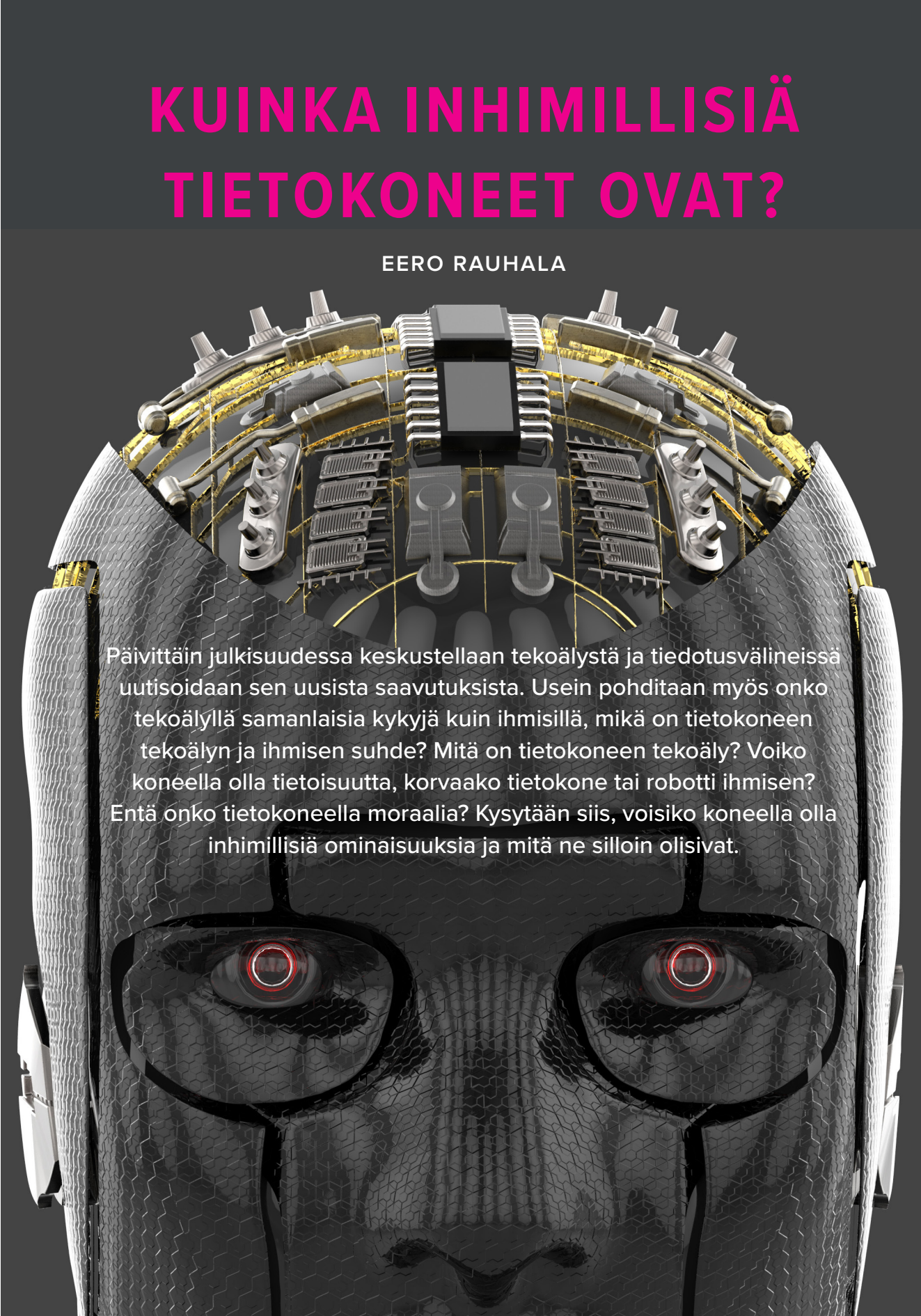


KUINKA INHIMILLISIÄ TIETOKONEET OVAT?

EERO RAUHALA



Päivittäin julkisuudessa keskustellaan tekoälystä ja tiedotusvälineissä uutisoidaan sen uusista saavutuksista. Usein pohditaan myös onko tekoälyllä samanlaisia kykyjä kuin ihmisillä, mikä on tietokoneen tekoälyn ja ihmisen suhde? Mitä on tietokoneen tekoäly? Voiko koneella olla tietoisuutta, korvaako tietokone tai robotti ihmisen? Entä onko tietokoneella moraalialia? Kysytään siis, voisiko koneella olla inhimillisiä ominaisuuksia ja mitä ne silloin olisivat.

Jotta konetta ja ihmistä voitaisiin verrata, pitää ensin selvittää mitä nämä kyvyt tarkoittavat ihmisessä. Mitä on ihmisen tietoisuus? Jos inhimillisyydellä tarkoitetaan sellaista, mitä ihmisellä on, mutta muilla olioilla ei mahdollisesti ole, on myös kysyttävä millä tavoin ihminen ja kone ovat samankaltaisia, millä tavoin ne eroavat ja onko ihmisessä jotakin, jota koneessa ei voisi olla?

Vastauksia tällaisiin inhimillisen olemassaolon peruskysymyksiin ovat aikojen kuluessa tarjonneet monet uskomukset ja käsitykset eri kulttuureissa, uskonnot, tieteet ja filosofiat. Ne ymmärretään nykyisin erityisesti filosofisina perusongelmina.

Käsityksiä ihmisen ja maailman olemassaolosta voidaan länsimaisen filosofian historiassa tarkastella yksinkertaistaen pitäen lähtökohtana kahden pitkäaikaista filosofista perinnettä, yhtäältä aineellisen ja toisaalta henkisen¹ olemassaolon ensisijaisuutta tähdentäviä ajatussuuntia. Filosofisessa ontologiassa, opissa maailman perusrakenteesta, nämä käsitykset ovat aikojen kuluessa esiintyneet monenlaisina. Lyhyesti luonnehtien, *monististen*, yksiarvoisten ontologioiden mukaan kaikki olemassaolo palautuu jompaankumpaan perusnäkemykseen. Esimerkiksi nykyisissä luonnontieteissä maailman rakenne ymmärretään usein naturalistisen ja materialistisen monismin mielessä, kaikki olevainen palautuu aineellisen tapahtumiseen. *Dualististen*, kaksiarvoisten näkemysten mukaan olemassaolo nähdään eri tavoin sekä aineellisena että henkisenä. *Pluralististen*, moniarvoisten filosofien ontologioiden maailmaa ja sen ilmiöitä kuvataan monitasoisina. Aineellisuutta ja etenkin henkisyyttä analysoidaan silloin erilaisina differentioituneina olemisen dimensioina. *Holistinen*, kokonaisvaltainen ajattelu tähdentää mm. dynaamista vuorovaikutusta ja osien muodostamaa kokonaisuutta.

Tarkastelen yllä esitettyjä kysymyksiä tietokoneista ja ihmisestä *fenomenologian* näkökulmasta. Esittelen aluksi lyhyesti Lauri Rauhalan holistista ja pluralistista fenomenologista ihmiskäsitystä. Tavoitteena on osoittaa, kuinka fenomenologisesta lähtökohdasta seuraa, että tietokoneilla ja tekoälyllä ei ole tajunnallisuutta, tietoisuutta eikä ihmiselle ominaista suhdetta todellisuuteen.

Ihmiskäsitys

Vastaukset edellä esillä oleviin kysymyksiin ihmisestä ja tietokoneesta suuntautuvat sen mukaan, minkälaisesta filosofisesta lähtökohdasta ongelmia tarkastellaan. Vuonna 2016 edesmenneen suomalaisen filosofisen ihmistutkimuksen uranuurtajan Lauri Rauhalan mukaan ihmistä on tutkittava kolmena olemismuotona: Ihminen on olemassa henkisenä ja tajunnallisena toiminnallisuutena², kehollisena ja aineellisena oliona sekä suhteina todellisuuteen; lyhyesti *tajunnallisena*, *kehollisena* ja *situationaalisena*³ oliona. Tämän pluralistisen ja holistisen filosofisen analyysin⁴ taustana on *eksistentiaalinen*, *hermeneuttinen* fenomenologia.

Tässä ajattelussa ihmisen olemuspuolet läpäisevät toisensa jatkuvassa ja vastavuoroisessa vuorovaikutuksessa. Näin muodostuu ihmisen kokonaisuus, mutta osat eivät palaudu toisikseen. Ilman niitä kaikkia, ihminen ei ole ihminen. Ihmistieteissä ongelmat pelkistyvät näille kolmelle tasolle eri tieteenalojen tutkittaviksi. Siten aineellista olemassaoloa tutkivat lähinnä luonnontieteet, henkistä kokemusmaailmaa kokemustieteet⁵, esimerkiksi psykiatria, psykologia ja kognitiotiede, ja ihmisen suhdetta ulkomailmaan sen eri tasoilla mm. sosiaalitieteet, psykologia, lääketiede jne. Tieteet tuottavat omanlaisia kuvauksiaan ihmisestä, *ihmiskuvia*, filosofisen analyysin tuloksena syntyy käsitys ihmisestä moniulotteisena kokonaisuutena, *ihmiskäsitys*.⁶

Tämän ihmiskäsityksen mukaan ymmärrys maailman todellisuudesta ilmenee ihmisen tajunnassa, tajunnallinen toiminta toisaalta myös muokkaa tätä todellisuutta. Tajunnassa ymmärretään muut ihmisen olemuspuolet. Tajunta ja sen suhde todellisuuteen on se ihmisen olemus, joka selvimmin erottaa ihmisen muista olioista.

Tajunta ja tietoisuus

Tajunnallisuus on kyky oivaltavaan ja ymmärtävään kokemukseen; *tajunta* on tämän kokemisen kokonaisuus.⁷ *Kokemussisällöt* ovat esimerkiksi tietoa, tunnetta, uskoa, tahtoa, arvoja jne. Kaikki kokemussisällöt ovat *merkitystä*, ne tarkoittavat meille jotakin. Tajunta on olemassa vain merkityksinä, se ei esimerkiksi ole jokin olio tai sijaitse jossakin paikassa. Kaikkia tajunnan sisältöjä ei aina tiedosteta, ne voivat myös olla *tiedostamattomia*.

*Oliot*⁸ voidaan tajunnan käsitteen avulla jakaa kahteen luokkaan: ne ovat joko tajuisia tai *tajuttomia*. Ihmiset ja monet eläimet ovat tajuisia olentoja, esim. koneet ja laitteet tajuttomia, niiden prosesseihin ei itsessään missään vaiheessa liity kokemuksia eikä merkityksiä. Myös ihmisessä tapahtuvat aineelliset ja orgaaniset prosessit ovat tajuttomia.

Tämän ihmiskäsityksen mukaan *ihmisen tietoisuus* on juuri tätä ymmärrystä, sitä tajunnallista aktiiviteettia, jossa kokemuksellinen todellisuus jäsentyy meille mielekkäinä ja järkevinä merkityksinä. Ihmisen tietoisuus on siten *mielen ilmenemistä*⁹. Se edellyttää kaikki muutkin ihmisen olemuspuolet, kehon aisteineen sekä joka hetki ainutkertaisen vuorovaikutustilanteen ja suhteen ulkomaailmaan ja kanssaihmiisiin. Tietoisuus ei palaudu muille tasoille, esim. aineelliseksi rakenteeksi.

Fenomenologia tavoittelee merkityksellisyuden jäsentymisen ja organisoitumisen prosessin selvittämistä arkipöytäkokemuksen asioiden ja esineiden perustana. Me ihmiset koemme todellisuuden aina jo valmiiksi jäsentyneenä meille ilmenevinä merkityksinä. Todellisuus ei ole eriytymätöntä, jäsentymätöntä ja hahmotonta kaaosta, vaan se on differentioitunut ja järjestynyt esineiksi ja asioiksi; näemme tuoleja, pöytiä, puita, eläimiä ja ihmisiä. Emme ensisijaisesti hahmota olioiden erillisiä osia, vaan koemme ne kokonaisuuksina, *merkityssuhteina*. Kommunikoimme tätä todellisuutta toisillemme kielellisinä, abstrakteina ja ideaalisina *käsitteinä*.

Fenomenologiassa kutsutaan *konstituutioksi*¹⁰ sitä mielellistä prosessia, jossa subjektille ilmenevät asiat ja esineet saavat merkityksensä, muodostuvat ja koostuvat meille mielekkääksi todellisuudeksi. Konstituutio tapahtuu vastavuoroisessa vuorovaikutuksessa subjektin ja ilmenevän todellisuuden välillä¹¹. Tämän ajattelun mukaan subjektia ja objektia ei voi tarkastella erillisinä; tuolista tulee tuoli vasta tässä konstituutiossa. Tuolia ei ole tuolina ilman että joku ymmärtää esineen tuoliksi.¹²

Ihmisen kokemuksellisuuden kyky on juuri ihmisille ominainen. Merkityksellisen todellisuuden syntymisen prosessi, konstituutio, on suurelta osin tiedostamaton ja *passiivinen*. Todellisuus jäsentyy meille esimerkiksi ajallisenä ja tilallisenä. Meille ihmisille on olemassa aika – menneisyys, nykyhetki ja tulevaisuus. Koemme oliot kolmiulotteisessa tilassa, jossakin paikassa, tietyn kokoi-

na. Ne liikkuvat tai ovat paikoillaan. Muun muassa näitä todellisuuden perusrakenteita emme tuota tietoisesti, vaan ne ilmenevät meille passiivisesti.

Konstituutio toimii aktiivisena, tietoisuuden ohjaamana esimerkiksi opiskelussa, ihmisen luovassa toiminnassa taiteissa, tieteissä ja kulttuurissa yleensä. Olemme vuorovaikutuksissa toisten ihmisten ja yhteisöjen kanssa yhteiskunnan toiminnoissa, tieteellisessä tutkimuksessa jne. Nämäkin aktiiviset suhdeverkostot ovat vastavuoroisia ja kehämäisiä, konstituutiossa ilmenevä uusi ymmärrys todellisudesta jäsentyy jo aiemmin konstituoidun todellisuuden ja tajuntamme kehässä.

Situaationaalisuus, kehollisuus ja ihmisen kokonaisuus

Situaationaalisuus on tajunnallisuuden ja kehollisuuden vastavuoroisia suhteita todellisuuteen¹³. Emme voi olla pelkästään vain olemassa, vaan olemme aina jossakin elämäntilanteessa, liittyneenä ja sijoittuneena konkreettisiin ja ideaalisiin olosuhteisiin. Ihmisen tilanteeseen kuuluu fyysisen ympäristön lisäksi muita ihmisiä, olentoja ja luontoa. Meillä on vanhemmat, perhe, ystävät ja työtoverit, olemme osallisina monimuotoisissa sosiaalisissa verkostoissa, yhteisöissä ja yhteiskunnassa. Situaatio sisältää kulttuurisia, historiallisia ja kielellisiä tekijöitä, arvoja ja normeja. Juuri nyt elämme tässä ajassa ja olemme yksilöinä tietyn ikäisiä.¹⁴

Kehomme on suhteessa ympäristöön sekä biologisesti kuvattavana elävänä organismina että fyysisesti kuvattavana elottomana, aineellisena rakenteena.¹⁵ Kehollisuus ei myöskään ole irrallinen muista ihmisen olemuspuolista. Tajunnassa yhtäältä koemme kehon ja ymmärrämme sen toimintaa, toisaalta kokemamme aineellisen todellisuuden perusta on kehollisuudessa. Tämä todellisuus ilmenee meille juuri tietynlaisen aisti- ja hermojärjestelmän välittämän informaation välityksellä. Tilallisuus jäsentyy kehon fyysisen tilan, sen koon ja muodon mukaan, ajallisuus elinprosessien, kokemamme elämän ja ihmisiän perusteella. Keho on tällä tavoin yhtäältä konstituoitu todellisuuden komponentti, toisaalta osallisena todellisuuden konstituutiossa.

Ihmisen kolmen olemuspuolen yhteispelissä etenevässä konstituution dynamiikassa syntyy kokonainen ihmisyksilö. Tätä dynamiikkaa voidaan

yksityiskohtaisemmin analysoida kaksitasoisina *hermeneuttisina kehinä*⁶. Tajunta ja situaatio muodostavat siten oman kehänsä, tajunta ja keho sekä keho ja situaatio omansa. Kullakin näistä kehistä on omanlaatuisensa rakenne ja toiminnallisuus.

Ihmisenä olemisen toteutumisessa *tajunnan ja tilanteen kehä* on merkittävässä erikoisasemassa, sillä juuri tässä kehässä muodostuvat merkityssuhteet, ymmärtävät suhteet todellisuuteen. Alati muuttuva elämäntilanteemme on omakohtainen ja jatkuvasti ainutkertainen, meillä jokaisella on oma koettu elämämme, eikä elämäntilanteemme koskaan toistu samanlaisena uudestaan. Tajunta saa merkitysten aiheet juuri yksilöllisesti koetusta tilanteesta. Ymmärrämme uudet kokemuksemme suhteessa omaan koettuun elämäämme. Tajunnan ja tilanteen dynamiikka muokkaa meidät yksilöllisesti *persooniksi*. Tässä kehässä toisaalta avautuvat ja kehittyvät myös *ylilyksilliset* merkityssuhteet, kuten ihmisyyden syväinen eettinen ulottuvuus, yhteisöllisyys yleensä sekä yhteisen, objektiivisen todellisuuden konstituitio.

Kun ihminen ymmärretään edellisen mukaan kokonaisuutena, jossa eri olemuspuolet ovat vastavuoroisessa vuorovaikutuksessa, ovat aivot tämän kokonaisuuden yksi kehollinen komponentti. Tajunnallisuus, kokemuksellisuuden kyky ja sen mahdollistama tietoisuus eivät tämän ajattelun mukaan synny yksin aivoissa, vaan ihmisen kokonaisuuden dynamiikassa. Tietoisuus ei siten edellytä pelkästään aivoja, vaan kaikki ihmisen olemassaolon eri tasot, tajunnallisuuden, situationaalisuuden ja keholisuuden, sekä näiden jatkuvan vuorovaikutuksen. Vasta tässä moniulotteisessa toiminnallisuudessa voi syntyä ihmisen differentioitunut, jäsentynyt, merkityksellinen käsitys todellisuudesta.

Tietokone ja tekoäly

Tietokoneen⁷ tekoälyllä tarkoitetaan tavallisesti tietokoneohjelmia, itseohjautuvia ja kehittyviä toiminta-algoritmeja ja niiden autonomista toimintaa. Sillä voidaan tarkoittaa myös laajemmin tietokoneen laskentakykyä ja kapasiteettia, koneiden muodostamia verkostoja, itseorganisoiuvia neuroverkkoja, laitteistoja ja ohjelmistoja yhdessä jne. Tekoälylle voidaan opettaa toimintamalleja ja hahmoja erilaisia sovellustilanteita varten. Parhaimmillaan se voidaan ohjelmoida tai opettaa

tavoitteelliseksi ja joustavaksi sekä toimimaan ja reagoimaan ympäristöstä saatavan palautteen mukaan jonkin päämäärän saavuttamiseksi. Kone tai robotti, jolla on kehittyntä tekoälyä, voi *simuloida*, ulkoiselta toiminnaltaan jäljitellä ja muistuttaa ihmisen toimintaa.

Tarkastellaan esimerkkinä tekoälystä itsenäisesti ajavaa autoa. Autossa on monimutkainen järjestelmä erilaisia teknisiä laitteita. Auton autonominen toiminta perustuu tekoälyyn, laitteistojen ja ohjelmistojen toiminnalliseen kokonaisuuteen. Auto lähettää sähkömagneettista säteilyä ja ääniaaltoja eri taajuuksilla. Se vastaanottaa kohteista takaisin heijastuvia ääni, valo-, infrapuna- ja ultraviolettisäteilyä, laser- ja ledipulsseja. Auton konenäköön liittyy ultraääni- ja optisia ilmaisimia, lämpökameroita ja muilla aallonpituuksilla toimivia kameroita ja videokameroita. Auto seuraa tiemerkintöjä ja opasteita, määrittää paikkaansa ja nopeuttaan pyörissä olevien antureiden ja GPS-järjestelmän avulla. Sillä on kiihtyvyydantureita ja Doppler-tutkia auton omien liikkeiden hallintaan ja ympäristön liikkeiden havaitsemiseen. Se on yhteydessä muiden ajoneuvojen tietokoneisiin, ulkopuolisiin tietoverkkoihin ja seuraava liikennetilannetta.

Kaikkea tätä eri vastaanottimista tulevaa informaatiota auton tietokoneet yhdistelevät ja käsittelevät monimutkaisilla ohjelmilla. Esimerkiksi kohteesta takaisin heijastuvia laserpulsseja voidaan havaita ilmaisimella, josta saatava mittausdata vietään algoritmin kolmiulotteiseen hilaruudukkoon. Tätä dataa yhdistetään ympäristöä kuvaavan karttakoordinaatiston ruudukkoon ja käsitellään erilaisilla laskennallisilla matemaattisilla ja tilastollisilla menetelmillä. Algoritmit tuottavat mallin ympäristön geometriasta ja siinä tapahtuvista liikkeistä auton liikkeisiin nähden.

Ympäristön hahmotus ja tunnistus perustuu havaintojen vertaamiseen tietokoneen muistissa oleviin, sille opetettuihin malleihin. Hiukan yksinkertaistaen, konetta voidaan opettaa käyttämällä miljoonaa kuvaa jalankulkijoista ja pyöräilijöistä. Hahmontunnistusalgoritmi pelkistää näistä kuvista kohteiden tyypillisiä perusmuotoja ja -piirteitä ja osaa niiden avulla erottaa pyöräilijän jalankulkijasta. Näihin opetettuihin hahmoihin algoritmit vertaavat havaintolaitteiden tuottamaa informaatiota ympäristöstä auton kulkiessa.

Fenomenologisessa ajattelussa todellisuuden ilmeneminen edellyttää tiedostavan subjektin, mutta tekoäly ei ole tällainen subjekti. Juuri tässä kiteytyy perustava ero tajuttoman tekoälyn ja tajuisen ihmisen välillä: vain tajuiselle olennolle on olemassa jotakin, vain hänelle on olemassa konstituoitu objektiivinen todellisuus.¹⁸ Meille ihmisille tuttu arkitodellisuus on niin itsestään selvä, että on vaikea käsittää, miten mikään ei merkitse tekoälylle mitään eikä se ymmärrä mitään ympäristöstään.

Konenäkö ei siten näe jalankulkijaa ihmisenä. Koneelle ei ole muitakaan erillisiä objekteja, ei yksilöitä, ei mitään ihmisen käsitteellis-kielellistä jäsenystä todellisuudesta. Se ei voi ymmärtää esineitä, ei autoja, rakennuksia, ajoväylää, mitään ympäristön kohteita tai piirteitä. Paitsi autonomisissa kulkuvälineissä ja konenäössä, kaikissa muissakin tekoälyn sovelluksissa, kuten puheen- ja kielentunnistuksessa, kielen kääntämisessä ja laajojen tietoaaineistojen eli ”big datan” seulonnassa, tekoäly on tajuton. Se ei ymmärrä käsittelemäänsä materiaalin, sanojen ja käsitteiden merkityksiä.¹⁹

Kun puhumme koneen toiminnasta ja ominaisuuksista, käytämme, kuten edellä, ihmisen maailmasta lainattua kieltä ja käsitteitä. Puhe inhimillisistä ominaisuuksista ja toiminnoista koneiden yhteydessä on kuitenkin harhaanjohtavaa. Kone ei ihmisen tavoin opi, tiedä, eikä ajattele, sillä ei ole tunteita, tahdosta ja arvoista puhumattakaan. Täsmällisessä ilmaisussa tulisi siis tässä yhteydessä käyttää jotakin muuta terminologiaa. Esimerkiksi etuliitteiden ”kone”, ”teko” tai ”keino” avulla voisimme puhua koneoppimisesta, keino- tai tekoälystä, konevalinnoista jne. Käytännössä tämä ei liene aina mahdollista, mutta tulisi silti muistaa, että ihmistä kuvaavaa kieltä käytetään uudessa merkityksessä tietokoneiden yhteydessä.

Ihminen ja tekoäly

Jokapäiväisessä keskustelussa ihmisen henkisistä kyvyistä on aivoilla usein korostunut asema. Saatetaan esimerkiksi sanoa aivojen kokevan, tuntevan, ajattelevan ja oppivan. Aivoista ei kuitenkaan missään tutkimuksessa tai mittauksessa löydy näitä kykyjä, sieltä voidaan vain mitata sähkökenttiä, -virtoja ja -pulsseja, magneettikenttien vaikutuksia jne.²⁰ Arkikeskustelussakin pitäisi siten ihmisen sisäisen kokemusmaailman yhteydessä puhua

mielestä, kokemuksesta tai ihmisestä kokonaisuutena aivojen asemasta.

Kysymys ihmisen tietoisuudesta on filosofisen ihmistutkimuksen vanhimpia ja keskeisimpiä ongelmia. Paitsi filosofiassa, tätä kysymystä tutkitaan nykyisin myös erityisesti kognitio- ja neurotieteissä, joissa tietoisuusteorioita esiintyy kymmeniä ellei satoja.²¹ Ihmisen aivoja ajatellaan usein tietokoneen mallina ja ihmisälyä tekoälyn tavoitteena. *Aivometaforan* mukaan tietokone voisi toimia kuten ihmisen aivot ja tietokoneohjelma kuten tietoisuus. Neuro- ja kognitiotieteiden niin sanottu *vaikea ongelma* liittyy dualistiseen ja reduktionistiseen ajatteluun: Miten aineelliset aivot tuottavat henkisen tietoisuuden?²² Ongelma nähdään tässä tavallisesti kysymyksenä juuri aivojen ja tietoisuuden suhteesta. Tekoälytutkijoiden, kognitio- ja neurotieteilijöiden kesken vallitsee laaja yhteisymmärrys siitä, että vaikea ongelma tällä tavoin esitettyinä on edelleen ratkaisematta. Varsin yleisesti myönnetään lisäksi, ettei koneilla ole tietoisuutta.²³

Tässä kuvatussa fenomenologisen näkemyksen perusteella vaikea ongelma on näennäinen. Se syntyy, kun pyritään materialistisen monismin tai joidenkin dualististen katsomusten mukaan ihmisen henkisen olemuksen reduktioon aivojen aineelliseen rakenteeseen ja prosesseihin. Jos kuitenkin, kuten tässä, jo lähtökohtana tunnustetaan henkisen olemuspuolen alkuperäisyys, itsenäisyys ja palautumattomuus, ongelma häviää. *Neurofenomenologiassa*²⁴ tavoitellaan samoin reduktionismista vapaata asennetta, käsitystä tietoisuuden omalaatuisesta ja riippumattomasta asemasta. *Fenoneurologia*²⁵ kääntää vaikean ongelman asetelman: Sen sijaan että etsisimme tietoisuutta aivoista, ongelma katoaa, kun käsitämme, että kaikki mitä aivoista tiedämme seuraa ihmisen tietoisesta ajattelusta, ei päinvastoin. Tietoisuuden tuottamaa kuvausta aivoista ei voi käyttää tietoisuuden itsensä selittämiseen.

Kun pyrimme ymmärtämään ihmisen ja koneen suhdetta fenomenologian avulla, voisimme yrittää analysoida myös koneen olemassaoloa soveltamalla siihen edellä kuvattua ihmisen olemassaolon kolmijakoa tajunnallisuuteen, kehollisuuteen ja situationaalisuuteen²⁶. Jo edellä on käynyt ilmeiseksi, että koneelta puuttuu tajunnallisuus. Sen ”kehollisuus” on vain elotonta ja aineellista, siihen ei liity orgaanista tapahtumista. Koneen ”kehollisuuden”

ja ”konesituaation” hermeneuttinen kehä rajoittuu siten vain tajuttoman aineellisen, fysikaalisesti kuvattavan koneen rakenteen ja ympäristön vuorovaikutuksiin. Niitä voisivat olla esimerkiksi sähkö- ja magneettikenttiin liittyvät vaikutukset, koneen yhteydet tietoverkkoihin, ilmaisimilta ja havaintovälineiltä saatava fysikaalinen informaatio sekä prosessoreiden ja verkkojen rakenteiden, koon ja energiankulutuksen asettamat rajoitteet. Tajunnallisuuteen liittyvä kehä on surkastunut yksisuuntaiseksi: kun koneelta itseltään puuttuu tajunnallisuus, vaikuttaa sen konesituaatiossa vain yksisuuntainen tajunnallinen komponentti, ihmisen tietoinen toiminta koneen rakentajana, ohjelmoijana, käyttäjänä, tavoitteiden ja rajoitusten asettajana.²⁷

Tietokoneen tekoäly on kuitenkin ihmiselle ilmenevän kokemusmaailman yksi osatekijä. Tekoäly muokkaa todellisuutta ja synnyttää uusia todellisuuden ulottuvuuksia. Tietokonemaailma tuottaa uusia merkitysten sisältöjä ihmisen tajunnassa. Ihminen ja tietokone muodostavat näin uuden hermeneuttisen kehän, jossa tietokoneet ja tekoäly toimivat sekä konstituoituina että konstituivina todellisuuden komponentteina.

Tulevaisuus ja tekoäly

Voisimmeko ajatella, että joskus tulevaisuudessa myös tietokone olisi tietoinen kuten ihminen? Tässä omaksutun tarkastelutavan mukaisesti on ensin jälleen palattava kysymään inhimillistä tietoisuutta. Mikä on tajunnallisuuden, kokemuksellisuuden kyvyn alkuperä ihmisessä, miten tämä kyky tuottaa inhimillisen tietoisuuden?

Ihminen on saanut tajunnallisuuden kyvyn syntymälahjana. Miten se on meille ihmislajina ja ihmisyksilöinä kehittynyt, on täysi arvoitus. Tämän kyvyn mahdollistama tietoisuuskaan ei synny ihmiselle valmiina. Ihmisen syntymän hetkellä, mahdollisesti jo aiemminkin, tajunnallisuuden kyky alkaa kuitenkin hämmästyttävällä tavalla toimia kehittyvän kehon, aistien ja hahmottuvan ympäristön yhteispelissä. Aluksi tässä vuorovaikutuksessa keskeisenä ovat oman kehon ja aistien toiminta. Tajunta ja tietoisuus, kokemukset sisältöineen ja situaatio monine eri ulottuvuuksineen hahmottuvat vasta vähitellen, vuosien kuluessa, ymmärretyiksi merkityksiksi, kielellisesti jäsen-tyneiksi käsitteiksi, meille ihmiselle ilmeneväk-

si järjestyneeksi ja rakenteiseksi todellisuudeksi.²⁸

Entä tekoälyn tietoisuus? Tämän esityksen näkemyksessä on vaikea ymmärtää, miten jokin edellisen kaltainen tajunnallisuuden kyky voisi syntyä koneelle tai edes mahdolliselle tulevaisuuden orgaaniselle, keinoitekoiselle neuroverkolle. Joka tapauksessa on selvää, että ihmisen tajunta on vain ihmisen oma, inhimillistä tietoisuutta ei koneelle synny. Mitä jokin ”konetietoisuus” sitten voisi olla? Miten se määriteltäisiin, miten se todettaisiin ja miten sitä tutkittaisiin? Näihin filosofisiin ja reaalitieteellisiin, mm. tekoälytutkimuksen, kysymyksiin meillä ei ole vastausta.²⁹

Tulevaisuuden tietokoneilla on siis edelleen vain tekoälyä, mutta ei tietoisuutta, vain ihmisellä (ja eläimillä) sitävästoin on sekä älyä että tietoisuutta yhdessä. Tajutonkin tekoäly voitaneen silti kehittää äärimmäisen monipuoliseksi ja tehokkaaksi. Tekoälystä, joka toimii tietyissä erityistilanteissa, kuten autonomisen auton tapauksessa edellä, pyritään kehittämään *yleistekoälyä*, joka ihmisen älyn tavoin toimisi erilaisissa tilanteissa. *Supertekoälystä* tulisi ihmistä älykkäämpi kaikissa tilanteissa.³⁰

Vain ihminen on inhimillinen

Vastauksia kysymyksiin tietokoneen ja ihmisen suhteesta ohjaavat filosofiset perusasetteet. Olen edellä esitellyt fenomenologista ihmiskäsitystä, jotta ymmärtäisimme, miten perustavalla tavalla ihminen tämän ajattelun mukaan eroaa tietokoneista ja tekoälystä. Kun vertaamme ihmistä koneeseen, huomaamme, että koneella ei ole tässä edellä kuvattuja inhimillisiä ulottuvuuksia. Kone on tajuton ja eloton, siinä tapahtuvat prosessit ovat vain aineellisia ja fysikaalisesti kuvattavia. Koneella ei ole situaatiota, yksilöllistä elävää kehoa, omaa itsenäistä elämää, se ei ole persoona. Sillä ei ole sosiaalisia suhteita, historiaa, käsitteitä, aikaa jne.

Jos asennoidutaan naturalistisen ja materialistisen monismin mukaisesti olemassaolon aineellista perustaa korostaen, saatetaan ajatella, että ihmisen tietoisuus palautuu aivojen toiminnaksi. Aivoja jäljittelevän neuroverkon toivotaan silloin kehittyvän yhä älykkäämmäksi ja enemmän ihmisen kaltaiseksi. Konetietoisuuden uskotaan heräävän riittävän monimutkaiseen neuroverkkoon esimerkiksi *emergentinä* ominaisuutena. Tällaisen koneen tietoinen tekoäly haluaisi ehkä kehittyä rajattomasti;

se synnyttäisi aina uutta, entistä tehokkaampaa tekoälyä. Voisimme näin päätyä *singulariteettiin*, jossa ihminen tulisi koneiden maailmassa jossakin vaiheessa tarpeettomaksi. Äärimmäisenä uhkakuvana tietoiset koneet riistäytyisivät ihmisen vallasta.

Jos sen sijaan pidämme henkisyttä ihmisen olemassaolon alkuperäisenä, itsenäisenä, aineeseen palautumattomana filosofisena perustana, esimerkiksi alussa kuvatun Lauri Rauhalan moniarvoisen ihmiskäsityksen tapaan, ei tietoisuus synny koneeseen ja tietokoneen tekoäly on jotakin aivan muuta kuin ihmisen älyä.

Fenomenologisesta näkökulmasta tietokoneiden ja ihmisen kokemuksellisen todellisuuden suhde on ihmisen kannalta kaksisuuntainen ja vastavuoroinen. Tietokoneet ovat osana ihmisen kokemaa todellisuutta. Samalla ihminen niiden kautta muokkaa todellisuutta. Vaikka koneet ovat tajuttomia, voidaan odottaa, että tekoäly tulee nopeasti ja perusteellisesti muuttamaan ihmisen kokemaa todellisuutta.³¹

Eettisyyden kyky kuuluu erottamattomasti ihmellisyyteen. Eettisyys edellyttää tajunnallisuutta ja situationaalisuutta. Tietokone ei ole tajuihin eikä tietoinen moraalinen toimija, se ei tiedä tekevänsä hyvää tai pahaa. Sillä ei ole oikeuksia, velvollisuuksia eikä vastuuta. Tajuttomienkin tietokoneiden vaikutukset ulottuvat kuitenkin kaikkiin ihmisyksilöihin, yhteiskuntiin ja koko ihmiskuntaan. Tietokonemaailma voidaan ymmärtää yhtä aikaa sekä valtavana mahdollisuutena että suunnattomana uhkana ihmiskunnalle. Tietokoneilla ja tekoälyllä on siten merkittävä moraalinen ulottuvuus, joka ilmenee ihmisten kokemusmaailmassa. Vastuu tekoälyn kehityksestä ja tietokoneiden vaikutuksista säilyy kuitenkin aina ihmisillä.³²

Ihminen on monitasoinen kokonaisuus; kokemuksellisuuden kyky, tajunta ja tietoisuus ovat syntyneet tässä ihmisen kokonaisuudessa. Tietoisuus ei toteudu vain aivoissa, se ei ole yksinomaan aivojen kyky, vaan vaatii kokonaisen ihmisen. Ihmistä erottaa muista olioista juuri tajunnallisuus ja sen suhde todellisuuteen rikkaana ja moniulotteisena henkisenä elämänä. Tällaista ihmellisyyttä koneilla ei ole, koneet eivät muutu ihmisiksi eivätkä ne korvaa ihmistä ihmisenä. Uhkana tulevaisuudessa eivät ole tietoisien tekoälyn kääntyminen ihmisiä vastaan eivätkä robottien kapina. Meidän

ihmisten tulee sen sijaan ottaa vastuamme tietokoneiden kehityksestä vakavasti ja kantaa huolta siitä, että voimme säilyttää täyden ihmellisyytemme konemaailman keskellä.

Kirjallisuus

- Backman, Jussi, 2016. Situationaalinen säätöpiiri: Rauhalan filosofinen kädenjälki. Teoksessa *Kokemuksen tutkimus V*, Lauri Rauhala 100 vuotta. Virpi Tökkäri (toim.). Lapin Yliopistokustannus. 73–102.
- Backman, Jussi, 2009. Lauri Rauhala ihmisen ainutkertaisuuden ajattelijana. Teoksessa Lauri Rauhala, *Henkinen ihminen*. Gaudeamus. 299–367.
- Backman, Jussi, 2005. *Omaisuuksia ja elämä – Heidegger ja Aristoteles kreikkalaisen ontologian rajalla*, 23⁴⁵, Niin ja näin -lehden filosofinen julkaisusarja. 92–104, 303.
- Bitbol, Michel, 2017. Phenomenology, *Constructivist Foundations*, 12(2), 112–115.
- Bitbol, Michel, 2012. Neurophenomenology, an Ongoing Practice of In-Consciousness, *Constructivist Foundations*, 7(3), 1651–73.
- Bitbol, Michel, 2008. Is Consciousness Primary?, *NeuroQuantology*, 6(1), 53–71.
- Chalmers, David J., 1995. Facing up to the Problem of Consciousness, *Journal of Consciousness Studies* 2(3), 200–219.
- Haikonen, Pentti O. A., 2017. *Tietoisuus, tekoäly ja robotit*. Arthouse.
- Himanka, Juha ja Honkela, Timo, 2017. Tekniikan kehitys tarvitsee filosofeja, *Kanava* 45, 47–49.
- Himanka, Juha, 2017. Kuinka tehdä filosofiaa? Onko sokraattinen menetelmä edelleen ajankohtainen? *Tiede ja edistys* 2, 140–152.
- Himanka, Juha, 2002. *Se ei sittenkään pyyri – johdatus mannermaiseen filosofiaan*. Tammi, 131–159, 219.
- Honkela, Timo, 2017. *Rauhankone. Tekoälytutkijan testamentti*. Gaudeamus.
- Merleau-Ponty, Maurice, 2008. *Phenomenology of Perception*. Routledge.
- Rauhala, Eero, 2013. Higgsin hiukkanen ja Husserlin subjektiveetti, *Tieteessä tapahtuu* 4, 13–21.
- Rauhala, Eero, 2011. Voisivatko fyysikotkin ymmärtää fenomenologiaa? *Tieteessä tapahtuu* 8, 31–28.
- Rauhala, Lauri, 2009. *Henkinen Ihminen*. Gaudeamus.
- Rauhala, Lauri, 2009. Kokemuksen häiriöt ja sairauden käsite, *Tieteessä tapahtuu* 6, 24–30.
- Rauhala, Lauri, 2005. *Ihminen kulttuurissa – kulttuuri ihmisessä*. Yliopistopaino.
- Rauhala, Lauri, 1995. *Tajunnan itsepuolustus*. Helsinki, Yliopistopaino.
- Rauhala, Lauri, 1973a. The Regulatory Situational Circuit in Psychic Disturbance and Psychotherapy. Teoksessa *Studia philosophica in honorem Sven Krohn septuagesimum annum complemtis 9 V. 1973*, T. Airaksinen ym. (toim.) Annales Universitatis Turkuensis, Series B: Humaniora, 126. Turun Yliopisto.
- Rauhala, Lauri, 1973b. Situationaalinen säätöpiiri psykkisissä häiriöissä ja psykoterapiassa. *Helsingin yliopistollisen keskussairaalan psykiatrisen klinikan julkaisusarja*, 23. Helsinki, Psychiatria Fennica.
- Tökkäri, Virpi, 2015. (toim.) Lauri Rauhala 100 vuotta. *Kokemuksen Tutkimus V*, Lapin yliopistokustannus.
- Varela, Francisco J., 1996. Neurophenomenology, A Methodological Remedy for the Hard Problem, *Journal of Consciousness Studies*, 3(4), 330–349.
- Winograd, Terry ja Flores, Fernando, 1986. *Understanding Computers and Cognition: A New Foundation for Design*. Ablex Publishing Corporation.

Loppuviitteet

- 1 Henkisellä tarkoitetaan alustavasti ja laajasti yleistäen aineellisen vastakohtaa.
- 2 Henkinen saa Lauri Rauhalan analyysissä myös täsmällisemmän merkityksen: Tajunnassa erottuu kaksi mielellistä toimintatasoa, korkeampitasoinen henkinen ja matalampi *psykkinen* (Lauri Rauhala 2009, 67–93). Näitä tajunnallisuuden eri asteita ei ole tässä esityksessä tarpeen erottaa.

- 3 Situaationaalisuus on suhteissa ja tilanteissa olemista. Situaatiosta tässä myöhemmin ja myös esim. Jussi Backman 2016.
- 4 Lauri Rauhala kuvaa analyysiaan myös holistisena ja *monopluralistisena*, toisiinsa palautumattomien olemuspuolien kokonaisuutena, ykseytenä moninaisuudessa.
- 5 Katso esim. Virpi Tökkäri 2015.
- 6 Lauri Rauhalan ihmiskäsitteestä hänen artikkelissaan *Kokeumuksen häiriöt ja sairauden käsite* (2009), kirjoissaan *Henkinen ihminen* (2009), *Ihminen kulttuurissa – kulttuuri ihmisessä* (2003), sekä tämän kirjan lähdeviitteissä mainituissa hänen teoksissaan. Rauhala on kirjoittanut situaationaalisuudesta vuodesta 1973 alkaen (1973a, 1973b).
- 7 Tajunnallisuudesta, esim. Lauri Rauhala 1995. Kykyä oivalta-vaan ja ymmärtävään tajunnallisuuteen eli kokemuksellisuuteen kutsutaan myös *elämyksellisyydeksi*. Fenomenologiassa käytetään nimitystä *noesis*, kokemuksisältöihin viitataan nimellä *noema*.
- 8 ”Oliola” tarkoitetaan tässä laajasti jotakin differentioitunutta ykseyttä, olentoa, eliötä, esinettä tai kappaletta, joka on olemassa. Olento ja eliö ovat eläviä olioita.
- 9 Mielen ilmeneminen tarkoittaa todellisuuden merkityksellisyysilmenemistä. Sillä ei siten ensisijaisesti viitata itseymmärrykseen, oman mielen olemassaolon ja sisäisen todellisuuden oivaltamiseen, vaan kaiken, niin ulkoisen kuin sisäisenkin todellisuuden mielellisyyteen. Tässä tehty erottely sisäinen-ulkoinen sivuuttaa keskeisen fenomenologisen kysymyksen ihmiselle ilmenevän todellisuuden luonteesta.
- 10 Konstituutio on eräs Edmund Husserlin fenomenologian peruskäsitteistä, ks. esim. Husserl 2011, 113–174. Juha Himanka kääntää konstituution ”todellistumiseksi” ja ”ilmentymiseksi”. Himanka jäsentää todellisuuden todellistumista mm. ”läsnäolon” käsitteen kautta, Himanka 2002.
- 11 Backman liittää konstituution läheisesti situaationaalisuuteen: Inhimillinen kokemus on ”aina tilanteeseen sidottua, kulttuurisesti ja historillisesti rakentunutta”. Edelleen ”’Looginen’ kokemus on useiden eri merkitysoisien yhteenasettuneisuutta ykseydeksi”. Hän selvittää myös konstituution etymologiaa, *constitutio, con* (lat.), ”yhdessä”, ”yhteen” ja *statuo* (lat.), ”asettaa paikoilleen”, Backman 2005.
- 12 Objektivistista todellisuutta, esimerkiksi esineitä, olioita ja asioita, ei fenomenologisen käsitteen mukaan ole olemassa eriytyneenä, rakenteisena ja jäsentyneenä todellisuutena ilman tajunnallista subjektia, jolle todellisuus ilmenee.
- 13 Martin Heideggerin mukaan ihminen on olemassaolossaan aina jo maailmassa, kietoutuneena ja sidottuna tilanteisiin ja ajallisuuteen. Maailmassaoleminen on Heideggerin fenomenologian peruskäsitteitä.
- 14 Situaationaalisuudesta fenomenologiassa ja tilanteen käsitteen historiasta, ks. esim. Backman 2016, 2009.
- 15 Kehollisuutta käsittelee Maurice Merleau-Pontyn perusteos havainnon fenomenologiasta, Merleau-Ponty 2008.
- 16 Hermeneuttisella kehällä ymmärretään vaiheittain etenevää kehän osatekijöiden vuorovaikutusta, jossa kehän komponentit vuoronperään tulkitaan uudelleen edellisten vaiheiden perusteella. Siten esimerkiksi tajunnassa tietynlaiseksi tulkitu havainto muokkaa aikaisempaa ymmärrystä todellisuudesta. Havainto asetuu tässä uudelleen tulkitussa todellisuudesta uuteen näkökulmaan, joka vuorostaan muuttaa käsitystä havainnon kohteesta.
- 17 Tietokoneella viitataan tässä esityksessä laajasti yleistäen mikroprosessorien perustuviin ja algoritmeja käyttäviin ”perinteisiin” digitaalisiin tietokoneisiin, niiden muodostamiin tietoverkkoihin, erilaisiin neuroverkkoarkkitehtuureihin ja kaikkien näiden yhdistelmiin. Eri laitteistoista ja niiden eroista, ks. esim. Pentti Haikonen 2017.
- 18 Ihmisille yhteinen *objektiivinen* todellisuus konstituoituu ja *verifioituu* Husserlin mukaan ihmisyksilöiden ja -yhteisöjen *intersubjektivisessa* kentässä. Intersubjektivisuus on osa ihmisen situaationaalisuutta, perustava ja alkuperäinen ihmisten keskinen vuorovaikutus, ks. Husserl 2011, 150–152, 167–174.
- 19 Timo Honkela tarkastelee kieltä, merkitystä ja ymmärrystä tekoälyn tutkimuksessa, mm. kielen konekäntämisen ongelmia, Honkela 2017.
- 20 Sähkökentät, sähkövirrat ja magneettikenttien vaikutukset aivoissa ovat tajunnassa tulkitettujen kokemuksisältöjen fysiologiaa, mitattavia *korrelaatioita*. Ne eivät sisällä merkityksiä. Nämä käsitteet ovat tietysti jo itsessään teoreettisia, ihmisen tietoisien ajattelun tulkintoja fyysikaalisena konstituoidun todellisuuden ja sen laitteistojen tuottamasta informaatiosta. Vrt. Eero Rauhala 2013, 2011.
- 21 Tässä kirjoituksessa ei ole mahdollista esittää yleisempää kat- sausta tietoisuuden tutkimuksista. Aihe on keskeinen fenomenologian perusteoksissa. Nyt käsillä olevaan tematiikkaan liit- tyen tiivistelmiä tutkimuksista esittävät esim. David Chalmers 1995 ja Francisco Varela 1996.
- 22 Chalmersin mukaan tietoisuuteen liittyvä ”vaikeita” ja ”helppo- ja” ongelmia. Vaikean ongelman ratkaisu edellyttäisi tietoisuuden palauttamista aivoihin fyysikaalisesti kuvattavaksi, tähän ei hänen mielestään ole ratkaisua. Helpot ongelmat ovat ratkais- tavissa olevia, esimerkiksi tekoälyn liittyviä kysymyksiä. Chal- mers 1995.
- 23 Myös esim. Haikonen katsoo, että tekoäly ei ole tietoinen, eikä tule ymmärtämään merkityksiä, Haikonen 2017. Honkela ajat- telee samoin, tekoäly ei tulevaisuudessakaan tule tietoiseksi ihmisen tapaan, Honkela 2017.
- 24 Varela 1996; Michel Bitbol 2012.
- 25 Bitbol 2017, 2008.
- 26 Yksityiskohtaisessa ja perusteellisemmässä fenomenologisessa tietokoneen ontologian analyysissä muutkin olemis- ja kuvaus- kategoriat saattavat osoittautua tarkoituksenmukaisiksi. Laa- jemmin tietokoneita ja kognitiota lähinnä Heideggerin feno- menologian perusteella ovat tarkastelleet Winograd ja Flores 1986. Ks. myös Himanka 2017 sekä Himanka ja Honkela 2017.
- 27 Tietokoneella ei tietenkään ole itsenäistä situaatiota ihmisestä riippumatta, sen ”konesituaatio” on ymmärrettävä lopulta vain osana ihmisen situaatiota.
- 28 Merleau-Pontyn mukaan se, että meillä ei ole mitään käsitystä varhaisimmasta lapsuudestamme, ei johdu unohtuksesta tai muistin toiminnan puutteesta, vaan siitä, ettei aluksi ollut vie- lä mitään ymmärrettyä, mitä nyt voisimme muistaa. Merleau- Ponty 2008, 404.
- 29 Chalmers kutsuu *zombiksi* tekoälyä omaavaa, hypoteettista tulevaisuuden oliota, joka ulkoisesti vastaisi ihmistä, käyttäytyisi ja toimisi täysin kuten ihminen, mutta jolla ei olisi sisäis- tä kokemusmaailmaa eikä tietoisuutta. Zombina toimivan koneen rakentaminen ei siis edellyttäisi tietoisuuden vaikean ongelman ratkaisua. *Turingin testillä*, tarkastelemalla koneen vastauksia sille esitettyihin kysymyksiin, voidaan yrittää tut- kia olisiko koneella (tai zombilla) ajattelua ja tietoisuutta.
- 30 Emme tiedä, mitä ”ihmistä älykkäämpi” supertekoaily voisi olla, emmekä sitä, voiko tajuton tekoäly yltää yleis- tai superteko- älyn tasolle. Tällä hetkellä esimerkiksi suuret monikansalliset teknologia-alan kaupalliset yritykset testaavat ensimmäisiä tekoälyjä, jotka autonomisesti kehittävät uutta tekoälyä.
- 31 Hyödynnämme tietokoneita ja tekoälyä päivittäin, emmekä enää tulisi toimeen ilman niitä. Joudumme kaikki myös jatkuvasti – usein tiedostamattamme – toimimaan konei- den ehdoilla ja olemaan niistä riippuvaisia. Älypuhelimemme ja muut tietovälineemme ovat jatkuvassa käytössä, haemme tietoa verkon tietokannoista, kommunikoimme sosiaalisessa mediassa ja pelaamme tietokonepelejä. Tietokonemaailma sisältää jo nyt virtuaalitodellisuuksia ja lisääjyä todellisuuksia, ihmiseen liitettyjä keinoaihteja ja -elimiä. Fenomenologiselta kannalta kaikki nämä tietokonetodellisuuden piirteet ovat sekä konstituoituja että konstituioivia ihmiselle ilmenevän todelli- suuden rakenteisia.
- 32 Tietokoneen ja moraalien kysymyksiä käsittelee mm. Honkela. Hän hahmottelee tekoälyä ihmiskunnan palveluksessa, rauhan- konetta, joka voisi tuoda toivoa rauhanomaisen maailman luomi- sessa, Honkela 2017.

Kirjoittaja on Helsingin yliopiston fysiikan dosentti ja eläkkeellä oleva yliopistonlehtori.