



ARENGUSEIRE
KESKUS

PIKKSILM
aprill 2022

Millisteks suurteks ohtudeks peale koroonakriisi peaksiime veel valmis olema?

Risto Uuk

Artikkel ilmus Arenguseire Keskuse trendiraportis „Pikksilm“.



Eksistentsiaalsed ohud on sellised ohud, mis võivad hävitada inimkonna kaug-tuleviku potentsiaali. Siia alla kuuluvad nii sellised sündmused, mis hävitaksid kogu inimkonna, kui ka sündmused, mis takistaksid inimkonnal oma potentsiaali realiseerimist.¹ Ohtudeks võivad olla asteroidid ja komeedid, supervulkaanid, täheplahvatused, tuumasõjad, kliimamuutus, muud keskkonnakahjud, looduslikud või tehislikud pandeemiad, väärtustega ühildamata tehisintellekt (1) ja muud võimalikud inimtekkelised ohud.¹ Nime-tatud ohtudest on kõige paremini arusaadavad looduslikud ohud, mida on ka kõige rohkem uuritud. Tõenäoliselt kõige spekulatiivsemad ja vähem tähelepanu saanud on väärtustega ühildamata tehisintellekti ohud.

Käesolevas essees keskendun kliimamuutusele, tehislিকে pandeemiatele ja väärtustega ühildamata tehisintellektile sellepärast, et need võivad olla kõige ohtlikumad ja tõenäolisemad eksistentsiaalsed ohud sellel sajandil ning kui Euroopa Liit (EL) tahab, saab nende puhul midagi ette võtta. **Eksistentsiaalsetele ohtudele keskendumine on oluline, sest need põhjustaksid kahju väga paljudele inimestele ning muudaksid tuleviku võimatuks või planeedi elamiskõlbmatuks tulevastele põlvkondadele.** Isegi kui selliste ohustsenariumide realiseerumisel inimkond välja ei sure, võivad need põhjustada tsivilisatsiooni kokkuvarisemise ja düstoopilise tuleviku.

Eksistentsiaalsete ohtude olulisus

Tulevik on potentsiaalselt väga pikk ja eksistentsiaalseid ohtusid vältides võiks tulevikus elada väga palju inimesi. Kui eeldada, et Maa jääb elamiskõlblikuks veel vähemalt üheks

miljardiks aastaks ja vähemalt üks miljard inimest saaks korraga jätkusuutlikult sellel elada, siis võiks vähemalt 10^{16} normaalse elueaga inimest tulevikus veel elada.³ See tähendab, et **eksistentsiaalse ohu realiseerumisel võiks elu elamata jääda kümnel tuhandel triljonil inimesel, mida on umbes sada tuhat korda rohkem, kui on inimesi kunagi üldse elanud.**⁴ See hinnang võib olla konservatiivne, sest see ei võta arvesse võimalikke tehnoloogilisi arenguid ja tsivilisatsioonide rajamist teistele planeetidele. Seega, inimkonna tulevik on potentsiaalselt palju pikem kui olevik ja minevik ning kaalul on palju rohkem inimelusid.

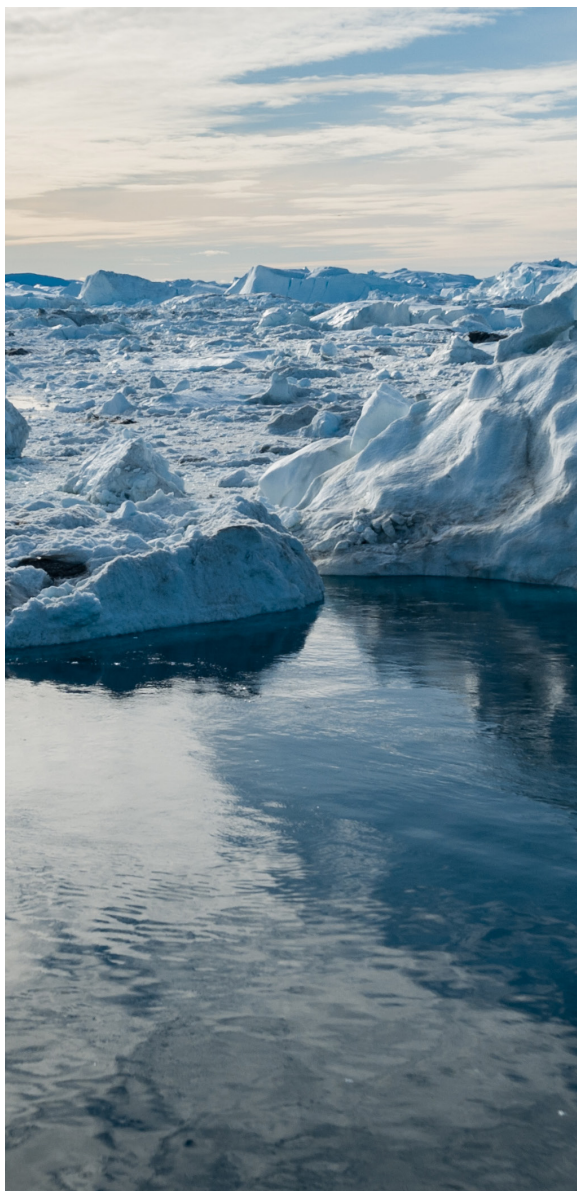
Inimkonna areng võib aga liiga vara katkeda, kuna kliimamuutusest, pandeemiast, uutest tehnoloogiatest ja teistest ohtudest tingitud olukorrad võivad senised saavutused ära rikkuda. Nimetatud ohtude esinemise tõenäosust sellel sajandil või umbes järgmise saja aasta jooksul peetakse mitme eksperdi (2) hinnangul väga kõrgeks. Ord (2020) hindab tõenäosuseks 17%¹, Cotton-Barratt ja teised (2016) hindavad tõenäosuseks 10%⁵ ning Sandberg ja Bostrom (2008) hindavad tõenäosuseks 19%.⁶ Kui võtta eksistentsiaalse ohu tõenäosuseks nende hinnangute põhjal umbes 0,1% aastas, siis inimesel on eksistentsiaalse ohu tõttu surma saamise tõenäosus viis korda suurem, kui saada surma autoõnnetuses.⁵ Neid tõenäosusi tuleb aga ettevaatlikult kasutada, sest need on ekspertide subjektiivsed hinnangud.

Hoolimata sellest, et eksistentsiaalsete ohtude puhul on kaalul väga palju ja nende esinemise tõenäosus paistab olevat väga suur, kulutatakse nende riskide vähendamiseks globaalselt väga vähe ressursse. Ord (2020) hindab otseselt eksistentsiaalsetele ohtudele tehtud

(1) Väärtustega ühildamata tehisintellekt tähendab seda, et tehisintellekt võib tuua ootamatuid ja soovimatuid tulemusi, kuna tehisintellekti arendajad (ja ühiskond üldiselt) ei suuda piisavalt hästi eesmärgi sõnastada ja tehisintellektile edasi anda.

(2) **T. Ord** – Oxfordi ülikooli vanemteadur, filosoofia doktorikraad Oxfordi ülikoolist ([toby-ord-cv.pdf](#) ([amirrorclear.net](#))); **O. Cotton-Barratt** – Oxfordi ülikooli Future of Humanity Institute'i teadusprogrammi direktor, matemaatika doktorikraad Oxfordi ülikoolist (Owen Cotton-Barratt – Future of Humanity Institute ([ox.ac.uk](#))); **A. Sandberg** – Oxfordi ülikooli Future of Humanity Institute'i vanemteadur, neuroteaduse doktorikraad Stockholmi ülikoolist (Anders Sandberg – [Wikipedia](#)); **N. Bostrom** – Oxfordi ülikooli filosoofiaprofessor, filosoofia doktorikraad London School of Economicsist ([cv.pdf](#) ([nickbostrom.com](#))).

globaalseid aastaseid kulutusi 100 miljoni dollari suurusjärgus.¹ Kõrgelt arendatud tehisintellektist tulenevatele ohtudele kulutatakse vaid mõnikümmend miljonit dollarit⁷ aastas, kuid tehisintellekti võimekamaks muutmisele kulutatakse miljardeid.⁸ Tehislikest pandeemiatest tulenevate ohtude vähendamisele kulutatakse suurusjärgus mõni miljon dollarit⁹, aga bioloogilistele riskidele üldiselt kulutatakse miljardeid dollareid, millest osa võib ka olla eksistentsiaalsete ohtude vähendamiseks kasulik.¹⁰ Kliimamuutusele kulutatakse igal aastal globaalselt sadu miljardeid dollareid¹¹, aga ainult väike osa sellest kulub äärmuslike ohtude uurimisele



ja ennetustööle.¹ Näiteks kulutatakse geoinseneeringule (3) – tehnoloogia, millega loodetakse vähendada globaalset soojenemist – mõnikümmend miljonit dollarit.¹² Seda peaks aga rohkem uurima, muu hulgas ka sellepärast, et sellise tehnoloogia ebaõnnestunud rakendamisel võib globaalne soojenemine hoopis kiirenedada.¹³

Samuti on oluline tõdeda, et paljud inimesed nõustuvad sellega, et eksistentsiaalsete ohtude vähendamine peaks olema üks peaesmärk. Näiteks Schubert ja teised (2019) uurisid kolme uuringu jaoks kokku üle 2500 Ameerika Ühendriikide ja Suurbritannia elaniku arvamuse ning leidsid, et enamik vastanutest tunneb veendunult, et eksistentsiaalseid ohtusid tuleks ennetada, nende ennetamine on moraalne kohustus ja nende ennetamisele raha kulutamine on olulisem, kui teiste valdkondade, näiteks haridus-, tervishoiu- ja sotsiaalkindlustuse valdkonna rahastamine.² Samas aga ei pea selle uurimuse järgi enamik inimesi eksistentsiaalseid ohtusid väga palju hullemaks muudest katastroofidest, mille korral sureks küll palju inimesi, kuid mitte kõik. Inimesed keskenduvad otsestele kahjudele, mida katastroofid põhjustavad, ning tunnevad tugevalt kaasa kannatanutele, kuid jätavad tähelepanuta kaugtuleviku tagajärjed.

Kliimamuutus

Ohu realiseerumise võimalikud stsenaariumid

Esimene võimalik stsenaarium kliimamuutuse katastroofiliseks ohuks on kontrollimatu kasvuhooneefekt. **Kontrollimatu kasvuhooneefekt tähendab seda, et tekib võimendatud ahel, kus atmosfääri soojenemine põhjustab vee aurustumist atmosfääri, mis omakorda põhjustab atmosfääri soojenemist, mille tulemusena ookeanid aurustuvad ja elu ei ole enam võimalik.**¹ Seda stsenaariumi peetakse teoreetiliselt võimalikuks, kuid see

(3) Geoinseneering on kliimamuutmise tehnoloogia, mille kõige tuntumad kasutusvaldkonnad on päikesekiirguse juhtimine ja süsinikdioksiidi eemaldamine atmosfäärist (*Climate engineering* – Wikipedia).

ei ole arvatavasti võimalik ainult inimtekkeliste heitgaasidega.^{14,15} Ka selline ahel, mis põhjustab ulatuslikku soojenemist, kuid ookeane täiesti ära ei aurusta, ei paista olevat kõigest inimtekkeliste heitgaaside tagajärjel võimalik, aga teadlased ei ole selles kindlad.¹

Teine võimalik stsenaarium kliimamuutuse katastroofiliseks ohuks on selline **soojenemise ahel, mis põhjustab igikeltsa sulamise ja sügavast ookeanist metaani eraldumise**. Igikeltsas võib olla kaks korda rohkem süsinikku, kui on praegu Maa atmosfääris kokku.¹⁶ Teadlased on veendunud, et järgmiste sajandite jooksul igikelts osaliselt sulab, aga pole täpselt teada, milline mõju võib sellel globaalsele soojenemisele olla.¹⁷ Samuti on sügaval ookeani põhjas jää sees palju metaani, aga pole täpselt teada, kui palju seal süsinikku kokku on.¹ Teadlased on veel vähem kindlad, mis juhtuks siis, kui ookeani põhjast hakkaks eralduma metaani, ja mis aja jooksul muutused toimuksid.¹

Kolmas võimalik stsenaarium kliimamuutuse katastroofiliseks ohuks on see, et **inimkond jätkab heitgaaside paiskamist atmosfääri samas tempos või kiirendab seda ja lõpuks tuleb kliimamuutuse peatamiseks rakendada äärmuslikke meetodeid, mis põhjustavad omakorda muu katastroofi**. Kui inimkond põletaks peaaegu kõik fossiilkütused, võib see aastaks 2300 põhjustada globaalset soojenemist umbes 9–13 kraadi võrra.¹⁸ See võib muuta suurema osa planeedist elamiskõlbmatuks ja laastada põllumajanduse täielikult.¹³

Neljas võimalik stsenaarium kliimamuutuse katastroofiliseks ohuks on see, et kliimamuutus ise ei põhjusta inimkonna väljasuremist, aga osutub riskiteguriks teistele eksistentsiaalsetele ohtudele.¹ **Selle stsenaariumi järgi suurendab kliimamuutus võimalust, et eksistentsiaalse katastroofi põhjustavad teised eksistentsiaalsed ohud, näiteks värtustega ühildamata tehisintellekt või tehnilik pandeemia**. Kliimamuutus riskitegurina võib kaasa aidata majanduslikule stagneerumisele, rahvusvahelise korra rikkumisele ja ühiskonna kokkukukkumisele.¹⁹

Ohu realiseerumise tõenäosus

Ord (2020) hindab kliimamuutusest tuleneva eksistentsiaalse ohu tõenäosuseks järgmise saja aasta jooksul ligikaudu 0,1%.¹ Ta selgitab, et seda arvu ei maksa võtta täielikult objektiivsena, vaid see on tema subjektiivne kindlushinnang, võttes arvesse kõike, mida ta seoses kliimamuutuse ohtudega teab. Duda ja Koehler (2020) hindavad tõenäosust sarnaseks.²⁰ Nende sõnul on kliimamuutusest tingitud eksistentsiaalse ohu võimalus väike, kuid mitte olematu, viidates võimalustele, et inimesed jätkavad või isegi suurendavad fossiilkütuste kasutamist ning põhjustavad mõne äärmusliku ettearvamatu ahela, mille tagajärjel kuumenemine kiireneb ja väljub kontrolli alt. Pamlin ja Armstrong (2015) hindavad kliimamuutusest tingitud eksistentsiaalse ohu tõenäosuseks järgmise kahesaja aasta jooksul 0,01%.²¹ Nad tõdevad, et need hinnangud on katse olemasolevaid hinnanguid kokku panna, et julgustada teisi nende hinnangute kindlust parandama.

Kas EL teeb piisavalt?

Euroopa Komisjon võttis 2021. aastal vastu uue kliimamuutusega kohanemise strateegia, et valmistuda kliimamuutuse tagajärgedeks, mille hulgas on kuumalained, põuad, metsade hävimine ja merepinna tõus.²² Strateegias tuuakse välja, et lisaks eespool mainitud mingil määral juba esinevate tagajärgedega kohanemisele tuleb ennetada võimalikke suuremaid tagajärgi, nagu igikeltsa sulamine, merejää kadu ja ulatuslik metsade hävimine. EL-i peamine algatus nimetatud tagajärgedega kohanemiseks on Euroopa roheline kokkulepe ja selle läbi lubadus saavutada kliimanetraalsus aastaks 2050.

Euroopa roheline kokkuleppe järgi on EL-i peamised poliitikasuunad ja meetmed järgmised: põhiliselt taastuvatel energiaallikatel põhineva energiaspektori loomine, tööstussektori kaasamine puhtaks ringmajanduseks, energiatõhus ehitamine, säästvale transpordile üleminek, keskkonnahoidliku toidusüsteemi kujundamine, ökosüsteemide säilitamine ja taastamine, erinevate kliimaeesmärkide

saavutamiseks lisainvesteeringute tegemine ning uute tehnoloogiate edendamine ja rakendamise.²³ 2021. aasta juunis plaanib Euroopa Komisjon üle vaadata ja vajaduse korral uuendada heitgaaside vähendamise meetmeid.²⁴ Euroopa Komisjon on teinud ka ettepaneku võtta vastu esimene Euroopa kliimaseadus, et tagada kliimaneutraalsuse saavutamine 2050. aastaks.²⁵

Oluline on rõhutada, et kliimamuutus on üleilmne probleem ja ainult EL-i pingutused süsinikku vähendada ei ole piisavad. EL peaks tegema koostööd Ameerika Ühendriikide ja teiste partneritega ning kehtestama ülemaailmseid standardeid.²⁶ Selleks, et EL-i tegevused kliimamuutuse ohtude vähendamiseks oleks võimalikult efektiivsed, tasub uurida analüüse kõige kuluefektiivsemate meetmete kohta. Organisatsioon Let's Fund on analüüsinud, millised kliimamuutuse meetmed võiksid olla kõige efektiivsemad vastavalt kolmele mõõdikule: ulatus ehk kui suurt osa kliimamuutuse mõjust võib meede lahendada, kõrvalejätetus ehk kui vähe ressursse meetmele praegu suunatakse ja lahendatavus ehk kuivõrd realistlik meetme rakendamine poliitiliselt on.²⁷ Nende hinnangul on **kõige efektiivsem suunata rohkem tähelepanu riigi kulutuste suurendamisele puhta energia teadus- ja arendustegevuseks ning mingil määral suurendada ka kliimamuutuste alusuuringute rahastust.** Eriti vähe on teada geoinseneeringu mõjust ja rakendusvõimalustest ning kliimamuutuste äärmuslikest tagajärgedest.²⁰ Mõttekoda Copenhagen Consensus Center on hinnanud kõige efektiivsemaks teadustööd geoinseneeringu kohta ja uuringuid puhta energia tehnoloogiate arendamiseks.²⁸ Organisatsioon Founders Pledge on leidnud, et **suur osa tähelepanust rohelse tehnoloogia arendamisel läheb päikese- ja tuuleenergia, energia efektiivsuse ning elektriautode peale, mistõttu oleks vaja rohkem ressursse suunata süsiniku sidumise ja ladustamise, tuumaenergia ning raskeveokite kütuste uurimisele.**²⁹

Tehislikud pandeemiad

Ohu realiseerumise võimalikud stsenaariumid

Esimene võimalik tehisliku pandeemia katastroofilise ohu stsenaarium on **patogeenide lekkimine teaduslaboritest.** Inimesed võivad tulevikus luua selliseid patogeene, mis on väga surmavad, väga kõrge nakatumisriskiga ja väga pika peiteperioodiga.³⁰ Selliste patogeenide puhul nakatunu suure tõenäosusega sureks, paljud inimesed suure tõenäosusega kannaksid patogeeni teistele edasi ja esimesed sümptomid ilmneksid alles siis, kui nakatunu on patogeeni paljudele teistele edasi kandnud. Näiteks polioviirust ja hobuserõugeid suudavad teadlased juba sünteesida ning on ka eksperimente, mis näitavad grippide ohtlikkuse suurendamise võimalikkust.³¹ Laboritest on patogeene palju kordi välja lekinud hoolimata sellest, et need on tavaliselt rangelt kaitstud. Näiteks 1971. aastal katsetas Nõukogude Liidu biorelvade labor rõugete kasutamist relvana ning lekkimise tulemusena nakatus kümme inimest ja suri kolm inimest enne, kui levik suudeti peatada.¹

Teine võimalik tehisliku pandeemia katastroofilise ohu stsenaarium on **bioloogiliste relvade kasutamine.** Bioloogiliste relvade kasutamine sõjas või terroristide poolt ei ole võrreldes looduslike pandeemiatega ajalooliselt eriti palju kahju tekitanud. Ühe kõige surmavama pandeemia, musta surma katkupandeemia tagajärjel suri vähemalt 75 miljonit inimest 450 miljonist inimesest kogu maailmas.³⁰ Biorelvade kasutamise tagajärjel aga sai Teises maailmasõjas ühes kõige surmavas juhtumis surma umbes 200 000 hiinlast.³² See aga ei tähenda, et tulevikus ei võiks biorelvade tagajärjel olla ohustatud väga suur osa rahvastikust sellepärast, et biotehnoloogia tõttu võivad sõjas või terrorismis kasutatavad patogeenid ohtlikumaks muutuda.

Kolmas võimalik tehisliku pandeemia katastroofilise ohu stsenaarium on **biotehnoloogia demokratiseerumine ja teabeohud.** Biotehnoloogia areneb kiirelt ja selle tõttu muutub biotehnoloogia kasutamine odavamaks ja kätte-

saadavamaks. Näiteks esimese inimese genoomi järjestuse loomise maksumus oli hinnanguliselt umbes üks miljard dollarit, praegu aga maksab terve genoomi järjestamine ligikaudu 600 eurot.¹ Tehisliku pandeemia oht suureneb ka sellepärast, et biotehnoloogia võib sattuda rohkemate pahatahtlike inimeste kätte või lihtsalt suureneb õnnetuste arv. Samuti on võimalus, et informatsioon selle kohta, kuidas luua väga surmavat, nakkavat ja pika inkubatsiooniperioodiga patogeeni, levib üksikute inimeste käest rohkemateni ning jõuab lõpuks valede inimesteni.

Ohu realiseerumise tõenäosus

Millett ja Snyder-Beattie (2017) ühe mudeli järgi on eksistentsiaalse ohu tõenäosus biotehnoloogia tagajärjel sellel sajandil ligikaudu 0,00013% kuni 0,0066%.³³ Autorid tõdevad, et nende mudeli tõenäosushinnang on väga konservatiivne, kuid võtab arvesse seda, et teadlased ise tahtlikult põhjustavad kahju, tsivilisatsioon kukub pandeemia tõttu kokku või tekib muid kõige hullemaid

stsenaariume. Nende teise mudeli järgi on eksistentsiaalse ohu tõenäosus biorelvade kasutamise tagajärjel sellel sajandil ligikaudu 0,005%. Autorid rõhutavad, et nende hinnangud on väga ebakindlad ja tuginevad mingil määral spekulatsioonile ning neid madalaid tõenäosusi ei peaks enesekindlalt kasutama sellepärast, et need ei võta arvesse potentsiaalselt kõige suuremat ohuallikat tulevikus, milleks on tehnoloogiline areng ja biorelvade lai kättesaadavus.

Sandberg ja Bostrom (2008) küsisid Oxfordi globaalse katastroofilise riski konverentsil osalejate käest, kui suur on nende arvates tõenäosus, et selle sajandi jooksul põhjustab tehislik pandeemia inimkonna väljasuremise.⁶ Vastanute mediaanhinnang oli 2%. Seda hinnangut peaks kasutama eriti ettevaatlikult, sest tegemist on subjektiivse hinnanguga ja raportis ei ole ülevaadet küsimustele vastanute kohta. Samas aga tasub mainida, et küsitletute mediaanhinnang sellele, kui suur on tõenäosus, et selle sajandi jooksul sureb vähemalt üks miljon inimest loodusliku pandeemia tagajärjel,



oli 60%. Praeguseks on COVID-19 tõttu juba üle ühe miljoni inimese maailmas surnud, seega võib seda kindlushinnangut pidada veenvaks.

Kas EL teeb piisavalt?

EL-is võeti 2013. aastal vastu EL-i terviseohutuse raamistik, milles toodi välja, kuidas parandada valmisolekut ja tugevdada suutlikkust koordineeritud reageerimiseks tõsistele piiriülestele terviseohtudele.³⁴

Raamistikus on nimetatud järgmised suunad: riikliku valmisoleku kavandamise koordineerimine, heade tavade ja teabe jagamise tugevdamine, kiirhoiatussüsteemi haldamine tõsistest terviseohtudest teatamiseks, riskihindamise ja -juhtimise parandamine, meditsiiniliste vastumeetmete ühishangete väljatöötamise ja kasutuselevõtu kehtestamine, parem reageerimise koordineerimine EL-i tasandil ning rahvusvahelise koostöö ja üleilmsete meetmete edendamine.³⁵ EL-i terviseohutuse raamistikus küll korra mainitakse bioterrorismi, kuid see ei ole otseselt suunatud sellistele tehislake pandeemiate stsenaariumidele nagu patogeene lekkimine teaduslaboritest, biorelvade kasutamine, biotehnoloogia demokratiseerumine ja teabeohud.

2015. aastal edastasid teaduskomiteed Euroopa Komisjonile teaduslikke nõuandeid selle kohta, mida poliitikakujundamisel arvesse võtta seoses sünteetilise bioloogia ohtudega ja milline peaks olema riskianalüüsi metodoloogia.³⁶ Raportis mainitakse pealiskaudselt patogeene lekkimist teaduslaboritest ja biotehnoloogia demokratiseerumist. Autorid tõdevad, et biotehnoloogia laiem kasutus võib ohtusid suurendada, kuid peavad seda võimalust vähetõenäoliseks, kui kasutajad on hästi informeeritud ja ettevaatlikud. Nad rõhutavad ka teadlikkuse suurendamise, koolitamise ja headest laboritavadest kinnihoidmise tähtsust. Kuna sünteetiline bioloogia areneb väga kiiresti ja sisaldab loomult teaduslikku teadmatust, tuleks selliseid riskianalüüse regulaarselt üle vaadata, täiendada ja korrata.³⁶ Selleks **oleks vaja kokku panna eksperdirühm, kes analüüsiks põhjalikult võimalikke ohustsenaariume ja meetmeid ohtude vähendamiseks. Näiteks**

Suurbritannias moodustati hiljuti selline eksperdirühm, et hinnata valitsuse lähene misviisi biotehnoloogia ohtudele.³⁷

Lisaks tuleb mainida COVID-19 pandeemiat kui kõige viimast juhtumiuuringut, mille näitel aru saada EL-i nõrkustest pandeemiateks ettevalmistumisel. 2021. aasta 24. märtsi seisuga on EL-is olnud kokku üle 590 000 koroonaviirusega seotud surma võrreldes Ameerika Ühendriikide üle 540 000 surma, Lõuna-Ameerika üle 520 000 surma ja Aasia üle 410 000 surmaga.³⁸ Miljoni elaniku kohta on Ameerika Ühendriigid surmade poolest esimesed üle 1600 surmaga, EL teine üle 1300 surmaga, Lõuna-Ameerika kolmas üle 1200 surmaga ja Aasia neljas kõigest üle 80 surmaga.³⁹ EL pole olnud koroonaviiruse pandeemiaks piisavalt valmis – puudu on olnud



meditsiinilistest vahenditest, tervisetöötajad on olnud ülekoormatud, haiglate võimekus on viidud üle piiri ning riskirühmad on kannatanud liiga palju.⁴⁰ Samuti on EL teinud vigu koroonaviiruse vastu vaktsineerimisel *pro* vaktsineerimise tegevuses.⁴¹ **Ainuüksi selle info põhjal võib öelda, et EL peaks paremini valmistuma tehisliseks pandeemiatega, millel võivad olla veel hullemad tagajärjed.**

Väärtustega ühildamata tehisintellekt

Ohu realiseerumise võimalikud stsenaariumid

Esimene võimalik tehisintellekti katastroofilise ohu stsenaarium on **suutmatus panna tehisintellekti inimeste eesmärkidele vastavalt tegutsema**. Selle stsenaariumi puhul õpetatakse tehisintellekti saavutama inimestele vajalikku eesmärki, aga see ei ole edukas ja tehisintellekt teeb vigu, millel on katastroofilised tagajärjed.³⁰ Inimväärtusi on keeruline masinatele tehniliselt edasi anda sellepärast, et nad ei pruugi aru saada inimeste eesmärkidest ning nad ei pruugi suuta neid eesmarke õigesti ellu viia ja arengu käigus säilitada.³⁰ Kui tegemist on robottoolmejaga, siis selle vead on üpris leebete tagajärgedega, näiteks ei saa tuba ühtlaselt puhtaks. Kui aga tehisintellekti süsteem on võimekam, näiteks isejuhtivate autode võrgustik või tark linn, võib vigade tagajärg olla palju hullem.

Teine võimalik tehisintellekti katastroofilise ohu stsenaarium on **inimväärtustest puudulik arusaamine ja seetõttu masinate valed eesmärgid kasutamine**. Selle stsenaariumi puhul masinaid küll suudetakse treenida tegema täpselt seda, mida inimesed tahavad, kuid nende kasutamine võib viia katastroofini, kui inimeste eesmärgid on valed ehk ei arvesta korrektselt kaasnevaid sotsiaalseid ja ühiskondlikke mõjusid.⁴² Kui riikide valitsejad jõuavad üksmeelele ühes õiges moraaliteoorias, mille järgi kogu maailma valitseda, võib ka see viia düstopiani, kui nende

moraaliteooria osutub tegelikult ikkagi ekslikuks (paljude inimeste olukorda halvendavaks) ja nende kasutuses on ülivõimekas tehisintellekt.

Kolmas võimalik tehisintellekti katastroofilise ohu stsenaarium on **inimeste pahatahtlik tehisintellekti kasutus**. Selle stsenaariumi puhul kasutatakse tehisintellekti otseselt pahatahtlikul eesmärgil, näiteks autonoomsete relvasüsteemide kasutamine massitapmisteks.³⁰ Samuti võib autonoomsete relvasüsteemide loomine põhjustada rahvusvahelist võidurelvastumist, mis suurendab suurte võimude vahelise sõja puhkemise võimalust.³⁰ Lisaks sõjalistele sekkumistele võidakse tehisintellekti kasutada ka demokraatia õõnestamiseks ning seda juba tehaksegi sotsiaalmeedias levivate libauudiste ja muude tegevuste abil.⁴³

Neljas võimalik stsenaarium on **tehisintellektile järkjärguline või järsk võimu äraandmine**. Selle stsenaariumi puhul on võimalus, et tehisintellektile antakse järk-järgult üle majanduslikud ja muud ülesanded, kuni lõpuks kontrollib tehisintellekt ühiskonda peaaegu täielikult.¹ Samuti on võimalus, et toimub teaduslik murrang, mille tulemusel muutub tehisintellekt, mida ühiskonnas mingite ülesannete täitmiseks kasutatakse, plahvatuslikult järjest võimekamaks ning inimestel kaob selle arengu ja käitumise üle kontroll.⁴⁴ Kui sellise ülivõimeka tehisintellekti käitumine ei ole inimeste väärtustega kooskõlas, viib see katastroofiliste tagajärgedeni.

Ohu realiseerumise tõenäosus

Müller ja Bostrom (2016) küsitlesid tehisintellekti eksperte selle kohta, millal saavutatakse inimtasemel tehisintellekt ning kas see saavutus toob ühiskonnale kasu või kahju.⁴⁵ Nende definitsiooni järgi on inimtasemel tehisintellekt selline tehisintellekt, mis suudab enamiku inimeste töid teha vähemalt sama hästi kui keskmine inimene. Nad küsitlesid nelja gruppi: tehisintellekti filosoofia konverentsil osalejad, üldintellekti (4) mõju ja ohtude konverentsil

(4) Üldintellekti all mõistetakse inimtasemel tehisintellekti, millel on üldine võimekus õppida lahendama paljusid erinevaid ülesandeid.

osalejad, Kreeka tehisintellekti assotsiatsiooni liikmed ning 100 kõige tsiteeritumat tehisintellekti autorit. Kokku vastas küsitlusele 170 inimest (31%). Kõikide gruppide peale kokku peeti **inimtasemel tehisintellekti saavutamist võimalikuks 50% tõenäosusega aastaks 2040 mediaani järgi ja aastaks 2081 keskmise järgi**. Küsimusele, et kui inimtasemel tehisintellekt saavutatakse, siis kas sellel oleks inimkonnale positiivne või negatiivne pikaajaline mõju, vastasid küsitletud, et **18% tõenäosusega (keskmise järgi) oleks sellel saavutusel väga halb mõju ehk see põhjustaks eksistentsiaalse katastroofi**. See arvamus erines grupiti – 1. grupp hindas tõenäosuseks 18%, 2. grupp 24%, 3. grupp 6% ja 4. grupp 8%. Kui usaldame kõige enam viimase grupi hinnangut, mille moodustasid kõige tsiteeritumad autorid, siis oleks eksistentsiaalse ohu võimalus inimtasemel tehisintellekti loomise korral ikkagi väga kõrge.

Grace ja teised (2018) küsitlesid tehisintellekti eksperte selle kohta, millal saavutatakse kõrgetasemeline masinintellekt ja kas see saavutus oleks inimkonnale kasulik või kahjulik.⁴⁶ Nende definitsiooni järgi on kõrgetasemeline masinintellekt saavutatud siis, kui masinad suudavad ilma inimeste abita sooritada kõiki ülesandeid paremini ja odavamalt kui inimtöölised. Nad küsitlesid 2015. aastal toimunud konverentsidel NIPS ja ICML (kaks enim hinnatud masinõppe konverentsi) osalenud eksperte. Kokku vastas küsitlusele 352 inimest 1634-st ehk 21%. Ekspertide keskmise hinnangu järgi saavutatakse kõrgetasemeline masinintellekt 50% tõenäosusega 45 aasta jooksul ja 10% tõenäosusega 9 aasta jooksul. Aasiast pärit teadlased arvasid, et see saavutatakse 30 aasta jooksul, Põhja-Ameerika teadlased aga, et 74 aasta jooksul. Teadlased hindasid, et kõrgetasemeline masinintellekt põhjustaks väga halva tagajärje, näiteks inimkonna väljasuremise, 5% tõenäosusega. Samuti arvasid teadlased, et ühiskond peaks pigem rohkem prioriseerima tehisintellekti ohtude minimeerimisega tegelevat teadustööd. Artikli juhtivautor aga ei arva, et ekspertide subjektiivseid hinnanguid peaks käsitletavas küsimuses tugeva tõendina kasutama.⁴⁷

Kas EL teeb piisavalt?

EL-is on tehisintellekti ohtudega tegelenud aastast 2018 kõrgetasemeline tehisintellekti eksperdirühm, mille pani kokku Euroopa Komisjon 52 liikmest ja mis tegutses kaks aastat.⁴⁸ Selle aja jooksul koostas eksperdirühm usaldusväärse tehisintellekti eetikajuhendi, milles toodi tehisintellekti süsteemide usaldusväärse saavutamiseks välja seitse peamist nõuet.⁴⁹ Nendeks nõueteks on (i) inimese toimevõime ja järelevalve ehk et inimestel oleks masinate üle kontroll, mitte vastupidi, (ii) tehniline töökindlus ja ohutus, (iii) privaatsus ja andmehaldus, (iv) läbipaistvus, (v) mitmekesisus, mittediskrimineerimine ja õiglus, (vi) ühiskondlik ja keskkonnavaline heaolu ning (vii) vastutuse võtmine. Selle eksperdirühma tööd on kõige teravamalt kritiseerinud eksperdirühma liige filosoofiaprofessor Thomas Metzinger, kes ütleb, et koostatud eetikajuhend ignoreerib pikaajalisi tehisintellektiga seotud ohtusid nagu tehisintellekti abil kodanike üle järelevalvepidamist ja sellise tehisintellekti kasutamist, mida inimesed ei suuda enam kontrollida ega võta tugevamat hoiakut surmavate autonoomsete relvasüsteemide suhtes.⁵⁰

Euroopa Komisjon tegeleb praegu tehisintellekti regulatsiooni loomisega⁵¹, mille uus ettepanek avaldati 21. aprillil 2021.⁵² Esimeses tehisintellekti regulatsiooni ettepanekus on järgmised põhilised mõtted: kehtivas õigusraamistikus tuleb teha võimalikud kohandused, mis puudutavad tehisintellekti; fookuses peab olema, kuidas viia kõige suuremate riskide võimalus miinimumini, ning kõrge riskitasemega tehisintellekti rakendustele peaksid olema määratud erinõuded standardite ja muude meetmete abil.⁵¹ Samuti on ettepanekus kirjas, et tuleb täpselt määratlada, kes vastutavad kõrge riskitasemega tehisintellekti rakenduste ohtude eest, kontrollida ja tagada, et kohustuslikud nõuded oleksid objektiivse vastavushindamise alusel täidetud, ning luua riiklike pädevate asutuste koostööraamistik, et vältida vastutuse killustamist ja suurendada liikmesriikide suutlikkust.

Ühendkuningriik on nüüdseks EL-ist välja astunud, aga pikaajalise EL-i liikmena on neil mitu tehisintellekti ohutusega seotud organisatsiooni, mida tasub siin nimetada. Cambridge'is asuv Centre for the Study of Existential Risk tegeleb muu hulgas ka tehisintellekti ohtude uurimisega. Nende **põhiline soovitus Euroopa Komisjonile on säilitada tehisintellekti regulatsiooni ettepanekus pakutud ideid ning mitte muuta neid näiteks korporatsioonide survele nõrgemaks.**⁵³ Sellega saab EL kujundada globaalseid standardeid ja võtta liidripositsiooni usaldusväärse tehisintellekti arendamises. Oxfordis asuv Future of Humanity Institute, mis tegeleb samuti muude teemade kõrval tehisintellektist tulenevate eksistentsiaalsete ohtude uurimisega, soovitas Euroopa Komisjonil võtta rohkem arvesse laiaulatuslikke ohtusid. Nende hinnangul tuleb pidada ülevaadet lisaks sellele, milline on kõrge ohuga tegevusvaldkonna võimalik kahju inimestele, ka selle üle, kuidas tehisintellekt võib ohustada avalikke huvisid, kui palju kasutajaid konkreetsel tehisintellekti rakendusel on ja kui sagedasti seda kasutatakse, kuid üldjoontes nõustuvad nad EL-i lähenemisega.⁵⁴

Kokkuvõte

Eksistentsiaalsed ohud peaksid olema EL-i prioriteet ühest küljest sellepärast, et nende realiseerumisel kannataksid väga paljud praegu elavad inimesed, kuid ka sellepärast, et nende tõttu jääks olemata kogu inimkonna tulevik koos kõikvõimalike inimkonna saavutuste ja hüvedega. Eksistentsiaalsete ohtude vähendamiseks kulutatakse nii globaalselt kui ka EL-is arvatavasti liiga vähe ressursse, kuigi paljud eksperdid ja tavainimesed nõustuvad, et see peaks olema eelisvaldkond. Peamised eksistentsiaalsed ohud, mida olen siin käsitlenud, on kliimamuutus, tehnilikud pandeemiad ja väärtustega ühildamata tehisintellekt, kuigi võimalikke ohtusid on veel teisigi. **Eksistentsiaalsed ohud paistavad olevat sellel sajandil väga suured – mõne eksperdi arvamusel isegi ligikaudu 20% –, kuigi enesekindlus nendes hinnangutes ei saa olla kõrge.**

Kliimamuutusest tulenevad katastroofilised stsenaariumid võivad olla kontrollimatu kasvuhoooneefekt, igikeltsa sulamine ja sügavast ookeanist metaani eraldumine, fossiilkütuste



kasutamise jätkamine või kiirendamine, äärmuslike meetodite rakendamine kliimamuutuse peatamiseks ning kliimamuutuse osutamine riskiteguriks teistele eksistentsiaalsetele ohtudele. Tehislikest pandeemiatest tulenevad katastroofilised stsenaariumid võivad olla patogeenide lekkimine teaduslaboritest, bioloogiliste relvade kasutamine ning biotehnoloogia demokratiseerumine ja teabeohud. Väärtustega ühildamata tehisintellektist tulenevad katastroofilised stsenaariumid võivad olla suutmatuse panna tehisintellekti inimeste eesmärkidele vastavalt tegutsema, masinate valed eesmärgil kasutamine inimväärtustest puuduliku arusaamise tõttu, inimeste pahatahtlik tehisintellekti kasutamine ja tehisintellektile järkjärguline või järsk võimu äraandmine.

Kui EL tahab, saab ta midagi nimetatud eksistentsiaalsete ohtudega ette võtta. Kliimamuutuse tagajärgede leevendamiseks ja vältimiseks on EL vastu võtnud uue kliimamuutusega

kohanemise strateegia, algatanud Euroopa rahelise kokkuleppe ja Euroopa kliimaseaduse. Siiski on **oluline üle vaadata, kas oleks võimalik kulutada rohkem ressursse kliimamuutuse äärmuslike ohtude uurimisele ja kõige kuluefektiivsematele, kuid tähelepanuta jäänud meetmetele.** Tehislike pandeemiate puhul on asjakohased algatused olnud EL-i terviseohutuse raamistik ja 2015. aastal süntetilise bioloogia ohtusid uuriva eksperdirühma loomine, kuid **vaja oleks uut eksperdirühma, kes hindaks eelkõige just kõige ohtlikumaid tehislike pandeemiate stsenaariume uute tehnoloogiliste arengute valguses.** Väärtustega ühildamata tehisintellekti ohtude vähendamiseks on EL-is asjakohased algatused olnud 2018. aastal loodud kõrgetasemeline tehisintellekti eksperdirühm ja praegu kavandatav tehisintellekti regulatsioon. **Samas peaks EL arvestama rohkem laiaulatuslike ohtude ja tuleviku ohtudega.**



Risto Uuk

Future of Life Institute'i poliitikaanalüütik

Kasutatud allikad:

- ¹ Ord, T. *The Precipice: Existential Risk and the Future of Humanity* (Hachette Books, 2020).
- ² Schubert, S., Caviola, L. & Faber, N. S. The Psychology of Existential Risk: Moral Judgments about Human Extinction. *Sci. Rep.* 9, 15100 (2019).
- ³ Bostrom, N. Existential Risk Prevention as Global Priority. *Glob. Policy* 4, 15–31 (2013).
- ⁴ Toshiko, K. & Haub, C. How Many People Have Ever Lived on Earth? <https://www.prb.org/howmanypeoplehaveeverlivedonearth/> (2020).
- ⁵ Cotton-Barratt, O., Farquhar, S., Halstead, J., Schubert, S. & Snyder-Beattie, A. *GLOBAL CATASTROPHIC RISKS 2016*. <https://globalchallenges.org/wp-content/uploads/2019/07/Global-Catastrophic-Risk-Annual-Report-2016.pdf> (2016).
- ⁶ Sandberg, A. & Bostrom, N. *Global Catastrophic Risks Survey*. 1–5. <https://www.fhi.ox.ac.uk/reports/2008-1.pdf> (2008).
- ⁷ Farquhar, S. Changes in funding in the AI safety field. *Centre for Effective Altruism*. <https://www.centreforeffectivealtruism.org/blog/changes-in-funding-in-the-ai-safety-field> (2017).
- ⁸ Worldwide Spending on Artificial Intelligence Systems Will Grow to Nearly \$35.8 Billion in 2019, According to New IDC Spending Guide. <https://www.businesswire.com/news/home/20190311005093/en/Worldwide-Spending-on-Artificial-Intelligence-Systems-Will-Grow-to-Nearly-35.8-Billion-in-2019-According-to-New-IDC-Spending-Guide> (2019).
- ⁹ *Biological Weapons Convention – Budgetary and Financial Matters* (2019).
- ¹⁰ Watson, C., Watson, M., Gastfriend, D. & Sell, T. K. Federal Funding for Health Security in FY2019. *Health Secur.* 16, 281–303 (2018).
- ¹¹ Buchner, B. *et al.* *Global Landscape of Climate Finance 2017*. <https://www.climatepolicyinitiative.org/publication/global-landscape-of-climate-finance-2017/> (2017).
- ¹² Geoengineering Research. *Open Philanthropy*. <https://www.openphilanthropy.org/research/cause-reports/geoengineering> (2013).
- ¹³ *Existential Risk Diplomacy and Governance*. <https://www.fhi.ox.ac.uk/wp-content/uploads/Existential-Risks-2017-01-23.pdf> (2017).
- ¹⁴ Goldblatt, C., Robinson, T. D., Zahnle, K. J. & Crisp, D. Low simulated radiation limit for runaway greenhouse climates. *Nat. Geosci.* 6, 661–667 (2013).
- ¹⁵ Ramirez, R. M., Kopparapu, R. K., Lindner, V. & Kasting, J. F. Can Increased Atmospheric CO₂ Levels Trigger a Runaway Greenhouse? *Astrobiology* 14, 714–731 (2014).
- ¹⁶ Zimov, S. A., Schuur, E. A. G. & Chapin, F. S. Permafrost and the Global Carbon Budget. *Science* 312, 1612–1613 (2006).
- ¹⁷ Tarnocai, C. *et al.* Soil organic carbon pools in the northern circumpolar permafrost region. *Glob. Biogeochem. Cycles* 23, (2009).
- ¹⁸ Tokarska, K. B., Gillett, N. P., Weaver, A. J., Arora, V. K. & Eby, M. The climate response to five trillion tonnes of carbon. *Nat. Clim. Change* 6, 851–855 (2016).

- ¹⁹ Richards, C. E., Lupton, R. C. & Allwood, J. M. Re-framing the threat of global warming: an empirical causal loop diagram of climate change, food insecurity and societal collapse. *Clim. Change* 164, 49 (2021).
- ²⁰ Duda, R. & Koehler, A. Report: Is climate change the biggest problem in the world? *80,000 Hours*. <https://80000hours.org/problem-profiles/climate-change/> (2020).
- ²¹ Pamlin, D. & Armstrong, S. *Twelve risks that threaten human civilisation*. <https://static1.squarespace.com/static/59dc930532601e9d148e3c25/t/5abf5a638a922d9f4625aff9/1522489978866/12+Risks+with+infinite+impact-Executive+Summary.pdf> (2015).
- ²² Kliimamuutuste suhtes vastupanuvõimelise Euroopa kujundamine – ELi uus kliimamuutustega kohanemise strateegia. <https://eur-lex.europa.eu/legal-content/ET/TXT/?uri=CELEX:52021DC0082> (2021).
- ²³ Euroopa roheline kokkulepe. <https://eur-lex.europa.eu/legal-content/ET/TXT/?qid=1576150542719&uri=COM%3A2019%3A640%3AFIN> (2019).
- ²⁴ EU climate action and the European Green Deal. *Climate Action – European Commission*. https://ec.europa.eu/clima/policies/eu-climate-action_en (2019).
- ²⁵ Euroopa kliimaseadus. *Kliimameetmed – Euroopa Komisjon*. https://ec.europa.eu/clima/policies/eu-climate-action/law_et (2020).
- ²⁶ Wolff, M. L., Jean Pisani-Ferry, Jeremy Shapiro, Simone Tagliapietra, Guntram. The geopolitics of the European Green Deal. *European Council on Foreign Relations*. <https://ecfr.eu/publication/the-geopolitics-of-the-european-green-deal/> (2021).
- ²⁷ Clean Energy Innovation. *Let's Fund*. <https://lets-fund.org/clean-energy/> (2019).
- ²⁸ Lane, L., Bickel, E., Galiana, I., Green, C. & Bosetti, V. *Fix The Climate: Advice for Policymakers*. <https://www.copenhagenconsensus.com/publication/fix-climate-advice-policymakers> (2009).
- ²⁹ Halstead, J. *Climate Change Cause Area* Report. <https://founderspledge.com/research/fp-climate-change> (2018).
- ³⁰ *Global Catastrophic Risks* 2017. <https://globalchallenges.org/initiatives/analysis-research/reports/> (2017).
- ³¹ Lewis, G. Reducing global catastrophic biological risks. *80,000 Hours*. <https://80000hours.org/problem-profiles/global-catastrophic-biological-risks/> (2020).
- ³² Harris, S. H. *Factories of Death: Japanese Biological Warfare, 1932–45 and the American Cover-Up*. (Routledge, 2002).
- ³³ Millett, P. & Snyder-Beattie, A. Existential Risk and Cost-Effective Biosecurity. *Health Secur.* 15, 373–383 (2017).
- ³⁴ EUROOPA PARLAMENDI JA NÕUKOGU OTSUS nr 1082/2013/EL, 22. oktoober 2013, tõsiste piiriüleste terviseohtude kohta ja millega tunnistatakse kehtetuks otsus nr 2119/98/EÜ. (2013).
- ³⁵ Health security and infectious diseases. *European Commission*. https://ec.europa.eu/health/security/overview_en (2020).
- ³⁶ *Opinion on Synthetic Biology II: Risk assessment methodologies and safety aspects*. https://ec.europa.eu/health/sites/health/files/scientific_committees/emerging/docs/scenihr_o_048.pdf (2015).
- ³⁷ Nelson, C. et al. *UK Government's approach to emerging infectious diseases and bioweapons*. <https://www.fhi.ox.ac.uk/wp-content/uploads/UKG-Biosecurity-Written-Submission-C-Nelson-et-al.pdf> (2019).

- ³⁸ Cumulative confirmed COVID-19 deaths. *Our World in Data*. <https://ourworldindata.org/coronavirus-data-explorer>.
- ³⁹ Cumulative confirmed COVID-19 deaths per million people. *Our World in Data*. <https://ourworldindata.org/coronavirus-data-explorer>.
- ⁴⁰ Bassot, É. *Ten opportunities for Europe post-coronavirus*. [https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/652035/EPRS_IDA\(2020\)652035_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/652035/EPRS_IDA(2020)652035_EN.pdf) (2020).
- ⁴¹ NICOLÁS, E. S. Von der Leyen concedes mistakes in vaccine rollout. *EUobserver*. <https://euobserver.com/coronavirus/150889> (2021).
- ⁴² Tegmark, M. *Life 3.0: Being Human in the Age of Artificial Intelligence*. (Knopf, 2017).
- ⁴³ *Global Catastrophic Risks 2020*. <https://globalchallenges.org/wp-content/uploads/Global-Catastrophic-Risks-2020-Annual-Report.pdf> (2020).
- ⁴⁴ Sotala, K. & Yampolskiy, R. V. Responses to catastrophic AGI risk: a survey. *Phys. Scr.* 90, 018001 (2014).
- ⁴⁵ Müller, V. C. & Bostrom, N. Future Progress in Artificial Intelligence: A Survey of Expert Opinion. in *Fundamental Issues of Artificial Intelligence* (ed. Müller, V. C.) 555–572 (Springer International Publishing, 2016). doi:10.1007/978-3-319-26485-1_33.
- ⁴⁶ Grace, K., Salvatier, J., Dafoe, A., Zhang, B. & Evans, O. When Will AI Exceed Human Performance? Evidence from AI Experts. *ArXiv170508807 Cs* (2018).
- ⁴⁷ Wiblin, R. & Harris, K. How well can we actually predict the future? *80,000 Hours*. <https://80000hours.org/podcast/episodes/katja-grace-forecasting-technology/> (2018).
- ⁴⁸ High-Level Expert Group on Artificial Intelligence. *Shaping Europe's digital future – European Commission*. <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence> (2018).
- ⁴⁹ *Eetikasunised usaldusväärse tehisintellekti arendamiseks*. https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI_ET.pdf (2019).
- ⁵⁰ Metzinger, T. Ethics washing made in Europe. <https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/24195496.html> (2019).
- ⁵¹ *Tehisintellekt: Euroopa käsitus tiptasemel ja usaldusväärsest tehnoloogiast*. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_et.pdf (2020).
- ⁵² Euroopa Parlamendi ja nõukogu määrus (ettepanek), millega nähakse ette tehisintellekti käsitlevad ühtlustatud õigusnormid (tehisintellekti käsitlev õigusakt) ja muudetakse teatavaid liidu õigusakte. https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0017.02/DOC_1&format=PDF (2021).
- ⁵³ Belfield, H. *et al.* Response to the European Commission's consultation on AI. *Centre for the Study of Existential Risk*. <https://www.cser.ac.uk/news/response-european-commissions-consultation-ai/> (2020).
- ⁵⁴ Torges, S. & Anderljung, M. *Consultation on the European Commission's White Paper on Artificial Intelligence: a European approach to excellence and trust*. <https://www.fhi.ox.ac.uk/wp-content/uploads/EU-White-Paper-Consultation-Submission-GovAI-Oxford.pdf> (2020).

