

Cum reflecta inteligența artificială natura umană: explorarea frontierelor eticii în era digitală



Recent, publicația The Guardian a acuzat compania Microsoft ca aceasta i-ar fi afectat reputația jurnalistică prin publicarea unui sondaj generat de inteligența artificială care specula asupra cauzei morții unei femei, plasat chiar lângă un articol al publicației de știri ([link](#)).

Serviciul de agregare de știri al Microsoft a publicat sondajul automat în imediata vecinătate a unui articol din The Guardian despre moartea unei antrenoare de polo în vârstă de 21 de ani, care a fost găsită decedată într-o școală din Sidney.

Sondajul, creat de un program de inteligența artificială, întreba: „Care credeți ca este cauza morții femeii?”. Cititorii erau apoi rugați să aleagă din trei opțiuni: crima, accident sau sinucidere.

Sondajul a generat indignarea cititorilor, dar și a conducerii cotidianului The Guardian, care a susținut că incidentul ar fi fost de natură să provoace suferința familiei îndurerate și că ar fi generat „prejudicii semnificative reputației” cotidianului The Guardian, precum și reputației jurnaliștilor care scriseseră articolul.

Acesta nu este nici primul și, probabil, nici ultimul incident de acest fel, și se circumscrie unui subiect mai amplu devenit extrem de important în ultimii ani în întreaga lume, și anume etica inteligenței artificiale - ce ar trebui să facem și ce nu ar trebui să facem cu inteligența artificială și ce ar trebui să permitem inteligenței artificiale să facă sau să nu facă pentru noi și în locul nostru.

Discuția aceasta aduce în prim-plan întrebări vechi de când lumea, care caută să identifice valorile umanității, ce este bine și ce este rău și cum ar trebui exprimate aceste valori pentru a le putea transmite și altora.

Experiența ne arată că nu ne putem rezuma doar la a introduce date în sistemele de inteligența artificială considerând că acest lucru este suficient ca ele să opereze singure. Dacă dorim ca aceste programe să fie instrumente cu adevărat utile și autonome și nu o sursă de riscuri și răspundere pentru cei care le creează și le utilizează trebuie, de asemenea, să le învățăm valorile noastre. Trebuie să introducem nu doar date, ci și compasiune, bunătate și respect pentru ceilalți. Succesul inteligenței artificiale pe viitor va depinde de cât de bine vom reuși noi, ființele umane, să găsim aceste valori în noi înșine și să le exprimăm în mod autentic.

Revoluția inteligenței artificiale reprezintă în mod clar o provocare tehnologică, pentru mediul de afaceri și pentru societate în general. Dar credem că, mai presus de orice, este o oportunitate de a privi în interiorul nostru și de a defini ce anume ne face să fim oameni. Și dacă însușirea acestor valori este cea care ne conferă natura umană, cum

putem cere inteligenței artificiale să le respecte, în timp ce noi înșine le încălcăm atât de des?

Unii cred că inteligența artificială este o amenințare, alții că este cea mai strălucită invenție umană, unii sunt entuziasmați de ea, alții sunt precauți, iar mulți alții simt toate aceste lucruri în diverse proporții. Dincolo de toate acestea și mai important decât ele, inteligența artificială este oglinda pe care o așezăm în fața noastră pentru a vedea cine suntem cu adevărat. Și poate că, în efortul nostru de a dezvolta și perfecționa inteligența artificială pentru a o transforma în instrumentul înțelept, plin de compasiune și bunăvoință pe care ni-l dorim, vom fi impulsionați să participăm în acest proces și să facem aceeași transformare în noi înșine.