

Ekku Sipilä

**AUTOMATISOITUIHIN AUTOIHIN LIITTYVÄT
EETTISET ONGELMAT**



JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2020

TIIVISTELMÄ

Sipilä, Ekku

Automatisoituihin autoihin liittyvät eettiset ongelmat

Jyväskylä: Jyväskylän yliopisto, 2020, 35 s.

Tietojärjestelmätiede, kandidaatintutkielma

Ohjaaja: Halttunen, Veikko

Automatisoidut autot ovat tänä päivänä yksi puhutuimmista kehittyvistä teknologian haaroista. Niillä voisi olla monia positiivisia vaikutuksia liikenteen, yhteiskunnan ja ympäristön näkökulmista. Niiden kehityksessä on kuitenkin vielä monia haasteita, joita tulee ottaa huomioon, ennen kuin laajempi käyttöönotto on mahdollista. Suuri osa näistä haasteista liittyy eettisiin ongelmiin, joita automatisoitujen autojen tuleminen liikenteeseen herättää. Tämän tutkielman tarkoituksena on kartoittaa näitä, automatisoituihin autoihin liittyviä eettisiä ongelmia. Tämän lisäksi taustoitetaan automatisoitujen autojen kehityksen nykytilaa, niiden mahdollisia vaikutuksia, niihin liittyvän lainsäädännön nykyistä tilaa sekä niiden eettiseen tarkastelun kannalta olennaisia teorioita ja etiikan kenttiä. Tuloksina havaittiin, että automatisoituihin autoihin liittyy moninaisia eettisiä ongelmia, joita on haastavaa ratkaista. Tulokset koottiin taulukkomuotoiseen viitekehukseen, joka toimii karkeana yleiskuvana automatisoituihin autoihin liittyvistä eettisistä ongelmista sekä lähtökohtana laajemmalle eettiselle pohdinnalle.

Asiasanat: automatisoidut autot, etiikka, robotiikka

ABSTRACT

Sipilä Ekku

Ethical issues related to Autonomous Cars

Jyväskylä: University of Jyväskylä, 2020, 35 p.

Information Systems, Bachelor's Thesis

Supervisor: Halttunen, Veikko

Automated cars are one of the most discussed topics of emerging technologies in today's world. They could bring many benefits to transportation, society and the environment. However, there are many challenges in their development that need to be addressed before it is possible to deploy them in a larger scale. Many of these challenges are ethical concerns that the deployment of autonomous cars brings up. The aim of this paper is to map the ethical problems associated with automated cars. In addition to this I will briefly introduce the following subjects: the current state of development of autonomous cars, the effects that autonomous cars might have, the current regulations for autonomous cars and the ethical theories and fields that are related to autonomous cars. The results of this study show that there are various ethical issues related to autonomous cars and that these issues are challenging to solve. The results were compiled into a tabular frame of reference, which serves as a rough overview of the ethical issues related to automated cars.

Keywords: autonomous cars, ethics, robotics

TAULUKOT

| | |
|---|----|
| TAULUKKO 1. Autojen automatisaation tasot SAE Internationalin standardia J3016 (2018) mukaillen. | 10 |
| TAULUKKO 2. Automatisoitujen autojen hyödyt ja haitat. | 13 |
| TAULUKKO 3. Koostava esitys automatisoitujen autojen eettisistä ongelmista | 27 |

SISÄLLYS

| | |
|--|----|
| TIIVISTELMÄ | 2 |
| ABSTRACT | 3 |
| TAULUKOT | 4 |
| SISÄLLYS..... | 5 |
| 1 JOHDANTO..... | 6 |
| 2 AUTOMATISOIDUT AUTOT..... | 9 |
| 2.1 Automatisoiduista autoista yleisesti..... | 9 |
| 2.2 Automatisoitujen autojen mahdolliset hyödyt ja haitat..... | 10 |
| 2.2.1 Käyttäjään kohdistuvat hyödyt ja haitat..... | 11 |
| 2.2.2 Muihin kuin käyttäjään kohdistuvat hyödyt ja haitat | 11 |
| 2.3 Automatisoitujen autojen laillisuus | 14 |
| 3 ETIIKKA JA ROBOTIIKKA | 16 |
| 3.1 Etiikasta yleisesti..... | 16 |
| 3.2 Robotiikan etiikka..... | 17 |
| 3.3 Teorioiden soveltuminen tutkielman aihepiiriin | 19 |
| 4 EETTISET ONGELMAT AUTOMATISOIDUISSA AUTOISSA | 21 |
| 4.1 Esiintyneet eettiset ongelmat automatisoiduissa autoissa | 21 |
| 4.1.1 Väistämättömien törmäysten ongelmat..... | 21 |
| 4.1.2 Ongelmat vastuiden jakamisessa..... | 24 |
| 4.1.3 Sekaliikenteestä koituvat ongelmat..... | 25 |
| 4.2 Viitekehys automatisoitujen autojen eettisten ongelmien hahmottamiseksi..... | 26 |
| 4.3 Esiintyneiden ongelmien pohdintaa | 28 |
| 5 YHTEENVETO | 30 |
| LÄHTEET | 32 |

1 JOHDANTO

Automatisoidut autot, usein myös itseohjautuvat autot tai robottiautot, ovat tänä päivänä useasti esillä oleva nouseva teknologia. Ne voisivat tuoda mukanaan monia hyötyjä, kuten esimerkiksi vähentää huomattavasti liikenteessä tapahtuvia onnettomuuksia, pienentää merkittävästi liikenteestä aiheutuvia päästöjä, keventää liikenteen ruuhkautumista, pienentää liikenteestä aiheutuvia kustannuksia sekä vähentää omistettujen autojen määrää (The National Highway Traffic Safety Administrationin, 2015; Pakusch, Stevens, Boden & Bossauer, 2018; Greenblatt & Shaheen, 2015). Yksittäisten käyttäjien näkökulmasta ne voisivat myös vähentää ajamisesta koituvaa stressiä sekä helpottaa liikkumista tarjoamalla uusia liikkuvuuden muotoja (Victoria Transport Policy Institute, 2020; Cohen & Hopkins, 2019). Automatisoituja autoja ei siltikään ole vielä laajamittaisesti markkinoilla (Victoria Transport Policy Institute, 2020) Useissa paikoissa, kuten esimerkiksi yli puolilla Yhdysvallan osavaltioista, niitä ei sallita vielä lainsäädännönkään puolesta edes testikäyttöön (National Conference of State Legislature, 2020). Tämä johtuu siitä, että automatisoitujen autojen kehityksessä on vielä monia haasteita, joita tulee ottaa huomioon. Osa haasteista on autojen teknologioiden kehitykseen liittyviä, mutta suuri osa on eettisiä kysymyksiä automatisoitujen toimintaan ja päätöksentekoon liittyen, joita on välttämätöntä pohtia ennen laajempaa käyttöönottoa.

Gerdes, Lenz, Maurer ja Winner (2016, s. 78-79) antavat esimerkkinä eettisestä ristiriidasta automatisoidun auton käyttäytymisessä seuraavan tilanteen: automatisoitu auto kohtaa pakollisen törmäämisen tilanteen, jossa kääntymällä vasemmalle se tappaa 8-vuotiaan tytön ja oikealle kääntyessään taas 80-vuotiaan rouvan. Miten auton kuuluisi toimia? Seurausoppien näkökulmasta katsottuna voitaisiin väittää loogista olevan törmätä 80-vuotiaan rouvan päälle, koska tällä on todennäköisesti vähemmän elinvuosia jäljellä. Tässä kuitenkin heräisi eettinen ristiriita siitä, että ihmistä arvoitettaisiin jonkin ominaisuuden, kuten iän, perusteella. Tämä on vain yksi esimerkki tilanteista, joita on välttämätöntä tarkastella automatisoitujen autojen kehityksessä ja joiden pohtimisessa on haastavaa löytää yhtä oikeaa ratkaisua. Myös ihmiskuljettaja voisi joutua edellä mainitun esimerkin kaltaiseen tilanteeseen, jolloin kuljettaja joutuisi itse pohtimaan samoja

eettisiä kysymyksiä. Automatisoitujen autojen tapauksessa olennainen ero tulee moraalisen pohdinnan kannalta siinä, että päätös toiminnasta ja moraalisesta arvottamisesta joudutaan tekemään etukäteen jo autoa kehittäessä, eikä reaaliaikaisesti törmäyksen tapahtuessa (Goodall, 2014).

Tutkielmani tarkoituksena on antaa yleiskuva automatisoituihin autoihin liittyvistä eettisistä ongelmista. En pyri tutkielmassani antamaan ratkaisuja kyseisiin ongelmiin vaan pohtia esiintyneitä ongelmia ja tämän pohjalta perustella, miksi eettinen tarkastelu on tärkeää automatisoitujen autojen kehityksessä. Ongelmia esitellessäni pyrin avaamaan niiden monimutkaisuutta esittelemällä niille useita eri ratkaisuvaihtoehtoja ja osoittamalla ratkaisujen välillä esiintyviä ristiriitoja.

Tutkimusongelmani tutkimuskysymys on: "Mitä eettisiä ongelmia automatisoituihin autoihin liittyy?" Pyrin tähän kysymykseen vastaamalla kartoittamaan mahdollisimman laajasti automatisoituihin autoihin liittyviä eettisiä haasteita. Vastatakseni kysymykseen mahdollisimman laajasti saatavien resurssien puitteissa, tutustun aihepiiriin kirjallisuuteen niin hyvin kuin mahdollista ja pyrin ottamaan kirjallisuudessa esiintyneitä näkemyksiä huomioon ongelmien tarkastelussa. Kokoan lopuksi esittämäni ongelmat viitekehykseen, jonka tarkoituksena on toimia automatisoitujen autojen eettistä tarkastelua ohjaavana ja tukevana työkaluna. Viitekehysten tarkoituksena ei siis ole olla ratkaisumalli nousseille eettisille ongelmille.

Tutkielma on toteutettu itsenäisenä kirjallisuuskatsauksena. Pyrin materiaalien hankinnassa keräämään ja käyttämään mahdollisimman luotettavaa sekä aihepiiriin soveltuvaa lähdemateriaalia. Lähdeaineistoa on pääasiassa kerätty JYKDOK- sekä Scopus-hakupalvelujen kautta. Vaihtoehtoisesti olen käyttänyt Google Scholar -hakupalvelua. Pääasiallisina hakusanoina käytin kyseisissä hakupalveluissa "autonomous vehicles" ja "ethics" -termejä tai näitä vastaavia sanoja. Tämän jälkeen valitsin esiintyneistä lähteistä niitä, jotka sopivat minun tutkielmani kohdistukseen ja olivat riittävän laadukkaita käytettäväksi. Laaduntarkastuksessa hyödynsin esimerkiksi Julkaisufoorumia. Tätä enemmän kuitenkin arvioin itse laadukkuutta arvioimalla niissä käytetyn argumentoinnin pätevyyttä.

Johdannon lisäksi tutkielma sisältää kolme sisältölukua sekä yhteenvedon. Ensimmäisessä sisältöluvussa avaan automatisoitujen autojen määritelmää ja niiden kehityksen nykytilaa. Avaan tässä luvussa myös niiden mahdollisia vaikutuksia sekä niitä koskevan laillisuuden tämän hetkistä tilaa maailmalla. Toisessa sisältöluvussa esittelen tutkielman kannalta olennaisia etiikan teorioita ja automatisoitujen autojen kannalta olennaisimman soveltavan etiikan kentän, robotiikan etiikan, yleisiä periaatteita sekä niiden toimivuutta automatisoitujen autojen eettisten ongelmien käsittelyyn. Viimeisessä sisältöluvussa vastaan tutkielman tutkimuskysymykseen eli esittelen, mitä eettisiä ongelmia automatisoituihin autoihin liittyy. Pohdin tässä luvussa esitettyjä ongelmia esimerkkien ja eri etiikan näkökulmien kautta sekä kokoan ongelmat lopuksi viitekehykseen. Yhteenvedossa tiivistän tutkielman olennaisimmat sisällöt ja tehdyt havainnot sekä

pohdin, miten hyvin tutkielmassa onnistuttiin vastaamaan tutkimuskysymykseen ja mihin tutkielmassa jää vastaamatta.

2 AUTOMATISOIDUT AUTOT

Esittelen tässä luvussa automatisoidut autot yleisesti. Avaan ensin automatisoitujen autojen määritelmää sekä niiden kehityksen nykyistä tilannetta. Tämän jälkeen esittelen karkeasti automatisoitujen autojen tuomia mahdollisia hyötyjä ja haittoja. Lopuksi avaan lyhyesti automatisoitujen autojen lainsäädännön nykyistä tilannetta maailmalla.

2.1 Automatisoiduista autoista yleisesti

Automatisoidut autot määritellään useimmiten niiden automatisaation tasojen mukaan. SAE Internationalin J3016 standardissa (2018) nämä on määritelty kuu-teen eri tasoon vaihdellen automatisaation puuttumisesta (taso 0) täyteen automatisaatioon (taso 5). Nämä tasot ovat tarkemmin esiteltynä Taulukossa 1. Tässä kirjallisuuskatsauksessa käsittelen automatisoituja autoja näiden tasojen pohjalta.

Kyseisen määritelmän ensimmäisessä ja toisessa tasossa puhutuilla tukijärjestelmillä tarkoitetaan nykyautoista jo löytyviä kaistalla pysymistä sääteleviä teknologioita sekä mukautuvaa vakionopeudensäädintä. Normaali, vain tiettyä nopeutta ylläpitävä vakionopeudensäädin, ei siis kuulu näihin tukijärjestelmiin. Ero ensimmäisellä ja toisella tasolla on siinä, että ensimmäisen tason autoissa on vain toinen näistä, kun taas toisen tason autoista löytyy molemmat. Kolmannessa tasossa auto pystyy ohjautumaan täysin itsenäisesti tietyissä olosuhteissa. Tästä esimerkkinä on moottoriteillä täysin itseohjautuva auto. Olennaisena erona tässä neljänteen tasoon on, että ihmiskuljettajan täytyy kuitenkin ottaa ohjat autosta pyydettyä, kun taas neljännessä tasossa tämä vaatimus häviää. Esimerkki neljännessä tasosta on vain tietyllä alueella toimiva kuskiton taksi. Viidennessä tasossa automatisoitu auto kykenee toimimaan itsenäisesti kaikissa olosuhteissa. (SAE International, 2018.)

Tällä hetkellä automatisoidut autot ovat kehitys- ja testivaiheissa. Moni autovalmistaja tarjoaa tasojen 2 ja 3 teknologioita, kuten mukautuvia vakionopeudensäätimiä ja automatisoitua taskupysäköintiä. Teslan Autopilot esimerkiksi tarjoaa automatisoitua ohjausta ja kiihdytystä tietyissä olosuhteissa. Usealla yrityksellä on myös tason 4 pilottiprojekteja, joissa autoja testataan tietyissä olosuhteissa. (Victoria Transport Policy Institute, 2020.) Johtavana tekijänä valmistajista voidaan nähdä Waymo, jonka neljännen tason autoa on testattu ajossa 10 miljoonaa mailia teillä ja 7 miljardia mailia simulaatiossa (Center for Sustainable Systems, 2019).

TAULUKKO 1. Autojen automatisaation tasot SAE Internationalin standardia J3016 (2018) mukailten.

| Myös ihmiskuljettaja havainnoi ympäristöä | |
|--|--|
| Taso 0 - ei automaatiota | Ihmiskuljettaja on täysin vastuussa auton kontrolloimisesta kaikissa tilanteissa. |
| Taso 1 - kuljettajan tukijärjestelmät | Yksi kuljettajan tukijärjestelmä, joka hoitaa joko ohjauksen tai kaasuttamisen/jarruttamisen tietyssä tilanteessa ajoympäristön tiettyyn olosuhteeseen perustuen. Oletus, että ihminen hoitaa kaikki muut dynaamiset ajotehtävät tarvittaessa. |
| Taso 2 - osittainen automaatio | Järjestelmä hoitaa tietyssä tilanteessa sekä ohjauksen että kaasuttamisen/jarruttamisen ajoympäristön tiettyyn olosuhteeseen perustuen. Oletus, että ihminen hoitaa kaikki muut dynaamiset ajotehtävät tarvittaessa. |
| Ihmiskuljettajan ei ole välttämätöntä havainnoida ympäristöä | |
| Taso 3 - ehdollinen automaatio | Järjestelmä havainnoi ympäristöä ja hoitaa kaikki dynaamiset ajotehtävät tietyissä olosuhteissa, sillä oletuksella, että ihminen ottaa ohjat pyydettyäessä. |
| Taso 4 - korkea automaatio | Järjestelmä havainnoi ympäristöä ja hoitaa kaikki dynaamiset ajotehtävät, vaikka ihminen ei reagoisi pyyntöön ottaa ohjat. Ihmisellä on kuitenkin vielä mahdollisuus ottaa ohjat. |
| Taso 5 - täysi automaatio | Täysi automaatio. Järjestelmä pystyy ajamaan itsenäisesti kaikissa olosuhteissa, eikä ihmisellä ole tarvetta ottaa ohjia missään tilanteessa. |

Yllä on koostettu tiivistys tasoista ja niiden välisistä eroista. Suurin ero on tasojen 2 ja 3 välillä. Tasossa 3 auton täytyy pystyä havainnoimaan ympäristöään kokonaisvaltaisesti ja tehdä päätöksiä dynaamisesti tämän mukaan, kun taas tasossa 2 dynaamisten päätösten tekeminen on vielä ihmisen vastuulla. (SAE International, 2018.) Suurin osa automatisoitujen autojen eettisistä ongelmistakin esiintyy vasta, kun siirrytään tasolta 2 ylöspäin (Borenstein, Herkert & Miller, 2017).

Automatisoitujen autojen trendi ja kehitys on kovassa vauhdissa, mutta siinä on vielä monia haasteita edessä ennen kuin korkeampien tasojen automatisointia saadaan yleisempään käyttöön. Victoria Transport Policy Institute (2020) sanoo automatisoitujen autojen teknologioiden tarvitsevan käydä läpi useampia vaiheita ennen kuin ne voivat tulla laajemmin saataville sekä luotettaviksi ja tarpeeksi edullisiksi.

2.2 Automatisoitujen autojen mahdolliset hyödyt ja haitat

Tässä alaluvussa käsittelemme hyötyjä ja haittoja, joita automatisoiduilla autoilla voitaisiin mahdollisesti saada aikaan. Useat automatisoiduista autoista

saatavista mahdollisista hyödyistä nostavat esiin myös uusia ongelmia tai haittoja, joten käsittelen esiintyviä kohtia molemmista näkökulmista. Jaottelen nämä auton käyttäjään kohdistuviin ja muihin kuin auton käyttäjään kohdistuviin näkökulmiin. Muita kuin käyttäjään kohdistuvia näkökulmia ovat esimerkiksi ympäristön ja yhteiskunnan kannalta olennaiset hyödyt ja haitat. Kokoan lopuksi esiteltyt hyödyt ja haitat taulukkoon.

2.2.1 Käyttäjään kohdistuvat hyödyt ja haitat

Automatisoidussa autossa henkilön ei tarvitsisi olla ohjaamassa autoa, joten se voisi toimia esimerkiksi työpisteenä tai nukkumapaikkana, mikä voisi vähentää stressiä ja lisätä tuottavuutta (Victoria Transport Policy Institute, 2020). Toisaalta ne voisivat tuoda mukanaan myös uusia stressin aiheita. Yllä mainittu etu vaatisi tason 5 automaation, mikä toisi mukanaan ahdistuksen siitä, pääseekö automatisoiduilla autoilla tiettyyn paikkaan vai jääkö auto esimerkiksi sääolosuhteiden takia matkan varrelle. Tämän stressin aiheen takia monet kuluttajat saattaisivat jättää täyden automatisaation auton ostamatta. (Grush, Niles & Baum, 2016.)

Toiseksi, automatisoidut autot voisivat lisätä liikkuvuuden saatavuutta henkilöillä, jotka eivät omista tai kykene ajamaan autoa, kuten vanhusten, lasten ja vammaisten keskuudessa (Cohen & Hopkins, 2019). Automatisoidut autot voisivat erityisesti edistää kysynnän mukaan käytettäviä liikkuvuuden muotoja, kuten kimpakyytejä ja yhteiskäyttöautoja, ja saada tämän avulla ihmisiä vaihtamaan yksityisistä autoista näihin. Yksilöillä tämä voisi vähentää liikkuvuuden kustannuksia ja aikaa sekä vähentää ajamisen stressiä. (Greenblatt & Shaheen, 2015.) Toisaalta, vaihdos kimpakyyteihin ja yhteiskäyttöautoihin voi tapahtua myös julkisen liikenteen muodoista, jolloin jo käytettyjen julkisten liikenteen muotojen hyvät puolet, kuten esimerkiksi ekologisuus, vähenisi. Pakusch ym. (2018) esittävät näiden epätoivottujen vaihdosten määrän voivan olla jopa kaksinkertainen yksityisistä autoista vaihtaviin verrattuna. Tämä voisi huonontaa muiden julkisen liikenteen muotojen saatavuutta ja näin ollen myös huonontaa liikkuvuuden saavutettavuutta monien keskuudessa.

Perinteisistä autoista eroavasti, automatisoidut autot toisivat mukanaan yksityisyyteen liittyviä ongelmia. Automatisoitujen autojen teknologiset vaatimukset datan keräämisen osalta tekevät niistä helposti seurattavia. Tämä myös herättää kysymyksen siitä, mitä kaikella kerätyllä datalla tehdään ja kuka niihin pääsee käsiksi. Autoissa kulkeva suuri henkilökohtaisten tietojen määrä myös houkuttelee hakkereita pyrkimään päästä tietoihin käsiksi. (Collingwood, 2017.)

2.2.2 Muihin kuin käyttäjään kohdistuvat hyödyt ja haitat

Yhtenä merkittävimmistä automatisoitujen autojen mahdollisuuksista on liikenteen turvallisuuden parantaminen. The National Highway Traffic Safety Administrationin (NHTSA) (2015) tekemän katsauksen mukaan 94 % USA:ssa tapahtuneista onnettomuuksista vuosien 2005 ja 2007 välillä johtui ihmisen virheestä.

The New York State Department of Motor Vehiclesin (NYSDMV) (2012) raportissa sama luku oli 78 %. Kummassakin näistä ihmisen tekemä virhe kattaa selvästi suurimman osan onnettomuuksiin johtavista tekijöistä. Automatisoiduilla autoilla olisi mahdollisuus vähentää onnettomuuksia poistamalla tämä inhimillinen tekijä yhtälöstä. Samalla ne kuitenkin toisivat mukanaan uusia riskejä, kuten esimerkiksi ihmisten ohjaamien ja automatisoitujen autojen välisen koordinaation keskinäisen sopimattomuuden (Nyholm & Smids, 2018), jotka tulisi ottaa huomioon suunnittelu- ja käyttöönottoprosessissa.

Automatisoiduilla autoilla voisi olla myös huomattavaa vaikutusta ympäristöön ja päästöihin. Suoria positiivisia vaikutuksia ympäristöön tulisi automatisoitujen autojen optimaalisemmasta toiminnasta sekä yksittäisinä autoina että osana liikennettä. Automatisoitujen autojen toiminta olisi tehokkaampaa polttoaineen kulutuksen ja päästöjen osalta. Ne myös mahdollistaisivat niin sanotun ”yhdistetyn ajamisen” eli datan jakamisen autojen kesken liikenteessä. Tämä optimoisi liikenteen kulkua, vähentäisi ruuhkautumista ja pienentäisi päästöjä muun muassa vähentämällä ”pysähdy ja mene”-ajamista. Automatisoidulla ajamisella olisi mahdollisuus keventää kulkuneuvoista aiheutuvia päästöjä merkittävästi. (Spieser, Treleaven, Zhang, Frazzoli, Morton & Pavone, 2014; Bischoff & Maciejewski, 2016; Fagnant & Kockelman, 2015; Brown, Gonder & Repac, 2014; Greenblatt & Saxena, 2015, viitattu lähteessä Pakusch ym. 2018.) Monet ”yhdistetyn ajamisen” tuomista eduista tosin perustuvat siihen oletukseen, että koko liikenne olisi automatisoitua, mikä ei ole odotettavaa vielä vähään aikaan.

Epäsuoria vaikutuksia ympäristöön automatisoiduista autoista tulee matkustamismuodon preferenssien muuttumisella ja niiden tuomien tarkoituksettomien vaikutusten kautta. Kuten aiemmin totesin, automatisoidut autot voisivat edistää kysynnän mukaan käytettäviä liikkuvuuden muotoja, kuten kimppekäyttöä ja yhteiskäyttöautoja, mutta vaihdos näihin voisi tapahtua myös julkisen liikenteen muodoista. Vaihdos kysynnän mukaan käytettäviin liikkuvuuden muotoihin voisi vähentää yksityisten autojen omistuksia ja niiden tuottamia päästöjä, mutta tarkoitukseton vaihdos muista julkisen liikenteen muodoista taas saisi aikaan vastakkaista vaikutusta. Muita huomioon otettavia asioita ympäristön kannalta ovat automatisoiduista autoista tulevat ”heijastusvaikutukset”, kuten esimerkiksi ajettujen kilometrien lisääntyminen, mitkä voisivat vähentää autojen tuomia positiivisia vaikutuksia (Pakusch ym., 2018). Thomopoulos ja Givoni (2015) sanovat automatisoitujen autojen tuovan uusia mahdollisuuksia julkiselle liikenteelle ja liikkumisen ympäristövaikutusten edistämiseksi. He mainitsevat tämän kuitenkin edellyttävän tarkoituksettomien vaikutusten huomioon ottamista sekä sitä, että teknologisten muutosten rinnalla tapahtuisi yhteiskunnallinen muutos siihen suuntaan, että julkinen ja jaettu liikenne nähtäisiin yksityisiä autoja parempana vaihtoehtona.

Edellä mainitut vaikutukset tuovat mukanaan myös muita seurauksia. Yksityisten autojen omistusten väheneminen ja kysynnän mukaan käytettäviin liikkuvuuden muotoihin siirtyminen pienentäisi parkkipaikkojen tarvetta (Greenblatt & Shaheen, 2015). Niistä aiheutuva sujuvampi ja tehokkaampi liikenne voisi myös vähentää ajoteiden kustannuksia. Toisaalta automatisointi myös veisi

työpaikkoja henkilöiltä, jotka ajavat kulkuneuvoja ammatikseen. Automatisoituihin autoihin liittyvät optimistiset ennusteet voivat myös saada aikaan muiden, vaihtoehtoisten liikkuvuuden muotojen, kehittämisen laiminlyöntiä. (Victoria Transport Policy Institute, 2020.)

TAULUKKO 2. Automatisoitujen autojen hyödyt ja haitat.

| | Hyödyt | Haitat/ongelmat |
|------------------------------------|--|---|
| Käyttäjään kohdistuvat | Stressin väheneminen ja tuottavuuden lisääntyminen. | Uusien stressin aiheiden syntyminen. |
| | Liikkuvuuden saatavuuden parantuminen. | Olemassa olevien liikkuvuuden muotojen saatavuuden, kuten bussien ja junien, heikkeneminen. |
| | Liikkumisen kustannusten väheneminen. | Uusien tietoturva- ja yksityisyyteen liittyvien ongelmien syntyminen. |
| Muihin kuin käyttäjään kohdistuvat | Liikenneonnettomuuksien väheneminen ja liikenneturvallisuuden paraneminen. | Uusien riskien esiintyminen. |
| | Kysynnän mukaan käytettävän liikkuvuuden, kuten kimpakyytien ja yhteiskäyttöautojen, suosion nousu ja vaihdosnäiden käyttämiseen. Tämä kontribuoi seuraaviin kolmeen esitettävään hyötyyn. | Tarkoituksettomat vaikutukset, kuten vaihdos muista julkisen liikenteen muodoista automatisoituihin kimpakyyteihin sekä ajettujen kilometrien nouseminen. |
| | Pienempi yksityisten autojen omistusten määrä. | Työpaikkojen väheneminen työkseen autoa ajavilta henkilöiltä. |
| | Päästöjen väheneminen. | Päästöjen mahdollinen nouseminen tarkoituksettomien vaikutusten seurauksena. |
| | Ruuhkautumisen helpottuminen, liikenteen tehostuminen ja parkkipaikkojen tarpeen väheneminen. | Mahdollisten muiden liikennettä ja liikkuvuutta edistävien vaihtoehtojen kehittämisen huomiotta jättäminen. |
| | Liikenteen aiheuttamien kustannusten väheneminen. | |

Yllä olen koontanut kahden edellisen kappaleen hyödyt ja haitat taulukkoon helpottaakseni niiden hahmottamista. Hyödyt ja haitat on jaoteltu kappaleiden mukaan auton käyttäjään kohdistuviin sekä muihin kohdistuviin näkökulmiin.

2.3 Automatisoitujen autojen laillisuus

Automatisoitujen autojen laillisuus vaihtelee paljon paikoittain ja alueittain. Yhdysvalloissa ei ole mitään valtakunnallisia rajoitteita automatisoiduille autoille vaan jokainen osavaltio saa määritellä rajoitteensa itse. Tästä syystä lakirajoitteet vaihtelevat hyvin paljon osavaltioiden kesken ja mitään selkeää yleistä linjaa ei ole. (Fagnant & Kockelman, 2015.) Tällä hetkellä Yhdysvalloissa 37 osavaltiota on asettanut lainsäädäntöä tai tehnyt toimeenpanoasetuksia automatisoituihin autoihin liittyen. Näistä osavaltioista 8 sallii autojen testikäytön ja 11 täyden käyttöönoton tiettyjen kriteerien täytyessä. Nämä kriteerit ja niiden yksityiskohtaisuus vaihtelevat hyvin paljon osavaltioiden kesken. Ne voivat vaihdella autojen tietyistä teknologisista vaatimuksista ja ihmiskuljettajan läsnäolon pakollisuudesta esimerkiksi Alabamassa määriteltyyn vaatimukseen siitä, että autoilla on oltava minimissään 2 miljoonan dollarin vastuuvakuutus. Autovalmistajien kannalta hyvin olennaisena vaatimuksena on se, kuuluuko valmistajien ottaa vastuu autojen aiheuttamista vahingoista. Tämäkin vaihtelee osavaltioiden kesken. Useimmilla valtioista näiden kriteerien täyttymisen lisäksi autojen valmistajien tulee saada erillinen lupa autojen testi- tai julkista käyttöönottoa varten. (National Conference of State Legislatures, 2020.)

Yhdysvalloista poiketen, Euroopan Unioni on määritellyt yhteisiä linjauksia jäsenmailleen automatisoitujen autojen laillisuuden suhteen. Euroopan Komissio on vuonna 2019 asettanut ohjenuorat, jotka sisältävät 38 kohtaa ja joita jokaisen EU:n jäsenmaan tulee toteuttaa automatisoitujen autojen kehityksen ja käyttöönoton suhteen. Ensimmäisessä näistä kohdista määritellään, että automatisoitu järjestelmä saa korvata kuljettajan kaiken ajotoiminnan osalta niissä olosuhteissa, joissa järjestelmän on oletettu pystyvän toimimaan. Tämä siis sallii automatisoitujen autojen liikenteeseen pääsemisen. Muuten kohtien esittämät vaatimukset vaihtelevat esimerkiksi ihmisen läsnäolon pakollisuudesta kyberturvallisuuden vaatimuksiin. (European Commission, 2019.) Näiden kohtien ohella EU:n jo olemassa oleva lainsäädäntö sopii sellaisenaan automatisoitujen autojen testaamiseen. Vuonna 2018 EU:n uusima yleisesti autojen kehitykseen ja käyttöönottoon liittyvä laki, jonka mukaan uusien autojen pitää läpäistä tyyppi hyväksyntä päästäkseen markkinoille, sopii myös automatisoitujen autojen testaamiselle. EU:n jäsenmaat voivat määritellä omat lakinsa automatisoitujen autojen osalta, kunhan ne eivät ole ristiriidassa edellä mainittujen lakien ja ohjenuorien kanssa. (European Commission, 2018.) Esimerkiksi Saksassa tason 3 ja tason 4 automatisoitujen autojen testaaminen on tällä hetkellä sallittua, mikäli ne täyttävät tietyt turvallisuuden kriteerit ja valmistajat saavat testaamisen erillisen luvan (Dentons, 2020). Myös Suomessa jo nykyinen tieliikennelainsäädäntö mahdollistaa automatisoitujen autojen kokeilut, mutta niihin täytyy saada Traficomilta erillinen lupa (Traficom, 2018).

Aasiassa lainsäädäntö vaihtelee paljon valtioiden välillä eikä yleistä linjausta näiden kesken ole. Esimerkiksi Australiassa ja Kiinassa lainsäädäntö tällä hetkellä sallii automatisoitujen autojen testikäytön liikenteessä, mikäli ne

täyttävät tietyt kriteerit ja saavat niistä vastaavalta taholta siihen luvan. Nämä turvallisuuden kriteerit vaihtelevat, mutta olennaisena yhtenäisenä tekijänä molemmilla näistä maista on, että autossa on oltava kuljettaja mukana testiajossa. (Dentons, 2020.) Aasian osalta Singapore on yksi maista, joka on lainsäädännön puolesta jo pitkällä kehityksessä. Vuonna 2017 maan lainsäädäntöön tehdyn muutoksen myötä, Singaporessa myönnetään lupia jo laajemmallekin käytölle kuin vain testaamiseen. Siellä vaatimus ihmisen läsnäololle autossa ei myöskään ole pakollinen eli lainsäädännön puolesta täysi automatisaatio on jo mahdollista. (Parliament of Singapore, 2017.)

3 ETIIKKA JA ROBOTIIKKA

Tässä luvussa käsittelen etiikkaa ja sen eri teorioiden soveltumista tutkielman aihepiiriin. Avaan ensin yleisiä tutkielman kannalta olennaisia, etiikan teorioita. Soveltavan etiikan kentistä avaan tarkemmin robotiikan etiikkaa, sillä sen katta-miin teknologioihin kuuluu automatisoidut autot, mikä tekee siitä olennaisim-man soveltavan etiikan kentän tämän tutkielman kannalta. Lopuksi tarkastelen teorioiden soveltumista tämän tutkielman eettiseen pohdintaan.

3.1 Etiikasta yleisesti

Etiikka on moraalisuutta tutkiva filosofinen tiede. Se on rationaalista tarkastelua ihmisten moraalisisista uskomuksista ja käyttäytymisestä. (Quinn, 2015, s. 79.) Etiikka on alana monitieteellinen ja se koostuu kolmesta pääsuuntauksesta, jotka ovat metaetiikka, normatiivinen etiikka ja soveltava etiikka (Marturano, 2002). Avaan seuraavaksi näitä kolmea suuntausta tarkemmin.

Metaetiikassa on kyse etiikan semanttisista ja loogisista käsitteistä. Se käsittelee yleistä eettistä analyysiä sekä tiettyjen normatiivisten teorioiden analysoimista ja niiden eettisen perustelun testaamista sekä eettisten käsitteiden selventämistä. (Marturano, 2002; Hopkins, 1997, s. 253.) Tässä tutkielmassa tarkastelen eettisiä ongelmia konkreettisemmin normatiivisten teorioiden näkökulmista, joita avaan lisää seuraavassa luvussa. Metaetiikan rooli tulee kuitenkin mukaan näiden teorioiden vertailussa.

John Hopkins (1997, s. 253) määrittelee normatiivisen etiikan olevan käsittelevän olennaisia kysymyksiä, kuten esimerkiksi, että mitkä lopputulemat ovat hyviä, mitkä teot ovat oikeita, mitkä säännöt ovat oikeudenmukaisia sekä mistä teoista toimijan pitäisi olla vastuussa. Normatiivisen etiikan näkemysten tavoitteena on siis tarkastella, mikä toiminta on oikein tai väärin ja mihin kuuluisi toiminnalla pyrkiä. Merkittävimmät normatiivisen etiikan suuntaukset ovat hyve-etiikka, deontologia ja seurausetiikka (Stahl, Timmermans & Mittelstadt, 2016). Keskityn tässä tutkielmassa pääasiassa tarkastelemaan eettisiä ongelmia velvollisuusetiikan ja seurausetiikan oppien näkökulmista, mutta määrittelen myös hyve-etiikan lyhyesti.

Hyve-etiikka pohjautuu antiikin Kreikan ajalta Aristoteleen opeista. Muista normatiivisista etiikan linjauksista poiketen, siinä oikeaa toimintaa ei pyritä määrittämään tiettyjen ohjeiden tai tavoitteiden kautta vaan enemmänkin joidenkin perustavanlaatuisien hyveiden kautta. Kun toiminta pohjautuu näihin hyveisiin, tulee oikeat teot luonnostaan ja väärin valintojen miettiminen saa jo itsessään aikaan epämiellyttävää oloa. Hyveiden kautta toimiville ihmisille oikeat teot tulevat kuin toiseksi luonnoksi. (Quinn, 2015, s. 117.)

Deontologinen etiikan (kutsutaan myös velvollisuusetiikka nimellä, juontuu Kreikan sanasta "deon", joka tarkoittaa velvollisuutta) moraalinen arvioiminen käsittelee käytännön toimintaa tärkeiksi katsottujen velvollisuuskiennäkökulmasta. Siinä tekojen moraalinen laatu määritellään sen mukaan, mikä toimijan intentio on ollut niitä tehdessä. Velvollisuusetiikan oppeja kiteyttää hyvin kuuluisimman deontologian suuntauksen edustajan, saksalaisen filosofin Immanuel Kantin, muodostama kategorinen imperatiivi. (Stahl ym., 2016). Sen mukaan ihmisten kuuluisi toimia aina semmoisten moraalisten sääntöjen mukaan, joiden haluaisi olevan universaaleja lakeja. (Quinn, 2015, s. 96). Kant myös mainitsee, että sillä ei ole väliä, mitä haluaa tehdä vaan sillä, mitä kuuluisi tehdä (Quinn, 2015, s. 96).

Seurauseetiikassa painotetaan toiminnan eettisen arvon tulevan ilmi vasta sen seurauksissa (Stahl ym., 2016). Tunnetuimpia seurauseetiikan suuntauksia on utilitarismi. Se jaotellaan vielä kahteen eri lajiin: tekoutilitarismiin ja sääntöutilitarismiin. Tekoutilitarismi perustuu John Stuart Millin ja Jeremy Benthamin esittämään teoriaan, joka on täysi vastakohta deontologialle. Siinä teon hyvyys määritellään sen mukaan, mitkä sen seuraukset ovat. Sen mukaan ihmisen tulisi aina tavoitella niitä tekoja, jotka tuottavat eniten hyvää kaikille sen toimijoille kokonaisuudessaan. Sääntöutilitarismi eroaa tästä siinä, että tarkasteltavana kohteena on tiettyjen sääntöjen pohjalta toimiminen eikä yksittäiset teot. Sääntöutilitarismissa siis ideana on se, että meidän pitäisi noudattaa sellaisia moraalisia lakeja, jotka tuottaisivat eniten hyvää yleisesti, mikäli kaikki noudattaisivat niitä. (Quinn, 2015.) Tästä eteenpäin puhun näistä kahdesta teoriasta yhtenä utilitarismin terminä, joilla tarkoitan yleisesti hyvien seurauksien toiminnan tavoittelemisen ideologiaa.

Soveltavassa etiikassa tarkastellaan käytännön ongelmia tietyllä tieteenalalla. Siinä sovelletaan nimensä mukaisesti monia eri etiikan teorioita, tavoitteena päästä parempaan ymmärrykseen ja saada eettisiä ohjenuoria tietyille aloille, jotka vaativat eettistä tarkastelua. Esimerkkejä soveltavasta etiikasta ovat bioetiikka, ympäristöetiikka, teknologian etiikka ja sodan etiikka. (Stahl ym., 2016; Marturano, 2002.) Soveltava etiikka itsessään ei siis tarjoa tiettyjä ohjeita tai tavoiteltavia asioita, vaan nämä riippuvat sovelletusta kentästä ja sen ongelmista. Tämän tutkielman kannalta olennaisin soveltuvan etiikan kenttä on robotiikan etiikka, sillä sen aihepiireisiin kuuluu automatisoidut autojen etiikka (Tzafestas, 2018). Avaan seuraavassa luvussa robotiikan etiikkaa tarkemmin.

3.2 Robotiikan etiikka

Robotiikan etiikka eli roboetiikka (engl. "roboethics") on tuore etiikan suuntaus, joka on vasta viime vuosien aikana vakiinnuttanut asemaansa soveltuvan etiikan kentällä (Veruggio, Solis & Van der Loos, 2011). Roboetiikkaa terminä on käytetty ensimmäisen kerran 2004 pidetyssä konferenssissa Gianmarco Veruggion toimesta (Tzafestas, 2018). Tzafestas (2018) listaa roboetiikan piiriin kuuluvan lääketieteellisten robottien etiikan, avustavien robottien etiikan, sosiaalisten

robottien etiikan, automatisoitujen autojen etiikan, sodankäynnin robottien etiikan sekä kyborgien etiikan. Roboetiikka kokonaisuudessaan siis haarautuu useampaan alemman tason, itsenäisenäkin tarkasteltaviin, etiikan piireihin. Näistä olennaisimpana tämän tutkielman kannalta on luonnollisesti automatisoitujen autojen etiikka.

Tarve roboetiikalle omana soveltavan etiikan kenttänä sekä sille, etteikö yleiset etiikan periaatteet riittäisi vastaamaan kyseisen alan eettisiin kysymyksiin, on ollut nousevana väittelyn aiheena viime vuosien aikana (Veruggio ym., 2011). Veruggio ym. (2011) esittävät muutamia argumentteja roboetiikan puolesta omana soveltavan etiikan kenttänä. Ensimmäiseksi he nojautuvat soveltavan etiikan yleiseen määrittelyyn siinä, että roboetiikan tarkoituksena on, kuten esimerkiksi bioetiikan, pyrkiä vastaamaan uusiin kysymyksiin, joita nousee uutta tieteen haaraa tai teknologiaa kehittäessä. Toisena argumenttina he esittävät ohjesääntöjen, joita jokainen robotti vaatii ja joilla pyritään ohjaamaan niitä oikeaan toimintaan, olevan nimenomaan roboetiikan kentän tuotoksia. Robotit ovat koneita, joiden päätöksiä ohjaavat niiden kehittäjät siihen asti, kunnes niiden kehitys on niin pitkällä, että niillä on oma tietoisuus ja vapaus tehdä päätöksiä. Näin edistyneen kehityksen taso on tällä hetkellä vain spekuloinnin tasolla. Tämän takia siis robottien tekojen ja toiminnan moraalinen vastuu on niiden kehittäjillä. Tätä vastuuta kehittäjät kantavat antamiensa ohjesääntöjen, roboetiikan kentän pohdinnan tuotosten, muodossa. (Veruggio ym., 2011.)

Yksityiskohtaisemmat eettiset ongelmat ja kysymykset vaihtelevat paljon roboetiikan haaran mukaan. Yleisellä tasolla katsottuna kuitenkin eettisiä huolia herättävät muun muassa yksityisyyden väheneminen, kysymykset siitä, kuka tai ketkä ovat vastuussa robottien toiminnasta sekä turvallisuuteen liittyvät huolet, joita herää, kun robotit toimivat ihmisten seurassa (Tzafestas, 2018). Wallach, Allen ja Smith (2008) esittelevät näiden ongelmien ratkaisemiseksi kolme eri lähestymistapaa. Ensimmäisenä näistä on ylhäältä-alas-tapa, jossa halutut ohjesääntöt sulautetaan robotin järjestelmään sen ohjailemiseksi. Toisena taas on alhaalta-ylös-tapa, jossa robotti kehittää moraalisen toimintansa oppimisen kautta. Viimeinen näistä on sekoittava lähestymistapa, jossa molempia edellä esiteltyjä tapoja käytetään sekaisin keskenään sopivissa määrin.

Ensimmäisessä edellä mainituista roboetiikan eettisten ongelmien ratkaisemisen lähestymistavoista tulee tiettyjä ongelmia vastaan. Näitä ongelmia demonstroivat roboetiikan kirjallisuudessa usein vastaan tulevat Asimovin lait, jotka Isaac Asimov (1942) esitteli osana novelliaan Runaround. Suomennettuna nämä kolme lakia kuuluvat seuraavasti:

1. Robotti ei saa vahingoittaa ihmistä eikä laiminlyönnin johdosta saattaa tätä vahingoittumaan.
2. Robotin on toteltava ihmisen sille antamia määräyksiä paitsi milloin ne ovat ristiriidassa ensimmäisen pääsäännön kanssa.
3. Robotin on varjeltava omaa olemassaoloaan niin kauan kuin tällainen varjeleminen ei ole ristiriidassa ensimmäisen eikä toisen pääsäännön kanssa.

Asimov ei esitellyt näitä lakeja toimimaan eettisinä ohjenuorina roboteille vaan demonstroidakseen, kuinka kyseisten lakien mukaan toimiva järjestelmä on viallinen ja tulee kohtaamaan konflikteja. Tietokoneet toimivat hyvin kirjaimellisesti ja järjestelmä, joka toimisi kyseisten kolmen lain mukaisesti, tarvitsisi maalaisjärkeä tai intuitiota, jotta se kykenisi toimimaan ja pystyisi selviämään konfliktitilanteista, joissa sille annetut säännöt ovat ristiriidassa keskenään (Goodall, 2014). Kuten Asimovinkin tarkoituksena oli lakeja esitellessään, monet, jotka käsittelevät kirjallisuudessaan Asimovin lakeja, kritisoivat niiden pohjalta robottien kykyä toimia tiettyjen sääntöjen mukaisesti (Goodall, 2014; Murphy & Woods, 2009). Asimovin lait, kuten muutkaan kiveen hakatut säännöt, eivät siis itsessään toimi hyvin etenkin nykyajan robotiikan etiikan ongelmien viitekehystenä. Asimovin lait toimivat enemmänkin perusteluna ja pohjana sille, että ongelmien eettinen tarkastelu on tarpeellista, ja että deontologisten lakien toiminnallisuutta roboteilla pitää muistaa tarkastella yksityiskohtaisemmin sekä useasta näkökulmasta.

Myös toisessa esitetyistä lähestymistavoista, moraalisuuden oppivassa järjestelmässä, nousee tiettyjä ongelmia. Tässä lähestymistavassa olennaista on se, että järjestelmä arvioi tekemiään päätöksiä tiettyjen tavoitteiden kautta. Ongelmia kuitenkin syntyy, kun kontekstit ja tilanteet vaihtuvat, jolloin on epäselvää, minkä tavoitteiden pohjalta oppimista tulisi arvioida. Kahdessa ensimmäisessä lähestymistavoista molemmissa on ongelmia, minkä takia kolmas lähestymistapa, näitä menetelmiä yhdistelevä lähestymistapa, on joustavin ja kattavin vaihtoehto eettisten robottien kehittämiseksi. (Wallach ym., 2008.) Tämä lähestymistapa voisi olla toimivin myös automatisoitujen autojen eettiselle päätöksenteolle, sillä niissä esiintyville ongelmille, joita esittelen luvussa 4.1, on haasteellista nimenomaan löytää tiettyä oikeaa toimintamallia deontologisten lakien tai tiettyjen tavoitteiden toimimisen välillä. Tällöin sääntöjä ja oppimista sekoittava järjestelmä voisi olla toimiva ratkaisu.

3.3 Teorioiden soveltuminen tutkielman aihepiiriin

Perinteisistä etiikan teorioista hyödynnän tässä tutkielmassa pääasiassa deontologiaa ja utilitarismia. Nämä sopivat kyseisen aihepiirin eettiseen tarkasteluun, koska ne tarjoavat rationaalisen pohjan päätöksenteolle. Deontologisissa algoritmeissa automatisoidun järjestelmän täytyy noudattaa tiettyjä ohjeita ja utilitaristisissa algoritmeissa taas tavoitella tiettyjä lopputuloksia. Nämä näkökulmat sopivat tietokoneiden toiminnan tarkastelulle, sillä tietokoneet toimivat tiettyjen ohjeiden tai tavoitteiden perusteella ja nämä suuntaukset tarjoavat nimenomaan ne. (Goodall, 2014.)

Deontologia ja utilitarismi eivät kuitenkaan sovi tai riitä jokaisen esiintyneen eettisen ongelman tarkastelulle. Näissä tilanteissa pohjaan tarkasteluani robotiikan etiikan periaatteisiin, automatisoitujen autojen etiikan kirjallisuudessa esiintyneisiin moraalisiin argumentteihin tai tutkielman kannalta olennaisiin moraalisiin koodistoihin/ohjenuoriin. Näitä ovat esimerkiksi Millerin (2011)

teknologioiden kehittämiseen ja suunnitteluun laaditut säännöt sekä The Institute of Electrical and Electronics Engineersin (tästä eteenpäin IEEE) (2019) eettinen koodisto. Esittelen näiden koodistojen olennaisia kohtia eettisten ongelmien pohdinnan yhteydessä.

4 EETTISET ONGELMAT AUTOMATISOIDUISSA AUTOISSA

Käsittelen tässä luvussa automatisoituihin autoihin liittyviä eettisiä ongelmia. Esittelen ensin nousseet eettiset ongelmat sekä tarkastelen niiden ratkaisemisessa nousevia ristiriitoja eri etiikan teorioihin, moraalisiin sääntöihin tai kyseistä aiheetta koskevaan kirjallisuuteen pohjaten. Lopuksi kokoan ongelmat ja niiden ratkaisuisissa nousevat ristiriidat viitekehukseen sekä esittelen olennaisimpia huomioita ongelmien pohdintaan liittyen.

4.1 Esiintyneet eettiset ongelmat automatisoiduissa autoissa

Tässä luvussa esittelen esiintyneet eettiset ongelmat. Olen jaotellut ongelmat seuraaviin luokkiin: väistämättömien törmäysten ongelmat, ongelmat vastuiden jakamisessa sekä sekaliikenteestä koituvat ongelmat. Useat eettisistä ongelmista kuitenkin sivuavat hieman useampaa kuin vain yhtä näistä luokista. Esimerkiksi eettiset kysymykset siitä, kuka on vastuussa auton päätöksenteosta, liittyvät usein väistämättömien törmäysten tilanteiden jälkitarkasteluun. Kyseiset luokat ja niiden sisäinen ongelmien käsittely ei siis ole täysin eroteltua toisistaan, mutta pyrin tällä jaottelulla hieman selkeyttämään esiintyneiden eettisten ongelmien tarkastelua.

4.1.1 Väistämättömien törmäysten ongelmat

Suuri osa automatisoitujen autojen eettisistä ongelmista kohdistuu erilaisiin versioihin "trolley problem":sta. Klassinen "trolley problem" esiteltiin vuonna 1967 Philippa Footin filosofisena ajatuskokeena (Foot, 1967). Tässä esitellään tilanne, jossa juna on ajamassa viittä työläistä päin ja he eivät kykene väistämään junan edestä pois. Jarruttamiseen aikaa ei myöskään ole. Vieressä on kuitenkin sivuraide, jossa on vain yksi työläinen. Koehenkilöille annetaan mahdollisuus vetää vivusta, jolloin juna siirtyy sivuraiteelle. Tekemättä mitään juna jatkaa suuntaansa ja tappaa viisi työläistä, mutta vivusta vetämällä juna ohjautuu toiselle raiteelle tappaen yhden työläisen. Onko tällöin moraalisesti oikein vetää vivusta? Eettinen ristiriita tässä herää aktiivisen teon ja passiivisuuden välillä. Vivusta vetäessä koehenkilö aktiivisesti tappaa yhden henkilön, mutta vetämättä hän jättää passiivisesti ihmisiä kuolemaan. Tätä ongelmaa on käytetty paljon eettisten pohdintojen perustana ja se on myös pohjana monien automatisoitujen autojen väistämättömien törmäystilanteiden eettisessä pohdinnassa.

Edellä mainittuun ilmiöön pohjautuvat kysymykset automatisoitujen autojen eettisessä pohdinnassa vaikuttavat herättävän kolmen tyyppisiä moraalisia ristiriitoja. Ensimmäisenä ovat ongelmat, jotka käsittelevät tilanteita, jossa pohditaan päätöksestä aiheutuvien menetettyjen ihmishenkien määriä vertailevia

kysymyksiä. Toisena taas ovat ongelmat, jossa pohditaan, tekevätkö jotkin tekijät tiettyä ihmishenkeä arvokkaammaksi kuin toinen. Kolmanteen luokkaan kohdistuvat ongelmat, jotka liittyvät auton kuljettajan uhraamiseen muiden pelastamiseksi. Käsittelen seuraavaksi näitä kolmea ongelmaa kirjallisuuden antamien esimerkkien kautta.

Ensimmäisestä luokasta Gerdes ym. (2016, s. 78-79) antavat seuraavan esimerkin: ihmiskuljettaja ajaa autoa manuaalisella ohjauksella ja kohtaa tilanteen, missä hän on, tarkoituksella tai vahingossa, ajamassa viiden jalankulkijan päältä. Tällöin auton onnettomuuden ehkäisemisjärjestelmä tunnistaa tilanteen, ottaa ohjat ihmiseltä ja väistää ainoaan mahdolliseen suuntaan, mutta tällöin törmää yhteen jalankulkijaan. Toimiiko järjestelmä tällöin oikein? Utilitaristisesta näkökulmasta voitaisiin sanoa suoraan, että kääntyminen on oikea teko, koska enemmän henkiä säästyy. Deontologisesta näkökulmasta voitaisiin kuitenkin tehdä moraalinen ero tappamisen ja kuoleman sallimisen välillä. Jos järjestelmä ei tekisi käännöstä, ei se (autonvalmistaja) olisi vastuussa teosta. Kuten klassisessa esimerkissä, molemmat vaihtoehdot ovat puolusteltavissa näkökulmasta riippuen. (Gerdes ym., 2016, s. 78-79.) Tilanne voitaisiin kääntää myös toisinpäin, jolloin ihmiskuljettaja ottaisi ohjat autosta ja kääntyisi kohti yhtä jalankulkijaa. Kumman kuuluisi tällöin olla vastuussa toiminnasta, kuljettajan vai autonvalmistajan?

Gerdes ym. (2016, s. 78-79) esittävät myös esimerkin toisen luokan ongelmasta: automatisoitu auto kohtaa pakollisen törmäyksen tilanteen, jossa oikealle kääntyessä auto törmää 80-vuotiaaseen rouvaan ja vasemmalle kääntyessä taas 8-vuotiaaseen tyttöön. Jos auto ei käännä kumpaankaan suuntaan, molemmat menehtyvät. Tässä voitaisiin väittää oikean teon olevan törmätä vanhempaan rouvaan, koska hänellä on todennäköisesti vähemmän elinaikaa ja koettavaa jäljellä. Tämä herättäisi kuitenkin eettisiä kysymyksiä. Yhdessä IEEE:n (2019) eettisissä ohjeista sanotaan, että kaikkia ihmisiä kuuluisi kohdella reilusti, ja että ihmisiä ei saisi syrjiä rodun, uskonnon, sukupuolen, vammaisuuden, iän, kansallisuuden, seksuaalisen suuntautumisen, sukupuoli-identiteetin eikä sukupuolen ilmaisemisen mukaan. 8-vuotiaan tytön preferoiminen 80-vuotiaan isoäidin edelle olisi siis suoraan tätä rikkovaa toimintaa, sillä silloin arvotettaisiin törmäykseen kuuluvia henkilöitä iän perusteella. Yksi vaihtoehto tässä olisi tehdä päätös täysin satunnaisesti, mutta sekin on moraalisesti kyseenalaista, koska tämän kaltaisissa tilanteissa voisi kuitenkin olla painavia syitä kohdistaa törmäys toiseen henkilöistä (Gerdes ym., 2016).

Bonnefon, Shariff ja Rahwan (2016) kysyivät tutkimuksessaan ihmisten mielipiteitä automatisoitujen autojen toiminnasta erilaisissa väistämättömien törmäysten tilanteissa. Näissä esiintyi moraalisia kysymyksiä kolmanteen luokkaan, itsensä uhraamiseen, liittyen. He esittävät tutkimuksessaan pääasiassa, hie-man muotoilua vaihdellen, seuraavat kolme tilannetta: ensimmäisessä auton täytyy päättää tappaako se kymmenen jalankulkijaa vai yhden ohikulkijan, toisessa auton täytyy päättää tappaako se yhden jalankulkijan vai uhraako kuljettajan ja kolmannessa auton täytyy päättää tappaako se kymmenen jalankulkijaa vai uhraako kuljettajan. Koehenkilöiden vastauksista ilmeni suurimman osan olevan sitä mieltä, että yleisellä tasolla olisi parempi, jos automatisoidut autot toimisivat

utilitarististen algoritmien perusteella. Tällöin harvempi haluaisi kuitenkin itse ostaa automatisoitua autoa.

Faulhaber ym. (2019) tekivät saman kaltaista tutkimusta VR-testien avulla, joiden tuloksena ilmeni myös ihmisten suosivan utilitaristisia algoritmeja jopa itsensä uhraamisen näkökulmasta katsottuna, mutta he huomioivat tämän saatavan johtua pelkästään sosiaalisesta paineesta eikä tämän takia itsensä uhraavat algoritmit olisi välttämättä se, mitä ihmiset olisivat valmiita itse ostamaan. Eroavaisuuksista siinä, mitä automatisoitujen autojen moraalisilta algoritmeilta toivotaan muiden osalta ja mitä itse oltaisiin valmiita ostamaan, syntyy sosiaalinen dilemma (Dawes, 1980; Van Lange, Joireman, Parks & Van Dijk, 2013, viitattu lähteessä Bonnefon ym., 2016). Jos autojen valmistajat implementoisivat auton sisällä olevia suojelevia algoritmeja, useammat ihmiset ostaisivat autoja, mutta ne toimisivat yleisesti turvallisuuden kannalta katsottuna turvattomammin. Jos taas valmistajat käyttäisivät utilitaristisia algoritmeja, harvempi ostaisi automatisoituja autoja, jolloin enemmän turvattomampia, perinteisiä autoja pysyisi liikenteessä. Tällöin olisi jo pelkästään yleisen turvallisuuden kannalta katsottuna kyseenalaista, kummasta algoritmista aiheutuisi suurempi hyöty, ottamatta edes huomioon esimerkiksi autonvalmistajien kärsimiä tappioita. Mikä on tällöin optimaalinen algoritmi?

Yhtenä ratkaisuna edellä mainittujen kaltaisten tilanteiden eettisen pohdinnan ohittamiseksi on antaa näissä tilanteissa ohjaus ihmiskuljettajalle, jolloin vastuu päätöksenteosta siirtyy hänelle (Gerdes ym., 2016, s. 71). Tämäkin ratkaisu on eettisesti ongelmallista esimerkiksi siksi, että siinä oletetaan kuljettajan olevan kykenevä reagoimaan tarpeeksi nopeasti ja toimimaan riittävän hyvin kyseisessä tilanteessa (Nyholm & Smids, 2018). Gerdes ym. (2016, s. 71) esittävät toteutettujen simulaatioiden osoittavan ihmisen voivan tarvita jopa 40 sekuntia tilannetaujan palauttamiseen, kun hän vaihtaa keskittymisen takaisin auton ohjaamiseen yllättäen muista tehtävistä, kuten lukemisesta tai torkkumisesta. Äkillisissä törmäystilanteissa ei siis olisi mahdollista olettaa ihmisen kykenevän reagoimaan tilanteeseen tarpeeksi nopeasti ja riittävällä keskittymiskyvyllä.

Edellä mainittujen kaltaisiin tilanteisiin voitaisiin törmätä myös perinteisillä autoilla, jolloin ihmiskuljettaja joutuisi arvioimaan samoja kysymyksiä törmäyksen hetkessä. Miksi tämän kaltaista eettistä pohdintaa tulisi siis tehdä vain automatisoitujen autojen näkökulmasta? Näiden törmäyksien todellisissa tilanteissa ihmiskuljettajien päätöksenteko tapahtuu reaaliajassa, jolloin heillä on harvoin aikaa tarkemmalle moraalille arvioinnille. Tämän takia kyseisten tilanteiden päätösten moraalinen pohdinta jää perinteisten autojen ja ihmiskuljettajien kannalta enemmän filosofisten ajatuskokeiden tasolle. Automatisoiduissa autoissa olennaisena erona näihin verrattuna tulee se, että näitä väistämättömien törmäyksien skenaarioita täytyy käsitellä etukäteen jo autoa ohjelmoidessa, jolloin moraalinen tarkastelu tulee olennaisemmin osaksi päätöksenteon määrittelyä (Goodall, 2014). Tällöin päätöksentekoon liittyvät eettiset ongelmat siirtyvät filosofisista ajatuskokeista käytännön toiminnan määrittelyn tasolle. Vaikka nämä tilanteet olisivat epätodennäköisiä, ovat ne silti mahdollisia, jonka takia automatisoitujen autojen kehityksessä on välttämätöntä ottaa kyseiset tilanteet

huomioon ennen kuin niitä voidaan ottaa laajamittaisemmin käyttöön (Bonnefon, Shariff, & Rahwan, 2016).

4.1.2 Ongelmat vastuiden jakamisessa

Yksi esiin nousevimpia kysymyksiä automatisoitujen autojen etiikkaa tarkastelevassa kirjallisuudessa on se, kenen pitäisi olla vastuussa törmäysten sattuessa. Kun auto päätyy törmäystilanteeseen, pitäisikö vastuussa olla auton kehittäjät ja valmistajat vai sen omistaja? Päätöksenteon törmäyksistä on kuitenkin tehneet kehittäjät sen valmistusvaiheessa, jonka perusteella voitaisiin sanoa heidän olevan vastuussa päätöksistä. Auton omistaja on kuitenkin auton ostaessaan hyväksynyt kyseiset toimintamallit, jolloin voidaan sanoa hänestä tulevan auton moraalisuuden edustaja ja näin ollen olevan vastuussa törmäyksistä. Entä jos törmäyksen aiheuttaakin jokin virhe auton ohjelmoinnissa tai se johtuu esimerkiksi siitä, että auton tietoturva on pettänyt, jonka johdosta joku on päässyt hakkeroitumaan auton sisälle ja muuttamaan päätöksentekoa? Pitäisikö vastuun muuttua? Vastuiden jakaminen autojen tekemistä päätöksistä on eettisesti monimutkaista ja eri näkökulmilta samaa ongelmaa katsoessa voidaan päästä eri lopputuloksiin. Tämänkaltaisiin kysymyksiin ei ole vielä vastattu tarpeeksi kattavasti automatisoitujen autojen eettisessä pohdinnassa. (Karnouskos, 2018.)

Borenstein ym. (2017) käsittelevät vastuiden jakamista Millerin (2011) esittämien sääntöjen pohjalta. Millerin (2011) säännöissä esitetään teknologian ennalta nähtävistä vaikutuksista olevan vastuussa ihmiset, jotka ovat olleet suunnittelemassa, kehittämässä tai ottamassa kyseistä teknologiaa käyttöön. Tämän mukaan siis voitaisiin sanoa autonvalmistajien olevan vastuussa auton toiminnasta. Borenstein ym. (2017) nostavat tästä automatisoitujen autojen kannalta olennaisena esiin sen, että tekoälyä sisältävää teknologiaa tarkasteltaessa ei voida varmaksi sanoa, mikä on ennalta nähtävää toimintaa. Autoon voidaan ohjelmoida algoritmit, joiden mukaan se toimii, mutta lopullista toimintaa on silti vaikea ennustaa dynaamisessa ajoympäristössä. Kuka auton toiminnasta on silloin vastuussa, kun se toimii eri tavalla kuin alun perin on tarkoitettu tai on ollut ennustettavissa?

Millerin (2011) kolmannessa säännössä sanotaan ihmisten, jotka tietoisesti käyttävät jotakin laskennallista artefaktia, olevan vastuussa kyseisestä käytöstä. Tämän perusteella taas voitaisiin sanoa autojen käyttäjien olevan vastuussa niiden toiminnasta. Automatisoituja autoja tarkasteltaessa ongelmallisena tulee kuitenkin se, että on vaikeaa sanoa, kuinka paljon tietoa auton toiminnasta on kohtuullista olettaa sen käyttäjiltä ja kuinka tietoisia käyttäjät ovat esimerkiksi siitä, miten ja milloin niiden kuuluisi ottaa ohjat autosta (Borenstein ym., 2017)? Jotta kyseisen säännön mukaan voitaisiin oikeudenmukaisesti antaa vastuu auton käyttäjille, pitäisi heillä olla hyvin perinpohjainen tietämys auton toiminnasta ja sen teknologioista, mikä on hieman epärealistinen oletus useimmilta autojen käyttäjistä.

Borenstein ym. (2017) väittävät, että autonvalmistajien tulisi lopulta olla vastuussa autojen toiminnasta, sillä ne hoitavat autojen testauksen. Vastuun siirtämisellä valmistajille saataisiin ainakin riittävän suuri paine autojen kunnolliselle testaamiselle. Tämä varmasti lisää autojen turvallisuutta, mutta selvästi hidastaisi automatisoitujen autojen tuleamista liikenteeseen. Kuten totesin kappaleessa 2.2.2, automatisoidut autot voisivat toimia paljon perinteisiä autoja turvallisemmin liikenteessä. Olisiko siis yleisen turvallisuuden kannalta katsottuna kuitenkin parempi vaihtoehto siirtää vastuita valmistajille? Utilitaristisesta näkökulmasta voitaisiin sanoa hieman turvattomampien, mutta kuitenkin perinteisiä autoja turvallisempien, autojen mahdollisimman nopean käyttöönoton olevan parempi vaihtoehto kokonaisuuden kannalta. Molemmat näkökulmat katsovat turvallisuutta ja molemmat ovat oikeassa omissa kannoissaan.

Kullekin vaihtoehdolle vastuiden jakamisesta löytyy perusteita riippuen tilanteesta ja näkökulmasta, josta ongelmaa tarkastellaan, minkä takia on vaikeaa päätyä yhteen oikeaan lopputulokseen siitä, kenen pitäisi olla vastuussa automatisoitujen autojen toiminnasta. Kyseistä ongelmaa käsittelevässä kirjallisuudessa painotetaan sitä, että vastuiden jakamisen pohdinnalle pitäisi antaa enemmän huomiota, ja että yleiselle, autonvalmistajien kesken jaettavalle, automatisoitujen autojen kehittämisen ohjesäännöille olisi tarvetta, jotta autojen toiminnan läpinäkyvyys ja yhdenmukaisuus parantuisi (Borenstein ym., 2017; Karnouskos, 2018).

4.1.3 Sekaliikenteestä koituvat ongelmat

Useasti automatisoituja autoja tarkastellessa, etenkin niiden hyötyjen näkökulmasta, katsotaan tilanteita, joissa koko liikenne olisi automatisoitua. Tämän saavuttamista ennen on kuitenkin pitkä vaihe, jossa liikenteessä esiintyy paljon sekä eri tasojen automatisoinnin autoja että perinteisiä autoja automatisoitujen autojen seassa. Nyholm ja Smids (2018) esittävät tutkimuksessaan automatisoitujen ja perinteisten autojen sekä eri automatisoinnin tasojen yhteensopimattomuuksista koituvia ongelmia, ratkaisuja näille ongelmille ja kyseisissä ratkaisuissa nousevia eettisiä ongelmia. Ratkaisuina autojen yhteensopimattomuuksista koituvalle korkeammalle törmäyksien riskille Nyholm ja Smids (2018) esittelevät kolme mahdollisuutta: muuttaa automatisoidut autot käyttäytymään enemmän ihmiskuljettajat kaltaisesti, antaa ohjat ihmisille tietyissä ajotilanteissa ja saada ihmiskuljettajat toimimaan enemmän automatisoitujen autojen kaltaisesti.

Ensimmäisessä kohdassa, automatisoitujen autojen käyttäytymisen muuttamisessa enemmän ihmisten käyttäytymisen kaltaiseksi, ongelmana nousee se, että ihmisten käyttäytyminen liikenteessä ei ole aina turvallista. Kuten kappaleessa 1.2.2 mainitsin, suurin osa liikenteessä aiheutuvista kolareista johtuu ihmisten virheistä. Olisi eettisesti kyseenalaista pyrkiä matkimaan tätä turvatonta käyttäytymistä. (Nyholm & Smids, 2018.)

Toisessa ratkaisuvaihtoehdoista, ohjauksen antamisessa ihmiskuljettajille tietyissä tilanteissa, nousee ongelmana se, onko kuljettaja kykenevä ottamaan

ohjat näissä tilanteissa (Nyholm & Smids, 2018). Kuten kappaleessa 4.1.1 todettua, olisi kyseenalaista olettaa ihmisen kykenevän toimimaan riittävän hyvin ja tarpeeksi nopeasti tietyissä tilanteissa sille annettaessa äkillisesti ohjat. Tämäkin vaihtoehto siis nostaa omat eettiset ongelmansa, eikä ainakaan toimi ratkaisuna kaikissa tilanteissa.

Nyholmin ja Smidsin (2018) esittämä kolmas vaihtoehto, pyrkimys saada ihmiskuljettajat toimimaan enemmän automatisoitujen autojen kaltaisesti esimerkiksi asettamalla nopeusrajoittimia autoihin, eroaa eniten yleisestä näkökulmasta automatisoitujen autojen eettisessä tarkastelussa. Se tarjoaa vaihtoehdon, että voisi olla hyödyllisempää muuttaa ihmisten nykyistä, monissa tilanteissa turvatonta ja eettisesti kyseenalaista, käyttäytymistä kuin sovittaa automatisoituja autoja toimimaan turvattomassa liikenteessä. Ihmisten ajokäyttäytymisen pakollisessa muuttamisessa kuitenkin nousee ongelma siinä, että se rikkoisi ihmisen vapautta. Nyholm & Smids (2018) sanovat Hartin ja Sachs (1994) vapaan lain toteuttamisen olevan olennainen osa ihmisarvoa. Tämä ratkaisumalli rikkoisi tätä arvoa. Nyholm & Smids (2018) esittävät tähän vasta-argumenttina, että vapaan toiminnan arvo ei välttämättä aina mene muiden, kuten turvallisemman liikenteen, arvokkuuden edelle, etenkin kun kyseessä on vapauden rajoittaminen siltä osin, minkä laki jo nyt kieltää. Kuitenkin, myös tässä vaihtoehdossa tulee vastakkainasettelu kahden eri arvon kanssa, joiden välisessä vertailussa voidaan päästä näkökulmasta riippuen moneen eri oikeaan ratkaisuun.

4.2 Viitekehys automatisoitujen autojen eettisten ongelmien hahmottamiseksi

Edellisessä kappaleessa esittelin ja pohdin automatisoituihin autoihin liittyviä eettisiä ongelmia sekä niiden mahdollisten ratkaisujen välisiä ristiriitoja. Seuraavaksi koostan ongelmat viitekehukseen (Taulukko 3) ongelmien kokonaisuuden helpommaksi hahmottamiseksi. Esittelen taulukossa ensin ongelman (sarake 1), jonka jälkeen esittelen vaihtoehtoiset ratkaisut kyseiselle ongelmalle (sarake 2). Kahdessa viimeisessä sarakkeessa esittelen kirjallisuudessa tai etiikan teorioissa esiintyneitä moraalisia argumentteja kunkin ratkaisun puolesta ja vastaan.

Alla esitelty taulukko antaa tiivistetyn koosteen tämän tutkielman tärkeimmistä tuloksista. Valitsin tutkielmaani tiettyjä aihepiirin kirjallisuudessa esiintyviä esimerkkejä, joilla pyrin demonstroimaan mahdollisimman laajasti automatisoitujen autojen erilaisia eettisiä ongelmia. Nämä esimerkit, jotka esiintyvät myös edellä esitellyssä viitekehyksessä, ovat vain pieni osa automatisoituihin autojen eettisistä ongelmista ja antavat siis vain yleiskuvan autojen eettisistä ongelmista. Kuten mainitsin aiemmin, tämän viitekehysten tarkoituksena ei ole antaa ratkaisuja esitetyille ongelmille, vaikka siinä esitetäänkin ratkaisuvaihtoehtoja. Vaihtoehtojen esittelemisellä pyrin nimenomaan osoittamaan ongelmien ratkaisemisen haastavuutta sekä niiden monimutkaisuutta.

TAULUKKO 3. Koostava esitys automatisoitujen autojen eettisistä ongelmista

| Eettinen ongelma | Ratkaisuvaihtoehto | Puolesta | Vastaan |
|---|--|--|---|
| Pitäisikö auton ajaa 8-vuotiaan tytön vai 80-vuotiaan rouvan päältä väistämättömän törmäyksen tilanteessa? (Gerdes ym., 2016, s. 78-79) | Ajaa 80-vuotiaan rouvan päältä | 8-vuotiaalla työllä on todennäköisesti enemmän elämää jäljellä (Utilitarismi) | Ihmisarvoa ei saa määritellä minkään ominaisuuden mukaan (IEEE Code of Ethics, 2019) |
| | Valitsee satunnaisesti kohteiden välillä | Päätös olisi eettisesti ongelmallista tehdä näiden välillä (IEEE Code of Ethics 2019) | Tilanteessa voisi olla painavia syitä suuntaan tai toiseen (Gerdes ym., 2016, s. 78-79) |
| Kuuluisiko auton ottaa ohjat ihmiskuljettajalta ja törmätä yhteen sivulliseen ohikulkijaan, kun ihminen on ajamassa kohti viittä henkilöä? (Gerdes ym., 2016, s. 78.79) | Kyllä | Oikea tapa on toimia niin, että kuolemien määrä on mahdollisimman pieni (Utilitarismi) | Tappaminen on aina väärin (Deontologia) |
| | Ei | Edellisen rivin Vastaan -argumentti | Edellisen rivin Puolesta -argumentti |
| Kuuluisiko autoissa käyttäjä itsensä uhraavia algoritmeja useampien ulkopuolisten henkien säästämiseksi? (Bonneton ym., 2016; Faulhaber ym., 2019) | Auton sisällä olevia suojelevia algoritmeja | Useampi ostaisi automatisoituja autoja, liikenteen turvallisuus parani (Utilitarismi) | Enemmän uhreja väistämättömien törmäysten tilanteissa (Utilitarismi) |
| | Itsensä uhraavia algoritmeja | Vähemmän uhreja väistämättömien törmäysten tilanteissa (Utilitarismi) | Edellisen rivin Puolesta -argumentti |
| Kenen kuuluisi olla vastuussa automatisoidun auton päätöksistä? (Karnouskos, 2018; Borenstein ym., 2017) | Vastuussa auton omistaja | Ostaessaan auton, omistajasta tulee sen moraalisen vastuun kantaja (Karnouskos, 2018) | Mitä jos auto käyttäytyy toisin kuin sen on oletettu toimivan? (Borenstein ym., 2017) |
| | Vastuussa auton valmistamiseen liittyvät tahot | Teknologioiden kehittäjät vastuussa sen ennalta nähtävästä toiminnasta (Miller, 2011) | Mitä jos toiminta on jotain muuta kuin ennalta nähtävää? (Borenstein ym., 2017) |
| Kuinka sekaliikenteestä koituvia yhteensopimattomuusongelmia tulisi | Automatisoitujen autojen toiminta enemmän | Vähentää yhteensopimattomuusongelmia | Matkisi ihmisten turvatonta käyttäytymistä |

| | | | |
|-----------------------------------|---|--|---|
| ratkaista? (Nyholm & Smids, 2018) | ihmistenkuljettajien kaltaiseksi | (Nyholm & Smids, 2018) | (Nyholm & Smids, 2018) |
| | Ohjat ihmiskuljettajalle tietyissä tilanteissa | Edellisen rivin Puolesta -kohta sekä vältetään eettisten ongelmien pohtiminen (Nyholm & Smids, 2018) | Ihminen ei kykene reagoimaan tarpeeksi nopeasti äkillisissä tilanteissa (Nyholm & Smids, 2018; Gerdes ym., 2016, s. 71) |
| | Pyritään saamaan ihmisiä toimimaan enemmän automatisoitujen autojen kaltaisesti | Paras vaihtoehto yleisen liikenteen turvallisuuden kannalta (Utilitarismi, Nyholm & Smids, 2018) | Riistäisi ihmisten vapautta (Nyholm & Smids, 2018) |

4.3 Esiintyneiden ongelmien pohdintaa

Tulosten perusteella voidaan todeta automatisoituihin autoihin liittyvän monimutkaisia eettisiä ongelmia, joiden ratkaisemiseksi on haastavaa, ellei mahdollista, löytää yhtä oikeaa ratkaisua. Tutkielmani tavoitteena oli antaa yleiskuva automatisoituihin autoihin liittyvistä eettisistä ongelmista ja perustella tarvetta aihepiirin laajemmalle eettiselle pohdinnalle. Väitänkin löydösten perusteella tälle laajemmalle pohdinnalle olevan selvä tarve. Esittelen seuraavaksi huomioita, joita olisi tulosten pohjalta mielestäni hyödyllistä ottaa huomioon laajemmassa eettisessä pohdinnassa.

Ongelmia esitellessäni kävi ilmi, että jokaiselle esitetyistä ongelmista löytyy useita ratkaisuvaihtoehtoja, ja että jokaiselle näistä vaihtoehtoista löytyy omat eettiset argumenttinsa eri näkökulmista ja eettisten teorioiden opeista katsottuna. Ratkaisuvaihtoehdot ja niiden argumentit ovat myös useimmiten ristiriidassa keskenään, kuten eettisille ongelmille onkin tyypillistä. Goodall (2017) väittää, että automatisoitujen autojen eettisten ongelmien ratkaisemisessa olevan olennaista yhdistää eri eettisten teorioiden, esimerkiksi deontologian ja utilitarismin, oppeja, sillä yhdestä etiikan suuntauksesta ei löydy vastauksia kaikkiin esiintyviin kysymyksiin. Hän myös mainitsee tällaisen lähestymistavan olevan jo toimivassa käytössä esimerkiksi elinluovutuksen jonojen priorisoinnissa. Löydösteni perusteella tuen tätä väitettä. Kuten edellä todettua, useimmissa ongelmista päästään ristiriitoihin niitä tarkasteltaessa eri etiikan suuntauksista, joten näitä suuntauksia yhdistelevä näkökulma voisi olla toimiva metodi automatisoitujen autojen laajemmassa eettisessä pohdinnassa.

Borenstein ym. (2017) mainitsevat, että automatisoiduille autoille ei tällä hetkellä ole riittävän kattavia yleisiä standardeja, minkä johdosta

autonvalmistajien suuntaukset kehityksessä vaihtelevat huomattavasti. Kuten todettua luvussa 4.1.1, autojen käyttäjien preferenssit siitä, haluaisivatko he autojen toimivan utilitarististen algoritmien vai itseä säästävien algoritmien mukaisesti, vaihtelevat huomattavasti, kun kysymyksessä on toivomukset muiden autojen toiminnasta verrattuna siihen, mitä he itse olisivat valmiita ostamaan. Yleisten standardien puuttuessa tämä ihmisten käytös voisi johtaa tilanteeseen, jolloin valtaosa autojen ostajista päätyisivät tiettyjen, esimerkiksi itseä suojelevia algoritmeja käyttävien, autonvalmistajien autoihin. Tämä voisi tehdä autoista liikenteen kokonaisuuden kannalta turvattomampia. Väitän siis, että autojen päätöksenteon ja niiden moraalisten algoritmien kehittämisessä olisi tarvetta tarkemmille yleisille standardeille. En sano, että juuri utilitaristiset algoritmit ovat tavoiteltava asia kaikilla autonvalmistajilla, vaan että linjaukset autoihin sisällytettävistä moraalisisista algoritmeista, mitä ikinä ne ovatkaan, ja muutenkin niiden kehittämisestä, tulisi perustua yhteisiin standardeihin, jotta edellä mainitun kaltaisen sosiaalisten dilemموjen mahdollisia haittavaikutuksia voitaisiin minimoida.

Kolmantena huomiona automatisoitujen autojen laajemmalle eettiselle pohdinnalle voisin nostaa luvussa 3.2 esitellyn, roboetiikassa yleisen, eri lähestymistapoja sekoittavan eettisten ongelmien ratkaisumallin hyödyntämisen. Automatisoidut autot voivat kohdata äärettömän monia eettisesti ongelmallisia tilanteita. Olisi mahdotonta ohjelmoida auton käyttäytymistä kaikkiin näihin tilanteisiin erillisinä tapauksina. Tämän takia joko tietyt toimintaa ohjaavat deontologiset säännöt tai tiettyjen tavoitteiden pohjalta oikeaa toimintaa oppiva järjestelmä on välttämätön automatisoiduille autoille. Kuten aiemmin todettua, kummassakin näistä nousee omat ongelmansa, joten näitä sekoittava lähestymistapa, jossa järjestelmä hyödyntäisi tekoälyä oppiakseen eettisesti oikeaa toimintaa, kuitenkin pohjaten ja tarkistaen toimintaansa säännöllisesti tiettyihin sääntöihin verraten, voisi olla mielestäni toimiva tapa automatisoitujen autojen eettisten ongelmien ratkaisemiselle. Tässä lähestymistavassa olennaista olisi testata autojen toimintaa säännöllisesti ja kattavasti eri parametreja hyödyntäen, jotta ne eivät oppisi vääränlaista toimintaa.

Edellä mainitut kolme kohtaa olisi mielestäni hyödyllistä ottaa huomioon automatisoitujen autojen eettisessä pohdinnassa. Väitän näiden kolmen kohdan olevan erityisesti hyödyllisiä, kun niitä käytetään toistensa tukena. Esimerkiksi eri etiikan teorioiden yhdistäminen voisi käytännössä tapahtua hyvin sekoittamalla niitä autojen oikeaa toimintaa oppivaan järjestelmään, jolloin auton tekoäly analysoisi päätöksentekoaan useiden näkökulmien kautta. Yleisiä linjauksia ja standardeja tähän taas olisi hyvä yhdistää määrittelemällä oppivien järjestelmien tueksi yleisiä deontologisia sääntöjä sekä yhdenmukaisia testitapauksia oikean oppimisen takaamiseksi. Muun muassa näitä menetelmiä hyödyntämällä olisi mielestäni hyvä lähestyä tarvittavaa automatisoitujen autojen eettisten ongelmien laajempaa tarkastelua.

5 YHTEENVETO

Tutkielmani tarkoituksena oli selvittää, mitä eettisiä ongelmia automatisoituihin autoihin liittyy. Pohjustin tähän tavoitteeseen pääsemistä avaamalla ensin ensimmäisessä sisältöluvussa automatisoitujen autojen määritelmää sekä niiden kehityksen nykyistä tilannetta. Esittelin tässä luvussa myös lyhyesti tämän hetkistä automatisoituihin autoihin liittyvää lainsäädäntöä sekä vaikutuksia, mitä ne voisivat saada aikaan. Toisessa sisältöluvussa taas esittelin tutkielman kannalta olennaisia etiikan teorioita ja kenttiä, kuten robotiikan etiikkaa, sekä pohdin niiden oppien soveltuvuutta automatisoitujen autojen eettiseen tarkasteluun. Kolmannessa sisältöluvussa esittelin tutkielman tärkeimmät tulokset eli toin esiin ilmenneet eettiset ongelmat ja havainnollistin niiden monimutkaisuutta.

Tutkimuskysymykseni oli: ”Mitä eettisiä ongelmia automatisoituihin autoihin liittyy?”. Kuten edellä totesin, vastasin tähän kysymykseen kolmannessa sisältöluvussa. Pyrin vastaamaan kyseiseen kysymykseen esittelemällä esiintyneitä eettisiä ongelmia havainnollistavien esimerkkien kautta. Perustelin ongelmien monimutkaisuutta esittämällä niiden ratkaisuvaihtoehtojen välillä nousevia ristiriitoja. Vaikka esittelinkin vaihtoehtoisia ratkaisuja ongelmille, tutkielmani tarkoitukseni ei ollut antaa ongelmille vastauksia, vaan nimenomaan osoittaa niiden ratkaisemisen haastavuutta sekä perustella tarvetta ongelmien laajemmalle pohdinnalle.

Tutkielmani tärkeimmät tulokset ovat kolmannessa sisältöluvussa esitetyt eettiset ongelmat ja niiden pohdintaan liittyvät huomiot. Konkreettisena tuotoksena kokosin luvussa 4.2 esiintyneet ongelmat sekä niiden ratkaisuvaihtoehdot ja näiden vaihtoehtojen välillä esiintyvät ristiriidat yhteen viitekehykseen. Tämän viitekehyksen, ja muutenkin tämän tutkielman, tarkoituksena on tarjota yleiskuva automatisoituihin autoihin liittyvistä eettisistä ongelmista ja toimia perustelevana tekijänä sekä lähtökohtana niiden laajemmalle pohdinnalle. Kolmannen sisältöluvun lopussa, luvussa 4.3, pohdin esiintyneitä ongelmia ja tein sen pohjalta huomioita, mitä automatisoitujen autojen etiikan laajemmassa tarkastelussa olisi hyödyllistä ottaa huomioon. Toin tässä esiin seuraavat kolme kohtaa, mitä voisi mielestäni hyödyntää laajemmassa tarkastelussa: etiikan teorioiden yhdistäminen ongelmien pohdinnassa, yleisten standardien luominen autojen kehitykseen ja eettiseen pohdintaan sekä oikeaa toimintaa oppivan ja deontologisia sääntöjä sekoittavan järjestelmän hyödyntäminen automatisoitujen autojen eettisessä päätöksenteossa.

Onnistuin mielestäni vastaamaan tutkielman tutkimuskysymykseen yleisellä tasolla. Tarkoitukseni oli antaa yleiskuva automatisoituihin autoihin liittyvistä eettisistä ongelmista ja tässä onnistuin. Tämä yleiskuva on kuitenkin kärkeä, eikä se kata läheskään kaikkia automatisoituihin autoihin liittyvistä eettisistä ongelmista. Etenkin yksityiskohtaisemmat ja automatisoitujen autojen tulehmissivuvaikutuksista nousevat eettiset kysymykset jäivät tämän tutkielman ulkopuolelle. Esimerkki tämmöisestä on esimerkiksi, onko oikein, että automatisoidut autot veisivät taksinkuljettajien työpaikkoja. En myöskään pureutunut

yksityiskohtaisemmin mihinkään esitellyistä ongelmista ja niiden vaihtoehtoisista ratkaisuista. Heikkoutena tutkielmassani näen siis rajalliset resurssit, jonka takia en pystynyt pureutumaan tarkemmin esiintyneisiin ongelmiin, enkä antamaan niin laajaa kuvaa automatisoituihin autoihin liittyvistä eettisistä ongelmista, kuin olisi voinut olla hyödyllistä antaa. Nämä kohdat jäävät jatkotutkimuksen ja laajemman eettisen pohdinnan varaan.

LÄHTEET

- Bischoff, J., Maciejewski, M. (2016) Simulation of city-wide replacement of private cars with autonomous taxis in Berlin. *Procedia Computer Science*, 83, 237–244. [doi:10.1016/j.procs.2016.04.121](https://doi.org/10.1016/j.procs.2016.04.121)
- Bonnefon, J., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science (New York, N.Y.)*, 352(6293), 1573-1576. doi:10.1126/science.aaf2654
- Borenstein, J., Herkert, J., & Miller, K. (2017). Self-driving cars: Ethical responsibilities of design engineers. *IEEE Technology and Society Magazine*, 36(2), 67-75. doi:10.1109/MTS.2017.2696600
- Brown, A., Gonder, J., & Repac, B. (2014). An analysis of possible energy impacts of automated vehicles. *Teoksessa Road vehicle automation* (s. 137-153). Springer, Cham.
- Center for Sustainable Systems. (2019). Autonomous vehicles factsheet. Haettu osoitteesta http://css.umich.edu/sites/default/files/css_doc/CSS16-18.pdf
- Cohen, S. A. & Hopkins, D. (2019). Autonomous vehicles and the future of urban tourism. *Annals of Tourism Research*, 74, 33-42. doi:10.1016/j.annals.2018.10.009
- Collingwood, L. (2017). Privacy implications and liability issues of autonomous vehicles. *Information & Communications Technology Law*, 26(1), 32-45. doi:10.1080/13600834.2017.1269871
- Dawes, R. M. (1980). Social dilemmas. *Annual review of psychology*, 31(1), 169-193.
- Dentons. (2020). *Global guide to autonomous vehicles 2020*. Haettu osoitteesta <https://www.dentons.com/en/insights/guides-reports-and-whitepapers/2020/january/29/global-guide-to-autonomous-vehicles-2020>
- European Commission. (2018). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee, the Committee of the Regions:: On the road to automated mobility: An EU strategy for mobility of the future*. Haettu osoitteesta https://ec.europa.eu/transport/sites/transport/files/3rd-mobility-pack/com20180283_en.pdf

- European Commission. (2019). *Guidelines on the exemption procedure for the eu approval of automated vehicles* (versio 4.1). Haettu osoitteesta <https://ec.europa.eu/docsroom/documents/34802>
- Fagnant, D. J. & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research Part A*, 77, 167-181. doi:10.1016/j.tra.2015.04.003
- Faulhaber, A., Dittmer, A., Blind, F., Wächter, M., Timm, S., Sütfeld, L., Stephan, A., Pipa, G. & König, P. (2019). Human decisions in moral dilemmas are largely described by utilitarianism: Virtual car driving study provides guidelines for autonomous driving vehicles. *Science and Engineering Ethics*, 25(2), 399-418. doi:10.1007/s11948-018-0020-x
- Foot, P. (1967). The problem of abortion and the doctrine of the double effect. *Virtues and vices* (pp. 19-33) doi:10.1093/0199252866.003.0002
- Gerdes, J., Lenz, B., Maurer, M. & Winner, H. (2016). *Autonomous driving: Technical, legal and social aspects..* Berlin: Springer. doi:10.1007/978-3-662-48847-8
- Goodall, N. J. (2014). Ethical decision making during automated vehicle crashes. *Transportation Research Record: Journal of the Transportation Research Board*, 2424(1), 58-65. doi:10.3141/2424-07
- Goodall, N. J. (2017). From trolleys to risk: Models for ethical autonomous driving. *American Journal of Public Health*, 107(4), 496. doi:10.2105/AJPH.2017.303672
- Greenblatt, J. & Shaheen, S. (2015). Automated vehicles, on-demand mobility, and environmental impacts. *Current Sustainable/Renewable Energy Reports*, 2(3), 74-81. doi:10.1007/s40518-015-0038-5
- Greenblatt, J. B., & Saxena, S. (2015). Autonomous taxis could greatly reduce greenhouse-gas emissions of US light-duty vehicles. *Nature Climate Change*, 5(9), 860-863. doi:10.1016/j.tra.2015.04.003
- Grush, B., Niles, J. & Baum, E. (2016). Ontario Must Prepare for Vehicle Automation: Automated vehicles can influence urban form, congestion and infrastructure delivery. RCCAO. Haettu osoitteesta https://rccao.com/research/files/RCCAO_Vehicle-Automation_OCT2016_WEB.pdf
- Karnouskos, S. (2018). Self-driving car acceptance and the role of ethics. *IEEE Transactions on Engineering Management*, , 1-14. doi:10.1109/TEM.2018.2877307

- Marturano, A. (2002). The role of metaethics and the future of computer ethics. *Ethics and Information Technology*, 4(1), 71-78. doi:10.1023/A:1015202319899
- Miller, K. W. (2011). Moral responsibility for computing artifacts: "The rules". *IT Professional*, 13(3), 57-59. doi:10.1109/MITP.2011.46
- Murphy, R. R. & Woods, D. D. (2009). Beyond asimov: The three laws of responsible robotics. *IEEE Intelligent Systems*, 24(4), 14-20. doi:10.1109/MIS.2009.69
- National Conference of State Legislatures. (2020). Autonomous vehicles | self-driving vehicles enacted legislation. Haettu osoitteesta <https://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx>
- Nyholm, S. & Smids, J. (2018). Automated cars meet human drivers: Responsible human-robot coordination and the ethics of mixed traffic. *Ethics and Information Technology*, , 1-10. doi:10.1007/s10676-018-9445-9
- Pakusch, C., Stevens, G., Boden, A. & Bossauer, P. (2018). Unintended effects of autonomous driving: A study on mobility preferences in the future. *Sustainability*, 10(7), 2404. doi:10.3390/su10072404
- Parliament of Singapore. (2017). Road Traffic (Amendment) Bill (No. 5/2017). Haettu osoitteesta <https://sso.agc.gov.sg/Bills-Supp/5-2017/Published/20170110?DocDate=20170110>
- Quinn, M. J. (2015). *Ethics for the Information Age, Global Edition, 6th edition*, Iso-Britannia: Pearson Education M.U.A.
- SAE International. (2018). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Standard J3016). Haettu osoitteesta https://saemobilus.sae.org/content/j3016_201806
- Stahl, B., Timmermans, J. & Mittelstadt, B. (2016). The ethics of computing. *ACM Computing Surveys (CSUR)*, 48(4), 1-38. doi:10.1145/2871196
- Spieser, K., Treleaven, K., Zhang, R., Frazzoli, E., Morton, D., & Pavone, M. (2014). Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in Singapore. *Teoksessa Road vehicle automation* (s. 229-245). Springer, Cham.
- The Institute of Electrical and Electronics Engineers (2019). IEEE Code of Ethics. Haettu osoitteesta <https://www.ieee.org/about/corporate/governance/p7-8.html>

- Traficom. (2018). Tieliikenteen automaatiokokeilut. Haettu osoitteesta <https://www.traficom.fi/fi/liikenne/tieliikenne/tieliikenteen-automatiokokeilut>
- Tzafestas, S. G. (2018). Ethics in robotics and automation: A general view. *International Robotics & Automation Journal*, 4(3)
doi:10.15406/iratj.2018.04.00127
- Van Lange, P. A., Joireman, J., Parks, C. D., & Van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, 120(2), 125-141. doi:10.1016/j.obhdp.2012.11.003
- Veruggio, G., Solis, J. & Van der Loos, M. (2011). Roboethics: Ethics applied to robotics, *IEEE Robotics & Automation Magazine*, 18(1), 21-22,
doi:10.1109/MRA.2010.940149
- Victoria Transport Policy Institute. (2020). Autonomous vehicle implementation predictions: Implications for transport planning. Haettu osoitteesta <https://www.vtppi.org/avip.pdf>
- Wallach, W., Allen, C. & Smit, I. (2008). Machine morality: Bottom-up and top-down approaches for modelling human moral faculties. *Ai & Society*, 22(4), 565-582. doi:10.1007/s00146-007-0099-0