

Tekoälyn teologiasta

Tekoälyyn keskittyvää teologiaa on tehty kansainvälisestikin hyvin vähän. Tämä on yllättävää, koska sille olisi suuri tarve: sosiaaliseen vuorovaikutukseen ja tekoälyyn liittyvät teknologiset kehityskäsitteet ovat jo mullistaneet ja ovat mullistamassa myös uskonnollisten ihmisten arkea. Tarkastelen tässä artikkelissa viimeaikaista keskustelua tekoälyn teologiasta lähinnä englantia puhuvassa maailmassa. Aloitan lyhyellä katsauksella tekoälyn käsitteeseen. Tämän jälkeen tarjoan yleiskuvan niistä muutamista teksteistä, joissa tekoälyn teologiaa systemaattisesti käsitellään. Näiden pohjalta esitän, että kaikkein hedelmällisin tekoälyn ja teologian vuoropuhelu käydään ihmiskäsitykseen, ei niinkään etiikkaan liittyen. Esittelen sitten joitakin teologisia ja tekoälytutkimuksesta nousevia näkökulmia älykkyyden käsitteeseen. Lopuksi tarkastelen lyhyesti, millaisen roolin teknologia ja tekoäly voisivat saada teologisessa narratiivissa ihmisestä ja hänen tulevaisuudestaan.

TEKOÄLY JA SEN LÄHIALUEET

Sana ”tekoäly” on monilta osin harhaanjohtava.¹ Se on harhaanjohtava siksi, että sillä voidaan viitata

niin tekoälyn tutkimuksen moniin aloihin – kognitiotieteeseen ja tietojenkäsittelytieteeseen – kuin myös näiden alojen lopputuotteisiin, esimerkiksi älykkäisiin ohjelmiin ja järjestelmiin. Sana on myös siinä mielessä harhaanjohtava, että ”älyllä” voidaan tarkoittaa monta eri asiaa. Tavallisesti sillä viitataan kuitenkin ihmisenkaltaiseen älykkyyteen, jolloin tekoälytutkimuksen päämääränä nähdään sellaisten järjestelmien luominen, jotka kykenevät ihmisenkaltaiseen älykkyyteen. Tämä ei kuitenkaan ole ainoa vaihtoehto: voihan olla olemassa monia erilaisia älykkyyden muotoja, joista ihmisen älykkyys on vain yksi. Keskityn jatkossa kuitenkin tekoälytutkimuksen siihen osaan, jossa tavoitellaan ihmisenkaltaista älykkyyttä.

Ihmisenkaltaisen älykkyyden suhteen erotetaan toisistaan kapea tekoäly ja laaja tai yleinen tekoäly. Kapeat tekoälyjärjestelmät keskittyvät johonkin tarkasti rajattuun tehtävään, esimerkiksi lääketieteellisen diagnoosin tuottamiseen tai shakin pelaamiseen. Monilla alueilla tällaiset järjestelmät ovat jo ylittäneet ihmisen suorituskyvyn: jopa luovuutta ja kekseliäisyyttä vaativassa Go-pelissä Googlen AlphaGo-ohjelma on peitonnut ihmispelaajan.²

Myös erilaiset oppivat järjestelmät, joita on sovellettu esimerkiksi kielenkääntämiseen, suurien datamassojen käsittelyyn sekä hahmontunnistukseen, ovat kehittyneet merkittävästi.

Laajan tekoälyn suhteen kehitys ei kuitenkaan ole ollut lainkaan näin nopeaa. Dartmouth Collegessa Yhdysvalloissa vuonna 1954 järjestettyä työpajaa pidetään usein tekoälytutkimuksen virallisena lähtölaukauksena. Dartmouthissa tutkimuksen tavoitteeksi otettiin ”yleisälykkyys” eli järjestelmä, joka on lähellä ihmisen suoritus- ja monissa, ellei kaikissa, ihmiselämän kannalta tärkeissä älykkyyden muodoissa. Ajateltiin, ettei tämä tavoite ollut kuin muutaman tutkimusprojektin päässä toteutumisesta. Optimismi osoittautui kuitenkin virheelliseksi. Vuonna 2018 olemme vielä kaukana yleisälykkyydestä: emme edes ole varmoja siitä, pystyvätkö nykyiset tietokoneet toteuttamaan tällaista järjestelmää.

Teologian piirissä on keskusteltu laajasti esimerkiksi bio- ja geeniteknologian kysymyksistä. Rajaamalla nämä aiheet artikkelini ulkopuolelle, koska niitä on käsitelty ansiokkaasti muissa yhteyksissä. Myös yleisempiä kysymyksiä teknologiasta ja siihen liittyvistä teologisista ja eettisistä ongelmista on käsitelty jonkin verran. Tekoäly on eräs teknologinen kehitysaskel, joten näiltä osin se kuuluu teknologian teologian ja etiikan alaan. Esimerkiksi yhdysvaltalainen teologi Philip Hefner on tullut tunnetuksi siitä, kuinka hän yhdistää teologisen ihmiskäsityksen ja teknologian.³ Palaan tähän tarkemmin alla.

Joiltakin osin keskustelu ”kyborgiteologiasta” on päällekkäinen tekoälyä koskevan teologisen keskustelun kanssa. Kyborgit ovat scifikirjallisuudesta filosofiaan hiipineitä biologisten ihmisten ja keinotekoisien järjestelmien hybridejä. Joiltakin osin kyborgit ovat todellisuutta jo nyt: erilaiset keinoraajat sekä havaintojärjestelmiä, kuten näköä ja kuuloa, tukevat tai ne mahdollistavat tekniset laitteet helpottavat monen elämää.⁴ Kyborgiteologiaksi nimitän sitä teologian aluetta, jossa pyritään tarkastelemaan transhumanismin sekä ihmisen biologisen, lääketieteellisen ja teknologisen parantelun kysymyksiä.⁵

Niin tekoälyn teologian kuin kyborgiteologian keskeiset kysymykset koskevat ihmiskäsitystä. Mikä

ihminen on ja mikä on hänen biologiansa rooli ja arvo? On kuitenkin hyvä pitää tekoälyn teologia ainakin alustavasti erillään kyborgiteologiasta. Kyborgien yhteydessä keskustellaan pääsääntöisesti järjestelmistä, jotka eivät edes tavoittele ihmisistä älykkyyttä. Kuulon ja näön mahdollistavat järjestelmät tai keinoraajat eivät pyri mallintamaan ihmisälyä, vaan lähinnä tukemaan tai korvaamaan joitakin kognitiivisia ja motorisia toimintoja. Ihmisen kykyjen palauttaminen, muokkaaminen ja kehittäminen ovat toki relevantteja myös tekoälyn kannalta, mutta tekoälyyn liittyy myös monia erityiskysymyksiä, jotka eivät tyhjenny ihmisen kehittämistä koskeviin teemoihin.

On selvää, että tekoälyn ja sitä ympäröivään tutkimukseen liittyy suuri joukko erilaisia eettisiä, yhteiskunnallisia ja poliittisia kysymyksiä. Mitä meidän tulisi ajatella esimerkiksi sodankäynnin ja liikenteen automatisoitumisesta? Millaista lainsäädäntöä meidän tulisi kehittää robotiikan ja tekoälyn osalta? Millaisia yhteiskunnallisia ja poliittisia muutoksia esimerkiksi palvelualueiden automatisoituminen voi tuottaa? Filosofisen etiikan alueella näitä kysymyksiä kutsutaan toisinaan robottietiikaksi.⁶ Muutamat teologit ja uskonnolliset yhteisöt ovat ottaneet kantaa myös robottietiikkaan. Kun pyritään ratkaisemaan liikenteeseen (esimerkiksi itseohjautuvat autot), työhön (esimerkiksi työvoimantarpeiden muutokset) sekä lääketieteeseen (esimerkiksi diagnooseissa avustavat järjestelmät) liittyviä, tarkasti rajattuja eettisiä ongelmia, teologisesti motivoidut ratkaisut eivät juuri poikkea muiden eetikkojen tarjoamista ratkaisuista. Teolo-

1 Suomeksi tekoälyn johdattaa esim. Haikonen 2017.

Aloittelijoille sopiva johdanto on Boden 2016.

2 Bostrom 2016, 15–16.

3 Tekoälyä koskevia kysymyksiä sivuaa esimerkiksi Hefner 2002

4 Hieman vanhentuneen mutta silti hyvän yleiskatsauksen tarjoaa Clark 2003.

5 Kyborgiteologiaa käsittelee esimerkiksi Hefner 2004.

Laaja katsaus teknologisiin tulevaisuusvisioihin teologian näkökulmasta on Burdett 2015.

6 Kattava ja tuore katsaus on Lin & Jenkins & Abney 2017.

gien erityispiirteeksi osoittautuu juuri tietynlainen ihmiskäsitys, joka vaikuttaa eettisten periaatteiden taustalla.

TEKOÄLYN TEOLOGIAN YLEISLINJOJA

Tekoälyn teologia kärsii selvästi alkukankeudesta. Joitakin merkkejä alan heräämisestä on kuitenkin nähtävissä: viime vuosikymmeninä teologian ja kognitiotieteen rajapinnalla on tapahtunut yhä enemmän.⁷ Vaikka kognitiotietiedettä ei voikaan täysin samastaa tekoälytutkimukseen, ei tekoälyä voida kognitiotieteestä erottaakaan. Kognitiotiede on poikkitieteellinen ala, jossa tarkastellaan tiedonkäsittelyä sen eri muodoissa.⁸ Kognitiotieteiden ja sen alojen, kuten kognitiivisen psykologian, kognitiivisen kielitieteen, kognitiivisen antropologian ja kognitiivisen neurotieteen, historia on tekoälyn historiaa. Kognitiotiede ja tekoäly perustuvat samaan ydinoivallukseen: älykkyys on tiedonkäsittelyä. Älykäs olio kerää ympäristöstään tietoa, käsittelee sitä ja ohjaa käyttäytymistään tämän tiedon nojalla. Tämä pätee niin biologisiin kuin keinoitekoisiin olioihin.

Erään ensimmäisistä systemaattisista katsauksista teologiaan ja kognitiotieteeseen laati yhdysvaltalainen teologi Gregory Peterson 2000-luvun alkupuolella.⁹ Kirjassaan Peterson esittelee joitakin kognitiotieteen tuloksia ja arvioi keskustelua tietoisuuden luonteesta. Lisäksi hän tarkastelee joitakin erityisteemoja, lähinnä syntiä ja moraalia, joissa hän katsoo kognitiotieteiden tulosten hyödyttävän teologiaa. Peterson tekee myöhemmän tekoälyn teologian kannalta keskeisiä huomioita.

Ensinnäkin hän kiinnittää huomiota siihen, mitä tekoälytutkimus kertoo inhimillisestä älykkyyydestä. Olemme oppineet esimerkiksi sen, etteivät tietokoneessa toimivat ohjelmat kykene mallintamaan kovinkaan hyvin ihmisälyä sen sosiaalisessa ja moraalisisessa kontekstissa. Jotta ihmisäly saavutetaan, on järjestelmän tavalla tai toisella kyettävä toimimaan fyysisessä maailmassa ja osallistuttava inhimillisen vuorovaikutuksen monimutkaisiin muotoihin. Tämä puolestaan edellyttää jonkinlaista kehoa, kieltä ja kulttuuria. Petersonin mukaan nämä tulokset sopivat hyvin yhteen 1900-luvun teologisen ihmiskäsityksen kanssa, jossa on painotettu ihmisen

kokonaisvaltaisuutta. Abstrakti äly ei ole erillään tunteista, ympäristöstä ja muista ihmisistä.¹⁰

Toiseksi Peterson nostaa esiin kysymyksen siitä, mitä Jumalan kuvana oleminen tarkoittaa. Jumalan kuva on Petersonin mukaan liitetty liian usein ajatukseen ihmisen erityisyydestä suhteessa muihin eläimiin ja olioihin. Tekoälytutkimus osoittaa kuitenkin sen, kuinka inhimillinen älykkyys on samalla jatkumolla muiden älykkyiden muotojen kanssa. Vaikka yleisälykkyys onkin vielä tekoälyn saavuttamattomissa, keinoitekoisten järjestelmien kyvyt osoittavat, ettei inhimillinen älykkyys ole mitenkään ylikuonnollista. Näistä syistä Jumalan kuvana oleminen kannattaisi arvioida uudelleen: ehkä ihmiset jakavat joitakin Jumalan kuvan osaluaita muiden älykkäiden olioiden kanssa tai sitten Jumalan kuva voitaisiin käsittää prosessiksi, johon ihmiset muiden olioiden kanssa osallistuvat.¹¹

Nämä kaksi teemaa, ihmisälyn luonne sekä Jumalan kuvana oleminen, ovat keskeisiä kaikessa myöhäisemmässä tekoälyn teologiassa. Petersonin työ tekoälyn teologian parissa jää kuitenkin alustavaksi pohdinnaksi. Hän lähinnä esittelee (vuoden 2018 näkökulmasta vanhentunutta) tutkimusta ja kiinnittää huomiota joihinkin päällekkäisyyksiin teologian ja tekoälytutkimusten tulosten kanssa. Verrattuna esimerkiksi psykologi Justin Barrettiin, Peterson ei muodosta systemaattista lähestymistapaa eikä kykene tekemään kovin yksityiskohtaisia johtopäätöksiä.¹²

Kristilliset kirkot ovat lausuneet tekoälystä hyvin vähän. Roomalaiskatolisessa kirkossa haaste on kuitenkin tunnistettu. Vatikaanin paavillinen tiedeakatemia (Pontifical Academy of Sciences) järjesti 2016 tekoälyä käsitelleen työpajan, jonne oli kutsuttu tunnettuja filosofi ja muita alan tuntijoita, kuten esimerkiksi Patricia Churchland ja Stephen Hawking. Työpajan tarkoituksena oli saada käsitys tekoälytutkimuksen nykytilasta ja tulevaisuuden näkymistä.¹³ Teologian näkökulmasta työpajan tulokset tuottavat kuitenkin pettymyksen: äänessä ovat filosofit, matemaatikot sekä muut tekoälyn tutkijat mutta juuri minkäänlaista tekoälyyn liittyvää teologista työtä ei työpajassa tehty.

Työpajan lopputuloksena syntyi kuitenkin eräänlainen julistus, jonka taustalla on selvästi

teologiseen ihmiskäsitykseen liittyviä huolenaiheita. Muutaman sivun mittaisessa dokumentissa tehdään useita kiinnostavia väitteitä.¹⁴ Ensinnäkin siinä todetaan tekoölyn mahdollistavan tulevaisuudessa monien ympäristöön ja terveyteen liittyvien ongelmien ratkaisemisen. Tekoöly voi mullistaa niin tieteellisen tutkimuksen kuin koulutuksenkin. Sen avulla ihminen voi ylittää omat kognitiiviset rajoituksensa ja ehkä jopa parantaa moraaliaan. Se voi auttaa hallitsemaan monimutkaisia kokonaisuuksia, kuten maailmantaloutta ja ilmastonmuutosta. Tekoöly on siis väline, työkalu, jonka avulla ihmisten hyvinvointia voidaan parantaa.

Tekoölyn kehityksessä on myös omat uhkansa. Julistuksen kirjoittajat eivät ole huolissaan maailmanlopun skenaarioista, joissa tekoöly ottaa vallan tai muuten tuhoaa ihmiselämän mahdollisuudet. Sen sijaan kirjoittajien huoli koskee yhä etenevää yhteiskunnan automatisaatiota, jolla saattaa olla merkittäviä vaikutuksia esimerkiksi työllisyyteen, työn luonteeseen sekä ihmisten asemaan yhteiskunnassa. Näistä syistä on tärkeää, että tekoölyn kehitystyössä fokus pidetään ihmisessä ja hänen tarpeissaan. Kehitystyötä ei voi jättää pelkästään taloudellista voittoa tavoittelevien yhtiöiden käsiin, vaan yhteiskuntien ja valtioiden on osallistuttava siihen. Valtioiden pitää säädellä tätä teknologiaa sekä pitää huolta sen kehittämisen arvopohjasta. Muuten vaarana on epätasa-arvoisuus ja teknologian hyötyjen päätyminen vain eliitin käsiin.

Näissä huomioissa ei ole mitään erityisen teologista. Teologisesti orientoitunut lukija voi kuitenkin löytää niiden taustalta katoliselle ihmiskäsitykselle tyypillisiä painotuksia. Ensinnäkin julistus korostaa ihmiskeskeisen teknologian tärkeyttä. Ihmisen moraaliset päämäärät ovat etusijalla teknologian kehityksessä, ei esimerkiksi taloudellinen voitto tai ihmiselämän ”vapauttaminen” biologisesta ruumiista. Keskipisteessä on ihminen biologisena, psykologisena ja sosiaalisena kokonaisuutena.¹⁵ Toiseksi julistuksessa korostetaan sitä, kuinka tekoölyn myötä teknologia tunkeutuu yhä lähemmäs ihmisyyden ydinalueita, vaikkapa moraalialueita, itsetietoisuutta, vuorovaikutusta ja jopa rakkautta. Jotta teknologiaa voitaisiin kehittää oikein, tarvitaan

laajempaa moraalista, yhteiskunnallista ja ihmiskäsitystä koskevaa keskustelua.

Vuonna 2017 joukko yhdysvaltalaisia tutkijoita julkaisi kirkoille tarkoitetun raportin ”Artificial Intelligence in Christian Thought and Practice”.¹⁶ Raportti on oikeastaan lyhyt maallikoille suunnattu ohjekirjanen tekoölyn nykytilasta ja kristittyjen asenteesta siihen. Kirjasta motivoi usko siihen, että tekoälyteknologiaa voidaan käyttää ihmisten hyödyksi ja kristillisten päämäärien edistämiseen. Tekoölyn kehittäjien joukossa on kristittyjä, jotka saavat moraalisen opetuksensa pääsääntöisesti omista kirkoistaan. Tekijät kutsuvat kristityt mukaan kehittämään teknologiaa sekä korostavat kristillisten kirkkojen mahdollisuuksia vaikuttaa siihen, miten tekoälyteknologiaa käytetään erityisesti Latalaisessa Amerikassa ja muualla globaalissa etelässä. Kristittyjä tutkijoita kutsutaan tekemään työtään avoimesti omista lähtökohdistaan. Tämä on tärkeää esimerkiksi siksi, että tekoälyjärjestelmät oppivat omasta ympäristöstään. Tällöin niillä on tapana jatkaa ympäristössä havaittuja ennakkoluoja, moraalisia (tai moraalittomia) taipumuksia ja ajatusvinoumia pikemminkin kuin korjata niitä. On siis tärkeää, että järjestelmillä on ohjaajia ja ”esikuvia”, jotka ovat sitoutuneet moraalisten järjestelmien kehittämiseen.

Kirjanen nostaa esiin monia positiivisia tapoja käyttää tekoälyä. Tekoälyjärjestelmät ovat edistyneet esimerkiksi kielenkääntämisessä ja sen tulkinnessa. Tulevaisuudessa ne mahdollistavat yhä tehokkaammin ja luotettavammin tapahtuvan

7 Yleiskatsauksen tarjoaa Visala 2015.

8 Hyvä yleiskatsaus on Frankish & Ramsey 2012.

9 Peterson 2003.

10 Peterson 2003, 142–146.

11 Peterson 2003, 146–149.

12 Vrt. Barrett 2011.

13 Batro & Dehaene 2016.

14 Batro & Dehaene 2016, 114–115.

15 Ks. esim *Communion and Stewardship* -dokumentti (2004).

16 Matias & Manikonda & Hale & Arnold 2017.

vuorovaikutuksen yli kieli- ja kulttuurirajojen.¹⁷ Kirjanen näkee myös mahdolliseksi kehittää esimerkiksi henkilökohtaista ohjausta, terapiaa sekä avustustyötä tekoälyn avulla.

Positiivisista mahdollisuuksista huolimatta tekoälyyn liittyy myös monia haasteita ja uhkia. Tekoäly mahdollistaa jo nyt esimerkiksi jatkuvan valvonnan, valtion kontrollin lisääntymisen ja yhä etenevän yksityisyyden kapenemisen maissa, joissa kansalaisyhteiskunta on heikko. Tämä kehitys ei osoita hidastumisen merkkejä. Tekoäly tulee myös muuttamaan työmarkkinoita. Automaatio ei korvaa ihmistä, vaan tekoälystä tulee ihmisen työtoveri. Työn luonne voi siitä huolimatta muuttua. Tekoälyn laajamittainen yleistyminen voi myös syventää käynnissä olevaa eriarvoistumiskehitystä. Kirjassa kiinnitetään huomiota myös siihen, miten tekoälyjärjestelmiä kohdellaan. Vaikka järjestelmä ei olisikaan persoona eikä tietoinen olio, sen kohtelulla on merkitystä oman itsemme kannalta. Palaan tähän tärkeään huomioon myöhemmin.

Edellä tarkastelluista teksteistä ei käsittäkseni ole löydettävissä systemaattista tekoälyn teologiaa. Teologia kiinnittää huomiota samankaltaisiin eettisiin ongelmiin ja mahdollisuuksiin kuin teologiasta riippumaton filosofinen ja eettinen keskustelu. Teologien tarjoamat eettiset ratkaisut eivät edellytä erityisiä teologisia oletuksia. Ihmiskeskeistä teknologian kehittämistä voidaan vaatia myös ei-teologisilla perusteilla. Teologisia perusteluja ei tarvita kritisoimaan utooppisia visioita teknologisesta tulevaisuudesta, jossa ihmiskunnan ongelmat ratkaistaan ja kuolema voitetaan. Monista raamatunlainauksista huolimatta edellä esittelemäni kirjanen sisältää hyvin vähän teologista aineistoa, vaikka sen tarkoitus on käsitellä tekoälyä juuri teologian valossa. Peterson osuu kuitenkin mielestäni oikeaan: tekoälyn ja teologian dialogissa teologisen ihmiskäsityksen tulee olla keskeisessä roolissa. Keskitynkin jatkossa kahteen teemaan: mitä tekoäly ja teologia kertovat älykkyydestä sekä siihen läheisesti liittyvään kysymykseen Jumalan kuvasta.

ÄLYKKYYS ONGELMANRATKAISUNA

Yhdysvaltalainen tietojenkäsittelytieteilijä ja teologi Noreen Herzfeld on edistänyt merkittävästi

tekoälyn teologiaa. Herzfeldin varhainen tutkimus vuodelta 2002 sekä hänen viimeaikaiset kirjoituksensa muodostavat kokonaisuuden, jonka keskiössä on teologinen ihmiskäsitys.¹⁸ Herzfeldin mukaan tekoälytutkimuksessa voidaan havaita kolme tapaa nähdä ihmisen älykkyyttä: (1) älykkyyttä ongelmanratkaisuna, jota edustaa ”vanha kunnan tekoäly”; (2) älykkyyttä kykyä toimia joustavasti ja järkevästi fyysisessä ympäristössä, johon liittyvät erityisesti ideat laajennetusta kognitiosta ja hermoverkoista sekä (3) älykkyyttä kykyä olla suhteessa ihmisten ja muiden olioiden kanssa. Käsittelen seuraavaksi kutakin älykkyyden lajia tarkemmin.

Tekoälyn ja kognitiotieteen lähtölaukaus oli se oivallus, että niin ihmismielet kuin keinotekoisetkin mielet prosessoivat informaatiota.¹⁹ Alun perin ajateltiin, että kaikki informaatio voidaan esittää symbolisesti ja sen käsittelyä ohjaavat loogiset säännöt. Symbolisen tiedonkäsittelyn periaatteilla on pitkä historia länsimaisessa filosofiassa, matematiikassa ja logiikassa. Nykyisen vaiheen voidaan kuitenkin katsoa alkaneen 1930-luvulla, jolloin uudesta matematiikasta ja logiikasta inspiroituneet tutkijat, kuten brittiläinen matemaatikko Alan Turing, kehittivät ajatusta koneesta, joka voisi periaatteessa ratkaista kaikki symbolisesti muotoiltavissa olevat ongelmat. Turing esitti ensimmäisten joukossa ajatuksen siitä, että kaikki älykkyyttä, myös ihmisälykkyyttä, on informaation käsittelyä. Mielen tietokonemetaforan mukaan ihmismieli toimii periaatteessa samalla tavalla kuin tietokone: se muodostaa tietoesityksiä eli representaatioita ja käsittelee niitä tiettyjen sääntöjen mukaisesti. Mieli on siis ikään kuin ohjelma (software), joka toimii aivoissa (hardware). Tällä idealla oli valtava jälkivaikutus niin ajatteluun kuin sovelluksiin.

Symbolien sääntöpohjaiseen prosessointiin perustuvilla järjestelmillä saavutettiin merkittävää edistystä aloilla, joita oli helppo käsitellä laskennallisesti ja joita koskevat säännöt olivat suhteellisen selvät. Tällaisia ovat esimerkiksi sääntöpohjaiset pelit (esimerkiksi shakki), matemaattiset todistukset ja matematiikka yleensäkin. Eräänlainen huipukohta saavutettiin vuonna 1996, jolloin IBM:n shakkitietokone Deep Blue voitti ensi kertaa shakin maailmanmestarin (ja suurmestarin), venäläisen Garri Kasparovin.

1970-luvulle tultaessa tekoälyprojekti oli kuitenkin jo ongelmissa. Nopeasti kävi ilmi, että symbolisen prosessoinnilla on suuria vaikeuksia suoriutua joistakin ihmisille varsin yksinkertaisista tehtävistä. Esimerkiksi arkikielen käsittely osoittautui hankalaksi: ohjelman on erittäin vaikea tunnistaa arki-ilmaisujemme merkitystä, koska tämä merkitys riippuu tilanteesta ja jaetusta taustatiedosta. Tässä on kysymys yleisen tekoälyn ehkä suurimmasta haasteesta eli ”kehyksen ongelmasta”. Jotta jokin ongelma voitaisiin ratkaista, pitää ensin tietää, mikä osa kaikesta mahdollisesta tiedosta voisi olla relevanttia ratkaisun kannalta. Jos tällaista rajausta ei tehdä, edes tavattoman suuri laskentateho ei riitä. Ihmismieltä auttaa tässä rajattu sosiaalinen konteksti sekä joukko erilaisia heuristiikkoja ja muita oikoteitä. Näitä on kuitenkin erittäin vaikea mallintaa koneellisesti muualla kuin jossakin rajatussa ongelmassa (esimerkiksi shakkiottelu). Kehyksen ongelma on yksi syy sille, miksi yleisykykyys on tietokoneiden jatkuvasta laskentatehon kasvusta huolimatta jäänyt saavuttamatta.²⁰

ÄLYKKYYS TOIMINTANA

Symboliselle prosessoinnille ja sääntöpohjaiselle ongelmanratkaisulle oli kuitenkin syntynyt vaihtoehto kauan ennen 1980-lukua. Jo 1940-luvun lopulla oli kehitetty ajatus oppivista verkostoista, hermoverkoista. Symbolisessa prosessoinnissa on tavallisesti keskusyksikkö, joka muuntaa merkkijonoja sääntöjen mukaan. Hermoverkot puolestaan koostuvat suuresta määrästä hyvin yksinkertaisia yksiköitä ilman minkäänlaista keskusyksikköä tai prosessoria. Hermoverkon toiminta ei edellytä ennalta ohjelmoituja sääntöjä, vaan se tunnistaa malleja siinä informaatiossa, mitä verkolle syötetään. Näiltä osin hermoverkko poikkeaa varsin paljon tietokoneesta: se ei muistuta laskukonetta keskusyksiköineen ja muisteineen, vaan siinä informaatio on hajautettuna koko järjestelmän alueelle.

Symboliseen prosessointiin liittyvien ongelmien tähden hermoverkkomallit tulivat yhä suosituimmiksi. Tästä seurasi erityisesti 1980-luvulla kognitiotieteissä ja mielenfilosofiassa käyty kuuma väittely kognitivismiin (symbolinen prosessointi) ja konnektionismiin (hermoverkot) välillä.²¹ Hermo-

verkoilla on monia etuja suhteessa kognitivismiin. Näistä tärkein on se, että hermoverkot oppivat joustavasti tilanteissa, joissa informaatio on fragmentaarista eikä selkeitä sääntöjä ole. Verkostojen toinen keskeinen etu on se, että niiden toimintaperiaatteet muistuttavat ihmisaivoja.

Hermoverkkojen avulla saavutettiin merkittävää edistystä. Nykyisin kehitystä ajaa ajatus syväoppimisesta. Kun verkostoja järjestetään useammaksi kerrokseksi, saadaan aikaan järjestelmä, joka kykenee tunnistamaan erilaisia yleisiä piirteitä ja taipumuksia suuressa informaation kokonaisuudessa. Syväoppimisen avulla on mallinnettu esimerkiksi puheentunnistusta, kasvo- ja hahmotunnistusta sekä luotu erilaisia asiantuntijajärjestelmiä, jotka auttavat esimerkiksi uusien lääkkeiden suunnittelussa. Joissakin asiantuntijuutta ja suunnittelua vaativissa tehtävissä tällaiset asiantuntijajärjestelmät ovat jo saavuttaneet ihmistä vastaavan kompetenssin tai ylittäneet sen.

Hermoverkkojen nousu liittyy laajempaan kehitykseen, jossa älykyys nähdään logiikan ja sääntöjen sijaan kyknä olla vuorovaikutuksessa ympäristön kanssa. Konnektionismi sai vaikutteita 1970-luvulta alkaen esitetyistä kognitivismin kritiikeistä. Filosofit Hubert Dreyfus väitti, ettei symboliseen prosessointiin perustuva tekoäly voi koskaan yltää samaan suoritukseen kuin ihmisäly.²²

17 Myös tekoälytutkija Timo Honkela (2017) on kiinnittänyt huomiota tähän seikkaan. Rauhaa ja yhteisymmärrystä voidaan tulevaisuudessa edesauttaa merkittävästi konekäntämisen lisäksi myös ”semanttisten koneiden” avulla. Nämä koneet kykenevät käntämisen lisäksi selittämään, mitä ihmiset sanoillaan tarkoittavat.

18 Herzfeld 2002, 2009. Hyödyllinen yhteenvetoon Herzfeld 2010.

19 Jatkossa esittämäni huomiot tekoälyn kehityksestä nojaavat Herzfeldin lisäksi Stan Franklinin (1995) kuvaukseen.

20 Ongelmasta on kirjoitettu paljon. Ks. esim. Dennett 1990.

21 Ks. esim. Bechtel & Abrahamsen 1991.

22 Dreyfus 1992.

Tämä johtunee siitä, ettei ihmisen älykkyys perustu informaation prosessointiin, vaan älykkyys on luonteeltaan toiminnallista. Se liittyy läheisesti siihen, millainen kehomme on ja kuinka tuo keho on vuorovaikutuksessa muiden mielen ja ympäristön kanssa.²³

Dreyfusin ja muiden symbolisen tekoälyn kriitikoiden seurauksena innostuttiin siitä, kuinka tekoäly voitaisiin yhdistää oppiviin ja fysikaalisiin koneisiin. Esimerkiksi australialainen robootikko Rodney Brooks puolustaa näkemystä, jonka mukaan ihmisälykkyys edellyttää havaintojärjestelmien ja motoriikan välityksellä tapahtuvaa toiminnallista vuorovaikutusta fysikaalisen ympäristön kanssa. Brooksin mukaan älykkyyttä ei voida luoda edes oppivilla ohjelmilla, koska nämä ovat vieläkin vankeina tietokoneen sisällä. Älykkyyttä syntyy vain silloin, kun oppiva ohjelma asetetaan jonkinlaiseen kehoon, joka kykenee havaitsemaan ja manipuloidaan ympäristöään.²⁴

ÄLYKKYYS SOSIAALISENA KYKYNÄ

Siinä missä toiminnallisessa älykkyudessa on kyse vuorovaikutuksesta fysikaalisen ympäristön kanssa, monet ovat korostaneet älykkyuden sosiaalista luonnetta. Ihmisäly on kyllä jatkuvassa vuorovaikutuksessa fysikaalisen ympäristönsä kanssa, mutta näin ovat kaikki muutkin eläimet. Ihmisälyä sen sijaan luonnehtii kyky toimia nimenomaan sosiaalisessa ja kulttuurillisessa ympäristössä. Tämän näkemyksen mukaan älykkyys edellyttää sosiaalista kontekstia, joka puolestaan edellyttää yhteistä kieltä, jaettuja normeja ja jatkuvaa dialogia.

Älykkyys sosiaalisena kompetenssina ei ole uusi idea. Turing esitti jo vuonna 1950 ajatuksen ”imitaatiopelistä”, jota myöhemmin alettiin kutsua Turingin testiksi.²⁵ Testin tehtävänä on mitata tekoälyn älykkyyttä. Testi voidaan järjestää esimerkiksi seuraavalla tavalla. Koehenkilö laitetaan huoneeseen, jossa hänellä on pelkkä tietokone ja näppäimistö. Hän voi kirjoittaa koneella viestejä toiselle henkilölle, joka ei ole läsnä huoneessa. Koehenkilö voi keskustella mistä tahansa haluamasta aiheesta ja esittää kysymyksiä. Koehenkilö ei kuitenkaan tiedä, onko hänen keskustelukumppaninsa oikea henkilö vai ei. Jos kyseessä on tekoäly ja koehenkilö

saadaan uskomaan, että kyseessä on oikea henkilö, on tekoäly saavuttanut riittävän kyvyn olla sosiaalisessa vuorovaikutuksessa. Tällöin sitä voidaan pitää älykkäänä.

Suosiostaan huolimatta Turingin testin ongelmat ovat yleisesti tunnettuja. Ensinnäkin arkinen sosiaalinen vuorovaikutuksemme edellyttää kykyä tunnistaa kasvoista tunteita ja reagoida asiaankuuluvalla tavalla. Tällaiset kehollisen vuorovaikutuksen muodot puuttuvat tietenkin Turingin testistä. Toiseksi Turingin testi ei mitenkään mittaa tekoälyn mahdollista tietoisuutta. Jos tekoälyyn on ohjelmoitu tarpeeksi suuri määrä mahdollisia vastauksia koehenkilön kysymyksiin, voisi se periaatteessa läpäistä testin kykenemättä millään tavalla esimerkiksi arvioimaan omia vastauksiaan tai ymmärtämään niitä.²⁶

Ongelmistaan huolimatta Turingin testi viittaa useisiin ihmisälylle olennaisiin sosiaalisiin edellytyksiin. Yhdysvaltalainen tekoälytutkija Terry Winograd vaikutti Dreyfusin symbolisen tekoälyn kritiikistä ja on 1970-luvulta alkaen vienyt tutkimuksiaan uuteen suuntaan.²⁷ Aiemmin Winograd oli pyrkinyt luomaan tekoälyä, joka pystyy kommunikoimaan jonkinlaisella inhimillisellä kielellä. Dreyfusiin tutustuttuaan hän kuitenkin luopui projektista, koska luonnolliset kielet ja niiden mahdollistama vuorovaikutus edellyttävät kokonaan tietokoneista poikkeavaa älykkyyttä. Tietokoneilla on niille sopivampia vuorovaikutuksen keinoja, joista inhimillinen vuorovaikutus poikkeaa varsin paljon. Ihmisten välinen vuorovaikutus edellyttää kykyä jakaa tietty konteksti ja elämänmuoto: tunnistaa kielien käyttöyhteydet. Tämän lisäksi tarvitaan kykyä tunnistaa sosiaalisia rooleja, odotuksia ja normeja ja ohjata käyttäytymistä asiaankuuluvasti.

Samansuuntaista kehitystä tapahtui myös neurotieteessä ja kognitiivisessa neurotieteessä, joiden pohjalta alkoi piirtyä uudenlainen kuva älykkyudesta. Jopa abstrakti, looginen ajattelu näyttää olevan läheisessä yhteydessä biologisiin ja kehollisiin kognition alueisiin, erityisesti tunteisiin. Esimerkiksi neurotieteilijä Antonio Damasio on tutkimuksissaan osoittanut, kuinka erilaiset sosiaalisten ja moraalisien tunteiden prosessointiin liittyvät häiriöt aiheuttavat vaikeuksia abstraktin ajattelun alueella.²⁸

Nykyään kognitiotiede käyttää joustavasti niin symbolista prosessointia, hermoverkkoja kuin näiden yhdistelmiä. Erilaiset hybridijärjestelmät ovat osoittautuneet tehokkaiksi: hermoverkot soveltuvat hyvin ihmisen intuitiivisten ja automaattisten mielen toimintojen mallintamiseen, kun taas sääntöpohjaiset järjestelmät mallintavat paremmin tietoista ajattelua ja loogista päättelyä. Tutkimuksen kärkenä on nimenomaan sosiaalisen älykkyyden, tunteiden ja kielen mallintaminen. Huolimatta lastentatohon valtavasta kasvusta ja arkkitehtuurien kehityksestä ihmisilyn sosiaaliset muodot ovat vielä tällä hetkellä keinoilyn tavoittamattomissa.

IHMISÄLY JUMALAN KUVANA

Olen nyt Herzfeldiä seuraten erotellut kolme lähestymistapaa älykkyyteen ja samalla edennyt 1950-luvun vanhasta tekoilyn mallista nykypäivään. Ennen kuin tarkastelen tarkemmin tekoilyn ja teologian dialogia, on syytä esitellä lyhyesti joitakin yleisiä seikkoja teologisesta ihmiskäsityksestä.

Teologisella ihmiskäsityksellä on paljon sanottavaa ihmisälystä. Teologiseksi ihmiskäsitykseksi voidaan kutsua sitä kristillisen teologian alaa, joka tarkastelee ihmisyksilöitä, yhteisöjä ja näiden keskinäisiä suhteita. Ihmiskäsitykseen kuuluvat tällöin esimerkiksi kysymykset siitä, mitä tarkoittaa Jumalan kuvana oleminen, mitä on vapaa tahto ja millaista on ihmisarvoinen elämä. Ihmiskäsitys niveltyy osaksi myös monia keskeisiä oppeja, kuten esimerkiksi oppeja synnistä, armosta ja pyhityksestä. Teologisten perinteiden välillä on jonkin verran variaatiota ihmiskäsityksiin liittyvissä kysymyksissä. Tunnettuja ovat erimielisyydet esimerkiksi tahdonvapaudesta ja armo-opista. Myös eettisissä kysymyksissä kirkkojen välillä on jonkin verran eroja. Erimielisyyksistä huolimatta voidaan kuitenkin tunnistaa kolme yhdistävää piirrettä.²⁹

Ensinnäkin teologinen ihmiskäsitys asettuu osaksi suurempaa kertomusta Jumalasta ja hänen suhteestaan ihmiseen ja koko luomakuntaan. Kertomuksen vakavasti ottamisesta seuraa, että ihmisyksilöiden ja ihmisyhteisön olemassaololla on objektiivinen tarkoitus. Ihminen on luotu sitä varten, että hän eläisi yhteydessä Jumalaan. Teologinen ihmiskäsitys antaa siis ihmiselle objektiivisen

päämäärän. Toiseksi teologia tekee vahvoja oletuksia siitä, millaisia olioita ihmiset ja muut persoonat ovat. Perinteisen näkemyksen mukaan ihmisellä joko on aineeton sielu tai hän on tällainen sielu. Monet nykyaikaiset teologiset ihmiskäsitykset ovat kuitenkin hylänneet dualismin ja pitävät ihmistä biologisena ja psykologisena kokonaisuutena. Kolmanneksi teologinen ihmiskäsitys tekee oletuksia siitä, millaisia psykologisia kykyjä ihmisillä on. Nämä oletukset koskevat tavallisesti ihmisen älykkyyden, moraalisen päättelyn ja tahdon luonnetta.

Nämä kolme teemaa kohtaavat, kun teologia pyrkii kuvaamaan Jumalan kuvan luonnetta.³⁰ Perinteisen näkemyksen mukaan Jumalan kuvana oleminen tarkoittaa sitä, että ihmisellä on erityisiä kognitiivisia kykyjä, jotka tekevät hänestä tavalla tai toisella Jumalan kaltainen. Kaikkein tärkeimpinä näistä on pidetty kykyä järkevään ja abstraktiin ajatteluun sekä moraaliseen käyttäytymiseen. Perinteisesti monet teologit ovat ajatelleet, että ihminen on Jumalan kuva, koska ihmisellä on järjellinen sielu eli järkewä mieli ja kyky tunnistaa moraalisia periaatteita. Järjellinen sielu on luonteeltaan ei-luonnollinen ja ei-fysikaalinen: Jumala lahjoittaa sen jokaiselle ihmisyksilölle erikseen. Tätä näkemystä voimme kutsua strukturaaliseksi näkemykseksi siitä, mitä Jumalan kuvana oleminen on.³¹

Kuluneen sadan vuoden aikana monet teologit ovat hylänneet strukturaalisen näkemyksen. Raamatuntutkimuksen piiristä nousi 1900-luvun puolenvälin paikkeilla ajatus siitä, että Jumalan

23 Keskeinen kognitivismin kriitikko 1980-luvulla oli yhdysvaltalaisfilosofi John Searle. Ks. esim. Searle 1984.

24 Ks. esim. Brooks 1991.

25 Turing 1990.

26 Keskustelua Turingin testistä, ks. Copeland 1993, 33–57.

27 Winograd & Flores 1987.

28 Damasio 2005.

29 Näihin johdattaa lyhyesti ja ytimekkäästi Cortez 2010.

30 Yhteenvetoni tästä teemasta pohjautuu seuraaviin lähteisiin: Cortez 2010, 14–40 ja van Huyssteen 2006, 111–162.

31 Eräs strukturaalisen mallin nykypuolustaja on Levering 2017.

kuvana oleminen viittaa pikemminkin ihmisen erityiseen rooliin luomakunnassa. Genesiksen tekstissä ihminen luodaan eräänlaiseksi luomakunnan kuninkaaksi ja käskynhaltijaksi. Tällöin ihminen on Jumalan kuva siinä määrin kuin hän edustaa Jumalan hallintavaltaa maan päällä ja osallistuu Jumalan luovaan toimintaan luomakunnassa. Jumalan kuvana oleminen on siis eräänlainen kutsumus ja rooli, johon ihmiskunta on kutsuttu. Tätä kutsutaan usein funktionaaliseksi näkemykseksi.³²

Strukturaalisen ja funktionaalisen lisäksi on vielä kolmas teoria *imago Deistä*. 1900-luvun eurooppalaiselle systemaattiselle teologialle on Karl Barthin vanavedessä ollut tyypillistä nähdä Jumalan kuvan toteutuvan ihmisen suhteissa muihin olioihin. Juuri pyrkimys suhteeseen muiden kanssa tekee ihmisestä kolmiyhteisen Jumalan kuvan. Tämän relationaalisen mallin suosio voidaan liittää kristologisen ja trinitaarisen teologian uuteen tulemiseen erityisesti 1900-luvun puolivälin jälkeen.³³

Strukturaalisen, funktionaalisen ja relationaalisen mallin lisäksi on esitetty myös muunlaisia painotuksia. Jotkut teologit ovat huomauttaneet, että erityisesti varhaisen kirkon teologiassa kristologian yhteys teologiseen antropologiaan oli hyvin läheinen. Juuri Kristus on tosi ja aito Jumalan kuva: ihmisissä Jumalan kuva näkyy vain heikosti. Ihmisen kutsumus onkin tulla Kristuksen kaltaiseksi.³⁴ Toiset ovat puolestaan korostaneet sitä, kuinka ihminen on luomakunnan eläimistä ainoa, joka kykenee muokkaamaan ja kehittämään itseään. Tällöin Jumalan kuvana oleminen tarkoittaisi jonkinlaista vapautta tai kykyä itsensä ylittämiseen.³⁵

On syytä korostaa, etteivät nämä mallit ole välttämättä toisensa poissulkevia. Ne voidaankin nähdä teologisen ihmiskäsityksen ulottuvuuksiksi. Perinteisessä teologisessa antropologiassa strukturaalisella ja kristologisella mallilla on ollut keskeinen rooli. 1900-luvun teologiassa puolestaan funktionaalista ja relationaalista mallia on korostettu ja strukturaalista mallia on pyritty välttämään.

KESKUSTELUA ÄLYKKYYDESTÄ JA JUMALAN KUVASTA

Noreen Herzfeldin mukaan tekoälyn tutkimuksessa ilmenneillä älykkyyden muodoilla ja teologisella

keskustelulla Jumalan kuvasta on yhteys. Älykkyys ongelmanratkaisuna voidaan liittää strukturaaliseen malliin, älykkyys järkevänä toimintana puolestaan funktionaaliseen malliin ja älykkyys sosiaalisena kompetenssina relationaaliseen malliin. Herzfeld ei itse tunnustaudu yhden mallin ehdottomaksi kannattajaksi, mutta on kuitenkin selvää, että hän suosii nimenomaan relationaalista tulkintaa.

Olen Herzfeldin kanssa samaa mieltä siitä, että joiltakin osin tekoälytutkimus tukee relationaalista ja sosiaalista käsitystä älykkyydestä. Nykyisen kognitiotieteen mukaan älykkyys on läheisesti kietoutunut aivojen, kehon ja ympäristön vuorovaikutukseen. Tekoälytutkijat eivät enää ajattele, että jokin tietokoneessa toimiva, ennalta ohjelmoituihin sääntöihin perustuva ohjelma kykenisi mallintamaan ihmisälyä. Älykkyys syntyy pikemminkin siitä, että olio on jatkuvassa vuorovaikutuksessa fyysikaalisen ympäristönsä kanssa kehon välityksellä. Lisäksi on selvää, että ihmisäly ja sen kehitys ovat riippuvaisia juuri tietynlaisesta sosiaalisesta ympäristöstä, toisin sanoen toisista ihmisistä, sosiaalisista normeista ja rooleista sekä kielestä. Tätä kokonaisuutta kutsumme tavallisesti kulttuuriksi. Ihmistä voidaan siis pitää kulttuurisena eläimenä. Yksilö syntyy osaksi ihmisyyhteisöä ja yhteisön biokulttuurista elämänmuotoa, joka on syntynyt pitkän evoluution seurauksena.³⁶

Herzfeld on oikeassa myös siinä, että strukturaalisella mallilla on ollut paha tapa yliarvioida ihmisen älykkyys. Sen sijaan, että pitäisimme ihmisälyä jonkinlaisena jumalallisena ja yliluonnollisena osana ihmismieltä, näyttää uskottavammalta pitää sitä aivojen, kehon ja ympäristön tuotteena. Tämä selittää myös sen, miksi niin usein epäonnistumme totuuden saavuttamisessa. Ihmismieli ei muutenkaan näytä niin erityiseltä kuin teologit ovat perinteisesti ajatelleet. Abstrakti ajattelu, moraalinen symbolinen kieli ja kulttuuri eivät edellytä Jumalan välittömästä toiminnasta syntyneitä sielua. Kullakin näistä toiminnoista on juurensa jo ei-inhimillisessä eläinkunnassa sekä biologisissa aivoissamme ja kehossamme. Myös lisääntynyt tietomme muiden eläinten älykkyydestä viittaa samaan suuntaan.

Mielestäni Herzfeld kuitenkin tekee näistä siinänsä tosista väitteistä liian pitkälle meneviä johto-

päätöksiä. Tässä yhteydessä tyydyn esittämään vain muutaman huomion väitettäni tukemaan.³⁷

Ensinnäkin tekoälytutkimus ei sulje pois ajatusta tietoisesta ja loogisesta älykkyydestä. Vanha kunnon tekoäly onnistuu mallintamaan ihmisen abstraktia ja tietoista ajattelua, jota luonnehtivat säännöt ja looginen päättely. On toki tunnustettava, että tällainen prosessointi on varsin vaivalloista ja hidasta verrattuna mielen automaattiseen toimintaan. Valtaosa mielen suorittamista tehtävistä, esimerkiksi sensomotorisesta toiminnasta, koordinaatiosta ja jopa käsitteellisestä ajattelusta, tapahtuu automaattisesti tietoisuudesta riippumatta. Symbolista prosessointia paremmin tätä automaattista toimintaa mallintavat juuri esimerkiksi hermoverkot. Tästä ei kuitenkaan seuraa se, etteikö tietoisella ja loogisella ajattelulla olisi mitään roolia. Tämä rooli on vain rajallisempi kuin aikaisemmin kuvittelimme. Ihmisälyn toimintaa voidaan pitää eräänlaisena yhteispelinä, jossa tietoinen ja looginen älykkyyden tavalla tai toisella valvoo ja ohjaa automaattisen mielen toimintaa.³⁸

Toiseksi ihmisen älykkyyden näyttää nykyisen kognitiotieteen ja tekoälytutkimuksenkin valossa erityiseltä. Ihmismieli ei ole vain hieman viritelty versio orangin, gorillan tai apinan mielestä. Joustava oppiminen ja kulttuuri ovat toimineet ihmisen selviytymisstrategioina, mikä puolestaan on synnyttänyt suuren joukon ihmiselle erityisiä kognitiivisia mekanismeja. Kyvyillemme symboliseen ajatteluun, kielen käyttöön ja tekojemme ohjaamiseen kulttuurin normien perusteella ei löydy vertaa muussa luomakunnassa.³⁹

ÄLYN TULEVAISUUS

Teologinen ihmiskäsitys on väline, jonka avulla tekoäly voidaan liittää osaksi kristillisen teologian narratiivia ihmisen päämääristä ja tarkoituksesta. Tällöin emme katso pelkästään nykyistä teknologiaa vaan myös tulevaisuutta. Tekoälytutkimuksessa tulevaisuuden ennakoinnilla onkin keskeinen rooli: tulisiko meidän esimerkiksi valmistautua mahdolliseen superälyn syntymiseen ja olisiko tämä paha vai hyvä asia?⁴⁰

Filosofinen keskustelu pyörii ”singulariteetin” käsitteen ympärillä. Käsitteelle ei ole yleisesti hy-

väksyttyä määritelmää. Eräs voisi olla tällainen: singulariteetti on hetki, jolloin superäly ilmaantuu. Myöskään superälykkyydelle ei ole yhtä ainoata määritelmää. Esimerkiksi filosofi Nick Bostrom ehdottaa, että viittaamme superälyllä tekoälyyn, joka ylittää ihmisen kyvyt monilla keskeisillä alueilla.⁴¹ Tällainen superäly voi kehittää uudenlaista teknologiaa, joka synnyttää toistaiseksi käsityskykymme ylittäviä mahdollisuuksia. Tämä ennustamattomuus ei ole estänyt tutkijoita väittelemästä kiivaasti siitä, mitä superälyn ilmaantuminen voisi tuoda tullessaan.

Kognitiotieteilijä Margaret Boden jakaa superälyä koskevat näkemykset kahteen ryhmään.⁴² Niitä, jotka uskovat singulariteetin tuloon, hän kutsuu singulariteettiuskovaisiksi erotukseksi singulariteettiskeptikoista. Skeptisismiä voidaan perustella monin eri tavoin. On mahdollista, ettei ihmisälykkyyttä vastaavaa keinoälyä pystytä periaatteessakaan luomaan. Vaikka ajattelisimmekin, että ihmisäly vastaava keinoäly tai jonkinlainen superäly olisivat periaatteessa mahdollisia, ei tästä vielä seuraa se, että se voitaisiin käytännössä toteuttaa. Superäly saattaa edellyttää niin paljon resursseja, ettei minkään organisaation tai maan kannata ottaa tällaista projektia harteilleen.

Singulariteettiuskovaiset puolestaan ajattelevat, että superäly on periaatteessa mahdollinen, jopa vääjäämätön seuraus teknologisesta kehityksestä. Singulariteettioptimistit, kuten keksijä Ray Kurzweil, näkevät singulariteetin jälkeisen tulevaisuuden

32 Ks. Middleton 2005.

33 Ks. esim. Pannenberg 1985.

34 Ks. esim. Crisp 2015.

35 Näin Hans Urs von Balthasaria seuraten esimerkiksi Tanner 2011.

36 Ks. esim. Barrett & Jarvinen 2015.

37 Perinpohjaisempia argumentteja löytyy Visala 2018a, 2018b.

38 Varsin suosittu ”kaksoisprosessiteoria” (esim. Kahneman 2011) tukee tätä ajatusta.

39 Ihmismielen erityisyyttä puolustavat esimerkiksi Baumeister 2005 ja Gazzaniga 2008.

40 Ks. esim. Bostrom 2016.

41 Bostrom 2016, 26.

42 Boden 2016, 147–153.

ruusuisena.⁴³ Superälykkyyden avulla voimme ratkaista ihmiskuntaa vaivaavat, jättimäiset ongelmat. Sodat loppuvat, ilmastokriisi saadaan ratkaistua ja pääsemme eroon ihmisluonnon varjopuolista. Parhaimmassa tapauksessa kehitymme biologisen olemassaolomme tuolle puolen: voimme ladata mieleemme tietokoneille ja elää ikuisesti. Kurzweilin mukaan tämä ei ole ainoastaan mahdollista, vaan myös moraalisesti välttämätöntä. Ihmiskunnan yhteisen hyvän tähden meidän on kaikin voimin pyrittävä voittamaan biologiseen ihmisyyteen sisältyvä kuolema, sairaus ja tuska. Vaikka kaikki transhumanistit eivät jaakaan Kurzweilin radikaalia visiota, he kuitenkin suhtautuvat myötämieelisesti ihmisen biologiseen ja kyberneettiseen kehittämiseen: ehkä lääkkeiden, geeniteknologian ja nanoteknologian avulla voisimme parantaa ihmisen ikää, suorituskykyä ja vastustuskykyä.⁴⁴

Myös singulariteettipessimistit ajattelevat, että teknologinen kehitys kohti tekoälyä on vääjäämätöntä. He eivät kuitenkaan jaa Kurzweilin ja kumppaneiden optimismia. Uutiskynnyksen ovat ylittäneet esimerkiksi fyysikko Stephen Hawkingin ja yrittäjä Elon Muskin varoitusten sanat. Kumppanin mukaan superälykkyys voi ollakin ihmiskunnan suurin uhka: tällainen älykkyys saattaa vahingossa tuhota ihmiselämän edellytykset. Vielä karmeampi olisi tulevaisuus, jossa superälykkyys tahallaan joko tuhoaa ihmisen tai orjuuttaa hänet.⁴⁵

Kognitiotieteilijä Pentti Haikonen maalaa synkeää kuvaa ”älyteknologian” todellisuudesta.⁴⁶ Lapsipuolen asemaan jää esimerkiksi yksityisyytemme. Yhä etenevä tiedonkeruun kaupallistuminen tekee tottumuksiamme, halujamme ja suunnitelmiamme koskevasta tiedosta kauppatavaraa, jonka avulla meitä ohjataan ja meille suunnataan mainontaa. Pahimmassa tapauksessa tällainen tieto ja sen käyttö voi olla uhka demokratialle: ihmiset voidaan esimerkiksi jakaa yhä tarkemmin hyödyllisten ja hyödyttömiin luokkiin ja heidän elämäänsä voidaan yhä tarkemmin ohjailla.

TEKOÄLYN TEOLOGIAN SUUNTAVIIVOJA

Tällaisiin huoliin on kiinnitetty huomiota myös tekoälyä koskevassa teologisessa keskustelussa. Sekä edellä esiteltyjä paavillisen tiedeakatemian

työpajaa että tutkijoiden opaskirjasta sävyttää huoli siitä, että inhimillisesti arvokkaat, moraaliset päämäärät unohtuvat tai häviävät tekoälyn kehityksen sivutuotteena. Kummassakin vaaditaan ihmisen ja hänen päämääriensä pitämistä tekoälyn kehityksen keskiössä. Tätä vaatimusta voidaan puolustaa myös edeltävän Jumalan kuvaa koskevan keskustelun perusteella.

Esimerkiksi Philip Hefnerin teknologian teologian lähtökohtana on se, että ihmiset ovat Jumalan kumppaneita luomistyössä. Ihminen on Jumalan kuva siinä määrin kuin hän toteuttaa Jumalan hänelle antamaa tehtävää. Ihmiselle on annettu erityistehtävä osallistua luomakunnan hyvän ja oman itsensä varjeluun.⁴⁷ Eräs tämän tehtävän ulottuvuus on, että ihminen on luonnostaan työkalujen käyttäjä ja luomakunnan muokkaaja. Muokkaus ja työkalut ovat siis osa ihmisenä olemisen mandaattia. Tällöin teknologian kehittäminen viljelemiseksi ja varjelemiseksi on oikeastaan velvollisuus, ei pelkästään vaihtoehto. Tästä lähtökohdasta käsin voidaan omaksua optimistinen asenne myös nykyteknologioihin. Yhdysvaltalainen teologi Ted Peters on soveltanut juuri tällaista ajattelutapaa teologisessa bioetiikassa. Meidän tulee pyrkiä kehittämään teknologiaa, joka mahdollistaa empatian ja lähimmäisenrakkautta lisääntymisen sekä erilaisten sairauksien ja vaivojen eliminoimisen.⁴⁸

Vastaavaa päättelyä voidaan soveltaa tekoälyyn. Jos tekoälyä kehittämällä on mahdollista ratkaista ihmiskunnan ongelmia ja auttaa sitä suojelemaan ympäristöä sekä kehittymään rauhanomaiseen suuntaan, tulee meidän tällöin tavoitella sitä. Tekoäly olisi siis ihmiskunnan käytössä oleva työkalu viljelemisen ja varjelemisen tehtävässä. Se ei kuitenkaan saa korvata ihmistä eikä riistää häneltä vastuuta. Siitä voi kuitenkin tulla jonkinlainen kumppani ja hyödyllinen avustaja.

Jos tekoälyn teologia ottaa lähtökohdakseen nämä väitteet, ei sillä ole muuta vaihtoehtoa kuin kritisoida joitakin tekoälyyn liittyviä odotuksia. Ensinnäkin tekoälyn teologia suhtautuu kriittisesti niihin transhumanismin muotoihin, joissa tekoälyn kehityksen katsotaan tuovan mukanaan teknologisen paratiisin ja vapautuksen biologian kahleista. Teknologista visiota voidaan pitää lähinnä kristil-

lisen eskatologian sekulaarina korvikkeena. Tekoälystä tulee sekulaarissa eskatologiassa se pelastaja ja lunastaja, jonka rooli kristillisessä teologiassa on varattu Jumalalle.⁴⁹

Tekoälyn teologia suhtautuu kriittisesti kaikkiin yrityksiin, joissa ihmisälyn kyvyt redusoidaan nykyisten tekoälyjen kykyihin. Kuten monet tekoälytutkijat myöntävät, nykyiset tai tällä hetkellä näköpiirissä olevat tekoälyt eivät kykene yleisälykkyteen. Ne eivät myöskään pysty moniin ihmiselämälle keskeisiin tehtäviin kuten merkitysten tunnistamiseen, viisauteen, ymmärtämiseen tai syvään sosiaaliseen vuorovaikutukseen. Nämä toisten kohteluun sekä hengelliseen ja moraaliseen elämään liittyvät mielen kyvyt ovat kuitenkin teologisen ihmiskäsityksen mukaan kaikkein tärkeimpiä ihmisälykkyyden muotoja. Jos ihmisäly on kykyä olla suhteessa toisten kanssa, eivät supertietokoneet juuri auta ihmisiä tässä tehtävässä – puhumattakaan siitä, että ne pystyisivät itse samaan.

Lisäksi tekoälyn teologian kehittelijät huomauttavat, että tekoälytutkimuksen ideaaliälykkyys vaikuttaa olevan hyvin kaukana siitä, mitä teologia pitää ihmisen ideaalina. Kristusta, jota teologit pitävät täydellisenä Jumalan kuvana, tuskin luonnehtii superälykkyys, vaikka jotkut teologit ovat pitäneetkin häntä kaikkietävänänä. Sen sijaan Kristusta luonnehtivat rakkaus ja myötätunto ihmisiä kohtaan sekä kuuliaisuus Jumalaa kohtaan. Tämä on pikemminkin superviisautta kuin superälykkyyttä. Tämä ideaali voisi olla eräs niistä seikoista, jonka teologit voisivat tuoda yhteiskunnalliseen keskusteluun tekoälystä ja etiikasta.

Edellä esitetyt teologiset huomiot Jumalan kuvasta ja ihmisen älykkyydestä sisältävät myös syvän yhteyden ihmisen ja muun luomakunnan välillä. Vaikka ihmistä on pidetty erityisenä muiden lajien keskuudessa, nykyteologit ovat pääsääntöisesti korostaneet ihmisen biologista luonnetta ja yhteyttä muuhun biologiseen luontoon.⁵⁰ Lisäksi funktionaalinen käsitys Jumalan kuvasta korostaa ihmisen roolia muun luomakunnan paimenena ja huolenpittäjänä: ihminen on saanut Jumalalta tehtäväksi pitää luonnosta ja muista eläimistä huolta. Näin ollen hän ei voi jättää luontoa taakseen ja muuttaa itse asumaan bittiavaruuteen. Tästä syystä tekoälyn teologia

suhtautuu kriittisesti yrityksiin irrottaa ihminen hänen biologiastaan ja hylätä viljelyn ja varjelen tehtävä.

TEKOÄLY JA IHMISEN PÄÄMÄÄRÄT

Edellä olen hahmotellut tekoälyn teologiaa, jossa Jumalan ihmisille antamat päämäärät ovat keskeisessä roolissa. Haluan lopuksi nostaa esiin vielä yhden keskeisen tekoälyn teologian haasteen, joka koskee nimenomaan näitä päämääriä itseään. Ranskalainen filosofi ja sosiologi Jacques Ellul väitti jo 1960-luvulla, että teknologiseen kehitykseen perustuva yhteiskunta tekee lopulta ihmisestä itsestään pelkän välineen.⁵¹ Teknologia ei siis pelkää mahdollista joidenkin inhimillisten päämäärien saavuttamista, vaan se vaikuttaa siihen, mitä pidetään tavoittelemisen arvoisena ja millaisia ihmiset haluavat olla. Tällöin unelma ihmiskeskeisestä teknologiasta, tässä tapauksessa tekoälystä, voi olla hyvin vaikea tai jopa mahdoton, saavuttaa.⁵² Tarkastelen seuraavaksi joitakin uhkakuvia, joita aiemmin hahmotellun tekoälyn teologian on kohdattava.

Eräs tapa ymmärtää tekoälyn kautta tapahtuvaa päämäärien muutosta on tarkastella niitä vaikutuksia, joita lisääntyvällä tekoälyjen ja ihmisten vuorovaikutuksella voi olla. Se, miten ihmiset suhtautuvat tekoälyyn, voi vaikuttaa siihen, millaisen moraalisen statuksen he antavat toisilleen. Voimme ajatella asiaa analogian kautta. Jotkut filosofit ovat väittäneet, että sillä, miten kohtelemme muita eläimiä, on vaikutus moraalipsykologiaamme. Vaikka muilla eläimillä ei olisikaan tietoisuutta, niitä on tämän argumentin mukaan syytä kohdella hyvin, koska julma kohtelu tukee taipumustamme kohdella toisia ihmisiä jul-

43 Kurzweil 2005.

44 Clark & Savulescu 2016 esittelee viimeaikaista keskustelua ihmisen kehittämisestä.

45 Superälyn syntymiseen liittyviä ongelmia tunnistaa myös Bostrom 2016, 140–154.

46 Haikonen 2017, 269–283.

47 Hefner 1993.

masti. Jos argumentti toimii, se voidaan laajentaa koskemaan myös tekoälyjärjestelmiä. Voiko käydä niin, että yhä enemmän sosiaalisesti ja ”ihmisenkaltaisesti” käyttäytyvän tekoälyjärjestelmän kohtelu vaikuttaa moraaliin taipumuksiimme negatiivisesti?

Vastaus riippuu siitä, millä tavalla tosiasiasa kohtelemme näitä järjestelmiä. Ongelmia voi syntyä, jos emme kohtele monin osin ihmisen kaltaisia keino-olioita inhimillisesti. Elokuvasa Blade Runner 2049 (2017) käsitellään juuri välineellistymisen tematiikkaa. Elokuvasa eletään ympäristötuhon jälkeistä aikaa, jossa ihmistä muistuttavat keino-oliot aina biologisista replikanteista erilaisiin hologrammeihin asti ovat yleisiä. Replikanteista on tullut eräänlainen orja- ja työväenluokka, jolla ei ole ihmisen statusta. Elokuvan ihmishahmot ovat tunnekyliä ja välinpitämättömiä. He suhtautuvat toisiin ihmisiin ja replikantteihin joko työkaluina, joiden avulla saadaan jotakin hyötyä, tai seksiojekteina. Koska he ovat tottuneet kohtelemaan monessa mielessä kaltaisiaan olentoja esineinä, he alkavat kohdella myös toisiaan esineinä, vain välineellisesti arvokkaina.

Toisaalta ongelmia voi syntyä siitäkin, jos ulotamme inhimillisen kohtelun koskemaan olioita, jotka ovat kaukana ihmisen moraalisisista ja sosiaalisista kyvyistä. Jo nyt jotkut ihmiset ulottavat inhimillisen ja intiimin kohtelun erilaisiin keino-oloihin. Esimerkiksi jotkut ihmiset ovat halunneet mennä naimisiin robottiensa kanssa. Monet robotit suunnitellaan tietoisesti niin, että ne herättävät vahvoja moraalisia ja sosiaalisia reaktioita ihmisissä, vaikka näillä olioilla ei olekaan kykyä syvään vuorovaikutukseen, rakkauteen ja autonomiaan.⁵³ Tällöin olemme vaarassa trivialisoida inhimillisen kohtelun ja sen edellytykset. Moraalisesti latautuneet, monimutkaiset ihmissuhteet voidaan ymmärtää yhä kapeammin.

Tästä esimerkkinä yhä nopeammin kehittyvät, seksuaalisten tarpeiden tyydyttämiseen suunnitellut nuket ja robotit. Vaikka näiden kyvyt ovat vielä alkeellisia, tulevaisuudessa on odotettavissa yhä kehittyneempiä versioita. Teologisen ihmiskäsityksen mukaan seksuaalisuus on osa laajempaa moraalisten arvojen, velvollisuuksien ja aikomusten

verkostoa. Seksuaalinen halu kuuluu rakkauden ja sitoutumisen yhteyteen: vain syvä yhteys toiseen persoonaan voi tyydyttää sen. Seksirobotti ei kykene tällaista halua tyydyttämään, koska minikäänlaista persoonienvälistä yhteyttä ja sen luomia velvollisuuksia ei pääse syntymään. Seksirobotti on vain itsetyydytyksen apuväline, ei persoona. Tämä puolestaan edesauttaa kehitystä, jossa seksuaalisuus nähdään irrallaan rakkaudesta, velvollisuuksista ja sitoumuksista. Samankaltaista kehitystä voi tapahtua myös muiden, moraalisisesti latautuneiden suhteiden, kuten hoivan ja huolenpidon suhteen.

Esimerkiksi empatiaa ja huolenpitoa simuloivia robotteja, kuten Paro-nimisiä, hylkeenpoikasia muistuttavia robotteja käytetään nykyisin vanhusten konekumppaneina. Tutkimuksissa erilaisten hoivarobottien on havaittu edistävän hyvinvointia ja niiden käyttöä voidaan joiltakin osin puolustaa.⁵⁴ Ongelma on kuitenkin itse kehityssuunta. Jos ja kun tällaiset keino-oliot yleistyvät, ne saattavat trivialisoida hoivaan liittyvät monimutkaiset vastuun ja välittämisen kysymykset. Saatamme huomaamattamme tunnustaa, etteivät hoiva, välittäminen ja empatia olekaan mitään muuta kuin asiaankuuluvien tunteiden tunnistamista ja jonkin välittävältä kuulostavan huokauksen tai sanan toistamista. Tämä voi puolestaan edesauttaa kehitystä, jossa hoiva kaikessa sen rikkaudessa unohtetaan: sen sijaan, että tarjoaisimme vanhuksillemme empatiaa, huolenpitoa ja välittämistä, käytämme yhteiskuntamme resurssit johonkin muuhun ja tarjoamme vain hoivan simulaatiota.

Uhkakuva voidaan esittää myös toisella tavalla. Kuvitellaan, että arkiympäristömme on tulevaisuudessa erilaisten tekoälyjen täyttämä. Kun soitan pankkiin, tekoäly vastaa. Tekoäly hoitelee aikatauluni sekä vapauttaa minut lomakkeiden kirjoittelusta ja veroilmoitusten tekemisestä. Kuvitellaan myös, että joissakin rajatuissa tilanteissa en ole varma, onko vastapuolella kehittynyt tekoäly vai ihminen. Koska kuvitelmamme pyrkii realismiin, ajatellaan myös, että nämä järjestelmät eivät ole yleisälykkäitä ja niiltä puuttuu kyky inhimillisen vuorovaikutuksen kirjoon ja moraaliseen vastuuseen. Mielestäni ei ole vaikea kuvitella, mitä tällaisessa maailmassa elämisestä voi seurata. Ihmiset voivat tottua

kohtelevaan yhä ”inhimillisempiä” järjestelmiä välipitämättömästi, koska niillä ei kuitenkaan ole ihmiskollegojensa vastuuta ja valtaa. Ihmiset voivat tulla myös yhä epäluuloisemmiksi myötätunnon osoituksia kohtaan, koska he herkistyvät niille seikoille, jotka paljastavat epäautenttisen, simuloitun myötätunnon.

Tällaisella kehityksellä olisi tekoälyn teologian kannalta ongelmallisia seurauksia. Moraalipsykologinen ongelma voidaan avata seuraavalla tavalla. Kyky toimia moraalisesti korkealla tasolla edellyttää pääsyä tilanteiden moraalisesti relevantteihin tosiseikkoihin. Näitä ovat esimerkiksi tieto siitä, millaiset asiat aiheuttavat kärsimystä toisille, millaista apua ihmiset tarvitsevat, miten henkilö tunnistaa omat negatiiviset tunteensa ja kykenee myötätuntoon. Kyky tunnistaa tällaisia tosiseikkoja ei ole automaattinen vaan edellyttää harjaantumista.⁵⁵ Kykyjen harjaantuminen riippuu kuitenkin varsin paljon kontekstista. On esimerkiksi paljon helpompia halveksua ja pilkata toista henkilöä sosiaalisessa mediassa kuin kahvipöydässä. Sosiaalinen media on muuttanut vuorovaikutuksen kontekstia ja vaikuttaa sitä kautta psykologiaamme. Tekoälyn yhä lisääntyvä läsnäolo arkielämässä voi tehdä samoin vielä suuremmassa mittakaavassa. Mitä vähemmän olemme tekemisissä muiden ihmisten kanssa, sitä vaikeammaksi harjaantuminen esimerkiksi moraalisten tosiseikkojen tunnistamisessa tulee. Tekoälyn yleistyminen voi siis muokata arkista ympäristöämme suuntaan, jossa tekoälyn teologian edellyttäminen päämäärien on yhä vaikeampi toteutua.

Huoli on siis lopulta se, että tekoälyn ja keinoelämän rajoitukset muuttavat ihmiskäsitystämme ja arvojamme. Tekoäly ei ole vielä onnistunut kovin hyvin juuri teologisen ihmiskäsityksen kannalta keskeisten kykyjen mallintamisessa. Kun tekoälyteknologia kuitenkin valtaa alaa, voi tällä olla katastrofaaliset vaikutukset: päämäärämme saattavat ohentua ja trivialisoitua. Kun tämä uhka otetaan huomioon, tulee tekoälyn teologian pitää yhä lujemmin kiinni ihmiskäsityksestään ja sitä kannattelevasta teologisesta kehiksestä. Tekoälyn ja keino-olioiden tulee olla ihmisen työkaluja, jotka palvelevat ihmisen päämääriä. Ne eivät saa muokata näitä päämääriä joksikin muuksi. Tekoälyn teolo-

gia voi osallistua laajempaan yhteiskunnalliseen keskusteluun tekoälystä. Se voi toimia välineellistymiskehityksen vastavoimana korostamalla ihmisen sosiaalisen ja moraalisen ulottuvuuden tärkeyttä ja sitä, kuinka nämä kietoutuvat elämänmuotoomme.⁵⁶

Aku Visala, uskonnonfilosofian dosentti,
akatemiaturkija, Helsingin yliopisto
aku.visala@helsinki.fi

KIRJALLISUUS

Barrett, Justin

2011 *Cognitive Science, Religion and Theology: From Human Minds to Divine Minds*. West Conshohocken: Templeton Press.

Barrett, Justin & Jarvinen, Matthew

2015 ”Cognitive Evolution, Human Uniqueness, and the *Imago Dei*”. *The Emergence of Personhood: A Quantum Leap?* Ed. M. Jeeves. Grand Rapids, MI: Eerdmans.

Battro, Antonio & Dehaene, Stanislas (ed.)

2016 *Power and Limits of Artificial Intelligence*. Pontificia Academiae Scientiarum Scripta Varia 132. http://www.pas.va/content/dam/accademia/pdf/scripta_varia_132.pdf (10.10.2018).

Baumeister, Roy

2005 *The Cultural Animal: Human Nature, Meaning, and Social Life*. Oxford: Oxford University Press.

48 Peters 2010.

49 Ks. Burdett 2015.

50 Hyvä esimerkki tästä on roomalaiskatolisen kirkon teologinen asiakirja *Communion and Stewardship* (2004), jossa hahmotellaan kokonaisvaltaista näkemystä teologiasta ihmiskäsityksestä.

51 Kts. esim. Ellul 1964.

52 Suomalaisessa filosofisessa keskustelussa tällaisen ”tieteellis-teknisen rationaalisuuden” kritiikkiä on harjoittanut esimerkiksi von Wright 1987.

53 Tätä kehitystä voidaan kritisoida eettisestä näkökulmasta; ks. esim. Bertolini 2017.

54 Ks. Meacham & Studley 2017.

55 Joitakin argumentteja tällaisen moraalisen harjaantumisen välttämättömyyden puolesta esittää Zoller 2017.

56 Haluan kiittää Suomen Akatemiaa tätä artikkelia varten tehdyn tutkimuksen rahoittamisesta ja nimettömiä arviointisijoita kommentteista ja parannusehdotuksista.

- Bechtel, William & Abrahamsen, Adele
1991 *Connectionism and the Mind: An Introduction to Parallel Processing in Networks*. Oxford: Basil Blackwell.
- Bertolini, Andrea
2017 ”Huijauis ihmisen ja robotin vuorovaikutuksessa”. Suom. Esa Erävalo. *Robotiikka, geeniteknikka ja etiikka*. Toim. Ajatushautomo Kompassi. <https://kompassiblog.files.wordpress.com/2018/02/robotiikka-geeniteknikka-etiikka-kirja.pdf> (10.10.2018).
- Boden, Margaret
2016 *AI: Its Nature and Future*. Oxford: Oxford University Press.
- Bostrom, Nick
2016 *Superintelligence: Paths, Dangers, Strategies*. New York: Oxford University Press.
- Brooks, Rodney
1991 ”Intelligence without Representation”. *Artificial Intelligence* 47, 139–159.
- Burdett, Michael
2015. *Eschatology and the Technological Future*. New York: Routledge.
- Clark, Andy
2003 *Natural-Born Cyborgs: Mind, Technologies and the Future of Human Intelligence*. New York: Oxford University Press.
- Clarke, Steven & Savulescu, Julian et al. (ed.)
2016 *The Ethics of Human Enhancement: Understanding the Debate*. New York: Oxford University Press.
- Communion and Stewardship
2004 *Communion and Stewardship: Human Persons Created in the Image of God*. http://www.vatican.va/roman_curia/congregations/cfaith/cti_documents/rc_con_cfaith_doc_20040723_communion-stewardship_en.html (10.10.2018).
- Copeland, Jack
1993 *Artificial Intelligence: A Philosophical Introduction*. Oxford: Basil Blackwell.
- Cortez, Marc
2010 *Theological Anthropology: A Guide to the Perplexed*. London: T & T Clark.
- Crisp, Oliver
2015 ”A Christological Model of the *Imago Dei*”. *Ashgate Research Companion to Theological Anthropology*. Ed. Charles Taliaferro & Joshua Farris. New York: Routledge, 217–229.
- Damasio, Antonio
2005 *Descartes’ Error: Emotion, Reason and the Human Brain*. London: Vintage.
- Dennett, Daniel
1990 (1984) ”Cognitive Wheels: The Frame Problem of AI”. *The Philosophy of Artificial Intelligence*. Ed. Margaret Boden. Oxford: Oxford University Press, 147–170.
- Dreyfus, Hubert
1992 *What Computers Still Can’t Do: A Critique of Artificial Reason* (Revised Edition). Cambridge, MA: The MIT Press.
- Ellul, Jacques
1964 *The Technological Society*. Toronto: Vintage Books.
- Frankish, Keith & Ramsey, William (ed.)
2012 *The Cambridge Handbook of Cognitive Science*. Cambridge: Cambridge University Press.
- Franklin, Stan
1995 *Artificial Minds*. Cambridge, MA: The MIT Press.
- Gazzaniga, Michael
2008 *Human: The Science Behind What Makes Your Brain Unique*. New York: Harper Collins.
- Haikonen, Pentti
2017 *Tietoisuus, tekoäly ja robotit*. Helsinki: Art House.
- Hefner, Philip
1993 *Human Factor: Evolution, Culture, and Religion*. Minneapolis, MN: Fortress.
- 2002 ”Technology and the Human Becoming”. *Zygon* 37:3, 655–666.
- 2004 ”The Created Co-Creator Meets Cyborg”. <http://www.metanexus.net/essay/created-co-creator-meets-cyborg> (10.10.2018).
- Herzfeld, Noreen
2002 *In Our Image: Artificial Intelligence and the Human Spirit*. Minneapolis, MN: Fortress.
- 2009 *Technology and Religion: Remaining Human in a Co-Created World*. West Conshohocken: Templeton Press.
- 2010 ”Human and Artificial Intelligence: A Theological Response”. *Human Identity at the Intersection of Science, Technology and Religion*. Ed. Nancey Murphy & Christopher Knight. Farnham: Ashgate, 117–130.
- Honkela, Timo
2017 *Rauhankone: Tekoälytutkijan testamentti*. Helsinki: Gaudeamus.
- Kahneman, Daniel
2011 *Thinking, Fast and Slow*. London: Allen Lane.

- Kurzweil, Ray
2005 *The Singularity Is Near*. New York: Viking.
- Levering, Matthew
2017 *Engaging the Doctrine of Creation: Cosmos, Creatures, and the Wise and Good Creator*. Grand Rapids, MI: Baker.
- Lin, Patrick & Abney, Keith & Jenkins, Ryan (ed.)
2017 *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. New York: Oxford University Press.
- Matias, Nathan & Manikonda, Lydia & Hale, Scott & Arnold, Ken
2017 *Artificial Intelligence in Christian Thought and Practice*. <https://medium.com/ai-and-christianity/artificial-intelligence-in-christian-thought-and-practice-20ec8635a94f>. (10.10.2018)
- Meacham, Darian & Studley, Matthew
2017 "Could a Robot Care? It's All in the Movement". *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. Ed. Patrick Lin & Keith Abney & Ryan Jenkins. New York: Oxford University Press, 98–111.
- Middleton, Richard
2005 *The Liberating Image: The Imago Dei in Genesis 1*. Grand Rapids, MI: Brazos Press.
- Pannenberg, Wolfhart
1985 *Anthropology in Theological Perspective*. Edinburgh: T&T Clark.
- Peters, Ted
2010 "Can We Enhance the *Imago Dei*?" *Human Identity at the Intersection of Science, Technology and Religion*. Ed. Nancey Murphy & Christopher Knight. Farnham: Ashgate, 215–238
- Peterson, Gregory
2003 *Minding God: Theology and the Cognitive Sciences*. Minneapolis, MN: Fortress.
- Rosenberg, Stan (ed.)
2018 *Finding Ourselves after Darwin: Implications for the Christian Doctrines of Original Sin, the Image of God and the Problem of Evil*. Grand Rapids, MI: Baker Academic.
- Searle, John
1984 *Minds, Brains and Science*. Cambridge, MA: Harvard University Press.
- Taliaferro, Charles & Farris, Joshua (ed.)
2015 *Ashgate Research Companion to Theological Anthropology*. New York: Routledge.
- Tanner, Kathryn
2011 "Creation and Salvation in the Image of an Incomprehensible God". *God of Salvation: Soteriology in Theological Perspective*. Ed. Ivor Davidson & Murray Rae. London: Routledge, 61–76.
- Turing, Alan
1990 (1950) "Computing Machinery and Intelligence". *The Philosophy of Artificial Intelligence*. Ed. Margaret Boden. Oxford: Oxford University Press, 40–66.
- van Huyssteen, Wentzel
2006 *Alone in the World: Human Uniqueness in Science and Theology*. Grand Rapids, MI: Eerdmans.
- Visala, Aku
2015 "Theological Anthropology and the Cognitive Sciences". *Ashgate Research Companion to Theological Anthropology*. Ed. Charles Taliaferro & Joshua Farris. New York: Routledge, 57–71.
- 2018a "Will the Structural Theory of the Image of God Survive Evolution?" *Finding Ourselves after Darwin: Implications for the Christian Doctrines of Original Sin, the Image of God and the Problem of Evil*. Ed. Stan Rosenberg. Grand Rapids, MI: Baker Academic, 64–78.
- 2018b "Human Cognition and the Image of God". *The Christian Doctrine of Humanity*. Ed. Oliver Crisp & Fred Sanders. Grand Rapids, MI: Zondervan, 91–109.
- von Wright, Georg Henrik
1987 *Tiede ja ihmisjärki*. Keuruu: Otava.
- Winograd, Terry & Flores, Fernando
1987 *Understanding Computers and Cognition: A New Foundation for Design*. New York: Addison-Wesley Professional.
- Zoller, David
2017 "Skilled Perception, Authenticity, and the Case against Automation". *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. Ed. Patrick Lin & Keith Abney & Ryan Jenkins. New York: Oxford University Press, 80–92.