



Meri Johnsson

Suomen eettisen tekoälykehityksen tila

Katsaus valtion Aurora AI –ohjelmaan ja sen eettiseen koodistoon suunnittelijan näkökulmasta

Metropolia Ammattikorkeakoulu

Medianomi

Viestinnän tutkinto-ohjelma

Opinnäytetyö

10.05.2021

Tiivistelmä

Tekijä(t):	Meri Johnsson
Otsikko:	Suomen eettisen tekoälykehityksen tila - Katsaus valtion Aurora AI –ohjelmaan ja sen eettiseen koodistoon suunnittelijan näkökulmasta
Sivumäärä:	34 sivua +2 liitettä
Aika:	10.05.2021
Tutkinto:	Medianomi
Tutkinto-ohjelma:	Viestinnän tutkinto-ohjelma
Suuntautumisvaihtoehto:	Digitaalinen viestinä
Ohjaaja(t):	Lehtori Mari Silver

Suomen tekoälykehitys on monissa yhteiskunnallisissa raporteissa noussut yhdeksi menestyneimmistä. Menestyksen mittarina voidaan osittain pitää julkisen hallinnon toteuttamia tekoälyohjelmia, jotka luontaisesti kiihdyttävät tekoälyn käyttöönottoa. Yksi tällainen on parhaillaan (2021) toimiva Aurora AI -ohjelma, jonka tavoitteena on kehittää erityisesti julkisia palveluja tekoälyn avulla. Tekoälyyn liittyy runsaasti sekä mahdollisuuksia että eettistä problematiikkaa. Opinnäytetyön lähtökohtana on tarkastella tekoälyn etiikkaa Aurora AI-ohjelman kautta ja sivuta suunnittelijan roolia ihmislähtöisen yhteiskunnan ja eettisten tekoälypalveluiden kehittäjänä.

Opinnäytetyössä kerätty tutkimusaineisto pohjautuu asiantuntijahaastatteluihin. Haastattelut toteutettiin teemahaastatteluina Zoom-videokonferenssisovelluksen välityksellä. Haastateltavat toimivat sekä palvelumuotoilijoina julkisella sektorilla että data-analytiikan asiantuntijoina (tekoälyn pohjautuvien) palveluiden parissa. Tutkimusaineistossa keskityttiin Aurora AI -ohjelman eettiseen koodistoon ja sen hyödyntämiseen.

Tulokset kiteytyvät yhteiseen ymmärrykseen siitä, että tekoälykehitys kaipaava monialaista ja poikkitieteellistä osaamista eettisten ja ihmislähtöisten yhteiskuntatavoitteiden toteutumiseksi. Eettiset suuntaviivat ulottuvat toistaiseksi vasta johdon tasolle eivät niinkään julkisen sektorin tekoälykehitykseen asti. Muotoiluajattelu, eettisten työkalujen lisäksi, osoittautui kehitystyössä keskeiseksi. Suunnittelijan rooli nähtiin pitkälti operatiivisena johtajana, kokonaiskuvan hahmottajana ja viestijänä.

Etiikan hyödyntäminen digitaalisten palveluiden suunnitteluratkaisuissa voidaan pitää ihmislähtöisen yhteiskuntakehityksen tärkeimpänä suunnannäyttäjänä. Sen avulla pystytään kommunikoimaan kansalaisten kanssa synnyttämällä avointa ja läpinäkyvää keskustelua ja tätä kautta kehittämään tekoälypalveluiden luotettavuutta.

Avainsanat:	tekoäly, tekoälyn etiikka, Aurora AI, julkinen sektori, suunnittelija
-------------	---

Abstract

Author(s): Meri Johnsson
Title: The State of AI in Finland - An overview of Aurora AI program and its code of ethics from a designer's perspective
Number of Pages: 34 pages + 2 appendices
Date: 05 May 2021
Degree: Bachelor of Culture and Arts
Degree Programme: Media
Specialisation option: Digital media
Instructor(s): Mari Silver, Senior Lecturer

Finland's artificial intelligence (AI) development has risen in many social reports as one of the most successful. The measure of success can in part be considered to be artificial intelligence programs implemented by public administrations that inherently accelerate artificial intelligence. One such is currently operating (2021) Aurora AI program explicitly aimed at developing public services using artificial intelligence.

There are plenty of opportunities in artificial intelligence, but also ethical problems. The starting point of this final project is to look at the ethics of artificial intelligence through the Aurora AI program and touch the role of designer in human-oriented society and ethical AI development.

The research material collected is based on expert interviews. The interview was conducted as thematic interviews through the Zoom video-conferencing application. Interviewees work as service designers in the public sector or otherwise work on (AI-based) services.

The results are crystallized in the common understanding that the use of artificial intelligence requires multidisciplinary and interdisciplinary expertise to achieve ethical and people-oriented societal goals. The field values the comprehensive perception and management of entities. The role of a designer in AI is still evolving and has not been taken shape yet.

The utilization of ethics in digital service design solutions can be considered the essential indicator of human-oriented social development. It enables communication with citizens, creating an open and transparent debate and finally develop better public administrations.

Keywords: artificial intelligence, ethics, AI ethics, public sector, designer

Sisällys

1	Johdanto	1
1.1	Tekoälyn määritelmä	3
1.2	Etiikan määritelmä	5
1.3	Tekoälyn moderni historia	6
2	Tekoälyavusteinen yhteiskuntakehitys	9
2.1	Ihmiskeskeinen yhteiskuntakehitys	11
2.2	Valtion Aurora AI –ohjelma	12
2.3	Lainsäädännön näkökulma	13
3	Eettiset periaatteet yhteiskuntakehityksen tukena	15
3.1	Kehityksessä mukana olevat toimijat	17
3.2	Suunnittelukäytännöt	19
3.2.1	Yhdenvertaisuus ja syrjimättömyys	20
3.2.2	Yksityisyydensuoja, tiedon laatu ja hallinta	22
3.2.3	Tiedollinen itsemääräämisoikeus ja omistajuus	23
3.3	Yhteenveto	24
4	Tutkimuksen toteutus	24
4.1	Tutkimusmenetelmä	25
4.2	Kohderyhmän kuvaus	26
4.3	Haastattelun rakenne ja aineiston analyysimenetelmät	26
5	Tutkimuksen tulokset	27
6	Pohdinta ja johtopäätökset	29
	Lähteet	31
	Liitteet	34
	Liite 1. Haastatteluteemat	34
	Liite 2. Haastattelurunko	35

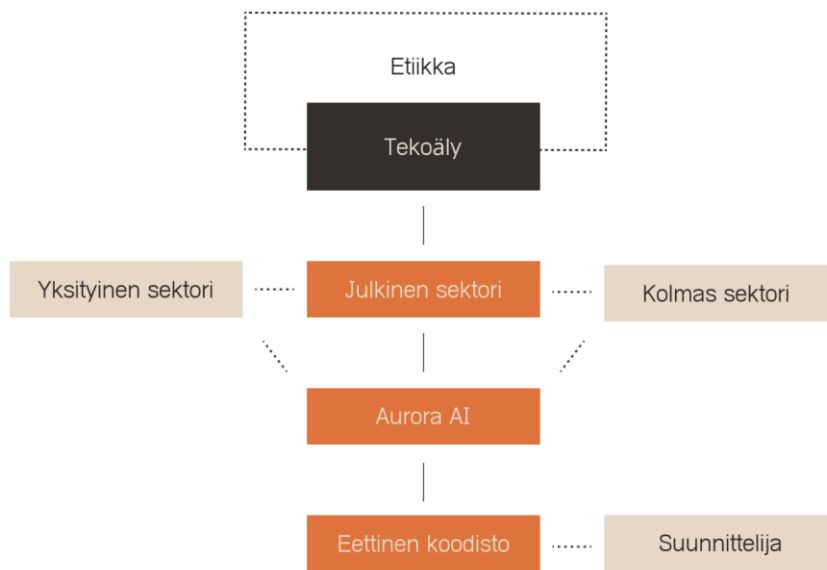
1 Johdanto

Digitalisaatio on luonut uudenlaisen perustan nykypäivän yhteiskunnille ja mahdollistanut modernin tavan lähestyä kehitystä. Erilaisten teknologioiden avulla voidaan digitalisoida palveluja ja helpottaa ihmisten jokapäiväistä elämää. Tekoäly lukeutuu uusien teknologioiden joukkoon ja se nähdään yhteiskuntakehitystä tukevana tekijänä, jonka tarkoitus on automatisoida rutiininomaisia tehtäviä ja siirtää niitä koneelle. Tekoälyn automatisaatioon ja adaptiivisuuteen liittyy mahdollisuuksien lisäksi myös haasteita, jotka luontaisesti vaikuttavat tekoälypohjaisten palveluiden, järjestelmien ja muiden toteutusten käyttäjäkuntiin. Tekoälyavusteisessa yhteiskuntakehityksessä arvopohjaiset ja kestävät ratkaisut nähdään kehityksen keskiössä. Tekoälyn tuottamat lopputulokset vaikuttavat sekä yksilö- että yhteiskuntatasolla, jonka takia tekoälylle on tehty linjauksia ja regulaatiota.

Suomi on listautunut viime vuosina tekoälykehityksen kärkijoukkoihin lukuisissa maailmanlaajuisissa tutkimuksissa. Yksi erottava tekijä kärkimaiden kehityksessä on ollut valtion ja hallinnon asettamat maakohtaiset strategiat. (Törnroth ym. 2020.) Suomen hallitus onkin ollut mukana kehityksessä Euroopan Unionin rinnalla ja ottanut selkeän roolin digitalisaation edistämisessä ja digitaalisten haasteiden ratkaisemisessa. (Valtioneuvosto 2020.) Valtion asettaman Digiohjelman tavoite on vauhdittaa julkisen sektorin digitalisaatiota ja saada kehityksestä kullekin toimijalle avointa ja eettistä. Digiohjelman rinnalla Suomi on keskittynyt tekoälyn tuottamiin mahdollisuuksiin myös muiden hallitusohjelmien kautta. Suomen tekoälyn kenttä alkoi muotoutua erityisesti vuonna 2017, kun työ- ja elinkeinoministeriö aloitti Tekoälyaika-ohjelman tavoitteena saada Suomi yhdeksi tekoälykehityksen esimerkkimaaksi. Menestystä kiihdyttää parhaillaan toimiva tekoälykehitykseen pohjautuva strategia, jonka pyrkimys on yhdistää julkisen, yksityisen ja kolmannen sektorin toimijuudet niin, että Suomesta saadaan globaalisti yksi toimivimmista ja elinvoimaisimmista julkisista hallinnoista. Aurora AI -ohjelma on keskittynyt rakentamaan Suomelle eettisempää tekoälykehityksen kenttää, ja huomioi toiminnassaan lainsäädännöllisen ja ihmisoikeudellisen

perustan. Aurora AI –ohjelman luoma eettinen koodisto on luotu suunnannäyttäjäksi tekoälyyn pohjautuvien palveluiden kehitysprosesseissa. (Valtioneuvosto 2020.)

Opinnäytetyö tarkasteleekin Suomen tekoälykehityksen tilaa, kehitykseen laadittuja eettisiä periaatteita ja parhaillaan toimivaa Aurora AI-ohjelmaa. Pyrin opinnäytetyölläni ymmärtämään ohjelman eettistä koodistoa ja suunnittelijan roolia sen hyödyntämisessä. Pyrin saamaan vastauksen siihen, miten Aurora AI -ohjelma asettamat tavoitteet ovat vaikuttaneet julkisen sektorin toimijoihin, erityisesti kehittäjiin ja suunnittelijoihin. Analysoin aihetta digitaalisten palveluiden suunnittelijan näkökulmasta.



Kuvio 1. Opinnäytetyön keskeiset käsitteet ja rakenne

Keskityn opinnäytetyössä aluksi tekoälyteknologian ja tekoälyyn sovellettavan etiikan määrittelyä (Kuvio 1), jonka jälkeen siirryn tarkastelemaan Suomen tekoälykehityksen tilaa ja syvemmin kansallista Aurora AI -ohjelmaa. Myöhemmissä luvuissa keskityn selvittämään suunnittelijan roolia tekoälykehityksessä. Tutki-

mustuloksia verrataan vuoroin opinnäytetyön teorian kanssa. Työn tutkimusosuudessa oletetaan, että lukijalla on jonkin tason ymmärrys digitaalisten palveluiden suunnittelutyöstä ja siihen liittyvistä menetelmistä. Suunnittelija on katto-termi, jolla viitataan digitaalisten palveluiden suunnittelijaan, kuten palvelumuotoilijaan tai käyttäjäkokemussuunnittelijaan.

Hypoteesini on, että digitalisaatio ja tekoäly muuttavat väistämättäkin suunnittelijan työnkuvaa. Asiointipalvelut, digitaaliset järjestelmät tai automaattiset tiedonvälitykset ovat esimerkkejä nykypäivän tekoälyavusteisista teknologioista. Näen, että tekoälyn lisääminen digitaalisiin palveluihin ja asiakasrajapintaan ovat työnkuvaa muuttavia vastuualueita. Tiedon kerääminen, ihmisen ja koneen välinen kommunikointi tai suosittelualgoritmit voisi oletettavasti olla suunnittelutyön keskiössä, jonka myötä esimerkiksi tuotetestaaminen ja suunnitteluajattelu olisi suunnittelutyön keskiössä. Suunnittelijan tulee näin olla tietoinen uusista teknologioista ja niiden vaikutuksesta loppukäyttäjiin.

1.1 Tekoälyn määritelmä

Opinnäytetyössä termillä ”tekoäly” viitataan joukkoon moderneja teknologioita, jotka sisältävät dataa hyödyntävät sovellukset, palvelut, oppivat algoritmit ja erilaiset tekoälyjärjestelmät.

Teko- eli keinoälyllä tarkoitetaan useita eri teknologioita, ja nimikkeen alle kuuluu ryhmä erilaisia menetelmiä, sovelluksia ja tutkimussuuntia. Sillä viitataan tietokoneohjelmiin, jotka jäljittelevät ihmisen tapaa ajatella, kommunikoida, päättellä ja tunnistaa puhetta tai hahmoja. Tekoäly jaetaan yleisesti kahteen eri ryhmään: yleiseen ja kapeaan tekoälyyn. (Merilehto 2018, 17–26.) Ihanteellisessa tapauksessa yleinen tekoäly kykenisi toimimaan täysin identtisesti ihmisen kanssa. Toistaiseksi tähän ei ole vielä päästy, vaan kaikki tekoäly on kapeaa tekoälyä. Kapea tekoäly ratkaisee ongelmia yksi kerrallaan ja pystyy tällä tavoin vain osittain jäljittelemään ihmisen luontaista älykkyyttä. (Merilehto 2018, 23.)

Koneoppiminen on tekoälyn sisälle sijoittuva osa-alue. Koneoppimisen avulla luodaan ohjelmisto, joka kykenee itsenäiseen toimintaan ja lopulta päätökseen

ilman ihmisen ja koneen välistä kommunikaatiota (Mäntylä n.d.). Tekoälyn ja koneoppimisen eroja on puitu pitkään tekoälyn kentällä. Ollila kommentoi kirjaan Tekoälyn etiikkaa seuraavasti: “Jos se on kirjoitettu Pythonilla, on se todennäköisesti koneoppimista. Jos puolestaan Powerpointilla on se mitä todennäköisimmin tekoälyä.” (Ollila 2019, 54.) Koneoppimisella pyritään automatisoimaan tiedon tulkintaa algoritmien avulla. Toisin sanoen koneoppimisessa kone oppii toimimaan itsenäisesti sille syötetyn datan pohjalta. (Martti 2019.)

Koneoppimista on kolmea eri tasoa: ohjattua, ohjaamatonta ja vahvistusoppimista. Ohjatussa oppimisessa koneelle syötetään oikeaoppinen vastaus opetusdatan avulla, jonka pohjalta saadaan tulokseksi tarkkaan määritelty oikea vastaus. Ohjaamattomassa oppimisessa kone pääättelee asioita ja tilanteita suhteiden ja säännönmukaisuuksien kautta. Tässä opetusmetodissa kone etsii samankaltaisia ryhmittymiä syötetystä datasta. Ohjaamatonta oppimista hyödynnetään silloin, kun valmiita luokkia ei ole vielä määritelty. Vahvistusoppimisessa puolestaan annetaan palautetta siitä, kuinka kone suoriutuu erilaisissa tilanteissa, ilman ennalta määritettyjä vastauksia. (Merilehto 2018, 19.) Sitä käytetään tilanteissa, joissa koneen täytyy operoida monimutkaisissa ympäristöissä havaitsemalla ja toimimalla havaintojen mukaisesti (Ollila 2019, 56).

Neuroverkko on keskeinen tekoälyn tutkimusalue. Se koostuu ryhmistä neuroneja, jotka pystyvät havainnoimaan ympäristöään. Neuronien alkukantainen tehtävä on ollut jäljitellä ihmisaivojen hermosoluja ja synapseja. (Lehtimäki 2021.) Syväoppimisessa hyödynnetään kyseisiä neuroverkkoja, aktivoidaan niitä halutuissa kohdissa ja optimoidaan niitä haastavien ongelmien ratkaisemiseksi. Syväoppiminen tapahtuu päällekkäisissä neuroverkkokerroksissa, jotka ovat tiiviisti yhteydessä toisiinsa ja havainnoivat opittua tiedon pohjalta. (Merilehto 2018, 20; Lehtimäki 2021.)

Data on tietoa, joka on koneellisesti luettavassa, viestittävässä tai käsiteltävässä muodossa. Se on järjestelmissä olevaa, sinne tulevaa tai sieltä lähtevää tietoa, jota puhekielessä kutsutaan myös informaatioksi. Dataa kerätään usein asiakirjoina, dokumentteina, symbolipohjaisina tai numeerisina tietoina. (Sparta

Consulting 2018; Sanastokeskus ry 2013.) Se on yksi keskeisimpiä ja arvokkaimpia osia tekoälykehityksessä. Data ei kuitenkaan ole arvoltaan mitään, jos sitä ei osata hyödyntää oikein. (Aakko 2021.) Kerrytettyjen datavarantojen ja koneoppimisen avulla voidaan löytää uusia yhteyksiä, jotka ihmiseltä jäävät usein huomaamatta, mutta joiden pohjalta voidaan saada eettisesti parempia lopputulemia (Merileho 2018, 33; Aakko 2021).

1.2 Etiikan määritelmä

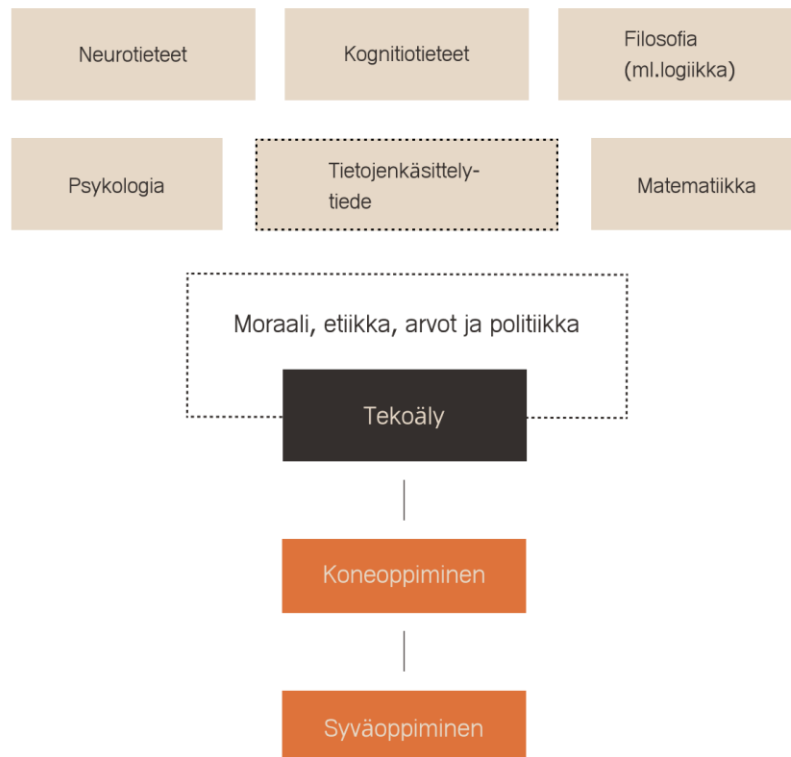
Etiikka on yksi filosofian osa-alue ja pohjautuu tarkemmin moraalifilosofiasta, jonka avulla teorioidaan moraalialia koskevia käsitteitä. Moraali ja etiikka ovat käsitteitä, jotka kulkevat, toisinaan myös erillään analysoitavia filosofian tutkimusaloja. (Ollila 2019, 32.) Tekoälyn etiikka on luonut poikkeuksellisen paljon keskustelua eettisellä kentällä, koska sen vaikutus yhteiskuntaan ja yksilöihin nähdään merkittävänä. Ollilan mukaan etiikka tulisi nähdä keskeisenä suunnannäyttäjänä suunniteltaessa tulevaisuuden toimivia yhteiskuntia (Ollila 2019, 12). Etiikan avulla useimmiten tutkitaan kysymyksiä oikeudenmukaisuudesta, hyvästä elämästä, hyvistä teoista tai ihmisarvosta (Siau & Wang 2018).

Mukaillen Helsingin yliopiston toteuttaman *Ethics of AI* -kurssia tekoälyn etiikka voidaan karkeasti jakaa kolmeen eri tutkimusalaan: metaetiikkaan, normatiiviseen etiikkaan sekä soveltavaan etiikkaan. Metaetiikalla tarkoitetaan käsitteiden tutkimusta. Sen avulla pyritään ymmärtämään eettisten käsitteiden tarkoituksenmukaisuutta, oikeutuksen analyysia, viittaussuhteita ja totuusarvoja. (Rusanen, Nurminen, Raisanen, Tarkoma & Halmetoja 2020; Ollila 2019, 32–34.) Normatiivisen eli preskriptiivisen etiikan avulla pystytään määrittelemään näitä arvoja ja kehittämään teorioita esimerkiksi oikeasta ja väärästä tai hyvästä ja pahasta. Sitä kutsutaan ajoittain myös kehottavaksi etiikaksi, sillä se pohjautuu tyypillisesti moraalisiin käskyihin. Soveltavan etiikan avulla voidaan mukauttaa normatiivisen etiikan teorioita tekoälyn käytännön tapauksiin. Se tarkastelee esimerkiksi sitä, miten tekoälyä on velvoitettu tekemään tietyissä tilanteissa. (Rusanen ym. 2020.)

Tavanomaisin ajatus on, että etiikkaa ei tule jakaa erikseen tekoälyn etiikkaan, sillä yleinen etiikkakäsitys ei ole riippuvainen tietystä toiminnasta, vaan sitä sovelletaan käytännön ilmiöihin (Kimppa 2019; Ollila 2010, 11). Ilmiöllä käsitetään esimerkiksi ihmisiä, yhteiskuntia, talousjärjestelmiä tai ympäristöjä koskevia kehityskaaria, joiden kehitykseen vaikuttavat useat eri tarkastelun alla olevat tekijät (Takala 2018, 19). Samalla tavalla tekoälyn etiikka ei ole tekoälyn itsensä luomaa etiikkaa, vaan etiikka on sovellettavissa tekoälyyn. Tekoälyn etiikka nähdään usein yhtenä soveltavan etiikan osa-alueena, jonka avulla keskitytään sen kehittämiseen, käyttöönottoon ja näihin liittyviin normatiivisiin kysymyksiin (Siau & Wang 2018). Tekoälyn eettiset kysymykset muotoutuvat hieman eri tavalla verrattain muihin teknologioihin, koska se on luonteeltaan autonomista ja adaptoituvaa. Autonominen tekoäly sanansa mukaisesti tarkoittaa kykyä toimia haasteellisissa ympäristöissä ilman käyttäjän jatkuvaa ohjausta. (Ollila 2019, 53.) Usein puhe tekoälyn eettisyydestä liittyykin tekoälyn luoman käyttötarkoituksen ja käyttökohteen sekä tekoälypohjaisten toteutusten ja tulosten eettisyyteen (Ollila 2019, 11–12).

1.3 Tekoälyn moderni historia

Tekoäly tieteenalana on kehittynyt muun muassa filosofian, logiikan, psykologian, kognitiotieteen, tietojenkäsittelytieteen ja matematiikan pohjalta (kuvio 2). Filosofian kautta on totuttu ymmärtämään inhimillisen tiedon luonnetta ja rajoja. (Suomen koodikoulu 2018.)



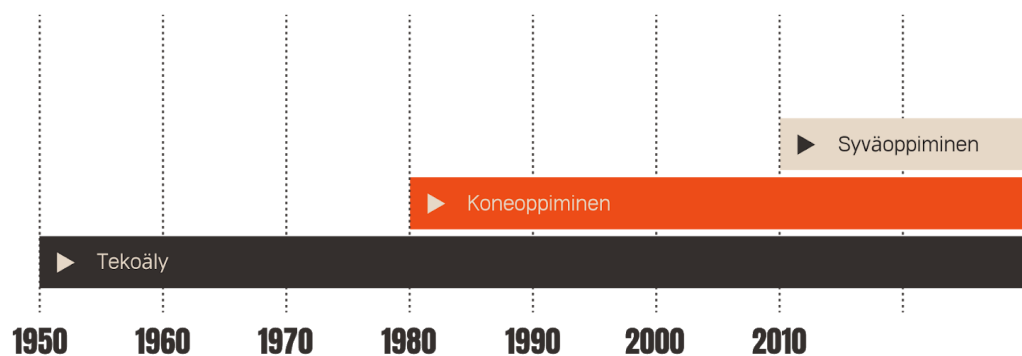
Kuvio 2. Tekoälyn taustatieteet ja sovellusalueet (hahmotelma teorian pohjalta)

Tekoälyn varsinainen kehittäminen alkoi 1950-luvulla Yhdysvalloissa. Ensimmäiset tekoälysovellukset pohjautuivat loogiseen ohjelmointiin ja algoritmeihin, jotka toimivat säännönmukaisuuksien kautta niille syötettyjen merkkien pohjalta. Ennustettiin, että kone pystyisi jäljittelemään ihmisen älykkyyttä jo yhden sukupolven päästä. (Suomen koodikoulu 2018; Merilehto 2018, 69.) Samaan aikaan kone läpäisi Alan Turingin testin, jonka tarkoituksena on mitata tekoälyteknologian kyvykkyyttä erilaisten tehtävien avulla. Tekoälyn tuottamaa tekstiä ei ensimmäistä kertaa pystynyt tunnistamaan koneen kirjoittamaksi. (Suomen koodikoulu 2018.)

Koneiden etiikasta tieteenalana on puhuttu jo 1940-luvun alussa, jolloin sitä käytettiin ennustamaan digitaalisten sotakoneiden tuottamien päätösten sosiaalisia ja eettisiä lopputulemia. Tekoälyn virallinen tutkimus alkoi kuitenkin myöhemmin 1950-luvun puolivälissä. (Borenstein, Grodzinsky, Howard, Miller & Wolf 2021.)

70-luvulla ymmärrettiin, että ihmisen tietoa ei voi rajata pelkkiin säännönmukaisuuksiin tai saatavilla olevaan dataan. Tekoälyn hyödyntäminen rajattiin vain tarkkoihin yksittäisiin ongelma-alueisiin, jonka seurauksena tietokoneet ja niiden laskentatehot kehittyivät ja datan monipuolinen käyttö kasvoi. Ihmisen aivotoinnasta saatiin enemmän tietoa, joka vauhditti edistynyttä tekoälykehitystä. (Suomen koodikoulu 2018.)

2000-luvulla siirryttäessä ymmärrys ihmisaivojen todellisesta toiminnasta avasi myös tekoälyn kentällä; koneelle ei enää tarvinnut syöttää dataa, vaan se pystyi oppimaan myös itsenäisesti. Neuroverkkojen kehitys alkoi ihmisaivojen toiminnasta, joissa yksittäiset aivosolut lähettävät signaaleja ja käytetyt yhteydet vahvistuvat. Tällöin syntyi koneoppiminen. (Suomen koodikoulu 2018.)



Kuvio 3. Tekoälyn moderni historia. (Hahmotelma lähteestä: Merilehto 2018. Tekoäly - Matkaopas johtajalle)

Vasta 2010-luvulla kehittäminen siirtyi koneen syväoppimiseen, jonka myötä myös eettiset kysymykset muotoutuivat tarkemmin. (Suomen koodikoulu 2018.)

Alan kehittäjät tekevät nopeita edistysaskeleita muun muassa oppimisen, päätelyn ja havainnon jäljittelemisessä. Osa asiantuntijoista uskoo, että innovaattorit saattavat kehittää pian järjestelmiä, jotka ylittävät ihmisen luontaisen oppimiskyvyn. Osa puolestaan epäilee, koska kognitiiviset toimet on sidottu arvokysymyksiin, jotka ovat ihmisen kokemuksen alaisia. (Suomen koodikoulu 2018.)

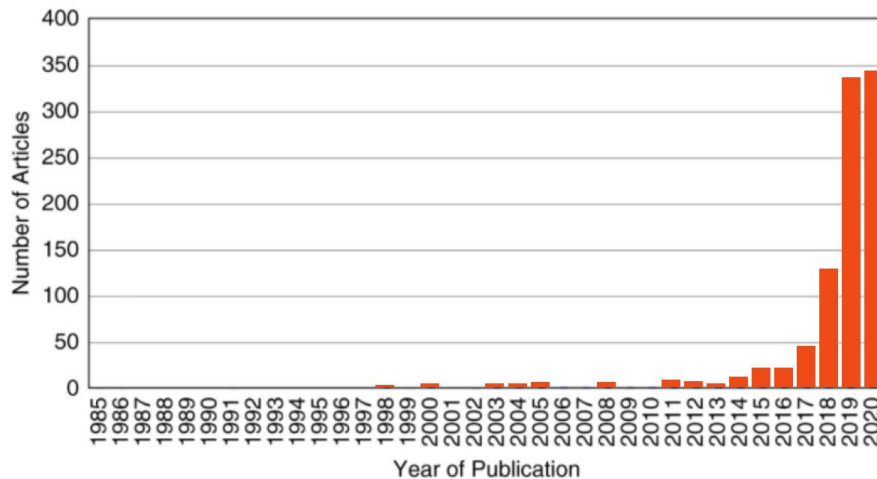


Figure 1. The counts of Google Scholar citations with (“AI” or “artificial intelligence”) and (“ethics” or “ethical”) in the title.

Kuvio 4. Google Scholarista löytyvien lähteiden määrä hakusanoilla “artificial intelligence” tai “AI” ja “ethics” tai “ethical”. (Hahmotelma lähteestä AI Ethics: A Long History and a Recent Burst of Attention)

Tekoälyn eettisyys on populaarikulttuurin myötä noussut merkittäväksi tieteelliseksi tutkimusalaksi. Tekoälylle on vasta 2010-luvulla ryhdytty tekemään asianmukaisia säädöksiä ja regulaatiota. Yllä oleva tilasto (kuvio 4) hahmottaa sitä, kuinka paljon Google Scholar -palvelusta on löytynyt tieteellisiä artikkeleita “etiikka”- ja “tekoäly”-avainsanoilla viimeisten vuosikymmenten aikana. (Borenstein ym. 2021.)

2 Tekoälyavusteinen yhteiskuntakehitys

Suomen Tekoälyaika oli työ- ja elinkeinoministeriön asettama tekoälyohjelma, joka toteutettiin vuosina 2017–2019. Tekoälyaika-ohjelman tavoitteena oli tehdä Suomesta tekoälykehityksen esimerkkinä. Toimijoita kannustettiin kehittämään kestäviä yhteiskunnalle edun mukaisia tekoälyratkaisuja. (Työ- ja elinkeinoministeriö 2017.)

Tekoälykehitys on ottanut merkittävän edistysaskeleen Tekoälyaika-ohjelman jälkeen, ja Suomi on listautunut maailmanlaajuisissa tutkimuksissa tekoälykehityksen kärkijoukkoihin. *McKinsey* -tutkimuksen mukaan Suomi on tekoälyvalmiuden kärjessä yhdessä Saksan ja Yhdysvaltojen kanssa ja sijoittuu näin muiden Pohjoismaiden yläpuolelle. Tutkimuksessa mitataan toimijoiden tekoälyn valmiustasoon vaikuttavia tekijöitä, kuten poliittisen ympäristön tukea, kyvyksien, tiedon ja infrastruktuurin laatua sekä maan investointikykyä. (Törnroth ym. 2020.) Tutkimuksessa selvitettiin myös tekijät, jotka erottavat kärkijoukon muiden toimijoiden kehityksestä. Näitä ovat parempi suorituskky, johtajuus sekä resurssien sitouttaminen tekoälyyn. (Balakrishnan, Chui, Hall & Henke 2020.)

Yhtäläisiä tuloksia esitettiin myös vuoden 2020 Oxford Insightsin ja kansainvälisen kehittämiskeskuksen IDRC:n tuottamassa *AI Readiness Index 2020* -tutkimuksessa, jossa mitattiin valtioiden valmiutta hyödyntää tekoälyä julkisella sektorilla. Suomi sijoittui uusimmassa tutkimuksessa kolmanneksi, heti Yhdysvaltojen ja Ison-Britannian jälkeen. Tutkimuksen mukaan tekoälyn potentiaali nähdään erityisesti julkisen sektorin terveydenhuollon, koulutuksen ja liikenteen eheyttäjänä. (Oxford Insights 2020.)

Globaalissa *Anticipating the impact of AI Ethics within the Public Sector* -raportin (2019) mukaan tekoäly parantaa julkisen sektorin tehokkuutta tekoälyavusteisten integraatioiden avulla. Sen mukaan julkisen sektorin työntekijät pystyvät kehittyvässä kuvassa siirtymään rutiininomaisista tehtävistä merkityksellisiin innovaatiokeskeisiin projekteihin. Nykyiset valtiot hyödyntävät tekoälyä esimerkiksi monimutkaisten petosten havaitsemiseen, terveysjärjestöt arviointitoimenpiteisiin ja poliisivirastot kasvojentunnistustapauksiin. Valtioiden yleisimmät tekoälyavusteiset palvelut ovat kansalaisten kanssa vuorovaikutuksessa olevat chatbot-robotit. (Hashmi 2019.)

2.1 Ihmiskeskeinen yhteiskuntakehitys

Tekoälyn todellinen potentiaali voidaan nähdä yhteiskunnan eheyttäjänä. Sen positiiviset vaikutukset hyvinvointiin, työllistymiseen, pääomankiertoon ja kestävään kehitykseen ovat merkittäviä. (Laitila 2018, 8–15.) Tekoälyn avulla voidaan keskittyä yhteiskunnallisiin epäkohtiin, kuten talouden epätasapainoon, eriarvoisuuteen, ilmastonmuutokseen ja arvomaailman vääristyneeseen kehitykseen. (Rusanen ym. 2020.) Asiantuntijoiden mukaan tekoälyä tulisi pitää positiivisena työkaluna, jonka avulla kehitetään ympäristöä ja ihmiskuntaa parempaan suuntaan. (Hulkko 2021; Halenius, haastattelu, 13.04.2021.) Suomen julkisella sektorilla keskitytään hyödyntämään erityisesti ohjelmistorobotiikka-, automaatio- ja tekoälypalveluita (Työ- ja elinkeinoministeriö 2019).

Työ- ja elinkeinoministeriön (2019) toteuttaman Tekoälyaika-ohjelman loppuraportin mukaan hyvinvointiyhteiskunnan kehittäminen vaatii rohkeita suuntia ja investointeja. Ihmiskeskeisessä yhteiskunnassa ihminen pystyy oppimaan hyvinvoinnistaan tekoälyavusteisten digitaalisten palveluiden avulla. Julkisen sektorin tekoälypalveluista pyritään rakentamaan eettisesti kestäviä, tietoturvallisia ja läpinäkyviä. Myös valtion päämäärä on luoda mahdollisuuksia tekoälyavusteille hyvinvointiyhteiskunnalle. (Työ- ja elinkeinoministeriö 2019.)

Ihmiskeskeisessä yhteiskunnassa yhdistyy ihmisarvoja kunnioittava perusta, jossa kansalaisella on jakamaton ja yhdenveroinen moraalinen asema. Ihmiskeskeisyys toimii tekoälykehityksen toiminnanohjaajana, jossa inhimillisten arvojen kunnioitus tekoälyn kehittämistyössä, käyttöönotossa, käytössä ja sen valvonnassa ovat keskeisessä osassa. Keskiössä ovat myös ympäristön, ekosysteemissä elävien olentojen ja tulevien sukupolvien huomioiminen. (Siau & Wang 2018, 47–53.) Takala on kommentoinut *Sitran Kohti Jatkuvaa Tilannekuvaa* -raportissa, että kehitystä tulee lähestyä systeemisenä kokonaisuutena ja asiakkaan eli kansalaisen kokonaisvaltaisen näkökulman kautta. (Takala 2018, 9.)

Asiantuntijoiden näkemykset tekoälyteknologiasta eroavat osittain muiden toimijoiden näkemyksistä, ja tekoälykehitystä hidastaa tekoälyn huono ymmärrettävyys kansallisella tasolla. Tekoälykeskustelua käydään enimmäkseen ministerien keskuudessa, kun sitä tulisi käydä yhä enemmän kunnilla ja kaupungeilla (Halenius, haastattelu 13.04.2021).

2.2 Valtion Aurora AI –ohjelma

Parhaillaan toimiva Aurora AI -ohjelma toimii esimerkkinä ihmislähtöisen yhteiskuntakehityksen tavoittelusta. Ohjelma on luonnollista jatkoa työ- ja elinkeinoministeriön perustamalle Tekoälyaika-ohjelmalle. Aurora AI -ohjelma on yksi opinnäytetyön avainkäsitteistä ja näin myös eettisen tarkastelun kohteena. Ohjelma on Sanna Marinin hallituksen tekoälyn kehitykseen pohjautuva strategia vuosille 2019–2022, jonka tavoitteena on yhdistää julkisen, yksityisen ja kolmannen sektorin toimijuudet niin, että Suomesta saadaan globaalisti yksi elinvoimaisimmista julkisista hallinnoista. Tämänkaltaisen sektorien välisen verkostotyöskentelyn tehtävä on helpottaa toimijoiden välistä yhteistyötä ja saada tekoälykehityksestä avointa ja eettistä pitäen ihminen keskiössä. (Valtiovarainministeriö 2020.)

Sektorien yhteistyöstä koostuvan Aurora AI -verkon tavoite on kehittää teknillistä sujuvuutta tekoälyyn pohjautuvien palvelujen, sovellusten ja alustojen keskinäiselle tiedonvaihdon ja yhteensopivuudelle (Valtiovarainministeriö 2020). Sitran data-asiantuntija Halenius puoltaa sektorirajat ylittävää tavoitetta, jossa pyritään edistämään ihmisten hyvinvointia ja helpottamaan arjen sujuvuutta (Halenius, haastattelu 13.04.2021). Verkostotyöskentelyn potentiaali nähdään siinä, että tekoälyssä hyödynnettyä dataa käytetään myös julkisen sektorin palveluissa. Ohjelmassa keskitytään erityisesti yhteiskunnan syrjäytymisvaarassa oleviin ryhmiin, kuten nuoriin, vanhuksiin ja maahanmuuttajiin. Ohjelman kolme keskeistä tavoitetta on luoda tekoälypalveluiden verkko, jonka avulla kehitetään kolmea elämäntapahtumaa: 1) nuoret kiinni yhteiskunnassa syrjäytymisen ehkäisyllä, 2) ulkomaalaisten opiskelijoiden kiinnittyminen suomalaiseen työelämään ja yhteiskuntaan sekä 3) kiinni työelämässä jatkuvan oppimisen kautta. (Valtiovarainministeriö 2020.)

Yleisenä haasteena kehityksessä on huomattu toimijoiden päällekkäiset ja suljetut kehitystyöt, jotka hidastavat tekoälyn käyttöönottoa ja kehitystä Suomen valtion tasolla. (Laitila 2020, 19.) Sosiaali- ja terveysalalla toimiva palvelumuotoilija Auer on kiinnittänyt huomiota samaan ilmiöön:

Uskon, että valtion rahalla tehdään paljon päällekkäistä toimintaa Suomen ja varsinkin EU:n sisällä. Se on ohjauksesta ja tietysti rahoituksesta kiinni, että jollain toimijalla on paljon resursseja ja järjestelmiä. (Auer, haastattelu, 08.04.2021.)

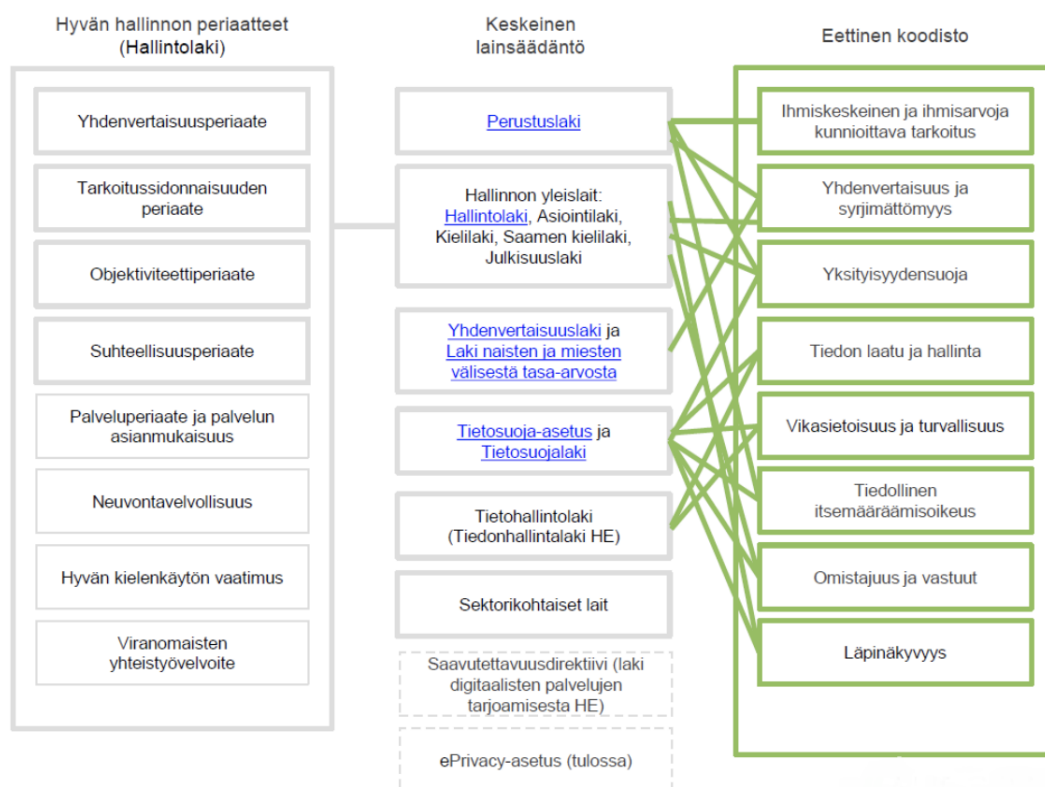
2.3 Lainsäädännön näkökulma

Euroopan Unionin säätelyt ja ratkaisut koskevat myös Suomen tekoälykehityksen toimintaa. Euroopan Unionin strategian mukaan yksilöiden perusoikeuksien, sananvapauden, yksityisyyden ja henkilötietojen suojele tulisi sisällyttää osaksi digitaalisten palveluiden kehitysprosesseja. Tiedon jakaminen sekä EU:n sisällä, että sen ulkopuolella tulisi olla täysin henkilötietoasetuksen ja yksilöiden oikeudet huomioitua. Tietoturvalliset määreet tehoavat vain, jos toiminta perustuu EU:n laatimiin perusoikeuksiin ja vapauksiin. (Euroopan Komissio 2021.)

Tekoäly nostaa esiin sosiaalipoliittisia haasteita ja eettisiä kysymyksiä liittyen moraalisiin, sosiaalisiin, taloudellisiin ja oikeudellisiin epäkohtiin. Tekoälyn soveltaminen julkisen sektorin hallintostrategioissa lisää huolta vastuun vastuu vapauttamisesta tekoälylle. Hallituksen ja yritysjohtajien on määritettävä eettiset riskit ja niiden ennaltaehkäisy ennen tekoälykehityksen alkua. Riskit voidaan minimoida, jos päättäjät määrittelevät ne uhkaaviksi. Suurimmalla osalla organisaatioista ei vielä vuonna 2019 ollut tehtynä minkäänlaisia eettisiä periaatteita kehityksen tueksi. (Hashmi 2019.)

Vuoden 2019 jälkeen Suomi on keskittynyt kehittämään eettisiä suuntaviivoja sekä valtio- että organisaatiokohtaisesti. Työ- ja elinkeinoministeriön Tekoäly-aika-ohjelmassa keskityttiin osittain eettisten periaatteiden määrittelyyn, ja sen jälkeläinen Aurora AI -ohjelma on perehtynyt Suomen kehitystä koskeviin lainalaisuuksiin. (Työ- ja elinkeinoministeriö 2019.) Laadittuja suuntaviivoja pidetään linjassa muun muassa Euroopan unionin perussopimusten ja perusoikeuskirjan kanssa (Siau & Wang 2018, 47–53).

Aurora AI –ohjelma on tuottanut perusteellisen etiikkaselvityksen, jonka mukaan tekoälyn eettisiä ehtoja säädellään nykyisin jo laajasti. Eettinen koodisto pohjautuu lainsäädännölliseen perustaan (kuvio 5), ja siinä on otettu huomioon hallintolain periaatteet ja keskeiset lainsäädännöt. (Haataja & Latvanen n.d.) Eettistä koodistoa tarkastellaan laajemmin opinnäytetyön seuraavassa luvussa.



Kuvio 5 Tekoälyn lainsäädännöllinen perusta (Haataja & Latvanen 2019)

Data-asiantuntija Haleniuksen mukaan digitaalisen palveluiden kentälle olisi hyödyllistä luoda sertifikaatteja, jotka informoivat käyttäjälle palvelun luotettavuudesta ja vastuullisuudesta. Tällaisten laatukriteerien avulla käyttäjä kykenisi helposti tarkistamaan täyttykö vaaditut laatukriteerit käytettävässä palvelussa:

Kun mennään digitaaliseen maailmaan, niin toimijalle ei ole samantaisia kriteereitä, kun esimerkiksi fyysisessä maailmassa. Jos on joku laite, niin sen tulee täyttää turvallisuuskriteerit, jotta sen voi tuoda markkinoille - valvontamekanismi niille jo syntynyt. Digitaalisessa maailmassa ei itseasiassa vielä ole sellaista mekanismia. (Halenius, haastattelu, 13.04.2021.)

Haastattelun jälkeen, vuoden 2021 alussa europarlamentti teki ensimmäisen lakiehdotuksen, joka löytyy *Regulation on a European approach for Artificial Intelligence* -raportista. Lakiehdotuksessa esitettiin, että tekoälyjärjestelmiä tulisi arvioida niiden turvallisuuden mukaan. Tavoitteena on kehittää tekoälyjärjestelmille CE-merkintä, joka kertoo sen luotettavuudesta. Sen tavoite on lisätä kansalaisten luottamusta tekoälyyn ja sen toteuttamiin ratkaisuihin. Viranomaisten on tarkoitus vastata markkinavalvonnasta yhdessä käyttäjien ja valmistajien kanssa. (Euroopan komissio 2021.)

3 Eettiset periaatteet yhteiskuntakehityksen tukena

Julkista sektoria koskevat eettiset näkökulmat *Anticipating the impact of AI Ethics within the Public Sector* -raportin mukaan liittyvät sääntelyyn ja hallintaan, laillisuuteen ja kiistämättömyyteen, turvallisuuteen, sosioekonomisiin vaikutuksiin ja moraalisiin kysymyksiin. Tekoälypohjaisia toteutuksia julkisen sektorin palveluissa voidaan tarkastella esimerkiksi taulukon (kuvio 6) esittämien kysymysten avulla. (Hashmi 2019.)

Sääntely ja hallinta	Miten valtio säätelee tekoälyä? Mikä on tekoälyjärjestelmien moraalinen asema? Kuka on vastuussa tekoälyn tekemistä päätöksistä? Kuinka tuomme kehitykseen avoimuutta?
Laillisuus ja kiistämättömyys	Kuinka tekoälyjärjestelmien laillisuus varmistetaan? Kuinka opetusdatan laadusta pidetään huolta?
Turvallisuus	Miten varmistetaan, että koneet eivät vahingoita ihmisiä? Kuka on vastuussa koneen tuloksista? Kuka korvaa vahingot?
Sosioekonomiset vaikutukset	Millä perusteella tekoäly saa luokitella tai profiloida ihmisiä? Miten lisäämme työllistymistä, kun tekoäly vie osan työpaikoista?
Moraali	Kuinka voimme hallita järjestelmää, joka on ylittänyt ihmisen ymmärryksen? Onko tekoälyn tuottama lopputulos ihmiselle hyödyllinen?

Kuvio 6. Tekoälyn eettisiä näkökulmia (Hashmi 2018)

Nykykäsityksen mukaan yhteiskunta on jäänyt tekoälykehityksen vierestä seuraajaksi, vaikka toimii lopulta sen kustantajana (Laitila 2018, 15). Auer toteaa, että kansalaiset ovat vielä lähtökohtaisesti eriarvoisessa asemassa:

Kansalaiset alkavat olla eriarvoisia ihan jo siinä, että mihin verorahansa käyttää tai minkälaisista palveluista maksetaan ja kuinka paljon. Kyllä kansallisen ohjauksen parantaminen olisi tärkeää. Todella hyvä jos pyritään siihen, että tehdään kansallisia asioita (Aurora AI) - se on kuitenkin kustannustehokasta pitkällä aikavälillä. (Auer, haastattelu 08.04.2021.)

Tekoälyn laaja käyttö, etenkin julkisen sektorin palveluissa, voi myötävaikuttaa siihen, että kansalainen voi joutua syrjäytetyksi. Palveluiden digitalisointi ja tekoälyn kasvava käyttöönotto väistämättäkin vähentää päivittäisi ihmiskontakteja, jonka myötä kansalaiset kuplautuvat. Tämä, etenkin vanhemmalla sukupolvella, saattaa vähentää autonomian-, vapauden ja omanarvontunnetta arkipäiväisessä elämässä. (Ojanen ym. 2019.)

Yhteisöllisyyden ja oikeudenmukaisuuden periaatteiden kautta voidaan tarkastella sitä, miten yhteisön jäsenten tulisi toimia, miten tekoälystä aiheutuvat haitat ja edut tulisi yhteiskunnassa jakautua. Tekoälyn tulee toimia yhteisten normien mukaisesti niin, että kansalaisia osallistetaan keskusteluun. (Leikas 2018.)

Julkisen sektorin sosiaali- ja terveystaloudissa toimiva palvelumuotoilija kommentoi tekoälyä hyödyllisenä kansanterveyttä ennustavana tekijänä:

Näen tämän kauhean arvokkaana, että voimme hallinnoida tavallaan kansanterveyttä ikään kuin ennaltaehkäisevään suuntaan. Tekoäly parhaimmillaan pystyisi ennustamaan ihmisen sairauksia, elämänpolkua, tai muuta - sehän voisi auttaa sitä järjestelmää systeemitasolla juuri oikea-aikaisesti ja ennaltaehkäisevästi. (Anonyymi, haastattelu 14.04.2021.)

Vaikka tekoälyyn pohjautuvat digitaaliset palvelut toimisivat ensisijaisesti kansalaisia varten, tulee nykytodellisuudessa ymmärtää, että digitaaliset palvelut voivat olla syrjäyttäviä. Tekoälypalvelut pyritään suunnittelemaan saavutettavasti, mutta ne ei todellisuudessa sitä vielä täysin ole. (Anonyymi, haastattelu 14.04.2021.)

3.1 Kehityksessä mukana olevat toimijat

Tekoälyä on mahdollista hyödyntää laajasti erilaisiin tehtäviin, kuten toiminnan ohjaukseen, luokitteluun tai itsenäiseen päätöksentekoon. Tekoälyllä pyritään usein tehostamaan ja tukemaan ihmisen tekemää työtä. Autonomiset päätöksenteot eivät välttämättä kuitenkaan ole eettisesti kestäviä, jos niille ole asetettu eettistä päämäärää. Läpinäkyvyyden kannalta sidosryhmien tulisi ymmärtää päätöksenteon kulku ja päätökseen vaikuttavat tekijät. Ymmärrys tekoälyn toimintamekaniikasta ja muuttujista ovatkin eettisesti varsin olennaisia. (Ojanen ym. 2019.)

Koneoppimisessa kone oppii sille syötetyn datan pohjalta, jolloin vastuuta tarkastellessa on tärkeää huomioida sille syötetyn datan laatua, määrää ja datan keräämisestä vastuussa olevat toimijat. Ihmisen osallisuus tekoälyn tekemisistä päätöksenteossa on yksi tärkein tarkastelun alla oleva eettinen tulokulma. Turku

AI Society on listannut toimintaperiaatteet vastuukysymyksille toiminnan tueksi. Sen mukaan suunnitteluprosessissa tulisi aina tarkastella toimijoita, jotka ovat vastuussa datan keräämisestä. Datan keräämiseen liittyy harjoitusdatan laatu, määrä, julkisuus, avoimuus, rajoitteet ja ajankohtaisuus. Tekoälylle tulisi heti alusta alkaen määrittää mallin toimintaperiaatteet, joita ovat muun muassa mallin tarkkuus, luotettavuus, käytetyt muuttujat sekä käsitteiden konstruointi. Käyttäjän tulisi ymmärtää missä määrin ja missä vaiheissa ihminen on tekoälyn ratkaisuihin mukana. (Ojanen ym. 2019.)

Tekoälypalveluiden suunnitteluun osallistuvat henkilöt ovat tyypillisesti koulutettuja tietojenkäsittelytieteisiin, tekniikkaan, tai niihin liittyviin kurinalaisuuksiin. Alojen tarjoama asiantuntemus on tärkeää, mutta ei kuitenkaan riittävää tekoälypalveluiden suunnittelussa. (Borenstein ym. 2021; Laitila 2018.). Kehityksessä tulisi olla laaja kirjo erilaisten taustojen ihmisiä ja näkökulmia, sillä sitä kautta varmistetaan mahdollisimman ihmiskeskeinen ja neutraali lopputulema (Bornstein ym. 2021; Halenius, haastattelu, 13.04.2021)

Laitilan mukaan tekoälyn kehitystä tukevia tieteenaloja ovat muun muassa systeemiajattelu, palvelumuotoilu ja digitalisaatio. (Laitila 2018, 49.) Palvelumuotoilijan yksi tärkeimmistä työkaluista on muotoiluajattelu, jonka tarkoituksena on toimia kriittisen ajattelun apuna (Borenstein ym. 2021). Muotoiluajattelu lukeutuu yhdeksi suunnitteluperiaatteeksi, ja sitä olisi hyödyllistä käyttää sekä tavanomaisissa, että tekoälyyn pohjautuvissa digitaalisissa palveluissa (Anonyymi, haastattelu 15.04.2021).

Helsingin kaupungilla työskentelevä Tiitto näkee suunnittelijan roolin keskeisenä tekoälykehityksessä:

“Suunnittelijalla on aika iso rooli selvittää, mikä on mahdollista. Siinä korostuu se, että suunnittelija hallinnoi myös riskejä. Uskon, että kun suunnittelija tekee valintoja, on mahdollisuus tehdä oikeita tai väärä valintoja. Suunnittelijan tulisi olla tietynlainen operatiivinen johtaja tekoälykehityksen keskiössä, ja olla tietoinen kaikkien osalueiden riskeistä ja mahdollisuuksista.” (Tiitto, haastattelu 01.04.2021.)

3.2 Suunnittelukäytännöt

Tekoäly on lähtökohtaisesti luonteeltaan autonomista, jonka myötä voi olla haasteellista ennakoida. Sen itseoppivuuteen liittyy riskejä ja sen tekemiä päätöksiä voi olla haasteellista perustella tai hyväksyä. Aurora AI –koodisto pyrkii määrittämään rungon, joka nähdään oikeudenmukaisena ja turvallisena. Suunnitteluprosessin alkuvaiheessa tulee alustaa mahdolliset riskit ja mahdollisuudet. Kartoituksen tarkoituksena on luoda vastustuskykyisiä palveluja hyökkäyksiä vastaan. (Ojanen ym. 2019.)

Tekoälyyn sovellettu etiikka voidaan karkeasti jakaa kolmeen eri tasoon, joiden avulla tekoälyyn pohjautuvia järjestelmiä kehitetään. (kuvio 7) Yksi taso on eettisten tekijöiden suunnittelu, jossa päättelykyky ja päätöksenteko liitetään osaksi autonomisia järjestelmiä. Toista tasoa kutsutaan suunnitteluetiikaksi, jonka kautta kehittäjät pyrkivät toteuttamaan tekoälyratkaisuja, jotka tukevat yhteiskunnallisia rakenteita, ja hyödyntävät uusimpia säännöksiä ja metodeja osana kehitysprosessia. Viimeinen taso on kehittäjien ammattietiikka: sitä tulee tukea kehityksen tutkimus-, suunnittelu-, toteutus- ja hallintavaiheissa, ja pyrkiä mahdollisimman arvoneutraaliin lopputulemaan. (Dignum 2018.)

Eettisten tekijöiden suunnittelu (Ethics by design)	Eettisen päättelyn ja päätöksenteon liittäminen osaksi autonomisia tekoälyjärjestelmiä teknisten ja algorimisten ratkaisujen avulla.
Suunnitteluetiikka (Ethics in design)	Suunnittelijajattelun, sääntely- ja suunnittelumetodien käyttö ja kehittäminen tekoälyjärjestelmien sulautuessa yhteiskunnallisiin rakenteisiin.
Ammattietiikka (Ethics for design/ers)	Tekoälyn kehittäjien ja hyödyntäjien eettisten ja moraalisten toimintamallien (standardien, sertifiointien) tukeminen tekoälyjärjestelmien tutkimus-, suunnittelu-, toteutus- ja hallintavaiheissa.

Kuvio 7. Tekoälyn etiikan kolme tasoa (Dignum 2018)

Julkisen sektorin Aurora AI -ohjelman etiikkaverkosto on keskittynyt kehittämään toimijoiden käyttöön suunnittelukäytännöt. Suunnittelukäytäntöjen tarkoitus on tehdä kullekin ohjelmassa työskentelevälle tai ohjelman kanssa vuorovaikuttavalle sitovat toimintatavat. (Haataja & Latvanen n.d.) Suunnittelukäytännöissä huomioidaan tekoälyratkaisun vaikutus yksilöön, lainsäädäntöön ja ympäristöön. (Haataja & Latvanen n.d.)

Oikein hyviä pointteja, koska mä näen tämän listan niin totta kai tuli heti mieleen data, että tätä tarkastellaan enimmäkseen datan kautta, mutta tottakai siihen tulee ihan kaikki mitä se prosessi sisältää alusta loppuun. Nää on niin monitulkintaisia, että ehdottomasti semmoista konkretiaa tarvis, mutta tosi hyvä edes niinku tämmöiselle maallikolle, joka ei siellä kentällä ole. (Halenius, haastattelu 13.04.2021.)

Aurora AI -ohjelman eettinen koodisto toimii alalukujen jakajana ja tarkastelun kohteena. Koodisto sisältää seuraavat määreet: ihmiskeskeinen ja ihmisarvoja kunnioittava tarkoitus, yhdenvertaisuus ja syrjimättömyys, yksityisyydensuoja, tiedon laatu ja hallinta, vikasietoisuus ja turvallisuus, tiedollinen itsemääräämisoikeus, omistajuus ja vastuut, sekä läpinäkyvyys. Luvuissa käydään vain osa koodiston sisältämistä periaatteista ja periaatteita on niputettu samaan lukuun loogisuuden lisäämiseksi. Luvuissa keskitytään teoriaan myös muista lähteistä ja peilataan niitä haastattelujen pohjalta syntyneisiin tuloksiin.

3.2.1 Yhdenvertaisuus ja syrjimättömyys

Aurora AI -ohjelma pyrkii rakentamaan ihmisten tarpeita hyödyttäviä tekoälyyn pohjautuvia digitaalisia ratkaisuja. Yhdenvertaisuus ja syrjimättömyys ovat tekijöitä, jotka ovat toiminnan keskiössä: se näkyy esimerkiksi kriittisten ryhmien kehitystyössä. Tekoäly ei kuitenkaan automaattisesti ole eettistä, vaikka sille kehitetään eettinen käyttökohde, tai se pyrkii olemaan kansalaisten edunmukaista.

Mietin, että otetaanko tavallaan heikossa asemassa olevat ryhmät huomioon? Niin ei varmasti oteta, että sehän on yks niistä vaa-roista, kun me digitalisoidaan meidän yhteiskuntaa, niin emme samalla huolehdi siitä, että kaikki pysyisi kehityksessä mukana. Tällä tavoin me itseasiassa lisätään sitä eriarvoisuutta yhteiskunnassa. (Halenius, haastattelu 13.04.2021.)

Tekoälyn eettisyys on osittain kiinni tekoälylle syötetyn opetusdatan laadusta, sen keräämismenetelmistä, organisoinnista, muokkaamisesta ja tallentamisesta. (Halenius 2019; Ojanen, Oljakka, Sahlgren, Tuikka & Vaiste 2019.) Yhteiskunnallisissa palveluissa erityisen tärkeäksi tulee käyttää dataa, joka on mahdollisimman eettisesti kerätty, käsitelty ja käytetty. Usein kun data on vinoutunutta, sen vinoutuminen vaikuttavat eniten heikossa asemassa oleviin ihmisiin. Jos opetusdata on vääristynyttä, voidaan olettaa, että lopputulemakin on vääristynyt. Tämän tyyppisiä esimerkkejä on paljon, kuten pankkipalvelut, jotka tietyn sukunimen perusteella tuottavat kielteisen lainapäätöksen. Vinoutumista ei välttämättä olla tietoisia, jonka takia ihminen tulee määritellä vastuulliseksi tekoälyn luomista tuotoksista. (Halenius, haastattelu 13.04.2021.)

Halenius toteaa, että datan vinoutumat näkyvät usein syrjivinä lopputuloksina:

Ajatellaan niin, että sitä kerättyä dataa esimerkiksi haavoittuneista ryhmistä ei voi käyttää, koska ihmisillä voi olla niin harvinaisia sairauksia, että datasta saisi helposti selville kuka henkilö on kyseessä. (Halenius, haastattelu 13.04.2021.)

Halenius on kuitenkin huomannut työssään, että usein ihmiset toivovat sujuvuutta arkeen profiloinnista riippumatta:

He usein todella toivovat, että sitä dataa käytetään. Ei heillä ole väliä tavallaan sen yksityisyyden suojan kannalta. Kunhan vaan saataisi helpotusta arkeen. (Halenius, haastattelu 13.04.2021.)

Kun kyseessä on aktiivisesta kansalaisesta, voidaan ajatella, että jokainen haluaa olla toimija, eikä niinkään toiminnan kohteena. Tekoälyä hyödyntävät palveluratkaisut tulisi suunnitella laadukasta ja syrjimätöntä dataa käyttäen. Tällaisten palvelukokonaisuuksien suunnittelijoilla tulisi olla vastuu perehtyä datan puutoksiin. Toisaalta tekoälyn hyödyntäminen on osittain kuin peili, joka näyttää ihmismielen ennakkoluulot ja olettamukset: ihmiseltäkään ei aina voi olettaa hyviä ratkaisuja. (Halenius, haastattelu 13.04.2021.)

3.2.2 Yksityisyydensuoja, tiedon laatu ja hallinta

Suomessa jokaisella henkilöllä on oikeus omaan yksityisyydenturvaan perustuslain mukaisesti. Alustatalous on mahdollistanut sen, että tieto on yksittäisten toimijoiden hallussa ja tarjoaa ihmisen toiminnan ennustamisen ja seuraamisen entistä kätevämmiin. Tekoälyn kohdalla yksityisyydensuojan toteutuminen sallii yksittäisen ihmisen oikeuden esimerkiksi tarkistaa ja hallita henkilötietojensa käyttöä. (Ojanen ym. 2019.)

Vuoden 2018 voimaan tulleen tietosuoja-asetuksen GDPR:n (*General Data Protection Regulation*) myötä Euroopan toimijoiden tulee pyytää käyttäjiltä lupa, kun tietoja kerätään, tai ihmisiä profiloidaan (Ojanen ym. 2019). Tiitto mainitsee haastattelussa GDPR-asetuksen olevan hyvä alku tiedon keräämiseen vaikuttavista tekijöistä käyttäjälle. Suunnittelutyössä tulisi keskittyä tarkemmin käytettävän tiedon selitettävyyteen:

Käyttäjiä voi varmasti epäilyttää, jos on kymmenien sivujen sanallinen kertomus siitä, miten tietoja käytetään. Läpinäkyvyys tarkoittaa myös sitä, että kommunikaation laatu niistä järjestelmistä ja käyttötapauksista pitää olla sellaista, että sitä ymmärtää kuka tahansa maallikko - se onkin aika hankalaa. (Tiitto, haastattelu 01.04.2021.)

Tietosuoja-asetuksen hyödyt näkyvät erityisesti kansalaisille läpinäkyvyytenä palveluiden käyttöönotossa ja helpottaa tiedonsaantia. Rekisterienpitäjien oletetaan noudattavan tietosuoja-asetuksen periaatteita ja velvoitteita, jotta pystytään varmistamaan henkilötietojen laillisuus. (Europarlamentti 2020.)

Suunnittelijan rooli kehityksessä tulee olla viestijänä kansalaiselle. Tietoa datan keräämisestä tulisi esittää asteittain kymmenien sivujen mittaisten tietosuojaselosteiden sijasta, joita vain yritysten lakimiehet kykenevät ymmärtämään. (Halenius, haastattelu 13.04.2021)

3.2.3 Tiedollinen itsemääräämisoikeus ja omistajuus

Kansalaisen itsemääräämisoikeuden kunnioittaminen on tärkeä ihmisarvoa ja vapautta kunnioittava periaate. Sen mukaan yksilöllä on vapaus päättää omista valinnoistaan, eikä sille kuuluvia oikeuksia tule rajoittaa. (Leikas 2018.)

Tulevaisuuden mahdollinen näkökulma yksilölliseen ja käyttäjälähtöiseen tietoon voi toteutua *MyData* -hankkeen myötä. Hanke on luotu edistämään ajatusta, että kansalainen itse omistaisi oman tietonsa. Kansalainen pystyisi esimerkiksi lisäämään tai poistamaan oman datan käyttöoikeuksia (Laitila 2018, 51). Haleniuksen mukaan tällainen oman tiedon omistajuus hyödyttäisi Aurora AI-ohjelmassa toimivien verkostojen välistä yhteistyötä myös niin, että viestintä loppukäyttäjille siitä, miten palvelua käytetään, on selkeästi loppukäyttäjän itsemääräämisoikeuden alla. Tällainen viestiminen kuuluisi esimerkiksi suunnittelijan työnkuvaan, jossa tieto tuodaan käyttäjän näkyville. (Halenius, haastattelu 13.04.2021.) Halenius vertaa tällaista tiedon liikuttelun hyödyntämistä konkreettisen esimerkin avulla:

Woltin kuskit eivät tämänhetkisen keskustelun mukaan ole ihan vahvimmassa asemassa. Mutta mahdollisuus, jossa Woltin huippukuski voisi kilpailuttaa omat keikkansa arvioinneista saatujen tilastojen pohjalta, ja tarkistaa kenellä työnantajalla on parhaimmat ehdot, mahdollistaisi sen, että tällaisia diilejä voisi napata useammasta eri firmasta riippuen tarjottavasta diilistä. (Halenius, haastattelu 13.04.2021.)

Haleniuksen mukaan suunnittelun tarkeys korostuu tällaisten palveluiden yhteydessä, jossa ihminen on itse palvelun keskiössä. (Halenius, haastattelu 13.04.2021.)

Tiedollista itsemääräämisoikeutta on myös pohdittu toisesta näkökulmasta:

Siinä on sellainenkin eettinen puoli, että jos tietoja rajoitetaan tosi paljon, siitäkin voi seurata negatiivisia vaikutuksia. Jos jokaisella on oikeus omaan tietoonsa ja itse päättää miten sitä käytetään, eli mahdollisuus pistää täysi stoppi, että mitään ei tavallaan missään näy - niin koen, että siitä ei seuraa ihmiselle pelkkää hyvääkään. (Tiitto, haastattelu 01.04.2021.)

Palvelumuotoilija Auer on kriittinen oman tiedon omistajuuteen liittyvistä seikoista:

Jotenkin ajattelen, että tiedon omistajuuteen liittyy aina joku niin kuin omistusoikeus; jos minä olisin säveltäjä ja sävellän teoksen, niin minulla on siihen omistusoikeus. Jos joku käyttää sitä minulta kysymättä, minulla on oikeus korvauksen saantiin ja toiselle velvollisuus korvauksen maksamiseen. Siinä vaiheessa, kun en enää halua, että kukaan soittaa sitä radiossa, niin se tulee ottaa pois soitosta. Jos peruutan suostumuksen, mihin se tieto on siinä vaiheessa jo mennyt? Saadaanko se sieltä pois? Tällä hetkellä epäilen. (Auer, haastattelu 08.04.2021.)

3.3 Yhteenveto

Ihmiskeskeisessä kehitystyössä suunnitteluun liittyy muun muassa moraalista, eettistä ja arvomaailmallista problematiikkaa, jotka tulee ottaa huomioon tekoälysuunnittelun eri vaiheissa. Suunnittelijan on hyödyllistä ymmärtää alan ammatti- ja suunnittelu-eettiset metodit ja hallita eettisten tekijöiden suunnittelumekanismeja varmistaakseen mahdollisimman eettinen ja arvoneutraali lopputulema.

Aurora AI-ohjelman koodistossa nousi esille esimerkiksi läpinäkyvyys, turvallisuus ja syrjimättömyys. Suomen julkisen sektorin tekoälykehitys näyttäytyi kankealta, eikä järjestelmät toistaiseksi kommunikoi keskenään. Suunnittelijan rooli vaihtelee projektikohtaisesti kehittäjästä palvelumuotoilijaan ja se nähdään kehityksen kokonaiskuvan hahmottajana, operatiivisena johtajana sekä lopputulosten arvioijana. Tekoälytoteutuksen kehityksen lopussa huomio tulee kiinnittää erityisesti yhdenvertaisuuteen ja syrjimättömyyteen

4 Tutkimuksen toteutus

Tutkimus toteutettiin kvalitatiivisena eli laadullisena tutkimuksena. Laadulliselle tutkimukselle tyypillistä on, että sen avulla ymmärretään tarkasteltavaa ilmiötä tutkittavan näkökulmasta. Laadullisessa tutkimuksessa analysoidaan tutkittavan ymmärrystä, ajatusmaailmaa, tunteita sekä kokemusta tietyistä ilmiöistä. (Juuti & Puusa 2020.)

Tämänkaltaisessa tiedonkeruun tavassa tutkija on ensisijainen tiedonkeruun väline, mikä tarkoittaa sitä, että tieto suodatetaan tutkijan, haastateltavan ja tutkimusympäristön välisen kanssakäymisen kautta. Vuorovaikutuksessa huomioidaan vastaajan omat erityispiirteet ja minimoida mahdolliset ympärillä olevat häiriötekijät. Tutkija analysoi jatkuvasti, miten puolueellisuus, joka voi liittyä esimerkiksi identiteettiin, kielikykyyn, kulttuuriin, tai tapoihin, voi vaikuttaa haastattelutilanteen tulkintoihin. (Lapan, Quartaroli & Riemer 2011, luku 4.)

4.1 Tutkimusmenetelmä

Tutkimusmenetelmänä käytettiin teemahaastattelua. Teemahaastattelu on yksi puolistrukturoidun haastattelun sovellettu osa-alue (Eskola & Vastamäki 2015). Teemahaastattelun hyödyt ovat sen vapaamuotoisuudessa, sillä se mahdollistaa ajatusten kehittymisen tutkimuksen edetessä (Puusa & Juuti 2020).

Teemahaastattelussa tutkija on etukäteen määrittänyt osan tutkimuksen lähtökohdista ja voi tätä kautta ohjata keskustelua haluttuun suuntaan. Haastateltavan oletetaan eläneen jonkin tietyn asian, ajan tai prosessin läpi, ja hän pystyy keskustelemaan siitä omakohtaisesti. (Puusa & Juuti 2020) Tässä opinnäytetyössä oletetaan, että kukin asiantuntija on jollakin tavalla tietoinen Suomen tekoälykehityksen tilasta ja valtion laatimasta Aurora AI -ohjelmasta.

Tutkimusteemoiksi valittiin Suomen nykyisen tekoälykehityksen tila ja Aurora AI-ohjelma, tekoälyyn kohdistuvat eettiset periaatteet ja niiden hyödyntäminen suunnittelutyössä. Jokainen haastattelu toteutettiin kunkin haastateltavan asiantunjuuden näkökulmasta ja tarkemmat apukysymykset rakentuivat yksilöllisesti niiden ympärille. Haastattelut toteutettiin etätapaamisena Zoom-videokonferenssisovelluksen avulla ja niiden kokonaiskesto vaihteli kolmesta vartista tuntiin. Ääntä ja kuvaa taltioitiin aineiston myöhempää läpikäyntiä, analyysia ja tiivistelmää varten. Tutkimusaineistoa peilattiin opinnäytetyössä esitettyyn teoriaosuuteen ja tutkijan omiin havaintoihin.

4.2 Kohderyhmän kuvaus

Tutkimukseen ei saatu suoraan Aurora AI -ohjelmassa työskenteleviä asiantuntijoita. Ohjelma vaikuttaa sen kansallisen tarkoituksensa ja valtion ohjelman mukaisesti kuitenkin laajasti myös muihin julkisen sektorin toimijoihin, joten haastatteluihin valittiin asiantuntijoita näillä kriteereillä.

Tutkittava kohderyhmä muodostui kolmesta senior-tason palvelumuotoilun ja yhdestä data-analytiikan asiantuntijasta. Aineiston pohjalta teoriaa saatiin lisättyä suunnittelijan roolin näkyvyyttä tekoälykehityksessä. Kunkin asiantuntijan työnkuva on hieman erilainen, mikä edesauttoi sitä, että tutkimustuloksista saatiin mahdollisimman moniulotteiset.

4.3 Haastattelun rakenne ja aineiston analyysimenetelmät

Tutkimuksen lähtökohta on kasvattaa tietoutta eettisistä linjauksista, menetelmistä ja toimintatavoista tekoälyyn pohjautuvien palveluiden suunnittelussa. Tavoitteena on kerätä kokonaisvaltaista ymmärrystä Suomen eettisen tekoälykehityksen tilasta, tarkemmin Aurora AI -ohjelmasta ja sille laaditusta eettisestä koodistosta.

Aineistoa analysoitiin käyttäen muutamaa teemahaastattelulle tyypillistä analysointimenetelmää. Äänitallenteena kerätyt aineistot aluksi litteroitiin, jonka jälkeen ne teemoitettiin asiayhteyden mukaisesti. Aineiston pohjalta tietoa heijastettiin haastateltavien keskeiseen vuoropuheluun ja tutkijan omaan ymmärrykseen aiheista.

Haastattelujen kysymysrunko ei ollut identtinen jokaisen haastateltavan kohdalla, sillä kysymykset muotoutuivat henkilökohtaisesti kunkin haastateltavan taustan ja osaamisen mukaan. Teemat ja haastattelukysymykset löytyvät tarkemmin opinnäytetyön liiteosiosta.

5 Tutkimuksen tulokset

Haastattelussa korostuivat yhteinen ajatus siitä, että julkisen puolen toimijoilla ei toistaiseksi ole yhtenäistä pohjaa, jolle tekoälykehitystä kehitetään. Kukin haastateltava oli tietoinen Aurora AI –ohjelman olemassaolosta ja sen lähtökohtaisista toimintaperiaatteista. Sen vaikutuksia ei kuitenkaan ole havaittavissa julkisen sektorin digitaalisten palveluiden kehitystiimeissä. Palvelumuotoilijat kokivat, että keskustelua käydään organisaatioiden johtohenkilöiden ja päättäjien kesken, ei niinkään vielä kehittäjien tai suunnittelijoiden. Aurora AI -ohjelman keskeiset haasteet nähtiin kankeiden järjestelmien ja rahoituksen puolelta. Tulokset kiteytyvät yhteiseen ymmärrykseen siitä, että tekoälykehitys kaipaa monialaista sekä poikkitieteellistä osaamista eettisten ja ihmislähtöisten yhteiskuntatavoitteiden toteutumiseksi. Kehityksessä arvostetaan kokonaisuuksien hahmottamista ja arvojohtamista. Viestimisessä toivottiin läpinäkyvyyttä ja, että eettinen kehityskulma ulottuisi laajasti koko organisaation toimijoille. Yhteisiä eettisiä päämääriä julkisen puolen organisaatioissa ei myöskään ole laadittu tasolle, jossa kehittäjät voisivat niitä konkreettisesti hyödyntää. Suunnittelijoiden epävarmuudet tiedon jakamisesta ja omistajuudesta julkisten palveluiden välillä yhtenevät keskenään. Tekoälykehityksen haasteet nähdään liittyvän henkilötietojen liikuttelun, tiedon omistajuuden ja turvallisuuden kannalta kriittisimpiä

Tekoäly voi lisätä suunnittelijan vastuuta prosessissa, sillä sille tyypilliset autonomisuus ja adaptiivisuus muokkaavat suunnittelijan työnkuvaa. Suunnittelijoiden toimenkuva julkisilla sektoreilla ei vielä toistaiseksi ole muuttunut tekoälykehityksen myötä, mutta se nähtiin varsin todennäköisenä lähitulevaisuudessa. Suunnittelijan roolia pidettiin tärkeänä ja se sijoittuu pitkälti asiakasrajapintaan. Kokonaiskuvan hahmottaminen ja kommunikointi niin kehitystiimin että kansalaisten välillä lukeutuu suunnittelijan kasvavaan työnkuvaan. Kommunikointi edesauttaa myös sitä, että ihmisten ennakkokäsityksiä puretaan tekoälyn ympäriltä, joka auttaa julkisen sektorin palveluita kehittymään. Erityisesti datan arvon ymmärtäminen prosessin eri vaiheissa koettiin keskeisenä. Sen avulla suunnittelija pystyy toimimaan kokonaisuuksien hahmottajana ja eettisten periaatteiden hyödyntäjänä.

Haastateltavat eivät tehneet suoraan töitä tekoälypohjaisten palveluiden parissa, mikä poikkeaa opinnäytetyön alkuperäisestä oletuksesta. Kolme haastateltavaa kokivat, että etiikkaa tulisi lisätä kehitystyöhön ja luoda eettisten periaatteiden yhteinen pohja tiimikohtaisesti. Tällä hetkellä etiikka tai arvot eivät ole keskeisessä osassa kehitysprosesseja. Kuitenkin palvelumuotoilijoille tyypillinen suunnitteluajattelu (; *design thinking*) tuo hyvät valmiudet suunnitella yhteiskunnan edunmukaisia ja ihmislähtöisiä tekoälyratkaisuja. Haastateltavat kokivat, että suunnittelijan rooli tulee olemaan keskeisempi tekoälypohjaisten suunnitteluratkaisuissa. Aurora AI-ohjelman tuottamaan eettiseen koodistoon toivottiin konkretiataason esimerkkejä niiden hyödyntämisestä prosessien eri vaiheissa. Haastateltavat näkivät, että suunnittelijan rooli tekoälykehityksessä voi olla keskeinen, jos suunnitteluajattelun mekanismeja hyödynnetään oikein.

Vaikka Suomi on toistaiseksi vasta määrittelyvaiheessa tekoälyn käyttöön-otossa ja sen eettisten periaatteiden laadinnassa, on se hyvässä vaiheessa Suomen strategian kannalta. Kuten aiemmin todettu, uusimman *McKinsey*-tutkimuksen mukaan Suomi on maailman mittakaavalla kolmantena tekoälykehityksessä. Raportissa parhaat erottaa siitä, että uskalletaan investoida tekoälyyn, nähdään sen potentiaali ja kehitetään tekoälyosaamista. Maat ovat laatineet tarkan kehitystä tukevan rungon (*framework*), hyödyntävät suunnitteluajattelun (*design thinking*) metodeja ja määrittäneen tarkan strategian. (Balakrishnan ym. 2021.) Pysyäkseen tekoälykehityksen kärkikahinoissa, Suomen on siis uskallettava investoida tekoälyyn, osajiin ja mahdollistaa kehitys arvojohtamisen avulla. Suomen julkinen sektori on hyvässä vaiheessa tekoälyn hyödyntämisessä ja eettisten periaatteiden määrittelyssä, ja Aurora AI -ohjelmaan on kirjattu miltei kaikki osiot *McKinsey*-tutkimuksen mukaisesti. Myös Tekoälykiihdyttämö, yhdessä eri toimijoiden kanssa, on luonut raportin, jonka tarkoituksena on toimia tiekarttana Suomen tekoälykehitykselle. Tiekartta sisältää verrattain samankaltaisia määritelmiä aikaisempiin raporttiin verrattuna:

- Vastuullisuus tulee sisällyttää kehitykseen ja se tulee olla rakennettuna kuhunkin kehitysvaiheeseen.
- Avoimuus eri toimijoiden välillä ja tekoälytarinoita tulee rakentaa ja jakaa eteenpäin.

- Suomessa tulisi olla monimuotoinen ja avoin tekoälyekosysteemi. Organisaatioiden tulisi uudistua ja tekoälyn tulisi olla osana jatkuvaa kehitystä.
- Kehitystä tulee kiihdyttää läpinäkyvällä ja avoimella toiminnalla, joita jokainen tekoälykehityksessä mukana oleva taho voi hyödyntää.

(Törnroth 2020.)

6 Pohdinta ja johtopäätökset

Opinnäytetyön tarkoituksena oli tarkastella Suomen tekoälykentän nykytilaa, tekoälyn etiikkaa kansallisen Aurora AI -ohjelman kautta ja sivuta suunnittelijan roolia eettisten tekoälypalveluiden suunnittelussa. Teoriaosuus antaa kattavan kuvan tekoälystä ja sen etiikasta. Tietoa kerättiin erityisesti julkisen sektorin Aurora AI -ohjelman kautta, joka vaikuttaa kansallisesti Suomen yhteiskuntakehitykseen ja digitalisoitumisen toteutumiseen. Suunnittelijan näkökulmaa kerättiin teemahaastattelujen pohjalta ja tuloksia sivuttiin opinnäytetyön teoriaosudessa. Tutkimuksen alkuperäinen suunnitelma oli haastatella Aurora AI -ohjelmassa työskenteleviä suunnittelijoita, joilla on suoraan omakohtaista kokemusta ohjelman eettisistä periaatteista. Lopulta haastatteluihin valikoitui eri sektorien toimijoita, joihin Aurora AI -ohjelma luontaisesti vaikuttaa.

Tutkimustuloksista ilmeni, että tekoäly ei toistaiseksi ole vaikuttanut digitaalisten palveluiden suunnittelijoiden työnkuvaan, jonka takia tuloksia on haasteellista analysoida. Aurora AI -ohjelmasta ollaan jollakin tasolla tietoisia, mutta sen vaikutus projekteihin ei ole merkittävä. Keskeisimpänä havaintona lähes kaikissa haastatteluissa ilmeni tekoälyn huono ymmärrettävyys yksilö- että yhteiskuntatasolla. Julkisen sektorin palveluissa hyödynnettävästä tekoälystä vain harva kansalainen tai suunnittelija ymmärtää täysin sen toimintamekanismit, vaikutukset ja heijastusvaikutukset. Merkittävänä haasteena haastatteluissa nähtiin myös tekoälyssä käytetyn datan eettisyys ja laatu. Suunnittelijoilla tulisi olla jonkin asteen ymmärrys tekoälyssä käytetyn datan merkityksestä, laadusta ja käsittelystä, jotta tekoälyn tuottamat päätökset olisivat yhteiskunnan edunmukaisia ja selitettävissä. Kehitykseen olisi hyödyllistä sisällyttää suunnittelijoille tyypillisiä muotoiluajattelun periaatteita.

Työn aihe oli haasteellinen ihmiseltä, joka ei opiskele filosofiaa, tietojenkäsittelytieteitä tai valtiotieteitä, mutta silti erityisen kiinnostava. Päällimmäinen ajatus tutkittavasta aiheesta on, että toimijat, jotka tekoälyä kehittävät ovat jokseenkin tietoisia siihen vaikuttavista eettisistä periaatteista. Kuitenkin kansalaiset, jotka tekoälyä käyttävät, ovat osittain vielä varsin epätietoisia sen toimintamekanismeista. Ihmislähtöisen yhteiskunnan ja digitaalisten palveluiden peruspilareita on palvella ihmistä. Koen, että tähän lopputulemaan ei ole vielä täysin päästy, mutta siihen ollaan hyvää vauhtia matkalla. Jotta julkinen sektori saa mahdollisimman eettisiä tekoälypalveluja, voisi nähdä, että tekoälyprojektien johtaminen tulisi olla arvojohtamista ja kunkin projektissa työskentelevän tulisi hyödyntää niitä projektien jokaisessa kehitysvaiheessa.

Ammatillinen osaamiseni kehittyi opinnäytetyöprosessin aikana merkittävästi. Keräsin kokonaisvaltaisen ymmärryksen tekoälystä, sen mahdollisuuksista, siihen liittyvät toimintamekanismeista, suunnittelutyössä vaadituista eettisistä periaatteista ja yleisesti nykyhetken tekoälyyn liittyvistä epäkohdista. Koen, että pystyn hyödyntämään opinnäytetyön aikana kerättyä tietoa myös työelämässä, ja oli mukava huomata, että pystyn keskustelemaan tekoälystä ja siihen liittyvistä aihealueista vaivattomasti. Opinnäytetyöprosessi oli ajoittain hyvin haasteellinen muiden velvoitteiden rinnalla. Siihen nähden voin olla tyytyväinen työn sisältöön. Koen kuitenkin, että aiheen parissa pystyisi pysymään huomattavasti pidempään, rajatumminkin ja yksityiskohtaisemmin.

Opinnäytetyö antaa hyvät valmiudet aiheen jatkotutkimukselle. Tässä opinnäytetyössä pohjustettiin tekoälyn sovellettavaa etiikkaa, mutta jatkotutkimuksessa olisi hyödyllistä keskittyä täysin suunnittelijan rooliin eettisessä tekoälykehityksessä, yksinomaan julkiseen tai yksityiseen sektoriin tai datan merkitykseen tekoälypohjaisten palveluiden muotoilussa.

Lähteet

Balakrishnan, Tara, Chui, Michael, Hall, Bryce & Henke, Nicolaus. The state of AI in 2020. McKinsey&Company. Luettavista osoitteesta <<https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020#>>

Borenstein, Jason, Grodzinsky, Frances S., Howard, Ayanna, Miller, Keith W., & Wolf, Marty J. 2021. AI Ethics: A Long History and a Recent Burst of Attention. Luettavissa osoitteesta <<https://www.computer.org/csdl/magazine/co/2021/01/09321834/1qmbkXCazy8>>

Dignum, Virginia 2018. Ethics in artificial intelligence: introduction to the special issue. Ethics Inf Technol.

Euroopan komissio 2021. Tekoäly – huippuosaamista ja luottamusta. Luettavista osoitteesta <https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_fi>

Eskola, Jari & Vastamäki Jaana 2015. Teemahaastattelu: Opit ja opetukset. 4. painos. Jyväskylä: PS-kustannus. (Luettu 15.2.2021)

Aakko, Juhani 2021. Future Care Finland 2021. Juhani Aakko kaivaa sote-dattaa päätöksenteon tueksi. Luettavissa osoitteesta <<https://futurecarefinland.fi/2021/02/juhani-aakko-ja-data-terveydenhuollossa>>

Haataja, Meeri & Latvanen, Marko n.d. AuroraAI-esiselvityshanke - Etiikka työkokonaisuuden suositukset. Luettavissa osoitteesta <<vm.fi/documents/AuroraAI+esiselvityshanke+Etiikka-suositukset.pdf>> (Luettu 23.3.2021)

Hashmi, Ali 2019. AI Ethics: The Next Big Thing In Government - Anticipating the impact of AI Ethics within the Public Sector. Luettavissa osoitteesta: <<https://www.worldgovernmentsummit.org/default-source/deloitte>>

Hulkko, Ville 2021. Etiikkapuhetta: hyvä vai huono tekoäly? Vieraana Silo Ai:n perustaja Ville Hulkko. Recording. (Kuunneltu 10.1.2021)

IEEE n.d. ETHICALLY ALIGNED DESIGN - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. Luettavissa osoitteesta <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_v2.pdf>

Leikas, Jaana 2018. Tekoälyn etiikka – mistä ylipäättään puhutaan? Vtt Research. Luettavissa osoitteesta <<https://www.vttresearch.com/fi/uutiset-ja-tarinat/tekoalyn-etikka-mista-ylipaataan-puhutaan>>

Laitila, Erkki 2018. Ihanteeksi Vastuullinen Tekoäly. Painos. Turku: Metayliopisto. (Luettu 28.1.2021)

Lehtimäki, Pasi 2020. eOppiva ja Valtiovarainministeriö 2020. Kouluttaja. Mistä tekoälyssä on kyse? Luettavissa osoitteessa: <<https://www.eoppiva.fi/koulutukset/mista-tekoalyssa-on-kyse/>> (Luettu 10.2.2020)

Lehtimäki, Pasi 2020. eOppiva ja Valtiovarainministeriö 2020. Tekoälyn toimintaperiaatteet. Luettavissa osoitteessa: <<https://www.eoppiva.fi/koulutukset/mista-tekoalyssa-on-kyse/>> (Luettu 10.2.2020)

Martti, Jyrki 2019. Tekoälyn vallankumous. Tivi. Luettavissa osoitteesta <<https://studio.tivi.fi/sofigate-the-business-technology-company/tekoalyn-vallankumous>>

Merilehto, Antti 2018. Tekoäly - Matkaopas Johtajalle. Painos. Helsinki: Alma Talent Oy. (Luettu 5.8.2020)

Mäntylä, Jussi n.d. Mitä on tekoäly? Tekoäly.info. Luettavissa osoitteessa <https://xn--tekoly-eua.info/mita_tekoaly_on/>

Ojanen, Atte, Oljakka Nea, Sahlgren, Otto, Tuikka, Anne-Marie & Vaiste, Juho 2019. Opas Tekoälyn Etiikkaan. Turku AI Society. Luettavissa osoitteesta <https://aisociety.fi/sites/aisociety.fi/files/opas_tekoalyn_etiikkaan_v1.pdf>

Ollila, Maija-Riitta 2019. Tekoälyn Etiikkaa. Painos. Helsinki: Otava. (Luettu 6.8.2020)

Oxford Insights 2020. AI Readiness Index 2020. Luettavissa osoitteesta: <<https://www.oxfordinsights.com/government-ai-readiness-index-2020>>

Puusa, Anu & Juuti, Pauli 2020. Laadullisen tutkimuksen näkökulmat ja menetelmät. Gaudeamus. (Luettu 15.2.2021)

Rusanen, Anna-Mari, Nurminen, Jukka K., Raisanen, Santeri, Tarkoma, Sasu, Halmetoja, Saara 2020. Ethics of AI 2020. Helsingin Yliopisto. Luettavissa osoitteesta <<https://ethics-of-ai.mooc.fi/22.2021>> (Luettu 2.11.2020)

Sanastokeskus ry 2013. Tietotekniikan termitalkoot. Luettavissa osoitteessa <http://www.tsk.fi/tsk/termitalkoot/fi/hakemistot-267.html?page=get_id&id=ID363&vocabulary_code=TSKTT>

Siau, Keng, Wang, Weiyu 2018. Building Trust in Artificial Intelligence, Machine Learning, and Robotics. CUTTER BUSINESS TECHNOLOGY JOURNAL. Luettavissa osoitteesta <https://www.researchgate.net/publication/324006061_Building_Trust_in_Artificial_Intelligence_Machine_Learning_and_Robotics>

Sparta Consulting 2018. Spartan tiedonhallinnan sanasto. Luettavissa osoitteesta <<https://spartaconsulting.fi/sanasto/>>

Stephen D.Lapan, Marilyn T. Quartaroli & Frances Julia Riemer 2011. Qualitative Research: An Introduction to Methods and Designs. San Francisco. Jossey-Bass. (Luettu 12.3.2021)

Suomen koodikoulu 2018. Johdatus tekoälyyn. Ladattavissa osoitteessa <<https://aoe.fi/api/download/Johdatustekolyyn-1576067361216.pdf>>

Takala, Petri 2018. Kohti jatkuvaa tilannekuvaa, Data-analytiikka yhteiskunnallisen päätöksenteon tukena. Luettavissa osoitteesta <<https://media.sitra.fi/2018/11/29153842/kohti-jatkuvaa-tilannekuvaa.pdf>>

Työ- ja elinkeinoministeriö 2019. Edelläkävijänä tekoälyaikaan. Helsinki. Luettavissa osoitteessa <https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161447/23_19_Tekoalyraportti.pdf> (Luettu 4.3.2021)

Työ- ja elinkeinoministeriö 2017. Suomen tekoälyaika. Helsinki. Luettavissa osoitteessa <https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/80849/TEMrap_41_2017_Suomen_teko%c3%a4lyaika.pdf?sequence=4&isAllowed=y> (Luettu 24.2.2021)

Törnroth, Alexander, Poikola, Antti, Hagstöm, Hanna, Ailisto, Heikki, Huttunen, Heikki, Purosto, Joanna, Himberg, Johan, Mäkinen, Jussi, Kutvonen, Markku, Keski-Äijö, Outi, Floréen, Patrik, Alanen, Pauliina, Myllymäki, Petri, Räsänen, Petri, Erkinheimo Pia, Lankiniemi, Sami, Kaukonen, Sini, Seppälä, Simo, Mucha, Tomasz & Saavalainen, Vera 2020. State of AI in Finland. Finland's Artificial Intelligence Accelerator. Luettavissa osoitteessa <<https://faia.fi/market-research/>> (luettu 19.2.2021)

Valtiovarainministeriö 2020. Asettamispäätös - Kansallinen tekoälyohjelma Aurora-Ai. Luettavissa osoitteesta <<https://vm.fi/documents/10623/16264993/aurora+asettamispaatos+korjattu.pdf/fd7831ba-d4b8-d0a8-e7f2-9cb38407c698/aurora+asettamispaatos+korjattu.pdf?t=1599049693743>> (Luettu 22.3.2021)

Haastattelut

Anonyymi, palvelumuotoilija. Helsinki. Videokonferenssihaastattelu 14.04.2021.

Auer, Tiina, palvelumuotoilija/käyttäjäkokeumus suunnittelija. Helsinki. Videokonferenssihaastattelu 08.04.2021.

Halenius, Laura, johtava asiantuntija. Sitra. Helsinki. Videokonferenssihaastattelu 13.04.2021.

Tiitto, Pirita, toimitilapäällikkö, Helsingin Kaupunki. Helsinki. Videokonferenssihaastattelu 01.04.2021.

Liitteet

Liite 1. Haastatteluteemat

Aineistonkeruu toteutettiin teemahaastatteluina, jotka muotoutuivat kunkin haastateltavan oman kokemuksen ja taustan kautta. Haastattelujen yhdistävät teemat olivat seuraavat:

- Aurora AI-ohjelman vaikutus yhteistyökumppaneihin
- Eettisten periaatteiden (Aurora AI-koodisto) hyödyntäminen työssä
- Eettisten tekoälypalveluiden kehitys tulevaisuudessa
- Suunnittelijan rooli eettisten periaatteiden hyödyntämisessä

Liite 2. Haastattelurunko

- 1) Oletko tietoinen Aurora AI-hankkeesta?
 - a) Mitä mieltä olet tällaisesta sektorirajat ylittävästä verkostotavoitteesta?
 - b) Miten Aurora AI on vaikuttanut / vaikuttaa omaan työnkuvaasi?
- 2) Aurora AI-hanke on luonut kehitystä tukevan eettisen koodiston. Se sisältää seuraavia asioita:
 - a) Ihmiskeskeinen ja ihmisarvoja kunnioittava tarkoitus
 - b) Yhdenvertaisuus ja syrjimättömyys
 - c) Yksityisyydensuoja
 - d) Tiedon laatu ja hallinta
 - e) Vikasietoisuus ja turvallisuus
 - f) Tiedollinen itsemääräämisoikeus
 - g) Omistajuus ja vastuut
 - h) Läpinäkyvyys
- 3) Mitä mieltä olet Aurora AI –ohjelman eettisestä koodistosta?
- 4) Minkälaisiin eettisiin periaatteisiin mielestäsi tulisi keskittyä, kun mietitään kansallisia tekoälypalveluita?
- 5) Onko yrityksessäsi yleistä eettistä koodistoa tekoälykehitykselle?
- 6) Kuinka hyödynnät eettisiä periaatteita työssäsi?
- 7) Mitkä asiat mielestäsi junnaavat jälkijunassa?
- 8) Mikä on suunnittelijan rooli tekoälypalveluiden kehityksessä?
- 9) Mitä suunnittelijan tulee ymmärtää tekoälystä, jotta voidaan tehdä eettisesti parempi yhteiskunta?