

UDK 004.8:27-181.5

2Herzfeld, N. L.

27-426:004.8

<https://doi.org/10.53745/ro.1.2.4>

Primljeno: 29. 12. 2025.

Prihvaćeno: 28. 1. 2026.

Izvorni znanstveni rad

© 2025 Barščevski, CC BY NC 4.0

Čovjek kao *imago Dei* i umjetna inteligencija

The Human Being as *Imago Dei* and Artificial Intelligence

Taras Barščevski¹

Sažetak

Rad kritički promišlja suvremeni razvoj i primjenu umjetne inteligencije u svjetlu kršćanske antropologije, polazeći od biblijsko-teološkog pojma čovjeka kao *imago Dei*. Polazeći od teze Noreen Herzfeld o nastojanju da se umjetnom inteligencijom stvori „drugi“ na našu sliku, rad najprije prikazuje kako suvremeni sustavi UI-ja, osobito konverzacijski modeli, prelaze granice tehničke asistencije i ulaze u područje odnosa, povjerenja i moralnog savjetovanja. U prvom poglavlju analizira se normativni i antropološki okvir ključnih etičkih smjernica za „pouzdanu UI“, uključujući crkvene dokumente, pri čemu se pokazuje da naglašeni antropocentrizam i redukcija čovjeka na imanentnu autonomiju otežavaju dosljednu primjenu načela i ostavljaju nerazriješenim pitanje izvora i mjerila dobra i zla. U drugom poglavlju pojam *imago Dei* razmatra se kroz supstancijalne, funkcionalne i relacijske interpretacije te se pokazuje kako suvremeni pristupi UI-ju analogno osciliraju između razuma, funkcionalnosti i odnosa. Treće poglavlje, polazeći od razumijevanja Post 3,5 i Post 3,22-24, tumači dinamiku „poznavanja dobra i zla“ kao pokušaj suobličjenja Bogu, što istodobno razotkriva čovjekovu ograničenost, poremećaj temeljnih odnosa prema Bogu, bližnjemu i stvorenju te sklonost uspostavi „četvrtog odnosa“ u kojem bi čovjek nastojao postati „bog“ vlastitim tehnološkim tvorevinama. Zaključni prijedlog rada jest da pouzdan odnos prema umjetnoj inteligenciji pretpostavlja obnovu čovjeka kao *imago Dei*,

¹ Sveučilište u Zagrebu Katolički bogoslovni fakultet

koji, po uzoru na Boga koji „počiva od svega djela svoga koji učini“, zna stati i usmjeriti svoju stvaralačku moć prema blagoslovu i posvećenju.

Ključne riječi: umjetna inteligencija, *imago Dei*, etičke smjernice, antropocentrizam, Knjiga postanka 1 – 3, sloboda, relacijska antropologija, ChatGPT

Abstract

This paper offers a critical reflection on the contemporary development and use of artificial intelligence from the perspective of Christian anthropology, grounded in the biblical and theological understanding of the human person as *imago Dei*. Taking as its point of departure Noreen Herzfeld's thesis that artificial intelligence seeks to create an "other" in our own image, the paper first shows how contemporary AI systems, especially conversational models, extend beyond mere technical assistance and move into the realm of relationships, trust, and moral guidance. The first chapter examines the normative and anthropological framework of the principal ethical guidelines for "trustworthy AI", including relevant Church documents, showing that an accentuated anthropocentrism and the reduction of the human being to immanent autonomy complicate the coherent application of these principles and leave unanswered the question concerning the origin and standard of good and evil. In the second chapter, the notion of *imago Dei* is examined through substantial, functional, and relational interpretations, showing how contemporary approaches to artificial intelligence fluctuate between reason, functionality, and relationality. The third chapter, drawing upon the interpretation of Gen 3:5 and Gen 3:22–24, interprets the dynamic of the "knowledge of good and evil" as an attempt at conformity to God that simultaneously reveals human limitation, the disturbance of fundamental relationships (with God, with others, and with creation), and the inclination to establish a "fourth relationship" in which the human being seeks to become a "god" through technological artefacts of his own making. The paper concludes by proposing that a trustworthy relationship with artificial intelligence presupposes the renewal of the human person as *imago Dei*, who, following the example of God who "rested from all the work that he had done", is capable of pausing and directing his creative power towards blessing and sanctification.

Keywords: artificial intelligence, *imago Dei*, ethical guidelines, anthropocentrism, Book of Genesis 1–3, freedom, relational anthropology, ChatGPT.

Uvod

Prema Noreen Herzfeld jedan od ciljeva umjetne inteligencije je stvaranje „drugoga“ na našu sliku (Herzfeld 2002b, 304). Na prvi pogled takva tvrdnja može djelovati preuveličano ili barem zahtjevati dodatno pojašnjenje. Ipak, razvoj znanosti i tehnologije u posljednjim desetljećima pokazuje da se upravo u tom smjeru ostvaruje iznimno snažan i ubrzan napredak. Međutim, koncept humanoidnih automata nije proizvod isključivo suvremenog tehnološkog doba. Nalazimo ga već u grčkim mitovima, primjerice

u pripovijesti o brončanom divu Talu (Talos), čuvaru Krete i izvršitelju Zeusove volje (Mayor 2018, 7), ili o Pandori, koju je Hefest načinio od gline i dao joj oblik žene, nakon što je Zeus, rezgnjevljen Prometejevim darivanjem vatre ljudima, odlučio kazniti čovječanstvo i naredio bogovima da načine ženu kao dar ljudima (Mayor 2018, 156). „Svime obdarena“ – što je doslovni prijevod imena Pandora – ona je ipak dio Zeusova plana za kaznu čovječanstvu, a njezina radoznalost, simbolizirana otvaranjem kutije ili posude, oslobađa sva zla. Slične predodžbe nalazimo i u religijskim i filozofskim tekstovima iz Kine, primjerice u daoističkom spisu Liezi, klasičnom filozofskom tekstu koji se tradicijski pripisuje Lie Yukouu (5. st. pr. Kr.), ali je u svojem konačnom obliku vjerojatno nastao u 3. ili 4. st. po Kr., te donosi pripovijest o umjetnom čovjeku sposobnom govoriti, pjevati i izvoditi ljudske pokrete (Mayor 2018, 121).

Prošla su mnoga stoljeća od tih ranih mitoloških i filozofskih refleksija do pojave prvih tehničkih pokušaja stvaranja humanoidnih automata. Kao važna prekretnica često se navodi da Vincijev „robotski vitez“ iz 1495. godine, kojeg mnogi smatraju prototipom humanoidnog robota (Pasek 2014, 4-5). Te se ideje u novije vrijeme ostvaruju u sve sofisticiranijim oblicima androidnih robota, koji ne samo da oponašaju ljudsku anatomiju i pokrete nego i posjeduju izrazito realistične crte lica i tijela. Primjerice, Geminoid HI-1 Hiroshija Ishigura iz 2006. godine osmišljen je upravo s ciljem postizanja što veće sličnosti između robota i njegova ljudskog „originala“ (Nishio, Ishiguro i Hagita 2007, 346-347), dok se daljnjim razvojem te linije danas stiglo do modela Geminoid HI-6, koji integrira naprednije sustave umjetne inteligencije i omogućuje znatno autonomniju komunikaciju i interakciju s ljudima (Intelligent Robotics Laboratory 2018). Još izraženiji iskorak predstavljaju tzv. socijalni roboti koji su dizajnirani ponajprije za interakciju s ljudima. Među njima se osobito ističe robot Sophia, koji ne samo da pokazuje sposobnost realistične mimike i komunikacije nego je 2017. godine stekao i državljanstvo Saudijske Arabije (Kouravanas i Pavlopoulos 2022, 141-142).

U tom se kontekstu nužno nameće i pitanje uloge sustava umjetne inteligencije poput ChatGPT-a. Iako on ne imitira čovjeka u njegovu tjelesnom izgledu, sve se češće predstavlja – ili doživljava – kao ravnopravan sugovornik: netko s kim se može raspravljati, dijeliti misli i osjećaje, tko može „razumjeti“ i „utješiti“, davati savjete u međuljudskim odnosima, pa čak postati sugovornik od povjerenja i intimni pratitelj kojem se povjeravaju najdublje misli i osjećaji. Upravo tu dolazi do kvalitativnog pomaka: umjetna inteligencija više ne ostaje ograničena na područje teh-

ničke asistencije, nego ulazi u sferu interpersonalnih odnosa, koji su se tradicionalno smatrali konstitutivno ljudskima.

Razvoj robotike dodatno se nastavlja u smjeru biohibridnih robota, koji povezuju biološke i mehaničke elemente, kao i autonomnih sustava sposobnih za samostalno funkcioniranje i donošenje odluka bez neposrednog ljudskog nadzora. Iako još ne postoje biohibridni roboti koji bi u potpunosti nalikovali čovjeku, a ni autonomni humanoidni roboti ne posjeduju svijest u ljudskom smislu, važno je naglasiti da se njihova unutar-nja funkcionalnost temelji na algoritmima, senzorima i programiranju, a ne na biološkim procesima i svijesti inherentnima ljudskoj osobi. Unatoč tomu, takvi sustavi sve više utjelovljuju želje, vrijednosti i ciljeve svojih dizajnera i konstruktora.

Današnji roboti i sustavi umjetne inteligencije sve više nalikuju čovjeku, ponajprije svojom funkcionalnošću. Oni preuzimaju ljudske poslove, zamjenjuju čovjeka na radnim mjestima te ulaze u područja koja su nekoć bila isključivo ljudska. Opremljeni umjetnom inteligencijom oni mogu „razmišljati“, „donositi odluke“, čak i „osjećati“ – ili barem vrlo uvjerljivo imitirati ljudsku inteligenciju, emocije i slobodne odluke – za nas, ili umjesto nas, ili čak bez nas. Bez obzira radi li se o stvarnim sposobnostima ili o sofisticiranoj simulaciji, posljedice su neosporne: one već sada duboko oblikuju život čovjeka i snažno uvjetuju njegovu budućnost.

Cilj ovoga rada jest kritički promisliti suvremeni razvoj i primjenu umjetne inteligencije u svjetlu kršćanske antropologije, polazeći od biblijsko-teološkog pojma čovjeka kao *imago Dei*. Nastanak i razvoj umjetne inteligencije pratile su i prate etičke smjernice za pouzdanu umjetnu inteligenciju, kojima se nastoji usmjeriti razvoj i spriječiti moguće zloporabe UI-ja. U temelju tih etičkih smjernica nalazi se naglašen antropocentizam iza čega se često krije koncepcija čovjeka koja zanemaruje njegovu transcendentalnu dimenziju bez koje je teško, ako ne i nemoguće, uspostaviti ispravan odnos prema tehničkom razvoju, a osobito prema umjetnoj inteligenciji.

U prvom poglavlju ukratko se analizira normativni i antropološki okvir etičkih smjernica za pouzdanu umjetnu inteligenciju s ciljem razotkrivanja implicitne slike čovjeka koja stoji u njihovoj pozadini. U drugom poglavlju polazi se od biblijsko-teološkog pojma *imago Dei* te se, kroz njegove supstancijalne, funkcionalne i relacijske interpretacije ispituju antropološke pretpostavke i granice suvremenih pristupa razvoju umjetne inteligencije. Na kraju se predlože biblijska slika čovjeka u njegovu odnosu prema Bogu, drugome i svemu stvorenomu, u čijem se kontekstu promatra i odnos čovjeka prema umjetnoj inteligenciji kao vlastitoj kreaciji.

1. Etičke smjernice

Zbog dobrobiti za čovječanstvo koja proizlazi iz razvoja umjetne inteligencije, ali i zbog opasnosti koju ona predstavlja, svjedoci smo brojnih inicijativa na različitim razinama usmjerenih na razvoj etičkih smjernica za umjetnu inteligenciju. Od crkvenih intervencija na najvišoj razini, osim dokumenta *Rome Call for AI Ethics*, koji je pod inicijativom Vatikana okupio predstavnike triju abrahamskih religija, kao i predstavnike velikih kompanija u IT industriji kao što su Microsoft i IBM te predstavnika FAO-a (*Ai Ethics: An Abrahamic Commitment to the Rome Call 2023*; *Rome Call for AI Ethics 2020*; *Etica dell'intelligenza artificiale: l'impegno delle religioni abramitiche nella Rome Call 2023*; *AI Ethics: An Abrahamic Commitment to the Rome Call. Joint Declaration 2023*; Wagle 2023; Weisz 2023), vrijedi spomenuti i više papških intervencija: pape Ivana Paval II., pape Benedikta XVI. te pape Franje (Tridente 2022, 2.1. L'insegnamento dei Pontefici), kao i enciklike *Laudato si'* (Franjo LS 2015, br. 102-107) i *Fratelli tutti* (Franjo FT 2020, br. 19, 33, 106), koje upozoravaju na rizike tehnokratske paradigme i potrebe etičkog upravljanja tehnološkom moći. Posebice je važan dokument COMECE pod nazivom „Robotizacija života: etika u svjetlu novih izazova“ (CEMECE 2019).

U tom kontekstu osobito je važno spomenuti i najnoviji dokument Dikasterija za nauk vjere i Dikasterija za kulturu i odgoj od 28. siječnja 2025. godine naslovljen *Antiqua et nova. Nota sul rapporto tra intelligenza artificiale e intelligenza umana* (Dicastero per la Dottrina della Fede i Dicastero per la Cultura e l'Educazione 2025). Iako je formalno kategoriziran kao „nota“, riječ je o autoritativnom dokumentu redovitog učiteljstva Crkve, s izričitim odobrenjem Svetoga Oca, koji se izravno bavi suvremenim izazovima i prilikama što ih pred čovječanstvo postavlja razvoj znanosti i tehnologije, a napose umjetne inteligencije. Polazeći od biblijsko-teološkog pojma čovjeka kao stvorenoga „na sliku Božju“ (Post 1,27), dokument naglašava potrebu cjelovite vizije ljudske osobe i poziv da čovjek „obrađuje“ i „čuva“ stvoreni svijet (Post 2,15) te predlaže pristup suvremenim tehnološkim izazovima „sa starom i novom mudrošću“ (*antiqua et nova sapientia*). Iako opći referentni okvir dokumenta *Antiqua et nova* jest kršćanska antropologija i viđenje tehničkog napretka u svjetlu biblijske objave, u razradi etičkih smjernica u velikoj se mjeri oslanja na opća etička načela. To, primjerice, pokazuje iznenadujuća sličnost između nekih etičkih načela predloženih u tom dokumentu i preporuka o umjetnoj inteligenciji koje su posljednjih godina oblikovale važne međunarodne institucije (Tanzella-Nitti 2025).

Pitanje etičkih smjernica umjetne inteligencije uvelike nadilazi crkveni kontekst te se razmatra u brojnim međunarodnim i institucionalnim dokumentima. Tako je Europska komisija 2019. godine s ciljem definiranja etičkih smjernica za umjetnu inteligenciju objavila dva ključna dokumenta: *Etičke smjernice za pouzdanu umjetnu inteligenciju* (AI HLEG 2019b) i *Definicija umjetne inteligencije: glavne mogućnosti i znanstvene discipline* (AI HLEG 2019a). Ti se dokumenti mogu sažeti u četiri temeljna etička načela – poštovanje ljudske autonomije, sprečavanje nastanka štete, pravednost i objašnjivost – koja su određena kao etički imperativi koji se moraju poštovati kako bi se osiguralo da se sustavi umjetne inteligencije razvijaju, uvode i upotrebljavaju na pouzdan način (Barščevski 2023, 38).

Uz dokumente Europske komisije valja spomenuti i niz drugih inicijativa nastalih u akademskom i istraživačkom kontekstu, poput *Asilomar AI Principles* (Future of Life Institute 2017) i *Montréal Declaration for Responsible Development of Artificial Intelligence* (Montréal 2018), kao i dokumente u izradi kojih su sudjelovali predstavnici tehnoloških kompanija ili koje su usvojile velike IT kompanije (International Telecommunication Union 2023; Google 2018; Microsoft 2021; Board 2023).

Opća etička načela u razvoju i primjeni UI-ja koja dokumenti Europske komisije navode proizlaze iz temeljnih ljudskih prava: poštovanja ljudskog dostojanstva, osobne slobode, demokracije, pravde, vladavine prava, ravnopravnosti, nediskriminacije, solidarnosti i prava građana (AI HLEG 2019b, br. 40-45). Možemo ih sažeti u četiri etička načela, određena kao etički imperativi, koja se moraju poštovati kako bi se osiguralo da se sustavi umjetne inteligencije razvijaju, uvode i upotrebljavaju na pouzdan način. To su načela poštovanja ljudske autonomije, sprečavanja nastanka štete te pravednosti i objašnjivosti (AI HLEG 2019b, br. 47-48): 1) Načelo poštovanja ljudske autonomije usmjereno je na osiguravanje poštovanja slobode i autonomije ljudskih bića (AI HLEG 2019b, br. 50). 2) Načelo sprečavanja nastanka štete podrazumijeva zaštitu ljudskog dostojanstva te duhovnog i tjelesnog integriteta prema kojem „sustavi umjetne inteligencije ne bi trebali uzrokovati ni pogoršati štetu ili na drugi način negativno utjecati na ljude“ (AI HLEG 2019b, br. 51). 3) Načelo pravednosti ima materijalnu i proceduralnu dimenziju. Materijalna dimenzija uključuje „osiguravanje ravnopravne i pravedne raspodjele i koristi i troškova te osigurava da pojedinci i skupine ne budu pod utjecajem nepravedne pristranosti, diskriminacije i stigmatizacije“. Proceduralna dimenzija uključuje mogućnost osporavanja i traženja učinkovite pravne zaštite protiv odluka koje donesu sustavi UI-ja i ljudi koji njima uprav-

ljaju, što zahtijeva i mogućnost identificiranja subjekata odgovornih za odluku, dok bi trebalo moći objasniti i sam proces odlučivanja (AI HLEG 2019b, br. 52). 4) Načelo objašnjivosti, prema kojem procesi trebaju biti transparentni, određuje da se „mogućnosti i svrha sustava UI-ja trebaju otvoreno priopćiti, a odluke se moraju objasniti, u mjeri u kojoj je to moguće, osobama na koje one utječu izravno i neizravno“ (AI HLEG 2019b, br. 53).

Konkretizirajući ta načela dokument navodi sedam zahtjeva koje bi sustavi umjetne inteligencije trebali ispuniti kako bismo ih mogli smatrati pouzdanima (AI HLEG 2019b, br. 55-91): 1) ljudsko djelovanje i nadzor nad njima, uključujući temeljna prava, ljudsko djelovanje i ljudski nadzor; 2) tehnička otpornost i sigurnost, uključujući otpornost na napade i sigurnost, rezervni plan i opću sigurnost, točnost, pouzdanost i ponovljivost; 3) privatnost i upravljanje podacima, uključujući poštovanje privatnosti, kvalitetu i cjelovitost podataka i pristupa podacima; 4) transparentnost, uključujući sljedivost, objašnjivost i komunikaciju; 5) raznolikost, nediskriminacija i pravednost, uključujući izbjegavanje nepravedne pristranosti, dostupnost i univerzalni dizajn i sudjelovanje dionika; 6) dobrobit za okoliš i društvo, uključujući održivost i pogodnost za okoliš, društveni učinak, društvo i demokraciju; 7) odgovornost, uključujući povjerljivost, smanjenje negativnih učinaka i izvještavanje o njima, kompromise i pravnu zaštitu.

Ta načela i zahtjevi, utemeljeni na temeljnim ljudskim pravima, ne doimaju se samo u potpunosti racionalnima i općeprihvatljivima, nego se nužno nameće pitanje njihove stvarne provedivosti. Jesu li oni u praksi u potpunosti ostvarivi i postoje li učinkoviti mehanizmi zaštite kako od zlonamjernih kršenja, tako i od nenamjernih pogrešaka? Osobito ostaje otvoreno pitanje postoji li mogućnost ispravljanja već nanesene štete te u kojoj je mjeri takva korekcija uopće izvediva. Uvođenje pravnih okvira, propisa i zakona, kao i različitih ograničenja pa čak i sankcija za prekršitelje, samo po sebi ne daje zadovoljavajući odgovor na problem mogućih negativnih posljedica koje mogu proizići iz zlonamjerne, neodgovorne ili nekompetentne uporabe umjetne inteligencije.

Unatoč velikoj važnosti i neupitnoj vrijednosti etičkih smjernica za razvoj i primjenu umjetne inteligencije, njihova se ograničenja sve jasnije očituju u praksi. Ta ograničenja ne proizlaze ponajprije iz nedostatka normativnih načela, nego iz antropoloških pretpostavaka na kojima su utemeljene. Naglašeni antropocentrizam, koji etičke smjernice postavlja isključivo u imanentni horizont ljudske autonomije, otežava njihovu do-

sljednu primjenu, osobito u kontekstu sustava umjetne inteligencije, čije je djelovanje obilježeno neprozirnošću, prilagodljivošću i nepredvidljivošću. Time čovjek ne nastupa više kao odgovorno stvorenje koje etičke norme prima i razlučuje u odnosu prema Bogu i stvorenom poretku, nego kao suveren zakonodavac koji sam sebi određuje mjerila dobra i zla, ne priznajući nikakvu instanciju izvan vlastita razuma i volje (Ravasi 1995, 49; Barščevski 2023, 47-54). Takvo samopostavljanje čovjeka u središte normativnog poretka dovodi do implicitnog raskida s biblijskim razumijevanjem čovjeka kao *imago Dei*, čija sloboda nije apsolutna, nego relacijska i odgovorna. Bez uvažavanja te relacijske i transcendentalne dimenzije ljudske egzistencije etičke smjernice ostaju nužne kao korektiv tehnološkog razvoja, ali nedostatne kao temelj pouzdanog odnosa prema umjetnoj inteligenciji. Na sličan problem upozorava i Odilon-Gbènoukpo Singbo, koji razvoj umjetne inteligencije promatra u širem kontekstu transhumanističkih nastojanja da se tehnološkim sredstvima prevladaju granice ljudske naravi. Prema njemu, suvremeni etički diskurs o umjetnoj inteligenciji često polazi od implicitne antropologije u kojoj se čovjek razumije kao samodostatan projekt, dok se dimenzija stvorenosti i odnosa prema Bogu potiskuje u drugi plan (Singbo 2024, 56-58). Time se etičke smjernice svode na pokušaj normativne kontrole tehničke moći bez dubljeg antropološkog i teološkog utemeljenja koje bi moglo osigurati njihovu dosljednu i dugoročnu primjenjivost.

Upravo ta napetost zahtijeva dublju antropološku refleksiju, koja će u nastavku biti razradena polazeći od biblijsko-teološkog pojma čovjeka kao *imago Dei*.

2. Čovjek kao *imago Dei*

Ograničenja etičkih smjernica za umjetnu inteligenciju upućuju na potrebu dublje i sustavnije antropološke refleksije, što potvrđuju i teološko-pastoralna promišljanja o umjetnoj inteligenciji koja upozoravaju da tehnološki razvoj postaje problematičan onda kada tehnička učinkovitost zanemaruje osobu i njezino dostojanstva, što za čovjeka može imati „tragične posljedice“ (Šegula 2024, 900). Zato se u nastavku rada polazi od biblijsko-teološkog pojma *imago Dei* jer se upravo u tom pojmu prelamaju temeljna pitanja o tome tko je čovjek, po čemu je jedinstven i kako se njegova sloboda i odgovornost trebaju razumjeti u odnosu prema Bogu, drugome i svijetu. Budući da svaka interpretacija slike Božje u čovjeku ostaje djelomična, ključno je razjasniti što je to u čovjeku što se (svjesno ili

nesvjesno) projicira na tehničke sustave, odnosno što bi računala morala pokazati da bi se mogla smatrati „djecom našega uma“ (*mind children*) (Herzfeld 2002b, 304; Moravec 2000, 125-126). Vrijedi spomenuti i to da se danas rasprave o umjetnoj inteligenciji često povezuju s transhumanizmom u osnovi kojega stoji vjera u mogućnost da ljudska vrsta transcendirira granice svojeg biološkog stanja pomoću tehnologije. Takva „vjera“ u neizbježan napredak čovjeka – prema riječima Singboa – nova je vrsta religije (Singbo 2024, 54), čime izražava svoje nastojanje ne samo „postati kao bog“ nego i stvoriti boga na svoju sliku.

Pitanje „što mi ljudi imamo zajedničko“ s takvim „drugim“ kršćanski su teolozi stoljećima promišljali upravo kroz koncept slike Božje u čovjeku. Otuda proizlazi i pitanje koje je za ovaj rad odlučujuće: Je li slika čovjeka, koju dijeli s Bogom, na neki način povezana sa slikom koju nastojimo projicirati na vlastitu tehnološku kreaciju umjetne inteligencije? U ovom se radu polazi od tvrdnje da jest – i da izbor onoga što projiciramo u umjetnu inteligenciju ima posljedice: posljedice za naše vlastito samorazumijevanje, ali i za buduću koegzistenciju s onime što proizvodimo.

Povijesno gledano, interpretacije *imago Dei* mijenjale su se, no većinu njih moguće je sistematizirati u tri dominantna usmjerenja: 1) supstancijalne interpretacije, koje *imago Dei* prepoznaju kao svojstvo pojedinca, bitan dio naravi (najčešće razum); 2) funkcionalne interpretacije, u kojima se *imago Dei* prepoznaje u djelovanju, osobito u čovjekovu pozivu da „vlada“ zemljom i da je „obrađuje i čuva“; 3) relacijske interpretacije, u kojima se slika Božja razumije polazeći od odnosa koje čovjek uspostavlja prema Bogu i prema drugome čovjeku. Zanimljivo je da se i pristupi razvoju umjetne inteligencije, u određenoj analogiji, kreću sličnim trajektorijama: od inteligencije shvaćene kao individualna sposobnost, preko inteligencije shvaćene kao funkcionalna uspješnost, prema inteligenciji koja se procjenjuje unutar odnosa i komunikacije.

2.1. „Mislim, dakle jesam“ – razum kao *imago Dei*

Na prvi pogled prepoznati *imago Dei* čovjeka u svojstvima razuma ili racionalnosti izgleda kao tumačenje Božje slike koja bi najviše odgovarala ideji umjetne inteligencije. Paul Ramsey zove je supstancijalističkom interpretacijom *imago Dei* – razumijevanje u kojem se Božja slika čini kao „nešto unutar supstancijalnog oblika ljudske naravi, neka sposobnost ili mogućnost koju posjeduje čovjek, a služila bi za razlikovanje čovjeka od prirode i drugih životinja“ (Ramsey 1993, 250). David Cairns u svojoj

studiji naslovljenoj *Slika Božja u čovjeku* tvrdi da kršćanski mislioci sve do sv. Tome, uključujući i njega, djelomično ovise od predkršćanske misli, shvaćajući Božju sliku u čovjeku kao razum (Cairns 1953, 58). Iako takva tvrdnja izgleda pretjerano, ipak moramo priznati da su mnogi rani kršćanski pisci promatrali *imago Dei* u terminima razuma ili racionalne inteligencije, tako da su je vidjeli kao „najviše božansku“ od svih ljudskih sposobnosti, kao upravu onu koja razlikuje *animal rationale* (Aristotel) od svih drugih vrsta (Herzfeld 2002a, 11). Među suvremenim teolozima supstancijalističku viziju čovjeka zastupa Reinhold Niebuhr. On vidi *imago Dei* u široko definiranom razumu kao razumu koji obuhvaća racionalnost, slobodnu volju i sposobnost uzdići se iznad sebe, koju definira kao samotranscendenciju (Niebuhr 1996, 161-162).

U području UI, supstancijalističkoj interpretaciji paralelan je pristup koji se najčešće naziva klasičnom ili simboličkom umjetnom inteligencijom. On se temelji na hipotezi da „fizički sustav simbola ima potrebna i dostatna sredstva za djelovanje zajedničke inteligencije“, gdje je sustav fizičkih simbola onaj u kojem se bilo koji proces na računalo može označiti nekim izrazom sastavljenim od fizičkih simbola, a svaki se takav izraz može mijenjati prema pravilima koja se mogu logički definirati (Herzfeld 2002a, 13). Drugim riječima, bilo koji problem koji se može prikazati kao formalna shema može predstavljati inteligenciju i računalo ga može rješavati na „inteligentan“ način.

Simbolička umjetna inteligencija postigla je velik uspjeh u domenama gdje su problemi jasno definirani i gdje je moguće izgraditi bazu eksplicitnih znanja zajedno s pravilima pretraživanja, zaključivanja, pohranjivanja i dohvaćanja znanja te umetanja novih pojmova u bazu (Putica 2018, 204). Primjeri su neke igre poput šaha ili određeni ekspertni sustavi u medicini, kemiji i srodnim područjima. No simbolički pristup pokazao se ograničenim ne toliko pred zahtjevnim problemima kao što je npr. izračunavanje neke složene jednadžbe, nego pred veoma jednostavnim stvarima koje može učiniti dvogodišnje dijete kao npr. prepoznati lice u različitim kontekstima ili razumjeti neku jednostavnu priču.

2.2. „*Biti znači djelovati*“ (*agere sequitur esse*) – funkcija/uloga kao *imago Dei*

Johannes Hehn je još davne 1915. godine sugerirao da Božju sliku treba razumjeti kao kraljevsku oznaku ili titulu, a ne kao atribut ljudske naravi (Hehn 1915, 36-52). Gerhard von Rad razvio je Hehnovo djelo u

funkcionalan, dinamičan pristup slici Božjoj, pristup koji sliku stavlja ne toliko u ono što jesmo, koliko u ono što smo pozvani činiti. On to argumentira time što prevodi s hebrejskog *b'ešelem 'elōhim* (Post 1,27) doslovno „kao sliku Božju“, a ne uobičajeno „na sliku Božju“, implicirajući time da je cijela osoba, a ne samo neka njezina osobina, na sliku Božju (von Rad 1973, 58). Iako bi prijevod *šelem* kao slika, replika ili idol mogao implicirati fizičku interpretaciju *imago Dei*, von Rad se njime koristi u podršku funkcionalne interpretacije. Čovjek je postavljen na zemlji na sliku Božju kao simbol Božje suverenosti. On je samo Božji predstavnik, pozvan očuvati i učvrstiti Božje vlasničko pravo nad zemljom (von Rad 1973, 60). Takvo razumijevanje *imago Dei* u terminima funkcije čovjeka vladati i upravljati zemljom djelujući kao Božji predstavnik na zemlji bilo je dosta zastupljeno među egzegetima Staroga zavjeta tijekom 20. stoljeća (Herzfeld 2002a, 16).

Sličan pomak dogodio se 80-ih godina u sferi umjetne inteligencije. Uočen je nedostatak napretka u razvoju zajedničke inteligencije korištenjem metoda simboličkog programiranja. To, međutim, ne znači da UI nije dao nikakve rezultate. Ako promatramo računalo u funkcionalnom smislu, tj. u njegovoj sposobnosti da djeluje kao predstavnik čovjeka, obavljajući zadatke koje su prije obavljali ljudi, možemo konstatirati određen uspjeh. Umjesto da pokušava reproducirati proces ljudskog razmišljanja funkcionalna umjetna inteligencija temelji se na prednostima svojstava računalne tehnologije i mjeri svoj uspjeh u praktičnim terminima. Ako promatramo UI u uskom smislu kao primjenu određenih tehnika programirana na rješavanje praktičnih problema, možemo reći da je dosegla uspjeh. Računalo ne mora pokazati cijelu gamu ljudskih sposobnosti, nego samo treba reproducirati ljudsku funkcionalnost u ograničenoj domeni. Kada Hans Moravec kaže da će „kompjuteri zamijeniti ljude u bitnim ulogama“, on na problem gleda iz funkcionalne perspektive. I doista, računala su već zamijenila ljude u mnogim sferama, iako ih možda više ne smatramo „bitnim“ ako ih može obavljati stroj. Velik dio toga što obično etiketiramo kao istraživanja umjetne inteligencije slijedi funkcionalni model, primjenjujući određene tehnike programiranja na rješavanje praktičnih problema. UI tako stvara strojeve koji preuzimaju ljudske zadatke. U tom smislu nije ni potrebno identificirati umjetnu inteligenciju kao inteligenciju: sve što strojevi moraju činiti je izvršiti neke iste zadatke koje izvršavaju ljudi.

Dva su problema povezana s definiranjem umjetne inteligencije u funkcionalnom smislu. Prvi je problem poteškoća u određivanju što ulazi u ka-

tegoriju umjetne inteligencije, a što je jednostavna računalna aplikacija. Funkcionalna definicija koja uključuje bilo koji program koji obavlja neke od funkcija koje ljudi inače obavljaju obuhvatila bi gotovo sve računalne aplikacije, ali bilo bi smiješno smatrati umjetno inteligentnim bilo koji program (Herzfeld 2002a, 17). Drugo, da bi funkcionalni pristup doveo do inteligencije koja bi bila slična ljudskoj, bilo bi potrebno ne samo specificirati koji pojmovi čine inteligenciju, nego se i uvjeriti da su sve te funkcije na odgovarajući način kompatibilne jedna s drugom (Herzfeld 2002a, 18). Stvaranje zajedničke inteligencije koja bi bila slična ljudskoj, bez obzira koliko njezino ostvarivanje ostaje teškim zadatkom, i dalje je san mnogih. Prema Johnu Haugelandu UI „samo želi pravi proizvod/artikl: strojeve s umovima u punom i doslovnom smislu tih riječi“ (Haugeland 1989, 2). Stoga, funkcionalni pristup, iako donosi dobre rezultate, kako u terminima financijske dobiti tako i poslovanja, ne opravdava nade i snove mnogih, što samih programera umjetne inteligencije što i šire javnosti.

2.3. *Ja jesam jer ti jesi/postojiš: odnos kao imago Dei*

Prema Karlu Barthu slika Božja „nije u nečemu što čovjek jeste ili čini“, nego je prepoznata u činjenici da je ljudsko biće „kopija/pandan Boga“, „counterpart of God“ (Barth 1958a, 185). Kao i funkcionalisti Barth argumentira na temelju tekstualne ekzegeze Post 1,26-27, no usredotočuje se ipak na dvama različitim odlomcima teksta: „Načinimo čovjeka na svoju sliku“ (Post 1,26) i „muško i žensko stvori ih“ (Post 1,27). Interpretira množinu „načinimo“ ne u odnosu na neki nebeski sud, već na narav samoga Boga, trojstvo koje uključuje i „ja“ koje može uputiti božanski poziv i „ti“ sposoban za božanski odgovor (Barth 1958b, 182). To suočavanje ja – ti unutar božanstva temelj je stvaranja čovjeka, tako da je naša vlastita ljudska narav ukorijenjena u odnosu s drugim. Taj odnos može imati dva oblika: odnos između čovjeka i Boga i odnos između čovjeka i čovjeka. Za Bartha, dakle, *imago Dei* postoji primarno u našem odnosu prema Bogu, a sekundarno u našem odnosu jednoga prema drugome. Iako se mnogi ne slažu s Barthom glede detalja što čini autentičnim jedan odnos ili je li razlikovanje muško – žensko u Post 1,26 adekvatan model za sve ljudske odnose, relacijski model *imago Dei* postao je dominantan u teologiji druge polovice i kraja 20. stoljeća. U svojoj kritici Bartha Gerrit Berkouwer ističe da je Isus postao kao mi više nego što smo mi postali kao on. Osim toga definira odnos koji konstituira *imago Dei* tvrdeći da analogija između ljudi i Boga nije ni analogija *entis* ni analogija *relationis*, nego analogija *amoris*.

Relacijski model postoji također i u umjetnoj inteligenciji. Kao što smo već imali priliku vidjeti, funkcionalna definicija inteligencije kao sposobnosti obavljanja zadataka ili niza zadataka je problematična jer je teško odrediti koji zadatci demonstriraju inteligenciju ili, u širem smislu, koji pružaju uvjerljiviju sliku nas samih (Herzfeld 2002a, 20). Tu je potrebu još prije pojava računala prepoznao engleski matematičar Alan Turing. U svom radu pod naslovom *Računalna tehnologija i inteligencija* iz 1950. godine pokreće pitanje postupaka koji bi bili bitni za pravi *imago Dei*, predlažući test koji omogućava provjeru inteligencije računala ili programske potpore. U testu ljudski ispitivač komunicira s igračima A i B tako što s njima razmjenjuje pisane poruke (u prozoru za razgovor, odnosno *chat*). Ako ispitivač ne može utvrditi koji je igrač (A ili B) računalo, a koji čovjek, smatra se da je računalo prošlo test. Glavni je argument da je računalo očito dostiglo razinu ljudske inteligencije ako ga tijekom razgovora na općem prirodnom jeziku nije moguće razlikovati od čovjeka. Neki kritičari Turingova testa kao testa inteligencije smatraju da se njime zapravo više utvrđuje ponaša li se računalo kao čovjek nego je li računalo inteligentno jer su test „prolazili“ računalni programi koji su neprestano mijenjali temu, radili brojne pravopisne pogreške, a katkad i odbijali odgovoriti na pitanje (University of Helsinki 2020).

Turingov test koristi relacionalnost kako bi odredio inteligenciju. Ipak Turing obraća pažnju i na važnost odnosa u stjecanju znanja ili inteligencije. On predlaže da, umjesto da pokušavamo oponašati um odrasle osobe, izgradimo um koji oponaša um djeteta, koji bi bio sposoban učiti i razviti se u odrasli um. To bi značilo da u istraživanju UI treba uzeti u obzir i socijalne aspekte znanja i djelovanja (Gasser 1991, 107-108). Ljudska inteligencija može biti nešto zajedničko, a ne strogo individualna osobina. Prema Johnu Haugelandu veliki dio onoga što smatramo ljudskom inteligencijom ne posjeduje pojedinac (Haugeland 1989, 26). Projekti, planovi i ponašanje proizlaze iz međusobnih odnosa, tako da, prema Haugelandu i Castelfranchi individualna inteligencija bila bi bez smisla: inteligencija ima smisao samo u susretu (Usp. Herzfeld 2002a, 23).

3. Poznavanje dobra i zla – prisvojena moć i povjerena zadaća čovjeka kao *imago Dei*

3.1. „Kao bogovi koji razlučuju dobro i zlo“ (Post 3,5)

U *Etičkim smjernicama za pouzdanu umjetnu inteligenciju* možemo uvidjeti spoznaju o neizmjernim mogućnostima čovjeka koje moraju biti kontrolirane, usmjerene i ograničene, inače čovjekova postignuća u sferi razvo-

ja umjetne inteligencije mogu postati prijatnija njegovu životu i opstanku. To nastojanje etičkih smjernica zaslužuje pozitivnu ocjenu. Istodobno, međutim, već na normativnoj razini postaju vidljiva i njihova ograničenja, koja proizlaze iz naglašenog antropocentrizma. Takav pristup stavlja u središte svih odnosa isključivo ljudsko biće, koje sebi pripisuje pravo određivati temeljna ljudska prava, definirati etička načela te konkretizirati zahtjeve koji bi trebali učiniti umjetnu inteligenciju „pouzdanom“. Čovjek se postavlja u središte stvaranja i nameće vlastite želje i potrebe kao najviše pravilo. Posljedica takva samopostavljanja nije samo gubitak osjećaja za transcendenciju nego i gubitak svijesti o vlastitim granicama jer se čovjek počinje doživljavati kao mjerilo svega o čemu ovisi ne samo njegova osobna egzistencija nego i opstanak drugih živih bića te same Zemlje.

Iz takve perspektive rada se potreba za potpunom kontrolom te duboki strah pred budućnošću: strah od proizvodnje sustava i strojeva koji bi mogli nadići čovjeka svojim mogućnostima; strah od gubitka kontrole nad njima i od toga da čovjek bude prepušten na milost i nemilost algoritamskih izračuna; strah, napokon, od proizvodnje (greškom ili namjerno) takvog sustava koji bi se mogao okrenuti protiv čovjeka. U tom smislu, etičke smjernice za umjetnu inteligenciju – kao i svi pokušaji određivanja pravila njezina razvoja – doista proizlaze iz svijesti da tehnološki napredak mora biti reguliran kako ne bi postao destruktivan. Sva nastojanja oko njihova oblikovanja, kao i kasniji pokušaji reguliranja njihove primjene, zaslužuju najviše priznanje. U praksi se, međutim, pokazuje da pojedina temeljna etička načela teško nalaze dosljednu primjenu. Ovdje nije riječ tek o lošem programiranju, tehničkim ograničenjima ili pogrešnim ciljevima, nego o odsutnosti upravo onih ljudskih vrлина koje oblikuju moralnu i egzistencijalnu kvalitetu ljudskog djelovanja – ljubavi, vjere i nade – vrлина po kojima čovjek ne samo računa i zaključuje nego voli i brine se, vjeruje i suosjeća, razumije i sanja.

Čovjek je stvoren za drugoga; stvoren je za zajedništvo s Bogom i ljudima, i srce mu ostaje nemirno dok se ne smiri u Bogu koji ga je stvorio za sebe (Aurelije Augustin 1973, 1.1). No etičke smjernice ne govore o Bogu, time se u njihovoj pozadini učvršćuje model autonomije lišen transcendentalne reference.

Taj antropocentrični obrat može se teološki čitati kao gubitak identiteta čovjeka kao Božje slike i prilike (Post 1,26), a time i gubitak njegove svrhe i uloge u svijetu. Kršeći zapovijed da ne jede sa stabla spoznaje dobra i zla (Post 2,17), čovjek odbacuje transcendentnu definiciju dobra

i zla koja dolazi od Boga, želeći sam odlučivati što je dobro, a što zlo (Ravasi 1995, 49).

Bog je dao čovjeku mjerilo dobra i zla radi očuvanja života u harmoniji Edenskog vrta. No čovjeku stablo izgleda dobro i zamamljivo, a osobito ga privlači obećanje da će postati kao bogovi koji razlučuju dobro i zlo (Post 3,5). Prevaren od zmije, uzevši i pojevši plod sa stabla spoznaje dobra i zla, čovjek ne postaje bog, nego otkriva vlastitu golotinju – vlastitu ograničenost kao stvorenja – koju zbog stida pokušava prikriti (Post 3,7.13) (Barščevski 2023, 47).

Teško je ne prepoznati u etičkim smjernicama, barem u njihovim dominantnim modernim formulacijama, odjek te iste temeljne dinamike: da čovjek sam određuje što je dobro, a što zlo; da sam postavlja temelje i etička načela koja bi mu jamčila blagostanje i sigurnost kao da bi mu najveća prijetnja životu bio sam Bog koji ga je stvorio. Stoga najveći izazov koji pred čovjeka postavlja umjetna inteligencija nije primarno u umnažanju normativnih okvira, nego u ponovnom otkrivanju čovjeka kao Božje slike. Stvarajući i razvijajući umjetnu inteligenciju, čovjek na nju, svjesno ili nesvjesno, projicira sliku koju ima o sebi i o svojoj ulozi u svijetu. Ta slika, međutim, ne može biti potpuna ako se izostavi čovjekova transcendentalna dimenzija i njegov odnos s Bogom.

3.2. „Kao jedan od nas – znajući dobro i zlo!“ (Post 3,22)

Nakon što su čovjek i njegova žena prekršili zapovijed da ne jedu sa stabla spoznaje dobra i zla (Post 2,17), Bog u svom sudu jasno izriče posljedice kršenja njegove zapovijedi koje pogađaju sva tri temeljna odnosa čovjeka: prema Bogu, prema bližnjemu i prema zemlji (Post 3,14-19). Ipak, nakon svega Bog neočekivano konstatira da „čovjek postade kao jedan od nas – znajući dobro i zlo!“ te iznosi „pravi“ razlog progona iz Edenskog vrta: „da ne bi sada pružio ruku, ubrao sa stabla života pa pojeo i živio navijeke“ (Post 3,22). Izgleda da se ovdje čovjek kao Božja slika, tj. „kao jedan od nas“ implicitno povezuje s Božjom osobinom poznavanja dobra i zla, a da bi se spriječila „totalna identifikacija“ čovjek biva udaljen od stabla života, pristup kojemu je zatvoren kerubinima i plamenim mačem (Post 3,24). No jedina spoznaja koju čovjek u tom trenutku ima je da su on i žena goli, što pokušavaju bezuspješno prikriti smokovim lišćem (Post 3,7) te prebacuju sa sebe odgovornost na drugoga: čovjek optužuje ženu – pa i Boga koji ju je stavio uz njega – da mu je ona dala jesti sa stabla, a žena optužuje zmiju da ju je ona prevarila.

Knjiga Postanka opisuje tri primarna odnosa čovjeka: prema Bogu, prema prirodi i prema drugom ljudskom biću. Između Stvoritelja i čovjeka prolazi „dah života“ (*nišmat hayyim*, Post 2,7), što se razlikuje od daha života (*nišmat-rûah hayyim*, Post 7,22) koji imaju ostala živa stvorenja. Riječ je o sposobnosti da pronikne otajstva svijesti, pa čak i podsvijesti. To je sposobnost introspekcije i intuicije, sposobnost ulaska u same sebe, samosvijest. Čovjek ne gleda samo prema gore, nego je privučen i onim što je dolje. On je „prah zemlje“ i zato se i zove „adam“. U njemu teče Božji dah života, no njegova materijalnost ga vuče prema dolje. To je odnos prema ostalim materijalnim bićima, živima i neživima. „Obradivati“ i „čuvati“ vrt (Post 2,15) znak je ljudskog rada. Dati imena životinjama simbol je svijesti i vlasti, znanosti i tehnike. No čovjek u kojem teče Božji dah i koji vlada, odnosno upravlja svijetom još nije kompletan. On sada gleda ispred sebe tražeći pomoć „kao on“ (*k^enegdô*) – hebrejski izraz ukazuje ne toliko na nešto „slično“, kao što se obično razumije, nego na neku stvarnost koja bi bila „ispred“ ili „nasuprot“, tj. s prednje strane i na istoj razini. Tu nalazimo ideju „partnerstva“, to jest odnosa s bližnjim, s drugim ljudskim bićem u kojem čovjek može pronaći samoga sebe. Taj odnos tipiziran je u ljubavnom jedinstvu muškarca i žene (Post 2,21-25) (Ravasi 1995, 45-46).

Dijalog sa stvarima je prema dolje i inferioran, dijalog s Bogom je prema gore i transcendentan, dijalog s bližnjim je horizontalan i neposredan. Pisao je ironično američki romanopisac Saul Bellow: „Čak i Adam, koji je mogao razgovarati s Bogom na ti, na kraju je zaželio malo ljudskog društva“ (Ravasi 1995, 46). Grijech će poremetiti te odnose. Odnos između muškarca i žene je poremećen: s jedne strane porodaj, koji je izvor života i radosti, od tada prati bol, dok s druge strane par koji je nekad bio „jedno tijelo“ sada razdire žudnja žene prema mužu koji će njome vladati (Post 3,16). Poremećen je odnos i sa zemljom, od koje će se Adam s trudom hraniti je će mu radati trnjem i korovom (Post 3,17-18). Poremećen je i odnos s Bogom, koji će istjerati Adama i Evu iz raja kako ne bi imali pristup stablu života, a ograničit će i prisutnost svog daha u čovjeku.

Pitanje koje se nameće u vezi s umjetnom inteligencijom je u koji od tih triju odnosa treba smjestiti odnos s tehničkim uređajima ili prema tehničkim uređajima? Problem je taj što se to pitanje često prešućuje. Kao da se ostavlja otvorenom mogućnost nekog četvrtog odnosa u kojem bi čovjek zapravo bio bog za stvorene uređaje s umjetnom inteligencijom. Tu svoju nadmoć čovjek želi pokazati i dokazati kontrolirajući umjetnu inteligenciju te ju čineći u potpunosti usmjerenim na svoje vlastite po-

trebe. S druge strane, svjestan je da, kako bi uređaji s umjetnom inteligencijom funkcionirali i ispunjavali svoju funkciju, moraju imati neku veću ili manju autonomiju, što može za njega predstavljati prijetnju ako strojevi izmaknu kontroli ili ako netko nedobronamjeran uspije preuzeti kontrolu nad njima.

No autonomija ne znači automatski slobodu i postizanje jednakosti i ravnopravnosti s čovjekom. Koliko god bili sofisticirani i osposobljeni činiti nešto što ni sam čovjek ne može, oni su uvijek proizvod njegovih ruku i čovjek se ne bi trebao odnositi prema njima kao prema ravnopravnim partnerima ili, još gore, kao prema strojevima/sustavima/bićima koji ga nadmašuju i od kojih bi u nekoj bližoj ili daljoj budućnosti ovisio njegov život. Stav prema umjetnoj inteligenciji mora biti stav proizvođača ili autora prema svom proizvodu ili djelu koje treba pomoći čovjeku, a ne zamijeniti ga ili ugroziti njegovu sigurnost i život.

Prihvatanje relacijske koncepcije ljudskog bića ima prednost u tome što izbjegava da se odnos prema uređajima umjetne inteligencije može razumjeti samo u terminima subordinacije ili kao pokušaj da se zadrži ili povrati kontrola nad njima. Za Adriana Fabrisa korektan odnos je taj koji uvijek uzima u obzir razlike među subjektima koji su uključeni u taj odnos (Fabris 2019, 107). Tu je riječ i o slobodi koju treba razumjeti pravilno. Ako se sloboda razumije kao sloboda individuuma, u smislu gospodara i vladara vlastitih odnosa, očigledno je da će najnaivniji biti zavedeni rastom moći koju sa sobom donose tehnički uređaji jer su nesvjesni cijene koju treba platiti upravo u slobodi. Kad netko želi pristupiti pitanju trudeći se sačuvati autoreferencijalnu središnjost ljudskog bića, završava time da ne razmatra na ispravan način aktivnost umjetnih uređaja i da je pokušava regulirati normama koje je lako prekršiti (Fabris 2019, 107).

Ljudsko biće danas oscilira između osjećaja svemoći i osjećaja bespomoćnosti, paradoksalno povezanih među sobom. Oba ne dopuštaju ispravan stav prema tehnologijama kojima se koristimo. Samo u relacijskoj perspektivi, to jest uzimajući u obzir ljudsko biće kao subjekt brojnih odnosa i kao onoga koji je odgovoran za te odnose, može se uspostaviti s etičke točke gledišta dobra interakcija s umjetnim agensima. No to je zadaća koju samo ljudsko biće može preuzeti na sebe. Riječ je o suživotu između čovjeka i stroja, o odnosu čovjeka prema onome što je sam proizveo, prema vlastitoj kreaciji, a ne prema prirodi ili društvu (Spadaro i Banchoff 2019, 435).

Zaključak

Danas se, čini se, više ne može izdvojiti samo jedan aspekt ili karakteristika tehničkog sustava koja bi bila najuvjerljivije analogna čovjeku. Dok se, s jedne strane, mnoge vrhunske tehnologije ograničavaju na izvodne iznimno preciznih i usko definiranih zadataka, s druge strane pojavljuju se sustavi koji u sebi ujedinjuju više različitih „funkcija“. Takvi sustavi umjetne inteligencije ne samo da odgovaraju na raznovrsna pitanja te pronalaze i nude tražene podatke nego se upuštaju u ono što se doživljava kao stvarna i kontinuirana interkomunikacija s korisnikom. Razina te komunikacije nije više isključivo informativna, nego sve češće poprima i verificirajući karakter, u smislu da se od umjetne inteligencije traži provjera dobivenih informacija, pa čak i savjetodavni, jer joj se povjeravaju osobna pitanja, osjećaji i sumnje koji su donedavno bili rezervirani za članove obitelji, prijatelje ili duhovnike. U trenutcima nesigurnosti, moralne dvojbe ili egzistencijalne tjeskobe od umjetne se inteligencije očekuje savjet, razumijevanje ili barem koherentan i strukturiran odgovor.

Može se stoga ustvrditi da se od umjetne inteligencije očekuje znatno više nego što ona po svojoj naravi može pružiti: od matematičkog algoritma traži se razumijevanje i suosjećanje, podrška i prijateljstvo, moralno ili duhovno vodstvo, pa čak i prava uzajamna ljubav. Time se ne iskazuje samo visoko povjerenje u točnost i navodnu dobronamjernost njezinih odgovora nego i zabrinjavajuća spremnost da se osobne odluke, pa i oblikovanje vlastita života, povjere sustavu koji ne posjeduje osjećaje, svijest, odgovornost ni sposobnost istinskog odnosa.

Na taj se način potvrđuje početna teza rada da je jedan od ciljeva razvoja umjetne inteligencije stvaranje „drugoga“ na našu sliku, no u znatno složenijem i ozbiljnijem obliku od puke imitacije pojedinačnih ljudskih sposobnosti. U konačnici ključno pitanje nije može li umjetna inteligencija zamijeniti čovjeka u izračunavanju i analizi podataka, izvršavanju složenih zadataka ili u interakciji s drugim osobama, nego može li samostalno donositi odluke, odnosno razlikovati dobro i zlo. Upravo se u toj točki razotkriva temeljna granica koju umjetna inteligencija po samoj svojoj naravi ne može prijeći: razlikovanje dobra i zla nije rezultat obrade podataka ili algoritamske optimizacije, nego čin slobodne i odgovorne osobe, koji pretpostavlja svijest, moralnu prosudbu i preuzimanje odgovornosti za donesene odluke.

Poznavanje dobra i zla nije prerogativa ljudske naravi, nego posljedica želje „biti kao bogovi koji razlučuju dobro i zlo“ (Post 3,5), potaknute

lukavstvom zmije koja izokreće Božju zapovijed da mogu slobodno jesti sa svakoga stabla u vrtu, osim sa stabla spoznaje dobra i zla (Post 2,16-17). Privučen obećanjem trenutačnoga užitka („dobro za jelo“), zavodljivošću izgleda („za oči zamamljivo“) i težnjom graditi svoj vlastiti život („za mudrost poželjno“) neovisno o Bogu, čovjek poseže za „zabranjenim plodom“ i razotkriva vlastitu golotinju – spoznaju ograničenosti kao stvorenja – koju pokušava prikriti smokvinim lišćem (Post 3,7). Pred „glasom“ (hebr. *qôl*) Boga „koji šeta vrtom za dnevnog povjetarca“ čovjek se bezuspješno sakriva među stablima (Post 3,8). Vjerojatno ni čovjek ni njegova žena nisu mogli naslutiti da će se njihovo nastojanje „biti kao bogovi koji razlučuju dobro i zlo“ započeti upravo spoznajem vlastite golotinje, strahom pred Božjim glasom, a na koncu i istjerivanjem iz Edemskog vrta.

Ipak Bog, koji potvrđuje čovjekovo „postignuće“ (“Evo, čovjek postade kao jedan od nas – znajući dobro i zlo”, Post 3,22), nakon što je čovjeka i njegovu ženu odjenuo u odjeću od krzna – darujući im time dostojanstvo koje uvelike nadilazi ono što čovjek može sam sebi osigurati – istjeruje ih iz vrta i nastanjuje istočno od njega, kako bi čovjek obrađivao zemlju iz koje je bio uzet. Time ga, moglo bi se reći, vraća na „tvorničke postavke“, tj. prvobitnom pozivu obrađivati i čuvati stvoreni svijet (Post 2,5.15), gdje *imago Dei* nije prisvojena moć, nego povjerena zadaća i odgovornost, a time i put obnove Božje slike na koju je stvoren (Post 1,27) – put koji koji ne leži u povratku u izgubljeni raj, nego u suobličavanju s Kristom, „slikom Boga nevidljivoga“ (Kol 1,15) i „trajnim likom Božjim“ (Fil 2,6), „novim Adamom“, koji je „objavljujući misterij Oca i njegove ljubavi potpuno otkrio i čovjeka njemu samome te mu objavio uzvišenost njegova poziva“ te „Adamovim sinovima vratio sličnost s Bogom“ (GS 22).

Božja „šetnja vrtom za dnevnog povjetarca“ podsjeća na ključni aspekt sličnosti čovjeka Bogu, koji dosad još nije bio razmotren: na subotu, sedmi dan stvaranja u kojem Bog dovršava svoje djelo i počiva od svega što je učinio te je blagoslivlje i posvećuje kao počinak posvećen Bogu (Post 2,2-3). Sklapajući Savez sa svojim narodom na gori Sinaj, Bog izričito zapovijeda održavanje i svetkovanje toga sedmog dana (Izl 20,8-11; Pnz 5,12-15), što treba podsjetiti čovjeka na Božje djelo stvaranja (Izl 20,11) i oslobođenje iz egipatskog ropstva (Pnz 5,15).

Božji odmor nije povezan s umorom, nego je znak završenosti djela stvaranja, čiji je vrhunac blagoslov i posvećenje. S druge strane „granica“ koju Bog stavlja blagoslovom sedmoga dana djelu stvaranja jasno pokazuje njegovu kontrolu nad vlastitom stvaralačkom moći. Drugim riječima, Bog upravlja svojom snagom, tj. jači je od nevjerojatne moći stvara-

nja kojom stvara svjetlost i rastavlja je od tame, stvara svod i dijeli vode, čini da zemlja klija zelenilom, postavlja Sunce i Mjesec na svod nebeski da vladaju danom i noću, stvara morske grdosije, svakovrsna živa bića i ptice krilate te, naspoljetku, i čovjeka da bude njihov gospodar, a ipak može stati i proslaviti blagoslovom kraj svojem djelu stvaranja. Stoga se sila stvaralačke moći ne mjeri ekspanzijom ili umnažanjem, nego snagom upravljanja tom moći i mogućnošću dovršenja započetog djela.

Simbol upravljanja moći da se zemlja podloži i da se vlada svim živim stvorovima na njoj bilo je upravo stablo spoznaje dobra i zla, s kojega je čovjek otkinuo plod i pojeo, prekršivši Božju zapovijed, ali još više pokazavši svoju nesposobnost da vlada snagom koju je primio od Boga. Upravo u tome se, moglo bi se reći, nalazi najveći izazov u stvaranju i razvoju umjetne inteligencije: hoće li čovjek biti sposoban kontrolirati vlastitu snagu ili će postati ovisan o onome što će stvoriti? Čovjek kao *imago Dei* pozvan je nasljedovati Boga koji zna zaustaviti se i odpočinuti od svoga djela te blagosloviti i posvetiti sedmi dan, odnosno vrijeme koje mu je Bog darovao.

Teško je zamisliti, ne samo teorijski nego i praktički, da bi se sav tehnološki razvoj u jednom trenutku mogao zaustaviti. No to samo još više ističe pitanje čovjekove stvarne sposobnosti upravljanja tehnološkim razvojem čak i uz nastojanja da se taj razvoj barem normativno regulira etičkim smjernicama. Riječ je pritom ne samo o teorijskim mogućnostima nego o konkretnim i dalekosežnim posljedicama koje sa sobom nosi za čovjeka biti *imago Dei* – znajući dobro i zlo. Ne radi se o tome da čovjek sam sebi postavi granice vlastita razvoja, nego da, svjestan svoje sličnosti s Bogom, usmjeri svoju stvaralačku moć na blagoslov i posvećenje prostora i vremena u kojem živi.

Popis literature

- AI Ethics: An Abrahamic Commitment to the Rome Call. 2023. U: <https://www.romecall.org/ai-ethics-an-abrahamic-commitment-to-the-romecall/> (pristupljeno 27. XII. 2025.).
- AI Ethics: An Abrahamic Commitment to the Rome Call. Joint Declaration. 2023. U: <https://www.romecall.org/wp-content/uploads/2023/01/AI-Joint-Declaration-5-Jan1-1.pdf> (pristupljeno 1. I. 2024.).
- AI HLEG. 2019a. Definicija umjetne inteligencije: glavne mogućnosti i znanstvene discipline. Bruxelles: Europska komisija. U: <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines> (pristupljeno 6. XII. 2023.).

- AI HLEG. 2019b. Etičke smjernice za pouzdanu umjetnu inteligenciju. Bruxelles: Europska komisija. U: <https://op.europa.eu/s/pGax> (pristupljeno 6. XII. 2023.).
- Aurelije Augustin. 1973. *Ispovijesti*. Vol. 3-4. Collectio patrum et scriptorum christianorum. Prijevod: Stjepan Hosu. Kršćanski klasici. Zagreb: Kršćanska sadašnjost.
- Barščevski, Taras. 2023. *Aktualne teme biblijske teologije*. Zagreb: Kršćanska sadašnjost.
- Barth, Karl. 1958a. Church Dogmatics III/1. *The Doctrine of Creation*. Bromiley, G. W. i Thomas F. Torrance (ur.). Edinburgh: T&T Clark.
- Barth, Karl. 1958b. Church Dogmatics III/2. U: Bromiley, Geoffrey W. i Thomas F. Torrance (ur.). *The Doctrine of Creation*. Edinburgh: T&T Clark.
- IBM AI Ethics Board. 2025. AI agents: Opportunities, Risks and Mitigations. U: <https://www.ibm.com/downloads/documents/us-en/1227c-12efb38b2b3> (pristupljeno 14. II. 2026.).
- Cairns, David. 1953. *The Image of God in Man*. New York: Philosophical Library.
- COMECE. 2019. Robotisation of Life. Ethics in view of new challenges (siječanj 2019.). Brussels. U: <https://www.comece.eu/wp-content/uploads/sites/2/2022/04/20190204-Robotization-of-Life-Ethics-in-view-of-new-challenges.pdf> (pristupljeno 26. XII. 2025.).
- Dicastero per la Dottrina della Fede i Dicastero per la Cultura e l'Educazione. 2025. Antiqua et nova. Nota sul rapporto tra intelligenza artificiale e intelligenza umana (28. I. 2025.). Roma: LEV.
- Etica dell'intelligenza artificiale: l'impegno delle religioni abramitiche nella Rome Call. 2023. (10. I. 2023.). U: <https://news.microsoft.com/it-it/2023/01/10/etica-dellintelligenza-artificiale-limpegno-delle-religioni-abramitiche-nella-rome-call/> (pristupljeno 27. XII. 2025.).
- Fabris, Adriano. 2019. Intelligenza artificiale e libertà umana. *Vita e Pensiero* 5: 101-107.
- Franjo, FT. 2020. *Fratelli tutti*. Enciklika o bratstvu i socijalnom prijateljstvu (3. X. 2020.). Zagreb: Kršćanska sadašnjost.
- Franjo, LS. 2015. *Laudato si'*. Enciklika o brizi za zajednički dom (24. III. 2015.). Zagreb: Kršćanska sadašnjost.
- Gasser, Les. 1991. Social conceptions of knowledge and action: DAI foundations and open systems semantics. *Artificial Intelligence* 47/1-3: 107-138. doi.org/10.1016/0004-3702(91)90052-L.
- Google. 2018. Artificial Intelligence at Google: Our Principles. U: <https://ai.google/responsibility/principles/> (pristupljeno 19. XII. 2025.).

- Haugeland, John. 1989. *Artificial Intelligence: The Very Idea*. Massachusetts: The MIT Press.
- Hehn, Johannes. 1915. Zum Terminus «Bild Gottes». U: Weil, Gotthold (ur.), *Festschrift Eduard Sachau zum siebzigsten Geburtstag*, 36-52. Berlin: Verlag von Georg Reimer.
- Herzfeld, Noreen. 2002a. Creare a nostra immagine: l'intelligenza artificiale e l'Imago Dei. *Annali di Studi Religiosi* 3: 9-25. Trento: EDB.
- Herzfeld, Noreen. 2002b. Creating in Our Own Image: Artificial Intelligence and the Image of God. *Zygon: Journal of Religion and Science* 37/2: 303-316. doi.org/10.1111/0591-2385.00430.
- Future of Life Institute. 2017. Asilomar AI Principles. U: <https://futureoflife.org/open-letter/ai-principles/> (pristupljeno 19. XII. 2025.).
- Intelligent Robotics Laboratory. 2018. Geminoid HI-6. Osaka University. U: <https://eng.irl.sys.es.osaka-u.ac.jp/robot> (pristupljeno 19. XII. 2025.).
- Ivezić, Bernard. 2023. Otac Appleove AI asistentice Siri tvrdi da ne razumijemo što AI jest. *Startup Report* 8: 36-37. U: <https://apps.jutarnji.hr/Startup-Report-2023Q4-No8.pdf> (pristupljeno: 19. XII. 2025.).
- Julia, Luc. 2019. *L'intelligence artificielle n'existe pas. 1 ressource en ligne (291 pages) vols.* Paris: Éditions First.
- Kouravanas, Nikolaos i Andreas Pavlopoulos. 2022. Social Robots: The case of Robot Sophia. *Homo Virtualis* 5/1: 136-165. doi.org/10.12681/homvir.30320.
- Mayor, Adrienne. 2018. *Gods and robots : myths, machines, and ancient dreams of technology*. Princeton: Princeton University Press.
- Microsoft. 2021. Responsible AI Principles from Microsoft. U: <https://www.microsoft.com/en-us/ai/responsible-ai> (pristupljeno 19. XII. 2025.).
- Université de Montréal. 2018. Montréal Declaration for Responsible AI. U: <https://montrealdeclaration-responsibleai.com/the-declaration> (pristupljeno 19. XII. 2025.).
- Moravec, Hans P. 2000. *Robot: mere machine to transcendent mind*. New York: Oxford University Press New York.
- Niebuhr, Reinhold. 1996. *The Nature and Destiny of Man: A Christian Interpretation*. Volume I: *Human Nature*. Luisville: Westminster John Knox Press.
- Nishio, Shuichi, Hiroshi Ishiguro, i Norihiro Hagita. 2007. *Geminoid: Teleoperated Android of an Existing Person*. U: *Humanoid Robots - New Developments*, Pina Filho, Armando Carlos de (ur.), 343-352. London: In-techOpen.

- Pasek, Anne. 2014. Renaissance Robotics: Leonardo da Vinci's Lost Knight and Enlivened Materiality. *Shift: Graduate Journal of Visual and Material Culture* 7: 1-25.
- Pavlič, Richard i Martina Vlahović. 2017. Čovjek – slika Božja. Teološka interpretacija i suvremeni izazovi. *Riječki teološki časopis* 25/2: 359-387.
- Putica, Marija. 2018. Umjetna inteligencija: dvojbe suvremenoga razvoja. *Hum* 13/20: 198-213.
- Ramsey, Paul. 1993. *Basic Christian ethics*. Louisville: Westminster/John Knox Press.
- Ravasi, Gianfranco. 1995. *Il racconto del cielo. Le storie, le idee, i personaggi dell'Antico Testamento*. Milano: Arnoldo Mondadori Editore.
- Rome Call for AI Ethics. 2020. Rome: RenAIssance Foundation. U: <https://www.romecall.org/> (pristupljeno: 14. II. 2026.).
- Singbo, Odilon-Gbènoukpo. 2024. Kršćanska antropologija i umjetna inteligencija u doba transhumanizma. *Bogoslovska smotra* 94/1: 53-74. <https://doi.org/10.53745/bs.94.1.6>.
- Spadaro, Antonio i Thomas Banchoff. 2019. Intelligenza artificiale e persona umana. Prospettive cinesi e occidentali. *Civiltà Cattolica* 170/2: 432-443.
- Šegula, Andrej. 2024. Umetna inteligencija v luči praktične pastoralne teologije. *Bogoslovni vestnik* 84/4: 899-908. doi.org/10.34291/BV2024/04/Segula
- Tanzella-Nitti, Giuseppe. 2025. Intelligenza artificiale: cosa è e come si usa. U: <https://disf.org/ai/2025> (pristupljeno 12. VIII. 2025.).
- Tridente, Giovanni. 2022. Anima digitale. La chiesa alla prova dell'Intelligenza Artificiale. Todi: Tau Editrice (ePub).
- International Telecommunication Union. 2023. AI for Good Global Summit. U: <https://aiforgood.itu.int/summit23/> (pristupljeno 19. XII. 2025.).
- University of Helsinki. 2020. Elements of AI: online tečaj. U: <https://course.elementsofai.com/hr/> (pristupljeno 20. XII. 2025.).
- von Rad, Gerhard. 1973. *Genesis: A Commentary*. Philadelphia: The Westminster Press.
- Wagle, Paul. 2023. Rome Call for AI Ethics: a Comprehensive Guide. U: <https://paulwagle.com/rome-call-for-ai-ethics-a-comprehensive-guide/> (pristupljeno 27. XII. 2025.).
- Weisz, Eliezer Simha. 2023. AI Ethics: An Abrahamic commitment to the Rome Call- reflections (10. I. 2023.). U: <https://www.romecall.org/wp-content/uploads/2023/01/RABBI-WEISZ-CONTRIBUTION.pdf> (pristupljeno 6. XII. 2023.).