Tehisajud, mis „mõtlevad“ ja „tunnevad“: miks me inimlikustame tehisaru mudeleid?

1 aasta tagasi - 04.10.2025 Autor: AM

The Conversation

Meie ees on ränikristallid, vaskviigud ja lõputud koodiread – see on tehisintellekti külm ja kalk reaalsus. Ometi räägime me sellest üha sagedamini kui olendist, kes „mõtleb“, „tunneb“ ja isegi „hoolib“. Tegemist on keelelise silmamoondusega, kus masinast saab meie endi peegelpilt, digitaalne teisik.

Ajakirjandus, püüdes keerulist tehnoloogiat rahvale arusaadavaks teha, on selle inimlikustamise laineharjal seilanud, kasutades võrdlusi meditsiinist, tuttavaid metafoore ja isegi düstoopilisi stsenaariume.

Aga mida me tegelikult võidame või kaotame, kui tehisintellekt lakkab olemast pelgalt tööriist ning muutub keeleliselt meie *alter ego´*ks?

Digitaalne aju

Hispaania ajaleht El País kirjeldas hiljuti Hiina tehisintellekti mudelit DeepSeek kui „digitaalset aju“, mis „näib üsna selgelt mõistvat oma sünni geopoliitilist konteksti“.

Selline keelekasutus asendab tehnilise žargooni – baasmudel, parameetrid, graafikaprotsessorid ja nii edasi – organiga, mida me kõik tunneme inimintellekti südamena. Tulemus on kahetine. Ühest küljest aitab see lugejal mõista masina poolt sooritatava ülesande („mõtlemise“) ulatust ja olemust. Teisest küljest aga loob see mulje, justkui oleks tehisintellektil oma „mõistus“, mis on võimeline tegema otsuseid ja mäletama konteksti – miski, mis on praegusest tehnilisest reaalsusest veel valgusaastate kaugusel.

See metafoor sobitub suurepäraselt George Lakoffi ja Mark Johnsoni klassikalise kontseptuaalse metafoori teooriaga, mis väidab, et me kasutame tuttavaid kontseptsioone selleks, et mõista abstraktset ja tundmatut. Tehisintellektist rääkides muudame me keerulise ja abstraktse võimekuse („statistiline arvutus“) millekski koduseks ja arusaadavaks („mõtlemine“).

Kuigi see on meie jaoks nähtuse mõistmisel abistav, on oht, et selline lähenemine hägustab piiri statistilise korrelatsiooni ja semantilise mõistmise vahel. See tugevdab illusiooni, et arvutisüsteemid suudavad midagi päriselt „teada“.

Tundevõimelised masinad

Veebruaris 2025 avaldas ABC uudisloo „emotsionaalsest tehisintellektist“, kus küsiti: „kas tuleb päev, mil nad on võimelised tundma?“ Artikkel kirjeldas Hispaania meeskonna edusamme, kes üritasid varustada vestlusroboteid „digitaalse limbilise süsteemiga“.

Siin muutub metafoor veelgi julgemaks. Algoritm ei oska enam pelgalt mõelda, vaid suudab ka kannatada või tunda rõõmu. See võrdlus dramatiseerib innovatsiooni ja toob selle lugejale lähemale, kuid kannab endas kontseptuaalseid vigu. Tunded on oma olemuselt seotud kehalise eksistentsi ja eneseteadlikkusega, mida tarkvaral olla ei saa. Esitledes tehisintellekti „emotsionaalse subjektina“, muutub lihtsamaks nõuda talt empaatiat või süüdistada teda julmuses. See aga nihutab moraalset fookust masina disaineritelt ja programmeerijatelt masinale endale.

Üks sarnane artikkel arutles: „kui tehisintellekt tundub inimlik, tunneb nagu inimene ja elab nagu inimene… mis tähtsust sel siis on, et ta on masin?“

Robotid, *kes* hoolivad

Humanoidseid roboteid esitletakse sageli just sellises võtmes. El País reportaaž Hiina püüdlustest luua eakate hoolduse androide ehk humanoidroboteid kirjeldas neid masinatena, mis (või kes?) „hoolitsevad oma vanurite eest“. Öeldes „hoolitsevad“, viitab artikkel perekondlikule kohustusele oma lähedaste eest hoolt kanda. Robotit esitletakse kui pereliiget, kes pakub emotsionaalset tuge ja füüsilist abi, mida varem pakkusid sugulased või hooldustöötajad.

See „hooldaja“ metafoor pole läbinisti halb. See legitimeerib innovatsiooni demograafilise kriisi kontekstis ning samal ajal leevendab tehnoloogilisi hirme, esitledes robotit mitte ohuna töökohtadele, vaid hädavajaliku abilisena personalipuuduse tingimustes.

Samas võib see varjutada eetilisi küsimusi, mis on seotud vastutusega, kui hooldustööd teeb erafirmade hallatav masin – rääkimata selle töövaldkonna niigi ebakindlast olemusest.

Arsti abiline

Ühes teises El País artiklis esitleti suuri keelemudeleid arsti abilise või „pikendusena“, mis on võimeline läbi vaatama meditsiinilisi andmeid ja pakkuma välja diagnoose. Metafoorid nagu „intelligentne skalpell“ või „väsimatu resident“ paigutavad tehisintellekti tervishoiusüsteemi usaldusväärse kaastöötaja, mitte asendaja rolli.

Selline hübriidne raamistik – mis pole ei elutu seade ega ka autonoomne kolleeg – soodustab avalikkuse heakskiitu, kuna see austab meditsiinilist autoriteeti, lubades samal ajal suuremat tõhusust. Kuid see avab ka arutelu vastutuse üle: kui „inimese pikendus“ teeb vea, kas süü lasub inimpersonalil, tarkvaral või seda turustaval ettevõttel?

Miks ajakirjandus metafoore armastab?

Need metafoorid on enamat kui lihtsalt ilukõne. Neil on vähemalt kolm eesmärki.

Esiteks, need lihtsustavad mõistmist. Sügavate närvivõrkude selgitamine nõuab aega ja tehnilist keelt, kuid „ajudest“ rääkimine on lugejale palju kergemini seeditav.

Teiseks, need loovad narratiivset pinget. Ajakirjandus elab lugudest, kus on peategelased, konfliktid ja lahendused. Tehisintellekti inimlikustamine loob need tegelased, koos kangelaste ja kurikaelte, mentorite ja õpipoistega.

Kolmandaks, metafoorid võimaldavad anda moraalset hinnangut. Ainult siis, kui algoritm sarnaneb subjektiga, saab teda vastutusele võtta või talle tunnustust avaldada.

Samas võivad just need samad metafoorid takistada avalikku arutelu. Kui tehisintellekt „tunneb“, siis on ju loogiline, et teda tuleks reguleerida sarnaselt kodanikega. Ja kui teda peetakse meist kõrgema intellektiga olendiks, tundub loomulik, et peaksime tema autoriteediga leppima.

Kuidas siis tehisintellektist rääkida?

Metafooridest täielikult loobuda oleks võimatu ja see pole ka eesmärk. Kujundlik keel on inimkonna viis tundmatut mõista, kuid oluline on seda kasutada kriitiliselt.

Sellepärast pakume välja mõned soovitused nii ajakirjanikele kui ka toimetajatele:

- **Lisage tehnilisi vastukaale.** See tähendab, et pärast metafoori tutvustamist selgitage lühidalt, aga selgelt, mida kõnealune süsteem tegelikult teeb ja mida mitte.
- **Vältige tehisintellektile absoluutse, inimlaadse tegutsemisvõime omistamist.** See tähendab, et fraase nagu „tehisintellekt otsustab“ tuleks täpsustada: kas süsteem „soovitab“? Kas algoritm „liigitab“?
- **Mainige vastutavaid inimallikaid.** Arendajate ja regulaatorite nimetamine tuletab meile meelde, et tehnoloogia ei teki tühjast kohast, vaid on inimeste arendatud.
- **Mitmekesistage metafoore.** Uurige vähem inimlikustavaid kujundeid – näiteks „statistiline mootor“, mis võivad lugu rikastada.

Kuigi tehisintellekti inimlikustamine ajakirjanduses aitab lugejatel keerulise tehnoloogiaga tuttavaks saada, muutub lihtsamaks projitseerida sinna ka oma hirme, lootusi ja vastutust, mida sarnasemaks masin meiega muutub.

Selle tehnoloogia edasise arengu käigus seisab nii ajakirjanike kui ka lugejate ees ülesanne leida õrn tasakaal metafoori köitva jõu ja mõistete täpsuse vahel, mida vajame, et pidada tuleviku üle informeeritud debatti.

Artikkel on avaldatud ajakirjast The Conversation Creative Commons´i litsentsiga.

Autorid:

Xosé López-García
Periodismo digital, comunicación digital, Universidade de Santiago de Compostela

Cristian Augusto Gonzalez Arias
Uurija, Pontificia Universidad Catolica de Valparaiso

- Tegijad

- Lahendused
- Tarkvara
- Tehisintellekt

Pilt