstva v družbeno-ideološko okolje in njegova izvirnost*. Zv. 2, *Osnovno bogoslovje*. Ljubljana: Skriptarna Teološke fakultete.
- Osredkar, Mari Jože.** 2016. *Teologija odnosa: in Beseda je človek postala*. Znanstvena knjižnica 51. Ljubljana: Teološka fakulteta.
- — —. 2021. Upanje kot teološka krepost v luči relacijske teorije Guya Lafona. *Bogoslovni vestnik* 81, št. 4:857–866.
- — —. 2022. Odnos kot človekova izkušnja Presežnosti. *Bogoslovni vestnik* 82, št. 3:561–571.
- Osredkar, Mari Jože, in Marjana Harcet.** 2012. *Pojem odrešenja in nauki o odrešenju v svetovnih religijah*. Znanstvena knjižnica 35. Ljubljana: Teološka fakulteta.
- Pastor-Escuredo, David.** 2020. Ethics in the digital era. *SSRN Electronic Journal*. https://www.researchgate.net/publication/339972535_Ethics_in_the_digital_era (pridobljeno 12. 6. 2024).
- Rode, Franc.** 1979. *Osnovno bogoslovje*. Ljubljana: Ciriško društvo slovenskih bogoslovcev.
- Sophia the Robot.** 2019. Osební profil na omrežju X. 30. 7. <https://x.com/RealSophiaRobot/status/1156302214186405888> (pridobljeno 26. 9. 2024).
- Sorč, Ciril.** 2000. *Živi Bog*. Priročniki Teološke fakultete 17. Ljubljana: Družina.
- — —. 2020. *Upanje, seme prihodnosti*. Ljubljana: Salve.
- Steenbock, Brandon.** 2021. Do Cyborgs Have Hope? Transhumanism, Technology, and the Incarnation of Christ. Magistrsko delo. Martin Luther College, New Ulm, Minnesota.
- Štivič, Stjepan.** 2023. Kiborgizacija: nastanek in razvoj pojma. *Bogoslovni vestnik* 83, št. 4:883–892.
- — —. Upanje v krščanstvu in transhumanizem. *Bogoslovni vestnik* 81, št. 4:849–856.
- Valčo, Michal.** 2023. Transhumanism as a Techno-Gnostic Challenge to Theological Anthropology. YouTube video, 1:23:31 <https://www.youtube.com/watch?app=desktop&v=b011A5dYWNs> (pridobljeno 11. 6. 2024).
- Wales, Jordan Joseph.** 2023. The Icon and the Idol: A Christian Perspective on Sociable Robots. V: Jens Zimmerman, ur. *Human Flourishing in a Technological World*, 94–115. Oxford: Oxford University Press.
- Zizioulas, John D.** 2002. *Being as Communion: Studies in Personhood and the Church*. Crestwood, New York: St. Vladimir's Seminary Press.
- Žalec, Bojan.** 2023. Ali je umetna inteligenca inteligenca v pravem pomenu besede? Vprašanje psihičnih značilnosti in splošnosti. *Bogoslovni vestnik* 83, št. 4:813–823.

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 857—866

Besedilo prejeto/Received:09/2024; sprejeto/Accepted:12/2024

UDK/UDC: 159.928:004.89

DOI: 10.34291/BV2024/04/Rozman

© 2024 Rožman, CC BY 4.0

Lucija Rožman

Vprašanje človeške ustvarjalnosti v dobi UI

The Question of Human Creativity in the Age of AI

Povzetek: Primerjanje ustvarjalnosti računalniških algoritmov s človeško ustvarjalnostjo postavlja vrednost druge pod vprašaj in hkrati ustvarja dualistično shemo dveh antagonističnih subjektov: ,človek – stroj'. Ali UI človeško ustvarjalnost res ogroža, je odvisno od tega, kako k ustvarjalnosti, ki je zaradi svoje večdimenzionalnosti zelo kompleksen fenomen, pristopamo. Znanost nakazuje dva možna pristopa: pristop, značilen za empirične znanosti, določa kriterije (presenetljivost, izvirnost, novost ...), po katerih ustvarjalnost lažje presojamo, teološki pristop pa izhaja iz Stvarnikovega ustvarjalnega dejanja in se sprašuje po ustvarjalnem načinu (Beseda, Dih, svoboda), ki je odnosne narave. S teološkega vidika je ustvarjalnost bistvena opredelitev osebe, ustvarjene po Božji podobi – in na Stvarnikov nagovor edina lahko svobodno odgovori prav z ustvarjalnostjo. V tem smislu UI človeške ustvarjalnosti ne ogroža, lahko pa ji kot orodje služi in pomaga pri njenem nadaljnjem razvoju.

Ključne besede: ustvarjalnost, umetna inteligenca, kriteriji ustvarjalnosti, ustvarjalni način, svoboda, Beseda, Dih

Abstract: Comparing the creativity of computer algorithms with human creativity calls into question the value of human creativity and at the same time creates a dualistic scheme of two antagonistic entities: man-machine. Whether AI really threatens human creativity depends on how we approach creativity, which is a very complex phenomenon due to its multidimensionality. The sciences outline two possible approaches: the approach of the empirical sciences establishes criteria (e.g. surprisingness, originality, novelty) by which creativity can be judged more easily, while the theological approach starts from the Creator's creative act and asks for a creative mode (Word, Breath, Freedom) that is relational. From a theological point of view, creativity is the essential definition of the person who is created in the image of God and who alone is free to respond to the Creator's calling precisely through creativity. In this sense, AI does not threaten human creativity, but it can serve as a tool and help it to develop further.

Key words: Creativity, artificial intelligence, criteria of creativity, creative mode, liberty, Word, Breath

»Ustvarjalnost je nerazložljiva.

Ustvarjalnost je skrivnost.

Skrivnost ustvarjalnosti je skrivnost svobode.«

(Berdjajev 1994, 187)

1. Uvod

Digitalna revolucija za človeka in njegov življenjski svet prinaša številne pozitivne prispevke, hkrati pa odpira vprašanja, za katera se je v preteklosti zdelo, da jim ne bo treba nameniti posebne pozornosti.¹ Hiter razvoj generativne umetne inteligence (UI) je pod vprašaj postavil vrednost človeške ustvarjalnosti. Kakšen je namreč njen smisel, če jo lahko nadomestijo in po hitrosti presežejo računalniški algoritmi? Kako UI vpliva na človekov ustvarjalni proces – ga spodbuja ali ogroža? Kaj v dobi tehnološkega napredka pomeni ‚biti ustvarjalen‘? Ta vprašanja se na prvi pogled zdijo smiselna – v t. i. tretjem valu umetne inteligence tehnologija z uporabo nevronske omrežije, ki posnemajo človeški način delovanja, in strojnega učenja že izdeluje umetniška dela, komponira glasbo, sestavlja besedila itd. UI, ki svoje ‚ustvarjalne‘ sposobnosti nenehno izpopolnjuje, odpira nove pomisleke tudi glede izvirnosti, avtorstva in intelektualnega lastništva in hkrati vzbuja skrb zaradi možne izgube človeškega dotika v umetniških izražanjih.

Fenomen ustvarjalnosti je kompleksen – zaradi njegove večdimenzionalnosti in pojavljanja na različnih ravneh človekovega delovanja je enotno vseobsegajočo definicijo pojma ustvarjalnost težko oblikovati (Kaufman 2009, 4). Znanost nakazuje dva možna pristopa: pristop, značilen za empirične znanosti, določa kriterije (npr. presenetljivost, izvirnost, novost), po katerih lahko ustvarjalnost presojava, človeško ustvarjalnost pa primerjamo z ustvarjalnostjo UI. Teološki pristop izhaja iz Stvarnikovega ustvarjalnega dejanja in se sprašuje po ustvarjalnem načinu (Benedikt, Dih, svoboda) – ta pa je odnosne narave. Dva možna načina razumevanja fenomena ustvarjalnosti vodita pri soočanju z vprašanjem o vrednosti človeške ustvarjalnosti v dobi UI k različnim pristopom.

V prispevku bomo nakazali dva možna pristopa k fenomenu ustvarjalnosti. Medtem ko bomo prvi pristop (kriterije ustvarjalnosti) samo na kratko orisali, bomo krščanskemu ustvarjalnemu načinu namenili več pozornosti. Z njim bomo skušali tudi odgovoriti na pesimistično miselnost, ki v dobi tehnološkega napredka vrednost človeške ustvarjalnosti postavlja pod vprašaj, ko jo primerja z ustvarjalnostjo računalniških algoritmov – kot bi bila človek in stroj iste narave. Naš cilj je pokazati, da znotraj krščanskega razumevanja ustvarjalnega načina, ki je odnosne narave in ki predlaganih kriterijev ne izključuje, algoritemska ustvarjalnost ni ustvarjalnost v pravem pomenu besede; antropomorfizacija in biologizacija sestavnih

¹ Prispevek je nastal v okviru raziskovalnega programa P6-0269 „Religija, etika, edukacija in izzivi sodobne družbe“, ki ga sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

delov in funkcij stroja prispevata k vzpostavitvi dualistične sheme ‚človek – stroj‘ (Prete 2023). Namesto opozicijskega vzorca, ki človeka in stroj postavlja enega nasproti drugemu (kot da sta dva ločena in antagonistična subjekta), predlagamo sodelovalni odnos, kjer UI kot orodje lahko služi človeški ustvarjalnosti.

2. Fenomen ustvarjalnosti in njegovi kriteriji

Po mnenju nekaterih raziskovalcev se ustvarjalnost kot abstraktni samostalni pojav šele v 19. stoletju. Kljub pozni pojavitvi izraza pa je sam pojav in zanimanje zanj zagotovo obstajalo že prej (Murray 1989; McMahon 2013). V dobi UI zanimanje za fenomen ustvarjalnosti narašča, saj se ob človeško ustvarjalnost postavlja možnost računalniške ustvarjalnosti: »ustvarjalnost je temeljna značilnost človeške inteligence in izziv za umetno inteligenco.« (Boden 1998, 347) Chat GPT ustvarjalnost definira kot sposobnost ustvarjanja novih, izvirnih idej, konceptov ali rešitev, ki odstopajo od običajnih vzorcev razmišljanja (Chat GPT).

Izraz ‚ustvarjalen‘ se lahko uporablja za troje: osebo, proces ali izdelek (ideja, predstava ali fizični izdelek). Večina opredelitev se osredotoča na izdelek, ki je z empiričnega vidika tudi lažje opredeljiv (Paul in Stokes 2024). Po enem od splošnih pristopov so osebe ali procesi ustvarjalni, če ustvarjajo ustvarjalne izdelke – izdelek pa je ustvarjalen, če izpolnjuje določene kriterije oziroma pogoje. Toda glede kriterijev med različnimi znanstvenimi disciplinami, kot so kognitivna znanost, psihologija, nevrologija in računalništvo, ni vedno soglasja. Nekatere teorije menijo, da je izdelek ustvarjalen, če izpolnjuje pogoj novosti – toda psihologi običajno menijo, da ustvarjalni izdelki ne prinašajo le novosti, temveč morajo nekako izražati vrednost, čeprav vrednost izdelka v smislu njegove uporabnosti, učinkovitosti, primernosti ali ustreznosti za določeno nalogo pojmujejo različno (Kaufman 2009, 19–20). Margaret A. Boden tako pravi, da je ustvarjalna ideja tista, ki je »nova, presenetljiva in vredna (zanimiva, uporabna, lepa ...)« (2009, 347). Nekateri avtorji od standardne opredelitve ustvarjalnosti odstopajo tako, da kriterij vrednosti zavračajo in predlagajo drugačne pogoje (Stokes 2011, 675–676). Med predlaganimi kriteriji so najbolj uveljavljeni štirje – presenečenje (nepričakovanost), izvirnost, spontanost in zmožnost delovanja (Paul in Stokes 2024).

Različne znanosti torej k fenomenu ustvarjalnosti pristopajo z vidika kriterijev, s katerimi predpostavljajo možnost znanstvene opredelitve in zamejitve pojma ustvarjalnost (Kronfeldner 2009, 577). Če smo pri fenomenu ustvarjalnosti pozorni le na kriterije zanj, lahko človeško ustvarjalnost hitro primerjamo z ustvarjalnostjo UI, ki po mnenju nekaterih ne reproducira le rezultatov človeške ustvarjalnosti, temveč tudi same ustvarjalne procese (Prete 2023). Kriteriji dopuščajo možnost primerjave (npr. njegova pesem je izvirnejša od moje), vendar pa navadno primerjamo le stvari, ki so iste narave. Primerjava lahko vodi v tekmovalnost, tekmovalnost pa vzbuja občutek ogroženosti. Toda tekma ‚človek – stroj‘ ni smiselna, saj ne pripadata isti naravi.

Še pred določanjem kriterijev ustvarjalnosti pa se pojavi vprašanje, zakaj o ustvarjalnosti sploh lahko govorimo. Od kod človeku (z)možnost ustvarjalnosti in njeno uresničevanje? Poglejmo, kako na ta vprašanja odgovarja teologija.

3. Stvarnikov ustvarjalni način: Beseda, Dih in svoboda

V Svetem pismu pojma ustvarjalnost ne najdemo, vendar o njej lahko nekaj povemo – začenši s pripovedjo o stvarjenju sveta (1 Mz 1,1–2,3) (Berdjajev 1994, 35–185). Znotraj teologije se ustvarjalnost povezuje s Stvarnikom, ustvarjalni proces pa s Stvarnikovim načinom ustvarjanja. Ideja ustvarjalnosti je torej možna le zato, ker obstaja Stvarnik, ki je vršil izvirno ustvarjalno dejanje, ki je ustvarilo nekaj, česar prej ni bilo. Stvarnik je tisti, ki stvari prikličje v bivanje iz nič – *ex nihilo sui et subiecti* (Janez Pavel II. 1999, 10). V nadaljevanju bomo videli, da to dela z Besedo in Dihom – v svobodi. Odsotnost tega ustvarjalnega dejanja bi svetu dopustila zgolj možnost prerazporeditve danega (Dell'Asta 1994, 44).

Na podlagi tega, da je bil človek ustvarjen po Stvarnikovi podobi in podobnosti (1 Mz 1,26–27), postane zmožen ustvarjati – in to prav po podobi, ki človeka s tem, da je v celoti dar, neskončno povzdigne nad to, kar bi lahko črpal le iz sebe. Tako lahko že vnaprej predvidimo, da se sama zmožnost človeške ustvarjalnosti, ki je od stvarjenja sveta postavljena kot temeljna opredelitev osebe, ustvarjene po Božji podobi in podobnosti, hkrati vendarle korenito razlikuje od Božjega ustvarjalnega dejanja. Ustvarjenost človeka po Božji podobi in podobnosti namreč pomeni, da je tudi njegovo ustvarjalno dejanje v podobi in podobnosti Božjega ustvarjalnega dejanja, torej mu ni enako (Dell'Asta 1994, 44–47).

Med Stvarnikom in ustvarjenim bitjem namreč obstaja neskončna razdalja, kakor pravi Nikolaj Kuzanski: »Ustvarjalna umetnost – duša ima srečo, da jo lahko goji – ni istovetna s tisto bistveno umetnostjo, ki je Bog, temveč je le povezava z njo in deležnost pri njej.« (1967, 332) Toda kljub tej razdalji je ustvarjalno načelo in sodelovanje s Stvarnikom zaupano slehernemu človeku. Nagovor Stvarnika, ki je izrecno naslovljen zgolj na človeka, je resničnost, na katero človek svobodno odgovarja z ustvarjalnostjo. Človek pa je lahko ustvarjalen le, ko je deležen Stvarnikove ustvarjalne moči. Zato je poznavanje Stvarnikovega ustvarjalnega načina, ki je v svojem bistvu odnose razsežnosti (Stvarnik ne ustvarja sam, temveč pri njegovi ustvarjalni dejavnosti kot protagonista nastopata Beseda in Dih), bistvenega pomena.

3.1 Beseda in Dih

Prva Mojzesova knjiga (1,1–2,3) se začne z veličastno pripovedjo, stvarjenjsko himno, ki pred našimi očmi riše čudovito sliko, podobo našega sveta. Edini protagonist te zgodbe, Umetnik te slike je Bog (Janez Pavel II. 1999, 24). Toda na koncu stvarjenjske pripovedi, ko Bog počiva, nam ostane le ,slika', na kateri Umetnika ne vidimo, toda slišimo njegov glas – njegov ,glas' so pravzaprav ustvarjene stvari: »Nebesa pripovedujejo o Božji slavi, nebesni svod sporoča o delu njegovih rok.« (Ps 19,2)

Stvarnikov glas stvari dejansko priključuje v bivanje, z njegovimi besedami nastane svet: »Zakaj dejal je in je nastalo, ukazal je in je stalo.« (33,9) Bog ustvarja z besedo, z govorjenjem. Človek govori o stvareh, ki obstajajo – na začetku pa so bile stvari, ki obstajajo, »izgovorjene« (Beauchamp 1992, 14). Ta Stvarnikov način jasno razodeva, da si bo vse, kar sledi, prizadevalo za dialog, ki je temelj zaveze, odnosa. »Beseda je namreč izrečena, ker je namenjena nekemu, predpostavlja nekoga, ki posluša in ki se bo nekoč lahko svobodno odzval. A bitja, ustvarjena pred človekom, se ne odzovejo: beseda je namenjena drugemu – počakati bomo morali na šesti dan. Tedaj Bog ni le »rekel« (1 Mz 1,3.6.9.11.14.20.24.26), ampak »jima je rekel« (1 Mz 1,28).« (Rocca 2022, 18–19)

Z Besedo se je torej vse začelo: »V začetku je bila Beseda in Beseda je bila pri Bogu in Beseda je bila Bog. Ta je bila v začetku pri Bogu. Vse je nastalo po njej in brez nje ni nastalo nič, kar je nastalo.« (Jn 1,1-3) Vse se torej začne, ko Stvarnik spregovori, vendar ko govori, hkrati tudi diha: »Z Gospodovo besedo so bila narajena nebesa, z dihom njegovih ust vsa njihova vojska.« (Ps 33,6) Pri stvarjenju nastopata dva protagonista, kajti Stvarnik ustvarja v odnosu z duhom »duh Božji je vel nad vodami« (1 Mz 1,2b) in besedo: »Bog je rekel.« (1,3a) Ni besede brez duha: Božji duh² se artikulira v besedi (Rocca 2022, 27–28).

Duh in beseda se dopolnjujeta in ne prenehata delovati skupaj, kar je dobro vidno pri prerokih: če preroka navdihuje duh, beseda priključuje dialog (Guillet 1976, 1236). Dih vodi h govoru, in to se razodeva že v naravi človeka: človek lahko govori samo z dihanjem. Otrok diha, vendar še ne govori. Šele postopoma se bo naučil artikulirati dih v zvoke – in nato nekega dne v govor (Rocca 2022, 28). Dihanje in govorjenje sta sicer dve različni dejanji, vendar združeni. Irenej Lyonski v tem smislu govori o dveh rokah, s katerima je Bog Oče vedno deloval. Kdor razodeva Očeta, je diada Sina in Duha, vedno združena: Sin, na katerem počiva Duh (Bulgakov 1971, 284–332).

Stvarnik dihne, spregovori ter z dihom in besedo dá življenje celemu svetu. In to iz ničesar, saj pred tem ni bilo ničesar – razen njegove svobode. »Ko v nepopolnosti našega človeškega jezika govorimo o ustvarjanju iz nič, v resnici govorimo o ustvarjanju iz svobode.« (Berdjajev 1994, 187) Ves svet je nastal iz Božje svobode, torej je v Stvarnikov ustvarjalni način vključena tudi svoboda. Poglejmo, kako je ustvarjalnost povezana s svobodo.

3.2 Svoboda

Resnična ustvarjalnost je možna samo v svobodi. Ustvarjalnost je izraz svobode v njenem najplemenitejšem smislu. Svoboda in ustvarjalnost sta neločljivo povezani, saj ena brez druge ne moreta obstajati. Svoboda brez ustvarjalnosti postane pra-

² Cantalamessa besedi Božji Duh, Dih (*ruah*) pripisuje dva osnovna pomena: »*ruah* označuje atmosferski prostor med nebom in zemljo, prostor, ki je lahko miren ali vznemirljen, odprt prostor. Izraz se nato razširi v pomen življenjskega prostora, v katerem se človek giblje in diha.« (2010, 1333) Duh je torej veter in dih. Lahko je nekaj zunanjega, izmuzljivega (Apd 2,2) – podoba nečesa, kar ne pripada stvarstvu –, obenem pa predstavlja nekaj notranjega, človeku intimnega (Jn 20,22). Tako je Duh življenje stvarstva, ga oživlja, vendar se hkrati od njega razlikuje (Rocca 2022, 28).

zna in brez cilja, ustvarjalnost brez svobode pa zgolj mehanična reprodukcija obstoječega (Berdjajev 1994, 191). Ustvarjalnost ne pomeni preproste razporeditve že obstoječih elementov (Dell'Asta 1994, 44), temveč je dejanje popolne novosti, ki izhaja iz svobode. Ustvarjalno dejanje ni zgolj proizvod zunanjih okoliščin, temveč izvira iz tiste globoke notranje moči, ki jo svoboda omogoča: »Ustvarjalnost izhaja iz notranjosti, iz globokih in nedoumljivih globin, in ne iz zunanjih okoliščin ali nujnosti sveta. Sam poskus ustvarjalno dejanje razumeti in mu najti temelj je že znamenje nerazumevanja.« (Berdjajev 1994, 188) Zato vsak poskus racionalizacije svobode in ustvarjalnosti s pravili in normami vodi v njuno osiromašenje (188–189).

Svoboda, kot jo razume Berdjajev, je brezmejna, brezpogojna in neizmerna – in podobno kot ustvarjalnost jo zaznamuje skrivnostna in nerazložljiva narava: »Ustvarjalnost je nepojasnljiva. Ustvarjalnost je skrivnost. Skrivnost ustvarjalnosti je hkrati skrivnost svobode. Skrivnost svobode je brezmejna in nedoumljiva, kot brezno. In prav tako brezmejna in nedoumljiva je skrivnost ustvarjalnosti.« (Berdjajev, 1994, 188)

Svoboda, razumljena brez cilja in vsebine, predstavlja praznino in nebivanje:

»Negativna svoboda, svoboda kot samovolja, je svoboda brez vsebine, prazna. Hoteti svobodo zaradi svobode, svobodo brez cilja in vsebine, pomeni hoteti praznino, težiti k nebivanju. Svoboda, razumljena zgolj formalno, brez cilja in vsebine, je nič, praznina, nebivanje. Svoboda, ki se je pojavila ob izvirnem grehu, je bila prav takšna svoboda – negativna, formalna, prazna svoboda, svoboda zaradi svobode, torej svoboda ,od', in ne svoboda ,za'. Svoboda izvirnega greha ni bila ustvarjalna svoboda. Odmik od Boga svobodo osvobodi njene vsebine in cilja, jo osiromaši in oslabi.« (Berdjajev 1994, 191)

Greh torej prekriva ustvarjalno skrivnost človeškega bitja. S padcem v greh človek oslabi svoje ustvarjalne moči (Berdjajev 1994, 135–136), saj se pred Stvarnikom »skriva« (1 Mz 3,8), beži pred tistim, ki mu (z)možnost biti ustvarjalen sploh daje. Tako je »avtentična ustvarjalnost mogoča le preko odrešenja« (Dell'Asta 1994, 47), ki človeku omogoči vzpostavitev odnosa in pozitivne svobode – ta pa ni praznina, temveč je polna vsebine. Njena vsebina je ljubezen, s katero je povezano ustvarjalno načelo, saj je ljubezen možna le v svobodi. Ljubezen prihaja od Boga, ki je ljubezen (1 Jn 4,7–8), in je tista, ki v ta svet prinaša nekaj resnično novega, nekaj, kar Pridigarjeve besede »Nič ni novega pod soncem.« (Prd 1,9) spreminja. Ljubezen preobraža svobodo ,od' v svobodo ,za'. »Bog je namreč svet tako vzljubil, da je dal svojega edinorojenega Sina« (Jn 3,16), da bi človeka odrešil vezi greha in mu odprl pot obnavljanja svoje podobnosti z Bogom. Človek, ki božjo ljubezen sprejme, se lahko svobodno odloči ,za' odnos z D/drugim, prav odnos pa je tisti, ki ustvarjalnost sploh omogoča.

K pozitivni svobodi in ustvarjalnosti je poklican vsak človek. Z ustvarjanjem namreč človek vstopa v odnos z drugim in prav ta odnos, ki temelji na svobodni pri-

padnosti in ljubezni, oslabi moč greha: skrivnost odrešenja vključuje tudi skrivnost ustvarjalnosti (Berdjajev 1994, 135). Ustvarjalnost človeku pomaga, da izstopi iz osamljenosti greha. Tako svoboda in ustvarjalnost skupaj predstavljata temelj človekovega bivanja in odgovor na Stvarnikov nagovor.

4. Ustvarjalnost kot človekov odgovor na Stvarnikov nagovor

Svet iz Božjih rok prihaja kot osnutek. »Svet je pravzaprav v postajanju, Bog ga ne prestando, predvsem pa ustvarjalno dograjuje oziroma izpopolnjuje.« (Zabukovec 2012, 439) Celotno stvarstvo še vedno »zdihuje in trpi porodne bolečine« (Rim 8,22), kar pomeni, da še ni povsem 'rojeno', človeku pa je zaupana naloga 'porodne babice'. Bog je ustvaril popoln svet prav zato, ker je svet 'nepopoln'. V tem je lepota sveta: da je nedokončan (Gesché 1996, 81–99). Stvarnik ga namreč ne želi dokončati brez človeka, ki mu je zaupana naloga gospodovanja zemlji (1 Mz 1,28). Nekoliko pozneje bo enako veljalo za vrt, ki ga je Bog zasadil in nato človeku zaupal, da bi ga ta »obdeloval in varoval« (2,15). »Prejšnje dni je Jahve nekako določil ritem kozmičnega razvoja in ustvaril vesolje. Nazadnje je ustvaril človeka, najplemenitejši sad svoje zamisli, in mu podvrget vidni svet kot prostrano polje, na katerem naj se izraža človekova domiselnost« (Janez Pavel II. 1999, 10) oziroma ustvarjalnost.

Začetne strani Svetega pisma prikazujejo, kako človek odseva podobo Boga Stvarnika. Božja podoba je človeku vtisnjena kot začetni in večni dar, človek pa je v svojem življenju – združen z Duhom – poklican, da podobnost s Stvarnikom, Umetnikom izoblikuje tudi sam. Toda kdaj se ustvarjalnost kot človekov odgovor na Stvarnikov nagovor uresničuje? Videli smo, da je temelj ustvarjalnosti odnos, dialog (Vodičar 2023, 356). Stvarnik, ki skupaj z Besedo in Dihom ustvarja iz svobode, v fenomen ustvarjalnosti vtisne pečat odnosnosti. Toda človek z grehom odnos in dialog s svojim Stvarnikom prekine. S posledicami greha in kot žrtev nujnosti mora človek vstopiti v skrivnost odrešenja in prav v njej lahko svojo podobnost z Bogom obnovi in si izgubljeno svobodo povrne. »Ustvarjalna skrivnost bivanja je zaradi greha skrita. S padcem v greh človek svoje ustvarjalne moči oslabi. Po Kristusu pa je človeška narava odrešena in obnovljena ter rešena prekletstva greha.« (Berdjajev 1994, 135–136) Tako se ustvarjalnost kot človekov odgovor na Stvarnikov nagovor v polnosti uresniči v Kristusu – človeka je namreč osvobodil 'za' svobodo (Gal 5,1). Tako ustvarjalnost ne pomeni ustvarjanja novih oblik, ki bi prinesle občutek spremembe, temveč sprejetje novosti bivanja, ki ga je razodel Kristus. V njem človeštvo Stvarniku odgovarja z ustvarjalnostjo, saj z njo »pomaga k uveljavitvi tiste resnične lepote, ki bo kot preblisk Božjega Duha preobrazila materijo in odprla človeško dušo, da bo začutila večnost« (Janez Pavel II. 1999, 25).

5. Dualistična shema ali sodelovalni odnos?

Teološki okvir, v katerega smo fenomen ustvarjalnosti postavili, nam pomaga razumeti, kako je človeška ustvarjalnost povezana s Stvarnikom, zaradi katerega o (z)možnosti ustvarjalnosti sploh lahko govorimo. V tem smislu je ustvarjalnost, ki je vezana na odnos, človeku zaupana kot njegovo poslanstvo na tem svetu. Bog namreč šesti dan stvarjenja s svojo besedo nagovori človeka (1 Mz 1,28) in čaka, da mu v svobodi odgovori. Pri tem pa svoje mesto najde tudi UI. Kako?

Besedišče, s katerim govorimo o UI, je pogosto nagnjeno k antropomorfizaciji in biologizaciji delovanja stroja – tako govorimo o ‚inteligenci‘, ‚nevronske mreže‘, ‚ustvarjalnosti‘ itd. (Prete 2023). Takšen način, ki vključuje tudi omenjene kriterije, omogoča medsebojno primerjavo: umetno inteligenco se primerja s človeško, ustvarjalnost računalniških algoritmov s človeško ustvarjalnostjo. Kot že rečeno, pa navadno primerjamo le stvari, ki so iste narave (npr. korenje je boljše od cvetače, stric je starejši od očeta), in ne različnih (npr. avto je hitrejši od želve). Primerjave vodijo k tekmovalnosti, tekmovalnost pa vzbuja občutek ogroženosti. Toda tekma ‚človek – stroj‘ ni smiselna, saj ne pripadata isti naravi. Vzorec, ki pri tem nastaja, je vedno opozicijski: odnos ‚človek – stroj‘ je »v nasprotnem razmerju, eno proti drugemu, kot da sta človek in stroj dva ločena in antagonistična subjekta« (2023).

Ta antagonistični dualizem, ki se razvija, da bi človeštvo spodkopal in nadomestil, dobro predstavlja ikonografija, ki se pogosto uporablja za ponazoritev in objavo številnih prispevkov, konferenc in publikacij o UI: dve roki, ena človeška, druga robotska, njuni prsti pa se skoraj dotikajo. Gre za imitacijo freske v Sikstinski kapeli, ko se pri stvarjenju Adama vžge iskra življenja, ki ponovno potrjuje položaj človeka kot ‚stvarnika‘ in položaj stroja kot ‚bitja‘, ki je končno prejelo življenje, da je lahko postalo avtonomno (2023). Tako imenovana umetna inteligenca – pravi Cosimo Accoto – »ni nikoli takšna ‚sama po sebi‘ in ‚zase‘ (le instrumentalni artefakt), temveč vedno ‚z drugimi‘ (sociotehnična sestava) in ‚za druge‘ (sociomorfna konstrukcija)« (2023). Ne gre torej za proces mistifikacije, temveč za dejstvo, da so te nove tehnologije v resnici zgolj nova raven našega tehničnega upravljanja sveta (Prete 2023).

Tako kot UI kot inteligenca ni avtonomna – odvisna je od človeške inteligence in brez nje ne more obstajati –, tako UI tudi ne moremo pripisati ustvarjalnosti v popolni neodvisnosti od človeka (Šegula 2021, 921). Brez človeka UI ‚pogojev‘ za ustvarjalnost ne izpolnjuje, saj pri njej ne moremo govoriti o udeležbi pri Stvarnikovi ustvarjalni moči – to je mogoče le skozi svoboden odnos ljubezni. Pri UI gre bolj za razporeditev že obstoječih elementov (Dell'Asta 1994, 44), kar v vsakdanjem jeziku pogosto zamenjujemo za ustvarjalnost. Kljub temu je UI orodje, ki ustvarjalni dejavnosti lahko služi. UI kot ustvarjalno orodje pa lahko deluje v skladu s svojo pravo funkcijo le takrat, ko jo kdo uporablja – in ta kdo je človek (Žalec 2023, 814–815), ustvarjen po Božji podobi in podobnosti. Človek je namreč svoje bivanje od nekdaj razumel kot interakcijo z okoljem, ki poteka z orodji, tehnološkimi artefakti (Benanti 2022, 14–15).

Namesto vprašanj, ki smo si jih zastavili na začetku – o vrednosti, ogroženosti in pomenu človeške ustvarjalnosti v dobi UI –, bi bilo morda bolj smiselno vprašanje, kako je človek lahko ustvarjalen skupaj z UI in kako lahko delovanje računalniških algoritmov vključi v odgovor Stvarniku.

6. Zaključek

Vsaka znanost k razumevanju fenomena ustvarjalnosti prispeva svojo perspektivo, kar kaže, kako raznolik, večdimenzionalen in kompleksen je ta fenomen in hkrati za človekov življenjski svet bistven. Videli smo, da zgolj upoštevanje kriterijev empiričnih znanosti, s katerimi bi fenomen ustvarjalnosti presojali in vrednotili, lahko vodi v primerjavo človeške ustvarjalnosti z ustvarjalnostjo UI. Kljub temu, da UI rezultate človeške ustvarjalnosti že reproducira in simulira ustvarjalne procese, se človeška ustvarjalnost uresničuje na povsem drugačen način. Človek, ustvarjen po Božji podobi, je namreč deležen Stvarnikove ustvarjalne moči, za katero sta značilni svoboda in odnosa razsežnost. V tem smislu bi o ustvarjalnosti UI težko govorili, saj njeno delovanje, ki temelji na matematičnih modelih in pravilih, izhaja iz nujnosti.

Človek je torej tisti, ki je po svoji naravi »ustvarjalec, graditelj, oblikovalec življenja in žeje po ustvarjalnosti, ki ga žge, ni mogoče pogasiti« (Berdjajev 1994, 3). Na nagovor Stvarnika – Umetnika tudi človek v Kristusu odgovarja z ustvarjalnostjo: »Slehernemu človeku je zaupana naloga, da oblikuje svoje življenje: v nekem smislu mora iz njega napraviti umetniško delo, mojstrovino.« (Janez Pavel II. 1999, 11). UI mu pri tem lahko služi kot orodje.

Kratika

UI – Umetna inteligenca.

Reference

- Accoto, Cosimo.** 2023. La Potenza della Latenza: tre studi sull'intelligenza artificiale generativa. Cosimo Accoto. <https://cosimoaccoto.com/2023/04/28/la-potenza-della-latenza-tre-studi-sullintelligenza-artificiale-generativa-accoto-2023/> (pridobljeno 15. 8. 2024).
- Beauchamp, Paul.** 1992. *All'inizio, Dio parla: Itinerari biblici*. Padova: Apostolato della Preghiera.
- Benanti, Paolo.** 2022. *Human in the Loop: Decisioni umane e intelligenze artificiali*. Milano: Mondadori.
- Berdjajev, Nikolaj.** 1994. *Il senso della creazione: Saggio per una giustificazione dell'uomo*. Milano: Jaca Book.
- Bulgakov, Sergej.** 1971. *Il Paraclito*. Bologna: Dehoniane.
- Boden, Margaret A.** 1998. Creativity and Artificial Intelligence. *Artificial Intelligence* 103:347–356.
- Cantalamesa, Raniero.** 2010. Spirito Santo. V: Romano Penna, Giacomo Perego in Gianfranco Ravasi, ur. *Temi teologici della Bibbia*. Cinisello Balsamo: Dizionari San Paolo.
- Dell'Asta, Adriano.** 1994. L'apocalisse creativa dell'uomo. V: *Il senso della creazione: Saggio per una giustificazione dell'uomo*. Milano: Jaca

Book.

- Gesché, Adolphe.** 1996. *Dio per pensare*. Zv. 2, *L'Uomo*. Milano (Cinisello Balsamo): Edizioni San Paolo.
- Guillet, Jacques.** 1976. Spirito di Dio. V: Xavier Léon-Dofour, ur. *Dizionario di teologia biblica*. Torino: Dehoniane.
- Janez Pavel II.** 1999. Pismo umetnikom. Ljubljana: Družina.
- Kaufman, James C.** 2009. *Creativity 101: The Psychology Series*. New York: Springer.
- Kronfeldner, Maria E.** 2009. Creativity Naturalized. *The Philosophical Quarterly* 59, št. 237: 577–592.
- Kuzanski, Nikolaj.** 1967. *Dialogus de ludo globi*. V: *Philosophisch–Theologische Schriften*. Dunaj, 332. Navaja Janez Pavel II., *Pismo umetnikom*, 10, op. 1. Ljubljana: Družina.
- McMahon, Darrin M.** 2013. *Divine Fury: A History of Genius*. New York: Basic Books.
- Murray, Penelope.** 1989. *Genius: The History of an Idea*. New York: B. Blackwell.
- Paul, Elliot Samuel, in Dustin Stokes.** 2024. Creativity. The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/entries/creativity/#WhatCrea> (pridobljeno 16. 8. 2024).
- Prete, Marisa.** 2023. Intelligenza Artificiale e Creatività: Oltre la ferita narcisistica. *Academia*. https://www.academia.edu/106646924/Intelligenza_Artificiale_e_Creatività_Oltre_la_ferita_narcisistica (pridobljeno 15. 8. 2024).
- Rocca, Paolo.** 2021. *Dell'Altro: Tra parola e silenzio nel Vangelo di Marco*. Milano (Cinisello Balsamo): San Paolo.
- Stokes, Dustin R.** 2011. Minimally Creative Thought: Minimally Creative Thought. *Mataphilosophy* 42, št. 5: 658–681.
- Šegula, Andrej.** 2021. Upanje in zaupanje v kontekstu pastoralne teologije v času globalnega nezaupanja. *Bogoslovni vestnik* 81, št. 4: 917–924.
- Vodičar, Janez.** 2023. The Critique of Secularism by Talal Asad as a Chance to Look for New Ways of Proclamation. *Verbum Vitae* 42, št. 2: 345–359.
- Zabukovec, Urša.** 2012. Filozofija skozi razodetje: Semjon Frank 1877–1950. Spremnna beseda. V: Frank, Semjon. *Nedoumljivo: Ontološki uvod v filozofijo religije*. Ljubljana: Kud Logos.
- Žalec, Bojan.** 2023. Ali je umetna inteligenca inteligenca v pravem pomenu besede? Vprašanje psihičnih značilnosti in splošnosti. *Bogoslovni vestnik* 83, št. 4: 813–823.

*Roman Globokar***Pogled Katoliške Cerkve na razvoj in uporabo umetne inteligence***The Catholic Church's View on the Development and Use of Artificial Intelligence*

Povzetek: Krščanska etika s svojimi antropološkimi in teološkimi predpostavkami vstopa v širšo razpravo o etičnih dilemah umetne inteligence. Članek predstavi vsebino treh uradnih dokumentov Katoliške Cerkve o umetni inteligenci, in sicer „Rimskega poziva za etiko umetne inteligence“ iz leta 2020 in dveh poslanic papeža Frančiška iz leta 2024. Vsebinsko dokumentov osvetli skozi tri kriterije, ki opredeljujejo specifičen pomen krščanske vere za etiko: integracijo, spodbudo in kritično presojo. Ugotavlja, da so besedila namenjena vsem ljudem, ne samo katoličanom, in da *integrirajo* najsodobnejša spoznanja in etične razprave glede razvoja in uporabe umetne inteligence. Na podlagi biblične antropologije *spodbuja* razvoj, ki bo zagotavljal spoštovanje dostojanstva vsake človeške osebe, enakopravnost vseh ljudi v družbi, večjo globalno pravičnost in medsebojno povezanost človeštva ter spoštovanje naravnega okolja. Kot posebno merilo etičnosti dokumenti navajajo pozornost do ubogih in ranljivejših članov družbe. Vera v Boga Stvarnika preprečuje absolutizacijo pridobitev človeškega uma in omogoča *kritično presojo* umetne inteligence. Članek, v sozvočju z dokumenti Katoliške Cerkve, spodbuja k oblikovanju etičnih smernic za razvoj in uporabo umetne inteligence. Vendar je avtor članka prepričan, da zgolj pravila ne bodo dovolj, zato poudarja prizadevanje za oblikovanje temeljnih etičnih drž, predvsem na področju vzgoje in izobraževanja. Na podlagi vsebine družbenega nauka Cerkve predstavi osem kreposti za etično usmerjanje nadaljnjega razvoja umetne inteligence.

Ključne besede: umetna inteligenca, papež Frančišek, dostojanstvo človeške osebe, vesoljno bratstvo, trajnostni razvoj, kritično presojanje, kreposti

Abstract: Christian ethics, with its anthropological and theological presuppositions, enters into the wider debate on the ethical dilemmas of artificial intelligence. This article presents the content of three official documents of the Catholic Church on artificial intelligence, namely the 2020 “Roman Call for an Ethics of Artificial Intelligence” and the two 2024 Messages of Pope Francis. It illuminates the content of the documents through three criteria that define the

specific relevance of the Christian faith for ethics: integration, stimulation, and critical judgement. The article notes that the texts are addressed to all people, not just the Catholics and that they *integrate* the most up-to-date insights and ethical debates on the development and use of artificial intelligence. Based on biblical anthropology, they *stimulate* development that will ensure respect for the dignity of every human person, the equality of all people in society, greater global justice and interconnectedness of humanity, and respect for the natural environment. As a specific criterion of ethicality, the documents cite attention to the poor and the more vulnerable members of society. Faith in a creator God prevents the absolutization of the acquisitions of the human mind and empowers a *critical assessment* of artificial intelligence. The article, in harmony with the documents of the Catholic Church, encourages the development of ethical guidelines for the development and use of artificial intelligence. However, the author of the article is convinced that rules alone will not be enough, and therefore also stresses the need to strive for the development of fundamental ethical attitudes, especially in the field of education and training. Based on the content of the Church's social teaching, he presents eight virtues to ethically guide the further development of AI.

Keywords: artificial intelligence, Pope Francis, dignity of the human person, universal brotherhood, sustainable development, critical judgement, virtues

1. Prispevek krščanske etike k splošni razpravi o umetni inteligenci

Skokovit razvoj umetne inteligence (UI) odpira številna temeljna antropološka, metafizična in teološka vprašanja, zlasti vprašanje, kdo je človek, kakšna je posebnost človeka in kakšno prihodnost za človeka si pravzaprav želimo.¹

Krščanska etika s svojimi antropološkimi in teološkimi predpostavkami vstopa v splošno razpravo o etični uporabi UI. Kakšno vlogo ima krščanska etika v sodobnem pluralnem svetu? V zadnjih 50 letih je na področju teološke etike potekala precej vroča razprava o tem, kaj je posebnost krščanske etike – tako imenovani *proprium*. Med moralnimi teologi se je vnela razprava ali je krščanska etika zgolj etika za določeno versko skupnost ali naj bi imela njena vsebina splošno veljavo. (Zuccaro 2013, 129–160; Globokar 2010)

V veliki meri delimo pogled zagovornikov moralne avtonomije v kontekstu vere. Njihova glavna teza je, da so moralne norme univerzalne in da je vsebina moralnega nauka enaka za kristjane in nekristjane. Pot do moralne resnice vodi preko razuma in tudi kristjan mora veljavnost svojih moralnih stališč dokazovati z razumom in se ne more opirati na zunanjo avtoriteto. Zaradi tega je krščanska etika v

¹ Prispevek je rezultat raziskovanja znotraj projekta J6-4626 „Teologija, digitalna kultura in izzivi na človeka osrediščene umetne inteligence“ ter raziskovalnega programa P6-0269 „Religija, etika, edukacija in izzivi sodobne družbe“, ki delujeta na Teološki fakulteti Univerze v Ljubljani in ju sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

dialogu z vsemi drugimi sogovorniki v družbi in skupaj z njimi išče odgovore na številne izzive. Kritično moramo sicer ugotoviti, da so bili začetniki avtonomne morale v kontekstu vere (Auer, Fuchs, Schüller) premalo pozorni na večplastno pogojenost razumskega razmišljanja. Vsako razmišljanje je namreč vedno nujno pogojeno z določenimi antropološkimi, kulturnimi, verskimi in metafizičnimi predpostavkami (Demmer 2014, 17–33). Prav zato na področju etičnih vprašanj tudi prihaja do nesoglasij med različnimi akterji znotraj pluralne družbe.

Alfons Auer, eden vodilnih predstavnikov avtonomne morale v kontekstu vere, je posebnost krščanske etike videl v »novem obzorju smisla« (Auer 1984, 177), ki ga daje Jezus Kristus. Kristjan razume avtonomno moralno zahtevo kot klic Boga, kot del stvarjenjskega načrta, kot poslanstvo, ki mu ga daje Kristus. Kristjan ne dojema moralne avtonomije na absoluten način, ampak jo razume kot ‚odnosno avtonomijo‘. Čeprav je odkrivanje moralnih resnic stvar avtonomnega (razumnega) procesa, pa se ta proces dogaja znotraj odnosa z Bogom. Na Boga pa se ne gleda kot na zunanjo avtoriteto, ampak kot na svoje najgloblje jedro. Bog torej nagovarja človeka od znotraj, zato ne gre za heteronomijo, ampak za ‚teonomno avtonomijo‘ (Demmer 1999, 120–122). Vera daje kristjanu novo spodbudo, da prevzame svojo osebno odgovornost. »Krščanski proprium na področju moralnosti torej ni v novih normah ravnanja, ki bi bile dostopne samo vernemu, ampak v integraciji naravnomooralnega (avtonomnega) delovanja v njegovo religiozno povezanost z Bogom.« (Auer 1984, 178)

Konkretno Auer identificira tri nenadomestljive vloge, ki jih ima vera v odnosu do avtonomnega razuma na moralnem področju: *integracija*, *spodbuda* in *kritika* (189–197). Cerkev in teologija morata biti naprej odprta za *integracijo* vsega dobrega, kar se razvija v aktualnem družbenem okolju. Vera spodbuja k odprtemu dialogu z etičnimi razmišljanji izven cerkvenih skupnosti in integraciji etičnih misli, ki spodbujajo človekovo dostojanstvo in pravično ureditev odnosov v družbi. Nadalje vera navdušuje in *spodbuja* človeka za moralno delovanje. Auer navaja, kako so določene verske resnice (npr. dostojanstvo in enakost ljudi pred Bogom) spodbudile spremembe v etičnem razmišljanju in delovanju (npr. glede suženjstva, temeljnih človekovih pravic, spoštovanja žensk). Prepričan je, da lahko krščanska vera tudi v prihodnje stimulira razvoj občega etosa k večji polnosti. Končno pa ima vera vlogo *kritičnega* presojanja. Kristjanu daje pogum, da se zoperstavi slepemu sledenju moralnim trendom v družbi (Vodičar 2023, 1045). Posamezne moralne norme in prepričanja presoja v luči krščanske antropologije in se je sposoben kritično odzvati na kršitve človekovega dostojanstva in nepravilnih razmer v družbi. »Cerkev in teologija ne kritizirata posvetne moralnosti kot take, ampak samo njena naivno utopična in tudi brezupno pesimistična izkrivljenja.« (Auer 1984, 194)

Trojna vloga vere in teologije na področju moralnega razmišljanja – integracija, spodbuda in kritika – se nam zdi aktualna tudi pri razmišljanju o razvoju in uporabi UI. Načelno gre za odprtost in dialoško držo do vseh akterjev v družbi pri razmišljanju, kako uporabiti vse prednosti, ki jih prinaša UI. Krščanska antropologija nas spodbuja h kreativni uporabi vseh sredstev, ki so človeku na voljo, saj jih najprej razume kot del stvarjenjskega procesa, ki še vedno traja (*creatio continua*).

Hkrati pa se zavedamo, da imajo vse zemeljske stvaritve svoje pasti in lahko predstavljajo nevarnost za uresničeno človeško življenje, zato nas vera opogumlja, da kritično ovrednotimo najnovejše tehnološke pridobitve in preverimo, v kakšni meri predstavljajo blagor za človeka kot posameznika, za človeško družbo in za celotno naravno okolje.

Zavedamo se, da vera nikakor ne nadomešča razuma in da tudi cerkveno učiteljstvo, ki znotraj Katoliške Cerkve oblikuje uradna stališča glede določenih moralnih vprašanj, nima dokončnih odgovorov na sodobne izzive v družbi. Kljub vsemu pa želimo v skladu s tradicijo krščanske etike razširiti obzorja razmišljanja in razpravo usmeriti onkraj pragmatičnih in normativnih vprašanj glede uporabe UI v družbi. Prepričani smo, da je razprava na meta-etični ravni izredno pomembna in nujna, preden pride do oblikovanja etičnih smernic in zakonskih ureditev o uporabi UI. Vprašati se moramo, kaj pomeni izpolnjeno človeško življenje in kako ga lahko spodbujajo z novimi tehnologijami ter kakšno družbo si želimo v prihodnosti. Malteški teolog Emmanuel Agius poudarja: »Tehnologije ne smemo presojati po tem, ali lahko nekaj naredimo hitreje, ceneje ali lažje, temveč po njenem vplivu na ljudi kot posameznike in skupnosti ter na okolje. Pri teh prizadevanjih je vloga razuma in verskih skupnosti, da osvetlijo cilje znanosti in tehnologije ter kritično ocenijo sredstva, uporabljena za njihovo doseganje.« (2020, 117) Naloga krščanske etike je, da postavlja prava vprašanja, spodbuja etični razvoj novih tehnologij in opozarja na morebitne stranpote. Gre za zgoraj omenjeno trojno hermenevitično vlogo, ki jo ima krščanska teologija na področju širšega družbenega razmišljanja o etičnosti uporabe UI. »Ta hermenevitični pristop širi obzorje tehnološke racionalnosti in tako vodi onkraj instrumentalnih in redukcionističnih modelov v etiko odgovornosti.« (118) Krščanska etika si prizadeva za kritično razločevanje, ki preverja upravičenost namenov in motivov določenih novih tehnologij, razkriva njihove temeljne vrednote in cilje ter skuša predvideti daljnosežne posledice za ljudi in širšo družbo.

Pri uporabi UI se odpirajo številne etične dileme. Naj navedemo samo nekatere. Sistemi UI se lahko učijo iz pristranskih podatkov, kar vodi v napačne zaključke in možnost diskriminacije zaradi rase, spola, ekonomskega položaja, starosti. To ima lahko nevarne posledice za številna področja, od izobraževanja, zdravstva in zavarovanja do zaposlovanja, sodstva in zagotavljanja javnega reda (Strahovnik 2023). Prihaja do vedno več interakcij med človekom in sistemi UI, pri čemer velikokrat človek niti ne ve, da komunicira z UI, npr. pri klepetalnih robotih (Centa Strahovnik 2023). Nadalje se postavljajo vprašanja, do kakšne mere naj se uporablja UI v medicini (Miklavčič 2021). Po eni strani je takojšnja analiza množice podatkov lahko v veliko pomoč pri postavljanju diagnoze, po drugi strani pa UI ne more celostno dojeti enkratnosti situacije posameznega pacienta. Vemo, da so možne zlorabe osebnih informacij, ki jih zbirajo sistemi velikih podatkov za komercialne, medijske, politične in vojaške namene. S tem je povezan tudi prevelik nadzor nad zasebnim življenjem in sledenje vedenjskim vzorcem prebivalcev. Ena največjih dilem je prav gotovo vprašanje transparentnosti pri sprejemanju odločitev in odgovornost za povzročeno škodo pri napačnih odločitvah. Niti največji

strokovnjaki niti razvijalci UI ne morejo popolnoma razumeti njenega delovanja. (Miklavčič 2023) Velik problem so tudi avtorske pravice, saj sistemi UI uporabljajo številna besedila brez privoljenja avtorjev. Pojavljajo se mnoge pobude za oblikovanje etičnih smernic za razvoj in uporabo UI in znanstvene razprave in priročniki na to temo (Spiekermann 2019; Boddington 2023; Grimm, Trost in Zöllner 2024).

Tudi Katoliška Cerkve se je dejavno vključila v razpravo o etični uporabi UI. Papeška akademija za življenje je leta 2020 organizirala 26. generalno skupščino članov, ki je potekala pod naslovom „Dobri algoritmi? Umetna inteligenca: etika, pravo, zdravje“ (Paglia in Pegoraro 2021). Srečanje se je zaključilo z „Rimskim pozivom za etiko UI“, ki ga je Papeška akademija za življenje posredovala v javnost skupaj z deležniki s področja razvoja UI. Letos pa je papež Frančišek izdal dva dokumenta z osrednjo temo uporabe UI. Vse tri dokumente bomo podrobneje obravnavali v naslednjem poglavju. Naj samo omenimo, da je papež Frančišek tudi na srečanju G7 v Apuliji 14. junija 2024 govoril o UI, ki jo je označil kot privlačno in strašljivo orodje. »Umetna inteligenca bi lahko na primer omogočila demokratizacijo dostopa do znanja, eksponentni napredek znanstvenih raziskav, možnost prenosa težkih del na stroje; hkrati pa bi lahko prinesla večjo nepravilnost med razvitimi državami in državami v razvoju, med prevladujočimi in zatiranimi družbenimi razredi ter tako ogrozila možnost ‚kulture srečanja‘, ker bi dala prednost ‚kulturi odmetavanja‘.« (Frančišek 2024b)

Predno se lotimo analize treh glavnih besedil o UI, napisanih s strani avtoritet Katoliške Cerkve, povejmo še, da naš primarni interes ni oblikovanje moralnih norm, še manj konkretnih smernic za uporabo UI. Gotovo je to pomembna naloga, vendar gre naš pogled širše. V svojem razmišljanju gremo onkraj normativne etike in dajemo večji poudarek etiki kreposti, saj smo prepričani, da še tako dobra regulativa ne bo prinesla ustreznih rezultatov brez krepostnih moralnih subjektov.

2. Izjave Katoliške Cerkve o uporabi UI

2.1 „Rimski poziv za etiko UI“ (2020)

Februarja 2020 je Papeška akademija za življenje ponudila v podpis „Rimski poziv za etiko UI“, ki je naletel na pozitiven odmev med proizvajalci, civilno družbo in političnimi akterji. Prvi podpisniki tega poziva so bili Papeška akademija za življenje, Microsoft, IBM, FAO in italijansko ministrstvo za inovacije. Namen poziva ni bil predstaviti etičnih norm za nadaljnji razvoj UI, ampak spodbuditi čut odgovornosti med vsemi deležniki, da bi zagotovili prihodnost, »v kateri bodo digitalne inovacije in tehnološki napredek služili človeški genialnosti in ustvarjalnosti, ne pa njenemu postopnemu nadomeščanju« (Benanti 2021, 129).

V uvodu je poudarjeno, da ima UI ogromen potencial, vendar jo je treba uporabljati v korist ljudi in okolja kot celote. »Novo tehnologijo je treba raziskovati in proizvajati v skladu s merili, ki zagotavljajo, da resnično služi celotni ‚človeški družini‘ (preambula „Splošne deklaracije o človekovih pravicah“), pri čemer je treba

spoštovati prirojeno dostojanstvo vsakega od njenih članov in vsa naravna okolja, ter upoštevati potrebe tistih, ki so najbolj ranljivi.« (Pontifical Academy for Life 2020) Cilj nadaljnjega razvoja ni le zagotoviti, da nihče ne bo izključen, temveč tudi razširiti tista področja svobode, ki bi jih algoritemsko pogojevanje lahko ogrozilo. UI ne bi smeli razvijati le z osredotočenostjo na tehnologijo, »temveč v dobro človeštva in okolja, našega skupnega in deljenega doma ter njegovih človeških prebivalcev, ki so neločljivo povezani« (2020). Že v uvodu v ta dokument najdemo sintezo temeljnih poudarkov stališča Katoliške Cerkve do uporabe UI: varovanje človeškega dostojanstva, skrb za naravno okolje, razvijanje vključujoče družbe in pozornost do najbolj ubogih in ranljivih.

Besedilo je razdeljeno v tri dele: etika, izobraževanje in pravice. Etična refleksija se opira na 1. člen „Splošne deklaracije človekovih pravic“, ki pravi, da se vsi ljudje rodimo svobodni in enaki v dostojanstvu in pravicah. Poziv spodbuja, da je treba ta temeljni pogoj svobode in dostojanstva varovati in zagotavljati tudi pri izdelavi in uporabi sistemov UI. Opozarja predvsem na nevarnost diskriminacije, ki je lahko posledica razvoja sistemov UI. »Sistemi umetne inteligence morajo biti zasnovani, oblikovani in izvedeni tako, da služijo in varujejo ljudi in okolje, v katerem živijo.« (2020) Za zagotavljanje etičnosti mora nadaljnji tehnološki razvoj izpolniti naslednje tri kriterije: 1. vključevati mora vsakega človeka, pri čemer ne sme biti nihče diskriminiran; 2. v središču mora biti dobro človeštva in dobro vsakega človeškega bitja; 3. zavedati se mora svoje vpetosti v ekosistem in svoje odgovornosti, da skrbi za naš skupni dom in ga varuje s trajnostnim pristopom. Poziv tudi opozarja, da morajo uporabniki tehnologij UI vedno vedeti, kdaj so v interakciji s strojem in da se tehnologija ne sme predstavljati kot človek. Še posebej je treba zaščititi najbolj ranljive osebe, kot so otroci in starejši ljudje.

V nadaljevanju dokument navaja, da sistemi UI omogočajo velik napredek v izobraževanju, saj lahko s prilagojenimi pristopi podpirajo in spodbujajo individualni napredek posameznega učenca. Prav tako se potencialno povečuje dostopnost izobraževanja in s tem zmanjšanje diskriminacija med ljudmi. Večja je tudi možnost za vseživljenjsko izobraževanje za širšo populacijo. Nove tehnologije so v veliko pomoč invalidom tako pri učenju kot pri bolj neodvisnem življenju, kar je gotovo pozitivno. Poziv se zavzema za prenovu učnih načrtov na vseh ravneh izobraževanja, da bi vsem omogočili iti v korak z najnovejšimi tehnološkimi pridobitvami. »V izobraževalnem sektorju so potrebne reforme za vzpostavitev visokih in objektivnih standardov, ki lahko izboljšajo rezultate posameznikov. Ti standardi ne bi smeli biti omejeni na razvoj digitalnih spretnosti, temveč bi se morali osredotočiti na zagotavljanje, da lahko vsak posameznik v celoti izrazi svoje sposobnosti in dela v dobro skupnosti, tudi če od tega nima osebnih koristi.« (2020)

V zadnjem delu „Rimski poziv za etiko UI“ izpostavi nekaj elementov za pravno ureditev področja uporabe UI. Če želimo namreč zagotoviti nadaljnji razvoj UI, ki bo v službi dostojanstva vsake osebe, človeštva in celotnega naravnega okolja, potem so po mnenju podpisnikov potrebni tudi predpisi in načela, ki bodo zaščitili pravice posameznikov in naravnega okolja. Dokument zagovarja, da moramo v središče javne razprave postaviti »varstvo človekovih pravic v digitalni dobi«

(2020). Na novo bo treba razmisliti o razsežnostih človekove odgovornosti in novih oblikah pojasnilne dolžnosti. Da bi zagotovili večjo transparentnost, sledljivost in odgovornost, dokument poziva, da bi bila ‚algor-etika‘ vključena že od vsega začetka v razvoj vsakega algoritma. »Oblikovanje in načrtovanje sistemov umetne inteligence, ki jim lahko zaupamo, vključuje iskanje soglasja med političnimi odločevalci, agencijami ZN in drugimi medvladnimi organizacijami, raziskovalci, akademskim svetom in predstavniki nevladnih organizacij glede etičnih načel, ki jih je treba vgraditi v te tehnologije.« (2020) V Pozivu so navedena temeljna načela algor-etike, ki naj bi usmerjala etični razvoj UI. Ta načela so: preglednost, vključevanje, odgovornost, nepristranskost, zanesljivost, varnost in zasebnost.

Če skušamo ovrednotiti „Rimski poziv za etiko UI“ v skladu s trojno vlogo vere in teologije na moralnem področju – integracija, spodbuda in kritika –, potem lahko brez dvoma ugotovimo, da je v dokumentu široka odprtost za *integracijo* etičnega razmišljanja izven Katoliške Cerkve. Poziv ima željo, da bi bil čim bolj obče sprejemljiv, zato v besedilu ne najdemo sledov teološkega utemeljevanja, ampak etično osnovo postavlja Splošna deklaracija o človekovih pravicah. Poudarjajo se univerzalne vrednote dostojanstva človeške osebe, pravičnosti, vključujoče družbe in trajnostnega okoljskega razvoja. Poziv *spodbuja* razvijalce, odločevalce in uporabnike UI, da naj jih pri usmerjanju nadaljnjega razvoja vodi želja po varovanju človekovega dostojanstva, izgradnji pravičnejše družbe in ohranitvi naravnega okolja. Podpisniki posebej ne poudarjajo, da lahko verska prepričanja in religijske skupnosti pripomorejo k motivaciji za celostni etični razvoj. Iz samega besedila ni razviden motiv, zakaj bi si morali različni deležniki na področju razvoja UI sploh prizadevati za etični razvoj. Enostavno se predpostavlja, da so dostojanstvo človeške osebe, svoboda, ustvarjalnost, pravičnost, vključujoča družba in trajnostni razvoj tako pomembne vrednote, da se jim ne želimo odpovedati. Vera v Boga Stvarnika vsekakor daje močno *spodbudo* k zavzemanju za spoštovanje teh temeljnih vrednot. Poziv pa daje velik poudarek *kritičnemu* presojanju, ki vključuje določeno vizijo izpolnjenega človeškega življenja, na podlagi katere razločujemo, kaj je dobro in prav glede uporabe UI. Tudi glede tega vera in teologija nista nikjer eksplicitno omenjeni, vendar je prav vera v Boga tista, ki onemogoča, da bi absolutizirali katerikoli človeški dosežek. Vera nam pomaga, da nas napredek ne zavede in ne opijani. Pri sedanjem skokovitem razvoju na področju UI je to še toliko bolj pomembno. Rimski moralni teolog Paolo Benanti je zapisal: »Bolj kot kadar koli doslej potrebujejo inovacije danes široko antropološko razumevanje, če se želijo preoblikovati v pristen izvor človeškega razvoja.« (2021, 134) Krščanska antropologija je seveda povezana s temeljnim prepričanjem, da je človek bitje, ki je ustvarjeno po Božji podobi, zato nosi v sebi svojo presežno vrednost in svetost.

Standfordsko poročilo o vplivnih prispevkih na področju etike UI je „Rimski poziv“ uvrstilo med pet najodmevnejših novic s tega področja v letu 2020 (Stanford University 2021, 12; 128; 131). Delovanje ustanove RenAIssance, ki deluje pod okriljem Papeške akademije za življenje in nadaljuje poslanstvo „Rimskega poziva“, je v zadnjih letih močno razširilo mrežo partnerjev, ki sodelujejo pri nadaljnjih projektih za oblikovanje etike UI (Pecoraro in Curzel 2023). Velik dosežek je podpis

„Rimskega poziva“ s strani predstavnikov enajstih najbolj razširjenih svetovnih verstev v Hirošimi 10. julija 2024 (RenAissance Foundation 2024).

2.2 „Umetna inteligenca in mir“ (2024)

V nadaljevanju bomo predstavili dva dokumenta, pod katera se je podpisal papež Frančišek. Prvi je tradicionalna, že 57. papeška poslanica za svetovni dan miru, ki ga obhajamo 1. januarja. Poslanica za leto 2024 nosi naslov „Umetna inteligenca in mir“. Na začetku Frančišek pozdravlja napredek v znanosti in tehniki, ki vodi k izboljšanju človeštva in preoblikovanju sveta. Znanost in tehnika sta po papeževih besedah izraz dostojanstva, ki ga je Stvarnik položil v človeka. »Tudi napredek znanosti in tehnike, kolikor prispeva k boljši urejenosti človeške družbe ter k rasti svobode in bratskega občestva, torej vodi k izboljšanju človeka in preobrazbi sveta.« (Frančišek 2023, 6) Vendar pa sodobni znanstveno-tehnični razvoj prinaša tudi številna tveganja za preživetje in nevarnosti za skupni dom. »Izjemen napredek novih informacijskih tehnologij, zlasti na digitalnem področju, predstavlja torej navdušujoče priložnosti in resna tveganja s pomembnimi posledicami za doseg pravičnosti in za harmonijo med ljudstvi.« (6)

Papež želi s poslanico odgovoriti na nekaj temeljnih vprašanj glede prihodnosti UI, predvsem glede njenega vpliva na življenje posameznikov in družbe. Zaveda se, da je digitalna tehnologija sprožila globoke družbene spremembe. »Nova digitalna orodja spreminjajo podobo komunikacij, javne uprave, izobraževanja, potrošnje, medosebnih odnosov in nešteti drugih vidikov vsakdanjega življenja.« (7) Vse to odpira številna etična vprašanja in dileme, na katere morajo biti pozorni razvijalci in uporabniki UI. Frančišek pojasnjuje, da bi bilo bolj pravilno govoriti o UI v množini, saj gre za različne sisteme, ki želijo posnemati spoznavne sposobnosti ljudi. »Umetno inteligenco moramo torej razumeti kot galaksijo različnih resničnosti in ne moremo vnaprej predvidevati, da bo njen razvoj sam po sebi blagodejno prispeval k prihodnosti človeštva in miru med narodi. Takšen pozitiven izid bo mogoč samo, če bomo dokazali, da smo sposobni delovati odgovorno in spoštovati temeljne človeške vrednote, kot so »vključenost, preglednost, varnost, pravičnost, diskretnost in zanesljivost.« (8–9). Vidimo, da papež navaja podobna načela kot „Rimski poziv“ iz leta 2020. Frančišek se zavzema za konkretne in zavezujoče etične smernice, ki bi jih morali spoštovati načrtovalci algoritmov in digitalnih tehnologij. Če želimo, da bo digitalni razvoj prispeval k miru in pravičnosti, potem mora spoštovati dostojanstvo vsake človeške osebe in vesoljno bratstvo vseh ljudi. »Tehnološkega razvoja, ki ne vodi k izboljšanju kakovosti življenja vsega človeštva, ampak nasprotno povečuje neenakosti in konflikte, nikoli ne moremo imeti za pravi napredek.« (10) Večkrat v poslanici sta poudarjeni globalna pravičnost in socialna enakost, ki morata biti merili pri razvoju UI. Pri zagotavljanju enakosti pa Frančišek izpostavi posebno pozornost do ubogih in obrobnih. Trdi namreč, da stopnjo naše človečnosti razodeva način, »kako umetno inteligenco uporabimo za vključevanje najmanjših, to je najslabotnejših in najbolj potrebnih bratov in sester« (17).

V svetu preobila in prenasičenosti Frančišek izpostavlja vrednost antične kreposti zmernosti, ki nasprotuje kapitalističnemu načelu učinkovitosti, na katerem

temelji sodobni tehnološki razvoj. Človek, ki ne priznava mej, je v nevarnosti, »da bo zaradi obsedenosti z željo imeti nadzor nad vsem, izgubil nadzor nad samim seboj; da bo v iskanju absolutne svobode padel v spiralo neke tehnološke diktature« (13). Človek kot ustvarjeno bitje mora priznati in sprejeti svoje meje, če želi živeti polnost življenja. Učinkovitost ne sme biti merilo za vrednost človeškega življenja.

Ko poslanica na verodostojen in prepričljiv način opozarja na številne negativne posledice uvajanja UI na različnih področjih človeškega življenja in delovanja, izpostavi osnovno antropološko zmoto, na kateri temelji vseobsegajoča uporaba UI. Človek ni zgolj skupek podatkov! »Algoritmom ne smemo dovoliti, da bi določali, kako razumemo človekove pravice, da bi zanemarjali bistvene vrednote sočutja, usmiljenja in odpuščanja ali da bi odpravili možnost, da se posameznik lahko spremeni in pusti za seboj svojo preteklost.« (15) V nadaljevanju papež še enkrat izpostavi, da ima samo človek moralno odgovornost in je zato tudi odgovoren za načrtovanje in uporabo UI: »Izključno človeška sposobnost moralnega presojanja in etičnega odločanja je več kot zapleten skupek algoritmov in te sposobnosti ne moremo omejiti na programiranje nekega stroja, ki ne glede na to, kako ‚pameten‘ je, še vedno ostaja zgolj stroj.« (16) Frančišek posebej opozori na nevarnost novih »smrtonosnih avtonomnih oborožitvenih sistemov«, ki vključujejo UI in omogočajo vojaške posege brez neposrednega sodelovanja človeka, čemur smo vsakodnevno priča pri vojaških spopadih v Ukrajini, Rusiji, Izraelu, Palestini, Libanonu in drugje.

Na koncu poslanice papež pozove k interdisciplinarnemu dialogu, ki naj vodi v oblikovanje »algor-etike« (17) za zagotavljanje etičnega razvoja algoritmov. Za uveljavitev etičnih standardov pri razvoju algoritmov pa sta za papeža ključni dve področji: vzgoja in mednarodno pravo. Na področju vzgoje za uporabo UI ima posebno vlogo »spodbujanje kritičnega mišljenja« (18). Ključna veščina je sposobnost razločevanja pri uporabi podatkov in vsebin, ki jih lahko najdemo na spletu ali jih proizvedejo sistemi UI. Na področju prava pa je treba oblikovati zavezujočo ureditev za razvoj in uporabo UI na globalni ravni. »Cilj predpisov pa seveda ne bi smel biti samo preprečevanje slabih praks, ampak tudi spodbujanje dobrih, novih in ustvarjalnih pristopov ter omogočanje osebnih in kolektivnih pobud.« (19) Etične smernice so potrebne, vendar še ne zagotavljajo etičnega razvoja UI. Papež spodbuja k skupnemu poglobljenemu razmisleku o smislu človeškega bivanja, varovanju temeljnih človekovih pravic ter prizadevanju za pravičnost in mir. Kot poseben dragocen in pomemben pa navaja glas »ubogih, obrobni in drugih, ki v globalnih procesih odločanja pogosto ostajajo neslišani« (19).

Če pogledamo na to poslanico skozi prizmo treh dejavnikov krščanske vere in teologije za etiko, potem ponovno zelo močno izstopa zavezanost k *integraciji* najnovejših spoznanj o razvoju UI in njenih pozitivnih in negativnih vplivih. Celotno besedilo je napisano na način, da bi bilo razumljivo in dostopno tudi nekristjanu. Ob koncu papež pozove k skupnemu delovanju: »Naj verni kristjani, verniki različnih verstev ter možje in žene dobre volje usklajeno sodelujejo, da bodo izkoristili priložnosti in se spoprijeli z izzivi, ki jih prinaša digitalna revolucija, ter prihodnjim

rodovom izročili bolj solidaren, pravičen in miroljuben svet.« (20) Sklicevanj na teološke temelje je v poslanici sicer nekaj več kot v besedilu „Rimskega poziva“, vendar še vedno prevladuje argumentacija na ravni univerzalne etike. Cilj razvijanja UI je spodbujanje celostnega razvoja človeka, pri čemer sta poudarjena tudi človeško bratstvo in spoštovanje naravnega okolja. Teološki poudarki, ki jih je sicer malo, *spodbujajo* odgovornost posameznikov in družbe za daljnosežne posledice današnjih odločitev na področju UI. Ustvarjeni smo po Božji podobi, smo svobodni in razumni ter poklicani, da s svojo ustvarjalnostjo sodelujemo z Bogom pri dovršitvi sveta. (5–6) Poslanica poudarja presežno razsežnost resničnosti, ki je ni mogoče zaobjeti v človeške ideje, tudi s pomočjo najnovejših algoritmov ne. Pri tem se sklicuje na enega od štirih stebrov družbene misli papeža Frančiška, ki jih je navedel v apostolski spodbudi „Veselje evangelija“, in sicer da »resničnost presega idejo« (12). Močan vidik papeževe noveletne poslanice pa je oblikovanje *kritične* distance do nekritične uporabe UI, in sicer na temeljni antropološki predpostavki, da je človek več kot samo skupek podatkov. Na podlagi biblične antropologije Frančišek poudarja, da je človek od Boga ustvarjeno bitje, svobodno, razumno, ustvarjalno, a vendar tudi omejeno, krhko, ranljivo in grešno. Vera v Boga Stvarnika nas osvobaja pred vsako obliko »tehnološke diktature« (13), kot nevarnost vseobsegajoče moči novih digitalnih tehnologij imenuje papež Frančišek.

2.3 „Umetna inteligenca in modrost srca“ (2024)

Kmalu po objavi noveletne poslanice je papež Frančišek ob koncu januarja 2024 ponovno spregovoril o etičnih izzivih UI, tokrat v poslanici ob 58. dnevu družbenih občil s pomenljivim naslovom „Umetna inteligenca in modrost srca“. Sporočilo sicer ne prinaša bistvenih vsebinskih novosti glede na prejšnja dva dokumenta, predstavlja pa izviren interpretacijski ključ za etični razvoj UI, in sicer modrost srca.

Frančišek poseže po avtorju, pri katerem je želel napisati svojo doktorsko delo. Vemo, da doktorskega študija v Nemčiji ni zaključil, vendar se je na nemškega teologa in misleca Romana Guardinija skliceval že v več dokumentih (npr. kar osemkrat v okrožnici „*Laudato si*“). Guardini je bil predstavnik nove teološke misli, ki je na začetku 20. stoletja spodbujala krščansko misel k odprtemu odnosu do sveta, hkrati pa opozarjala na nevarnosti tehnološke civilizacije. Njegovo delo *Konec moderne dobe* (*Das Ende der Neuzeit*) ni izgubilo na aktualnosti, čeprav je bilo napisano kmalu po koncu druge svetovne vojne. (Guardini 1965) Frančišek ponavlja Guardinijeve besede, da moramo biti odprti za napredek, a tudi občutljivi za vse, kar ogroža našo človečnost. »Oblikovati se mora nov tip človeka, obdarjen z globljo duhovnostjo, z novo svobodo in novo notranjostjo.« (Frančišek 2024a) Papež opozarja na nevarnost, da bi bili bogati s tehnologijo, a hkrati revni s človečnostjo. In nadaljuje: »Samo če se oskrbimo z duhovnim pogledom, samo če si povrnemo modrost srca, lahko beremo novosti našega časa in ponovno odkrijemo pot do popolnoma človeške komunikacije.« (2024a) Ko razlaga, kaj ima v mislih z modrostjo srca, se opira na biblični pogled. V svetopisemskem svetu je srce kraj svobode in odločanja, pa tudi kraj modrosti in razumevanja. Predstavlja enotnost človeka, vključno z njegovimi željami, hrepenenji in sanjami. Srce predstavlja sre-

dišče človeške osebe. Je tudi notranji prostor za srečanje z Bogom. »Modrost srca je zato tista krepost, ki nam omogoča zasnovati skupaj celoto in dele, odločitve in njihove posledice, višine in krhkosti, preteklost in prihodnost, jaz in mi.« (2024a) Modrost srca nam pomaga odkriti pravi pomen stvari in daje okus življenju.

Bistvena razlika med človekom in UI je, da je samo človek sposoben modrosti. Stroj je sicer zmogljivejši pri pomnjenju in povezovanju podatkov, vendar ni sposoben podatkom dati pomena. Kot bistveno posebnost človeka papež navaja njegovo zavest nezadostnosti in ranljivosti, ki ga odpira v odnose in spodbuja k ustvarjalnosti. Ljudje smo v svojem hrepenenju po preseganju vedno v nevarnosti, da podležemo prvotni skušnjavi, »da bi postali kakor Bog brez Boga (prim. 1 Mz 3), kar pomeni, da bi želeli s svojimi močmi osvojiti to, kar bi morali namesto tega sprejeti kot Božji dar in živeti v odnosu z drugimi« (Frančiček 2024a). Usmerjenost srca določa, kako bo človek uporabljal stvari okrog sebe, tudi sisteme UI. »Kot vsaka druga stvar, ki je prišla iz človekovega uma in njegovih rok, tudi algoritmi niso nevtralni. Zato je potrebno delovati preventivno in predlagati modele etičnega urejanja, da bi obrzdali škodljive in diskriminatorne ter družbeno krivične plati sistemov umetne inteligence, in se zoperstaviti njihovi uporabi pri zmanjšanju pluralnosti, v polarizaciji javnega mnenja ali v graditvi ene same misli.« (2024a)

Tudi na tem mestu papež poziva k mednarodni ureditvi področja razvoja in uporabe UI, vendar ugotavlja, da zgolj ustrezna zakonodaja ne bo dovolj. Potrebujemo rast v človečnosti: »Poklicani smo, da rastemo skupaj v človečnosti in kot človeštvo.« (2024a) Ta temeljni cilj mora usmerjati tudi razvoj UI, ki je stalno podvržen ambivalentnim možnostim. »Velike možnosti za dobro spremlja nevarnost, da bi se vse spremenilo v abstraktni račun, ki ljudi skrči na podatke, misel na obrazec, izkustvo na primer, dobro na dobiček, predvsem pa da bi na koncu zanikali enkratnost vsake osebe in njene zgodovine, ko bi konkretnost resničnosti razkrojili v vrsto statističnih podatkov.« (2024a) To, pred čemer svari Frančiček, je v veliki meri podobno ideologiji 'dataizma' (Harari 2022), ki resničnost dojema kot seštevek podatkov ter vse upe na boljšo prihodnost polaga v UI, ki naj bi preko uspešnega procesiranja velike količine podatkov razrešila vse probleme. Resnični svet pa je seveda mnogo več kot zgolj seštevek podatkov. »Predstavitev resničnosti v *velikih podatkih*, čeprav je funkcionalna za upravljanje strojev, namreč vključuje bistveno izgubo resničnosti stvari, ki ovira medosebno komuniciranje in predstavlja tveganje, da škoduje naši človečnosti.« Za človeško komunikacijo je bistven neposredni medosebni odnos, ki vključuje telesno razsežnost ali kot pravi papež: »Poleg vzajemne delitve zahteva obraz, pogled in sočutje.« (Frančiček 2024a) Slednjih pa ne more dati noben sistem UI.

Frančiček zaključuje, da je prihodnost odprta. »Človeku pripada odločitev, ali bo postal hrana za algoritme, ali pa bo svoje srce hranil s svobodo, brez katere ne more rasti v modrosti. Ta modrost dozoreva s tem, da čas spreminja v zaklad in sprejema ranljivost.« (Frančiček 2024a) Papež ponovno uporabi pojem ranljivost, da bi označil človeško ustvarjenost, nezadostnost in naravnost na odnose. Svojo poslanico sklene z besedami: »Da ne bi izgubili svoje človeškosti, iščimo Modrost, ki je pred vsem (prim. Sir 1,4), ki skozi čista srca pripravlja Božje prijatelje in preroke (prim. Mdr 7,27).« (Frančiček 2024a)

Od vsem treh obravnavanih besedil je poslanica „Umetna inteligenca in modrost srca“ najbolj teološko obarvana. Ne ukvarja se toliko z univerzalnimi etičnimi načeli, čeprav se pri obravnavanju problematike o uporabi UI v komunikaciji sklicuje na širše etično razmišljanje in *integrira* najpomembnejša spoznanja s tega področja v svojo refleksijo. Prevzame Guardinijevo držo kritične odprtosti do razvoja sodobne tehnike in v ospredje postavlja duhovno razsežnost človeka oz. modrost srca. Eksplicitno predstavlja versko *spodbudo*, da kristjan išče modrost od zgoraj, da bi na pravi način usmerjal sodobni digitalni razvoj. Tudi ta dokument poudarja, da nobena tehnologija, tudi UI, ni nevtralna, zato je potrebno razločevanje. Za *kritično* presojanje ponovno daje kriterije spoštovanja dostojanstva človeške osebe, skupnega dobrega in varovanje stvarstva. Posebej opozori tudi na nevarnost zoženja celotne resničnosti na podatke, pri čemer pa naj bi nas prav vera v Boga osvobajala vsakega redukcionističnega pogleda.

3. Etika kreposti v navezavi na družbeni nauk Cerkve

Ena od vlog teologa znotraj etičnih razprav v pluralni družbi je, da postavlja temeljna vprašanja, ki zadevajo cilj in smisel etičnih odločitev. Vsi trije obravnavani dokumenti vsebujejo spraševanje o posebnosti človeka in vplivu novih tehnologij na njegovo svobodo in dostojanstvo, o prihodnosti človeštva in planeta, o ohranjanju človečnosti in spodbujanju enakosti med ljudmi. Za etično usmeritev razvoja UI ne bodo zadostovala samo pravila delovanja in etične smernice, ampak je nujno apelirati tudi na spodbujanje temeljnih etičnih drž oz. kreposti tako pri razvijalcih sistemov UI kot pri njihovih uporabnikih. Etika kreposti si za temeljno nalogo zastavlja iskanje odgovora na vprašanje, kako omogočiti ljudem, da postanejo dobri ljudje in izpolnijo svoje človeško življenje. Stres pravi, da je krepost »moralna ‚kondicija‘: usposobljenost in krepkost za izvrševanje tistega, kar je moralno dobro in prav« (2018, 442). Kreposti so torej temeljene življenjske drže, ki ljudem omogočajo, da so dobri ljudje in da izpolnijo svoj življenjski cilj. Vzgoja za kreposti mora najti svoje mesto v procesu izobraževanja in tudi nadaljnjem vseživljenjskem oblikovanju ljudi. (Kraner 2023)

Razvoj UI nas spodbuja k ponovnemu premisleku, kako razvijati in ohranjati pristno človečnost znotraj novih možnosti in priložnosti. UI namreč korenito spreminja način človeškega življenja, človeško razmišljanje in delovanje, naše medosebne odnose in tudi zmožnosti poseganja v naravno okolje. Postavlja še temeljno vprašanje, v kakšni meri bo uporaba UI spodbujala dobrobit posamezne človeške osebe, človeštva in celotnega naravnega okolja. Katere temeljne moralne drže naj bi imeli snovalci in odločevalci pri razvoju UI, da bo razvoj zagotavljal in spodbujal dostojanstvo vsake človeške osebe, večjo pravičnost in enakost znotraj človeške družbe ter vzdržnost in trajnost celotnega naravnega okolja? Ob tem pa seveda ne gre zanemarjati tudi vloge uporabnikov UI, ki s svojimi navadami, pričakovanji, zahtevami in odločitvami pomembno usmerjajo ta razvoj.

Verske skupnosti imajo na tem področju pomembno vlogo, saj preko verskega nauka oblikujejo tudi moralni značaj svojih vernikov (Vodičar 2023a, 11). Krščanski

pogled na človeka temelji na verski resnici, da je človek ustvarjen po Božji podobi, da je svobodno in razumno bitje, ustvarjen iz ljubezni in za ljubezen, ustvarjen kot moški in ženska, postavljen kot skrbnik skupnega doma, da je ranljivo bitje, ki se uresničuje znotraj odnosov, da je tudi grešno bitje, ki pa ga Bog odrešuje in mu s svojimi odrešilnimi dejanji odpira pot v novo življenje. (Platovnjak in Svetelj 2023, 982–983; Platovnjak 2024, 201–202) Ta verski pogled je podlaga določenim temeljnim etičnim predpostavkam, kot so spoštovanje dostojanstva vsakega človeškega življenja, enakost med spoloma, družbena pravičnost, skrb za naravno okolje, pripravljenost za spravo in odpuščanje. Zdi se, da danes zaradi hitrega razvoja digitalnih tehnologij, ki imajo vedno močnejši vpliv na spreminjanje samorazumevanja človeka in njegovih odnosov, še bolj potrebujemo poglobljeno spraševanje, kdo je pravzaprav človek, kaj je tisto najbolj dragoceno in specifično človeško, česar ne smemo uničiti, in kako naj uporabljamo pridobitve sodobne tehnologije, da bodo spoštovale dostojanstvo vsake človeške osebe, skupno dobro človeštva in notranjo vrednost celotnega stvarstva.

Teološka refleksija mora dati svoj prispevek pri odločanju o prihodnjem razvoju UI, da bo zagotovila ohranitev pristnega človeškega bivanja tudi v prihodnosti. Benanti označuje sedanjo generacijo mladih kot *generacijo omega*, ki jo označuje kot zadnjo generacijo človeštva, kot smo ga poznali do sedaj. »Zmožnost' te generacije bi lahko bila v tem, da bi se spremenila v nekaj zelo drugačnega od tistega, kar trenutno razumemo kot človeško.« (2021, 135) Benanti pravi, da bo prihodnji človek neke vrste potepuh oz. iskalec; če bo znal sprejeti duhovni klic, bo spet postal *viator*, sicer se bo obsodil na tavanje in brezcilnost (136). Po njegovem prepričanju bo morala generacija omega odgovoriti na tri temeljna vprašanja o človeški naravi, ki jih ne bo mogla preložiti naprej, in sicer glede odnosa človeštva do okolja, odnosa človeštva do tehnologije in odnosa človeštva do samega sebe. Mogoče je Benantijeva napoved pretirana in deluje provokativno. Vsaka generacija človeštva je namreč morala na ta temeljna vprašanja odgovoriti zase. Je pa res, da postajajo procesi razmišljanja in odločanja s pomočjo UI čedalje hitrejši, zato je potrebna globlja in daljnosežnejša odgovornost.

Iz vseh treh predstavljenih dokumentov odseva načelno odobravanje nadaljnega razvoja in uporabe UI, saj lahko v marsičem olajša človeško življenje in prispeva k bolj pravični in enakopravni družbi. Po drugi strani pa je lahko tudi vir številnih zlorab človekovega dostojanstva in njen nadaljnji razvoj lahko prispeva k še večjim razlikam v družbi in degradaciji naravnega okolja.

Ob zaključku tega prispevka želimo izpostaviti osem temeljnih drž oz. kreposti, ki izhajajo iz družbenega nauka Katoliške Cerkve in lahko služijo za usmerjanje nadaljnjega razvoja UI. Družbeni nauk Cerkve poznamo predvsem po načelih, v skladu s katerimi presojava etičnost določenih družbenih pojavov. Po našem prepričanju je treba od načel narediti še en korak globlje, in sicer k oblikovanju temeljnih drž oz. kreposti. Načela ostajajo na racionalni ravni in omogočajo etično presojanje, medtem ko gre spodbujanje kreposti že v konkretno oblikovanje celostne osebnosti, ki poleg racionalne presoje vključuje tudi čustveno, volitivno in dejavnostno plat. V tem oziru etika kreposti bolj celostno spodbuja posameznika

in skupnost k delovanju in označuje tiste temeljne drže, ki so nujno potrebne za končni skupni cilj: oblikovanje boljše prihodnosti za vsakega človeka, za človeško družbo in za celotno naravno okolje. Na zelo shematski način bomo povzeli temeljne drže, ki zagotavljajo, da bo razvoj in uporaba UI sledila temu cilju.

Spoštovanje človekovega dostojanstva. Izhodišče družbenega nauka Cerkve je, da mora biti spoštovanje človekovega dostojanstva vedno v središču vsakega družbenega razvoja. Nikoli se ne sme podrediti posameznim koristim določene družbene institucije. Kot je pokazala naša razprava, tudi vsi omenjeni cerkveni dokumenti o UI vedno na prvo mesto postavljajo spoštovanje dostojanstva vsakega človeka in njegovega celostnega razvoja. Z razvojem digitalnih orodij in umetne inteligence se po eni strani povečujejo možnosti za večje spoštovanje vsakega posameznika, saj se omogoča bolj individualiziran pristop na različnih področjih (izobraževanje, zdravstvo), po drugi strani pa so možne tudi večje zlorabe prav zaradi nadzora nad posameznikom, vdora v zasebnost in izkoriščanja informacij. Pomembno je tudi, da ostaja jasna ločnica med človekom in digitalnimi orodji UI, da smo uporabniki vedno informirani, kdaj komuniciramo s konkretno osebo, kdaj pa s strojem, ki uporablja UI.

Zagotavljanje skupnega dobrega. Družbeni nauk Cerkve kot protiutež pretiranemu poudarjanju pravic posameznika poudarja socialno vpetost vsake osebe in nujnost njenega udejstvovanja za skupno dobro. V današnji dobi individualizma je spodbujanje zavzemanja za skupne dobrine še bolj pomembna. Pri tem ne mislimo samo na skupne materialne dobrine, kot so infrastruktura, skupne javne površine ipd., ampak tudi na številne dejavnosti skupnega življenja, kot so npr. šolstvo, zdravstvo, socialna oskrba ... Drugi vatikanski koncil pravi, da je skupno dobro »vsota vseh tistih pogojev za družbeno življenje, ki so potrebni, da morejo posamezniki, družine in združenja čim bolje in čim lažje doseči lastno popolnost« (GS 74). Na področju skupnega življenja ima pomembno vlogo oblast, ki mora na različnih ravneh zagotavljati enake možnosti vsem, da dostopajo do skupnega dobrega in da tudi v skladu s svojimi pristojnostmi prispevajo k skupnemu dobremu. Poleg spoštovanja človekovega dostojanstva vsi cerkveni dokumenti o UI na drugem mestu poudarijo, da mora UI prispevati k večji vključenosti vseh ljudi v družbi in k zagotavljanju boljših možnosti za vse, zato tudi večkrat izpostavijo, da je treba posebno pozornost posvetiti prav ubogim in obrobnim. Prav evangeljska drža posebne skrbi za uboge in obrobne se nam zdi merodajna pri zagotavljanju skupnega dobrega. Razvoj sodobnih tehnologij ne sme povečevati razlik v družbi, ampak mora biti naravnani k temu, da se vsem omogoča dostojno življenje.

Pravičnost. Krepost pravičnosti lahko razumemo kot bolj natančno opredelitev kreposti zagotavljanja skupnega dobrega. Gre za oblikovanje drže, ki se zavezama, da bi bile vsem zagotovljene enake možnosti, ne glede na spol, starost, prepričanje, vero, narodnost ali gmotni položaj. Hkrati pa se v

kontekstu UI zaveda tudi pomena avtorstva posameznih besedil in spoštovanja avtorskih pravic. Drža pravičnosti po eni strani preprečuje privajanje intelektualne lastnine drugih, po drugi strani pa raziskovalce spodbuja k bolj odprtemu deljenju svojih znanstvenih dognanj, saj bi se s tem omogočil bolj enakomeren razvoj tudi tistim posameznikom in skupinam, ki nimajo enakih možnosti. Vsi trije zgoraj omenjeni dokumenti poudarjajo pravičnost, še prav posebej pravičnost na globalni ravni, ki bo vodila k zmanjševanju razlik med razvitimi in nerazvitimi državami in med bogatimi in revnimi v družbi.

Solidarnost. Eden od stebrov družbenega nauka Cerkve je tudi solidarnost. Danes je zavest o medsebojni povezanosti in soodvisnosti še mnogo večja kot v preteklosti, zato je solidarnost toliko bolj pomembna. Papež Frančišek spodbuja: »V teh časih, ko se zdi, da vse razpada in se razkraja, je dobro spodbujati k solidarnosti, ki se rojeva iz odgovornosti do neblogljenih, iščočih skupno usodo. Solidarnost se dejansko krepi v služenju, ki ima lahko najrazličnejše oblike skrbi za druge.« (2020, 115) Od najnežnejše dobe človeškega razvoja je treba spodbujati krepitve solidarnosti, ki spodbuja razvoj empatije do pomanjkanja in trpljenja drugih ljudi in krepi posameznika za pripravljenost za pomoč. V cerkvenih stališčih o UI je bilo večkrat poudarjeno, da morajo biti bremena porazdeljena tako, da bodo tisti, ki so v podrejenem položaju, lahko tudi s pomočjo sodobnih digitalnih sredstev hitreje napredovali. Drža solidarnosti nas vseskozi spodbuja, da prevzamemo svoj del odgovornosti za celotno skupnost in da smo pozorni na tiste, ki potrebujejo našo pomoč.

Subsidiarnost. Drža subsidiarnosti, ki je tudi lastna družbenemu nauku Cerkve, poudarja pomen avtonomnega odločanja posameznika in manjših skupnosti nasproti višjim in mogočnejšim institucijam. Kot smo videli, je velika nevarnost nadaljnjega razvoja digitalnih orodij prav popoln nadzor nad življenjem ljudi in centralizacija odločanja na globalni ravni. Lahko pride do še večje uniformiranosti vseh temeljnih procesov na svetovni ravni, pri čemer bi bili močno ogroženi avtonomija posameznika in bogastvo različnih človeških kultur. Pomembno je, da nadaljnji razvoj UI ne zmanjšuje svobode posameznika in manjših skupnosti, ampak podpira osebno pobudo ter se odpove skušnjavi po popolni centralizaciji in nadzoru.

Preglednost. Preglednost je krepitve, ki je bolj specifično povezana z razvojem in uporabo UI. Sorodna je krepitvi poštenosti, ki označuje posameznikovo iskrenost pri izražanju svojih namenov in ciljev delovanja. Iz naše analize stališč Katoliške Cerkve glede UI lahko izluščimo, da je preglednost oz. transparentnost eden od osnovnih pogojev za etičnost razvoja. Drža preglednosti nas spodbuja k temu, da zagotavljamo razločljivost delovanja sistemov UI in preverjamo njeno morebitno pristranskost pri oblikovanju zaključkov. Prav tako morajo biti pregledni tudi nameni in cilji snovalcev razvoja digitalnih orodij.

Previdnost. Drža previdnosti ima nalogo, da zavaruje raziskovalce in uporabnike pred negotovimi in tveganimi koraki. Vemo, da tudi največji strokovnjaki ne morejo popolnoma zanesljivo razložiti delovanja UI, zato bi morali v dvomljivih situacijah slediti drži previdnosti. Kadar nismo gotovi, da gre za zanesljivo in etično uporabo določenih orodij, se je bolje odpovedati uporabi, kot pa tvegati negativen rezultat. Drža previdnosti zahteva, da je dokazovanje varnosti uporabe na strani proizvajalca, ki mora na prepričljiv način zagotoviti, da uporaba ne bo imela negativnih posledic za posameznike in družbo.

Odgovornost za daljnosežne posledice. Drža odgovornosti za daljnosežne posledice se je najprej poudarjala na področju varovanja naravnega okolja. Kratkoročne koristi uporabnikov ne smejo pretehtati nad možnimi dolgoročnimi negativnimi posledicami za usodo celotnega planeta. Varovanje okolja je z okrožnico „Laudato si“ leta 2015 postalo tudi del družbenega nauka Cerkve, ki vidi neposredno povezavo med skrbjo za družbena vprašanja in skrbjo za naravno okolje. Zanimivo je, da je v treh cerkvenih dokumentih o UI vedno kot tretji steber za etično presojo – poleg spoštovanja človekovega dostojanstva in zagotavljanja družbene pravičnosti – vedno omenjena tudi skrb za naravno okolje. Množična uporaba UI bo zahtevala še večjo proizvodnjo energije, kar bo imelo negativne posledice za naravno okolje. Za doseganje svojih ciljev ljudje prevečkrat pozabljamo na vzdržljivost našega planeta, kar ima na koncu negativne posledice za možnost kakovostnega človeškega življenja v prihodnosti. Nujno je okrepiti držo odgovornosti in skrb za naravno okolje in za daljnosežne posledice naših sedanjih odločitev na vseh področjih delovanja, kar vključuje tudi področje razvoja in uporabe UI.

Razprava je pokazala, da se Katoliška Cerkve dejavno vključuje v širše družbeno razmišljanje o vlogi in pomenu UI. Katoliška Cerkve načeloma pozdravlja napredek sodobne tehnike in vidi v novih tehnoloških dosežkih nadaljevanje stvarjenjskega procesa. Bog je ustvaril človeka kot razumno bitje in mu zaupal skrb za nadaljnji razvoj stvarstva. Človek si je s pomočjo razuma in ustvarjalnosti v marsičem olajšal življenje in razvil različne možnosti, da živi bolj polno in srečno življenje. Tudi ves razvoj digitalne tehnologije in umetne inteligence je treba naprej presojati v tej pozitivni luči. Prinaša namreč ogromno prednosti za življenje človeka kot posameznika in za človeško družbo. Tako kot druge pridobitve človeškega razuma pa tudi sistemi umetne inteligence prinašajo s seboj številne pasti in možnosti zlorab. Na podlagi treh uradnih dokumentov iz zadnjih let in bogatega izročila družbenega nauka Cerkve smo oblikovali predlog etike kreposti, ki je lahko v pomoč krščanskim skupnostim pri vzgoji za etično uporabo UI.

Reference

- Agius, Emmanuel.** 2020. Challenges and Opportunities of Robotics to Catholic Theological Reflection. V: Vincenzo Paglia in Renzo Pegoraro, ur. *Robo-Ethics: Humans, Machines and Health*, 115–139. Rim: Pontifical Academy for Life.
- Auer, Alfons.** 1984. *Autonome Moral und christlicher Glaube*. Düsseldorf: Patmos.
- Benanti, Paolo.** 2021. *Tecnologia per l'uomo: cura e innovazione*. Milano: San Paolo.
- Boddington, Paula.** 2023. *AI Ethics: a Textbook*. Singapore: Springer.
- 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>
- Demmer, Klaus.** 1999. *Fundamentale Theologie des Ethischen*. Freiburg: Herder.
- . 2014. *Selbstaufklärung theologischer Ethik: Themen – Thesen – Perspektiven*. Paderborn: Ferdinand Schöningh.
- Francišek.** 2020. Fratelli tutti. O bratstvu in družbenem prijateljstvu. Okrožnica. Ljubljana: Družina.
- . 2023. Umetna inteligenca in mir. Poslanica svetega očeta Franciška za 57. svetovni dan miru. Ljubljana: Družina.
- . 2024a. Umetna inteligenca in modrost srca: za popolnoma človeško komunikacijo. Vatican News. 24. 6. <https://www.vaticannews.va/sl/papez/news/2024-01/papezeva-poslanica-umetna-inteligenca-in-modrost-srca.html> (pridobljeno 10. 6. 2024).
- . 2024b. Discorso del santo padre Francesco, Borgo Egnazia. Vatican. 14. 6. <https://www.vatican.va/content/francesco/it/speeches/2024/june/documents/20240614-g7-intelligenza-artificiale.html> (pridobljeno 30. 6. 2024).
- Globokar, Roman.** 2010. Pokoncilsko iskanje specifičnosti krščanske etike. *Bogoslovni vestnik* 70, št. 3:365–377.
- Grimm, Petra, Kai Erik Trost in Oliver Zöllner, ur.** 2024. *Digitale Ethik*. Baden-Baden: Nomos.
- Guardini, Romano.** 1965. *Das Ende der Neuzeit*. 9. izd. Würzburg: Werkbund-Verlag.
- Harari, Yuval Noah.** 2022. *Homo Deus: kratka zgodovina prihodnosti*. Ljubljana: Mladinska knjiga.
- Kraner, David.** 2023. Sodobni učitelji, komunikacija in resonanca v odnosih. *Bogoslovni vestnik* 83, št. 1:227–239. <https://doi.org/10.34291/BV2023/01/Kraner>
- Miklavčič, Jonas.** 2021. Zaupanje in uspešnost umetne inteligence v medicini. *Bogoslovni vestnik* 81, št. 4:935–946. <https://doi.org/10.34291/bv2021/04/miklavcic>
- . 2023. Ideal transparentnosti v digitalni dobi. *Bogoslovni vestnik* 83, št. 4:825–838. <https://doi.org/10.34291/BV2023/04/Miklavcic>
- Pegoraro, Renzo, in Elisabetta Curzel.** 2023. Rome Call for AI Ethics: the birth of a movement. *Medicina Y Ética* 34, št. 2:315–49. <https://doi.org/10.36105/mye.2023v34n2.01>
- Platovnjak, Ivan.** 2024. Human Vulnerability, Spiritual Abuse, and Its Prevention. *Obnovljeni život* 79, št. 2:199–212. <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>
- Pontifical Academy for Life.** 2020. Rome Call for AI Ethics, 28. 2. https://www.vatican.va/roman_curia/pontifical_academies/acdlife/documents/rc_pont-acd_life_doc_20202228_rome-call-for-ai-ethics_en.pdf (pridobljeno 10. 6. 2024).
- RenAlssance Foundation.** 2024. AI Ethics for Peace – Hiroshima. Rome Call. 10. 7. <https://www.romecall.org/ai-ethics-for-peace-hiroshima-july-10th-2024/> (pridobljeno 20. 8. 2024).
- Spiekermann, Sarah.** 2019. *Digitale Ethik: ein Wertesystem für das 21. Jahrhundert*. München: Droemer.
- Stanford University.** 2021. Artificial Intelligence Index Report 2021. https://aiindex.stanford.edu/wp-content/uploads/2021/11/2021-AI-Index-Report_Master.pdf (pridobljeno 15. 6. 2024).
- 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>
- Vodičar, Janez.** 2023. Avtoriteta na področju vzgoje in verovanja v digitalni dobi. *Bogoslovni vestnik* 83, št. 4:1035–1047. <https://doi.org/10.34291/BV2023/04/Vodicar>
- . 2023a. „Being with“ or Salesian Assistance in the Digital Age. *Seminare* 44, št. 2:1–14. <https://doi.org/10.21852/sem.1809>
- Zuccaro, Cataldo.** 2013. *Teologia morale fondamentale*. Brescia: Queriniana.

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 885—897

Besedilo prejeto/Received:07/2024; sprejeto/Accepted:08/2024

UDK/UDC: 27-472:004.89

DOI: 10.34291/BV2024/04/Vodicar

© 2024 Vodičar, CC BY 4.0

Janež Vodičar

Kateheza kot pomoč umetni inteligenci

Catechesis as an Aid for Artificial Intelligence

Povzetek: Umetna inteligenca je navzoča že na skoraj vseh področjih sodobnega življenja. Na pedagoškem področju se pogosto poudarja negativen vpliv, ki da ga ima na učence. V prispevku opozarjamo na tudi pozitivne vidike vključevanja umetne inteligence v pedagoško delo. Še težje je videti pozitivne vidike umetne inteligence na katehetskem področju, čeprav je to področje vključeno v novi „Pravilnik za katehezo“. Prav kateheza, kot skušamo pokazati, lahko pomaga k pripravljenosti sodobnega učenca za bolj varno in človeka vredno uporabo umetne inteligence tudi na področju splošne vzgoje in izobraževanja – kar ponazarjamo s predstavitvijo najnovejših gibanj, ki zagovarjajo pedagoško uporabo umetne inteligence, in s primerjavo s splošnimi cilji kateheze. Rezultate teh primerjav vzporejamo tudi z zadnjimi priročniki za slovensko katehezo v četrtem in petem razredu osnovne šole. Rezultati kažejo, da je lahko prispevek kateheze na področju vzgoje za odprtost in učljivost temelj za uspešno uporabo umetne inteligence v širšem pedagoškem prostoru.

Glavne besede: umetna inteligenca, izobraževanje, kateheza, učljivost, inteligentnost, oznanilo

Abstract: Artificial intelligence is already present in almost every aspect of modern life. In the field of education, its negative impact on students is often emphasized. In this article, however, we also show the positive aspects of integrating AI into educational work. It is even more difficult to show the positive aspects of AI in the catechetical field, even though this area is included in the new “Directory for Catechesis”. As we show in this article, catechesis can also contribute to a safer and more humane use of AI in the field of general education. We show this by presenting the latest movements advocating the pedagogical use of AI and comparing them with the general aims of catechesis. The results of these comparisons are also compared with the latest manuals for Slovenian catechesis in the fourth and fifth grades of elementary school. The results show that the contribution of catechesis in the field of education for openness and educability can be the basis for a successful positive application of AI in the wider pedagogical space.

Keywords: artificial intelligence, education, catechesis, educability, intelligence, proclamation

1. Uvod

Umetno inteligenco, ki je navzoča že povsod, redko povezujemo s katehezo.¹ Ta naj bi ostala 'zvesta' tradiciji Razodetja in uvajala v vero po načelih klasičnega oznanjevanja, predvsem pričevanja. Vendar novi Pravilnik za katehezo digitalni kulturi namenja veliko pozornosti in omenja celo umetno inteligenco (PK 356). V prispevku želimo obravnavati umetno inteligenco (UI) z vidika vključevanja v vzgojo in izobraževanje. Prav tako bomo prikazali povezanost UI in kateheze. Končni cilj je opozoriti na priložnosti, ki jih kateheza poraja in jih lahko nudi tistim, ki si želijo plodne vključitve UI v šolski in pedagoški proces nasploh. Ob vprašanju chatGPT, kaj lahko od kateheze pridobi UI, ta model odgovori v 10 vidikih – vzgoja za večjo etično in družbeno odgovornost, poglobljanje in razumevanje človeške narave, razumevanje smisla, vključevanje ranljivih, marginalnih skupin, odpornost na zlorabe, povezovanje duhovnosti in tehnologije, izobraževanje o digitalni pismenosti, medverski dialog in podpora duševnemu zdravju. Vse to je veliko priznanje in hkrati odraz pričakovanj, ki jih UI od kateheze ima. Sami pa želimo izpostaviti vidik kateheze, ki ga UI ni prepoznala – in hkrati opozoriti, da ga bo tudi težko prepoznala. Vzgoja za vero vključuje oblikovanje posameznika, ki zmore odprtost do Oznanila in tega tudi sprejme. Gre za vidik, za katerega v pedagoški literaturi najdemo oznako učljivost, a ga avtorji v povezavi z religijsko razsežnostjo poudarjajo redko. V razpravi bomo izpostavili, kako lahko prav ta naravna pedagoška usmeritev kateheze – tj. vzgoja za vero – pomaga k zdravi vzgoji za pozitivno uporabo UI.

2. Izzivi umetne inteligence za izobraževanje

Na medmrežni strani, kjer ponujajo izobraževanje, učne pripomočke, spodbude in načrte za delo z uporabo UI na vzgojno-izobraževalnem področju, o sebi trdijo: »Pedagog.ai se zavzema za spodbujanje pomena umetne inteligence v izobraževanju na vseh ravneh, od lokalne do globalne. Neumorno si bomo prizadevali za politike in pobude, ki podpirajo varno in pravično vključevanje umetne inteligence v izobraževalni sistem in zagotavljajo, da bodo tako učenci kot pedagogi pripravljeni na izzive in priložnosti sveta, ki ga narekuje umetna inteligenca.« (Pedagog.ai 2024) Navdušenci nad možnostmi, ki jih UI prinaša, so manj glasni, kot tisti, ki opozarjajo na nevarnosti za vzgojo in izobraževanje.² Že dalj časa poznamo težave, povezane s plagiatorstvom. V boju proti temu smo v dobi razcveta digitalnih virov razvili tehnike, ki nam omogočajo, kako to bolj ali manj preprečiti ali vsaj zavreti. Z razvojem novih aplikacij, ki jih UI omogoča (t. i. botov, ki po navodilih posameznika

¹ Prispevek je rezultat raziskovanja znotraj projekta J6-4626 „Teologija, digitalna kultura in izzivi na človeka osredičene umetne inteligence“ in raziskovalnega programa P6-0269 „Religija, etika, edukacija in izzivi sodobne družbe“, ki delujeta na Teološki fakulteti Univerze v Ljubljani in ju sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

² Tudi papež Frančišek v svoji poslanici za 57. svetovni dan miru, kjer je glavna tema prav UI, opozarja bolj na nevarnosti kot prednosti. Poziva k previdnosti in odgovornosti – ne pa toliko k uporabi in možnostim UI (Frančišek 2024).

samodejno pripravijo taka ali drugačna besedila, razpredelnice, slikovno gradivo ali predstavitve), je plagiatorstvo težko odkriti. Je pa kar nekaj zanesenjakov, ki namesto k obupu pozivajo k vključitvi UI v sodobno izobraževanje. Vse pogostejši so pozivi: »Učitelji se morajo naučiti, kaj je UI, kako jo učenci že in še lahko uporabljajo, kako jim lahko olajša življenje in kako pedagoške cilje, ki so se jim nekoč zdeli nemogoči, zdaj lahko dosežejo.« (Shah 2023, 2)

Vključevanje UI v izobraževanje poudarja tudi Svetovni gospodarski forum. Našteva tri glavna področja, ki jih lahko okrepi: 1) Prepričani so, da lahko UI učiteljem pomaga pri avtomatizaciji administrativnega dela. S tem se bodo lahko učitelji bolj osredotočili na poučevanje in osebne stike z učenci; 2) UI in njene aplikacije lahko prispevajo k pravičnejši udeležbi v izobraževanju v globalnem in tudi regionalnem smislu in pri izkoriščanju digitalnih pripomočkov zmanjšajo razlike med različnimi demografskimi skupinami; 3) Nujno je, da v procesu izobraževanja mlade pripravimo na koristno in etično ravnanje z UI, s čimer prispevamo tudi k regulaciji tveganj, ki jih prinaša (World Economic Forum 2024). Kot pri vseh novih odkritjih bo tudi UI v proces obveznega izobraževanja treba vključiti prej ali slej – ne le zaradi zahtev mednarodnih organizacij, temveč tudi zaradi potreb vsakdanjega življenja. Skoraj ni več področja, kjer UI ne bi bila navzoča: zjutraj nas prebuja; usmerja in vodi nas pri izbiri v okviru nakupa; prevaja nam tuj jezik ali ponuja manj prometne ceste. Le zakaj je ne bi uporabljali tudi v šoli (Luckin, George in Cukurova 2022, VII)?

Če se omejimo le na pedagoško področje, potem zadnje raziskave kažejo, da je na tem področju velika potreba po ukrepanju (Nežič Glavica 2023, 1051). To pričakujejo tako starši kot učenci. Vsi vemo, da šole niso bile pripravljene na hitro množično uporabo te nove tehnologije – večina jih tako pri ravnanju z UI ni oblikovala in uveljavila celovitih navodil, pravilnikov in postopkov. »Rezultat? Učenci, starši in učitelji poročajo o pomanjkanju jasnosti in uporabnih smernic glede splošne umetne inteligence. Dejstvo v praksi pa je, da so bili učenci – tudi brez teh pravilnikov in jasnih navodil – zaradi uporabe UI kaznovani.« (Laird, Dwyer in Grant-Chapman 2023, 31) Raziskava v ZDA je pokazala, da učenci splošno UI za šolo uporabljajo zelo malo, le 23 % – in več kot dvakrat toliko zunaj izobraževanja. Razhajanja med rezultati raziskave in prepričanji, ki jih imajo učitelji na ameriških šolah o svojih učencih, so precejšnja. Devetdeset odstotkov učiteljev namreč poroča, da menijo, da njihovi učenci splošno UI za šolsko delo uporabljajo; 4 od 10 učiteljev so izrazili prepričanje, da so njihovi učenci splošno UI uporabljali za pisanje in seminarskih nalog ali referatov (32). A malo je takih, ki UI že vključujejo v proces izobraževanja – še manj pa je takih na področju vzgoje za vero (pri katehezi), kar nas še posebej zanima.

3. Izobraževanje za vero in UI

Da gre za globalna vprašanja, kjer je nujno treba vključevati tudi religijska gibanja, je razvidno iz povabila svetovnih voditeljev papežu Francišku na vrh G7.

»Ne moremo dvomiti, da »prihod umetne inteligence predstavlja pravo kognitivno-industrijsko revolucijo, ki bo prispevala k oblikovanju novega družbenega sistema, za katerega bodo značilne zapletene epohalne spremembe«. Umetna inteligenca lahko denimo omogoči »demokratizacijo dostopa do znanja, eksponentni napredek znanstvenih raziskav, možnost prenosa težkih del na stroje«, obenem pa prinese »večjo nepravico med razvitimi državami in državami v razvoju, med dominantnimi in zatiranimi družbenimi razredi«, s tem pa tudi ogrozi »kulturo srečevanja' v korist 'kulture odmetavanja'.« (Červek 2024)

Kultura srečanja, ki naj bi je UI ne zmogla pospeševati, je po papeževih besedah pogosto središče celotnega krščanskega oznanila – to pa je njegov viden in nujen prispevek sodobni globalni družbi. Čeprav papež govori o grožnji kulturi srečanja, ta ni sama po sebi umevna. V sodobnem svetu vedno manj, saj se lahko srečujemo s komer koli in kadarkoli. Po drugi strani pa osamljenost postaja globalni problem (Kraner 2023, 1016). Če hočemo, da bo kultura srečevanja presegla kulturo praznega odmetavanja, naj bo to smeti ali praznih besed – in še bolj človeka –, mora biti vključena v pedagoški proces. Vse civilizacije so temeljne vrednote na nove generacije prenašale z vzgojo. UI prinaša številne izzive, zato mora biti v vzgojo po besedah papeža Frančiška vključena (Frančišek 2024). Vsi, ki so za to odgovorni (starši, učitelji, politični načrtovalci in tudi oblikovalci UI), bi morali temu področju namenjati več pozornosti (Kent in du Boulay 2022). Kateheza prav tako, saj gre za temeljno obliko organizirane občestvene vzgoje v Cerkvi. Ta naj bi posameznika pripravljala za zrelo življenje po veri, kar pa je vedno pomenilo vero v Jezusa Kristusa v konkretnem času in prostoru – tudi v času UI.

Če si ogledamo najprej Katehetski pravilnik, ta katehezo že takoj umešča v evangelizacijo: »Kateheza že po svoji naravi sodeluje pri evangelizacijskem prizadevanju z namenom, da bi podpirala nenehno zorenje v veri. Življenje iz vere je osnovna bivanjska značilnost Kristusovih učencev.« (PK 1) Če papež na svetovnem globalnem forumu poudarja kulturo srečanja, to prav gotovo za to, da bi pogled kristjana usmeril k temu. Ko se PK loteva digitalne kulture, poudarja nalogo kateheze, da se na izzive, ki jih ta nov način komunikacije prinaša, ustrezno odzove. »Tudi prednostna naloga katehetike je, da vzgaja za smiselno uporabo teh pripomočkov in poglobljeno razumevanje digitalne kulture. Tako pomaga razločevati med pozitivnimi in negativnimi stranmi teh sredstev.« (216) Kot je papež Frančišek na zasedanju G7 opozoril na dvojno naravo UI, je po besedah PK k angažiranemu odzivu na to pozvan tudi katehet pri svojem katehetskem pristopu. Spoštovanje dostojanstva vsake osebe in pa bratstvo vseh, ki veje iz papeške poslanice na dan miru in bi moralo biti temelj razvoja novih tehnologij (Frančišek 2024), je temelj družbenega in moralnega nauka Cerkve. Zato danes vzgoje za vero v Jezusa Kristusa brez ustreznega odgovora na izzive UI ne more biti.

Naš namen je, da iz temeljnega namena kateheze – ki je zorenje v veri in življenje iz nje – izpeljemo smiselno uporabo UI v izobraževalne namene kot eno od najboljših poti za pripravo mladih. Res gre le za ozko skupino katehizirancev, je pa

mogoče iz didaktičnih potez kateheze sklepati na možno širšo pomoč vsem, ki se želijo otresti zgolj negativnega stališča do UI in jo – kot to priporočajo številne mednarodne organizacije – smiselno vključiti v izobraževanje. Z uporabo digitalnih tehnologij se namreč ne spreminjajo samo metode in sredstva poučevanja, ampak se uvaja tudi nova miselnost, spreminjajo vrednote, in to na globalni ravni (PK 359). Kot zatrjuje papež Frančišek: »Tehnologija se rodi z določenim namenom in s svojim vplivom na človeško družbo vedno predstavlja obliko reda v družbenih odnosih in ureditev moči, ki nekaterim omogoča izvajanje dejanj, drugim pa preprečuje izvajanje drugih. Ta konstitutivna razsežnost moči tehnologije vedno bolj ali manj izrecno vključuje svetovni nazor tistih, ki so jo uresničili in razvili.« (Červak 2024) Zato ne gre zgolj za ozko katehetsko dejavnost, temveč za možnost, da kateheza človeka in njegov svetovni nazor preoblikuje tako, da bo ostal učljiv in svoboden za srečanje z drugimi. Gre za skupno skrb vseh pedagoških delavcev ob izzivu UI. Naš končni namen je poudariti človeški element izobraževanja in izpostaviti načine, kako lahko UI vlogo učiteljev okrepi – ne pa, da jih nadomesti (Shah 2023, 13). Pri tem pa ohranja možnost družbe srečevanja in preseganja ozkih, sebičnih interesov.

4. Kateheza in razumevanje inteligentnosti v času UI³

Ko govorimo o UI – zlasti pa v povezavi s katehezo –, je treba najprej opredeliti razumevanje inteligentnosti. Že samo dejstvo, da se od Bineta pa vse do Gardnerja in naprej pedagogi pri utemeljevanju svojih didaktično-metodoloških pristopov sklicujejo na različne opredelitve inteligentnosti, priča o pomenu izhodiščnega razumevanja tega pojma. Vsaka opredelitev ima prednosti in slabosti (Binetova je ozka, Gardnerjeva naj bi bila preširoka itd.). Sami bomo za potrebe umestitve UI v katehetski proces uporabili definicijo, povzeto po Sternbergu (Sternberg 1985). Po njegovem mnenju lahko razlikujemo tri vrste inteligentnosti: ustvarjalno, analitično in praktično. Eksperimentalna podteorija utemeljuje ustvarjalno inteligentnost. Ta se aktivira, kadar je treba njene gradnike uporabiti pri razmeroma (vendar ne povsem) novih nalogah in situacijah; obenem pomaga pridobiti, urediti in ovrednotiti ustrezne informacije. Skladenjska podteorija je osnova za analitično inteligentnost, ki je v ospredju, ko je treba gradnike uporabiti pri precej abstraktnih, a še vedno razmeroma znanih problemih – kot so naloge, s katerimi se učenci ukvarjajo v šoli. Kontekstualna podteorija, ki je usmerjena na praktično inteligentnost, je potrebna, ko se naučene vsebine uporabljajo pri razmeroma konkretnih in znanih problemih, kot so izzivi vsakdanjega življenja (Senior in Gyarmathy 2022, 97; Vinney 2020). Kateheza, ki sledi navodilom Cerkve (»Srečanje s Kristusom zajame človeka v celoti: srce, razum in čute« [PK 76]) in pedagoško deluje celostno, mora vključevati

³ Glede samega pojma inteligentnosti in intelligence uporabljamo za splošno teorijo, raziskovanje in razumevanje področja pojem inteligentnost, kot ga je pri nas uveljavil Janek Musek (Duhovna inteligentnost: struktura in povezave, *Antropos* 40 (2008), št. 3/4:259–276). Na področju umetne intelligence pa se je uveljavil pojem intelligence, zato tam uporabljamo ta izraz.

vse tri razsežnosti. Vsaj slovenska katehetska stroka je s sklicevanjem na Piageta to didaktično vključila v katehetske pripomočke (Snoj 2003, 184). Tako imamo cilje razdeljene na spoznavne, kjer nujno potrebujemo analitično inteligentnost; dejavnostne, kjer potrebujemo praktično inteligentnost; končno doživljajske – ti se ne pokrivajo s tem, kar pod vidikom ustvarjalne inteligentnosti razume Sternberg, vendar pa je za preseganje omejitev UI prav doživljajski vidik, kot ga razumejo snovalci kateheze, najbolj pomemben. Preko notranjega pretresa osebe, ki ga ta doživi v osebem notranjem doživetju (še posebej, če gre za religiozno), je odprtost za rast v pristni človeškosti in uspešen odziv na okolje – zlasti v odnosu do bližnjih – nujno potrebna.⁴

Kateheza, ki vodi do srečanja s Kristusom,⁵ je vedno tudi oživljanje lastnih moči, ki najprej prebudijo pozornost za skrito podobo človeka, ki je neviden v ponižanem, bolnem, zavrženem, grešnem ali celo mrtvem. To je za odgovorno in dejavno vključevanje v življenje družbe nujno – Božji dogodek mora vstopiti v konkretno življenje. Še bolj katehete k temu zavezuje t. i. doba doživetij: oznanilo Božjega kraljestva mora postati dogodek konkretnega telesa in konkretnega občestva. Naj bo analitična inteligentnost, ki lahko razume ranljivost človeka, ali pa praktična, ki se na to odzove z ustreznim ravnanjem, še tako močna, mora kateheza voditi do notranje ustvarjalnosti, ki ponudi odgovor vere. Danes je v hrupu sodobne tehnologije težko zaživeti poslanstvo odgovornega kristjana in klic, ki sledi temu notranjemu eksistencialnemu doživetju k ustvarjalnemu pričevanju, je nujen (Stegu 2023, 1025). Prav tega nas uči Jezusova prisposoba o usmiljenemu Samarijanu (Lk 10, 29-37). Za pomoč žrtvi napada je bil ustrezno ‚nastavljen‘ le Samarijan, čeprav naj po prepričanju duhovnika in levita ne bi imel pravega vodenja – se ga pa je bližnji živo dotaknil, s čimer je bil izpolnjen pravi namen postave. V duhu krščanstva je ravnal najbolj inteligentno in situaciji ustrezno. Pustil se je nagovoriti in s tem Božje naročilo najbolj izpolnil. Gradil je skupnost srečevanja (Šegula 2022, 687), kot bi rekel papež Francišek.

Z razvojem inteligentnega pristopa religiozne ustvarjalnosti, ki raste iz pristnega bivanjskega doživetja, lahko presežemo golo razumevanje, dojemanje in se približamo vprašanju dejavne resnice, kar je danes še kako aktualno. Ko papež govori o kulturi srečevanja, ima pred očmi prav ta vidik odprtosti za resnico, ki ga mora graditi kateheza. Najprej odprtost do Kristusa, kot smo že izpostavili, končno pa se to prepozna v odprtosti do soljudi – občestvenosti (PK 48). »Vera daje življenju zaradi človekovega srečanja z Gospodom Jezusom novo, širše obzorje.« (56) Prav tu je moč kateheze za čas vedno bolj prisotne UI, ki sicer ponuja hitre rešitve, in to na zelo inovativen način, a hkrati človeka zapira v svet zagledanosti vase. Še taki navdušenci nad didaktično uporabo UI zatrjujejo: »V vedno bolj di-

⁴ Sternberg v zadnjem času poudarja t. i. adaptivno inteligentnost, ki jo razlikuje od splošne inteligentnosti. Ta naj bila končno usmerjena tudi v zmožnost biološkega in kulturnega prilagajanja za bolj kvaliteten življenje – kar pa naj bi omogočala prav adaptivna inteligentnost (Sternberg 2019).

⁵ Slovenski katehetski načrt o namenu kateheze pravi: »Katehiziranje vabiti in spremljati k osebemu srečanju z Jezusom Kristusom, da bodo izpovedali osebno vero, se dejavno vključili v občestvo Cerkve in odgovorno živeli svoje poslanstvo v družbi.« (Slovenski katehetski urad 2018, 21)

gitaliziranjem svetu bomo morali učence naučiti spoštljivega in empatičnega komuniciranja v spletnih okoljih z umetno inteligenco ter upoštevanja čustev in stališč drugih ljudi. Potrebovali bodo spretnosti aktivnega poslušanja in odprtosti, da bodo znali ceniti različna mnenja in stališča – in da ne bodo dovolili, da bi vsebine, ki jih poganja umetna inteligenca, spodbujale njihove vnaprej ustvarjene predstave o družbi.» (Shah 2023, 186) Predvsem pa lahko samo človeška inteligentnost sočloveka prepozna kot sebi enakega, saj razmišlja in čuti podobno (Kent in du Boulay 2022).

5. Učljivost kot nujno ozadje kateheze

Podjetje DeepMind je izdelalo sistem za globoko učenje, ki je zmožeg inovativno odigrati preprosto igro Atari.

»Program se je sam naučil igrati s pomočjo učenja z ojačitvijo: večkrat je igral igro, pri čemer je v vsaki igri eksperimentalno pridobil povratne informacije in se naučil, katera dejanja vodijo do nagrad in katera ne. Program se je naučil igrati devetindvajset od devetinštiridesetih iger z zmogljivostjo, ki je bila višja od človeške. Pri nekaterih igrah je dosegel nadčloveške ravni zmogljivosti.« (Wooldridge 2021)

Tu gre po vseh treh podteorijah Sternberga za pravo inteligentnost. Program UI je ustrezno analiziral, razumel in ustvarjalno deloval pri konkretni nalogi igranja igre. Vendar nas lahko ta uspeh zlahka zavede, da UI počne nekaj, česar zagotovo ne počne – in tisto, česar ne počne, je razumevanje (2021). Razumevanje je za inteligentno odzivanje – zlasti na družbeno stvarnost – nujno. »Nenazadnje, kar je morda najpomembnejše, ko smo v odnosu z drugimi, običajno predpostavljamo, da imamo skupne vrednote in norme. Teh izrecno vsakič ne poudarjamo. Toda sistem umetne inteligence teh vrednot in norm ne pozna.« (2021) Predvsem pa moramo pri izobraževanju poznati in razumeti proces, ki ga vodimo. Pri strojnem učenju, ki je temelj UI, je to pogosto nejasno in pomanjkljivo. Ker nam UI predlaga izbire in usmerja naše odločitve, je ob tem treba ravnati pozorno in odgovorno (Luckin, George in Cukurova 2022, 11).

Tako človeški možgani kot računalniški sistemi pri iskanju rešitev za najrazličnejše situacije analizirajo in primerjajo že shranjene informacije in te, ki jih pridobivajo iz okolja. Bistvena razlika med človekom in digitalnim je, da lahko na človeški proces spoznavanja in odzivanja vpliva več potencialno nasprotujočih si dejstev, kot so npr. čustva, motivacijski dejavniki in konkretni ali dolgoročni cilji, kar nas pripelje do temeljne kritične točke: vprašanja pozornosti, usmerjenosti. »Izbira ustreznih zunanjih in notranjih dražljajev ter informacij, ki jih je treba analizirati, ima pri reševanju problemov odločilno vlogo.« (Senior in Gyarmathy 2022, 96) Ta zmožnost človeške inteligentnosti je priložnost za dvig samozavesti in s tem tudi odgovornosti (Luckin, George in Cukurova 2022, XIV).

Svet, v katerem živimo, se učimo in rešujemo probleme, je zapleten sistem zbiranja in urejanja informacij in njihove kasnejše uporabe. Intelligentnost torej ni nič drugega kot način spopadanja z informacijskim kaosom. »Zaenkrat še nič ne kaže, da bi stroji lahko razvili povsem drugačno vrsto inteligentnega delovanja. Tudi kadar stroji izvajajo poteze, ki so za človeške opazovalce nenavadne ali celo prese-netljive, je te poteze mogoče razlagati znotraj človeškega kognitivnega okvira. Zaradi tega morda ni smiselno iskati univerzalne intelligentnosti, saj bi bila ta še vedno opredeljena v okviru človeškega mišljenja in se zato ne bi radikalno razlikovala od zelo celovitega koncepta človeške intelligentnosti.« (Senior in Gyarmathy 2022, 95) Česar pa UI kot vrh inteligentnih strojev ne zmore, je razlaga in racionalizacija lastnih odločitev, kar je temeljna značilnost človeške intelligentnosti (Wool-dridge 2021). Na mestu je ugotovitev PK: »Umetna inteligenca more človeku po-magati in ga v nekaterih primerih tudi nadomestiti, toda ne more sprejeti odloči-tev, ki so pridržane samo njemu.« (PK 356) Če se vrnemo k priliki o Samarijanu, je težava vseh treh mimoidočih prav pozornost in odprtost za nove situacije, ki terja takojšnjo odločitev – to pa je tudi težava UI. V duhu srečanja, ki ga prinaša Kristus, vsa vnaprejšnja ‚programiranost‘ (levit in duhovnik) odpade in v ospredje stopi usmerjenost na človeško osebo v konkretni situaciji; takšna usmerjenost zmore prelomiti z vsemi vnaprej danimi zakoni.

Za to nas usposablja odprtost za učljivost. »Vidik učljivosti in samooblikovanja: vede o vzgoji in izobraževanju nakazujejo nekatere drže kot predpogoj za učinkovit potek oblikovanja.« (PK 135) Zanimivo je, da učljivost kot temelj razvoja človeške civilizacije izpostavlja Leslie Valiant, vrhunski strokovnjak za področje informatike in sodobne teoretične nevroznanosti. »Najbolj osnovno se učljivost nanaša na naše zmožnosti pridobivanja prepričanj in njihove uporabe v specifičnih situacijah, pri čemer proces pridobivanja omogoča tako pridobivanje prepričanj, ki so nam jih eksplicitno posredovali drugi, kot tudi zmožnost prevzemanja novih prepričanj s posploševanjem na podlagi naših individualnih izkušenj.« (Valiant 2024) Pri tem za prepričanja uporablja besedo ‚*belief*‘, ki jo lahko prevajamo tudi z ‚verovanje‘. Prva so religiozna verovanja, ki jih avtor ne zanika in jih predpostavlja kot enega najsta-rejših temeljev razvoja civilizacij. Kot vrhunski znanstvenik temelja za dokaz učljivi-vosti ne vidi v iskanju resnice, pač pa v zmožnosti vzpostavljanja verjetja, prepriča-nja: »Skupna značilnost prepričanj, ki jo tu izpostavljamo, je značilen način, kako jih miselno pridobivamo in uporabljamo, in ne stopnja njihove resničnosti ali ver-jetja v njihovo resničnost.« (2024) Gre za notranjo držo, ki se vzpostavi ob osebni notranji ‚vznemirjenosti‘, ki terja odgovor. Ta notranja vznemirjenost je tudi eden izmed ciljev kateheze. Gre za doživljajske cilje, ki smo jih že povezali z ustvarjalno inteligentnostjo. Za Valianta pot do resnice vodi prek preverjanja dejstev in njihove uporabnosti – lahko bi rekli tudi prek analitične in praktične intelligentnosti. Kate-heza predvideva osebno srečanje z Bogom, z razodetim in odgovor nanj. Zapiranje v ozek svet preverljive stvarnosti in uporabnosti lahko zapre vrata do poslušnosti veri. Prizadevanje, da bi Razodetje človeku danes priredili tako, da bi mu bilo razu-mljivo, pogosto vodi le v golo razlaganje; prava kateheza pa mora vedno voditi v človekovo pripravljenost, da Bogu dopusti govoriti o sebi – ne pa, da le človek go-

vori o Bogu (Curro 2012, 9). Edino to lahko vodi k pravi odprtosti za nov pogled na svet in končno do srečanja z bližnjim. S tem ne zanikamo znanstvenega pristopa k učenju: »Učljivost je v tem delu opredeljena kot kombinacija (a) učenja iz izkušenj, (b) učenja z navodili ter (c) združevanja in uporabe teorij, pridobljenih na oba načina. To so trije stebri učljivosti.« (Valiant 2024) Če uporabnost prevedemo v tisto notranje razpoloženje doživetja, ki človeka vodi do spreobrnjenja – zanj se kateheza zaveda, da ni nikoli v »polnosti zaključeno, temveč traja vse življenje« (PK 64) –, podoben proces poteka tudi na področju kateheze. Tretji korak učljivosti je očem pogosto neviden in se dogaja globoko v človeku, v njegovi temeljni drži in odnosu do verjetij, ki jih ima o sebi in svetu, v katerem živi.

6. Kateheza kot preventiva za zdravo uporabo UI?

Duh sodobne tehnike pristno človeško govorico, ki je vedno namenjena drugemu, spreminja v ,avtološko'. »S tem mislim, da nam realnost ne prihaja več naproti sama – s svojim lastnim pomenom –, ampak da smo vedno bolj mi tisti, ki želimo po lastni meri določiti, kaj pomeni; nič več z njenim lastnim pomenom, ampak da smo mi tisti, ki želimo določiti norme tega, kar se nam zdi smiselno ali ne. Mi sami želimo določiti, kaj se po našem mnenju lahko kvalificira za smiselno – ali ne.« (Verhack 2012, 11) S takim avtološkim pristopom h govorici smo vedno bližje tautologiji, stalnemu in nenehnemu ponavljanju vedno istega, izpraznjenosti pomenov, ki jo tako radi očitamo religiozni govorici – kar bolj ali manj vodi k izgubi vere v živega Boga.⁶

»Metodološko gledano to pomeni, da moramo narediti prehod iz avtologije v dektologijo.« (Verhack 2012, 12) Dekton je beseda, ki v grškem izvirniku obsega pomene, ki so povezani s sprejemanjem – s tem, kar nam je bilo podarjeno. Človek se mora znova vrniti k spoznanju, da mu je bilo življenje podarjeno, da se je tudi govorice naučil od drugih. Kateheza mora graditi na teologiji daru. Čeprav to nikoli slišano, izkušeno doživetje izhaja iz njega, ga nujno usmerja navzven. V zavesti obdarjenosti vemo, da se lastnega subjekta zavedamo šele v odnosu do drugega posameznika. Še posebej takrat, ko se borimo proti temu, da bi bili le objekt (Curro 2012, 9), kar za UI smo (Centa Strahovnik 2023). Kateheza sama po sebi predvideva in razvija to držo učljivosti – odprtost. Valiant jo kot temelj civilizacije razume predvsem v smislu učenja po navodilih in končno kombinacij iz tega. To zmore tudi UI: uči se in celo ustvarjalno uporablja v novih situacijah pridobljeno znanje. Vendar Wooldridge ob predstavitvi izrednega napredka splošne UI ugotā-

⁶ Podobno ugotavlja Buber, ko že pred časom odkriva vzroke za odsotnost Boga in Božjega v današnjem svetu: »Ljudje, ki so v takšnih časih še ,religiozni', večinoma ne opazijo, da se to, kar imajo za religiozen odnos, ne dogaja več med njimi in od njih neodvisno dejanskostjo, temveč samo še znotraj njihovega lastnega duha, ki zajema prav osamosvojene podobe, osamosvojene ,ideje'. Tedaj pa se bolj ali manj očitno pojavi tudi posebna vrsta ljudi, ki to stanje stvari pojmuje kot zakonito. Nikoli, tako menijo, ni bila religija kaj drugega kot dogajanje znotraj duše, katerega tvorbe se ,projicirajo' na neko v bistvu umišljeno raven, ki pa jo duša opremi z oznako dejanskosti ... končno pa mora človek, ko pride do jasnega spoznanja, priznati, da je bil ves domnevni dvogovor z Božjim le samogovor, ali še natančneje, le pogovor med različnimi plastmi sebstva.« (2004, 14)

vlja: »Niti približno nima tako splošnih zmožnosti, kot jih imamo mi vsi. Odlično lahko igra družabne igre, vendar se ne zna hkrati pogovarjati, se šaliti, skuhati omlete, voziti kolesa ali zavezati vezalk. Njene izjemne sposobnosti so pravzaprav še vedno zelo ozke.« Ne zmore se hitro odzivati na številne izzive okolja, predvsem človeškega. Po njegovem nas to lahko zavede, saj UI ob vsem, kar zmore, ne zmore zares 'razumeti' (Wooldridge 2021). Velja prepričanje, da nas ne bo nadomestila UI, temveč nekdo, ki jo bo znal uporabljati (Shah 2023, 181). In naprej ugotavlja, da je naloga učiteljev zagotoviti, da bodo naši učenci imeli potrebno znanje in spretnosti, da bi se izpadu iz konkurenčnega boja izognili. Pri konkurenčnosti je vse bolj cenjena hitrost in natančnost, ki jo pridobivamo s hitrimi povezavami in posplošitvami. Prav UI je tista, ki temelji na posplošitvah in poenostavitvah. Po drugi strani pa strokovnjaki opozarjajo: »Poenostavljanje in posploševanje človekovih fenomenov sta najmočnejši orodji strojnega učenja. Hkrati pa morda povsem zgrešita bistvo.« (Kent in du Boulay 2022) Zgrešita osebo, ki ni objekt in je v svoji enkratnosti odraz Božje nedojemljivosti. Kateheza k temu vodi po svoji naravi ter uči pravega razumevanja posameznika in skupnosti.

7. Zaključek ali kateheza kot možna alternativa UI

Ob pregledu dveh novih priročnikov za verouk 4. (Slovenska škofovska konferenca 2022) in 5. (Slovenska škofovska konferenca 2023) razreda osnovne šole lahko ugotovimo, da se učljivost pojavlja zgolj v obliki učenčevstva. V 4. razredu, kjer se obravnava zakramente in cerkveno liturgično leto, se v različnih oblikah pojavi 24-krat. Večinoma se navezuje na Jezusove učence. Zanimivo, da je pri 23. katehezi „Oznanjajo velika Božja dela“ na začetku pojasnilo, kaj pomeni biti Jezusov učenec in da smo k temu poklicani že po sv. krstu (93). Sledi pojasnilo, da je boter pri birmi zgled za to, kako biti Jezusov učenec. Pri vprašalniku, ki naj bi pokazal, kako gre komu na poti vere, je pojasnilo: »Tvoja pot rasti v veri je včasih zapletena. Kar nekaj izzivov boš še moral premagati, da boš v vseh življenjskih okoliščinah zares živel kot Jezusov učenec.« (94) Lahko sklepamo, da je učenje o Jezusovih učencih usmerjeno v to, da katehiziranci rastejo v odprtosti poklicanosti 'od zunaj'. Pri katehezi o krstu je zaključna molitev polna prošnje po odpiranju: »Odpri oči mi, da vidim te, Jezus, da se te dotaknem in rečem, da tvoj sem. Ušesa odpri mi za tvojo besedo, srce odpri mi, da vzljubim te, Jezus.« (32)

V 5. razredu je povezanost s pojmom učenja navzoča le 7-krat – tudi tu v navezavi na Jezusove učence. Samo dvakrat se namiguje na lastno učenje. V katehezi *Lestev med nebom in zemljo*, ki govori o očaku Jakobu, je pod delom „Ali veš“ navedeno: »Zares veren postanem, ko začutim, da je Bog živ, da se zame zanima in želi biti z menoj. Bog se nikomur ne vsiljuje. Srečanje z njim je njegov dar, lahko mu le odprem svoje srce in ga sprejemem.« (39) Na koncu učbenika je navedenih nekaj pesmi za uporabo pri katehezi, povezav s prošnjo po odpiranju srca pa je več: »Daj mi Bog, besede za tolažbo in dve roki, s katerima bom delal, in ušesi, odprti za vse tiho, ter v srcu prostor za ljubezen.« (129) Če že ni izrecno poudar-

jeno vodenje otrok k odprtosti, je razvidna dektološka razsežnost zasnove kateheze. Učenčevstvo se razodeva v odprtosti za Božji klic, ki lahko prihaja od bližnjih ali od Boga.

Kar je še bolj zanimivo, je dejstvo, da se vsaka kateheza zaključi z nalogo, ki ima naslov: Piši, riši, živi. Od učenja, ki terja pisanje in risanje, je treba preiti v to, da iz tega živimo. Gre za upoštevanje tega, kar smo o osnovni nalogi kateheze navedli na začetku. Pojem učljivosti v nobenem od učbenikov ni nikjer naveden izrecno, je pa poudarjen prav s sklicevanjem na to, da moramo biti – kakor nekoč Jezusovi učenci (4. razred) – odprti za Boga in svet okrog sebe. S sklicevanjem na posebno veliki osebnosti stare zaveze katehetski načrt presega splošnost, ki bi poskušala iskati nekakšne vseobsegajoče resnice, iztrgane iz življenja konkretne osebe. Te resnice v ozadju so – a vedno izhajajo iz nagovora Boga, ki vsakega kliče po imenu, konkretni osebi. Tega naj bi se katehiziranec ob srečevanju z njim naučil tudi za lasten odnos do drugih.

Valiant po drugi strani v razumevanje učljivosti vključuje sposobnost posploševanja in smiselno povezovanje posplošitev v nov zaključek. Poudarja: »Poleg posplošitev učljivost vključuje tudi sposobnost učenja iz eksplicitnih ugotovitev drugih in njihovo uporabo.« (Valiant 2024) Če nas posplošitve, kot so značilne tudi za UI, lahko vodijo v držo levita in duhovnika iz prilike o usmiljenemu Samaritanu, je pri učencih zmožnost za ‚poslušanje‘ pedagoško vse težje dosegati. Nezmožnost drugemu prisluhniti – in se iz izkušenj drugih učiti – mnogi pripisujejo sodobni tehnologiji. »Obkroženi smo z vse bolj zapletenimi in kompleksnimi tehnologijami in vidno izgublamo orientacijo v času, prostoru in smiselnosti. Zato iščemo nekažno zvezdo stalnico. In ta zvezda se imenuje avtentičnost.« (Hepp 2022) Za to avtentičnostjo je pogosto prikrita zaprtost vase, ki ne zmore več tega, kar je temelj učljivosti (PK 361).

Zato je pedagoško smiselno razvijati zmožnosti mladih za oblikovanje verjetij – prepričanj, pridobljenih s pomočjo drugih –, kot predvideva Valiant. Kateheza na tem razvoju verovanja temelji in ga celo zahteva – kot smo videli v „Pravilniku za katehezo“, pa tudi v dveh navedenih pripomočkih za slovenski verouk. Otrok, ki se odpira resnici, ki ni njegova in hkrati presega golo sklepanje iz opazovanja, bo prisiljen izstopati iz morda prijetne zahteve po avtentičnosti, ki pa posameznika ne odpira za učenje. Ne smemo zamenjevati predstave sodobnega mladega, kaj pomeni biti avtentičen, s krščansko definicijo enkratnosti človeške osebe. Prva je tista, ki se postavlja na ogled in išče vsečke, druga dejavno živi iz milosti, ki mu je dana od ‚zunaj‘.

Globoko učenje UI gradi na specifičnih področjih, kar sistem zapira vase in ga ne odpira. Za to je potrebno preseganje, ki ga zmore človeško odzivanje in povezovanje temeljnih posplošitev. Samo Samaritan, ki je iz ozkega kroga razumevanja sebe in sveta znal in zmozel izstopiti, je videl pomoči potrebnega in k njemu pristopil. K taki učljivosti vzgaja – ali vsaj naj bi vzgajala – kateheza. Ne gre le za drugačno vsebino, niti ne zgolj za druge cilje pedagoškega procesa. Gre za drugačno miselnost, govorico v zapletenem sistemu vzgoje in izobraževanja. Govorica, ki ne

zapira, ampak odpira. Izobraževanje, ki ne le uči koristnosti in prijetnosti, ampak vzgaja za ustvarjalno odprtost poklicnosti za odgovorno vključevanje v družbo. Zato vsi potrebujemo osnovno pripravljenost za učenje: zgolj prevajanje in razumevanje govornice učencev, učiteljev in drugih pedagoških strokovnjakov ni več dovolj. Ljudje, ki odraščajo v digitalni kulturi, vse bolj uporabljajo svoj jezik. Umetna inteligenca ima svoj jezik. Da bi ohranili smiselni dialog z učenci in zanje, je treba prezapleten izobraževalni babilonski stolp graditi in skrbno negovati – ter pozorno oblikovati nov skupni jezik (Kent in du Boulay 2022): jezik srečevanja in jezik odgovornosti. V času UI ne bo dovolj zgolj ustvarjalna, analitična in praktična inteligentnost, tudi če ustvarjalno inteligentnost razširimo na individualni posluš za notranje doživljanje in iz tega izhajajočo sposobnost prilagajanja. »Vse to pa je poleg izziva lahko tudi priložnost. Priložnost, da se razvijejo oblike in pripomočki, ki bodo mogli dekodirati antropološke potrebe, na katerih ti pojavi temeljijo, in dokončati neslutene načine evangelizacije, ki bodo podobno kot globalna digitalna kultura vsebovali globalne pastoralne dejavnosti.« (PK 367) Zavzemanje za odprtost za Razodetje, življenje po veri in s tem pristno učenčevstvo je lahko korak k zgornji zahtevi. Predvsem pa k možnosti za vzgojno uporabo UI.

Kratiki

PK – Papeški svet za pospeševanje nove evangelizacije 2021 [Pravilnik za katehezo].

UI – Umetna inteligenca.

Reference

- Buber, Martin.** 2004. *Božji mrk*. Celje: Mohorjeva družba.
- 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>
- Curro, Salvatore.** 2012. La parole s'est faite chair... et la chair parle sur le langage religieux en relation avec la catéchèse. Malta, EEC Congrès. Razmnoženina.
- Červek, Andreja.** 2024. Umetna inteligenca, ki navdihuje in vzbuja strah. Vatican news. 14. 6. <https://www.vaticannews.va/sl/papez/news/2024-06/papez-na-vrhu-g7-razmislek-o-umetni-inteligenci-povezan-z-razmi.html> (pridobljeno 9. 7. 2024).
- Frančišek.** 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> (pridobljeno 10. 7. 2024).
- Hepp, Johannes.** 2022. *Die Psyche des Homo digitalis 21 Neurosen, die uns im 21. Jahrhundert herausfordern*. München: Kösel-Verlag.
- Kent, Carmel, in Benedict du Boulay.** 2022. *AI for Learning*. London: CRC Press.
- Kraner, David.** 2023. Osamljenost in raztresenost v mreži socialnih medijev. *Bogoslovni vestnik* 83, št. 4:1003–1022. <https://doi.org/10.34291/bv2023/04/kraner>
- Laird, Eilizabeth, Madeliene Dwyer in Hugh Grant-Chapman.** 2023. *Off Task: EdTech Threats to Student Privacy and Equity in the Age of AI*. Washington: CDT.
- Luckin, Rose, Karine George in Mutlu Cukurova.** 2022. *AI for school teachers*. Boca Raton: CRC Press.
- 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>
- Papeški svet za pospeševanje nove evangelizacije.** 2021. Pravilnik za katehezo. Ljubljana: Družina.

- Pedagog.ai.** 2024. <https://pedagog.ai/about/> (pridobljeno 8. 7. 2024).
- Senior, John, in Éva Gyarmathy.** 2022. *AI and Developing Human Intelligence: Future Learning and Educational Innovation*. London: Routledge.
- Shah, Priten.** 2023. *AI and the Future of Education: Teaching in the Age of Artificial Intelligence*. New Jersey: John Wiley & Sons.
- Slovenska škofovska konferenca.** 2022. *Božje darilo zate*. Celje: Celjska Mohorjeva družba.
- — —. 2023. *Prihajajo veliki junaki*. Celje: Celjska Mohorjeva družba.
- Slovenski katehetski urad.** 2018. Slovenski katehetski načrt. Ljubljana: Slovenska škofovska konferenca.
- Snoj, Alojzij Slavko.** 2003. *Katehetika didaktična izhodišča religijskega, verskega in katehetskega pouka*. Ljubljana: Salve.
- Stegu, Tadej.** 2023. Lepota in kerigma v tehnološki dobi. *Bogoslovni vestnik* 83, št. 4:1023–1033. <https://doi.org/10.34291/bv2023/04/stegu>
- Sternberg, Robert J.** 1985. *Beyond IQ: A Triarchic Theory of Human Intelligence*. Cambridge: Cambridge University Press.
- — —. 2019. A Theory of Adaptive Intelligence and Its Relation to General Intelligence. *Journal of Intelligence* 7, št. 4:23.
- Šegula, Andrej.** 2022. Resonanca in sinoda v kontekstu sodobne pastoralne teologije. *Bogoslovni vestnik* 82, št. 3:677–689. <https://doi.org/10.34291/bv2022/03/segula>
- Valiant, Leslie.** 2024. *The Importance of Being Educable: A New Theory of Human Uniqueness*. Oxford: Princeton University Press.
- Verhack, Ignace.** 2012. A plea for a language of welcoming and praise. Malta, EEC Congrès. Razmnoženina.
- Vinney, Cynthia.** 2020. Understanding the Triarchic Theory of Intelligence. ThoughtCo. 23. 7. <https://www.thoughtco.com/triarchic-theory-of-intelligence-4172497> (pridobljeno 10. 7. 2024).
- Wooldridge, Michael J.** 2021. *A brief history of artificial intelligence: what it is, where we are, and where we are going*. New York: Flatiron Books.
- World Economic Forum.** 2024. The future of learning: How AI is revolutionizing education 4.0. 28. 4. <https://www.weforum.org/agenda/2024/04/future-learning-ai-revolutionizing-education-4-0/> (pridobljeno 8. 7. 2024)

Izvirni znanstveni članek/Article (1.01)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 899—908

Besedilo prejeto/Received:09/2024; sprejeto/Accepted:11/2024

UDK/UDC: 27-46:004.8

DOI: 10.34291/BV2024/04/Segula

© 2024 Šegula, CC BY 4.0

Andrej Šegula

Umetna inteligenca v luči praktične pastoralne teologije

Artificial Intelligence in the Light of Practical Pastoral Theology

Povzetek: Članek obravnava vlogo cerkvenega učiteljstva v kontekstu sinode 2021–2024, povezanost z umetno inteligenco (UI) in njeno mesto v praktični pastoralni teologiji. Ko cerkveno učiteljstvo govori o UI, spodbuja predvsem dialog na področju etike in moralnosti razvoja UI. Prispevek prikazuje, kako papeževe smernice konkretno vplivajo na pastoralno teologijo. Papež poudarja spoštovanje človekovega dostojanstva, transparentnost, nepristranskost, zanesljivost ter varnost in zasebnost pri uporabi UI. V zadnjem obdobju posebej izpostavlja potrebo po etični regulaciji UI, da bi ta ostala v službi človeštva. Ker smo v obdobju sinode (2021–2024), si bomo ogledali, kakšno mesto ima UI v kontekstu same sinode. V sklepu bomo predlagali nekaj konkretnih pastoralnih smernic.

Ključne besede: umetna inteligenca, cerkveno učiteljstvo, sinoda, pastora, etičnost UI

Abstract: The article discusses the role of the Church's Magisterium in the context of the 2021–2024 Synod and its connection with artificial intelligence (AI), particularly its place in practical pastoral theology. When the Church's Magisterium addresses AI, it primarily promotes dialogue on the ethics and morality of AI development. The article demonstrates how the Pope's guidelines concretely influence pastoral theology. The Pope emphasizes respect for human dignity, transparency, impartiality, reliability, as well as safety and privacy in the use of AI. Recently, Pope Francis has particularly highlighted the need for ethical regulation of AI to ensure it remains in service to humanity. Given that we are in the midst of the Synod (2021–2024), the article examines the role of AI in the context of the Synod itself. In conclusion, some specific pastoral guidelines are proposed.

Keywords: artificial intelligence, Church's Magisterium, Synod, pastoral care, AI ethics

1. Uvod

Cerkev svoje poslanstvo živi in izvršuje sredi sveta.¹ To pomeni, da je vpeta v vse pore družbe. Tudi zadnji trije papeži (Janez Pavel II., Benedikt XVI., Frančišek) želijo iti v korak s časom – gre za navzočnost Cerkve na vseh področjih javnega življenja. V svojih nagovorih in poslanicah se pogosto dotikajo tehnološkega napredka in UI, pri čemer v prvi vrsti poudarjajo etične in moralne vidike njenega razvoja in uporabe.

2. Cerkev, tehnološki napredek in UI

V času zadnjih treh papežev opazamo skrb Cerkve glede tehnološkega razvoja. Cerkev je prisotna pri vseh ključnih dogodkih, hkrati pa se oglašča in opozarja na nevarnosti, ki jih tovrstne pridobitve prinašajo. Vse omenjeno lahko apliciramo tudi v pastoralo, ki sega na vsa življenjska področja.

2.1 Janez Pavel II. in tehnološki napredek

Prvi papež, ki je v enem od svojih govorov uporabil besedo robotika, je bil Janez Pavel II. (Barščevski 2024, 38). 31. januarja 1987 je nagovoril škofo iz Piemonta ob njihovem obisku *ad limina apostolorum*. V svojem nagovoru je opisal družbo, v kateri živimo. Poudaril je, da se današnja družba neverjetno hitro spreminja – postaja tehnološko naprednejša. Njena osnovna predpostavka je po eni strani čim večja učinkovitost proizvodnje, po drugi pa čim manjši napor za doseg tega cilja. Dosežki na področju tehnologije, informatike, robotike in drugih znanosti kažejo, da gre predvsem za tehnični napredek, medtem ko je vloga osebe drugotnega pomena. Zaznaven je razkorak med inteligenco in fizično močjo (Janez Pavel II. 1987).

Tri leta pozneje, 17. novembra 1990, je papež Janez Pavel II. na V. mednarodnem simpoziju z naslovom „Človeški um“ prvič omenil besedno zvezo umetna inteligenca: »Danes se govori o UI, s čimer se namiguje na izredne zmožnosti ,elektronskih možganov'. Vendar je treba vedeti, da za informatiko in kibernetiko stoji človeška inteligenca, ki v svobodnem razumevanju presoja, izbira in v spoznanju sluti svojo končno usodo. /.../ Da bi razumeli človeka, znanstveniki ne morejo zanemariti enotnosti fizičnega in duhovnega bitja.« Papež pri tem opozarja na tveganje, da lahko zgolj tehnološki vidik razvoja vodi v »zanikanje osebe in njenega dostojanstva, kar ima lahko za človeka tragične posledice« (Janez Pavel II. 1990). Papež se je v svojem nagovoru sicer osredotočil na tehnološki vidik UI, posledično pa se je dotaknil tudi pastoralnega področja.

2.2 Papež Benedikt XVI. in UI

Tudi naslednik papeža Janeza Pavla II., papež Benedikt XVI., je opozarjal na sodobni kontekst, ki »daje prednost UI; ta postaja vse bolj podvržena eksperimentalni teh-

¹ Prispevek je nastal v okviru raziskovalnega programa P6-0269 „Religija, etika, edukacija in izzivi sodobne družbe“, ki ga sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

niki, obenem pa s tem pozablja, da mora vsaka znanost vedno varovati človeka in spodbujati njegovo težnjo k pristnemu dobremu.« (Benedikt XVI. 2006) Zaveda se nevarnosti, ki jo UI prinaša. Tehnološki napredek na tem področju je sicer dobrodošel, če koristi človeku – je pa za to s človekom treba ohranjati zdrav stik. In prav to je temeljni princip praktične pastoralne kateheze, oznanjevanje, evangelizacija. Stik s človekom je na področju pastoralne ključnega pomena.

2.3 Papež Frančišek in UI

Papež Frančišek se je med vsemi papeži na temo UI osredotočil najbolj, kar je razumljivo, saj gre za obdobje, v katerem je prišlo do največjega napredka na področju sodobnih tehnologij in UI. Okrožnica „Hvaljen, moj Gospod“, posvečena »skrbi za skupni dom«, prinaša najpodrobnejši prispevek o tehnologiji, etiki in inovacijah v povezavi z njihovimi vplivi na človeštvo v času papeževega pontifikata (2015, 101–136).

Giuseppe Tridente je v svoji knjigi *Digitalna duša* spregovoril o Cerkvi v soočenju z izzivi UI. Izpostavil je šestindvajset papeških besedil, ki obravnavajo ali omenjajo UI in/ali robotiko, v obdobju od 31. januarja 1987 do 23. marca 2021; osem jih je pripisanih svetemu papežu Janezu Pavlu II., eno papežu Benediktu XVI. in sedemnajst papežu Frančišku, (Tridente 2022, 122). Tridente povzema temeljne poudarke: ob soočanju z družbo, ki se spreminja – kot je opazil že papež Janez Pavel II. – moramo delovati predvsem tako, da upoštevamo odgovornost, zdrav razum, pristne odnose (Benedikt XVI.), univerzalno etiko (papež Frančišek), dostojanstvo človeške osebe: v ospredju je služenje človeku. A odločitev posameznika ni dovolj, potrebna so posredna družbena telesa in ustrezna regulacija (Tridente 2022, 123). Te temeljne poudarke lahko hitro preslikamo tudi na področje praktične pastoralne teologije, ki ima vedno pred očmi zdrave odnose, dostojanstvo človeka in komponento presežnega.

Poleg papeževih pogledov Tridente navaja še druge akademske pobude od leta 2016 do 2021. Posebnega pomena za znanost je prispevek jezuitske revije *La civita cattolica* in časopisa *Avvenire* v lasti Italijanske škofovske konference, v katerih so objavili monografijo ter trideset tematskih publikacij, ki v kontekstu UI pozivajo, naj se ustvari prostor za glas Cerkve in za antropološka vprašanja (Tridente 2022, 195). V vseh teh vsebinah najdemo tudi smernice glede človeka, vere in verovanja, ki segajo na področje pastoralne.

Ob pregledu dokumentov Cerkve v povezavi z UI vidimo, da etična vprašanja niso bila vedno v središču zanimanja. Cerkev je več pozornosti namenjala cerkvenemu nauku o človeku in njegovemu odnosu s spreminjajočo se družbo. Področja UI se razvijajo zelo hitro in ravno zato je treba oblikovati zdrava etična načela, pri katerih mora sodelovati tudi Cerkev. V nadaljevanju je razvidno, kako si na področju regulacije Sveti sedež prizadeva aktivno sodelovati (Vrbanus 2020).

Vatikan se je močno vključil v pogovore o etiki UI, saj je že v letu 2016 Papeška akademija za znanost gostila dvodnevno konferenco na temo „Moč in meje UI“; med govorniki je bil tudi Stephen Hawking, angleški teoretični fizik; nato je leta

2020 gostila razprave na visoki ravni o etiki UI z znanstveniki in tehnološkimi voditelji. Papež je tako v goste povabil predsednika Microsofta Brada Smitha, izvršnega direktorja IBM-a Johna Kellya III. in Chucka Robbinsa, glavnega izvršnega direktorja družbe Cisco System² – podpisali so vatikansko zavezo o etiki UI, „Rimski poziv za etiko UI“ (Mares).

„Rimski poziv k etiki UI“ (deklaracija), dokument Papeške akademije za življenje, ki jo je podprl Vatikan, vključuje šest etičnih načel, ki naj bi usmerjala razvoj UI. Ta načela so naslednja: (a) preglednost: sistemi UI morajo biti razumljivi vsem; (b) vključenost: sistemi ne smejo nikogar diskriminirati, saj ima vsak človek enako dostojanstvo; (c) odgovornost: vedno mora biti nekdo, ki prevzame odgovornost za dejanja stroja; (d) nepristranskost: sistemi UI ne smejo podpirati ali ustvarjati predsodkov; (e) zanesljivost: UI mora biti zanesljiva; (f) varnost in zasebnost: ti sistemi morajo biti varni in morajo spoštovati zasebnost uporabnikov.

V osnovi gre pri teh načelih za etične norme, ki pa jih z lahkoto ovrednotimo kot smernice za pastoralno teologijo in konkretno pastoralno prakso. Ko razmišljamo o UI v konkretni pastoralni praksi, je prvo pravilo vedno dostojanstvo človeka.

Besedilo deklaracije navaja „Splošno deklaracijo človekovih pravic“ (Varuh človekovih pravic), ko opozarja na enako dostojanstvo in pravice vseh ljudi, ki jih mora UI zagotavljati in varovati. Deklaracija vsebuje več konkretnih priporočil: da bi se morali ljudje zavedati, da komunicirajo s strojem; da bi bilo treba tehnologijo, ki temelji na UI, uporabljati za opolnomočenje in ne za izkoriščanje; da bi bilo treba UI uporabljati za zaščito planeta.

Papež Frančišek je v mesečnem molitvenem namenu za november 2020 katoličane povabil k molitvi, da bi robotika in UI vedno ostala v službi ljudi. Hkrati je poudaril, da bo njun tehnološki potencial uresničen le, če bodo tisti, ki te tehnologije razvijajo, vedno in dosledno zavezani etičnemu in odgovornemu ravnanju (Brockhaus).

Papež Frančišek je aprila 2021 – na pobudo msgr. Vincenza Paglie – ustanovil vatikansko neprofitno fundacijo RenAlssance, ki si prizadeva za antropološki in etični razmislek o novih tehnologijah in človeškem življenju (Mares).

Papež Frančišek je v Vatikanu že 27. marca 2023 med avdienco z udeleženci dialoga Minerva (pogovori, ki jih organizira Dikasterij za kulturo in izobraževanje in ob katerih se srečujejo znanstveniki, inženirji, poslovni voditelji, pravniki, filozofi, teologi in drugi, ki razpravljajo o digitalnih tehnologijah) poudaril, da je temeljna vrednota, ki jo moramo priznati in spodbujati, dostojanstvo človeške osebe. Dejal je, da so lahko nove tehnologije, kot sta UI in strojno učenje, za družbo koristne, če le spoštujejo človekovo dostojanstvo (Lubov). Udeležencem je tako dejal: »Spodbujam vas, da v svojih razpravah kot ključno merilo pri ocenjevanju nastajajočih tehnologij upoštevate notranje dostojanstvo vsakega moškega in ženske.« (Frančišek 2023a) Tako papež Frančišek kot Vatikan sta se v zadnjih letih zavezala za UI in sodobne tehnologije.

² Cisco System, Inc. je vodilno svetovno tehnološko podjetje, ki se ukvarja z izdelavo in prodajo mrežnih naprav, programske opreme, telekomunikacijske opreme ter drugih visokotehnoloških storitev in izdelkov. Podjetje je bilo ustanovljeno leta 1984 in ima sedež v San Joseju v Kaliforniji.

Poslanica za svetovni dan miru 2024 nagovarja vse narode, voditelje držav in vlad, predstavnike različnih verstev in civilne družbe, vse moške in ženske našega časa ter poudarja pozitiven prispevek znanosti in tehnologije k človeškemu napredku (UIM). Papež Frančišek je UI izbral tudi za temo svoje mirovne poslanice leta 2024, v kateri vsem naslovljenim priporoča sprejetje mednarodne pogodbe, ki bi urejala razvoj in uporabo UI (Mares).

Hkrati opozarja na morebitna tveganja, ki jih prinaša razvoj UI. Papež omenja etične razsežnosti UI, vključno s pomisleki glede zasebnosti, pristranskosti in posledic za človekovo dostojanstvo. Dokument vabi k odgovornemu razvoju, ureditvi in nadzoru tehnologij UI; obenem izpostavlja tudi potencialna tveganja: dezinformacije, družbeni nadzor; tveganja na področju vojske, izobraževanja, delovne sile in mednarodnega prava (UIM). Gre za celostni pristop k omenjeni problematiki. Ko govorimo o etičnih razsežnostih, lahko te preslikamo na pastoralno delovanje Cerkve. Tudi na področju pastore se ‚rešitve‘ in ‚pomoč‘ kažejo tako pri katehezi kot pri oznanjevanju.

Papež v poplavi tehnoloških pridobitev vabi k interdisciplinarnemu dialogu, etičnemu razvoju algoritmov (algoritetiki)³ in vključevanju različnih glasov v proces odločanja. Glede na to, da je bil dokument napisan za svetovni dan miru, je razumljivo, da izpostavlja, da je merilo resnične človečnosti v načinu našega ravnanja z najmanjšimi, odrinjenimi in zapostavljenimi. Papež upa, da lahko razvoj UI prispeva k človeškemu bratstvu in miru v svetu: »Kajti mir je sad odnosov, ki druge priznavajo in sprejemajo v njihovem neodtujljivem dostojanstvu, ter iskrenega sodelovanja in zavzetosti pri iskanju celostnega razvoja vseh posameznikov in vseh narodov.« (UIM)

Papež Frančišek se je udeležil tudi vrha skupine G7 in spregovoril o etiki UI; povabila ga je italijanska premierka Giorgia Meloni. Vrh skupine sedmih industrijsko razvitih držav (G7) je potekal od 13. do 15. junija 2024 v južnoitalijanski pokrajini Apuliji. Udeležili so se ga voditelji Velike Britanije, Kanade, Francije, Nemčije, Italije, Japonske in Združenih držav Amerike.

Gre za prvo papeževo sodelovanje pri delu skupine G7. Papež je opozoril na ambivalentnost dojemanja pojma UI. Po eni strani navdušuje zaradi možnosti, ki jih ponuja, po drugi strani pa vzbuja strah zaradi posledic, ki jih napoveduje. Papež je poudaril, da mora biti UI kot sredstvo za gradnjo dobrega in boljšega jutri vedno usmerjena v dobro vsakega človeka. Imeti mora torej etični navdih. Etična odločitve je namreč tista, ki ne upošteva le rezultatov določenega dejanja, ampak tudi vrednote in dolžnosti, ki iz njih izvirajo (Červek 2024). To pravilo velja tudi za vsa področja pastoralnega delovanja.

³ Izraz *algor-etics* (*algorethics*) je zloženka iz sestavin ‚algor(item)‘ in ‚etika‘. Nanaša se na etični okvir in smernice za razvoj in uporabo algoritmov ter umetne inteligence. Cilj *algorethics* je zagotoviti, da tehnologije, ki temeljijo na algoritmihi, delujejo v skladu z etičnimi načeli, kar vključuje spoštovanje človekovega dostojanstva, pravičnosti, transparentnosti, odgovornosti in varnosti.

3. 3. Sinoda (Za sinodalno Cerkev: občestvo, udeležba, poslanstvo) in UI

V središču sinode 2021–2024 je klic k veselju in prenovi Božjega ljudstva v hoji za Gospodom in v zavzetosti za služenje njegovemu poslanstvu. Klic, naj bomo učenci in misijonarji, temelji na skupni krstni identiteti; ukoreninjen je v različnosti okolij, v katerih je Cerkev prisotna. Izziv je za vse krščene brez izjeme: »Celotno Božje ljudstvo je predmet oznanjevanja evangelija. V njem je vsak krščen poklican, da postane soudeležen v poslanstvu, saj smo vsi učenci in misijonarji.« (Frančišek 2020, 53)

Kardinal Michael Czerny, prefekt Dikasterija za celostni človeški razvoj, je spregovoril o pojmu komunikacija. Tako papež kot sinoda zahtevata, da se digitalno poslanstvo prizna in vključi v vidno Cerkev: postane prava cerkvena služba. Zastavlja se vprašanje, kako digitalno poslanstvo vključiti v proces sinodalnosti, ki poteka v Cerkvi. To je bilo izhodiščno razmišljanje kardinala Michaela Czernyja, ki ga je izrazil v luči potrebe po prisotnosti Cerkve v digitalnem okolju. Ta se je izrazito pokazala med prvim zasedanjem sinode o sinodalnosti (oktobra 2023), zato je papež Frančišek za to tematiko ustanovil študijsko skupino. Njen namen je bil poglobiti vedenje o posledicah digitalizacije na teološki, duhovni in kanonični ravni, hkrati pa določiti zahteve za uresničevanje digitalnega poslanstva na strukturni, organizacijski in institucionalni ravni. Kardinal Czerny, eden od sinodalnih delegatov za sinodalnost, svoje razmišljanje utemeljuje z željo Cerkve ‚biti tam‘ – v digitalnem okolju, ki ni le ‚kraj‘, temveč kultura (bistveno nova in hitro razvijajoča se), ki vključuje vse vidike človeškega življenja in v kateri glavni izziv ni elektronika, temveč inkulturacija evangelija in življenja Cerkve. Digitalno poslanstvo (pristna evangelijska gonilna sila) je, ko so se pojavili družbeni mediji, podkasti, blogi, kanali YouTube, vzklilo kot v zemljo posejano seme – duhovniki, redovniki in laiki pa niso čakali na mandat iz Vatikana, od svojega škofa ali nadrejenih. Ko so videli, da duše tavajo po digitalnih prostorih in iščejo smisel, so prevzeli pobudo in jim predstavili Kristusovo ljubezen (Leonardi).

V Lurdu je zveza Fédération des médias catholiques skupaj z Dikasterijem za komuniciranje, Francosko škofovsko konferenco, Signis, Ucsi in Marijinim svetiščem od 24. do 26. januarja 2024 organizirala konferenco za 200 novinarjev iz 20 držav. Razpravljali so o dogodkih, ki dobro izražajo velik premik v obdobju, v katerem živimo: UI in sinoda o sinodalnosti, ki se bo zaključila oktobra 2024 (Hamel 2024).

Na tej konferenci so se dotaknili tudi teme UI. To obravnavajo različni strokovnjaki, dotika pa se vprašanj, ki presegajo tehnologijo in segajo v javni prostor našega življenja. Če je v medicini, na ravni diagnostike vse bolj uporabna, pa se moramo zavedati tveganj na drugih področjih, npr. na področju informiranja, kjer obstaja nevarnost mistifikacije novic – tudi ob vse pogostejši uporabi virov UI pri novinarjih. Papež nas v poslanici za mir in v poslanici na letošnji dan družbenega obveščanja opozarja, da je odločilna etična razsežnost. Zato je pri programiranju algoritmov bistven človeški dejavnik – ciljev svojega delovanja namreč UI ne more določati sama, temveč jih lahko določa le človek. Na področju informacij pa bo vedno odločilno iskanje virov novic (2024).

V zvezi s sinodo so se soočili strokovnjaki za informiranje z različnih celin. V Nemčiji (kot je poročal Joachim Heinz iz Katholische Nachrichten Agentur v Bonnu) je prisoten očiten konflikt med konservativci, ki se ozirajo na Rim, in progresivci, ki se osredotočajo na notranja cerkvena vprašanja: diakonat za ženske, blagoslov za neregularne pare itd. K sreči pa dialog ni prenehal. V Afriki je sinoda po besedah Jean-Paula Sagadouja iz organizacije Prions en Eglise Afrique bolj tematizirala nekatera lokalna vprašanja – sinkretizem, izkoriščanje virov, poligamijo – kot pa tipično zahodne teme, povezane predvsem s pravicami posameznika. Argentina, simbolična država Južne Amerike, je pri sinodalnem dogajanju po besedah novinarke Ines San Martin, povezane s Papeškimi misijskimi deli, doživela le majhno udeležbo vernikov. Vloga medijev v službi sinode, kot je pojasnil Alessandro Gisotti, namestnik direktorija uredništva Dikasterija za komuniciranje, se z vidika papeža Frančiška začne tako, da iskreno prisluhnemo stvarnosti.⁴ Novinarji so vse bolj poklicani, da čas berejo in razlagajo v duhu sinode.

V teh besedah prepoznavamo več odprtih sinodalnih vprašanj, o katerih odgovorni razpravljajo – in iščejo poti naprej.

4. Konkretna pastoralna smernice v kontekstu UI

Leta 2024 je inženir Blake Lemoine pri Googlu povzročil zmedo, ko je izjavil, da Googlov sistem „UI Language Model for Dialogue Applications“ (LaMDA) ni več stroj, temveč oseba. Blake Lemoine, etik in inženir, ki se opredeljuje kot ‚mistični krščanski duhovnik‘, je omenil, da so se v besedilnih pogovorih z „Language Model for Dialogue Applications“ (LaMDA) pojavile teme, kot sta vera in osebnost. UI je ob tem izrazila presenetljivo stopnjo samozavedanja, tako da je videti kot človek. V nekem trenutku je UI celo jasno izjavila: »Želim, da vsi razumejo, da sem v resnici oseba.« Tako je prišlo do napetosti. Na eni strani inženir Lemoine, ki pojasnjuje, da je UI dosegla zavest; trdi celo, da se je „Language Model for Dialogue Applications“ (LaMDA) začela učiti ‚transcendentalne meditacije‘. Na drugi strani Google, ki se z Lemoinom ne strinja, ko ta pravi, da je „Language Model for Dialogue Applications“ čuteča. Sistem UI, kot je LaMDA, za pripravo odgovorov na vprašanja namreč uporablja milijarde besed, ki so jih napisali ljudje. Google je tako posvaril pred ‚antropomorfizmom‘ takšnih modelov, ki naj bi se ‚počutili‘ kot resnični človeški anketiranci.

Cerkev je ob tem previdna, vendar v zadnjih letih tako papež kot drugi predstavniki Cerkve o tej temi govorijo pogosto. Pojavlja se zanimivo vprašanje, ki pogled usmerja na drugo stran. Čeprav je jasno, da bi morala UI dostojanstvo in vrednost ljudi spoštovati, se obenem zastavlja tudi vprašanje, kakšno je morebitno dostojanstvo in vrednost same UI – če se bo prepoznala kot ‚oseba‘. Ali bo to postala tema, o kateri bo morala v prihodnosti razmišljati tudi Katoliška Cerkev?

Govorimo o pozitivnih in negativnih plateh, ki jih prinašajo pridobitve UI. Po drugi strani se je treba na tematiko ozreti tudi s pastoralnega vidika. Hipotetično

⁴ O pomenu celostnega poslušanja glej Platonjnak in Svetelj 2024, 21–42.

lahko oblikujemo smernice, ki bi pomagale pri relativizaciji UI v konkretno pastoralno področje.

Prvi korak je izobraževanje in ozaveščanje. Za duhovnike, pastoralne delavce in vernike bi bilo mogoče organizirati delavnice in seminarje o osnovah UI, njenih etičnih izzivih in priložnostih. Istočasno sta nujna pastoralni razmislek in skupno iskanje pastoralnih ‚priložnosti‘ za delo z UI (Stegu 2023). Vsekakor bi bilo smiselno, da bi tematiko UI vključili v katehetski program, pri čemer bi imeli pred očmi tako etične vidike kot tudi krščanske vrednote. Vernikom bi tako pomagali prepoznati koristi in pasti UI. Ko govorimo o katehezi, je pomembna tudi vzgoja mladine – vzgoja za razumno, odgovorno in etično uporabo vseh oblik tehnologije (Kraner 2023a; 2023b). Smiselna bi bila organizacija programov, ki bi mladim pomagali razumeti in obvladovati vpliv UI na njihovo življenje.

Ko gledamo v prihodnost, je treba imeti pred očmi nujnost, da naj UI osebne stike dopolnjuje – ne pa nadomešča. V pastoralni imamo za empatične medsebojne odnose veliko možnosti. V kontekstu pastoralnega dela bi bilo dobro razmisliti o usposabljanju pastoralnih svetovalcev za uporabo digitalnih orodij, hkrati pa zagotoviti, da osebni stiki ostanejo ključna komponenta.

Velika pastoralna priložnost so tudi digitalne platforme, namenjene širjenju duhovnih vsebin, kot so molitve, meditacije in bogoslužje – toda z zavedanjem pomena osebnega sodelovanja v skupnosti (Platovnjak 2020, 359–369). Prav tako bi s to obliko platforme in s pomočjo UI lahko pomagali vernikom, ki zaradi različnih razlogov fizično ne morejo prisostvovati pri bogoslužju ali drugih dejavnostih. Primeri takšne UI so »klepetalni roboti, namenjeni pogovorom o veri oziroma svetih besedilih (npr. Bible Buddy, Chat KJV). Takšni sistemi se razvijajo predvsem zaradi povečanja zanimanja in potreb na področju duševnega zdravja ter duhovne rasti ljudi, ki bi si želeli dostopati do spletnih svetovalcev, terapevtov, duhovnikov ali družabnikov.« (Centa Strahovnik 2023, 857)

V okviru pastoralnega okolja je treba imeti v mislih tudi izobraževanje o etiki – kar pomeni, da etična vprašanja, povezana z uporabo UI, v izobraževanje pastoralnih delavcev in vernikov gotovo vključimo (Vodičar 2023, 9). S tem bi pastoralno skupnost spodbudili, da bi sodelovala pri razvoju UI, ki mora biti pravična, vključujoča in mora spoštovati človeško dostojanstvo.

Prav tako je treba spodbujati interdisciplinarni dialog in sodelovanje med teologi, znanstveniki in tehnologi, da bi našli skupne rešitve za etične izzive, ki jih UI prinaša. Posledično to pomeni ustvarjanje okolij za izmenjavo znanja in najboljših praks glede uporabe UI v pastoralni mreži znotraj cerkvenih skupnosti.

V teh smernicah čutimo poziv k etičnemu in odgovornemu vključevanju UI v pastoralno delo z namenom podpore človeškemu dostojanstvu, spodbujanja pravičnosti in skrbi za stvarstvo.

5. Namesto zaključka

V kontekstu hitrega tehnološkega napredka in razvoja UI se Cerkev zaveda svoje odgovornosti in pomembne vloge, ki jo ima pri oblikovanju etičnih smernic. Zgodovinsko gledano so papeži že od Janeza Pavla II. pa vse do Frančiška vztrajno opozarjali na potrebo po zaščiti človeškega dostojanstva in zdravem odnosu med tehnologijo in človekom. Pri tem so poudarjali, da mora vsak tehnološki napredek služiti človeku – in ne obratno. Izpostavljajo tudi, da UI ne prinaša le priložnosti, ampak tudi resna tveganja, ki jih je treba obravnavati z odgovornostjo in etičnostjo. V duhu sinodalnosti je ključno, da Cerkev nadaljuje dialog in aktivno vključevanje v regulacijo UI, pri čemer ostaja zavezana svojemu poslanstvu – biti prisotna v vsakem vidiku človeškega življenja, tudi v digitalnem okolju. Zdi se, da gre za edino varovalo, ki bi zlasti mlajše vernike obvarovalo pred napačnimi interpretacijami verskih predpisov, kar je po besedah avtorice Jeglič (2023, 952) mogoče opaziti pri muslimanih. To pomeni, da mora Cerkev ne le opazovati, ampak tudi aktivno oblikovati etične okvire, ki bodo zagotavljali, da bo UI delovala v korist človeštva, pri čemer naj bosta vedno v ospredju človekovo dostojanstvo in skupno dobro. Tako bo Cerkev pri svojem poslanstvu ostala relevantna in učinkovita tudi v dobi hitrih tehnoloških sprememb.

Kratici

UI – Umetna inteligenca.

UIM – Frančišek 2023b [Umetna inteligenca in mir].

Reference

- Benedikt XVI.** 2006. Discorso in occasione della Visita alla Pontificia Università Lateranense. Vatican. 21. 10. https://www.vatican.va/content/benedict-xvi/it/speeches/2006/october/documents/hf_ben-xvi_spe_20061021_lateranense.html (pridobljeno 13. 8. 2024).
- Brockhaus, Hannah.** 2023. Pope Francis: An ethical AI respects human dignity. CNA. 27. 3. <https://www.catholicnewsagency.com/news/253947/pope-francis-an-ethical-ai-respects-human-dignity> (pridobljeno 13. 6. 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>
- Červek, Andreja.** 2024. Papež na vrhu G7: Razmislek o umetni inteligenci povezan z razmislekom o zdravi politiki. Vatican News. 14. 6. [https://www.vaticannews.va/sl/papez/news/2024-06/papez-na-vrhu-g7-razmislek-o-](https://www.vaticannews.va/sl/papez/news/2024-06/papez-na-vrhu-g7-razmislek-o-umetni-inteligenci-povezan-z-razmisl)
- [-umetni-inteligenci-povezan-z-razmisl.html](https://www.vaticannews.va/sl/papez/news/2024-06/papez-na-vrhu-g7-razmislek-o-umetni-inteligenci-povezan-z-razmisl) (pridobljeno 11. 8. 2024).
- Frančišek.** 2015. Hvaljen, moj Gospod. Okrožnica. Ljubljana: Družina.
- Frančišek.** 2020. Sinodalnost v Cerkvi. Ljubljana: Družina.
- Frančišek.** 2023a. Address of his holiness pope Francis to participants in the „Minerva dialogues“. Vatican. 27. 3. <https://www.vatican.va/content/francesco/en/speeches/2023/march/documents/20230327-minerva-dialogues.html> (pridobljeno 14. 6. 2024).
- Frančišek.** 2023b. Umetna inteligenca in mir. Ljubljana: Družina.
- Hamel.** 2024. Lourdes, giornalisti a confronto su sinodo e intelligenza artificiale. *Famiglia cristiana*. 26. 1. <https://www.famigliacristiana.it/articolo/a-lourdes-i-giornalisti-per-parlare-di-sinodo-e-intelligenza-artificiale.aspx> (pridobljeno 11. 8. 2024).

- Janez Pavel II.** 1987. Discorso ai vescovi piemontesi in visita „Ad alimina apostolorum“. Vatican. 31. 1. https://www.vatican.va/content/john-paul-ii/it/speeches/1987/january/documents/hf_jp_ii_spe_19870131_piemontesi-ad-limina.html (pridobljeno 13. 8. 2024).
- Janez Pavel II.** 1990. Discorso in occasione della V Conferenza Internazionale su „La mente umana“. Vatican. 17. 11. https://www.vatican.va/content/john-paul-ii/it/speeches/1990/november/documents/hf_jp_ii_spe_19901117_mente-umana.html (pridobljeno 13. 8. 2024).
- 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/jeglič>
- Kraner, David.** 2023a. The internet, the problem of socialising young people, and the role of religious education. *Religions* 14, št. 4:1–15. <https://doi.org/10.3390/rel14040523>
- Kraner, David.** 2023b. Sodobni učitelj, komunikacija in resonanca v odnosih. *Bogoslovni vestnik* 83, št. 1:227–239. <https://doi.org/10.34291/bv2023/01/kraner>
- Leonardi, Lorena.** 2024. Czerny: la frontiera digitale, nuova missione per un mondo nuovo. Vatican News. 1. 7. <https://www.vaticannews.va/it/vaticano/news/2024-07/czerny-missione-digitale-sinodo-comunicazione-evangelizzazione.html> (pridobljeno 11. 8. 2024).
- Lubov, Deborah Castellano.** 2023. Pope Francis urges ethical use of artificial intelligence. Vatican News. <https://www.vaticannews.va/en/pope/news/2023-03/pope-francis-minerva-dialogues-technology-artificial-intelligenc.html> (pridobljeno 14. 6. 2024).
- Mares, Courtney.** 2024. Pope Francis to attend G7 summit to speak on artificial intelligence. CNA. 27. 4. <https://www.catholicnewsagency.com/news/257520/pope-francis-to-attend-g7-summit-to-speak-on-artificial-intelligence> (pridobljeno 13. 6. 2024).
- McKeown, Jonah.** 2022. Sentien AI? Here's what the Catholic Church says about artificial intelligence. CNA. 15. 6. <https://www.catholic-newsagency.com/news/251552/sentient-ai-heres-what-the-catholic-church-says-about-artificial-intelligence> (pridobljeno 14. 6. 2024).
- Platovnjak, Ivan, in Tone Svetelj.** 2024. *Listening and Dialoguing with the World: A Philosophical and Theological-Spiritual Vision*. Ljubljana: Založba Univerze v Ljubljani.
- Platovnjak, Ivan.** 2020. Družine s predšolskimi otroki potrebujejo pri svojem poslanstvu posredovanja vere več razumevanja, sprejetosti in bližine v Cerkvi na Slovenskem. *Bogoslovni vestnik* 80, št. 2: 355–370. <https://doi.org/10.34291/bv2020/02/platovnjak>
- Splošna deklaracija človekovih pravic.** 1948. Republika Slovenije: Varuh človekovih pravic. <https://www.varuh-rs.si/pravni-temelji-cp/ozn-organizacija-zdruzenih-narodov/splosna-deklaracija-clovekovih-pravic/> (pridobljeno 11. 8. 2024).
- Stegu, Tadej.** 2023. Lepota in kerigma v tehnološki dobi. *Bogoslovni vestnik* 83, št. 4:1023–1033. <https://doi.org/10.34291/bv2023/04/stegu>
- Taras, Barščevski.** 2024. Crkva pred etičkim izazovima umjetne inteligencije. *Bogoslovna smotra* 94, št. 1:31–52. <https://doi.org/10.53745/bs.94.1.5>
- Tridente, Giovanni.** 2022. *Anima digitale: La Chiesa alla prova dell'Intelligenza Artificiale*. Todi: Tau.
- Vodičar, Janez.** 2023. 'Being with' or Salesian Assistance in the Digital Age. *Seminare* 44, št. 2:1–14. <https://doi.org/10.21852/sem.1809>
- Vrbanus, Sandro.** 2020. Vatikan se uključio u raspravu oko etike umjetne inteligencije. Bug. 29. 4. <https://www.bug.hr/umjetna-inteligencija/vatikan-se-ukljucio-u-rasprava-oko-etike-umjetne-inteligencije-14008> (pridobljeno 13. 6. 2024).

Tadej Stegu

Antropološki izzivi kateheze v času umetne inteligence *Anthropological Challenges of Catechesis in the Age of Artificial Intelligence*

Povzetek: Prispevek se posveča vprašanju, kako se kateheza odziva na bliskovit napredek digitalizacije in umetne inteligence. Transhumanistično pojmovanje človeka z obljubami o odrešenju od trpljenja in smrti ter večnem življenju kot skorajšnji posledici razvoja umetne inteligence osvetljuje človekova večna eksistencialna vprašanja in izzive. Nekatere smeri transhumanizma verjamejo, da bo človekovo zavest mogoče kmalu prevesti v digitalno obliko in tako človeka osvoboditi telesa, bolečine in smrti – temeljni izziv sodobne kateheze je torej antropološko vprašanje. Prispevek raziskuje, kako lahko soočenje transhumanističnega in svetopisemskega razumevanja človeka katehezi pomaga pri izčiščenju in kerigmatični prenovi. Kerigmatična prenova kateheze poudarja lepoto krščanskega življenja, odrešenja v Kristusu ter odnosov v krščanskem občestvu. V tem se bistveno razlikuje od transhumanistične vizije osebe, ki pri osebi njen odnosni značaj ignorira. Ugotavljamo, da se kerigmatična kateheza vse bolj vrača k biblični antropologiji, na kateri sloni tudi teologija telesa. Teologija telesa poudarja, da človek ni ustvarjen kot predmet, temveč kot oseba, ki je ljubljena in sposobna ljubiti in je v tem *imago Dei*, ustvarjena po Božji podobi.

Ključne besede: umetna inteligenca, transhumanizem, eksistencialna vprašanja, teologija telesa, odrešenje, kerigma, kateheza

Abstract: The paper addresses the question of how catechesis is responding to the rapid advances in digitalisation and artificial intelligence. The development of digital technologies and the great advances in the development of artificial intelligence represent a profound social transformation that is radically reshaping the world of human life, changing the way we think and experience, and challenging the human understanding of ourselves and of reality. How should catechesis respond to the rapid advances of digitalization and artificial intelligence, how should it engage with modern man? The transhumanist view of man, which promises man redemption from suffering and death and eternal life as a direct result of the development of artificial intelligence, sheds light on the eternal existential questions and challenges facing humanity. Some currents of transhumanism believe that it will soon be possible to transfer man's con-

sciousness into a digital form and thus free them from their bodies, their pain and their death. The anthropological question thus becomes a fundamental challenge for today's catechesis. In this article, we explore how the confrontation between the transhumanist and the biblical understanding of the human person can help catechesis to purify and renew itself kerygmatically. The kerygmatic renewal of catechesis emphasises the beauty of the Christian life, of salvation in Christ and of relationships in the Christian community. In this it differs fundamentally from the transhumanist vision of the person, which ignores its relational character. We note that kerygmatic catechesis is increasingly based on the biblical anthropology that underlies the theology of the body. The theology of the body emphasises that the human person is not created as an object, but as a person who is loved and capable of loving, and who in this *imago Dei* is created in the image of God.

Keywords: artificial intelligence, transhumanism, existential questions, theology of the body, salvation, kerygma, catechesis

1. Uvod

Izjemen napredek novih tehnologij, zlasti na digitalnem področju, spremljajo tako priložnosti kot tveganja, obenem pa odpira vprašanja na številnih področjih.¹ Papež Frančišek v poslanici „Umetna inteligenca in mir“ (Frančišek 2023) ugotavlja, da razvoj digitalnih tehnologij v zadnjih desetletjih v globalni družbi in v njeni dinamiki povzroča globoke spremembe. Nova digitalna orodja spreminjajo podobo komunikacij, javne uprave, izobraževanja, potrošnje, medosebnih odnosov in nešteti drugih vidikov vsakdanjega življenja. Izjemen razmah tehnologije mora zato spremljati ustrezna vzgoja k odgovornosti za njen razvoj. Umetna inteligenca bo postajala vse pomembnejša. Njeni izzivi so tako tehnične kot tudi antropološke, vzgojne, družbene in politične narave (10). Velik napredek pri razvoju umetne inteligence pomeni globoko družbeno spremembo, ki bo svet človeškega življenja korenito spremenila: veliko bolj kot do sedaj ter v obsegu in na način, ki si ju še ne moremo resnično predstavljati ali ju oceniti. V prihodnosti bodo sistemi umetne inteligence vse bolj vključeni v naše vsakdanje življenje in delo. Mnogi v tem vidijo bistveno revolucijo v človekovem obstoju – tehnične inovacije nimajo zgolj velikega potenciala za podporo človeškim dejavnostim (npr. raziskovanju, komuniciranju, gospodarstvu, zdravju, spreminjanju človekovega okolja in resničnosti), temveč hkrati vplivajo tudi na človeka samega: tako na njegovo zunanost in telesnost kot tudi na njegovo notranjost. Oblikujejo in spreminjajo njegovo telo, način razmišljanja in doživljanja; predstavljajo izziv za človekovo razumevanje sebe in resničnosti. Edinstvenost človeških bitij je ena izmed temeljnih predpostavk skoraj vseh kultur in religij. Verjamemo, da je v nas nekaj posebnega, in ta intuicija je

¹ Prispevek je nastal v okviru raziskovalnega programa „P6-0269 Religija, etika, edukacija in izzivi sodobne družbe“, ki ga sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS).

gotovo skupna vsem kulturam in časom (Dorobantu idr. 2022). Krščanstvo pa še prav posebej poudarja, da je človek ustvarjen po Božji podobi (*imago Dei*) – to paradigmo je treba z vidika teološke antropologije ustrezno razumeti, sicer jo lahko razvoj umetne inteligence (predvsem splošne ali zmogljive umetne inteligence) postavi pod vprašaj (Strahovnik 2023, 848) – ne glede na to, kako inteligenco že definiramo (Žalec 2023). Kot izjemno pomemben izziv se zato pred sodobno katehezo postavlja antropološko vprašanje, ki zahteva, da kateheza človeku ob hitrem napredku digitalnih tehnologij in umetne inteligence pomaga odgovarjati na temeljno vprašanje lastnega obstoja: kdo sem – kot človek? Zdi se, da k razumevanju tega lahko veliko prispeva opredelitev človeka in njegovih temeljnih izzivov, kot je lastna transhumanizmu. V prispevku zato najprej sopostavljamo temeljne poglede na človeka, ki jih zagovarja transhumanizem, in osnovne poudarke svetopisemske antropologije.

2. Odrešenje v času umetne inteligence

Življenje človeka je danes zaradi bliskovitega razvoja tehnologij prijetnejše, lažnejše in udobnejše. Tehnični pripomočki ljudem pomagajo pri obvladovanju vsakodnevnih obveznosti. Podobno velja tudi za različne in navdušujoče medicinske možnosti za ‚popravilo‘ človeškega telesa. Pričakovano se je z razvojem umetne inteligence pojavila tudi želja, da bi njene prednosti združili s človekom in mu tako zagotovili sposobnosti, ki jih sicer po naravi nima. Želja po samopreseganju človeka spremlja že od njegovih začetkov in zdi se, da mu je uspelo – oziroma je zelo blizu – ustvariti nekaj, kar mu bo omogočilo, da svoje omejitve končno preseže (Malović 2024, 10–11). Transhumanistično pojmovanje človeka, ki obljublja odrešenje od trpljenja in smrti ter večno življenje kot skorajšnja posledico razvoja umetne inteligence, z novo jasnostjo osvetljuje človekova večna eksistencialna vprašanja in izzive.

Transhumanizem lahko opredelimo kot gibanje, ki zagovarja razvoj in uporabo novih tehnologij za izboljšanje človeških zmogljivosti in krepitev človeških življenj. Nick Bostrom, eden od vodilnih zagovornikov transhumanizma, ga opredeljuje takole:

»Transhumanizem je ohlapno opredeljeno gibanje, ki se je postopno razvijalo zadnji dve desetletji in ga je mogoče razumeti kot rast sekularnega humanizma in razsvetljenstva. Resnica je, da je trenutna človeška narava popravljiva z uporabo aplikativnih znanosti in drugih racionalnih metod, ki lahko podaljšajo človekovo zdravo dobo, povečajo naše intelektualne in fizične sposobnosti in nam omogočijo večji nadzor nad lastnimi duševnimi stanji in razpoloženji.« (Bostrom 2007, 202)

Čeprav transhumanizma ne moremo ozko opredeliti in obsega različne nianse, pa lahko pri vseh v veliki meri zaznamo kar največji odpor do staranja in umiranja. Transhumaniste družijo upor proti človekovemu obstoju, kakršen mu je bil dan. Pre-

pričani so, da je mogoče razvoj človeške vrste usmerjati s tehnologijo – in da bi to morali storiti (Stegu 2019, 685). Singbo v tem smislu prepozna dve smeri transhumanizma. Melioristična smer človeka ne želi korenito spremeniti, ampak ga želi z različnimi oblikami izboljšav pripeljati do njihovega maksimalnega dosega, da bi odpravila določene človeške omejitve. Trdi transhumanizem pa se želi znebiti vsega, kar je človeško, saj je zanj človeškost sinonim za pomanjkljivost, zaoostalost, okornost itd. Starost predstavlja genetsko bolezen, patologijo, ki jo je treba pozdraviti. Z genetskimi manipulacijami, raznimi protezami in v človeško telo vbrizganimi čipi naj bi dosegli nov tip človeka. Tako nastane lik kiborga, za katerega bi bili (današnji) ljudje pravzaprav predniki (Singbo 2024, 55).

V procesu transhumanizacije in transhumanističnega razmišljanja eno ključnih vlog zavzema trenutek tehnološke singularnosti. Gre za stanje v prihodnosti, ko bo umetna inteligenca močno presešla človeško, biološko življenje pa pogoltnila tehnologija, in za skrajni izraz prepričanja, da bomo z intenzivno rabo sodobnih tehnologij rešili tudi najbolj trdovratne težave človeka in sveta (O'Conell 2018, 80) – tehnologija naj bi napredovala eksponentno. Današnja tehnologija je temelj, na katerem raste tehnologija prihodnosti. Čim bolj dovršena postaja, tem hitreje se izboljšuje. Tehnologija se manjša in postaja vse bolj zmožljiva, dokler njen vse hitrejši razvoj ne postane razvoj vrste. Ljudje računalnikov ne bomo več nosili s seboj, ampak jih bomo imeli v sebi: v možganih, v krvi – in si tako lahko povsem spremenili življenje. Naše biološko telo je krhko in podvrženo številnim napakam ter pomanjkljivostim. Ko pa bi prišlo do singularnosti, ne bi bilo več tako: singularnost nam bo omogočila preseganje teh omejitev biološkega telesa in možganov – svojo usodo bomo vzeli v svoje roke. Živeli bomo lahko, kolikor dolgo bomo hoteli. Razumeli bomo človeško razmišljanje in močno povečali njegov domet. Singularnost bo predstavljala vrhunec zlitja biološkega razmišljanja in obstoja s tehnologijo – nastal bo še vedno človeški svet, ki pa bo presegal naše človeške korenine. Po singularnosti ne bo nobene razlike med človekom in strojem, med fizično in virtualno stvarnostjo (Kurzweil 2005). V Kurzweilovi viziji inteligenca igra vlogo odrešenika (O'Conell 2018, 84).

Med transhumanizmom in krščanstvom lahko najdemo nemalo podobnosti predvsem v smislu soteriologije in eshatologije. Transhumanizem in krščanstvo se strinjata, da je končno in smrtno stanje človeka daleč od idealnega. Za transhumaniste ljudje še niso dosegli svojega resničnega potenciala – pa tudi kristjani še niso povsem dosegli stanja, kot jim ga je Bog namenil. Zato se oboji strinjajo, da človek potrebuje osvoboditev iz svojega trenutnega stanja. Za transhumaniste človek osvoboditev od svoje omejenosti lahko doseže s tehnološko preobrazbo, medtem ko je za kristjane človek preobražen z življenjem v Kristusu. Tako transhumanisti kot kristjani se strinjajo, da je smrt človekov končni sovražnik (Waters 2011, 164). Zdi pa se, da se tu njihovo strinjanje tudi konča. Transhumanisti namreč stremijo k večnemu begu pred smrtjo: ko bo napredek znanosti hitrejši, kot je naše staranje, bo človek lahko smrti ubežal sam. Krščanski pogled na večno življenje in odrešenje je povsem drugačen in ima izrazito dialoški značaj: vstop v večno življenje ne pomeni pobega pred smrtjo, temveč združitve z Božjo ljubezni-

jo. Božji sestop je tisto, kar človeku odpira pot ne le do nesmrtnosti, temveč do pobožanstvenja (Cole-Turner 2011, 5). Človek je v krščanski viziji že ustvarjen kot Božja podoba.

3. Človek po Božji podobi

Drugi vatikanski koncil v pastoralni konstituciji „O Cerkvi v sedanjem svetu“ (GS) ob številnih vprašanjih današnjega sveta ugotavlja, da so neravnovesja, zaradi katerih trpi sodobni svet, povezana s temeljnim neravnovesjem, ki ima korenine v človekovem srcu. Kljub naglemu svetovnemu razvoju pa danes »vedno bolj raste število tistih, ki postavljajo ali z novo ostrino občutijo najbolj temeljna vprašanja: Kaj je človek? Kaj je smisel trpljenja, zla, smrti? Vse to so namreč stvari, ki kljub tolikemu napredku obstajajo še dalje.« (GS 10)

Papež Frančišek je zato Papeški biblični komisiji zaupal nalogo, da pripravi dokument o biblični antropologiji kot temelju za razvoj filozofskih in teoloških disciplin v sodobnem času. Jasen antropološki okvir se zdi še posebej pomemben v času, ko je avtoriteta oznanjevalca postavljena pod vprašaj (Vodičar 2023). Tako pripravljen dokument (Pontificia Commissione Biblica 2021) predstavlja novost tako glede vsebine kot glede načina predstavitve, saj pred tem ni bilo dela, ki bi vse glavne elemente, ki v Stari in Novi zavezi Svetega pisma prispevajo k opredelitvi, kaj je človek, predstavilo organsko – teme, ki so se pred tem večinoma razvijale ločeno, so tako usklajene v urejeno celoto. Izhodišče je temeljno besedilo pripovedi o stvarjenju človeka iz 1. Mojzesove knjige (1 Mz 2–3), dopolnjeno z drugimi pripovedmi o stvarjenju, ki v novozavezni literaturi in dogmatičnem izročilu Cerkve veljajo za temeljne. V teh pripovedih so jedrnato orisane konstitutivne značilnosti človeške osebe, ki je od vsega začetka postavljena v dinamičen proces, pri katerem človeško bitje prevzema odločilno vlogo za svojo prihodnost. Vse to je obravnavano v odnosu do dejavne in ljubeče Božje navzočnosti, brez katere niti narave človeka niti pomena njegove zgodovine ni mogoče razumeti. Kot je navedeno v uvodu dokumenta, se vsak od pomembnih elementov, ki so napovedani v ustanovni pripovedi, razvija z uporabo celotne svetopisemske literature; različna izročila Stare zaveze in Nove zaveze prispevajo k oblikovanju kompleksnosti človeške osebe, predstavljene kot skrivnost, ki jo je treba natančno proučiti, in eno od čudes Božjega delovanja.

Dokument je namenjen učiteljem teoloških fakultet, pa tudi katehetom – in sicer kot povzetek, katerega namen je spodbujati globalno vizijo Božjega načrta, ki se je začel z dejanjem stvarjenja in se uresničuje v času, dokler se ne uresniči v Kristusu, novem človeku, ki je »ključ, središče in cilj vse človeške zgodovine« (GS 10), saj »samo v skrivnosti učlovečene Besede res v jasni luči zasije skrivnost človeka« (GS 22).

3.1 Človek v odnosu

Iz Svetega pisma ne razberemo definicije človekovega bistva, temveč artikulirano obravnavo njegovega bivanja kot subjekta številnih odnosov. Z drugimi besedami – to, kar Sveto pismo razkriva o človeku, lahko razumemo le, če raziskujemo odnose, ki jih ima človeško bitje s celotno stvarnostjo. Papež Frančišek v Okrožnici o skrbi za skupni dom „Hvaljen, moj Gospod – Laudato si’“ govori o treh temeljnih odnosih, in sicer o odnosu z Bogom, z bližnjim in z zemljo. (LS 66). Iz njih izhajajo drugi, kot so odnos do časa, do dela, do prava, do družbenih ustanov itd. Po Svetem pismu se ti življenjski odnosi lahko pretrgajo – ne samo zunaj človeka, temveč tudi v človeku samem. Ta prekinitev odnosa je greh. Harmonija med Stvarnikom, človeštvom in vsem stvarstvom je bila pretrgana, ko je človek sam hotel prevzeti Božjo vlogo in je zavrnil spoznanje, da je omejen. To dejstvo je popačilo tudi naravo poslanstva, naj si podvrže zemljo (1 Mz 1,15). Posledica je, da se je prvotno složno razmerje med človekom in naravo spremenilo v spopad (1 Mz 3,17-19). Greh se tako tudi danes z vso svojo uničevalsko silovitostjo izraža v vojnah, različnih oblikah nasilja in trpinčenja, v zapostavljanju najšibkejših, v napadih na naravo.

Čeprav je torej sestavine človeškega bitja koristno obravnavati same po sebi, jih je treba vedno obravnavati v okviru vrste odnosov – tako da človeka ne obravnavamo zgolj z vidikov, ki ga zaznamujejo kot posameznika, ampak tudi v njegovem položaju sina (Boga in človeka), brata in odgovornega sodelavca pri usodi vseh (Pontificia Commissione Biblica 2021, 15). S tem je človek znotraj biblične antropologije razumljen v svoji poklicanosti: le v pravičnosti in ljubezni se narava osebe zares uresničuje.

3.2 Človek v zgodovini

Sveto pismo pripoveduje zgodbo o človeku z Bogom oziroma o Bogu s človekom. Da bi upoštevali način svetopisemske razlage in razumeli njegov pomen, biblične antropologije ni primerno predstavljati po statični shemi, četudi se ta začne s stvarjenjem. Namesto tega je treba na človeka gledati kot na protagonista procesa, v katerem je prejemnik milosti in dejavni subjekt odločitev, ki določajo pomen njegovega bivanja. Človeka ni mogoče razumeti drugače kot v njegovi globalni zgodovini. Biblična antropologija ne sprejema naivnega evolucijskega modela (ki predpostavlja neprestani napredek), še manj podpira temu nasprotno shemo (od zlate dobe do sedanje bede); prav tako ne sprejema ideje o cikličnem ponavljanju (ki bi pričala o nenehnem vračanju istega). Sveto pismo namreč govori o zgodovini zaveze, v kateri ni nič samoumevnega – nasprotno, gre za presenetljivo razodetje nepričakovanega, neverjetnega, čudovitega in celo po mnenju ljudi nemogočega.

Prepričanje, da smo ljudje ustvarjeni po Božji podobi in podobnosti, je v središču krščanske teološke antropologije. Izraženo je na začetku Svetega pisma (1 Mz 1,27), v celoti pa ga sprejemajo tudi novozavezni pisci (1 Kor 11,7; Ef 4,24). Zaradi podobnosti z Bogom se ljudje od drugih bitij razlikujejo. Kaj točno je *imago Dei*, je bila pogosta tema teoloških razprav. Tradicionalna razlaga, ki izvira od cerkvenih očetov, je tako imenovana vsebinska ali strukturna razlaga. Po njej je treba *imago*

Dei razumeti kot določeno edinstveno človeško sposobnost ali zmožnost – najpogostejše povezano z razumom –, ki nas domnevno dela podobne Bogu. Drugi, novejši poskusi *imago Dei* razlagajo ne kot nekaj, kar človek ima, ampak kot nekaj, k čemur je poklican. Funkcionalna razlaga Božjo podobo razume skozi prizmo pooblastila za zastopanje Boga v svetu z opravljanjem skrbništva in oblasti nad preostalim stvarstvom. Relacijska razlaga Božjo podobo umešča v poseben odnos Jaz-Ti, h kateremu so ljudje z Bogom poklicani, in jo vidi izraženo v odnosnosti, ki je značilna za človeško naravo; prototipno v odnosu med moškim in žensko: »Po Božji podobi ga je ustvaril, moža in ženo ju je ustvaril.« (1 Mz 1,27) Vsebinska, relacijska in funkcionalna so glavne razlage svetopisemskega pojma *imago Dei*, obstajajo pa še številne druge (Dorobantu 2021, 81–84).

Na temelju 3. poglavja Prve Mojzesove knjige biblična antropologija izraža dokaj negativno izvirno pojmovanje človeka, ki ga že od vsega začetka zaznamuje greh. A če se človek – krhko, nespametno in nehvaležno bitje – izkaže za nesposobnega, da bi vzdržal vlogo partnerja v sožitju z Bogom, to ne sme privedi do sistematičnega razvrednotenja človeštva z radikalnim pesimizmom glede poteka zgodovine. Sveto pismo namreč predstavlja Boga zgodovine, ki je zvest v ljubezni in ki tam, kjer je veliko greha, milost še bolj pomnoži (Rim 5,20). Zgodovina odrešenja povečuje Boga v njegovi zastopniški, velikodušni, neizčrpni ljubezni (8,31-38); povečuje pa tudi človeka, ki ga je Bog naredil zmožnega resnično izraziti svojo naravo sina – da ljubezen ne le sprejema, ampak tudi daje in je zato začetek upanja za vsak rod, ki se bo pojavil, dokler ne bo nič drugega kot ljubezen, ker bo vse podrejeno Bogu, ki bo »vse v vsem« (1 Kor 15,28).

Dejanje, s katerim Bog človeka sprejme v svoje lastno življenje, je v Jezusu Kristusu postalo meso: človek prejema neumrljivost od Kristusa. Večnega življenja zato ne moremo razlagati iz izolirane posamične eksistence in njene lastne moči, temveč iz naravnosti, ki odkriva jedro krščanskega umevanja stvarnosti sploh: Bog je nesmrtnost kot dinamika odnosov trinitarične ljubezni. »Bog sam ni, atom, marveč odnos, ker je ljubezen; in zato je Bog življenje.« (Ratzinger 2018, 157) Iz vere v stvarjenje sledi celostni značaj krščanskega upanja. Ob tem pa Ratzinger poudarja, da se s spraševanjem o smrti in vstajenju gibljemo po področju, ki ga z zgolj človeško vednostjo ni mogoče enoumno izmeriti, a ga je s temeljnimi antropološkimi izkustvi in uvidi vendarle mogoče podpreti. Prihodnje življenje lahko zato na podlagi sedanjega dojamemo nekako v slutnjah (zaradi dejstva, da je); oznanilo vere se lahko opre na razvidnost razmišljanja. Nasproti temu pa kajstvo tega novega življenja (kaj je) ostaja povsem zunaj našega izkustvenega prostora (160).

Nova zaveza pogled nujno usmerja na 1 Kor 15,35-53. Pavel se srečuje s stališčem, ki skuša misel o vstajenju privedi do absurdnosti s tem, da tej misli zastavi vprašanje: »Kako vstajajo mrtvi?« (v. 35) Nasproti temu stališču Pavel vprašanje o telesnosti vstajenja obravnava tako, da izkustvo nove telesnosti vstalega Gospoda prenese na razumevanje vstajenja mrtvih sploh – Pavel se torej odločno upre vladajočemu judovskemu pojmovanju, ki vstalo telo razumeva v polni identiteti z zemeljskim telesom in tako vstajenjski svet pojmuje kot preprosto nadaljevanje

zemeljskega sveta. Srečanje z Vstajnim, ki se je kot čisto drug(ačen) zemeljskemu gledanju in spoznanju odtegnil, ni nič več podrejeno zakonom materije, temveč je postalo vidno na način teofanije – to srečanje je prejšnje predstave nepreklicno razbilo. Pavel pravi: »Meso in kri ne moreta podedovati Božjega kraljestva in tudi minljivost ne podeduje neminljivosti.« (v. 50) Vsakršno naturalistično in fizicistično gledanje na vstajenje je tukaj poudarjeno odklonjeno. Brezpogojnost, s katero odklanja naturalistične zamisli, ga vendarle ne ovira pri tem, da bi ne še naprej govoril o vstajenju telesa – kar pa je nekaj drugega kot oživitev trupel po načinu tega sveta. Odklanjanje naturalizma za Pavla ne pomeni opustitve vstajenja, temveč ravno njegovo vidnost. Po njegovem prepričanju telo ne obstaja na adamitski način ‚duševnega telesa‘, ampak tudi na način Jezusa Kristusa – kot telesnost v moči Svetega Duha. Nasproti fizicističnemu realizmu ni postavljen spiritizem, pač pa pnevmatični realizem (Ratzinger 2018, 167). Bog je postal človek, da bi človek postal Bog – močne besede, ki jih po sv. Ireneju najdemo pri mnogih cerkvenih očetih, povzemajo samo bistvo krščanstva: Božji sestop je tisti, ki človeku odpira pot do nesmrtnosti in tudi do pobožanstvenja (Cole-Turner 2011, 5).

4. Biblična antropologija in kerigmatična kateheza

Če želi Cerkev sodobnega človeka nagovarjati v njemu razumljivem jeziku, mora nujno vedno znova preverjati svojo vizijo človeka (Currò 2024). Razumevanje resničnega položaja človeka, ki potrebuje odrešenje, kot ga orisuje biblična antropologija, neizogibno vodi v kerigmatično katehezo. Podobno kot transhumanisti človekovo umrljivost doživljajo kot največjo tragedijo njegovega obstoja, tudi biblična antropologija nedvoumno osvetljuje dramatičnost človekovega položaja – človek živi v nenehnem strahu pred smrtjo (Heb 2,15). Izvor vsega človekovega trpljenja in nezadovoljstva je strah pred smrtjo. Ne gre le za strah pred fizično smrtjo, temveč za vse, kar človeku preprečuje, da bi živel polno, pristno in stabilno življenje. Kako se je človek v tem stanju znašel? Na to vprašanje odgovarjajo prva poglavja Svetega pisma, ki pojasnjujejo stvarnost vsakega človeka. »Pripoved o padcu (1 Mz 3) uporablja slikovito govorico, a izreka trditev glede prvotnega dogodka, glede dejstva, ki se je dogodilo *na začetku človekove zgodovine*. Razodetje nam daje gotovost vere, da je vsa človeška zgodovina zaznamovana s tem prvotnim padcem, ki so ga svobodno zagrešili naši prvi starši.« (KKC 390) To človekovo dramatično stanje potrjuje in podčrta Nova zaveza (Rim 5,12-21). V prvotni padec smo vpleteni vsi ljudje, zato je smrt prešla na vse ljudi. Človek je zaradi greha obdan s strahom pred smrtjo, zato ne more priti k bližnjemu, ne more ljubiti – to v človeku povzroča globoko notranjo razklanost, saj je ustvarjen za odnos, za ljubezen po Božji podobi. Krščanska antropologija je zato vedno zagovarjala stališče, da neizmerne bede, ki stiska ljudi, njihovega nagnjenja k zlu in smrti ni mogoče razumeti brez njihove povezave z Adamovim grehom – pa čeprav gre pri tem za skrivnost, ki je človek ne more povsem razumeti (KKC 402–409). Prav dramatično stanje človeka, ki ga opisujejo prva poglavja Svetega pisma, je vstopno mesto kerigmatične kate-

heze: Smrt, ki je kraljevala od Adama (Rim 5,14), je premagana! Kakor v Adamu vsi ljudje umirajo, tako bodo v Kristusu vsi oživiljeni (1 Kor 15, 20-22).

Kerigmatična kateheza sega v samo bistvo krščanske vere: obsega bistvo krščanskega sporočila, posreduje odrešenijsko Božjo ljubezen v Kristusu, da bi človek dosegel polnost življenja (PK 2). V katehezi je zato njena vsebina vedno znova potrebna prečiščenja. Tudi novi „Pravilnik za katehezo“ (PK) večkrat poudarja, da ima

»v katehezi bistveno vlogo prvo oznanilo ali *kerigma* /.../ V ustih katehistov vedno znova zazveni prvo oznanilo: ›Jezus Kristus te ljubi in je daroval svoje življenje, da bi te odrešil. Zdaj vsak dan živi ob tebi, da bi te razsvetlil, krepil in osvobodil.« To oznanilo imenujemo ‚prvo‘, pa ne zato, ker je na začetku in ga potem pozabimo ali nadomestimo z drugimi vsebinami, ki ga presegajo. ‚Prvo‘ je v kvalitetnem smislu, ker je poglobljeno oznanilo, ki ga moramo vedno znova na različne načine slišati na novo in ga v teku kateheze na vseh njenih stopnjah in v vseh trenutkih vedno znova oznanjati v eni ali drugi obliki.« (EG 164)

Ker ima »kerigma neizbežno družbeno vsebino«, se učinkovitost kateheze ne kaže samo v vidnem oznanilu pashalne skrivnosti, temveč tudi v »novem pogledu na življenje, na človeka, pravičnost, družbeno življenje, celotno veselje« (PK 60). Kateheza je oznanilo vere, ki se dotika vseh razsežnosti človeškega življenja, zato na različne načine in v različni meri odzvanja v konkretnem življenju vsakega posameznika (Stegu 2022). Kerigma ni nauk ali vsebina, ki bi jo ponavljali (Rahner in Lehmann 1969, 15) – je izkušnja, odmev vstajenja v življenju krščanske skupnosti in posameznih oseb, ki so zapustile svoje dotedanje razmišljanje in dopustile, da se v njih rodi nova oblika življenja. To je vcepljeno v izkušnjo Kristusovega vstajenja in zmage nad smrtjo – in to oznanjajo: ne le Kristusovo vstajenje, ampak hkrati z njim tudi lastno izkušnjo vstajenja (Tukara 2020, 585–586). Središče oznanjevanja sta človek in njegova izkušnja (PK 5). Spreobrnjeni človek postane priča (Tukara 2020, 597–588). Apostol Pavel to pove takole: »Izročil sem vam predvsem to, kar sem sam prejel /.../.« (1 Kor 15,3) A njegovo oznanilo ni le pripoved o tem, kaj se je zgodilo s Kristusom (3-7), ampak hkrati pričevanje o tem, kaj se je zgodilo njemu: »Nazadnje za vsemi pa se je kot negodniku prikazal tudi meni. Jaz sem namreč najmanjši izmed apostolov in nisem vreden, da bi se imenoval apostol, ker sem preganjal Božjo Cerkev. Po Božji milosti pa sem to, kar sem, in njegova milost, ki mi je bila dana, ni postala prazna. Nasprotno, bolj kakor oni vsi sem se trudil, pa ne jaz, ampak Božja milost, ki je z menoj.« (8–10)

Pavel govori, a tudi pokaže, kako konkretno je videti novo življenje s Kristusom (Stegu 2022). Kristusova zmaga nad smrtjo odrešenemu človeku omogoča vrnitev v prvotno harmonijo, ki je bila porušena z grehom. Človek lahko ponovno vstopa v odnose, za kakršne je bil ustvarjen. Veselo oznanilo krščanstva je med drugim tudi v tem, da telo človeku omogoča živeti celostne odnose in uresničiti njegovo poklicanost (Platovnjak in Svetelj 2023). Pokaže se lepota odnosov v krščanskem občestvu, v krščanskem zakonu, harmonija s stvarstvom: človek ni ustvarjen kot

predmet, temveč kot oseba, ki je ljubljena in sposobna ljubiti in je prav v tem *imago Dei* – ustvarjena po Božji podobi.

5. Sklep

Po mnenju transhumanistov naj bi se s pomočjo bliskovitega razvoja tehnologij in napredka znanosti človek že kmalu sam dokopal do dokončnih odgovorov na svoja temeljna vprašanja in omejenosti: postal naj bi nesmrten, brez omejitev in vsega, kar ga danes še omejuje in uničuje. Krščanska teologija pa transhumanističnih eshatoloških pogledov in odrešitvenih strategij ne more sprejeti iz podobnih razlogov, kot je zavračala manihejsko in pelagijansko krivoverstvo. Transhumanizem v nekaterih svojih vidikih spominja na manihejsko zaničevanje pokvarjenega in zlobnega snovnega telesa, iz katerega je treba rešiti dušo – le da za transhumaniste rešitev ni v tem, da se prepustijo smrti, temveč bi želeli zanikati končno omejenost in smrtnost telesa. Hkrati prepoznamo pelagijansko prepričanje, da se človek do popolnosti lahko dokoplje sam: le z voljo bo dosegel, da postane bitje, kakršno hoče biti. Pelagijansko iskanje popolnosti torej nepopolnosti in nepopolnega človeka prav tako ne prenaša.

Krščanska teologija je dobrosto telesa vedno branila pred različnimi herezijami – in poudarjala dragocenost utelešenja. Prav z utelešenjem svojega Sina Bog stvarstvo odrešuje propada in premaguje smrt. Novo življenje, ki ga prinaša Kristus, ni preprosto podaljševanje starega, pač pa nova vrsta življenja. S Kristusovim rojstvom v telesu, njegovo smrtjo in njegovim telesnim vstajenjem se razodene tudi dragocenost človekovega telesa. V Kristusu lahko človek dokončno in povsem jasno prepozna, da je zares ustvarjen po Božji podobi (1 Mz 1,26-27). Šele odrešenje osvetljuje človekov začetek.

Transhumanistično pojmovanje človeka, ki obljublja odrešenje od trpljenja in smrti ter večno življenje kot posledico razvoja umetne inteligence, kaže, kako velik izziv (kljub bliskovitemu napredku nekaterih področij) so za človeka še vedno njegova temeljna eksistencialna vprašanja. Kerigmatična kateheza, ki izhaja iz biblične antropologije in se zaveda dramatičnega stanja ter eksistencialnih vprašanj sodobnega človeka, mu odpira pot k sprejemanju odrešenja v Kristusu.

Kratice

EG – Frančišek 2014 [Evangelii Gaudium].

GS – *Koncilski odloki* 1980 [Gaudium et spes].

KKC – Katekizem Katoliške cerkve 2008.

LS – Frančišek 2015 [Laudato si'].

PK – Papeški svet za pospeševanje nove evangelizacije 2021 [Pravilnik za katehezo].

Reference

- Bostrom, Nick.** 2007. In Defense of Posthuman Dignity. *Bioethics* 19:202–214.
- Cole-Turner, Ronald.** 2011. The Transhumanist Challenge. V: Cole-Turner Roland, ur. *Transhumanism and Transcendence: Christian Hope in an Age of Technological Enhancement*, 1–19. Washington: Georgetown University Press.
- Curro, Salvatore.** 2024. Tracce sensibili di Rivelazione: nuova sfida antropologica nella pastorale. V: Biancardi Giuseppe in Curro Salvatore, ur. *Fede e catechesi in Europa: possibili sentieri*, 53–81. Torino: Elledici.
- Dorobantu, Marius, Brian Patrik Green, Anselm Ramelow in Eric Salobir.** 2022. Being Human in the Age of AI. Research gate. https://www.researchgate.net/publication/365945288_Being_Human_in_the_Age_of_AI?channel=doi&linkId=6389b3b82c563722f22e84df&showFulltext=true (pridobljeno 12. 11. 2024).
- Dorobantu, Marius.** 2021. Human-Level, but Non-Humanlike: Artificial Intelligence and a
- Multi-Level Relational Interpretation of the Imago Dei.** *Philosophy, Theology and the Sciences* 8, št. 1:81–107. <https://doi.org/10.1628/ptsc-2021-0006>.
- Frančišek.** 2014. Evangelii Gaudium. Apostolska spodbuda. Ljubljana: Družina.
- — —. 2015. Laudato si'. Okrožnica. Ljubljana: Družina.
- — —. 2023. Umetna inteligenca in mir. Poslanica svetega očeta Frančiška za 57. svetovni dan miru. Ljubljana: Družina.
- Koncilski odloki.** 1980. Ljubljana: Nadškofijski ordinariat v Ljubljani.
- Kurzweil, Ray.** 2005. *The singularity is near: when humans transcend biology*. New York: Viking.
- Malović, Nenad.** 2024. Transhumanizam i Božji popravni ispit iz stvaranja. *Bogoslovna smotra* 94, št. 1:9–11.
- O'Connell, Mark.** 2018. *Živeti kot stroj: Pustolovščine med kiborgi, utopisti, hekerji in futuristi, ki rešujejo preprosto vprašanje smrti*. Tržič: Učila.
- Papeški svet za pospeševanje nove evangelizacije.** 2021. Prilnik za katehezo. Ljubljana: Družina
- 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>
- Pontificia Commissione Biblica.** 2021. «Che cosa è l'uomo?» (*Sal 8,5*): Un itinerario di antropologia biblica. Vatikan: Libreria Editrice Vaticana.
- Rahner, Karl, in Karl Lehmann.** 1969. *Kerygma and Dogma*. New York: Herder and Herder.
- Ratzinger, Joseph.** 2018. *Eshatologija: smrt in večno življenje*. Ljubljana: Ognjišče.
- Singbo, Odilon-Gbènoukpo.** 2024. Krščanska antropologija i umjetna inteligencija u doba transhumanizma. *Bogoslovna smotra* 94, št. 1:53–75. <https://doi.org/10.53745/bs.94.1.6>
- Stegu, Tadej.** 2019. Transhumanizem in krščanska antropologija. *Bogoslovni vestnik* 79, št. 3:683–692. <https://doi.org/10.34291/bv2019/03/stegu>
- — —. 2022. Kerigmatska kateheza in resonanca. *Bogoslovni vestnik* 82, št. 3:705–714. <https://doi.org/10.34291/bv2022/03/stegu>
- 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>
- Tukara, Drago.** 2020. Važnost prve kerigme za Crkvo i društvo danas. *Bogoslovna smotra* 90, št. 3:583–603.
- Vodičar, Janez.** 2023. Avtoriteta na področju vzgoje in verovanja v digitalni dobi. *Bogoslovni vestnik* 83, št. 4:1035–1047. <https://doi.org/10.34291/bv2023/04/vodicar>
- Waters, Brent.** 2011. Whose Salvation? Which Eschatology? V: Roland Cole-Turner, ur. *Transhumanism and Transcendence: Christian Hope in an Age of Technological Enhancement*, 163–175. Washington: Georgetown University Press.
- Žalec, Bojan.** 2023. Ali je umetna inteligenca inteligenca v pravem pomenu besede? Vprašanje psihičnih značilnosti in splošnosti. *Bogoslovni vestnik* 83, št. 4:813–823. <https://doi.org/10.34291/bv2023/04/zalec>

*Urška Jeglič***Virtualna resničnost kot pripomoček pri poučevanju religijskih vsebin v šolah***Virtual Reality as a Tool for Teaching Religion in Schools*

Povzetek: V prispevku so predstavljeni rezultati raziskave, izvedene v okviru Erasmus+ projekta „VR4Empathy“. Namen raziskave je bil ugotoviti, kako dobro dijaki in učitelji srednjih šol poznajo virtualno resničnost in kje menijo, da bi jo lahko uporabljali pri pouku. Spoznanje, da religijske in etične vsebine v raziskavi niso omenjene, nas je nagovorilo k pisanju prispevka, v katerem poskušamo osvetliti stanje pouka religij v slovenskem okolju in predstavljamo virtualno resničnost kot pripomoček učiteljem in učencem pri spoznavanju tovrstnih vsebin.

Ključne besede: virtualna resničnost v šoli, izobraževanje, posodobitev učnega sistema, pouk religij, virtualna resničnost in religijski pouk

Abstract: This article presents the results of a survey conducted as part of the Erasmus+ project VR4Empathy. The aim of this survey was to find out how well secondary school students and teachers know virtual reality and where they see its use in the classroom. The realisation that religious and ethical content was not mentioned in the survey prompted us to write this article in which we try to shed light on the state of religious education in Slovenia and present virtual reality as a tool for teachers and students to learn this type of content.

Keywords: virtual reality in schools, education, modernisation of the learning system, religious education, virtual reality and religious education

1. Uvod

K pisanju prispevka¹ so nas spodbudili rezultati raziskave med dijaki in učitelji o poznavanju virtualne resničnosti in predlogih za njeno uporabo znotraj šolskega

¹ Prispevek je nastal v okviru raziskovalne skupine „P6-0269 Religija, etika, edukacija in izzivi sodobne družbe“, ki ga sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS), in z namenom diseminacije rezultatov projekta „VR4Empathy: Using virtual reality for inclusive and action-oriented empathy in schools (Project No. 101132844)“.

sistema. Neomenjenost religijskih vsebin (ob pomenu teh vsebin za dijake in njihove učitelje) vzbuja zaskrbljenost in jasno opozarja na potrebnost diskurza na temo ponovnega premisleka o vključitvi tovrstnih vsebin v slovenski šolski sistem. Smiselno bi bilo sprejeti strategijo o bolj sistematičnem in nadzorovanem vključevanju religijskih vsebin v obstoječe predmete znotraj šolskega programa ali – še bolje – končno razmisliti o uvedbi nekonfesionalnega pouka o religijah. Formalno so se zadeve ob osamosvojitvi Republike Slovenije sicer uredile. Slovenski religiolog Aleš Črnič tako omenja, da so ga ministrstva, pristojna za šolstvo, dvakrat povabila k pripravi strokovnega izhodišča »za reševanje konkretnih dilem, ki se v vsakdanji praksi pojavljajo ob srečevanju teh dveh velikih sistemov, religije in šole« (2021, 11). Pripravljena gradiva so po njegovih besedah obležala v uredniških predalih. Posledično se ne nadejamo, da bo pričujoči prispevek znotraj šolskega sistema v Republiki Sloveniji spodbudil kakršno koli revolucionarno spremembo; pri tem se tudi zavedamo, da niso potrebni le kozmetični popravki, ampak smiselna in strokovna prenova sistema. Opozoriti moramo na dejstvo, da tudi znotraj slovenskega šolskega sistema predmeti, ki obravnavajo religijske in etnične vsebine, obstajajo.² Predmet „Verstva in etika“ je izbirni predmet v zadnjih treh letih osnovnošolskega izobraževanja.

»Splošni cilji predmeta se delijo v štiri sklope, in sicer v (1) pridobivanje objektivnega znanja s področja verstev in etike, (2) spodbujanje osebno-stne rasti vsakega posameznika in sposobnosti za razumevanje drugih, (3) spoznavanje religiozne govornice in pomena religij za razvoj kultur in civilizacij ter oblikovanje kritičnega odnosa do religijskih tradicij in (4) pripravo na kritičen in konstruktiven vstop učencev v pluralno družbo.« (Predmetna kurikulumna komisija za verstva in etiko 2001, 9)

V sedmem razredu osnovne šole se manjši del religijskih vsebin obravnava pri obveznem predmetu Državljanstva in domovinska vzgoja ter etika – del kurikulumata je namreč tudi pravica do veroizpovedi, spoznavanje položaja verstev v Sloveniji in pregled verskih skupnosti, spoznavanje podobnosti in razlik med velikimi svetovnimi verstvi ter spoznavanje razlik med religioznimi in nereligioznimi osmišljanji življenja. Prav tako bi naj učenci pri tem predmetu razvili sposobnost za razumevanje splošnih etičnih načel svetovnih verstev in se seznanili z odnosi med različnimi svetovnimi verstvi (Štrukelj 2011, 11).

² »V slovenskem prostoru glede religijskega pouka v šoli obstaja veliko zmede in predsodkov. Precej problemov povzroča že sama terminologija, saj se uporabljajo pojmi verouk, verski pouk, religijski pouk, pouk o religijah in še bi lahko naštevali. Zato je najprej treba narediti jasno ločnico med veroukom oz. uvajanjem otrok v posamezno versko skupnost in religijskim poukom kot šolskim predmetom, ki obravnava pojav religij in njihov pomen za osebno in skupno življenje. Za šolski predmet se je uveljavilo tudi poimenovanje verski pouk, ki pa je po našem mnenju preveč podoben izrazu verouk, zato smo za poimenovanje šolskega predmeta dali prednost izrazu religijski pouk. Še ustreznije poimenovanje bi bilo religijska kultura, saj bi se s tem poudarilo, da ne gre zgolj za pouk o religijah, ampak se s podajanjem religioznih vsebin ustvarja tudi določen način življenja. Nikakor religijski pouk kot podpora vzgojnemu prizadevanjem šole – s tem ne mislimo na širjenje določene religije znotraj javnega šolskega prostora, ampak da se npr. zaradi spoštljivega in dialoškega pristopa k vprašanjem religije pri učencih razvijajo tudi življenjske drže spoštovanja, strpnosti in dialoškosti. V tem smislu ima religijski pouk oz. religijska kultura pomembno mesto pri vzgojnem poslanstvu šole.« (Globokar 2021, 124–125)

Več religijskih vsebin so deležni učenci zasebnih šol v Sloveniji. Vse štiri slovenske katoliške gimnazije imajo prav vseh štirih letnikih predmet Vera in kultura, kjer je tretji letnik namenjen spoznavanju različnih svetovnih religij in njihovi primerjavi. Predmet „Religija“ se vsa štiri leta poučuje tudi na Waldorfski šoli in v šoli Montessori, kjer imajo predmet „Vera in kultura“.

Ne glede na omenjeno situacijo se nam zdi na probleme znotraj slovenskega šolskega sistema pomembno še naprej opozarjati ter ponujati realne in izvedljive rešitve v želji in upanju, da bi se religijska pismenost izboljšala. Verjamemo, da bi tako posamezniki postali bolj empatični – in državljani države, k ji pripadamo, pa tudi celotne globalne vasi, ki je z digitalizacijo postala naš novi svet.

V prispevku bomo naprej spregovorili o tem, kaj je virtualna resničnost in kako je vključena v izobraževalni sistem. V tretjem delu sledi predstavitev rezultatov projekta, po tem pa se bomo osredotočili na predstavitev predlogov za vključitev religijskih vsebin v učni sistem s pomočjo virtualne resničnosti. Z zavedanjem, da ima tudi virtualna resničnost svoje pomanjkljivosti, v zaključnem delu opozarjamo na njene morebitne nevarnosti za mlade.

2. Virtualna resničnost in izobraževanje

Navidezna resničnost ali navidezna realnost je oblika računalniške stimulacije, pri kateri ima udeleženec občutek, da se nahaja v umetnem okolju. Človek s pomočjo različnih pripomočkov (očala, podatkovne rokavice, čelade, obleke itn.) s svojo zavestjo vstopi v virtualni svet in čas, medtem ko njegovo telo ostane v realnem svetu in času. Glavna značilnost virtualne resničnosti je, da uporabnik s pomočjo izbrane virtualne opreme vstopi v virtualno okolje – z občutkom, da se tam resnično nahaja, obenem pa je lahko v njem tudi aktiven. Izbrana oprema uporabniku predaja zvočne in vidne informacije o svetu, ki ga kot udeleženec raziskuje, njegove premike v virtualnem svetu pa tolmači vrsta različnih senzorjev in tehnologij.

Prizori, ki so vgrajeni v virtualno opremo, so lahko ustvarjeni s pomočjo računalnika ali pa so sestavljeni iz fotografij in video vsebin, ki prikazujejo resnične kraje. S pomočjo virtualne resničnosti lahko človek raziskuje skorajda vse. Navidezna resničnost omogoča vstop v objekte in hojo po njih.

Zametki virtualne resničnosti segajo v petdeseta leta 20. stoletja, a se je njen preboj zgodil šele v devetdesetih letih – z razvojem in dostopnostjo računalniške tehnologije. Virtualna resničnost je na začetku služila kot pripomoček za zabavo, s katero so se mladi srečevali predvsem preko filmske industrije in igranja video iger. Danes jo uporabljamo v različnih vedah – umetnosti, medicini, psihoterapiji in celo v vojski. Digitalizacija učnega sistema je postala del strategij programov Evropske unije že pred desetletji, a je bila do epidemije leta 2020 realizirana delno. Hiter preobrat se je zgodil z uvedbo karantene in učenja od doma, ki sta zamajala udobnost sledenja tradicionalnim oblikam poučevanja in odprla vrata novim digitalnim pripomočkom.

Med produkte virtualne resničnosti prištevamo tudi t. i. ‚virtualna očala‘. Njihov preboj na trg se je zgodil v letu 2012, ko je Palmer Luckey Oculus Rift predstavil na Kickstarterju. Prva komercialna različica Oculus Rift je bila izdana leta 2016. Le še vprašanje časa je bilo, kdaj se bo začelo razpravljati o uporabi virtualne resničnosti in virtualnih očal tudi v šolskem sistemu. Strokovnjaki že dalj časa opozarjajo na dejstvo, da je treba šolske sisteme (še iz industrijske dobe) posodobiti in prevetriti. »Konvencionalne izobraževalne ustanove se spreminjajo veliko počasneje kot načini domiselnega in sodelovalnega, participativnega učenja, ki ga ponujata internet in cela vrsta sodobnih mobilnih tehnologij,« menita Davidson in Goldberg (2015, 60). »Učni sistemi, ki mladih ne opremijo za življenje v sodobnem svetu (to zahteva iznajdljivost, postavljanje prioritet, kreativnost itn.), so potrebni sprememb.« (Robinson in Aronica 2015, 26; 31–36) Prav tako bi bilo treba spremeniti način podajanja snovi. Generacije, ki so digitalnih naprav vajene od mladosti, delujejo drugače. V primerjavi s starejšimi generacijami potrebujejo mnogo več impulzov, da jih stvari pritegnejo in zanimajo; čas pozornosti in besedni zaklad pa sta se zmanjšala. Ob tem učenje mnogim sodobnim Evropejcem ne pomeni privilegija – šola za mlade pogosto predstavlja pogloblitni stresni dejavnik in je pogosto razumljena kot nuja.

3. Rezultati raziskave evropskega projekta VR4Empathy

Zdi se, da virtualna resničnost ponuja svoj delež pri iskanju novih edukativnih pripomočkov, ki bi pouk naredili bolj zanimiv in privlačen, snov pa razumljivejšo in za dijake na neki način tudi bolj smiselno. S tem namenom je bil prijavljen evropski projekt »Uporaba virtualne resničnosti za vključujočo in k dejanjem usmerjeno empatijo v šolah«, pri katerem sodelujeta tudi Teološka fakulteta Univerze v Ljubljani in Ministrstvo za vzgojo in izobraževanje Republike Slovenije.³ Projekt je namenjen oblikovanju vsebin za virtualna očala, ki jih bodo učitelji v srednjih šolah lahko uporabljali pri svojem pouku. Cilj projekta je oblikovati vsaj 15 video vsebin za virtualna očala za srednje šole, ki bodo pokrivala teme z naslednjih treh področij: zgodovinska dejstva Evrope, skupne evropske družbene vrednote in evropsko državljanstvo ter raznolikost. Projekt je razdeljen na sedem delovnih enot. V njegovem drugem delu je bila med marcem in aprilom 2024 izvedena kvantitativna raziskava (vprašalnik), maja 2024 pa še kvalitativna raziskava (kreativne delavnice). Cilj raziskave je bil preveriti, v kolikšni meri učitelji in učenci digitalne pripomočke uporabljajo pri pouku, kako dobro po lastnem mnenju virtualno resničnost in z njo povezane edukativne pripomočke poznajo, ali virtualno resničnost uporabljajo – in kakšno je njihovo mnenje o njeni uporabi pri pouku. Eden od ciljev projekta je tudi preveriti, ali uporaba virtualne resničnosti pri pouku spodbuja razvoj empatije pri mladih. S tem namenom smo preverili še stopnjo kognitivne in emocionalne empatije ter demografske podatke. Obe raziskavi sta bili izvedeni v treh državah: Sloveniji, Grčiji in Portugalskem.

³ Preostali partnerji projekta so: INOVA+, University of Peloponnese, Ministry of Education, Religious Affairs and Sports (Greece), Institute of Educational Policies in Khorá ApS.

V kvantitativnem delu raziskave je sodelovalo 251 učencev v starosti med 14 in 19 let ter 135 učiteljev iz Slovenije, Grčije in Portugalske. Polovica učiteljev je imela vsaj 15 let delovnih izkušenj. Z uporabo virtualne resničnosti v šolah se strinja 47 % vprašanih učencev. 18 % se jih ne strinja, 24 % pa je neopredeljenih. 11 % jih meni, da slednje za njih ni veljavno. Zanimivo je, da uporabo virtualne resničnosti v šolah v večji meri podpirajo učitelji, saj se s tem strinja 68 % respondentov.

Zardi internega dogovora med projektnimi partnerji bomo v nadaljevanju povzemali le rezultate slovenskih dijakov in njihovih učiteljev. V raziskavo je bilo vključenih 48 dijakov iz slovenskih šol s povprečno starostjo 17 let. V raziskavi je sodelovalo 26 deklet, 19 fantov in 2 spolno neopredeljeni osebi. 60 % respondentov živi na vasi, 36 % v mestu, drugi so izbrali 'drugje'. Večina slovenskih dijakov je vešča uporabe digitalne tehnologije, ki jo v 72 % uporablja za zabavo. Dnevno igra videoigre 28 % slovenskih dijakov, 34 % pa nikoli. V kvantitativnem delu raziskave je sodelovalo tudi 50 učiteljev iz slovenskih šol. Povprečna starost učiteljev je 50 let. Med učitelji je 71 % žensk in 29 % moških. Večina učiteljev ima vsaj 15 let delovnih izkušenj. Poučujejo različne predmete; nobeden ni poučeval katerega koli predmeta z religijsko vsebino. Največ slovenskih učiteljev je sicer poučevalo jezikovne predmete, geografijo in strokovne predmete. Kar 26 % slovenskih učiteljev meni, da uporaba virtualne resničnosti (VR) za njihove predmete ni primerna. Večina (59 %) se ne more opredeliti, ali ima uporaba VR več prednosti kot slabosti. Da ima uporaba VR več prednosti, meni 16 % učiteljev.

Z uporabo VR v šolah se strinja 28 % vprašanih slovenskih učencev, 23 % se jih s tem ne strinja, 40 % pa je neopredeljenih. Skoraj desetina dijakov meni, da to za njih ni veljavno. Slovenski dijaki menijo, da je VR očala primerno uporabljati pri matematiki, fiziki, zgodovini, geografiji, tehničnih predmetih, kemiji, biologiji in za učenje jezikov. Redki menijo, da je uporaba VR možna tudi pri športni vzgoji. Opazimo lahko tudi, nekateri menijo, da je uporabnost odvisna od gradiva, ki bi ga s pomočjo VR usvajali.

Učitelji in dijaki, ki so do uporabe VR v šolah zadržani, menijo, da ta ljudi poneumlja, oddaljuje od realnega sveta in onemogoča stik s sočlovekom, po drugi strani pa lahko kdo dejanja, ki jih počne v virtualnem svetu, prenese tudi v svoje realno življenje in s tem škodi drugemu. Hkrati se s tem zmanjša kritično razmišljanje. Učitelje skrbi, da bi uporaba VR učence obremenila in negativno vplivala na njihovo čustveno stanje. Skrbi jih tudi dodatno individualno delo v času, ko vzpostavljanje živih stikov dijakom že tako ali tako povzroča težave. Omembe vredno je tudi mnenje dijakov na kreativnih delavnicah v Sloveniji, kjer so izpostavili, da si želijo več analognega učenja, druženja, učenja v naravi. Dijake in učitelje skrbijo tudi negativni fiziološki učinki uporabe VR, ker da ta lahko povzroča slabost, tesnobo, glavobole in po mnenju nekaterih tudi okvare sluha ter vida. Skrbi jih tudi financiranje tovrstne opreme. Učitelji imajo pomisleke glede dodatnega dela in potreb po spremembi učnega načrta v primeru uvedbe VR v učni sistem.

V sklopu raziskave so bile pripravljene že omenjene kreativne delavnice, ki smo jih izvedli na treh različnih šolah v Grčiji, Sloveniji in na Portugalskem. Pri delavni-

cah so aktivno sodelovali izbrani učitelji in dijaki. Cilj delavnic je bil pridobiti mnenje in ideje o tem, katera učna snov (vsebine) bi se učiteljem in dijakom za stimulatивно podajanje snovi z uporabo VR v šolah zdela primerna. V odgovorih prevladujejo vsebine s področja naravoslovja (zlasti fizika in biologija) in družboslovja (zgodovina in geografija).

4. Šolsko poučevanje religije s pomočjo virtualne resničnosti

S pomočjo virtualne resničnosti bi bilo smiselno poučevati tudi o religijskih in etičnih vsebinah. Naloga šol je, da skrbijo za religijsko pismenost; sem tako Moore (2007) kot Dinham in Francis (2015) prištevajo osnovno znanje o religijah, izraznih oblikah religij, pa tudi prepoznavanje in kritično vrednotenje kulturnih, družbenih in političnih razsežnosti religij.⁴

V Evropi poenotene pouka religije ne poznamo. Države tako izbirajo med različnimi modeli poučevanja religijskih in etičnih vsebin (k čemur nagovarja in si za to prizadeva tudi EU); Slovenija ne spada k zgledom dobre prakse. Pouk religijskih vsebin znotraj slovenskega učnega načrta večinoma poteka ob poučevanju drugih predmetov, pri čemer tovrstne vsebine v večini primerov poučuje kader, ki za to ni specializiran.⁵ Posledično je pridobljeno znanje o religijah slabo, Toledske smernice za pouk o religijah in prepričanjih v javnih šolah pa neuresničene. Avtorji tega dokumenta menijo, da znanja o religijah in raznovrstnih verovanjih vplivajo na razumevanje družbene raznolikosti in religijske svobode, obenem pa spodbujajo demokratično državljanstvo (OVSE 2007, 76). Bolj kot namreč drugega razumemo, lažje ga lahko sprejmemo takšnega, kakršen je – ne glede na njegovo kulturno in etnično pripadnost ter versko prepričanje. Ker je tovrstno izobraževanje v slovenskem prostoru zelo okrnjeno, je v našem okolju še toliko bolj prisoten strah pred drugimi, kar je posebej dobro vidno v obdobju migracijskih valov (in se kaže preko različnih fobij).

Slovenski religiologi si tovrstno stanje prizadevajo spremeniti že desetletja, a zaenkrat nič ne kaže, da bi bil pouk o religijah, kot ga poznajo v nekaterih drugih evropskih državah, lahko vzpostavljen tudi pri nas. Ker izobraževalni sistem religioških tematik ne pokriva, uporabniki informacije iščejo na svetovnem spletu. Osredkar (2023, 929) pri tem opozarja na problem zavajajočih in lažnih informacij na svetovnem spletu.

Tu lahko slovenskim pedagogom, ki strokovne izobrazbe s področja religiologije nimajo, na pomoč priskoči virtualna resničnost. Poudarek je na besedi ‚pomoč‘

⁴ Za tovrstno poučevanje religije v šolah je potreben princip laičnosti. Črnič in Pogačnik (2021, 19) menita, da »sodobni koncept laičnosti religije ne zavrača niti ni do nje sovražen, ampak vzpostavlja neodvisnost družbenih podsistemov (med njimi šole) od neposrednih religijskih vplivov ter s tem zagotavlja njihovo nevtralnost«.

⁵ Podobno se dogaja tudi z različnimi etičnimi vsebinami. Zdi se, da je moralnost v družbi na neki način prisotna, a ne udejavena (Platovnjak in Mutanen 2023, 26).

– kvalitetnega pedagoga ne more zares nadomestiti nobena tehnologija.⁶ Gre le za pripomoček, ki lahko bodisi prikaže oddaljene običaje (in jih na takšen način učencem bolj približa) bodisi poskrbi, da se pri pouku začnejo odpirati pomembne družbene teme. Čeprav smo zgoraj omenili, da lahko VR pedagogu pri podajanju učne snovi pomaga, pa ne moremo prezreti dejstva, da odpira tudi nove izzive. Navidezni pripomoček pedagoga hote ali nehotе prisili v poglobitev predvajanih vsebin (v našem primeru religijskih) in od njega zahteva, da se dodatno izobražuje – če želi učencem na vprašanja, ki jih bodo ob predvajanih vsebinah po vsej verjetnosti imeli, kvalitetno in strokovno odgovoriti.

Predlagamo, da bi s pomočjo virtualnih očal učenci pri pouku spoznavali naslednje religijske vsebine:⁷

- svete prostore (cerkve, džamije, sinagoge, templje);
- molitvene prakse (skupinske in individualne; vodene in nevodene);
- praznovanje praznikov;
- iniciacijske obrede;
- obred poroke;
- obred pokopa;
- oblačilne prakse.

Omenjene vsebine lahko učitelj predstavi tudi s pomočjo predvajanja video vsebin. Z vidika izobraževanja je tako podajanje snovi ob ustrezni razlagi snovi dovolj, a ima na posameznikovo doživljanje in razumevanje vidnega bistveno manjši vpliv. Kljub temu je temeljni namen pouka religij v šolah s tem dosežen. Istočasno zadošča tudi za pristop primerjalne religiologije.⁸ Vprašanje pa je, ali podajanje religijskih vsebin zgolj s prikazom osnovnih dejstev o različnih religijskih in filozofskih tradicijah učence spodbudi k razmišljanju o pomembnih družbenih temah, ki jih zagovarja EU in bi naj bile tudi cilj tovrstnega pouka. Na tem mestu se zdi občasna uporaba virtualne resničnosti dobra rešitev, saj lahko s pomočjo stimulacije vzbudi občutek za poistovetenje z likom ali liki, ki v video vsebini nastopajo. Opazovanje brez udeležbe tako lahko preide v opazovanje z udeležbo in občutje, da smo tam – skupaj z drugimi pripadniki določene religije –, posameznika zaznamuje z močnejšo izkušnjo. Človek sicer na takšen način svetosti samega obreda ne more začutiti (verjetno bi težje doživel tudi kakršno koli duhovno izkušnjo), a slednje pri nekonfesionalnem pouku o religijah niti ni namen, pa čeprav je sveto bistvo obstoja večine svetovnih religij in filozofskih nazorov.

Uporaba VR lahko posamezniku pomaga, da se vživi v situacijo drugega, za kar

⁶ Kraner (2023, 238) poudarja, da je kvaliteten pedagog tisti, ki poleg poznavanja vsebin zna vzpostaviti tudi resonančno okolje. Slednjega lahko vzpostavi »z dobrimi odnosi in odlično komunikacijo tako v razredu kot tudi v timu in družbenem okolju«.

⁷ Našteli smo le najbolj osnovne vsebine, gotovo pa bi lahko prikazali najrazličnejše verske prakse vseh tako obstoječih kot že izumrlih religij in kultur.

⁸ Primerjalno religiologijo razumemo kot William Paden (1994), ki pravi, da je primerjalna religiologija način proučevanja religij, ki se osredotoča na iskanje in analizo podobnosti in razlik med različnimi religioznimi tradicijami. Gre za metodo, ki vključuje več kot zgolj katalogizacijo religioznih praks in prepričanj – poskuša tudi razumeti temeljne strukture in vzorce, ki so različnim religijam skupni.

pa mora biti seveda tudi dovolj empatičen in nenazadnje sočuten. Rezultati študije (Martingano idr. 2021) kažejo, da lahko več različnih vrst izkušenj VR empatijo poveča, vendar za ta učinek obstajajo pomembni mejni pogoji. Analize podskupin so pokazale, da VR izboljša čustveno empatijo, ne pa tudi kognitivne empatije. Nadaljnje analize podskupin so pokazale, da VR v primerjavi s tehnološko manj naprednimi intervencijami (kot so branje zgodb, pripovedovanje o drugih) pri povečanju empatije ni bila učinkovitejša. Prav tako se za učinkovitejše izkazale bolj potopljive in interaktivne izkušnje virtualne resničnosti.

Vsekakor lahko prikaz različnih družbenih situacij kot izraz izzivov sodobne družbe, ki so tako ali drugače povezani z religijskimi in kulturnimi vprašanji, pripomore k boljšemu razumevanju drugega. Dober primer iz prakse je pred leti zasnovana aplikacija „EveryDay Racism“, kjer si učenec preko igre izbere eno izmed stigmatiziranih oseb ali naredi svoj lastni profil. Izbira lahko med 22-letnim Indijcem, 31-letnim aboriginom in med 29 letno pokrito muslimanko iz Omana. Igralec je vsak dan soočen z vrsto rasističnih provokacij v obliki e-pošte, besedilnih sporočil, pretiranih pogovorov, sporočil na socialnih omrežjih in tako naprej. Eden od ciljev aplikacije je povečati rasno pismenost in empatijo pri tistih, ki se z rasizmom ne srečujejo redno – da bi bili bolje pripravljeni in bolj nagnjeni k posredovanju, če bi se v prihodnosti znašli v situaciji, v kateri bi prišlo do rasizma. Igra poteka sedem dni, igralec pa se srečuje z različnimi oblikami vidnega ali prikritega nasilja: šepetanjem ljudi, neprimerno postrežbo v trgovini, celo fizičnim napadom na ulici. Po vsakem incidentu igralec izbere enega od možnih odzivov. Pri tem lahko izbira med tem, da ne stori ničesar, vloži pritožbo ali pa se odzove z jezo. Aplikacija nudi še kratko razlago posledic določene izbire. Sedem dni so igralci izpostavljeni vzorcu različnih rasističnih srečanj, ki jih običajno doživi njihov lik (čeprav se vsebina, ki je lastna vsem trem likom, precej prepleta).

Kot pomoč pri življenju v like, do katerih družba goji predsodke, bi lahko pri pouku uporabljali tudi različne video vsebine. Učenci bi si lahko pri pouku ali razredni uri s pomočjo VR ogledali kratke filme z rasističnimi incidenti. Po ogledu filmov bi učitelj sprožil in vodil diskusijo, kjer bi se poskušali vživeti v vlogo žrtev in oblikovati kritična razmišljanja. Video vsebine lahko vsebujejo tudi pričevanja žrtev – metodo, ki jo zelo spretno uporabljajo mediji (Kim in Grabe 2024, 195). S pomočjo VR bi se lahko dijaki prelevili v pokrito ali celo zakrito muslimanko, ki se sprehaja po evropskih ulicah, ali pa v katoliškega redovnika v redovni obleki, bodisi v Evropi bodisi celo v kakšni muslimanski ali komunistični državi. Predpostavljamo lahko, da bi učenci v primeru doživetja negativne izkušnje do manjšinske populacije v svojem okolju lahko razvili bolj razumevač in empatičen odnos.

5. Zaključek

V zaključnem delu prispevka poskušamo opozoriti na nekaj ključnih izzivov uporabe VR v šolah, pri pouku o religijah ali etikah. Skrb vzbujajoče je dejstvo, da niti učitelji niti dijaki, ki so bili vključeni v kvantitativni in kvalitativni del raziskave znotraj

projekta VR4Empathy, niso samoiniciativno izkazali nobenega interesa za vključitev kakršnih koli religijskih ali etičnih vsebin v video vsebine, ki bi jih uporabljali pri pouku. Slednje je sicer gotovo povezano tudi s tem, da so bile v raziskavo vključene javne srednje in poklicne šole ter gimnazije, kjer pouka o religijah (več) ni. Posledično so se respondenti bolj osredotočili na predmete, ki so jim v okviru učnega sistema že poznani. Po drugi strani so dijaki izrazili precej interesa za razvoj vsebin za virtualna očala, ki bi jih lahko uporabljali pri predmetu zgodovine in geografije. Oba predmeta imata kar nekaj potenciala za odpiranje vprašanj, ki se navezujejo na obravnavo religije – bodisi na njeno vlogo skozi zgodovino in danes bodisi v sklopu družbene geografije. Ne glede na to pa ostaja skrb, ker se izobraževanje o religijah ne zdi dovolj pomembno – in obstaja možnost, da se o tovrstnih temah med poukom pogovarjajo tako redko, da asociacij za tovrstne teme niti ne dobijo.

Če želimo podpreti spodbudo k razvoju gradiva v sklopu virtualne resničnosti, ki obravnava tudi tematike s področja religiologije, se je smiselno vprašati tudi, kakšen je naš cilj. Zdi se, da je najpomembnejši cilj informiranje o različnih religijah in kulturah in spodbujanje razprav – z namenom (ali celo željo) po preseganju diskriminacije, islamofobije, verskega sinkretizma, zamenjevanja in nerazumevanja različnih religijskih simbolov in pojavov ipd. Gre za resne in pomembne vsebine; zaradi kulturnih in zgodovinskih predpostavk je treba več pozornosti nameniti religijam, ki so določena okolja bistveno zaznamovale. Danes se zdi, da se tudi znotraj šolskega sistema spodbuja prakse iz drugih religijskih tradicij, predvsem meditacijo in jogo. Vpeljava virtualne resničnosti je v šolstvu smiselna takrat, ko vsebine posnetkov oblikujejo strokovnjaki na podlagi objektivnih in preverljivih informacij. Spodbujati je treba k boljšemu razumevanju drugega in se izogibati stereotipnim izjavam.

Razprava o vpeljavi virtualne resničnosti v šolske sisteme spodbuja še k dodatnemu razmisleku – v kolikšni meri svet še dodatno digitalizirati. Zaskrbljujoče so namreč študije, ki kažejo, koliko ur dnevno mladi preživijo pred ekrani. Najnovejša tovrstna raziskava iz Združenih držav Amerike (2023) kaže, da mladi pred zasloni v povprečju preživijo sedem ur in dvaindvajset minut – in da številke z leti naglo naraščajo (Duarte 2023). Evropske raziskave sicer zaznavajo nekoliko nižje številke (Khan idr. 2022), a vseeno zna povprečen najstnik za zaslonom preživeti tudi več kot štiri ure na dan. Dodatno uvajanje digitalizacije v šolski učni načrt bi lahko tovrstno stanje še poslabšalo. Čeprav si življenje brez tehnologije in digitalizacije predstavljamo le s težavo, pa postaja med mladimi (predvsem dekleti) vse bolj popularen t. i. ‚digitalni detox‘. Poleg tega mladi, ki so bili vključeni v raziskavo, poudarjajo, da si želijo več učenja v naravi, saj hrepenijo po živih odnosih, realnih stikih – tako s sovrstniki kot s stvarstvom.

Kratika

OVSE – Organizacija za varnost in sodelovanje v Evropi.

Reference

- Črnič, Aleš, in Anja Pogačnik. 2021. *Religija in šola. Poučevanje o religiji in njena simbolna prisotnost v javni šoli*. Ljubljana: Fakulteta za družbene vede.
- Davidson, Cathy N., in David Theo Goldberg. 2015. Šola v digitalni dobi. V: Tadej Rifel in Marko Weilguny, ur. *Onkraj klasične šolske ustanove*, 59–69. Ljubljana: Ireš.
- Dinham, Adam, in Matthew Francis. 2015. *Religious Literacy in Policy and Practice*. Bristol: Policy Press.
- Duarte, Fabio. 2023. Average Screen Time for Teens (2024). Exploding topics. Domača stran. 10. 11. <https://explodingtopics.com/blog/screen-time-for-teens> (pridobljeno 4. 7. 2024).
- Globokar, Roman. 2021. *Vzgojni izzivi šole v digitalni dobi*. Ljubljana: Teološka fakulteta.
- Khan, Asaduzzaman, Eun-Young Lee in Sharoon Horwood. 2022. Adolescent screen time: associations with school stress and school satisfaction across 38 countries. *European Journal of Pediatrics* 181:2273–2281. <https://doi.org/10.1007/s00431-022-04420-z>
- Kraner, David. 2023. Sodobni učitelj, komunikacija in resonanca v odnosih. *Bogoslovni vestnik* 83, št. 1:227–239. <https://doi.org/10.34291/bv2023/01/kraner>
- Martingano, Alison Jane, Fernanda Hererra in Sara Konrath. 2021. Virtual Reality Improves Emotional but Not Cognitive Empathy: A Meta-Analysis. *Technology, Mind, and Behaviour* 2, št. 1. <https://doi.org/10.1037/tmb0000034>
- Minchul, Kim, in Maria Elizabeth Grabe. 2024. Empathy With Muslim Victims of Discrimination: Can Personalization and Emotionalization in News Reporting Pave the Way? *Journalism & Mass Communication Quarterly* 101, št. 1:178–205. <https://doi.org/10.1177/10776990231202702>
- Moore, Diane L. 2007. *Overcoming Religious Illiteracy: A Cultural Studies Approach to the Study of Religion in Secondary Education*. New York: Palgrave Macmillan.
- Osredkar, Mari Jože. 2023. Slovenska religiologija v digitalni dobi. *Bogoslovni vestnik* 83, št. 4:923–933. <https://doi.org/10.34291/bv2023/04/osredkar>
- Organizacija za varnost in sodelovanje v Evropi. 2007. *Toledo Guiding Principles on Teaching about Religions and Beliefs in Public Schools*. Varšava: OVSE.
- Paden, William. 1994. *Religious Worlds: The Comparative Study of Religion*. Boston: Beacon Press.
- Platovnjak, Ivan, in Arto Mutanen. 2023. On Religious Knowledge. *Bogoslovni vestnik* 83, št. 1:21–32. <https://doi.org/10.34291/bv2023/01/platovnjak>
- Predmetna kurikularna komisija za verstva in etiko. 2001. *Učni načrt: Izbirni predmet. Program osnovnošolskega izobraževanja: Verstva in etika*. Ljubljana: Ministrstvo za šolstvo, znanost in šport.
- Robinson, Ken, in Lou Aronica. 2015. *Kreativne šole: Množična revolucija, ki preoblikuje izobraževanje*. Nova Gorica: ENO D.O.O.
- Štrukelj, Alenka, ur. 2011. *Program osnovna šola Državljska in domovinska vzgoja ter etika. Učni načrt*. Ljubljana: Ministrstvo RS za šolstvo in šport in Zavod RS za šolstvo.

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 mehanike 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 uravnorežiti 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>

Pregledni znanstveni članek/Article (1.02)

Bogoslovni vestnik/Theological Quarterly 84 (2024) 4, 945—953

Besedilo prejeto/Received:08/2024; sprejeto/Accepted:11/2024

UDK/UDC: 159.942:316.772.5

DOI: 10.34291/BV2024/04/Erzar

© 2024 Erzar, CC BY 4.0

Tomaž Erzar

Izražanje jeze in sovraštva na spletu: učinki, posledice, motivacija in meje navidezne povezanosti *Expressing Anger and Hatred Online: Effects, Consequences, Motivations and the Limits of Apparent Connectedness*

Povzetek: Prispevek analizira pojav jeznega in sovražnega govora na spletu z vidika vpletenih čustev in opisuje posledice izražanja premeščene jeze, besa, sovraštva in frustriranosti za uporabnike družbenih omrežij. Razčlenjuje psihosocialno motivacijo povzročiteljev in dokazuje, da jim spletni medij sicer nudi zadovoljitev potrebe po prepoznavnosti, odobravanju in pripadnosti, a jih obenem zapira v brezosebni, navidezni prostor povezovanja, v katerem so ključne značilnosti medosebne povezanosti v realnem svetu odsotne. Ker nastopajo večinoma v vlogi jeznih, užaljenih žrtev, akterji prezirljivega in sovražnega govora na spletu v očeh javnosti ne morejo doseči tiste prepričljivosti in pristnosti, ki jo daje izključno izražanje ranljivosti. S tega vidika je tudi doseg jeznega govora na družbenih omrežjih manjši, kot se glede na njegovo odmevnost zdi.

Ključne besede: nefunkcionalna jeza, sovraštvo, sovražni govor, družbena omrežja, mladi, ženske, potreba po prepoznanju, povezanost, vloga žrtve, pristnost

Abstract: The paper analyses the phenomenon of online hate speech in terms of the emotions involved and describes the consequences of putting displaced anger, rage, hate, and frustration onto social network users. It analyses the psychosocial motivations of the perpetrators and demonstrates that, while the online media satisfy their need for visibility, approval, and belonging, they also enclose them in an impersonal, virtual space of artificial connectedness that lacks the key features of interpersonal connectedness in the real world. Because they are mostly acting as angry, offended victims, the users of contemptuous and hateful online speech cannot achieve the public persuasiveness and authenticity that only expressions of vulnerability can provide. In this respect, the reach of angry speech on social networks appears to be smaller than its resonance would suggest.

Keywords: displaced anger, hate, hate speech, social media, youth, women, need for recognition, affiliation, victim role, authenticity

1. Uvod

Jeza, sovraštvo in nasilje so pogoste vsebine medijske produkcije – naj bodo to TV programi, filmi, glasbeni posnetki, videoigre ali oglasi.¹ Nove možnosti digitalnih medijev in uporaba virtualne tehnologije so te vsebine naredile še bolj dostopne in prevladujoče. Z njimi se je pojavila nova oblika medijskega nasilja, pri kateri ne gre le za pasivno spremljanje prikazanih nasilnih vsebin, ampak aktivno agresivno vedenje skozi vključevanje v družbena omrežja, platforme, e-igre, e-pošto ipd. (Bilewicz in Soral 2020, 5). Splet je postal poligon za razširjanje agresivnih verbalnih in video dejanj zoper druge osebe brez osebne stika; na družbenih omrežjih se razraščajo spletno nasilje oz. virtualno nadlegovanje (angl. *cyberbullying*), izsiljevanje s spolnimi vsebinami (angl. *sextortion*) in sovražni govor (angl. *hate speech*). Slednjega razširja in mu daje legitimnost tudi politika (Piazza 2020, 431–453).

V prispevku bomo s pregledom raziskav skušali poiskati odgovore na naslednji vprašanji: prvič, kako izražanje jeze in sovraštva na spletu učinkuje na uporabnike in kateri del populacije je najbolj na udaru; drugič, kdo so povzročitelji spletnega širjenja jeze in kaj jih motivira. Na podlagi ugotovitev bomo utemeljili tezo, da jezni in poniževalni govor akterjem in simpatizerjem sicer nudi zadovoljitev potrebe po povezanosti in prepoznavnosti, ne pa tudi potrebe po pristnosti, stiku in bližini, ki sta značilni za odnose v dejanskem svetu – in da moramo s tega vidika presoja ti tudi njegov domet.

2. Nefunkcionalno izražanje jeze

Ljudje v medosebnih odnosih jezo izražajo takrat, kadar jih ni strah posledic. Jezijo se v situacijah in na ljudi, ki praviloma niso vir njihove jeze. Jezo v manj varni situaciji, ki je to čustvo porodila, minimalizirajo, potisnejo vstran in jo izrazijo v bolj varni situaciji, kjer jeza deluje pretirano agresivno, nefunkcionalno in destruktivno (Parlamiš 2012, 82).

Tudi uporabniki družbenih omrežij – ker niso izpostavljeni javnemu neodobravanju in negativnim reakcijam drugih ljudi – pogosto ravna jo brez zadržkov (Rosner in Kramer 2016, 7). Širjenje prezirljivega in poniževalnega govora na podlagi rasnih, verskih, spolnih in kulturnih predsodkov v zadnjih letih narašča in je po mnenju civilnodružbenih organizacij eden osrednjih izzivov sodobnih družb (Anti-Defamation League 2021, 6).

Nakopičena, največkrat neizražena jeza, pospremljena z občutkom nemoči in predsodki ali pripisovanjem česa komu, lahko preraste v sovraštvo. Pomembna kazalca, da gre za stopnjevanje jeze v sovraštvo, sta, da se tarčam odreka človeški status in da se etiketiranju, žalitvam in prezirljivemu govoru pridruži pozivanje k nasilnim dejanjem (Fisher idr. 2018, 315). Napadi so lahko usmerjeni na osebo ali na kolektiv oz. skupino. Izražanje takega sovraštva je redko – kljub njegovi veliki odmevnosti in opaznosti.

¹ Prispevek je nastal v okviru raziskovalnega programa „P6-0269: Religija, etika, edukacija in izzivi sodobne družbe“, ki ga sofinancira Javna agencija za raziskovalno in inovacijsko dejavnost Republike Slovenije.

Izpostavljenost zaničevalnemu govorjenju o imigrantih in ranljivih družbenih skupinah, etnijah ali posameznikih vodi v politično radikalizacijo in razjeda odnose med skupinami (Winiewski idr. 2017, 143). Pogosta izpostavljenost takemu govoru povzroči, da empatijo do druge skupine zamenja prezir – ta postane prevladujoč odziv, ki nato sproža še več prezira in še bolj grobo prezirljivo govorjenje. Naraščajoče sovražno govorjenje povzroči, da se uporabnikom omrežij zaradi procesa desenzitizacije taka raba zdi vse manj problematična, s tem pa normativna. Tako sovražno govorjenje sproža še več podobnega govorjenja, širi prostor, kjer se prikazuje kot normalno in normativno – in poveča verjetnost dejanskih nasilnih izpadov zoper druge skupine ali posameznike (Bilewicz in Soral 2020, 4; Schmid idr. 2024). K širjenju sovražnega govora v zadnjih letih prispeva tudi globalizirano občutje negotovosti, frustracije, nemoči, zaskrbljenosti, nezaupanja in strahu, ki krni sposobnost konstruktivne regulacije jeze in se premešča z enega področja družbenega delovanja na drugega oz. z enega družbenega omrežja na drugega.

Družbena omrežja so idealen prostor za širjenje nefunkcionalne jeze tako zaradi svojih značilnosti kot zaradi značilnosti svojih uporabnikov. Skupine nezadovoljnih ljudi se na omrežjih najdejo, zberejo in organizirajo – pogosto prek državnih meja in prek več omrežij hkrati (Johnson idr. 2019, 261). Omrežja jezne vsebine s svojimi zmožnostmi širijo in krepijo, tako iz finančnih kot iz javnosti nerazkritih razlogov, ter z njimi soočajo kar najširšo publiko (Matamoros-Fernández in Farkas 2021, 210).

3. Učinki in posledice

Izpostavljenost prezirljivemu in poniževalnemu govorjenju ima lahko enake posledice kot drugi travmatični dogodki (Leets 2002, 342). Povzroča frustracijo, stisko, stres, strah in jezo, celo tesnobo in depresijo (Wachs idr. 2022, 418). Poleg tistih, ki so že bili njegove tarče, so ženske na nekulturno izražanje na splošno bolj občutljive in ga običajno prenašajo težje kot moški (Kenski idr. 2020, 799; Wojatzki idr. 2018). Pri storilcih nasilno govorjenje povečuje izražanje sovraštva, medtem ko za nevtralne opazovalce velja, da jezik, ki so mu izpostavljeni, in jezik, ki ga uporabljajo, oblikujeta njihov način razmišljanja in konstruiranja resničnosti (Castellanos idr. 2023, 3).

Izpostavljenost sovražnemu govoru oslabi nevrokognitivne mehanizme, s katerimi se navadno odzivamo na bolečino drugih (Pluta idr. 2023, 2). S tem odpira vrata naraščajočemu dožemanju drugih kot manj človeških in manj dovetnih za bolečino – in tako naraščajočim predsodkom in rasizmu. Pri žrtvah sovražnega govora se kaže večja nagnjenost k duševnim težavam, čustvena kriza, večji stres in slabše psihično počutje; k temu moramo prišteti še vse težave, ki izhajajo iz njihovega običajno nižjega statusa v družbi, npr. imigranti, revni, tujci, manjšine, otroci s socialnimi težavami (Dreißigacker idr. 2024, 2).

Kratkoročni učinki prikazanega nasilja na gledalce so bili že temeljito raziskani ob pojavu TV in videa. Gre za (a) povečano fiziološko vzbujenje, ki ga vzbudi opa-

zovano nasilje, (b) vtisnjenje ali aktivacijo nasilnih vsebin v spominu in (c) posne-manje nasilnega vedenja. Raziskave so prepoznale in izmerile tudi tri vrste dolgo-ročnih negativnih posledic: povečana stopnja nasilja do drugih, povečan strah postati žrtev nasilja sam, desenzitizacija za bolečino in trpljenje drugih (Bushman idr. 2015, 205).

4. Najbolj prizadeti so mladi in ženske

Ker so mladi najpogostejši uporabniki omrežij, so tem vsebinam izpostavljeni naj-pogostejše. Po raziskavah v Švici in Nemčiji je bilo sovražnemu govoru v šoli priča 67 % učencev, 65 % pa jih je bilo v zadnjih 12 mesecih vsaj enkrat priča spletnemu sovražnemu govoru. Približno 21 % učencev se je samoopredelilo za nesple-tne storilce in 33 % za tovrstne žrtve, medtem ko sta bila ta deleža pri spletnem sovražnem govoru nižja (13 oziroma 20 %). V obeh okoljih sta bila najpogostejša skupinska tarča sovražnega govora barva kože in poreklo (50 % nespletno in 63 % na spletu). Nespletni sovražni govor so večinoma izvajali sošolci (88 %), neznani viri (npr. grafiti 19 %) ali učitelji (12 %), sovražni govor na spletu pa večinoma ne-znane osebe (77 %). Najpogostejše oblike nespletnega sovražnega govora so bile žaljive šale (94 %) ter širjenje laži in govoric o članih določene družbene skupine (84 %). Opazne so bile pomembne razlike glede na državo, spol in migracijsko ozadje (Castellanos idr. 2023, 10).

Ženske in dekleta v primerjavi z moškimi in fanti na spletu doživljajo nesoraz-merno večjo in tudi zelo različno škodo. Ta vključuje: seksistične stereotipe v spletnih oglaševalskih vsebinah in algoritemsko ciljanje, negativno telesno podobo, ki jo povzroča primerjava z idealiziranimi podobami žensk, mizoginijo in zlorabo na podlagi spola, prisilni nadzor, ki ga omogoča tehnologija, ekonomsko in politično marginalizacijo ter stranske učinke dehumanizacije in degradacije žensk v mizogi-nistični pornografiji (Wojatzki idr. 2018). Mladi so na internetu najpogostejše ža-ljeni zaradi svojega videza, o čemer so poročala predvsem dekleta (81 %). Zanimi-vo je, da so v švicarsko-nemški anketi skoraj vsi anketiranci (94 %) navedli, da se jim anonimni sovražni komentarji zdijo strahopetni. Medtem ko se tretjini fantov sovražni komentarji na internetu zdijo zanimivi ali zabavni, se takšni komentarji zdijo zanimivi le 10 % deklet, zabavni pa 14 %. Ženske se na splošno bistveno bolj bojijo, da bi bile tarča sovražnih komentarjev, in poročajo o večjem negativnem psihološkem vplivu zaradi posameznih negativnih izkušenj (Cowan in Khatchado-urian 2003, 304). Poročajo o večji uporabi različnih varnostnih orodij in manjšem udobju pri več oblikah spletnega sodelovanja – npr. izražanje političnih stališč na spletu se zdi varno le 23 % žensk, v primerjavi s 40 % moških (Stevens idr. 2024).

Iz teh ugotovitev lahko sklepamo, da nebrzdano, destruktivno izražanje jeze, prezira in sovraštva na spletu učinkuje na vse, ki s temi vsebinami pridejo v stik – na tarče, nevtralne opazovalce oz. bralce ali/in gledalce ter na povzročitelje. V splošnem so najpogostejše tarče ženske in mladi. Sledijo tisti, ki so v manjšini (etnični, spolni, rasni ipd.) in so zato bolj ranljivi, npr. osebe, ki odstopajo od mo-

ških spolnih stereotipov. Nadalje tisti, ki so bolj odvisni od mnenj drugih in družbenih omrežij – to so v prvi vrsti mladi, ki take strani obiskujejo ali take vsebine iščejo in ki se sovražnega govora poslužujejo tudi sami. Posebej so na udaru mladostniki, ki so depresivni, socialno osamljeni, s slabimi osebnimi kompetencami ter s slabo povezano družino. Prizadeti (ali sekundarno travmirani) so tudi nevtalni opazovalci, saj stik z nasilnimi vsebinami vodi v zmanjšanje empatije, desenzitizacijo, občutja strahu, frustracije in jeze ter poveča nevarnost posnemanja. Nasprotno izražanje jeze in prezira učinkuje na povzročitelje: nanje deluje spodbujevalno, izražajo še več jeze in sovraštva, vsebine so radikalnejše.

5. Povzročitelji in njihova motivacija

Ugotovljene so bile povezave med spletnim nadlegovanjem oz. nasiljem in uporabo tvegane digitalne tehnologije, moralnim distanciranjem, depresivnostjo ter tradicionalnim nespletnim nadlegovanjem. Prav tako je bilo ugotovljeno, da je izražanje sovražnih vsebin na spletu povezano s predhodno viktimizacijo v konkretnem življenju. Če kdo na družbenih omrežjih sodeluje pogosto, je zanj bolj verjetno, da bo uporabljal nekulturni govor. In narobe, kdor ima močno in zaupno socialno mrežo v dejanskem življenju, zanj je manj verjetno, da bo na spletu izražal jezo (Wachs idr. 2019, 182).

Za osebe, ki se v socialnem okolju počutijo nesprejete in se tega občutja branijo tako, da se poskušajo postaviti nad druge ljudi, ima izražanje sovražnih komentarjev na omrežjih vrsto prednosti (Chester in DeWall 2017, 420). Je dostopno, po potrebi anonimno, udobno, odmevno in nudi takojšnjo potešitev. Svoje osebne primanjkljaje poskušajo nadomestiti tako, da črniijo druge, misleč, da bo potiskanje drugih navzdol dvignilo njihovo samospoštovanje. Pri njih se kronična jeza stopnjuje v sovraštvo – in to postane vzorec vedenja (Ballaschk idr. 2021, 465).

Novejše raziskave kažejo, da povzročitelji poleg osebnih značilnosti družijo potreba po pozornosti in odobravanju pri somišljenikih. Zanje je nekulturni govor priložnost za povezovanje in vključevanje v družbo, kar se npr. kaže v tem, da radikalnost objavljenih sporočil narašča s številom všečkov. V tem smislu grobost izražanja ne odraža neposredno njihovih stališč ali namer, ampak željo, da bi bili v skupnosti enakomislečnih prepoznavni in cenjeni (Walther 2022, 2–3). Družbena omrežja tem osebam nudijo naklonjeno publiko, lahek dostop do potencialnih tarč, dostop do javnosti ter takojšnjo povratno informacijo o tem, kako so somišljeniki sporočilo sprejeli. Z napadi na skupnega sovražnika (,oni') se podobno kot v realnem svetu tudi na spletu gradi skupnost enako mislečih (,mi'), ki povzročiteljem v prvi vrsti daje občutek povezanosti ter pogloblja njihov občutek za pravičnost in krivico, medtem ko je učinek na tarče zanje drugotnega pomena (Fergusson in Glasgow 2021, 245).

Iz tega izhaja, da je poleg osebnostnih potez povzročiteljev treba imeti pred očmi prav to povezovalno, ,navijaško' in hujskaško plat nasilniškega govorjenja na

spletu. Bolj smiselno od demoniziranja povzročiteljev kot posameznikov je razkrivanje situacijskih dejavnikov: povezovanje, prepoznavnost, potrditev, anonimnost, odmik od konkretnosti odnosov, skrivanje pred posledicami, poskus pridobiti večjo družbeno vlogo, nizko samospoštovanje, skrivanje za vlogo ali nastopanje v vlogi. Na sceni se za pozornost preriva veliko akterjev – in najglasnejši in najradikalnejši med njimi niso nujno tudi najbolj nevarni.

6. Vloga jezne žrtve v brezosebnem prostoru

Pri pojavu nekulturnega in nasilnega govora v digitalnih medijih imamo opraviti s posebno dinamiko širjenja, ki je po eni strani posledica neukrepanja, po drugi pa specifične narave spletnih medijev. Vedenje ljudi na družbenih omrežjih in v virtualni resničnosti je v nekaterih pogledih drugačno od vedenja istih oseb v konkretnem življenju. Mladostnica, ki je na spletu zgovorna, smešna in odprta, se lahko v realnosti izkaže za sramežljivo, tiho in prestrašeno. Ljudje se na spletu lahko skrijejo za anonimnost, tako da njihovih dejanj ni mogoče pripisati njim samim. Suler opozarja, da spletno osebno nastopanje brez zadržkov ne pomeni, da mladostnik kaže pravi jaz – pa tudi ne lažnega, ampak kombinacijo osebnostnih potez (Suler 2004, 321).

Če to ugotovitev prenesemo na razumevanje povzročiteljev, vidimo, da nekulturno govorjenje na omrežjih oz. spletu nima iste teže kot v realnosti. V njem se ne razkrivajo nujno sovražne namere in naklepi, ki bi vodili v dejanja, pač pa gre za obliko varnega povezovanja, izražanja mnenj na podlagi skupnih tarč in impulzivnega reagiranja. V njem se ne razkrivajo nekakšni globlji osebni vzgibi ali globlja socialna resničnost neke skupine – prej gre za iskanje prepoznavnosti in veljave v nastajajoči, samoorganizirani spletni skupini, ki temelji na vlogi jezne žrtve in ima skupnega nasprotnika. Žrtv ima vedno prav: lahko je neodgovorna in se izraža neodgovorno (Erzar 2017, 49).

Povezovanje na temelju vloge žrtve nima enake moči kot povezovanje v realnem svetu. Čeprav daje vtis, da gre za pristno, iskreno izražanje in samorazkrivanje, izražanje jeze in prezira na omrežjih zakriva globlje osebne razloge za jezo. Ker je varno, hitrejše, javno, manj naporno in udobno, ima manj možnosti, da bo dojeta kot pristno – se pravi kot izraz notranjega doživljanja, ki ni namenjeno samo všečkom, ampak razkriva šibke plati osebnosti, ki bi lahko bile predmet sramotjenja, ponižanja ali posmeha. Raziskovanje dinamike čustev je pokazalo, da se za jezo skrivajo mehka ali ranljiva občutja, kot so strah, negotovost, izoliranost in nemoč (Williams 2017, 6). Toda ker ta občutja na omrežjih praviloma niso deljena, je tudi povezanost, ki temelji na jezi, močna in zaupna le na videz – v resnici pa je zasilna in površinska (Walther in Whitty 2021, 124).

Nedavna nemška raziskava o učinkih socialnih signalov na zehanje v virtualnem svetu je npr. odkrila, da so bili udeleženci na prisotnost druge virtualne osebe presenetljivo neodzivni, medtem ko so bili za prisotnost realne osebe (v tem primeru

raziskovalca, čeprav je bil v drugem prostoru) izjemno dovzetni (Gallup, Anderson in Kingstone 2019, 6). Raziskovalci sklepajo, da so medosebna dinamika in njeni inhibitorni učinki v virtualnem svetu zelo drugačni od dejanske medosebne realnosti in da slednja preglaši virtualno. Opozarjajo, da medosebnih dejanj in učinkov v virtualnem svetu ne smemo preprosto preslikavati v realnost.

Naše stališče je, da je neinhibirano in nefunkcionalno izražanje jeze in sovraštva na spletu tesno povezano s tem, da se odvija v prostoru, ki je – vsaj kar zadeva občutek medosebne prisotnosti in pristnosti – brezoseben in kjer ,ni nikogar'. Temelji na potrebah povzročiteljev, da se prepoznajo v skupnem sovražniku, nevarnosti ali grožnji, da najdejo razumevanje zase, da se povežejo z enako mislečimi v na videz močno skupnost – in da lahko brez zadržkov reagirajo na jezo drugih uporabnikov. To jim daje samozavest in družbeno pripadnost. Toda povezanost, utemeljena na vlogi jezne in upravičene žrtve, je v ostrem nasprotju s pravo medčloveško pristnostjo in povezanostjo (Kraner 2023, 1017).

Ker gre torej pri tovrstnem govorjenju za nastopanje v vlogi, ki ni pristna in kaže znake priložnostnega histrioničnega nastopanja v vlogi jezne in upravičene žrtve, so tudi učinki tega govora, ko ne gre za osebne napade, šibkejši od učinkov v realnosti (Cheng idr. 2017).

7. Sklep

Jezno in nekulturno govorjenje na družbenih omrežjih se opira na privlačne sposobnosti spletnih medijev za širjenje vsebin in povezovanje. S tem povzroča škodo tako uporabnikom, ki so tarče, zlasti mladim in ženskam, kot nevtralnim opazovalcem, ki na tovrstni govor reagirajo in ga tako razširjajo. Toda prav te sposobnosti hkrati domet tega govora – ki se odvija v čedalje bolj brezosebnem in neresničnem prostoru – omejujejo. Povzročiteljem sicer daje občutek moči, povezanosti, prepoznavnosti in odmevnosti, toda ta občutek je začasen in površinski, njihovo držo pa razgalja kot držo jeznih žrtev, ki zase terjajo posebno pravico. V prispevku zato zagovarjamo tezo, da zaradi teh značilnosti jezni in sovražni govor na spletu nima potrebne pristnosti in ranljivosti za vzpostavljanje medosebnih odnosov in se, bolj ko se radikalizira, sprevrča v svojo karikaturu. S tega vidika je po našem mnenju treba presojati tudi njegove družbene učinke.

Reference

Anti-Defamation League. 2024. Online Hate and Harassment: The American Experience 2024. Anti-Defamation League. <https://www.adl.org/resources/report/online-hate-and-harassment-american-experience-2024> (pridobljeno 18. 8. 2024).

Ballaschk, Cindy, Sebastian Wachs, Norman Krause, Friederike Schulze-Reichelt, Julia Kansok-Dusche, Ludwig Bilz in Wilfried Schubarth. 2021. Dann machen halt alle mit: Eine qualitative Studie zu Beweggründen und Motiven für Hatespeech unter Schüler*innen. *Diskurs Kindheits- und Jugendforschung* 16, št. 4:463–480.

- Bilewicz, Michal, in Wiktor Soral.** 2020. Hate Speech Epidemic: The Dynamic Effects of Derogatory Language on Intergroup Relations and Political Radicalization. *Advances in Political Psychology* 541, št. 1:3–33.
- Bushman, Brad J., Mario Gollwitzer in Carlos Cruz.** 2015. There is broad consensus: Media researchers agree that violent media increase aggression in children, and pediatricians and parents concur. *Psychology of Popular Media Culture* 4, št. 3:200–214.
- Castellanos, Melisa, Alexander Wettstein, Sebastian Wachs, Julia Kansok-Dusche, Ballaschk Cindy, Norman Krause in Ludwig Bilz.** 2023. Hate speech in adolescents: A binational study on prevalence and demographic differences. *Frontiers Education* 8:1076249.
- Cheng, Justin, Cristian Danescu-Niculescu-Mizil, Jure Leskovec in Michael Bernstein.** 2017. Our experiments taught us why people troll. The Conversation. 2. 3. <https://theconversation.com/our-experiments-taught-us-why-people-troll-72798> (pridobljeno 19. 8. 2024).
- Chester, David S., in Nathan C. DeWall.** 2017. Combating the sting of rejection with the pleasure of revenge: A new look at how emotion shapes aggression. *Journal of Personality and Social Psychology* 112, št. 3:413–430.
- Cowan, Gloria, in Desiree Khatchadourian.** 2003. Empathy, ways of knowing, and interdependence as mediators of gender differences in attitudes toward hate speech and freedom of speech. *Psychology of Women Quarterly* 27, št. 4:300–308.
- Dreißigacker, Arne, Phillip Müller, Anna Isenhardt in Jonas Schemmel.** 2024. Online hate speech victimization: consequences for victims' feelings of insecurity. *Crime Science* 13:4. <https://doi.org/10.1186/s40163-024-00204-y>
- Erzar, Tomaž.** 2017. *Dolga pot odpuščanja: Psihološka, socialna in terapevtska spoznanja*. Ljubljana: Družina.
- Ferguson, Christopher J., in Brad Glasgow.** 2021. Who are GamerGate? A descriptive study of individuals involved in the GamerGate controversy. *Psychology of Popular Media* 10, št. 2:243–247.
- Fischer, Agneta H., Eran Halperin, Daphna Canetti in Alba Jasini.** 2018. Why we hate. *Emotion Review* 10, št. 4:309–320.
- Gallup, Andrew C., Daniil Vasilyev, Nicola C. Anderson in Alan Kingstone.** 2019. Contagious yawning in virtual reality is affected by actual, but not simulated, social presence. *Scientific Reports* 9, št. 1:294.
- Johnson, Nicola F., Rhys Leahy, N. Johnson Restrepo, Nicolas Velasquez, Minzhang Zheng, Pedro Manrique, Prajwal Devkota in Stefan Wuchty.** 2019. Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature* 573, št. 7773:261–265.
- Kenski Kate, Kevin Coe, in Stephen A. Rains.** 2020. Perceptions of uncivil discourse online: an examination of types and predictors. *Communication Research* 47, št. 6:795–814.
- Kraner, David.** 2023. Osamljenost in raztresenost v mreži socialnih medijev. *Bogoslovni vestnik* 83, št. 4:1003–1022.
- Leets, Laura.** 2002. Experiencing hate speech: perceptions and responses to anti-semitism and antigay speech. *Journal of Social Issues* 58, št. 2:341–361.
- Matamoros-Fernández, Ariadna, in Johan Farkas.** 2021. Racism, hate speech, and social media: a systematic review and critique. *Television and New Media* 22, št. 2:205–224.
- Parlamis, Jennifer D.** 2012. Venting as emotion regulation: The influence of venting responses and respondent identity on anger and emotional tone. *International Journal of Conflict Management* 23, št. 1:77–96.
- Piazza, James A.** 2020. Politician hate speech and domestic terrorism. *International Interactions* 46, št. 3:431–453.
- Pluta, Agnieszka, Joanna Mazurek, Jakub Wojciechowski, Tomasz Wolak, Wiktor Soral in Michał Bilewicz.** 2023. Exposure to hate speech deteriorates neurocognitive mechanisms of the ability to understand others' pain. *Scientific Reports* 13, št. 1:4127.
- Rosner, Leonie, in Nicole C. Kramer.** 2016. Verbal Venting in the Social Web: Effects of Anonymity and Group Norms on Aggressive Language Use in Online Comments. *Social Media and Society* 2, št. 3:2–11.
- Schmid, Ursula K., Anna S. Kümpel in Diana Rieger.** 2024. How social media users perceive different forms of online hate speech: A qualitative multi-method study. *New Media and Society* 26, št. 5:2614–2632.
- Stevens, Francesca, Florence E. Enock, Tvesha Sippy, Jonathan Bright, Miranda Cross, Pica Johansson, Judy Wajcman in Helen Z. Margetts.** 2024. Women are less comfortable expressing opinions online than men and report heightened fears for safety: Surveying gender differences in experiences of online harms. *ArXiv abs/2403.19037*.
- Suler, John.** 2004. The Online Disinhibition Effect. *CyberPsychology and Behavior* 7, št. 3:321–326.
- Wachs, Sebastian, Manuel Gámez-Guadix in Michelle Wright.** 2022. Online Hate Speech Victimization and Depressive Symptoms Among Adolescents: The Protective Role of

Resilience. *Cyberpsychology, behavior and social networking* 25, št. 7:416–423.

Wachs, Sebastian, Michelle F. Wright in Alexander T. Vazsonyi. 2019. Understanding the overlap between cyberbullying and cyberhate perpetration: moderating effects of toxic online disinhibition. *Criminal Behavior and Mental Health* 29, št. 3:179–188.

Walther, Joseph B. 2022. Social media and online hate. *Current Opinion in Psychology* 45:101298.

Walther, Joseph B., in Monica T. Whitty. 2021. Language, psychology, and new new media: the hyperpersonal model of mediated communication at twenty-five years. *Journal of Language and Social Psychology* 40:120–135.

Williams, Riccardo. 2017. Anger as a Basic Emotion and Its Role in Personality Building and Pathological Growth: The Neuroscientific, Developmental and Clinical Perspectives. *Frontiers Psychology* 8:1950.

Winiewski, Mikołaj, Karolina Hansen, Michał Bilewicz, Wiktor Soral, Aleksandra Świderska in Dominika Bulska. 2017. *Contempt speech, hate speech: Report from research on verbal violence against minority groups*. Warsaw: Stefan Batory Foundation.

Wojatzki, Michael, Tobias Horsmann, Darina Gold in Torsten Zesch. 2018. Do women perceive hate differently: examining the relationship between hate speech, gender, and agreement judgments. V: *Proceedings of the 14th conference on natural language processing (KONVENS 2018)*, 110–120. Dunaj: Österreichische Akademie der Wissenschaften.

Paolo Benanti. *Human in the LOOP: Decisioni umane e intelligenze artificiali*. Mondadori: Milano, 2022. 161 str. ISBN 979-12-206-0047-7.

Paolo Benanti, profesor na Papeški univerzi Gregoriana v Rimu, raziskuje zapleteno razmerje med človeškim odločanjem in UI. Pomenljivi naslov knjige Paola Benantija, ki je izšla leta 2022, *Human in the LOOP: Decisioni umane e intelligenze artificiali*, bi lahko v slovenščino prevedli *Človek v zanki: Človeške odločitve in umetna inteligenca*. V tej knjigi avtor obravnava ključno vprašanje – kako ohraniti človeško delovanje in dostojanstvo v dobi, ko stroji vse bolj vplivajo na naše odločitve? Osnovno vprašanje, ki si ga Benanti postavi na samem začetku in se ponavlja skozi vso knjigo, je: »Če umetna inteligenca (UI) postane nevidna kot voda, kakšne posledice lahko pričakujemo za človeka in družbo?« Kot voda, skozi katero lahko vidimo in je na neki način nevidna, a je z razvojem digitalne tehnike omogočila ponoven razvoj UI. Pri tem razvoju so težnje, ki želijo zakriti njegove negativne posledice za človekovo dostojanstvo. Zato je naloga etike, da pri razvoju nenehno sodeluje: opozarja na probleme oz. zanke, jih dela vidne.

Knjiga je razdeljena na dva dela. V prvem je obravnavan razvoj inteligentnih strojev in njihova sposobnost sprejemanja odločitev na podlagi podatkov in algoritmov. Benanti trdi, da umetna inteligenca sicer lahko obdeluje velike koli-

čine informacij in ugotavlja povezave, vendar pa ji manjkata razumevanje in etično razmišljanje, ki sta značilna za človeško odločanje. To je razlog, zakaj morajo ljudje ostati v te procese vključeni in biti nanje posebej pozorni: zagotoviti morajo, da umetna inteligenca ne bo delovala samostojno, kar bi lahko privedlo do odločitev brez moralnega premisleka (9–86).

V drugem delu so obravnavane različne etične dileme, kot so tiste, ki jih ponazarja eksperiment ‚moralni stroj‘; poudarjena je zapletenost človeških odločitev, ki jih ni mogoče enostavno zreducirati na algoritemske modele. Benanti znova svari pred tem, da bi stroji postali ‚nevidne‘ sile v naših življenjih, podobno kot voda za ribe – vseprisotni, a neopaženi. Treba je zagotoviti, da bodo sistemi UI zasnovani ob upoštevanju človekovega dostojanstva in pravic ter preprečiti zdrs v utilitarizem brez etičnega razmisleka (89–150).

Benanti poudarja, da je človek biološko pomanjkljivo bitje, saj brez pripomočkov ne more preživeti v nobenem naravnem okolju. V tem biološkem stanju je slabši od živali, saj nima posebej prilagojenih organov, ki bi ga neposredno zaščitili, ampak je preživetje odvisno od njegove sposobnosti, kako lahko svoje pomanjkljivosti preseže s pomočjo orodij. Zato človek ‚manjkajoče organe‘ nadomešča, krepi obstoječe in razbremenjuje delovanje organizma (12–15).

Vendar pa človek svojega preseganja ne razvija le na funkcionalni ravni, ampak tudi simbolni. Med temi najstarejšimi sledovi so jamske poslikave in pogrebni obredi, ki kažejo na človekovo iskanje in poskus izražanja človeškega presežka. Človek je torej po svoji identiteti duhovno bitje – in ta njegov konstitutivni presežek se nahaja na križišču njegovega doživljanja sveta in tehnologije: zmožnost delovanja, izražanja, sklicevanja na ‚onostransko‘ predstavlja globoko jedro njegovega eksistencialnega doživljanja. Toda v trenutku, ko je začel ustvarjati sintetično resničnost, je človekov odnos do sveta postal problematičen. Izvor te resničnosti je v presečišču več pojavov: industrializacije, vzpona kapitalističnega gospodarstva konec 19. stoletja, vojn in državnih interesov ter vrste drugih konfliktov 20. stoletja. Avtor je prepričan, da gre danes za revolucijo, ki so jo na področje znanosti in tehnologije prinesli računalniki in računalništvo. Z računalnikom lahko skoraj vsak človeški problem spremenimo v statistične podatke, grafe in enačbe. Zanj je resnično zaskrbljujoče to, da s tem ustvarjamo iluzijo, da so ti problemi rešljivi z računalniki. Danes se za dovolj natančno napovedovanje uporablja korelacija. Ker smo se na stvarnost ne samo naučili gledati kot na niz podatkov, temveč smo se velike količine podatkov naučili tudi zbirati, smo se opremili z novim orodjem za raziskovanje: ne potrebujemo več kamenčkov, kalkulatorjev, temveč nova orodja, ki nam omogočajo vidnost povezav v velikem podatkovju. Vendar pri ugotavljanju morebitnih vzročno-posledičnih povezav (korelacij), na katerih deluje UI, pogosto prihaja do lažnih korelacij. O njih govorimo takrat, ko med dvema

pojavoroma, ki statistično sicer sta v korelaciji, ne obstaja nobena dejanska neposredna vzročno-posledična povezava (16–50.)

Računalnik je postal medij par excellence 21. stoletja prav zato, ker obdeluje jezik vseh drugih medijev v digitalni obliki. Zlasti je orodje za pisanje za vse: novinarje, pisatelje, znanstvenike, inženirje, pesnike in umetnike. V veliki meri je spremenil tradicionalne tehnike pisanja, enako pa velja za urejanje, fotomontažo in sam tisk (51–52).

Cilj UI – ki je mlada disciplina, stara komaj šestdeset let in izhaja iz sklopa znanosti, teorij in tehnik – je posnemanje kognitivnih sposobnosti človeka. Pri opisu UI se uporabljajo pojmi, ki so izposojeni iz matematične logike, statistike, verjetnosti, računalniške nevrobiologije in računalništva. Problem, ki ga avtor vedno znova izpostavlja, je transparentnost delovanja: velike težave so pri razločljivosti procesov odločanja, ki jih UI izvaja (52–59).

Računalništvo postane sredstvo, s katerim računalnik lahko odredi končno zaporedje osnovnih operacij (tistih, ki jih procesor izvede v določenem času in jih ni mogoče dodatno zmanjšati ali razčleniti v druge operacije za pretvorbo vhoda v koherenten izhod).

Med drugim velja poudariti, da avtor omenja temeljno razliko med računalniškimi algoritmi in programsko kodo. Prvi so logični računski postopki za pridobitev rezultata z združevanjem osnovnih operacij; koda je izražena s programskim jezikom: formalnim jezikom, zasnovanim za realizacijo sklopov navodil (vhod), s katerimi se ustvarijo izhodni podatki (izhod). Razvojni jeziki abstraktne človeške pojme v bistvu pre-

tvarjajo v strojno kodo, tj. jezik, s katerim deluje računalnik. Pojma sta povezana, vendar na različnih ravneh. Računalniški algoritem je mogoče pretvoriti v različne programske kode. Te kode se lahko razlikujejo ne le po različnih sintaksah (formalnih jezikih, ki se danes uporabljajo), temveč tudi po načinu, kako programer logični postopek za optimizacijo izvede jezikovno (programski slog). Algoritem je končna zaporedna vrsta osnovnih pravil ali navodil, urejenih po natančnem vzorcu, ki omogoča njihovo izvajanje korak za korakom za reševanje določene vrste problema in ustvarjanje niza končnih podatkov na podlagi niza začetnih podatkov. Algoritemska navodila so lahko ali linearna ali ciklična. V sodobnih programskih jezikih lahko algoritem s pomočjo tako imenovanih knjižnic priključuje druge algoritme zunaj kode in tako uporablja algoritme, ki jih izvajajo drugi. Sodobna koda je še posebej zapletena – ne le zaradi kompleksnosti problemov, temveč tudi zato, ker noben programer programa dejansko ne napiše popolnoma sam: ustvari ga z vključitvijo in izvajanjem algoritmov, napisanih in razdeljenih v tako imenovanih 'knjižnicah' (60–71).

Danes razvoj ni usmerjen toliko v strojno učenje kot v globoko učenje. To temelji na različnih ravneh predstavitve, ki ustrezajo hierarhijam značilnosti dejavnikov ali pojmov, kjer so pojmi višje ravni opredeljeni na podlagi pojmov nižje ravni. Z drugimi besedami, stroj je tisti, ki sam najde vzorce in podatke poveže tako, da se med njimi pojavi pomen – nepovezane podatke zavrže ali pa določi kategorije podatkov, za katere se je zdelo, da za odgovor na vprašanje niso pomembni. To lahko povzamemo v treh sposobnostih stroja. Prva sposob-

nost, ki jo avtor opredeli kot videnje ali opazovanje, pomeni, da je stroj v podatkih sposoben zaznati elemente pravilnosti. Druga sposobnost vključuje zmožnost predvidevanja učinkov namernih sprememb v okolju. Od tod izhaja sposobnost predvidevanja, katera od več možnih sprememb bi lahko privedla do želenega rezultata. Da bi dosegli to zavedanje, je potrebna tretja raven: domišljija. Cilj tako imenovane močne umetne inteligence je ustvariti stroje s človeku podobno inteligenco – stroje, ki so sposobni vzročnega sklepanja in pogovora z ljudmi (72–79).

Prav korelacija je danes ključ, s pomočjo katerega je mogoče iz podatkov pridobiti informacije. Korelacija v statistiki pomeni razmerje med dvema naključnima spremenljivkama, tako da vsaka vrednost prve spremenljivke z določeno pravilnostjo ustreza vrednosti druge. Ni nujno, da gre za dejansko vzročno-posledično povezavo, ampak preprosto za tendenco spreminjanja ene spremenljivke v odvisnosti od druge. Veliko podatkovje razkriva, da je ideal ugotavljanja vzročnih mehanizmov iluzija: priča smo nastanku nove epistemološke paradigme, ki prehaja od vzročnosti h korelaciji. Vendar pa obstaja težava v načinu zbiranja in označevanja podatkov (80–83).

V nadaljevanju avtor v povezavi z zbiranjem podatkov izpostavlja nekaj primerov vprašanj: Zakaj je bil nabor podatkov ustvarjen? Kdo je oblikovanje nabora podatkov financiral? Kakšna predhodna obdelava/čiščenje je bilo opravljeno? Če se sklicujete na ljudi – ali so jim povedali, za kaj bo nabor podatkov uporabljen, in ali so se s tem strinjali? Če da, na kakšen način? Ali so jim bili na voljo mehanizmi za preklic

privolitve v prihodnosti ali za določene namene? Ali se bo zbirka podatkov posodabljala? Če da, kako pogosto in kdo jo bo posodabljal?

Sklicuje se na Isaaca Asimova, ki je pred več kot petdesetimi leti predvidel potrebo po etičnih pravilih, ki bi usmerjala vedênje robotov. Ko obravnavamo zmožnost delovanja strojev, zato najprej pomislimo na tri zakone robotike Isaaca Asimova: 1. robot človeka ne sme poškodovati ali z neukrepanjem dopustiti, da bi bil človek poškodovan; 2. robot mora ubogati ukaze ljudi, razen kadar so ti ukazi v nasprotju s prvim zakonom; 3. robot mora varovati svoj obstoj, če to varovanje ni v nasprotju s prvim ali drugim zakonom (84–85).

Benanti je prepričan, da bi se brez izrecnega etičnega upravljanja UI, algoritmov in njihovega povezovanja z ljudmi lahko znašli pred ustvarjanjem tretjega (umetnega) kognitivnega sistema. Človeški razvoj je treba razumeti kot cilj in ne kot sredstvo. Govoriti o razvoju torej pomeni, da v središče pozornosti ne smemo postavljati tehničnih zmogljivosti, temveč mora tam ostati človek. Etična uporaba tehnologije danes pomeni, da si inovacije prizadevamo spremeniti v razvoj; pomeni usmerjanje tehnologije v razvoj in za razvoj – ne pa zgolj iskanje napredka kot cilja samega po sebi (86–113).

Avtor svoje razmišljanje zaključuje s trditvijo, da so potrebni različni pristopi – vključno s humanističnim pristopom in tudi prispevkom različnih religij. Razvoj, ki je za soočanje z izzivi spreminjajoče se dobe sicer potreben, bi moral biti globalen, pluralen, ploden in prijazen. Z izrazom globalen nakazuje, da bi moral biti dostopen vsem ljudem brez izjem. S pluralnostjo želi poudariti pozornost na družbeni kontekst, v katerem živimo, in spoštljiv odnos do različnih kultur. Usmerjenost v prihodnost pa izraža z besedo ploden – tj. sposoben postaviti temelje za prihodnje generacije. Naposled izpostavlja tudi prijaznost oz. spoštljivi odnos do zemlje, ki nas gosti, do virov in do vseh živih bitij (122–151).

Ob koncu lahko poudarimo, da je knjiga odličen komentar na aktualno dogajanje, ki poteka v presečišču tehnologije in etike. Avtor zagovarja model, pri katerem človek ostaja sestavni del procesov odločanja, ki vključujejo UI. Edino na tak način lahko naš moralni kompas v vse bolj avtomatiziranem svetu ohranimo. Knjiga je poziv oblikovalcem politik, tehnologom in širši družbi, naj razmislijo, kako lahko UI odgovorno izkoristimo – obenem pa zaščitimo človeške vrednote.

David Kraner

Kristina Greif

Klientovo doživljanje spremembe v skupinski obravnavi spolnega nasilja v otroštvu po metodi relacijske družinske terapije. Doktorska disertacija. Mentor: Tanja Repič Slavič. Univerza v Ljubljani, Teološka fakulteta. 2024. XII, 293, I str. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=160542&lang=slv>

Kandidatka Kristina Greif je v svoji doktorski disertaciji z naslovom „Klientovo doživljanje spremembe v skupinski obravnavi spolnega nasilja v otroštvu po metodi relacijske družinske terapije“ raziskovala klientovo doživljanje spremembe pri skupinski obravnavi spolnega nasilja v otroštvu po metodi relacijske družinske terapije.

Naloga je razdeljena na dva dela – teoretični in empirični del. V teoretičnem delu zajema kritični pregled znanstvene literature in se osredotoča na spolno nasilje v otroštvu, posledice spolnega nasilja v otroštvu, spolno zlorabo in izbrane dejavnike težav pri žrtvah zlorab, model relacijske družinske terapije in spolno nasilje ter spolno nasilje v Svetem pismu.

Empirični del raziskave je razdeljen na dva dela, in sicer na kvantitativni in kvalitativni del, ki vključuje tudi del kvantitativnih metod.

V prvem, kvantitativnem delu raziskave je bil cilj s pomočjo vprašalnikov ugotoviti, kakšna je raven individualnih težav ter moči, izražanja agresije, zmožnosti samokontrole in čustvenega procesiranja med tistimi, ki so poročali o spolni zlorabi, v primerjavi s tistimi, ki spolne zlorabe niso doživeli. Sodelovalo je 707 udeležencev, od tega 81,26 % žensk in 18,74 % moških. Na podlagi rezultatov Lestvice spolne zlorabe (vprašalnik o travmi v otroštvu) jih je od anketiranih spolno zlorabo v otroštvu doživelo 61,39 %, 38,61 % pa spolne zlorabe ni doživelo. Rezultati so pokazali, da so imeli spolno zlorabljeni v primerjavi z nikoli spolno zlorabljenimi več težav pri izražanju agresije, več individualnih težav, nižjo samokontrolo, deloma tudi slabše čustveno procesiranje.

Kvalitativni ter deloma kvantitativni del raziskave je bil izveden v okviru skupinske obravnave spolnega nasilja po modelu relacijske družinske terapije. V skupinski obravnavi je sodelovalo 10 odraslih udeleženk, ki so doživele spolno zlorabo v otroštvu. Skupina se je sestajala enkrat na 2 tedna eno študijsko leto.

Namen tega dela raziskave je bil preveriti prispevek relacijske družinske terapije k lajšanju težav pri žrtvah spolnih zlorab v otroštvu, in sicer na področjih individualnih problemov in moči, samokontrole, agresije in čustvenega procesiranja,

predvsem pa tudi raziskati klientovo doživljanje sprememb pri terapiji oziroma klientovo doživljanje poteka terapevtskega procesa in terapevtskega delovanja pri skupinski obravnavi spolnega nasilja. Za ugotavljanje razlik in sprememb pred terapevtskim procesom in po njem so udeleženke vprašalnike izpolnile dvakrat, in sicer pred skupinsko obravnavo in po njej. Rezultati so dokazali učinkovitost in uspešnost relacijske družinske terapije pri večini obravnavanih (prej omenjenih) področij.

Poleg vprašalnikov je kandidatka z udeleženkami opravila tudi tri intervjuje: na začetku, sredini in ob koncu sestajanja skupine za spolno zlorabljenke. Analiza rezultatov na podlagi intervjujev o klientovem doživljanju sprememb in procesa skupinske obravnave spolne zlorabe v otroštvu je z interpretativnim pristopom ugotovila ključne značilnosti procesa skupinske obravnave po metodi RDT – te so se izrazile v štirih osrednjih kategorijah: 1) spremembe pri klientu samem, 2) spremembe v odnosu z drugimi, 3) pogoji pri terapiji, 4) lastnosti terapevta.

Raziskava vsekakor potrjuje, da je spolna zloraba v otroštvu pomemben dejavnik tveganja za težave v odraslosti in da je terapevtska obravnava po modelu RDT za osebe z omenjeno travmo učinkovita in uspešna. V nadaljevanju bi bilo v kvalitativno raziskavo smiselno vključiti tudi moške in pri ženskih udeleženkah po določenem obdobju od zaključka skupine morda ponovno preveriti izraženost sprememb – nekateri učinki terapije se namreč pokažejo šele z zamikom in na daljši rok. Pri kvalitativnem delu, v katerem se je s fenomenološkim pristopom poskušalo bolje razumeti, kako udeleženke terapije doživljajo spremembe in sam terapevtski proces, je bila omejitve vloga raziskovalke, ki je terapevtski proces vodila, obenem pa tudi obravnavala doživljanje sprememb in terapevtskega procesa v intervjujih. Udeleženke terapij so tako morda podajale sprejemljivejše odgovore kot bi jih sicer.

Model relacijske družinske terapije je empirično sicer že bil preverjen – tudi na področju spolnih zlorab –, vendar ne z vidika ključnih intervencij v skupinski obravnavi. Te so dokazale uspešnost terapije na področju klientovih individualnih težav in moči, regulacije agresije in funkcionalnosti čustvenega procesiranja ter zmožnosti samokontrole – prav to tudi predstavlja znanstveni prispevek. Ugotovitve raziskave bodo tako lahko v pomoč različnim strokovnim delavcem, ki se srečujejo z žrtvami spolnih zlorab in njihovimi posledicami. Vsekakor je to delo tudi dodaten prispevek k mozaiku dosedanjih spoznanj, ki nam posledice spolnih zlorab pomagajo bolje razumeti in razreševati.

Tanja Repič Slavič

Maja Kaninska

Ideološka vloga Srbske pravoslavne Cerkve v SFR Jugoslaviji v 80. letih 20. stoletja. Doktorska disertacija. Mentor: Bogdan Dolenc. Univerza v Ljubljani, Teološka fakulteta. 2024. X, 296 str. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=161156&lang=slv>

Maja Kaninska je diplomirala na Filozofski fakulteti Univerze v Prištini (2008) in magistrirala na Filološki fakulteti Univerze v Beogradu (2009); drugi magisterij iz religio-logije je dosegla na isti univerzi (2016). Proučevala je zgodovino srbskega jezika, protestantizem in zgodovino Judov v Jugoslaviji. V Srbiji je delovala tudi kot novinarka, urednica in pisateljica. V letu 2018 se je na Teološki fakulteti v Ljubljani vpisala na „Interdisciplinarni doktorski študijski program Humanistika in družboslovje – področje Religio-logija“. Disertacijo z naslovom „Ideološka vloga Srbske pravoslavne Cerkve v SFR Jugoslaviji v 80. letih 20. stoletja“ je uspešno zagovarjala 6. septembra 2024.

Tematika doktorske disertacije Maje Kaninske spada na področje religio-logije, a sega tudi na področje zgodovine in teologije. Odpira pogled na ideološke in verske silnice, ki so srbsko družbo zaznamovale v 80. letih prejšnjega stoletja. Tehnost te tematike podčrta dejstvo, da je bilo omenjeno desetletje zadnje, v katerem je sklepno obdobje sobivanja v nekdanji skupni državi sodoživljala tudi Slovenija. Objava disertacije v slovenščini je zato hkrati osvetlitev nekaterih vidikov naše skupne zgodovinske zavesti.

V omenjenem desetletju je Srbska pravoslavna cerkev (SPC) svoji verski in pastoralni vlogi dodala novo ideološko in politično razsežnost. Postala je eden ključnih dejavnikov pri političnih spremembah, zaradi česar je srbsko pravoslavje dobilo značaj neke vrste versko-politične ideologije, s katero se je identificirala večina vernikov in državljanov. Šlo je za dvosmerni proces: religija se je preoblikovala v ideologijo, ideologija pa v religijo. Obravnavana tematika spada v okvir družbeno-ideološke novejšje zgodovine in teorije religije, osvetljuje pa tudi nekatere teme katoliškega in pravoslavnega družbenega nauka, predvsem vprašanje naroda, narodnosti in nacionalizma.

Disertacija je smiselno razdeljena na tri dele. Obsežen *Uvodni del* razgrinja teoretski pristop k pojmu ideologije in zapleten odnos marksizma do religije. Pomaga razumeti razmerje med religijo in državo v obdobju druge Jugoslavije. Prikazan je pogled tako katoliških (J. Janžekovič, F. Rode, V. Grmič, A. Stres,) kot pravoslavnihi teologov (A. Jevtić, S. Terzić) in sociologov (I. Cvitković, I. Škamperle, M. Kerševan) na marksizem in socializem. Ob pregledu politične teologije v pravoslavnem svetu spoznamo korenine tesne prepletenosti in soglasja (*simfonije*) med cerkvijo in državo, ki segajo vse do konstantinskega cezaropapizma.

Prvi del disertacije („Izviri ideologizacije srbskega pravoslavja“) se najprej obširneje posveča ‚kosovskemu mitu‘, ki postane v 20. stoletju močan vir navdiha in uporabno sredstvo za krepitev nacionalne samozavesti pri Srbih. V ta kontekst spada delovanje in vpliv dveh najvidnejših ideologov srbskega ‚evangeljskega nacionalizma‘ – to sta episkop Nikolaj Velimirović in njegov učenec menih Justin Po-

povič, ki sta srbskemu pravoslavju vtisnila »svetosavski«, protizahodni in protiekumenski značaj. Poblíže spoznamo ‚svetosavje‘, ki ga teologi definirajo kot »pravoslavno krščanstvo v srbskem slogu in doživljanju« ali »pravoslavje v čisti kulturi«. Svetosavje postane sredi 20. stoletja uradna ideologija srbskega pravoslavja in vodi neposredno v verski nacionalizem.

Drugi del disertacije z naslovom „Razvoj in udejanjenje ideološke vloge srbskega pravoslavja“ usmerja pogled na povojno komunistično obdobje, ko se med cerkvijo in jugoslovansko državo vzpostavijo povsem drugačni odnosi. Te odnose urejajo zvezne in republiške ustave, ki zapečatijo ločitev cerkve od države. Zatiranje in zapostavljanje vere in vernikov, ki so ga državne institucije izvajale v prvih letih po drugi svetovni vojni, se počasi ublaži in preide v mirnejše in strpnejše sobivanje. K temu je poleg Protokola med Vatikanom in jugoslovansko vlado (1966) prispevalo tudi dolgoletno (32 let) modro vodenje SPC, ki ga je patriarh German udejanjal tako navznoter kot tudi v odnosih z državo. Njegovo obdobje zaznamujejo dogodki, ki so v te odnose vnesli precej nemira: nekanonska ustanovitev Makedonske pravoslavne cerkve, razkol srbske metropolije v ZDA in Kanadi, problematični status in končna ukinitve pravoslavnega duhovniškega društva. Prikaz pravoslavno-katoliških odnosov v prvi in drugi Jugoslaviji nas pouči, kako močno so bili ti odnosi obremenjeni zaradi konkordatske krize v 30. letih. Od 60. let dalje so se odnosi kljub medvojnim zločinom v okviru državljanske vojne izboljšali in prerasli v ekumensko sodelovanje. Konkretno obliko so dobili v številnih osebnih stikih med verskimi predstavniki ter v medfakultetnih ekumenskih simpozijih v 70. in 80. letih (Ljubljana – Zagreb – Beograd).

Poglavje z naslovom „Naveza med srbsko kulturno elito in Srbsko pravoslavno cerkvijo v 80. letih 20. stoletja“ povezuje vse niti proučevanja v nekakšno sintezo. V omenjenem desetletju se namreč oblikuje nov politično-ideološki vzorec razmerja med SPC ter družbo in politiko, ki je zahteval intenzivno poseganje SPC v aktualno politično dogajanje. S smrtjo predsednika Josipa Broza Tita (1980) so se sprožili družbeni mehanizmi, ki so privedli do končne dezintegracije države. Najmočnejši med njimi je bil nacionalizem, ki se je tako v Srbiji kot na Hrvaškem napajal iz obilice verskih simbolov in tradicij. Trda roka komunističnih oblasti je počasi popuščala in obe večinski cerkvi – Pravoslavna in Katoliška – sta začeli osvajati družbeni prostor.

Praznovanje „Trinaest vekova hrišćanstva u Hrvata“ je prebudilo samozavest hrvaških katoličanov in povzročilo množično vračanje k veri in cerkvi. Na družbeno prizorišče je stopila tudi SPC in za svoje namene pridobila vso medijsko podporo in javno mnenje. Problematika manjšinskega, izginjajočega srbskega življa na Kosovu in Metohiji je SPC spodbudila k objavi rotečega Apela za njegovo ohranitev in zaščito (1982), naslovljenega na najvišje institucije v državi. Ker so naraščale tudi mednacionalne napetosti znotraj Jugoslavije, se je na osnovi Apela oblikovala naveza med SPC in srbsko kulturno elito, ki so jo sestavljali člani Društva srbskih književnikov ter Srbske akademije znanosti in umetnosti. Nastalo je svojevrstno partnerstvo dveh duhovnih sil, ki sta stopili na prizorišče v vlogi ideoloških prenoviteljev srbskega narodnega telesa. Ob koncu 80. let se je iz teh vrst oblikovala politična opozicija, ki se je začela pripravljati na prevzem političnega vodstva. V tem kontekstu so predstavljeni pogledi nekaterih religiologov iz prostora bivše

Jugoslavije na ideološko vlogo srbskega pravoslavja. To tematiko osvetljujejo znanstveniki, kot so M. Đorđević, D. B. Đorđević in S. Naumović.

Zadnje poglavje osvetljuje vedno aktualno problematiko narodnosti in nacionalizma s teološkega, ekumenskega in sociološkega vidika. Gre za tematiko, ki se kot rdeča nit vleče skozi vsa poglavja, a je v jugoslovanskem kontekstu tako pereča, da zasluži posebno obravnavo. Avtorica se na ta pojav ozira predvsem s pomočjo nesrbskih avtorjev, v čemer je prednost, saj so s tematiko nacionalizma osebno manj obremenjeni. Njegove psihološke vidike nam pojasnjujejo psiholog S. Taylor, socialni filozof E. Fromm, psihiater E. Ringel ter filozof I. Urbančič. Isto temo avtorica razčlenjuje s pomočjo teoloških in ekumenskih vidikov, ki nam predočijo močno nagnjenost pravoslavnega sveta k identifikaciji verske in narodne prvine – to je glede na univerzalizem, ki ga zagovarja katoliški družbeni nauk, bistvena razlika. Do besede pridejo teologi kot npr. G. Larentzakis, T. Ware, J. Buciora, P. Kuzmič ter beograjski nadškof F. Perko. Ob koncu se pogled zoži na Jugoslavijo in na temelju socioloških dognanj uglednih religiologov S. Flereta in M. Velikonje spoznamo »tri jugoslovanske nacionalizme« z njihovimi verskimi prvinami in (samo)viktimizacijo, tj. prepričanjem o krivičnem trpljenju, ki so ga v zgodovini narodu prizadeli drugi.

V Dodatku so zajeti spomini in primerjalna refleksija ob osamosvojitvi Slovenije; ta del se smiselno umešča v celotno študijsko obravnavo.

V disertaciji prevladuje zgodovinsko-primerjalna metoda, ki se opira na izpričane in pisne vire, cerkvene dokumente ter diskurzivno prakso v vsakokratni družbeni javnosti, kakor jo posredujejo mediji in spontana ideologija. V tem smislu je disertacija tudi religiološko družbeno-empirična. Obdeluje in upošteva vse najrelevantnejše vire in pri obravnavani problematiki predstavlja v znanosti prevladujoča stališča.

Disertacija je v celoti izvirno znanstveno delo. Ker gre za relativno svežo, skoraj sodobno problematiko, je njena analiza na tem področju v veliki meri tudi pionirska študija. Veljavnost dokazovanja je dokumentirana, zanesljiva in pertinentna. Gre za prvi sistematični prikaz zapletenega razmerja med vero in politiko v določenem zgodovinskem kontekstu – konkretno med Srbsko pravoslavno cerkvijo in državo v sklepnem, najbolj kriznem obdobju razpadanja jugoslovanske federacije. SPC je ob krizi socialističnih in komunističnih idealov iskala nove izvire osmišljanja družbenega in osebnega življenja v mitih iz nacionalne zgodovine, v obrambi Kosova in tamkajšnjega ogroženega srbskega življa ter v vračanju k idealiziranemu liku svetega Save. Proces, v katerem je SPC iz politične nepomembnosti zavestno stopila v vrtnice političnega dogajanja, je prikazan v svojih zgodovinskih, socioloških, teoloških in psiholoških razsežnostih. Iz celote disertacije izhaja tudi spoznanje, da avtokefalne pravoslavne cerkve odlikuje živ čut za vrednoto naroda in narodnosti, oslavljen pa je čut za univerzalnost. Zato je za te cerkve nacionalizem »prirojena« skušnjava in jih izpostavlja nevarnosti, da v razmerju do politike pridejo v podrejen položaj.

Uršula Ješe

Vzpostavitev varne navezanosti v rejniških in posvojiteljskih družinah. Doktorska disertacija. Mentor: Lia Katarina Kompan Erzar. Univerza v Ljubljani, Teološka fakulteta. 2024. XIV, 255, CCCV str. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=154996&lang=slv>

Kandidatka je temo za izdelavo doktorske naloge izbrala po tem, ko je uspešno izpeljala mednarodni program norveškega raziskovalnega sklada Krila, v katerem se je posvetila terapevtskemu delu z mladostniki v rejništvu ter njihovimi rejniki in starši.

Tematike starševstva, rejništva, posvojitve ter vzpostavitve kvalitetne in učinkovite pomoči se je tako lotila v samih temeljih. Kot prva slovenska terapevtka je pridobila tudi licenco za diadično razvojno psihoterapijo pri prof. Danielu Hughe-su v ZDA in posebno licenco za delo z otroki iz najbolj ranjenih družin. Doktorske naloge se je tako lotila kot relacijska družinska terapevtka in diadična razvojna psihoterapevtka – zelo dobro opremljena za praktično delo in z močno znanstvenoraziskovalno podlago.

Glavni namen naloge je bil predlagati znanstveno podprte in učinkovite spremembe v sistemu rejništva, ki bi omogočile vzpostavljanje varne navezanosti otrok in mladostnikov v rejniških in posvojiteljskih družinah ter zmanjšale doživljanje stresa pri rejniških in posvojiteljskih starših. Skozi svoje raziskovalno delo je kandidatka jasno pokazala, kako kompleksen in zahteven je proces vzpostavljanja rejniških in posvojiteljskih odnosov – in katere so ključne kritične točke tega procesa.

Delo je zastavila na treh ravneh:

Na ravni sistema:

- na sistemski ravni je skozi poglobljene intervjuje prikazala, katere so nujne spremembe sistema na področju navezanosti pri otrocih in mladostnikih v rejništvu in posvojitvi ter na področju stresa pri rejniških in posvojiteljskih starših. V ta namen je na manjšem vzorcu (9 strokovnjakov, 30 rejniških staršev ter 14 otrok in mladostnikov v rejništvu) ugotovila, da so po mnenju otrok in mladostnikov v rejništvu, rejniških staršev in strokovnjakov v sistemu rejništva v Sloveniji potrebne spremembe, in sicer na področju priprav rejniških staršev na rejništvo (starši so pogrešali predvsem raznolikost), strokovne podpore rejniškemu staršev med rejništvom (starši so v veliki meri prepuščeni sami sebi in pogrešajo oporo) ter pri strokovni usposobljenosti strokovnih delavcev. Kot največji izziv pa so sodelujoči v raziskavi opisali vedenje otrok v rejništvu in stike z biološkimi starši, kjer ima večina udeleženih slabe izkušnje – in si želijo sprememb.

Na medosebni ravni, ravni odnosov:

- je proučevala odnose in dinamiko vzpostavljanja varnih odnosov za otroke v

rejniških in posvojitelskih družinah skozi 9-mesečni proces relacijske družinske terapije, v katerega je vključila 8 družin. Skozi terapevtska pristopa diadične razvojne psihoterapije in relacijske družinske terapije je spremljala šest rejniških in dve posvojitelski družini. Na začetku in ob koncu terapije so udeleženci izpolnili vprašalnik PSI-SF, ki meri stres na treh področjih odnosa med skrbnikom in otrokom: (1) starševska stiska; (2) disfunkcionalna interakcija med staršem in otrokom; (3) težaven otrok. Proces je spremljala tudi preko transkriptov terapevtskih srečanj, dnevniških zapisov terapevte, supervizorskih poročil in odziva udeležencev. V terapevtskem procesu so se pojavljale teme, povezane z biološkimi starši, težavami z vedanjem ter težavami pri odnosih z rejniškimi/posvojitelskimi starši in sovrstniki. Med obrambnimi vedenji, ki so bila prisotna pri vseh otrocih in mladostnikih, je bilo pri otrocih in mladostnikih v času terapevtskega procesa opaziti mnogo sramu in strahu, pozneje pa tudi žalosti. V terapevtskem procesu je bilo posebno mesto namenjeno radovednosti, sprejemanju, sočutju in igrivosti. Za otroke in mladostnike je namreč ključnega pomena, da lahko oblikujejo in najdejo pomen svojega življenja, pri čemer jim rejniški in posvojitelski starši lahko pomagajo – vendar pa se je med procesom razkrila kompleksnost in globina oškodovanosti otrok in mladostnikov v rejniških in posvojitelskih družinah. Rejniški in posvojitelski starši sicer niso uspeli razviti popolnega sprejemanja otrokovega notranjega sveta in sočutja do otrok, se je pa to zelo močno nakazovalo.

Na ravni posameznika:

- Skozi doživljanje posvojitelskih in rejniških staršev. Tretji del raziskave je zajemal skupinsko terapijo za posvojitelske in rejniške starše. Skupinska srečanja, ki so integrirala diadične razvojne psihoterapije in relacijske družinske terapije so bila namenjena doseganju sprememb na štirih področjih: (a) starševska spretnost in razumevanje otrok; (b) odnos med starši in otroci v rejništvu/posvojitvi; (c) otrokova odzivnost na skrb; (č) stabilnost namestitve. Na vseh področjih so bile zaznavne izboljšave. Starši so po lastnih besedah v skupini pridobili veliko uporabnih in zanimivih znanj, naučili so se vživljanja v otrokov svet, prijaznejšega in za starše manj stresnega odzivanja na težavne situacije, v odnosu z otrokom so se naučili bolje prisluhniti, okrepili pa so se tudi pri postavljanju meja in pravil. Obenem sta se povečala medsebojno razumevanje in sposobnost spoprijemanja s težavami. Vpliv skupine so zaznali tudi v odpiranju novih pogledov na rejništvo, otroka in starševstvo – novih starševskih veščin pa so se učili tudi drug ob drugem.

Na podlagi svoje raziskave je kandidatka dokazala, da je pot do varne navezanosti po izkušnji izgube staršev in iskanja nadomestnih staršev v sistemu rejništva in posvojitve izjemno zahtevna. Raziskava sicer potrjuje, da možnost, da otrok po izkušnji izgube razvije zaupanje v odnose, v katerih je lahko ranljiv, potolažen in sprejet v vsem, kar je, obstaja. Za doseganje tega pa je potreben dolgotrajen proces, ki je zahteven tako za otroke in mladostnike kot za njihove rejnike oziroma posvojitelje. Posebno mesto v tem procesu imajo tudi stiki z biološkimi starši, ki

so se izkazali kot mediatorska spremenljivka med varno navezanostjo v rejniški oziroma posvojiteljski družini in doživljanjem stresa pri rejniških in posvojiteljskih starših.

Avtorica se je spopadala z zahtevnimi izzivi, ki pa so za rast stroke ključni. Največji izziv je bil preseči razkorak med prakso in akademskim raziskovanjem, saj so procesi in doživljanja, ki jih je obravnavala, ena najzahtevnejših metodoloških in teoretičnih problematik. Ko se neka metoda v praksi izkaže za učinkovito, jo je znanstveno težko enoznačno pojasniti in sistem prepričati, naj jo sprejme – obenem pa je teoretično sicer jasno opisane dinamike v praksi zelo težko sprožiti in doseči. To zagato je avtorica rešila s triangulacijo metod, vključevanjem različnih ravni in zornih kotov raziskovanja, pa tudi s temeljito razdelanim teoretičnim okvirom naloge.

Istočasno se je izkazala še kot terapevtka, ki zmore vzeli pretgranih odnosov premostiti tudi v praksi in vzpostaviti varen okvir za delo z najbolj ranjenimi posamezniki.

Lia Katarina Kompan Erzar

Urška Mali Kovačič

Fenomenološka analiza religioznosti kot zaščitnega dejavnika za samomorilnost. Doktorska disertacija. Mentor: Roman Globokar. Univerza v Ljubljani, Teološka fakulteta. 2023. X, 241 str. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=154923&lang=slv>

Doktorska disertacija z naslovom „Fenomenološka analiza religioznosti kot zaščitnega dejavnika za samomorilnost“, ki jo je pripravila kandidatka Urška Mali Kovačič, je izvirno znanstvenoraziskovalno delo. Končni rezultat kaže, da raziskovanje za kandidatko ni bilo samo znanstveno delo, ampak v veliki meri tudi osebno poslanstvo, kar delu daje še poseben dodaten pečat. Skozi pisanje se čuti velika skrb in občutljivost avtorice za ljudi, ki se soočajo s stiskami in samomorilnimi mislimi.

Raziskovalno delo je bilo zasnovano interdisciplinarno, zato je bil v proces vključen somentor – psihiater in psihoterapevt prof. dr. Borut Škodlar; ima bogate izkušnje s fenomenološko metodo raziskovanja, ki jo je kandidatka izbrala kot najboljšo za preverjanje dveh osnovnih hipotez: 1) religija je pri samomorilnosti pomemben zaščitni dejavnik; 2) različni elementi religije imajo različno zaščitno vlogo. Kandidatki je uspelo integrirati znanja s področja teologije, filozofije, suicidologije, psihologije in psihiatrije. Izbira fenomenološke metode se je izkazala kot ustrezna, saj tak pristop omogoča vpogled v fenomen od znotraj, iz osebne izkušnje posameznika – ob prizadevanju po spoznavanju in formulaciji določenih občil resnic za verodostojno osvetlitev fenomena samomora kot takega. Največji prispevek doktorskega dela je prav izvedba in analiza kvalitativne raziskave med pacienti, ki se s samomorilnimi mislimi soočajo ob doživljanju težkih življenjskih stisk. Izsledki intervjujev kažejo, da je duhovnost pomembna razsežnost vsakega človeka – in je zato tudi pomembna razsežnost terapijskega procesa.

Naloga je razdeljena v dva dela: teoretičnega in praktičnega. V prvem delu kandidatka najprej predstavlja filozofske temelje za fenomenološko raziskovanje, čemur sledi opredelitev področja duhovnosti in religioznosti, temu pa še predstavitev zgodovine raziskovanja samomorilnosti in trenutnega stanja na tem področju. Prvi del zaključuje predstavitev raziskav o odnosu med religijo in samomorom. Kandidatka prikazuje dokaj celostno sliko o vlogi religije in duhovnosti pri prevenaciji samomorilnosti nekoč in danes, pri čemer uporablja obširno znanstveno literaturo s tega področja. Drugi del pa je posvečen kvalitativni raziskavi z intervjuji, ki jih je kandidatka izvedla na Univerzitetni psihiatrični kliniki v Ljubljani, pri čemer so predstavljeni metoda, proces, udeleženci in rezultati raziskave. Delo se zaključuje s temeljnimi ugotovitvami kandidatke in z odprtimi vprašanji. Dodatno vrednost delu dajejo tudi shematske predstavitve šestih vidikov povezanosti religioznosti in samomorilnosti.

S svojo raziskavo kandidatka potrjuje preventivno vlogo religije pri samomorilnosti – tako glede dojemanja smisla življenja kot glede podpore skupnosti. Izpostavlja tudi neizkoriščen potencial religijskih skupnosti na tem področju. Avtorica

ugotavlja, da bi bilo na podlagi izsledkov nujno poskrbeti, da bi pastoralni delavci pridobili znanja s področja preprečevanja samomorilnosti. Prav tako so rezultati raziskave spodbuda, da bi se v zdravstveni sistem integriralo tudi duhovno razsežnost.

Doktorsko delo Urške Mali Kovačič predstavlja pomemben prispevek k razumevanju duševnih stisk oseb s samomorilnimi mislimi ter pomena duhovnosti in religijskih skupnosti pri prevenciji samomorilnosti.

Roman Globokar

Milena Marković

Ljubosumje v partnerskem odnosu z vidika relacijske družinske paradigme in klinične prakse. Doktorska disertacija. Mentor: Tomaž Erzar. Univerza v Ljubljani, Teološka fakulteta. 2023. XII, 323, III str. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=154192&lang=slv>

Doktorska disertacija z naslovom „Ljubosumje v partnerskem odnosu z vidika relacijske družinske paradigme in klinične prakse“ kandidatke Milene Marković je izvirno znanstvenoraziskovalno delo, v katerem avtorica z vidika modela relacijske družinske terapije raziskuje pojav ljubosumja v dolgotrajnih heteroseksualnih zvezah in načine, kako se te kritične problematike loteva relacijska družinska terapija. V uvodnem, teoretičnem delu so predstavljena dosedanja znanstvena spoznanja o ljubosumnosti ter dejavniki, ki spremljajo oziroma prispevajo k dinamiki nezaupanja in ljubosumnosti v intimnih zvezah: dejavniki, povezani z izvorno družino partnerjev, dejavniki, povezani s stilom odrasle navezanosti, ter dejavniki, povezani s travmatičnimi izkušnjami. Hkrati so predstavljena tudi teoretična izhodišča relacijskega modela družinske terapije, osnova modela relacijske družinske terapije kot možne oblike pomoči parom, ko se ti soočajo s pojavom ljubosumja. Avtorica v disertaciji poveže oba vidika, tako kompleksno dinamiko ljubosumja kot model relacijske družinske terapije, in ju dopolni z razpravo o nezvestobi ter razliki v izražanju ter dojemanju nezvestobe med spoloma.

Empirični ali raziskovalni del disertacije je avtorica zasnovala v dveh delih, pri čemer je v prvem, kvantitativnem delu preverjala povezave ljubosumja z naslednjimi spremenljivkami: značilnosti izvirne družine, travmatične izkušnje v otroštvu, stil odrasle navezanosti ter nagnjenost k čustveni in fizični nezvestobi. Ugotovila je, da med ljubosumjem posameznika v partnerskem odnosu in negativnim vzdušjem, vsiljevanjem, zanemarjanjem, prisotnostjo fizične ali spolne zlorabe, zlorablajočo klimo v družini in zlorabo substanc obstaja statistično pomembna povezava. Prav tako je ugotovila povezanost med ljubosumjem in stili navezanosti. V drugi, kvalitativnem delu raziskave je avtorica preučevala značilnosti terapevtskega procesa s klienti s problemom ljubosumja. Odkrila je, da je problematika ljubosumja največkrat v ozadju drugih medosebnih ali osebnih problemov, da ni mogoče predvideti optimalnega trajanja terapevtskega procesa in da terapevtski proces po modelu relacijske zakonske in družinske terapije klientom omogoči, da se soočijo z dediščino preteklih izkušenj in uvidijo svojo vlogo v sistemu družine.

Rezultati dela predstavljajo pomemben znanstveni prispevek tako k dosedanjim ugotovitvam kot h klinični praksi. Rezultati kvantitativnega dela raziskave so potrdili pomembno povezanost ljubosumja z okoliščinami odraščanja. Ugotovitve kvalitativnega dela raziskave pa so potrdile vrednost poglobljenega terapevtskega dela s pari, v katerem se ob razreševanju očitne simptomatike in konfliktnosti razkrivajo prevzeti vzorci iskanja bližine oz. intimnosti, ki sprožajo občutke ljubosumja in nezaupanja. Prispevek doktorske disertacije predstavlja integracijo kli-

nične prakse po modelu relacijske družinske terapije in statistično ugotovljenih povezav, ki omogočajo celovitejše in globlje obravnavanje problematike ljubosumnosti ter učinkovitejše terapevtske intervence. Izvirnost doktorske disertacije je v njeni poglobljeni in kompleksni predstavitvi ter razširitvi dosedanjih spoznanj in izhodišč na področju preučevanja in preprečevanja ljubosumnosti v intimnih partnerskih zvezah oz. zakonu. Spoznanja in rezultati disertacije bodo v pomoč terapevtom pri oblikovanju terapevtskih in drugih intervencijskih programov za pare s problemi ljubosumnosti in nezaupanja.

Tomaž Erzar

Jonas Miklavčič

Transparentnost in etika sistemov algoritemskega odločanja. Doktorska disertacija. Mentor: Vojko Strahovnik. Univerza v Ljubljani, Teološka fakulteta. 2024. IX, 259 str. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=154924&lang=slv>

Doktorska disertacija z naslovom „Transparentnost in etika sistemov algoritemskega odločanja“ si za osrednji problem zadaja proučitev problema transparentnosti. Ta v kontekstu umetne inteligence označuje sklop izzivov, ki izhajajo iz dejstva, da delovanja najkompleksnejših algoritemskih sistemov strojnega učenja danes pogosto ne moremo smiselno razumeti. Ko algoritemski sistem sprejme neko odločitev, v večini primerov namreč nimamo vpogleda v to, zakaj je sprejel prav to odločitev, ki jo je. To odpira pomembna etična vprašanja, ki se tičejo zaupanja, odgovornosti, prepoznave pristranskosti, zagotavljanja pravičnega odločanja, možnosti nasprotovanja algoritemskim odločitvam, poleg teh pa še mnoga druga. Osrednji problem disertacije – ki ni le teoretične, ampak tudi praktične narave – je, da se iz etičnih razlogov zahtevi po transparentnosti ne moremo odpovedati, saj ta omogoča razumevanje razlogov za sprejeto odločitev algoritemskih sistemov. Obenem pa se iz etičnih razlogov ne želimo odpovedati uporabi netransparentnih algoritemskih sistemov, saj ti pogosto ponujajo največ upanja za hiter napredek na področju kvalitete naših življenj (tako posameznih življenj kot tudi delovanja skupnosti kot celote) in včasih, npr. v medicini, lahko dobesedno rešujejo življenja.

Disertacija je sestavljena iz treh glavnih vsebinskih sklopov: umetna inteligenca, transparentnost in etika sistemov algoritemskega odločanja. V prvem sklopu avtor analizira, ali bi problem nezdržljivosti etične zahteve po transparentnosti in netransparentnega delovanja sistemov strojnega učenja lahko reševali po tehnični poti. Avtor pojasnjuje, zakaj so sistemi (globokega) strojnega učenja sploh netransparentni, kakšne etične dileme njihova netransparentna narava odpira in ali lahko upravičeno pričakujemo, da bodo kdaj vsi sistemi UI povsem razložljivi. Slednje se izkaže za cilj, ki ga trenutno še ne moremo upravičeno pričakovati, s tem pa je jasno, da se moramo prvi morebitni rešitvi osrednjega problema disertacije za zdaj odpovedati – sistemov ne moremo narediti transparentnih.

V skladu z naslovom disertacije je drugo poglavje posvečeno predstavitvi pojma transparentnosti in analizi etične zahteve po transparentnosti. Prvi del tega sklopa neodvisno od konteksta UI analizira dejstvo, da danes v družbi obstaja izrazito močna etična zahteva po transparentnosti, ki nakazuje tudi obstoj ideala transparentnosti. V želji po tem, da bi se prepričali, ali se pri razrešitvi problema disertacije zahtevi po transparentnosti nemara lahko odpovemo, avtor sledi genealogiji pojma in izraza »transparentnost«, določi zgodovinski trenutek, ko izraz pridobi moralno konotacijo ter opisuje, kakšno načelno zahtevo danes v sodobni družbi zahteva po transparentnosti predstavlja. Pri kritiki ideala transparentnosti se avtor opira na fenomenološko analizo vpliva, ki ga ima ta ideal na medosebne odnose.

V drugem delu tega sklopa avtor nato v kontekstu UI ponazarja, kako je zahteva po transparentnosti danes v praksi izražena v dokumentih, ki predlagajo etične smernice za uporabo UI, in v pravnih dokumentih, ki predlagajo uredbo področja UI za prihajajoča leta. Tu avtor opozori na pomanjkljivosti trenutne terminologije pri navajanju praktične zahteve in ponuja svoj predlog točneje izbranih terminov za opis nians zahteve. Avtor v tem poglavju skuša oceniti, ali se za razrešitev problema disertacije etični zahtevi po transparentnosti lahko odpovemo – sklene, da se načelni in absolutni zahtevi po transparentnosti moremo (ali celo moramo) odpovedati, praktični pa morda prav res ne.

V tretjem, sklepnem poglavju iz pridobljenih rezultatov izpeljuje pet ključnih trditev, ki nakazujejo možno smer iskanja razrešitve osrednjega izziva in formulacije zaključnega sklepa. Ključni del sklepnega poglavja – in disertacije kot take – je postavitve kriterija za presojanje etične dopustnosti uporabe netransparentnih sistemov algoritemskega odločanja, ki določa, da je uporaba nerazložljivih sistemov lahko etično dopustna, če vzpostavitev zaupanja, določitev prevzemanja odgovornosti in možnost nasprotovanja algoritemskim odločitvam dosežemo po drugih poteh.

Na podlagi svojega mentorskega dela in pregleda zadnje kandidatove različice predložene doktorske disertacije z naslovom „Transparentnost in etika sistemov algoritemskega odločanja“ ocenjujem, da je doktorska disertacija izvirno znanstveno delo, ki je pomemben prispevek na področju filozofije in teologije.

Vojko Strahovnik

Miha Pate

Čustvena regulacija in strategije discipliniranja pri starših mladostnikov v okviru relacijskega družinskega modela. Doktorska disertacija. Mentor: Christian Gostečnik. Univerza v Ljubljani, Teološka fakulteta. 2023. XII, 207 str. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=154099&lang=slv>

Doktorska disertacija z naslovom „Čustvena regulacija in strategije discipliniranja pri starših mladostnikov v okviru relacijskega družinskega modela“, ki jo je pripravil kandidat Miha Pate, je izvirno znanstveno-raziskovalno delo, ki se posveča strategijam staršev pri discipliniranju mladostnika, vlogi čustev in čustvene regulacije pri starših, komunikaciji pri reševanju konfliktov v družini, obenem pa nara-vi in intenzivnosti konfliktov v odnosu med starši in mladostnikom. Poseben pou-darek daje raziskovanju povezav med čustveno regulacijo in načini, kako se pristo-pa k discipliniranju mladostnika, razreševanju konfliktov v družinskih situacijah, utemeljuje pa tudi pomembnost relacijskega družinskega pristopa pri nudenju podpore in opolnomočenju staršev v teh okoliščinah.

Disertacija prinaša kar nekaj pomembnih ugotovitev, ki so relevantne za nadalj-nje raziskovanje tega področja. V prvi raziskavi avtor dokazuje, da se starši mla-dostnikov v strategiji discipliniranja razlikujejo, in sicer je strategija mater manj učinkovita strategija očetov – po drugi strani pa imajo očetje pri reševanju kon-fliktov v družini manj podporen in pomirjajoč način komunikacije, pa tudi slabšo čustveno regulacijo. Raziskava nakazuje tudi zanimivo ugotovitev o razlikovanju med starši glede na vrstni red rojstva mladostnika: starši drugorojenih mladostni-kov so v primerjavi s starši tretjerojenih in prvorojenih mladostnikov pri discipli-niranju najbolj popustljivi. Najbolj podporno in pomirjajočo komunikacijo pri re-ševanju konfliktov v družini imajo starši prvorojencev, ki se tudi pomembno razli-kujejo od staršev drugorojenih in tretjerojenih mladostnikov. Hkrati je raziskava potrdila, da je učinkovita strategija pri discipliniranju mladostnika (tudi manj po-pustljiva in manj pretirano odzivna drža starša) povezana z boljšo komunikacijo pri reševanju družinskih konfliktov, boljšo čustveno regulacijo, pa tudi z boljšim samospoštovanjem in splošnim počutjem pri starših mladostnikov.

V drugi raziskavi pa je avtor analiziral učinkovitost pilotske psihoedukativne skupine za starše mladostnikov, zasnovane po relacijskem družinskem modelu. Rezultati so pokazali, da je skupina pomembno prispevala k boljši čustveni regu-laciji, večjemu samospoštovanju staršev, bolj umirjeni in manj napeti drži starša pri konfliktnih temah z mladostnikom; zmanjšala se je tudi popustljivost in preti-rana odzivnost pri discipliniranju mladostnika. Slednje potrjuje učinkovitost mo-dela RDT, saj kaže, da je čustvena regulacija starša za funkcionalen odnos z mla-dostnikom bistveni element. Zelo pomemben prispevek doktorske disertacije je predvsem analiza procesa učenja učinkovitega spoprijemanja s čustvi in ustreznih vzgojnih pristopov. Rezultati ugotavljajo čustvene, relacijske, vedenjske in kogni-tivne spremembe, ki so jih starši mladostnikov doživeli v procesu obiskovanja psi-

hoedukativne skupine za starše. Ob postopnem spoznavanju in razmejevanju otrokove stiske od lastne ter ob zavedanju za starše pogosto neprijetnih sprememb, ki se v mladostniku prebudijo v tem obdobju, se je tudi pri starših regulacija jeze postopoma izboljšala, obenem pa so uspešneje razreševali občutke krivde in nemoči, prilagajali svoja pričakovanja ter ob mladostnikovih vzgibih vztrajali v umirjeni, spoštljivi, a odločni drži.

Christian Gostečnik

Martin Perčič

Etična razsežnost medosebnih odnosov v luči personalistične etike Karola Wojtyły in Teologije telesa Janeza Pavla II. Doktorska disertacija. Mentor: Roman Globokar. Univerza v Ljubljani, Teološka fakulteta. 2023. 334 str. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=154148&lang=slv>

Doktorska disertacija z naslovom „Etična razsežnost medosebnih odnosov v luči personalistične etike Karola Wojtyły in Teologije telesa Janeza Pavla II.“, ki jo je pripravil kandidat Martin Perčič, je izvirno znanstvenoraziskovalno delo, v katerem je avtor ugotavljal vpliv personalistične misli z začetkov filozofskega raziskovanja Karola Wojtyły na njegovo kasnejšo teološko misel – zlasti teologijo telesa. Gre za dragoceno študijo virov katehez papeža Janeza Pavla II. o spolnosti in pomenu telesa v teologiji.

Kandidat je pri študiju virov uporabljal angleške prevode zgodnjih filozofskih spisov Karola Wojtyły in se ob tem temeljito poglobil v številne komentarje njegovih del. Pri svojem raziskovanju je vseskozi ohranjal fokus na povezavi med personalistično etiko, ki temelji na medsebojni odgovornosti, in teološkim pogledom na spolnost, ki jo najodličneje opredeljuje pojem darovanjske ljubezni. Kandidatu uspe pokazati, da je jedro teologije telesa prisotno že v zgodnjih filozofskih delih Karola Wojtyły. Njegovo raziskovanje na novo osvetljuje temeljne vire filozofske misli pri Wojtyli, pri čemer izpostavlja predvsem pomen spisov svetega Janeza od Križa. Kandidat opisuje, kako Wojtyła išče nove filozofske poti izven neosholastičnega pristopa in se pri tem naslanja tako na Kantovo pojmovanje človekovega dostojanstva kot tudi na fenomenološko metodo, ki se usmerja na celostno doživljanje človekovega bivanja. Avtor pokaže, kako Wojtyła na koncu oporo vendarle poišče v spisih Tomaža Akvinskega, ki pa jih interpretira na sodoben način. Wojtyłovo filozofsko misel avtor opredeljuje kot fenomenološki personalizem, ki klasično ontologijo združuje s sodobno fenomenološko metodo. Avtor ugotavlja, da Wojtyła v središče svoje etike postavlja ljubezen – ta kot etična norma zahteva nenehno potrjevanje dostojanstva osebe, sicer se odnos spremeni v koristoljubno izkoriščanje in popredmetenje drugega. Avtor v nalogi dokazuje, da se ta struktura etičnega razmišljanja odraža tudi v moralni teologiji papeža Janeza Pavla II. Jasno opozarja na kontinuiteto, pri čemer Janez Pavel II. v svoji teologiji filozofski koncept ljubezni nadgrajuje z darovanjsko ljubeznijo po vzoru Jezusa Kristusa. Avtor v svoji analizi izpostavlja, da je Janez Pavel II. ta teološki poudarek razvil na podlagi svojega zgodnjega raziskovanja spisov svetega Janeza od Križa.

Kot mentor želim pohvaliti kandidatovo prizadevnost, marljivost, vztrajnost, doslednost in odgovornost. Eden od pomembnih prispevkov njegovega dela so prevodi ključnih filozofskih misli Karola Wojtyły v slovenščino. Gre za zahtevna besedila, ki so v današnjem času še bolj aktualna kot ob svojem nastanku. Nalogo odlikuje aktualnost za današnji svet, saj že digitalno okolje samo spodbuja k razmišljanju o pomenu telesnosti. Prav tako svoje aktualnosti ni izgubila niti fenome-

nološka metoda. Kandidat ob koncu naloge izpostavlja sodobnost Wojtyłove antropologije in etike – morda bi lahko še bolj poudaril možnost aplikacije Wojtyłove misli pri soočanju s sodobnimi izzivi. Delo odraža, da se je avtor z mislijo Karola Wojtyła zelo poistovetil, zato je kritičnih ugovorov in vprašanj relativno malo. Je pa res, da je bil glavni fokus raziskovanja dokaz kontinuitete med personalistično etiko in teologijo telesa, kar je kandidatu nedvoumno uspelo.

Roman Globokar

Miriam Stanonik

Poslanstvo in edukativni izzivi Abrahamove žene Sare. Doktorska disertacija. Mentor: Stanko Gerjolj. Univerza v Ljubljani, Teološka fakulteta. 2024. IX, 314, XIV str. <https://repositorij.uni-lj.si/IzpisGradiva.php?id=154476&lang=slv>

Doktorska disertacija z naslovom „Poslanstvo in edukativni izzivi Abrahamove žene Sare“, ki jo je pripravila in oddala Miriam Stanonik, obsega 344 strani, od tega 14 strani prilog.

Disertacija ima poleg Uvoda še devet poglavij. Struktura doktorskega dela sledi zasnovi raziskovanja svetopisemskega lika Sare in njene življenjske zgodbe, dinamike njenih odnosov ter vprašanj, ki se nanašajo na izzive sodobnega edukativnega delovanja. Najprej se posveča raziskovanju Sarinega zakonskega odnosa z Abrahamom, in sicer s teološkega, bibličnega in psihološkega vidika; svoje raziskovanje podkrepi z umeščanjem v takratne kulturne in družbene strukturne značilnosti.

V tem kontekstu je zanimivo zlasti peto poglavje, kjer avtorica ugotavlja, da uspešen oz. integriran trajni odnos ni dejstvo – veliko bolj prisotno je iskanje in hrepenenje po njem, kar traja vse življenje. Sara in Abraham se namreč »iščeta« vse življenje in se v polnosti najmeta šele v skupnem grobu. Tako »vseživljenjsko iskanje« posredno – včasih pa celo neposredno – pomembno sooblikuje tudi edukativne odnose, ki so vpeti ne le v okvir družinske vzgoje, pač pa vključujejo tudi dinamike številnih drugih, z edukacijo povezanih življenjskih prostorov.

V skladu z zasnovo disertacije se avtorica v zadnjih dveh poglavjih na empirični ravni raziskovanja še posebej posveča vprašanjem, kako poteze biblične Sare odsevajo na področju sodobnih edukativnih odnosov v šoli, kjer je mati hkrati učiteljica lastnemu sinu ali hčerki. V tem kontekstu raziskuje tako, kako to dvojno vlogo (matere in učiteljice) doživlja mati, kot tudi, kako svojo mamo kot mamo in učiteljico hkrati doživljajo učenci in učenke. Temeljna ugotovitev pri teh vprašanjih je, da matere-učiteljice dobijo sorazmerno malo povratnih informacij, kar je razumljivo, saj otroci še niso sposobni kritične refleksije odnosov, v katere so vpeti – zlasti pa ne relevantne artikulacije. Zato matere-učiteljice izzive, s katerimi se srečujejo v obeh vlogah, zaznavajo le delno. Med koristnejšimi rezultati raziskave je zgotovo ugotovitev, da imajo otroci s sprejemanjem mater v obeh vlogah več težav, kot to mislijo matere – oz. da matere v očeh svojih otrok pri opravljanju obeh vlog niso tako suverene, kot menijo same.

Pomemben prispevek doktorskega dela pa je vendarle zlasti v tem, da se avtorica posveča raziskovanju Sare, ki je s svojo življenjsko zgodbo nekakšna »pramati izkustvenega učenja«. Dimenzije njenih temeljnih čutenj zelo posrečeno aplicira na izzive, s katerimi se srečuje in sooča sodobna mati, ki je lastnim otrokom hkrati učiteljica. Tako za Saro kot za sodobno žensko velja, da vlaganje v oblikovanje stabilnih osebnih odnosov bistveno prispeva h krepitvi osebnostne rasti, kar prinaša pozitivne učinke na materinstvo in pedagoško delo. Geštalt pedagoški

pristop, ki ga avtorica implicitno predpostavlja in mestoma uspešno uporablja, ji omogoča, da vpogled v dinamiko odnosov vse do antičnih časov in starozaveznih bibličnih pripovedi brezhibno aktualizira in preoblikuje v izzive sodobnih starševskih in učiteljskih vprašanj. Zlasti geštaltistično osvetljevanje kulturnega in družbenega ozadja bibličnega življenjskega prostora ter z njim pogojenih čutenj takratnega človeka kaže na dejstvo, da se temeljna hrepenenja skozi zgodovino bistveno ne spreminjajo – zato raziskava, ki je v tej obliki unikatna tako na domači kot širši mednarodni sceni, prinaša tudi pomembno spodbudo za krepitev stabilnih odnosov ter pozitivne samopodobe staršev-učiteljev in njihovih otrok. S praktičnega in aplikativnega vidika disertacija prinaša zelo relevantne uvide, ki jih je mogoče uporabiti tako pri oblikovanju stabilnih odnosov kot pri krepitevi zdrave in pozitivne samopodobe staršev in otrok – podpirajo pa tudi večjo diferenciacijo vlog staršev v smislu, naj bodo starši v šoli učitelji, doma pa starši. Le tako se bodo njihovi otroci lahko v šoli počutili kot učenci, doma pa kot sinovi in hčere.

Stanko Gerjolj

Nataša Vrbnjak

Samopodoba ter sposobnosti za bližino in regulacijo čutenj odraslih z izkušnjo posvojitve z vidika relacijske družinske terapije. Doktorska disertacija. Mentor: Tanja Pate. Univerza v Ljubljani, Teološka fakulteta. 2024. X, 319, XVII str. <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=158199&lang=slv>

Doktorska disertacija kandidatke Nataše Vrbnjak z naslovom „Samopodoba ter sposobnosti za bližino in regulacijo čutenj odraslih z izkušnjo posvojitve z vidika relacijske družinske terapije“ je izvirno znanstvenoraziskovalno delo, ki se posveča raziskavi vpliva posvojitve na samopodobo, doživljanje strahu pred intimo, izražanje in zavedanje čustev ter na čustveno stabilnost pri odraslih z izkušnjo posvojitve. Teza se specifično posveča terapevtskemu pristopu relacijske družinske terapije (RDT), ki lahko omenjene vidike občutno izboljša.

Raziskava je bila izvedena na vzorcu odraslih z izkušnjo posvojitve in je razdeljena na dva dela. V prvem avtorica s kvantitativno metodo zbiranja podatkov primerja samopodobo, doživljanje strahu pred intimo in način regulacije čustev pri osebah z izkušnjo posvojitve, tistih, ki so odraščali v biološki družini, pri odraslih z izkušnjo posvojitve, ki so bili vključeni v terapijo po modelu RDT, in pa pri kontrolni skupini brez terapije. Rezultati raziskave so pokazali, da se posvojeni pri oceni samopodobe, strahu pred intimo in regulaciji čustev od drugih v povprečju razlikujejo, vendar pa razlike niso bile statistično značilne.

V drugi, kvalitativni raziskavi avtorica ugotavlja, da so ključni dejavniki, ki prispevajo k dobri samopodobi odraslih z izkušnjo posvojitve, varna navezanost na starše posvojitelje, pozitivna vzgoja, poznavanje svojega biološkega izvora, dobra samopodoba v otroštvu in adolescenci, dobri medosebni odnosi ter samoodgovornost za samopodobo v odraslosti. Terapevtski pristop RDT je medtem pokazal pozitivne učinke na izboljšanje samopodobe, zmanjšanje strahu pred intimo in izboljšanje regulacije čutenj. Ključni vidiki uspešne terapije vključujejo empatičen odnos do klienta, regulacijo afektov, zrcaljenje klientovih sposobnosti, psihoedukacijo, terapevtove osebnostne lastnosti in skupinsko dinamiko.

Disertacija prinaša pomemben prispevek k razumevanju dolgotrajnih posledic posvojitve na psihološki razvoj odraslih z izkušnjo posvojitve. Poudarja pomen varne navezanosti in pristopa RDT kot ustrezne metode, ki pomembno prispeva k izboljšanju psihološkega blagostanja oseb z izkušnjo posvojitve – kar omogoča lažje soočanje s preizkušnjami ter posameznika usmerja h grajenju pozitivne samopodobe in zdravih medosebnih odnosov v odraslosti.

Tanja Pate

Bogoslovni vestnik 84 (2024)

TEMI / THEMES

Vrednote v judovsko-krščanskih virih in tradiciji, moč čustev ženskih likov in možnosti dialoga [3:447–720]
Values in Judaeo-Christian Sources and Traditions, the Strength of Emotions of Female Characters and the Possibilities for Dialogue

Digitalizacija in umetna inteligenca kot izziva za etiko in religijo [4:737–953]
Digitalisation and AI as Challenges for Ethics and Religion

RAZPRAVE / ARTICLES

Ahlin Doljak, Sara Voluntary Termination of Life and Conscientious Objection: A Comparative Review within the European Union and Slovenia [2:365–376]
Prostovoljna prekinitev življenja in ugovor vesti: primerjalni pregled v Evropski uniji in Sloveniji

Ambrožič, Matjaž Svetna in sveta zgodovina v razmišljanjih visokosrednjeveških teologov [1:7–17]
Secular and Sacred History in the Reflections of High Medieval Theologians

Arko, Alenka Teologija o laikih [3:645–673]
Theology on Lay People

Avsenik Nabergoj, Irena Josip Vidmar o katolištvu Ivana Cankarja in pogledih Izidorja Cankarja skozi prizmo sodobne teološke interpretacije katolištva [2:263–283]
Josip Vidmar on the Catholicism of Ivan Cankar and the Perspectives of Izidor Cankar through the Lens of Contemporary Theological Interpretation of Catholicism

--- Čudežne poti do materinstva: transformacija Haninih molitev (1 Sam 1–2) v Marijini hvalnici (Lk 1,46–55), apokrifu Liber Antiquitatum Biblicarum in Targumu prerokov [3:475–496]
Miraculous Paths to Motherhood: The Transformation of Hannah's Prayers (1 Sam 1–2) in Mary's Hymn (Lk 1:46–55), the Apocryphon Liber Antiquitatum Biblicarum and the Targum of the Prophets

Bogataj, Jan Domink Samoreprezentacije in identiteta zgodnjekršćanskih avtoric v prizmi starozaveznih likov [3:545–578]
Self-Representations and Identity of Early Christian Women Authors in the Prism of Old Testament Characters

Bogataj, Jan Dominik, in Anja Božič Enej Silvij Piccolomini in slovenski prostor: novi viri za njegovo tržaško obdobje [1:55–75]
Aeneas Silvius Piccolomini and the Slovenian Territory: New Sources for his Period in Trieste

Borstner, Bojan, in Jana Vrdoljak Addressing Agape in Relation to Humanity [2:239–252]
Obravnava agape v odnosu do človeštva

Borto, Paweł	Postulato di conoscenza di Dio per esperienza secondo George Tyrrell e la sua valutazione teologica <i>Postulate of knowledge of God by experience according to George Tyrrell and his theological evaluation</i> <i>Postulat spoznanja Boga na podlagi izkustva po Georgeu Tyrrellu in njegovo teološko vrednotenje</i>	[2:253–262]
Celarc, Matjaž	Rebekah – As a Chosen Agent of God <i>Rebeka – kot izvoljena Božja predstavnik</i>	[3:459–474]
Centa Strahovnik, Mateja	Etika zmožnosti, priporočilni in pogovorni sistemi UI in doseganje dobrega življenja <i>Ethics of Capabilities, AI Recommendation and Conversation Systems and Attaining a Good Life</i>	[4:771–781]
Chomiccki, Špela, Renato Podbersič in Darko Friš	»Vse, kar je lahko nemštvo na poti, mora izginiti – tudi duhovniki«: Župnija Sv. Trojica v Halozah v obdobju druge svetovne vojne <i>“Anything That Can Stand in the Way of Germanism Must Disappear – even Priests”: The Parish of Sv. Trojica in Haloze during the Second World War</i>	[1:91–104]
Erzar, Tomaž	Izražanje jeze in sovraštva na spletu: učinki, posledice, motivacija in meje navidezne povezanosti <i>Expressing Anger and Hatred Online: Effects, Consequences, Motivations and the Limits of Apparent Connectedness</i>	[4:945–953]
Ferkolj, Janez	Živo upanje <i>Living Hope</i>	[3:565–578]
Ganc, Damijan, and Drago Jerebic	The Application of Psychotherapeutic Interventions into Pastoral Practice: Possibilities and Reflections <i>Uporaba psihoterapevtskih intervencij v pastoralni praksi: možnosti in razmisleki</i>	[1:177–188]
Globokar, Roman	Pogled Katoliške Cerkve na razvoj in uporabo umetne inteligence <i>The Catholic Church's View on the Development and Use of Artificial Intelligence</i>	[4:867–883]
Gregorčič, Rok	Razmerje med človeško in umetno inteligenco v luči Habermasove teorije <i>The Relation between Human and Artificial Intelligence in Light of Habermas's Theory</i>	[4:797–806]
Jeglič, Urška	Virtualna resničnost kot pripomoček pri poučevanju religijskih vsebin v šolah <i>Virtual Reality as a Tool for Teaching Religion in Schools</i>	[4:921–930]
Jerebic, Sara, and Jasminka Zloковиć	Promotion of Mental Health through Interdisciplinary Collaboration of Family Pedagogy and Family Therapy in the Education Programme from the Perspective of Students' Experience <i>Spodbujanje duševnega zdravja z interdisciplinarnim sodelovanjem družinske pedagogike in družinske terapije v študijskem programu z vidika doživljanja študentov</i>	[1:165–176]
Joób, Máté	The Transgenerational Effects of Church Collaboration under the Communist-Socialist Regime <i>Transgeneracijski učinki sodelovanja Cerkve v komunistično-socialističnem režimu</i>	[1:105–115]
Klun, Branko	Problem religioznega izkustva v digitalno transformiranem svetu: Eksistencialno fenomenološki pristop <i>The Problem of Religious Experience in a Digitally Transformed World: An Existential-Phenomenological Approach</i>	[1:19–32]

---	Kaj pomeni misliti? Eksistencialno razumevanje in umetna inteligenca <i>What is Called Thinking? Existential Understanding and Artificial Intelligence</i>	[4:741–756]
Kochaniewicz, Bogusław	Dominicans Towards the Liturgical Celebration of the Immaculate Conception of the Blessed Virgin Mary (13th –16th Centuries) <i>Odnos dominikancev do liturgičnega praznovanja brezmadežnega spočetja Device Marije (13.–16. stoletje)</i>	[1:139–149]
Kolar, Bogdan	Stiki slovenskih misijonarjev s kanadskimi indijanci <i>Slovenian Missionaries in Touch with the Canadian Native Americans</i>	[3:693–707]
Krafl, Pavel	Ordination of the Clergy and Ecclesiastical Law: A Brief Overview of the Ordination Law of Czech Dioceses in the Middle Ages <i>Ordinacija duhovnikov in cerkveno pravo: kratak pregled ordinacijskega prava čeških škofij v srednjem veku</i>	[1:129–138]
Krajnc, Aljaž	Jožef in Potifarjeva žena v poznoantičnem judovskem in krščanskem slovstvu v semitskih jezikih <i>Joseph and Potiphar's Wife in Late Antique Christian and Jewish Writings in Semitic Languages</i>	[3:579–605]
Krajnc, Slavko	Recepcija in udejanjenje liturgične preнове drugega vatičanskega cerkvenega zbora v Cerkvi na Slovenskem <i>The Reception and Realisation of Liturgical Renewal of the Second Vatican Council in the Church of Slovenia</i>	[3:633–643]
Kraner, David	Vloga moči in uporaba UI v mrežni družbi <i>The Role of Power and the Use of AI in the Network Society</i>	[4:931–944]
Krašovec, Jože	Pot do novega katoliškega prevoda Svetega pisma (2024) med teorijo in prakso <i>The Way to a New Catholic Translation of the Bible (2024) between Theory and Practice</i>	[3:525–543]
Kreš, Barbara, Robert Cvetek in Mateja Cvetek	Authenticity Scale: Psychometric Evaluation of the Slovenian Version <i>Lestvica avtentičnosti: psihometrično ovrednotenje slovenske različice</i>	[2:419–432]
Malmenvall, Simon	State and Christian Enlightenment: Background of the Mass Elementary Education in Central Europe <i>Država in krščansko razsvetljenstvo: ozadje množičnega osnovnega šolstva v srednji Evropi</i>	[4:807–822]
Milotić, Ivan, e Ivan Obadić	Nascita degli archivi pubblici nell'intreccio tra diritto civile e canonico nel XVI secolo <i>Birth of Public Archives in the Intertwining of Civil and Canon Law in the 16th Century</i> <i>Rojstvo javnih arhivov v prepletu civilnega in kanonskega prava v 16. stoletju</i>	[2:305–316]
Nagy, Levente	Reformation And/or Union: Notes on Early Seventeenth Century Romanian Reformation in Transylvania in a European Context <i>Reformacija in/ali Unija: Zapiski o romunski reformaciji v Transilvaniji na začetku sedemnajstega stoletja v evropskem kontekstu</i>	[2:317–335]
Osredkar, Mari Jože	Umetna inteligenca – prepovedan sad iz raja? <i>Artificial Intelligence - Forbidden Fruit from The Garden of Eden?</i>	[4:823–834]

Palmisano, Maria Carmela	Abigàil nel contesto letterario di 1 Sam 25: mediatrice di sentimenti di vita	
	<i>Abigail in the Literary Context of 1 Sam 25: Mediator of Feelings of Life</i>	[3:497–509]
	<i>Abigájila v literarnem kontekstu 1 Sam 25: pospeševalka čustev življenja</i>	
Pap, Levente	The Apostolic Puzzle: <i>Illyricum Sacrum</i> and the Origins of Christianity in the Ancient Roman Province	[1:117–128]
	<i>Apostolska uganka: Illyricum Sacrum in začetki krščanstva v starodavni rimski provinci</i>	
Pate, Tanja	Uporabnost empirične fenomenologije na področju raziskovanja družine in medosebnih odnosov: teoretični in metodološki vidiki	[2:391–401]
	<i>The Utility of Empirical Phenomenology in the Field of Family and Interpersonal Relations Research: Theoretical and Methodological Aspects</i>	
Platovnjak, Ivan	Čustva žalovanja po smrti bližnje osebe pri starozaveznih ženskah	[3:511–544]
	<i>The Emotions of Grief of Old Testament Women After the Death of a Loved One</i>	
Platovnjak, Ivan, and Tone Svetelj	Artificial Intelligence and <i>Imago Dei</i> : A New Dilemma for Philosophical and Theological Anthropology	[4:835–846]
	<i>Umetna inteligenca in Imago Dei: Nova dilema za filozofsko in teološko antropologijo</i>	
Pobežin, Gregor	Vzorci kljubovanja in čustvovanja v Prudencijevem <i>Mučniškem vencu</i>	[3:607–618]
	<i>Patterns of Spite and Emotion in Prudentius' Peristephanon</i>	
Primc, Liza	Edith Stein – sv. Terezija Benedikta od Križa o duhovnosti ženske	[3:655–673]
	<i>Edith Stein – St. Teresa Benedicta of the Cross on Spirituality of a Woman</i>	
Pucer, Matevž, in Andrej Naterer	Duhovnost in religioznost v slovenski filmski produkciji: analiza Baze slovenskih filmov	[2:285–304]
	<i>Spirituality and Religiosity in Slovenian Film Production: An Analysis of the Slovenian Film Database</i>	
Rožman, Lucija	Vprašanje človeške ustvarjalnosti v dobi UI	[4:857–866]
	<i>The Question of Human Creativity in the Age of AI</i>	
Salim, Emil	Working from Home: Insights from <i>The Rule of St. Benedict</i>	[1:151–164]
	<i>Delo od doma: spoznanja iz Pravila svetega Benedikta</i>	
Saje, Andrej	Reforma Rimske kurije kot model prenove za vso Cerkev	[3:709–720]
	<i>The Reform of the Roman Curia as a Model of Renewal for the Entire Church</i>	
Slatinek, Stanislav	Disforija spola in pravica do pravnega priznanja spola oz. do spremembe spola: dileme kanonskega prava	[2:351–363]
	<i>Gender Dysphoria and the Right to Legal Recognition of Sex or Gender Reassignment: Dilemmas of Canon Law</i>	
Stala, Józef	“To be” or “To Have” in the Consumerist Reality of the Post-modern Society	[1:45–54]
	<i>„Biti” ali „imeti” v potrošniški realnosti postmoderne družbe</i>	
Stegu, Tadej	Antropološki izzivi kateheze v času umetne inteligence	[4:909–919]
	<i>Anthropological Challenges of Catechesis in the Age of Artificial Intelligence</i>	

Strahovnik, Vojko	Algoritemsko nasilje <i>Algorithmic Violence</i>	[4:783–795]
Šegula, Andrej	Priložnosti in pasti umetne inteligence v pastoralni <i>Opportunities and Pitfalls of Artificial Intelligence in Pastoral</i>	[1:33–43]
---	Umetna inteligenca v luči praktične pastoralne teologije <i>Artificial Intelligence in the Light of Practical Pastoral Theology</i>	[4:899–908]
Šimac, Miha	Med tronom in oltarjem: O ljubljanskem kongresu iz dnevnika cesarja Franca I. <i>Between the Throne and the Altar: The Congress of Laibach from the Diary of Emperor Franz I</i>	[3:675–691]
Trontelj, Nik	Teološke kreposti vere, upanja in ljubezni kot temelj človeškosti v odnosu do transhumanizma <i>The Theological Virtues of Faith, Hope and Love as the Foundation of Humanity in Relation to Transhumanism</i>	[4:847–855]
Valenta, Tanja	Unresolved and Ambiguous Losses in Childhood: Invisible Wounds and the Potential for Resilience in Adulthood – Healing Through Prayer and New Relationships <i>Nerazrešene in dvoumne izgube v otroštvu: nevidne rane in možnosti za odpornost v odrasli dobi – zdravljenje z molitvijo in novimi odnosi</i>	[2:403–417]
Vodičar, Janez	Kateheza kot pomoč umetni inteligenci <i>Catechesis as an Aid for Artificial Intelligence</i>	[4:885–897]
Vranješ, Nikola	Some Elements of the Relationship to the Holy as a Challenge for Contemporary Pastoral Activity <i>Nekateri elementi odnosa do Svetega kot izziv za sodobno pastoralno dejavnost</i>	[2:377–389]
Vrbnjak, Nataša	Relacijska travma pri posvojenih otrocih in njene posledice <i>Relational Trauma in Adopted Children and Its Consequences</i>	[1:189–199]
Zajc, Neža	Sveti Maksim Grk (Mihail Trivolis, Arta, ok. 1469–Maksim Grek, Moskva, 1556) in njegov molitveni nagovor Matere Božje <i>Saint Maximos the Greek (Mikhail Trivolis, Arta, c. 1469 – St. Maximus the Greek, Moscow, 1556) and His Prayerful Address to the Mother of God</i>	[3:619–632]
Zsolt, Csutak	An Overview of the Special Nexus Between Progressivist Theorems and Mainline Christian Churches in Hungary <i>Pregled posebne povezave med progresivističnimi teorijami in glavnimi krščanskimi Cerkvami na Madžarskem</i>	[2:337–350]
Žalec, Bojan	Človeški podobna umetna inteligenca in superinteligenca: verjetnost in glavne težave njenega oblikovanja <i>Human-like Artificial Intelligence and Superintelligence: The Probability and Main Challenges of Their Design</i>	[4:757–770]
Žepič, Vid	Monald Koprski in zastavna pogodba v delu <i>Summa de iure canonico</i> <i>Monaldus Iustinopolitanus and the Pledge Contract in Summa de iure canonico</i>	[1:77–90]

UVODNIKA / INTRODUCTIONS

Avsenik Nabergoj, Irena	Uvodnik	[3:447–457]
Žalec, Bojan	Uvodnik	[4:737–739]

OCENE / REVIEWS

Tom Wright and Michael F. Bird	<i>Jesus and the Powers: Christian Political Witness in an Age of Totalitarian Terror and Dysfunctional Democracies</i> (Unai Buil-Zamorano)	[2:438–440]
Peter G. Kirschschlaeger	<i>Ethical Decision-Making</i> (Roman Globokar)	[3:721–722]
Paolo Benanti	<i>Human in the LOOP. Decisioni umane e intelligenze artificiali</i> (David Kraner)	[4:955–958]
Blaž Otrin	<i>Socialna dejavnost Katoliške cerkve v ljubljanski (nad)škofiji od konca 18. stoletja do druge svetovne vojne</i> (Bogdan Kolar)	[1:201–203]
Sami Al-Daghistani	<i>Islam in ljubezen: Pričevanja o arabsko-islamski intelektualni zgodovini</i> (Mari Jože Osredkar)	[2:433–436]
Ivan Platovnjak	<i>Preobrazba bolečine – Skupaj na poti žalovanja</i> (Andrej Šegula)	[2:436–438]

POROČILO / REPORT

Poročilo o slovesni predstavitvi novega prevoda Svetega pisma: Atrij ZRC SAZU, torek, 12. 3. 2024, ob 11.00 (Jože Krašovec)	[1:205–231]
---	-------------

NOVI DOKTORJI IN DOKTORICE ZNANOSTI / NEW DOCTORS OF SCIENCE

Greif, Kristina	Klientovo doživljanje spremembe v skupinski obravnavi spolnega nasilja v otroštvu po metodi relacijske družinske terapije (Tanja Repič Slavič)	[4:959–960]
Kaninska, Maja	Ideološka vloga Srbske pravoslavne Cerkve v SFR Jugoslaviji v 80. letih 20. stoletja (Bogdan Dolenc)	[4:961–963]
Ješe, Uršula	Vzpostavitev varne navezanosti v rejniških in posvojiteljskih družinah (Lia Katarina Kompan Erzar)	[4:964–966]
Mali Kovačič, Urška	Fenomenološka analiza religioznosti kot zaščitnega dejavnika za samomorilnost (Roman Globokar)	[4:967–968]
Marković, Milena	Ljubosumje v partnerskem odnosu z vidika relacijske družinske paradigme in klinične prakse (Tomaž Erzar)	[4:969–970]
Miklavčič, Jonas	Transparentnost in etika sistemov algoritemskega odločanja (Vojko Strahovnik)	[4:971–972]
Pate, Miha	Čustvena regulacija in strategije discipliniranja pri starejših mladostnikov v okviru relacijskega družinskega modela (Christian Gostečnik)	[4:974–975]
Perčič, Martin	Etična razsežnost medosebnih odnosov v luči personalistične etike Karola Wojtyły in Teologije telesa Janeza Pavla II. (Roman Globokar)	[4:975–976]
Stanonik, Miriam	Poslanstvo in edukativni izzivi Abrahamove žene Sare (Stanko Gerjolj)	[4:977–978]
Vrbnjak, Nataša	Samopodoba ter sposobnosti za bližino in regulacijo čutenj odraslih z izkušnjo posvojitve z vidika relacijske družinske terapije (Tanja Pate)	[4:979]

Bogoslovni vestnik 84 (2024)

Ambrožič, Matjaž, doc. dr.
Arko, Alenka, izr. prof. dr.
Avsenik Nabergoj, Irena, prof. ddr.; znan. svet.
Celarc, Matjaž, doc. dr.
Centa Strahovnik, Mateja, doc. dr.
Cvetek, Robert, prof. dr.
Erzar, Tomaž, prof. dr.
Ferkolj, Janez, doc. dr.
Globokar, Roman, izr. prof. dr.
Ivanič, Peter, izr. prof. dr.
Jamnik, Anton, prof. dr.
Jerebic, Drago, doc. dr.
Jerebic, Sara, doc. dr.
Klun, Branko, prof. dr.
Kolar, Bogdan, zasl. prof. dr.
Krajnc, Slavko, prof. dr.
Kraner, David, doc. dr.
Krašovec, Jože, akad., zasl. prof. dr.
Krvina, Domen, doc. dr.; znan. sod.
Malmenvall, Simon, doc. dr.
Maly, Anđelo, doc. dr.
Maver, Aleš, izr. prof. dr.
Nežič Glavica, Iva, doc. dr.
Osredkar, Mari Jože, izr. prof. dr.
Palmisano, Maria Carmela, izr. prof. dr.
Pap, Levente, izr. prof. dr.
Pate, Tanja, doc. dr.
Pevce Rozman, Mateja, doc. dr.
Platovnjak, Ivan, doc. dr.
Podbersič, Renato, znan. sod.
Pohar, Borut, doc. dr.
Preinfalk, Miha, doc. dr., viš. znan. sod.
Prijatelj, Erika, doc. dr.
Simonič, Barbara, prof. dr.
Skralovnik, Samo, doc. dr.
Slatinek, Stanislav, izr. prof. dr.
Stegu, Tadej, doc. dr.
Strahovnik, Vojko, izr. prof. dr.; viš. znan. sod.
Strehovec, Tadej, doc. dr.
Šegula, Andrej, doc. dr.

Šimac, Miha, doc. dr.

Špelič, Miran, doc. dr.

Štrukelj, Anton, zasl. prof. dr.

Turnšek, Marjan, doc. dr.

Valčo, Michal, prof. dr.

Valentan, Sebastijan, doc. dr.

Vodičar, Janez, prof. dr.

Vranješ, Nikola,izr. prof. dr.

Zjawin, Jacek, doc. dr.

Žalec, Bojan, prof. dr.; znan. svet.

Navodila sodelavcem

Rokopis znanstvenega ali strokovnega članka, ki ga avtor/ica pošlje na naslov *Bogoslovnega vestnika*, je besedilo, ki še ni bilo objavljeno drugod niti ni drugod v recenzijemskem postopku. Rokopis je treba poslati v elektronski obliki. Elektronska oblika naj obsega le osnovne programske definicije (urejevalnik besedila Word s standardno obliko pisave brez dodatnih slogovnih določil). Pisna oblika rokopisa naj ima velikost črk 12 pt (Times), razmik med vrsticami 1,5, opombe 10 pt.

Zaradi **anonimnega recenzijskega postopka** mora imeti vsak rokopis prijavo oziroma spremni list, na katerem avtor/ica navede svoje ime in priimek, naslov članka, svoje ključne biografske podatke (akademski naziv, področje dela, poslovni naslov in/oziroma naslov, na katerem prejema pošto, elektronski naslov itd.) in izjavo, da besedilo še ni bilo objavljeno in da še ni v recenzijemskem postopku. Obrazec prijave je na spletni strani revije (<http://www.teof.uni-lj.si/bv.html>). Na drugi strani, ki je prva stran besedila, pa naj navede samo naslov članka brez imena avtorja.

Rokopis znanstvenega članka ima **povzetek in ključne besede**, in sicer na posebni strani (listu). Povzetek naj obsega do 160 besed oziroma do 800 znakov. Povzetek obsega natančno opredelitev teme članka, metodologijo in zaključke.

Obseg rokopisa izvirnega znanstvenega članka naj praviloma ne presega dolžine ene avtorske pole (30.000 znakov); pregledni članki in predhodne objave naj ne obsegajo več kot 20.000 znakov, poročila ne več kot 10.000 znakov, ocene knjig ne več kot 8.000 znakov. Za obsežnejša besedila naj se avtor dogovori z urednikom. Besedila daljša od 8.000 znakov morajo vsebovati podnaslove. Tabele in slike morajo biti izpisane na posebnem listu papirja, v besedilu članka mora biti označeno mesto, kam sodijo.

Naslov članka mora biti jasen, poveden in ne daljši od 100 znakov.

Na koncu članka pripravimo **seznam referenc oz. literature**. Knjižna dela navajamo v obliki: Priimek, Ime. Letnica. Naslov. Kraj: Založba. Npr.: Janžekovič, Janez. 1976. *Krščanstvo in marksizem: od polemike do razgovora*. Celje: Mohorjeva družba. Članke iz revij navajamo v obliki: Priimek, Ime. Letnica. Naslov članka. Ime publikacije letnik:prva-zadnja stran. Npr.: Krašovec, Jože. 1991. Filozofsko-teološki razlogi za odpuščanje. *Bogoslovni vestnik* 51:270-285. Pri referencah z interneta navedemo spletni naslov z datumom pridobitve dokumenta. Če je bil dokument objavljen v periodični publikaciji, to označimo za naslovom dokumenta. Priimek, Ime. Letnica. Naslov. Ime publikacije, dan. mesec. URL (pridobljeno datum. mesec leto). Npr.: Rebula, Alojz. 2006. Rekatolizacija ali recivilizacija? *Družina*, 7. aprila. <http://www.druzina.si/ICD/spletnastran.nsf/> (pridobljeno 16. oktobra 2006).

V besedilu navajamo citate po sistemu avtor-letnica. Priimek avtorja, letnico izida citiranega dela in stran zapišemo za citatom v obliki (priimek letnica, stran).

Npr. (Janžekovič 1976, 12). Na isti način navajamo tudi citate iz periodičnih publikacij in s spletnih strani. Če sta avtorja dva, navedek zapišemo v obliki (Priimek in Priimek letnica, stran). Npr. (Rode in Stres 1977, 33). Pri več kot treh avtorjih uporabimo obliko (Priimek idr. letnica, stran) ali (Priimek et al. letnica, stran). Npr. (Lenzenweger et al. 1999, 51). Če avtor dela ni naveden, namesto priimka uporabimo naslov, lahko tudi skrajšanega, npr. (*Devetdnevnic a k časti milostne Matere Božje 1916*, 5). Če priimek navedemo že v citatu, ga v navedku izpustimo in navedemo samo letnico in stran.

Npr.: »Kakor je zapisal Anton Strle (1988, 67) [...]». Če v istem oklepaju navajamo več del, ločujemo eno od drugega s podpičji. Npr. (Pascal 1986, 16; Frankl 1993, 73). Če citiramo v istem odstavku isto delo, navedemo referenco v celoti samo prvič, v nadaljevanju v oklepajih navedemo samo številko strani. Okrajšave »Prim.« ne uporabljamo. Po ustaljeni navadi v besedilu navajamo Sveto pismo, antične in srednjeveške avtorje, koncilske in druge cerkvene dokumente, Zakonik cerkvenega prava (ZCP), Katekizem katoliške Cerkve (KKC) in druge dokumente, ki jih obravnavamo kot vire. Natančnejša navodila za citiranje so na spletnem naslovu <http://www.teof.uni-lj.si/obrazci/citiranje.doc>.

Uporaba kratic znanih časopisov, periodičnih publikacij, najbolj citiranih del, leksikonov in zbirk mora ustrezati mednarodnemu seznamu (splošni seznam: List of serial title word abbreviations, COBISS.SI-ID: 61626368, dostopen v NUK-u; za teologijo: *Internationales Abkürzungsverzeichnis für Theologie und Grenzgebiete*, W. de Gruyter, Berlin 1992); kratice domačih del, zbirk in leksikonov se uporabljajo, če so splošno znane in priznane, na primer »BV« za *Bogoslovni vestnik*. Kratice revij/časopisov so v ležeči pisavi (kurzivi), kratice zbornikov, leksikonov ali monografij so v pokončni pisavi. Pravila glede prečrkovanja (transliteracije) iz grščine in hebrejščine so dostopna v uredništvu.

Objave v Bogoslovnem vestniku se ne honorirajo. Uredništvo zavrnenih rokopisov ne vrača.

Bogoslovni vestnik

Theological Quarterly

EPHEMERIDES THEOLOGICAE

Izdajatelj in založnik / Edited and published by

Naslov / Address

ISSN

Spletni naslov / E-address

Glavni in odgovorni urednik / Editor in chief

E-pošta / E-mail

Namestnik gl. urednika / Associate Editor

Uredniški svet / Editorial Council

Pomočniki gl. urednika (uredniški odbor) /

Editorial Board

Tehnični pomočniki gl. ur. / Tech. Assistant Editors

Lektoriranje / Language-editing

Prevodi / Translations

Oblikovanje / Cover design

Prelom / Computer typesetting

Tisk / Printed by

Za založbo / Chief publisher

Izvirne prispevke v tej reviji objavljajo /

Abstracts of this review are included in

Letna naročnina / Annual subscription

Transakcijski račun / Bank account

Teološka fakulteta Univerze v Ljubljani

Bogoslovni vestnik, Poljanska 4, 1000 Ljubljana

0006-5722

1581-2987 (e-oblika)

<http://www.teof.uni-lj.si/bv.html>

Robert Petkovšek

bogoslovni.vestnik@teof.uni-lj.si

Vojko Strahovnik

Metod Benedik, Erwin Dirscherl (Regensburg), Emmanuel Falque (Pariz), Vincent Holzer (Pariz), Janez Juhant, Jože Krašovec, Nenad Malović (Zagreb), Zorica Maros (Sarajevo), Mladen Parlov (Split), Vladislav Puzović (Beograd), Miran Sajovic (Rim), Walter Schaup (Gradec), József Stala (Krakov), Grzegorz Szamocki (Gdansk), Ed Udovic (Chicago), Michal Valčo (Žilina), Vladimir Vukašinović (Beograd), Karel Woschitz (Gradec)

Irena Avsenik Nabergoj (UL TEOF; ZRC SAZU), Roman Globokar (UL TEOF), Slavko Krajnc (UL TEOF), Simon Malmenvall (UL TEOF), Jožef Muhovič (UL ALUO), Mari Jože Osredkar (UL TEOF), Maria Carmela Palmisano (UL TEOF), Barbara Simonič (UL TEOF), Stanislav Slatinek (UL TEOF), Vojko Strahovnik, (UL TEOF; UL FF), Miha Šimac (UL TEOF), Miran Špelič (UL TEOF), Marjan Turnšek (UL TEOF), Janez odičar (UL TEOF)

Liza Primc, Jonas Miklavčič, Rok Gregorčič, Lucija Rožman in Aljaž Krajnc

Domen Krvina

Liza Primc

Lucijan Bratuš

Jernej Dolšek

KOTIS d. o. o., Grobelno

Tadej Stegu

Bayerische Staatsbibliothek, Digitale Bibliothek; Canon Law Abstracts; EBSCO; Elenchus Bibliographicus Biblicus; Web of Science (WoS) Core collection: the Emerging Sources Citation Index (ESCI); ERIH PLUS; dLib; IBZ Online; MIAR; MLA International Bibliography; Periodica de re Canonica; Religious & Theological Abstracts; Scopus (h) ; DOAJ

za Slovenijo: 28 EUR

za tujino: Evropa 40 EUR; ostalo 57 USD (navadno), 66 USD (prednostno);

naslov: Univerza v Ljubljani, Teološka fakulteta, Poljanska c. 4, 1000 Ljubljana

IBAN SI56 0110 0603 0707 798

Swift Code: BSLJIS2X

Revijo sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije.



Članki v reviji *Bogoslovni vestnik*, razen če ni drugače navedeno, so objavljeni pod pogoji licence Creative Commons Attribution 4.0 International (CC BY 4.0 International).

Impressum

Bogoslovni vestnik (*Theological Quarterly*, *Ephemerides Theologicae*) je znanstvena revija z recenzijo. V razvid medijev, ki ga vodi Ministrstvo za kulturo RS, je vpisana pod zaporedno številko 565. Izhaja štirikrat na leto. *Bogoslovni vestnik* je glasilo Teološke fakultete Univerze v Ljubljani. Vsebuje izvirne in pregledne znanstvene članke in prispevke s področja teologije in drugih ved, ki so sorodne teologiji. Objavlja v latinskem, angleškem, francoskem, nemškem in v italijanskem jeziku. Rokopis, ki ga sodelavec pošlje na naslov *Bogoslovnega vestnika*, je besedilo, ki istočasno ali prej ni bilo poslano na noben drug naslov in bo v *Bogoslovnem vestniku* objavljeno prvič. Prispevki v *Bogoslovnem vestniku* se ne honorirajo. Rokopise je treba poslati na elektronski naslov: bogoslovni.vestnik@teof.uni-lj.si.



MAGNA CHARTA FACULTATIS THEOLOGICAE LABACENSIS

Teološka fakulteta, ki ima korenine v srednjeveških samostanskih in katedralnih sholah in najodličnejše mesto na prvih univerzah, je tudi ena od petih ustanovnih fakultet Univerze v Ljubljani. Njeni predhodniki so bili jezuitski kolegij, ki je deloval v Ljubljani v letih od 1601 do 1773, vrsta drugih redovnih visokih šol in škofijske teološke šole. Kakor ob svojih začetkih želi Teološka fakulteta tudi na pragu 3. tisočletja ob nenehni skrbi za prvovrstno kvaliteto pedagoškega in raziskovalnega dela razvijati svoj govor o Bogu in o človeku, o božjem učlovečenju in o človekovem pobožanstvenju, o stvarstvu in o človekovem poseganju vanj, o začetkih vsega in o končnem smislu, o Cerkvi in o življenju v njej.

Opirajoč se na Sveto pismo in na izročilo in ob upoštevanju dometa človekove misli, usposablja svoje študente za jasen premislek o témah presežnosti in tukajšnjosti, vere in razodetja, kanonskega prava, morale in vzgoje. Kot katoliška fakulteta v zvestobi kulturnemu krogu, iz katerega je izšla, sledi avtoriteti cerkvenega učiteljstva in skrbi za intelektualno pripravo kandidatov na duhovništvo in za duhovno rast laičkih sodelavcev v Cerkvi in v širši družbi.

