

Manta Räsänen

VOIKO KONE AJATELLA?
Tekoälyn varhaishistoriaa Descartesin, Lovelacen ja Turingin
silmin

Yhteiskuntatieteiden tiedekunta
Kandidaatintutkielma
Huhtikuu 2024

TIIVISTELMÄ

Manta Räsänen: Voiko kone ajatella?: Tekoälyn varhaishistoriaa Descartesin, Lovelacen ja Turingin silmin
Kandidaatintutkielma
Tampereen yliopisto
Historian tutkinto-ohjelma
Huhtikuu 2024

Tämän tutkimuksen tarkoituksena oli selvittää, mitä kolme tekoälyn kehityksen kannalta merkittävää tieteentekijää on ajatellut vahvasta tekoälystä, eli koneiden kyvystä ajatella, sekä miten ja miksi käsitykset ovat muuttuneet eri aikakausina, ja miten ne suhtautuvat nykikäsitteisiin vahvasta tekoälystä. Aineistona käytettiin René Descartesin, Ada Lovelacen ja Alan Turingin tieteellisiä tekstejä julkaisuista *Metodin esitys*, *Notes by the Translator* ja *Computing Machinery and Intelligence*. Tutkimus oli vertailevaa tutkimusta ja sen toteutustapana oli aineistolähtöinen laadullinen sisällönanalyysi. Aikaisempaa tutkimusta vastaavista aiheista on melko vähän, varsinkin historian tieteenalalla, ja etenkin vertaileva tutkimus eri aikakausien välillä on uusi keskustelunavaus.

Tarkastelussa olleet tieteentekijät ovat kolmelta eri aikakaudelta: René Descartes 1600-luvulta, Ada Lovelace 1800-luvulta ja Alan Turing 1900-luvulta. Tämän lisäksi tutkielmassa taustoitetaan lyhyesti tekoälyn syntyä ilmiönä antiikissa. Viidentenä ajallisena kontekstina toimii nykyaika, jota vasten tekoälyn kehityshistoriaa peilataan. Tieteentekijöiden ajattelun lisäksi jokaisen käsitellyn aikakauden tieteellisen kehityksen tilaa ja materiaalisia mahdollisuuksia kuvaillaan ja arvioidaan suhteessa ajattelijoiden näkemyksiin.

Tutkimuksen pohjalta voidaan todeta, että tarkastellut tieteentekijät eri aikoina ovat olleet yksimielisiä siitä, että vahvaa tekoälyä ei heidän aikanaan ole onnistuttu kehittämään. Myös nykyajassa tieteentekijät uskovat, että vahvaa tekoälyä ei ole onnistuttu kehittämään. Descartes ja Lovelace uskoivat myös, että vahvaa tekoälyä ei ole mahdollista saavuttaa. Turing sen sijaan uskoi, että vahva tekoäly on tulevaisuudessa mahdollista saavuttaa. Eri aikakausien väliset ajatteluerot kumpuavat osittain tieteen kehitysvaiheiden eroista eri aikakausina, mutta myös sillä, miten ajattelijat ovat suhtautuneet ihmisyyteen, älykkyyteen ja ajatteluun, on merkitystä. Tutkimuksen pohjalta voidaan myös tunnustaa tekoälyn kehityshistoria pitkäksi ja vaiherikkaaksi prosessiksi, joka on alkanut antiikista ja jatkunut läpi vuosisatojen vaihtelevalla aktiivisuudella tähän päivään saakka.

Avainsanat: Ada Lovelace, Alan Turing, Charles Babbage, jäljittelypele, René Descartes, tekoäly

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin Originality Check -ohjelmalla.

Sisällys

1 JOHDANTO.....	1
1.1 TUTKIMUKSEN TEKEMINEN	2
1.2 KÄSITTEET	4
1.3 ANTIIKKI – MISTÄ KAIKKI ALKOI?	5
2 RENÉ DESCARTES JA 1600-LUKU	7
2.1 DESCARTES JA TEKOÄLY	8
3 ADA LOVELACE & CHARLES BABBAGE JA 1800-LUKU	11
4 ALAN TURING JA 1900-LUKU	14
4.1 TURING JA TEKOÄLY.....	15
5 VERTAILUA	18
6 JOHTOPÄÄTÖKSET.....	20
7 LÄHTEET	22
8 TUTKIMUSKIRJALLISUUS	22

1 Johdanto

Tällä hetkellä kaikki puhuvat tekoälystä. Suuret kansainväliset teknologiayritykset, yliopistot, valtiot ja pääomasijoittajat ympäri maailmaa investoivat miljardeja tekoälyn kehittämistyöhön. Tekoälyä käytetään jo esimerkiksi asiakaspalveluroboteissa, terveydenhuollon diagnoosien tekemisessä, liikenteenohjauksessa, turvallisuusjärjestelmissä, tietotyön apuvälineenä ja palveluiden personoimisessa kuluttajille. Kehittyneiden kielimallien, kuten ChatGPT:n, avulla tekoäly esittäytyy lähes kaikille ihmisryhmille, ja tavallisetkin kansalaiset pääsevät testaamaan uusinta tekoälytekniikkaa.

Kaikilla ilmiöillä on historiansa. Ajatus tekoälystä ei ole syntynyt tyhjästä. Se on vaatinut vuosisatojen ja jopa vuosituhansien pituisen historian, jonka aikana tekoälyä on kehitetty eri muodoissaan. Aluksi ajatus tekoälystä on ollut puhdas ajatusleikki. Jo antiikissa oli tarinoita ja myyttejä, joissa esiintyy ihmismäisiä koneita. Mitä pidemmälle matematiikan ja mekaniikan osaaminen eteni, sitä todellisemmaksi ajatus ajattelevista koneista muuttui. 1800-luvulla syntyi ensimmäisiä yrityksiä tuottaa varhainen tietokone. Tuolloin *analyttiseksi koneeksi* kutsutun koneen esimerkkinä toimi edistynyt kudontakone.¹ 1900-luvun puolivälissä ensimmäiset tietokoneet ja 1900-luvun lopussa ensimmäiset tekoälyn sovellukset syntyivät. 2000-luvulla tekoäly on jo osa arkeamme, halusimme tai emme. Ajatus koneiden mahdollisuudesta ihmismäiseen ajatteluun on vaivannut ihmisiä tuhansia vuosia. Nykyisen kehityksen myötä kysymys siitä, voiko kone ajatella, on noussut pinnalle ajankohtaisempana kuin koskaan aiemmin.

Tämä kandidaatintutkielma keskittyy tekoälyn historiaan, ja etenkin siihen, mitä tieteenekijät eri aikakausina ovat vastanneet kysymykseen: voiko kone ajatella? Aihe on ajankohtainen ja erittäin tärkeä. Tutkimukseni keskittyy tieteenhistoriaan², jonka tarkastelu on hyödyllistä useista eri syistä. Tieteenhistorian tutkiminen auttaa ymmärtämään, kuinka eri tieteenalat ovat kehittyneet ajan myötä. Tämä auttaa hahmottamaan nykytieteiden taustalla olevia periaatteita ja ajattelutapoja. Tieteenhistoria sisältää myös monia esimerkkejä siitä, miten tieteellinen ymmärrys on kehittynyt

¹ Kyseessä oli Joseph-Marie Jacquardin kehittämä Jacquard-niminen kuviokudontaan erikoistunut kone, jota ”ohjelmoitiin” reikäkorttien avulla. Monet saattavat hieman liioitellustikin kutsua tätä kudontakonetta jopa maailman ensimmäiseksi tietokoneeksi. (Davis 2003, s.178–179.)

² Suomen kielessä tieteenhistoria niputetaan usein *aate-* tai *oppihistorian* (käännös tulee ruotsin kielen *idé- och lärdomshistoria*sta) käsitteiden alle, ja etenkin oppihistoriaa pidetään tieteenhistorian synonyymina. Tieteenhistoriasta puhutaan välillä myös tiedon historiana, mutta näiden kahden käsitteen väliltä löytyy myös eroavaisuuksia. Puhun tässä tutkielmassa tieteenhistoriasta, mutta myös oppihistoria-käsitteen käyttö olisi aivan yhtä perusteltua.

virheiden ja väärinkäsitysten kautta. Virheistä oppiminen voi auttaa nykytutkijoita välttämään samoja sudenkuoppia ja edistämään tieteen laatua. Tieteenhistorian tuntemus voi myös tarjota ideoita ja inspiraatiota nykyajan tutkijoille. Sen ymmärtäminen, miten menneet tutkijat ovat ratkaisseet ongelmia ja haasteita, rohkaisee uusia ideoita ja innovaatioita.³ Kaiken kaikkiaan tieteenhistorian tutkiminen ei ole vain menneisyyden ymmärtämistä vaan myös avain nykytieteen ja tulevaisuuden tutkimuksen parantamiseen. Näin ajankohtaisen ilmiön historian linjojen tuntemus auttaa myös suhtautumaan pinnalla olevaan ilmiöön osana historiallista kehitystä. Tieteellisen diskurssin historian tunteminen antaa näkökulmia tekoälykeskustelun syvälliselle ymmärtämiselle.

1.1 Tutkimuksen tekeminen

Tutkimuskysymyksenäni toimii, mitä kolme tekoälyn kehityksen kannalta merkittävää tieteentekijää on ajatellut vahvasta tekoälystä, eli koneiden kyvystä ajatella, sekä miten ja miksi käsitykset ovat muuttuneet eri aikakausina, ja miten ne suhtautuvat nykykäsityksiin vahvasta tekoälystä. Ihmismäisesti ajattelevien koneiden mahdollisuutta on pohdittu paljon eri näkökulmista filosofiassa niin mielenfilosofian, tieteenfilosofian, etiikan kuin metafysiikankin puolella. Myös monet tekoälyä tai sen ensiasteita kehittämässä mukana olleet matemaatikot, loogikot ja tietojenkäsittelytieteilijät ovat muodostaneet aiheesta omia filosofisia näkemyksiään.

Tässä tutkielmassa otan lähempään tarkasteluun René Descartesin, Ada Lovelacen ja Alan Turingin, ja heidän ajatuksensa vahvasta tekoälystä. Olen valinnut nämä kolme tieteentekijää siitä syystä, että he edustavat eri aikakausia, erilaisia materiaalisia todellisuuksia, joissa he ovat muotoilleet ajatuksensa, sekä hieman toisistaan eroavia näkemyksiä. Näiden kolmen tieteentekijän näkemykset antavat siis lukijalle perspektiiviä usealta eri aikakaudelta ja näkökulmasta. Alkuperäislähteitä on kolmen tieteellisen artikkelin/julkaisun verran: Descartesin *Metodin esitys* vuodelta 1637, Lovelacen *Notes by the Translator* vuodelta 1843 sekä Turingin *Computing Machinery and Intelligence* vuodelta 1950. Tutkimus on vertailevaa tutkimusta ja menetelmänäni toimii aineistolähtöinen laadullinen sisältöanalyysi. Eri aikakausien tieteen tilaa ja alkuperäislähteistä nousevia tieteentekijöiden ajatuksia vertaillaan suhteessa ajalliseen kontekstiinsa, toisiinsa sekä nykyajan tieteen tilaan ja ajatteluun.

Keskityn siis kolmen alkuperäislähteen pohjalta kolmen eri tieteentekijän ajatuksiin eri aikakausilta, ja kerron tieteen kehityksen tilasta jokaisen tieteentekijän aikakauden osalta. Lovelacen luvussa pohdin myös Charles Babbagen yritystä tuottaa ensimmäinen mekaaninen tietokone, sillä Babbagen

³ Kts. Morus 2017 s.12–17.

työ toimi innoituksena Lovelacelle, ja hänen ajatuksilleen koneista, sekä niiden mahdollisuuksista ja rajoituksista. Tämän lisäksi johdatan lukijan lyhyesti tekoälyn historian alkuun ja pohdin ilmiön syntyä antiikissa tutkimuskirjallisuuden pohjalta. Aikajana kulkee siis tekoälyn syntysijoilta antiikista 1600-luvun kautta 1800-luvulle ja sieltä lopulta 1900-luvun puoliväliin. Viidentenä ajallisena kontekstina toimii nykyaika, jota vasten peilaan tekoälyn kehityshistoriaa.

Aihettani tarkastellessa on hyvä erottaa kaksi erillistä, mutta toisiinsa kietoutunutta historiallista prosessia: tekoälyn kehityshistoria ja tieteellisen diskurssin historia vahvan tekoälyn mahdollisuudesta eli siitä, voiko kone ajatella vai ei. Vaikka keskityn pääasiassa alkuperäislähteiden kautta tieteellisen diskurssin historiaan, ei materiaalista ja teknologista tekoälyn kehityksen prosessia voi ohittaa. Tieteen tila ja materiaallinen todellisuus kuitenkin vaikuttavat väkisininkin siihen, miten eri aikakausilla on ajateltu. On kuitenkin hyvä tiedostaa, että kyseessä ei ole kaikenkattava perustutkimus tekoälyn kehityksen historiasta, enkä keskity kaikkiin tieteen kannalta olennaisiin yksityiskohtiin, joita tekoälyn kehitys on vaatinut historiansa aikana. Perustelen rajaukseni ja näkökulmani puhtaasti sillä, että niin laaja-alaiseen tutkimukseen minulla ei olisi mahdollisuutta tämän kandidaatintutkielman puitteissa. Tämä tutkielma ohjaa pääasiassa katseensa filosofisen diskurssin historiaan ja kertoo, mitä kolme tekoälyn kehityksen kannalta tärkeää tieteentekijää eri aikakausilta on vastannut kysymykseen: ”Voiko kone ajatella?”

Tekoälyn kehityksen historiasta ei oikeastaan ole aikaisempaa puhdasta perustutkimusta historian tieteenalalla. Tekoälyn historiaa on kuitenkin tarkasteltu jonkin verran muiden tieteenalojen puolella. Tietotekniikan, matematiikan, tietojenkäsittelytieteen, neurotieteen ja filosofian puolella on käsitelty kunkin tieteenalan historiaa jonkin verran, ja näiden kirjoitusten kautta piirtyy se tekoälyn historiakäsitys, jonka nykyään omaamme. Poikkeuksena tietenkin tieteentekijöiden elämäkerralliset julkaisut, joita on tuotettu myös historian tieteenalan puolella. Esimerkiksi Alan Turingin elämän kohdalla tukeudun Andrew Hodgesin teokseen: *Alan Turing, arvoitus* (suom. Kimmo Pietiläinen, Terra cognita, Helsinki, 2000) ja Ada Lovelacen sekä Charles Babbagen elämän osalta käytän ensisijaisesti James Essingerin teosta: *Adan algoritmi; kuinka lordi Byronin tytär Ada Lovelace käynnisti digiajan* (suom. Tapani Kilpeläinen, Vastapaino, Tampere, 2016). Filosofian tieteenalan puolella on tarkasteltu jonkin verran filosofisen diskurssin historiaa siitä, mitä eri aikakausien filosofit ovat ajatelleet koneiden ajattelukyvystä. Suomessa Descartesin ajattelua on tästä näkökulmasta tutkinut filosofian tohtori Renne Pesonen, ja Turingin ajattelua on ruotinnut etenkin filosofian professori, ja teoreettisen filosofian dosentti, Panu Raatikainen.

1.2 Käsitteet

Mitä on tekoäly?⁴ Helsingin yliopiston tietojenkäsittelytieteen professorin Hannu Toivosen mukaan tekoäly voi olla esimerkiksi metafora, itsenäinen tietokoneohjelma, jolla on vaikutusta johonkin itsensä ulkopuoliseen, tieteenala, tietokoneohjelman tarkoituksenmukaista toimintaa, suuri joukko erilaisia tietokoneohjelmia, mainostermi, meidän tulevaisuutemme, keino tutkia mieltä ja ajattelua, unelma tai painajainen.⁵ Käsite tekoäly esiintyy nykyään hyvin monenlaisissa erilaisissa konteksteissa, eikä aina ole aivan selvää, mitä käsitteellä nykyajassa tarkoitetaan. Tässä kandidaatintutkielmassa tekoälyllä tarkoitetaan kuitenkin pääasiassa tieteenalaa ja johonkin tarkoituksenmukaiseen toimintaan kykenevää tietokoneohjelmaa. Tarkoituksenmukainen toiminta voi sisältää muun muassa tiedon käsittelyä, ongelmien ratkaisemista, oppimista, puheentunnistusta, tekstintuottamista, kuvantunnistusta ja paljon muuta. Tieteenalana tekoäly on pääosin tietojenkäsittelytiedettä, mutta siihen sisältyy myös filosofiaa ja kognitiotiedettä. Tieteenalan tarkoituksena on tutkia älykkään oloisesti toimivia koneita, niiden toimintaperiaatteita, sitä, miten koneet saadaan toimimaan älykkäästi sekä sitä, mitä älykkyys ylipäättään tarkoittaa koneen näkökulmasta.⁶

On tärkeää huomata, että käsite *tekoäly* on käsitteenä melko nuori, ja syntynyt vasta vuonna 1956.⁷ Tätäkin aiemmin historiassa on ollut ajatuksia tekoälyn kaltaisesta koneiden ajattelusta ja älykkyydestä. Tämä tutkielma keskittyy kokonaan tekoälyn varhaishistoriaan, eli aikaan ennen ”tekoäly”-käsitteen keksimistä. Tieteentekijät ennen vuotta 1956 eivät tietenkään puhuneet nimenomaisesti tekoälystä käsitellessään aihetta. Esimerkiksi Alan Turing käytti kirjoituksissaan käsitettä *machine intelligence* eli koneäly. Voidaan kuitenkin oikeutetusti sanoa, että jo antiikista alkava, ajattelevaa konetta ja sen mahdollista kehittämistä koskenut keskustelu, tutkimus ja työ lopulta mahdollistivat tekoälyn ja sen käsitteen synnyn.⁸ Kun tarkastelen alkuperäislähteitäni, joista kaikki on kirjoitettu ennen vuotta 1956, koen myös silloin oikeutetuksi puhua tekoälyn varhaishistoriasta vaikka, varsinaisissa lähteissä kyseistä käsitettä ei esiinnykään. Käytän tässä tutkielmassa pääasiassa käsitettä tekoäly, mutta kirjoitan myös ajattelevista koneista tai koneiden ajattelukyvyistä, sillä alkuperäislähteissä käytetään näitä ilmaisuja.

⁴ Eng. *Artificial Intelligence*, lyhenne AI.

⁵ Toivonen 2023, s. 13, 33, 53, 73, 93, 133, 173 & 193.

⁶ Toivonen 2023, s.73.

⁷ Termin luojana pidetään yhdysvaltalaisista tietojenkäsittelytieteilijää John McCarthyä.

⁸ Perustelen tämän luvussa 1.3.

Tekoäly on perinteisesti jaoteltu filosofien toimesta⁹ kahteen kyvykkyysluokkaan toimintansa perusteella: heikkoon tekoälyyn¹⁰ ja vahvaan tekoälyyn.¹¹ Heikolla tekoälyllä viitataan tekoälyjärjestelmiin, jotka ovat rajoitettuja tiettyihin tarkkaan määriteltyihin tehtäviin. Ne eivät kykene ajattelemaan ihmisten tavoin itsenäisesti vaan ne perustuvat tiettyihin sääntöihin ja algoritmeihin, jotka ohjaavat niiden toimintaa. Heikon tekoälyn järjestelmät voivat kuitenkin sisältää koneoppivia keinotekoisia neuroverkkoja, ja ovat täten kykeneväisiä oppimaan uusia asioita ja kehittymään jatkuvasti tarkoituksenmukaisemmiksi. Nykyiset tekoälyjärjestelmät ovat kaikki heikkoa tekoälyä, vaikka ne voivatkin olla erittäin kehittyneitä ja hyödyllisiä tietyissä sovelluksissa. Heikko tekoäly voi myös suoriutua joistain tarkkarajaisista tehtävistä paremmin kuin ihminen.

Vahva tekoäly viittaa tekoälyjärjestelmiin, jotka omaavat ymmärryksen ja älykkyyden tason, joka vastaa ihmisen älykkyyttä. Ne pystyvät suorittamaan tehtäviä, jotka vaativat älykkyyttä, oppimaan ja kehittämään uusia asioita, sopeutumaan erilaisiin tilanteisiin ja ymmärtämään maailmaa monimutkaisella tavalla. Vahva tekoäly kykenee ratkaisemaan ongelmia itsenäisesti ja kehittämään omia luovia strategioitaan ilman ennalta määrättyjä sääntöjä tai rajoituksia. Tällainen tekoäly vastaisi ihmisen älykkyyttä.¹² Vahvaa tekoälyä ei nykykäsitysten mukaan olla kyetty kehittämään.

Käytän vahvan ja heikon tekoälyn käsitteitä havainnollistamaan erilaisia tapoja puhua koneiden ajattelukyvyistä. On kuitenkin hyvä tiedostaa, että tekoälyn käsitteen tavoin tämä jako on historiallisesti melko tuore. Se on syntynyt vasta tarkastelussani olevien alkuperäislähteiden ilmestymisen jälkeen. Nykykäsitteiden käyttämisessä menneistä ilmiöistä on oltava erityisen tarkkana, jotta kykenee välttämään niiden harhaanjohtavan tai anakronistisen käytön. Koen jaottelun kuitenkin hyödylliseksi, sillä se erottelee toisistaan kaksi erilaista tapaa hahmottaa tekoälyä. Nämä toisistaan eroavat tavat hahmottaa koneiden ajattelukykyä näkyvät jollain tasolla myös alkuperäisaineistoissa, vaikka käsitteellisesti jako onkin tuore.

1.3 Antiikki – mistä kaikki alkoi?

Ihmiset ovat lähes koko historiansa ajan käyttäneet apunaan erilaisia työkaluja ja apuvälineitä kuten tulta, keppejä, kiviä, köysiä, sirpejä, vasaroita ja miekkoja. Tieteen kehittyessä myös apuvälineet ovat kehittyneet yhä moderneimmiksi ja monimutkaisemmiksi. Ihminen on kehittänyt kelloja,

⁹ Jaon on luonut yhdysvaltalainen John Searle vuonna 1997. (Kts. esim. Bringsjord & Govindarajulu 2022, 8).

¹⁰ Eng. Weak AI tai Narrow AI.

¹¹ Eng. Strong AI tai General AI.

¹² Bringsjord & Govindarajulu 2022, 8.1.

kutomakoneita, lentokoneita, junia ja lopulta myös laskimia, tietokoneita ja tekoälyohjelmia. Tekoälyn historian kannalta onkin tärkeää ymmärtää tekoälyn suhde materiaaliseen kehitykseen ja todellisuuteen. Jos tarkastelee tieteen kehitystä, joka on lopulta johtanut nykyisiin tekoälyjärjestelmiin, voi rajanveto olla haasteellista sen suhteen, mistä tekoälyn kehityshistoria alkaa.

Historioitsija ja folkloristi Adrienne Mayorin mukaan historioitsijoilla on yleensä ollut taipumuksena aloittaa teknologian ja automatisaation historian tarkastelu keskiajan käsityöläisistä, jotka kehittivät ensimmäisiä alkeellisia itsestään liikkuvia koneita, kuten kelloja.¹³ Mayorin mukaan olisi kuitenkin hyödyllistä suunnata katse myös tuosta pari tuhatta vuotta taaksepäin ja antiikin Kreikkaan. Vaikka materiaalisessa todellisuudessa antiikin Kreikassa ei nähty teknologian suurta riemukulkua, kiinnittäisi Mayor huomion ajan myytteihin ja tarinoihin. Nuo myytit ja tarinat loivat valtavan määrän teknologisia ideoita ja visioita.¹⁴ Antiikin mytologiaa, tarinoita ja myyttejä tarkastelemalla voimme löytää historian varhaisimmat aavistukset siitä, mitä tiede ja teknologia tulevaisuudessa tuovat tullessaan.

Myös tekoälyn juuret johtavat monilta osin antiikin Kreikkaan. Ensimmäiset kirjallisesti säilyneet ideat ihmismäisestä koneesta löytyvät Kreikan mytologiasta. Tarun mukaan Kreetan kaupunkivaltiota vartioi pronssista veistetty jättiläinen, Talos. Talos oli yksi kolmesta ihmeellisestä lahjasta, jotka tulen ja takomisen jumala Hefaistos teki Zeuksen pojalle Minokselle, Kreetan ensimmäiselle kuninkaalle. Taloksen tehtävänä oli suojella Kreetan saarta merirosvoilta ja muilta uhilta.¹⁵ Talos on historian ensimmäisiä kirjallisia esimerkkejä ajatuksesta, jossa koneeseen yhdistetään ihmismäisiä ominaisuuksia ja piirteitä. Ajattelevan koneen, eli tekoälyn, ajatus on siis kirjaimellisesti tuhansia vuosia vanha.

Mytologian lisäksi myös tieteen kehitys antiikin Kreikassa on omalta osaltaan vaikuttanut tekoälyn kehitykseen tieteenalana. Filosofia, matematiikka, lukuteoria, mekaniikka ja logiikka ottivat isoja edistysaskelia antiikin Kreikassa. Esimerkiksi Aristoteles kehitti syllogistisen logiikan, joka oli ensimmäinen deduktiivinen päättelysystemi. Nykyiset tekoälyjärjestelmät ja niiden takana olevat algoritmit hyödyntävät pitkälle sovellettua deduktiivista logiikkaa. Aristoteleen kehittämää analytiikkaa ja syllogistiikkaa alettiin kuitenkin kutsuaan logiikaksi vasta 500 vuotta Aristoteleen jälkeen.¹⁶

¹³ Mayor 2018, s.1

¹⁴ Mayor 2018, s.1.

¹⁵ Mayor 2018, s.7.

¹⁶ Niiniluoto 2015, 1.

2 René Descartes ja 1600-luku

Rene Descartes oli ranskalainen filosofi, matemaatikko ja tieteentekijä, joka tunnetaan jopa yhtenä modernin filosofian isähahmoista.¹⁷ Hän vaikutti merkittävästi monilla eri tieteenaloilla, kuten filosofiassa, matematiikassa ja fysiikassa. Descartes syntyi Ranskassa, Indre-et-Loiressa vuonna 1596 lääkärin ja asianajajan lapseksi.¹⁸ Vuoden ikäisenä hän menetti äitinsä. Descartes opiskeli La Flechen jesuiittojen kuninkaallisessa oppilaitoksessa sekä Poitiersin yliopistossa. Valmistuttuaan Descartes toimi sotilaallisissa tehtävissä ja osallistui esimerkiksi Prahan valloitukseen kolmikymmentävuotisessa sodassa vuonna 1620.¹⁹

Descartes arvosti rauhallista elämäntyyliä ja vaali yksinäisyyttä ja riittävää lepoa. Hän julkaisi melko vähän ja verkkaisella tahdilla. Descartes ei halunnut kasata itselleen liikaa paineita töidensä suhteen.²⁰ Descartesin tunnetuimmat teokset ovat *Metodin esitys* (1637), *Mietiskelyjä ensimmäisestä filosofiasta* (1641) ja *Filosofian periaatteet* (1644). Metodin esitys pitää sisällään myös tieteen historiassa klassikkoaseman saaneet *Geometrian* ja *Optiikan* tutkielmat. Metodin esitys ei ilmestyessään herättänyt suurta huomiota. Siitä tehtiin vain 500 kappaleen painos, joista 200 jäi Descartesille itselleen ja 300 myyntikappaleta ei Descartesin kuollessakaan vielä ollut saatu myytyä loppuun.²¹ Descartesin tuottamat tieteelliset kirjoitukset oli suunnattu tuohon aikaan pienelle, aiheeseen vihkiytyneelle, tiedeyhteisölle, eikä Descartesin kirjoitukset levinneet laajalle koko hänen elinaikanaan. Tiedeyhteisöjen sisällä Descartesin teokset kuitenkin nauttivat arvostusta, ja moni hänen jälkeensä elänyt filosofi on ammentanut ajatuksiaan Descartesin mukaan nimetystä *kartesiolaisesta perinteestä*. Descartesin teokset noteerattiin myös kirkon toimesta ja kartesiolaisuus nimitettiin harhaopiksi niin katolisen kuin protestanttisenkin kirkon piirissä. Descartesin teokset joutuivat *Index librorum prohibitorum*²² vuonna 1663.²³

Filosofian lisäksi Descartes kehitti analyyttistä geometriaa, muunsi monia analyyttisen geometrian ongelmia algebran kielelle sekä rakensi matemaattista perustaa differentiaali- ja integraalilaskennalle.²⁴ Filosofian historiassa Descartes tunnetaan dualismin suurena kannattajana.

¹⁷ Ainakin G.W.F. Hegel nimitti häntä sellaiseksi (Alanen 2018).

¹⁸ Alanen 2018, 1.

¹⁹ Alanen 2018, 1.

²⁰ Saarinen 1985, s.120.

²¹ Saarinen 1985, s.120.

²² Index librorum prohibitorum (suom. Kiellettyjen kirjojen lista) oli roomalaiskatolisen kirkon laatima luettelo kielletyistä kirjoista. Se julkaistiin vuonna 1559, ja se sisälsi kirjallisia teoksia, joiden lukemista tai omistamista pidettiin harhaoppisina tai moraalisesti haitallisina. Luettelo päivitettiin säännöllisesti, ja se oli voimassa aina vuoteen 1966 asti, jolloin paavi Paavali VI kumosi sen.

²³ Saarinen 1985, s.122.

²⁴ Domski 2022.

Hänen mukaansa voimme epäillä kehomme olemassaoloa, mutta emme mieleemme olemassaoloa. Tämän takia mieli ja keho eivät voi olla sama asia, sillä toisen olemassaoloa voi kyseenalaistaa, mutta toisen ei. Descartes uskoi vahvasti mielen olemassaolon välttämättömyyteen, ja jopa siihen, että mieli voisi olla olemassa ilman ruumista.²⁵ Descartesille oleminen oli ajattelua ja tämä näkyy myös hänen kuuluisimmasta lausahduksestaan ”Cogito ergo sum”.²⁶

2.1 Descartes ja tekoäly

Mielen ensisijaisen aseman lisäksi Descartesin mielenfilosofista ajattelua leimasi myös vahva usko mekanistiseen filosofiaan. Descartesin mukaan kaikki eläimet, ihmistä lukuun ottamatta, olivat ikään kuin mekaanisia automaatteja tai koneita. Hän ajatteli, että eläimillä ei ollut mieltä, vaan niiden käyttäytyminen pystyttiin selittämään täysin fysiologisia syy-seuraussuhteita tulkitsemalla. Descartes uskoi myös ihmisen refleksien perustuvan jonkinlaiseen mekaanisuuteen ja ihmisen olevan kykenevä toimimaan osittain myös ilman mieltä.²⁷

Tämä on huomionarvoista etenkin tarkasteltaessa Descartesin ajatuksia ihmisen ja koneen eroista ja yhtäläisyyksistä. Koska mielellä, ja nimenomaan ihmisen mielellä, on Descartesin ajattelussa niin kohotettu asema, ei Descartes uskonut koneiden olevan koskaan kykeneväisiä ihmisen kaltaiseen ajatteluun, eli vahvaan tekoälyyn. Hän kirjoittaa teoksessaan Metodien esitys, muun muassa seuraavasti:

Jos sen sijaan olisi sellaisia koneita, jotka muistuttavat meidän ruumistamme ja jäljittelevät meidän toimintojamme siinä määrin kuin on asiallisesti mahdollista, meillä olisi aina kaksi täysin varmaa keinoa todeta, etteivät ne suinkaan ole aitoja ihmisiä.²⁸

Nuo kaksi erottavaa tekijää olivat Descartesin mukaan kielellinen ilmaisu ja järki. Ehkä tarpeellista pureutua kumpaankin argumenttiin hieman tarkemmin. Kielellinen argumentti kuuluu seuraavasti:

Voidaan kyllä hyvin kuvitella puhetta tuottava kone, joka tuottaa sanoja ja vieläpä sellaisia, jotka vastaavat sen elimissä muutosta aiheuttavien esineiden aktioihin: jos sitä vaikkapa kosketetaan johonkin kohtaan, niin se kysyy, mitä sille halutaan sanoa, tai jos sitä kosketetaan johonkin toiseen kohtaan, niin se huutaa, että sille tehdään pahaa, ja niin edelleen. Mutta se tuskin osaisi järjestellä sanojaan eri tavoin vastatakseen mielekkäästi

²⁵ Pesonen 2021, s.46.

²⁶ Suom. ”Ajattelen siis olen”.

²⁷ Pesonen 2021, s.46.

²⁸ Descartes 2001 (1637), s.154–155 (suom. Sami Jansson).

*kaikkeen siihen, mitä sen läsnäollessa puhutaan, mihin kaikkein typerimmätkin ihmiset pystyvät.*²⁹

Descartes ei uskonut, että koneet pystyisivät käyttämään lauseita tai yhdistelemään kirjaimia niin kuin ihmiset tekevät ilmaistessaan toisilleen ajatuksiaan. Descartesin tulevaisuuden ennustuksiin ja niiden paikkansapitävyyteen nykyajassa voi suhtautua monin eri tavoin. Vaikka nykyiset tekoälyjärjestelmät eivät vielä olekaan tavoittaneet ihmisen kaltaista älykkyyttä, on tekstin ja puheentuotto yksi pisimmälle kehittyneistä tekoälysovelluksista tänä päivänä. ChatGPT:n kaltaiset kehittyneet kielimallit tuottavat sanoja ja lauseita jo melko virheettömästi monissa tapauksissa. Ne pystyvät vastaamaan niille esitettyihin kysymyksiin, pitämään keskustelua yllä ja toimimaan apuna esimerkiksi tekstin rakenteen hahmottelussa. Descartesin puolustukseksi on kuitenkin myönnettävä, että ihmisen tasolle ne eivät keskustelussa yllä.

Sen sijaan Descartesin toinen argumentti on vielä vaikeampi kumota jopa näin 2020-luvulla.

*Toiseksi, vaikka sellaiset koneet suorittaisivat useita tehtäviä yhtä hyvin kuin me tai jopa paremmin kuin kukaan meistä, ne taatusti epäonnistuisivat joissakin muissa, mikä paljastaisi sen, ettei niiden toiminta perustu tietämiseen vaan pelkästään elimien järjestykseen.*³⁰

On totta, että monet tekoälysovellukset suoriutuvat jo meitä paremmin tietyistä tarkkarajaisista tehtävistä kuten monimutkaisten matemaattisten yhtälöiden ratkaisemisesta, manuaalisesta tiedonjärjestelystä sekä selkeäsääntöisistä peleistä, kuten shakista. Tekoälyjärjestelmillä on kuitenkin taipumus tehdä myös aivan käsittämättömiä virheitä joissain ihmiselle itsestään selvissä asioissa. Tällaisia virheitä voi syntyä esimerkiksi kuvantunnistuksessa, jossa tekoäly saattaa erehtyä luulemaan esimerkiksi kuvaa täynnä pelkkiä viivoja joksikin eläimeksi, tai tiedonhaussa, jossa tekoälyjärjestelmä voi kiven kovaan väittää itsestään selvästi väärää tietoa täysin oikeaksi.

Tällä ”arkijärjen” puuttumisella perustellaan ihmisten eroa suhteessa tekoälyyn nykyajassakin. Ihmiselle ominainen arkinen päättely on usein koneelle täysin käsittämätöntä. Hannu Toivonen esittää kirjassaan *Mitä on tekoäly?* asian yksinkertaisella esimerkillä: jos tehtävänä olisi piknikille meno ja piknikviltti olisi hukassa, mehu lopussa sekä nuha vaivaisi, miten ihminen selvittäisi tilanteen? Ihminen korvaisi viltin vanhalla päiväpeitolla, pakkaisi mehun sijaan vettä mukaan ja jos

²⁹ Descartes 2001 (1637), s.154–155 (suom. Sami Jansson).

³⁰ Descartes 2001 (1637), s.154–155 (suom. Sami Jansson).

nuha olisi allergista, välttäisi allergisoivaa heinikkoa. Ratkaisut tuntuvat meille itsestään selviltä, mutta miten tällaisen luovan ongelmanratkaisun voisi ikinä koodata tekoälysovellukseen? Jos tekoälylle antaisi ohjeeksi korvata piknikviltin vanhalla päiväpeitolla, se voisi selviytyä yksittäisestä tilanteesta, mutta mitä, jos päiväpeittoa ei olisikaan saatavilla? Ihminen valitsisi tällöin vanhan pöytäliinan, maton, pyyhkeen tai vaikka pressun. Tämänkaltaisen päättelyketjun koodaaminen on kuitenkin mahdotonta. Tekoäly menisi tällaisesta täysin sekaisin, eikä kykenisi suoriutumaan mitenkään tästä ihmiselle erittäin helposta tehtävästä.³¹ Ihmisten älykkyys on erilaista kuin koneiden älykkyys.

Descartesin argumentti koneiden ja ihmisen järjen erottelusta vaikuttaa siis edelleen melko pätevältä. Tekoäly tarvitsee usein edelleen erityisen ohjeistuksen eri toimintojaan varten, eikä tavoita ihmisen järjen kaltaista luovuutta ja ongelmanratkaisukykyä. Vaikka Descartes ajatteli argumentoidessaan luultavasti ihmisen näköistä konetta, jonkinlaista robottia, eikä voinut kuvitellakaan nykyaikaisen tietokoneen tai tekoälysovelluksen ulkoista muotoa, on hänen ajattelunsa koneiden ajattelukyvyn suhteen kestänyt yllättävän hyvin aikaa. Monessa muussa asiassa, kuten eläinten mekaniikassa, Descartesin ajatukset on kumottu aikoja sitten tieteen kehityksen myötä. Myös koneiden kielellisen tuottamisen argumentti on haastettu viime vuosien tekoälykehityksen myötä aivan uudella tavalla. Voiko tiede vielä kumota Descartesin järkiargumentinkin? Se vaatisi vahvan tekoälyn kehittämistä.

³¹ Toivonen 2023, s.49–50.

3 Ada Lovelace & Charles Babbage ja 1800-luku

Englantilaista matemaatikkoa Ada Lovelacea pidetään yleisesti maailman ensimmäisenä tietokoneohjelmoijana ja hänen mukaansa on nimetty 1970-luvulla myös ohjelmointikieli ”Ada”. Ada Lovelace syntyi vuonna 1815 lordi Byronin ja tämän vaimon, Annabellan, tyttäreksi. Adan ollessa kuukauden ikäinen, hänen äitinsä otti tytön mukaansa ja he pakenivat lordi Byronin luota. Annabellan ja lordi Byronin avioliitto oli ollut onneton, ja sitä olivat varjostaneet Byronin rahankäyttöongelmat ja pettäminen.³² Pakenemisen jälkeen Ada ei nähnyt isäänsä enää koskaan.³³ Ada kiinnostui jo nuoresta pitäen erilaisten luovien keksintöjen suunnittelusta ja etenkin matematiikasta. Hänestä kasvoi lopulta niin taitava matemaatikko ja älykäs ajattelija, että hänen äidillään oli haasteita löytää hänelle tarpeeksi tasokas yksityisopettaja.

Lovelacen elämän ja tieteellisten kirjoitusten kannalta merkittävä henkilö on Charles Babbage. Babbage oli englantilainen matemaatikko ja filosofi, joka tunnetaan parhaiten yrityksistä rakentaa höyrykäyttöinen ”analyttinen kone”, joka oli ensimmäinen varhainen yritys rakentaa ohjelmitava tietokone. Babbage rakensi myös analyttisen koneen edeltäjänä mekaanista differenssikonetta³⁴, jonka avulla olisi voinut laskea logaritmitaulukoita.³⁵ Babbagen suunnittelemat koneet eivät kuitenkaan koskaan valmistuneet differenssikoneen kapeaa mallikappaletta lukuun ottamatta.³⁶ Hänen piirustustensa pohjalta on kuitenkin myöhemmin onnistuttu rakentamaan myös laaja, toimiva differenssikone. Lovelace ja Babbage tutustuivat toisiinsa, kun Ada oli 17-vuotias.³⁷ He päätyivät lopulta työskentelemään yhdessä muutaman vuoden tuntemisen jälkeen, kun Lovelace tarjoutui kääntämään italiankielisen, Babbagea ja tämän työtä käsittelevän artikkelin.³⁸

Älykkyydestään ja urauurtavasta matemaattisen ajattelun osaamisestaan huolimatta Lovelace ei koskaan täysin uskonut kykenevänsä itsenäiseen tieteentekijän työhön tai julkaissut omalla nimellään mitään, sillä naisena 1800-luvun Englannissa siihen ei kannustettu tai annettu minkäänlaisia mahdollisuuksia. Lovelace ottikin tehtäväkseen vain kääntää Luigi Federico Menabrean kirjoittaman artikkelin Babbagen analyttisestä koneesta. Hän kirjoitti artikkelin oheen kuitenkin Babbagen

³² Essinger 2016, s.19.

³³ Essinger 2016, s.54.

³⁴ Differenssikone oli mekaaninen laskentalaitte. Babbagen tavoitteena oli luoda siitä kone, joka pystyisi automaattisesti suorittamaan monimutkaisia laskutoimituksia. Differenssikoneen ensisijainen tarkoitus oli tuottaa matemaattisia taulukoita, erityisesti polynomeja. Babbage suunnitteli koneen käyttäen hammaspyöriä ja muita mekaanisia komponentteja, jotka mahdollistivat polynomilaskujen suorittamisen.

³⁵ Essinger 2016, s.118–119.

³⁶ Mallikappale oli noin 1/7 laajasta differenssikoneesta. (Essinger 2014, s. 154).

³⁷ Essinger 2016, s.134.

³⁸ Essinger 2016, s.195–196.

avustuksella myös laajan osion nimeltä *Kääntäjän merkintöjä* (*Notes by the Translator, 1843*). Lovelace oli taitava matemaatikko ja hän kykeni näkemään analyyttiselle koneelle sellaisia sovellusmahdollisuuksia, joita Babbage ei itse osannut nähdä.³⁹ Huomautus G on tunnetuin osa *Kääntäjän merkintöjä*. Siinä Lovelace kirjoitti tietokoneohjelman, jota monet pitävät maailman ensimmäisenä laatuaan. Ohjelma oli suunniteltu analyyttiselle koneelle ja sen tarkoituksena oli kyetä laskemaan Bernoullin lukuja.

Lovelace pohti *Kääntäjän merkinnöissä* koneiden ohjelmoinnin lisäksi myös niiden ajattelukykyä. Hän kirjoitti seuraavasti:

*Analyyttinen kone ei suinkaan väitä synnyttävänsä mitään. Se pystyy tekemään mitä tahansa, minkä osaamme määrätä sen tekemään. Se pystyy noudattamaan analyysia; mutta se ei osaa ennakoita mitään analyyttisiä suhteita tai totuuksia. Sen alaa on auttaa meitä tuomaan saataville se, minkä jo tunnemme.*⁴⁰

Voimme huomata, että Lovelace oli Descartesin tapaan skeptinen koneiden ajattelukyvyyn mahdollisuudesta. Hänen ajatuksensa koneiden kyvyistä ja taidoista ovat melko vastaavat nykytutkijoiden käsitysten kanssa. Heikko tekoäly kykenee toimimaan sille annetun algoritmin mukaisesti. Tekoäly tietää sen, minkä me jo tiedämme tai pystymme ennustamaan, mutta se ei kykene ennakoimaan uutta tietoa tai uusia totuuksia. Niin kauan, kuin vahvaa tekoälyä ei ole onnistuttu kehittämään, Lovelacen lausunto näyttää pitävän pintansa.

Tekoälyjärjestelmien kehityshistorian näkökulmasta analyyttinen kone on merkittävä virstanpylväs, vaikka kone ei vielä 1800-luvulla Babbage ja Lovelacen toimesta valmistunutkaan. Analyttinen kone oli ensimmäinen ajatus ohjelmoitavasta tietokoneesta, ja tietokoneohjelmien myötä myös tekoälyohjelmien kehittäminen on nykyään mahdollista. Babbage myönsi saaneensa innoitusta analyyttiseen koneeseen Jacquard-kutomakoneesta, jolla kudottiin kuviollisia kankaita.⁴¹ Jacquard-kutomakone perustui reikäkorttijärjestelmään, jonka Babbage ja Lovelace halusivat analyyttisessä koneessa valjastaa kuviokudonnan sijaan matematiikkaan.

Materiaaliset mahdollisuudet eivät kuitenkaan 1800-luvulla olleet helpot analyyttisen koneen rakentamiseksi. Babbage hylkäsi mekaanisen differenssikoneensa kehittelyn heti keksittyään idean höyrykäyttöisestä analyyttisestä koneesta. Hän oli saanut differenssikoneeseen rahoitusta Britannian

³⁹ Essinger 2016, s.213–216.

⁴⁰ Lovelace 1889 (1843), s.44. (Suom. Tapani Kilpeläinen.)

⁴¹ Essinger 2016, s.158.

hallitukselta, mutta jättäessään sen kesken ja aloittaessaan analyyttisen koneen suunnittelun, hän tuotti rahoittajilleen pettymyksen.⁴² Babbagen oli vaikeaa saada uutta rahoitusta analyyttiseen koneeseen, sillä rahoittajat eivät ymmärtäneet koneen toimintaideaa ja hyötyä tai uskoneet sen relevanssiin. Myöskään uskoa siihen, että Babbage saisi koneen valmiiksi ei enää löytynyt.⁴³

Babbage koitti rakentaa konettaan hammaspyöristä, mutta ei onnistunut projektissaan, sillä analyyttinen kone oli rakenteeltaan varsin monimutkainen. Analyyttistä konetta vastaavalla suorituskvyyllä varustettujen koneiden rakentaminen, kyettiin aloittamaan kunnolla vasta 1930-luvulla, sähkömekaniikan kehityksen myötä.⁴⁴ Tämä kertoo Babbagen idean ainutlaatuisuudesta: vaikka materiaaliset mahdollisuudet eivät vielä sataan vuoteen analyyttisen koneen keksimisestä olleet riittävällä tasolla sen rakentamiseen, hänen suunnitelmansa olivat oikeansuuntaisia ja matemaattisesti päteviä. Myös Lovelacen ajatukset koneen ohjelmoimisesta olivat reilusti edellä aikaansa. Lovelacea pidetäänkin usein maailman ensimmäisenä tietokoneohjelmoijana ja hänen jälkeensä kesti vielä hyvä tovi, ennen kuin tietokoneiden ohjelmoinnista tuli tieteen valtavirtaa.

⁴² Babbage oli saanut hallitukselta differenssikoneen kehitykseen valtavan summan, yhteensä 17 500 puntaa (nykyrahassa 60 miljoonaa puntaa). (Essinger 2016, s.123).

⁴³ Essinger 2016, s.199–205.

⁴⁴ Davis 2003, s.143.

4 Alan Turing ja 1900-luku

Modernin tietokoneen isänäkin pidetty, matemaatikko ja loogikko Alan Turing syntyi Lontoossa 23. kesäkuuta 1912. Hän opiskeli Sherbornen sisäoppilaitoksessa ja loisti kouluaikoinaan etenkin luonnontieteissä ja matematiikassa.⁴⁵ Turing jatkoi yliopistoon *King's Collegeen*, josta valmistui vuonna 1934 hyvin arvosanoin. Vuonna 1938 hän suoritti tohtorintutkinnon Princetonin yliopistosta. Toisen maailmansodan aikana Turing toimi brittiläisen armeijan palveluksessa. Hänellä oli ratkaiseva rooli saksalaisten Enigma-salakirjoitusjärjestelmän murtamisessa, jonka myötä liittoutuneet saivat purettua saksalaisten viestejä.⁴⁶ Sodan jälkeen Turing suunnitteli yhtä ensimmäisistä digitaalisista ja ohjelmoitavista tietokoneista, ja rakensi sen prototyypin.⁴⁷ Alan Turing kuoli syanidimyrkytykseen 7. kesäkuuta 1954. Kuolinsyytutkimuksessa kuolema tulkittiin itsemurhaksi, mutta siitä ei luultavasti saada koskaan täyttä varmuutta.⁴⁸

Turing yhdisteli tieteessään matemaattista logiikkaa, sodanaikaista kryptologian osaamista, käytännön elektroniikkaosaamista ja luovaa loogisfilosofista pohdintaa. Turing ei kuitenkaan itse koskaan pitänyt itseään filosofina, vaikka hänen ajatuksensa ja kirjoituksensa ovat nousseet myöhemmin mielenfilosofian klassikoiksi.⁴⁹ Turingilla oli sodan jälkeen vaikeuksia saada rahoitusta ja tukea suunnitelmilleen rakentaa digitaalinen tietokone. Hänen suunnitelmansa jäivät tiedeyhteisössä muiden hankkeiden jalkoihin, sillä Turing perusteli työnsä merkityksen puhtaasti tieteellisillä motiiveilla, teollisten tai kaupallisten sijaan.⁵⁰ Turingin vahingoksi koitui myös hänen sota-aikaisten saavutustensa salailu kansallisen turvallisuuden nimissä; hän ei voinut esitellä sota-aikansa saavutuksia Enigman murtamisesta, eikä täten todistaa hänen digitaalisen tietokoneen idean moniulotteisuutta ja omaa osaamistaan.⁵¹

Turing kirjoitti jo huomattavasti laajemmalle tiedeyhteisölle kuin Descartes ja Lovelace, mutta hänen aikansa tiedettä leimasi vahva hyötyajattelu, johon Turingilla oli vaikeuksia sopeutua. Turingilla oli vaikeuksia sopeutua myös yhteiskuntaan ja erilaisiin yhteisöihin eksentrisen persoonansa ja

⁴⁵ Davis 2003, s.146.

⁴⁶ Davis 2003, s.172–174.

⁴⁷ Turingin luomuksen nimi oli ACE (= Automatic Computing Engine). (Kts. Davis 2003, s.188–191.)

⁴⁸ Muun muassa Alan Turingin äiti uskoi, että Alanin kuolema oli onnettomuus, joka johtui syanidin huolimattomasta tutkimuskäytöstä. Turingin elämäkerran kirjoittaja Andrew Hodges kuitenkin spekuloi, että on myös mahdollista, että Turing halusi toteuttaa itsemurhansa tavalla, jonka hänen äitinsä voisi tulkita vahingoksi, ettei tämä järkyttyisi. (Hodges 2000, s.506–507.)

⁴⁹ Hodges 2019, 1.

⁵⁰ Hodges 2019, 1.

⁵¹ Hodges 2019, 1.

homoseksuaalisuutensa vuoksi.⁵² Tiedepiireissä etenkin hänen laskennan teorian, matematiikan ja logiikan saavutukset kuitenkin tunnettiin, ja häntä arvostettiin. Turingin sodanaikaiset saavutukset, digitaalisen tietokoneen prototyyppi ja siihen liittyvä teoriatyö sekä tekoölyajattelu saivat laajan yleisön arvostuksen kuitenkin suurelta osin vasta Turingin kuoleman jälkeen.

4.1 Turing ja tekoöly

Turing pohti uransa aikana aktiivisesti koneiden ajattelukyvyyn mahdollisuutta ja ihmisen ja koneiden välisiä eroja. Vuonna 1950 hän julkaisi merkittävän artikkelin *Computing Machinery and Intelligence*. Nykyään tekoölykeskustelun klassikoksi muodostuneessa artikkelissaan Turing esitteli myöhemmin Turingin testiksi nimetyn ajatuskokeen, jota hän itse kutsui nimellä ”jäljittelypeli” (eng. imitation game).⁵³ Jäljittelypelin ideana on, että kuulustelija (ihminen) asetetaan vuorovaikutukseen sekä koneen että ihmisen kanssa ilman tietoa siitä, kumpi on kumpi. Vuorovaikutus käydään tekstin välityksellä, jotta äänensävy ei auttaisi kuulustelijaa. Koneen tehtävänä on saada kuulustelija vakuuttuneeksi siitä, että se on koneen sijasta ihminen. Ihmisen tehtävänä on auttaa kuulustelijaa erottamaan ihminen koneesta.⁵⁴ Turing ennusti, että 50 vuoden kuluttua, eli noin 2000-luvulla, kone onnistuisi jäljittelemään ihmistä niin hyvin, että kuulustelijalla olisi koneen ja ihmisen erottelussa vain korkeintaan 70 % onnistumismahdollisuus.⁵⁵ Voimme todeta, että Turingin ennustus oli aivan liian optimistinen.⁵⁶

Jäljittelypelin lisäksi Turing pohti artikkelissaan laajasti erilaisia perusteluja uskoa tai olla uskomatta koneiden ajattelukykyyn. Hän eritteli yhdeksän tunnettua argumenttia koneiden ajattelukykyä vastaan ja argumentoi itse niitä vastaan. Hän myös mainitsee artikkelissa siirtyvänsä tarkastelemaan omien mielipiteidensä vastaisia argumentteja siirtyessään yhdeksään koneiden ajattelukyvyyn mahdollisuutta vastustavaan argumenttiin.⁵⁷ Jo yksittäin tästä voi päätellä Turingin uskoneen ainakin osittain koneiden ajattelukyvyyn mahdollisuuteen. Turing muotoili omia ajatuksiaan koneiden ajattelukyvyistä seuraavasti:

⁵² Turing eli aikana, jolloin homoseksuaaliset teot olivat laittomia ja homoseksuaalisuutta pidettiin Britanniassa mielenterveyden häiriönä. Turing saikin tuomion homoseksuaalisuudesta vuonna 1952. (Hodges 2000, s.476.)

⁵³ Turing oli sivunnut aihetta jo muutama vuosi aiemmin ilmestyneessä teknisessä työraportissa ja muutamissa julkisissa esiintymisissä, mutta 1950 ilmestyneessä artikkelissa hän esittelee *jäljittelypelin* ensimmäistä kertaa kokonaisuudessaan ja sellaisena, kun sen nykyään tunnemme. (Kts. esim. Copeland 2000 ja Oppy&Dove 2019.)

⁵⁴ Turing 1950, 1.

⁵⁵ Turing 1950, 6.

⁵⁶ Raatikainen 2021, s.71.

⁵⁷ Turing 1950, 6.

Uskon alkuperäisen kysymyksen: "Voivatko koneet ajatella!" olevan liian merkityksetön ansaitakseen keskustelua. Silti uskon, että vuosisadan lopussa sanojen käyttö ja yleinen koulutettu mielipide ovat muuttuneet niin paljon, että koneiden ajattelusta voidaan puhua ilman, että on odotettavissa vastalauseita.⁵⁸

Hän siis uskoi ainakin tulevaisuudessa tekoälyyn, joka osaisi ajatella itse. Mutta mitä Turing piti ajatteluna? Miten hän vertaa koneen ja ihmisen ajattelua toisiinsa? Nämä kysymykset aukeavat paremmin tarkastelemalla muutamaa noista yhdeksästä argumentista ja Turingin vastauksia niihin.

Ensimmäinen huomionarvoinen Turingin vasta-argumentti nähdään kolmannen argumentin kohdalla. Kolmas argumentti on nimeltään *Matemaattinen argumentti* (engl. The Mathematical Objection), ja sen pääteesinä on, että on matemaattisesti ja loogisesti todistettu, että diskreettien tilojen koneilla on tiettyjä rajoituksia.⁵⁹ On siis olemassa tiettyjä asioita, joita kone ei voi tehdä.⁶⁰ Turingin vastaus tähän argumenttiin oli, että on totta, että koneiden kyvyille on todistettavasti tiettyjä rajoituksia. Kuitenkin sen lisäksi on vain väitetty ilman minkäänlaisia todisteita, ettei sellaiset rajoitukset pätsisi inhimilliseen älykkyyteen.⁶¹ Pääajatuksena toimii, että jos ihmisenkin älykkyys ja ajattelukyky voitaisiin edes periaatteessa todistaa matemaattisesti ja loogisesti rajalliseksi, miksi rajoitukset olisivat argumentti koneiden, mutta ei ihmisten ajattelukykyä vastaan? Hän uskoi, että niin koneilla kuin ihmisillä on ajattelukyvyssään tiettyjä rajoituksia, mutta siitä huolimatta niiden toimintaa voidaan kutsua ajatteluksi. Turingin mukaan tämän argumentin esittäjä luultavasti myös hyväksyisi jäljittelypelin älykkyyden mittariksi.

Toinen huomionarvoinen kohta Turingin tekoälyajattelun kannalta tulee esiin neljännen argumentin eli *Argumentti tietoisuudesta* (engl. The Argument from Consciousness) kohdalla. Argumentti tietoisuudesta on tiivistetysti argumentti jäljittelypeliä vastaan, sillä siinä koneelta vaaditaan tunteiden tuntemista, jotta sen voidaan uskoa kykenevän ajattelemaan kuten ihminen.⁶² Turing kuitenkin puolusti jäljittelypeliä argumentoimalla, että emmehän voi tietää, tunteeeko yksilö tunteita, mikäli emme itse ole juuri tuo yksilö. Tämän vuoksi tuntuu mahdottomalta arvioida myös koneen

⁵⁸ Turing 1950, 6. (Suom. Manta Räsänen).

⁵⁹ Turing 1950, 6.3.

⁶⁰ Argumentti on kokonaisuudessaan laajempi ja tarkemmin määritelty kuin annan tässä ymmärtää. Argumentissa mainitaan esimerkkejä, kuten Gödelin lause, jonka avulla tuodaan esiin diskreettien tilojen koneiden rajoituksia. Siinä myös pohditaan tämän argumentin sopivuutta esim. *jäljittelypeleihin*. Tässä yhteydessä ja Turingin tekoälyajattelun ymmärtämisen kontekstissa lukijan ei kuitenkaan ole tarvetta ymmärtää matemaattista argumenttia tämän syvällisemmin.

⁶¹ Turing 1950, 6.3.

⁶² Turing 1950, 6.4.

ajattelukykyä, mikäli emme itse ole juuri tuo kone.⁶³ Emme siis voi esimerkiksi tietää, onko tunteellisesti kirjoitettu runo tai ääneen lausuttu rakkaudentunnustus oikeasti kirjoitettu tai lausuttu tunteen vallassa vai ei. Ulospäin näkyvä tuotos on se, jonka avulla joudumme tekemään arviomme. Jäljittelypelissä ajattelun prosessilla ei ole merkitystä. Ulospäin näkyvä ilmaisu on se, joka ratkaisee, pidämmekö testihenkilöä ihmisenä vai koneena. Tämän perusteella voi tulkita Turingin ajatelleen tuloslähtöisen lähestymistavan olevan ainoa mahdollisuus tutkia objektiivisesti ajattelua. Emme voi päästä kenenkään pään sisään ja havaita ajatusprosesseja, jotka ovat käynnissä tai tunteita, joita ihminen kokee. Miksi siis koneiden tapauksessa ajattelua pitäisi tarkastella prosessin, eikä lopputuloksen kautta?

Turing argumentoi myös luvussa 2 esiteltyä Ada Lovelacen näkemystä vastaan kuudennessa argumentissa nimeltään ”Lady Lovelace’s Objection”. Turingin mukaan Lovelacen epäusko koneiden ajattelukykyä kohtaan kumpuaa siitä, että hänellä ei omana aikanaan, ole ollut riittävää näyttöä uskoa koneiden ajattelukykyyn.⁶⁴ Hän muotoilee Lovelacen argumentin myös yleisempään muotoon: ”*kone ei koskaan voi tehdä mitään todella uutta.*”⁶⁵ Tähän Turing vastaa sanomalla, että ihmisen ajattelusta on myös mahdotonta sanoa, milloin kyseessä on jotain täysin uutta ja milloin seurausta ihmiselle opetetusta asioista tai yleisesti tunnettujen periaatteiden seuraamisesta syntyvästä ”uudesta” ajatuksesta.⁶⁶ Turing siis väitti, että koneille asetettujen kriteerien pohjalta, myöskään ihminen ei kykene keksimään mitään todella uutta.

Tämän lisäksi Turing argumentoi vielä kuutta muuta koneiden ajattelukykyyn mahdollisuutta vastustavaa argumenttia vastaan. Yllä esiteltyjen kolmen argumentin ja vastausten perusteella saa kuitenkin jo melko kattavan kuvan siitä, mitä Turing ajatteli koneiden ajattelukykyvyydestä. Turing nosti tekoälykeskusteluun aktiivisesti myös ihmisen ajattelukykyyn rajoitteita. Hän vertasi koneiden ajattelua ihmisen ajatteluun ja koki usein ongelmalliseksi asettaa koneiden ajattelun arvioinnille kriteerejä, joita ihmisen ajattelun tarkastelussa ei kyetä käyttämään. Turing uskoi vahvan tekoälyn mahdollisuuteen. Hän arvioi, että tulevaisuudessa kone kykenisi ihmisen silmissä vaikuttamaan siltä, että se ajattelee kuten ihminen. Turing pitäisi tätä riittävänä todisteena koneiden kyvystä ajatella, sillä Turing tarkasteli ajattelua tuloskeskeisesti ajatteluprosessin tarkastelun sijaan.

⁶³ Turing 1950, 6.4.

⁶⁴ Turing 1950, 6.6.

⁶⁵ Turing 1950, 6.6. (Suom. Manta Räsänen).

⁶⁶ Turing 1950, 6.6.

5 Vertailua

Descartesin ajatuksista on löydetty yhtymäkohtia Turingin ajatteluun ja tarkemmin jäljittelypeiliin. Hänen koneiden ajattelukykyä pohtivien ajatustensa on sanottu jopa ennakoineen, tai ennalta kuvanneen, jäljittelypeiliä.⁶⁷ Tätä on perusteltu sillä, että Descartes ei pelkästään vastaa kielteisesti kysymykseen: ”Voiko kone ajatella?”, vaan myös perustaa kielteisen vastauksensa siihen, että kone ei kykene ihmismäiseen ajatteluun ja kielen tuottamiseen. Tästä on siis pääteltävissä, hieman anakronistisesti, että Descartes hyväksyisi jäljittelypeilin välineeksi varmistamaan hänen oletuksensa siitä, että kone ei voi ajatella.⁶⁸ Descartes ja Turing näyttävät siis tämän päättelyn pohjalta olevan yksimielisiä siitä, että koneen ajattelukykyä tulee arvioida sen mukaan, miten ihmismäisesti ne pystyvät käyttäytymään.

Merkittävin ero Turingin ja Descartesin tekoälyajattelussa on, että Turing uskoi koneiden ajattelukyvyyn olevan mahdollista tulevaisuudessa, mutta Descartesin mukaan voimme aina erottaa koneen ja ihmisen toisistaan. Kone ei Descartesin mukaan tule koskaan olemaan ihmisen kaltainen ajattelija. Tähän näkemyseroon on varmasti monia eri syitä, mutta yksi mahdollisesti merkittävä ero löytyy siitä, miten erilaisilla aikakausilla Descartes ja Turing elivät. Descartes eli 1600-luvulla, joten hänen ajattelunsa tekoälystä perustuu puhtaaseen ajatusleikkiin. Tuolloin ensimmäiset tietokoneiden prototyyppitkin olivat vielä kaukainen utopia. Turing sen sijaan 1900-luvun puolivälissä oli jo kehittelemässä ensimmäisiä prototyyppisiä ohjelmoitavista digitaalisista tietokoneista. Hänen uskonsa tulevaisuuden teknologian kehitykselle oli luultavasti suurempi kuin Descartesin 1600-luvulla. Hänen oli myös helpompi nähdä, mihin koneet ainakin teorian tasolla voisivat pystyä.

Turingin käsitykset ihmisen ajattelusta olivat myös erilaisia kuin Descartesilla. Descartesille ihmismieli ja ”järki” olivat kohotteisessa asemassa. Hän uskoi, että mieli voisi mahdollisesti olla olemassa myös ilman ruumista, ja että ihmisen järki on täysin uniikki verrattuna kaikkiin muihin luomakunnan olentoihin. Turing sen sijaan rinnasti ihmisen ajattelun suoraan koneiden toimintaan ja vertasi näitä lopputuloskeskeisesti. Ajattelun prosessilla ei sinänsä ollut Turingille merkitystä, vaan sen lopputulos ratkaisi. Tämä ratkaisi Turingin mielestä myös kaikki ajattelun määrittelyyn liittyvät kysymykset ja epäselvyydet. Descartesin ajatuksissa on nähtävissä yhtymäkohtia jäljittelypeiliin, ja hänkin perustelee argumenttejaan lopputulosten pohjalta. Descartesin mielenfilosofian pohjalta on kuitenkin vaikea kuvitella, että hän uskoisi koneiden yltävän lähelläkään ihmismieltä tai inhimillistä

⁶⁷ Oppy & Dove 2019.

⁶⁸ Oppy & Dove 2019.

järkeä, vaikka molemmat käyttäytyisivätkin samalta vaikuttavalla tavalla. Tätä taustaa vasten, on mielestäni kyseenalaista uskoa, että Descartes olisi hyväksynyt jäljittelypelin ajattelun mittarina sellaisenaan. Toki vertailu on hyvin vinoutunutta, sillä Descartesille ei koskaan esitelty jäljittelypelejä, tai edes prototyyppejä sellaisesta koneesta, joka voisi pärjätä siinä. 1600-luvun koneet, joista Descartes puhui, olivat hyvin erilaisia kuin koneet, joista Turing puhui 1900-luvulla.

Descartesin lisäksi myös Lovelace suhtautui koneiden ajattelukykyyn hyvin nihkeästi. Hänen mukaansa kone ei kykene tuottamaan mitään sellaista, mitä emme ohjeista sitä tekemään. Kone voi tuoda tietoomme sellaisia asioita, joita jo tiedämme, mutta ei mitään uutta. Turing uskoi aikakauden teknologisilla mahdollisuuksilla olleen vaikutusta Lovelacen ajatuksiin tekoälystä. Hän uskoi, että Lovelacen on ollut helppo vastustaa vahvan tekoälyn mahdollisuutta, sillä hänen oman aikansa teknologia ei ole antanut hänelle riittävästi perusteita uskoa toisin. Tämä argumentti ei kuitenkaan poista sitä tosiasiaa, että Lovelacen ajattelu kuvaa tämänhetkistä tekoälyn kehityksen tilaa melko hyvin. Heikko tekoäly toimii Lovelacen muotoilemalla tavalla: kone kykenee toteuttamaan sille ohjelmoituja tarkkarajaisia käskyjä, mutta ei kykene keksimään mitään uutta. Lovelacea pidetään maailman ensimmäisenä ohjelmoijana, joten hän on luultavasti nähnyt koneiden ajattelukykyyn rajat juuri ohjelmoinnin rajojen kaltaisina.

Toki nykyaikanakin on hyvin perusteltuja mielipiteitä sen puolesta, että kone kykenee luomaan uusia asioita. Esimerkiksi koneoppivat laajat kielimallit kykenevät hallusinoimaan aineiston pohjalta uusia yhdistelmiä kirjaimille, sanoille ja asiakokonaisuuksille. Tällöin kone näyttää luovan uusia asioita, kun se yhdistelee erilaisia asioita ja muotoilee niitä uudelleen erilaisilla lauserakenteilla. Kone ei kuitenkaan tuota välttämättä mitään relevanttia, ja täysin virheellisesti tai käsittämättömästi hallusinoivat koneet ovat välillä hyvinkin ongelmallisia tuotekehittäjille. Kone ei myöskään ymmärrä, mitä se luo, vaan se vain yhdistelee aineiston pohjalta uudenlaisia merkkijonoja. Kone ei siis toimi erityisen ”luovasti” tai kykene esimerkiksi kehittyneeseen ongelmanratkaisuun. Kuitenkin jos ajatellaan lopputuloskeskeisesti, kuten Turing, kone näyttää luovan uutta. Jos unohdetaan prosessin seuraaminen tai se, että kone toimii vailla minkäänlaista ymmärrystä, se näyttäytyy uutta luovana toimijana. On myös todettava, että usein ihmisenkin ajattelussa uuden luominen on vanhan tiedon ja ajatusten yhdistelemistä uudella tavalla. Vastaus kysymykseen ”Voiko kone luoda mitään uutta?” on siis nykyajassakin melko kompleksinen ja riippuu käsitteiden määritelmästä sekä siitä, keskitytäänkö ajattelun prosessiin vai lopputulokseen. Samalla relevantiksi nousee myös kysymys ”Voiko ihminen luoda mitään uutta?”

6 Johtopäätökset

Elämme parhaillaan keskellä tekoälykesää.⁶⁹ Uusia tekoälyä hyödyntäviä sovelluksia ilmestyy jatkuvasti kuluttajien käyttöön ja etenkin koneoppivien järjestelmien, kuten laajojen kielimallien, kehittäminen on ottanut valtavia edistysaskeleita viime vuosina. Voimme kuitenkin todeta, että tekoäly ei ole ilmiönä uusi. Jo antiikissa omaksuttiin idea ajattelevista ja ihmismäisistä koneista, sekä polkaistiin käyntiin modernin deduktiivisen logiikan kehitys. Olen tässä tutkielmassa tarkastellut kolmen eri tieteen tekijän ajatuksia tekoälystä eri aikakausilta. Heistä kaikki ovat yksimielisiä siitä, että heidän aikakaudellaan vahvaa tekoälyä ei vielä ole ollut. Descartes ja Lovelace uskoivat myös, että vahvaa tekoälyä ei ole mahdollista saavuttaa; kone ei ikinä kykene ajattelemaan kuin ihminen. Turing sen sijaan uskoi, että vahva tekoäly on tulevaisuudessa mahdollista saavuttaa.

1600-luvulla elänyt René Descartes kehitti pohjaa monille matemaattisille sovelluksille, joita teknologia nykypäivänä hyödyntää. Hän myös esitti omat näkemyksensä koneiden ajattelukyvyyn mahdollisuuksista. Descartes ei uskonut vahvaan tekoälyyn. Hän esitti kaksi argumenttia, joiden vuoksi kone ei hänen mielestään kykenisi koskaan ihmisen kaltaiseen ajatteluun. Toinen argumentti oli kielellinen argumentti ja toinen niin kutsuttu järkiargumentti. Kielellistä argumenttia vastaan on helpompi argumentoida nykypäivänä, sillä laajojen kielimallien tekstintuotto on kehittynyt viime vuosina valtavasti. Järkiargumenttia vastaan sen sijaan on haasteita asettua vielä nykypäivänäkin: emme ole onnistuneet kehittämään vahvaa tekoälyä, joten kone ei kykene ihmisen järjen kaltaiseen toimintaan.

1800-luvulla Babbagen ja Lovelacen työn kautta otettiin ensimmäisiä askelia kohti tietokoneiden kehittämisen konkretiaa. Babbage kehitti ensimmäisen mekaanista tietokonetta muistuttavan koneen, *analyttisen koneen*, suunnitelman, mutta ei koskaan saanut konetta valmiiksi. Lovelace oli ensimmäinen, joka pohti mahdollisuuksia ohjelmoida analyttistä konetta, ja kirjoitti maailman ensimmäisenä pidetyn ohjelmointikoodin. Lovelace kommentoi myös koneiden ajattelukykyä ja suhtautui siihen, Descartesin tapaan, hyvin epäileväisesti. Hän ei uskonut koneiden ajattelukyvyyn mahdollisuuteen, eikä siihen, että kone voisi luoda jotain uutta.

Turing kehitti ensimmäisen digitaalisen tietokoneen prototyypin sekä laajasti laskennan teoriaa ja matemaattista logiikkaa, jotka toimivat nykyistenkin tietokonejärjestelmien taustalla. Hän eli aikana,

⁶⁹ Tekoälykesällä (*Summer of AI*) viitataan tässä yhteydessä ajanjaksoon, jolloin tekoäly tieteenalana ottaa suuria harppauksia eteenpäin.

jolloin tietokoneiden, ja mahdollisesti myös ajattelevien koneiden, konsepti sekä olemassaolo oli huomattavasti helpommin kuviteltavissa kuin Descartesin ja Lovelacen aikana. Turingin tekoälyajattelua leimasi koneiden ajatteluprosessin rinnastaminen ihmisen ajatteluprosessin kanssa ja näiden lopputulosten vertailu. Kehittelemässään jäljittelypelissä Turing asetti ulospäin näkyvän käyttäytymisen ajattelun mittariksi ja päätteli, että tulevaisuudessa koneet kykenisivät ihmisen kaltaiseen ajatteluun.

Descartesilla, Lovelacella ja Turingilla on kaikilla hieman toisistaan eroavia ajatuksia ja vastauksia kysymykseen: Voiko kone ajatella? Nämä erot voidaan nähdä osittain selittyvän sen aikakauden ja materiaalisuuden kautta, jossa heistä kukin on elänyt. Kuitenkin myös monet muut tekijät, kuten suhtautuminen ajatteluun, tieteeseen, ihmisyyteen ja teknologiaan, ovat vaikuttaneet ajattelun eroihin. Toki myös tässä mielessä jokainen ajattelija on oman aikakautensa tuote; ajattelu ei koskaan tapahdu ajallisessa tai kulttuurillisessa tyhjiössä. Erojen lisäksi kaikkien kolmen tietentekijän ajattelussa on myös paljon yhtäläisyyksiä. Descartes ja Lovelace eivät kumpikaan uskoneet vahvan tekoälyn mahdollisuuteen. Descartes ja Turing asettivat molemmat koneen ajattelukyvyn mittariksi ihmisen. Kaikki kolme olivat yhtä mieltä siitä, että kone ei ole vielä kykeneväinen ajattelemaan.

Aiheesta voisi tehdä jatkotutkimusta monesta eri näkökulmasta. Tekoälyn historiassa olisi vielä paljon kirjoitettavaa niin tieteellisen diskurssin kuin tieteenalan kehityksenkin puolella. Historiantutkimuksen puolella aiheita on käsitelty hyvin vähän, ja jatkotutkimusideoita ainakin allekirjoittaneella on todella paljon. Aiheesta ei ole kattavaa perustutkimusta, ja nopeasti kehittyvän tutkimuksen vuoksi uutta tekoälyhistoriaa syntyy vuosittain. Myös filosofisen diskurssin historiassa, ja etenkin sen suhteuttamisessa materiaalisuuteen maailmaan ja tieteenhistoriaan, olisi paljon paikkoja jatkotutkimukselle.

Diskurssi koneiden ajattelukyvyn mahdollisuudesta on jatkunut jo pitkään. Tulevaisuudessa se luultavasti nousee vain koko ajan merkittävämpään asemaan. Kone- ja syväoppivat järjestelmät ovat kehittäneet valtavasti heikkoa tekoälyä viime vuosina, ja monet odottavatkin jo kuumeisesti ihmisen kaltaisen vahvan tekoälyn syntyä. Tieteellisen diskurssin historian tunteminen voi auttaa nykyisessä tekoälyvillityksessä eläviä ihmisiä suhtautumaan aiheeseen pitkän, jo 1600-luvulta alkaneen ja maltillisesti, mutta kiihtyen kehittyvän prosessin osana. Descartes ja Lovelace luultavasti kehottaisivat meitä varovaisuuteen liian optimistisissa arvioissamme vahvasta tekoälystä. Turing puolestaan muistuttaa, että mitään mahdollisuutta ei kannata sulkea pois. Tulevaisuudessa tieteen kehityksen myötä lähes kaikki on mahdollista.

7 Lähteet

- Descartes, René, suomentanut Jansson, Sami, ”Metodin esitys” (1637). Teoksessa Aho Tuomo, Jansson Sami (toim.): *Teokset. 1: Yksityisiä ajatelmia, Järjen käyttöohjeet, Metodin esitys, Optiikka, Kirjeitä 1619–1640*. Gaudeamus: Helsinki, 2001.
- Lovelace, Ada, ”Notes by the Translator” (1843). Teoksessa Charles Babbage (kirjoittanut), Henry P. Babbage (editoinut): *Babbage’s Calculating Engines, Being a Collection of Papers Relating to Them; Their History, and Construction* 1889.
- Turing, Alan, ”Computing machinery and intelligence”. *Mind* 49, 1950, s. 433–460. <<https://redirect.cs.umbc.edu/courses/471/papers/turing.pdf>> Katsottu 2.3.2024.

8 Tutkimuskirjallisuus

- Alanen, Lilli, suomentanut Roinila, Markku, ”Descartes, René”. Ensyklopedia Logos, 2018. <<https://filosofia.fi/fi/ensyklopedia/descartes-rene>> Katsottu 29.2.2024.
- Bringsjord, Selmer & Govindarajulu, Naveen, Sundar, ”Artificial Intelligence”. Teoksessa Edward N. Zalta & Uri Nodelman (toim.): *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition) <<https://plato.stanford.edu/archives/fall2022/entries/artificial-intelligence/>>. Katsottu 4.3.2024.
- Copeland, Jack B, ”The Turing Test”. Teoksessa *Minds and Machines* 10, 2000, 519–539.
- Davis, Martin, suomentanut Vilkkö, Risto, ”Tietokoneen esihistoria Leibnizista Turingiin”. Art House: Helsinki, 2003.
- Domski, Mary, ”Descartes’ Mathematics”. Teoksessa Zalta, Edward N. & Nodelman, Uri (toim.): *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition). <<https://plato.stanford.edu/archives/fall2022/entries/descartes-mathematics/>>. Katsottu 2.3.2024.
- Essinger, James, suomentanut Kilpeläinen, Tapani, ”Adan algoritmi; kuinka lordi Byronin tytär Ada Lovelace käynnisti digiajan”. Vastapaino: Tampere, 2016.
- Hodges, Andrew, suomentanut Pietiläinen, Kimmo, ”Alan Turing, arvoitus”. Terra cognita: Helsinki, 2000.
- . ”Alan Turing”. Teoksessa Zalta, Edward N. (toim.): *The Stanford Encyclopedia of Philosophy* (Winter 2019 edition). <<https://plato.stanford.edu/archives/win2019/entries/turing/>>. Katsottu 16.4.2024.
- Mayor, Adrienne, ”Gods and Robots: Myths, Machines, and Ancient Dreams of Technology”. 1st ed. United States: Princeton University Press; 2018.
- Morus, Iwan, Rhys, ”The Oxford illustrated history of science”. Oxford University Press: Oxford, 2017.
- Niiniluoto, Ilkka, ”Logiikka”. Ensyklopedia Logos, 2015. <<https://filosofia.fi/fi/ensyklopedia/logiikka>>. Katsottu 29.2.2024.
- Oppy, Graham & David, Dowe, ”The Turing Test”. Teoksessa Zalta, Edward N. (toim.): *The Stanford Encyclopedia of Philosophy* (Spring 2019 edition). <<https://plato.stanford.edu/archives/spr2019/entries/turing-test/>>. Katsottu 29.2.2024.
- Pesonen, Renne, ”Miksi pitäisi miettiä, voiko kone ajatella?” Teoksessa Raatikainen, Panu (toim.): *Tekoäly, ihminen ja yhteiskunta: filosofisia näkökulmia*. Gaudeamus: Helsinki, 2021, 39–67.

- Raatikainen, Panu, ”Voiko tietokone todella ajatella – filosofisia ajatuskokeita” Teoksessa Raatikainen, Panu (toim.): *Tekoäly, ihminen ja yhteiskunta: filosofisia näkökulmia*. Gaudeamus: Helsinki, 2021, 69–86.
- Saarinen, Esa, ”Länsimaisen filosofian historia huipulta huipulle Sokrateesta Marxiin”. WSOY: Helsinki, 1985.
- Toivonen, Hannu, ”Mitä tekoäly on? : 100 kysymystä ja vastausta”. Kustannusosakeyhtiö Teos: Helsinki, 2023.