



Eettinen tekoäly autonomisissa järjestelmissä

Tuisku Salonen | Nina Rilla | Pietari Pikkuaho

Eettinen tekoäly autonomisissa järjestelmissä

Tuisku Salonen, Nina Rilla & Pietari Pikkuaho

Teknologian tutkimuskeskus VTT Oy



ISBN 978-951-38-8780-3

VTT Technology 418

ISSN-L 2242-1211

ISSN 2242-122X (Verkkójulkaisu)

DOI: 10.32040/2242-122X.2023.T418

Copyright © VTT 2023

JULKAISIJA – PUBLISHER

VTT

PL 1000

02044 VTT

Puh. 020 722 111

<https://www.vtt.fi>

VTT

P.O. Box 1000

FI-02044 VTT, Finland

Tel. +358 20 722 111

<https://www.vttresearch.com>

Sisällysluettelo

1	Alkusanat	1
2	Autonominen liikkuminen	2
2.1	Nykypäivän autonomisuuden tasot ajoneuvoissa	6
3	Käyttäjien odotukset autonomista liikkumista kohtaan	8
3.1	Odotusten sosiologia	8
3.2	Arvot ja etiikka	9
3.3	Käyttäjät ja arvot	11
3.4	Arvokartta	12
4	Tekoälyyn liittyvien arvojen tutkiminen	15
5	Käyttäjien arvot: Raitiovaunuliikenne	16
5.1	Hyväntahtoisuus	18
5.2	Turvallisuus	19
5.3	Itseohjautuvuus	20
5.4	Universalismi	20
6	Käyttäjien arvot: Itseohjautuvat autot	22
6.1	Hyväntahtoisuus	25
6.2	Turvallisuus	25
6.3	Itseohjautuvuus	26
6.4	Universalismi	27
7	Käyttäjien arvot: itseohjautuvat laivat	28
7.1	Hyväntahtoisuus	31
7.2	Turvallisuus	31
7.3	Valta	32
7.4	Hedonismi	32
8	Keskustelu	33

Lopuksi	36
Kirjallisuus	37
Liite 1: Kuljettamattomaan raitiovaunuun liittyvät arvot.	1
Liite 2: Autonomisiin autoihin liittyvät arvot.	2
Liite 3: Autonomiseen merenkulkuun liittyvät arvot.....	3

1 Alkusanat

Miten itseohjautuvat autot ja muut liikennevälineet muuttavat liikennettä ja miten ne voisivat lisätä liikenneturvallisuutta? Missä määrin autonomisten järjestelmien tulee olla eettisiä? Miten eettisyys voi toimia kilpailuetuna autonomisten järjestelmien kehityksessä?

Autonomisilla järjestelmillä voidaan viitata itseohjautuviin autoihin, laivoihin, juniin, erilaisiin koneisiin sekä kokonaisiin teollisuusjärjestelmiin. Ne voivat lisätä turvallisuutta, tehokkuutta ja kestävyyttä, mutta järjestelmien ohjelmointiin ja käyttöön liittyy kuitenkin monia eettisiä kysymyksiä.

Kuinka vastuullisuus todentuu monimutkaisten toimijajärjestelmien ekosysteemeissä? Miten jaettu ymmärrys eettisistä arvoista muotoutuu ja yhteinen arvonluonti mahdollistuu liikenteen kaltaisissa toimijaympäristöissä, joissa autonomisuuden aste kasvaa kohti täyttä autonomisuutta?

Tässä katsauksessa tarkastellaan tekoälyyn liittyvien arvojen muodostumista kolmessa autonomiseen liikkumiseen liittyvässä kontekstissa: raitiovaunuliikenteessä, yksityisautoilussa ja meriliikenteessä.

Katsaus on toteutettu strategisen tutkimuksen neuvoston rahoittamassa ETAIROS-tutkimushankkeessa, jonka keskeinen tavoite on vahvistaa julkisen ja yksityisen sektorin yhteistä ymmärrystä eettisten käytäntöjen ja pelisääntöjen välttämättömyydestä. ETAIROS kehittää eettisesti kestäviä tekoälyn suunnittelumenetelmiä ja tuottaa näkemystä tekoälyä koskevista koordinaation ja ohjauksen malleista yhdessä sidosryhmien kanssa. Autonomiset järjestelmät on yksi ETAIROS:n tutkimusviitekehyksen viidestä tema-alueesta.

2 Autonominen liikkuminen

Autonomiseen liikkumiseen sekä autonomisiin liikennejärjestelmiin kohdistuu voimakasta kiinnostusta ympäri maailman. Isoja odotuksia on ladattu esimerkiksi tie-, raide-, laiva- ja lentoliikenteeseen, joissa hyötyjä toivotaan näkyvän esimerkiksi korkeamman turvallisuuden, tehokkuuden sekä pienempien päästöjen muodossa.^{1,2,3,4,5,6} Autonomisten järjestelmien etuja pidetäänkin usein niin huomattavina ja itsestään selvinä, ettei esimerkiksi järjestelmien laajempaa

-
- ¹ GEAR 2030 (2017). High Level Group on the Competitiveness and Sustainable Growth of the Automotive Industry in the European Union. FINAL REPORT
<https://www.europarl.europa.eu/cmsdata/141562/GEAR%202030%20Final%20Report.pdf> (20.3.2030)
- ² GEAR 2030 (2016). Roadmap on Highly automated vehicles. DISCUSSION PAPER.
<https://circabc.europa.eu/sd/a/3e06f7bf-2719-4be1-9f24-ba6b2975d7eb/Discussion%20Paper%20-%20rev.1%2004-05-2015.pdf> (20.3.2023)
- ³ Singh et al., (2021). Deployment of Autonomous Trains in Rail Transportation: Current Trends and Existing Challenges. IEEE Access, VOLUME 9, 91427- 91461.
- ⁴ Karetnikov, et al. (2021). Technology Level and Development Trends of Autonomous Shipping Means. In: Murgul, V., Pukhkal, V. (eds) International Scientific Conference Energy Management of Municipal Facilities and Sustainable Energy Technologies EMMFT 2019. EMMFT 2019. Advances in Intelligent Systems and Computing, vol 1258. Springer, Cham. https://doi.org/10.1007/978-3-030-57450-5_36
- ⁵ Kim, et al. (2020). Autonomous shipping and its impact on regulations, technologies, and industries, Journal of International Maritime Safety, Environmental Affairs, and Shipping, 4:2, 17-25, DOI: 10.1080/25725084.2020.177942
- ⁶ National Research Council (2014). Autonomy Research for Civil Aviation: Toward a New Era of Flight. Washington, DC: The National Academies Press.
<https://doi.org/10.17226/18815>.

yleistymistä tulevaisuudessa kyseenalaisteta juuri lainkaan.^{7,8,9} Suomessakin autonomisen liikenteen saapumiseen on valmistauduttu mm. Valtioneuvoston ja Traficomien antamien linjausten kautta, jotka koskevat liikenteen automaation edistämistä lähivuosina.^{10,11}

Autonomisen liikkumisen yleistymisen odotetaan tapahtuvan noin vuosien 2025–2030 välillä, jolloin esimerkiksi kuluttajapuolella itseajavien autojen uskotaan saapuvan vahvemmin markkinoille. Nämä odotukset perustuvat pitkälti tiekarttoihin, valmistajien lupauksiin sekä ihmisten odotuksiin.^{12,13} Toisaalta samassa yhteydessä on huomattava, että itseajavia autoja povattiin markkinoille jo alun perin 1970-luvulla, kuin myös 2010-luvulla ja 2020-luvun alkupuolella. Itseasiassa, itseajaviin autoihin kohdistuvia odotuksia voi verrata monessa suhteessa fuusioreaktioon tai grafeeniin, joiden on uskottu mullistavan maailman jo useat viimeisistä vuosikymmenistä, mutta joihin kohdistuvat odotukset eivät aina ole linjassa niiden käyttöön liittyvien haasteiden kanssa.

Onkin todennäköistä, että automaatio ei ole tulossa laajemmaksi osaksi tie-, raide, laiva- tai lentoliikennettä lähivuosina. Viivästyminen johtuu pitkälti autonomisten järjestelmien käyttöönottamiseen liittyvistä haasteista ja niiden vaatimista hyvin merkittävistä investoinneista esimerkiksi koulutuksen, osaamisen, infrastruktuurin sekä nykyisen kaluston puolella.^{14,15} Itseajavien ajoneuvojen

⁷ European Commission (2019). STRIA Roadmap on Connected and Automated Transport: Road, Rail and Waterborne. https://trimis.ec.europa.eu/sites/default/files/2021-10/stria_roadmap_2019-connected_and_automated_transport.pdf (20.3.2023)

⁸ Pfothenauer, S., & Jasanoff, S. (2017). Panacea or diagnosis? Imaginaries of innovation and the 'MIT model' in three political cultures. *Social Studies of Science*, 47(6), 783–810. <https://doi.org/10.1177/0306312717706110>

⁹ Gandia, et al. (2019). Autonomous vehicles: scientometric and bibliometric review, *Transport Reviews*, Vol. 39(1), 9-28. <https://doi.org/10.1080/01441647.2018.1518937>.

¹⁰ <https://www.lvm.fi/-/valtioneuvostolta-periaatepaatos-liikenteen-automaation-edistamisesta-1592000> (20.3.2023)

¹¹ <https://www.lvm.fi/-/tavoitteena-ihmiskeskeinen-liikenteen-automaatio-lainsaadanto-ja-to-menpidesuunnitelma-lausunnoille-1246107> (20.3.2023)

¹² Devi, et al. (2020). A Comprehensive Survey on Autonomous Driving Cars: A Perspective View. *Wireless Personal Communications: An International Journal*, Vol. 114(3), 2121–2133. <https://doi.org/10.1007/s11277-020-07468-y>

¹³ Faisal et al., (2019). Understanding autonomous vehicles: A systematic literature review on capability, impact, planning and policy. *The Journal of Transport and Land Use*, Vol.12(1), 45–7.

¹⁴ Karetnikov, et al. (2021). Technology Level and Development Trends of Autonomous Shipping Means. In: Murgul, V., Pukhkal, V. (eds) *International Scientific Conference Energy Management of Municipal Facilities and Sustainable Energy Technologies EMMFT 2019. Advances in Intelligent Systems and Computing*, vol 1258. Springer, Cham. https://doi.org/10.1007/978-3-030-57450-5_36

¹⁵ Anderson, M. (2020). The road ahead for self-driving cars: The AV industry has had to reset expectations, as it shifts its focus to level 4 autonomy. *IEEE Spectrum*, Vol. 57 (5), 8 – 9. DOI: 10.1109/MSPEC.2020.9078402

yleistyminen tulee siis varmasti tuomaan mukanaan merkittäviä muutoksia useille toimialoille, kuin myös luomaan kokonaan uusia toimialoja.

Samalla muutokset sekä kuljetuksen, logistiikan ja yksityisautoilun puolella voivat laukaista uusia megatrendejä. Kehityspoluista riippuen, hyötyjen toteutuminen voi vaatia liikennejärjestelmän uudelleensuunnittelua, tieinfrastruktuurin päivittämistä, kuin myös autokannan nopeampaa uudistumista.^{16,17,18} Autonomisiin liikennevälineisiin liittyvien odotusten tarkastelu voikin itsessään olla arvokasta. Asetettaessa vierekkäin vielä ratkaistavana olevat ongelmat autonomisten järjestelmien kehityksessä sekä itse autonomisiin järjestelmiin kohdistuvat odotukset, muodostuu nähtävillä jossain määrin vinoutunut kuva.

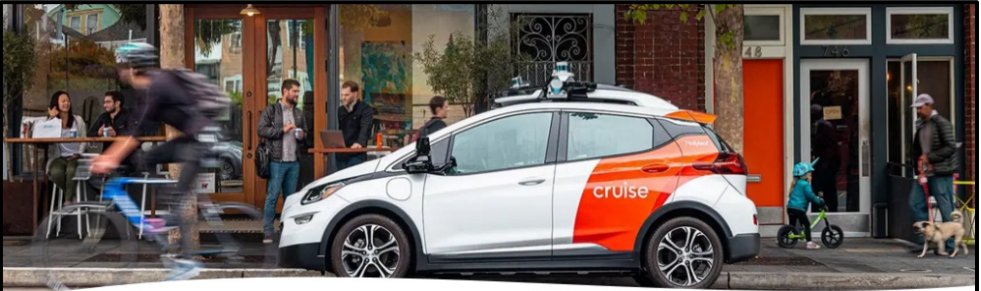
Tiedämme historiallisesti olevamme myös melko huonoja ennustamaan uusien teknologioiden vaikutusta yhteiskuntaamme. Tästä hyviä esimerkkejä ovat vaikkapa internet, digitalisaatiokehitys ja sosiaalinen media. Voimme melko hyvällä varmuudella todeta, että autonomisen liikenteen laajempi yleistyminen tulee vaikuttamaan yhteiskuntaamme, mutta se miltä tämä muutos tulee näyttämään ja kuinka pitkään se tulee viemään, ovat vaikeampia arvioida. Todennäköistä kuitenkin on, että yhden suuren automaatioajan sijaan tulemme näkemään hitaan ja vaiheittaisen, hyvin asteittaisensiirtymän, esimerkiksi vihreän- tai digisiirtymän tapaisen. Samalla monitahtisuus ja eri maiden väliset erot tulevat epäilemättä olemaan merkittävä osa siirtymää. Liikkumisen edelläkävijät tulevat ajamaan teknologian sovellutuksia nopeatempoisen ja joustavan regulaation sekä merkittävien investointien kautta (kuten Kuvan 1 Autonominen taksi, Cruise¹⁹), toisten järjestelmien pitäytyessä maltillisemmassa etenemisessä.

¹⁶ Othman, K. (2021). Impact of Autonomous Vehicles on the Physical Infrastructure: Changes and Challenges. *Designs*, Vol. 5(3). <https://doi.org/10.3390/designs5030040>

¹⁷ ERTRAC Task Force (2015). Automated Driving Roadmap. https://erticonetwork.com/wp-content/uploads/2015/03/www.etrac.org_uploads_documentsearch_id35_ERTRAC_Automated-Driving_draft3-web.pdf (20.3.2023)

¹⁸ Dolata, U. (2011). "Radical change as gradual transformation: Characteristics and variants of socio-technical transitions," *Research Contributions to Organizational Sociology and Innovation Studies*, SOI Discussion Papers 2011-03, University of Stuttgart, Institute for Social Sciences, Department of Organizational Sociology and Innovation Studies.

¹⁹ <https://www.cnbc.com/2022/06/02/cruise-gets-green-light-for-commercial-robotaxis-in-san-francisco.html>; <https://www.getcruise.com/news/were-going-commercial>; <https://rodnevbros.com/no-front-seat-occupants-adventures-in-autonomous-ride-services/> (20.3.2023)



**Esimerkkitapaus:
Autonominen taksi
Cruise, San Francisco,
[Taso 4 autonomia]**

- Cruise saavutti merkittävän paulun 2.6.2022 tullessaan ensimmäiseksi suurkaupungissa toimivaksi autonomisia taksikyytejä tarjoavaksi kaupalliseksi palvelukseksi.
- Myös muut yritykset, kuten Googlen omistama Waymo testaavat autonomisia takseja useissa kaupungeissa, San Francisco mukaanlukien, mutta yksikään niistä ei vielä tarjoa palvelua avoimeen käyttöön.
- Cruise alkoi tarjoamaan kyytejä kaupallisesti 2.6.2022. Takseissa ei ole turvakuljettajaa. Taksi tilataan appilla, Uberin tyyliin.
- Tällä hetkellä Cruise tarjoaa palvelua vain yö-aikaan. Rajoitteita ovat että sen taksit eivät liiku ruuhkaisilla kaduilla, toimivat vain tietyissä osissa kaupunkia, välttämällä erityisesti valtteita, eivät saa ajaa yli 80km/h ja että ne ovat liikenteessä vain sääolosuhteiden ollessa selkeät.
- Cruisen saavutus on kuitenkin merkittävä siinä suhteessa että se on onnistunut luomaan itseajavan autopalvelun vähimmäisvaatimukset täyttävän tuotteen, eli siis MVP:n (minimum value productin).

Kuva 1. Autonominen taksi, Cruise ¹⁹

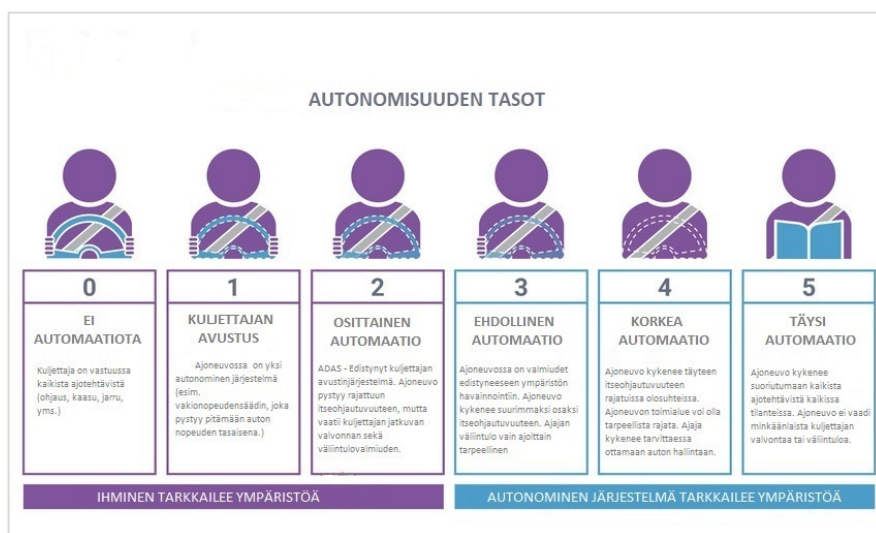
Keskeisimmät syyt sille, että autonomisia järjestelmiä ei tulla näkemään laajassa käytössä nopeasti eivät kuitenkaan johdu regulaatiosta, vaan siitä, että kyseessä ei ole yksinomaan teknologinen haaste. Itseasiassa itseajavien autojen yleistyminen tulee todennäköisesti vaatimaan samanaikaisesti useisiin yhteiskunnallisiin, teknologisiin, taloudellisiin ja sosiaalisiin ongelmiin vastaamista. Samalla eri näkökulmat ovat vahvasti keskinäisriippuvaisia, mikä tarkoittaa, että niiden ratkaisu tulee todennäköisesti vaatimaan myös suurempaa yhteiskunnallista koordinaatiota.

Yhteiskunnallisesta näkökulmasta autonomisten järjestelmien säätely, turvallisuuden, eettisyyden ja vastuun kysymykset ovat keskeisiä kipukohtia. Taloudellisella puolella sekä autonomisia järjestelmiä tukevan infrastruktuurin kustannukset, autokannan uudistumisnopeus, sekä autonomisilla ominaisuuksilla varustetun kaluston suuremmat kustannukset voidaan mainita muutamina keskeisinä haastekohtina. Teknologisesta näkökulmasta ajoympäristön havainnointi ja sensoriteknologian kehittyminen, erilaisissa ajo-olosuhteissa ja -ympäristöissä selviäminen, autonomisten mallien kehitys, kuten myös turvallisuuden, eettisyyden, kaupallisuuden ja säätelyn asettamat lisävaatimukset teknologiselle kehitykselle nousevat puolestaan esille. Lähes jokainen suuremmista

haasteellisuudesta sisältää myös vahvoja sosioteknisiä sidoksia ja näkökulmia.
20,21,22

2.1 Nykypäivän autonomisuuden tasot ajoneuvoissa

Itseohjautuvien ajoneuvojen autonomisuutta kuvataan usein asteikolla (Kuva 2), jossa tasolla 0 automaatiota ei ole lainkaan, ja puolestaan korkeimmalla tasolla 5 ajoneuvo kykenee täyteen automaatioon kaikissa olosuhteissa²³. Vaikka autonomisuuden tasojen määrittely eri liikennevälineissä on pääpiirteittäin sama, voi kuitenkin asteikossa olla vaihtelua (vrt. meriliikenne, Kuva 8).



Kuva 2. Autonomisuuden tasot (Lähde: Synopsys²³)

²⁰ Sifakis, J. (2021). Why is it so hard to make self-driving cars? (Trustworthy autonomous systems). WI-IAT '21: IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. <https://doi.org/10.1145/3486622.0000002>

²¹ Ernst, R. & Hu, X. S. (2021). Autonomous Systems Design—A Virtual Roundtable. Computer, vol. 54(11), 17-25. doi: 10.1109/MC.2021.3101914.

²² Ali Assaad, M.; Talj, R.; Charara, A. (2019). Autonomous Driving as System of Systems: roadmap for accelerating development. 2019 14th Annual Conference System of Systems Engineering (SoSE), 102–107. <https://doi.org/10.1109/SYSOSE.2019.8753886>

²³ <https://www.synopsys.com/automotive/autonomous-driving-levels.html> (20.3.2023)

Autonomisuuden tasoja käytetään autonomisten ajoneuvojen karkeaan luokitteluun, tyypillisesti henkilöautojen kontekstissa. Tasot 0–2 kuvaavat tilanteita, joissa ihmisen rooli ajamisessa on vahvempi ja autojen sensoriteknologia sekä autonomisuus matalampaa. Lähes kaikki markkinoilla olevat henkilöautot myös sisältyvät näihin kategorioihin. Muun muassa jo monesta nykyautosta löytyvät adaptiivinen vakionopeudensäädin ja kaista-avustin ovat tason 1 autonomiaa, ja tietyissä tilanteissa toimivat auton ohjausta tukevat toiminnot, esim. auton pitäminen kaistalla automaattisesti, on puolestaan tason 2 autonomiaa.²⁴ Tasoilla 3–5 autonomisen järjestelmän tulee pystyä itsenäisempään toimintaan sekä päätöksentekoon. Kuluttajamarkkinoilla on nähtävissä joitain rajattuja esimerkkejä tason 3 autonomisuuteen kykenevistä autoista. Tällä hetkellä korkeimmalla autonomisella tasolla olevat ajoneuvot on rajattu tason 4 laitteisiin (Ks. Esim. Cruise-tapauseimerkki). Tason 5 autonomisen ajoneuvon tulisi kyetä selviämään kaikista olosuhteista sekä ajoympäristöistä ilman ihmisen väliintuloa. Tällä hetkellä ei kuitenkaan ole täysin selvää edes, miten tällaiseen kykenevän tekoälyn kehityksessä olevia ongelmia tulisi lähteä ratkaisemaan, puhumattakaan sen saavuttamisesta lähitulevaisuudessa, tai pidemmällä aikavälillä.²⁵

²⁴ <https://moottori.fi/ajoneuvot/jutut/niin-mika-oli-mita-eri-autonomisen-ajamisen-tasot-tarkoittavat-ja-missa-nyt-mennaan/> (20.3.2023)

²⁵ Wang et al. (2021). Towards the Unified Principles for Level 5 Autonomous Vehicles. *Engineering*, Vol. 7(9), 1313-1325. <https://doi.org/10.1016/j.eng.2020.10.018>.

3 Käyttäjien odotukset autonomista liikkumista kohtaan

3.1 Odotusten sosiologia

Tulevaisuuteen suuntautuvaa keskustelua esimerkiksi autonomisesta liikenteestä voidaan lähestyä muun muassa odotusten sosiologian näkökulmasta ^{26,27,28,29}. Näkökulma keskittyy tieteen ja teknologian tulevien muutosten ymmärtämiseen odotusten tuotannon ja niiden kierrättämisen kautta. Odotusten sosiologia on tiede- ja teknologiatutkimukseen liittyvä käsite, joka on kiinnostunut esimerkiksi uusien teknologioiden ja innovaatioiden leviämisestä sekä näiden uutuuksien yhteiskunnalle ja taloudelle aiheuttamista muutoksista ³⁰.

Odotukset, lupaukset, kuvitelmat ja visiot maalaavat kuvia toivotusta tulevaisuudesta ³¹. Meidän on hyvä ymmärtää odotuksiin liittyvää dynamiikkaa, jotta ymmärrämme teknologioiden luomaa muutosta yhteiskunnassa, sillä odotukset ohjaavat eri toimijoiden toimintaa tieteen ja teknologian alalla. Odotukset legitimoivat, herättävät kiinnostusta, ohjaavat rahoitusta ja auttavat tunnistamaan

26 Brown, N. and Michael, M. (2003). A Sociology of Expectations: Retrospecting Prospects and Prospecting Retrospects. *Technology Analysis and Strategic Management*, 15(1), 3-18.

27 Berkhout, F., (2006). Normative expectations in systems innovation. *Technology Analysis & Strategic Management*. Vol. 18 (3-4), 299-311.

28 Borup et al., (2006) The Sociology of Expectations in Science and Technology. *Technology Analysis & Strategic Management*. Vol 18 (3-4), 285-298.

29 van Lente, Harro (2012) Navigating foresight in a sea of expectations: lessons from the sociology of expectations, *Technology Analysis & Strategic Management*, 24:8, 769-782, DOI: 10.1080/09537325.2012.715478

30 Brown, N. & Michael, M. (2003). A Sociology of Expectations: Retrospecting Prospects and Prospecting Retrospects. *Technology Analysis and Strategic Management*, 15(1), 3-18.

31 Borup et al., (2006) The Sociology of Expectations in Science and Technology. *Technology Analysis & Strategic Management*. Vol 18(3-4), 285-298.

mahdollisia uhkia ja mahdollisuuksia ³². Odotukset voivat olla positiivisia tai negatiivisia, sekä toiveita että pelkoja. Ne vähentävät epävarmuutta markkinoilla, joilla toimijoiden on tehtävä tulevaisuutta koskevia päätöksiä ja valittava potentiaalisimmat teknologiat.

Useista mahdollisista kuljetusteknologian kehityspoluista autonomiset järjestelmät on valittu yhdeksi merkittävimmistä uusista teknologioista. Ennustus autonomisista ajoneuvoista on mobilisoitunut rahoitusta ja jälkeen jäämisen pelossa alan yritysten on otettava odotukset vakavasti.

Vaikka visiot ja odotukset eivät toteudu niin teknologisen kehityksen varmistamiseksi ja eteenpäin viemiseksi tarvitaan myönteisiä odotuksia, lupauksia, ja kuvitelmia maalaamaan kuvia toivotusta tulevaisuudesta ja siitä, miten sinne päästään. Näitä sosioteknisiä kuvitelmia pitää rakentaa ja myös vaalia, jotta ihmisten, teknologian kehittäjien ja rahoittajien kiinnostus säilyy, ja kehityskulku jatkuu myös tulevaisuudessa. Odotusten sosiologian näkökulmasta tekoälyn käyttöönoton onnistumiseksi täytyy rakentaa kollektiivisia odotuksia, jotta kaikki sidosryhmät saadaan tekoälyn puolelle.

Jaettu hyöty: Tekoälyteknologioiden tulisi hyödyttää ja antaa voimaa mahdollisimman monelle ihmiselle.

(Asilomar tekoälyn periaatteet No 14)

3.2 Arvot ja etiikka

Odotuksiin liittyy oleellisesti myös arvot ja etiikka. Etiikka muodostuu moraalisisista periaatteista, jotka ohjaavat ihmisen käyttäytymistä ja toimintaa. Etiikka ja arvot ovat yhteiskunta- ja kulttuurisidonnaisia, ja muodostuvat sosiaalisessa kanssakäymisessä. Koska monet tekoälyratkaisut ja -teknologiat ovat vielä tulevaisuuden teknologioita, kuten täysin autonomiset autot tai laivat, näiden ratkaisujen hyväksyttävyyden ja haluttavuuden edellyttävät laajaa ja monipuolista eettistä keskustelua yhteiskunnassa. Tekoälyn etiikassa pyritään hahmottamaan ja ratkaisemaan tekoälyn kehitystä ja hyödyntämistä koskevia eettisiä kysymyksiä³³.

³² van Lente, Harro (2012) Navigating foresight in a sea of expectations: lessons from the sociology of expectations, *Technology Analysis & Strategic Management*, 24:8, 769-782, DOI: 10.1080/09537325.2012.715478

³³ Pöntinen, H. (2019). Tekoälyn etiikkaa - Valtiot tekoälyn hyödyntäjinä. Helsingin yliopisto, Teologinen tiedekunta, Pro gradu.

Tekoälyteknologiat ovat hiljattain saavuttaneet pisteen, jossa niillä on laajempi yhteiskunnallinen merkitys ja samalla huolenaihe. Laajempi merkitys johtuu esimerkiksi siitä, että julkishallinto on paremmin kykenevä hyödyntämään valtavia datamääriä automatisoidussa päätöksenteossa ja yritykset uusissa liiketoimintamalleissa.³⁴ Monia ihmisiä kuitenkin huolettaa esimerkiksi arkaluonteisten henkilötietojen kerääminen ja niiden käyttö tekoälyn suunnittelussa.

Arvojen kohdistaminen: autonomiset tekoälyjärjestelmät tulee suunnitella siten, että niiden tavoitteet ja käyttäytyminen voidaan taata ihmisten arvojen mukaisiksi koko niiden toiminnan ajan

(Asilomar tekoälyn periaatteet No 10)

Tekoälyn eettinen keskustelu liittyy muun muassa vinoutumiin ohjelmointivalinnoissa ja tekoälyjärjestelmän kouluttamiseen käytettävässä datassa, esim. sukupuoleen tai etniseen taustaan liittyen, jotka saattavat johtaa epätasa-arvoa tuottaviin ja toisintaviin järjestelmiin³⁵. Useat tekoälyjärjestelmät kasvojentunnistuksesta puheentunnistukseen ovat usein epätarkempia ei-valkoihoisten, naispuolisten tai nuorten ihmisten kohdalla^{36,37,38,39}. Vaikka eettisen ja ihmislähtöisen teknologian suunnittelun suuntaviivat ja periaatteet ovat laajalti

³⁴ Stahl, B.C., (2022). Responsible innovation ecosystems: Ethical implications of the application of the ecosystem concept to artificial intelligence. *International Journal of Information Management*, 62, 102441.

³⁵ Stahl, et al. (2021). Organisational responses to the ethical issues of artificial intelligence. *AI & SOCIETY*, 37, 23–37. <https://doi.org/10.1007/s00146-021-01148-6>

³⁶ Criado Perez, C. (2019). *Invisible Women: Data Bias in a World Designed for Men*. Abrams Press.

³⁷ Waelen, R.A. (2023). The struggle for recognition in the age of facial recognition technology. *AI and Ethics*, Vol. 3, 215–222. <https://doi.org/10.1007/s43681-022-00146-8>

³⁸ Klare, et al. (2012). Face Recognition Performance: Role of Demographic Information. *IEEE Transactions on Information Forensics and Security*, Vol. 7(6), 1789-1801. doi: 10.1109/TIFS.2012.2214212.

³⁹ Tatman, R. (2017). Gender and Dialect Bias in YouTube's Automatic Captions. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, p. 53–59, Valencia, Spain. Association for Computational Linguistics. DOI: 10.18653/v1/W17-1606

saatavilla (esim. EU:n eettiset suuntaviivat⁴⁰; OECD⁴¹; Asilomar: tekoäly periaatteet⁴²), edelleen on haastavaa lisätä yhteiskunnallista keskustelua tekoälyn etiikasta ja arvoista. Tätä keskustelua tarvitaan, jotta uudet tekoälyratkaisut, kuten autonominen liikenne, olisivat haluttuja ja hyväksytyjä ratkaisuja yhteiskunnassa.

3.3 Käyttäjät ja arvot

Uusi teknologia voi herättää ennakkoluuloja ja pelkoja käyttäjissä, jolloin riskinä on järjestelmän käyttämättömyys. Uusien teknologisten ratkaisuiden, kuten tekoälyn avulla saavutettava autonomisuus, kehityksessä on hyvä pitää mielessä ihmislähtöisen teknologian periaatteet⁴³ sillä teknologian taustalla vaikuttava ihmiskuva määrittelee teknologian kehityssuuntaa. Autonomisten järjestelmien suunnittelussa yksi haaste liittyy tavallisten käyttäjäryhmien, kuten julkisen liikenteen matkustajien tai yksityisautoilijoiden, toiminnan ja tarpeiden huomioimiseen.

Onkin tärkeää tutkia ennakkoon teknologian vastaanottoa potentiaalisissa käyttäjissä, jotta uudenlaiset autonomiset järjestelmät olisi mahdollista suunnitella käyttäjien ajatuksia, tarpeita ja mahdollisia pelkoja kuunnellen, käyttäjien hyväksynnän ja ratkaisun houkuttelevuuden lisäämiseksi. Keskeisiä kysymyksiä ovat esimerkiksi mitä arvoja tekoälyn hyödyntämisellä halutaan tavoitella? Miten tekoälyn käyttöönotto vaikuttaa yksilöön, yhteisöihin ja laajemmin yhteiskuntaan?⁴⁴.

Kuten Kuva 3 osoittaa, arvokeskustelu erityisesti autonomisen liikenteen ratkaisuihin on tärkeää mm. niin teknologisten ratkaisuiden luotettavuuden kuin käyttäjien kokeman luottamuksen osalta.

Vuoropuhelun puuttuminen teknologioiden kehittäjien ja käyttäjien väliltä voi johtaa tilanteeseen, jossa vanhojen teknologioiden korvaaminen uusilla autonomisilla järjestelmillä koetaan tarpeettomaksi ja ylhäältä annetuksi.

⁴⁰ Euroopan Komissio (2019). LUOTETTAVAA TEKÖÄLYÄ KOSKEVAT EETTISET OHJEET. https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI_FI.pdf (20.3.2023)

⁴¹ OECD (2020). Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449> (20.3.2023)

⁴² Asilomar AI Principles. <https://futureoflife.org/2017/08/11/ai-principles/?cn-reloaded=1> (20.3.2023)

⁴³ Leikas, J. (2009). Life-Based Design: A holistic approach to designing human-technology interaction. VTT Technical Research Centre of Finland. VTT Publications No. 726, <https://publications.vtt.fi/pdf/publications/2009/P726.pdf>

⁴⁴ Tällöin puhutaan hyödyntämisen etiikasta, ks. esim. Pöntinen (2019).



Kuva 3. Liikenteen autonomisuuden arvokeskustelun näkökulmat. Lähde: kirjoittajat.

Kun mahdolliset ongelmat ovat tiedossa jo varhain, on niihin puuttuminen helpompaa ja edullisempaa vielä kehitysvaiheessa. Kehittäjäpuolen ratkaisujen ja käyttäjien tarpeiden kohtaamisen tulisikin olla suunnittelun lähtökohtana. Näin ollen yksi tapa selvittää käyttäjien tarpeita on tutkia, mitkä arvot ovat käyttäjilleen tärkeitä.

Inhimilliset arvot:
Tekoälyjärjestelmät tulee suunnitella ja käyttää siten, että ne ovat yhteensopivia ihmisarvon, oikeuksien, vapauksien ja kulttuurisen monimuotoisuuden ihanteiden kanssa.

(Asilomar tekoälyn periaatteet No 11)

3.4 Arvokartta

Yksi tunnettu lähestymistapa arvojen tutkimukseen on Schwartzin teoria arvojen yleismaallisesta sisällöstä ja rakenteesta vuodelta 1992⁴⁵. Pitkän empiirisen, monia eri maita ja kulttuureita kattavan työn tuloksena on pyritty löytämään universaaleja arvoja, jotka ovat merkitykseltään sekä keskinäisiltä suhteiltaan samanlaisia eri kulttuureissa. Teoriaan lukeutuu kymmenen eri arvoluokkaa, jotka

⁴⁵ Helkama, K., Myllyniemi, R., & Liebkind, K. (2010). Johdatus sosiaalipsykologiaan. Helsinki, Suomi: Edita.

ovat seulonnan tulosta – teoriasta siis puuttuvat arvot, joilla on eri merkitys eri kulttuureissa. Teorian arvoluokkia ovat valta, suoriutuminen, mielihyvä, virikkeisyys, itsenäisyys, universalismi, hyväntahtoisuus, perinteet, yhdenmukaisuus ja turvallisuus. Luokat ja niiden sisältämät arvot on esitelty tarkemmin Taulukossa 1.

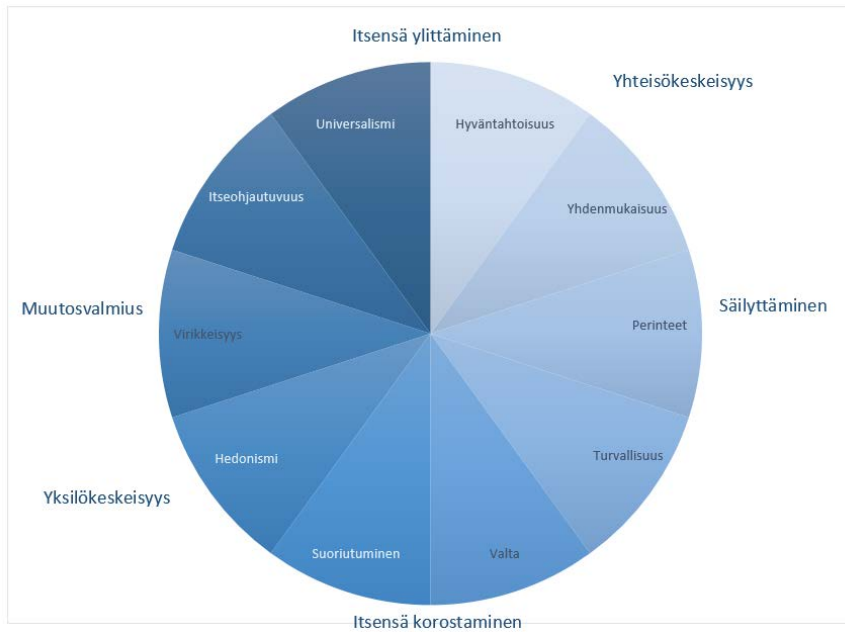
Taulukko 1. Schwartzin teorian arvoluokat ja niihin sisältyvät arvot⁴⁶

Arvoluokka	Arvot
Hyväntahtoisuus	Auttavaisuus, rehellisyys, anteeksiantavaisuus, vastuuntunto
Yhdenmukaisuus	Tottelevaisuus, vanhempien kunnioittaminen, kohteliaisuus, itsekuri
Perinteet	Oman elämänosan hyväksyminen, nöyryys, perinteiden kunnioittaminen, kohtuullisuus
Turvallisuus	Puhtaus, kansallinen turvallisuus, palvelusten vastavuoroisuus, yhteiskunnan järjestys, perheen turvallisuus
Valta	Yhteiskunnallinen valta, arvovalta, varallisuus
Suoriutuminen	Menestys, kyvykkyys, kunnianhimo, vaikutusvalta
Hedonismi	Mielihyvä, elämästä nauttiminen
Virikkeisyys	Uskaliaisuus, vaihteleva elämä, jännittävä elämä
Itseohjautuvuus	Uteliaisuus, luovuus, vapaus, omien tavoitteiden valitseminen, itsenäisyys
Universalismi	Ympäristönsuojelu, yhteys luontoon, luonnon ja taiteen kauneus, suvaitsevaisuus, viisaus, tasa-arvo, maailmanrauha

Universaalia rakennetta koskeva oletus liittyy arvojen keskinäiseen suhteeseen eli siihen, että tietyt arvot ovat keskenään lähekkäisiä tai vastakkaisia kaikissa kulttuureissa. Näitä suhteita havainnollistetaan arvokehikolla (Kuva 4), jolle arvoluokat asettuvat sen mukaisesti, minkälaisia tarpeita ja päämääriä ne palvelevat.

⁴⁶ Helkama, K., Myllyniemi, R., & Liebkind, K. (2010). Johdatus sosiaalipsykologiaan. Helsinki, Suomi: Edita.

Kehikkoa jäsentävät päämäärät ovat toistensa vastakohtia: itsensä ylittäminen vs. itsensä korostaminen, muutosvalmius vs. säilyttäminen ja yhteisökeskeisyys vs. yksilökeskeisyys. Kehikossa ylhäällä oikealla olevat arvoluokat liittyvät yhteisökeskeisiin seikkoihin, kun taas vasemman puolen alempana sijaitsevat arvot ovat yksilökeskeisiä. Arvoluokista turvallisuus ja universalismi ovat luonteeltaan sekaluokkia, niihin sisältyy sekä yksilöllisiä että yhteisöllisiä arvoja.⁴⁷



Kuva 4. Schwartzin (1992) teorian arvoluokat ja niiden väliset suhteet kehällä.

⁴⁷ Helkama, K., Myllyniemi, R., & Liebkind, K. (2010). Johdatus sosiaalipsykologiaan. Helsinki, Suomi: Edita.

4 Tekoälyyn liittyvien arvojen tutkiminen

Tutkimusaineisto arvojen tarkastelua varten on kerätty virtuaalisesti järjestetyissä neljässä käyttäjätyöpajassa vuosina 2021 ja 2022. Työpajoihin osallistui yhteensä kaksikymmentä henkilöä. Virtuaaliset työpajat pidettiin tarkoituksellisesti pieninä keskustelun sujuvuuden vuoksi. Työpajoissa osallistujat keskustelivat erilaisista heille esitetyistä tulevaisuuskenaariosta, osallistuivat arvokeskusteluun sekä kuvasivat omia näkemyksiään ja suhtautumisiaan esitettyihin tulevaisuuksiin. Tulevaisuuskenaariot kuvasivat esimerkkitapauksia ja mahdollisia tulevaisuuden kehityskulkuja, jotka rakennettiin työpajan teeman mukaan. Raitiovaunuliikennettä ja autoilua koskevissa työpajoissa osallistujat olivat demografisilta piirteiltään työssäkäyviä, pääosin naispuolisia henkilöitä eri puolilta Suomea, laivaliikennettä koskevassa keskustelussa oli puolestaan ainoastaan miespuolisia henkilöitä. Osallistujat rekrytoitiin työpajoihin sähköpostitse ja / tai sosiaalisessa mediassa jaettujen hakuilmoitusten kautta. Osallistujia ei kompensoitu työpajoihin osallistumisesta mutta heidän kesken arvottiin elokuvalippuja. Keskustelut nauhoitettiin ja kirjattiin myöhempää analyysia varten.

Schwartzin arvoteoriaa on käytetty työpajoista kerätyn aineiston analyysissa löyhänä viitekehyksenä kuvaamaan niitä ajatuksia ja arvoja, joiden pohjalta keskustelua kuljettajattomasta raitiovaunusta ja autonomisista autoista sekä laivoista käytiin, ja jotka ovat näin ollen merkityksellisiä tulevaisuuden liikenteen kontekstissa käyttäjilleen. Aineistonkeruussa ei ole käytetty Schwartzin teorialle tyypillisiä arvomittareita, vaan tulokset perustuvat laadullisesta aineistosta tehtyyn sisällönanalyttiseen tulkintaan ja vastausten temaattiseen luokitteluun.

5 Käyttäjien arvot: Raitiovaunuliikenne

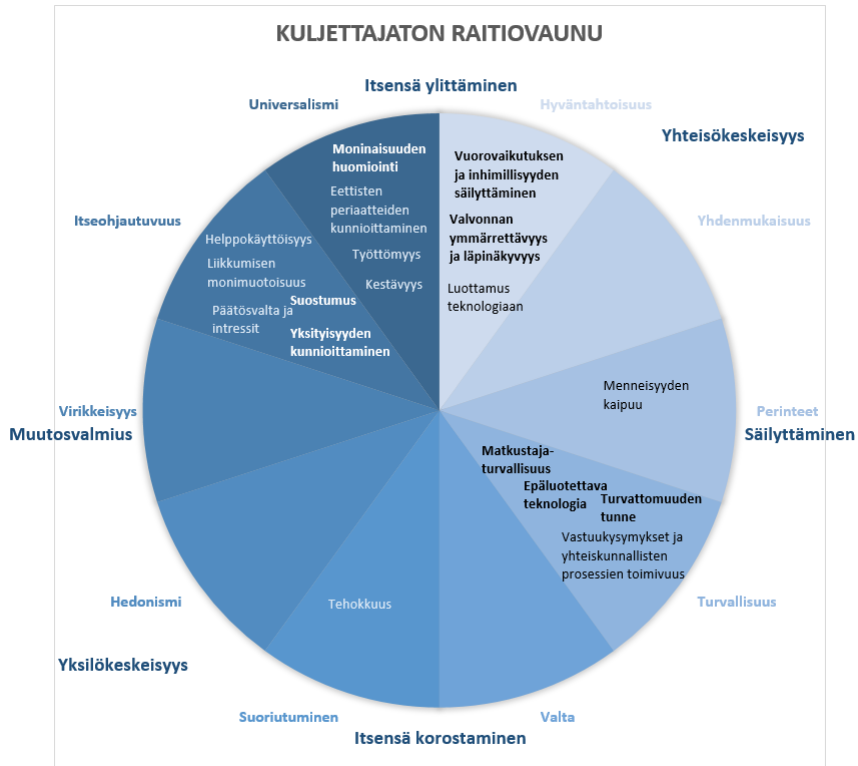
Raitiovaunu on keskeinen osa kaupunkiliikennettä ja julkisen liikenteen toimivuutta. Suomessa raitiovaunuliikennettä on käytössä Tampereella sekä Helsingissä. Raitiovaunut ovat tällä hetkellä Suomessa matalan automaation liikennevälineitä. Kuskien apuna ei ole merkittäviä ohjaus- tai automaatiojärjestelmiä. Keskeisimmät teknologiset apuratkaisut liittyvät näkyvyyden parantamiseen kameroilla sekä informaation antamiseen kuskille raitiovaunun nopeudesta ja aikataulusta. Raitiovaunut ovat kuitenkin yksi tulevaisuuden automaation kohde. Yksi itseohjautuvan raitiovaunun ensimmäisistä esimerkeistä oli Siemensin vuonna 2018 Saksassa demonstroima prototyyppi, jonka odotetaan valmistuvan markkinoille 2026. Raitiovaunu kykenisi täysin itsenäiseen ohjautuvuuteen ja kuljettajan rooli siirtyisi etävalvomoon.

Työpajan osallistajat keskustelivat erilaisista raitiovaunun tulevaisuuskuvista ja siitä millaisia ajatuksia ja suhtautumisia autonominen raitiovaunu herättää. Työpajassa keskityttiin erityisesti Tampereen raitiovaunuliikenteeseen, joka oli aloittanut toimintansa muutama kuukausi ennen työpajaa vuonna 2021. Työpajassa esitettiin kolme raitiovaunuliikenteen skenaariota muutaman minuutin pituisina videoina:

- Skenaario 1: Nykypäivän matkustaminen. Tässä skenaariossa mietittiin erityisesti sitä miltä raitiovaunumatkustaminen tuntuu, kun raitiovaunussa on suhteellisen vähän autonomisuutta ja kuljettaja on paikalla.
- Skenaario 2: Lisääntynyt autonomisuus. Tässä skenaariossa raitiovaunussa on lisääntynyttä matkustajien ja kuljettajan valvontaa mutta kuljettaja on paikalla.
- Skenaario 3: Autonominen raitiovaunu. Viimeisessä skenaariossa keskusteltiin siitä miltä matkustaminen autonomisessa etävalvotussa raitiovaunussa tuntuisi.

Esitetyt kolme skenaariota muodostavat yhdessä kuljettajaton raitiovaunu tapauksen, jota avataan Schwartzin arvokehän avulla. Keskusteluista tunnistettiin erilaisia arvoja, jotka asetoitiin eri arvoluokkiin. Kaikki keskusteluissa nousseet arvot on esitetty Kuvassa 5. Seuraavaksi käymme läpi eniten keskustelua

herättäneet arvot, jotka näyttäytyivät merkityksellisimpinä keskustelijoille. Kaikkien arvojen kuvaukset löytyvät Liitteestä 1.



Kuva 5. Kuljettajattomaan raitiovaunuun liittyvät arvot Schwartzin arvokehällä.

Kuljettajatonta raitiovaunua koskevat arvot sijoittuivat suurelta osin kehällä itsensä ylittämisen ja säilyttämisen ulottuuksille, ja arvot olivat luonteeltaan pitkälti yhteisiä päämääriä edistäviä. Arvoluokista korostuivat keskusteluissa eniten **hyväntahtoisuus, turvallisuus, itseohjautuvuus** ja **universalismi** (Taulukko 2).

Taulukko 2. Arvoluokat kuljettajattomassa raitiovaunussa

HYVÄNTAHTOISUUS <ul style="list-style-type: none">• Vuorovaikutus ja inhimillisuus• Valvonnan ymmärrettävyys ja läpinäkyvyys	TURVALLISUUS <ul style="list-style-type: none">• Matkustajaturvallisuus• Epäluotettava teknologia• Turvattomuuden tunne
ITSEOHJAUTUVUUS <ul style="list-style-type: none">• Yksityisyyden kunnioittaminen• Suostumus	UNIVERSALISMI <ul style="list-style-type: none">• Moninaisuuden huomioiminen

5.1 Hyväntahtoisuus

Hyväntahtoisuus on toisten ihmisten ja yhteisön hyvinvoinnin edistämistä ja priorisoimista. Luokkaan sisältyy arvoja kuten auttavaisuus, rehellisyys, anteeksiantavuus, vastuullisuus, luotettavuus, tosiystävyyden ja kypsä rakkaus. Keskusteluissa kuljettajattomasta raitiovaunusta käyttäjille tärkeimpiä arvoja olivat vuorovaikutuksen ja inhimillisyyden säilyttäminen sekä valvonnan ymmärrettävyys ja läpinäkyvyys.

Vuorovaikutus ja inhimillisuus ovat matkustuskokemuksen kannalta tärkeitä arvoja, jotka halutaan säilyttää tulevaisuudenkin joukkoliikenteessä. Ihmiskuljettaja nähtiin keskusteluissa merkityksellisenä osana julkista liikennettä, ja kuljettajan puuttumisen tai siirtymisen etäyhteyden päähän uskottiin tekevän kokemuksesta kalsean ja etäännyttävän. ”Heiluttelevat kuljettajat” ovat jättäneet ”inhimillisen kosketuksen”, eikä pelkkien vartijoiden läsnäolon uskottu tuovan samaa myönteistä tunnetta vaan päinvastoin synnyttävän ikävän tunteen matkustajille. Käyttäjät siis toivoivat jatkossakin matkustuskokemukselta vuorovaikutteisuutta, ja ettei teknologia korvaisi kaikkea inhimillistä kohtaamista.

Valvonnan ymmärrettävyys ja läpinäkyvyys olivat asioita, joita keskustelijat arvostivat kuljettajattomassa raitiovaunussa. Erityisesti autonomisen teknologian myötä lisääntyvä valvonta ja tiedonkeruu herättivät runsaasti kysymyksiä valvonnan perusteltavuudesta, ja tärkeänä pidettiin sitä, että tiedonkeruu on kansalaisille ymmärrettävää ja läpinäkyvää. Käyttäjille tulisi siis selkokielellä kertoa, mitä tietoa heistä kerätään ja miksi. Valvonnan kohteena olemiseen ei suhtauduttu ehdottoman kielteisesti, mikäli valvonta katsottiin perustelluksi ja matkustajille itselleen hyödylliseksi.

5.2 Turvallisuus

Turvallisuus on oman ja läheisten elämän tasaisuutta ja harmonista jatkuvuutta sekä yhteiskunnan vakautta. Luokkaan sisältyy arvoja kuten yhteiskunnallinen järjestys, perheen turvallisuus, kansallinen turvallisuus, puhtaus ja palvelusten vastavuoroisuus. Kuljettajattoman raitiovaunun turvallisuutta pidettiin tärkeänä ja se näkyikin monipuolisena keskusteluna, jota turvallisuudesta käytiin sekä yhteiskunnallisella että henkilökohtaisella tasolla. Tärkeimpiä esiin nousseita arvoja olivat matkustajaturvallisuus, epäluotettava teknologia ja turvattomuuden tunne.

Matkustajaturvallisuus näyttäytyi tärkeänä arvona keskusteluissa ja sitä käsiteltiin monelta kantilta raitiovaunun etäoperointia ajatellen. Keskustelijat samastuivat matkustajien asemaan ja olivat huolissaan, pystyykö kuljettaja todella valvomaan raitiovaunua riittävän hyvin etäältä ja miten matkustaja saa vaivattomasti yhteyden henkilökuntaan tai apua, jos jotakin sattuu. Turvallisuuden pelättiin kuljettajan läsnäolon puuttumisen myötä heikkenevän, mutta uuden edistyneemmän teknologian uskottiin toisaalta jopa lisäävän turvallisuutta. Tekoäly saattaisi jopa mahdollistaa nopeamman reagoinnin sairaskohtauksen sattuessa ja autopilotti vähentäisi inhimillisiä virheitä. Vertailtaessa autonomisiin autoihin, kuljettajaton raitiovaunu tuntui turvallisemmalta ja hyväksyttävämmältä.

Epäluotettava teknologia huoletti turvallisuuden näkökulmasta. Teknologiaa pidettiin alttiina virhetulkinnoille ja sen kykyä yhtäältä reagoida eteen juokseviin lapsiin, ja toisaalta suojautua haastavilta sääolosuhteilta ja ilkeiltä, epäiltiin. Tekoälyn ei katsottu osaavan tulkita muiden kulkijoiden aikeita eikä sen paremmin ratkaista muitakaan ongelmia, joihin kuljettaja on osannut reagoida nopeasti. Turvaa toisi mahdollisesti teknologian ja ihmisen yhteistyö, johon edelleen turvaudutaan lentokoneissakin:

“kyllä me pystytään lentokonekin lentämään jo täysin autonomisesti, mutta kyllä edelleen backuppi halutaan pitää, ja niin tässä ratikassakin haluaisin backupin, että siellä on teknologiaa ja siellä on ihminen”.

(Raitiovaununutyöpajan osallistuja)

Turvattomuuden tunne nähtiin erillisenä varsinaisesta turvallisuudesta, mutta sitä pidettiin yhtä lailla tärkeänä kuljettajattonta raitiovaunua arvioitaessa. Etäkuljettajan ei uskottu tuovan samanlaista turvallisuuden tunnetta kuin mitä kuljettaja omalla läsnäolollaan tuo. Vaikka kuljettajan ei uskottukaan voivan paljoakaan vaikuttaa väkivaltatilanteessa, tunnetta siitä, että joku on tilanteessa vastuussa, arvostettiin.

5.3 Itseohjautuvuus

Itseohjautuvuudella tarkoitetaan toiminnan ja ajattelun vapautta, kuten omien päämäärien valintaa sekä vapautta kehittää omia ideoita ja luoda uutta. Luokkaan kuuluu arvoja kuten luovuus, vapaus, omien tavoitteiden valitseminen, uteliaisuus, riippumattomuus. Keskusteluissa tärkeimpiä itseohjautuvuuteen liittyviä arvoja olivat yksityisyyden kunnioittaminen ja suostumus.

Yksityisyyden kunnioittamista pidettiin keskusteluissa tärkeänä käyttäjiin kohdistuvan lisääntyvän monitoroinnin vuoksi. Pelko kerättyjen tietojen väärinkäytöstä herätti ahdistusta ja tietojen keruun tarkoituksesta oltiin kiinnostuneita. Käyttäjille oli merkityksellistä ymmärtää, kenen hyväksi tietoja kerätään, mitä tiedolla tehdään ja kuka niihin pääsee käsiksi. Keskusteluissa luotiin reunaehtoja tiedonkeruulle, pohdittiin muun muassa sitä mitä tietoja oltiin valmiita itsestä luovuttamaan ja millä ehdoin. Henkilökohtaisen ja tunnistettavan tiedon jakamisen sekä tallentavan valvonnan suhteen oltiin varautuneempia. Keskustelijat pohtivat myös valvonnan rajoja:

“rajoittuuko valvonta ratikan sisälle vai myös sen ulkopuolelle, esim. pysäkeille? Entä ohikulkijat tai penkeillä nukkuvat?”
(Raitiovaunutyöpajan osallistuja)

Suostumus ja mahdollisuus kieltäytyä valvonnasta nousi tärkeäksi arvoksi keskusteluissa. Kerättyjen tietojen väärinkäyttöä pidettiin merkittävänä riskinä lisääntyvän monitoroinnin vuoksi, ja täten osallistujille oli tärkeää pohtia mahdollisuutta kieltäytyä valvonnasta. Huolta herätti kysymys, voiko profiloinnista kieltäytyä ja miten. Jos henkilötietoja kerättäisiin, missä vaiheessa hyväksyntä annettaisiin ja miten laajalle raitiovaunun ympäristöön valvonta ulottuisi? Siinäkin tapauksessa, että suostumusta kysyttäisiin, ei siihen vastaamista pidetty yksinkertaisena. Keskustelijoiden omista aiemmista kokemuksista kumpusi ajatus, että lupa seurata käyttäjää annetaan liian helposti, sillä sopimuksia ei jakseta lukea tai kiireessä tulee tehtyä virhevalintoja evästeiden suhteen.

5.4 Universalismi

Universalismi on laajempaa, myös muiden kuin omaan lähipiiriin kuuluvien ihmisten sekä luonnon arvostamista. Universalismiin liittyviä arvoja ovat esimerkiksi laajakatseisuus, ta-sa-arvo, sosiaalinen oikeudenmukaisuus, maailmanrauha, kauneuden maailma, yhteys luontoon, kypsä elämänymmärrys ja ympäristön suojeleminen. Keskeisimpänä arvona universalismissa nähtiin moninaisuuden huomiointi.

Moninaisuuden huomiointia pidettiin tärkeänä, ja keskustelijat tarkastelivatkin eri käyttäjäryhmien näkökulmasta kuljettajattoman raitiovaunun saatavuutta ja käytettävyyttä. Tietoa vapaista paikoista ei pidetty kovin tärkeänä tavalliselle

matkustajalle, mutta lastenvaunujen kanssa kulkevalle tieto saattaisi helpottaa kulkemista. Ikäihmisille taas tiedonkeruun epäiltiin tuntuvan vieraalta ja hyötyjen vaikeilta ymmärtää, kuten eräs työpajaosallistuja asian ilmaisi:

“ennen kaikkea ajattelen ikäihmisiä, monikaan ei varmaan ilahdu, että kerätään tietoa koska tuntuu aika vieraalta ja kaukaiselta, ja hyödyt on vaikea ymmärtää”.

Yleisenä huolenaiheena pidettiin erityisryhmien jäämistä katveeseen asioissa, joita pidetään itsestäänselvyyksinä.

6 Käyttäjien arvot: Itseohjautuvat autot

Itseajavien autojen historia on pitkä, jo 1920-luvulla kehitettiin ensimmäinen kauko-ohjattava auto (Kuva 6). Ensimmäinen versio itseajavasta autosta ilmeni 1960-luvulla, jolloin erityisesti rakennettuun moottoritiehen upotetut sensorit mahdollistivat yksittäisen auton itseohjautuvuuden. LIDARin eli valotutkan käyttöä itseajavien autojen ohjauksessa kokeiltiin puolestaan 1980-luvulla, jolloin Yhdysvaltojen asevoimien tutkimusorganisaatio DARPA⁴⁸ ALV-projekti (autonomous land vehicle) pilotoi ensimmäisen tietokonenäköä, erilaisia sensoreita ja LIDAR-skannausta hyödyntävän auton. Tämä prototyyppi kykeni navigoimaan noin 30 kilometrin tuntinopeuksilla suljetussa kaupungissa, ja väistämään yksinkertaisia esteitä.

DARPA:n vuonna 2005 laukaisema Grand Challenge-haaste itseajaville autoille voidaan nähdä modernien itseajavien autojen kilpajuoksun laukaisijana. Se myötävaikutti tutkimuksen siirtymisestä pääasiassa yliopistoista ja valtion tutkimuslaitoksista kaupalliselle puolelle, esimerkiksi Googlen perustaessa tutkimusyksikkönsä 2009.

Saavuttaessa 2010-luvulle innostus itseajavia autoja kohtaan räjähti. Odotukset nopeasta kehityksestä ja yleistymisestä levisivät laajalle ja useat autonvalmistajat ilmoittivat itseajavien autojen olevan tuotantolinjoilla vuosikymmenen loppupuolella ja julkiset demonstraatiot herättivät merkittävää kiinnostusta teknologiaa kohtaan.

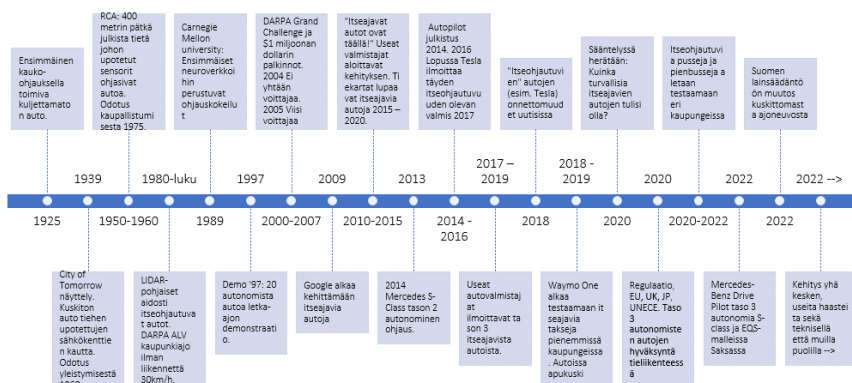
Kuitenkin toistuvat viivästymiset, valistuneempi käsitys teknisistä ongelmista sekä 2017 jälkeen alkanut sekä laajentunut uutisointi itseajavien autojen onnettomuuksista ovat hidastaneet odotuksia. Esimerkiksi Tesla on toistuvasti tehnyt korkealentoisia lupauksia vuodesta 2016 lähtien, mutta tästä huolimatta itseajavat Teslat eivät ole materialisoituneet.

Toisaalta vuosi 2020 merkitsi ajankohtaa, jolloin lainsäädäntö useiden kansainvälisten sekä kansallisten toimijoiden parissa alkoi edistyä merkittävästi ja esimerkiksi Euroopan Unioni, Yhdistyneet Kansakunnat (YK), Iso-Britannia ja Japani tekivät linjauksia. Myös Suomi päivitti lainsäädäntöään kuljettajatonta ajoneuvoa varten. Samalla itseohjautuvat pienbussit ja taksit ovat olleet rajatuissa testeissä useissa maissa. Suomessa vuoden 2022 keväällä Starshipin pienet kuljetusrobotit ovat tulleet osaksi katukuvaa tietyissä Espoon ja Helsingin osissa.

⁴⁸ Defense Advanced Research Projects Agency (DARPA)

Samana vuonna esim. Cruise sai luvan kuljettajattoman robotti-taksin kaupalliseen operointiin San Franciscossa ja Saksassa Mercedes-Benzin Drive Pilotista tuli ensimmäinen tason 3 hyväksytty autonominen kuluttaja-auto.

Itseajavien autojen historia



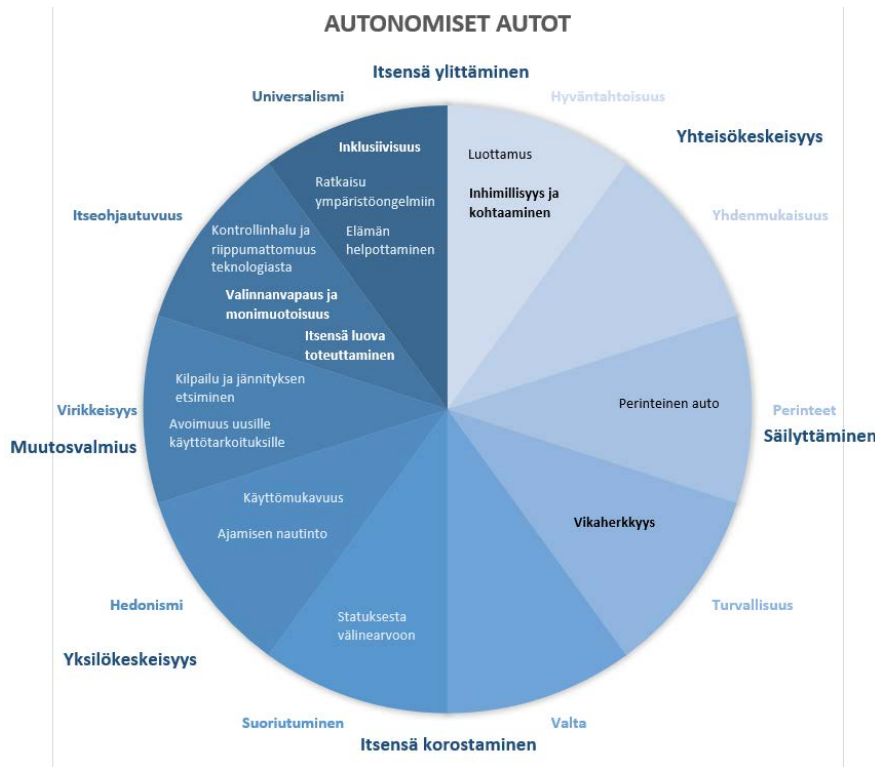
Kuva 6. Itseajavien autojen historia. Lähde: kirjoittajat

Itseajavia autoja käsittelevässä työpajassa osallistajat keskustelivat kahdesta skenaariosta, ja siitä millaisia ajatuksia ja suhtautumisia nämä herättivät.

- Skenaario 1: Nykypäivän autoilu ja historia. Työpajakeskustelu aloitettiin pohtimalla nykyisenkaltaiseen, matalan autonomian autoiluun liittyviä mielikuvia.
- Skenaario 2: Autonominen autoilu unelma: Tulevaisuuden itseajavat sähköautot, joissa käsitys autosta tilana muuttuu, eikä ihmisen osallistumista auton ohjaukseen enää tarvita.

Keskusteluista tunnistettiin erilaisia arvoja, jotka asemoitiin Schwartzin arvokehäille eri arvoluokkiin. Kaikki keskusteluissa nousseet arvot on esitetty Kuvassa 7. Seuraavaksi käymme läpi eniten keskustelua herättäneet arvot, jotka näyttäytyivät merkityksellisimpinä keskustelijoille. Kaikkien arvojen kuvaukset löytyvät Liitteestä 2.

Autonomisiin autoihin liittyvät arvot painottuivat itsensä ylittämiseen ja muutosvalmiuteen, mutta myös itseä korostavia ja muutosta vastustavia arvoja nousi työpajassa esille. Merkittävimpiä arvoluokkia olivat **hyväntahtoisuus**, **turvallisuus**, **itseohjautuvuus** ja **universalismi** (Taulukko 3).



Kuva 7. Autonomisiin autoihin liittyvät arvot Schwartzin arvokehällä.

Autonomisiin autoihin liittyvät arvot painottuivat itsensä ylittämiseen ja muutosvalmiuteen, mutta myös itseä korostavia ja muutosta vastustavia arvoja nousi työpajassa esille. Merkittävimpiä arvoluokkia olivat **hyväntahtoisuus, turvallisuus, itseohjautuvuus** ja **universalismi** (Taulukko 3).

Taulukko 3. Arvoluokat itseohjautuvissa autoissa

HYVÄNTAHTOISUUS <ul style="list-style-type: none">• Inhimillisyyden ja kohtaaminen	TURVALLISUUS <ul style="list-style-type: none">• Vikaherkkyys
ITSEOHJAUTUVUUS <ul style="list-style-type: none">• Itsensä luova toteuttaminen• Valinnanvapaus ja monimuotoisuus	UNIVERSALISMI <ul style="list-style-type: none">• Inklusiivisuus

6.1 Hyväntahtoisuus

Keskusteluissa autonomisista autoista käyttäjillä korostui inhimillisyyden ja kohtaamisen merkitys. Samanlaisia vuorovaikutuksen ja inhimillisyyden arvoja tuotiin esille myös raitiovaunuliikenteen kontekstissa.

Inhimillisyyden ja kohtaaminen näyttäytyivät keskusteluissa tärkeinä arvoina, ja niiden osalta autonomisten autojen suhteen katsottiin olevan kaksi mahdollista kehityssuuntaa. Parhaimmillaan autonomiseen autoon syntyisi tunneside ja siitä tulisi vielä enemmän tasa-arvoinen kaveri hellittelynimineen. Autojen liiallista inhimillistämistä ei kuitenkaan katsottu pelkästään hyvänä asiana, sillä niiden ei haluttu korvaavan ihmisten välistä vuorovaikutusta:

“kyllä me varmaan inhimillistetään niitä autoja jotenkin sitten. En tiedä onko se nyt hyvä asia vai ei, että kyllä meidän pitäisi tämä ihmisten välinen vuorovaikutus säilyttää ja pyrkiä säilyttämään kaikin tavoin”.

(Itseohjautuvien autojen työpajan osallistuja)

Huolenaiheeksi muodostui myös yksinäisyyden lisääntyminen ja muista ihmisistä eristäytyminen teknologian avulla, jos autonominen auto korvaisi lisäksi myös riippuvuutta muiden ihmisten avusta, kuten lasten harrastuksiin viemisen.

6.2 Turvallisuus

Autonomisten autojen suhteen turvallisuutta käsiteltiin pääasiassa vikaherkkyiden näkökulmasta, kun puolestaan raitiovaunuliikenteen suhteen esille nousivat myös turvattomuus ja matkustajaturvallisuus.

Vikaherkkyyden myötä heikentynyt turvallisuus herätti laajalti huolta keskustelijoissa. Turvallisuutta uskottiin ennen kaikkea rapauttavan elektroniikan lisääntyminen, joka puolestaan lisäisi järjestelmiä ja osia, jotka saattaisivat vikaantua. Vikaherkkyys näyttäytyi pahimmillaan kontrollin menettämisenä hätätilanteessa. Arkisemmin vikaherkkyys ilmeni järjestelmän kyseenalaisena kykyinä toimia eri sääoloissa tai väärinä vikailmoituksina, jotka lamauttavat koko auton. Hakkerit loivat myös uudenlaisen uhan autojen ollessa tulevaisuudessa vielä enemmän tekoälystä riippuvaisia. Yleisesti esillä ollut argumentti autonomisten autojen turvallisuutta lisäävästä vaikutuksesta tunnistettiin, mutta tilalle pelättiin tulevan muita ongelmia kuten eräs työpajaosallistuja totesi:

“jos inhimillinen tekijä poistetaan, niin tulee todella iso määrä teknologisia tekijöitä, joissa saattaa joku klikkaus tapahtua, että ei välttämättä se riskiherkkyys tai onnettomuusherakkyys mene yhtään mihinkään”.

6.3 Itseohjautuvuus

Autojen osalta keskusteluissa nousi useita itseohjautuvuuteen liittyviä arvoja, kuten valinnan vapaus ja monimuotoisuus sekä itsensä luova toteuttaminen. Itseohjautuvuus raitiovaunuliikenteessä keskittyi yksityisyyden kunnioittamiseen ja suostumukseen tulla valvotuksi.

Valinnanvapaus ja monimuotoisuus liitettiin laajemmin siihen, missä määrin kuluttajilla olisi vapaus valita ajoneuvonsa tulevaisuudessa. Ajatus autonomisesta autosta tuntui uhkaavan yksilön vapautta valita itselleen mieluisin auto, ja yhteen muottiin pakottamista turvallisuuden nimissä pelättiin:

“autoilusta tulee tylsää, sellaista kontrollinomaista, vähän niin kuin, että Pohjois-Koreassa ollaan, ja nyt on kaikki pakotettua ja tietylleen samalla tavalla turvallisuuden nimissä”.
(Itseohjautuvien autojen työpajan osallistuja)

Ajoneuvojen monimuotoisuutta toivottiin, jottei autoilusta tulisi tylsää ja kontrollinomaista. Markkinoiden epäiltiin eriytyvän uusiin autonomisiin ja vanhoihin perinteisiin autoihin, mikä monipuolistaisi liikkeellä olevaa autokantaa.

Itsensä luovaa toteuttamista pidettiin suuressa arvossa myös autoilla. Auto ei näyttäytynyt käyttäjilleen vain henkilökohtaisena välineenä, vaan myös persoonallisuuden jatkeena, jota haluttiin luovasti käyttää erilaisiin tarkoituksiin ja muokata omannäköiseksi. Käyttäjät haluavat valita tietyn näköisen ja värisen auton, sillä auton katsottiin heijastavan käyttäjänsä identiteettiä. Äärimmillen standardoituna autonominen auto pahimmillaan tukahduttaisi luovuuden.

6.4 Universalismi

Keskusteluista universalismin alle sopivia arvoja oli inklusiivisuus, samoin kuin raitiovaunuliikenteessä.

Inklusiivisuus eli monimuotoisuuden huomioiminen liikenteessä herätti runsaasti keskustelua. Parhaimmillaan autonomisten autojen toivottiin tuovan helpotusta toimintarajoitteisten ihmisten omaehtoiseen liikkumiseen, helpottaen monen nykyisestä yksityisautoilusta ulkopuolelle jäävän käyttäjäryhmän elämää. Toiveen kuitenkin pelättiin jäävän kauniin ajatuksen tasolle autonomisten autojen oletettavasti kalliiden hintojen vuoksi. Korkeiden hintojen epäiltiin sulkevan autonomisista autoista erityisesti hyötyviä käyttäjäryhmiä ulkopuolelle. Autonomisten autojen kehitystä verrattiin myös keskusteluun hoivaroboteista:

“miten demokraattinen kehitys yhteiskunnassa menee, että onko vanhoilla ihmisillä käymässä niin, että ne, joilla ei ole rahaa riittävästi, ne ei saa näitä robotteja, ja ne on huonolla hoidolla, vai onko niin, koska niillä ei ole varaa palkata oikeaa ihmishoitajaa, niin ne on sitten robottien armoilla niin sanotusti”.
(Itseohjautuvien autojen työpajan osallistuja)

7 Käyttäjien arvot: itseohjautuvat laivat

Laivojen, ja meriliikenteen laajempaa autonomisoitumista ajavat tällä hetkellä erityisesti meriturvallisuuteen liittyvät argumentit. Autonomisten teknologioiden ajatellaan ehkäisevän onnettomuuksia ja siten suojelevan ympäristöä. On esimerkiksi osoitettu, että 65,8% onnettomuuksista (2011-2018) johtuivat ihmisen toiminnasta, ja 20% järjestelmävioista.⁴⁹ Erityinen turvallisuuden painopistealue on alusten yhteentörmäyksen välttäminen (Collision Avoidance), jossa alus tekee väistämispäätökset merenkulun kansainvälisten COLREG-sääntöjen mukaan ottaen huomioon muiden alusten liikkeet. Traficom (2019) selvityksen mukaan, autonomisten alusten tarvitsemat algoritmit ovat kuitenkin tänä päivänä riittävän kehittyneitä ennakoimaan ja suoriutumaan tarvittavista väistöliikkeistä. Näin ollen ne voisivat liikennöidä lyhyillä vakioiteillä vähäliikenteiseen aikaan.

Toinen autonomisuutta ajava asia on ilmastotavoitteiden tiukkeneminen sillä automatisoitu navigointi tehostaa kuljettuja merimatkoja, mikä puolestaan vähentää hiilidioksidipäästöjä pienemmän polttoainekulutuksen vuoksi.

Vaikka laivojen autonomisuuden lisääntymisestä on esitetty selkeitä positiivisia argumentteja, ei autonomisia laivoja vielä ole laajasti käytössä sillä meriliikenne on hyvin säädeltyä, ja autonomisuuden lisääntyminen vaatii useita lakiuudistuksia. Esimerkiksi Suomessa luotsauslakia muutettiin vuonna 2019 siten, että etäluotsaus tuli mahdolliseksi, ja vuonna 2022 suoritettiin ensimmäinen etäluotsauspilotointi. Autonomisia laivoja on kuitenkin pilotoitu esim. Japanissa alkuvuodesta 2022^{50,51} ja Norjassa on otettu käyttöön jopa täysin autonominen ro-ro alus vuonna 2021⁵².

Merenkulun autonomisuus etenee asteittain kuten muissakin liikennemuodoissa. Nykypäivänä laivojen autonomisuus on Kuvan 8 osoittamilla tasoilla AL1-AL2, joissa laivoilla on autonomisia järjestelmiä päätöksenteon tueksi sekä laivoilla että

⁴⁹ <https://www.wartsila.com/insights/whitepaper/the-future-of-smart-autonomy-is-here>

⁵⁰ [Targeting 2025 for Fully Autonomous Navigation | The Nippon Foundation \(nippon-foundation.or.jp\)](#) (20.3.2023)

⁵¹ https://wwwcdn.imo.org/localresources/en/OurWork/Safety/PublishingImages/Pages/MASS/09_Hideyuki%20Ando_Development%20and%20Demonstration%20of%20Autonomous%20Ships%20in%20Japan.pdf (20.3.2023)

⁵² [Electric, Autonomous Ro-Ro Freight Ferries Begin Service in Norway \(maritime-executive.com\)](#) (20.3.2023)

etäyhteydellä.⁵³ Täysin autonomisiin ilman valvontaa liikkuviin AL5-AL6 tason laivoihin on vielä matkaa. Merenkulun automaatiotasojen määrittelyssä tulee ottaa huomioon aluksen matkan eri vaiheet, kuten aluksen satamamanöveeraus ja navigointiosuudet, joista jälkimmäinen hoidetaan jo tänä päivänä automaatiotasojen hyödyntäen. Laivojen toimintaympäristö on esimerkiksi raiteilla liikkuvaa liikennettä huomattavasti monimutkaisempaa, toisaalta autonomisuudesta on myös paljon apua jo nykypäivän laivaliikenteessä.

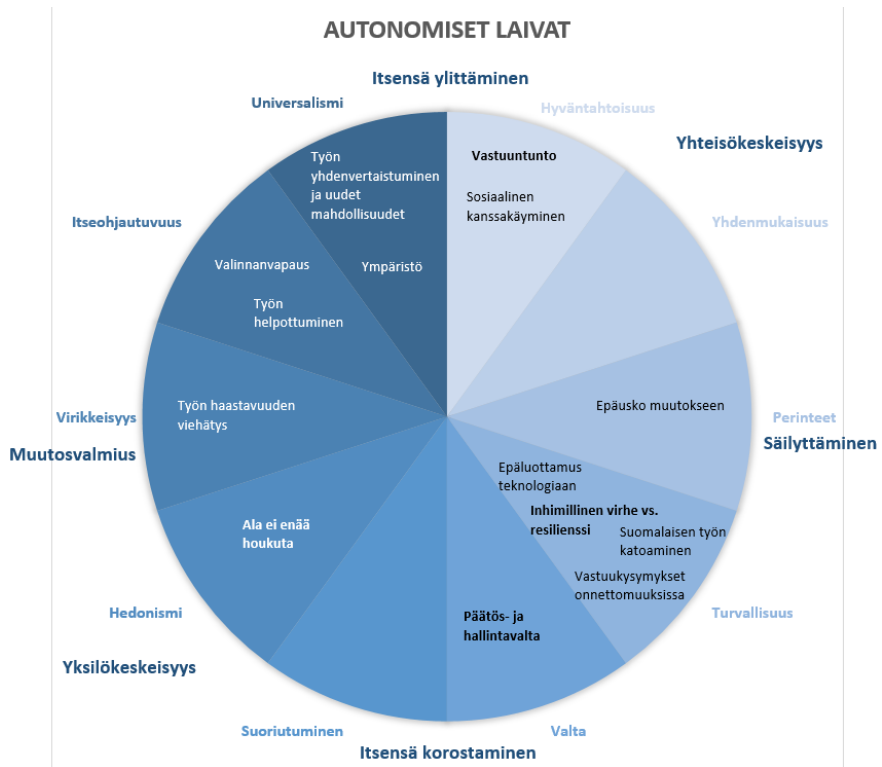


Kuva 8. Autonomian tasot meriliikenteessä. Lähde: Wärtsilä white paper, 2021.

Itseohjautuvien laivojen työpajassa keskusteltiin merikapteeniopiskelijoiden kanssa miltä autonomiset laivat näyttävät tulevien merenkulun ammattilaisten silmissä, sekä mitä ajatuksia ja arvoja autonomisuus ja tekoäly herättävät merenkulussa kahden skenaarion avulla.

⁵³ Wärtsilä white paper (2021). The future of Smart Autonomy is here. DISCOVER A STEPWISE APPROACH. <https://www.wartsila.com/insights/whitepaper/the-future-of-smart-autonomy-is-here> (20.3.2023)

- Skenaario 1: Nykypäivän merenkulku. Ensimmäisessä skenaariossa pohdittiin nykypäivän meriliikennettä, ja siihen liittyviä haasteita.
- Skenaario 2: Merenkulkua maalla. Toisessa tulevaisuuteen liittyvässä skenaariossa käytiin läpi sitä miltä merenkulkuammatti näyttäisi maista käsin, ja mitä haasteita lisääntynyt automaatio tuo laivoihin.



Kuva 9. Autonomisiin laivoihin liittyvät arvot Schwartzin arvokehällä

Autonomisiin laivoihin liittyvät arvot sisälsivät sekä itsensä ylittämisen että korostamisen arvoja (Kuva 9). Arvoja löytyi myös vaaka-akselilla molemmiin puolin, mutta painotus oli muutoksen vastustamisessa. Merkittävimpiä arvoluokkia olivat **hyväntahtoisuus**, **turvallisuus**, **valta** ja **hedonismi** (Taulukko 4).

Taulukko 4. Arvoluokat autonomisissa laivoissa

HYVÄNTAHTOISUUS <ul style="list-style-type: none">• Vastuuntunto	TURVALLISUUS <ul style="list-style-type: none">• Inhimillinen virhe vs. resilienssi
VALTA <ul style="list-style-type: none">• Päätös- ja hallintavalta	HEDONISMI <ul style="list-style-type: none">• Ala ei enää houkuttele

7.1 Hyväntahtoisuus

Keskusteluissa autonomisesta aluksesta käyttäjät pitivät tärkeänä vastuuntuntoa, kun itseohjautuvien autojen kontekstissa korostuivat inhimillisuus ja kohtaaminen sekä autonomisen raitiovaunun osalta erityisesti valvonnan ymmärrettävyys ja läpinäkyvyys nousivat arvokeskustelussa esille.

Vastuuntunto näyttäytyi tulevaisuuden merenkulkijoille tärkeänä arvona. Laivan päällikkö on vastuussa laivan turvallisuudesta, eikä vastuun antamista tietokoneelle nähty mahdollisena. Tietokoneella ei uskottu olevan vastuuntuntoa siinä missä ihmisistä koostuva miehistö pyrkii tekemään kaikkensa laivan, matkustajien, lastin ja miehistön turvallisuuden eteen. Ihminen nähtiin myös tulevaisuuden meriliikenteessä vastuunkantajana koko operaatiolle. Päällikön vastuun merkitys merenkulussa on korostunut, ja sen pakoilua pidetään häpeällisenä, kuten erään merenkulun työpajakeskustelijan pohdinnasta il-menee:

”Tietokone itsessään ei koe mitään vastuuntuntoa. No, ehkei kaikki päällikötkään, jos ’putoaa pelastusveneeseen’ niin ei välttämättä ole hirveän vastuuntuntoinen kaveri muutenkaan.”

7.2 Turvallisuus

Autonomisen merenkulun suhteen turvallisuutta käsiteltiin pääasiassa ihmisen resilienssin näkökulmasta, joka asettui vastakkain inhimillisistä virhettä käsittelevien välttämien kanssa. Myös raitiovaunujen ja itseohjautuvien autojen työpajoissa nousi turvallisuuden kontekstissa esille tekoälyn vikaherkkyys, ja siitä aiheutuva turvattomuus.

Inhimillinen virhe vs. resilienssi -keskustelu asetti vastakkain yhtäältä väitteet lisääntyvän teknologian mahdollisuuksista vähentää inhimillisestä virheestä

johtuvia onnettomuuksia ja toisaalta ihmisen resilienssin eli joustavan kyvyn sopeuttaa omaa toimintaansa ja estää näin onnettomuuksia. Ensin mainittu väittäjä tunnistettiin, mutta sitä kyseenalaistettiin eikä sille löydetty todisteita, kuten eräs meriliikennetyöpajan keskustelija asian muotoili:

"Voi sanoa, että 0 % [onnettomuuksista] on tekoälyn aiheuttamia, kun ei sitä ole vielä käytössä".

Sen sijaan ihmiset nähtiin kehityskelpoisina ja autonomista teknologiaa turvallisempina sekä pystyvämpinä korjaamaan esimerkiksi mekaanisen laiterikon, jolle tekoäly ei pystyisi tekemään mitään. Ihmisten joustavuus ulottuu myös vikatilanteiden ratkaisemista pidemmälle. Eri alusten miehistöt voivat keskenään poiketa väistämissäännöistä ja luoda tilannekohtaisia sopimuksia, jotka mahdollistavat tehokkaamman ajon.

7.3 Valta

Valta on hallintavaltaa muihin ihmisiin ja resursseihin sekä yhteiskunnallisen arvostuksen ja statuksen kunnioitusta. Tähän luokkaan voidaan sijoittaa sellaisia arvoja kuin arvovalta, varakkuus ja yhteiskunnallinen valta. Autonomiseen merenkulkuun liittyy päätös- ja hallintavalta.

Päätös- ja hallintavalta liittyivät tulevaisuuden merenkulkijoilla sekä ihmisen päätösvaltaisuuteen että aluksen hallintaan. Aluksen päälliköllä on paljon valtaa ja sen myötä myös vastuuta päätöksenteosta. Tekoälyn määritelmään miellettiin kuuluvan tekoälyn itsenäinen päätöksenteko, jolloin ihminen syrjäytettäisiin päätöksenteosta. Etenkin tiukan paikan tullen toivottiin, että ihminen on vastuussa päätöksenteosta. Päätösvallan siirtymisen myötä herää myös vastuukysymyksiä. Jos ihmisen sijaan päätöksen tekisikin kone, pohdittiin kuka tällöin olisi vastuussa. Katsottiinkin, että aina on oltava ihminen, joka tarvittaessa ottaa ohjat tilanteesta. Päätösvallan lisäksi laivan hallintaan liittyvä valta kiehtoi työssä. Eräs keskustelija kuvailikin, miten *"laivan ajaminen on aika siistiä puuhaa"*.

7.4 Hedonismi

Hedonismi on mielihyvän tavoittelua sekä itsensä hemmottelua. Hedonistisia arvoja ovat esimerkiksi mielihyvä, elämästä nauttiminen ja itsensä hemmottelu. Autonomisista aluksista käydyssä keskustelussa nousi esille hedonistisia arvoja liittyen alan muutokseen kielteisessä merkityksessä.

Ala ei enää pidetty houkuttelevana, jos laivoista tulisi autonomisia. Jos työ suoritettaisiin maista käsin, lähtisi työstä ne syyt, minkä takia alaa alun perin oli lähdetty opiskelemaan. Työn hyvinä puolina pidettiin erityisesti pitkiä vapaita, joiden arveltiin poistuvan työn muututtua valvomo-operaattorin työksi. Jos työ laivalla jatkuisi, mutta työnkuva olisi aktiivisen toiminnan sijasta vain passiivista järjestelmien valvontaa, uskottiin siitä koituvan taas muita ongelmia ja kuormitusta.

8 Keskustelu

Arvot palvelevat ihmisten tarpeita ja tavoitteita, ja sikäli niiden avulla on mahdollista ymmärtää paremmin mistä lähtökohdista ja millä ehdoin käyttäjät ovat valmiita hyväksymään autonomisen liikenteen tulevaisuudessa. Arvot kertovat siitä, mikä käyttäjille on tärkeää.

Sekä autonomisten autojen että kuljettajattoman raitiovaunun osalta tärkeimmiksi nousivat samat arvoluokat: hyväntahtoisuus, turvallisuus, itseohjautuvuus ja universalismi. Vaikka liikkumismuodot erosivat merkittävästi toisistaan, teknologian lisääntymisen pelättiin kummassakin tapauksessa asettuvan kiilaksi ihmisten väliseen vuorovaikutukseen, vähentäen inhimillistä kohtaamista. Vaikka yksityisautoiluun ei itsessään liity palvelukokemusta kuten raitiovaunussa, lisääntyvän automaation pelättiin vähentävän vuorovaikutusta ja pahentavan yksinäisyyttä. Vaikutus näkyisi arkisissa tilanteissa, joissa autonomisuuden myötä auton käyttömahdollisuudet monipuolistuisivat, ja korvaisivat ennestään sosiaalisia tilanteita. Näin ollen ihannetilanteessa autonominen liikenne eristämisen sijaan yhdistäisi ihmisiä, ja helpottaisi ihmisten keskinäistä vuorovaikutusta. Voisiko autonominen autokyyti lisätä aikaa yhteydenpitoon läheisten kanssa, kun ei tarvitsekaan itse keskittyä ajamiseen?

Myös kaikesta inhimillisestä riisuttua teknologiaa pidettiin kalseana. Raitiovaunuliikenteessä vuorovaikutusta voitaisiin vahvistaa esimerkiksi suoralla ja helposti saatavilla olevalla kanavalla etävalvomoon (tai muuhun henkilökuntaan). Lähestyttävyyttä puolestaan voisi lisätä inhimillisten piirteiden tuominen teknologian olemukseen ja toimintalogiikkaan. Autonomisia autoja koskevassa työpajassa nousikin robotti-imurit esimerkiksi hellyttävästä, toverinomaisesta autonomisesta teknologiasta.

Autonomisen laivan osalta vuorovaikutus tai inhimillinen puoli ei noussut yhtä vahvasti esille. Tulevaisuuden merikapteenit olivat kyllä huolissaan, jos lisääntyvän teknologian myötä miehistön koko pienenesi, ja laivalla tarvittaisiin enää vain muutaman ihmisen läsnäoloa. Tästä pelättiin aiheutuvan sosiaalisia ongelmia pitkille merimatkoille. Lisäksi työn loiston nähtiin häviävän työn muuttuessa teknologian valvonnaksi tai maista käsin tehtäväksi tavalliseksi vuorotyöksi. Näitä seikkoja tarkasteltiin hedonistisesta kulmasta. Hyväntahtoisuusarvoista vastuuntunto kuitenkin korostui, mikä johtunee merikapteenin työhön liittyvästä suuresta vastuusta.

Liikkumistavasta riippumatta lisääntyvä automaatio herätti huolta turvallisuuden puolesta. Kun on kyse turvallisuuskriittisestä toimintaympäristöstä, kuten liikenteestä, turvallisuus nousee ymmärrettävästi yhdeksi tärkeimmistä arvoista. Teknologiaa pidettiin alttiina vioille, sekä kohtalokkaille että turhille, tavallisesti helposti ratkaistavissa oleville ongelmille, jotka kuitenkin ovat teknologialle ylivoimaisia voittaa. Inhimillisen virheen riskistä huolimatta ihmisen kykyyn luoda turvallisuutta uskottiin teknologiaa enemmän. Teknologiaa ei tulisikaan puskea puolivalmiina markkinoille. Epäonniset kokemukset ja onnettomuus uutiset vielä keskeneräisestä teknologiasta nakertavat luottamusta sen toimivuudesta ja turvallisuudesta.

Itseohjautuvuusarvot koskivat raitiovaunun osalta yksityisyyttä ja suostumuksen antamista, autonomisten autojen osalta valinnanvapautta ja itsensä toteuttamista. Yhteistä näille lienee kokemus siitä, että yksilönvapautta rajoitetaan ylhäältä alaspäin pakottamalla käyttäjät omaksumaan uuden teknologian käyttäjien tarpeita kuulematta. Käyttäjille on tärkeää läpinäkyvyys ja mahdollisuus tehdä itse tietoinen valinta, oli kyse sitten omien tietojen luovuttamisesta tai auton valinnasta. Valvonnan läpinäkyvyyttä ja ymmärrettävyyttä käyttäjille pidettiin erittäin tärkeänä, ja jonkinlaisena ehtona lisääntyvän valvonnan hyväksymiselle. Tältä osin tulisikin panostaa ymmärrettävään viestintään matkustajien moninaisuus huomioiden – raitiovaunun käyttäjissä on monen taustaisia ja ikäisiä ihmisiä, jotka viestinnän tulisi tavoittaa tasa-arvoisesti.

Raitiovaunu on julkinen liikenneväline toisin kuin auto, joka koetaan henkilökohtaisena tilana, ja näin ollen mukautuvampana omiin tarkoitukseen. Auto voi rakentua osaksi omaa identiteettiä ja harrastuksia toisin kuin raitiovaunu - yksityisautoissa lisääntyvän automaation tulisikin jättää tilaa käyttäjän omalle persoonalle. Lienee myös luontevaa, että raitiovaunua käsiteltiin kollektiivisesta tulokulmasta, itsensä ylittämisen kautta ja muita ajatellen, erilaisiin käyttäjäryhmiin samaistuen. Toisaalta kuljettajaton autokin herätti pohdintaa yhteisökeskeisistä näkökulmista, kuten siitä, toteutuuko autonomisen auton potentiaali eri tavoin rajoitteisten ihmisten elämää helpottavana välineenä vai tuleeeko siitä lopulta vain varakkaiden luksustuote. Universalismiarvoihin liittyen yhteistä työpajojen välillä oli tasa-arvoinen suhtautuminen tulevaisuuden teknologian saavutettavuuteen kaikkien erilaisten käyttäjäryhmien välillä.

Toisin kuin muissa työpajoissa, autonomisten laivojen osalta keskeisinä arvoluokkina nousivat valta- ja hedonismiarvot. Merenkulun työpajassa käyttäjät katsoivatkin aihetta oman työnsä näkökulmasta. Merikapteeniopiskelijat arvioivat muutosta laajemmin oman uravalintansa mielekkyyden kautta, kun taas autoilijat ja raitiovaunumatkustajat arvioivat autonomisuutta liikkumisen tapojensa näkökulmasta. Valta-arvoja taas ei noussut esille muissa työpajoissa. Merikapteeniopiskelijoiden valtaorientaatio johtunee niin ikään työn luonteesta, sillä laivan päälliköllä on viime kädessä kaikki valta ja vastuu laivan ajamisesta.

Käyttäjien arvojen tarkastelu Schwartzin arvokehikolla paljastaa myös uuteen teknologiaan kohdistuvia ristiriitaisia odotuksia. Teorian mukaan keskenään ristiriitaisia arvoluokkia ovat kehikolla toisiinsa nähden vastakkain asettuvat arvoluokat, kuten itseohjautuvuus ja turvallisuus. Kaikissa työpajoissa näihin

luokkiin liittyi arvoja, jotka synnyttivät ilmeisen ristiriidan. Käyttäjien näkökulmasta autonomiseen raitiovaunuun ja autoihin liittyi vaihtokauppa: parhaimmillaan autonomiset ratkaisut lisääisivät liikenteen turvallisuutta, mutta yksityisyyden ja valinnanvapauden kustannuksella. Autonomisten laivojen osalta ristiriita näkyi käyttäjien pelkona teknologian turvallisuutta heikentävästä vaikutuksesta, vaikka samalla käyttäjien nykyiset kokemukset jo olemassa olevasta automaatiosta laivalla olivat positiivisia. Nykyisen teknologian nähtiin jopa helpottavan työskentelyä laivalla. Sisäinen ristiriita oli myös ilmeinen siinä, miten osallistujat tunnustivat enemmistön onnettomuuksista olevan ihmisen syytä, mutta he samalla näkivät inhimillisen resilienssin olevan tärkeimpiä tekijöitä turvallisuuden ylläpitämisessä. Resilienssin voi nähdä olevan myös yhdenlaista itseohjautuvuutta, jota erityisesti arvostettiin ja pidettiin merenkulun turvallisuuden ja sujuvuuden kannalta keskeisenä.

Lopuksi

Tämä raportin tarkoituksena on tuoda esille tekoälyn kehitykseen ja käyttöönottoon liittyvää arvokeskustelua, erityisesti käyttäjien näkökulmasta. Tarkastelussa oli kolme autonomisen liikkumisen muotoa, raitiovaunuliikenne, itseohjautuvat autot ja autonomiset laivat. Täysin kuljettajatonta ei liikkuminen vielä ole, mutta autoissa pääsemme jo nauttimaan, tai huolehtimaan, autonomisuuden tuomista iloista eikä kuljettajaton raitiovaunuliikennekään ole kaukana. Myös merenkulussa on jo paljon automaatiota ja älyä. Autonomisia laivoja, sekä merenkulun muita osa-alueita kuten luotsausta, kehitetään ja testataan kiivaasti.

Työpajakeskusteluissa kävi esille se, että tekoälyyn ja autonomisuuteen liikenteessä liittyy vahvoja tunteita ja arvoja, sekä paljon niin positiivisia kuin huolestuttavia tulevaisuuden odotuksia. Liikenne ja liikkuminen ovat osa ihmisten päivittäistä arkea, eikä keskusteluissa noussut esille kriittistä ja täysin tekoälyä ja autonomiaa tyrmääviä arvoja. Käyttäjät kokivat tekoälyn kehityksen enemmänkin mahdollisuutena, toki odotukset tekoälyä kohtaan olivat hieman huolestuneita erityisesti turvallisuuden näkökulmasta.

On tärkeä ymmärtää mikä tekoälyteknologioissa on tärkeää käyttäjille, ei ainoastaan miten käyttäjät ottavat muutoksen vastaan. Tästä näkökulmasta turvallisuus on mielenkiintoinen arvo sillä se selkeästi dominoi tekoälykeskustelua niin kehityksen kuin käyttöönoton näkökulmista - ihan syystäkin onhan teknologian luotettavuus yksi keskeisistä asioista sen käyttöönoton kannalta.

On myös huomionarvoista, että vaikka tekoäly ja autonomisuus liikenteessä nähtiin positiivisena, keskusteluissa ei noussut vahvasti esille ilmasto- ja ympäristöarvot siitä huolimatta, että tekoälyn ja autonomisuuden kehityksen kanssa kulkee samanaikaisesti vahva liikenteen sähköistämisen trendi. Vaikka ympäristökestävyys ei noussut puheenaiheeksi, keskustelu käsitteli yhteiskunnallista kestävyyttä monista eri näkökulmista. Toisaalta työpajojen kysymyksen asettelu oli luonnollisesti enemmän yhteiskunnallisessa kontekstissa, jossa tekoälyä käsiteltiin hyvin henkilökohtaisesta näkökulmasta.

Kirjallisuus

- Anderson, M. (2020). The road ahead for self-driving cars: The AV industry has had to reset expectations, as it shifts its focus to level 4 autonomy. *IEEE Spectrum*, Vol. 57 (5), 8 – 9. DOI: 10.1109/MSPEC.2020.9078402
- Ali Assaad, M.; Talj, R.; Charara, A. (2019) Autonomous Driving as System of Systems: roadmap for accelerating development. 2019 14th Annual Conference System of Systems Engineering (SoSE), 102–107. <https://doi.org/10.1109/SYSE.2019.8753886>
- Berkhourt, F., (2006). Normative expectations in systems innovation. *Technology Analysis & Strategic Management*. Vol. 18 (3-4), 299-311.
- Borup, M.; Brown, N.; Konrad, K. & Van Lente, H. (2006). The Sociology of Expectations in Science and Technology. *Technology Analysis & Strategic Management*. Vol 18 (4), 285-298.
- Brown, N. & Michael, M. (2003). A Sociology of Expectations: Retrospecting Prospects and Prospecting Retrospects. *Technology Analysis and Strategic Management*, 15(1), 3-18.
- Criado Perez, C. (2019). *Invisible Women: Data Bias in a World Designed for Men*. Abrams Press.
- Devi, S.; Malarvezhi, P.; Dayana, R.; Vadivukkarasi, K. (2020) A Comprehensive Survey on Autonomous Driving Cars: A Perspective View. *Wireless Personal Communications: An International Journal*, Vol 114(3), 2121–2133. <https://doi.org/10.1007/s11277-020-07468-y>
- Dolata, U. (2011). "Radical change as gradual transformation: Characteristics and variants of socio-technical transitions," *Research Contributions to Organizational Sociology and Innovation Studies, SOI Discussion Papers 2011-03*, University of Stuttgart, Institute for Social Sciences, Department of Organizational Sociology and Innovation Studies.

- ERTRAC Task Force (2015). Automated Driving Roadmap. https://erticonetwork.com/wp-content/uploads/2015/03/www.ertrac.org_uploads_documentssearch_id35_ERTRAC_Automated-Driving_draft3-web.pdf (20.3.2023)
- Ernst, R. & Hu, X. S. (2021). Autonomous Systems Design—A Virtual Roundtable. *Com-puter*, vol. 54(11), 17-25. doi: 10.1109/MC.2021.3101914.
- European Commission (2019). STRIA Roadmap on Connected and Automated Transport: Road, Rail and Waterborne. https://trimis.ec.europa.eu/sites/default/files/2021-10/stria_roadmap_2019-connected_and_automated_transport.pdf (20.3.2023)
- Faisal, A.; Yigitcanlar, T.; Kamruzzaman, M.; Currie, G. (2019). Understanding autonomous vehicles: A systematic literature review on capability, impact, planning and policy. *The Journal of Transport and Land Use*, Vol.12(1), 45–7.
- Gandia, R.M.; Antonialli, F.; Habib Cavazza, B.; Miranda Neto, A.; Alves de Lima, D.; Yutaka Sugano, J.; Nicolai, I.; Zambalde, A.L. (2019). Autonomous vehicles: scientometric and bibliometric review, *Transport Reviews*, Vol. 39(1), 9-28. <https://doi.org/10.1080/01441647.2018.1518937>
- GEAR 2030 (2017). High Level Group on the Competitiveness and Sustainable Growth of the Automotive Industry in the European Union. FINAL REPORT. <https://www.europarl.europa.eu/cmsdata/141562/GEAR%202030%20Final%20Report.pdf> (20.3.2030)
- GEAR 2030 (2016). Roadmap on Highly automated vehicles. DISCUSSION PAPER. <https://circabc.europa.eu/sd/a/3e06f7bf-2719-4be1-9f24-ba6b2975d7eb/Discussion%20Paper%20-%20rev.1%2004-05-2015.pdf> (20.3.2023)
- Helkama, K., Myllyniemi, R., & Liebkind, K. (2010). *Johdatus sosiaalipsykologiaan*. Helsinki, Suomi: Edita.
- Karetnikov, V., Ol'Khovik, E., Ivanova, A., Butsanets, A. (2021). Technology Level and Development Trends of Autonomous Shipping Means. In: Murgul, V., Pukhkal, V. (eds) *International Scientific Conference Energy Management of Municipal Facilities and Sustainable Energy Technologies EMMFT 2019*. EMMFT 2019. *Advances in Intelligent Systems and Computing*, vol 1258. Springer, Cham. https://doi.org/10.1007/978-3-030-57450-5_36
- Kim, M.; Joung, T-H.; Jeong, B. & Park, H-S. (2020). Autonomous shipping and its impact on regulations, technologies, and industries, *Journal of International Maritime Safety, Environmental Affairs, and Shipping*, 4:2, 17-25, DOI: 10.1080/25725084.2020.177942

- Klare, B. F.; Burge, M. J.; Klontz, J. C.; Vorder Bruegge, R. W. & Jain, A. K. (2012). Face Recognition Performance: Role of Demographic Information. *IEEE Transactions on Information Forensics and Security*, vol. 7(6), 1789-1801. doi: 10.1109/TIFS.2012.2214212.
- Leikas, J. (2009). Life-Based Design: A holistic approach to designing human-technology interaction. VTT Technical Research Centre of Finland. VTT Publications No. 726, <https://publications.vtt.fi/pdf/publications/2009/P726.pdf>
- National Research Council (2014). *Autonomy Research for Civil Aviation: Toward a New Era of Flight*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/18815>
- Othman, K. (2021). Impact of Autonomous Vehicles on the Physical Infrastructure: Changes and Challenges. *Designs*, Vol. 5(3). <https://doi.org/10.3390/designs5030040>
- Pfotenhauer, S., & Jasanoff, S. (2017). Panacea or diagnosis? Imaginaries of innovation and the 'MIT model' in three political cultures. *Social Studies of Science*, 47(6), 783–810. <https://doi.org/10.1177/0306312717706110>
- Pöntinen, H. (2019). *Tekoälyn etiikkaa - Valtiot tekoälyn hyödyntäjinä*. Helsingin yliopisto, Teologinen tiedekunta, Pro gradu.
- Sifakis, J. (2021). Why is it so hard to make self-driving cars? (Trustworthy autonomous systems). *WI-IAT '21: IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. <https://doi.org/10.1145/3486622.0000002>
- Singh, P.; Dulebenets, M.A.; Pasha, J.; Santibanez Gonzalez, E.; Lau, Yui-yip; Kampmann, R. (2021). Deployment of Autonomous Trains in Rail Transportation: Current Trends and Existing Challenges. *IEEE Access*, VOLUME 9, 91427- 91461.
- Stahl, B.C., (2022). Responsible innovation ecosystems: Ethical implications of the application of the ecosystem concept to artificial intelligence. *International Journal of Information Management*, 62, 102441.
- Stahl, B.C. B.C; Antoniou, J.; Ryan, M.; Macnish, K. & Jiya, T. (2021). Organisational responses to the ethical issues of artificial intelligence. *AI & SOCIETY*, 37, 23–37. <https://doi.org/10.1007/s00146-021-01148-6>
- Tatman, R. (2017). Gender and Dialect Bias in YouTube's Automatic Captions. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, p. 53–59, Valencia, Spain. Association for Computational Linguistics. DOI: 10.18653/v1/W17-1606

van Lente, H. (2012). Navigating foresight in a sea of expectations: lessons from the so-ciology of expectations. *Technology Analysis & Strategic Management*, 24:8, 769-782. DOI: 10.1080/09537325.2012.715478

Waelen, R.A. (2023). The struggle for recognition in the age of facial recognition technology. *AI and Ethics*, Vol. 3,215–222.
<https://doi.org/10.1007/s43681-022-00146-8>

Wang, J.; Huang, H.; Li, K.; Li, J. (2021). Towards the Unified Principles for Level 5 Au-tonomous Vehicles. *Engineering*, Vol. 7(9), 1313-1325.
<https://doi.org/10.1016/j.eng.2020.10.018>

Liite 3: Autonomiseen merenkulkuun liittyvät arvot

Hyväntahtoisuus on toisten ihmisten ja yhteisön hyvinvoinnin edistämistä ja priorisoimista. Luokkaan sisältyy arvoja kuten auttavaisuus, rehellisyys, anteeksiantavuus, vastuullisuus, luotettavuus, toistuvuus ja kypsä rakkaus. Keskusteluissa autonomisesta aluksesta käyttäjät pitivät tärkeänä vastuuntuntoa.

Vastuuntunto näyttäytyi tulevaisuuden merenkulkijalle tärkeänä arvona. Laivan päällikkö on tavallisesti vastuussa laivan turvallisuudesta, eikä vastuun antamista tietokoneelle nähty mahdollisena. Tietokoneella ei uskottu olevan vastuuntuntoa siinä missä ihmisistä koostuva miehistö pyrkii tekemään kaikkensa laivan, matkustajien, lastin ja miehistön turvallisuuden eteen. Ihminen nähtiin myös tulevaisuuden meriliikenteessä vastuunkantajana koko operaatiolle. Päällikön vastuun merkitys merenkulussa on korostunut, ja sen pakollua pidetään häpeällisenä, kuten erään keskustelijan pohdinnasta ilmenee: "tietokone itsessään ei koe mitään vastuuntuntoa. No, ehkei kaikki päällikötkään, jos putoaa pelastusveneeseen niin ei välttämättä ole hirveän vastuuntuntoinen kaveri muutenkaan."

Sosiaalinen kanssakäyminen koettiin tärkeäksi. Tämä nousi arvona esille etenkin pohdittaessa ihmisen mahdollista tulevaa roolia laivalla valvojana teknologialle. Jos miehistö koostuisi vain muutamasta henkilöstä, uskottiin sen olevan raskasta, sillä "ihmisille ne sosiaaliset kanssakäymiset on kuitenkin tosi merkityksellisiä", kuten eräs merikapteeniopiskelija asian ilmaisi.

Inhimillinen virhe vs. resilienssi -keskustelu asetti vastakkain yhtäältä väitteet lisääntyvän teknologian mahdollisuuksista vähentää inhimillisestä virheestä johtuvia onnettomuuksia ja toisaalta ihmisen joustavan kyvyn sopeuttaa omaa toimintaansa ja estää näin onnettomuuksia. Ensinnäkin mainittu väittämä tunnustettiin, mutta sitä kyseenalaistettiin eikä sille löydetty todisteita, kuten eräs keskustelija asian muotoili: "voi sanoa, että 0% [onnettomuuksista] on tekoälyn aiheuttamia, kun ei sitä ole vielä käytössä". Sen sijaan ihmiset nähtiin kehityskelpoisina ja autonomista teknologiaa turvallisempina sekä pystyvämpinä korjaamaan esimerkiksi mekaanisen laiterikon, jolle tekoäly ei pystyisi tekemään yhtään mitään. Ihmisten joustavuus ulottuu myös vikatilanteiden ratkaisemista pidemmälle. Eri alusten miehistöt voivat keskenään poiketa väistämissäännöistä ja luoda tilannekohtaisia sopimuksia, jotka mahdollistavat tehokkaamman ajon.

Epäluottamus teknologiaan heijasteli huonoja kokemuksia teknologiasta. Teknologiaa ei pidetty pettämättömänä, sen ei uskottu selviävän vaikeilla vesialueilla ja jo pienten ongelmien pelättiin aiheuttavan suuria vahinkoja. Tulevaisuuden merenkulkijat eivät täysin luottaneet teknologiaan, muttei sen uskottu herättävän luottamusta risteilymatkustajissakaan: "kuinka moni on valmis lähtemään ristelyle, jossa ei ole yhden ainutta ihmistä komentosillalla ohjaamassa laivaa, kuinka moni ihminen on valmis luottamaan oman henkensä ja turvallisuutensa täysin tietokoneen laskennan varaan?". Näkemys oli, että ihmisen tulisi valvoa teknologian toimivuutta ja ottaa tarvittaessa kontrolli itselleen. Teknologian katsottiin olevan myös sabotoinnille altista.

Vastuukysymykset onnettomuuksissa herättivät keskustelua merenkulun automaatioissa. Jos päätöksenteko ei olekaan enää ihmisen käsissä, pohdittiin kuka silloin vastaa, jos sattuu onnettomuus. Onko vastuussa silloin operaattori, järjestelmien valmistaja vai niiden asentaja. Haastavaksi myös uskottiin voivan muodostua vastuukysymykset onnettomuustilanteessa, jossa toinen laiva olisi miehitetty ja toinen miehittämätön.

Suomalaisen työn katoaminen nähtiin yhtenä uhkakuvana. Pidettiin todennäköisenä, että laivojen ohjaamisen siirtyessä komentosillalta valvomoon, operaattoritoiminta keskitettäisiin alueellisesti. Suomea ei pidetty todennäköisenä laivojen ohjaamisen keskuksena Itämeren alueella, jolloin työnäkymät Suomessa heikkenisivät.

Yhdenmukaisuus on mukautumista vallitseviin normeihin ja odotuksiin sekä muita häiritseviä teoista pidättäytymistä. Luokkaan sisältyviä arvoja ovat esimerkiksi velvollisuudentuntoisuus, itsekuri, kohteliaisuus, vanhempien ihmisten kunnioittaminen ja mukautuminen normeihin. Keskusteluissa ei noussut tähän arvoluokkaan sopivia arvoja.

Suorittuminen on menestyksen tavoittelua, ja saavutuksia arvioidaan vallalla olevan kulttuurin arvostamia standardeja noudattaen. Suorittumisarvoja ovat esimerkiksi kunnianhimo, menestys, kyvykyys ja vaikutusvalta. Autonomisten laivojen osalta suorittumista ei käsitelty.

Turvallisuus on oman ja läheisten elämän tasaisuutta ja harmonista jatkuvuutta sekä yhteiskunnan vakautta. Luokkaan sisältyy arvoja kuten yhteiskunnallinen järjestys, perheen turvallisuus, kansallinen turvallisuus, puhtaus ja palvelusten vastavuoroisuus. Autonomisen merenkulun suhteen turvallisuutta käsiteltiin pääasiassa ihmisen **resilienssin** näkökulmasta, joka asetti vastakkain inhimillistä virhettä käsittelevien väittämien kanssa. Turvallisuus-arvoja heijastelivat myös epäluottamus teknologiaan, sabotoinnin mahdollisuus, vastuukysymykset onnettomuuksissa ja pelko suomalaisen työn katoamisesta.

Perinteillä tarkoitetaan kulttuuriin tai uskontoon liittyvien tapojen noudattamista ja säilyttämistä sekä sääntöjen hyväksymistä. Luokkaan sopii arvot kuten perinteiden kunnioitus, nöyryys, hurskaus, oman elämänsä hyväksyminen ja mukautuminen vallitseviin kulttuuriin/uskonnonnormeihin. Tähän arvoluokkaan keskusteluista nousi epäusko muutokseen.

Epäusko muutokseen virisi keskustelijoissa pohdittaessa täysin autonomisen laivan mahdollisuutta tulevaisuudessa. Muutosta ei nähty realistisena tulevaisuutena vaikkakin joidenkin tehtävien korvaaminen teknologialla olisikin mahdollista.

Valta on hallintavaltaa muihin ihmisiin ja resursseihin sekä yhteiskunnallisen arvostuksen ja statuksen kunnioitusta. Tähän luokkaan voidaan sijoittaa sellaisia arvoja kuin arvovalta, varakkuus ja yhteiskunnallinen valta. Autonomiseen merenkulkuun liittyy päätös- ja hallintavalta.

Päätös- ja hallintavalta liittyvät tulevaisuuden merenkulkijoilla sekä ihmisen päätösvaltaisuuteen, että aluksen hallintaan. Aluksen päälliköllä on paljon valtaa ja sen myötä myös vastuuta päätöksenteosta. Tekoälyn määrittelemään miellettiin kuuluvan tekoälyn itsenäinen päätöksenteko, jolloin ihminen syrjäytettäisiin päätöksenteosta. Etenkin tiukan paikan tullen toivottiin, että ihminen on vastuussa päätöksenteosta. Päätösvallan siirtymisen myötä herää myös vastuukysymyksiä. Jos ihmisen sijaan päätöksen tekisikin kone, pohdittiin kuka tällöin olisi vastuussa. Katsottiinkin, että aina on oltava ihminen, joka tarvittaessa ottaa ohjat tilanteesta. Päätösvallan lisäksi laivan hallintaan liittyvä valta kiehtoi työssä. Eräs keskustelija kuvailikin, miten "laivan ajaminen on aika siistiä puuhaa".

Itseohjautuvuudella tarkoitetaan toiminnan ja ajattelun vapautta, kuten omien päämäärien valintaa sekä vapautta kehittää omia ideoita ja luoda uutta. Luokkaan kuuluu arvoja kuten luovuus, vapaus, omien tavoitteiden valitseminen, uteliaisuus, riippumattomuus. Keskusteluissa nousi useita itseohjautuvuuteen liittyviä arvoja, kuten valinnanvapaus ja työn helpottuminen.

Valinnanvapaus arvona ilmeni pohdinnassa täydellisestä autonomian tilasta. Kysymystä pidettiin mielipidevaltaisena, sillä merikapteenissa katsottiin olevan eroja sen suhteen, että toiset haluavat ohjata laivaa itse ja toiset taas antaisi tehtävän mieluummin tietokoneelle. Valinnanvapaus ilmeni myös teknologian mahdollistamana lisääntyvästä vapaa-aikana, jonka puitteissa ihmiset saisivat tehdä vapautuvalla ajallaan mitä ikinä halusivat.

Työn helpottumisen katsottiin olevan yksi seuraus lisääntyvästä teknologiasta. Nykyiset älykkäät järjestelmät koettiin hyödyllisiksi ja työtä helpottaviksi. Järjestelmät nähtiin avuksi muun muassa ihmisen havainnointikyvyn ylittävissä tilanteissa. Parhaimmillaan eri järjestelmät kaikki tukevat toisiaan. Kontrollia järjestelmiin pidettiin kuitenkin edelleen keskeisenä, ihmisen tulisi olla "luopin keskellä eikä luopin ulkopuolella".

Hedonismi on mielihyvän tavoittelua sekä itsensä hemmottelua. Hedonistisia arvoja ovat esimerkiksi mielihyvä, elämästä nauttiminen ja itsensä hemmottelu. Autonomisista aluksista käydyssä keskustelussa nousi esille hedonistisia arvoja liittyen alan muutokseen kielteisessä merkityksessä.

Ala ei enää pidetty houkuttelevana, jos laivoista tulisi autonomisia. Jos työ suoritettaisiin maista käsin, lähtisi työstä ne syyt, minkä takia alaa alun perin oli lähdetty opiskelemaan. Työn hyvinä puolina pidettiin erityisesti pitkiä vapaita, joiden arveltiin poistuvan työn muututtua valvomo-operaattorin työksi. Jos työ laivalla jatkuisi, mutta työnkuva olisi aktiivisen toiminnan sijasta vain passiivista järjestelmien valvontaa, uskottiin siitä koituvan taas muita ongelmia ja kuormitusta.

Virikkeisyys on jännityksen ja uusien elämysten etsimistä sekä halua kohdata haasteita. Luokalle tyypillisiä arvoja ovat vaihteleva ja jännittävä elämä, uskaliaisuus ja avoimuus uusille kokemuksille. Autonomisista laivoista virikkeisyyteen sopivia arvoja oli työn haastavuuden viehätyksellä.

Työn haastavuuden viehätyksellä ilmeni nykyisestä merikapteenin työstä puhuttaessa. Työtä pidettiin haastavana muun muassa ympäristöstä ja työstä itsestään johtuvien seikkojen vuoksi. Haastetta pidettiin kuitenkin myös mielenkiintoisena työn piirteinä.

Universalismi on laajempaa, myös muiden kuin omaan lähipiiriin kuuluvien ihmisten sekä luonnon arvostamista. Universalismiin liittyviä arvoja ovat esimerkiksi laajakatseisuus, tasa-arvo, sosiaalinen oikeudenmukaisuus, maailmanrauha, kauneuden maailma, yhteys luontoon, kypsä elämänymmärrys ja ympäristön suojeleminen. Keskusteluista universalismin alle sopivia arvoja olivat työn yhdenvertaistuminen ja uudet mahdollisuudet sekä ympäristö.

Työn yhdenvertaistuminen ja uudet mahdollisuudet todennäköisesti toteutuisivat työn muutoksen myötä, jos aikaisemmin raskaasta työstä laivalla tulisikin vuorotyötä valvomossa. Fyysisille rajoituksille, jotka nykyään ovat toimineet karavana tekijänä ammattiin pääsystä, ei olisi enää perusteita, ja näin ollen työstä tulisi yhdenvertaisempaa. Myös sairauden tai vammaan sattua kykenisi edelleen jatkamaan "haaveammattissa". Toisaalta maista käsin ohjaaminen nähtiin laajentavan uravaihtoehtoja, sillä kun ikää ja perhettä tulee, laivalla työskentely nähtiin mahdollisuudeksi.

Ympäristöä arvostettiin, ja autonomisten alusten pelättiin aiheuttavan ympäristölle vahinkoa onnettomuuden sattuessa. Alusten ajautuessa karille ympäristökatastrofit ovat mahdollisia laivan päästäessä haitallisia aineita mereen. Seuraukset voivat olla kohtalokkaita herkille ekosysteemeille, kuten eräs merikapteeniopiskelija asian ilmaisi: "kun se öljyläiva on siellä korallien päällä ja ihmetellään mihin tämä ekosysteemi meni".

Nimeke	Eettinen tekoäly autonomisissa järjestelmissä
Tekijä(t)	Tuisku Salonen, Nina Rilla & Pietari Pikkuaho
Tiivistelmä	<p>Autonomiset järjestelmät voivat lisätä turvallisuutta, tehokkuutta ja kestävyyttä, mutta järjestelmien ohjelmointiin ja käyttöön liittyy kuitenkin monia eettisiä kysymyksiä. Tässä katsauksessa tarkastellaan tekoälyyn liittyvien arvojen muodostumista kolmessa autonomiseen liikkumiseen liittyvässä kontekstissa: raitiovaunuliikenteessä, yksityisautoilussa ja meriliikenteessä. Tarkoituksena on tuoda esille tekoälyn kehitykseen ja käyttöönottoon liittyvää arvokeskustelua, erityisesti käyttäjien näkökulmasta.</p> <p>Tutkimusaineisto on kerätty virtuaalisesti järjestetyissä neljässä käyttäjätyöpajassa vuosina 2021 ja 2022. Tulokset perustuvat laadullisesta aineistosta tehtyyn sisällönanalyttiseen tulkintaan ja vastausten temaattiseen luokitteluun. Liikenne ja liikkuminen ovat osa ihmisten päivittäistä arkea, eikä keskusteluissa noussut esille kriittistä ja täysin tekoälyä ja autonomiaa tyrmääviä arvoja. Tekoälyn ja autonomisuuteen liikenteessä liittyy kuitenkin vahvoja tunteita ja arvoja, sekä paljon niin positiivisia kuin huolestuttavia tulevaisuuden odotuksia. Käyttäjät kokivat tekoälyn kehityksen enemmänkin mahdollisuutena, toki odotukset tekoälyä kohtaan olivat hieman huolestuneita.</p> <p>Kun on kyse turvallisuuskriittisestä toimintaympäristöstä, kuten liikenteestä, turvallisuus nousee ymmärrettävästi yhdeksi tärkeimmistä arvoista. Käyttäjien näkökulmasta autonomiseen raitiovaunuun ja autoihin liittyy vaihtokauppa: parhaimmillaan autonomiset ratkaisut lisäisivät liikenteen turvallisuutta, mutta yksityisyyden ja valinnanvapauden kustannuksella. Autonomisten laivojen osalta ristiriita näkyi käyttäjien pelkona teknologian turvallisuutta heikentävästä vaikutuksesta.</p> <p>Sekä autonomisten autojen että kuljettajattoman raitiovaunun osalta tärkeimmiksi nousivat samat arvoluokat: hyväntahtoisuus, turvallisuus, itseohjautuvuus ja universalismi. Vaikka liikkumismuodot erosivat merkittävästi toisistaan, teknologian lisääntymisen pelättiin kummassakin tapauksessa asettuvan kiilaksi ihmisten väliseen vuorovaikutukseen, vähentäen inhimillistä kohtaamista. Ihannelilanteessa autonominen liikenne eristämisen sijaan yhdistäisi ihmisiä, ja helpottaisi ihmisten keskinäistä vuorovaikutusta.</p> <p>Katsaus on toteutettu strategisen tutkimuksen neuvoston rahoittamassa ETAIROS-tutkimushankkeessa, jonka keskeinen tavoite on vahvistaa julkisen ja yksityisen sektorin yhteistä ymmärrystä eettisten käytäntöjen ja pelisääntöjen välttämättömyydestä.</p>
ISBN, ISSN, URN	ISBN 978-951-38-8780-3 ISSN-L 2242-1211 ISSN 2242-122X (Verkkójulkaisu) DOI: 10.32040/2242-122X.2023.T418
Julkaisuaika	Huhtikuu 2023
Kieli	Suomi, englanninkielinen tiivistelmä
Sivumäärä	40 s. + liitt. 4 s.
Projektin nimi	ETAIROS
Rahoittajat	
Avainsanat	autonomiset järjestelmät, tekoäly, etiikka, vastuullisuus, arvot
Julkaisija	Teknologian tutkimuskeskus VTT Oy PL 1000, 02044 VTT, puh. 020 722 111, https://www.vtt.fi/

Title	The ethics of Autonomous Systems
Author(s)	Tuisku Salonen, Nina Rilla & Pietari Pikkuaho
Abstract	<p>Autonomous systems can increase safety, efficiency and sustainability, but there are many ethical issues involved in their development and use. This review examines the formation of AI-related values in three contexts related to autonomous mobility: tram transport, private cars and maritime transport. The aim is to highlight the debate on values related to the development and deployment of AI, especially from the perspective of users.</p> <p>The research data was collected in four virtual user workshops in 2021 and 2022. The results are based on a content analysis of the qualitative data and a thematic classification of the responses. Transport and mobility are part of people's daily lives, and no critical values that completely reject AI and autonomy emerged from the discussions. However, there are strong feelings and values associated with AI and autonomy in transport, as well as many positive and worrying expectations for the future. Users perceived the development of AI more as an opportunity, although there were some concerns about expectations of AI. When dealing with safety-critical environments such as transport, safety understandably becomes one of the most important values. From a user perspective, there was a trade-off between autonomous trams and cars: at best, autonomous solutions would increase transport safety, but at the expense of privacy and freedom of choice. For autonomous boats, the contradiction was reflected in users' fears of the safety impact of the technology.</p> <p>For both autonomous cars and driverless trams, the same values emerged as the most important: benevolence, safety, autonomy and universalism. Although the two modes of transport differed significantly, in both cases the increase in technology was feared to be a wedge between people, reducing the human encounter. Ideally, instead of isolating people, autonomous transport would unite them and facilitate their interaction.</p> <p>The review was carried out in ETAIROS research project, funded by the Strategic Research Council, with the key objective of strengthening the common understanding between the public and private sectors on the need for ethical practices and codes of conduct.</p>
ISBN, ISSN, URN	ISBN 978-951-38-8780-3 ISSN-L 2242-1211 ISSN 2242-122X (Online) DOI: 10.32040/2242-122X.2023.T418
Date	April 2023
Language	Finnish, English abstract
Pages	40 p. + app. 4 p.
Name of the project	ETAIROS
Commissioned by	
Keywords	autonomous systems, artificial intelligence, ethics, responsibility, values
Publisher	VTT Technical Research Centre of Finland Ltd P.O. Box 1000, FI-02044 VTT, Finland, Tel. 020 722 111, https://www.vttresearch.com

Eettinen tekoäly autonomisissa järjestelmissä

Autonomiset järjestelmät voivat lisätä turvallisuutta, tehokkuutta ja kestävyyttä, mutta järjestelmien ohjelmointiin ja käyttöön liittyy kuitenkin monia eettisiä kysymyksiä. Miten itseohjautuvat autot ja muut liikennevälineet muuttavat liikennettä ja miten ne voisivat lisätä liikenneturvallisuutta? Kuinka vastuullisuus todentuu monimutkaisten toimijajärjestelmien ekosysteemeissä? Miten jaettu ymmärrys eettisistä arvoista muotoutuu ja yhteinen arvonluonti mahdollistuu liikenteen kaltaisissa toimijaympäristöissä, joissa autonomisuuden aste kasvaa kohti täyttä autonomisuutta? Tässä katsauksessa tarkastellaan tekoälyyn liittyvien arvojen muodostumista kolmessa autonomiseen liikkumiseen liittyvässä kontekstissa: raitiovaunuliikenteessä, yksityisautoilussa ja meriliikenteessä. Tutkimusaineisto on kerätty virtuaalisesti järjestetyissä neljässä käyttäjätyöpajassa vuosina 2021 ja 2022. Tulokset perustuvat laadullisesta aineistosta tehtyyn sisällönanalyttiseen tulkintaan ja vastausten temaattiseen luokitteluun. Katsaus on toteutettu strategisen tutkimuksen neuvoston rahoittamassa ETAIROS-tutkimushankkeessa, jonka keskeinen tavoite on vahvistaa julkisen ja yksityisen sektorin yhteistä ymmärrystä eettisten käytäntöjen ja pelisääntöjen välttämättömydestä.

ISBN 978-951-38-8780-3
ISSN-L 2242-1211
ISSN 2242-122X (Verkkajulkaisu)
DOI: 10.32040/2242-122X.2023.T418



beyond the obvious