

UDK

004.8+004.896]:008

004.8+004.896]:179

<https://doi.org/10.53745/bs.94.1.4>

Primljeno: 6. 12. 2023.

Izlaganje sa skupa

KAKO (PRE)ŽIVJETI S PAMETNIM STROJEVIMA*

Bojan JERBIĆ

Fakultet strojarstva i brodogradnje Sveučilišta u Zagrebu
Ivana Lučića 5, 10 000 Zagreb
bojan.jerbic@fsb.hr

Sažetak

Robotika i umjetna inteligencija (UI) neupitno poboljšavaju ljudske potencijale. Djeluju u rasponu od jednostavnog zaključivanja, izvođenja jednostavnih radnji do imitacije složenih ljudskih kognitivnih i radnih funkcija. Njihovi zadatci mogu varirati od zabave ili igre do pomaganja ljudima u teškim ili zamornim zadaćama. Kako bismo razumjeli utjecaj umjetne inteligencije, važno je učiti iz prošlih uspjeha i neuspjeha, kao i predvidjeti njezine buduće smjerove i potencijalne pravne, etičke i socioekonomske implikacije. U bliskoj će budućnosti inteligentni roboti sve više postajati sastavni dio života običnih ljudi. Potencijalne koristi su velike, ali tu su i značajni rizici i etički izazovi. Naša autonomija može biti ugrožena, a društvena interakcija opstruirana. Proširena uporaba robota može dovesti do smanjenog kontakta među ljudima i mogućih ograničenja osobnih sloboda. Pametni strojevi oblikuju radikalno nov svijet, dovodeći do značajnih ekonomskih i kulturalnih promjena, stvarajući i pobjednike i gubitnike na globalnoj razini. Umjetna inteligencija razvija se uzimajući čovjeka kao uzor. Unatoč tomu, umjetna inteligencija i ljudska bića ostaju fundamentalno različiti zbog svoje različite prirode. Možemo li onda dijeliti iste moralne koncepte? Iako postoji zabrinutost oko toga što robotika i umjetna inteligencija znače za našu budućnost, postoji i vjerovanje koje dolazi iz znanstvene ili pak filozofijske perspektive, da robotika i UI predstavljaju dio te evolucije. Napredujemo usput redefinirajući čovječanstvo i neprestano preispitujući što nas kao ljude čini jedinstvenima.

Ključne riječi: robotika, umjetna inteligencija, roboetika.

* Izlaganje sa skupa prerađena je i skraćena verzija članka: Bojan JERBIĆ – Marko ŠVAČO, Artificial Intelligence and Robotics as the Driving Power of Modern Society, u: *Tehničke znanosti*, 554 (2023.) 22, 1-55.

Uvod

Napredak u digitalnim tehnologijama i srodnim područjima se ubrzava i širi na sve ljudske djelatnosti. Na raskrižju znanosti, tehnologije i metafizičkih uvjerenja ubrzani razvoj nas potiče da istražimo dublje značenje i svrhu naše potrage za poboljšanjem ljudskog društva. Već duže vrijeme pitamo se na koji način tehnološki razvoj oblikuje našu budućnost i razumijemo li budućnost koju stvaramo. Iskustvo nas uči kako nerazumijevanje razvojnih procesa uzrokuje gubitak prilika i pozicija u društvenom i ekonomskom kontekstu. Sve je to još više naglašeno zbog brzine kojom se događaju promjene u znanosti i onda posljedično tomu i u cijelom društvu, što ne ostavlja prostor za popravak ili dugotrajno prilagođavanje. Čini se da je ključno razumjeti budućnost ne samo u tehničkom smislu nego i u filozofijskoj konotaciji te znati kako će se budućnost razvijati i u skladu s tim kako će se mijenjati naše društvo, kultura i ekonomija.¹ Iako svjedočimo napretku u brojnim znanstveno-tehnološkim disciplinama, čini se da razvoj robotike i umjetne inteligencije (UI) ponajviše intrigira, iako su to tehnologije s poviješću od prije šezdeset godina. To je zato što se spomenute tehnologije izravno upliću u naš antropološki smisao i u svojoj osnovi otvaraju kozmičko pitanje našeg postojanja. Čovjek je oduvijek na sebe preuzimao odgovornost za donošenje odluka i njihovo provođenje. Danas se sve češće prepuštamo strojevima pa se postavlja legitimno pitanje gubimo li kontrolu nad svojom budućnošću.

Znanstveno-tehnološku revoluciju kojoj svjedočimo ponajprije karakterizira transformacija digitalne tehnologije iz tipično virtualne domene u fizičku. Najbolji primjer te digitalne transformacije je robot, jer se može shvatiti kao računalo s rukama koje informacije i/ili znanje pretvara u (fizičku) radnju/djelovanje. Robot je ono što možemo tumačiti kao »kibernetičko-fizički sustav«, odnosno fizičku metamorfozu digitalne (izvorno informacijske) tehnologije. Zbog složenosti interakcije robotske mehanike s okolnim svijetom ili s čovjekom neizbježna je primjena metoda upravljanja koje se oslanjaju na elemente inteligencije. Umjetna inteligencija je stoga često korišten termin koji se odnosi na područje znanosti koje ima ambiciju omogućiti strojevima (robotima) obavljanje naprednih funkcija kao što su percepcija, učenje, planiranje, rješavanje problema i donošenje odluka. Zato su robotika i umjetna inteligencija neraskidivo isprepletene i često impliciraju jedna drugu. Podrazumijevaju tehnički sustav koji može osjetiti stanje

¹ Usp. Michael D. LANGFORD, *A Theological Framework for Reflection on Artificial Intelligence*, u: *SPU Works*, 171 (2022.), u: <https://digitalcommons.spu.edu/works/171> (25. X. 2023.).

okoline, prikupljati informacije iz različitih izvora, razmišljati, učiti i djelovati u smislu programiranih ciljeva. U sinergiji s drugim tehnologijama umjetna inteligencija značajno ubrzava znanstveni i tehnološki razvoj i napredak ljudskog društva. Stalno se otvaraju nove mogućnosti primjene povezanih tehnologija, kako u industriji, zdravstvu tako i u svakodnevnom životu. Sve to počiva na ideji da se čovječanstvo može razvijati i poboljšavati kroz tehnološki napredak. Težnja za ljudskim poboljšanjem potiče pitanja o prirodi našeg postojanja kao i etičkim razmatranjima koja se javljaju kada se ljudi aktivno upliću u vlastitu biologiju i svijest. Možda najveća važnost umjetne inteligencije leži u njezinu utjecaju na širenje našeg znanja o sebi, posebno u područjima kognitivnih znanosti i neuroznanosti. Može se, dakle, reći kako se neuroznanost i umjetna inteligencija razvijaju komplementarno. Unatoč tomu, imamo malo stvarnog znanja o tome kako ljudski mozak stvarno funkcionira. Stoga neuroznanost ima važnu ulogu u razvoju umjetne inteligencije. Nove računalne tehnologije omogućavaju proučavanje mozga pomoću računalnih modela i simulacija, pridonoseći otkrivanju novih teorija o kognitivnim procesima. U isto vrijeme postavljaju se brojna pitanja povezana sa sigurnošću primjene, ugrožavanjem etičkih standarda ili pak posthumanističkim i distopijskim scenarijima.

Najveća opasnost tehnološkog razvoja leži u potencijalno negativnom utjecaju strojeva na ljudsku psihologiju. Još uvijek ne znamo dovoljno o tome kako će se razvijati suživot ljudi i inteligentnih strojeva. Koje su psihološke implikacije umjetne inteligencije na naš vlastiti društveni i emocionalni razvoj? Naprimjer, kako će se naša samopercepcija, naš osjećaj tko smo, promijeniti s pojavom strojeva s ljudskim kvalitetama i sposobnostima? U skoroj budućnosti možemo očekivati da će inteligentni strojevi nadmašiti ljudsku inteligenciju, ili barem određene kompleksne vještine, kod većine populacije. Hoće li to onda proizvesti našu nelagodu ili strah u odnosu na pametne strojeve? Stoga se razvoj pametnih strojeva mora temeljiti na interdisciplinarnim pristupima koji će uzeti u obzir ljudsku psihologiju kao polazište kojem se strojevi trebaju prilagoditi, a ne obrnuto.

1. Izazovi i socioekonomski utjecaj

Umjetna inteligencija službeno je priznata kao akademsko područje tijekom konferencije u Dartmouthu 1956. godine², a pet godina kasnije George Devol

² Usp. John McCARTHY – Marvin MINSKY – Nathaniel ROCHESTER – Claude E. SHANNON, A Proposal for the Dartmouth Summer Research Project on Artificial In-

patentirao je prvi industrijski robot nazvan Unimate.³ Pojam umjetna inteligencija usko je povezan s područjem tzv. simboličke umjetne inteligencije, koja je bila raširena do kasnih 1980-ih.⁴ Kako bi se prevladala neka od ograničenja simboličke umjetne inteligencije, »subsimboličke« (ili nesimboličke) metodologije, kao što su neuronske mreže, neizraziti sustavi, evolucijski računalni modeli i drugi računalni modeli slični mozgu počele su dobivati na popularnosti, što je dovelo do pojave tzv. računalne inteligencije kao dijela umjetne inteligencije.⁵

Prava i potpuna umjetna inteligencija još ne postoji. Na toj visokoj razini umjetna inteligencija trebala bi oponašati ljudsku kogniciju do te mjere da strojevima omogućiti da sanjaju, razmišljaju, osjećaju, doživljavaju emocije te oblikuju vlastite namjere i ponašanja.⁶ Iako još uvijek nema dokaza da bi ta vrsta prave umjetne inteligencije mogla postojati prije 2050. godine, načela računalne znanosti koja pokreću umjetnu inteligenciju brzo napreduju i važno je procijeniti njezin učinak, ne samo s tehnološkog nego i s društvenog aspekta, uključujući etičke i pravne perspektive.

Temeljna ideja razvoja umjetne inteligencije je stvaranje autonomnih tehničkih sustava. Autonomija tehničkog sustava primarno znači isključivanje čovjeka ili njegove pomoći u procesu primjene. Autonomni tehnički sustav, odnosno inteligentni robotski sustav, ne ovisi o vanjskim čimbenicima i može raditi bez umora dvadeset i četiri sata dnevno. Općenito, njegovo djelovanje ne može se ometati emocionalnim utjecajem, umorom ili drugim vanjskim utjecajima na koje su tipično osjetljivi ljudi. Međutim, upravo autonomija umjetne inteligencije sustava izaziva strahove povezane s gubitkom ljudske kontrole nad rezultatima vlastita tehnološkog razvoja. Već danas ljudi nekritički vjeruju u računalne algoritme. Što se tiče čovjeka i stroja, ljudi obično procjenjuju da je stroj manje pristran. Međutim, kada se uzme u obzir da se većina UI-aplika-

telligence (31. VIII. 1955.), u: <http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf>. (25. X. 2023.).

³ Usp. Bob MALONE – George DEVOL: A Life Devoted to Invention, and Robots (26. IX. 2011.), u: <https://spectrum.ieee.org/george-devol-a-life-devoted-to-invention-and-robots> (25. X. 2023.).

⁴ Usp. Paul SMOLENSKY, Connectionist AI, symbolic AI, and the brain, u: *Artificial Intelligence Review*, 1 (1987), 95-109.

⁵ Usp. Pinar ÖZTÜRK – Axel TIDEMANN, A review of case-based reasoning in cognition–action continuum: a step toward bridging symbolic and non-symbolic artificial intelligence, u: *The Knowledge Engineering Review*, 20 (2014.) 1, 51-77.

⁶ Usp. Geng TU – Jintao WEN – Cheng LIU – Dazhi JIANG – Erik CAMBRIA, Context-and Sentiment-Aware Networks for Emotion Recognition in Conversation, u: *IEEE Transactions on Artificial Intelligence*, 3 (2022.) 5, 699-708.

cija temelji na statističkim i klasifikacijskim metodama, legitimno je pitanje u kojoj je mjeri »pametni« sustav uspio usvojiti objektivno znanje ili mudrost o domeni problema na temelju (uvijek) ograničenog skupa podataka. Dokazano je da mnoge UI -aplikacije »pate« od neke vrste pristranosti (eng. bias). Stoga automatizirano donošenje odluka može dovesti do pogrešnih procjena koje ponavljaju i pojačavaju postojeće pristranosti. Ignoriranje te činjenice o umjetnoj inteligenciji može imati opasne implikacije u slučajevima kao što su medicinska dijagnoza, upravljanje strojevima (npr. robotska vozila, kolaborativni ili kirurški roboti) ili onima koji imaju neku vrstu društvenog utjecaja. Posljedično tomu potrebno je imati formalne alate za provjeru i ocjenu sigurnog rada autonomnih sustava. Validacija se može usredotočiti na proces razmišljanja i provjere baze znanja inteligentnog sustava te osigurati da je formulacija inteligentnog ponašanja unutar sigurnosnih granica. Treba imati na umu da mnogi autonomni sustavi svoju inteligenciju stječu nenadziranim učenjem ili koriste nesimboličke UI-metode koje je teško objasniti, odnosno čije je znanje nedovoljno transparentno ili čitljivo da bi se moglo provjeriti. Autonomni sustavi ne ponašaju se točno, već samo u skladu s našim očekivanjima. Stoga neovisnost umjetne inteligencije i posljedično tomu njezina određena nepredvidljivost, ili neka vrsta neposlušnosti, predstavljaju prijetnju. Strah od takve nekontrolirane strojne inteligencije nameće različita pitanja, posebno u smislu »singularnosti« umjetne inteligencije. Ako inteligentni sustavi dođu u stanje samosvijesti koja se samostalno razvija bez naše kontrole, UI-singularnost podrazumijeva prekretnicu nakon koje tehnološki razvoj postaje nekontroliran i ireverzibilan.

Stoga se danas na međunarodnoj razini intenzivno proučavaju novi pravni okviri u pogledu razvoja i primjene robotike i umjetne inteligencije. Trebalo bi poticati povezana istraživanja kako bi se taj problem riješio ne samo s tehnološkog gledišta nego i osigurao odgovoran razvoj umjetne inteligencije, uzimajući u obzir fundamentalna filozofijska i kulturološka pitanja.⁷ Unatoč tomu autonomni sustavi, odnosno inteligentni roboti i računalne aplikacije, zbog brojnih potencijalnih koristi koje donose, šire svoju primjenu u industriji, uslužnim djelatnostima, obrazovanju, zdravstvu i medicini, industriji zabave, prometu, vojnim primjenama, financijama i energetici, nudeći svoje kompetencije u smislu samostalnog donošenja odluka ili automatiziranog zaključivanja.

Usvajanje umjetne inteligencije dokazano pridonosi gospodarskom rastu, posebice u spororastućim razvijenim gospodarstvima. Pokretački meha-

⁷ Usp. Michael D. LANGFORD, *A Theological Framework for Reflection on Artificial Intelligence* (30. X. 2023.).

nizmi ovise o širem razumijevanju društvenih pojava kao i o razini tehničkog znanja.⁸ Nužna je transformacija obrazovnih okvira i komunikacije sa svim dionicima društva koji će promijeniti poimanje tehnologije kao prijetnje koja nas vodi u posthumanističku budućnost. Tehnologija je proizvod ljudske kulture. Stoga treba shvatiti da su robotika i umjetna inteligencija ravnopravan evolucijski dio naše kulture za koji smo osobno odgovorni.

Treba dodatno istaknuti snažan analitički potencijal umjetne inteligencije, poglavito u analizi internetskih izvora ili nestrukturiranih podataka. Zahvaljujući globalnoj umreženosti, dostupnost podataka u svim segmentima ljudskog interesa raste, a umjetna inteligencija nam omogućuje da nestrukturirane i neobrađene podatke pretvorimo u korisno znanje. Analitički alati koji se koriste strojnim učenjem bit će sve važniji u budućnosti za razumijevanje novih značenja, prepoznavanje skrivenih problema i stvaranje novih rješenja. Takve sustave danas znamo kao velike jezične modele (eng. Large Language Models – LLM) ili generativne sustave koji se koriste neuronskim mrežama za prepoznavanje obrazaca i struktura unutar postojećih podataka za generiranje novog i originalnog sadržaja. Temelje se na različitim algoritmima učenja, uključujući nadzirano, nenadzirano ili polunadzirano učenje. Takvi sustavi traže pristup velikim količinama podataka, pa i našim osobnim podacima, uključujući biometriju, osobne medicinske podatke, ono što volimo, što čitamo, što kupujemo, koju glazbu slušamo, i tako dalje. Je li to potencijalni izazov našoj privatnosti? Trebamo li dopustiti dijeljenje naših podataka kako bismo umjetnu inteligenciju učinili još boljom i učinkovitijom? Sve su to pitanja na koja moramo imati odgovore i koja su važna za predviđanje i pripremu odgovarajućih regulatornih, pravnih i političkih okvira. Stoga se ovdje mora naglasiti da će istodobno zakonska regulativa i tehnologije za sprječavanje raznih oblika zlorabe osobnih podataka, kojima nažalost često svjedočimo, sve više dobivati na značaju. Zahvaljujući velikim jezičnim modelima značajno su unaprijeđeni sustavi za razumijevanje govora. Zahvaljujući razvoju metoda dubokog učenja⁹ i vjerojat-

⁸ Usp. Peter STONE – Rodney BROOKS – Erik BRYNJOLFSSON I DR., *Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report (2021.)*, u: https://ai100.stanford.edu/sites/g/files/sbiybj18871/files/media/file/AI100Report_MT_10.pdf. (30. X. 2023.).

⁹ Usp. Yanming GUO – Yu LIU – Ard OERLEMANS – Songyang LAO – Song WU – Michael S. LEW, *Deep learning for visual understanding: A review*, u: *Neurocomputing*, 187 (2016.), 27-48; Daniele RAVI – Charence WONG – Fani DELIGIANNI – Melissa BERTHELOT – Javier ANDREU-PEREZ – Benny LO – Guang-Zhong YANG, *Deep learning for health informatics*, u: *IEEE Journal of Biomedical and Health Informatics*, 21 (2017) 1, 4-21; Neil C. THOMPSON – Kristjan GREENEWALD – Keeheon LEE – Gabriel F. MANSO, *The computational limits of deep learning* (10. VI. 2020.), u: <https://arxiv.org/abs/2007.05558>

nosnih modela,¹⁰ danas je moguće komunicirati s računalom bez većih poteškoća. Često je teško razlučiti razgovaramo li s računalom ili s čovjekom. U budućće takve aplikacije iz etičkih razloga morati jasno upozoravati korisnika da komunicira s računalom, odnosno robotom (eng. chatbotom).

Vrlo važan trend je razvoj robotske motoričke inteligencije. To je vrsta inteligencije koje često nismo svjesni, a omogućuje nam da kontroliramo svoje tijelo i krećemo se u jednostavnim svakodnevnim situacijama. Npr. pranje čaše ili spremanje košulje u ormar za nas je gotovo automatska radnja, no pokušamo li istu implementirati na robotu, susrećemo se s nizom iznimno složenih funkcija, koje je vrlo teško programirati. Razvojem motoričke inteligencije omogućit ćemo robotima integraciju u složenu ljudsku okolinu i njihovu primjenu u potpuno novim zadaćama.

Računalni vid već je danas bolji od ljudskog u mnogim primjenama, a osim pojačanog nadzora javnih prostora, sve će se više koristiti za podršku raznim vrstama usluga. Stoga će pitanja privatnosti postajati sve važnija. Roboti će u sinergiji s naprednim računalnim vidom značajno poboljšati svoje performanse u dinamičnom i promjenjivom okruženju. Računalni vid posebno će utjecati na unaprjeđenje mobilnih robota i autonomnih vozila.

Nasuprot autonomiji, poseban izazov predstavlja suradnja čovjeka i stroja. Suradnički robotski sustavi postaju novi koncept robotizacije.¹¹ Zahvaljujući umjetnoj inteligenciji moguće je programirati robote tako da mogu osjetiti ljudsku prisutnost i razumjeti ljudske namjere. U takvim primjenama robot će

(5. XI. 2023.); Xizhao WANG – Yanxia ZHAO – Farhad POURPANAH, Recent advances in deep learning, u: *International Journal of Machine Learning and Cybernetics*, 11 (2020.) 4, 747-750; Chigozie NWANKPA – Winifred IJOMAH – Anthony GACHAGAN – Stephen MARSHALL, Activation functions: Comparison of trends in practice and research for deep learning (8. XI. 2018.), u: <https://arxiv.org/abs/1811.03378> (5. XI. 2023.).

¹⁰ Usp. Robin MANHAEVE – Giuseppe MARRA – Thomas DEMEESTER – Sebastijan DUMANČIĆ – Angelika KIMMIG – Luc DE RAEDT, Neuro-Symbolic AI= Neural+ Logical+ Probabilistic AI, u: Pascal HITZLER – Md Kamruzzaman SARKER (ur.), *Neuro-Symbolic Artificial Intelligence: The State of the Art*, Manhattan-Hartford, 2022., 173-191; Vishwas RAO – Sandeep MADIREDDY – Carlo GRAZIANI – Puning XUE – Romit MAULIK, Probabilistic Machine Learning and Data Assimilation (2021.), u: https://www.ai4esp.org/files/AI4ESP1110_Rao_Vishwas.pdf (5. XI. 2023.); Vinicius Veloso Eleuterio NOGUEIRA – Luiz Edival de SOUZA, UAV mission planning and execution via non-deterministic AI planning on ROS (2018.), u: https://www.google.hr/url?sa=t&rcrt=j&q=&esrc=s&source=web&cd=&ved=2ahUKewi7i5apnNODAxWMrvEDHRKpBMYQFnoECBIAQ&url=https%3A%2F%2Fwww.sba.org.br%2Fopen_journal_systems%2Findex.php%2Fcba%2Farticle%2Fdownload%2F537%2F499%2F&usq=AOvVaw3sg8mtSiLCkpkUcaEKp33t&opi=89978449 (5. XI. 2023.).

¹¹ Usp. Human-machine Collaboration Moves Towards a Safe and Smart Factory, (31. IV. 2022.), u: <https://www.market-prospects.com/articles/human-machine-collaboration-and-smart-factory> (5. XI. 2023.).

blisko surađivati s ljudima u njihovom svakodnevnom okruženju (kuće, uredi, tvornice, bolnice). To znači da je potrebno stvoriti modele za prirodnu i intuitivnu komunikaciju između čovjeka i robota. Općenito, moramo biti spremni prihvatiti da roboti nikada neće u potpunosti imati perceptivne sposobnosti poput ljudi i stoga možda neće moći otkriti značajke okoline s kojima ljudi žele imati posla tijekom komunikacije. Multimodalna sučelja trebala bi biti korisna zbog svoje potencijalno visoke redundantnosti, veće osjetljivosti, povećane točnosti i mogućih sinergijskih učinaka različitih pojedinačnih načina komunikacije, ako se uzmu zajedno.

Uvođenje robota u našu svakodnevicu, izvan determinističkog industrijskog okruženja, otvara potpuno nove mogućnosti primjene, posebice u uslužnim poslovima. Inteligentni roboti mogu igrati važnu ulogu u poboljšanju kvalitete ljudskog života kao i u poboljšanju različitih korisničkih usluga.¹² Sadašnje generacije robota teško je primijeniti u poslovima koji zahtijevaju intuiciju, prosudbu, empatiju, a posebno motoričku inteligenciju. Razvoj uslužnih robota iznimno je složen zadatak jer se temelji na interdisciplinarnoj znanosti, a ne samo na robotskoj tehnologiji.¹³ Zahtijeva novo i inovativno razmišljanje o odnosima među različitim entitetima.¹⁴ Uslužni roboti predstavljaju konvergenciju i primjenu mnogih tehnologija, uključujući elektroniku, računala, mehaniku, senzoriku, navigaciju, primijenjenu psihologiju, sociologiju, obradu jezika i biologiju.¹⁵ Zbog izravne interakcije s ljudima, uslužna robotika uključuje sofisticiranu primjenu inteligencije,¹⁶ uz pretpostavku senzorne sposobnosti (verbalne, vizualne, fizičke, kinezičke i parajezične), kognitivnih kapaciteta (autonomija, interaktivnost) i afektivnih doživljaja (emocija).

¹² Usp. Li XIAO – Vijay KUMAR, Robotics for Customer Service: A Useful Complement or an Ultimate Substitute?, u: *Journal of Service Research*, 24 (2019.) 2, 9-29.

¹³ Usp. Ai-Hsuan CHIANG – Silvana TRIMI, Impacts of service robots on service quality, u: *Service Business*, 14 (2020.) 2, 439-459; Eduardo ZALAMA – Jaime Gómez GARCÍA-BERMEJO – Samuel MARCOS – Salvador DOMÍNGUEZ – Raul FELIZ – Roberto PINILLOS – Joaquin LÓPEZ, Sacarino, a service robot in a hotel environment, u: *ROBOT2013: First Iberian Robotics Conference Advances in Intelligent Systems and Computing* (2014.), 3-14.

¹⁴ Usp. Thomas NOVAK – Donna HOFFMAN, Relationship journeys in the Internet of Things: a new framework for understanding interactions between consumers and smart objects, u: *Journal of the Academy of Marketing Science*, 47 (2019.) 2, 216-237.

¹⁵ Usp. Eduardo ZALAMA – Jaime Gómez GARCÍA-BERMEJO – Samuel MARCOS – Salvador DOMÍNGUEZ – Raul FELIZ – Roberto PINILLOS – Joaquin LÓPEZ, Sacarino, a service robot in a hotel environment, 3-14.

¹⁶ Usp. Sebastian LECHEVALIER – Junichi NISHIMURA – Cornelia STORZ, Diversity in patterns of industry evolution: how an intrapreneurial regime contributed to the emergence of the service robot industry, u: *Research Policy*, 43 (2014.) 10, 1716-1729.

Općenito, korisnici pozitivno reaguju na uslužne robote, čak i u pružanju emocionalne interakcije. Međutim, pokazalo se da je potrebno poboljšati vještine socijalne interakcije, jer s druge strane, društveno prihvaćanje inteligentnih strojeva i umjetne inteligencije općenito uvelike ovisi o javnoj komunikaciji i obrazovnim konceptima koji oblikuju javno mnijenje.¹⁷ Ti koncepti moraju biti formulirani na način koji je razumljiv i prihvatljiv svima bez obzira na obrazovanje. Primjerice, robot je fizička inkarnacija umjetne inteligencije i njegov izgled može biti iznimno važan u postizanju kvalitetne komunikacije. Ako robot ima humanoidni izgled i pokazuje ljudsko ponašanje, ljudi mogu s njim komunicirati na prirodniji način. Ne smijemo zaboraviti ni tzv. neverbalnu komunikaciju kojom se ljudi intenzivno i često nesvjesno koriste u svakodnevnoj komunikaciji. Ponašanje antropomorfnog ili humanoidnog robota mogu lakše razumjeti čak i ljudi koje robotska tehnologija ne zanima niti ju prihvaćaju. Humanoidna veličina i oblik robota važni su za integraciju u čovjekov okoliš i dijeljenje prostora s ljudima.

Stoga je potreban interdisciplinarni napor za rješavanje integracije, uzimajući u obzir ljudske čimbenike, kognitivne znanosti, robotske tehnologije, programske tehnike i UI-sučelja, ali i psihologiju i sociologiju. Kombinirajući najbolje od svakoga, dobivamo komplementarnu sinergiju dvaju svjetova, ljudskoga i tehničkoga. To je izuzetno složen zadatak, imajući u vidu kako tu suradnju ostvariti na siguran način. Kao rezultat, pojavila se potpuno nova istraživačka domena pod nazivom »HRI – Human Robot Interaction« (Interakcija ljudi i robota). Neke relevantne istraživačke discipline vezane uz to su senzorna informatika i sučelja čovjek-računalo ili čovjek-robot. To uključuje razvijanje metoda koje upravljačkom sustavu omogućavaju da objasni svoje »razmišljanje« i svoje namjere na ljudima razumljiv način i da u isto vrijeme razumije ponašanje ljudi i konotaciju radnog okruženja.

Suradnja čovjeka i stroja također otvara potpuno novu i neočekivanu dimenziju. Intelligentni roboti mogu poboljšati ne samo naše fizičke nego i kognitivne sposobnosti. Ljudski kreativni potencijali uglavnom su ograničeni našim obrazovanjem, emocijama, resursima i kulturom. Naša sposobnost probijanja stečenih granica predstavlja snagu inovacije. Kako bismo riješili goruće probleme našeg vremena, inovacije se moraju ubrzati barem za faktor deset.¹⁸

¹⁷ Usp. Isto; Dan HUANG – Quiroung CHEN – Jiahui HUANG – Shaojun KONG – Zhiyong LI, Customer-robot interactions: Understanding customer experience with service robots, u: *International Journal of Hospitality Management*, 99 (2021.), 103078.

¹⁸ Usp. Alessandro BRANDOLISIO – Michaels LEITL – Karel J. GOLTA, *The AI Toolbook. Mit Künstlicher Intelligenz die Zukunft sichern*, Hamburg, 2021.

S druge strane suočeni smo desetljećima s ozbiljnim padom istraživačke produktivnosti.¹⁹ Razlog je sve veća kompleksnost novih tehnologija s kojima se bavimo. Opće je prihvaćeno kako su vremena takozvanog »lako dohvatljivog voća«, u smislu izuma, prošla (eng. Low-hanging fruits are over). Stoga danas moramo uložiti daleko više napora u dosizanju novih znanstvenih horizonata. S druge strane, umjetna inteligencija nema kulturu, emocije i/ili osjećaj za estetiku. Umjetna inteligencija može stvarati rješenja bez predrasuda i ograničenja, samo na temelju objektivnih informacija. Što bolje podatke pružamo, to bolje rezultate dobivamo. Stoga nam umjetna inteligencija može pomoći otvoriti nove još neotkrivene domene mogućih rješenja i proširiti horizonte kreativnog razmišljanja. Prema Davidu Rotmanu »Najveći utjecaj umjetne inteligencije bit će pomoći ljudima da dođu do otkrića do kojih ne bismo mogli doći sami.«²⁰ To je ogroman potencijal za ubrzanje inovativnog razvoja modernog društva.²¹ Dakle, sljedeća faza umjetne inteligencije je doba proširene inteligencije, koja neprimjetno povezuje čovjeka i stroj.

2. Roboetika

Masovna digitalizacija ljudskog društva povećava mogući utjecaj umjetne inteligencije na našu percepciju društvenih, etičkih i kulturoloških vrijednosti kroz brojne aplikacije kojima komuniciramo. Sve naše predrasude i slučajne ili pak namjerne pogreške uključene su u različite baze podataka, publikacije, web-stranice i općenito sve raspoložive digitalizirane izvore. Važno je pritom naglasiti da su to sve informacije koje smo mi kao ljudi izravno ili neizravno stvorili te odražavaju sve naše nesavršenosti kao i kompleksnost ljudskog društva u različitim aspektima. Svako odstupanje u uzorku podataka za učenje umjetna inteligencija će prepoznati kao značajku koju će nastojati pojačati. Stoga bilo kakva pogrešna procjena koja proizlazi iz nesavršenosti podataka ili UI-metoda može pridonijeti nekoj vrsti spoznajne devijacije ili dezinformiranja korisnika ili korisničke skupine. Dezinformacije u globalnom digitalnom

¹⁹ Usp. Nicholas BLOOM – Charles I. JONES – John Van REENEN – Michael WEBB, *Are Ideas Getting Harder to Find*, u: *American Economic Review*, 110 (2020.) 4, 1104-1144.

²⁰ David ROTMAN, *AI is reinventing the way we invent* (15. II. 2019.), u: <https://www.technologyreview.com/2019/02/15/137023/ai-is-reinventing-the-way-we-invent/> (8. XI. 2023.).

²¹ Usp. Ricardo VINUESA – Hossein AZIZPOUR – Iolanda LEITE – Madeline BALAM – Virginia DIGNUM – Sami DOMISCH – Anna FELLÄNDER – Simone Daniela LANGHANS – Max TEGMARK – Francesco Fuso NERINI, *The role of artificial intelligence in achieving the Sustainable Development Goals*, u: *Nature Communications*, 11 (2020.) 233, u: <https://doi.org/10.1038/s41467-019-14108-y>, www.nature.com/naturecommunications (8. XI. 2023.).

prostoru predstavljaju ozbiljnu prijetnju demokraciji ili društvu u cjelini jer učinkovito mijenjaju osjećaj objektivne istine i na taj način manipuliraju javnim mnijenjem. Stoga se mora spriječiti svaka zloporaba umjetne inteligencije koja može negativno utjecati na demokraciju ili na bilo koji način ograničiti ljudske slobode.²²

U budućnosti očekujemo povećanje složenosti interakcije s robotima, autonomnim strojevima i inteligentnim računalnim aplikacijama. Istodobno, ubrzan razvoj bioničkih sustava pojačat će integraciju ljudi i tehnologija, ponajprije kroz niz protetskih pomagala i aktivnih implantata. To otvara pitanja kiborgizacije, odnosno pitanja gdje prestaje naš biološki temelj, a počinje tehnologija, i obrnuto. To nedvojbeno izaziva značajne etičke izazove.²³

Temeljna filozofijska pitanja bave se kategorijama svijesti, kulture, morala, kao i zasebnim kategorijama poput snova, intuicije i emocija. Sve su to kategorije koje je teško definirati, a još teže matematizirati, odnosno pretvoriti u računalne programe. Stalno se pitamo mogu li roboti koje gradimo jednog dana biti svjesni i iskusiti unutarnje ili subjektivne svjetove slične onima koje doživljavaju ljudi. Vjerojatno ne, ali sigurno ćemo moći razviti neke oblike tehničke svijesti i umjetnih emocija koje bi mogle sličiti ljudskim. No, moramo biti svjesni da takva robotska inkarnacija čovječanstva vjerojatno nikad neće biti doslovno usporediva s ljudskim karakteristikama. Morat ćemo prihvatiti UI-robote kao zasebnu vrstu, neku vrstu umjetnog života. Zatim ćemo morati razumjeti unutarnji svijet strojne psihologije i njihovu subjektivnost.

Takozvana »roboetika« naglašava dva osnovna stava: odnos ljudi prema strojevima i odnos strojeva prema ljudima.²⁴ Odnos ljudi prema strojevima/robotima (materijalnim dobrima) oduvijek je uključivao svojevršno poštovanje nečega vrijednog i bio je društveno reguliran (uništavanje materijalnih dobara nije dopušteno). Danas se taj odnos ističe humanizacijom tehnologije i implementacijom umjetne inteligencije (antropomatska robotika, pametna računalna sučelja, biomimetički sustavi itd.). Kao rezultat toga, možemo primijetiti

²² Usp. Peter M. ASARO, *The Liability Problem for Autonomous Artificial Agents* (2016.), u: <https://cdn.aaai.org/ocs/12699/12699-56141-1-PB.pdf> (8. XI. 2023.).

²³ Usp. Seth BAUM, *Social Choice Ethics in Artificial Intelligence*, u: *AI & Society*, 35 (2020.) 1, 165-176.

²⁴ Usp. Gianmarco VERUGGIO, *The Birth of Roboethics* (2005.), u: <http://www.roboethics.org/icra2005/veruggio.pdf> (8. XI. 2023.); Gianmarco VERUGGIO, *The EURON Roboethics Roadmap* (2006.), u: <https://www3.nd.edu/~rbarger/ethics-roadmap.pdf> (8. XI. 2023.); Gianmarco VERUGGIO – Fiorella OPERTO, *Roboethics: Social and Ethical Implications of Robotics*, u: Bruno SICILIANO – Oussama KHATIB (ur.) *Springer Handbook of Robotics*, Berlin, 2008., 1499-1524.

neku vrstu emocionalnog odnosa ljudi prema stvarima, koji uključuje različite razine interakcije:

- pojednostavljen odnos – nema iznenađenja, nema razočaranja, odnos je jednostran tako da čovjek nema moralnu obvezu prema stroju/robotu kao osobi;
- složen odnos – uključuje »sociotehničku« dimenziju koja remeti postojeće društvene odnose. Robot ili pametni uređaj pojavljuje se kao uljez, interaktivni društveni stroj koji se često postavlja između nas stvarajući emocionalne smetnje. Iako se čini neupitnim naš izbor »očuvanja« između čovjeka i robota, sve ostalo postaje upitno;
- robot kao elektronička osoba (?) – takvi bi roboti trebali imati određenu vrstu tehničke svijesti i emocija. Tada bi im bilo potrebno priznati moralna prava, ali i obveze.

Odnos strojeva/robota prema ljudima postaje značajniji posebno danas, a posebno će biti važan u budućnosti, zahvaljujući razvoju robotike te inteligentnih i autonomnih sustava. Zbog napretka umjetne inteligencije, intenzitet i složenost interakcije stroja s ljudima raste, ističući sljedeća pitanja:

- Moralne dužnosti robota uglavnom proizlaze iz upravljačkog programa koji je napravio čovjek u skladu sa svrhom kojoj je robot namijenjen. Stoga se čini da je odgovornost za moralno ponašanje robota na ljudskoj strani, ili bismo mogli reći da su u ovom slučaju roboetički problemi preuveličani.
- Autonomni sustavi temeljeni na evolucijskim i probablističkim metodama omogućuju kreiranje ponašanja koje se temelji samo na očekivanjima. Mogu li takvi roboti potkopati etička načela unatoč Asimovim zakonima robotike? Složeni interaktivni odnosi mogu dovesti do moralno upitnih odluka.
- Mogu li inteligentni roboti razviti vlastite etičke standarde? Kad bi mogli razviti određenu vrstu svijesti i emocija, to bi bilo moguće.

Imajući u vidu trenutni stupanj znanstvenog i tehnološkog razvoja, roboti nisu moralni agenti, ali bi to mogli postati u budućnosti.²⁵

²⁵ Usp. Bernhard IRRGANG, Ethical Acts in Robotics, u: *Ubiquity* (2006.) 1, 2-16; usp. Wendell WALLACH – Colin ALLEN, *Moral Machines: Teaching Robots Right from Wrong*, New York, 2009.; Selmer BRINGSJORD – Joshua TAYLOR, *The Divine-Command Approach to Robot Ethics*, u: Partrick LIN – Keith ABNEY – Georg A. BEKEY (ur.), *Robot Ethics – The Ethical and Social Implications of Robotics*, Cambridge – London, 2012.; Selmer BRINGSJORD – Konstantine ARKOUDAS – Paul BELLO, *Toward a general logicist methodology for engineering ethically correct robots*, u: *IEEE Intelligent Systems*, 21 (2006.) 4, 38-44.

- Moralno ponašanje je naučeno ponašanje razvijeno iz lokalne kulture. Stoga je podjednako instinktivno i racionalno (»međutim, da bismo bili moralno odgovorni i djelovali, potreban je sudionik kojeg karakterizira osobnost ili subjektivnost.«).²⁶
- Moralna odgovornost podrazumijeva razumijevanje konteksta vlastitog ponašanja. Može li robot razumjeti kulturu? Npr. ako je moralno načelo povrijeđeno, moralna odgovornost uzrokuje sram. Sram je očito emocionalna kategorija. Možemo li onda moralna uvjerenja, a zatim i etička načela »matematizirati«?

Razvoj etičke umjetne inteligencije i robotike nalaže sljedeće metode i pristupe:²⁷

1. uključivanje etike u istraživanje i razvoj umjetne inteligencije i robotike – uključujući istraživačku etiku i etiku dizajna;
2. uključivanje etike u implementaciju i korištenje UI i robotike – usmjereno na organizacije koje implementiraju i koriste se umjetnom inteligencijom i robotskom tehnologijom;
3. kreiranje politike i kulture korporativne odgovornosti koje podupiru etički razvoj i korištenje umjetnom inteligencijom i robotikom – usmjerene na programere, korisnike i organizacije za podršku;
4. definiranje nacionalnih i međunarodnih standarda za certifikaciju etičke umjetne inteligencije i robotike – usmjerene na institucije za standarde;
5. kreiranje politika i propisa za podršku etičkim praksama u umjetnoj inteligenciji i robotici – usmjereni na institucije za standarde;
6. obrazovanje, obuka i podizanje svijesti o etičkim i društvenim aspektima umjetne inteligencije i robotike – usmjereno na obrazovne institucije i medije.

Prema tome, trebat ćemo definirati što očekujemo u smislu morala od inteligentnih i autonomnih sustava te koje su međusobne odgovornosti, posebno u smislu stvaranja novog biotehnološkog društva s odgovarajućim etičkim kodeksom.

²⁶ Usp. Bernhard IRRGANG, Ethical Acts in Robotics, u: *Ubiquity*, (2006.) 1, 2-16.

²⁷ Usp. Philip BREY – Philip JANSEN – Jonne Maas – Bjorn LUNDGREN – Anais Resseguier, D4.7: An Ethical framework for the development and use of AI and robotics technologies. WP4: Artificial Intelligence and Robotics – Ethical, Legal and Social Analysis, SIENNA project, H2020, No 741716, 2020., u: https://www.sienna-project.eu/digitalAssets/801/c_801912-1_1-k_d4.7_ethical-framework-for-ai--robotics_v2.pdf. (10. XI. 2023.).

Zaključak

Masovna digitalizacija ljudskog društva donosi brojne koristi, ali istodobno otvara brojna pitanja povezana s potencijalnim rizicima. Umjetna inteligencija, kao jedna od najsnažnijih suvremenih tehnologija, može imati presudan utjecaj na našu percepciju društvenih, etičkih i kulturoloških vrijednosti kroz brojne računalne servise preko kojih komuniciramo. Postoje etička pitanja o tome kako se koristiti tim tehnološkim alatima za rad prema općem dobru, zajedno s drugim filozofijskim pitanjima, da ne spominjemo politička. Robotika i umjetna inteligencija otvaraju novu platformu za preispitivanje starih pitanja, tjerajući nas da razmišljamo o našoj ljudskosti i našoj jedinstvenosti ili barem o prirodi naše inteligencije. Stoga etički izazovi zahtijevaju redefiniranje postojećih zakona, društvenih navika, pa i postojećih civilizacijskih koncepata u cilju zaštite ljudskih prava i sloboda. Istraživačka UI- zajednica ima ključnu ulogu u razvoju novog biotehnološkog društva. Naš uspjeh mjerit će se po tome koliko ćemo poboljšati kvalitetu života cijele ljudske populacije, a ne samo pojedinih skupina, te koliko ćemo uspjeti sačuvati naš planet za buduće generacije. Svjesni smo brojnih prednosti koje proizlaze iz tehnološkog razvoja, ali i izazova koje taj tehnološki razvoj predstavlja za nas. Stoga, robotika i umjetna inteligencija ne smiju biti ljudski »konačni izumi«.

Abstract

HOW TO SURVIVE WITH SMART MACHINES

Bojan JERBIĆ

Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb
CRTA – Regional Center of Excellence for Robotic Technologies
Ivana Lučića 5, HR – 10 000 Zagreb
bojan.jerbic@fsb.hr

Robotics and artificial intelligence (AI) are unquestionably enhancing human potential. They operate in a range from simple reasoning, performing simple actions to imitating complex human cognitive and labour functions. Their tasks can vary from fun or games to helping people with difficult or tedious tasks. In order to understand the impact of artificial intelligence, it is both important to learn from past successes and failures, and to predict its future directions and potential legal, ethical and socio-economic implications. In the near future, intelligent robots will increasingly become an integral part of ordinary people's lives. The potential benefits are great, however there

are also significant risks and ethical challenges. Our autonomy can be threatened, and social interaction obstructed. The widespread use of robots can lead to reduced contact between people and possible restrictions on personal freedoms. Smart machines are shaping a radically new world, bringing significant economic and cultural changes, creating both winners and losers on a global scale. Artificial intelligence is developed by taking humans as a model. Despite this, artificial intelligence and human beings remain fundamentally different due to their different natures. Can we then share the same moral concepts? Although there are many concerns about what robotics and AI mean for our future, there is a belief from a scientific or philosophical perspective that they represent a part of the evolution. We progress, redefining humanity along the way, constantly questioning what makes us unique as humans.

Key words: *robotics, artificial intelligence, roboethics.*