BORUT KRŽIŠNIK

Bratov Učakar 46, SI-1000 Ljubljana, Slovenia
Tel: +386/(0)5913-8118, M: +386/(0)41-905-139, borut.krzisnik@guest.arnes.si, www.borutkrzisnik.si

STROJ, KI PLEŠE

Bolj kot navdušenje nad spektakularnimi odkritji, ki jih dosega umetna inteligenca (UI), je zgodovinsko gledano pomembnejša perspektiva, ki se sprašuje, kakšna bo vloga človeka v prihodnosti, kjer bo UI vse bolj dominantna. Ker so zametki prihodnosti prisotni skozi vso zgodovino, si to vprašanje z enako veljavnostjo lahko zastavljamo tudi za sedanjost, ki jo živimo.

UI se velikokrat omenja v kontekstu, kjer nastopa kot tujek, vsiljen od zunaj, kot sovražna sila, ki je naravnana proti nam. To demoniziranje je pogosto prisotno tako v znanstveni-fantastiki kot v vsakdanjem življenju, vendar pa ne ustreza resničnem stanju - UI ni ne hudobna – je le neprimerno bolj kompetentna - niti demoni ne prežijo za vsakim vogalom, ampak so del nas samih. UI je naše delo in je podaljšana perspektiva človekovega intelekta. Čeprav temelji na enakih funkcionalnih principih kot naši biološki možgani, jih s svojo neprimerno večjo zmogljivostjo na vse večjem številu področij daleč presega in jih pušča za seboj ujete v svoje biološke omejitve. V tem smislu je UI nič drugega kot naslednja razvojna stopnja v evoluciji človekovih možganov.

Prav tako ne moremo omejevati UI zgolj na področje tehnologije in partikularnih odkritij. UI je nova paradigma, ki zajema praktično vse vidike človekovega obstoja in s katero človek dosega razumevanje najglobjih principov delovanja sveta. To novo obdobje bo uvedlo temeljite spremembe v delovanju družbe in življenju posameznika, spremembe, ki bojo po obsegu primerljive z industrijsko revolucijo. Tokrat se bo moral človek prilagoditi novim standardom, saj se bo prvič v zgodovini znašel v družbi bolj kompetentnih bitij ali strojev. UI bo vse bolj dominantna in bo prevzela mnoga delovna mesta, večinoma tista, ki zahtevajo rutino. In
eprav to predstavlja gospodarski problem, se lahko upravičeno vprašamo, ali niso ta delovna mesta itak bolj primerna za stroje in da bi za človeka bilo bolj primerno, da se ukvarja z bolj ustvarjalnimi in zanimivimi deli. Na ta način bo sobivanje z UI zahtevalo od človeka redefiniranje tega, kar smatramo kot “človeško”. V tem smislu konflikt med

UI in človekom ni boj za absolutno zmago z izključujočim izidom, ampak proces, ki vodi k integraciji. Tako si človek s tem, ko sam ustvarja UI in sobiva s čutečimi stroji odkriva, kdo in kaj pravzaprav je in je nenehno postavljen pred vprašanje:"Kaj bi rad postal?"

UI v svoji osnovi ni ne zlohotna ne dobra. Čeprav je pri specifičnih nalogah superiorna in obvladuje celotno zapuščino človekovega intelekta, je na lestvici t. i. »človeške inteligence«, kjer prevladuje splošna inteligentnost, vrednostni sistem in etika, še vedno na zelo nizki stopnji - nič višje kot na ravni insekta. Za razliko od bolj eksaktnih nalog, kjer se UI odlično izkaže tudi na najvišjih težavnostnih stopnjah, se pri razumevanju in definiranju moralnosti pojavijo hude težave. Oblikovanje ciljev, vrednostnega sistema in moralnih norm je največji kamen spotike pri oblikovanju varne in zanesljive UI, če naj ta postane človekov partner, in ne eksekutor. Poglavitni razlog za to so abstraktni pojmi, ki so pri tem v rabi in ki jih ni mogoče natančno in enoznačno opredeliti. Že Bertrand Russell je poudaril, da se nejasnosti teh pojmov niti ne zavedamo, preden jih ne poskusimo točno in nedvoumno definirati. Bolj kot to, da UI na zdajšnji stopnji razvoja še ni sposobna zadovoljivo razumeti človekove razlage moralnosti tako, kot si želi človek, je pomembno to, da s tem nerazumevanjem UI postavi človeka pred ogledalo. Pojmi, ki konstituirajo tisto, kar je človeku najpomembnejše, kot so smisel, pravica, empatija, vera, svoboda, ljubezen ipd., se izkažejo kot pojmi, ki jih ni mogoče natančno definirati. Vrednote, v katere človek verjame in ki oblikujejo pomenski okvir njegove lastne eksistence, se zmeraj in znova izkazujejo kot arbitrarne in nedoločljive – še več, inherentno neizrekljive.

Z bolj praktične perspective, če bi hoteli ustvariti moralno UI, nikakor ne bi mogli postaviti za vzgled človeka samega oziroma njegovo zgodovino, kjer kri teče v potokih. Če bi ji za kriterij ponudili današnje globalne medije in socialna omrežja, bi ta UI v roku 24 ur postala desničarski skrajnež, ki podpira fašistične prijeme. Glede na izjemno učinkovitost globalnih medijev vključno s socialnimi omrežji danes, vsako mnenje lahko zakotali sneženo kepo in povzroči plaz, neprimerljivo večji kot v predinternetnem času. Delež algoritemsko generiranih objav se lahko povzpne tudi do 50 % (npr. med volitvami). Svoboda govora v tovrstnem kontekstu dobi zelo vprašljiv predznak, saj ogromen delež vseh objav zavzamejo tiste, ki so elektronsko generirane, ki širijo neresnične, pristranske in premišljeno provokativne informacije in ki služijo enemu ali majhnemu številu političnih ali korporativnih subjektov. Ob epohalni manipulaciji javnega mnenja s strani Cambridge Analitice med zadnjimi ameriškimi volitvami smo bili priča popolni, v nebo vpijajoči nemoči pravnega sistema, da kakorkoli omeji tovrstna dejanja. Pomenljivo je – in alarmantno – da ni videti nobene politične volje, da bi se tovrstne manipulacije pravno regulirale, saj služijo interesom najvišjih finančnih in korporativnih elit, ki se aktivno borijo proti kriminalizaciji tega najmočnejšega manipulacijskega orodja v zgodovini.

Kako naj se UI potemtakem nauči človeške etike? Ker človek svojega moralnega kodeksa ni zmogel zadovoljivo definirati, obstaja realna verjetnost, da bi to veliko bolje opravila UI in sicer s samo-učenjem. Tako bi s svojo superiorno inteligenco nemara sama lahko ugotovila, kaj je za človeštvo najbolje. Nemara bi problem obravnavala z globalne in daljnoročne perspective - in ne parcialne in kratkoročne - ter ugotovila, da vojne, genocidi in uničevanje planeta niso dobri za človeštvo. Nemara bi ugotovila, da

je najbolje, če konstruktivno sodeluje med seboj, če pravično razdeli bogatstvo in živi v miroljubni koeksistenci. Nekaj podobnega je pokazala naslednja raziskava.

Leta 1981 sta Richard Axelrod in William Hamilton v reviji Science objavila članek z naslovom “Evolucija sodelovanja” (“Evolution of the Coolaboration”), ki je postal (tako Wikipedia) največkrat citiran članek na polju političnih znanosti. Raziskava v obliki specializirane računalniške simulacije proučuje kakšen način človekovega sodelovanja je najstabilnejši in dolgoročno najučinkovitejši in rezultati so pokazali, da je to strategija “konstruktivnega pogojnega sodelovanja”. Izsledki te raziskave neposredno naslavljajo vprašanje o najprimernejšem družbenem sistemu. “Pogojno sodelovanje” pooseblja državno regulirani kapitalizem – ta upošteva uspešnost prejšnjih sodelovanj in dolgoročno sodelovanje. Pri tem je pomembno, da je bolj stabilen tisti sistem, ki daje prednost dolgoročnemu sodelovanju z manjšim osebnim zaslužkom (lahko bi ga imenovali “konstruktivni kolektivizem”), kot pa sistem, ki omogoča večji zaslužek na račun zavrnitve sodelovanja (strategija “zgrabi denar in zbeži”). Strategija slednjega je sicer lahko uspešna, vendar samo kratkoročno; namenja prednost hitremu zaslužku in ne dolgoročnem sodelovanju ter pooseblja divji, liberalni kapitalizem. Podobno velja za pozitivno utopično strategijo »vedno sodeluj«.

Pred vsakim razmislekom o človekovi prihodnosti se moramo najprej vprašati, ali naj vzamemo za njen kriterij slepo evolucijo ali pa UI, ki razpolaga s celotno človekovo intelektualno doktrino in zmožnostmi, ki daleč prekašajo vsak človekov intelekt. Današnji razvoj z UI kot gonilno silo dere naprej z eksponentnim pospeškom – v primerjavi z njim se evolucija izkaže kot neskončno počasna, inertna in vse bolj očitno nerelevantna. Evolucija enostavno ni več aktualen način razvoja – to vlogo je prevzela UI. Prvič v zgodovini se je namreč razvilo bitje, ki bo s svojim intelektom sposobno samostojno upravljati z lastno evolucijo.