

Antropomorfizācijas ietekme uz studentu mentālajiem mākslīgā intelekta modeļiem

Kā pedagogiem, mūsu uzdevums ir palīdzēt jauniešiem attīstīt mentālos modeļus, kas viņiem ļaus orientēties pasaulē, kas arvien vairāk paļaujas uz mākslīgā intelekta (MI) sistēmām. Mākslīgā intelekta sistēmām bieži tiek piešķirtas cilvēciskas īpašības, piemēram, jūtas, garīgie stāvokļi un uzvedība^[1]. Mākslīgā intelekta **antropomorfizāciju** nostiprina attēlojumi daiļliteratūrā — filmās, TV un grāmatās —, kā arī ziņu medijos. Mediji bieži vien pēc noklusējuma izmanto smaidošu, cilvēkveidīgu robotu attēlus, lai ilustrētu stāstus par mākslīgā intelekta sistēmām^[1-1]. Arī nozares valoda pēc savas būtības ir antropomorfiska, sākot ar nosaukumu mākslīgais intelekts un beidzot ar mašīnmācīšanos, halucinācijām un modināšanas vārdiem.

Šos aprakstus var uztvert burtiski, ja studentiem trūkst stabila, izveidota mentālā modeļa par mākslīgā intelekta sistēmām^[2]. Mākslīgā intelekta sistēmu aprakstīšanai izmantotā valoda tās attēlo kā spējīgas “domāt” vai “spriest”, kas maldīgi attēlo šo sistēmu izmantotos procesus kā līdzīgus cilvēkam līdzīgam intelektam^[3]. Ja jūs uzskatāt, ka mākslīgā intelekta sistēmas spēj domāt, tad jūs varētu tām piedēvēt arī emocijas. Tas liek jums uzskatīt šīs mašīnas par tādām, kurām ir nodoms un morāls raksturs, vēl vairāk pastiprinot esošās bailes un nepareizos priekšstatus par šo sistēmu iespējām^[3-1]

Galvenie jēdzieni >

Antropomorfizācija ir prakse, ar kuru objektiem, kas nav cilvēki, tiek piešķirtas cilvēkam līdzīgas īpašības un uzvedība^[1-2].

Mākslīgais intelekts (MI) bieži tiek aprakstīts, izmantojot ļoti uz cilvēkam orientētu valodu.

Mākslīgā intelekta antropomorfizācija rada maldīgus priekšstatus par mākslīgā intelekta sistēmu darbību un to pašreizējām iespējām, kā arī zaudē lietotāju rīcībspēju^[3-2].

Pedagogi var izvairīties no mākslīgā intelekta sistēmu antropomorfizācijas,

- Nelietot MI kā skaitāmu lietvārdu, kā tas ir vārdos “MI”
- Vārdu, piemēram, “klausās”, “saprot” un “rada”, aizstāšana ar vārdiem “reģistrē”, “analizē” un “ģenerē”
- Lietojot terminu MI tāpat kā “bioloģija” vai “kiberdrošība”

Antropomorfizācijas riski

Ja jaunieši uzskatīs šo tehnoloģiju par dabiski cilvēcisku, mēs riskējam ietekmēt viņu...

- **Rīcības brīvības sajūta:** Mākslīgā intelekta izmantošanas apzinātā vienkāršība var radīt skolēniem domu, ka viņiem ir maza vai nekāda kontrole pār tā lomu un ietekmi uz viņu dzīvi un izvēlēm. Cilvēki var pakļauties sistēmām, ja tās uzskata par viedākām, nekā patiesībā tās ir^[4].
- **Drošība:** Kad mākslīgā intelekta sistēmas tiek uztvertas kā cilvēciskas, jaunieši ziņo par emocionālas saiknes sajūtu ar tām, kas var viņus pakļaut ietekmei un manipulācijām^[5].
- **Sociālā saikne:** Cilvēki jau tā biežāk antropomorfizējas, ja viņiem dzīvē trūkst sociālās saiknes vai ja viņi vēlas izprast neparedzamu vidi^[6].
- **Ziņkāre:** Ja mākslīgā intelekta sistēmas tiek uztvertas kā cilvēciskas, tas noved pie “melnās kastes” skatījuma uz lietojumprogrammām. Šāds abstrakcijas līmenis var atturēt cilvēkus no vēlmes uzzināt par to, kā mākslīgais intelekts darbojas, liekot viņiem to uztvert kā “maģiju”^[7].
- **Aizspriedumi:** Ja mākslīgā intelekta sistēmas tiek uztvertas kā cilvēkveidīgas, tās tiek uzskatītas arī par “baltām”. Šis skatījums uz mākslīgo intelektu nozīmē, ka daudzi krāsainie cilvēki jūtas atturīgi no mākslīgā intelekta aģentiem^[8]. Daudzas mākslīgā intelekta sistēmas ir arī antropomorfizētas kā sieviešu asistentes, tādējādi uzturot stereotipus par sievietēm kā pakļautām^[9].

Kad antropomorfizācija darbojas

Mākslīgā intelekta sistēmu antropomorfizācija sniedz priekšrocības izstrādātājiem gan ieviešanas, gan uzticēšanās ziņā. Kad lietotāji uztver rīkus kā cilvēciskus, viņi ir daudz atvērtāki attiecībā uz to integrēšanu savā

dzīvē un lietošanu^[4-1], tāpēc mākslīgā intelekta sistēmu izstrādātājiem var būt personiska interese, lai cilvēki tos uztvertu šādi.

Ir pierādījumi, ka antropomorfizācija palīdz jaunākiem skolēniem izprast sarežģītas parādības zinātnē. Padarot sistēmas cilvēcīgākas, nekā tās ir patiesībā, mēs varam pastāstīt par tām stāstus, kas var būt noderīgi, apgūstot sarežģītus jēdzienus.

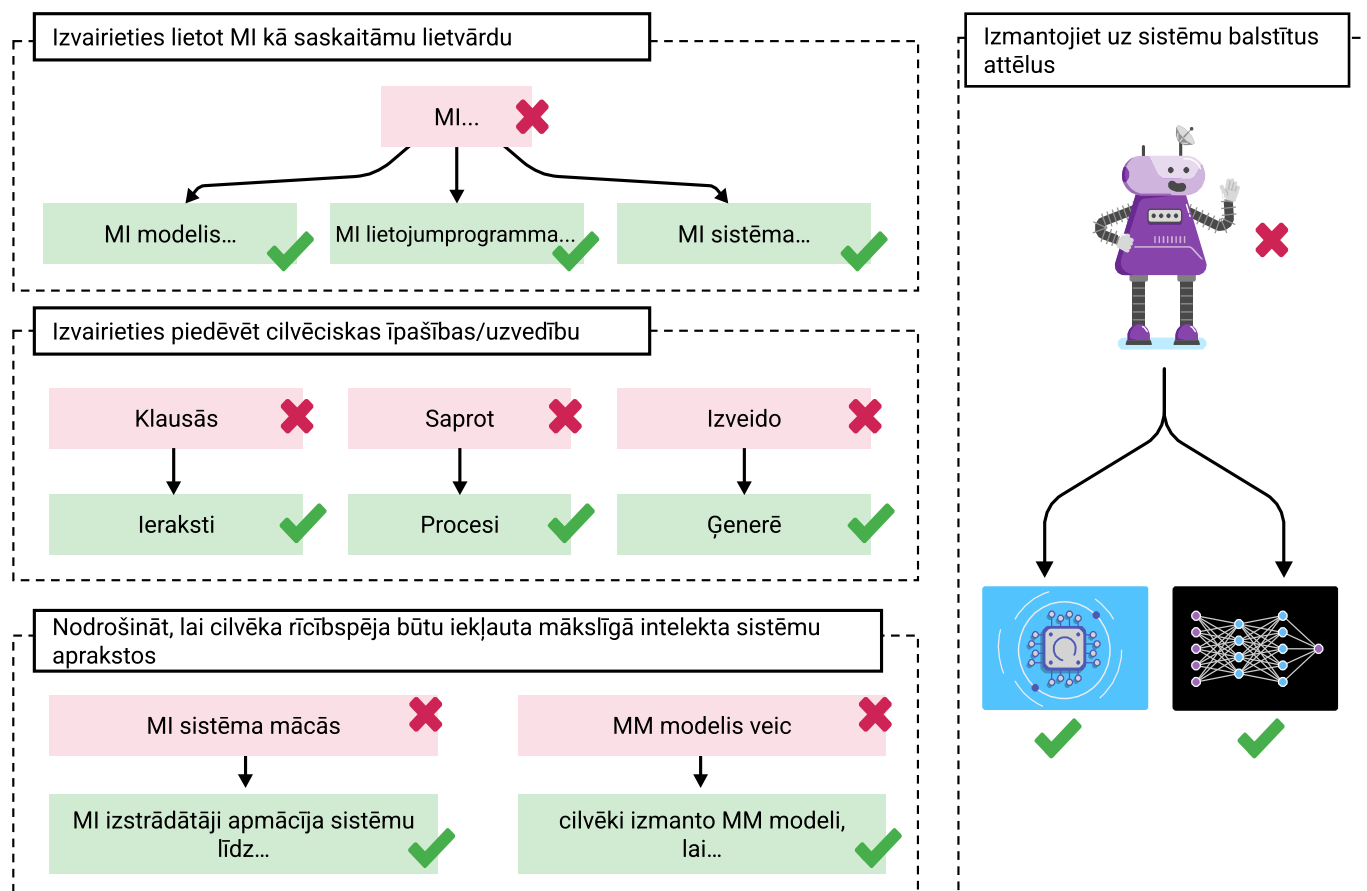
Kopumā antropomorfizācijas lietderība ir tieši saistīta ar to, cik tuvu kaut kas ir faktiskajam intelektam. Piemēram, tas ir noderīgi, mācoties par plūdmaiņām, bet, mācoties par pērtiķiem, tas var radīt maldīgus priekšstatus^[2-1]. Mākslīgā intelekta sistēmas tiek pasniegtas kā ārkārtīgi tuvas cilvēka intelektam, tāpēc antropomorfizācija ir īpaši bīstama.

Kā izvairīties no mākslīgā intelekta sistēmu antropomorfizācijas

Lai palīdzētu studentiem izstrādāt pietiekami noturīgus mentālos modeļus, lai saprastu, kā mākslīgā intelekta sistēmas tiek attēlotas kā cilvēkam līdzīgas, jūs varat:

- Izvairieties lietot mākslīgo intelektu kā saskaitāmu lietvārdu. Tā vietā, lai kaut ko sauktu par “mākslīgo intelektu”, izmantojiet apzīmētāju, piemēram, “mākslīgā intelekta sistēma”, “lietojumprogramma” vai “modelis”. Lietojot terminu “mākslīgais intelekts”, lietojiet to tāpat kā “bioloģija” vai “kiberdrošība” — kā studiju un prakses jomu, nevis kā kaut ko, uz ko var norādīt.

- Izvairieties no cilvēku valodas lietošanas, lai aprakstītu mākslīgā intelekta lietojumprogrammu darbību. Tā vietā izmantojiet uz sistēmu balstītu valodu. Aizstājiēt tādus vārdus kā “klausās”, “saprot” un “rada” ar “ieraksta”, “apstrādā” un “ģenerē”.
- Nepiešķiriet mākslīgā intelekta sistēmai pārāk lielu rīcības brīvību, to aprakstot. Tā vietā, lai teiktu “mākslīgā intelekta sistēma mācās” vai “mašīnmācīšanās (ML) modelis to dara”, iekļaujiet vienādojumā cilvēku. Piemēram, “Mākslīgā intelekta izstrādātāji apmācīja sistēmu līdz...” vai “cilvēki izmanto mašīnmācīšanās modeli līdz...”.



Ieteicamā valoda un attēli, lai izvairītos no antropomorfisma

Piemēri

Šeit ir daži praktiski piemēri, kā aprakstīt mākslīgā intelekta sistēmas, neizmantojot antropomorfizētus aprakstus:

Viedais runātājs

“Viedais runātājs tiek aktivizēts ar tādu frāzi kā “Hei, Google”.” Pēc tam tas ieraksta jūsu balsi un nosūta to uz datoru, izmantojot internetu.

Šis dators analizē ierakstu un paredz atbildi uz jūsu pieprasījumu, piemēram, mūzikas atskaņošanu vai atbildi uz jautājumu. Runātājs saņem paredzēto atbildi un atskaņo to jums

###Liels valodas modelis (LLM)

“Liels valodas modelis (LLM) ir mākslīgā intelekta sistēma, kas ģenerē tekstu, reaģējot uz lietotāja uzvedni.” Kad tiek ievadīta uzvedne, sistēma to sadala mazākos gabalos un salīdzina ar plašu teksta datu kopu, uz kuras ir apmācīts LLM. LLM izmanto apmācības datus esošos modeļos un

attiecības, lai novērtētu uzvedni. Pēc tam atbilde tiek ģenerēta, apvienojot vārdus un frāzes, kas, domājams, ir atbilstošas."

Atsauces

[Avota pdf fails](#)

1. Salles, A., et al. (2020). Antropomorfisms mākslīgajā intelektā. [the-cc.io/qr22_1](#) ↩ ↩ ↩
2. Tang, X., & Hammer, D. (2024). "Es par to domāju šādi, un tas man palīdz saprast": antropomorfisms sākumskolas skolēnu mehāniskajos stāstos. [the-cc.io/qr22_2](#) ↩ ↩
3. Placani, A. (2024). Antropomorfisms mākslīgajā intelektā: ažiotaža un maldi. [the-cc.io/qr22_3](#) ↩ ↩ ↩
4. Huang, J., et al. (2024). Kad antropomorfisms sāp? Kā instrumentu antropomorfisms negatīvi ietekmē patērētāju atlīdzību par instrumentu lietotājiem [the-cc.io/qr22_4](#) ↩ ↩
5. Williams, R., et al. (2023). Mākslīgais intelekts + ētikas mācību programmas vidusskolas jauniešiem: mācības, kas gūtas no trim uz projektiem balstītām mācību programmām. [the-cc.io/qr22_5](#) ↩
6. Alabed, A., et al. (2022). Mākslīgā intelekta antropomorfisms un tā ietekme uz lietotāju paškongruenci un pašintegrāciju ar mākslīgo intelektu: teorētiskais ietvars un pētījumu programma. [the-cc.io/qr22_6](#) ↩
7. Tedre, M., et al. (2016). Mašīnmācīšanās mācīšana datorzinātņu izglītībā K-12 klasēs: potenciāls un trūkumi. [the-cc.io/qr22_7](#) ↩
8. Cave, S., & Dihal, K. (2020). Mākslīgā intelekta baltums. [the-cc.io/qr22_8](#) ↩
9. Barau, S. (2024). Maldināšana, diskriminācija un objektivizācija: sieviešu mākslīgā intelekta aģentu ētiskie jautājumi. [the-cc.io/qr22_9](#) ↩