

Kuidas võib tehisaru muuta teadlase elu?

Marek Tamm

Suured keelemudelid ja nende rakendused, nagu ChatGPT või Gemini, on kujundamas ümber paljusid elu- ja töövaldkondi. Üha teravamalt on üleval küsimus, mil moel ja määral mõjutavad suured keelemudelid teadlase igapäevast tööd.

De Gryteri teaduskirjastus viis mullu suvel läbi ulatusliku küsitluse tehisaru kasutamisest teadustöös (AI Adoption Trends Among Scholars), millele vastas 748 teadlast 82 riigist. Üle poole vastanutest (57 protsenti) kinnitas, et tunneb „mõnevõrra“ tehisaru tehnoloogiat. „Väga hästi kursis olevaks“ hindas ennast 28 protsenti loodus- ja täppisteaduste valdkonna esindajat, samas kui humanitaar- ja sotsiaalteadlaste seas oli see näitaja 19 protsenti.

Küsimusele, mil moel ChatGPTd oma teadustöös kasutatakse, vastas 31 protsenti teadlastest, et ilmunud teadusartiklite kokkuvõtete koostamiseks, 27 protsenti kasutab seda kirjandusülevaadete tegemiseks, 24 protsenti uurimisideede ja -hüpoteeside arendamiseks, 23 protsenti õppetöös (nt loenguplaanide koostamiseks), 22 protsenti artikli akadeemiliseks vormistamiseks ning 20 protsenti teadusartiklite, uurimistaotluste jt akadeemiliste tekstide kirjutamiseks.

Praeguseks on valminud ka esimesed empiirilised uurimused suurte keelemudelite kasutamisest teaduspublikatsioonides. Ühes tänavu aprillis avaldatud artiklis uuriti pea miljonit loodus- ja täppisteaduslikku artiklit, mis on ilmunud viimasel neljal aastal. Eesmärk oli statistilise analüüsi toel tuvastada jälgi suurte keelemudelite kasutamisest. Kuigi tulemused on ositi oletuslikud, kinnitab uuring, et nt arvutiteaduste valdkonnas on suurte keelemudelite kasutamine kasvanud ligemale 17 protsenti, samas kui nt matemaatikas on tõus 6 protsenti kandis.

Analoogne uurimus ligemale 30 000 trükis ilmunud masinõppe alase konverentsiettekande kokkuvõtte ja 1000 suulise ettekande põhjal kinnitas, et suurte keelemudelite kasutamine on valdkonnas laialt levinud.

Ridamisi leiab viimasest paarist aastast artikleid, mis testivad ühes või teises teadusvaldkonnas suurte keelemudelite kasutamise võimalusi. Ühe varase näitena võib mainida Signe Altmäe ja kolleegide 2023. aasta juulis ilmunud artiklit „Tehisaru teadustöö kirjutamisel: sõber või vaenlane“, mis jõuab järeltulejatele, et ChatGPT-l on vähemalt biomeditsiini valdkonnas suur potentsiaal, et aidata teadlasi uurimisprojekti kavandamisel, analüüside teostamisel ja tulemuste vormistamisel.

Need näited, mille nimekirja saaks veel pikalt jätkata, osutavad kõik põhimõttelisele küsimusele – milline saab olema teadlase roll tehisaru ajastul? Ma näen ise terendamas kahte stsenaariumi: tehisaru on teadlase abiline, kes võimestab tema senist tööd, või tehisaru on teadlase asemik, kes jätab teadlase enda assistendi rolli.

Kui siiani on tehisaru toiminud ennekõike abilisena, aidates teadlasel koguda vajalikku informatsiooni, sünteesida senist uurimisseisu, toimetada ja tõlkida tekste, puhastada ja sorteerida andmeid, viia läbi erisuguseid kvantitatiivseid analüüse jms, siis üha rohkem näeme katseid delegeerida järjest ulatuslikum osa

uurimistööst generatiivse tehisaru rakendustele.

Tallinna Ülikooli kolleeg Andres Karjus on näiteks hiljuti avaldanud artikli masinvoimestatud meetodikast humanitaar- ja sotsiaalteadustes, mis näitab, et suure osa senisest valdkondlikust uurimistööst saab väga edukalt delegeerida tehisarule.

Üks ambitsioonikamaid katseid samast vallast on OpenAi arvutiteadlase Chris Lu ja tema kolleegide projekt The AI Scientist (Tehisaru Teadlane), mille eesmärk on välja arendada tehisarupõhised toimijad (agendid), kes suudaksid ise välja pakkuda uusi uurimisideid, kirjutada vajaliku koodi, teha eksperimente, visualiseerida tulemusi ja vormistada töö tulemused nõuetekohases teadusartiklis, mida rakendus on võimeline ühtlasi ise ka retsenseerima. Projekti esimesed tulemused on võrdlemisi paljulubavad.

Need uued väljavaated viitavad minu hinnangul selgele vajadusele leppida võimalikult kiiresti ja üleilmselt kokku tehisaru kasutamise reeglid teadustöös, alustades läbipaistvuse põhimõttega, mis lubab lugejatel mõista, millist rolli on tehisaru konkreetses teadustöös etendanud. Sellega seonduvalt on vajalik välja töötada tehisaru eetika, seda laiemas teaduseetika raamistuses.

Eriti väiksemate keelte ja kultuuride seisukohalt on kõige pakilisem küsimus tehisaru keelelis-kultuurilisest joondamisest, s.t. tehisaru suutlikkus arvestada kohalikku konteksti ja suhelda teisteski keeltes samal tasemel kui inglise keeles. Eriti humanitaar- ja sotsiaalteadustes sõltub suurte keelemudelite viljakas kasutamine otseselt kasutatud treeningandmete iseloomust. Ja viimaks on kõige selle juures vajalik senisest paremini teadvustada tehisaru kasutamise ökoloogilist jalajälge, et me uute teadmiste järel ei muudaks oma elu planeedil võimatuks.

[Ilmunud ajalehes Postimees 4. jaanuaril 2025](#)