

Razkritje umetne intelligence

Razumevanje njenih družbenih in antropoloških posledic

■ **Octavian M. Machidon**, dr. znanosti s področja elektrotehnike in telekomunikacij ter diplomirani teolog, je docent na Fakulteti za računalništvo in elektrotehniko Univerze v Ljubljani.



Ali ste danes uporabljali umetno inteligenco (UI)? Morda ste poslušali glasbo, si ogledali videoposnetek, ki vam ga je priporočil algoritem UI, morda ste za načrtovanje poti in izogibanje zastojem uporabili navigacijsko aplikacijo v avtu ali na pametnem telefonu ali pa ste preverili e-pošto, ne da bi opazili, da je UI filtrirala neželeno pošto. UI oblikuje naše vsakodnevno življenje, vendar to ni novost. Ti primeri uporabe UI so prisotni že dlje časa in smo se nanje že navadili. V zadnjem času pa smo priča novemu navdušenju nad UI. Zakaj se to dogaja? Kaj je zdaj drugače?

Razlog za pogostejše pojavljanje UI v medijih je njena najnovejša različica – 'generativna' UI – inovativni proizvod, ki lahko počne stvari, ki so videti, zvenijo in se zdijo srhljivo človeške (Monahan, 2023). Večina ljudi se strinja, da lahko UI znatno koristi človeštvu in preoblikuje različne panoge, od financ in izobraževanja do transporta in zdravstvenega varstva. UI lahko med mnogimi drugimi stvarmi avtomatizira ponavljajoča se opravila, izboljša procese odločanja ter poveča natančnost in hitrost analize podatkov. Kljub ogromnim potencialnim koristim pa UI prinaša tudi pomembne etične in družbene izzive. Enako kot vsako orodje se lahko UI uporablja tako za dobro kot za slabo. Poleg tega – glede na svojo naravo posnemanja človeške intelligence in komunikacije z uporabniki na človeški način – prinaša edinstvene antropološke izzive.

Preden se posvetimo njenim glavnim prednostim in pogostim kritikam, začnimo s kratkim uvodom v to, kaj točno je UI in kakšno je njeno trenutno stanje razvoja.

Kaj je UI?

Umetna inteligenca je trenutno najbolj vroča tehnologija. Kaj pa sploh je UI? Pa-

radoksalno je to zelo težko vprašanje, na katerega ni enoznačnega odgovora. Kratka šolska razlaga bi zvenela tako: UI je krovni izraz za sklop tehnologij, ki omogočajo računalnikom izvajanje nalog, ki bi običajno zahtevale človeško inteligenco. To vključuje prepoznavanje obrazov, razumevanje govora, vožnjo avtomobilov, pisanje stavkov, odgovarjanje na vprašanja in ustvarjanje slik.

Raziskave na področju UI so se začele v petdesetih letih prejšnjega stoletja, v šestdesetih letih pa je ameriško ministrstvo za obrambo UI uporabilo za treniranje računalnikov, da posnemajo človeško razmišljanje (Scott, 2024). Zgodnji sistemi UI so bili stroji, ki so prejeli vhodne podatke in nato podali priporočila na podlagi usposabljanja in algoritmov, ki so določali, kaj naj sistem stori s podatki. Predstavljajte si računalnike, ki lahko igrajo damo ali šah.

Razvojni preboj, ki smo mu bili priča v zadnjih dveh desetletjih, je posledica napredka na področju strojnega učenja (Simonite, 2023). To vključuje treniranje sistemov UI za izvajanje nalog na podlagi primerov namesto neposrednega programiranja njihovega vedenja. Tehnika strojnega učenja, imenovana globoko učenje,

je to metodo še okrepila – zdaj lahko računalniki, ki igrajo damo ali šah, premagajo celo najboljše človeške igralce v teh igrah.

Kje srečujemo UI?

UI ima ogromno transformativno moč pri izboljševanju produktivnosti in učinkovitosti v različnih sektorjih. Izjemna je pri avtomatizaciji ponavljajočih se nalog, kot so vnos podatkov, odgovarjanje na e-pošto in generiranje računov. UI učinkovito obdeluje velike količine podatkov, hitro pridobiva relevantne informacije za analizo ter olajšuje nadaljnjo interpretacijo in transformacijo podatkov. Poleg tega omogoča hitrejšo in bolj učinkovito odločanje s hitrim zbiranjem zanesljivih vpogledov in uporabo algoritmov za napovedovanje.

Ta potencial se že uresničuje na kritičnih področjih, kot je zdravstvo. Tam UI pomaga pri diagnostiki z identifikacijo majhnih nepravilnosti na slikah, triangulacijo diagnoz iz simptomov in vitalnih znakov bolnikov, klasifikacijo bolnikov, vodenju zdravstvenih kartotek in obdelavi zdravstvenih zahtevkov.

Medtem ko specializirani sistemi UI močno vplivajo na področja, kot je zdravstvo, so na voljo tudi bolj splošna orodja UI za širšo javnost. Posledično mnogi ljudje vsakodnevno uporabljajo UI, pogosto nezavedno. Študija, izvedena v Sloveniji leta 2023, je pokazala, da 42 % anketirancev uporablja tehnologije UI, pri čemer se uporaba povečuje z višjo izobrazbo (Kako Slovenci gledamo na umetno inteligenco?, 2024). Za primerjavo: raziskava Pew Research iz istega obdobja poroča, da 55 % Američanov redno uporablja storitve UI (AI Statistics 2024, 2024). Vendar vsi ti podatki temeljijo na samoocenah. Dejan-

ske številke bi lahko bile veliko višje, saj mnogi uporabljajo storitve UI, ne da bi se tega zavedali.

Vzemimo za primer pametni telefon: vsak pametni telefon je opremljen z več funkcionalnostmi, ki jih poganja UI. Glasovni asistenti, kot sta Siri in Google Assistant, uporabljajo obdelavo naravnega jezika za razumevanje in odgovarjanje na uporabniške ukaze. UI izboljšuje delovanje kamer pametnih telefonov z lastnostmi, kot sta prepoznavanje prizorov in portretni način. Funkcije za predvidevanje besedila in samodejno popravljanje pomagajo pri sestavljanju sporočil. Poleg tega UI omogoča prepoznavanje obrazov za odklepanje naprav in upravljanje varnosti. Te vgrajene funkcionalnosti UI naredijo pametne telefone bolj intuitivne in učinkovite. V zadnjem času pa vse več ljudi uporablja generativna orodja UI, kot je ChatGPT, ki lahko na podlagi uporabniških poizvedb ustvarijo besedila, slike ali celo videoposnetke.

Vidimo, da UI postaja vse bolj prisotna v naših vsakodnevnih življenjih in se je razvila do te mere, da razume človeške poizvedbe, izražene v naravnem jeziku. Ustavimo se za trenutek in se vprašajmo: Kaj pomeni ta nova realnost? Kaj pomeni za naprave, da razumejo govor ali napišejo stavek? Katere naloge bi lahko zahtevali od takšnih naprav? Koliko bi jim morali zaupati, da jih opravijo?

UI ustvarja nov način interakcije s tehnologijo

Če gledamo na njeno vseprisotnost in način, kako se nanjo zanašamo, lahko rečemo, da med ljudmi in UI vlada zaupanje. Zaupamo in sledimo priporočilom vsebin, ki jih ponujajo algoritmi UI na družbenih omrežjih, kot so videoposnetki na YouTubu, pesmi, ki jih poslušamo, in izdelki, ki jih kupujemo. Zaupamo navigacijskim aplikacijam UI, ki nas med vožnjo usmerjajo po najboljših poteh, in se zanašamo na sisteme UI za varovanje svojih telefonov s pomočjo prepoznavanja obrazov ali biometrične avtentikacije. Poleg tega zaupamo rezultatom, ki jih ustvarjajo generativna orodja UI, in odgovorom, ki jih ponujajo na naša vprašanja. Na primer, promocijski video

za ChatGPT 4.0 prikazuje UI, ki pomaga otroku pri matematičnih nalogah, kar nakazuje, da bi lahko kmalu zaupali UI tudi pri izobraževanju otrok.

Nikoli prej nismo mogli govoriti o vzpostavitvi zaupanja z napravo ali strojem, vendar z UI presegamo zgolj uporabo tehnoloških orodij in vstopamo v 'zaupljiv' odnos z umetnimi agenti. Po integrativnem modelu organizacijskega zaupanja Mayerja, Davisa in Schoormana (1995) je zaupanje pripravljenost na ranljivost ene strani za dejanja druge strani, kar temelji na pričakovanju, da bo druga stran brez nadzora prve izpolnila obljube in dejanja. To se morda sliši strašljivo, a je res: delovanje ali 'razmišljanje' sistemov UI, s katerimi komuniciramo, je zunaj naše zmožnosti nadzora, zato moramo zgolj zaupati njihovim rezultatom. Ko se naš odnos z UI pogloblja, sistemi UI posredujejo naše doživljanje sveta, oblikujejo človeško vedenje. Ta novi način interakcije s tehnologijo prinaša pomembne antropološke izzive in etična vprašanja, ki so neločljivo povezani s tehnično zasnovo kateregakoli sistema UI.

Izzivi umetne inteligence

Poleg antropoloških vprašanj prinaša hiter napredek UI tudi druge pomembne izzive, kot so trajnost, etika, pristranskost, dezinformacije in avtorske pravice (Simonite, 2023). Nekateri kritiki trdijo, da so zmožnosti UI precenjene, medtem ko drugi napovedujejo globoke družbene spremembe, ki bi lahko spremenile človeštvo, kot ga poznamo.

Pristranskost ostaja osrednja težava. Ker se sistemi UI zanašajo na podatke, ki odražajo naš pristranski svet, te pristranskosti tudi ohranjajo, kar vodi v nepoštene in diskriminatorne rezultate. Poleg tega se pojavljajo etične skrbi zaradi uporabe UI za ustvarjanje globokih ponaredkov in širjenje dezinformacij. Veliko ljudi se boji vpliva UI na zaposlovanje, saj se bojijo izgube delovnih mest in finančne negotovosti. Pravna odgovornost za dejanja, ki jih povzročajo sistemi UI, kot so nesreče avtonomnih vozil, je še ena izmed izstopajočih skrbi. Ljudje pogosto nimajo pregleda nad tem, kako se ravna z njihovimi podatki.

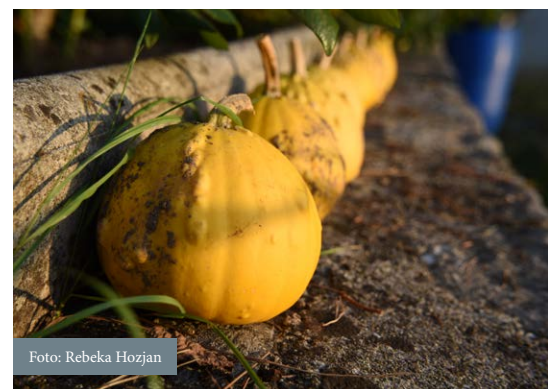


Foto: Rebeka Hozjan

Prav tako se postavlja vprašanje, ali bo čezmerno zanašanje na UI pri izvajanju nalog vodilo v zmanjšanje človeških intelektualnih in kognitivnih sposobnosti. Ne nazadnje način, kako UI močno podaljšuje čas, preživet v digitalnem svetu (npr. na družbenih omrežjih), povzroča pereče težave z duševnim zdravjem, zlasti pri otrocih in mladostnikih.

Kljub tem strahovom praktične koristi UI vodijo v njeno postopno sprejemanje. Izobraževanje uporabnikov o tem, kako UI deluje, kakšne so njene koristi in omejitve, je ključno za varno in smiselno vključevanje. Ker je UI tukaj in bo ostala z nami, je bistveno razumeti, v čem je dobra, kakšne so njene omejitve in kako jo odgovorno uporabljati. Ozaveščenost o njenih potencialno negativnih učinkih je ključna za zmanjšanje tveganj in povečanje koristi. V prihodnjih številkah bomo podrobneje obravnavali te izzive in iskali načine, kako varno pluti med čermi umetne inteligence. ◀

Literatura

- AI Statistics 2024 (2024). AIPRM. Pridobljeno 13. 7. 2024 s spletne strani: <https://www.aiprm.com/ai-statistics/>.
- Mayer, Roger C.; Davis, James H.; Schoorman, David F. (1995): An Integrative Model Of Organizational Trust. *Academy of Management Review* 20, no. 3 (July): 709–734. Pridobljeno 10. 7. 2024 s spletne strani: <https://doi.org/10.5465/amr.1995.9508080335>.
- Monahan, Jennifer (2023): *Artificial Intelligence, Explained*. Carnegie Mellon University Heinz College. Pridobljeno 5. 7. 2024 s spletne strani: <https://www.heinz.cmu.edu/media/2023/July/artificial-intelligence-explained>.
- Kako Slovenci gledamo na umetno inteligenco? (2024). *Računalniške novice*. Pridobljeno 12. 7. 2024 s spletne strani: <https://racunalniske-novice.com/kako-slovinci-gledamo-na-umetno-inteligenco/>.
- Scott, Gordon (2024): What Is Artificial Intelligence (AI)? *Investopedia*. Pridobljeno 13. 7. 2024 s spletne strani: <https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp>.
- Simonite, Tom (2023): What Defines Artificial Intelligence? The Complete WIRED Guide. *WIRED*. Pridobljeno 13. 7. 2024 s spletne strani: <https://www.wired.com/story/guide-artificial-intelligence/>.