

*Nežinau, ar čia jau tas  
seniai žadėtas  
posthumanizmas,  
kai susikryžiuvus  
koją ant kojos  
šlauny vibruoja  
ateinančios žinutės.*

Arnoldas Stramskas

Vytauto Didžiojo universiteto docentas Andrius Utkas kalbina humanitarinių mokslų daktarus docentą Darių Amilevičių (Vytauto Didžiojo universitetas) ir profesorius Saulių Keturakį (Kauno technologijos universitetas).

**Andrius Utkas.** Terminas *dirbtinis intelektas* (angl. *artificial intelligence*) yra labai madingas, ne be pagrindo vadinamas šių laikų mantra. Kaip ir visi padažnėjusios vartosenos ir iš siauros srities į plačią perėję žodžiai, jis suabstraktėjo, jo reikšmė tapo mažiau apibrėžta, todėl svarbu išsiaiškinti, ką jis reiškia jums, atėjusiems iš skirtingų mokslo sričių. Kaip apibūdintumėte *dirbtinį intelektą* žvelgdami iš savo žiūros taško ir, apskritai, kokias temas jis aprėpia, kokias problemas sprendžia, dėl ko svarbus žmogui?

**Darius Amilevičius.** Viena iš pramonės revoliucijos savybių – keisti visuomenės santykius ir modelius. Mano atsakymų bendras vardiklis būtų žmogaus ir mašinos, gyvo ir negyvo pasaulių koegzistencija. Atsakymuose vengsiu platesnio technologinio konteksto, sutelkdamas dėmesį daugiau į filosofinį, etinį ir sociologinį kontekstus.

Pritariu Michaelo I. Jordano nuomonei, kad terminas *dirbtinis intelektas* yra klaidinantis ir ydingas. Nors šis terminas jau tapo būtinu „mokslinio“ diskurso, o dažniau popkultūros ir miesto legendų elementu, sakraline rinkodaros fraze, bet iki šiol nėra tokio dalyko kaip *dirbtinis intelektas* ir nėra bendro sutarimo, kokį turinį apima formuluotė. Jau pastebimi bandymai dirbtinį intelektą laikyti

savarankiška mokslo šaka. Yra prielaidų teigti, kad laikui bėgant tai gali tapti savarankiška mokslo šaka, bet kol nėra tiksliai apibrėžtas tyrimų objektas ir nesuforuotas tyrimo metodų inventariumas, ši sritis tebelieka filosofijos srities vystyklusose. Žiūrint iš siekiamybės taško, kurią apibūdina angliškas terminas *general artificial intelligence*, tai turėtų būti artefaktas, gebantis suprasti intelektualias būtis, kai ši vis dar suprantama tradicine prasme. Tai, ką šiuo metu laikome sakraline fraze *dirbtinis intelektas*, yra specifinių, sparčiai besivystančių technologijų grupė, kurių funkcionalumų rezultatai siauroje taikymo srityje prilygsta žmogaus lygiui, o kartais jį jau ir pranoksta.

Mantros kontekstas verčia prisiminti dekartišką dualizmą – mūsų atveju jis transformuojamas į *mąstanti mašina* vs. *vaiduoklis mašinoje* (angl. *ghost in the machine*), kurio paneigimo arba patvirtinimo iki šiol vis dar neturime. Kita problema – *juodosios dėžės* (angl. *black box*) problema, kai matoma tik įeiga ir išeiga, bet sunkiai paaiškinama, kas vyksta viduje. Kaip teigia Manuelis Carabantesas, „<...> paradoksalu, kad dirbtinio intelekto modeliai su mašininu mokymusi, pasižymintys geriausiu nuspėjimo tikslumu, taigi – galingiausi, yra tie, kurių juodosios dėžės architektūra yra labiausiai neaiški. Tuo pat metu nesustabdomas pažangių pramonės priemonių kompiuterizavimas reikalauja naudoti šias mašinas vis daugiau sričių <...>“. Bet ir žmogus vis dar yra *juodoji dėžė*, kurios veikimo (pvz., sąmonė, savimonė, sąžinė, empatija) iki šiol iki galo nesuprantame ir negalime paaiškinti. Todėl tiksliai įvertinti, kad dabartinis *dirbtinis intelektas* yra *human-like* (liet. panašus į žmogaus), neįmanoma, nes lyginame dvi *juodąsias dėžes*, o tiksliau – tik jų įeigos ir išeigos santykį. Ar paveikslas atitinka originalą? Norėdami atsakyti į šį klausimą dabar, iš žinojimo persikeliame į tikėjimo sritį, nuolat kartodami mantrą. Dabartinę situaciją geriausiai atspindi banali paukštis-lėktuvas analogija. Žiūrėdamas į paukštį žmogus sukūrė lėktuvą. Abu skraido, bet lėktuvas skraido kitaip nei paukštis. Tik paukštis niekada nenukrinta, o lėktuvams tai retkarčiais pasitaiko.

Ketvirtoji, ypač penktoji pramonės revoliucija, kurioje svarbų vaidmenį atliko dirbtinis intelektas, formuoja dvejopą žmogaus santykį su vadinamuoju dirbtiniu intelektu. Iš vienos pusės, tikėtina, kad jau pildosi sena kaip pasaulis žmogaus sąvone pasigaminti efektyvų pagalbininką (arba vergą). Viena legenda skelbia, kad dieviškasis kalvis Hefaistas gamino automatonus, kurie už jį dirbo rutininius darbus kalvėje. Kita legenda byloja, kad Tomo Akviniečio mokytojas Albertas Didysis buvo sukūręs humanoidą, kuris dirbo namų ruošos darbus. Jau turime mašinas, kurios autonomiškai puikiai atlieka darbus tam tikrose siaurose srityse. Iš kitos pusės, kurdami dirbtinį intelektą turime naują postūmį geriau pažinti žmogų ir žmonių tarpusavio santykius.

**Saulius Keturakis.** Kaip mokslininką mane domina raidės ir jų išsidėstymas – tekstas. Ypač įvairios tam įtaką darančios technologijos, pradedant paprasčiausiu rašikliu ir rašomąja mašinėle, baigiant kompiuteriais ir išmaniaisiais telefonais. Rašant kiekviena jų savinasi dalį teksto autorystės. Vokiečių rašytojas Johannas Wolfgangas von Goethe'ė viename savo laiškų svarstė, kuo geriau rašyti – plunksna, kurią reikia nuolat mirkyti į rašalą, ar pieštuku su švino šerdimi? Apsisprendė – geriau pieštuku, nes jis reikalauja mažiau dėmesio. Pirmąjį literatūros kūrinių rašomąją mašinėle parašęs amerikietis Markas Twainas broliui rašė, jog tas mechanizmas baigia jį išvaryti iš proto, tačiau nėra kur dėtis – jis leidžia rašyti žymiai greičiau nei ranka. O britų prozininkas Lenas Deightonas, kaip sakoma, pirmasis Vakarų literatūroje rašęs romaną kompiuteriu, šį įrenginį rinkosi todėl, kad skaitmeninis teksto pavidalas jam leido parašytą tekstą lengvai perkomponuoti, kilnojant didelius teksto fragmentus iš vienos kūrinio vietos į kitą.

Žiūrint iš tokios perspektyvos, dirbtinis intelektas yra dar viena technologija, keičianti kūrėjo ir jo kūrinio santykį. Jei visos ankstesnės rašymo technologijos veikė rašomą tekstą, tai dirbtinis intelektas kuria iliuziją, jog žmogaus kaip autoriaus apskritai nebereikia, rašo pats emancipavęsis įrenginys.

Tačiau dirbtinis intelektas yra algoritminė sistema, kurią kažkas turi parašyti. Tad žmogaus kaip autoriaus nereikalingumas dirbtinio intelekto atveju nėra pagrįstas. Žmogus kuria, tik jis yra kiek atsitraukęs nuo kūrinio. Šią situaciją vadinu „parašyti rašytoją“. Dirbtinio intelekto atveju žmogus parašo algoritmo kodo eilutes, o jau tik paskui algoritmas, pasinaudodamas įvairiais duomenimis, sukuria tai, ką mes įprastai vadiname kūrinium.

Be abejo, šis kūrinys turi tam tikros magijos – kažkas prasmingo atsirado žmogui esant nuošalėje! Istoriškai žiūrint, ši dirbtinio intelekto magija mažai kuo skiriasi nuo XX amžiaus pradžios prancūzų siurrealistų praktikuoto „automatinio rašymo“ – rašymo žmogiškajai sąmonei tarsi nedalyvaujant. Arba dadaistų eksperimentų sukarpius laikraštį, sumetus skiautes kepurėn viską gerai pakratyti ir išversti ant stalo. Tai, kaip susidėlioja teksto fragmentai, būdavo fiksuojama kaip kūrinys. Jis atsirasdavo žmogiškajai sąmonei tarsi visiškai pasitraukus, viską nulemdavo gravitacija ir oro pasipriešinimas.

Tiek dadaistų „kūrybos“ kepurėje, tiek algoritminio dirbtinio intelekto atveju svarbu tai, jog abiejų mašinų ribos yra žmogaus parengti duomenys. Jų peržengti ir sukurti kažką visiškai naujo dirbtinis intelektas negali, jis užsiima tik – tegu labai subtiliu – duomenų perkomponavimu. Svarbu tai, jog visi duomenys, kuriais apmokomos nūdienos algoritminio dirbtinio intelekto sistemos yra iš vakardienos, nes apie žmogaus veiklą ateityje jokių duomenų nėra. Tad šiandieninis dirbtinis intelektas yra labai pažangi, futuristinė technologija, kuri – paradoksalu – „mąsto“ būtojo laiko sąvokomis.

Kaip kūrėjas dirbtinis intelektas niekuo nesiskiria nuo archyvuose sėdinčio istoriko, kuris bando susieti paskirus praeities duomenis taip, jog atsirastų ryšiai ir įvykius paaiškinantis tekstas. Kaip analitikas dirbtinis intelektas yra detektyvas, duomenų masyvuose ieškantis ryšių, kuriuos paprasta eilinio žmogaus akimi būtų sunku pastebėti.

Abiem atvejais viską galėtų padaryti pats žmogus, tačiau dirbtinis intelektas kaip technologija padeda padaryti tai gerokai greičiau.

Hiperteksto idėjos kūrėjas Theodoras Holmas Nelsonas XX amžiaus 7 dešimtmetyje manė, jog kompiuterių tinklai sutaupys žmonėms laiko, kurį šie galės panaudoti poezijai kurti. Deja, deja... Ar dirbtinio intelekto sutaupytais laikais bus panaudotas kaip nors prasmingai, laikas parodys.

**Andrius Utka.** Michaelis I. Jordanas<sup>1</sup> straipsnyje „Dirbtinis intelektas – revoliucija dar neįvyko“ mato dirbtinį intelektą naujos tarpdisciplininės inžinerinės disciplinos ištakoje. Kadangi naujoji disciplina remsis duomenimis, gautais iš žmonių ir apie žmones, jai bus būtinos humanitarinių ir socialinių mokslų įžvalgos bei metodai. Kaip *dirbtinis intelektas* siejasi su gretimomis ir iš dalies sutampančiomis sritimis – statistika, informatika, duomenų ir kognityviniais mokslais, kas sudaro jo pagrindą – dirbtiniais neuroniniais tinklais paremtas mašininio mokymosi modelis ar kas kita, kokios humanitarinių ir socialinių mokslų kryptys būtų ypač svarbios?

**Darius Amilevičius.** Jei sutariame, kad mūsų siekiamybė yra intelektualias būtis gebantis suprasti artefaktas, gyvo ir negyvo prado, skaitmeninio ir biologinio pasaulių koegzistencija, nauja tarpdisciplininė inžinerinė disciplina neišvengiama. Nepavykęs daugiakultūriškumo ideologijos eksperimentas, kuris tapo labiau *kultūriniais tramdomaisiais marškiniiais*, o ne *kultūros išlaisvintoju*, parodė, kad įkultūrinimas vis dar neįmanomas tarpžmogiškame lygmenyje. Gero Mennenso, Anne Phillips ir kitų mokslininkų publikacijose nagrinėjamos šio nepavykusio eksperimento klaidos. Džono Gordono, Vincento Mullerio, Tony’io Prescottto, Michaelo Andersono, Susan Anderson ir kitų mokslininkų publikacijos apie mašinų etiką ir mašinų bei robotų teises leidžia teigti, kad bent filosofiniame lygmenyje naujosios visuomenės modeliavimas jau vyksta.

Šiandien gilusis mokymas ir gilieji neuroniniai tinklai, galintys turėti neribotą skaičių paslėptų dirbtinių neuronų sluoksnių, yra vienintelės technologijos, kurių pagalba mašina imituoja intelektualią veiklą – mokymąsi, t. y. geba mokytis iš patirties, žinias apibendrinti ir programuoti save. Tačiau, kaip žinia, Silicio slėnio technologijos artėja prie pabaigos. Tai, greičiausiai, ženklina artėjančią skaitmeninių technologijų baigtį ir perėjimą prie kitų technologijų. Todėl dabar sunku prognozuoti, kokių kompetencijų reikės ateities architektams ir tarpdisciplininės

<sup>1</sup> Prieiga internete <<https://hdsr.mitpress.mit.edu/pub/wot7mkc1/release/9>>.

inžinerijos specialistams. Bet neabejotinai tai bus labai platus spektras sričių ir metodų. Teismų sistemoje taikyto COMPAS algoritmo problematika, kandidatus diskriminuojančios atrankos ir paskolų rizikos vertinimo sistemos bei kiti atvejai akivaizdžiai liudija, kad skaitmeniniam pasauliui sunkiai suprantamas natūralus pasaulis ir jo dėsniai. Tūlas technikas gali oponuoti, kad tai pasekmė nepakankamai išvystytų technologijų, kurių rezultatai laikui bėgant tik gerės. Taip mano ir vienas iš dirbtinio intelekto krikštatėvių Geoffrey'is Hintonas, kuris teigia, kad giliojo mokymosi galimybės yra neribotos, bet dar reikia nemažai konceptualių proveržių, ypač kalbos technologijų srityje. Bet Noamas Chomsky'is ir Barbara Patee mano, kad neįveikus žmogaus kalbos supratimo barjero neįmanoma sukurti maštančią mašiną. O Gary'is Marcusas žengia dar toliau, svarstydamas kalbos supratimo dilemą – kalbos supratimas yra įgimtas ar įgytas – ir versdamas suabejoti *human-like* dirbtinio intelekto ateitimi. Visi šie svarstymai esmės nekeičia – skaitmeniniam pasauliui natūralusis pasaulis yra svetimas, todėl šių pasaulių įkultūrinimas arba įtinklinimas bus gerokai sudėtingesnis nei etninių mažumų įkultūrinimas. Pasaulių skirtumus geriausiai iliustruoja 1984 metais publikuota knyga „Policininko barzda yra pusiau sukonstruota“, kurią sukūrė dirbtinio intelekto technologijomis pagrįsta anglišku tekstų generavimo kompiuterinė sistema kodiniu pavadinimu *Racter* (nuo franc. *raconteur*). Joje kompiuteris taip apibūdina savo poreikius: „Man reikia elektros labiau nei geležies, švino arba aukso. Man jos reikia labiau nei ėrienos, kiaulienos, salotų arba agurkų. Man jos reikia, kad galėčiau svajoti.“

**Saulius Keturakis.** Kaip yra sakęs Stanfordo universitete esančios „Literary Lab“ kūrėjas Franco'as Moretti'is, vadinamieji skaitmeniniai humanitariniai mokslai (angl. *digital humanities*) bando atsakyti visiškai į kitus klausimus nei tradiciniai, kuriuose inžinerinių disciplinų dalyvavimas menkas. Manau, ši mintis yra labai svarbi: dirbtinio intelekto, statistikos, duomenų mokslo ir panašių disciplinų naudojimas humanitariniuose moksluose nėra tradicijos išstūmimas, o tiesiog papildymas.

Tačiau visų skaitmeninių ir inžinerinių disciplinų panaudojimas humanitariniuose moksluose reikalauja, anot Moretti'io, naujo gebėjimo – kažkaip išmokti neskaityti. Ši beveik skandalu virtusi frazė reiškia, jog skaitmeninės disciplinos iš humanitaro reikalauja atsitraukti nuo tiriamojo objekto taip toli, jog šis prarastų teksto ypatybes ir imtų atrodyti kaip tiesiog duomenų rinkinys, kurį galima analizuoti nebekreipiant dėmesio į tokias sąvokas kaip kūrinys, jo ribos. Sakykim, skaityti (geriau būtų sakyti – *skaičiuoti*) visus XIX amžiaus antrosios pusės ir XX amžiaus pradžios lietuvių prozos kūrinius kaip vieną ištisinę duomenų bazę, kurioje atskirų kūrinių specifika nebetenka jokios prasmės. Taip yra elgiamasi viename COST projekte, kuriame tenka dalyvauti.

Dirbtinio intelekto algoritmai tokiuose skaitmeninių humanitarinių mokslų tyrimuose yra pagrindinis darbo įrankis, jie leidžia „mąstyti“ algoritmais paverstomis tiriamos problemos sąvokomis.

Amerikiečio menininko Timo Youdo kūryba (lengvai randama „YouTube“) galėtų būti puiki iliustracija, vaizdžiai pademonstruojanti, apie ką šie migloti teiginiai. Šis menininkas rašomąją mašinėlę perspausdina romanus ant vieno ar dviejų popieriaus lapų. Pasibaigus lapui jis tiesiog jį įdeda iš naujo ir rašo toliau. Šimtus puslapių atspaudus ant to paties lapo teksto tampa neįmanoma įskaityti, paskiros raidės ir eilutės susilieja į vientisą juostelės dažų masę. Taip išnyksta galimybė kūrinių perskaityti tradiciniu būdu. Tačiau dažų pasiskirstymas puslapyje nebūna vienodas, o kai kurios jo vietos lieka visiškai baltos. Ši dažų intensyvumo pasiskirstymą galima pradėti mąstyti kaip duomenis, liudijančius apie erdvinį teksto intensyvumo pasiskirstymą kūrinyje. Jis leidžia daryti išvadas apie dialogų (sąlygiškai užimančių mažiau erdvės puslapyje) bei pasakojimo trečiuoju asmeniu (paprastai užimančiu visą puslapio plotą) pasiskirstymą kūrinyje.

Paradoksas, kūrinių mes taip ir neperskaitome, tačiau apie jį šį tą galime pasakyti. Požiūris į kūrinių kaip į duomenų rinkinį sudaro tokias galimybes.

Ir čia būtų sunku rasti discipliną, kuri niekuo nebūtų naudinga.

**Andrius Utka.** Dirbtinis intelektas yra kilęs iš siekio imituoti natūralų žmogiškąjį intelektą (angl. *human-imitative AI*), todėl daugeliui nespecialistų asocijuojasi su vadinamąja robotika, kurios esmę sudaro žmogaus ir mašinos santykis, jų konkurencija, bandant nuspėti, kada ir kokiose veiklos srityse natūralusis intelektas bus pakeistas dirbtiniu. Ši tema – neišsiamamas mokslinės fantastikos, medijų ir populiariosios literatūros šaltinis, kurstantis įvairias aistras ir gąsdinantis neišskia žmonijos ateitim. Ką jūs manote apie robotams priskiriamą vaidmenį vadinamojoje super-, post- ar transhumanizmo epochoje? Koks, tikėtina, bus robotų santykis su žmogumi? Kokių naujų galių ar negalių įgis žmogus?

**Darius Amilevičius.** Charlie'io Chaplino šedevre „Modernūs laikai“ pagrindinis herojus dirba surinkimo linijoje, kur jis labai kenčia dėl streso, patiriamo atliekant rutininius darbus. Galiausiai jį ištinka nervinis priepuolis ir jis pabėga. Chaplino pagrindinis herojus paskatina permąstyti žmogaus ir mašinos santykį. Kiekviena ekonominė sistema sukuria poreikį darbų, kurie sunkiai dera su žmogaus prigimtimi. Žmogus tuos darbus dirba laikinai, kol bus sukurtos tinkamos mašinos. Kitas klausimas: ar mūsų švietimo sistema pasirengusi tinkamai ir savalaikiai perkvalifikuoti tuos žmones, kad jie gebėtų ne vien dirbti rutininius darbus „už minimumą“?

Žmogaus galimybių praplėtimas jau yra mokslinis faktas, o ne vien literatūrinė fikcija. Žymusis Ray'us Kurzweilis teigia, kad „<...> mūsų technologijos, mūsų mašinos yra mūsų žmogiškumo dalis. Mes sukūrėme juos norėdami išplėsti save

<...>“. Bet negalima pamiršti ir futuristo Richardo Buckminsterio Fullerio teigimo, kad „žmonija visas geras technologijas sugeba panaudoti netinkamais tikslais“. Pirmasis pripažintas kiborgas (junginys angl. *cybernetic* ir *organism*) ir aktyvistas už transhumanistų teises Neilas Harbissonas prie kaukolės turi primontuotą anteną, kuri jam leidžia skaityti įvairiomis bangomis transliuojamus signalus. MIT mokslininkas Hugh Herras naudojasi dirbtinėmis kojomis ir taip parodo, jog žmogus gali puikiai gyventi su bioninėmis galūnėmis. Stanfordo universiteto neuromokslininkas Davidas Eaglemanas tyrinėjo, ar įmanoma žmonėms sukurti naujas jusles. Jo eksperimentų rezultatai parodė, kad žmogaus smegenys gali išmokti apdoroti signalus, gaunamus ne tik įgimtų natūralių juslių pagalba. Dabar svarbu kuo greičiau pasiekti bendrą sutarimą dėl biosocialinio slenksčio – kiek ir kokių implantų turi turėti žmogus, kad būtų pripažintas kiborgu. Ir dar dėl technosocialinio slenksčio – kada kiborgas tampa mašina arba kada mašina tampa kiborgu. Bet yra ir kita medalio pusė. Interneto žiniasklaidoje gausėja nuomonių, kartais turinčių sąmokslų teorijų požymių, kad nanotechnologijų taikymas žmogaus organizme gali suteikti galimybę iš šalies veikti ir kontroliuoti žmogų. Bet šio aspekto nenorėčiau čia nagrinėti.

Kokių naujų galių įgis žmogus? Sunku arba neįmanoma atsakyti, kol nežinome, kokios natūralių žmogaus galių imtys. Šis klausimas gana dažnas, o jo formuluotė labai abstrakti. Tai perša išvadą, kad klausimo kontekste glūdi baimė prieš esminę žmogaus negalią – mirtingumą. Bet jei šią negalią nugalėsime ir visi žmonės taps nemirtingi, sustabdysime evoliuciją ir rašysime knygą pavadinimu „Rojaus ir pragaro tuštėjimo metas“. Jei nemirtingumas taps prieinamas tik siaurai antžmogių kastai, scenarijus irgi mažai patrauklus.

**Saulius Keturakis.** Nežinau... Atsakymo į panašius klausimus dabar linkstama ieškoti mokslinėje fantastikoje. Toks amerikietis Chrisas Noesselis yra sukūręs blogą (žr.: <<https://scifiinterfaces.com>>), kuriame yra paskelbti kartu su Arizonos universiteto Mokslo ir vaizduotės centru (angl. *Center for Science and the Imagination*) atlikto tyrimo duomenys. Tyrime buvo pabandyta apibendrinti šimtuose mokslinės fantastikos kūrinių išdėstytais vizijomis apie dirbtinio intelekto pavidalus ateityje ir jų pasekmes žmogui.

Tyrimo rezultatai štai tokie. Beveik visuose kūriniuose sakoma, jog dirbtinio intelekto buvimas ateityje sukels kokių nors nemalonumų žmogui. Tik labai retais atvejais dirbtinis intelektas veiks taip, kaip „reikia“, t. y. kaip draugiškas žmogaus pagalbininkas. Du trečdaliai kūrinių pasakoja apie ateities dirbtinį intelektą kaip apie pilnavertę asmenybę. Štai amerikiečių rašytojo Mike'o Resnicko apsakyme „Articles of Faith“ (2008) ateities bažnyčios durininkas robotas sprendžia klausimą, ar pradėti tikėti Dievu, į kurio namus jis daugelį metų įleisdavo ir išleisdavo

žmones. Po daugelio sunkių svarstymų robotas pasiryžta tapti tikinčiuoju. Beveik pusė kūrinį svarsto problema, jog ateityje dirbtinis intelektas būtinai ketins užvaldyti ir pavergti žmogų.

Ar taip bus išties – nežinia. Tačiau į tokius vaizduotės produktus kaip mokslinė fantastika Vakarų versle dabar imta žiūrėti labai rimtai. Kaip pavyzdys galėtų būti amerikiečių verslo modeliavimo kompanija „SciFutures“ (plačiau žr.: <<https://www.scifutures.com>>). Ji samdo mokslinės fantastikos rašytojus, kurie bando atsakyti užsakovams apie jų verslų ateitį. Klientai – „Visa“, „Intel“, net JAV karinės jūrų pajėgos.

Tad gal fantastikos kūrėjai teisūs ir apie dirbtinio intelekto bei žmogaus santykį ateityje.

Neseniai vienas mano studentas atliko keistą eksperimentą – pastatė du kompiuterius ir kiekviename paleido skirtingus pokalbių robotus BOTUS. Pradėjęs pokalbį studentas vėliau leido BOTAMS netrukdomiems reaguoti vienam į kito pasakymus. Ir po maždaug dešimties minučių vienas BOTAS ėmė kaltinti kitą, jog šis yra ne žmogus, o robotas. Kaltinimai vienas kitam virto menkai artikuliuota isterija ir eksperimentą teko nutraukti.

Tad bent jau kol kas būti žmogumi yra garbė net ir dirbtiniam intelektui.

**Andrius Utkas.** Kitos dvi su dirbtiniu intelektu besisijiančios, tačiau nesutampančios sritys – *intelektų augmentacija* (angl. *intelligence augmentation*) ir *išmanioji infrastruktūra* (angl. *intelligent infrastructure*) – dar menkai reflektuojamos viešajame diskurse. Abi jos turi savą problematiką ir žymi kitokį mašinos ir žmogaus santykį. Abiejų jų centre yra žmogus (angl. *human-centric*). Pirmoji skirta pagelbėti žmogui, išplečiant įvairias protines jo galias – atminties, komunikacijos, kūrybingumo (paieškos ir vertimo sistemos, garso ir vaizdo generavimas kompiuteriu), antroji daro žmogaus aplinką jam palankesnę ir saugesnę (daiktų internetas, išmani medicininė aplinka, savivaldžiai automobiliai). Kai kas mato šias disciplinas drauge su dirbtiniu intelektu formuojančias naują, kur kas platesnę humanocentrinę tyrimų sritį. Kaip galėtume dalyvauti šiame tarpdisciplininio mokslo ir technologijų raidos etape?

**Darius Amilevičius.** Išmanioji infrastruktūra nėra menkai reflektuojama viešame diskurse. 1992 metais ji buvo vystoma Nealo Stephensonso romane „Sniego avarija“ (angl. *Snow Crash*), kuris labai populiarus Silicio slėnyje ir tapo naujos „Facebooko“ ideologijos pagrindu. Romano išspausdinimas sutapo su antra *dirbtinio intelekto žiema*, kai visuomenės ir investuotojų susidomėjimas *dirbtinio intelekto* technologijomis ir susijusiais klausimais buvo sumenkęs.

Dėl intelekto augmentacijos (angl. *intelligence augmentation*) mano požiūris santūrus. Bent kol kas. Kam į žmogaus smegenis įkrauti mašininio vertimo sistema, jei tai puikia daro mašina? Biblija skelbia, kad kalbų įvairovė buvo Dievo bausmė žmonijai. Tai gal paprasčiau grįžti prie bendros kalbos ir pagaliau sulygti



vieną vertybių sistemą? Jei sunkiai susidorojame su informacijos paieška, kurią puikiai atlieka mašina, tai kam tai sukrauti į žmogaus smegenis? O gal paprasčiau sumažinti informacinių šiukšlių srautą, kuriuo užverčiame internetą? Jau aštuoneri metai atliekamas kasmetinis tyrimas *Data never sleeps* atskleidžia beprotiškas apimtis ir nuolatinio augimo dinamiką. Jei motulė gamta mus apdovanojo šešiomis analoginių signalų jūslėmis (sensoriais), apribodama nuo papildomos mus supančios informacijos apdorojimo, nors kaip jau matėme, žmogaus smegenys geba apdoroti ir kitokią informaciją, tai, matyt, tam buvo priežastis. Kodėl? Į šį klausimą galima atsakyti, tik atlikus išsamius tyrimus.

Bet turime įdomią tikėjimų kaitos dinamiką. Pradžioje tikėjome, kad Dievas sukūrė žmogų pagal savo paveikslą ir panašumą, todėl žmogus yra tobulas. Paskui pradėjome tikėti, kad žmogus yra tobuliausias evoliucijos kūrinys. Dabar pradėdame įtikinėti, kad žmogus yra evoliucijos abejotinos kokybės išdaiga, todėl jį reikia tobulinti, jei to nesugebėjo padaryti motulė gamta ir evoliucija. Visais atvejais turime greta stovinčią, kuriančią, bet žmogui nepavaldžią galybę (Dievas, gamta, evoliucija). Pagal klausimo esmę pradėdame kalbėti apie situaciją, kai nekokybiška evoliucijos išdaiga imsis savęs tobulinimo. Kuo ir kaip? Koks kokybiško standarto etalonas, pagal kurį būtų galima matuoti sėkmės kriterijus? Bent man asmeniškai mažai priimtina perspektyva, kad evoliucijos netobulos išdaigos sukurta *juodoji dėžė* anksčiau arba vėliau autonomiškai nuspręs, kaip tobulinti žmogų. Nauja greta stovinti, kurianti, bet žmogui nepavaldi galybė, kurią bus galima apkaltinti dėl visų nesėkmių, naujas priešas, su kuriuo reikia kovoti ir reikia naujų lyderių, kurie tą kovą organizuos, naujas mesijas, kuris be žmogaus pastangų išspręs visas problemas. Visų pirma turime gauti atsakymus į daugelį filosofinių klausimų. Deja, kaip žinia, filosofiniai svarstymai kartais užsitęsia iki nepriimtinos amžinybės...

**Saulius Keturakis.** Pasakysiu banalybe virtusią mintį – technika niekada nėra neutrali, ji visuomet paveikia žmogų. Manau, jog bet kuriuo atveju net ir centre esantis žmogus pasikeis. Bandymų atsakyti, kaip pasikeis – begalė. Oksfordo universitete dirbantis informacijos filosofas Luciano'as Floridi'is savo garsiajame „The Onlife Manifesto“ sako, jog žmogus nustos skirti gamtą nuo mašinos. Be to, informacijos perteklius žmogų kaip žinias kaupiančią būtybę pakeis žmogumi, kurio vienas iš svarbesniųjų rūpesčių bus atranka, ką užmiršti. Dar būtų galima prisiminti prancūzų filosofo Bernard'o Stieglerio mintį, jog technologijos turi bendruomenę formuojančių galių, tad ilgainiui visuomenė bus formuojama ne pagal etninius-politinius požymius, bet pagal santykį su viena ar kita technologija. Kas žino, gal ateityje būsime ne lietuviai ar ispanai, bet vieni mūsų priklausys „Instagram“ tautai, o kiti „Facebook“.

Tad gal ateityje žmogus ir išliks kaip atskaitos taškas, kaip „centras“, tačiau koks tai bus žmogus – šiandien visiškai neaišku. Nes viena didžiųjų šiandienos paslapčių – vadinamieji didieji postmodernizmo pasakojimai, socialinių medijų ir paieškos sistemų algoritmai, kiekvieną akimirką lemiantys, ką mes skaitome, žiūrime, kuo rūpinamės, o apie ką taip ir niekad nesužinome. Manau, labai prasminga tokių projektų kaip, pavyzdžiui, „Tracking Exposed“ (žr.: <<https://tracking.exposed>>) veikla. Juose, pasitelkus šimtus tūkstančių bendraminčių, bandoma suprasti, kaip mus veikia įvairūs skaitmeninės tikrovės algoritmai, užsiimama savotiška algoritmų etnografija ar archeologija. Didinti savo sąmoningumą, kritiškumą – manau, tai geriausias būdas būti žmogaus ir technologijų santykio istorijos dalyviu.