

Nina Cvar

**Aleš Završnik, Katja Simončič (ur.):  
Artificial Intelligence, Social Harms and Human Rights.  
London: Palgrave Macmillan, 2023.  
290 str., (ISBN 978-3-031-19148-0), 106,99 EUR**

Inovacije in tehnološki razvoj spreminjajo družbeno vez, toda razvojni premiki se ne odvijajo ločeno od družbenoekonomskega, političnega in kulturnega okolja, kar nam na primeru tehnologije umetne inteligence skozi pravni in etični okvir nazorno pokaže znanstvena monografija *Artificial Intelligence, Social Harms and Human Rights*, ki je v začetku leta 2023 izšla pri založbi Palgrave Macmillan, uredila pa sta jo Aleš Završnik in Katja Simončič.

V zadnjem času lahko spremljamo hiter razvoj strojnega sistema umetne inteligence, širša javnost pa se je s to tehnologijo morda prvič поблиže srečala ob t. i. velikih jezikovnih statističnih modelih, kot je npr. Chat GPT. Delovanje Chata GPT sloni na velikih količinah podatkov, na podlagi katerih ta podaja predvidevanja verjetnih zaporedij besed, četudi ne razume jezika v človeškem smislu.

Orodja z umetno inteligenco lahko pozitivno podpirajo številna družbena področja, po drugi strani pa lahko sistemi umetne inteligence kršijo temeljne človekove pravice, kodificirane v ustavah in zavezujočih mednarodnih pravilih. Tudi sicer ti sistemi ne posedujejo t. i. etičnega razumevanja sveta, vendar pa lahko o njih morda najlažje razmišljamo prav z vidika naravne (človeške) inteligence in njenih specifičnosti.

Čeravno se monografija, ki vključuje deset prispevkov, analize mesta umetne inteligence v sodobni družbi loteva skozi prizmo človekovih pravic – zanima jo, v kolikšni meri jih umetna inteligenca krši oz. kakšne so implikacije njenega delovanja za človekove pravice – vse analize prinašajo svež, predvsem pa dragocen in pomemben prispevek k razumevanju strojnega sistema umetne inteligence.

Delo, katerega osrednja stava je, da naj se pri razvoju pričujoče tehnologije upošteva interdisciplinarni pristop, je v grobem razdeljeno na dva velika epistemološka dela: prvi del obravnava različna področja, na katerih se lahko v povezavi z umetno inteligenco pojavijo kršitve človekovih pravic v sklopu kazenskega pravosodja in kazenskega pregona, sodobnega vojskovanja in horizontalnih zasebnopravnih razmerij, kot je npr. zaposlovanje, drugi del pa se osredotoča na politike in modele upravljanja z umetno inteligenco; ob tem vključuje široko mrežo vključenih deležnikov, raziskovalni poudarek pa namenja tistim pristopom, ki bi lahko zagotovili kar največjo splošno javno korist, vključujoč poglobljen pretres učinkov na različne zakonske pravice.

Glede na to, da gre pri umetni inteligenci za strojni sistem, ki deluje tako, da lahko z uporabo strojnih in/ali človeških podatkov in vhodnih podatkov vpliva na okolje z izdelavo rezultata za določen nabor ciljev, avtorice in avtorji monografije izpostavljajo oblikovanje ustreznega pravnega in etičnega okvira, v katerem naj se odvija razvoj tega sistema.

Tudi sicer je morda največja odlika dane monografije, da ji uspe ujeti edinstveno ravnovesje med obravnavo tematik, povezanih z umetno inteligenco, človekovimi pravicami in etiko, obenem pa prek analiz sama podaja zanimivo mešanico družboslovnih in humanističnih diskurzov skupaj z računalniškimi znanostmi in tehniko.

Knjigo odpre prispevek *Artificial Intelligence and Sentencing from a Human Rights Perspective*, ki se ukvarja z vpeljavo umetne inteligence v kazenskopravni postopek, ob tem pa tematizira številna kompleksna pravna vprašanja, ki so bodisi nadaljevanje obstoječih problematik bodisi gre za povsem nove izzive, pri čemer se morda za najbolj lucidnega izkaže poskus analitičnega preseka z analizo UI v kazenskem pravu. Sledi besedilo *Technical and Legal Challenges of the Use of Automated Facial Recognition Technologies for Law Enforcement and Forensic Purposes*, ki podaja premislek o tehničnih in pravnih izzivih uporabe tehnologij samodejnega prepoznavanja obrazov za namene kazenskega pregona in forenzike, a ob tem opozarja na nezanesljivost takšnih metod, saj je avtomatizirano prepoznavanje obrazov po svoji naravi verjetnostno, zaradi česar je treba vzpostaviti ustrezne celovite zakone, s katerimi bi se zapolnila regulativna praznina. V nadaljevanju monografija postreže z ekskurzom v sodobno vojskovanje, in sicer s poglavjem *Artificial Intelligence, International Law and the Race for Killer Robots in Modern Warfare*, ki ponudi kritičen premislek o vplivu umetne inteligence na vojaško logistiko, obveščevalno dejavnost in nadzor. Poleg izrisa geopolitičnega konteksta posebno pozornost namenja rabi brezpilotnih letal, podobno kot analiza tehnologije samodejnega prepoznavanja obrazov pa tudi pričujoči prispevek poziva k oblikovanju skladnega, zavezujočega mednarodnega pravnega okvira za uporabo brezpilotnih letal, ki naj med drugim določa, da morajo vsa brezpilotna letala z umetno inteligenco sloneti na ustreznem človeškem nadzoru. Z diskriminatornimi učinki avtomatiziranega odločanja na področju zasebnega prava, natančneje na področju zaposlovanja ter dostopa do blaga in storitev ter njihove dobave, se ukvarja poglavje *Artificial Intelligence and the Prohibition of Discrimination in the EU: A Private Law Perspective*, ki se sklene s predstavitvijo mehanizmov t. i. algoritemske vladnosti, izpostavljajoč, da so najbolj problematične tiste odločitve, ki jih sprejmejo algoritmi, pri katerih zaradi uporabe neustreznega merila ali niza meril – samoučeči se algoritmi kot merila izberejo npr. narodnost, rasno ali etnično poreklo ter invalidnost – prihaja do prikrite diskriminacije.

Prvemu sklopu monografije, ki ga sestavljajo štiri poglavja, sledi drugi sklop s šestimi besedili, ki pod drobnogled postavljajo različne problematike, povezane s politikami upravljanja z umetno inteligenco in njenimi različnimi oblikami uporabe. *In Defence of Ethics and the Law in AI Governance: The Case of Computer Vision* obravnava presečišče pravne in etične skladnosti sistemov umetne inteligence na področju raziskav in razvoja s ciljem določiti predhodna etična in pravna tveganja na izbranem področju in z osrednjim pozivom, naj bosta pravo in etika vključena v oblikovanje politik in pripravo postopkov odločanja, pri čemer pa ne pozabi podati niti premisleka o odnosu med pravom in etiko. Odlika poglavja je njegova izrazita interdisciplinarnost, ki se med drugim odraža v navedbi družbenih in etičnih posledic orodij UI na številnih področjih, od bančništva in finančnega sektorja, zaposlovanja in sodnega sistema pa vse do izobraževanja, vključno z velikimi jezikovnimi modeli. Najpomembnejši poudarek izbranega poglavja je kritika

stališč, ki tehnologijo razumejo nevtralno. Tudi naslednje besedilo, *What Role for Ethics in the Law of AI?*, obdeluje področje etike, in sicer v navezavi na procese algoritemskega sklepanja ter na razmerje med etiko in pravom. Ob tem npr. podaja poglobljeno analizo evropskega pravnega okvira, za katerega avtorica pravi, da sloni na zaščiti temeljnih pravic, katerih skupno jedro je koncept človekovega dostojanstva. Temu doda zanimivo razpravo o razmerju med etiko, moralo in prevpraševanjem artificialnosti inteligence oz. uma, sklene pa ga s premislekom o razmerju med nezavezujočnostjo etičnih smernic in pravom.

Ali lahko govorimo o komputacijski etiki, katere interes so algoritmi za sprejemanje etično sprejemljivih odločitev? Raziskava *Introduction to Computational Ethics* prek dveh glavnih usmeritev – pristopov od spodaj navzgor in pristopov od zgoraj navzdol, ki jih tudi sicer najdemo v policy modelih digitalizacije – naslavlja prav vprašanje komputacijske etike, pri čemer podaja taksonomijo komputacijskih pristopov, analizo pa zaključí s predlogom hibridnih metod, ki združujejo pristopa od spodaj navzgor in pristope od zgoraj navzdol, kar naj bi omogočilo učinkovite modele odločanja. Razmerje zasebnih podjetij, ki so v aktualni ekonomskopolitični konstelaciji ključni razvijalci umetne inteligence, do mednarodnega prava človekovih pravic naslavlja študija *Artificial Intelligence and Human Rights: Corporate Responsibility Under International Human*. Četudi, kot nas pouči izbrano poglavje, je bil s sprejetjem (pravno nezavezujočih) mednarodnih standardov za varstvo človekovih pravic s strani zasebnih podjetij dosežen določen napredek v smeri pravne varnosti, še naprej ostajajo odprta nekatera vprašanja glede odgovornosti podjetij, ki razvijajo UI. Pričujoče besedilo sloni na normativnem pristopu, ki mu omogoča sistematično in dobro strukturirano analizo odnosa med zasebnimi podjetji, UI in človekovimi pravicami. Glede na to, da je – tako avtorica poglavja – možnost, da UI negativno vpliva na človekove pravice, očitna, po drugi strani pa mednarodni pravni okvir človekovih pravic (še) ne določa jasnih standardov za podjetja, ki razvijajo UI, je moč reči, da je osrednji poziv avtorice članka oblikovanje sodelovanja med praktiki, oblikovalci prava in politike ter akademiki – slednje naj bi bilo po njenem ključnega pomena za razjasnitev odgovornosti podjetij s področja umetne inteligence in prenos spoštovanja človekovih pravic podjetij v zvezi z umetno inteligenco v izvedljive postopke v praksi.

Svojevvrsten, a pomemben odmik od drugih vsebin monografije predstavlja predzadnje poglavje *As Above so Below: The Use of International Space Law as an Inspiration for Terrestrial AI Regulation to Maximize Harm Prevention*, v katerem se avtorja osredotočata na potencial UI za preprečevanje škode, ki izhaja iz degradacije okolja, vključujoč podnebne spremembe, onesnaževanje okolja, ekstremne vremenske razmere in naravne nesreče. Ker mnogi fenomeni naštetih področij obravnave segajo tako rekoč onkraj planeta, poglavje odpira tudi izzive vesoljskega prava. V tem oziru besedilo deluje celo multidisciplinarno, a poudarja, da bi lahko nekatera načela vesoljskega prava služila kot koristno vodilo pri razvoju mednarodnega pravnega okvira za uporabo umetne inteligence na Zemlji.

Rdeča nit monografije *Artificial Intelligence, Social Harms and Human Rights* sta res pravo in etika, a njeno teoretsko propulzivnost gre iskati prav v interdisciplinarni

metodološki nameri, ki jo kot tako smelo sklene sklepno, sociološko osnovano poglavje z naslovom *Democratizing the Governance of AI: From Big Tech Monopolies to Cooperatives*. Poleg navedbe teze o UI kot mestu kulminacije alienacije avtorici v besedilu na eni strani ponudita kritičen pregled različnih vidikov monopolizacije UI, npr. problem koncentracije raziskav in inovacij pod okriljem korporacij in elitnih univerz, na drugi strani pa s predlogom o preoblikovanju velikih tehnoloških podjetij v zadruga z lastninskimi deleži njihovih članic in članov poudarita nujnost demokratizacije UI. Prav demokrati-zacija in javno dobro ter ustrezen pravni in etični okvir pa so tako rekoč konceptualna veziva predstavljene monografije.