

UDK
004.8:272
004.8:17
27-732.2:004.8
(044.6)
<https://doi.org/10.53745/bs.94.1.5>
Primljeno: 19. 1. 2024.
Prihvaćeno: 23. 2. 2024.
Pregledni rad

CRKVA PRED ETIČKIM IZAZOVIMA UMJETNE INTELIGENCIJE

Taras BARŠČEVSKI

Katolički bogoslovni fakultet Sveučilišta u Zagrebu
Vlaška 38, p.p. 5, 10 000 Zagreb
taras.barscevski@kbf.unizg.hr

Sažetak

Brz razvoj umjetne inteligencije (UI) koja je ušla u gotovo sve sfere ljudskog života – od kulture i umjetnosti do obrazovanja, zdravstva, socijalnih pitanja, ekonomije, politike i svakodnevnog života – stvara nove izazove i prilike, no također nosi sa sobom značajne rizike koji mogu imati dalekosežne posljedice. Osim što UI može olakšati i ubrzati mnoge svakodnevne aktivnosti i podići kvalitetu života na višu razinu, postoji i rizik od prijetnji privatnosti, lošeg utjecaja na psihičko zdravlje i obiteljske odnose, povećanja društvenih i ekonomskih nejednakosti, kao i potencijalne zloporabe u nadzoru i vojnim primjenama, što bi moglo dovesti do tragičnih posljedica kako za planet Zemlju tako i za čovječanstvo u cjelini. Crkva, kao Majka i Učiteljica ne može ostati po strani tamo gdje se »stvara« nova slika čovjeka te »kroji« budućnost Zemlje i čovječanstva. Zato, prepoznajući svoju ulogu u brzo mijenjajućem svijetu tehnologija i umjetne inteligencije, Crkva naglašava odgovorno korištenje tehnologije u skladu s Božjom voljom i na dobrobit cijelog čovječanstva i svakog čovjeka, a posebice onih najpotrebnijih i najsiromašnijih. U ovom radu pokazujemo neke od scenarija razvoja umjetne inteligencije i važnost etičkih smjernica za sve sudionike tog procesa te kako crkveno učiteljstvo – od sv. Ivana Pavla II., pape Benedikta XVI., pa sve do pape Franje, kao i različitih vatikanskih institucija – odgovara danas na sve nove etičke izazove koje pred čovječanstvo stavlja umjetna inteligencija.

Ključne riječi: umjetna inteligencija (UI), ChatGPT, etičke smjernice, crkveno učiteljstvo, algoretika, Rome Call.

Uvod

Tema umjetne inteligencije (UI) i različita pitanja povezana s njom svakim danom postaju ne samo sve prisutnija među znanstvenicima i stručnjacima u sferi tehnologije nego uvelike nadilaze granice tradicionalnih tehnoloških domena, prožimajući različite aspekte ljudskog života od kulture i umjetnosti do obrazovanja, zdravstva i socijalnih pitanja do ekonomije, politike i svakodnevnog života. Ta sveprisutnost UI-a stvara nove izazove i prilike, no također sa sobom donosi značajne rizike koje mogu imati dalekosežne posljedice, uključujući prijetnje privatnosti, psihičkom zdravlju i obiteljskim odnosima, povećanje društvenih i ekonomskih nejednakosti, kao i potencijalnu zlorabu u nadzoru i vojnim primjenama, što bi moglo dovesti do tragičnih posljedica kako za okoliš i ljudski habitat tako i za sam ljudski život.

Već danas u mnogim sferama svjedočimo velikom doprinosu tehnologija povezanih s umjetnom inteligencijom koje uvelike olakšavaju i ubrzavaju mnoge svakodnevne aktivnosti, podižući kvalitetu života na višu razinu. No, s druge strane, upravo ubrzan rast i donekle neograničene mogućnosti UI-a potiču ne samo znanstvenike i stručnjake nego i filozofe, humaniste, pisce i filmske režisere da zamišljaju kako razvoj i suradnja čovjeka i stroja može služiti na dobrobit čovječanstva, ali često prognoziraju i najmračnije apokaliptične scenarije u kojima bi roboti mogli zavladati svijetom i uništiti čovječanstvo.

Pred sve većim utjecajem umjetne inteligencije na svakodnevni život, ali i pred promjenama paradigmi ljudskog života, Crkva ne može ostati po strani, nego je pozvana dati svoje svjedočanstvo vjere, nade i ljubavi, naviještajući evanđelje spasenja. Pred sve rasprostranjenijima osjećajima ravnodušnosti i nemoći ona mora posvjedočiti svoju vjeru u Boga koji je »tako ljubio svijet da je dao svoga Sina Jedinorođenca da nijedan koji u njega vjeruje ne propadne, nego da ima život vječni« (Iv 3,16). Pred razočarenjem i apokaliptičkim scenarijima kraja svijeta ona je pozvana probuditi u ljudima sigurnost nade u Božji naum i bolji svijet. Pred bešćutnošću prema boli i stradanjima drugih, hladnoćom ekonomskih računica te egoizmom, Crkva je pozvana pokazati milosrđe i ljubav, koji su nadahnuti istinom.

U ovom radu pokazujemo neke od scenarija razvoja umjetne inteligencije i važnost etičkih smjernica za sve sudionike tog procesa te kako crkveno učiteljstvo odgovara na etičke izazove koje pred čovječanstvo stavlja umjetna inteligencija.

1. Umjetna inteligencija između stvarnosti i »realne imaginacije«

U svom članku iz 2000. godine naslovljenom »Why the Future Doesn't Need Us« računalni znanstvenik i inženjer Bill Joy, jedan od osnivača tvrtke Sun Microsystems, upozorio je na činjenicu da bi do 2030. godine samoreplicirajuća računala i napredak u nanotehnologiji mogli dovesti do računalne tehnologije koja bi mogla imati značajne posljedice za ljudsku vrstu i društvo općenito s obzirom na to da bi roboti mogli postati superiorniji u odnosu na ljude.¹ Istraživač robotike i umjetne inteligencije Hans Moravec pretpostavlja da će se do 2040. godine pojaviti četvrta generacija robota koja bi imala sposobnost rasuđivanja i vještine nalik ljudskima te bi na neki način nama i nalikovala, ali će u drugim aspektima biti kao nešto što svijet dosad nije vidio, jer kako roboti budu dizajnirali svoje nasljednike, tako će i svijet postajati sve »čudniji«.² Boljim i jeftinijim izvođenjem poslova roboti će istisnuti ljude iz njihovih uloga. »Prilično brzo«, piše Moravec, »mogli bi nas istisnuti iz samog postojanja.«³

Od pojave umjetne inteligencije u 50-im godinama 20. stoljeća čini se da je teško procijeniti jesu li se ostvarile sve prognoze znanstvenika, ali i nade i strahovi ljudi. Velik utjecaj na stvaranje predodžaba o budućnosti svijeta, ulozi umjetne inteligencije te interakcije između ljudi i robota imaju ne samo znanstvenici nego, možda još i više, pisci znanstvene fantastike i filmski režiseri. Niz knjiga o robotima (Ja, Robot,⁴ Kompletni robot,⁵ Čelične spilje,⁶ Golo Sunce,⁷ Roboti zore,⁸ Roboti i Carstvo,⁹ Pozitronski čovjek)¹⁰ i Zakoni robotike¹¹

¹ Usp. Bill JOY, Why the Future Doesn't Need Us (I. IV. 2000.), u: *Wired*, u: <https://www.wired.com/2000/04/joy-2/> (29. XI. 2023.).

² Usp. Hans P. MORAVEC, *Robot: mere machine to transcendent mind*, New York, 2000., 108-110.

³ *Isto*, 13.

⁴ Usp. Isaac ASIMOV, *Ja, robot*, Beograd, 1977.

⁵ Usp. Isaac ASIMOV, *The complete robot*, London, 2018.

⁶ Usp. Isaac ASIMOV, *Čelične spilje*, Zagreb, 1957.

⁷ Usp. Isaac ASIMOV, *Golo sunce*, Zagreb, 1959.

⁸ Usp. Isaac ASIMOV, *Roboti zore*, Beograd, 1985.

⁹ Usp. Isaac ASIMOV, *Roboti i carstvo*, Beograd, 1990.

¹⁰ Usp. Isaac ASIMOV, *Pozitronski čovek*, Beograd, 1993.

¹¹ Tri zakona robotike dio su kratke priče »Runaround«, koja se pojavila 1942. godine u američkom znanstveno-fantastičnom časopisu *Astounding*: 1) robot ne smije naškoditi čovjeku ili svojom pasivnošću dopustiti da se čovjeku naškodi; 2) robot mora slušati ljudske naredbe, osim kad su one u suprotnosti s prvim zakonom; 3) robot treba štiti svoj integritet, osim kad je to u suprotnosti s prvim ili drugim zakonom«, Isaac ASIMOV, Runaround, u: *Astounding Science – Fiction*, 29 (1942.) 1, 100. Kasnije se pojavljuje u knjizi *Ja, robot* iz 1950. godine, gdje ih je Asimov naveo kao pisane zakone za (fikcijsko) izdanje *Priručnika robotike* iz 2058. godine. Usp. Isaac ASIMOV, *I, robot*, New York, 1950., 40. U kasnijim djelima Asimov je dodao i četvrti, ili nulti zakon koji prethodi drugima: »0. Robot ne smije naškoditi čovječanstvu ili svojom pasivnošću dopustiti da se čovječanstvu naškodi«.

Isaaca Asimova postali su temeljni žanra znanstvene fantastike o umjetnoj inteligenciji tako da ih nalazimo u mnogim kasnijim knjigama i filmovima. Tu je i roman Philipa K. Dicka *Do Androids Dream of Electric Sheep?*,¹² koji je postao inspiracija za film *Blade Runner*,¹³ koji istražuje granice između čovjeka i umjetne inteligencije. Nadalje, tu je roman Arthura C. Clarkea *2001: Odiseja u svemiru*,¹⁴ koji postavlja pitanja o svijesti i osjećaju mašina i ljudskoj interakciji s mašinama, i Williama Gibsona *Neuromancer*,¹⁵ što je ključno djelo novog književnog smjera unutar znanstvene fantastike, cyberpunka, u kojem osmišljava pojam »kibernetičkog prostora i kriminala« te istražuje umreženu budućnost i umjetnu inteligenciju davno prije nego je internet kao takav bio uopće osmišljen i poznat široj javnosti.

Mnoge tehnološke inovacije iz filmova 70-ih i 80-ih godina nisu se ostvarile (*Terminator*,¹⁶ *Robocop*,¹⁷ *Povratak u budućnost*¹⁸), a umjetna inteligencija, često prikazivana u ranim filmovima kao prijetnja čovječanstvu u smislu inteligentnih strojeva koji bi mogli uništiti čovječanstvo (Dan kada je zemlja stala,¹⁹ *Kolos: Projekt Forbin*²⁰) postaje sve sofisticiranija prikazujući robote koji mogu komunicirati s ljudima na realan način (*Ratovi zvijezda IV: Nova nada ili Zvezdani ratovi*,²¹ *Terminator*²²). U 90-im se godinama 20. stoljeća pojavljuju filmovi koji prikazuju robote dizajnirane da služe ljudima i postanu njihovi prijatelji (*Kratki spoj*,²³ *Mojih 200 godina*²⁴), a već u našem 21. stoljeću filmovi istražuju složene odnose koji se mogu razviti između ljudi i strojeva (*Umjetna inteligencija*,²⁵ *Ona*,²⁶ *Ex Machina*²⁷) ili se bave temama samosvijesti i emocija

¹² Usp. Philip K. DICK, *Do androids dream of electric sheep?*, New York, 1996.

¹³ Usp. Ridley SCOTT, redatelj, *Blade Runner*, Warner Bros., 1982.

¹⁴ Usp. Arthur C. CLARKE, *2001: Odiseja u svemiru*, Zagreb – Opatija, 1993.

¹⁵ Usp. William GIBSON, *Neuromancer*, Zagreb, 2001.

¹⁶ Usp. James CAMERON, redatelj, *The Terminator*, Orion Pictures – Metro–Goldwyn–Mayer, 1984.

¹⁷ Usp. Paul VERHOEVEN, redatelj, *RoboCop*, Orion Pictures, 1987.

¹⁸ Usp. Robert ZEMECKIS, redatelj, *Back to the Future*, Universal Pictures, 1985.

¹⁹ Usp. Robert WISE, redatelj, *The Day the Earth Stood Still*, 20th Century Fox, 1951.

²⁰ Usp. Joseph SARGENT, redatelj, *Colossus: The Forbin Project*, Universal Pictures, 1970.

²¹ Usp. George LUCAS, redatelj, *Star Wars Episode IV: A New Hope*, 20th Century Fox – Lucasfilm, 1977.

²² Usp. James CAMERON, redatelj, *The Terminator*, Orion Pictures – Metro–Goldwyn–Mayer, 1984.

²³ Usp. John BADHAM, redatelj, *Short Circuit*, Tri–Star Pictures, 1986.

²⁴ Usp. Chris COLUMBUS, redatelj, *Bicentennial Man*, Columbia – Touchstone, 1999.

²⁵ Usp. Steven SPIELBERG, redatelj, *A. I. Artificial Intelligence*, Warner Bros. – DreamWorks, 2001.

²⁶ Usp. Spike JONZE, redatelj, *Her*, Warner Bros. Pictures, 2013.

²⁷ Usp. Alex GARLAND, redatelj, *Ex Machina*, Universal Pictures, 2014.

umjetne inteligencije (Uzvišenost,²⁸ Chappie²⁹) dok filmovi Matrix³⁰ i WALL–E³¹ prikazuju distopijsku budućnost u kojoj su strojevi preuzeli primat nad ljudima.

U isto vrijeme stvarna primjena umjetne inteligencije postaje sve raširenija. UI se sada koristi u svemu, od glasovnih asistenata, robota za čišćenje, kuhinjskih aparata i pametnih domova preko personaliziranog digitalnog marketinga i automatiziranog korisničkog servisa, do samovozećih automobila i zdravstvenih dijagnoza, kao i u optimizaciji lanca opskrbe, inteligentnom upravljanju energijom i razvoju pametnih gradova, te čak i u kreiranju personaliziranih obrazovnih programa i progresivnim istraživanjima u biotehnologiji.

Granica između znanstvene fantastike i stvarnosti se briše, a filmaši se koriste najnovijom tehnologijom za stvaranje realističnijih i nijansiranih prikaza umjetne inteligencije u filmovima.³² Čini se da više ne postoje granice između naših mogućnosti i mašte. Pitanje je samo vremena kada će to što je zasad predstavljeno kao plod mašte postati stvarnost. I koji će zamišljeni (izmišljeni) scenarij postati stvarnost? Matrix ili WALL–E? Terminator ili Ex Machina? Iako smo svjesni da je riječ o mašti, ne možemo se ipak oti dojmom da se može ostvariti neki od tih scenarija, a da naše buduće perspektive uvelike ovise od sadašnjosti. Upravo na tome i naglašava Andrea Ciucci, koordinator središnjeg ureda Papinske akademije za život u svom intervjuu prilikom prezentacije svoje knjige *Oprostite, ali zašto ste vi ovdje?* Priče o ljudskim i umjetnim inteligencijama:³³ »Izraz umjetna inteligencija na prvi pogled ukazuje na nešto u budućnosti. No nije tako. To je sadašnjost našeg postojanja [...] pred ogromnom snagom strojeva ne smijemo prestati biti ljudi.«³⁴

Mašta pisaca, kao doprinosi i pretpostavke teoretičara još uvijek nadilaze realne mogućnosti ostvarivosti raznih scenarija razvoja umjetne inteligencije, no filmski scenariji kao i primjena najnovijih tehnologija u snimanju filmova sve više nadilaze jaz između mašte i stvarnosti. Kombinirajući već ostvarena

²⁸ Usp. Wally PFISTEE, redatelj, *Transcendence*, Waner Bros. Pictures – Summit Entertainment, 2014.

²⁹ Usp. Neill BLOMKAMP, redatelj, *Chappie*, Columbia Pictures, 2015.

³⁰ Usp. Lana WACHOVESKI – Lilly WACHOWSKI, redatelji, *The Matrix*, Warner Bros., 1999.

³¹ Usp. Andrew STANTON, redatelj, *WALL–E*, Walt Disney Pictures – Pixar Animation Studios, 2008.

³² Usp. Jan MACAROL, *Beyond Human: Najbolji AI filmovi svih vremena* (3. III. 2023.), u: *City Magazine*, u: <https://citymagazine.si/hr/onkraj-covjeka-najbolji-filmovi-o-umjetnoj-inteligenciji-svih-vremena/> (2. XII. 2023.).

³³ Usp. Andrea CIUCCI, *Scusi, ma perché lei è qui? Storie di intelligenze umane e artificiali*, Milano, 2022.

³⁴ REDAZIONE, *Intervista a Andrea Ciucci. Storie di intelligenze umane e artificiali* (14. IV. 2022.), u: *Terre di mezzo. Editore*, u: <https://www.terre.it/interviste/scrittori-illustratori/intervista-a-andrea-ciucci-storie-di-intelligenze-umane-e-artificiali/> (2. XII. 2023.).

postignuća u sferi umjetne inteligencije i robotike s usavršenim prototipovima i naprednim vizualnim efektima, filmska industrija uspijeva stvoriti uvjerljive i realistične prikaze budućih tehnologija koji kod gledatelja stvaraju ili osjećaj razotkrivanja nekih već postojećih tajnih tehnologija i eksperimenata ili sasvim bliske u budućnosti njihove realizacije te oblikuje našu percepciju budućnosti, posebno u kontekstu razvoja umjetne inteligencije i tehnologije. Sve što čovjek zamisli izgleda ostvarivo, no svako ostvarenje sa sobom vuče i konkretne posljedice, koje mogu biti ne samo ozbiljne i nepovratne nego razarajuće i smrtonosne za pojedince, ali i za cijelo čovječanstvo i svijet. Upravo zato smo više nego ikad svjesni potrebe etičkih principa u razvoju umjetne inteligencije i jasnim pravilima za njezino korištenje, no također i poteškoća u njihovoj izradi, a posebice u primjeni. Istodobno postavlja se pitanje hoće li etičke smjernice riješiti sve probleme i što ako ih netko prekrši ili ne prihvati? Drugim riječima, osim etičkih smjernica za umjetnu inteligenciju, što je neophodno za njezin pouzdan razvoj, koji su to temelji na kojima se one osnivaju i kakvu ulogu i doprinos u razvoju umjetne inteligencije mogu imati vjernici i Crkva.

2. Crkva i umjetna inteligencija

Crkva kao Majka i Učiteljica (usp. KKC, 2030–2046) koja okuplja ljude »iz svakog vremena, i plemena, i puka, i jezika« (Otk 7,9) ne može po svojoj naravi kao sakrament »najprisnijeg sjedinjenja ljudi s Bogom« i »jedinstva ljudskoga roda« te kao »znak i sredstvo« potpunog ostvarenja tog jedinstva u budućnosti (usp. KKC 775) ostati po strani tamo gdje se »stvara« nova slika čovjeka te »kroji« budućnost čovječanstva i Zemlje.

Za razliku od iskrivljenog mišljenja da je Crkva neka zaostala institucija koju je pregazilo vrijeme, koja ograničava čovjeka i uopće razvoj društva, zapravo ona je oduvijek išla ukorak s vremenom tražeći uvijek najbolja, često i inovativna, rješenja za propovijedanje i širenje evanđelja te imajući uvijek na umu svoju misiju naučavati, posvećivati i upravljati za dobro svakog pojedinca i cijele ljudske zajednice. Tako i danas pred novim izazovima znanstvenih i tehnoloških inovacija Crkva, osim što podsjeća da digitalna inovacija i tehnološki napredak moraju pridonositi razvoju i dobrobiti društva, posebice onih koji bi mogli potencijalno biti isključeni, te zaštititi i razvoju okoliša, također pomaže u integraciji različitih glasova u aktualan dijalog o umjetnoj inteligenciji i etici strojeva.³⁵

³⁵ Usp. Edoardo SINIBALDI – Chris GASTMANS – Miguel YÁÑEZ – Richard M. LERNER – László KOVÁCS – Carlo CASALONE – Renzo PEGORARO – Vincenzo PAGLIA, Con-

Stav Crkve danas prema umjetnoj inteligenciji kao i prema novim tehnologijama produžetak je zauzimanja i nastojanja Crkve oko razvoja i napretka »svega što je dobro i Bogu milo« (usp. Rim 12,2) ne samo kroz primjenu u svakodnevnom životu novih izuma nego ponajprije kroz inkulturaciju evanđelja u povijesti. No uvijek s veoma izričitom brigom za dobrobit cijelog čovječanstva i svakog čovjeka, a posebice onih najpotrebitijih i najsiromašnijih. Takav stav prema razvoju novih tehnologija uvelike ovisi o pogledu Crkve na čovjeka, o kojem se, diveći se Božjem stvorenju, nebesima, mjesecu i zvijezdama, psalmist pita: »Pa što je čovjek da ga se spominješ, sin čovječji te ga pohodiš?« (Ps 8,5), svjestan toga da ga je sam Bog učinio malo manjim od sebe, okrunivši ga slavom i sjajem te mu dao vlast nad djelima njegovih ruku i podložio mu sve pod noge (usp. Ps 8,4-7).

Crkva je danas svjesna da je »čovječanstvo ušlo u novo doba«³⁶ u kojem je »umjetna inteligencija sve prisutnija u svim aspektima svakodnevnog života«³⁷ te da se »tehnologijom doskočilo nebrojenim nevoljama koje su mučile i sputavale ljudsko biće« (LS 102), dok »inovacije u tom području ne samo utječu na to kako razumijemo ovaj svijet nego znače da su ta sredstva sve više odlučujuća za ljudsko djelovanje, čak i u donošenju odluka«³⁸. Papa Franjo konstatira da su »postignuti veliki napredci, osobito u medicini, inženjerstvu i komunikacijama, a mnogi znanstvenici i stručnjaci ulažu velike napore u osmišljavanju alternativa kako bi razvoj postao održiv« (LS 102) te izražava svoju radost što su uključeni »muškarci i žene dobre volje [...] kako bi 'algoretika', tj. etičko promišljanje o korištenju algoritama, bilo sve prisutnije ne samo u javnoj raspravi nego i u razvoju tehničkih rješenja«³⁹.

2.1. Pristup tehnologijama i umjetnoj inteligenciji u crkvenom učiteljstvu

Svjesna svog poslanja u svijetu koji se neprestano mijenja, posebno u kontekstu brzog razvoja umjetne inteligencije, a s obzirom na to da je Bog povjerio čovjeku skrb nad svojim stvorenjem, Crkva naglašava važnost odgovornog korištenja

tributions from the Catholic Church to ethical reflections in the digital era, u: *Nature Machine Intelligence*, 2 (2020.) 5, 243.

³⁶ Papa FRANJO, *Laudato si'*. Enciklika o brizi za zajednički dom (24. III. 2015.), Zagreb, 2015., br. 102 (dalje: LS).

³⁷ Papa FRANJO, Audience with participants in the meeting Rome Call for A.I. Ethics organized by the Pontifical Academy for Life and the RenAIssance Foundation (10. I. 2023.), u: <https://press.vatican.va/content/salastampa/en/bollettino/pubblico/2023/01/10/230110a.html> (2. XII. 2023.).

³⁸ *Isto.*

³⁹ *Isto.*

tehnologije na način »da ju se koristi, ali ne zlorabi, dokuči njezine tajne i razvije njezin potencijal«⁴⁰. U tom se duhu Crkva angažira u razmatranju kako tehnologija UI-a može služiti čovječanstvu na etički i moralno odgovoran način.

Prvi papa koji je u jednom svom govoru upotrijebio riječ robotika je Ivana Pavao II. Obraćajući se 31. siječnja 1987. biskupima iz talijanskog Pijemonta u pohodu Ad limina apostolorum papa izražava svijest o društvu koje se mijenja i postaje »tehnološki naprednije«, što je preokret »za koji ne znamo kamo će odvesti i kada će završiti«. Papa kaže: »Tehnika utječe na ekonomiju, koja je postala znanost te zahtijeva autonomiju od politike i etike. Osnovna pretpostavka postaje maksimalna efikasnost proizvodnje s minimalnim naporom. Zahvaljujući dostignućima tehnike, telematike, informatike, robotike, uloga osobe čini se uglavnom suvišnom, sekundarnom: u isto vrijeme, međutim, više se cijeni inteligencija nego fizička snaga.«⁴¹

Isti papa, Ivan Pavo II., tri godine kasnije u svom govoru prigodom V. Međunarodnog simpozija pod nazivom »Ljudski um« 17. studenoga 1990. godine prvi put spominje sintagmu umjetna inteligencija: »Danas se govori o 'umjetnoj inteligenciji' aludirajući na izvanredne mogućnosti 'elektroničkih mozgova'. Međutim, uvijek valja zapamtiti da na temelju informatike i kibernetike stoji nadređen podatak ljudske inteligencije koja, upravo zbog svog duhovnog karaktera i posljedične neprilagodljivosti samo fizičko-kemijskim fenomenima, u slobodnom razumijevanju prosuđuje, u shvaćanju može i birati, u spoznaji naslućuje svoju konačnu sudbinu.« Kako bi razumjeli čovjeka, znanstvenici ne mogu zanemariti »njegovo kompaktno jedinstvo fizičkog i duhovnog bića« i zato papa upozorava na rizik da mehanicistička i materijalistička tumačenja vode prema »poricanju osobe i njezina dostojanstva s posljedicama koje su bacile teške i tragične sjene na ljudsku povijest prošlosti, ali i na naše vrijeme«⁴².

Također i papa Benedikt XVI. upozorava na suvremeni kontekst koji »daje primat umjetnoj inteligenciji koja postaje sve više podložna eksperimentalnoj tehnici i na taj način zaboravlja da svaka znanost uvijek mora čuvati čovjeka i promicati njegovu težnju prema autentičnom dobru«. Potrebno je

⁴⁰ IVAN PAVAO II., Udienza generale (26. VI. 2002.), u: https://www.vatican.va/content/john-paul-ii/it/audiences/2002/documents/hf_jp-ii_aud_20020626.html (30. XII. 2023.).

⁴¹ IVAN PAVAO II., Discorso ai vescovi piemontesi in visita »Ad alimina apostolorum« (31. I. 1987), u: https://www.vatican.va/content/john-paul-ii/it/speeches/1987/january/documents/hf_jp-ii_spe_19870131_piemontesi-ad-limina.html (31. XII. 2023.).

⁴² IVAN PAVAO II., Discorso in occasione della V Conferenza Internazionale su »La mente umana« (17. XI. 1990.), u: https://www.vatican.va/content/john-paul-ii/it/speeches/1990/november/documents/hf_jp-ii_spe_19901117_mente-umana.html (31. XII. 1990.).

održavati ravnotežu kako bi svaka egzistencija imala »solidan temelj i valjanu svrhu« uz težnju »dubljeg razumijevanja života«⁴³.

Od svih se papa ipak najviše osvrnuo na temu umjetne inteligencije papa Franjo, što je sasvim razumljivo s obzirom na to da su se najveća otkrića u sferi robotike i umjetne inteligencije umnožila upravo u posljednjem desetljeću. Enciklika *Laudato sì* posvećena »brizi za zajednički dom« donosi najdetaljniji doprinos o tehnologiji, etici i inovaciji u doba Papina pontifikata u vezi s njezinim implikacijama na čovječanstvo. Cijelo treće poglavlje pod naslovom »Ljudski korijeni ekološke krize« usredotočeno je na »tehnokratsku paradigmu« u svjetlu širih promišljanja o ekološkoj krizi (usp. LS 101 – 136).

Giuseppe Tridente u svojoj knjizi *Digitalna duša*. Crkva suočena s izazovima umjetne inteligencije izdvojio je dvadeset i šest papinskih tekstova u razdoblju od 31. siječnja 1987. do 23. ožujka 2021. godine od kojih osam sv. Pape Ivana Pavla II., jedan pape Benedikta XVI. i sedamnaest pape Franje u kojima se govori ili spominje tematika umjetne inteligencije i/ili robotike.⁴⁴ Osnovne naglaske učenja tih papa G. Tridente sažima na sljedeći način: »Suočeni s društvom koje se mijenja – što je tražio već i Ivan Pavao II – moramo djelovati prije svega s: 1) odgovornošću; 2) zdravim razumom, ako doista težimo slobodi prosuđivanja i izbora; 3) autentičnim dobrom (Benedikt XVI.); 4) univerzalnom etikom (papa Franjo) i očuvanjem; 5) dostojanstvom ljudske osobe. Unošenjem u tu sredinu 6) malo duše (tal. *supplemento d'anima*) omogućit će se da strojevi rade ljudima pravu 7) uslugu, tj. da su u službi čovjeku. Ali odluka pojedinca nije dovoljna, 8) potrebna su posredna društvena tijela i 9) odgovarajuća regulacija.«⁴⁵

Osim učenja trojice papa, Tridente navodi još i druge akademske inicijative u razdoblju od 2016. do 2021. godine u kojima se raspravljalo o umjetnoj inteligenciji, a koje su organizirali Papinska akademija znanosti u suradnji s Papinskom akademijom društvenih znanosti, Papinskom akademijom za život i Papinskim vijećem za kulturu.⁴⁶ Od posebnog značenja je znanstveni doprinos isusovačkog časopisa *La civiltà cattolica* i novina u vlasništvu Talijanske biskupske konferencije *Avvenire* u kojima je objavljena monografija⁴⁷ te trideset

⁴³ BENEDIKT XVI., *Discorso in occasione della Visita alla Pontificia Università Lateranense* (21. X. 2006.), u: https://www.vatican.va/content/benedict-xvi/it/speeches/2006/october/documents/hf_ben-xvi_spe_20061021_lateranense.html (31. XII. 2023.).

⁴⁴ Usp. Giovanni TRIDENTE, *Anima digitale. La chiesa alla prova dell'Intelligenza Artificiale (e-book)*, Todi, 2022.

⁴⁵ *Isto*

⁴⁶ Usp. *Isto*.

⁴⁷ Usp. RAZNI AUTORI, *AI*, Roma 2020., u: <https://www.laciviltacattolica.it/prodotto/ia/> (2. I. 2024.).

i pet tematski namijenjenih publikacija koje u kontekstu umjetne inteligencije »pozivaju na potrebu da se napravi 1) prostor za glas Crkve, zahvaljujući njezinoj sposobnosti tumačiti i prenositi, i postavi 2) antropološko pitanje, koje leži u temelju znanstvenih pitanja. Ne može se isključiti iz rasprave, zapravo, 3) jedinstvenost ljudske osobe protiv zamke i iluzije neograničene supermoći. I u ovom slučaju ističe se mogućnost da se iz tehnologije izvuče neka vrsta 4) 'duhovne vrijednosti', kako bi se, i u ovom slučaju, sačuvalo nedodirljivo dostojanstvo čovjeka.«⁴⁸

Kao što se može vidjeti iz ovog veoma kratkog osvrtu na intervencije papa i druge inicijative Svete Stolice povezane s umjetnom inteligencijom, etičko pitanje nije uvijek bilo u središtu zanimanja, nego se više pažnje posvećivalo crkvenom nauku o čovjeku i njegovoj interakciji s društvom koje se mijenja. Iako postoje različita viđenja glede razvoja umjetne inteligencije među znanstvenicima i predstavnicima IT industrija, svi se oni ipak slažu glede nužnosti izrade etičkih smjernica. Sveta Stolica se posljednjih godina veoma aktivno pridružila tim nastojanjima za bolju regulativu koja bi opisala dozvoljene načine uporabe umjetne inteligencije i s njom povezanih tehnologija,⁴⁹ tražeći saveznike i istomišljenike ne samo među kompanijama iz IT industrije nego i među predstavnicima i pripadnicima drugih religija.

2.2. *Rome Call for AI Ethics*

Logičnim i konkretnim nastavkom najdetaljnijeg doprinosa o tehnologiji, etici i inovacijama u odnosu na posljedice koje tehnološki razvoj može imati na čovječanstvo enciklike pape Franje o brizi za zajednički dom *Laudato si* te različitih inicijativa vatikanskih institucija možemo smatrati organizaciju kongresa »RenAIssance. For a Human-centric Artificial Intelligence«, koji se održao u Rimu 28. veljače 2020. godine, a kulminirao je potpisivanjem dokumenta *Rome Call for AI Ethics*, koji je nastao u Papinskoj akademiji za život, a koji je uredila Zaklada RenAIssance za algor-etiku, odnosno za promicanje etičkog razvoja umjetne inteligencije.⁵⁰ Dokument su tada uz predsjednika Papinske akademije za život mons. Vincenza Paglija potpisali još i predsjednik Microsofta Brad

⁴⁸ Giovanni TRIDENTE, *Anima digitale. La chiesa alla prova dell'Intelligenza Artificiale (e-book)*.

⁴⁹ Usp. Sandro VRBANUS, Vatikan se uključio u raspravu oko etike umjetne inteligencije (29. II. 2020.), u: <https://www.bug.hr/umjetna-inteligencija/vatikan-se-ukljucio-u-raspravu-oko-etike-umjetne-inteligencije-14008> (2. I. 2024.).

⁵⁰ Usp. *Rome Call for AI Ethics* (28. II. 2020.), u: <https://www.romecall.org/the-call/> (14. I. 2024.).

Smith, vicepredsjednik IBM-a John Kelly III, glavni direktor FAO-a Dongyu Qu te talijanska ministrica za inovaciju Paola Pisano. Potpisnici su se obvezali promicati etički pristup umjetnoj inteligenciji (= algoretiku), koji obuhvaća tri područja utjecaja: etiku, obrazovanje i pravo te se temelji na šest načela: 1) transparentnost: UI mora biti razumljiv svima; 2) inkluzivnost: UI ne smije diskriminirati nikoga jer svi ljudi imaju jednako dostojanstvo; 3) odgovornost: uvijek mora postojati netko tko preuzima odgovornost za ono što radi stroj; 4) nepristranost: UI ne smije biti pristana ni stvarati pristranost; 5) pouzdanost: UI mora biti pouzdana; 6) sigurnost i privatnost: UI mora biti sigurna i poštivati privatnost korisnika.⁵¹

Dana 10. siječnja 2023. godine uz potpisnike prvog Rome Call susreta⁵² dokument Rome Call for AI Ethics potpisali su uz prvog promicatelja događaja »AI Ethics: An Abrahamic Commitment to the Rome Call« Vincenzo Paglia, također i glavni rabin i član Vijeća Glavnog rabinata Izraela Eliezer Simha Weisz te predsjednik Abu Dhabi Peace Foruma i predsjedavajući Vijeća za šerijatsku fetvu Ujedinjenih Arapskih Emirata šaik Abdallah bin Bayyah.⁵³ Involviranjem u promicanje etičkog razvoja umjetne inteligencije predstavnikâ abrahamskih religija možda se željelo uključiti u cijeli proces što više pojedinaca koji pripadaju jednoj od triju monoteističkih religija, a mogu biti investitori, programeri, odnosno konstruktori ili korisnici UI proizvoda, kako bi oni iz vlastitih religijskih uvjerenja prihvatili etičke smjernice razvoja umjetne inteligencije i primijenili ih u svom radu i životu.

Obraćajući se 10. siječnja 2023. godine sudionicima susreta »Rome Call«, papa Franjo govori o suradnji i suglasnosti sudionika u promicanju kulture koja stavlja umjetnu inteligenciju u službu općeg dobra svih i čuvanja zajedničkog dobra te podsjeća da je »bratstvo među svima ... uvjet da i tehnološki

⁵¹ Usp. Rome Call for AI Ethics (28. II. 2020.), u: https://www.academyforlife.va/content/dam/pav/documenti%20pdf/2020/CALL%2028%20febbraio/AI%20Rome%20Call%20x%20firma_DEF_DEF_con%20firme_.pdf (8. I. 2024.); AI Ethics: An Abrahamic Commitment to the Rome Call. Joint Declaration, Rome (10. I. 2023.), u: <https://www.romecall.org/wp-content/uploads/2023/01/AI-Joint-Declaration-5-Jan1-1.pdf>. (14. I. 2024.). Također usp. Rome Call for AI Ethics (28. II. 2020.), u: <https://www.romecall.org/the-call/> (14. I. 2024.).

⁵² Predstavnik Microsofta je kao i 2020. godine bio njegov predsjednik Brad Smith, dok je IBM predstavljao globalni dopredsjednik Dario Gil, a FAO glavni ekonomist te organizacije Maximo Torero Cullen.

⁵³ Usp. Maria Milvia MORCIANO, Le tre fedi abramitiche per un'etica dell'intelligenza artificiale: firmata la Call Rome (10. I. 2023.), u: *Vatican News*, u: <https://www.vaticannews.va/it/vaticano/news/2023-01/etica-dell-intelligenza-artificiale-l-impegno-delle-religioni.html> (6. XII. 2023.).

razvoj služi pravdi i miru svugdje u svijetu«⁵⁴. Mons. Vincenzo Paglia istaknuo je da su provedena istraživanja o jednom od najvažnijih pitanja UI-a našeg vremena pokazala hitnu potrebu za etikom u razvoju i primjeni tehnologije pogotovo s obzirom na to da je napredak u području UI-a brži i širi nego što se očekivalo. Pozvao je one koji dizajniraju UI da implementiraju i rade na njezinu prihvaćanju i stave ljudsko biće u središte svoje i njihove brige.⁵⁵ Crkva time pokazuje ne samo interes za brojna pitanja povezana s razvojem UI-a nego i odgovornost i brigu da napredak »ne isključi nikoga« i da se »diskriminirajuća uporaba tih instrumenata ne ukorijeni na štetu najosjetljivijih i najisključenijih... imajući na umu da način na koji se odnosimo prema posljednjima i najmanje cijenjenima među našom braćom i sestrama pokazuje vrijednost koju prepoznajemo u ljudskom biću«⁵⁶. Naglašavajući važnost razlikovanja između numeričkih vrijednosti s kojima rade algoritmi i moralnih vrijednosti koje su u domeni etike te ističući potrebu za razvojem jezika koji bi mogao prevesti moralne vrijednosti u format prihvatljiv za strojeve, znanstveni direktor Zaklade RenAIssance Paolo Benanti stavlja naglasak na prvenstvo čovjeka, koji je u odnosu prema stroju jedini »pravi poznavatelj i nositelj vrijednosti« te da je u tom kontekstu potrebno zaštititi ljudsko dostojanstvo i ljudska prava.⁵⁷

Čovjek je u središtu i nastupa druge dvojice predstavnika abrahamskih religija rabina Eliezera Simhe Weisza te šeika Abdallaha bin Bayyaha. U svom promišljanju rabin Weisz je istaknuo da židovstvo veliča mudrost čovjeka, stvorenog na Božju sliku i priliku, koja se očituje u ljudskoj inovativnosti općenito, a posebno u umjetnoj inteligenciji. Istodobno ukazuje na to da mora postojati neka kreativna tenzija između ljudske kreativnosti i moralnog autoriteta koja proizlazi iz naše veze sa Stvoriteljem svemira i izvorom svake mudrosti, jer bez nje u opasnosti su jedinstvenost i poziv ljudske osobe stvorene na sliku i priliku Božju te postoji opasnost za egzistenciju čovječanstva u cjelini.⁵⁸

⁵⁴ Papa FRANJO, Audience with participants in the meeting Rome Call for A.I. Ethics.

⁵⁵ Usp. Vincenzo PAGLIA, Morning Welcome. Archbishop Paglia speaks with HE Sheikh Al Mahfoudh Bin Bayyah and Chief Rabbi Eliezer Simha Weisz (10. I. 2023.), u: https://www.romecall.org/wp-content/uploads/2023/01/DISCORSO-MONS-PAGLIA_BENVENUTO_ENG.pdf (14. I. 2024.).

⁵⁶ Papa FRANJO, Audience with participants in the meeting Rome Call for A.I. Ethics.

⁵⁷ Usp. Etica dell'intelligenza artificiale: l'impegno delle religioni abramitiche nella Rome Call (10. I. 2023.), <https://news.microsoft.com/it-it/2023/01/10/etica-dellintelligenza-artificiale-limpegno-delle-religioni-abramitiche-nella-rome-call/> (6. XII. 2023.).

⁵⁸ Usp. Eliezer Simha WEISZ, 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> (6. XII. 2023.).

Sa svoje strane šeik Abdallah bin Bayyah rekao je da religije u svojoj biti nastoje čovjeku osigurati sreću kako na ovom svijetu tako i u onom zagrobnom, potičući stjecanje znanja i razumijevanje svemira. Naglasio je da islamska religija podupire proširivanje znanja koje obuhvaća i religijsku i svjetovnu spoznaju te posebice istaknuo važnost etičkih okvira u znanstvenim i tehnološkim inovacijama kako bi se sačuvalo dostojanstvo i plemenitost čovjeka te zaštitio njegov život.⁵⁹

Svi sudionici Rome Call susreta kao i predstavnici abrahamskih religija istaknuli su u svojim nastupima potrebu za etičkim smjernicama u razvoju i primjeni tehnologije kako bi se zaštitilo ljudsko dostojanstvo i autonomija te značaj stjecanja i primjene znanja, kako religijskog tako i svjetovnog, u unapređenju ljudskog iskustva i društva. Također svi naglašavaju važnost čovjeka kao središnjeg elementa u razmatranju i primjeni tehnologije, a posebno umjetne inteligencije, no dok rabin Weisz i šeik Bin Bayyah naglašavaju religijske aspekte antropologije, mons. Paglia više naginje prema interreligijskom i globalnom humanističkom pristupu s fokusom, u duhu pape Franje, na ljudsko bratstvo. Drugim riječima, iako svi ističu središnji položaj čovjeka u kontekstu tehnološkog razvoja i potrebi za očuvanjem ljudskog dostojanstva, razlike se pojavljuju u shvaćanju uloge čovjeka u svijetu i njegova odnosa prema tehnologiji. Rabin Weisz i šeik Bin Bayyah pristupaju pitanju UI iz specifične religijske perspektive, naglašavajući jedinstvenost svake osobe stvorene na Božju sliku, religijski pogled na znanje i njegovu ulogu u društvu te izražavaju zabrinutost zbog potencijalnih opasnosti nastalih uporabom nove tehnologije. Mons. Paglia sklon je više naglasiti potrebu za integriranjem tehnologije u društvo s poštovanjem prema ljudskim vrijednostima, a papa Franjo promiče univerzalni pristup koji bi uključivao sve religije i kulture.

2.3. *Etičke smjernice i ChatGPT*

U svom govoru od 27. ožujka 2023. godine sudionicima inicijative Minerva Dialogues, koju organizira Dikasterij za kulturu i obrazovanje, a na kojoj se okupljaju znanstvenici, inženjeri, upravitelji poduzeća, pravnici, filozofi, teolozi te predstavnici Crkve, papa Franjo je zahvalio onima koji se zauzimaju da tehnologija bude usredotočena na čovjeka i usmjerena na dobro, izražavajući

⁵⁹ Usp. Abdallah BIN BAYYAH, *Speech Given at The Ethics of Artificial Intelligence: The Commitment of the Abrahamic Religions to the Rome Call* (10. I. 2023.), u: <https://www.romecall.org/wp-content/uploads/2023/01/Bin-Bayyah-AI-Conference-Speech.pdf> (6. XII. 2023.).

uvjerenje da je »dijalog između vjernikâ i nevjernikâ o temeljnim pitanjima etike, znanosti i umjetnosti, te o traženju smisla života, put za izgradnju mira i za cjelovit ljudski razvoj«⁶⁰. Papa upozorava da »ne možemo dopustiti da algoritmi ograničavaju ili uvjetuju poštovanje ljudskog dostojanstva, niti da isključuju suosjećanje, milosrđe, oprost i, ponajviše, otvorenost nadi u promjenu osobe« te je istaknuo da već postoji »suglasnost o razvojnim procesima osmišljenima u skladu s vrijednostima inkluzivnosti (inclusion), transparentnosti (transparency), sigurnosti (security), jednakosti (equity),⁶¹ povjerljivosti (privacy) i pouzdanosti (reliability)«⁶², indirektno se referirajući na načela iz Rome Call for AI Ethics, (transparency, inclusion, accountability, impartiality, reliability, security, privacy),⁶³ no ipak unosi neke manje promijene.⁶⁴

Identična ili slična načela danas susrećemo u etičkim smjernicama različitih IT kompanija kao npr. IBM,⁶⁵ Microsoft⁶⁶, Facebook (Meta)⁶⁷ ili OpenAI,⁶⁸ no referentnom točkom su ipak dva dokumenta Europske komisije iz 2019. godine:⁶⁹ Etičke smjernice za pouzdanu umjetnu inteligenciju⁷⁰ i Definicija umjetne inteligencije: glavne mogućnosti i znanstvene discipline⁷¹. Oni proizlaze iz temeljnih ljudskih prava: poštovanja ljudskog dostojanstva, osobne

⁶⁰ Papa FRANJO, Discorso del Santo Padre Francesco ai partecipanti all'incontro dei »Minerva Dialogues« promosso dal Dicastero per la cultura e l'educazione (27. III. 2023.), u: <https://www.vatican.va/content/francesco/it/speeches/2023/march/documents/20230327-minerva-dialogues.html> (31. XII. 2023.).

⁶¹ U prijevodu na hrvatski jezik na službenoj *web*-stranici Svete Stolice »equity« (tal. l'equità; fr. l'équité; španj. la equidad; polj. bezstronność; ali njem. Gerechtigkeit; ukr. Справедливість).

⁶² Papa FRANJO, Discorso ai partecipanti all'incontro dei »Minerva Dialogues«.

⁶³ Usp. RENAISSANCE FOUNDATION, The six principles, u: <https://www.romecall.org/the-call/> (13. I. 2024.).

⁶⁴ Postoje i razilaženja u hrvatskim prijevodima (inkluzivnost – uključenost; jednakost – pravda). U izvornom tekstu Papa stavlja na prvo mjesto inkluzivnost, dodaje načelo jednakosti te razdvaja načelo sigurnosti i povjerljivosti/privatnosti na dva načela, no zato izostavlja načela odgovornosti (*accountability*) i nepristranosti (*impartiality*).

⁶⁵ Usp. What is AI ethics?, u: <https://www.ibm.com/topics/ai-ethics> (3. I. 2024.).

⁶⁶ Usp. Microsoft Responsible AI principles in practice, u: <https://www.microsoft.com/en-gb/ai/responsible-ai?activetab=pivot1%3aprimar6> (3. I. 2024.).

⁶⁷ Usp. Facebook's five pillars of Responsible AI, u: <https://ai.meta.com/blog/facebook-five-pillars-of-responsible-ai/> (3. I. 2024.).

⁶⁸ Usp. OpenAI Supplier Code of Conduct (10. X. 2023.), u: <https://openai.com/policies/supplier-code> (14. I. 2024.); Developing safe & responsible AI, u: <https://openai.com/safety> (14. I. 2024.).

⁶⁹ Usp. Taras BARŠČEVSKI, *Aktualne teme biblijske teologije*, Zagreb, 2023., 38–39.

⁷⁰ AI HLEG, *Etičke smjernice za pouzdanu umjetnu inteligenciju* (8. IV. 2019.), Bruxelles, 2019., u: <https://op.europa.eu/s/pGax> (6. XII. 2023.).

⁷¹ AI HLEG, *Definicija umjetne inteligencije: glavne mogućnosti i znanstvene discipline* (8. IV. 2019.), Bruxelles, 2019., u: <https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines> (6. XII. 2023.).

slobode, demokracije, pravde, vladavine prava, ravnopravnosti, nediskriminacije, solidarnosti i prava građana.⁷² Možemo ih sažeti u četiri etička načela: 1) poštovanje ljudske autonomije; 2) sprečavanje nastanka štete; 3) pravednosti i 4) objašnjivosti⁷³, koji se konkretiziraju u sedam zahtjeva koje bi sustavi umjetne inteligencije trebali ispuniti kako bismo ih mogli smatrati pouzdanima: 1) ljudsko djelovanje i nadzor, 2) tehnička otpornost i sigurnost, 3) privatnost i upravljanje podacima; 4) transparentnost; 5) raznolikost, nediskriminacija i pravednost; 6) dobrobit za okoliš i društvo; 7) odgovornost.⁷⁴

Ova dosta visoka razina suglasnosti u odnosu na etičke smjernice temeljene uglavnom na antropocentričnim postulatima između toliko različitih subjekata govori nam – što papa Franjo ocjenjuje veoma pozitivno – da je moguće naći suglasnost o etičkim smjernicama te surađivati na dobrobit društva i svakog pojedinca, ne zanemarujući i održivi razvoj Zemlje. S druge strane konsenzus glede etičkih smjernica ukazuje da je riječ upravo samo o dogovoru koji ovisi od vlastitih interesa zainteresiranih subjekata te da nije moguće nekog prisiliti na pridržavanje etičkih smjernica bez njegove suglasnosti.

Puštanje u rad 30. studenoga 2022. godine prototipa sustava umjetne inteligencije ChatGPT-a, koji je u samo dva mjeseca dostigao 100 milijuna aktivnih korisnika i tako postao aplikacija s najbržim rastom u povijesti interneta,⁷⁵ razotkrilo je još više važnost i urgentnost etičnosti i odgovornosti korištenja umjetne inteligencije,⁷⁶ ali i konačnost jasne regulacije i pravila za umjetnu inteligenciju,⁷⁷ koji bi bili obvezujući za sve, od naručitelja i investitora, dizajnera i programera, pa sve do dobavljača i korisnika. Upravo brz napredak UI-a

⁷² Usp. *Povelja Europske unije o temeljnim pravima*, (7. VI. 2016.), Bruxelles, 2016., u: <https://eur-lex.europa.eu/legal-content/HR/ALL/?uri=CELEX:12016P/TXT> (15. VIII. 2021.). Ima se ipak osjećaj da *Etičke smjernice za pouzdanu umjetnu inteligenciju* pristupaju selektivno izboru temeljnih ljudskih prava, preferirajući jedne te izostavljajući druge kao npr. »pravo na život« (čl. 2 §1), »zabranu ropstva i prisilnog rada« (čl. 5) te »pravo na stupanje u brak i pravo na osnivanje obitelji« (čl. 9), kao i »slobodu mišljenja, savjesti i vjeroispovijedi« (čl. 10). Usp. Taras BARŠČEVSKI, *Aktualne teme biblijske teologije*, 43–45.

⁷³ Usp. AI HLEG, *Etičke smjernice za pouzdanu umjetnu inteligenciju*, br. 48.

⁷⁴ Usp. *Isto*, br. 58.

⁷⁵ Usp. Lorenzo TIROTTA, *ChatGPT è l'app con la crescita più veloce di sempre! Raggiunti 100 milioni di utenti* (3. II. 2023.), u: https://www.hwupgrade.it/news/web/chatgpt-e-l-app-con-la-crescita-piu-veloce-di-sempre-raggiunti-100-milioni-di-utenti_113840.html (1. I. 2024.).

⁷⁶ Usp. Borivoje DOKLER, *ChatGPT umjetnu je inteligenciju prezentirao na dosad neviđen način* (12. VI. 2023.), u: <https://www.poslovni.hr/sci-tech/chatgpt-umjetnu-je-inteligenciju-presentirao-na-dosad-neviden-nacin-4393922> (1. I. 2024.).

⁷⁷ Usp. *Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI* (9. XII. 2023.), u: <https://www.europarl.europa.eu/news/hr/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai> (2. I. 2024.).

te mogućnosti ChatGPT-a, njegov širok spektar primjene, dostupnost, interaktivnost i prilagodljivost postali su i uzrok još veće zabrinutosti u vezi s mogućom zlorabom tehnologije.

Jasno je da su potrebni neki regulatorni okviri za umjetnu inteligenciju u sklopu kojih bi bili analizirani i klasificirani sustavi umjetne inteligencije prema riziku koji predstavljaju za korisnike, a prekršitelje čekale bi odgovarajuće sankcije. Kako bi se osiguralo da sustavi umjetne inteligencije koji se stavljaju na europsko tržište i upotrebljavaju u EU-u budu sigurni i da se njima poštuju temeljna prava i vrijednosti EU-a, Vijeće EU-a i Europski parlament postigli su 9. prosinca 2023. godine privremeni dogovor o prijedlogu usklađenih pravila o umjetnoj inteligenciji, takozvani Akt o umjetnoj inteligenciji.⁷⁸ Riječ je o prvom zakonodavnom prijedlogu te vrste u svijetu, tako da on može postati svjetski standard za regulaciju umjetne inteligencije.⁷⁹

Papa Franjo je dosad najiscrpnije progovorio o umjetnoj inteligencije u svojoj poruci od 8. prosinca 2023. godine napisanoj u povodu obilježavanja 57. Svjetskog dana mira 1. siječnja 2024. godine pod naslovom »Umjetna inteligencija i mir«. Napominjući da bi »napredak znanosti i tehnologije, u mjeri u kojoj pridonosi boljem uređenju ljudskog društva, rastu i razvoju slobode i bratskog zajedništva, trebao voditi čovjekovu boljitku i preobrazbi svijeta«, papa Franjo podsjeća da je »Bog ljudima dao svoga Duha kako bi ih resila 'umješnost, sposobnost i razumijevanje u svim poslovima' (Izl 35,31)« te da je ljudska »inteligencija izraz dostojanstva kojim nas je obdario Stvoritelj, koji nas je stvorio na svoju sliku i priliku (usp. Post 1,26) i osposobio nas da na njegovu ljubav odgovorimo slobodom i spoznajom«⁸⁰. No uz zadivljujuće mogućnosti u novim informacijskim tehnologijama značajni napredci donose također i ozbiljne opasnosti s ozbiljnim implikacijama, posebno u smislu socijalne pravde i harmonije među narodima. Zato razvoj umjetne inteligencije mora biti popraćen odgovarajućom odgovornošću jer »tehnološki razvoji koji ne vode do poboljšanja kvalitete života cijelog čovječanstva, nego napro-

⁷⁸ COUNCIL OF THE EU, Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world, (9. XII. 2023.), u: <https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/> (2. I. 2024.).

⁷⁹ Usp. VIJEĆE EU-a, Akt o umjetnoj inteligenciji: Vijeće i Parlament postigli dogovor o prvim pravilima za umjetnu inteligenciju u svijetu (9. XII. 2023.), u: <https://www.consilium.europa.eu/hr/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/> (2. I. 2024.).

⁸⁰ PAPA FRANJO, *Poruka Svetoga oca Franje za 57. svjetki dan mira (1. siječnja 2024.) »Umjetna inteligencija i mir«* (8. XII. 2023.), br. 1, u: <https://ika.hkm.hr/dokumenti/papina-poruka-za-57-svjetski-dan-mira/> (1. I. 2024.).

tiv pogoršavaju nejednakosti i sukobe, nikako se ne mogu smatrati pravim napretkom⁸¹. Kako bi umjetna inteligencija bila u službi najboljih ljudskih potencijala i naših najviših ciljeva, potrebno je da u njezinu razvoju budu ispoštovane »temeljne ljudske vrijednote kao što su inkluzivnost, transparentnost, sigurnost, pravda, povjerljivost i pouzdanost«⁸². S obzirom na to da nije dovoljno pretpostaviti »da će se oni koji razvijaju algoritme i digitalne tehnologije ponašati etično i odgovorno, potrebno je ojačati ili, ako je potrebno, uspostaviti tijela zadužena za razmatranje novih etičkih pitanja i zaštitu pravâ onih koji koriste neki od oblika umjetne inteligencije ili su izloženi njihovu utjecaju«⁸³. Osim toga potrebno je imati i »osjećaj ograničenja«, jer kada »čovjek, koji je po definiciji smrtnan, misli da tehnologijom može nadići svaku ograničenost, u opasnosti je da, opsjednut željom da sve kontrolira i u potrazi za apsolutnom slobodom, izgubi kontrolu nad samim sobom i upadne u spiralu tehnološke diktature«⁸⁴. »Ne smije se«, nastavlja Papa, »dopustiti da algoritmi diktiraju način na koji shvaćamo ljudska prava, da se ostavi po strani temeljne vrijednote suosjećanja, milosrđa i oprostjenja ili ukloni mogućnost da se pojedinac može promijeniti i ostaviti prošlost iza sebe.«⁸⁵ Umjetna inteligencija može biti koristan alat za uvođenje važnih inovacija u poljoprivredi, obrazovanju i kulturi, tj. za promicanje cjelovitog ljudskog razvoja u svijetu, uključujući najmanje i najslabiju braću i sestre koji su u najvećoj potrebi.⁸⁶ Više nego jasno je da takva etika ne može biti kompatibilna s upotrebom te tehnologije u ratovima i zato jamčenje odgovarajućeg, smislenog i dosljednog ljudskog nadzora nad oružanim sustavima predstavlja imperativ.⁸⁷ Od ključne je važnosti da vlade i međunarodne organizacije reguliraju upotrebu umjetne inteligencije obvezujućim zakonima, uzimajući u obzir da je u potrazi za normativnim propisima koji mogu pružiti etičke smjernice onima koji rade na razvijanju digitalnih tehnologija, »bitno je prepoznati i utvrditi ljudske vrijednosti, dok u radu na izradi etičkih smjernica za proizvodnju oblikâ umjet-

⁸¹ Papa FRANJO, *Umjetna inteligencija i mir*, br. 2. Također usp. LS 194; papa FRANJO, *Discorso ai partecipanti al Seminario »Il bene comune nell'era digitale«* (27. rujna 2019.), u: https://www.vatican.va/content/francesco/it/speeches/2019/september/documents/papa-francesco_20190927_eradigitale.html (14. I. 2024.).

⁸² Papa FRANJO, *Umjetna inteligencija i mir*, br. 2; Papa FRANJO, *Discorso ai partecipanti all'incontro dei »Minerva Dialogues«*.

⁸³ *Isto*.

⁸⁴ Papa FRANJO, *Umjetna inteligencija i mir*, br. 4.

⁸⁵ *Isto*, br. 5.

⁸⁶ Usp. *Isto*, br. 6.

⁸⁷ Usp. *Isto*.

ne inteligencije ne može se zanemariti dublja pitanja o smislu ljudskog postojanja, zaštiti temeljnih ljudskih prava i težnji za postizanjem pravde i mira⁸⁸.

3. Ograničenja etičkih smjernica

Sve češće upućivanje na umjetnu inteligenciju u crkvenim dokumentima ukazuje na to da je Sveta Stolica prepoznala važnost i značenje umjetne inteligencije za napredak i bolje uređenje ljudskog društva, kao i odgovornosti svih vjernika za razvoj tako moćnog instrumenta od upotrebe kojega može ovisiti budućnost čovjeka ili Zemlje. Inicijativa Rome Call for AI Ethics, koju su osim predstavnika tehnoloških tvrtki koje se bave razvojem umjetne inteligencije potpisali i predstavnici abrahamskih religija, kao i poruka svetog oca Franje na 57. Svjetski dan mira upućena »narodu Božjem, narodima, čelnicima država i vlada, predstavnicima različitih vjera i civilnog društva, te svim muškarcima i ženama našeg vremena«⁸⁹ žele doprijeti do svih ljudi s obzirom na to da umjetna inteligencija dotiče danas čitavu ljudsku obitelj. Svi su svjesni da tehnološki napredak može učiniti ovaj svijet boljim u kojem će biti više solidarnosti, pravde i mira,⁹⁰ no isto tako može dovesti ne samo do povećana nejednakosti i nepravde među ljudima, nego i do novih ratova i sukoba u svijetu, pa čak i zaprijetiti opstanku čitavog čovječanstva. Drugim riječima, cilj regulacije razvoja i korištenja umjetne inteligencije u njezinim različitim oblicima nije samo sprečavanje štetnih praksa nego i poticanje dobrih praksa stvaranjem novih i kreativnih pristupa i olakšavanjem pojedinačnih i grupnih inicijativa.⁹¹ No bez obzira na donošenje etičkih smjernica za umjetnu inteligenciju na različitim razinama, pa čak i prvog regulatornog okvira za umjetnu inteligenciju u obliku Akta o umjetnoj inteligenciji Vijeća EU i Europskog parlamenta, konkretna primjena i najnoviji razvoji umjetne inteligencije pokazuju ne samo poteškoće u primjeni etičkih smjernica nego često i njihovu »suvišnost«. Drugim riječima, jedanput pušten u rad neki softver ili aplikacija ili robot više »ne sluša« etičke smjernice korištene u njegovoj izradi, nego samo algoritme koji mogu pružiti statističke aproksimacije s obzirom na to da se ne može sve predvidjeti i sve izračunati te da će uvijek postojati nedostupan ostatak koji izmiče svim pokušajima mjerenja.⁹² Kao primjer možemo navesti samo neka pitanja nastala s puštanjem u rad ChatGPT-a.

⁸⁸ *Isto*, br. 8.

⁸⁹ *Isto*, Uvod.

⁹⁰ Usp., *Isto*, Zaključak.

⁹¹ Usp. *Isto*, br. 8. Također usp. LS 177.

⁹² Usp. Papa FRANJO, *Umjetna inteligencija i mir*, br. 4.

ChatGPT je učinio umjetnu inteligenciju dostupnom širokim masama, tako da je mnogo više korisnika dobilo pristup novim tehnologijama koji nam mogu pomoći u svakodnevnom životu i olakšati mnoge poslove,⁹³ no omogućio je također dotad običnim korisnicima već gotovih tehnoloških proizvoda postati aktivnim tvorcima novih i originalnih sadržaja, ali, po vlastitoj želji ili igrom slučaja, i zlonamjernim akterima na internetu u stvaranju uvjerljivijih prevara i dezinformacija, izlažući riziku sigurnost i privatnost ostalih korisnika. To nam govori o slabostima i ograničenjima ChatGPT-a koji, bez obzira na to da je veoma moćan alat u stalnom samousavršavanju, može ponekad pružiti netočne informacije ili generirati dvosmislenosti, preopćenito interpretirati informacije ili ih pojednostaviti te dati pomalo pristrane odgovore. To da može ne razumjeti neko pitanje, da nije sposoban za kritičko razmišljanje tako da ne može procijeniti informacije temeljem moralnih ili etičkih vrijednosti te nema empatiju i ne može prepoznati emocionalne potrebe ili osjećaje korisnika, što znači da nije prikladan za rad ni s osjetljivim temama ni s pitanjima koja zahtijevaju dublje analize ili moralne odluke,⁹⁴ ne treba ni spominjati.

Možemo ipak pretpostaviti da je ChatGPT stvoren pridržavajući se temeljnih etičkih načela za umjetnu inteligenciju, što ipak ne jamči njegovo besprijekorno funkcioniranje, niti ga štiti od zlorababa. Osim toga koliko i na koga je moguće primijeniti regulatorne okvire te kazniti prekršitelje, a u slučaju nanesene štete je li moguća reparacija? Već je bilo rečeno da etičke smjernice proizlaze iz temeljnih ljudskih prava, kao npr. ljudskog dostojanstva, osobne slobode, pravde, ravnopravnosti itd., no tu postajemo svjesni da postoje različita tumačenja i razumijevanja istih tih temeljnih ljudskih prava, kao i toga na čemu se osnivaju ta ljudska prava? Drugim riječima, prihvaćajući ista etička načela i temeljna ljudska prava iz kojih oni proizlaze, nemamo ista shvaćanja o čovjeku, o njegovoj naravi, dostojanstvu i slobodi?

Na nepostojanje konsenzusa glede nekih temeljnih pojmova ukazao je Gianfranco Ravasi, istaknuvši da nekad prihvaćen pojam ljudske naravi prema kojem *agere sequitor esse* kao i istine koja nam prethodi i koja nas nadilazi, danas su često odbačeni i relativizirani, tako da čovjek smatra da može sam odlučivati tko će biti,⁹⁵ dok je danas istina fluidna i subjektivna te je svako sam

⁹³ Usp. Borivoje DOKLER, ChatGPT umjetnu je inteligenciju prezentirao na dosad neviđen način (12. VI. 2023.), u: <https://www.poslovni.hr/sci-tech/chatgpt-umjetnu-je-inteligenciju-prezentirao-na-dosad-neviden-nacin-4393922> (3. I. 2024.).

⁹⁴ Usp. Umjetna inteligencija ChatGPT – prednosti i nedostaci (7. XI. 2023.), u: <https://www.fer-projekt.com/hr/blog-post/97/umjetna-inteligencija-chatgpt-prednosti-i-nedostaci> (4. I. 2024.).

⁹⁵ Tu u diskurs ulazi i pitanje *gendera*.

konstruira, za razliku od Platona za kojeg je istina bila kao »polje« prema kojoj se kreće dvopreg duše koju vuku dva konja, jedan smrtni, a drugi besmrtni, jedan naših osjetila, a drugi uma, kako bi jedanput, pristigavši u to »polje istine« počeli otkrivati nove krajolike (Platon, Fedar). Prema Ravasiju zadaća sinodalne Crkve je upravo ući zajedno kao Crkva u svijet, u izazove koji su toliko važni i u promjene paradigma, koje su toliko velike da ne mogu biti riješene unutar jedne partikularne zajednice, a tim manje pojedinaca.⁹⁶ Potrebno je izaći – kao što ponavlja često papa Franjo – iz sakristije, no treba izaći na današnje trgove, današnje areopage, gdje se svijet okuplja »na pripovijedanje i slušanje novosti« (usp. Dj 17,21) i to ne samo kako bi čuli o problemima, strahovima ili dilemama svijeta, nego kako bi mu navijestili novost Boga koji je postao čovjekom, da nas učini bogovima⁹⁷ i usrksnuo kako bi nas uveo u vječni život. Reakcija Atenjana na Pavlove riječi o Kristovu uskrsnuću svima nama je dobro poznata: »Još ćemo te o tom slušati« (Dj 17,32), što nam daje jasno na znanje da kršćanska poruka neće uvijek biti prihvaćena, zato što ne dolazi od svijeta i onih koji iz svijeta govore (usp. 1 Iv 4,5).

Zaključak

Etičke smjernica za umjetnu inteligenciju kao i svi pokušaji odrediti pravila njezina razvoja proizlaze iz spoznaje o mogućnostima napretka umjetne inteligencije, koje moraju biti kontrolirane, usmjerene, limitirane, inače čovjekova postignuća u sferi razvoja umjetne inteligencije mogu postati prijetnja za njegov život i opstanak na Zemlji. Sva nastojanja oko stvaranja etičkih smjernica i kasnijih pokušaja regulirati njihovu primjenu zaslužuju naše najviše priznanje. S druge strane postoje i ograničenja etičkih smjernica koje su uvelike izašle na vidjelo s puštanjem u rad ChatGPT-a. Izgleda da nema temeljne ljudske vrednote u razvoju umjetne inteligencije koja bi bila ispoštovana besprijekorno. Što znači, npr. transparentnost ChatGPT-a prema kojoj UI bi trebao biti razumljiv svima, kada je malo kome jasno na koji način ChatGPT uopće formulira svoje konkretne odgovore, što ga čini u potpunosti nepredvidljivim? Što znači nepristranost s obzirom na to da se ChatGPT prilagođava svom sugovorniku i nudi mu prihvatljive odgovore, izbjegavajući ukazati na očigled-

⁹⁶ Usp. Gianfranco RAVASI, *Una cultura della Sinodalità nella Chiesa* (video) (23. X. 2020.), u: https://www.youtube.com/watch?v=rVs0ynTnl1g&ab_channel=DiscerningLeadership (4. I. 2024.).

⁹⁷ Usp. Atanazije ALEKSANDRIJSKI, *De Incarnatione*, 54,3, u: Jacques Paul MIGNE (ur.), *Patrologiae cursus completus. Series Graeca*, Paris, 1857-1866., 25, 192B.

ne greške u mišljenju i zaključcima svog sugovornika? Što znači pouzdanost kada je veoma teško doći do izvora informacije iz kojeg UI iznosi činjenice ili na kojima temelji svoje zaključke? I nije riječ o lošem programiranju ili krivim ciljevima, nego upravo o odsutnosti ljudskih vrлина kao što su ljubav, vjera i nada, koje čine čovjeka čovjekom koji voli i brine se, vjeruje i suosjeća, razumije i sanja.

Čovjek je stvoren za drugog, 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.⁹⁸ Etičke smjernice ne govore o Bogu. One, nasuprot, naglašavaju antropocentrizam, koji stavlja u središte svih odnosa isključivo ljudsko biće koje pripisuje sebi pravo određivati temeljna ljudska prava i definirati etička načela. Čovjek se postavio u središte stvaranja i nameće svoje želje i potrebe kao najviše pravilo. Izgubio je ne samo osjećaj za transcendenciju nego i potrebu za transcendencijom jer se čovjek smatra mjerilom svega o čemu ovisi ne samo njegova osobna egzistencija nego i opstanak svih ostalih živih bića te i same Zemlje. Čovjek je odbacio Boga i svoj identitet slike i prilike Boga (usp. Post 1,26), ali time i svoju svrhu i ulogu u ovom svijetu. Kršeći zapovijed ne jesti sa stabla spoznaje dobra i zla (usp. Post 2,17), čovjek odbacuje bilo kakvu transcendentnu definiciju koja bi dolazila od Boga želeći sam odlučivati što je dobro, a što zlo. Bog je dao čovjeku definiciju dobra i zla u svrhu očuvanja njegova života u harmoniji Edenskog vrta, no čovjeku stablo izgleda dobro i zamamljivo, a posebice ga privlači želja da bude kao bogovi koji razlučuju dobro i zlo (usp. Post 3,5). Prevaren od zmije, ubravši i pojevši plod sa stabla spoznaje dobra i zla, čovjek neće postati bog, nego će zapravo otkriti vlastitu golotinju – svoju ograničenost kao stvorenja – koju će zbog osjećaja stida jedno pred drugim pokušati prekriti smokvinim lišćem (usp. Post 3,7,13).

Teško je ne primijetiti u etičkim smjernicama tu želju čovjeka da sam odlučuje što je dobro, a što zlo, da sam definira temelje i etička načela koji bi mu jamčili blagostanje i sigurnost, kao da bi mu najveća prijetnja za život bio sam Bog koji ga je stvorio. Možda je upravo zato i najveći izazov koji pred čovjekom stavlja umjetna inteligencija ne toliko stvaranje etičkih smjernica i donošenje regulatornih okvira za umjetnu inteligenciju, nego ponovno otkrivanje sebe kao Božje slike. Stvarajući i razvijajući umjetnu inteligenciju, čovjek zapravo reflektira na nju sliku koju ima o sebi i svojoj ulozi u svijetu, koja nikada ne može biti potpuna, ako se izostavi čovjekova transcendentalna dimenzija te njegov odnos s Bogom.

⁹⁸ Usp. Aurelije AUGUSTIN, *Ispovijesti* (1,1), Zagreb, 1973., 7.

Abstract

The Church in the Face of Ethical Challenges of Artificial Intelligence

Taras BARŠČEVSKI

Catholic Faculty of Theology, University of Zagreb
Vlaška 38, p.p. 5, HR – 10 000 Zagreb
taras.barscevski@kbf.unizg.hr

The rapid development of artificial intelligence (AI), penetrating nearly every aspect of human life, from culture and arts to education, healthcare, social issues, economics, politics, and daily life, presents new challenges and opportunities. However, it also carries significant risks with far-reaching consequences. Beyond facilitating daily activities and enhancing the quality of life, AI could pose threats to privacy, mental health, and family relationships, exacerbate social and economic inequalities, and potentially be misused in surveillance and military applications, thereby leading to tragic outcomes for the humanity and the planet. In its role as Mother and Teacher, the Church cannot remain indifferent when a new vision of humanity is being forged and the future of Earth and mankind is shaped. Recognising her role in the rapidly changing world of technology and AI, the Church underscores the responsible use of technology aligned with the Divine will and for the benefit of all humanity, particularly those in the greatest need and the most impoverished. The paper showcases the scenarios of AI development and the importance of ethical guidelines for all participants in this process, and how Church teachings from Saint John Paul II, Pope Benedict XVI, to Pope Francis, as well as various Vatican institutions, respond to the new ethical challenges posed before the humanity by AI.

Key words: *artificial intelligence, AI, ChatGPT, ethical guidelines, church teaching, algorithms, Rome Call.*