

Toni Pitkänen

**IHMISEN JA ROBOTIN VÄLINEN YHTEISTYÖ:  
AUTONOMISEN RAHTILAIVAN VAATIMIEN  
PROSESSIEN JA KÄYTTÖJÄRJESTELMIEN  
SUUNNITTELU VUOROVAIKUTUKSEN JA  
KÄYTETTÄVYYDEN NÄKÖKULMASTA**



JYVÄSKYLÄN YLIOPISTO  
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA  
2019

## TIIVISTELMÄ

Pitkänen, Toni

Ihmisen ja robotin välinen yhteistyö: autonomisen rahtilaivan vaatimien prosessien ja käyttöjärjestelmien suunnittelu vuorovaikutuksen ja käytettävyyden näkökulmasta

Jyväskylä: Jyväskylän yliopisto, 2019, 89 s.

Kognitiotiede, pro gradu -tutkielma

Ohjaaja: Saariluoma, Pertti

Rahtilaivan toiminta on tulevaisuudessa yhä autonomisempaa. Laivan kapteenin odotetaan pystyvän tarkkailemaan yhä enemmän itsenäistä laivaa sekä tarpeen tullen ohjaamaan sitä maista käsin. Tämän pro gradu -tutkielman tarkoituksena on tarkastella, minkälaisia seikkoja tulee varsinkin kapteenin suhteen ottaa huomioon autonomisen rahtilaivan etäohjaamisprosesseja suunniteltaessa. Näitä kokonaisuuksia ovat muun muassa vuorovaikutuksen luominen, automatisoitavien alueiden löytäminen, autonomian asteen päättäminen, ihmismäisen päätöksenteon huomioiminen sekä käyttöjärjestelmien luominen vuorovaikutukselle. Osa-alueiden luontiin ja puutteiden hahmottamiseksi on hyödynnetty aihetta tukevien tutkimusten tarkastelua. Tutkimuksista käy ilmi, että nimenomaan autonomisen rahtilaivan kannalta kognitiotieteellistä tutkimusta ei ole liiaksi, mutta siitä on kuitenkin havaittavissa joitakin vastauksia yllä mainittuihin asiakokonaisuuksiin: Esimerkiksi käyttöjärjestelmät ja -liittymät voivat olla samankaltaisia kuin itse laivan kannella ohjailtaessakin, ja jos uusia ominaisuuksia on hyvä luoda, voimme hyödyntää käytettävien järjestelmien ohjenuoria. Haastavaa automatisoinnissa lienee ihmismielen syöttäminen laivan järjestelmän ymmärtämään muotoon korvaten kapteenin toimintaa merenkulussa. Toinen haastava mutta tärkeä osa-alue on olosuhteellisen tietoisuuden hankkiminen ja analysointi päätöksenteon kannalta. Sitä tarvitsee niin laiva kuin sen ohjaajakin onnistuneeksi vuoropuheluksi.

Asiasanat: autonomia, rahtilaiva, kapteeni, merenkulku, vuorovaikutus, käytettävyys, kognitio, operointi, käyttöjärjestelmä, käyttöliittymä, robotti

## ABSTRACT

Pitkänen, Toni

Collaboration of Human and Robot: Designing Processes and Interfaces of Autonomous Cargo Ship from Interaction and Usability Point-of-View

Jyväskylä: University of Jyväskylä, 2019, 89 pp.

Cognitive Science, Master's Thesis

Supervisor: Saariluoma, Pertti

Cargo ships are expected to become increasingly autonomous and their captain to assess and, in need, intervene the situation of the process from land. The purpose of this master's thesis is to examine which kind of subjects must be taken into notice to create these processes of autonomous cargo ship utilizing tools for the captain. Examples of these ensembles are creating interaction, finding areas of automatization, deciding the level of autonomy, regarding human-action when making it a key value of automatic robots, and creating usable interfaces for interaction. To delve deeper into these process fields—finding ways of creating these and point out the missing links—research on these topics has been gathered and further analyzed. Based on existing research, we find not too many topics have considered all these exact points of interest in autonomous cargo ships' processes on cognition, but there is sufficiently data to obtain conclusions to some of them brought up above: for example, interfaces could be very similar to those seen on the ship's navigation bridge, and if new innovations are needed, we could use guidelines of existing interface usability research. The toughest parts, though, are expected to be “pouring” human mind into systems of autonomous ships to replace captain behavior in seafaring, as well as obtaining and evaluating information of situational awareness for decision making; both the ship and the captain, and between them.

Keywords: autonomy, cargo ship, captain, seafaring, interaction, usability, cognition, operation, GUI, interface, robot

# SISÄLLYS

1	JOHDANTO.....	6
2	TUTKIMUSKYSYMYKS JA TUTKIMUSMETODIT .....	9
	2.1 Tutkimuskysymys ja kirjallisuuskatsaus metodina.....	9
	2.2 Tutkimusmenetit .....	10
	2.2.1 Tutkimuskysymyksen asettelu.....	11
	2.2.2 Valittavat tietokannat .....	11
	2.2.3 Asiasanapohjainen haku .....	11
	2.2.4 Lähteiden kriteerit.....	12
	2.2.5 Hakutulosten seulonta metodologisesti .....	13
	2.2.6 Kirjallisuuskatsauksen tekeminen .....	13
	2.2.7 Syntetisointi.....	14
3	IHMISEN JA ROBOTIN VÄLINEN VUOROVAIKUTUS .....	15
	3.1 Tutkimusalan historiaa ja eri näkökulmia .....	15
	3.2 Interaktiivinen ja autonominen robotti .....	17
	3.2.1 Autonomian eri kasvot.....	17
	3.2.2 Robotti ja Turingin kone.....	18
	3.2.3 Interaktiivisuuden ja autonomisuuden toteutus .....	19
	3.3 Vuorovaikutus.....	20
4	AUTONOMISEN JÄRJESTELMÄN OMINAISUUKSIA .....	23
	4.1 Neljä pääominaisuutta .....	24
	4.2 Automatisoitavien toimintojen ominaisuuksia järjestelmissä .....	25
5	AUTOMAATION ERI TASOT SEKÄ TÄYDELLINEN AUTONOMISUUS.....	28
	5.1 Automaation eri tasot miehittämättömissä järjestelmissä.....	28
	5.2 Täydellinen autonomisuus? .....	31
6	IHMISMIELIEN MALLINTAMINEN.....	35
	6.1 ACT-R.....	36
	6.2 ACT-R/E.....	37
	6.3 Neuroverkot .....	38
	6.4 Mihin pisteeseen saakka ihmismieltä tulee mallintaa? .....	39
7	TIETOISUUS YMPÄRÖIVÄSTÄ TODELLISUUDESTA .....	41
	7.1 Tietoisuuden eri osa-alueet .....	42
	7.2 Tietoisuuden saavuttaminen.....	45
	7.3 Tietoisuuden mittaaminen .....	46
	7.4 Olosuhteellinen tietoisuus etäohjattavissa roboteissa .....	48

8	IHMISEN JA TEKNOLOGIAN VÄLISEN VUOROVAIKUTUKSEN ROOLIT .....	51
	8.1 Työnjohtaja, valvoja.....	51
	8.2 Operoija, toimija.....	52
	8.3 Mekaanikko .....	52
	8.4 Työtoveri, kumppani.....	53
	8.5 Sivustakatsoja .....	54
9	JÄRJESTELMÄT JA KÄYTTÖLIITTYMÄT ROOLIEN VÄLISINÄ TULKKEINA .....	55
	9.1 Keille järjestelmiä kehitetään?.....	55
	9.2 Graafiset käyttöliittymät.....	56
	9.3 Käytettävien järjestelmien ja käyttöliittymien suunnittelu etäohjattavissa järjestelmissä.....	57
	9.4 Ymmärrettävyys järjestelmän ja ihmisen välillä .....	60
	9.5 Järjestelmien mukautuminen eri tilanteissa.....	61
10	MERENKULUN AUTONOMISET ROBOTIT TUTKIMUKSISSA .....	64
	10.1 Tutkimuksia autonomisesta merenkulusta soveltuen löydettyihin asiakokonaisuuksiin .....	64
	10.2 Sovellutuksia aiemmista tutkimuksista autonomisiin laivoihin .....	68
	10.3 Ympäristölliset attribuutit autonomisessa laivassa .....	69
11	VIRHEISTÄ OPPIMINEN MUISTA LIIKENNÖINTITAVOISTA .....	71
12	JOHTOPÄÄTÖKSET JA POHDINTAA .....	74
13	LÄHTEET .....	79

# 1 JOHDANTO

Tekoäly yleistyy alinomaan monella tieteenalalla. Nykypäivänä jo esimerkiksi autonomisesti reittinsä määrittävien ja ajavien autojen etiikkaa pohditaan (Lin, 2016), miten esimerkiksi työpaikan ja kodin välillä liikumme. Työelämässä työskentelemme koko ajan enemmän automatiikan parissa sen korvaten työtehtäviämme, mutta se ei ainoastaan korvaa ihmistyöntekijöitä, vaan sen lisäksi automatisoitavista järjestelmistä tehdään ihmisen työkumppaneita (Acemoglu & Restrepo, 2018), missä yhteistyön sujuvuutta on analysoitava. Jopa taskuisiamme löytyvät älypuhelinmuodossa kontrollit, jotka voidaan liittää arkipäiväiseen ympäristöömme, kuten älykodin kodinkoneiden hallinta sekä niistä automaattisesti ja välittömästi saatavilla oleva tieto. Toisaalta niiden tietoturvalisyyden haavoittuvuuteen on kiinnitettävä erityishuomiota nopeasti muuttuvassa teknologiassa. (Fernandes, Jung & Prakash; 2016.)

Suomalaistutkijoiden mukaan (HS, 2017) tähän tekoälyn ripeään kehitykseen olisi alusta alkaen hyvä ottaa huomioon kognitiiviset näkökulmat, missä osallistetaan tekoälyn käyttäjien kyvyt ja tarpeet – niin työssä kuin vapaa-ajallakin tapahtuvassa tekoälyn hyödyntämisessä. Mikäli kehittäisimme enemmän kognitiotieteellisiä ominaisuuksia tekoälyihimme, voisimme saavuttaa käyttäjäystävällisen ja turvallisen tekoälyn, joka palvelee ihmiskuntaa paremmin kuin ilman kognitiivisia taitoja (HS, 2017) niin sanotusti pelkkänä koneena. Useinhan ihminen voi perustaa omat valintansa tunteisiin niiden sekoittuessa loogisen päättelyn kanssa, mikä vaikuttaa koko kognitioon eli käsitykseen haavoittavissa olevista kokonaisuuksista.

Ihmisen ja robottien, mukaan lukien tekoäly, välisen vuorovaikutuksen kehityksessä onkin tärkeässä roolissa kansainvälinen yhteistyö monen eri tutkimusalan kanssa. Koska kognitiotiede itsessään jo käsittää jo muun muassa psykologisen, teknologisen, sosiologisen ja kielitieteellisen näkökulman, tarvitsemme kansainvälistä yhteistyötä kattaaksemme vaikkapa kulttuurieroista muodostuvat nyanssit ihmisen ja robotin yhteistyön vakuuttamiseksi. Esimerkiksi Rolls-Roycen ja VTT:n autonominen laivateknologia on yksi malli tällaisesta monialaisesta tutkimuksesta, jossa kehitetään itseohjautuvia, autonomisia laivoja yhdessä monen eri tutkimusalan kanssa (VTT, 2017).

Autonomisen rahtilaivan kehittäminen voidaan tätä projektia ajatellen jakaa karkeasti ottaen kahteen kategoriaan: tekoälyn kehittämiseen (esim. Zhang, Wu & Zhu; 2018) sekä käytettävyydeltään ihmisläheisten, kognitiivisten järjestelmien suunnitteluun (mm. Yanco, Drury & Scholtz; 2004) ja laivan ohjauksvälineiden hallinnointiin muuttuvissa olosuhteissa (Dorneich, Whitlow, Rogers, Feigh & Robert; 2015). Täten katamme monitieteellisen älylaivaprosessin kognitiotieteellisen aspektin tarkentuessamme lähinnä jälkimmäiseen osa-alueeseen kahdesta. Älyllisesti laivan eli toisin sanottuna sen sisältämän järjestelmän tulisi itsessään kyetä esimerkiksi autonomisesti lähtemään satamasta, väistellä kulureitillään esteitä ja hallita eri sääolotiloja sekä saapua satamaan turvallisesti (esim. Moreira & Soares, 2010). Tätä varten voidaan kehittää erinäisten teknologisten ratkaisujen avulla toimivia loogisia, ihmismielen kaltaisia malleja, joissa tarvitaan monenkaltaista osaamista ja havainnointia, mistä kerrotaan tarkemmin luvuissa 6 ja 7.

Kaiken tämän toiminnan – satamasta poistuminen, avomerellä liikkuminen, satamaan saapuminen – aikana etäällä toimivalla kapteenilla tulee olla käsitys tapahtumien kulusta. Tästä syystä teknistä toteutusta tärkeämpiä osa-alueita tämän kognitiotieteen tutkimuksen kannalta ovat ihmisen ja etätyön mahdollistavien järjestelmien välinen vuorovaikutus ja selkeä käytettävyys (Man, Weber, Cimbritz, Lundh & MacKinnon; 2018), ihmisen tietoisuus laivaa ympäröivästä todellisuudesta, sekä ne operaatiot, joissa ihmisen toimintaa voitaisiin automatisoida ja tehostaa. Arviolta 85–96 % laivojen törmäyksistä toisiin kappaleisiin on johtunut ihmisestä (Rothblum, 2000), eli vaikkapa tässä autonomia on juuri omiaan auttamaan ihmisen määrittämien tehtävien saavuttamista luomaan tehostetumpaa yhteistyötä automaattisoidun laivajärjestelmän ja kapteenin välillä. Myös Yhdysvallat on nostanut tekoälystrategiaansa esiin sen epäkohdan, jossa laivojen operoijilla ei ole aina tarpeeksi olosuhteellista tietoisuutta ympäristössä tapahtuvista seikoista. Tähän tekoälystrategia aikoo tehdä parannuksia esimerkiksi antamalla kapteeneille tekoälyllä varustetuin keinoin hankittua tietoa nopeasti muuttuvissa, monimutkaisissa ympäristöissä, joissa välttämättä pelkkä ihminen ei havaitsisi vaaroja. (Yhdysvaltain puolustusministeriö, 11, 2018.)

Vaikka kyseessä on kognitiotieteen pro gradu -tutkielma, aihetta tarkastellaan aluksi hieman teknisistä lähtökohdista (Spong, Hutchinson & Vidyasagar; 2004) yhdistäen lähihistorian tutkimuksia (Asimov, 1942; Asimov & McKeever, 1982) pääpainon ollessa kuitenkin ihmismieltä tulkitsevien toimintaperiaatteiden tarkastelua projektin kokonaistavoitteiden eli autonomisen merenkulun onnistumisen kannalta. Kokonaisuonnistumisen pääosa-alueita ovat nimenomaan vuorovaikutus (Adams, 2005) laivan järjestelmän ja sen operoijan välillä etenkin kriittisimpinä hetkinä, kuten häiriötilanteissa (esim. Reason, 2000). Ihannetilanne toteutettavan järjestelmän kannalta olisi se, että laivan järjestelmät olisivat niin pitkälle kehitettyjä ja autonomisoituja, että laivan olisi tarpeen olla yhteydessä ihmiseen vain hädän hetkellä epävarmoissa tilanteissa. Tällöin se kykenisi matkustamaan satamasta toiseen vailla suurempaa ihmisen puutumista mutta niin, että etäältä käsin toimiva kapteeni pystyy kuitenkin seurata laivan toimintaa ja sen liikkeisiin vaikuttavia parametreja koko ajan. (Esim. Niculescu ym., 2007.) Halutessaan kapteeni voisi ottaa ohjat omiin käsiinsä, mutta

mitkä ovat ne toteutuksen kannalta olennaiset keinot, joilla tämänkaltaisia toimintoja voitaisiin lähteä toteuttamaan käytännössä? Entä millaisia tutkimuskysymyksiä kaiken kaikkiaan näinkin laajaan kokonaisuuteen on hyvä luoda? Seuraavassa luvussa kerrotaan aluksi tarkemmat tutkimuskysymykset ja -metodit, joilla kysymyksiin syvennyttään.



## 2 TUTKIMUSKYSYMYS JA TUTKIMUSMETODIT

Tämä pro gradu -tutkimus on toteutettu kirjallisuuskatsauksena. On aluksi syytä selventää, miksi kyseinen metodi on valittu sekä minkälaisia seikkoja tutkielmassa on huomioitu aineiston keruun ja analysoinnin ollessa kyseessä. Tässä luvussa kerrotaan tutkielman tutkimuskysymykset sekä ne metodit, joilla avoimena oleviin kysymyksiin poraudutaan. Ensimmäiseksi alaluvussa 2.1 kerrotaan tutkimuskysymykset sekä se, miksi aiheita varten on luotu kirjallisuuskatsaus. Toiseksi alaluvussa 2.2 esitetään tutkimuskysymysten luonteen pohjalta valittu metodi tutkimusaineiston hankkimiseen ja analysointiin.

### 2.1 Tutkimuskysymys ja kirjallisuuskatsaus metodina

Salminen (2011, 3) tiivistää Baumeisterin ja Learyn (1997) näkemykset kirjallisuuskatsauksen tärkeyden pääpiirteistä viiteen kategoriaan:

1. olemassa olevan teorian kehittäminen ja uuden luominen
2. teorian arviointi
3. kokonaiskuvan rakentaminen
4. ongelmien tunnistaminen
5. mahdollisuus kuvata teorian historiaa ja sen kehitystä.

Nämä osa-alueet ovat tiiviisti tämän pro gradu -tutkielman keskiössä tarkasteltaessa autonomisen rahtilaivan vaatimien prosessien määrittelyä ja luomista käytettävyyden näkökulmasta. Autonomisen – tai ehkäpä realistisemmin pitkälle autonomisoidun – rahtilaivan toteutus on suhteellisen nuori käsite niin ikään suhteellisen nuoren tekoälyn eli koneoppimisen aikakauden aikana, ja siksi tämä arviointi olemassa olevan tutkimuksen pohjalta toimii eräänlaisena kokonaiskatsauksena tähän prosessiin. Tekoälyn kehitys on usein keskittynyt ohjelmoimillisiin, matemaattisiin ja loogisiin teknologisiin kehityskohteisiin, joita tässäkin tutkimuksessa käsitellään hieman, mutta pääasiallisesti tämän tutkielman tarkoituksena on syventyä teknologian vastapuolelle: käytettävien

järjestelmien määrittelyyn, käyttämiseen ja ihmisen huomioon ottamiseen näissä etäohjattavissa, maista käsin operoitavissa käyttöjärjestelmissä varsinkin kapteenin kannalta. Tutkielman tutkimuskysymys siis kuuluu: Millaisia asiantuntemuksien kokonaisuuksia on huomioitava autonomisen rahtilaivan vaatimien prosessien määrittelyssä ja luomisessa käytettävyyteen pohjautuen?

Baumeisterin ja Learyn (1997) viittä kategoriaa tarkastelemalla voidaan paljastella kokonaistutkimuskysymys pienempiin osakysymyksiin, jotka on lueteltu seuraavassa. Erityisesti kohdat 1, 3 ja 4 ovat tämän pro gradu -tutkielman kannalta tärkeitä.

1. Millaista on olemassa oleva teoria ja tutkimus, ja kuinka sen kautta voidaan luoda uutta?
2. Kuinka aiemman tutkimuksen teoriapohja on sovitettavissa kyseessä olevaan aiheeseen?
3. Miten voimme rakentaa kokonaiskuvaa niinkin laajahkosta aiheesta kuin autonominen rahtilaiva? Kokonaiskuva rakentuu pienemmistä palasista, ja siksi nämä pienemmät kokonaisuudet on tunnistettava ja löydettävä ensin, minkä jälkeen voimme koota niistä kokonaisuutta.
4. Minkälaisia ongelmia kohtaamme autonomisten rahtilaivojen vaatimien prosessien suunnittelussa, määrittelyssä ja lopulta luomisessa? Ongelmat on tunnistettava ja luokiteltava eri kokonaisuuksiinsa, jotta edellisen kohdan kokonaiskuvasta tulee vakaampi eivätkä ongelmat yllätä vasta esimerkiksi järjestelmien luomisvaiheessa.
5. Kuinka esimerkiksi käytettävyyden ja tekoälyn teorian ja tutkimuksen historia on kehittynyt ja on vertailtavissa nykyhetkeen?

## 2.2 Tutkimusmenetelmät

Salminen (2011) on tiivistänyt osuvasti myös ne kirjallisuuskatsauksen keinot, joita edellä mainittuihin tutkimuskysymyksiin voidaan hyödyntää löytääksemme vastauksia niihin tässä kirjallisuuskatsauksessa. Tämän tutkimuksen kirjallisuuskatsauksen metodina toimii Salmisenkin (2001, 10) mainitsema Finkin (2005, 3–5) malli, jonka tarkoituksena on systemaattisena kirjallisuuskatsauksena luoda yhteenvedon aiempien tutkimusten tärkeimmistä aihepiireistä sekä runsaan tutkimusaineiston pohjalta luoda katsaus niin tutkimusalan historiaan kuin tuoreimpiinkin tutkimustuloksiin. Lisäksi systemaattisella kirjallisuuskatsauksella saadaan usein selville niitä puutteita ja aukkoja, joita ei ole vielä tutkittu, mikä on varsinkin tässä kyseessä olevassa autonomisen rahtilaivan järjestelmien luomisessa käytettävyyden suhteen tarpeen. (Salminen, 9.)

Salmisen (2011, 10) mainitseman Finkin (2005, 3–5) mallin seitsemänportaisen kirjallisuuskatsauksen perusteella on toteutettu tämä pro gradu -tutkielma metodologisesti. Seuraavassa on lueteltu nämä seitsemän kohtaa tämän tutkielman kannalta.

### 2.2.1 Tutkimuskysymyksen asettelu

Tutkimuskysymyksen asettelu on koko tutkimuksen lähtökohtana. Tutkimuskysymysten asettamisen pohjalta on rakentavampaa lähestyä aihetta jo tietyllä asiakokonaisuudella. Tutkimuskysymykset on asetettu tätä edeltävässä alaluvussa 2.1.

### 2.2.2 Valittavat tietokannat

Kun kysymys ja kysymykset on asetettu siten, että ne ovat selkeitä ja niihin voidaan löytää vastauksia, voidaan valita tietokannat, joista tietoa haetaan. Tietokannana on käytetty lähinnä *Google Scholar* -hakukonetta, jossa aineiston luotettavuutta voi arvioida lähes välittömästi viittausmäärän perusteella, tosin ai-noastaan sitä ei ole käytetty arviointiin, mistä on kerrottu lisää alaluvussa 2.2.5. Välillä on ajoittain tähän menetelmään yhdistetty tavallisia *Google*-hakuja, joista on löydetty hakutermeillä muun muassa uutisartikkeleja tai muita viittauksiin tarvittavia verkkosivuja, jotka ovat käsitelleet tutkielman aihetta spesifimmin.

### 2.2.3 Asiasanapohjainen haku

Tämän jälkeen, kolmantena kohtana, on valittu hakutermit, joilla aineistoa kerätään. Aineistonkeruu olennainen osa tämän tutkimuksen metodologiaa, sillä sen perusteella luomme kokonaisuuden, jota tarkastelemme ja jonka perusteella luomme yleistyksiä ja vastauksia tutkimuskysymyksiin. Ensiksi on tarkasteltu kokonaisaihepiiristä löytyvää tutkimusta, kirjallisuutta ja niitä ympäröiviä asiakokonaisuuksia, jotka ovat suhteellisen laajoja syleillen niin tekoälyä kuin kieli-tiedettäkin. Näin kokonaiskuva hahmotuu paremmin ja asiasanojakin on vai-vattomampaa hahmotella. Aineistoa on sitä varten kerätty lähinnä asiasanapoh-jaisena hakuna, koska aina avainsanoja ei ole välttämättä linkitetty tutkimuksiin, ja tähän *Google Scholar* toimii vaivatta asiasanapohjaisena hakukoneena, sillä se näyttää tulokset riippumatta siitä, ovatko hakusanat avainsanoina vai tutki-muksessa itsessään tekstin seassa.

Asiasanat eli hakusanat on kirjoitettu perusmuodossaan, sillä taivutettui-na hakutulokset voivat olla vähempimääräisiä. Asiasanoihin on kussakin kate-goriassa haun yhteydessä liitetty termejä, joilla selvennetään asiakokonaisuuksien kytkeytymistä autonomisiin rahtilaivoihin sekä niiden etäohjausta varten käytettäviin järjestelmiin. Näitä spesifejä asiasanoja ovat olleet muiden muassa *autonomous ship*, *autonomy ship*, *remote control ship*, *teleoperation ship*, *teleoperation autonomous*, *autonomous ship captain*, *ship interface*, *ship usability*, *ship interface usa-bility* sekä *autonomous ship interface*. Kokonaisuuksia on löydetty esimerkiksi samalla tavalla kuin Adamidesin, Christoun, Katsanoksen, Xenoksen ja Hadzi-lacoksen (2015; 256 & 257) metodiosiossa, mutta toisin kuin heidän tutkimuk-sessaan tässä tutkielmassa asioita on lähestytty asialähtöisesti, vaikkakin ajoit-tain tietyt tutkijat ovat enemmän edustettuina kuin toiset. Seuraavassa on mää-ritelty haussa käytettyjä asiasanoja kokonaisuuksittain:

- a. Asiasanat on valittu juurikin huolellisesti rajaamaan materiaalia vastaamaan tutkimuskysymyksiin, ja asiasanoina on käytetty taustateoriaa tukemaan muun muassa *human computer interaction*, *human robot interaction*, näiden lyhenteitä *HCI* ja *HRI* sekä *autonomous*, *autonomy*, *HRI interaction*, *automation*, *human robot usability*, *automation level*, *autonomy level*, *unmanned system autonomy* ja *perfect autonomy*.
- b. Seuraavaksi ihmismielen mallintamista autonomisten järjestelmien kannalta asiasanoina on hyödynnetty seuraavia: *modeling human mind*, *human mind model*, *modeling human mind neural network*, *total human mind modeling*.
- c. Tämän jälkeen yhtä tärkeimmistä aiheista, tietoisuutta ympäröivästä todellisuudesta, on tutkittu termeillä, kuten *situation awareness*, *situation awareness gather*, *situation awareness measure*, *situation awareness*, *situation awareness remote control system*, *situation awareness robot*, sekä yhdistelmillä, joissa *situation awareness* on korvattu sen lyhenteellä *SA*.
- d. Luvun 8 roolit ihmisen ja teknologian välisessä vuorovaikutuksessa on huomattu asiasanoilla *HRI roles*, *autonomous ship roles*, *autonomy roles* ja *autonomous system roles*.
- e. Sitten ehkäpä tämän tutkimuksen kannalta tärkeimmät luvut 9 ja 10 ovat muodostuneet seuraavanlaisten asiasanojen pohjalta: *user interface*, *graphical user interface*, *remote control interface*, *remote control interface*, *robot interface*, *ship autonomous interface*, *design autonomous system interface*, *autonomous system understand*, *autonomous system adapt*, *empirical autonomous ship*, *autonomous ship test*, *autonomous ship interface test* sekä *autonomous ship system environment*. Termit *user interface* ja *graphical user interface* on lähes aina haettu myös termeinä *UI* ja *GUI*.

#### 2.2.4 Lähteiden kriteerit

Neljäntenä vaiheena on lähteiden seulonta. Lähteiden kriteereinä on käytetty esimerkiksi juurikin *Google Scholarissa* lähteeseen viittaavien artikkelien määrää, artikkelin julkaisuvuotta sekä kieltä. Lähtökohtaisesti jokaisen artikkelin viitatusmäärän on ollut oltava vähintään useita kymmeniä sekä julkaisuvuoden tarpeeksi tuore muttei liian tuore, jotta se on löytänyt jo todistettavaa jalansijaa tutkimusalalta. Lisäksi kielen rajoituksena on ollut pääsääntöisesti englanti, jonka alaisuudesta löytyy enemmän ja viitatumpaa tutkimusta, mutta myös muutama suomenkielinen artikkeli on hyväksytty mukaan tärkeyden ja ajankohtaisuuden perusteella.

Vaikka tähänkin kirjallisuuskatsaukseen on haettu näkemyksiä niin vanhemmista tutkimuksista koskien teoriaa ja historiaosuutta kuin uudemmistakin, tekoälyä ja järjestelmien käytettävyyttä käsitteleviä tutkimuksia muodostaen niistä yhteen sitovia kokonaisuuksia täydentäen itse kutakin, tämän pro gradu -tutkielman kirjoitushetkellä moni asia ja tutkimus on varmasti edistynyt huo-

mattavasti koskien automaattista merenkulkua. Tutkimus on kirjoitushetkellä parhaimmista tutkimuksista löytyviä tematiikkoja punoen niitä toisiinsa, mutta aivan tuoreimpia, kirjoitushetkellä samalta vuodelta löytyviä tutkimuksia ei ole niin paljoa hyödynnetty tässä kokonaisuuskatsauksessa lähinnä kognitiivisesta näkökulmasta. Koska osa tutkimuksista on suhteellisen nuoria, niiden täydellistä validiutta ei ole aukottomasti pystytty todistamaan. Kun kirjallisuuskatsauksen valittujen artikkelien tärkeys on voitu todistaa jo usean vuoden ajan ja tutkimuksiin löytyy useampia vahvistavia viittauksia, se vankistaa löydöksiä eikä horjuta tutkimustuloksia vaikkapa piakkoin ilmestyvien tutkimuksien kumo- tessa aivan tuoreiden akateemisten tutkimusten tuloksia. Tämä olisi voinut käydä, jos sitä ei olisi otettu harkintaan.

## 2.2.5 Hakutulosten seulonta metodologisesti

Viides vaihe Finkin (2005, 3–5) systemaattisessa kirjallisuuskatsauksen mallissa on hakutulosten seulonta metodologisesti. Tämä tarkoittaa käytännössä sitä, että tutkimukseen valitaan kaikista laadukkain materiaali. Tässä pro gradu - tutkielmassa tutkimusten laatua on tarkasteltu Julkaisufoorumin (2018) sanoin sen ”arvioimia tieteellisiä lehtiä, sarjoja, konferensseja ja kirjakustantajia” sil- mällä pitäen. Arviointijärjestelmällä voidaan määrittää julkaisujen julkaisufoo- rumien luotettavuutta. Näin yksittäisten artikkelien toteaminen arvostetuksi julkaisuksi vahvistaa myös koko tämän tutkimuksen paikkansa pitävyyttä ar- tikkeliin yhdistyessä luotettavaksi kokonaisuudeksi. Kaiken kaikkiaan tämän tutkimuksen ollessa kirjallisuuskatsaus lähteiden muodostamien kokonaisuuk- sien tärkeys korostuu. Huomioitavaa on kuitenkin se, että menetelmällä voi- daan arvioida sitä foorumia, jossa tutkimus on julkaistu, ei julkaisua itsessään, mutta oletettavasti korkeatasoisilla julkaisijoilla on tarkat kriteerit tutkimuksille, joten siksi laadukkaammat tutkimukset mitä luultavimmin myös pääsevät vas- taavalle tasolle julkaisijan suhteen.

Julkaisufoorumista (2018) löytyvien artikkelien tasot on luokiteltu as- teikolla yhdestä kolmeen, joista 1 = perustaso, 2 = johtava taso sekä 3 = korkein taso. Tasolle 1 yltämättömät julkaisukanavat on merkitty tasolla 0. Lopuille tut- kimuksen lähteille ei löytynyt akateemista tasoa, kuten esimerkiksi järjestöjen tai yritysten verkkosivuille, sanomalehdille taikka ministeriöille. Kaikkiaan täs- sä pro gradu -tutkielmassa käytetyt lähteet lukeutuvat näille tasoille seuraavan- laisesti:

- a. taso 0: 4 tutkimusta
- b. taso 1: 35 tutkimusta
- c. taso 2: 30 tutkimusta
- d. taso 3: 11 tutkimusta.

## 2.2.6 Kirjallisuuskatsauksen tekeminen

Toiseksi viimeisenä vaiheena on itse kirjallisuuskatsauksen tekeminen. Tutki- muksesta on saatu suhteellisen kronologinen, loogisesti etenevä ja lopulta tut-

kimuskysymyksiin vastaava siten, että lähteet on kategorisoitu asiansanojensa, ajankohtaisuuksiensa ja aihepiiriensä perusteella omiin kokonaisuuksiinsa. Ensiksi on pureuduttu teoriataustaan ja tutkimusalan määrittelyyn liittyvään historiaan, jonka kautta aihetta pohjustetaan muun muassa robottien käytettävyyden lähtökohdista. Tämän jälkeen kerrotaan, mitä ylipäätään autonomialla sekä autonomisilla järjestelmillä tarkoitetaan, minkä myötä päästään autonomian tai automaation eri tasojen määritelmiin.

Sen jälkeen voimme syventyä niihin keinoihin ja päämääriin, miten ja kuinka pitkälle ihmismieltä voimme mallintaa. Tähän sisältyy siten olennaisesti tutkimusaineistossa erittäin tärkeäksi aiheeksi noussut kysymys: kuinka ihminen havainnoi ympäröivää todellisuutta ja käyttää hyväkseen olosuhteellista tietoisuutta? Sen seurauksena on tärkeää tarkastella niitä rooleja, jotka liittyvät autonomisten järjestelmien käyttäjiin ja kuinka he kaikki käsittelevät olosuhteellista tietoisuutta vuorovaikutteisesti keskenään.

Ja totta kai tähän tarvitaan käyttöjärjestelmiä varsinkin kapteenin kannalta, joka on vastuussa kokonaisprosessin onnistumisesta, mistä päästään yhteen tärkeimmistä asiakokonaisuuksista vastaamaan tämän tutkielman tutkimuskysymyksiin, eli lukuun 9; käyttöjärjestelmien ja -liittymien luominen eritoten laivan ja kapteenin välille käytettävällä tavalla yhteen vetäen aiempien lähteiden löydöksiä liittyen muun muassa ihmisen ja robotin väliseen vuorovaikutukseen, robotiikkaan sekä olosuhteellisen tietoisuuden hankkimiseen. Lopulta päästään siihen, miten nykyään esimerkiksi käytännön tutkimukset ovat asiaa lähestyneet. Koska tutkielma käsittää monia asiakokonaisuuksia eikä läheskään kaikkiin löydy vastauksia, eri osa-alueiden pohdinnat ovat usein samassa yhteydessä kuin itse asian käsittely eivätkä pelkästään tutkielman lopussa. Näin on ehkä helpompaa käsittää pohdinnan liittyvän läheisyydessä keskusteltuun kokonaisuuteen ja sitoa pohdinta tuoreessa muistissa olevaan kontekstiin.

### 2.2.7 Syntetisointi

Viimeinen, seitsemäs vaihe Finkin (2005, 5) kirjallisuuskatsauksen portaikossa on syntetisointi. Sillä tarkoitetaan sitä, kuinka tuloksia vedetään yhteen ja luodaan kokonaisuudesta yhtenäistävät tulokset ja pohdinnat. Tämä vaihe on ehkäpä tämän tutkimuksen kannalta haastavin vaihe: kuten Salminen (2011, 10) toteaa Baumeisterin ja Learyn (1997) sanoin syntetisoinnin olevan haasteellista, siitä ei todennäköisesti tee helpompaa se, että kyseisen pro gradu -tutkielman aihepiiri on laaja käsittäen monia eri tieteenaloja, useiden eri vuosikymmenten tutkimustuloksia sekä samalla sen ollen kuitenkin nuorehko tieteenala itsessään tekoälyn saralla. Siitä syystä syntetisointiin on alkuun otettu muutamia luotettavimpia artikkeleja olevia katsauksia nykypäivän autonomisten rahtilaivojen käytännöntutkimuksista, joissa voidaan nähdä, miten esimerkiksi järjestelmien käytettävyyttä on tutkittu suhteellisen autonomisissa, etäohjattavissa järjestelmissä, kuten laivoissa. Sen jälkeen syntetisointia on pohdittu erikseen, voitaisiinko esimerkiksi tutkimusaloja yhdistämällä eli virheitä havaitsemalla etukäteen muista automaattisista ohjausjärjestelmistä oppia jotain – tässä tapauksessa lentokonejärjestelmistä.

### 3 IHMISEN JA ROBOTIN VÄLINEN VUOROVAIKUTUS

Ihmisen ja robotin välisen vuorovaikutuksen luonti oikeellisella tavalla on kaikesta lähtökohtana automatisoitujen järjestelmien kanssa toimimisessa. Robotti tai järjestelmä on usein tärkeässä asemassa sijaitseva välillinen keino saavuttaa ihmisen toivoma lopputulema eikä lainkaan itseisarvo sinänsä, joten on oleellista tarkastella niin ihmisen ja robotin välisen vuorovaikutuksen tutkimusalan määritelmiä kuin robotin tai järjestelmän eli välikappaleen määritelmääkin yksityiskohtaisemmin. Alaluvussa 3.1 perehdytään aluksi ihmisen ja robotin välisen tutkimuksen historiaan ja sen siirtymiseen nykyaikaiseen robotiikkaan; alaluvussa 3.2 puolestaan selvennetään mutta samalla kyseenalaistetaan tutkimuksessa käytettyjen termien merkityksiä, kuten robotti, botti, automatiikka ja autonomia, sekä viimeisenä alaluvussa 3.3 pureudutaan vuorovaikutuksen luomiseen ihmisen ja robotin eli automatisoidun järjestelmän välille.

#### 3.1 Tutkimusalan historiaa ja eri näkökulmia

Ihmisen ja robotin välinen vuorovaikutus (IRV, engl. *Human-Robot Interaction, HRI*) pohjautuu jo 1940-luvulla Isaac Asimovin muodostamaan kolmeen sääntöön robotiikasta:

1. Robotti ei saa vahingoittaa ihmistä.
2. Robotin täytyy totella ihmistä.
3. Robotin täytyy paeta, mikäli joku tai jokin yrittää satuttaa sitä.  
(Asimov, 1942; Asimov & McKeever, 1982, 1.)

Vielä tänäkin päivänä näiden kolmen ohjeistuksen voidaan katsoa pitävän pääpiirteittäin paikkansa ihmistä varten toteutettavien, autonomistenkin robottien perustoimintaperiaatteita muodostettaessa. Robotin halutaan usein olevan vaaraton, tottelevainen ja itsesuojeluvaiston omaava entiteetti.

Eritoten tämä tottelevaisuus oli myös lähtökohtana IRV:ssa jo vuonna 1992, jolloin ihmisen ja tietokoneen välinen vuorovaikutus sai ensimmäisen mainintansa tehdasteollisuuden kontekstissa (Scholtz, 2002b). Tehtaissa ihminen nähtiin aiemmin etänä operoivan järjestelmän suorana valvojana, joka jatkuvasti tarkkailee kokonaistoimintaa. Tähän telerobotiikan käsitteeseen sisältyi, ja yhä sisältyykin suurelta osin, se että kone osaa tietyllä tasolla matkia ihmisen aistimiskykyjä ulottaen ne omaan ympäristöönsä eli etäälle operoijasta. (Sheridan, 1992.) Tämä voidaan nähdä ikään kuin ihmisen ja tietokoneen välisen vuorovaikutuksen (ITKV, engl. *Human-Computer Interaction, HCI*) sekä IRV:n välisenä murroksena (Sheridan, 1992). Vuodesta 1992 on joka tapauksessa pitkä matka ihmisen ja robotin väliseen vuorovaikutukseen, joka on nuorehko käsite eikä lainkaan yksikäsitteinen selite tänäkään päivänä.

IRV-tutkimus on kuitenkin saanut lyhyehkön historiansa myötä monet eri kasvot ja sitä on tarkasteltava lukuisilta muiltakin eri kanteilta kuin vain Asimovin kolmen säännön tai tehdasteollisuuden lähtökohdan perusteella. Sen voidaan katsoa muotoutuneen vasta myöhemmin niin ITKV-tutkimuksen kuin ihmisen ja teknologian väliseen vuorovaikutukseen (ITEV, engl. *Human-Technology Interaction, HTI*) nojautuvan katsantokannan myötä. IRV on lähtökohtaisesti ITKV:n ja ITEV:n rinnalla oma itsenäinen tutkimusala, johon yhdistyvät muun muassa sosiologiset, kulttuuriset, psykologiset, kognitiiviset, kielelliset, teknilliset sekä käytettävyydelliset aspektit. (Esim. Saariluoma & Leikas, 2010; Walters ym., 2009.)

Tosin IRV eroaa määritelmällisesti ja merkitykseltään ITKV:sta ja ITEV:sta, vaikkakin se usein on palanen niitä ja niiden sisältämiä osa-alueita. Toisin kuin nämä kaksi muuta tutkimusalaa, IRV:ssa ihminen ikään kuin irtaannutetaan kauemmaksi fyysisistä roboteista, samalla kun roboteista tehdään edeltäjiäänkin autonomisempia, enenevässä määrin itsenäisempiä vailla vaadittavaa ihmiskontaktia. (Breazeal, 2004, 181.) Pitää yhä paikkansa, ja on mitä luultavimmin yhä yleistyvämpää, että robotit toimivat tulevaisuudessa ihmisen kanssa työkumppaneina samoissa ympäristöissä (esim. Bruemmer ym., 2005), kuten vaikkapa ruohonleikkurit ja vanhustenhoitorobotit, mutta samalla niille annetaan jalansijaa myös sellaisissa konteksteissa ja miljöissä, joissa ihmisen työpanosta halutaan usein vähennettävän, kuten merenkulussa (VTI, 2017).

Lisäksi IRV-suuntauksen lisäominaisuuksia ITKV-kantaan nähden on se, että IRV voidaan määritellä ihmisten ja robottien sekä niiden vuorovaikutuksen tutkimiseksi (Fong, Thorpe & Baur; 2001, 256). Tällöin tarkasti ITKV:n määritelmää, "ihmisen mielestä hyödyllisten järjestelmien käytettävyyden tutkiminen liittyen muun muassa ergonomiaan ja käyttöliittymään" (Dix, Finlay, Abowd & Beale; 2009), tulkiten IRV eroaa tästä siten, että käytettävyys ei ole kirjaimellisesti niin suuressa roolissa. Mutta pohtiessa vaikkapa sitä periaatetta, että robottia käyttävällä ihmisellä täytyy olla suhteellisen helppokäyttöinen käyttöliittymä robotin kanssa toimimiseen, käytettävyys on usein joka tapauksessa merkittävässä osassa IRV-tutkimusta. Ehkäpä täysin yksiselitteistä tai yksimielistä määritelmää IRV-alasta ei voi välttämättä sanoa, vaan sen muovatessa montaa muuta eri tieteenalaa on tilannekohtaisesti punnittava, mitä kussakin toteutettavassa projektissa IRV:lta vaaditaan.



## 3.2 Interaktiivinen ja autonominen robotti

Robotin, automatiikan ja autonomian määritelmät voivat joskus olla hieman häilyviä. Robotti on usein kuvitettu ihmisenmuotoiseksi, metallinväriseksi fyysiseksi kappaleeksi, joka on tehty liikehtimään kankeasti ja olemaan minimalistisessa vuorovaikutuksessa ihmisen kanssa, ehkäpä puhuen robottimaisella äänellä, mutta nykypäivänä robotit harvoin kuitenkaan ovat enää sellaisia. Ne ovat usein integroituja jokapäiväiseen elämäämme siten, että emme välttämättä edes huomaa niitä. Tästä on esimerkkinä älylaitteena toimiva automatisoitu jääkaappi, joka osaa antaa palautteena tietoa terveydestämme hektisen arkemme keskellä: kun moderni elämäntyyli on yhä enemmän sitä, että emme viettä yhtä paljon aikaa keittiössämme kuin aiemmin, tällainen älyjääkaappi on omiaan antamaan dataa valintojemme terveellisyydestä (Luo, Jin & Li; 2009). Miel-lämmekö tällaista laitetta edes robotiksi? Vai voiko sitä jopa kutsua autonomiseksi robotiksi?

Siksi automatiikan, autonomian ja robotiikan käsitteitä tässä tutkimuksessa on selvennetty tässä luvussa. Alaluvussa 3.2.1 tehdään pesäeroa automaation ja autonomian välille, kun taas alaluvussa 3.2.2 puntaroidaan robotin tai jopa botin toimintoja ja merkityksiä nykypäivänä liittyen esimerkiksi Turingin (1950) koneen määritelmiin. Lopulta alaluvussa 3.2.3 pohditaan sitä, kuinka robotista, botista tai järjestelmästä voitaisiin luoda mahdollisimman automatisoitu, jopa autonominen, ja vuorovaikutukseen kykenevä ihmisen eli valvojan tai operoijan kanssa.

### 3.2.1 Autonomian eri kasvot

Automaatio ja autonomia voivat sekoittua joskus keskenään, ja siksi onkin tärkeää tehdä niiden välille selvä ero tämänkin tutkielman kannalta sekä täsmentää autonomian määritelmää. Autonomian voidaan katsoa olevan eräänlainen automaation ylin taso, jolla jokin järjestelmä on automatisoitu eli itsenäistetty toiminnaltaan niin pitkälle, että se osaa toimia jossakin määrin itsevaltaisesti: esimerkiksi määrittää omia tavoitteitaan, tehdä päätelmiä sekä toimia niiden mukaisesti (Huang, Pavek, Novak, Albus & Messin; 2005). Huang ja muut (2005) määrittävätkin, että ylin eli viides automaation taso on autonominen merkiten sitä, että esimerkiksi miehittämättömissä ajoneuvoissa vähin määrä vuorovaikutukseen on vaadittu ja enin määrä itsenäiseen työskentelyyn on onnistuttu toteuttamaan eli automatisoimaan. Näin laite kykenee pääosin itsenäisesti toimimaan vailla suurta tarvetta kommunikoida.

Castelfranchi ja Falcone (2003) ovat samoilla linjoilla mutta täsmentävät autonomian määritelmää: autonomia voidaan yksinkertaisesti ilmaista funktiona, jossa jok'ikinen laitteen arkkitehtuurin osa tai muu ehto onnistuneelle toiminnalle voidaan ilmaista abstraktina resurssina tai voimana päämäärää varten. Näin kyetään myös kohdentamaan heikkoudet ja puutteet sekä riippuvuudet. (Castelfranchi, 2000; Castelfranchi & Falcone, 2003.) Lisäksi Castelfranchi ja Falcone (2003) määrittävät kahdeksan eri automatisoitua, autonomista osa-

aluetta, joista seuraavassa on lueteltu tämän pro gradu -tutkielman kannalta tärkeimmät:

- epäsosiaalinen autonomia, jotta autonominen agentti eli toimija kykenee mahdollisimman itsenäisesti liikkumaan ympäristössään eli niin satamissa kuin avomerelläkin
- toiminnallinen autonomia, jotta agentti kykenee itsenäiseen päätöksentekoon, suunnitteluun ja ongelmanratkaisuun
- päämääräautonomia, joka ei tarkoita itsessään pelkästään ulkopuolisten päämäärien tavoittelua, vaan agentin on kyettävä luomaan omiaan, pienempiä päämääriä
- tietoinen autonomia, eli agentin tulee olla tietoinen ympärillään tapahtuvasta toiminnasta sekä havainnoitava muuttuvia ympäristöjä.

Nämä autonomian eri tavoin määritellyt osa-alueet ovat suhteellisen yksikäsitteisesti yhdistettävissä autonomisen rahtilaivan toimintoihin. Laivan tulee aivan optimaalisessa tilanteessa olla melko itsenäinen liikkumaan erilaisissa vesistöissä, kyettävä huomaamaan ympäristössä ja päämäärissään puutteita sekä asettamaan uusia välipäämääriä itsevaltaisella tavalla. Näistä itse autonomisen, arkkitehtonisen järjestelmän ominaisuuksista on kerrottu enemmän luvussa 4.

### 3.2.2 Robotti ja Turingin kone

Sanalla *robotti* on usein tarkoitettu fyysistä robottia. Robotin voidaan katsoa olevan mekaaninen kone koostuen teknisistä elektronisista kappaleista sisältäen ihmisen syöttämiä parametreja ja funktioita, joiden perusteella se toimii ihmisen määritysten mukaisesti. Sitä voidaan fyysisiltä ominaisuuksiltaan luonnehtia jäyhäliikkeiseksi, olomuotoaan vaihtamattomaksi, kinemaattiseksi, konenäön taitavaksi sekä ohjailtavaksi yksikökseen. (Esim. Spong ym., 2004.) Mutta robotti ei ole kuitenkaan pelkästään teknisten osiensa summa: sen toimintaa ja muun muassa kinemaattisia liikkeitä ohjaa joukko algoritmeja, ohjelmia ja muita loogisia joukko-operaatioita, jotka määräävät sen jokaista liikettä. Näin robotti palvelee ihmistä sen alkuperäisessä merkityksessään (tsek. *robot*; työ). (Spong ym., 2004.)

Mekaanisen robotin sisällä puolestaan on usein *agentti* eli *botti*: kohtalaisen itsenäisesti työskentelevä järjestelmä, joka toimii ihmisen tavoin ihmisen syöttämien ohjeistuksien mukaisesti. Agentin on kuvailtu ohjailevan itselleen osoitettuja robotin mekaanisia kappaleita kinemaattisesti, jotta fyysinen ulkokuori eli niin sanottu kehikko pystyy suorittamaan tarvitsemansa tehtävät. Kuitenkin *agentti* on oikeastaan ainoastaan kattotermi laaja-alaiselle ja yksipuoliselle tutkimus- ja kehityskohteelle, sillä terminä se kyetään purkamaan auki yksityiskohtaisemmin sekä määrittelemään muun muassa koskemaan vähemmän käsinkosketeltavia asioita. (Nwana, 1996.) Näitä ovat nykyaikana vaikkapa tietokonejärjestelmiä kontrolloivat botit, jotka esimerkiksi ohjailevat käyttöliittymää; alituisesti oppivat älybotit tai yhteistyössä ihmisen kanssa toimivat botit (Nwana, 1996). Botit voivat olla siis mihinkään fyysiseen laitteeseen kytkemättömiä,

ikään kuin kaikkialla ympärillämme olevia botteja, joita ihminen ei välttämättä edes miellä boteiksi, tai joista ihminen ei osaa ajatella, että ne voivat tarkkailla käyttäytymistämme koko ajan. Tällaisia botteja ovat esimerkiksi sosiaaliset botit: ne voivat vaikuttaa niin ihmismäisiltä kuin mahdollista, ja ihmisen voi olla vaikea jopa erottaa oikeaa ihmistä tietokoneesta. (Ferrara, Varol, Davis, Menczer & Flammini; 2016.)

Nykyaikaisen tekoälyn ja tällaisten robottien loogisen toimintakyvyn kehityksessä käytetään ainakin jossain muodossa usein vuosikymmentenkin takaa kantautuvia ideologioita. Esimerkiksi Turingin (1950) koneen malli on yksi siteeratuimmista teorioista tekoälyn kehittämisen kannalta. Se pohjaa ajatuksensa sille, että kone lukee nauhalta symboleja eli merkkejä, joita kone tulkitsee sille syötettyjen säännösten perusteella. Tällöin muun muassa robotin voidaan sanoa olevan eräänlainen tietokoneohjelma; ehkäpä joskus jopa Turingin kone?

Kuitenkaan kaikki robotit ja agentit eivät ole suoranaisesti Turingin koneita, ja Turingin (1950) alkuperäinen määritelmä tällaisesta koneesta onkin todettu suurelta osin puutteelliseksi modernien laskennallisten ongelmien ratkaisuun (Leeuwen & Wiedermann, 2001). Turingin kone ei osaa esimerkiksi vuorovaikutuksessa ihmisen kanssa ottaa huomioon tämän sukupuolta, sillä se tulkitsee vain olemassa olevaa "raakaa" dataa, eikä se täten kykene olemassa olevan, analysoitavan tiedon tulkitsemiseen muuten kuin sille määritellyllä merkitysjoukolla. (Hayes & Ford, 1995, 977.) Toisin sanottuna Turingin koneet eivät siis kykene vuorovaikutukseen ihmisen kanssa, koska ne ovat järjestelmältään ja parametreiltaan sulkeutuneita (esim. Wegner, 1998, 347).

### 3.2.3 Interaktiivisuuden ja autonomisuuden toteutus

Kuitenkin robottien halutaan toimivan niin itsenäisesti kuin ihmistenkin kanssa yhteistyössä. Tästä syystä robottien autonomisuudeksi sekä ulkoisten, alati vaihtuvien muuttujien huomioimisen ohjelmoiminen on paljon muuta kuin pelkkiä yksinkertaisia algoritmeja ja sulkeutuneita järjestelmiä. Mitä siis vaaditaan, jotta robotista saadaan interaktiivinen ja autonominen, ja miten se eroaa niin kutsutusta tavallisesta palvelurobotista?

Koska epäautonomisten eli Turingin (1950) koneen toimintaa noudattavien robottien järjestelmät ovat sulkeutuneita, niiden parametrit eivät ole uudelleenmääriteltävissä. Interaktiivisten robottien ja niiden järjestelmien eli agenttien tulee olla avoimia ulkopuolisille muuttujille, minkä myötä ohjelmallisesti ja ohjelmoinnillisesti interaktiivisuuden toteuttaminen muodostuu huomattavasti haasteellisemmaksi operaatioksi. Interaktiivisten käyttöjärjestelmien onkin osattava mallintaa reaaliaikaisesti ohjelmoinnillisia *objekteja* eli muuttujia, joilla on omia sisäisiä muuttujiaan; kuten merenkulussa vaikkapa sää, joka itsessään sisältää jo tuulen ja merenpinnan aaltoilun – puhumattakaan näiden omista muuttujistaan ja lukuisista muista toisiinsa vaikuttavista tekijöistä. (Wegner, 1998, 347.) Tästä syystä järjestelmä ei voi olla sulkeutunut, koska kaikkien mahdollisten objektien ja parametrien määrittäminen voi koitua mahdottomaksi. Vaikka monessa teknologisessa ratkaisussa Turingin koneet ovat tärkeässä roolissa ja niitä soveltamalla voidaan saada toimivia ratkaisuja, pelkästään

niiden varaan ei vuorovaikutteista, autonomista järjestelmää voi rakentaa, koska Turingin kone sulkee ulkopuolisen maailman suorittaessaan laskuoperaatioita. (Wegner, 1998, 349.)

Ei ole välttämättä olemassa yhtä tai yhtäkään ainoaa keinoa toteuttaa järjestelmälle vuorovaikutus ja autonomisuus täysin toimivalla tavalla, koska kaikkien mahdollisten muuttujien luetteloiminen lienee mahdotonta. Kuitenkin lähelle interaktiivisuutta ja autonomisuutta voidaan päästä eri menetelmin, joista eräässä vaihtoehdossa on Turingin (1950) koneen idea taustalla siinäkin. Yhtämittaisten Turingin koneiden (engl. *Persistent Turing Machine, PTM*) suorittamien operaatioiden myötä voidaan arvioida monen eri vaihtoehdon arvokkuutta tai hyödyllisyyttä vuorovaikutteisesti. (Goldin, Smolka, Attie & Sonderegger; 2004, 125.) Tässäkin ajatusmallissa autonomisuuden aspekti jää melko vaillinaiseksi, sillä vaikka järjestelmästä saadaan vuorovaikutteinen peräkkäisten operaatioiden myötä, se ei vielääkään poista sitä tosiseikkaa, että reaali maailmassa on otettava huomioon monta tekijää yhtäaikaaisesti; ei peräjälkeen. Laajentamalla tätä ajattelua rinnakkaisten rajoitteiden ohjelmointiin (engl. *Concurrent Constraint Programming*) tai siihen erikoistuneeseen ohjelmointikielen *CHR* (engl. *Constraint Handling Rules*; rajoitteidenhallintasäännöt) voidaan päästä jo suhteellisen lähelle haluttua ideologiaa; robotti kykenee tekemään autonomisia, samanaikaisia päätöksiä vertaillen eri vaihtoehtoja avoimessa reaaliaikaisessa ympäristössä sekä ollen vuorovaikutuksessa ihmiseen päätöksenteon lomassa. (Goldin ym., 2004, 125; Sneyers, van Weert, Schrijvers & De Koninck; 2010, 29.) Kaikki ohjelmointikieliet eivät kuitenkaan kykene synkroniseen laskentaan: on tärkeää kyetä vertailemaan eri ohjelmointikielten hyötyjä ja haittoja suhteessa toivottuihin toiminnallisuuksiin (esim. Goldin ym., 2004, 122). Ehkäpä yksi suosituimmista ja helppokäyttöisimmistä ohjelmointikielistä koneoppimisen toteuttamiseksi on *Python* (2018). Esimerkiksi *Pythonin* eri moduuleilla voidaan ratkaista vähemmänkin hallittavissa olevia matemaattisia ja funktionaalisia haasteita tehokkaasti, helppokäyttöisesti ja suhteellisen riippumattomasti analysoiden muun muassa suuria määriä dataa (Pedregosa ym, 2011). Toisaalta tässäkin, kuten kaikissa ohjelmointikielissä, päästään siihen alkupisteeseen, että tiedon pitäisi olla tiedettävissä ja laskettavissa, jotta jokin ohjelmointikieli ja täten autonominen robotti pystyisi tekemään jonkinlaisia päätelmiä ja raportoida selkeästi siitä ihmiselle – puhumattakaan siitä, onko tieto aina oikeellista.

### 3.3 Vuorovaikutus

Ihmisen ja robotin välisessä (IRV) tutkimuksessa vuorovaikutus on kriittisessä roolissa. Robotti tai autonominen robotti on olemassa lähes ainoastaan palvelukseen ihmistä, eikä ihminen olekaan luonut robottia apulaisekseen turhaan. Siksi on osattava tunnistaa käyttäjän tarpeet ja vaatimukset, jotta vuorovaikutus kyetään luomaan tarkasti räätälöidyllä tyylillä. Robotti ei voi tietää tai aina arvatakaan oikein hyväksytyjä käyttäytymissääntöjä tai -normeja, ellei niitä ole tarkasti määritelty sille. (Esim. Adams, 2005.) Robotti on myös lähes kaikilta aisteiltaan ja ymmärryksen tasolla ihmistyöntekijää heikompi, ja näiden kriittis-

ten osa-alueiden luomiseksi tarvitaan aivan uudenlainen vuorovaikutusmekanismi. Autonominen robotin tulee myös osata arvioida, minkä tason vuorovaikutusta kyseinen kommunikointi on; kenen kanssa, millaisessa kontekstissa ja milloin se tapahtuu. (Adams, 2005, 450.)

Myös vuorovaikutuksen määrä ihmisen suhteessa robottiin on elintärkeä. Muun muassa Parasuraman, Galster, Squire, Furukawa ja Miller (2005) huomasivat tutkimuksessaan ihmisen valvoessa kahdeksaa samanaikaista autonomista miehittämätöntä ajoneuvoa, että neljän valvottavan robotin kohdalla ihmisen tarkkaavaisuuskyky kaikkia kohtaan oli äärimmäisen korkea. Kun tarkkailtavien autonomisten robottien määrä kasvoi kahdeksaan, ihminen herpaantui huomattavasti enemmän eikä kyennyt kasvaneen työtaakan ja sen tuoman kognitiivisen työkuorman vuoksi suoriutumaan tehtävästään enää läheskään yhtä hyvin. Tämä alleviivaa sitä tosiasiaa, että autonomisten robottien tulisi olla yhä enemmän vuorovaikutuksessa toistensa kanssa esimerkiksi valvoen toistensa toimintaa kuin myös antaa ihmiselle tiiviimpää, tarkempaa ja helpommin analysoitavaa dataa niiden havainnoimista tekijöistä. (Parasuraman ym., 2005, 491.)

Ihmisen vuorovaikutus autonomiseen robottiin ei olekaan ainut keino lähestyä ja tutkia IRV:ta. Robotti kykenee usein antamaan ihmisestä tarkempaa analyysia, kuin ihminen itse siihen pystyisi tarkkailemalla itseään. Ihmisen tarkkaillessa robotin toimintaa ei ole välttämättä mahdollista samalla yhtä tehokkaasti tarkkailla omaa toimintaansa, joten robotti voi antaa arvokasta tietoa esimerkiksi siitä, milloin ihmisen keskittyminen herpaantuu sekä milloin vaaditaan jonkinlaisia toiminnanmuutosta. Tällöin itse robotin ja autonomisen robotin yksi ominaisuuksista tulee olla vuorovaikutus ihmisen suuntaan juuri oikealla hetkellä. (Devin & Alami, 2016.) Autonominen robotin tulee osata havainnoida ihmisen ja muiden agenttien eli toimijoiden käyttäytymistä näiden päämäärien, suunnitelmien ja toimenpiteiden kautta. Tämän kerätyn datan robotti käsittelee niin, että vain oleellinen osa jää jäljelle. Täten kognitiivinen kuormitus ihmisellä vähenee. Näin myös karsitun, eksaktin informaation linkityessä aiempaan dataan kokonaisvirhealttius pienenee, kun analysoitavaa tietoa on määrällisesti vähemmän kuin jatkuvassa vuorovaikutuksessa syntyneen toimintatavan myötä. (Devin & Alami, 2016, 326.) Ihminen kykenee siis käsittelemään kokonaisuutta ja tärkeitä tietoja tarkemmin, kun sen määrä on rajattu nimenomaan tärkeimmäksi katsottuun. Toisen robotin kanssa vuorovaikutuksessa tiedon määrä puolestaan voi korreloida vaikkapa ohjelmallisesti todennäköisyyksien laskemisessa, kun jonkin skenaarion tapahtumistodennäköisyyttä pohditaan.

Yksi vuorovaikutuksen osa-alueista on robottien keskinäinen vuorovaikutus. Usein niin kutsuttu perinteinen IRV on käsitellyt aihetta pääsääntöisesti ihmisen ja robotin välisenä vuorovaikutuksena eli isäntä-renkiperiaatteena (engl. *master-slave*-asetelma), mutta yhä tiiviimmin käy niin, että roboteilta vaaditaan autonomista älykkyyttä olla vuorovaikutuksessa samanlaisen kumppaninsa kanssa. (Fong ym.; 2005, 2.) Tämä puolestaan luo yhden vanhimmista ongelmista tekoälyn kanssa: miten robotin järjestelmä kykenee tunnistamaan, että se on vuorovaikutuksessa toisen robotin eli järjestelmän kanssa; ei nimenomaan ihmisen (Turing, 1950).

Autonomisen robotin on tiedettävä tarkasti myös, miten sen on sopivaa olla vuorovaikutuksessa ihmiseen tai toiseen autonomiseen robottiagenttiin. Ihmiset ovat koko elinaikansa ajan kehittäneet uusia vuorovaikutustaitoja toistensa kanssa: pärjätäksemme olemme joutuneet omaksumaan sosiaalista älykkyyttä sekä tunneälyä kanssaeläjiimme nähden. Tämä sama on edessä autonomisten robottien ja ihmisten välisessä maailmassa: kun robotit osaavat ikään kuin lukea ihmistä ja tämän tunnetiloja, ne osaavat toimia ihmisen kanssa preferoiduin ja sosiaalisesti hyväksytyin keinoin. (Breazeal, 2004, 185.) Autonomisen robotin tuleekin tunnistaa esimerkiksi se, milloin ihminen toimii väärin tai tarvitsee lisätietoja, mutta robotti ei saa olla luonteeltaan tungetteleva tai ärsyttävä, vaan sen on toimittava "lempeän" ihmisen lailla (Devin & Alami, 2016). Tämä vuorovaikutuksen tematiikka on linjassa myös Asimovin (1942) kahden ensimmäisen robotiikan säännön kanssa: robotin autonomian tasosta huolimatta sen tulee totella ihmistä, eikä se saa myöskään vahingoittaa tätä.

## 4 AUTONOMISEN JÄRJESTELMÄN OMINAISUUK- SIA

Useimmat robotit ovat aikojen saatossa toimineet siten, että ihminen on voinut hallita ja nähdä robotin toimintaa etänä esimerkiksi videokuvana tai muiden erinäisten sensoreiden kautta saada tietoa olosuhteellisesta tietoisuudesta. Valvojana toimiva ihminen on voinut määrittää pitkälle automatisoiduille, puoli-autonomisille roboteille välietappeja, joita pitkin robotti on voinut saavuttaa määränpäänsä. Useimmat ihmisen ja robotin väliseen vuorovaikutukseen tarkoitettut ominaisuudet ja järjestelmät on kehitetty robotiikan eksperteille niin ikään robotiikan eksperttien toimesta, mikä ei ole nykyaikana aina haluttu lopputulema: järjestelmistä halutaan yhä enemmän arkipäiväisempiä, jolloin ne eivät ole käytettävissä pelkästään robotiikan ja autonomisten järjestelmien asiantuntijoilla, vaan myös niillä, jotka ovat substanssiasiantuntijoita. (Scholtz, 2003.) Esimerkiksi tämän tutkimuksen kontekstissa merenkävijät, kuten merikapteenit, ovat enimmäkseen asiantuntijoita kyseisessä asiassa eli merenkulussa mutteivät välttämättä robotiikan asiantuntijoita. Lisäksi kun järjestelmistä tulee yhä enemmän autonomisia ja ne kykenevät vielä itsenäisempään päätöksentekoon kuin aiemmin, on pohdittava tarkasti, millaista ihmisen ja robotin välinen yhteistyö on: Kuinka käyttöliittymä sidotaan yhteen autonomiseen robottiin ja sen sisältämän järjestelmän toimintoihin? Mitkä ovat ihmisen ja robotin väliset toiminnot ihmisen asettaman määränpään saavuttamiseksi, kuinka ne voivat muuttua järjestelmän ajon aikana? (Scholtz, 2003.) Autonomisen järjestelmän ominaisuudet joutuvatkin suurennuslasin alle heti halutessa automaatiota ihmisen suorittamiin tehtäviin, ja tällaisten järjestelmien ominaisuudet voitaneen jakaa karkeasti ensiksi neljään pääominaisuuteen, joista kerrotaan alaluvussa 4.1, sekä sen jälkeen alaluvussa 4.2 niistä järjestelmän osista, jotka voidaan automatisoida tiettyyn pisteeseen saakka.

## 4.1 Neljä pääominaisuutta

Betty ja muut (2009) ovat yhteensummaavassa katsauksessa luoneet neliosaisen kategorisoinnin pääominaisuuksista, jotka autonomisen järjestelmän voidaan katsoa sisältävän. Ensimmäisenä yläkategoriana ovat **vaatimukset**. Autonomisilta järjestelmiltä vaaditaan uudenlaisia ominaisuuksia pelkästään automatisoituihin järjestelmiin nähden, joten tarvitaan moderneja ja innovatiivisia lähestymistapoja luoda esimerkiksi kieli, jolla järjestelmän toiminnot tapahtuvat. Lisäksi, mikäli vaatimuksissa on epävarmuutta, kuten järjestelmän päämäärien asettamisessa, kieleltä vaaditaan taipuvaisuutta eli dynaamisuutta ajonaikaisiin muutoksiin. Tällaista vaatimuksien uudelleenmäärittelyä voi tulla vastaan esimerkiksi olosuhteellista tietoisuutta hankittaessa eli tarkkailtaessa ympäristöä. (Betty ym., 2009.) Esimerkki tilanteesta voi tulla vastaan, kun laivan järjestelmän ilmoittama operaatio ei ole suoritettavissa saatavilla olevilla keinoilla tai koodillisilla syntakseilla, jolloin ihmisen on annettava jokin uudelleenmäärittely tai sallittava laivan toteuttaa jokin epävarmempi operaatio itsenäisesti.

Toinen ominaisuus, joka on ominainen autonomisille järjestelmille, on **mallintaminen**. Malleilla voidaan kuvantaa sitä maailmaa sekä sen sisältämiä laskennallisia ja parametrillisia määritelmiä, joilla järjestelmä toimii. Mitä tarkempia ja yksityiskohtaisempia järjestelmän sisältämät mallit ovat, sitä varmemmin voidaan tehdä päätelmiä todellisuudesta. Mallit siis kuvastavat parhaimmillaan maailman käyttäytymistä järjestelmän ympärillä, mutta kuitenkin on tehtävä yksinkertaistuksia: mikään malli ei voi olla täysin kattava kuvaus totaalisesta ympäröivän maailman toiminnasta, ja siksi mikään malli ei voi sisältää kaikkia mahdollisia parametreja ympäristöstä. (Betty ym., 2009.) Tästä syystä on tehtävä tarpeeksi kattavia ja detaljisia malleja mutta ei liian, sillä muuten ihmisen voi olla mahdotonta luoda tai ymmärtää sellaisia sekä järjestelmällä kulua liikaa aikaa ja tehoa laskennallisesti analysoida malleja. Malleille tarvitaan siis kevyitä, tarkkoja määritelmiä – aivan kuin kevyitä teknisiä toimenpiteitäkin ihmisen puuttumistarpeen vähentämiseksi ja autonomian toimimiseksi. (Betty ym., 2009.)

**Tekniikka** tai teknologia (engl. *engineering*) on siis kolmas ominaisuus autonomisella järjestelmällä. Niin malleille, parametreille, toiminnoille, vuorovaikutukselle kuin muullekin autonomisen järjestelmän osa-alueelle on oltava tapa tai syntaksi, jolla kyseinen toiminta-ala toteutetaan. Mallien pohjalta voidaan muodostaa niin kutsuttuja toimintasyklejä ja selkeästi havaittavia toimintaoperaatioita, joiden läpinäkyvyydellä ja yksinkertaistettavuudella voidaan luoda selkeästi havaittavia ja hallinnoitavia toimintoja, jotka on arvioitava, minkälaisella teknologisella ratkaisulla kukin toteutetaan. (Betty ym., 2009.) Tekniikan on oltava autonomisissa järjestelmissä myös taipuisa siinä mielessä, että sen on mukauduttava aina kyseiseen, olosuhteelliseen tilanteeseen (Betty ym., 2009). Tekniikan tasosta riippuen voidaan päättää, mitä voidaan lopulta automatisoida ja mitkä seikat ovat sellaisia, joita vähiten voi itsenäistää.

Tästä syystä neljäntenä pääominaisuutena autonomisessa järjestelmässä ovat **vakuukset** tai **vakuutukset** (engl. *assurances*): on oltava olemassa varmuus siitä, että mallit ja parametrit ovat toimivia myös alati muuttuvissa ympäristöis-



sä. Ympäristö muuttuu merelläkin usein, eikä usein kuljettu merireitti aina ole samanlainen, mihin järjestelmä voi varautua tuntemalla ja tiedostamalla todennäköisyyksiä, joilla tietyt tapahtumat tapahtuvat. Todennäköisyyksien avulla voidaan myös sanoa, millä varmuudella jokin omaksuttu malli toimii toisenlaisessa, muuttuvassa ympäristössä, ja tässäkin avainasemassa sijaitsevat kepeät algoritmit sekä teknologia. (Betty ym., 2009.)

## 4.2 Automatisoitavien toimintojen ominaisuuksia järjestelmissä

Parasuraman, Sheridan ja Wickens (2000, 288) luovat artikkelissaan kategorisoinnin neljälle eri automatisoitavissa, tai autonomisoitavissa, olevalle toiminnalle: informaation hankkimisen, analysoinnin, päätöksenteon ja toiminnan. Vaikka nämä neljä ominaisuutta ovat tärkeitä ja olennaisia osia myös autonomisissa järjestelmissä, niitä ei voida kaikkia välttämättä omaksua täysin autonomisen rahtilaivan järjestelmien automaattisuutta luodessa. He painottavatkin, että järjestelmät voivat sisältää kaikkia näitä automaation aloja mutta eri tasoilla. Siksi tämänkaltaisen projektin, kuten autonominen rahtilaiva, automaattisten tai autonomistenkin järjestelmien ominaisuuksia on pohdittava äärimmäisen tarkkaan: pienikin liiallinen salliminen automaattisuudeksi voi johtaa kohtalokkaihin seurauksiin, kuten pahimmillaan rahtilaivan uppoamiseen sekä ympäristön vahingoittumiseen.

Ensimmäisenä automatisoitavissa olevana ominaisuutena on **informaation hankinta** (engl. *information acquisition*). Tietoa voidaan hankkia varsinkin sensorien avulla automatisoidusti, kuten kameroin visuaalisesti tai karttajärjestelmin sijaintiin pohjaten. Informaation hankinta voitaneen siis heijastaa ainakin jossain määrin ihmisen aisteilla havaittavin keinoin: tiedonhankinta on ensimmäinen kerros järjestelmän automatisoinnissa, ja aivan kuten ihminen ensimmäistä kertaa alkaa havaita ja hahmottaa asioita, automaationkin ensimmäisiä havaitsemiskeinoja on näkeminen. (Parasuraman ym., 2000, 288.) Samalla tavalla kuin vaikkapa laivan kapteeni näkee silmillään tai kartasta huomiotavia seikkoja, järjestelmä voi kerätä tietoa kameroin tai tulkiten karttaa sillä tavalla, jolla se on toteutettu. Parasuraman ja muut (2000, 288) noteeraavat sen tosiseikan, että tietoa kerätessä sitä on myös kyettävä suodattamaan oikeellisesti: ei voida kerätä kaikkea mahdollista dataa ympäristöstä, sillä silloin sen analysointi järjestelmällisesti ei ole mahdollista, vaan on osattava valikoida juuri se oikea, haluttu informaatio, joka tuodaan operoijan tietoisuuteen.

Toinen osa-alue, joka voidaan automatisoida tiettyyn pisteeseen saakka, on **tiedon analysointi**, mutta siihen liittyy ehkäpä eniten kysymyksiä ja epävarmuutta. Tieto kasvaa ja varmistuu luonnollisesti robotin kerätessä sitä yhtä enemmän sekä toimintaiän mukaan, mikäli samantyyppisten havaintojen lukumäärä robotin analysoimassa datassa kasvaa: kuten ihmisenkin aivot, robotti on oppiva, eikä se voi tietää kaikkea välttämättä etukäteen (tosin robotin järjestelmään voidaan syöttää sopiva data, mikäli se on toisessa robotissa havaittu toimivaksi). Jo 1989 Compton ja Jansen pohtivat tiedon analysointia: Niin kuin ihminenkin ei voi induktiivisesti yleistää kaikkea kokemaansa tai havain-

noimaansa heti ensikättelyssä, robottikaan ei voi tehdä hätiköityjä johtopäätöksiä ja yleistyksiä suoralta kädeltä. Kun jokin autonominen robotti kerää ja kerryttää tietoaan, on erittäin tärkeää muistaa, että sitä ei yleistetä tai yritetä ymmärtää täysin heti ensimmäisellä käsittelykerralla. Ihmisen on pystyttävä antamaan robotille valmiudet siihen, että se esimerkiksi analysoi tiedon vastaavuutta aiempaan dataan sekä kykenee tunnistamaan kontekstin, jossa tieto on saavutettu. (Compton & Jansen, 1989, 10.) Samankaltainen tieto voi olla merkitykseltään hyvinkin erilaista riippuen siitä ympäristöstä, jossa se on hankittu. Data on myös suunniteltava siten, että pohja, jolle se on alunperin suunniteltu toimivaksi, voi vaihdella ja muuttaa muotoaan ajan saatossa, jotta robotti voi analysoida yhä uudelleen aiempaakin dataa. (Compton & Jansen, 1989, 10.)

Tiedon analysointi on olennainen osa halutun päämäärän tavoittamiseksi, jotta autonomiset robotit voivat valita oikean toimintatavan oikealle tavoitteelle oikealla tavalla oikein parametrein. Robotti voi oppia muun muassa ympäristöään 3D-mallintamalla sen sekä tunnistamalla erilaisia objekteja esimerkiksi analysoiden relaatioita ympäristössä olevista objekteista mm. syvyyden ja position suhteen, onko robotti niiden yllä tai vaikkapa sisällä. Jos esimerkiksi toisen objektin uloisimmat koordinaatit ovat kaikki pienempiä kuin toisen objektin, objektin voidaan täten laskea olevan toisen sisällä. (Tenorth & Beetz, 2009.) Lisäksi data voidaan järjestää toiminnallisuuksien mukaan kategorisoiden, jotta robotti voi helposti päästä niihin käsiksi. Järjestelmällisen datan avulla robotti ottaa tarvitsemansa parametrit tarvitsemaansa suoritukseen sekä toimii niiden perusteella, sekä sen kautta myös ihmisen on helpompi omaksua robotin toimintaperiaatteet ja tarvittaessa muokata niitä ymmärtääkseen järjestelmän kokonaisu toimintaa automaattisuuden tai autonomisuuden taustalla. (Tenorth & Beetz, 2009.)

Kolmas ala, joka voidaan katsoa olevan automatisoitavissa ja autonomisoitavissa, on **päätöksenteko**. Päätöksentekoa voidaan helpottaa ihmisen toiminnan kanssa tehden järjestelmä antaen sille avaimet omiin päätelmiin tai ehdotelmiin. Eksperteille suunnatut järjestelmät, kuten kyseinen autonominen tai automaattinen laivajärjestelmä, on suunniteltu ehdollisella logiikalla: esimerkiksi tuotannossa ja oikeassa ympäristössä tapahtuvilla toiminnoilla toisin kuin itse suunnittelu- tai tekovaiheessa tapahtuvilla skenaarioilla, kuten sääolosuhteiden vaikutus navigaatioreitin valintaan. (Parasuraman ym., 2000; 288, 289.) Usein voikin olla haastavaa simuloida etukäteen päätöksentekoon vaikuttavia seikkoja virtuaalisessa tai rajoitetussa ympäristössä. Päätöksenteko on kuitenkin usein mielletty yhdeksi autonomisen ajoneuvon agentinkin pääpiirteeksi, jossa muun muassa mallinnetaan ympäristöä muodostaen samalla symbolisia merkityksiä sekä suorittaen loogista prosessointia niiden avulla. Päätöksenteko kytkeytyy myös todistettavaan suoritus- ja oppimiskykyyn: vaikka oppiminen aina sisältää epävarmuutta, se on pääroolissa kaikissa autonomisissa järjestelmissä ja aina taustalla kehittyneen päätöksenteon kanssa, mikä on osaltaan paradoksi. (Veres, Molnar, Lincoln & Morice; 2011, 185.)

Viimeisimpänä kohtana Parasuramanin ja muiden (2008) neliosaisessa mallissa automatisoitavan järjestelmän ominaisuuksista on **toiminnan automatisointi**. Kun tietoa on kerätty tarpeeksi ja oikeellisella tavalla sekä päätös huolella analysoitu ja tehty siitä toimintokohteesta, joka tullaan toteuttamaan, voi-

daan lopulta puhua toiminnallisuuksien automatisoinnista tai jopa autonomisoinnista. Automatisointi tässä mielessä tarkoittaa ihmiskäden tai -äänien korvaamista mekaanisella tai teknisellä metodilla, joka osittain tai jopa kokonaan korvaa ihmisen lopullisia toimia yksittäisessä toiminnossa, ja tämä voi tapahtua vaikkapa siten, että järjestelmä kuitenkin vahvistaa ihmiseltä, tehdäänkö sen päättelemä toiminto, ja valvoja voi painaa yhtä painiketta hyväksyäkseen sen. (Parasuraman ym., 2000, 289.) Mitä kompleksisemmiksi järjestelmien automatisointi etenee, sitä enemmän on järjestelmän tehtävä itsenäisiä päätöksiä, eikä se voi aina kysyä valvojalta tai operoijalta, mitä tehdään, vaan on osattava erottaa ne tilanteet, joissa järjestelmä kysyy vahvistusta ja joissa se toimii täysin itsenäisesti. Jos järjestelmä kyselisi valvojalta vahvistuksen kaikkiin suoritettaviin operaatioihin monimutkaisissa ja seikkaperäisissä toimintaympäristöissä, kuten laivan merellä seilattaessa, operoijan, kuten merikapteenin, kognitiivinen kuormitus olisi varmasti ylittämättömän. Sitä varten tämä viimeinen kohta eli toimintojen automatisointi vaatii aina kyseisen projektin kohdalla kriittistä analysointia, mitä toimintoja kapteenin on yhä aina tai millä toimintavarmuudella vahvistettava. Lisäksi toiminnan automatisointiin tarvitaan lopulta agenteja, jotka toimivat ihmisen toimintaa matkien, (Parasuraman ym., 2000, 289), mistä on kerrottu enemmän luvussa 6.

## 5 AUTOMAATION ERI TASOT SEKÄ TÄYDELLINEN AUTONOMISUUS

Automaation ja autonomian välillä on eroja. Automaatiolla tarkoitetaan usein automatisoitua järjestelmää tai robottia, joka kykenee toimimaan itsenäisesti sekä suorittamaan ennalta määrättyjä operaatioita usein teknisistä lähtökohdista. Toisaalta autonomiallakin voidaan tarkoittaa korkeimman tason automaatiota, ja termit voivatkin hämärtyä toisiinsa. Siksi tämän tutkimuksen kannalta alaluvussa 5.1 selkiytetään näiden kahden termin välistä merkityseroa, minkä jälkeen alaluvussa 5.2 punnitaan mieltämystä täydellisestä autonomiasta, jossa ihmisen ei tarvitsisi tehdä järjestelmän kanssa oikeastaan juuri lainkaan yhteistyötä.

### 5.1 Automaation eri tasot miehittämättömissä järjestelmissä

On usein tärkeää osata määrittää, kuinka automaattiseksi tai jopa autonomiseksi toteutettavat järjestelmät halutaan. Joskus järjestelmältä saatetaan vaatia vain minimaalista automaattisuutta, jolloin järjestelmän ohjaus ja hallinta jää pääosin sitä valvovan roolin, usein ihmisen, käsiin. Joskus taas järjestelmästä halutaan niin automaattinen kuin vain mahdollista, jolloin ihmisen ei tarvitsisi puuttua järjestelmän suorittamiin toimintoihin kuin hätätapauksessa. Tässä voidaan päästä lähelle autonomian käsitettä: järjestelmä on lähestulkoon itsenäinen yksikkönsä, joka suoriutuu sille määritetyistä tehtävistä omalla päättelyllään – joskus voiden jopa määrittää omat päätavoitteensa tai välitavoitteensa.

Huang, Pavek, Novak, Albus ja Messin (2005, 64) ovat määritelleet tämän tutkimuksen kannalta tärkeimmät viisi eri automaation tasoa miehittämättömille järjestelmille. Tasot määrittelevät, minkälaisen osuuden ihmisen työpanoksesta järjestelmän osat korvaavat sekä mitkä sellaisia automatisoituja ominaisuuksia voivat olla. Myös autoalan standardisointiin perehtynyt järjestö SAE International (*Society of Automotive Engineers*, 2018) on luonut oman määritelmänsä kuudesta tai oikeastaan viidestä automaation tasosta, joista 0. on ihmisen täysin itsensä ohjailema. Nämä kuusi tasoa on luotu lähtökohtaisesti maa-

ajoneuvoille, mutta niitä voitaneen vertailukohtana käyttää tälläkin maaperällä eli merellä, koska ne koskevat erityisesti fyysisiä kontroleja, kun taas Huangin ja muiden (2005 64) tasot on tässä kuvattu enemmänkin olosuhteelliseen tietoisuuteen ja informaation vaihtoon liittyen. Seuraavassa käydään läpi Huangin ja muiden (2005, 64) määrittämät tasot sekä vertaillaan niitä SAE:n (2018) kategorisoiimiin tasoihin.

Ensimmäisellä tasolla, **etäohjauksessa**, ihminen tekee käytännössä kaiken. Ihminen toimii järjestelmän operoijana, joka aistii ympäristössä tapahtuvat muutokset, tarkkailee kokonaistoimintaa ja yksittäisiä operaatioita, sekä analysoi ja suorittaa toimintoja. Etäohjaustasolla järjestelmä saattaa suorittaa yksinkertaisia tai äärimmäisen rajattuja toimintoja sekä kommunikoida ihmisen suuntaan, mutta muuten kaikki on operoijan eli ihmisen hallinnassa. SAE:n (2018) määritelmän mukaan tällä ensimmäisellä tasolla ajoneuvo voi omatoimisesti esimerkiksi jarruttaa, kiihdyttää tai kääntyä hätätilanteissa. Tämä etäohjauksen taso ei ehkä tule kyseen autonomisen rahtilaivan suunnittelussa, sillä laivalta toivotaan autonomia operaatioita.

Toinen taso, **teleoperointi**, sallii miehittämättömälle järjestelmälle jo enemmän itsenäistä toimintaa kuin ensimmäinen taso. Teleoperointi mahdollistaa järjestelmälle muiden muassa kyvyn aistia ympäristöään erinäisten antureiden, kuten kameran tai lämpömittarin, kautta. Keräämäänsä dataa se voi nyt tällä toisella tasolla raportoida valvojalleen ja tunnistaa omia tilojansa. Ihminen tekee saamallaan datalla kuitenkin samat päätökset tavoitteista ja toiminnasta kuin ensimmäiselläkin tasolla. Toinen taso on ikään kuin hieman luottavampi versio etäohjauksesta, joka ei tuota valvojalleen juuri itse järjestelmän keräämää tietoa. SAE:n (2018) mukaan puolestaan tällä tasolla fyysiset kontrollit voivat olla jo ajoneuvon hallinnassa: kapteeni voisi irroittaa kätensä ja jalkansa ohjaukskontroleilta laivan itsensä ohjailemana, mutta ihmisen tulee olla valppaana ottamaan ohjat omiin käsiinsä hetkenä minä hyvänsä. Tämä toinen taso on jo lähempänä autonomisen laivan toimintaperiaatteita, sillä voidaan tehdä oletta- ma, että laivan järjestelmällä tulee olla erilaisia monitorointikeinoja ympäristöönsä, jotta vaikkapa merikapteeni voi tehdä ympäristöllistä tarkkaavaisuutta vaativia päätöksiä.

Kolmatta tasoa voitaisiin kuvailla **ihmisen ohjailemaksi**. Tällä automaation tasolla miehittämätön järjestelmä voi jo pienissä määrin suunnitella omaa toimintaansa sekä tehdä karkeita johtopäätöksiä ympäristöstään ja toiminnostaan. Täten järjestelmä pystyy havainnoimaan ja raportoimaan toimintaansa, joka on hahmottamisen kanssa jaettu kanssa ihmisvalvojalle. SAE:n (2018) kolmatta tasoa katsoen laiva voisi liikehtiä omin avuin ilman operoijan valvontaa silmiä räpäyttämättä, mutta hädän hetkellä esimerkiksi vaihtuvan sääolosuhteen myötä se voisi ottaa yhteyden valvojaansa, jonka tulee hallita tilanne vastedes. Lisäksi tämä taso on SAE:n (2018) määritelmässä ensimmäinen taso, jolla todellisesti viitataan tilaan, jossa ei tarvitse enää aina manuaalisesti ohjata tai ajaa, eli autonomia terminä nousee jo puheenaiheeksi. Autonomisen rahtilaivan ollessa kyseessä tämä taso on luultavasti lähellä sellaista realistisesti kuviteltavissa olevaa lähiaikoina.

**Ihmisen avustama** taso on neljäs automaation taso miehittämättömissä järjestelmissä. Tällä tasolla järjestelmä on kykeneväinen tarkkailemaan ympäris-

töään ja toimintojaan sekä raportoimaan niistä valvojalle. Järjestelmän analysointi, suunnittelu ja päätöksenteko ovat pääosin jaettu ihmisen kanssa, eli järjestelmä on jo suhteellisen itsenäinen, pitkälle automatisoitu. Järjestelmä on siis mahdollistettu suorittamaan suurimman osan hahmottamisesta ja toimimisesta itsenäisesti. Kolmanteen tasoon nähden tämä on hieman kunnianhimoisempikin tavoite: laivan järjestelmät tulisi automatisoida niin pitkälle, että ihmisen, kuten merikapteenin, työtehtävät olisivat jo huomattavasti vähentyneet. SAE:n (2018) määritelmässä jo tällä neljännellä tasolla viitataan jossain määrin autonomiaan: ajoneuvo kykenee omatoimisesti ja turvallisesti liikehtimään koko reitin ajan huomioiden ympäristön ja olosuhteet. Jopa ohjausvälineet, kuten ratti ja polkimet, voivat olla jo tarpeettomat eikä ihmisen tarvitse enää ajaa sitä, mutta jos niitä varmuuden vuoksi tarvitaan, voidaan automaation tasoksi määrittää vaikkapa kolmas. Kolmantena huomiona standardissa on se, että tämä neljäs automaation taso on toimiva ainoastaan sille suoduissa tarkkoissa olosuhteissa ja ympäristössä, eikä ajoneuvo voi omaksua muunlaisten ympäristöjen toiminnallisuuksia, jos ympäristö vaihtuu usein. (SAE, 2018.) Näin voi käydä muun muassa tyynen kelin vaihtuessa erittäin myrskyisään.

Viimeinen, ylin taso on niin Huangin ja muiden (2006, 64) kuin SAE:nkin (2018) määritelmässä **autonomia**. Tätä tasoa voitaisiin kutsua täydellisen itsenäisyyden tasoksi: järjestelmälle on toteutettu keinot ja myönnetty oikeudet toimia niin automaattisesti ja itsenäisesti kuin suinkin mahdollista. Järjestelmä on parhaansa mukaan itsemääräämisoikeuden omaava entiteetti, johon ihmisen tarvitsisi puuttua minimaalisesti. Miehitämätön järjestelmä, kuten vaikkapa rahtilaiva, kykenisi hoitamaan sille määrättyt tehtäväkokonaisuudet ja niiden osa-alueet omalla päätöksenteollaan kaikissa mahdollisissa olosuhteissa (SAE, 2018) täydellisellä ihmismäisellä päätöksenteolla, mutta kuitenkin se neuvottelisi ihmisen kanssa tilanteen tullen. Järjestelmän tulee siis tällä autonomian tasolla myös kyetä yhteistyöhön ihmisen kanssa, eikä se saakaan alkaa toimia niin kutsuttuna despoottina torjuen kaikki ihmisen antamat ehdotukset tai ukaasit esimerkiksi suunnanmuutoksista. Järjestelmän on siksi kyettävä vuorovaikutukseen ja neuvottelukykyyyn tällä ja kaikilla muillakin tasoilla. Autonomisen rahtilaivan tuskin edes odotetaan saavuttavan tätä tasoa lähivuosina, sillä todellisen autonomian kehittyminen niin itsenäiseen pisteeseen saakka vaatinee lisätutkimusta vielä useiden vuosien tai jopa vuosikymmenten ajan teknologian ja ihmismielen mallintamisen saralla. Lopulta voidaankin kritisoida itse ajatusta *autonomisesta rahtilaivasta*: ehkä on turvallisempaa ennakoida, että laivan automaation taso voitaisiin nykytutkimuksen valossa saavuttaa realistisesti tasojen kaksi ja neljä välille, missä vähintään tasosta kolme alkaen voisimme puhua autonomiasta.

Myös Veres ja muut (2011, 170) ovat samoilla linjoilla automaation tasojen merkityksistä. Tosin he rajaavat tasot vain kolmeen, mutta niiden pääperiaatteet ovat samankaltaisia kuin Huangin ja muiden (2005, 64) laajentaen autonomisuuden ajatusta koskemaan nimenomaan liikehdintää ja kulkureittejä. Ensimmäinen tasoista määritetään siten, että ajoneuvo ottaa vastaan ihmisen määrittämän ajoreitin ja osaa navigoida sitä myöten määrittäen omat välietappinsa matkalla. Ajoneuvo tietää myös kartan ja sijaintinsa sillä. Toisella tasolla valvojan tarvitsee määrittellä kartalla yleiset tavoitteet, jotka kulkuneuvo tulkitsee ja

pilkkoo tarkempiin osiin reitillä välttämällä esimerkiksi vastaantulevia kulkuneuvoja lähtien ja saapuen itsenäisesti. Tätä toista tasoa tarvitaan varsinkin autonomisissa rahtilaivoissa. Viimeisenä eli kolmantena tasona ihmisen tehtävänä on ainoastaan määrittää tehtävien yleiset tavoitteet abstraktilla tasolla ja ajoneuvolla on äärimmäinen päätöksenteon kyky liittyen kyseiseen kokonaistavoitteeseen. Lisäksi tällä kolmannella tasolla on tietysti kahden aiemman tason ominaisuudet. (Veres ym., 2011, 170.)

## 5.2 Täydellinen autonomisuus?

Automatisoidusta robotista tuskin ikinä halutaan täysin autonomista, tai se ei välttämättä tottele enää ihmistä eli sen omistajaa. Tällaisesta skenaariosta on jo vuosikymmenten ajan luotu mielikuvaa, että robotit nousisivat ikään kuin barrikadeille ja ottaisivat ihmiskunnan haltuunsa, jolloin ihmisen antamat komennot eivät enää pätisikään. Robotin tuleekin ajon aikana mukautua ja toimia oikeellisesti mm.

- ihmisen asettamiin uusiin määrityksiin
- järjestelmän virhe-/häiriötilanteisiin tai tunkeutumisyrittäisiin
- toimintaympäristön muutoksiin sekä
- resurssien vaihtelevuuteen (esim. Cheng ym., 2009, 11.)

Ihminen on lopulta se, joka on toteuttanut robotin toimimaan haluamallaan tavalla, joten on vaikeaa uskoa, että tällainen kauhuskenaario tapahtuisi, jossa robotit toimisivat täysin niille toteuttamattomalla tavalla. Toisin sanottuna niiden tulisi kyetä toimimaan niille syötetyn koodillisen tai syntaktisen määrittelyn ulkopuolelta, mikä tarkoittaisi sellaista teknistä määrittystä, että järjestelmä kykenisi tekemään itsenäisiä päätöksiä missä tahansa asiassa. Tällainen vaatisi täysin kokonaisvaltaista koodillista määrittystä, jonka toteutus nykytekniikan valossa lienee haastavaa.

Kuitenkin joskus saattaa käydä niin, että ihminen onkin se erehtyväisempi osapuoli, jolloin järjestelmän on hyvä pitää pintansa ja olla ottamatta ihmisen vaatimia muutoksia vastaan. Autonomisen robotin olisikin hyvä osata havaita, milloin ihmisellä tulee virheitä toiminnassaan ja haluamissaan muutoksissa autonomisen järjestelmän määrittämiin tavoitteisiin. Ihminen on niin kutsutuissa rutiinitoimenpiteissä usein robottia herpaantuvampi ja joskus erehtyväisempikin, joten robotti voi selkeästi huomauttaa, mikäli valvojan päätöksissä on puutteita. (Dekker, 2017 & Reason, 1990.) Jos ihminen haluaa vahingossa esimerkiksi rahtilaivalta muuttaa sen suuntaa mutta järjestelmä huomaa muutoksen olevan väärä ja toisen laivan olevan vaihdettavalla kulkureitillä, miehittämättömän autonomisen järjestelmän tulisi osata ilmaista ihmiselle, miksi se ei suostu tekemään tämän toivomaa toimintoa. Tällaisetkin yksityiskohtaiset toimintatavat ovat alkujaan ihmisen määrittämiä, eli voitaisiinkin sanoa, että robotti on juuri sellainen, kuin miksi ihminen on sen luonut.

Vaikeimmat algoritmit ja määritykset liittyen autonomisuuteen voivat olla lähes mahdottomia toteuttaa koodin kannalta, eikä haastavimpia autonomisiin, miehittämättömiin ajoneuvoihin liittyviä algoritmeja olekaan välttämättä saatu edes toteutettua. Esimerkiksi useat liikenteeseen liittyvät algoritmit ovat koskeneet esineiden välttelyä tai ihmisen kontrollin irtaannuttamista autonomisen robotin toiminnasta, ja tämä kuvastaa sitä vaillinaisuutta, joka meillä vielä on täydellisen autonomisuuden toteutuksessa. (Goerzen, Kong & Mettler; 2010, 94.) Mikä meissä ihmisissä esimerkiksi laskee sen, miten kävelemme portaita? Vai laskeeko sitä mikään? Olemme oppineet kävelemään portaita pienestä pitäen, mutta onko oppiminen ollut tietoinen oppimisprosessi, vai onko se syntynyt vain alitajuisesti? Voiko roboti puolestaan oppia kävelemään portaita itsenäisesti, jos sille annetaan ainoastaan edellytykset eli pääperiaatteet siihen kokonaistavoitteen kannalta: vaikkapa kulku alakerrasta yläkertaan?

Kuitenkin tällainen pienehkö kokonaisuus portaiden kulussa on vain yksityiskohtainen, rajatussa toimintaympäristössä tapahtuva prosessi. Usein tehtävät, jotka halutaan saavuttaa, eivät ole yksinkertaista tavoitteenmäärittämistä ja reittiohjausta, vaan vaaditaan kokoaikaista ympäristöntarkkailua, jota autonominen roboti peilaa koko ajan järjestelmästäan löytyviin vastaavuuksiin. Kaikessa tässä ihminen on viime kädessä valvovana osapuolena, joka voi keskeyttää tai antaa lisätietoja missä vaiheessa tahansa prosessin suorittamista. (Goerzen ym., 2010, 95.) Täydellinen ihmismielen mallintaminen täysautonomiseksi robotiksi voikin olla mahdotonta tai ei välttämättä ainakaan lähivuosien aikana tapahtuva muutos. Olemmeko edes valmiita luottamaan kaiken toiminnan robotin käsiin? Niin koodillisesti kuin periaatteellisestikaan.

Jos emme ymmärrä aina edes itseämme tai toisiamme, miten voimme olettaa, että kykenemme luomaan robotin, joka toimii täysin itsenäisesti sekä selkeästi ymmärtää ihmistä ja ihminen sitä? Autonominen robottikaan ei ole välttämättä lopulta kuin osa ihmisen tekemiä, todellisesta maailmasta irtaannutettuja asiakokonaisuuksia yhdessä. Miten roboti voi tietää tai ymmärtää enemmän kuin olemme kertoneet sille tai antaneet sille avaimia tiedonkeruuseen? Tämä dilemma kiteytyy ajattelutapaan *kiinalaisesta huoneesta* (engl. *Chinese room*; Searle, 1980 & 1990), jossa tietokoneelle on annettu tietty määrä symboleja, joiden mukaan toimia. Kiinalaisessa huoneessa tietokoneen tulee pystyä ymmärtämään kiinan kieltä täydellisesti pelkästään sille annettujen merkkien pohjalta ja luomaan niistä kokonaisia merkityksiä toisiinsa, mutta miten se onnistuu siinä, jos sille ei ole annettu edellytyksiä siihen? Toisin sanoen, mikäli ihminen keskustelisi tällaisen koneen kanssa, ihminen ei pystyisi välttämättä erottamaan, keskusteleeko toisen, täydellistä kiinaa osaavan henkilön vai oppineen tietokoneen kanssa. (Searle, 1980 & 1990.) Loppupeleissä olemme vain antaneet robotille enemmän tai vähemmän itsenäisesti suoritettavia, joko hyvin tai heikosti määritettyjä toimintoja, eivätkä syötetyt toimintamallit ole vielä tänäkään päivänä viimeisteltyjä kokonaisuuksia esimerkiksi virheettömän kiinan kielen taitamisessa. Samanlaisia ongelmallisuuksia voi tulla vastaan ihmisenkin kanssa työskennellessä, mikäli esimerkiksi laivassa työskentelevälle apulais- tai aloittelevalle kapteenille ei ole annettu tarpeeksi yksityiskohtaisia määrityksiä, kuinka laivaa ohjataan tai kuinka erikoistilanteissa toimitaan.



Projektien kokonaistavoitteet ja välietapit ovatkin usein joko hyvin määritettyjä (engl. *well-defined*) tai huonosti määritettyjä (engl. *ill-defined*; Simon, 1973). Hyvin määritetty ongelma voi olla esimerkiksi annettuna muodossa "Laske yhteen  $1 + 1$ ", jolle löytyy oma komentonsa ohjelmointikielestä kuin ohjelmointikielestä, ja se voidaan toteuttaa pilkkomatta suhteettoman moneen välitavoitteeseen. Huonosti määriteltävissä oleva kysymys puolestaan voi olla esimerkiksi laivalle annettuna muodossa "Aja Oslon satamasta New Yorkin satamaan", joka sisältää itsessään jo lukuisia muitakin välitavoitteita, jotka on osaltaan määriteltävä tarkasti, kuten satamasta lähteminen, objektien välttäminen avomerellä sekä satamaantulo. Simonin (1973, 183) määrittelemien kuuden kohdan perusteella voidaan löyhästi tarkastella, onko kyseessä hyvin vai huonosti määritetty tai määriteltävissä oleva ongelma:

1. On olemassa selkeät kriteerit minkä tahansa ratkaisun testaamiselle sekä mekanisoitavissa oleva prosessi kriteerin käytäntöönpanolle.
2. On olemassa vähintään yksi ongelma-avaruus (engl. *problem space*), jossa voidaan ilmaista alkuperäinen ongelmatila (engl. *problem state*), päämäärätila (engl. *goal state*) sekä kaikki muut saavutettavissa tai vähintään kuviteltavissa olevat tilat, kun pyritään kohti ratkaisua.
3. Saavutettavissa olevat tilamuutokset voidaan esittää ongelma-avaruudessa, eli tietystä tilasta voidaan päästä suoraan tiettyyn tilaan tietyllä operaatiolla.
4. Mikä tahansa tieto, jonka ongelmanratkaisija voi saada pulmasta, voidaan esittää joko yhtenä tai useampana ongelma-avaruutena.
5. Jos kyseessä oleva ongelma vaatii määrittelytoimintaympäristön ulkopuolella toimimista, tilamuutosten ja vaikutusten määritelmiä voidaan peilata niin kutsuttuihin luonnonlakeihin tai yleisesti hyväksytyihin toimintaperiaatteisiin.
6. Kaikki oletetut toiminnallisuudet ovat toteutettavissa käsillä olevien teknologioiden avulla sekä tarvittava informaatio on olemassa tai vähäisellä vaivalla hankittavissa. (Simon, 1973, 183.)

Tällaisissa huonosti määriteltävissä olevissa projektikokonaisuuksissa, kuten autonomisen rahtilaivan toteuttaminen ja kulkureitillä navigoiminen, autonomian riskit voivat olla suuremmat kuin sen hyödyt, mikäli kaikkia projektin tavoitteita ja niitä saavuttavia toimintoja ei ole kyetty määrittämään äärimmäisen hyvin. Jokainen automatisoitavaksi haluttu ominaisuus on valittava äärimmäisen tarkasti sekä myös toteutettava äärimmäisellä huolella, eikä kaikkea välttämättä halutakaan autonomiseksi. Koska erilaisissa projekteissa vaihtelevat tavoitteet ja metodit ovat ratkaisevat tekijät onnistumisessa, Turner ja Cochrane (1993) määrittävät projektit neljään kategoriaan koskien hyvin ja huonosti määritettyjä kokonaisuuksia, ja autonominen rahtilaiva koskee näistä kategorioista pääasiassa neljättä: päämäärät ja tavoitteet ovat heikosti määritettyjä. Autonominen rahtilaivaliikenne voisi periaatteessa pudota näistä kategorioista toiseenkin, jossa päämäärät ovat hyvin spesifioituja mutta metodit eivät, sillä yleensä laivalla on selkeä päämäärä, mutta se voi muuttua ja saada välitavoitteita esimerkiksi määrittelemättömien sääolosuhteiden tai arvaamattoman muun vesi-

liikenteen toimesta. Tällöin autonomian tulee määrittää itse omat tavoitteensa ja ke, ja tällainen tavoitteenmäärittely voi olla tilastollista päättelyä tai tilanteellista huomioimista eikä lainkaan hyvin määriteltyä. Koska autonomia vaatii hyvin määritetyt tavoitteet ja toimenpiteet, joilla ne saavutetaan, on hankala kuvitella täysin automatisoitua laivaliikennettä arkipäivään ainakaan lähiaikoina. Kuten Cochrane ja Turnerkin (1993, 101) toteavat; kun tavoitteiden ja metodien määrittely on haastavaa, ei voida toteuttaa projekteja tavallisella tavalla. Ja mikä olisikaan tavallinen toteutustapa autonomiselle rahtilaivalle.

Voi olla siis lopulta mahdotonta luoda täydellistä autonomisuutta (Goerzen ym., 2010) ja päästä ylimmälle automaation tasolle eli autonomiaan (Huang ym., 2005). Tästä syystä on oletettavasti tärkeää päättää suunnitteluvaiheessa, mille automaation tasolle jokin järjestelmä on tarkoitettu toteutettavaksi, jos ylin eli viides taso ei ole käytännössä toteutettavissa. Tämän päämääränä olevan tason lähtökohtana on kyettävä tulkitsemaan ihmisen mieltä siihen pisteeseen saakka, jolla taso voidaan tyydyttyneesti sanoa olevan toteutettavissa taikka toteutettu. Ihmismielen mallintaminen täydellisesti ylimmän tason saavuttamiseksi onkin kokonaisuudessaan erittäin huonosti määriteltävissä oleva tehtävä, mutta ihmismieli itsessään voidaan jakaa useisiin lohkoihin, joita kyetään tarkastelemaan sekä itsenäisesti että yhteydessä toisiinsa. Täten voimme saada kokonaiskäsitystä siitä, miten ihmismieli toimii, ja voimme paremmin määrittää projekteillemme, minkälaiset tavoitteet ja toiminnot tyydyttävät vaatimuksemme. Näin kykenemme toteuttamaan yhä autonomisempia järjestelmiä toimien ihmismielen lailla, jotta ymmärrämme niiden toimintaa ja pystymme parempaan vuorovaikutukseen.

## 6 IHMISMIELLEN MALLINTAMINEN

Ihmisellä ja robotilla on kummallakin omat vahvuutensa ja heikkoutensa useiden tehtävien suorittamisessa. Onkin erittäin olennaista tiedostaa se lähtökohta, että ihmisen ja robotin toimintatavat ovat usein tyystin erilaiset: ihminen toimii tyypillisesti useamman eri aistinsa varassa yhdistäen niitä. Lisäksi ihminen osaa ottaa huomioon monia eri päätöksentekoon vaikuttavia seikkoja sekä improvisoida (Bradshaw, Dignum, Jonker & Sierhuis; 2012, 10). Robotti puolestaan kykenee usein toimimaan ainoastaan sille suodussa toimintaympäristössä ja niillä parametreilla, jotka sille on annettu (esim. Spong ym., 2004). Tietyt fyysiset ja psyykkiset aspektit, kuten herpaantuminen tai kestävyys, ovat ihmisellä usein robottiin nähden heikompia. Robotille taas on hankalampaa luoda kyky päätellä induktiivisesti, sillä asioiden ja merkityksien suhde toisiinsa sekä symbolien monikäsitteiset merkitykset (esim. Harnad, 1990) ovat haastavia toteuttaa dynaamisesti eli ympäristöstä riippumatta sen sijaan, että huomioon otettavat seikat on määrätty etukäteen (mm. Wegner, 1998, 349). (Bradshaw ym., 2012, 10.)

Tietyissä määrin ihminen kykenee siis yhdistelemään tiedonpalasia toisiinsa huomattavasti robotin syntaktista määrittelyä paremmin. Toisaalta esimerkiksi kognitiivisen kuormituksen haittojen vuoksi moni tehtävä voi vaarantua. Ilman robotin täydentävää osaamista esimerkiksi monen asian samanaikaisesta työskentelystä, keskeytymättömyydestä ja suuren tietomäärän hallinnoimisesta nopealla syötöllä (Kirsh, 2000) ihminen olisi tavoitteiden onnistumisen kanssa suhteellisen heikossa asemassa monessa automatisoidussa työtehtävässä. Siksi robotteja luodaan ihmiselle ikään kuin apukäsiksi täydentämään sellaisia ominaisuuksia, joita ihmisellä eli esimerkiksi valvojalla ei luonnostaan ole tai jotka vaativat vaikkapa aivotyöskentelyä ja kapasiteettia suuressa mittakaavassa: tässä robotti ei niin kutsutusti ylikuormitu yhtä helposti kuin ihminen. Mutta miten voisimme lopulta luoda toimivan vuorovaikutuksen ihmisen ja täydellisen autonomisen tai pitkälle automatisoidun robotin välille, elleimme ymmärrä ihmisen mieltä, mutta kuitenkin haluamme robotin toimivan ihmisen mielen lailla? Jotta robotti siis saadaan toimimaan ihmisen mielenliikkeiden kaltaisella tavalla ja mahdollisimman itsenäisesti eli autonomisesti, on osattava mallintaa ihmismieltä ja täydentää näillä tutkituilla malleilla robotin puutteellisia ihmismäisiä kykyjä.

Seuraavissa aliluvuissa on esitelty muutama malli tai keino, jonka mukaan ihmismieltä voidaan mallintaa tai simuloida. Ensimmäiseksi alaluvussa 6.1 tutustutaan ehkäpä yhteen tunnetuimpaan malliin, *ACT-R*, minkä jälkeen alaluvussa 6.2 esitellään tästä johdettua, täydennettyä mallia *ACT-R/E*. Tämän jälkeen alaluvussa 6.3 puntaroidaan ehkäpä modernimpia ja joissain tapauksissa tehokkaampiakin käyttötapauksia tekoälyn mallintamiseksi neuroverkkoihin pohjautuvilla konsteilla. Lopuksi alaluvussa 6.4 pohditaan sitä, mihin pisteeseen saakka ihmisen mieltä pitää tai ylipäättään kannattaa mallintaa. Kussakin ihmisen mielen mallintamista koskevassa alaluvussa on otettu huomioon niitä seikkoja, joita autonomisen rahtilaivan vaatimat prosessit ja skenaariot todennäköisesti koskisivat, kuten olosuhteellinen tietoisuuden saavuttamiseksi ja hyödyntämiseksi tarvittavat taidot. Alaluvuissa on myös tuotu ilmi niitä tutkimuksia, joissa mallintamiskeinoja on jo hyväksi havaitulla keinolla onnistuttu hyödyntämään: näissä soveltuvuus kyseisen pro gradu -tutkimuksen rahtilaivan toteuttamiseksi voisi olla harkitsemisen arvoinen seikka.

## 6.1 ACT-R

*ACT-R*-malli (engl. *Adaptive Control of Thought-Rational*) on laskennallinen kognitiivinen arkkitehtuuri ihmismielen havainnollistamiselle, joka konnektionistista ja symbolisista merkityksistä luo oletuksen sille, että jokainen ihmisen mielen korkean tason tapahtuma voidaan mallintaa erotelluin operaatioin. Eri-tyisesti mallia on tutkittu – tätäkin tutkimusta koskien – visuaalisten käyttöliittymien käytössä: visuaalinen tarkkaavaisuus näytöllä liikkeessa voidaan kuvata mallina, jonka *ACT-R* osaa prosessoida. (Anderson, Matessa & Lebiere; 1997.) Siksi mallin on katsottu olleen hyvä esimerkki sille, että ihmisen ja tietokoneen välisen vuorovaikutuksen operaatiot esimerkiksi käyttöliittymissä on mahdollista määritellä ja siten mukailla automatisoidusti. Sitä onkin usein käytetty simulaattorina sille, miten ihminen ajattelee, havainnoi ja käyttäytyy ympäristössään. (Anderson ym., 1997.)

Malli ei kuitenkaan välttämättä voi olla nykyaikaiseen, ihmisen ja tietokoneen tai robotin väliseen vuorovaikutukseen täysin holistisesti sopeutettavissa oleva. Moni asia on Andersonin ja muiden (1997) aikaisesta tutkimuksesta muuttunut ja monimutkaistunut robottien tullessa arkipäiväiseen elämäämme, ja käyttöliittymien merkitys on vaihtunutkin siinä mielessä, että moni vuorovaikutus voi tapahtua esimerkiksi äänikomennoin ilman silmin havaittavaa käyttöliittymää. Yksinään malli ei siis välttämättä suoraan pädekään ihmisen ja robotin välisen vuorovaikutuksen tai autonomisen rahtilaivan järjestelmien ja toimintojen suunnitteluun siten, että jokin järjestelmä tai robotti osaisi yhdistää kaikki kompleksiset merkitykset toisiinsa itsenäisesti sekä tehdä esimerkiksi itsenäisiä päätöksiä koskien päämääriä merellä navigoidessa – saati määrittää näitä tavoitteita ja välitavoitteita merkityksien pohjalta. Mutta sitä voitaneen kuitenkin tarkastella esimerkiksi yksittäisten käyttöliittymien tai niiden näkymien suhteen; miten autonomisen rahtilaivan järjestelmät kykenisivät havain-

noimaan ja tulkitsemaan vaikkapa ympäristöllisiä attribuutteja rajatuissa skenaarioissa.

Yksi esimerkki *ACT-R*-mallin hyödyntämisestä todellisuudessa onkin se, kuinka ympäristön ja mielen väliset prosessit voidaan mallintaa. Gaissmaier, Schooler ja Mata (2008, 288) päätyivät johtopäätökseen, jossa he toteavat rajattujen kognitiivisten toimintojen ympäristön tarkkailussa olevan muutettavissa *ACT-R*:n mukaisiksi proseduureiksi. Paino eritoten sanalla *rajatuissa*: he ymmärtävät ihmismielen rajatut kapasiteetit ja painottavat, että vähemmän on usein enemmän (Gaissmaier ym., 2008, 283) sekä liika tieto voi satuttaa (Gaissmaier ym., 2008, 280). Ehkäpä samanlaista mentaliteettia voitaisiin soveluttaa esimerkiksi siihen, miten kapteeni tulkitsee pala palalta etäältä sensoreiden avulla laivan ympäristöä ja miten myös laivan järjestelmä itse osaisi tehdä samankaltaisia päätöksiä olosuhteellista tietoisuutta vaativassa päätöksenteossa. Luultavasti siten esimerkiksi juuri *ACT-R*-mallin pätevyyttä autonomisen rahtilaivaliikenteen ja siihen kytkeytyvien prosessien kanssa voitaisiin testata pie-niin asiakokonaisuuksiin pilkottuina prosesseina.

Tämä Gaissmaierin ja muiden (2008) havainto ympäristön tarkkailussa voisi osoittautua tutkimisen arvoiseksi seikaksi autonomisen rahtilaivan toiminnallisuuksien suunnittelussa esimerkiksi yhdistettynä Fleetwoodin ja Byrnen (2006) löydöksiin ikoneista. He huomasivat, että ihminen pystyy tunnistamaan ikoneja käyttöliittymistä jo ilman, että niihin tarvitsee kiinnittää itsessään huomiota, eli ikään kuin niin kutsutusti sivusilmällä. Ihminen etsii usein vilkkaalta ja ahtaalta, useita elementtejä sisältävältä ruudulta juuri oikeaa tietoa, ja siksi ikonin sijainnin loogisuus usein auttaa siinä, että ikonia ei tarvitse havainnoida jokaisella kerralla uudelleen: ikonin tuttuus auttaa käyttämään sitä hyödyksi yhä uudelleen. Lisäksi ikonin löytäminen käyttöliittymästä on vaivattomampaa, mikäli ihmisellä on jo jokin ennako-olettama, miltä ikoni näyttäisi ulkoasultaan. (Fleetwood & Byrne, 2006.) Nämä ihmismielelle tyypillisen toiminnallisuudet ovat Fleetwoodin ja Byrnen (2006; 189, 190) mukaan mahdollista ilmaista *ACT-R*-mallina, jossa robotti osaa matkia ihmisen ikonin paikantamista. Ehkäpä mallia voitaisiin jossain määrin autonomisessa rahtilaivaliikenteessä soveltaa vaikkapa siihen, kuinka ihminen havaitsee sääatutkista ikonein ilmaistavia vaihtelevia sääolosuhteita, tai kuinka monitoreista voidaan nähdä ympäristössä tekijöitä, jotka eivät sinne esimerkiksi normaalisti kuulu. Näin autonomisen rahtilaivan järjestelmä olisi yhä autonomisempi tehdessään samanlaisia päätöksiä, kuin mitä ihminenkin tekisi ottaessaan huomioon päätöksentekoon vaikuttavia olosuhteellisia tekijöitä.

## 6.2 ACT-R/E

*ACT-R/E* (engl. *Adaptive Character of Thought-Rational/Embodied*) puolestaan on laajennettu ja ehkäpä modernimpi arkkitehtuuri *ACT-R*:sta siinä mielessä, että se lisää muun muassa kaksi uutta moduulia tälle mallintamiskeinolle: avaruudellisen hahmotuskyvyn eli kolmiulotteisuuden sekä erinäisten objektien ja asioiden muuttamiskyvyn. *ACT-R* on siksi osaltaan hieman puutteellinen, sillä sen

avulla on tyypillisesti mallinnettu ihmisen mieltä sangen rajallisessa ympäristössä, jossa ei ole otettu huomioon havainnointikyvyn ja toiminnan tai johtopäätösten välistä vuorovaikutusta. Tämän *ACT-R/E* taklaa paremmin tutkien nimenomaan havainnoinnin ja toiminnan välisiä syy-seuraussuhteita. (Trafton ym., 2013.)

Mallin laajentaessa perinteisen *ACT-R*:n ideaa sillä on useita hyötyjä. Suurella uskottavuudella malli kykenee löytämään ne elementit, joita ihminenkin etsii, eli vaikkapa merikapteenin suorittamat objektien etsinnöt ruudulta lienee mahdollista simuloida. Malli voi myös osoittaa ne kohdat, joissa ihminen on unohtanut tehdä jotakin, ja auttaa palaamaan niihin, mikäli mahdollista. Tällainen voisi olla vaikkapa merikapteenin epähuomio rutiininomaisissa tehtävissä, kuten jokin prosessi satamasta lähdössä tai sinne saapuessa. *ACT-R/E* myös helpottaa kommunikointia ja yhteisymmärrystä ihmisen ja robotin välillä siten, että ihmisen uskomukset ovat selkeämmin tulkittavissa. (Trafton ym., 2013.)

Trafton ja muut (2013, 34) kumminkin nostavat kritisoinnin arvoiseksi seikaksi sen, että malli ei ole mikään täydellisen optimaalisen toiminnan malli. Kognitiivinen arkkitehtuuri simuloituna vaatii laitteistoilta suhteellisen paljon tehokkuutta, eikä se olekaan optimoiduin nopeudessa tai laskennan tarkkuudessa antaen aina täsmällisen tarkat laskutulokset. Kuitenkin sitä voitaisiin käyttää kommunikointityökaluna ihmisen ja robotin välillä yhtenä parhaimmista keinoista. (Trafton ym., 2013.) Toinen soveltamiskeino voisi olla – alaluvussa 6.1:kin mainitulla tavalla – kommunikointi tulevista sään muutoksista sekä laivan itsensä arviot suunnanmuutoksistaan niihin nojaten, sillä sään muutokset tuskin ikinä tapahtuvat silmänräpäyksessä. Tällaiseen nopeaan tai täsmälliseen ihmisen mielen kaltaiseen toimintaan tarvitaan tehokkaampia ja täsmällisempiä mallinnuskeinoja, ja siksi esimerkiksi matemaattisen ihmismielen tai nopean loogisen päättelykyvyn mallintamiseksi tarvitaan erilaisia lähestymiskeinoja, mistä on kerrottu seuraavassa alaluvussa.

### 6.3 Neuroverkot

Yhdeksi moderneimmista ihmisen mielen mallinnuskeinoiksi sekä tekoälyn muodostamismahdollisuuksiksi on kehkeytynyt idea neuroverkoista (McCulloch & Pitts, 1943) objektien tunnistamiseksi ja havaitsemiseksi monessa visuaalisessa tehtävässä (esim. Zhang ym. 2018). Näille malleille ei tarvitse antaa avaimia jokaisen suoritettavan tehtävän ratkaisemiseksi, vaan mallit oppivat esimerkiksi havainneaineiston kautta, minkälaisia yhtäläisyyksiä asioilla on. Autonomisen rahtilaivan ja sen järjestelmien toimintoja mallintaessa havainneaineiston tulisi siis olla jo olemassa, ja se voitaisiin teoriassa saavuttaa esimerkiksi jo taltioidusta meriliikenteen datasta, kuten kameralla taltioiduista sääolosuhteiden vaikutuksista laivan liikkeisiin, odottamattomien tapahtumien seikkaperäiset kulut ja kapteenien lopulliset päätökset näihin tietoihin nojaten.

Juuri tämänkaltaiset ominaisuudet ovat bravuureja autonomian tai mahdollisimman automaattisen toiminnallisuuden edistämisessä vaikeimmalla saralla eli ihmisen mielen toiminnallisuuden matkimisessa. Näiden menetelmien so-

veltuvuutta voitaisiin tutkia yhä edelleen autonomisten rahtilaivojen sekä niiden vaatimien järjestelmien suunnittelussa. *ACT-R* tai *ACT-R/E* tuskin yksinään voivat ikinä toimia täydellisenä ihmisen mielen kopiona autonomisella järjestelmällä, vaikkakin niillä voidaan saavuttaa tiettyjen asiakokonaisuuksien kautta ihmismielen havainnointia tukevia johtopäätöksiä; vaikkapa merikapteenien toiminnasta käyttöliittymien kanssa sekä olosuhteellista tietoisuutta vaativien tehtävien suoritusta päätöksentekoon vaikuttavien parametrien summana. Lisäksi modernimmat mallintamiskeinot ovat reaaliaikaiseen päätöksentekoon soveltuvaisempia kuin ehkäpä raskaammat mallintamiskeinot, sillä olosuhteet vaihtuvat merellä melko usein ja autonomisen järjestelmän on tehtävä päätelmiä tulevista ja nykyisistäkin toiminnoistaan alati kiihtyvällä tahdilla. Nykypäivänä esimerkiksi *Python*-ohjelmointikielellä on suhteellisen helposti saatavilla olevia neuroverkkomodulleja eli paketteja, joilla voidaan tekoälyn ja neuroverkkojen avustuksella luoda oppivia malleja. Syvien neuroverkkojen avulla vaikkapa tutkien kuvia voidaan saavuttaa luotettavia ja merkittäviä oppimistuloksia objektien luonteesta antamatta juurikaan esitietoja tehtävänsuoritukseen. (Krizhevsky, Sutskever & Hinton; 2012. Pedregosa ym., 2001.) Krizhevsky ja muut (2012, 8) kuitenkin nostavat esiin sen tosiseikan, että kuvien analysointi on erilaista kuin liikkuvan kuvan analysointi ja videoiden analysointiin tulisi kiinnittää enemmän huomiota jatkotutkimuksissa, ja koska se on autonomisen rahtilaivan yksi tärkeimpiä ominaisuuksia, olisikin se suotavaa. Koska neuroverkkojen toimintaperiaatteet ja niiden soveltamiskeinot ovat suhteellisen laajoja eikä tämän tutkimuksen pääpainona ole tietotekniset lähtökohdat, neuroverkkojen toimintaan ei tässä tutkimuksessa syvennyttä sen enempää.

#### 6.4 Mihin pisteeseen saakka ihmismieltä tulee mallintaa?

Onko mikään kognitiivinen malli lopulta tarpeeksi kokonaisvaltainen kuvantamaan ihmisen mieltä? Pystytäänkö kaikki niin sanotut mielen sopukat mallintamaan joksikin yksikäsitteiseksi toimintokokonaisuudeksi? Jos ihmismieli ei edes nykypäivänä kykene päättämään, pystytäänkö ihmismieli mallintamaan, pystytäänkö ihmisen mieli mallintamaan lopulta ikinä täydellisesti, jos ihmisen mielikään ei siihen pysty ajatuksen tasolla? Olisiko paras keino päästä ihmismielen sisälle pelkästään haastatella ihmisiä sekä tutkia ja analysoida näiden käytöstä soveltuen tutkittavaan aiheeseen, kuten tässä tapauksessa pitkälle automatisoitu rahtilaiva, ja toteuttaa hienoista tekoälyä järjestelmien automatisoinnilla – ei siis lähteä ajatuksella, että laiva itse olisi täysin autonominen tai saavuttaisi luvussa 5 mainitun ylimmän automaation tason eli autonomian?

Tämän pro gradu -tutkimuksen ollessa kyseessä nämä henkilöt voisivat olla esimerkiksi merikapteeneja, jotka tietävät, kuinka laivan tulisi käyttäytyä tietyissä tilanteissa ja kuinka erikoistilanteissa toimitaan. Pelkästään merikapteenin kokonaisvaltaisen ajatusmaailman simuloiminen jonkin laskennallisen mallin raameihin kuulostaa tehtävältä, jota ei tulla täysin saavuttamaan lähi-vuosina, joten voisi olla realistisempaa lähteä ihmismielen mallintamiseen mentaliteetilla, jossa pieni määrä tekoälyä toimii rinta rinnan ihmisen mielenmaail-

man kanssa. Tutkitaan ensin esimerkiksi olosuhteellista tietoisuutta mittaavissa testeissä, kuinka ihminen saavuttaa olosuhteellista tietoisuutta, minkä jälkeen voidaan pohtia, mille automaation tasolle vietäisiin minkätyyppistä automatiisaatiota, jos ylin taso eli autonomia ei ole realistisesti saavutettavissa. Tämän myötä ihmistä palvelee robotti, joka osaa hieman ajatella loogisesti, matemaattisesti ja ihmismäisesti muttei kuitenkaan läheskään täydellisesti: kuten Baker ja Tenenbaumkin (2014, 198) toteavat, ihminen päihittää tekoälyn ylivoimaisesti sosiaalisissa tilanteissa vielä tänäkin päivänä. Vaikka sosiaalista tekoälyä onkin viime vuosina pystytty edistämään, sellaisen sosiaalisen älykkyyden kehitys, joka voittaisi ihmisen, vaatisi todella paljon edistysaskelia.





## 7 TIETOISUUS YMPÄRÖIVÄSTÄ TODELLISUUDESTA

Ihmisen päätöksentekoon vaikuttavat lähes aina olosuhteisiin liittyvät tekijät. Esimerkiksi ihmisen siirtyessä sijainnista toiseen tämä mitä luultavimmin ottaa huomioon vaikkapa kulkuvälineen, reittivaihtoehdot ja vallitsevan sääolosuhteen. Tällainen valinta saattaa vaikuttaa yksinkertaiselta, mikäli kulkuvälinevaihtoehtoja ei ole montaa tai sääolosuhteet eivät juurikaan muutu, mutta varsinkin monimutkaisempiin järjestelmiin ja päätöksiin siirtyessä ihmisen on otettava huomioon ja puntaroitava lukuisia tehtävään suoraan tai epäsuoraan vaikuttavia muuttujia sekä niiden syy- ja seuraussuhteita, missä virheen mahdollisuudet kasvavat huomattavasti. (Endsley, 1995a.) Tämä sama ongelmallisuus heijastuu siihen tosiasiaan, kun ihmisen on määriteltävä autonomisen järjestelmän tai robotin kannalta, mitä seikkoja sen tulisi huomioida ympäristöstään. Olosuhteellinen tietoisuus (engl. *situational awareness, SA*) lopulta määrää suuren osan ihmisen päätöksestä monisyisissä järjestelmissä sekä on suoraan kytköksissä toiminnan tehokkuuteen dynaamisissa ympäristöissä, mikä tekee siitä merkittävimpiä vaikuttajia kokonaisuunnistumiseen (Endsley, Bolté & Jones; 2003).

Robotin on usein tarkkailtava ympäristöään, sillä robottia ympäröivä todellisuus on usein ihmisen mielenkiinnon kohteena. Robotti voi valvoa vallitsevia sääolosuhteita, liikkua kinemaattisesti paikasta toiseen kuljettajana tai vaikkapa avustaa ihmistä samassa toimintaympäristössä samalla antaen tärkeää tietoa ympärillä sijaitsevista, työhön vaikuttavista muuttujista. Etenkin ihmisen yhteistyö etäohjauksena robotin kanssa auttaa monessa kompleksisessä työtehtävässä, jossa esimerkiksi ihmisen läsnäolo voi olla vaaraksi itselleen, kuten pelastustehtävissä tai sotilaallisissa toimitissa, tai sellaisissa tehtävissä, joissa ihmisen etävalvonta on koneellisen havainnoinnin kautta yhtä tehokasta tai jopa tehokkaampaa kuin paikan päällä seuraaminen. (Esim. Riley & Endsley, 2005; Salas, Prince, Baker & Shrestha; 1995.) Varsinkin autonomisen robotin tai järjestelmän tulee osata havainnoida ympäristöään kaikin mahdollisin keinoin, jotta se pystyisi esimerkiksi Asimovin (1942) kolmannen lain mukaan suojelemaan itseään tai tehdä juuri oikeat päätökset mahdollisesti operoijaa konsultoiden. Siksi olosuhteellinen tietoisuus on usein haluttu ominaisuus sekä oleellinen

osuus sen kokonaistoimintaa, mutta sen määrittäminen ei ole välttämättä yksioikoista (Drury, Keyes & Yanco; 2007). Esimerkiksi yksi hyväksytyimmistä määritelmistä on ollut Endsleyn (1988): havainnointi elementeistä kyseisessä tilassa kuluvana aikana, elementtien tarkoitusperän sisäistäminen sekä elementtien toiminnan ja olemuksen ennustus tulevaisuuteen. Drury ja muut (2007) kuitenkin huomauttavat, että tämä yleisesti käytetty määrittäminen ei ole tarpeeksi kattava tai yksityiskohtainen taklataksaan nykyajan monitahoiset haasteet olosuhteellisessa tietoisuudessa, vaan sitä on tarkasteltava palastellummin, ja sitä koskeva kirjallisuus on tästä syystä hyvin laajaa (mm. Yanco & Drury, 2004b, 2836).

Olosuhteellisen tietoisuuden määrittäminen ei olekaan aivan yksiselitteistä. Akateemista, yleispätevää määritelmää ei ehkä voida täysin sanoa olevan, vaan sille on akateemikoiden myötä muotoutunut monet eri kasvat. Salmonin ja muiden (2009) mukaan se voidaan nähdä varsinkin työmuistia tarkastelevana konstruktiona (Bell & Lyon, 2000), kognitiivisena informaatioprosessina (Endsley, 1995b) tai täysin ulkopuolisesti ohjailtuna tietoisuutena (Smith & Hancock, 1995). Lisäksi on epäselvyyttä siitä, tarkoitetaanko sillä tietoisuuden hankinnan prosessia (Fracker, 1991), tässä prosessissa muodostunutta tulosta (Endsley, 1995b) vaiko kenties näiden kahden yhdistelmää (Smith & Hancock, 1995). Myöskään olosuhteellisen tietoisuuden mallintamismenetelmissä ei ole päästy yksimielisyyteen: malleina voidaan käsittää muun muassa Endsleyn (1995b) kolmitasoinen kognitiivinen informaatioprosessimalli, aisteihin perustuva ekologinen malli (Niesser, 1976) taikka aktiviteettiteoriamalli (Bedny & Meister, 1999). Voidaan myös kiistellä jopa siitä, onko termi täysin oma psykologinen käsitteensä (esim. Bell & Lyon, 2000), vai onko se vain summa muista kognitiivisista komponenteista, kuten havainnointikyky ja työmuisti. (Salmon ym., 2009.) Yhtä kaikki; tässä luvussa kuitenkin lähestytään olosuhteellista tietoisuutta lähinnä kognitiivisesta sekä informaatioprosessimaisesta näkökulmasta, jota edustanee parhaiten muun muassa Endsley (1995b & 2000).

## 7.1 Tietoisuuden eri osa-alueet

Olosuhteellinen tietoisuus voidaan siis jakaa kolmeen osaan: havainnointiin, ymmärrykseen ja ennustamiseen (Endsley, 1995b). **Ensimmäinen taso** sisältää sensoritasolla havaittavissa olevia merkkejä tai vihjeitä esimerkiksi aistittavissa olevista seikoista, kuten ympäristön näkeminen tai äänimerkkien kuuleminen. (Endsley & Connors, 2008, 1.) Robotille nämä kaksi voivat olla helpohkoja toteuttaa toimiviksi ominaisuuksiksi, mutta ihmisen ohjatessa robottia etänä muita oleellisia yksityiskohtia aistitasolla ei voi välttämättä mitata yhtä hyvin kuin ihmisen ollessa paikan päällä. Miten robotti voi aistia muun muassa poikkeavan hajun, joka on seurausta epäkunnossa olevasta laitteistosta tai sen osasta, jos järjestelmä ei jostain syystä kykene muulla tavalla poikkeavuutta tunnistamaan? Huomaamatta jäänyt syy toimimattomuudesta voi vaikuttaa oleellisesti ihmisen tekemään päätökseen, mitä järjestelmä tekee seuraavaksi tai jättää tekemättä, kun valvoja- tai toimijaroolissa työskentelevä ihminen ei ole paikan

päällä havaitsemassa ongelmaa, jolloin viallisuus saattaa kumuloida muitakin ongelmallisuuksia. Autonomisen järjestelmän ollen kyseessä robotti saattaa vielä herkemmin tehdä virheellisiä johtopäätöksiä, sillä se ei kykene välttämättä huomioimaan hajusta johtuvia muita ongelmia, jotka paikan päällä oleva ihminen kykenisi yhdistämään tuoksuun. Biologisten aistien matkiminen robottitasolla lienee haastavaa, eikä hajuaisti ole niistä lainkaan helpoiten replikoitavissa, mutta jonkin verran on olemassa tutkimusta siitä, kuinka robotti voi esimerkiksi aistia tuoksua sekä paikantaa niitä suhteellisen tarkasti (Song, Liu & Wang; 2010).

**Toinen taso** olosuhteellisessa tietoisuudessa muodostuu kyseisellä hetkellä vallitsevan tilanteen hahmottamisesta sekä sen perinpohjaisesta ymmärtämisestä. Toisella tasolla ainoastaan ensimmäisen tason elementtien havainnointi ei yksinään riitä muodostamaan kokonaiskuvaa tilanteesta, vaan toimivien autonomisten järjestelmien toteuttamiseksi on sekä ihmisellä että robotilla oltava äärimmäisen laaja-alainen ymmärtämys myös näiden elementtien tärkeydestä ja merkityksestä. Tämä merkitys ulottuu niin elementteihin itseensä kuin niiden vuorovaikutukseensa toisten elementtien kanssa kohdistuen alati relevantteihin robotin operoijan määrittämiin tavoitteisiin. (Endsley, 1995b, 37.) Tämä puolestaan voidaan käsittää Normanin (1986, 41) seitsemänä vuorovaikutuksen tasona, joita Scholtz (2002) avaa syvällisemmin:

1. Tavoitteen määrittäminen: korkean tason termein, mitä ihminen haluaa toiminnallaan saavuttaa.
2. Aikomuksen määrittely: syventyen ensimmäiseen kohtaan pohdinta tarkemmin siitä, mikä tyydyttää tavoitteen saavuttamisen.
3. Toiminnan erittely: mitkä ovat ne toiminnot, joiden avulla toisen kohdan aikomus toteutuu niin itsessään kuin järjestyksessä muihin toimintoihin nähden.
4. Kolmannen kohdan toiminnan suorittaminen, eriteltyjen toimintojen fyysinen täytöntöönpano: tietokonetermein oikeiden komentojen valinta niitä vastaavien toiminnallisuuksien toteuttamiseksi.
5. Järjestelmän tilan hahmotus: kolmannessa kohdassa määritellyn ja neljännessä suoritettujen toiminnan arviointi, tapahtuneen käsittäminen.
6. Järjestelmän tilan tulkinta: viidennessä kohdassa hahmotettujen tilan laaja-alaisempi ymmärtäminen sekä tapahtuneen analysointi.
7. Lopputuleman arvioiminen: käyttäjä tai operoija peilaa viidennessä ja kuudennessa kohdassa havaitun ja analysoidun tilan vastaavuutta ensimmäisessä ja toisessa kohdassa määritettyyn tavoitteeseen. Tämän jälkeen käyttäjä määrittää seuraavan päätepisteen toiminnalle palamalla ensimmäiseen kohtaan.

Täten järjestelmää operoiva henkilö iteroi näitä seitsemää vuorovaikutuksen osa-aluetta koko ajan, kunnes on päästy tavoitteeseen. Saavuttaakseen lopullisen, ylimmän tason tavoitteen operoija asettaa ja toteuttaa koko ajan välitavoitteita näiden kohtien mukaisessa järjestyksessä tai muokkaa aiempia tavoitteita suorittaakseen tehtävän pienemmissä osuuksissa, mikäli mahdollista. Näin järjestelmän toimintaa ja tilaa kyetään havainnoimaan ja tulkitsemaan tarkemmal-

la tasolla sekä yhdistämään sitä toisiin informaation osasiin, mikä edesauttaa saavuttamaan tilanteellista hahmotuskykyä tällä toisella tasolla. (Norman, 1986; Scholtz, 2002b; Scholtz, Antonishek & Young; 2004.)

Scholtzin (2002) mukaan Norman (1986) näkee näissä seitsemässä tasossa kuitenkin kaksi ongelmallisuutta. Ensimmäinen ongelmallisuus liittyy ihmisen rajallisiin kykyihin järjestelmässä toteuttaa haluamansa toiminnot. Tämä yhteentörmäys tapahtuukin silloin, kun operoija haluaa suorittaa toiminnon, jollaista järjestelmä ei kykene tarjoamaan. Toinen kompastuskivi tämän toisen tason alemmassa seitsemässä tasossa piilee evaluoinnissa: järjestelmän oletettu syöte saattaa poiketa ihmisen käsityksestä sen tilasta, jolloin väärät johtopäätökset voivat aiheuttaa ongelmatilanteita, jotka mahdollisesti myöhemmin kumuloiduvat ja kilpistyvät vääriksi johtopäätöksiksi. (Norman, 1986; Scholtz, 2002b.) Tällainen tilanne voi käydä esimerkiksi autonomista laivajärjestelmää operoidessa silloin, kun laivan tarkkaillessa ympäristöään se huomaa – tai ei huomaa – erikoistilanteen, josta se ei ilmoita valvojalle tai operoijalle, tai josta se ilmoittaa, mutta järjestelmää ei ole konfiguroitu juuri tähän erikoistilanteeseen mukautuvaksi. Tämän seurauksena ihmisellä ei ole konsteja taltuttaa vikatilannetta, vaan vaaditaan esimerkiksi määrittelijää uudelleenarvioimaan ja -spesifioimaan kyseessä oleva toiminnallisuutta vastaava funktio (Saariluoma, 2010), jonka mekaanikko osaltaan korjaa ja jonka operoija tai valvoja vahvistaa validiksi korjaukseksi, jotta tilanteellinen tarkkailu voi jatkua onnistuneesti.

Viimeisimpänä **kolmas taso** edustaa ennustuskykyä tulevaisuuteen (Endsley, 1995b, 37). Ympäristössä vallitsevan kyseisen tilan ja sen elementtien tunnistaminen ja analysointi voi olla riittävää kyseisen tilanteen hallitsemiseksi, mutta on varauduttava jatkossakin siihen, että olosuhteet sekä toimintoihin ja toistensa välisiin vuorovaikutuksiin vaikuttavat objektit ja tekijät voivat muuttua oleellisesti. Jos tätä ennustusta ei ole tehty aiemmin mahdollisissa puitteissa, silloin kyseessä vallitsevaan tilanteeseen varautuminen saattaa olla vajavaista, mikä voi heijastua ympäristöllisen ja tilanteellisen toimintavarmuuden heikkenemiseen muun muassa vajavaisen, kriittisen tiedon aiheuttamissa päätöksissä. Endsleyn (1995b, 37) esimerkkiä lentokoneesta mukailen on tärkeää osata arvioida, mihin suuntaan esimerkiksi autonominen laivajärjestelmä seuraavaksi ohjaa laivaa. Perustuen aiempiin kokemuksiin ja dataan laivan käyttäytymisestä ihminen voi tarkemmin nimenomaan tietää, minkälaiset tilanteeseen havaitut muuttujat tulevat vaikuttamaan laivan liikehdintään, jolloin operoijalle tai valvojalle jää enemmän aikaa päättää, minkälaisia tulevia toimenpiteitä laivalle tulisi antaa. Kyky ennustaa myötävaikuttaa siis verrattain paljon ihmisen asettamien kokonaistavoitteiden saavuttamisessa pala palalta tehostaen ajankäyttöä, tiedon määrää sekä toiminnan tehokkuutta. (Endsley, 1995b, 37.) Tähän liittyen esimerkiksi laivateollisuuteen ja merenkulkuun liittyen onkin kehitetty menetelmä, jonka avulla järjestelmän ja laivan käyttäytymisen pohjalta voidaan tehdä ennusteita sen tulevasta sijainneista. Järjestelmä oppii ennustemalleja autonomisesti ja nopeasti hyödyntämällä yhtä tai useampaa ennalta määrättyä, senhetkistä horisonttia. Jatkuvan ja aiempaan dataan pohjautuvan toimintatavan myötä tehokkuus paranee entisestään, kun alus tarkkailee ja arvioi omaa toimintaansa, mikä tekee siitä efektiivisen merenkulun ja lii-

kehdoennän ennakoinnin kannalta tilanteellisessa tietoisuudessa. (Bomberger, Rhodes, Seibert & Waxman; 2006; 6 & 7.)

## 7.2 Tietoisuuden saavuttaminen

Olosuhteellinen tietoisuus voitaneen kiteyttää ihmisen ymmärrykseksi siitä, mitä robotin ympärillä on tapahtumassa ja tulee tapahtumaan sekä mitkä tilanteelliset elementit vaikuttavat tapahtumiin, lopputulemaan ja johtopäätöksiin. Ja jotta ihminen voisi ymmärtää tämän parhaiten, täytyy operoijalla sekä robotilla itsellään olla keinonsa saada selvää, mitkä seikat voivat vaikuttaa robotin toimintaan, tiedonkeruuseen sekä täten haluttuihin lopputuloksiin. (Adams, 2005, 448.) Ihminen voikin saada kolmella eri keinolla yksityiskohtaista tietoa siitä, minkälainen olotila robotin läheisyydessä eli miljöössä vallitsee, mikäli robotille on suotu tämänkaltaiset kommunikointi- ja tiedonkeruumahdollisuudet. (Hjelmfelt & Pokrant, 1998; Yanco & Drury, 2004b):

1. eksplisiittinen: kysytään suoraan järjestelmän tai esimerkiksi ympäristön tolasta niin kauan kuin autonominen robotti voi vastata oikeellisesti ja yksiselitteisesti
2. implisiittinen: tutkimuksen järjestäjät mittaavat koehenkilöiden tehtävnsuoritusta olettamalla, että koehenkilön suorituskyky korreloi olosuhteellisen tietoisuuden kanssa: kun olosuhteellisen tietoisuuden määrä kasvaa, samoin kasvaa suorituskyky
3. subjektiivinen: koehenkilöt arvioivat itse oman olosuhteellisen tietoisuuden tasonsa.

Kun ihminen työskentelee pääsääntöisesti yhden robotin kanssa, myös kolme osa-aluetta tulee kyseeseen (Drury ym., 2007, 2). Ensimmäinen näistä on ihmisen tietoisuus robotista: ihmisellä tulee olla tietoisuus koko ajan siitä, minkälaiset olosuhteet ja muuttujat robotin ympärillä vallitsevat. Näihin muuttujiin kuuluvat muun muassa sijainti, toiminnot, status ja muut ympäristöön vaikuttavat tekijät. Tarkemmin sanottuna sen voitaisiin sanoa olevan varmuus siitä, kuinka ihminen eli operoija tunnistaa nämä elementit.

Toinen osa-alue on robotin tietoisuus ihmisestä; informaatio siitä, minkälainen kapasiteetti robotilla on tiedostaa ja ymmärtää ihmisen antamat komennot. Tämän tietoisuuden avulla robotti osaa kohdentaa komennot juuri oikeisiin toimintoihin. Tietoisuuteen liittyvät myös ne erikoistilanteet, joissa robotti ei kykene toimimaan komentojen mukaan tai komennot vaativat muutoksia. Robotti voikin havainnoida ihmisen toiminnassa puutteita muun muassa Devinin ja Alamin (2016) kehittämän viitekehyksen avulla. Viitekehyksen myötä robotti osaa arvioida ja ylläpitää eri toimijoiden mentaalaisia malleja (engl. *mental model*) ei pelkästään ympäristöllisten attribuuttien vaan myös tavoitteiden tilan havainnoinnin kautta.

Kolmas ja viimeinen osa-alue on ihmisen tietoisuus tehtäväkokonaisuuden kokonaistilanteesta: ihminen on aina ylin valvoja autonomisiin järjestelmiin

ja robotteihin kytkeytyvissä verkostoissa, joten tällä on oltava korkeimman tason käsitys siitä, mitä kokonaisuudessaan on tapahtumaisillaan (mm. Yanco & Drury, 2004a, 2843). Ihminen ohjaa koko ajan robottia ja sen toimintoja kohti asettamiaan päämääriä, jolloin tilanteellinen hahmotus auttaa toimimaan oikein pienemmissä, palastelluissa operaatioissa. Näiden korkeamman tason tavoitteiden pienempien osatavoitteiden tilanteellinen hahmottaminen auttaa käsittämään kokonaistilanteen osissa helpommin.

### 7.3 Tietoisuuden mittaaminen

Vaikka olosuhteellisen tietoisuuden kolme tasoa vaikuttavat tehokkuuteen, olosuhteellisen tietoisuuden vuorovaikutus tehokkuuden tasoon ei ole kuitenkaan täysin korreloitunutta. Ei voida siis sanoa, että henkilön saavutettua kolmannen tason tietoisuuden tehokkuus olisi huipussaan, vaan ihminen voi silti suoriutua heikosti. Tämä nähdään jo Normanin (1986) seitsemästä tasosta ja niihin liittyvistä sudenkuopista, mutta muita heikkouksia olosuhteellisessa tietoisuudessa ovat juurikin jo aiemmin mainitut käyttöliittymien puutteet ja toimintahäiriöt sekä operoijan eli ihmisen itsensä kognitiiviset puutteet tai heikkenemiset. (Scholtz, 2002b.) Eritoten näiden kahden heikkouden yhteisvaikutus voi johtaa jopa onnettomuuksiin saakka. Yksi tärkeimpiä ominaisuuksia autonomisissa järjestelmissä onkin se, kuinka virhetilanteisiin asti päätyminen voitaisiin ehkäistä ja miten häiriötilanteessa autonominenkin järjestelmä otetaan jälleen haltuun.

Negatiivisten seurausten eli häiriötilanteiden ehkäisyyn ja olosuhteellisen tietoisuuden mittaamiseen on olemassa erinäisiä keinoja, joista ensimmäisenä on ehkäpä tyypillisin ja yleisin: suoraan kysyminen. Tämä *SAGAT*-menetelmä (engl. *Situation Awareness Global Assessment Technique*; tilanteellisen tietoisuuden yleinen hahmotusmenetelmä; Endsley, 2000) on eräs yksinkertaisimmista tavoista saada selville se, millaisena operoija havaitsee ohjaamansa järjestelmän kyseisen tolan. Operoijilta tiedustellaan *SAGAT*-menetelmän ripeillä kyselyillä tavoitekeskeisesti järjestelmän keskeisistä olosuhteelliseen tietoisuuteen liittyvistä vaatimuksista siten, että järjestelmässä kyseisellä hetkellä tapahtuvat toiminnot keskeytetään. Näin operoijalle jää aikaa vastata intuitiivisesti ja nopeasti kaikkiin kolmeen tasoon (havainnointi, oivallus, ennustus) liittyviin kysymyksiin, jolloin tulee ilmi järjestelmän hyvät ja huonot puolet – mitä enemmän operoija saa vastauksia oikein, sitä parempi olosuhteellinen tarkkaavaisuus on sillä alueella. (Endsley, 2000; Scholtz, 2002b.) *SAGAT*in hyvä puoli muihin mittausmenetelmiin nähden lieneekin siinä, että se pohjautuu juuri kyseisellä hetkellä tapahtuviin, heti operoijan tuntemuksista mitattuihin näkemyksiin kyseisessä ympäristössä kohdentuviin elementteihin. Näin käytöstä saadaan arvokasta tietoa eikä olosuhteellisen tietoisuuden arvioinnissa operoijan tarvitse tarkastella ja muistella syvällisesti, mitä on tapahtunut aiemmin, jolloin kokemusten myöhäisempi linkittäminen toisiinsa voi aiheuttaa vääristymiä todellisuudesta ja näkökulmista sekä niiden tulkinnoista. (Endsley, 2000.) *SAGAT* kuitenkin vaatii nimenomaan toiminnan keskeytystä kyselyihin vastaamiseksi,

minkä vuoksi se ei sopeudu kaikkiin tilanteisiin sovellettavaksi keinoksi. Tästä syystä menetelmän käyttö voisi olla perustelluinta vaikkapa simulaatioympäristöissä, joissa mallinnetaan tosielämän käytöstä, mutta tällöin yleistäminen itse varsinaiseen autenttiseen tilanteeseen ei välttämättä ole suoraan verrannollista vaan ehkäpä pikemminkin suuntaa antavaa.

SAGAT-mittausmenetelmän ohella toinen käytetty keino olosuhteellisen tietoisuuden mittaamiseksi on SART (engl. *Situational Awareness Rating Technique*; Selcon & Taylor, 1990). SART mittaa tarkalla tasolla ja yksinkertaisesti suoraan operoijalta tämän näkemyksiä kustakin esitetystä, tarkastelun kohteena sijaitsevasta olosuhteellisen tietoisuuden eri ulottuvuudesta, ja sitä on käytetty varsinkin ilmailussa olosuhteellisen tietoisuuden mittaamiseen (esim. Durso ym., 1998, 1), mutta se ei ole välttämättä ainut ala, jossa sitä voidaan soveltaa. SART-menetelmää onkin käytetty myös muun muassa puoliautonomisten kuljetussaattueiden operoijien olosuhteellista tietoisuutta mitatessa (Davis, Animashaun, Schoenherr & McDowell; 2008). Tutkimuksen tuloksista huomattiin, että miehitettyjen ajoneuvosaattueiden osittaisella automatisoinnilla erityisesti suorituskyky parani monella olosuhteellisen tietoisuuden osa-alueella: kohteiden tunnistaminen tehostui, välimatkojen arviointi parani sekä toimintakyky odottamattomissa seisahduksissa kohentui. Tutkimuksessa SARTia on käytetty yhdessä muiden metodien kanssa, kuten NASAn tehtävänkuormitusindeksi (engl. *NASA Task Loading Index [NASA-TLX]*; Hart & Staveland, 1988), matkapaahoinvoinnin arvioinnin kyselymenetelmä (engl. *Motion Sickness Assessment Questionnaire [MSAQ]*; Gianaros, Muth, Mordkoff, Levine & Stern; 2001) sekä asteikko arvioimaan luottamusta automatisoituun järjestelmään (Jian, Bisantz & Drury; 2000), jotka kaikki osaltaan voivat olla hyödyllisiä kokonais-suorituskykyä mitatessa.

SART ja SAGAT eivät siis suinkaan olekaan ainokaisia tai ainoita oikeita tapoja tutkia olosuhteellista tietoisuutta tai siihen vaikuttavia alueita, vaikkakin nämä kaksi voivat olla tunnetuimpia. Vaikka ne ovat yleisiä menetelmiä, esimerkiksi SARTia on kritisoitu muun muassa siitä, että se mittaa olosuhteellista tietoisuutta liian subjektiivisella tasolla. Subjektiivisten mittausmenetelmien ongelmakohta on siinä, että ne sekoittavat tuotteen ja prosessin, jolloin operoijan olosuhteellista tietoisuutta mittaavat tulokset eivät ole täsmälleen kysymystä vastaavat, vaan vastaukset saattavat sekoittaa muihin, jopa mielenkiinnon ulkopuolella sijaitseviin osa-alueisiin. (Gilson ym., 1994, 121.) Toisaalta päätöksenteon hetkellä olosuhteellisen tietoisuuden mittaaminen absoluuttisen objektiivisesti ei liene edes mahdollista, sillä päätös perustuu aina jonkin, henkilön taikka järjestelmän, muodostamaan johtopäätökseen. Oman hankaluutensa luo myös se, mitä tarkoitetaan *yhteisellä* olosuhteellisella tietoisuudella, jolloin ihmisen ja robotin yhteisesti muodostaman kokonaiskäsityksen (engl. *shared situational awareness*) mittaaminen voi olla osaltaan entistä haastavampaa. (Esim. Nofi, 2000, 71.) Valitsemalla oikean mittausmenetelmän olosuhteellisen tietoisuuden puutteesta tai sen vääränlaisesta tulkinnasta johtuvien ongelmallisuuden määrää voidaan vähentää, ja tässä valinnassa voi auttaa esimerkiksi Salmonin, Stantonin, Walkerin ja Greenin (2009) luoma kattava katsaus eri mittaustavoista koostuen niiden eri ominaisuustyypeistä, kuten sovellutusalueista, vaativuudesta ja vaatimuksista sekä hyvistä ja huonoista puolista.

## 7.4 Olosuhteellinen tietoisuus etäohjattavissa roboteissa

Drury ja muut (2007) tutkivat, miten olosuhteellista tietoisuutta saadaan etäohjattavien robottien avulla. He jakavat olosuhteellisen tietoisuuden viiteen eri alakohtaan, joita he vertailevat keskenään kahden eri järjestelmän välillä. Koehenkilöitä rohkaistiin ajattelemaan ääneen, mikä on olosuhteellisen tietoisuuden mittaamisessa yksi tehokkaita konsteja, sillä tietoisuuden tasoa voidaan mitata reaaliaikaisesti juuri tarvittavalla hetkellä (Drury ym., 2007, 2). Seuraavia löydöksiä voitaisiin hyödyntää eritoten autonomisen rahtilaivan järjestelmien kehityksessä olosuhteellista tietoisuutta kerätessä, sillä heidän tutkimuksensa pääpainona ovat kartta-, tutka- ja videopohjaisen käyttöliittymät, jotka molemmat lienevät laivaliikenteenkin tärkeimmät apuvälineet (Yle, 2017). Kokonaisuudessaan voitaisiin sanoa, että etäohjattavissa roboteissa karttapohjainen käyttöliittymä on videopohjaista parempi antamaan sijaintiin ja olosuhteisiin liittyviä tietoja, kun taas videopohjainen käyttöliittymä antaa paremmin tietoa ympäristöstä ja tietoisuudesta ympärillä tapahtuvista aktiviteeteista (Drury ym., 2007).

Koehenkilöiltä arvioitiin ensimmäisenä olosuhteellisen tietoisuuden tasoa sijainnista. Koehenkilöt pitivät koko ruudun kokoista karttaa, joka osasi merkitä maamerkkejä ja robotin itsensä sille, parempana vaihtoehtona kuin karttaa, jonka sai halutessaan pois päältä. Paremmassa vaihtoehdossa oli myös mahdollista lisätä itse välietappeja kartalle sekä seurata sitä, missä robotti oli ollut milläkin ajanhetkellä. Tämä mahdollistaa sen, että menneistä sijainneista on helppo suunnistaa kokonaan näkyvällä kartalla seuraavia sijainteja kohden, joita ei ole välttämättä vielä löydetty. Koehenkilöt myös huomauttivat, että vaikka kartta oli hyvä antamaan tietoa kokonaissijainnista, he käyttivät sen lisäksi usein hyväkseen videokuvaa: koehenkilöt tunnistivat kameran kautta kartan kanssa yhteistyössä sellaisia sijainteja, joissa he olivat olleet jo aiemmin, mikä voi olla merkittävä keino suunnistaa tietyillä, tutuilla merialueilla.

Toisena seikkana olosuhteellisessa tietoisuudessa on tietoisuus ympärillä tapahtuvista aktiviteeteista ja robotin liikehdinnässä. Koehenkilöt kokivat kameran paljon parempana vaihtoehtona kartalle seurata robotin liikettä. Kun robotti oli jostakin syystä hidastanut vauhtiaan tai pysähtynyt täysin, he hankkivat tietoisuutta kameroiden avulla, mitä ympärillä tapahtuu. Se onkin huomattavasti kartan robotti-ikonin liikkumattomuutta parempi tapa saada selvää, miksi esimerkiksi autonominen rahtilaiva on hidastanut vauhtiaan; oli se sitten vaikkapa meressä kelluva jätelautta tai robotin sensoreiden havaitsema toinen alus, mistä syystä laiva on kääntymässä. Edistyneemmät koehenkilöt osasivat ottaa hyötyä irti myös robotin lasertunnistustekniikoista: kun robotti oli jumissa, osa koehenkilöistä kykeni tulkitsemaan tilannetta niin, että esimerkiksi robotin kaksi sivua olivat vapaana mutta kaksi muuta esteiden ympäröimänä. Samankaltaisia edistyneempiä ympäristönhavaitsemistekniikoita tarvitaan autonomisen rahtilaivan kehityksessä, missä ympäröivien aktiviteettien havaitseminen edes lämpökameroiden tai erikoisvalaistujen tuplakameroiden kautta ei välttämättä ole mahdollista, kuten Drury ja muiden (2007) tutkimuksessa on, mikäli laiva on jo jostain syystä jumiutunut, ei pysty liikkumaan eikä ole pystynyt ha-



vaitsemaan poikkeavuuksia ympäristössään tarpeeksi ajoissa. Näin laivan moni eri varotoimi osaisi varoittaa laivan järjestelmää yhdessä muiden kanssa, ja varmuuden vuoksi ihminen eli operoija tai valvoja voisi itsekin varmistaa ympäristön turvallisuuden, jotta liikkuminen merellä on mahdollisimman mutkaton.

Kolmas mielenkiinnon kohde on tietoisuus ympäristöstä. Toisessa tutkittussa, robottia etäohjaavassa järjestelmässä oli asennettuna kamera, jonka kirkkaan valon sai päälle halutessaan. Lamppu auttoi varsinkin navigoimaan pimeissä lokaatioissa, joissa kameroiden zoomaus ei onnistunut tarkentamaan ympäristöjään valon puutteen vuoksi. Autonomisessa rahtilaivassa kirkkaiden lamppujen sijainti etenkin laivan keulassa voitaneen katsoa olevan jo itsestäänselvyys, jotta laiva osaa itse havainnoida eteensä ja myös valvoja tai operoija kykenee siihen. Drury ja muut (2007) noteerasivat myös sen, että videokuvan näkeminen robotin takaa auttoi havainnoimaan ympäristöä yhä tehokkaammin kuin pelkällä yhdellä kameralla. Autonomisessa rahtilaivassakin kameroiden määrän – näkeminen moneen suuntaan haluamassaan kuvakulmassa – voidaan taten katsoa olevan olosuhteellista tietoisuutta kasvattava mahdollisuus. Tässä on kuitenkin hyvä osata huomioida sensorien määrä, sillä joskus liian suuri määrä erinäisiä ympäristöä havainnollistavia sensoreita voi olla olosuhteellista tietoisuutta hankaloittava seikka, kun käyttäjä ei osaa keskittyä enää olennaiseen (esim. Yanco ym., 2004). Teknologisesti kuitenkin sensorien fuusioima data voi olla tärkeää, ja sen tärkeys on otettava huomioon: miten voimme yhdistää monen eri sensorin aistiman muutoksen, jos pelkästään yksi ei riitä, ja näyttää siitä tapahtumasta selkeästi tarkka yhteenveto kapteenille (Yle, 2017)?

Neljännessä osa-alueessa, tietoisuudesta statuksesta eli tilasta, ei juurikaan havaittu eroavaisuuksia karttapohjaisen tai videopohjaisen järjestelmän välillä. Tämä tutkimustulos puoltaa sitäkin ideaa, että autonomisessa rahtilaivassa on sekä videopohjainen että karttapohjainen käyttöliittymä ja että molemmilla voidaan saavuttaa laivan tilaa koskevaa tietoisuutta toisiaan täydentäen. Eräs keino havaita esimerkiksi etäohjatun laivan tilaa on laivan järjestelmien niin kutsuttujen energiatasojen tai suorituskykyyn liittyvien tekijöiden havainnollistaminen: kuten autonomisen tai pitkälle automatisoidun järjestelmän toimintaan vaikuttavien akkutasojen ilmaisu käyttöliittymissä.

Ehkäpä mielenkiintoisimpana yksityiskohtana sekä viidentenä ja viimeisenä Drury ja muiden (2007) tutkimuskohteena on kokonaistehtävän suoritukseen liittyvä tietoisuus, kuten autonomisen rahtilaivan perille pääsemisen tila. Kummankaan järjestelmän, kartta- tai videopohjaisen, avulla ei ole havaittavissa suurta merkitsevyyttä sille, että kokonaistehtävän tietoisuuden taso olisi kasvava. Kummankaan järjestelmän koehenkilöt eivät noteeranneet ääneen sitä, kuinka kokonaistehtävänsuoritus olisi jollain tapaa helpompaa ja silmin nähden havaittavampaa etäohjattaessa robottia, vaan he ainoastaan mainitsivat, että olivat unohtaneet tehdä jonkin tietyn asian. Tämä voi olla autonomisen tai pitkälle automatisoidun rahtilaivankin tapauksessa asia, jossa on hyvä tutkia lisää; miten voitaisiin havaita kokonaisuutta paremmin käyttöliittymien ja järjestelmien kautta ja kuinka tilanteellista hahmotuskykyä voitaisiin parantaa antaa esimerkiksi kartalla näkyvin keinoin kokonaissuoritusta koskevia vinkkejä.

Rahtilaivan suurimpia automatisoinnin tai autonomian myötä ilmaantuvia huolenaiheita on törmääminen vieraisiin kappaleisiin avomerellä. Tähän tarvitaan olosuhteellista tietoisuutta niin kokonaisuudessaan kuin yksityiskohteisellakin tasolla. Muun muassa Riveiro, Falkman ja Ziemke (2008) ovat tutkineet, miten olosuhteellisen tietoisuuden tasoa voidaan parantaa vierasperäisten esineiden tunnistuksessa ja kuinka meriliikenteessä havaitaan poikkeamia ja parannetaan olosuhteellista tietoisuutta interaktiivisella visualisoinnilla. Riveiro ja muut (2008, 53) uskovat, että kun tiedonlouhintaa (engl. *data mining*) suuresta määrästä dataa yhdistetään ihmisen vuorovaikutukseen järjestelmän kanssa, kuten kapteenin toimintaan laivan järjestelmissä, voidaan saavuttaa kokonaisuus, jossa ihmisen tietämys sekä järjestelmän automaattisuus kohtaavat. He ovat luoneet käyttöliittymän, jossa on esitettyä vasemmalla kartta; oikealla laivan kontrollit, kuten moduuli, joka huomaa erilaisia poikkeamia ympäristössä, sekä alalaidassa notifikaatiolista ja listaus tarkemmista tehtäviä koskevista tiedoista (Riveiro ym., 2008; 51 & 52). Lopuksi he testasivat luomuksensa pätevyyttä pienimuotoisella simulaatiolla: Ensiksi he määrittivät, minkälaisia aluksia meriliikenteessä voi tulla vastaan, ja niitä olivat laiva tai alus, kalastusvene sekä rahtilaiva. Sen jälkeen he pystyivät havaitsemaan ympäristössä erikoisuuksia, kuten esimerkiksi aluksia, jotka kulkivat opetettuun lähtödataan nähden suuremmalla nopeudella, sekä epätavallisia ja normaalista poikkeavia aluksien kulkusuuntia, kuten vieraalla alueella rannikon lähestyminen. (Riveiro ym., 2008, 53.) He kuitenkin huomauttavat, että käyttäjätestien tekeminen on yhä tarpeen testaamaan käyttöliittymän kattavampi toiminta juuri halutulla käyttäjäjoukolla (2008, 53), kuten tämän tutkimuksen autonomisen rahtilaivan ollessa kyseessä. Riveiron ja muiden (2008) tutkimus kuitenkin luo hyvää pohjaa sille, miten poikkeamia olosuhteellisessa tietoisuudessa ympäröivässä todellisuudessa voidaan havaita tekoälyn avulla antaen käyttäjälle visuaalisella käyttöliittymällä tarpeeksi tietoa juurikin tästä havainnoinnista.

## 8 IHMISEN JA TEKNOLOGIAN VÄLISEN VUORO- VAIKUTUKSEN ROOLIT

Kuka tällaisia autonomisille roboteille toteutettuja käyttöliittymiä siis käyttää taikka robotteja robotteja? Entä kuka käyttöliittymiä puolestaan toteuttaa? Millaisia tietoja ja taitoja kaikilla eri autonomisiin järjestelmiin kytkeytyvillä rooleilla tulee olla? Onko taas autonominen robotti itsessään jo rooli? Tässä luvussa tuodaan vastauksia muun muassa näihin kysymyksiin. Onkin tärkeää osata arvioida roolit esimerkiksi sen vuoksi, että halutaanko autonomisesta robotista ihmisen työtoveri eli kumppani vaiko hierarkiassa alempana tai ylempänä toimiva tekijä, missä toinen osapuoli on enemmän määräävä (vrt. engl. *master-slave*-asetelma; isäntä-renkiperiaate). Tällöin robotin ja sen sisältämän järjestelmän toimintaa voidaan alusta alkaen kehittää haluttuun suuntaan, kun tiedetään, mikä on kunkin tahon tehtävä autonomisiin robotteihin kytkeytyvissä rooliverkostoissa. Seuraavassa on kerrottu nämä viisi tarkasti määriteltyä roolia sekä niiden toimenkuvia. (Scholtz, 2002, 5; Scholtz, 2003; 7.)

### 8.1 Työnjohtaja, valvoja

Ensimmäisessä ryhmässä ovat työnjohtajan eli valvojan (engl. *supervisor*) roolisissa olevat roolit. Ihminen on loppupeleissä autonomisen robotin päävalvoja. Ihmisen on seurattava reaaliaikaisesti robotin liikkeitä siltä varalta, että robotin tekemät arvovalinnat eri vaihtoehtoista ovat juuri ne oikeat valvojan määrittämiä tavoitteita edistävät. Tämän toteuttamiseksi ihmisen hyödyntämät järjestelmät ovat väline seurata ja valvoa autonomisten järjestelmien toimintoja ja tehtäviä. Tärkeimpiä mielenkiinnon kohteita ihmisen valvonnan alla ovat autonomisten järjestelmien suorittamat tehtävät: kun autonominen järjestelmä ilmoittaa selkeästi ihmiselle, mitä se on tehnyt, tällä hetkellä tekemässä sekä seuraavaksi tekemässä, se mahdollistaa ihmisen toimimaan valvojaroolissaan. Tällöin toiminta on läpinäkyvää, mikä on roolien toimimisen kannalta olennainen aspekti. (Miller & Parasuraman, 2007, 71–73.)

Valvoja voi olla myös mekaanisesti toteutettu rooli. Tällöin valvoja toimii järjestelmän sisäisesti tarkkaillen tiedonsiirtoa ja järjestelmässä tapahtuvia toimenpiteitä, mutta se ei kuitenkaan keskeytä tiedonsiirtoa. Valvoja välittää tietoa arkkitehtonisella tasolla moduulilta toiselle samalla pitäen huolta siitä, että tieto välittyy sallitun aikarajan puitteissa: näin valvoja on jatkuvasti tietoinen koko järjestelmän tilasta kuitenkin mahdollisesti ilmoittaen ihmisvalvojalle poikkeamista. (Fritsch, Kleinhagenbrock, Haasch, Wrede & Sagerer; 2005, 3412.)

## 8.2 Operoija, toimija

Toisena roolina on toimija (engl. *operator*), joka nimensä mukaisesti operoi robotin toimintaa ja sen kanssa. Toimijarooli ei siis valvojalle tyypillisesti pelkästään valvo autonomisen järjestelmän toimintaa, vaan tämä voi kontrolloida yhden tai useammankin robotin toimintaa: muuttaa suoritettavia toimintoja tai esimerkiksi määrittää seuraavat toiminnat. Ihmisen on usein tarkkailtava robottia ja operoida robotin kanssa siten, että tämä valvoo itse robotin liikehdintää, mutta ihmisellä on kuitenkin rajallinen kognitiivinen kapasiteetti eli kyky käsitellä yhdenkin robotin suorittamia proseduureja – puhumattakaan, kun operoija käsittelee useaa robottia taikka järjestelmää ja niiden toimintoa samanaikaisesti. (Chien, Lewis, Mehrotra, Brooks & Sycara; 2012.) Tämä voi johtaa siihen, että ihmiselle jää toisaalta vähemmän aikaa keskittyä avoimessa jonossa (engl. *open-queue*) sijaitseviin tehtäviin, sillä niiden priorisointi myös kuluttaa aikaa ja kognitiivista kapasiteettia. Haluttujen toimintojen filteröiminen ei-halutuista saattaa osoittautua harhauttavaksi ja häiritseväksi prosessiksi kauemmas itse tärkeimmistä tehtävistä. (Chien ym., 2012, 478 & 479.)

Kun robotin tehtävät on järjestetty puolestaan prioriteettien mukaiseen jonoon, ihmisen keskittymiskyky on vakaampaa sekä robottikin kykenee auttamaan ihmistä organisoidummin. Paineen alla sekä kognitiivisen kuormituksen lisääntyessä ihminen yleensä alkaa luottaa robotin tai järjestelmän valintoihin tehtävienratkaisussa. Siksi tässä mielessä prioriteettijonossa olevat robotin tehtävät voivat auttaa toimijaroolissa robotin kanssa työskentelevää paremmin kuin avoimet tehtävät. (Chien ym., 2012; 478 & 479.)

## 8.3 Mekaanikko

Kolmantena roolina on mekaanikko, jolla on pääsy robotin mekaanisiin, fyysisiin osiin. Mekaanikko voi siis virittää vaikkapa kameraa tai navigointiin liittyvää laitteistoa komponenttien ja mekanismien osalta. Roolin toimenkuvaan kuuluu esimerkiksi keskeyttää jokin operaatio sen vakavuuden vuoksi samalla arvioiden, minkälaisia vaikutuksia väliintulolla voi olla vuorovaikutukseen jatkossa. (Scholtz, 2003.)

Jos mekaanikko esimerkiksi korjaa jonkin toimimattoman komponentin muttei tiedä, miten korjaus tulee vaikuttamaan muihin osiin sekä täten koko-

naistoiminnallisuuteen, niin tällöin korjaukseen vaadittavia toimenpiteitä on mahdollisesti varmistettava testaajilta tai muilta käyttäjiltä. Testaajat lienevätkin mekaanikon, esimerkiksi ohjelmoijan tai muun korjaajan, läheisimpiä työkumppaneita: korjausten jälkeen toiminnallisuudet on lähes väistämättä testattava uudestaan, jotta kokonaisuus toimii oikeellisesti. Mekaanikon rooli poikkeaa operoijan roolista siten, että muutokset on tehty suoraan laitteistotasolla, johon operoija ei pelkästään toimijaroolilla pääse välttämättä käsiksi, ja juuri tämä muokkaus luo testaustarpeen. (Scholtz, 2003.)

## 8.4 Työtoveri, kumppani

Usein ihmisen ja robotin välinen yhteistyö määritellään yhdensuuntaisesti siten, että ihminen antaa robotille käskyt, joita robotti tottelee isäntä-renkiperiaatetta mukaillen. Konventionaalisesti robotti vain ottaa käskyjä vastaan, eikä se kysy täsmennyksiä, vaan se ainoastaan luottaa ihmisen päättelykykyyn. Tämä toisaalta sitoo operoijan kyseessä olevaan käyttöjärjestelmään, sen käytettävyyteen sekä sen ymmärtämiseen perinpohjaisesti. Robotin toiminnallisuudelta odotetaan kuitenkin yhä tiiviimmin sitä, että se toimii ihmisen kanssa rinnan, työkumppanina (engl. *peer*) toimien ihmismäisellä tavalla, mikä tekeekin siitä neljännen roolin. (Fong ym., 2005.)

Esimerkiksi portaita kävelevän robotin luomiseksi fyysiset ja designhaasteet huomioon ottaen se voitaneen toteuttaa suhteellisen yksinkertaisesti, mutta heti, kun robotti halutaan ihmisen rinnalle työtoveriksi, on operoijan osattava tulkita robotin kyseisiä toimintatiloja sekä kyettävä ennustamaan seuraavia tarkoituksenmukaisia asennoitumisia (engl. *intentional stance*) peräkkäisiin tehtäviin robotin rationaalisen agentin kannalta. Esimerkiksi operoijan toimiessa kumppanina robotin kanssa, kun molemmat näkevät ikään kuin robotin silmillä kameran avulla tehtävien suorittamista, yhteistyökumppanuus toimii tarkoituksensa mukaisesti. Täten molemmat kumppanit pystyvät yhdessä analysoimaan kyseistä tapahtumaa samanaikaisesti: robotti omilla analysointimethodillaan sekä ihminen omillaan kumpikin vahvistaen roolien onnistumisen astetta. (Bruemmer ym., 2005, 503.)

Kahdensuuntaista työskentelytapaa lähestytäänkin usein siten, että robotti tai autonominen järjestelmä kykenee tavoittelemaan ihmisen määrittämiä päätöspisteitä samalla tavalla kuin inhimillinen työntekijä: vaikkapa välillä kysellen ihmisen antamista käskyistä lisätietoja tai varmistusta, haluaako operoija välttämättä tehdä kyseisen toimenpiteen, jos se vaikuttaa inhimilliseltä virheeltä. (Fong ym., 2005.) Haaste piilee juuri siitä, miten käyttöjärjestelmät mukautuvat kahdenväliseen vuorovaikutukseen ja kumppaneina työskentelemiseen: mikä on sellaista tärkeää ja oleellista informaatiota, jota kumppaneina työskentelevät nimenomaan välittävät toisilleen? Eritoten robotin tieto siitä, mitä tietoa ja milloin se kysyy tai antaa ihmistoimijalle, sekä miten kommunikaationa toimivan käyttöliittymän toiminnallisuudet tukevat vuoropuhelua. (Kaupp, Makarenko & Durrant-Whyte; 2010; 446, 454 & 455.)

Lopulta, jos käyttöliittymän toiminnot ja robotin toiminnallisuus kumppanina eivät toimi käsi kädessä, robotin kanssa yhteistyössä toimiva ihmiskumppani voi harkita alkaa toimivansa valvojan roolissa. Näin palataan asetelmaan, jossa robotti vain vastaanottaa käskyjä eikä vuoropuhelulle ole edellytyksiä. Ei voidakaan olettaa, että robotti ja ihminen toimisivat tasavertaisina työtovereina, vaan robotin antaessa toiminnastaan viite, että työkuppanuus ei toimi, ihminen voi ottaa kontrollin omiin käsiinsä. (Scholtz, 2003.)

## 8.5 Sivustakatsoja

Viidentenä ja viimeisimpänä roolina on sivustakatsoja (engl. *bystander*). Scholtz (2003) alleviivaa, että tämä rooli on ehkäpä haastavin kaikista viidestä roolista, vaikka sivustakatsojalla onkin kaikista vähiten vuorovaikutusta muiden roolien tai robotin kanssa. Sivustakatsojan määrittelemisen voi olla haastavaa, sillä usein sillä tarkoitetaan juurikin jotakin sellaista tahoja, joka on samassa tilassa robotin tai siihen liittyvän käyttöliittymän kanssa (Scholtz, 2003). Esimerkiksi autonomisen konttilaivan ohitse ajava toinen laiva on sivustakatsoja, mutta sillä on kuitenkin rajatut oikeudet: Mitä sivustakatsojan oikeudet ja rooliin kytkeytyvät toiminnot ovat? Minkälaisiin osiin käyttöliittymää sivustakatsojalla on oikeus nähdä? Entä miten robotin tulisi suhtautua tähän sivustakatsojarooliin? Ohitse ajava laiva sivustakatsojana on kuitenkin oikeutettu aavistamaan laivan liikehdintää – niin kuin sitä olisi ohjaamassa ihminen – jotta vaikkapa yhteentörmäyksiltä vältyttäisiin, joten sivustakatsojalla olisi hyvä olla jonkinlainen mahdollisuus ennakoita laivan käyttäytymistä.

Tätä varten on tärkeää pohtia, mikä on sivustakatsojalle suotu käyttöliittymä sekä mitä tietoja robotti voi antaa määrittelemälleen sivustakatsojalle rajoitetun käyttöliittymän kautta; oli se sitten pelkkä hienoinen vinkki laivan suunnanmuutoksesta valon avulla, niin se voi olla elintärkeä vieressä ajavalle sivustakatsojalaivalle, etteivät laivat päädy yhteentörmäykseen. Sivustakatsoja usein olettaakin robotin antavan joitain signaaleja siitä, että tietty mekaaninen liike tai järjestelmän sisäinen toiminto on alkamassa, alkanut ja loppunut. Muutoin robotin kanssa samassa tilassa sijaitseva sivustaseuraaja voi tuntea olonsa epäluottavaiseksi robotin ennakoimattomuudesta johtuen. (Scholtz & Bahrami, 2003; 3216 & 3217.) Autonomisen laivankin on varauduttava näihin ja useisiin muihin vaihteleviin skenaarioihin oikeellisiin rooleihin liittyen, sillä ne voivat olla ratkaisevassa asemassa liittyen vaikkapa myös siihen, miten laiva erottaa sivustakatsojan luotetusta valvojasta ja miten se kommunikoi sivustakatsojan kanssa.

## 9 JÄRJESTELMÄT JA KÄYTTÖLIITTYMÄT ROOLIN VÄLISINÄ TULKKEINA

Autonomiset robotit eivät ole kuitenkaan pelkästään algoritmiensa sekä staattisten taikka dynaamisten parametriensa summa. Tapa, jolla näitä sisäisiä toimintoja kuvastetaan käyttäjälle, on järjestelmän käyttöliittymä, joka on kaiken lähtökohtana käytettäväksi, autonomiseksikin robotiksi sekä linkki käyttäjän ja järjestelmän välille. Käyttöliittymä toimii ikään kuin peilinä järjestelmän sisäiselle toiminnalle, jota käyttäjä voi tulkita helppolukuisesti vailla välttämättä totaalista subjektiosaamista kyseessä olevasta aiheesta tai sen sisäisistä hankalistakin elementeistä. Tuskin useinkaan käyttäjä on tietoinen tai edes haluaa perinpohjaisesti ymmärtää käyttöjärjestelmän kanssa toimivaa tietokantaa, vaan usein kaiken ainoastaan toivotaan toimivan moitteetta ja sillä tavalla, jollaiseksi käyttäjät toiminnot kuvittelevat tai mitä he mieltävät intuitiivisiksi. (Esim. Galitz, 2007.)

### 9.1 Keille järjestelmiä kehitetään?

Juuri siksi ihmisten elämänmuotoihin on pureuduttava kunnollisesti, jotta tiedämme, mitä olemme kehittämässä ja kelle sekä järjestelmistä saadaan niin helpposti ymmärrettäviä ja käytettäviä kuin mahdollista (mm. Saariluoma & Leikas, 2010). Kuten tässäkin tutkimuksessa on usein nostettu esiin, käyttöliittymät ja robotit eivät liene olemassa juurikaan muusta syystä kuin siitä, että ihminen tarvitsee niitä. Ihmisen on mahdollista saavuttaa haluamansa tehtävät automaation avulla, missä ihmisen omaa työpanosta kevennetään tai pienennetään. Tämän tutkimuksen paikalla käyttäjäkunta ovat merikapteenit, joiden työn avainkohtiin on kiinnitettävä erityisen paljon huomiota. Tarkemmat käyttäjätutkimukset merikapteenien päivittäisestä työstä ja sen eksaktista sisällöstä lienevät tarkentuvan myöhempien tutkimuksien myötä, mutta oletettavasti tämänkin tutkimuksen kolme päävaihetta – satamasta poistuminen, häiriö- tai ongelmatilanteisiin vastaaminen avomerellä seilattaessa etäjärjestelmien avulla sekä satamaan saapuminen – ovat ne meriliikenteen prosessien vaiheet, joita ollaan au-

tomatisoimassa ja jotka ovat kolme kriittisintä vaihekokonaisuutta kokonaisuutensa onnistumisen kannalta. Karkeasti sanottuna autonomisia tai pitkälle automatisoituja järjestelmiä voitaneen kehittää niille järjestelmien asiantuntijoille, joiden työtaakkaa halutaan keventää, sekä helpottaakseen päätöksentekoa, missä virhealttiutta halutaan madaltaa vaikkapa puutteellisen ihmistiedon osalta. (Esim. Reason, 2000.)

## 9.2 Graafiset käyttöliittymät

Graafinen käyttöliittymä (engl. *Graphical User Interface, GUI*) on yksi tärkeimmistä, ellei jopa tärkein, autonomisen robotin kanssa työskentelevän henkilön työväline. Siksi sen tulee olla muun muassa käytettävä, monipuolinen sekä selkeästi havaittavissa ja tulkittavissa oleva keino robotin toiminnan tulkintaan ja ohjaamiseen. Käyttöliittymien kautta etänäkin työskentelevä, autonomisen järjestelmän operoija voi saavuttaa olosuhteellista tietoisuutta, joka on tarpeen ja ehkäpä tärkein osio päätöksenteossa. Kun riittävä taso on siinä saavutettu, operoija tai valvoja voi samaisten käyttöliittymien avulla toteuttaa tarpeelliseksi katsomansa operaation, hyväksyä robotin ehdottaman automatisoidun operaation taikka vaihtoehtoisesti katsella järjestelmän syötteistä, onko robotin täysin itsensä valitsema toiminto ollut onnistunut.

Käyttöliittymä on kuitenkin vain yksi, pienehkö osa kokonaisvaltaista ideaa siitä, että ihminen ja robotti voivat toimia saumattomasti yhdessä toistensa kanssa. Käyttöliittymien luomisen pohja makaa sillä fundamentalistisella lähtökohdalla, että robotit ja järjestelmät ovat olemassa lähinnä vain palvellakseen ihmistä. Käyttöliittymien toteuttamiseksi tarvitaan perinpohjaista ymmärtämisen tasoa etenkin suunnittelun (engl. *design*) ongelmakohdista, jotka muodostuvat biologisista, psykologisista ja sosiokulttuurisista aspekteista. (Saariluoma & Leikas, 2010, 22.) Käyttöliittymäsuunnittelussa voidaan luottaa myös usein intuitiivisuuteen, niin kutsuttuihin nyrkkisääntöihin sekä jo yleisesti otaksuttuihin psykologisiin toimintatapoihin: esimerkiksi painiketta alaspäin vaihtaessa jokin pienenee tai liikkuu alaspäin ja ylöspäin taas suurenee tai puolestaan siirtyy ylöspäin, missä luotetaan siihen, että ihminen pitää järjestelmää helposti ymmärrettävänä. Kuitenkaan kaikessa ei voida luottaa yleisesti päteviin tai todettuihin toimintatapoihin, vaan niissäkin saattaa ilmaantua variaatioita, kun esimerkiksi selvitetään, miten ihmisen mieli käsittelisi jonkinlaista käyttöliittymään liittyvää tilannetta. Kun ihmisen mentaaliset mallit on huomioitu, käyttäjä mitä todennäköisimmin pitääkin järjestelmää intuitiivisena, kun se kuvastaa omaa ajattelumaailmaa. (Galitz, 2007, 103.) Pelkästään käyttöliittymän parametrien tai laitteiden ymmärtäminen ei ole tarpeeksi luomaan käytettävyyttä, vaan on tärkeää osata erotella ihmismieleen vaikuttavat psykologiset tekijät omiksi yksiköikseen sekä analysoida niitä. Tähän on olemassa keino: kognitiivinen malli *ACT-R/E*, josta on kerrottu luvussa 6.2.



### 9.3 Käytettävien järjestelmien ja käyttöliittymien suunnittelu etäohjattavissa järjestelmissä

Kuten ihmisen ja tietokoneen välisessä vuorovaikutuksessa (ITKV), siihen tiiviisti ja olennaisesti kytkeytyvässä ihmisen ja robotin välisessä vuorovaikutuksessa (IRV) olennaiseen rooliin nousee käytettävyys. Vaikka ITKV-tutkimus jonkin määritelmän mukaan voikin olla sisältämättä käytettävyyden termin, se sisältyy siihen eittämättä kuitenkin olennaisesti, sillä nämä kaksi käsitettä ovat lähekkäin toisiaan. (Hüttenrauch, 2006, 13 & 14.) Järjestelmän avulla operoija tai valvoja saa tuiki tarpeellista tietoa robotin toiminnasta koko ajan, minkä johdosta päätöksenteko voi olla varmempaa. Kun kyseinen järjestelmä ja sen käyttöliittymienkin käytettävyys ovat moitteettomalla tasolla ja tarkkaan harkittuja, kokonaistehtävän onnistumismahdollisuuksien voitaneen yleisesti katsoa olevan vakaammalla pohjalla kuin käytettäessä järjestelmiä, joissa käytettävyyttä ei ole otettu huomioon. Esimerkiksi hätätilanteen tullessa etäohjattavien käyttöjärjestelmien käytettävyys ja intuitiivisuus ovat keskeisessä asemassa, kun reaktionopeus esimerkiksi törmäyksen ehkäisemiseksi sekä ajan riittämättömyys tarkkaan pohdintaan toimintojen käytöstä ratkaisevat lopputuleman.

Adamides ja muut (2015) tarkastelevat kattavahkossa kirjallisuuskatsauksessa ihmisen ja robotin välistä yhteistyötä käsittelevissä tutkimuksissa esille nousseita käytettävyyden eri аспекteja. He luovat kahdeksanosaisen taksonomian koostuen vuorovaikutuksen ja käyttöliittymien viitteellisistä suunnitteluohjenuorista, joita edelleen voidaan hyödyntää käytettävien järjestelmien kehityksessä etäohjattaviin robotteihin. Esimerkkinä Adamides ja muut (2015) käyttävät etäohjattavaa maataloudessa toimivaa ruiskerobottia, johon taksonomia on heuristisesti sovellettu, mutta sovellettavuus ulottunee yleisesti muihin robotteihin. Seuraavassa on esitelty Adamidesin ja muiden (2015) luoman taksonomian kahdeksan osa-aluetta sekä muutamia niihin liittyviä tutkimuksia:

1. Ensimmäisenä ja mitä luultavimmin kaikkien järjestelmien tulevaisuuden toimivuuden kulmakiviä on alustan arkkitehtuurin ja skaalautuvuuden tärkeys. Toiminnan vaativuuden perusteella mukautuvan käyttöliittymän pohjana on myös eri tilanteisiin sopeutumaan kykenevä arkkitehtuuri, sillä sovellusten eriävät käyttötavat ovat muuttuneet ja tulevat muuttumaan oleellisesti koko ajan (esim. Fritsch ym., 2005). Käyttöliittymän mukautumiskyky onkin lähes suoraan verrannollinen sen alla piilevään ohjelmistoarkkitehtuuriin, joten käyttöliittymien skaalautuvuudessa ja monipuolisuudessa voidaan hyödyntää robotiikasta ja tietojenkäsittelytieteestä periytettyjä oppeja. Näitä evaluoidaan muun muassa erilaisten olosuhteellisen tietoisuuden mittareiden avulla, jotta käyttöliittymien evoluutio ja uudelleenkäyttö toteutuvat parhaalla mahdollisella keinolla. (Scholtz, 2002a.)
2. Toisena kohtana käytettävän käyttöliittymän suunnittelussa on virhelanteiden ehkäisy ja niistä palautuminen. Käyttöliittymän tulee antaa selkokielellä tarpeeksi palautetta käyttäjän toiminnoista, jotta esimerkiksi robotin operoija kykenee huomaamaan toimintojensa seurauksia.

Ja jos ne päätyvät kriittiselle tasolle asti huomautuksista huolimatta, virhetilanteista toipumiselle tulisi myös antaa täydet mahdollisuudet osoittaen virheen sijainnin ja korjausmahdollisuudet. (Esim. Nielsen, 1992; Weiss ym., 2009.) Kuten Adamides ja muutkin (2015) havainnollistavat, esimerkiksi tekstinkäsittelyohjelmassa toiminnon peruminen on vaivatonta ja siihen on olemassa helppo ratkaisu, mutta todellisessa tilanteessa, kuten autonomisen laivan ajaessa merellä, yhdenkin virhe liikkeen peruminen sitä etäohjattaessa voi olla mahdotonta ja sen seuraukset kohtalokkaat.

3. Kolmas alue käyttökelpoisten käyttöliittymien suunnittelussa etäohjattavissa roboteissa on esteettisen, selkeän ja yksinkertaisen ulkoasun tarjoaminen. Koska käyttöliittymä on avain onnistumiseen operoijan ja robotin vuorovaikutukselle, sen tulee tarjota ainoastaan oleellista robotin sensorien muodostamaa informaatiota muun muassa johdonmukaisella ja tehokkaan minimalistisellä tyyllillä. (Adamides ym., 2015.; Devin & Alami, 2016, 326) Tärkeimpien ja useimmiten käytettyjen informaationlähteiden tulosteet tulee sijoittaa käyttöliittymän näkymään siten, että operoija näkee suurimmat päätöksentekoon vaikuttavat elementit nopealla vilkaisulla, jolloin myös olosuhteellisen tietoisuuden taso säilyy mahdollisimman korkealla. Harvemmin tarvittavan tiedon tulisi saada mahdollisimman vähän huomiota käyttöliittymässä, mutta siihenkin tulisi olla selkeät pääsymahdollisuudet. (Mankoff ym., 2003.)
4. Neljäntenä aspektina korostuu oikean informaation esittäminen: oikeassa kontekstissa, oikealla hetkellä sekä oikealla modaliteetilla. Operoijan tulee kyetä käsitellä monesta eri lähteestä, kuten monitoreista ja sensoreista, saapuvaa informaatiota, joten informaation esillepano graafisissa käyttöliittymissä tulee nostaa myös valokeilaan. Kun operoija tavoittaa haluamansa informaation nopeahkosti ja vaivattomasti, voidaan toiminnan ja arvioinnin välistä aikaa vähentää.
5. Robotin tilan tunteminen on äärimmäisen tärkeää toimintojen onnistumisen sekä yhteistyön toimivuuden kannalta, koska ihminen ja robotti toimivat usein eri tiloissa. Siksi viides ominaisuus onkin mahdollisuudet robotin tilojen tarkkailuun, mistä hyötöy niin robotti itse kuin ihminen. (Keyes, Micire, Drury & Yanco; 2010, 183.)
6. Mahdollista hyötösuhteeltaan tehokkaita sekä tuloksia tuottavia toiminnallisuuksia. Ihmisen ja robotin välisessä vuorovaikutuksessa tehokkuutta mitataan ajalla, joka ihmisellä menee tehtävän suorittamiseen, kun taas tuloksia tuottavalla tarkoitetaan sitä, kuinka hyvin tehtävänsuoritus on onnistunut. Molempiin tulisi tarjota helppokäyttöisiä toimintoja käyttöliittymässä, jotta esimerkiksi autonomisen rahtilaivan valvojalla tai operoijalla ei menisi kauaa etsiä toimintoa haluamaansa tehtävään; vaikkapa evätessä tai suomassa tiettyä ominaisuutta etäope- roitavan järjestelmän toiminnallisuuksissa.
7. Robotin ympäristöä tarkkailtaessa tulee olla erilaisia monitoreita olosuhteellisen tietoisuuden saavuttamiseksi. Esimerkiksi kameroiden avulla saadaan tietoa siitä, miksi vaikkapa etäohjattu autonominen rahtilaiva ei liiku meressä kelluvan roskalautan vuoksi. Muita yksin-

kertaisia keinoja lisätä olosuhteellista tietoisuutta ovat esimerkiksi kompassi ja kartta, jotka jo voitaneen olettaa kuuluvan autonomisen tai pitkälle automatisoidun rahtilaivan toimintoihin.

8. Viimeisenä taksonomian kohtana Adamides ja muut (2015) suosittelevat hyödyntämään mentaalaisia malleja sekä kognitiivisia arkkitehtuurreja. Kognitiiviset tekijät vaikuttavat tehtävänsuorituksessa mentaaliin osa-alueisiin, kuten oppimiseen ja suorituskykyyn. Teleoperoitujen, ihmisen ja robotin välisten käyttöjärjestelmien käyttöliittymät tulisi suunnitella siten, että operoijan tai valvojan keskittyminen ohjautuu aina tärkeimpään, kyseiseen robotin suorittamaan tehtävään. Näin useiden eri sensoreiden avulla saatava toisiinsa kietoutuva informaatio auttaa vähentämään kognitiivisen kuormituksen aiheuttamia haittoja, kun käyttäjä keskittyy olennaisimpaan. (Adamides ym., 2015.)

Yanco ja muut (2004) huomauttavat kuitenkin, että vaikka järjestelmän sensorien avulla kerättävä tietoisuus vallitsevista olosuhteista on tärkeää, samaten on sensorien määrä. Yanco ja muut (2004) tutkivat kilpailuasetelmassa tapahtunutta operoijan ja robotin välisen yhteistyön onnistumisen avaimia järjestelmissä, joissa oli sensoreina niin videotointia kuin esimerkiksi infrapuna- tai lämpökameratunnistustakin. He löysivät tutkituista teleoperoiduista ja puoliautonomisista järjestelmistä, että sensorien määrä ei ole suoraan verrattavissa olosuhteellisen tietoisuuden tasoon: sensorien kautta saatu data, joka ei ole yhdistetty tiukasti ja sulautettu toisiinsa, voi olla haitaksi. Jotta ihminen saisi siis oleellista tietoa oleellisilla hetkillä, Yanco ja muut (2004; 146 & 147) ovat osaltaan luoneet ohjenuoria siitä, mitä tulisi ottaa huomioon käyttöliittymäkehityksessä ihmisen ja robotin välisen vuorovaikutuksen luontevaksi toiminnaksi:

- Kartta siitä, missä sijaintikoordinaateissa robotti on ollut kullakin ajanhetkellä, sillä sijaintien muistaminen on kuormittavaa. Autonomisen tai pitkälle automatisoidun rahtilaivan ideassa operoijan tai valvojan saamat tiedot GPS-sijainnista ovat tärkeitä saavuttamaan kokonaiskäsitystä siitä, kuinka esimerkiksi rahtauksessa ja satamissa toimitaan kansainvälisillä vesillä suhteessa toisiin laivoihin (Chen ym., 2016).
- Kognitiivisen kuormituksen vähentämiseksi anna tietoa yhdistettynä siten, että esimerkiksi video on integroitu muun analysoitavan datan kanssa yhdeksi kokonaisuudeksi. Näin käyttäjän ei tarvitse yhdistellä tiedonjyviä toisiinsa.
- Esitä tieto käyttöliittymässä kompaktissa koossa. Kun analysoitavia ja tarkkailtavia näkymiä tai ikkunoita on vähäinen määrä, käyttäjän on yksinkertaista pysyttäytyä yhdessä, isossa näkymässä, jossa näkyy oleellisin tieto, eikä aikaa kulu näkymästä toiseen siirtymisessä.
- Ilmaise robottia, kuten laivaa, ympäröivää tilaa moniulotteisesti. Mikäli käyttöliittymistä ei näe tarpeeksi laaja-alaisesti tai moneen suuntaan, esimerkiksi automatisoitu laiva voi törmätä esteisiin tai sivullisiin veneisiin.
- Auta robottia päättämään, minkäasteista autonomiaa käytetään kyseessä olevassa tilanteessa. Usein saattaa tapahtua niin, että robotin järjestelmä

ei ole varma, kuinka itsenäisesti se saa suorittaa tehtävää, joten ihmisen avustus on välillä tarpeen. Autonominenkin rahtilaiva saattaa haastavan tilanteen tullen olla kykenemätön toimimaan, jolloin sen järjestelmä voisi ilmaista avuntarpeen, johon operoija tai valvoja voi vastata halutulla toimintavaihtoehdolla. Laiva voi siis usein vaihdella autonomisen tai automatisoidun roolin välillä, minkä pitäisi näkyä jollakin tapaa myös käyttöliittymien ulkonäössä. Siksi käyttöjärjestelmistä on toteutettava muuntumiskykyisiä eli dynaamisia.

Kuten huomataan näistäkin kahdesta suuntaviivauksesta järjestelmien kehitystä varten, ohjeistukset käyttöliittymäsuunnitteluun kietoutuvat usein toisiinsa. Pääprioriteettina autonomista rahtilaivaa ajatellen on usein olosuhteellista tietoisuutta vaativien toimenpiteiden onnistuminen, mikä mahdollistetaan sensoreilla, joihin on kytketty tarkasti räätälöidyt käyttöliittymät. Käyttöliittymiä ei kuitenkaan voi luoda ainoastaan käytettävyyden tai yleisesti hyväksi katsottujen toimintatapojen näkökulmasta, vaan käytettävyyden myötä on huomioitava myös tulkittavuus. Käyttöliittymien kehityksessä etäohjatuissa järjestelmissä on siksi loppujen lopuksi kiinnitettävä erityishuomiota ymmärrettävyyteen, millä tavalla mikäkin järjestelmän osa on toteutettu. Ehkäpä pääpiirteittäin autonomisen rahtilaivan vaatimien järjestelmien toteutus on hyvin samantyyppistä kuin itse laivallakin toimiessa tarvittavat toimenpiteet: hyvin samankaltaiset monitorit, sensorit ja aistit olisi hyvä saada mallinnettua keinotekoisesti järjestelmän kautta operoijalle eli valvojalle, jolloin etänä laivan operointi olisi mahdollisimman samankaltaista kuin paikan päälläkin eikä työn luonne muuttuisi olennaisesti. Siksi järjestelmän ymmärrettävyyttä on hyvä pitää silmällä etäohjattavien järjestelmien suunnittelu- ja toteutusvaiheessa, mistä on kerrottu hieman seuraavassa alaluvussa.

#### 9.4 Ymmärrettävyys järjestelmän ja ihmisen välillä

Järjestelmän oletettu syöte saattaa toisinaan poiketa ihmisen käsityksestä sen tilasta. Järjestelmän operoija tai valvoja saattaa kuvitella tietyn laivan antaman ilmoituksen tai syötteen tarkoittavan jotain aivan muuta kuin mikä sen todellinen tarkoitus on, minkä vuoksi laivan kapteeni voi tehdä omat johtopäätöksensä eri tavalla kuin itse paikan päällä laivalla. Esimerkiksi painike, jossa lukee "Vaihda salasana", voidaan tulkita monisyisesti: onko kyseessä komento eli imperatiivi vai valinta, jonka voi tehdä halutessaan? Väärät johtopäätökset (Norman, 1986) lienevätkin yksi suurimmista syistä virheisiin ylipäänsä järjestelmien käytössä, eivätkä niistä johtuvat seuraukset ole autonomisen rahtilaivan kyseessä mitenkään vähäpätöiset, vaan ne voivat johtaa vaikkapa jopa laivan uppoamiseen. Yksikin pieni virhe voi johtaa toiseen virheeseen, jotka yhteisummana mahdollistavat kohtalokkaat seuraukset. Myöskään järjestelmän ilmoittamatta jättäminen epävarmassa tilanteessa ei ole sen parempi vaihtoehto, sillä silloin ihminen voi jäädä pimentoon asian laitoihin vaikuttavista tekijöistä,

jolloin seuraukset voivat olla yhtä kohtalokkaat puutteellisen tiedon myötä. (Esim. Scholtz, 2002b.)

Autonomisen robotin tulee tarpeen tullen toimia ihmisen haluamalla tavalla. Mikäli järjestelmän tai robotin hallinnoija, kuten laivakapteeni, huomaa vikatilanteen tai näkee järjestelmän tuottamista tiedoista, että esimerkiksi rahtilaiva on autonomisesti päättänyt olla toteuttamatta erästä käskyä, sen tulee olla vastaanottavainen kapteenin käskyille. Jotta autonominen robotti voi tietää ihmisen antamien parametrien ja käskyjen merkityksen, niiden tulee olla tarpeeksi yksikäsitteisiä ja tulkinnanvarasta riippumattomia. Kuinka todellisuutta siis voidaan havainnoida symbolisin merkityksin, joita tämän tutkielman kyseessä autonominen laivajärjestelmä kykenee tulkitsemaan sekä tulkkaamaan laivan kapteenille etäjärjestelmien kautta? Tätä ovat sivunneet muun muassa Man ja muut (2018, 241), jotka huomauttavat, että etäohjattavat laivajärjestelmät eivät ole ainoastaan välikappale ja monitori todellisen maailman näkemiseen, vaan järjestelmien on kyettävä tulkitsemaan ympäröivää todellisuutta erinäisten symbolisten merkitysten kautta. Tämän jälkeen järjestelmien on annettava teknisistä tulkinnoista todellisen valehtelemattomat kuvaukset, joista ihminen saa selkeästi selvää. Ihminen on siis sitoutettava ymmärrettävin keinoin itse kokonaisympäristöön, eikä siihen riitä välttämättä yksi monitorointikeino, kuten kamera, vaan yksittäisestä symbolisesta tulkintatavasta muodostuu laajempi ja paljon haasteellisempi käsite, kun koko ympäristöä on keinoitekoisesti – symbolistisia merkityksiä tulkiten ja toisiinsa vertaillen – pystyttävä tarkkailemaan samalla tavalla kuin vaikkapa todellisesti laivan kanneltakin. Tällöin selkeiden käyttöliittymien avulla operoija, tässä tapauksessa laivan kapteeni, pystyy hahmottamaan ympäristön kokonaistilaa yksittäisten selkeästi tulkittavien sensoreiden avulla sekä toimimaan haluamallaan tavalla oikeisiin kohteisiin, kun järjestelmän tulkitsema, symbolien välisistä vuorovaikutuksista koostuva hahmotus toimii. (Man ym., 2018.)

## 9.5 Järjestelmien mukautuminen eri tilanteissa

Järjestelmien on usein mukauduttava ympäristöön, jossa ne toimivat. Järjestelmä voi toisena ajanhetkenä tai sijainnista riippuen alkaa muuttaa itse omia toimintaperiaatteitaan: esimerkiksi yhä enemmän automaattisesti toimivien kuluneuvojen on kyettävä mallintamaan ympäristöään niin valoisassa kuin pimeämmässäkin mutta myös niin ruuhkaisessa kuin vapaammassakin ympäristössä, ja jotta ne kykenevät siihen, niiden mukautumiskyvyn on oltava turvallisuuden takaamiseksi lähes saumaton. Usein käyttäjä voi myös itse päättää, kuinka automaattisesti tai autonomisesti jokin järjestelmä toimii, sekä vaihtaa kontrollia käyttöliittymän kautta, ja näihin ongelmakohtiin ovat ottaneet kantaa muun muassa Dorneich ja muut (2015).

Heidän patenttinsa ottaa huomioon puoliautonomisten järjestelmien käytön mukautumiskyvyn käyttöliittymien osalta, mikä on juuri tämän tutkielman kannalta keskeisimpiä kohtia. Keskeisimpiä kohtia heidän tutkimuksessaan puolestaan on kuvio 5, joka on suomennettu tämän tutkimuksen kuviossa 1.

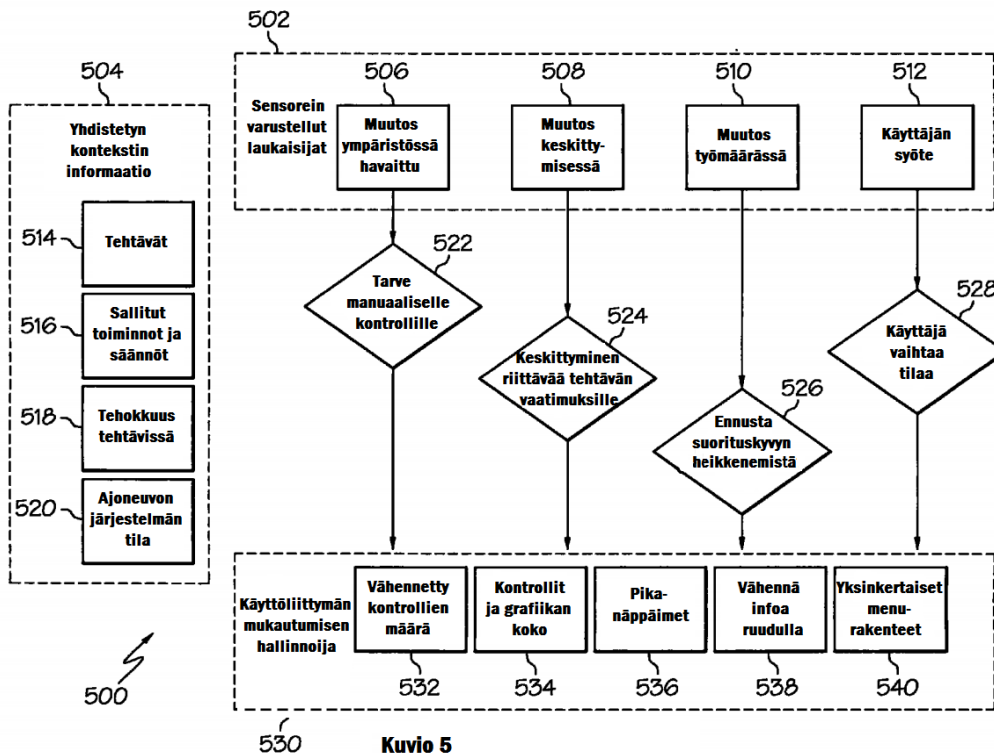
Dorneich ja muut (2015) havainnollistavat kuviolla, minkälaisia kokonaisuuksia, täsmällisempiä ominaisuuksia, toimintoja ja syklejä miehittämättömien, puoliautonomisten ja etäohjattavien järjestelmien käyttöliittymissä tulisi ottaa tarkasteluun. Ensiksi erinäiset muutoksia havaitsevat sensorit (502) huomaavat poikkeavuuksia niille määrättyissä tehtävissä. Tällaiset sensorit voivat olla esimerkiksi ympäristön muutoksia havainnoivia kameroita (506), ihmisen eli tässä tapauksessa kapteenin käytöstä mittaavia sensoreita vahtien muutoksia keskittymisessä tai työmäärässä (508 & 510), taikka käyttäjän syötettä analysoivia sensoreita (512).

Toiseksi nämä sensorit osaltaan antavat lävitseen syötteet, jotka kytkeytyvät yhdistetyksi tietomassaksi koskien kokonaistehtäviä (504; 514). Tieto kerätään yhteen, ja sen jälkeen sitä jalostetaan, miten kustakin ympäristöstä saatu data kietoutuu etäältä käytettävän valvontajärjestelmän vaatimuksiin ja miltä osa-alueelta tieto on kerätty. Yhdistelty tieto tarkastellaan eksaktisti, miten esimerkiksi sallituilla laivajärjestelmän operaatioilla (516) saataisiin mahdollistettua haluttu (osa)tehtävä ja kuinka paljon tehokkuutta järjestelmältä vaaditaan tehtävän suhteen (518). Yhdistetyllä tiedolla voidaan selvittää myös, mikä on kyseessä olevan järjestelmän, kuten laivajärjestelmän, tila.

Taustalla toimivat, päätöksentekoon puuttuvat järjestelmän osat kykenevät hankitun ja yhdistellyn tiedon perusteella muuttamaan juurikin sitä autonomisuuden tilaa tai astetta, jolla järjestelmä operoi. Kuitenkaan itse valvoja tai operoija, tässä tapauksessa laivajärjestelmän kapteeni, ei voi ummistaa silmiään järjestelmän toiminnalta, mistä syystä Dorneich ja muut (2015) ovat ottaneet huomioon ne operaatiot, joilla käyttöliittymän käytettävyys ja ulkoasu voivat muuttua sen mukaan, minkälainen operoinnin tai autonomisuuden taso on kyseisellä hetkellä kyseessä. Järjestelmän havainnoidessa ympäristöään autonomisen päätöksenteon voidaan katsoa olevan hankalasti tehtävissä, joten se voi antaa automaattisesti valvojalle, kuten kapteenille, keinon ottaa ohjat käsiinsä. Tällainen skenaario voi olla vaikkapa satamasta lähteminen tai sinne saapuminen, tai vaikka avomerelläkin tapahtuva odottamaton käänne sääolosuhteissa, liikennöintiajoneuvoissa tai kelluvissa roskalavoissa, joita ei ole osattu ottaa huomioon automaattiseen päätöksentekoon käytetyissä datamalleissa. Tässä tapauksessa on hyvä olla olemassa keinoja, joiden ansiosta esimerkiksi huomattu tarve manuaaliselle kontrollille (522) voidaan mahdollistaa. Samanlaisia huomattavia toimintoja järjestelmän käytettävyyden tueksi ulkoasun muuttamiselle voidaan tehdä myös, kun operoija itse haluaa vaihtaa tilaa manuaalisesta automaattiseksi tai toisinpäin (528), taikka kun järjestelmä havaitsee automaattisesti, että operoijan keskittymisessä tehtävän suhteen tapahtuu muutos (524) tai vaihtoehtoisesti se ennustaa heikkenemän järjestelmän tai käyttäjän suorituskyvyssä (526).

Näin hallitsemalla muutoksia käyttöliittymää pystytään muokkaamaan visuaalisesti, minkä Dorneich ja muut (2015) ovat visualisoineet kuvion 1 alalaidassa näkyvällä käyttöliittymän mukauttamisen hallinnoijalla (530). Käyttöliittymien käytettävyyden kannalta – eritoten tällaisissa etäältä hallittavissa, puoliautonomisesti käyttäytyvissä ajoneuvoissa – käyttöliittymän mukautuminen kulloiseenkin tilanteeseen sopivaksi on lähes välttämätöntä, ja siksi tilanteeseen sopivalla tavalla voidaan varautua muutoksiin. Konkreettisia muutoksia

käyttöliittymän ulkoasussa näissä tilanteissa ovat esimerkiksi kontrolloitavien objektien määrän vähentäminen ruudulla (532), jotta olennaisimmat kontrollit olisivat helpoiten aktivoitavissa; kontrollien ja ylipäänsä ruudulla näkyvien grafiikoiden koko (534), jotta niitä voitaisiin mahduttaa oikea määrä ruudulle parhaiten havaittavissa olevalla tavalla; pikanäppäinten (536) piilotus tai näyttäminen ripeiden toimintojen mahdollistamiseksi, informaation vähentäminen ruudulla (538) oikean informaation korostamiseksi sekä yksinkertaistetut menurakenteet (540) selkeimmän ja nopeimman mahdollisen navigoinnin helpottamiseksi. (Dorneich ym., 2015.)



Kuvio 1. Patenti käyttöliittymän mukauttamisesta puoliautonomisissa, etäohjattavissa ajoneuvoissa (Dorneich ym., 2015).

## 10 MERENKULUN AUTONOMISET ROBOTIT TUTKIMUKSISSA

Autonomisten, miehittämättömien rahtilaivojen katsotaan parantavan useaa eri seikkaa meriliikenteen suhteen. Kun laiva ei vaadi enää itseohjautuvuuden myötä niin paljoa fyysistä henkilökuntaa itse laivalle sen navigointireitin ajaksi, laivaan voidaan lastata aiempaa enemmän rahdattavaa tavaraa, useimmiten kontteja. Kustannukset vähenevät miehistön kannalta sekä samalta nousevat yksittäistä rahtauskertaa katsoen. (Tekniikka&Talous, 2007.) Myös turvallisuuden katsotaan mahdollisesti parantuvan sen myötä, kun laivaa ohjataan yhä enemmän autonomisella tavalla, jotteivät inhimilliset virheet saa jalansijaa enää niin paljoa, vaan järjestelmä voikin varoittaa jo etukäteen mahdollisista kömmähdyksistä ihmisen eli operoijan tai valvojan toimissa. Autonomisten rahtilaivojen ei kuitenkaan odoteta valtaavan markkinoita kansainvälisillä vesillä kuin vasta vuodesta 2025 alkaen, eikä silloinkaan risteilyaluksien henkilökuntaa olla näillä näkymin korvaamassa autonomisuudella, sillä niiden ideana on nimenomaan palvelu, vaan automatisointi ja autonomisointi on tarkoitettu juurikin rahtilaivoille. (Tekniikka&Talous, 2007.) Seuraavissa alaluvuissa on tuotu ilmi muutamia tälle pro gradu -tutkielmalle ominaisimpia tutkimuksia: ensiksi alaluvussa 10.1 tutkimuksia sitoen tämän tutkielman päälöytöjä sekä alaluvussa 10.2 muutama mahdollinen yleistys muista tutkimuksista automatisoituun merenkulkuun.

### 10.1 Tutkimuksia autonomisesta merenkulusta soveltuen löydettyihin asiakokonaisuuksiin

Merenkulku on jo nykypäivänäkin yksi osa-alue, jossa sovelletaan jo jossain määrin autonomisten robottien toimintaa ainakin osittain, tai ainakin voitaisiin soveltaa alustavin tuloksin. Tutkimuksissa – tosin hieman vajavaisissa tuoreuden vuoksi – on löydetty, että laivojen suunniteltu pitkälle automatisointi voi olla mahdollista: Esimerkiksi Moreira ja Soares (2010) tarkkailivat ja loivat tankkerille liikeratoja, joiden mukaan kulkea. Tankkeri kykeni seuraamaan näitä ennalta määrättyjä reittejä suhteellisen hyvin matemaattisten, muun muassa MATLAB-ohjelmalla luotujen mallien pohjalta. Tankkerille annettiin tarpeen



tullen komentoja maista TCP- ja IP-yhteydellä, mutta pääsääntöisesti tankkeri kykeni tahdittamaan nopeuttaan autonomisesti sekä asettamaan myös päämääräänsä itsenäisesti esiohjelmoituja tehtäviä mukailleen. Tutkimuksen hyvinä puolina vuoropuhelu tankkerin ja sen hallinnoijan välillä oli tehokasta, sillä järjestelmät ohjeistuksiin, kontrolleihin ja tarkkailuun olivat eroteltuina omissa yksiköissään, eli käytettävyyden kannalta tällainen järjestelmä on mahdollista toteuttaa. Toisaalta testit tehtiin tyynellä kelillä, joten tarvittaisiin vielä varmistusta epävakaaomissa sääolosuhteissa. (Moreira & Soares, 2010.)

Onneksi ailahtelevimmissakin sääolosuhteissa on tutkittu autonomisten laivojen simuloituja vaikutuksia. Muun muassa Nicolescu ja muut (2007) ovat tätä tutkimusta sitoen löytäneet merkitseviä tutkimustuloksia. He simuloivat laivan etenemistä hankalilla, liikennöidyillä vesialueilla, joissa törmäyksenesto ja onnistunut navigointi ovat tärkeimpiä ominaisuuksia. He huomasivat, että laiva kykeni väistämään toisia laivoja melko onnistuneesti sekä tunnistamaan maa-alueita ja ehkäisemään niihin törmäämistä tehokkaasti. Järjestelmän katsotaan olevan toimiva jopa laivajoukoilla tunnistaen toisia aluksia ja toimien itsenäisesti suhteessa niihin. Kuitenkin ehkäpä mielenkiintoisinta ja eniten olennaista Nicolescun ja muiden (2007) tutkimuksessa on se, että laivan etäohjausjärjestelmän oppiminen oli helpohkoa. Laivan ohjaajat oppivat etäohjauksen ilman suurempia opetussessioita: he omaksuivat kontrollit ripeästi ehkäpä sen ansiosta, että järjestelmä antoi palautetta koko ajan reaaliaikaisesti. Lisäksi kontrollit laivan ohjaamiseen olivat uskottavia. (Nicolescu ym., 2007.)

Kolmantena ja viimeisimpänä esimerkkinä olemassa olevista, etäohjattaviin aluksiin liittyvistä tutkimuksista tämän tutkielman kannalta, kognitiivisesta ja ihmisläheisestä näkökulmasta, on tärkeää nostaa Manin, Lundhin, Porathin ja MacKinnonin (2015) tutkimus. Heidän tutkimuksensa kokooa yhteen tämänkin pro gradu -tutkielman löydöksiä autonomisten, etäohjattavien alusten ohjaamisesta maista käsin. Man ja muut (2015) huomaavat, että aluksen ohjaajan kannalta ilmoitusjärjestelmä esimerkiksi laivan automaation epätoiminnallisuuksista on tärkein väline aluksen hallintointiin, jotta kapteenikin voi suhtauttaa olosuhteellisen tietoisuutensa päätöksiinsä. Käytännössä heidän testaamansa ilmoitusjärjestelmä kategorisoi ilmoitukset väreinä, jotka kapteeni voi halutessaan nähdä lajiteltuina tyyppin tai vakavuuden mukaan sekä auttaa laivan navigointia ongelmakohtissa tavallisin navigointikonstein, kuten karttajärjestelmällä.

Man ja muut (2015) totesivat myös sen tässäkin tutkimuksessa esille nousseen tärkeimmän huomion, että olosuhteellinen tietoisuus on oleellisin osa-alue onnistuneeseen tehtävänsuoritukseen. Ongelmana maista käsin etäohjattavilla järjestelmillä operoivilla merenkulkueksperteillä oli se, että koska he eivät olleet fyysisesti paikan päällä aluksella, he eivät kokeneet samanlaista realistista harmoniaa asioiden syy-seuraussuhteista maissa. Maissa olevat operoijat tarvitsisivat paljon ennakoivamman järjestelmän siihen, että voisivat ennakoita tapahatumia järjestelmien aavistavan teknologian avulla sekä pääsemällä sisään olosuhteelliseen tietoisuuteen etuajassa jo ennen ongelmia. Itse kannelta operoijat puolestaan kokivat, että osasivat ennustaa laivan liikkeitä lähitulevaisuuteen siten, että he saivat olosuhteellista tietoisuutta heti konkreettisesti. (Man ym., 2015, 2680.) Tämä tulos vain vahvistaa sitä tosiasiaa, joka tässäkin pro gradu -

tutkielmassa on tuotu useasti esille, että etäohjattavien järjestelmien käytettävyyteen sekä kognitiivisessa mielessä toteuttamiseen ja olosuhteellisen tietoisuuden hankkimiseksi on käytettävä huomattava määrä aikaa ja osattava ottaa niin monia eri seikkoja huomioon, että tuskin mikään tutkimus on täydellisen kattava siihen, kun emme tunne vielä täydellisesti edes ihmismielen toimintaa, saati laivakapteenien mielenmaailmaa. Kognitiivisen osaamisen huomioon ottaminen kaikissa autonomisen rahtilaivan prosessin vaiheissa on äärimmäisen tärkeää, sillä loppujen lopuksi kapteeni, joka on vain ihminen, on vastuussa tietoisuutensa varassa etäohjattavien järjestelmien kautta koko operaation onnistumisesta.

Vaikka autonomisen rahtilaivan voidaan katsoa olevan tulevaisuudessa monen eri työntekijäroolin helpottava tekijä, lähitulevaisuudessa puolestaan tarvitaan varmasti usean muunkin roolin panosta onnistuneeksi laivan seilaimiseksi satamasta satamaan, kun autonomia ei ole vielä kehittynyt huippuunsa. Man ja muut (2015, 2680) painottavat, kuinka eri roolit, kuten tässäkin tutkielmassa on mainittu luvussa 8, ovat yhtä olennaisia kokonaisuonnistumiselle kuin pelkkä kapteenin tietoisuus tai päätökset. Tieto varsinkin olosuhteellisesta tietoisuudesta voi kulkea tiedon vastaanottajalta tai välittäjältä eli operoijalta valvojalle, joka toimii ikään kuin kapellimestarina tai koordinoijana; valvojalta tekniselle konsultille, sekä lopulta teknikolta itse kapteenille, jonka oletetaan tekevän päätös monen eri roolin kautta tullee tiedolle. Man ja muut (2015) huomioivat siksi, että mikäli missä tahansa vaiheessa tapahtuu esimerkiksi väärinarviointi tai tietoa jää uupumaan, tieto voi muuttua ja muuttaa muotoaan olennaisesti, mistä syystä tehdyt johtopäätökset voivat olla vääriä kapteenin toimesta, joka on jäänyt tietoisuuden ulkopuolelle. Tämän pohjalta on erittäin huomattavaa tarkastella järjestelmää niiden kaikkien roolien osalta, jotka siihen vaikuttavat: automaatiota ja tekoälyä voi olla hyvä kehittää muidenkin kuin pelkän kapteenin toiminnan osalta. Kuinka montaa roolia itse kapteeni lopulta edes edustaa? (Man ym., 2015.)

Autonomisen rahtilaivan prosesseissa yksilölliset erot on otettava huomioon eri rooleilla, varsinkin jos kapteeni on kaikkia näitä. Kaiken kukkuraksi kapteeni ei ole myöskään aina sama henkilö joka kerralla, kun laiva seilaa samaa, tuttua reittiä, jolla laivakaan ei ole aina samaa tuttua mallia. Luvussa 11 kerrotaan väärinymmärryksistä lentoliikenteessä, kuinka esimerkkinä lentokoneen automaattisen lentojärjestelmän koettiin olevan samanlainen kaikissa lentokoneissa, mutta koneen malli olikin erilainen eikä ilmoituksia vaaratilanteista otettu todesta lentokapteenin toimesta. Samaten kuin laivan tyyppi voi vaihtua ja järjestelmän on sulauduttava kyseessä olevan laivan toimintoihin ja ominaisuuksiin, etäohjattavan järjestelmän valvojien tai operoijien ominaisuudet voivat vaihdella matkojen välillä riippuen, kuka kapteeni on kulloinkin roolissa. Tämä luo vaatimuksen järjestelmille, joissa on otettu yleisimmät kapteenien yhteiset ominaisuudet, sekä jotka kykenevät huomaamaan aina kulloisenkin kapteenin tyypillisimmät toimintatavat muuttaen esimerkiksi ulkoasuaan ja toiminnallisuuksiaan vaikkapa alaluvussa 9.5 mainituissa kohdissa. Näin järjestelmät kykenisivät antamaan esimerkiksi Manin ja muiden (2015) mainitsemin keinoin ilmoituksia siten, että samalla tavalla kuin kapteeni hankkisi olosuhteellista tietoisuutta yksilöllisesti laivan kannella, laivan automaattiset järjes-

telmät olosuhteelliseen tietoisuuteen kuvastaisivat työvuorossa olevan kapteenin toimintatapoja sekä edesauttaisivat ilmoituksilla niistä havaituista ennustemerkeistä, joiden perusteella maista käsin toimiva kapteeni voi luoda päätelmiä tulevaisuuden tapahtumista.

Fyysiset tuntemukset ympäristöllisistä tekijöistä laivalla paikalla ollessa ovat tärkeitä huomioida ennustukseen liittyen. Värinä on yksi olennaisimpia merkittäviä muutoksia laivan seilattaessa: kun kapteeni on fyysisesti laivan komentosillalla ja tuntee laivan liikehtivän oudosti, tämä voi vaikkapa tehdä päätelmänsä väärästä kulkusuunnasta ja muuttaa laivan kurssia. Ennakointiin tarvitaan aikaa, koska laiva ei käänny yhtä nopeasti kuin esimerkiksi samalla tavalla itseohjautuva auto. (Yle, 2017.) Pelkkien käyttöjärjestelmien avulla tällaisia fyysisiä, haptisia palautteita on haastavampaa luoda samankaltaisesti. Siksi – ja ylipäänsä koko siirtymävaiheen helpottamiseksi laivan komentokeskuksesta etäohjattavuuteen maista käsin – laivan etäohjaus tapahtunee juurikin samankaltaisesti kuin laivan kannelta mutta vain ikään kuin simuloitussa tilassa, kuten kuviossa 2 on esitetty ja Rolls-Royce on jo tehnyt (YBW, 2017). Rolls-Roycen kontekstissa, johon tämä kirjallisuuskatsaus liittyy, autonomisen rahtilaivan toimintoja on tutkittu jo jonkin verran, kuten Degerholmin (2017) ja Levanderin (2017) tutkimuksissa, mutta eritoten järjestelmien käytettävyyden sekä kapteenin ja robotin välisen vuorovaikutuksen tutkimus on vielä vajavaista.



Kuvio 2. Rolls-Royce esittelemässä ensimmäistä kaupallista etäohjattavaa alustaan (YBW, 2017).

## 10.2 Sovellutuksia aiemmista tutkimuksista autonomisiin laivoihin

Parasuramanin ja muiden (2005) tutkimustulosta soveltaen voidaan ajatella, että ihmisen ohjatessa vain yhtä tai muutamaa robottia kerrallaan tämä onnistuu fokuksomaan työkapasiteettinsa erinomaisesti tehtäväänsä. Tällöin robotille jää itselleen enemmän kapasiteettia keskittyä omaan työtehtäväänsä eikä sillä kulu kallisarvoista aikaa analysoimaan ja raportoimaan ihmisen herpaantumista tai virheliikkeitä valvojan tai operoijan roolissa työskentelevälle. Tässäkin mielessä laivaliikenteen robotisaatio ja autonomisaatio on oiva alue käyttää itseohjautuvuutta. Ehkäpä harvoin kuitenkaan laivan kapteenin tulisi ohjata montaa eri laivaa kerrallaan, sillä silloin fokus voisi herpaantua olennaisesti. Kapteenin valvoessa ja ohjaessa ainoastaan yhtä laivaa kerrallaan vuoropuhelu yhden laivan järjestelmän ja kapteenin kanssa lienee parhaimmillaan, ja kapteeni voi tarvittaessa keskittyä vaikkapa antamaan delegoitavia tehtäviä laivan järjestelmälle niin halutessaan, vaikka teoriassa etäohjattavuus voisikin mahdollistaa samanaikaisesti monen rahtilaivan toiminnan tarkkailun.

Miller ja Parasuraman (2007) osaltaan ovat luoneet mallin, jolla voidaan palastella kukin suoritettava tehtävä osatehtäviin sekä suoda jokaiselle osatehtävälle oma vuorovaikutuskeinonsa niin järjestelmän automatisoinnin kuin jokaisen osatehtävänkin suhteen. Heidän malliaan mukaillen voidaan sanoa, että kapteenin tarkastellessa yhden järjestelmän muodostamia, pienempiä tehtäväkokonaisuuksia kapteeni voi kokea olevansa paremmin kontrollissa koko laivankulkumatkan ajan. Kontrollintunne ja delegointimahdollisuudet järjestelmän ja valvojan tai operoijan välillä voivat myös pienentää stressin määrää ja turhaantumista sekä kohentaa käyttäjän hyväksyntää järjestelmää kohtaan. (Miller & Parasuraman, 2007.)

Nicolescun ja muiden (2007) tulosta ajatellen voidaan pohtia, että kun luodaan kapteeneille ja muulle laivan henkilökunnalle toimivia järjestelmiä ja käyttöliittymiä etäohjattaviksi, kontrollien kuvastaminen oikean laivan kontroleja on ratkaisevassa roolissa. Esimerkiksi etäohjattava ratin tärinä voi olla tärkeä kontrollintunne kapteenille, josta voidaan päätellä laivan kulkuun vaikuttavia seikkoja. Ehkä ei ole järkevää luoda kontroleja täysin vastakkaisella tavalla kuin oikeassa laivassa, vaan sen kannalta, että etäohjattavat järjestelmät ovat helposti omaksuttavissa ja käytettävissä, ne on hyvä tehdä kuvastamaan oikeita laivan kontroleja ja järjestelmiä. Tämä vaatii vahvistusta käyttäjäkyselyin ja -tutkimuksin oikeilta laivakapteeneilta, millä tavalla etäohjattavan laivan järjestelmät ja kontrollit tulee toteuttaa niin, että ne tuntuvat oikeilta ja intuitiivisilta etäohjaajalle. Tässä vaiheessa, kun olemme luomassa uutta, voimme järjestelmiä luodessamme mukailla hyväksi todettuja navigointimenetelmiä mutta samalla murtaa jo olemassa olevien toimintatapojen kaavoja.

Intuitiivisuuden ja luonnollisen vaikutuksen luominen etäohjattavissa roboteissa voi olla vankalla pohjalla siinä mielessä, miten kapteenit ja muut etäohjattaviin laivajärjestelmiin liittyvät roolit suhtautuvat automatisointiin taikka robotisointiin. Etäohjattavat robotit eivät välttämättä luo yhtä paljon epäluotamusta tai jopa pelkoa ihmisille kuin fyysiset, läsnä olevat robotit: Woodsin,

Waltersin, Koayn ja Dautenhahnin (2006) mukaan huomattiin, kuinka fyysisesti samassa tilassa sijaitsevat robotit aiheuttivat käyttäjilleen epämukavan olotilan, joten merenkulussa etänä toimiva laiva ei luo tätä epäkohtaa, kun niin kutsuttu robottilaiva ei ole uhkaamassa operoijan omaa turvallisuutta. Näin koko onnistumisprosessi satamasta satamaan navigoinnissakin saattaa onnistua paremmin, kun kapteeni voi turvallisesti mielin tarkkailla ja operoida järjestelmän kautta kaukana toimivan laivan toimintaa eikä pelkotila vaikuta toimintaan. Näin robotti on naamioitu ei-robottimaiseksi: se on integroitu jokapäiväiseen todellisuuteemme, kuten vaikka nykyaikana älypuhelimistammekin löytyvät, äänikomennoin toimivat robotit, joita ei edes välttämättä mielletä roboteiksi. Tietenkin vertailukelpoisia tutkimuksia itse laivalta ohjaamisen ja etänä maista ohjaamisen välillä on tehtävä erikseen.

### 10.3 Ympäristölliset attribuutit autonomisessa laivassa

Ympäristö ja konteksti, jossa hankittu tieto on saavutettu, on tärkeää tunnistaa. Data ei ole olemassa ainoastaan massana, josta voidaan erottaa faktoja, vaan sen alkuperä on pystyttävä tiedostamaan sen tarkemman analysoinnin ja liitettävyyden kannalta toisiin datoihin tai järjestelmiin. (Esim. Compton & Jansen, 1990.) Huang ja muut (2005, 32–35) ovatkin kategorisoineet ympäristöön liittyvät attribuutit autonomisen toiminnan vakauden vahvistamiseksi miehittämättömissä järjestelmissä (engl. *unmanned systems, UMS*). Eritoten merenkulussa nämä faktorit on syytä tarkastella yksityiskohtaisesti, sillä autonomisen robotin logiikassa on otettava huomioon esimerkiksi sään epävakaudesta, jolloin laivan on tehtävä itsenäisesti päätöksiä muun muassa suunnanvaihdossa sekä sulauttaa toimintansa etäohjattaviin toimintoihin käytetyiksi järjestelmiksi ja käytettävyyden kannalta laivakapteenille. Yksittäisistä parametrikokonaisuuksista laivan automaattiseen tai autonomiseen päämäärien asettamiseen sekä navigointiin ovat varmasti tärkeimmistä päästä ympäristölliset, ulkoiset seikat. Siksi seuraavassa onkin arvioitu, mitkä näistä tekijöistä liittyvät nimenomaan merenkulkuun ja autonomisten laivojen järjestelmien toimintakykyyn.

Ensimmäisenä Huang ja muut (2005, 32) ovat määritelleet staattisen ympäristön eli navigaatioalueen, joka ei juurikaan muutu. Merenkulussa tällainen on muun muassa meren syvyys, eli laivan on tiedettävä, voiko se vaikkapa kuljettaa painavan lastin matalissa kanavakohdissa; toisin sanoen onko sen syväys tarpeeksi suuri. Laivan on siis tunnettava oma staattinen painolastinsa sekä suhteellisen samana pysyvä veden syvyys sekä arvioitava, voiko se navigoida kyseisestä karttapisteestä toiseen. Suhteellisen staattinen on myös siis koko kulku-ympäristö, jolla laiva liikkuu, eli vesi.

Jotta voidaan määrittää vaihtelevat, dynaamiset ympäristötekijät, on tunnettava ensiksi staattiset tekijät. Tiedetään että vesi elementtinä ja maastotyyppinä on suhteellisen paikoillaan pysyvä eivätkä merenkulkureitit juuri vaihtele sen materiaalin muuttumisen vuoksi, mutta se voi olla dynaaminenkin ja aiheuttaa monia konfliktitilanteita laivan järjestelmän autonomisuudessa. Veden muuttaessa liikehtelemistään myrskyn aikana laivan järjestelmän on kyettävä

tekemään ratkaisuja omista päämääristään ja siitä, miten ne saavutetaan. (Huang ym., 2005, 33.) Päätöksenteossa on täten minimoitava pääte pisteiden saavuttamisen riskit: Milloin navigoida suoraan myrskyn läpi, eli milloin sään epävakaus on tarpeeksi minimaalinen tuottaakseen mahdollisimman vähän haittaa lastin saattamiseksi perille? Milloin taas myrsky on ylitsepääsemätön, jolloin on kierrettävä se, ja mitkä vaihtoehtoiset reitit silloin ovat? Milloin puolestaan autonomisen laivajärjestelmän on annettava ilmoitus ihmiselle siitä, ettei se kykene tekemään päätöstä itse? Tällöin tarvitaan myös käyttöjärjestelmän suhteen muutoksia, mistä on kerrottu esimerkiksi alaluvussa 9.5.

Yleisimmät muutokset sääolosuhteista ovatkin yksi suurimmista merenkulkuun vaikuttavista dynaamisista muuttujista. Muita säähän liittyviä olosuhteita ovat

- tuulen nopeus (mitä reittiä pitkin päämäärä saavutetaan nopeiten; lyhyttä mutta hidasta, vaiko kenties pidempää mutta nopeampaa)
- ilmanpaine (kuinka lastina oleva materiaali reagoi näihin)
- näkyvyys ja kirkkaus
- muut harvinaisemmat sään ääri-ilmiöt (Huang ym., 2005, 34).

Jokaiseen näistä muuttuvista tekijöistä laivan on osattava reagoida tilanteen tullen, mikä luo haasteen autonomisten robottien kehittäjille, eikä harvinaisempiakaan sään muotoja voida täysin jättää ulkopuolelle järjestelmien toimintalogiikasta. Robotin tulee koko ajan myös antaa selkeää ulosantia toiminnastaan, seuraavista liikkeistä sekä vallitsevista parametreista, jotka vaikuttavat sen toimintaan, jotta teknikko voi tarvittaessa muuttaa järjestelmän sisäisiä algoritmeja tai kapteeni laivan itsensä täysin tai osin asettamia osatehtäviä ja päämääriä.

Juuri tällaisiin odottamattomiin tapahtumiin ympäristössä (engl. *Autonomous Response to Unexpected Events; ARTUE*) on osattava varautua robotin ja sen järjestelmän kehityksen yhteydessä. Esimerkiksi tavoitteellisissa autonomisissa järjestelmissä (engl. *Goal-Driven Autonomy, GDA*) – jotka pyrkivät saavuttamaan tietyn asetetun maalin – tällainen ympäristöllisten, muuttuvien tekijöiden huomioiminen on onnistuneesti toteutettu jollakin tasolla jo armeijan laivaston järjestelmien funktionaalisuudessa, johon liittyvässä autonomisuudessa on otettu huomioon muiden muassa niin jäävuorten kuin lähistöllä sijaitsevien laivojen koordinaatit. Samaisessa autonomisuudessa on tarkasteltu myös sitä vaihtoehtoa, mikäli myrsky iskee kesken matkan: laivalle on annettu tietty uusi päämäärä hakeutua suojaan, mikäli myrskyn arvioidut parametrit ovat suurta riskiluokkaa, jolloin siitä täytyy ilmoittaa kapteenillekin. (Molineaux, Klenk & Aha; 2010, 1552–1554.)

## 11 VIRHEISTÄ OPPIMINEN MUISTA LIIKENNÖINTITAVOISTA

Laivaliikenne, merenkulku, on siinäkin mielessä turvallinen alue kokeilla tällaisten autonomisten robottien toimintaa, että esimerkiksi rahtilaivan kyydissä ei ole yhtä montaa ihmishenkeä kuin risteilylaivan tai lentokoneen. Samalla tavalla kuin lentoliikennekin on automatisoitu suurimmaksi osaksi eli ihminen ei välttämättä edes tiedosta robotin ohjaavan lentoa, samaa voitaisiin soveltaa laivaliikenteeseenkin. Lentoliikenteen tekoälyn epäkohdista voitaneen ottaa oppia autonomista laivaliikennettä rakentaessa, kun olemme ehkäpä kypsempiä nyt tekniikan ja teknologian kanssa lentoliikenteen suhteen.

Mutta autonominen rahtilaivaliikenne ei ole yhtä autuutta. Kun siirrytään yhä enemmän autonomiseen tai pitkälle automatisoituun rahtilaivaliikenteeseen, riskien määrä kasvaa. Koska ihminen päihittää robotin vielä useilla eri osa-alueilla, kuten sosiaalisen vuorovaikutuksen saralla kommunikoinnissa, on osattava tunnistaa alueet, joita haluamme automatisoida ja autonomisoida ensimmäiseksi ja mahdollisella keinolla. Esimerkiksi laivaliikenteessä yksi tällaisia osa-alueita voi olla pitkäkestoinen, yhtäjaksoinen liike, jossa ei välttämättä havaita muutoksia yleensä niin usein kuin eksaktimmeissa tilanteissa, ja tällainen voi olla vaikkapa seilaaminen avomerellä laivan ollessa kyseessä. Samanlaista automatisointia yhtäjaksoisessa liikkeessä on käytetty jo lentoliikenteessä pidemmän aikaa: kun lentokone on ilmassa, sitä voi ohjailta pääosin järjestelmä lennon ajan (esim. Nelson, 1998). Lentoliikenteen voidaankin ehkäpä katsoa olevan hieman samantyyppinen kuin laivaliikenteen: kuten lentokoneessa nousu ja lasku ovat kriittisimmät vaiheet lentomatkan onnistumisen suhteen ja ilmassa matkanteko suhteellisen suoraviivaista, laivaliikenteessäkin satamasta lähtö ja sinne saapuminen ovat varmasti haastavimmat matkan osa-alueet sekä avomerellä liikkuminen suoraviivaisempaa ja automatisoitavissa olevampaa. Parhaimmassa tapauksessa lentokapteenin tulee ottaa kantaa lennon vaiheisiin ainoastaan epävarmoissa tilanteissa lentokoneen järjestelmän siitä ilmoittaessa, ja samanlaista automatisointia, nyt jopa autonomisointia, olisi tulevaisuudessa käytössä enemmänkin laivaliikenteessä. Kuitenkaan lentoliikenne ei ole vielä niin pitkälle mennyt, että lentokapteeni toimisi täysin etäältä, joten laivaliikenne on siinä edellä aikaansa rahtiliikenteen suhteen. Tietenkään näiden kahden lii-

kennöintitavan liikennöintialustat ovat erilaiset eikä yleistys ole selkeää, mutta voitaneen joka tapauksessa noteerata, mikä lentoliikenteessä on onnistunut ja mikä puolestaan ei ja mistä nimenomaan virheet ovat johtuneet välttämättä samantyyppisiä kompastuksia laivaliikenteessä.

Endsley (1995a) tarkastelee kolmea eri virhetasoa lentokoneonnettomuuksissa, jotka liittyvät olennaisesti olosuhteelliseen tietoisuuteen: **virhe nähdä tilanne oikeellisesti**. Tämä on käynyt lentoliikenteessä esimerkiksi silloin, kun valvoja tai operoija ei ole saanut olosuhteellista tietoa johtuen joko suunnitteluvaiheessa huomiotta jääneestä järjestelmävirheestä tai epäkohdasta kommunikoinnissa. Tällainen tilanne on käynyt muun muassa silloin, kun lentokoneessa on jäänyt jokin ovi väärään asentoon, joko auki tai kiinni, eikä siitä ole tullut ilmoitusta. Toisenlainen vakava virhe tapahtui silloin, kun polttoaineen ollessa vasta kriittisellä tasolla siitä tuli ilmoitus, ja lentokone ehti jo syöksyä alas. Toisin sanoen data on ollut saatavilla, mutta pääsy siihen on ollut erinäisistä syistä rajoittunutta. (Endsley, 1995a; 289 & 290.) Aivan samantyyppisiä ongelmia voi tulla vastaan laivaliikenteessäkin: laivasta loppuu polttoaine kesken matkan siitä ilmoittamatta, tai järjestelmä ei vaikkapa ilmoita satamasta poistuttaessa, että alus onkin vielä ankkuroituna paikoilleen. Toisena syynä virheiden huomaamisen puuttumiseen on suuri työkuorma: mitä enemmän huomioitavia asioita operoijalla tai valvojalla, kuten kapteenilla on, sitä enemmän tältä voi jäädä huomaamatta tärkeitä seikkoja ympäristössä aiheuttaen unohtamisen (Endsley, 1995a, 290). Tässä ongelmatilanteessa voitaisiin ehkäpä soveltaa alaluvussa 9.5 mainittua käyttöliittymän mukauttamisen hallinnoijaa: kun huomataan, että ympäristössä on useita tarkkailtavia muuttujia eikä kapteeni pysty hallitsemaan kaikkia, kapteenille voidaan suoda kontrollien avulla mahdollisuus vaihtaa operointitapa manuaalisesta autonomisempaan tasoon vaikkapa niiden tehtävien suhteen, jotka järjestelmä on katsonut vähemmän kriittisiksi ja jotka automaattisesti kytetään ratkaisemaan; antaen kapteenille vähemmän kontrolloitavia tehtäviä kerralla vähentäen työkuormaa sekä ehkäisten vaaratilanteiden syntymistä.

Toinen taso on **virhe hahmottaa tilanne**. Siitä huolimatta, että tapahtumiin vaikuttavat tekijät nähdään ja tilanne päästään havaitsemaan, voi käydä niin, että tietoa ei osata jäsenellä ja ymmärtää oikein. Sen merkitys saattaa jäädä ontumaan, sekä sen syy- ja seuraussuhteista ei välttämättä päästä jyvälle, mikä voi johtua heikosta mielentilasta tai mentaalista mallista. (Endsley, 1995a, 290.) Eräässä tapauksessa lentokapteeni hylkäsi ilmoituksen jäänmuodostumisesta, sillä oletti, ettei kyseisellä lentokonemallilla oikeasti ollut jäämisiongelmaa. Yksi ongelmallisuus piili lento-onnettomuuksissa siinä, että lentäjät ollettivat toisenmallisen lentokoneen olevan samankaltainen ohjattavan koneen kanssa, mutta kyseessä oleva ohjattava lentokone olikin huomattavasti erilainen verrattuna aiempiin vastaaviin koneisiin, joihin lentäjät olivat tottuneet. (Endsley, 1995a, 290.) Samalla tavalla ongelmia laivaliikenteessä voi ilmetä: kun laivan kapteeni ohjaa laivaa etäältä, on erittäin tärkeää tiedostaa ja ymmärtää laivan ominaisuudet peilautuen järjestelmän syötteisiin: joskus järjestelmävirheiden myötä käyttäjälle näkyvät ilmoitukset tai syötteet voivat olla vääriä ja niihin voi olla reagoimatta, mutta olettaen suurimmassa osassa tapauksista, että järjestelmä on rakennettu oikein ja kapteenin mentaalinen malli kuvastaa järjes-



telmän sisäistä ohjelmoinnillista mallia, sekä järjestelmä että kapteeni voivat tehdä yhtenäisiä päätelmiä ja päätöksiä asianlaidoista. Myös kyseenalaistaminen järjestelmän toiminnallisuudesta voi olla tarpeen: kun kapteeni huomaa, että jokin syöte tai toiminnallisuus on väärin, se voidaan korjata jatkon kannalta, mutta jos kaikki järjestelmän toiminnot pidetään totena, sekin voi virhetilanteessa johtaa monen seikan kautta vakaviin häiriötilanteisiin. Kuitenkaan järjestelmän toiminta ei ole mikään erehtymätön, absoluuttinen fakta, vaan sekin on vain ikään kuin tietyllä tavalla toimiva robotti. Endsley (1995a, 291) nostaa esiin sellaisenkin seikan, jossa sumussa järjestelmä ei kyennyt huomioimaan toisen lentokoneen olemassaoloa lainkaan. Sama ongelmallisuus piilee myös laivaliikenteessä: vaikka kapteenilla voi olla kaikki olemassa oleva, järjestelmän kautta saavutettu tietoisuus, sekään tietoisuus ei kerro aina kaikkea, kuten sankassa sumussa seilaavaa toista alusta sen valoistakaan huolimatta.

Kolmas ja viimeinen olosuhteellisen tietoisuuden virhetaso on **tilanteen ennustus tulevaisuuteen**. Järjestelmien operoijat tai valvojat voivat olla täysin tietoisia ymmärtäen tapahtuvien tapahtumien syyt ja yksityisseikat, mutta ennustuskyky tietyn ajanhetken päähän voi olla olematonta tai heikkoa. Esimerkiksi lennolla mukana olevan henkilökunnan tehtävänä oli ennustaa polttoaineen riittävyys, mutta se ei sujunut täysin ongelmitta ennustuksien jääden joko kokonaan puuttumatta tai niiden aliarvioiminen, mistä seurasi äkillinen kriittinen tila polttoaineessa jo aiemmin mainitun järjestelmän ilmoittamisvian yhteis-tekijänä. Yleisimmin ihmisen mentaaliset mallit ennustuskykyyn ovat heikot, ja sitä tehostaakseen järjestelmät voivat nykypäivänä yhä enemmän tehdä tilastollisia ennustuksia kunkin mahdollisen tilanteen todennäköisyydestä ja tapahtumisajankohdasta laivaliikennettäkin ajatellen (Endsley, 1995a, 291.) Näin laivan kapteenin ei tarvitse itse käyttää kallisarvoista olosuhteellisen tietoisuuden keräämiseen ja perinpohjaiseen ymmärtämiseen tarvittavaa aikaa siihen, että ennustaa yksin kuormittuneella aivokapasiteetilla lähitulevaisuuteen vaikuttavia tekijöitä ja tapahtumia, vaan ihmisen toimintaa on hyvä tukea tekoälyllä, mistä on kerrottu enemmän esimerkiksi alaluvussa 6.3.

## 12 JOHTOPÄÄTÖKSET JA POHDINTAA

Kirjallisuuskatsauksen pohjalta nousee niin selkeitä asiakokonaisuuksia kuin epäselvempiäkin, kirjoitushetkellä vailla kysymyksiä olevia kokonaisuuksia. Kuten tutkimuksen alussa huomataan, esimerkiksi vanhemmat vuosikymmenten takaiset huolenaiheet ovat ajankohtaisia nykypäivänäkin: Mitä robotilla tarkoitetaan? Mitkä ovat robotin ominaisuudet? Ottavatko robotit ihmisistä vallan? Näitäkin peruskysymyksiä on tarkasteltu liittyen toisiinsa varsinaisen tutkimuskysymyksen lomassa: miten luomme käytettäviä järjestelmiä etäohjattavalle laivalle eritoten kapteenin kannalta, ja mitä asioita on otettava huomioon kaikkiaan tällaisia prosesseja suunnitellessa, luodessa ja testatessa luodaksemme vakaat, käyttäjäystävälliset prosessit?

Kirjallisuuskatsauksen perusteella saadaan vastauksia näihin kysymyksiin. Kronologisesti edettynä – niin tutkielman teemojen kuin historiallisestikin – aloittaen siitä, mitä robotilla ylipäänsä tarkoitetaan rahtilaivan kontekstissa. On hyvä purkaa mystiikkaa sanan *robotti* ympäriltä siinä mielessä, että robotti tuskin on minkään kuvittelemamme robotin ulkomuodossa, ihmisruumismaisessa kehikossa lainehtiva metallimöykky, vaan se on yhä virtaviivaisesti seilaava laiva, jonka järjestelmät ovat vain entistä itsenäisempiä. Laivan automaattiset järjestelmät on ihannemaailmassa kytketty fyysisiin, laivaa eteenpäin vieviin moottoreihin ja muihin apuvälineisiin, jotka esimerkiksi kääntävät laivaa. Voi olla, että vaaditaan teknologisia innovaatioita toteuttamaan tätä käytännössä. Etäältä operoiva kapteeni puolestaan ottaa vuorovaikutukseen tarvittavilla järjestelmillä siihen yhteyttä tarvittaessa, ja laiva myös vuorovaikuttaa automaattisesti operoijaan päin sekä koko ajan että kriittisillä hetkillä erityisellä huomiolla. Vuorovaikutus tapahtuu autonomisten järjestelmien pääominaisuuksien periaatteiden mukaisesti sekä mahdollisesti niin selkeäkielisesti kuin mahdollista välttääksemme kommunikaatiokatkot. Toinen vaihtoehto on luoda kieli sillä tavalla, jolla kapteenit ovat tottuneet paikan päälläkin komentoja tekemään ja ottamaan vastaan, jotta vuorovaikutus on autenttisinta eikä poikkea itse paikan päällä tapahtuvasta operoinnista.

Toiseksi tutkimuskysymyksiin nojaten on tehtävä selväksi, kuinka autonomiseksi eli automaattiseksi järjestelmämme haluamme. Osa-alueet, jotka automatisoimme, on päätettävä tarkasti etukäteen. Näinkin tuoreessa aiheessa

kuin autonominen rahtilaiva ei ole varmastikaan perusteltua lähteä sillä oletuksella, että kaikki on automaattista heti alkuunsa, vaan on valittava tarkasti ne kohteet, joista haluamme aloittaa. Näitä voivat olla juurikin laivan itse päättämät kulkureitit ja järjestelmät, joilla automatisoimme ja valvomme prosesseja. Tämänäntyypistä automaatiota ja päättelyä on olemassa tutkimuksissakin jo jonkin verran hyödynnettävänä. Itse kohteiden valinnan jälkeen on lisäksi päätettävä se taso, jolle kohteet automatisoidaan. Kirjallisuuskatsaus antaa luvussa 5 hyvin ne vaihtoehdot, joihin tällaisessakin autonominen rahtilaivan prosesseissa voidaan pyrkiä. Mutta sitä varten tarvitsemme keinot luoda autonomiaa.

Todellinen haaste lienee siis piilevän autonomiaa varten luotavissa älykässä malleissa, jotka on luotava siten kuin ihminenkin ajattelisi. Vaadimme kaikille osa-alueille, joissa ihmisen mieltä korvataan, sitä tarkkaa tietoa, miten esimerkiksi kapteeni toimisi kyseisissä tilanteissa. Kuinka saamme kerättyä tietoisuutta kapteenien toiminnasta käytännössä, ja miten sitä voitaisiin ammentaa robotille jonkinnäköisen mallin kautta? Erilaisia lähestymiskeinoja on useita, kuten ehkäpä varmimpana keinona käyttäjäkohtaiset tutkimukset, mutta yksi asia on selvää: mitä enemmän dataa meillä on analysoitavana, sitä tarkempia päätelmiä voimme tehdä automaattisesti. Siksi olisi hyvä alkaa kerätä dataa jo nyt varhaisessa vaiheessa autonominen rahtilaivan prosesseja sekä osallistaa kapteenit ja muut merenkulun erityisosajaajat kaikkien prosessin vaiheiden suunnittelun, toteutuksen ja testauksen yhteydessä. Teemmehän lopulta järjestelmiä ja automaattisia toimintoja vain ihmistä varten, ja kuka olisikaan parempi kertomaan onnistumisen asteesta kuin ihminen.

Ihmisen olosuhteellinen tietoisuus on siis tärkeintä koko automatisoidun laivan kulun kannalta. Tämä on varmastikaan liioittelematta kaikista olennaisin osa-alue onnistumiseksi: kapteeni on vastuussa epävarmojen, laivan automaatiikan mahdollistamien vaihtoehtojen puntaroinnissa. Tällainen tilanne voi tulla vastaan erityisesti virhetilanteissa, joissa automaatiikka ei kykenekään itsenäisesti tekemään päätöksiä, vaan kapteenin on otettava ohjat käsiinsä samalla tavalla kuin aiemmin on totuttu manuaalisesti laivaa operoimaan. Kuitenkaan olosuhteellinen tietoisuus ei ole pelkästään kapteenin harteilla, vaan automaattisen laivajärjestelmän on samalla tavalla huomioitava ympäristössään tapahtuvia asioita oman ja muiden turvallisuuden kannalta. Olosuhteellisen tietoisuuden hankkiminen ja analysointi ovat siis monen eri tahon yhteinen tehtävä, ja niin kapteeni kuin järjestelmä voi tehdä omat päätelmänsä hankitusta tiedosta. Käyttöliittymät toimivat olosuhteellisen tietoisuuden hankinnan ja analysoinnin välitykselle ja näkemysten vertailuun, ja jos niissä on ristiriitoja, kapteeni voi niin sanotusti yliajaa automaattisesti tehtyjä tai suunniteltuja toimintoja. Näissä – niin kuin kaikissa muissakin etäohjaukseen kytkeytyvissä tilanteissa – käyttöliittymien toimivuus, selkeys, käytettävyys ja käyttäjäystävällisyys ovat erittäin tärkeitä.

Sitä varten on esimerkiksi luvussa 10 kerrottu niistä seikoista, jotka on ohjenuorina otettava pohjalle käyttöliittymien suunnittelun ja toteutuksen kannalta. Käyttöliittymien kautta olosuhteellista tietoisuutta voidaan hankkia, arvioida ja hyödyntää päätöksentekoon, mikä on kriittistä koko merenkulun operatiolle eli laivan kulkuun avomerellä, satamissa sekä muiden epävarmojen karrikkojen, roskalauttojen tai muiden tällaisten odottamattomien tilanteiden tul-

lessa kyseen. Voi lisäksi olla, että kapteenin kannalta käyttöjärjestelmät ovat samankaltaisia kuin laivallakin olevat käyttöliittymät, kuten ohjausmekaniikat ja maista käsin tarkkaillessa laajakuvanäytöt luomaan tunnelmaa oikeasta laivan ohjausympäristöstä taikka komentosillalta. Kunhan käyttöjärjestelmät ja -liittymät ovat kapteenien käyttötestausten pohjalta tehtyjen jatkotutkimusten perusteella onnistuneita, voimme sanoa onnistuneemme kriittisimmässä osaluueessa eli olosuhteellisen tietoisuuden hyödyntämisessä etäältä.

Voi myös olla pohtimisen arvoista, millaisen tekoölyn haluamme luoda järjestelmien avustuksella toimivaksi. Ei ole enää ainoastaan Turingin koneen päänsä vaiva erottaa sitä, onko vastapuolena toimivan kommunikoidijan rooli ihminen vai tietokone. Koska ihmiset ovat yhä enemmän sidoksissa järjestelmiin, joiden avulla tieto välittyy ihmisen ja robotin välillä, ihmisen on hyvä tietää, kenen tai minkä kanssa keskustelu tapahtuu. Pitäisikö järjestelmiä tehdessä niistä muovata niin ihmismäisiä, että ihminen keskustelisi järjestelmän kanssa kuin toisen ihmisen kanssa, kuten autonominen rahtilaivan järjestelmä vaikuttaisi ihmiseltä? Ehkäpä niin suurta tekoölyä ei välttämättä kannata tehdä, vaan voisi olla kannattavaa ja kustannustehokkaampaa pelkästään naamioida järjestelmä ystävälliseksi, käyttäjäystävälliseksi ja intuitiiviseksi eikä kokonaisuudessaan toteuttaa sille täydellistä, ihmismäistä kommunikaatiota, sillä kapteeni kuitenkin tietää usein keskustelewansa järjestelmän kanssa – ei toisen ihmisen. Tämän projektin ollessa kyseessä laivan kapteeni voi lähes sataprosenttisella varmuudella tietää keskustelewansa laivan teknologisen järjestelmän kanssa eikä luonnollisen henkilön. Ihmismäisen laivarobotin teko ei ole välttämättä perusteltua, sillä vaikka järjestelmä olisi kuinka robottimainen ja autonominen, se on silti vain ihmisen eli kapteenin apukäsiksi tehty välikappale päättämään laivan suuntaviivauksia ja toimintaa pyrkien ajattelemaan ja toimimaan kuin sitä ohjailisi ihminen.

Kuitenkin arkipäivään, miksei työpäiväänkin, sulautetut suorastaan autonomiset robotit auttavat ihmistä usein jopa tämän tiedostamatta itse siitä. Esimerkkinä automatkan päätepisteeseen reitin päättävät autonavigaattorit, jotka laskevat reitin aina itse. Silti ihminen luottaa niihin suhteellisen vankasti ja tekee niiden pohjalta usein ratkaisunsa. Samanlaisesti rahtilaiva voi päätellä oman reittinsä ja kapteeni sen vahvistaa käyttöjärjestelmien kautta, kun järjestelmää pidetään luotettava työpartnerina.

Viimeisimpänä tässä tutkimuksessa on nostettu esiin se seikka, miten näinkin tuoreessa kokonaisuudessa voitaisiin ottaa oppia muilta aloilta. Se on varmasti haasteellista: mitä voimme tarkastella ja hyödyntää toisista järjestelmistä, toisilta toimialoilta, kuten esimerkiksi lentoliikenteestä. Voi olla huomion arvoista käydä läpi niitä samankaltaisia aloja, joilla autonomiaa on hyödynnetty. Niistä voisimme kategorisoida niin hyvin onnistuneet automatisoinnit kuin epävarmoissa tilanteissa esimerkiksi onnettomuuksiin johtaneet virhetilanteet, joissa voimme välttää samanlaiset kompastuskivet. Näitä ovat muun muassa virhetilanteiden puutteellinen tai virheellinen ilmaisu käyttöliittymien kautta. Lisäksi, koska dataa tarvitaan väistämättä paljon autonomian tai pitkälle automatisoidun laivanjärjestelmän mallien toimintaan, voisimmeko analysoida sitä jo olemassa olevasta muusta automatisoidusta liikenteestä?

Tutkimusta yhteen vetäen pohtimisen arvoisia kysymyksiä ovat sen lisäksi seuraavat: Ovatko tässä kirjallisuuskatsauksessa löydetyt asiakokonaisuudet myös itse merenkulun eksperttien mielestä ne suuret kokonaisuudet, jotka on ratkaistava? Ja jos ne ovat, niin vaikuttavatko tutkittavat osakokonaisuudet siltä, että ne voidaan myös ratkaista samankaltaisin keinoin kuin aiemmassa tutkimuksessa on tehty? Tässä tutkimuksessa itse kapteenit tai merenkulun ekspertit eivät ole edustettuina, sillä tutkimus on tarkoitettu pääasiassa hahmottamaan kokonaisprosessia autonomisen rahtilaivan ja sen järjestelmien kannalta. Jatkossa heidät on otettava tiiviisti osaksi suunnitteluprosesseja ja kehitysvaiheita, jotta kykenemme luomaan järjestelmiä juuri tälle oikealle kohderyhmälle ja ottamaan tuoreeltaan heidän ajatuksensa ja mielipiteensä huomioon. Tutkimushenkilöiden hankkiminen määrällisesti alkuun lienee haastavaa, sillä kapteeneja ja muita merenkulun ammattilaisia ei ole välttämättä liialti saatavilla, joten kokonaisuutta voitaisiin lähestyä laadullisesti strukturoiduin haastatteluin. Ja jos tutkimushenkilöiden määrä kyseisessä projektissa voi olla vähäinen, samaa voisi sanoa olemassa olevan tutkimuksen määrästä käyttöliittymien käytettävyydestä etäältä ohjattavissa laivoissa.

Tämänkin epäkohdan paikkaamiseen tarvittaisiin lisätutkimusta ja haastatteluja laivahenkilökunnan – varsinkin kapteenien – kanssa. Tässä kirjallisuuskatsauksessa ilmi tulleet kokonaisuudet on hyvä varmistaa haastatteluin, jotta voimme edetä esimerkiksi olosuhteellisen tietoisuuden hankkimisen metodin analysoinnilla ja valinnalla sekä käyttöliittymien ohjenuorien perusteella luotavilla demoilla. Haastatteluista voidaan saada elämäpohjaisia faktoja: näiden perusteella voitaisiin suunnitella käyttöliittymiä, joita sitten puolestaan testata ja hyödyntää itse käytännössä. Kirjallisuuskatsauksen pohjalta voitaisiin aiemmista tutkimuksista koskien etäohjattavien ajoneuvojen käyttöliittymiä suunnitella prototyyppinäkyviä erilaisille sidosryhmille ja rooleille, jotta kokonaisuudessaan vuoropuhelu laivan järjestelmän ja sen käyttäjien välillä olisi mahdollisimman vaivatonta ja selkeää, sillä se on olennaisimpia asioita rahtilaivan onnistumisessa suunnitellulla kulkureitillä.

Kaiken kaikkiaan vaikuttaa mahdolliselta, joskin haasteelliselta, luoda autonominen rahtilaiva, jos autonomisella tarkoitetaan täydellisen automaattista prosessia. Prosessiin kuuluu muun muassa tässä pro gradu -tutkielmassa ilmi tuotuja asiakokonaisuuksia niin teknologian, keinoälyn, kielitieteen, sosiologian kuin ihmisen olosuhteellisen tietoisuuden kannalta sekä olosuhteellisen tietoisuuden ylläpito käytettävien, etäohjattavien järjestelmien kautta, mutta tämäkään kirjallisuuskatsaus ei liene mikään täydellisen kattava kokonaisuus ottamaan kaikki seikat huomioon. Kirjallisuuskatsauksen aiheena on luoda katsaus rahtilaivan prosesseihin osallistuvien tahojen näkökulmasta niihin prosesseihin, joita varsinkin laivan kapteeni tulee tarvitsemaan, kun autonominen tai pitkälle automatisoitu rahtilaiva toteutuu käytännössä ja automatisoinnin ongelmakohtiin on otettava kantaa. Kuitenkin siihen, että rahtilaivat kykenevät täysin itsenäisesti seilaamaan satamasta satamaan vailla suurempaa ihmisohjausta, saattaa mennä vielä pitkään, eikä sanana *autonomia* ole välttämättä alkuvaiheessa relevantti, vaan parempi sana voisi olla pitkälle automatisoitu laiva. Kuten Asimov (1942) jo puhui tekoälystä ja roboteista sekä niiden vaaroista, nykypäivänä 2010-luvulla taistelemme usein vielä samankaltaisten ongelmien kanssa:

mitä robotti tarkoittaa ja mitä vaaroja siinä voi olla sekä kuinka ihmismieltä voimme mallintaa. 1940-luvulta tähän päivään tultaessa moni asia on muuttunut mutta moni kysymys säilynyt yhä samana vain kontekstin muuttuessa. Siksi yllättävää tekoälyn kehitystä ihmismielen kopioijana tuskin vielä lähivuosina tullaan keksimään, kun emme osaa edes vielä sanoa, kuinka ihmismieli toimii. Pitkälle automatisoitu rahtilaiva vaatii lähitulevaisuudessa yhä kapteenin panosta onnistuneeksi reitinkuluksi satamasta satamaan juuri niissä kohdissa, joita teknologia ei vielä kykene paikkaamaan ihmisen ajattelumaailman toiminnassa, sekä valvomaan kokonaisprosessia. Samalla tavalla kuin ihminenkin ei välttämättä tiedä kaikkea ja saattaa kysyä apua työtoveriltaan, rahtilaiva voi myös kysyä sitä tarkkailevalta, maista käsin etäohjattavien järjestelmien kautta operoivalta kapteenilta, mikäli sille on suotu seuraavanlaiset kyvyt: "Mitä teemme nyt?"

## 13 LÄHTEET

- Acemoglu, D. & Restrepo, P. (2018). *Artificial Intelligence, Automation and Work*. National Bureau of Economic Research.
- Adamides, G.; Christou, G.; Katsanos, C.; Xenos, M. & Hadzilacos, T. (2015). Usability Guidelines for the Design of Robot Teleoperation: A Taxonomy. *IEEE Transactions on Human-Machine Systems*, 45 (2), 256–262.
- Adams, J. (2005). Human-Robot Interaction Design: Understanding User Needs and Requirements. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 49 (3), 447–451.
- Anderson, J. R.; Matessa, M. & Lebiere, C. (1997). ACT-R: A Theory of Higher Level Cognition and Its Relation to Visual Attention. *Human-Computer Interaction*, 12 (4), 439–462.
- Asimov, I. (1942). Runaround. *Astounding Science Fiction*, 29 (1), 94–103.
- Asimov, I. & McKeever, L. (1982). *The complete robot*. New York: Doubleday.
- Baker, C. L. & Tenenbaum, J. B. (2014). Modeling Human Plan Recognition. *Plan, Activity, and Intent Recognition: Theory and Practice*, 177–204.
- Baumeister, R. F. & Leary, M. R. (1997). Writing Narrative Literature Reviews. *Review of General Psychology*, 1 (3), 311–320.
- Bedny, G. & Meister, D. (1999). Theory of activity and situation awareness. *International Journal of Cognitive Ergonomics*, 3 (1), 63–72.
- Bell, H. H. & Lyon, D. R. (2000). *Using Observer Ratings to Assess Situation Awareness*. New Jersey: Laurence Erlbaum Associates.
- Betty, H. C.; Rogério, D. E.; Holger, G.; Inverardi, P. & Magee, J. (2009). Software Engineering for Self-Adaptive Systems. *Lecture Notes in Computer Science*, 5525.

- Bomberger, N. A.; Rhodes, B. J.; Seibert, M. & Waxman, A. M. (2006). Associative Learning of Vessel Motion Patterns for Maritime Situation Awareness. *IEEE International Conference on Information Fusion*, 1–8.
- Bradshaw, J. M., Dignum, V., Jonker, C., & Sierhuis, M. (2012). Human-Agent-Robot Teamwork. *IEEE Intelligent Systems*, 27 (2), 8–13.
- Breazeal, C. (2004). Social Interactions in HRI: The Robot View. *IEEE Transactions on Systems, Man, and Cybernetics (Part C: Applications and Reviews)*, 34 (2), 181–186.
- Bruemmer, D. J.; Few, D. A.; Boring, R. L.; Marble, J. L.; Walton, M. C. & Nielsen, C. W. (2005). Shared understanding for collaborative control. *IEEE Transactions on Systems, Man, and Cybernetics (Part A: Systems and Humans)*, 35 (4), 494–504.
- Castelfranchi, C. (2000). Founding Agent's "Autonomy" On Dependence Theory. *Proceedings of the 14th European Conference on Artificial Intelligence*, 353–357. IOS Press.
- Castelfranchi, C. & Falcone, R. (2003). From Automaticity to Autonomy: The Frontier of Artificial Agents. *Agent Autonomy*, 103–136. Boston: Springer.
- Chen, L.; Zhang, D.; Ma, X.; Wang, L.; Li, S.; Wu, Z. & Pan, G. (2016). Container Port Performance Measurement and Comparison Leveraging Ship GPS Traces and Maritime Open Data. *IEEE Transactions on Intelligent Transportation Systems*, 17 (5), 1227–1242.
- Cheng, B. H. C.; de Lemos, R.; Giese, H.; Inverardi, P.; Magee, J.; Andersson, J.; Becker, B.; Bencomo, N.; Brun, Y.; Cukic, B.; Di Marzo Serugendo, G.; Dustdar, S.; Finkelstein, A.; Gacek, C.; Geihs, K.; Grassi, V.; Karsai, G.; Kienle, H.; Kramer, J.; Litoiu, M.; Malek, S.; Mirandola, R.; Müller, H.; Park, S.; Shaw, M.; Tichy, M.; Tivoli, M.; Weyns, D. & Whittle, J. Software Engineering for Self-Adaptive Systems: A Research Road Map (Draft Version). *Dagstuhl Seminar*, 10431.
- Chien, S. Y.; Lewis, M.; Mehrotra, S.; Brooks, N. & Sycara, K. (2012). Scheduling Operator Attention for Multi-Robot Control. *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 473–479.
- Compton, P. & Jansen, R. (1990). A philosophical basis for knowledge acquisition. *Knowledge Acquisition*, 2 (3), 241–258.
- Davis, J.; Animashaun, A.; Schoenherr, E. & McDowell, K. (2008). Evaluation of Semi-Autonomous Convoy Driving. *Journal of Field Robotics*, 25 (11–12), 880–897.



- Degerholm, K. (2017). *Teknisiä innovaatioita, jotka tukevat autonomisten laivojen kansipuolen operointia (Opinnäytetyö)*. Satakunnan ammattikorkeakoulu, Merenkulun koulutusohjelma.
- Dekker, S. (2017). *The Field Guide to Understanding "Human Error"* (2. painos). Ashgate.
- Devin, S. & Alami, R. (2016). An Implemented Theory of Mind to Improve Human-Robot Shared Plans Execution. *11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2016*, 319–326.
- Dix, A.; Finlay, J.; Abowd, G. D. & Beale, R. (2009). Human-computer interaction. *Encyclopedia of database systems*. Springer US.
- Dorneich, M. C.; Whitlow, S.; Rogers, W.; Feigh, K. & Robert, E. (2015). *Yhdysvaltain patentti 8,977,407*. Washington, DC: U.S. Patent and Trademark Office.
- Drury, J. L.; Keyes, B. & Yanco, H. A. (2007). LASSOing HRI: Analyzing Situation Awareness in Map-Centric and Video-Centric Interfaces. *2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 279–286.
- Durso, F. T.; Hackworth, C. A.; Truitt, T. R.; Crutchfield, J.; Nikolic, D. & Manning, C. A. (1998). Situation Awareness As a Predictor of Performance in En Route Air Traffic Controllers. *Air Traffic Control Quarterly*, 6 (1), 1–20.
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society Annual Meeting*, 32 (2), 97–101. Los Angeles: SAGE Publications.
- Endsley, M. R. (1995a). A taxonomy of situation awareness errors. *Human Factors in Aviation Operations*, 3 (2), 287–292.
- Endsley, M. R. (1995b). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, 37 (1), 32–64.
- Endsley, M. R. (2000). Direct measurement of situation awareness: Validity and use of SAGAT. *Situation Awareness Analysis and Measurement*, 10.
- Endsley, M. R.; Bolté, B. & Jones, D. G. (2003). *Designing for Situation Awareness: An approach to User-Centered Design*. London: Taylor & Francis.
- Endsley, M. R. & Connors, E. S. (2008). Situation awareness: State of the art. *IEEE Power and Energy Society General Meeting–Conversion and Delivery of Electrical Energy in the 21st Century*, 1–4.

- Fernandes, E.; Jung, J. & Prakash, A. (2016). Security Analysis of Emerging Smart Home Applications. *IEEE symposium on security and privacy (SP)*, 636–654.
- Ferrara, E.; Varol, O.; Davis, C.; Menczer, F. & Flammini, A. (2016). The Rise of Social Bots. *Communications of the ACM*, 59 (7), 96–104.
- Fink, A. (2019). *Conducting Research Literature Reviews: From the Internet to Paper*. Sage Publications.
- Fleetwood, M. D. & Byrne, M. D. (2006). Modeling the Visual Search of Displays: A Revised ACT-R Model of Icon Search Based on Eye-Tracking Data. *Human-Computer Interaction*, 21 (2), 153–197.
- Fong, T.; Thorpe, C. & Baur, C. (2003). Collaboration, dialogue, human-robot interaction. *Robotics Research*, 255–266. Berlin: Springer-Verlag.
- Fong, T.; Nourbakhsh, I.; Kunz, C.; Flückiger, L.; Schreiner, J.; Ambrose, R.; Burridge, R.; Simmons, R.; Hiatt, L. M.; Schultz, A.; Trafton, J. G.; Bugajska, M. & Scholtz, J. (2005). The Peer-to-Peer Human-Robot Interaction Project. *AIAA Space 2005*.
- Fracker, M. L. (1991). Measures of Situation Awareness: Review and Future Directions.
- Fritsch, J.; Kleinhagenbrock, M.; Haasch, A.; Wrede, S. & Sagerer, G. (2005). A Flexible Infrastructure for the Development of a Robot Companion with Extensible HRI-Capabilities. *IEEE International Conference on Robotics and Automation (ICRA)*, 3408–3414.
- Gaissmaier, W.; Schooler, L. J. & Mata, R. (2008). An ecological perspective to cognitive limits: Modeling environment-mind interactions with ACT-R. *Judgment and Decision Making*, 3 (3), 278–291.
- Galitz, W. O. (2007). *The Essential Guide to User Interface Design*. Kanada: John Wiley & Sons, Inc.
- Gianaros, P. J.; Muth, E. R.; Mordkoff, J. T.; Levine, M. E. & Stern, R. M. (2001). A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, space, and environmental medicine*, 72 (2), 115–119.
- Gilson, R. D.; Hitchcock, L.; Pew, R. A.; Endsley, M. R.; Smith, K.; Taylor, R. M.; Adam, E. C. & Sarter, N. B. (1994). *Situational Awareness in Complex Systems*. Orlando: University of Central Florida.
- Goerzen, C. Kong, Z. & Mettler, B. (2010). A Survey of Motion Planning Algorithms from the Perspective of Autonomous UAV Guidance. *Journal of Intelligent and Robotic Systems*, 57 (1–4), 65.

- Goldin, D. Q.; Smolka, S. A.; Attie, P. C. & Sonderegger, E. L. (2004). Turing machines, transition systems, and interaction. *Information and Computation*, 194 (2), 101–128.
- Compton, P. & Jansen, R. (1990). A philosophical basis for knowledge acquisition. *Knowledge acquisition*, 2 (3), 241–258.
- Harnad, S. (1990). The Symbol Grounding Problem. *Physica D: Nonlinear Phenomena*, 42 (1–3), 335–346.
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139–183.
- Hayes, P. & Ford, K. (1995). Turing test considered harmful. *14th international joint conference on Artificial intelligence*, 1, 972–977.
- Helsingin Sanomat (HS, 2017, 20. kesäkuuta). Superäly päihittää ihmisen jo vuonna 2050 – tekoälystä tulee moniosaaja. Haettu 6.10.2018 osoitteesta <https://www.hs.fi/tiede/art-2000005260827.html>
- Hjelmfelt, A. T. & Pokrant, M. A. (1998). Coherent tactical picture. *CNA RM*, 97–129. Alexandria, Virginia: Center for Naval Analyses.
- Huang, H. M.; Pavek, K.; Novak, B.; Albus, J. & Messin, E. (2005). Autonomy Levels for Unmanned Systems (ALFUS). *AUVSI's Unmanned Systems North America*, 849–863.
- Hüttenrauch, H. (2006). *From HCI to HRI: Designing Interaction for a Service Robot* (Väitöskirja). Kuninkaallinen teknillinen korkeakoulu.
- Jian, J. Y.; Bisantz, A. M. & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, 4 (1), 53–71.
- Julkaisufoorumi (2018). Haettu vuosina 2017 ja 2018 osoitteesta <https://www.tsv.fi/julkaisufoorumi/haku.php>
- Kaupp, T.; Makarenko, A. & Durrant-Whyte, H. (2010). Human–robot communication for collaborative decision making – A probabilistic approach. *Robotics and Autonomous Systems*, 58 (5), 444–456.
- Keyes, B.; Micire, M.; Drury, J. L. & Yanco, H. A. (2010). Improving Human-Robot Interaction through Interface Evolution. *Human-robot Interaction*, 183–202. InTech.
- Kirsh, D. (2000). A Few Thoughts on Cognitive Overload. *Intellectica*, 1 (30), 19–51.

- Krizhevsky, A.; Sutskever, I. & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 1097–1105.
- Van Leeuwen, J. & Wiedermann, J. (2001). The Turing machine paradigm in contemporary computing. *Mathematics unlimited – 2001 and beyond*, 1139–1155. Heidelberg: Springer Berlin.
- Levander, O. (2017). Autonomous ships on the high seas. *IEEE Spectrum*, 54 (2), 26–31.
- Lin, P. (2016). Why Ethics Matters for Autonomous Cars. *Autonomous Driving*, 69–85. Springer: Berlin, Heidelberg.
- Luo, S.; Jin, J., & Li, J. (2009). A Smart Fridge with an Ability to Enhance Health and Enable Better Nutrition. *International Journal of Multimedia and Ubiquitous Engineering*, 4 (2), 69–80.
- Man, Y.; Weber, R.; Cimbritz, J.; Lundh, M. & MacKinnon, S. N. (2018). Human factor issues during remote ship monitoring tasks: An ecological lesson for system design in a distributed context. *International Journal of Industrial Ergonomics*, 68, 231–244.
- Mankoff, J.; Dey, A. K.; Hsieh, G.; Kientz, J.; Lederer, S. & Ames, M. (2003). Heuristic evaluation of ambient displays. *ACM SIGCHI conference on Human factors in computing systems*, 169–176.
- McCulloch, W. S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5 (4), 115–133.
- Miller, C. A. & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors*, 49 (1), 57–75.
- Molineaux, M.; Klenk, M. & Aha, D. (2010). Goal-Driven Autonomy in a Navy Strategy Simulation. *AAAI Conference on Artificial Intelligence*.
- Moreira, L. & Soares, C. G. (2011). Autonomous Ship Model to Perform Manoeuvring Tests. *Journal of Maritime Research*, 8 (2), 29–46.
- Nelson, R. C. (1998). *Flight Stability and Automatic Control* (2). New York: WCB/McGraw Hill.
- Nicolescu, M.; Leigh, R.; Olenderski, A.; Louis, S.; Dascalu, S.; Miles, C.; Quiroz, J. & Aleson, R. (2007). A training simulation system with realistic autonomous ship control. *Computational Intelligence*, 23 (4), 497–516.
- Nielsen, J. (1992). Finding Usability Problems Through Heuristic Evaluation. *ACM SIGCHI Conference on Human Factors in Computing Systems*, 373–380.

- Niesser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. WH Freeman / Times Books / Henry Holt & Co.
- Nofi, A. A. (2000). *Defining and measuring shared situational awareness*. Alexandria, Virginia: Center for Naval Analyses.
- Norman, D. A. (1986). Cognitive Engineering. *User Centered System Design*, 31-61.
- Norman, D. (2013). *The Design of Everyday Things*. Arizona: Basic Books.
- Nwana, H. S. (1996). Software agents: An overview. *The knowledge engineering review*, 11 (3), 205-244.
- Oksa, M. (2016). Web API development and integration for microservice functionality in web applications (Pro gradu -tutkielma). Jyväskylän yliopisto.
- Parasuraman, R.; Galster, S.; Squire, P.; Furukawa, H. & Miller, C. (2005). A Flexible Delegation-Type Interface Enhances System Performance in Human Supervision of Multiple Robots: Empirical Studies With RoboFlag. *IEEE Transactions on Systems, Man, and Cybernetics (Part A: Systems and Humans)*, 35 (4), 481-493.
- Parasuraman, R.; Sheridan, T. B. & Wickens, C. D. (2000). A Model for Types and Levels of Human Interaction with Automation. *IEEE Transactions on Systems, Man, and Cybernetics (Part A: Systems and Humans)*, 30 (3), 286-297.
- Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M. & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Python (2018). Haettu 14.6.2018 osoitteesta <https://www.python.org/>
- Reason, J. (1990). *Human error*. Cambridge University Press.
- Riley, J. M. & Endsley, M. R. (2005). Situation awareness in HRI with collaborating remotely piloted vehicles. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 49 (3), 407-411. Los Angeles: SAGE Publications.
- Riveiro, M.; Falkman, G. & Ziemke, T. (2008). Improving maritime anomaly detection and situation awareness through interactive visualization. *11th International Conference on Information Fusion*.
- Rothblum, A. M. (2000). Human Error and Marine Safety. *US Coastguard Research and Development Centre*, 1-10.

- Saariluoma, P. & Leikas, J. (2010). Life-Based Design - An Approach to Design for Life. *Global Journal of Management and Business Research*, 10 (5), 17–23.
- SAE International (Society of Automotive Engineers, 2018). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. Standard J3016\_201806. Warrendale, Pennsylvania.
- Salas, E.; Prince, C.; Baker, D. P. & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors*, 37 (1), 123–136.
- Salminen, A. (2011). Mikä kirjallisuuskatsaus? Johdatus kirjallisuuskatsauksen tyyppeihin ja hallintotieteellisiin sovelluksiin. *Opetusjulkaisuja 62. Julkisjohtaminen 4*. Vaasa: Vaasan yliopiston julkaisuja.
- Salmon, P.; Stanton, N.; Walker, G. & Green, D. (2006). Situation Awareness Measurement: A Review of Applicability for C4i Environments. *Applied Ergonomics*, 37 (2), 225–238.
- Salmon, P. M.; Stanton, N. A.; Walker, G. H.; Jenkins, D.; Ladva, D.; Rafferty, L. & Young, M. (2009). Measuring Situation Awareness in Complex Systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, 39 (3), 490–500.
- Scholtz, J. C. (2002a). *Evaluation Methods for Human-System Performance of Intelligent Systems*. Maryland, Gaithersburg: National Institute of Standards and Technology.
- Scholtz J. C. (2002b). Human-Robot Interactions: Creating Synergistic Cyber Forces. *Multi-Robot Systems: From Swarms to Intelligent Automata*. Dordrecht: Springer.
- Scholtz J. C. (2003). Theory and Evaluation of Human Robot Interactions. *Proceedings of the Annual Hawaii International Conference on System Sciences*.
- Scholtz, J. C.; Antonishek, B. & Young, J. (2004). Evaluation of a Human-Robot Interface: Development of a Situational Awareness Methodology. *Proceedings of the Annual Hawaii International Conference on System Sciences*.
- Scholtz, J. C. & Bahrami, S. (2003). Human-robot interaction: Development of an evaluation methodology for the bystander role of interaction. *IEEE International Conference on Systems, Man and Cybernetics*, 4, 3212–3217.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3 (3), 417–424.
- Searle, J. R. (1990). Is the Brain's Mind a Computer Program? *Scientific American*, 262 (1), 25–31.

- Selcon, S. J. & Taylor, R. M. (1990). Evaluation of the Situational Awareness Rating Technique (SART) as a tool for aircrew systems design. *Situational Awareness in Aerospace Operations*, 8.
- Sheridan, T. B. (1992). *Telerobotics, Automation, and Human Supervisory Control*. Cambridge, Massachusetts: MIT Press.
- Simon, H. A. (1973). The Structure of Ill Structured Problems. *Artificial Intelligence*, 4 (3-4), 181-201.
- Smith, K. & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors*, 37 (1), 137-148.
- Sneyers, J.; van Weert, P.; Schrijvers, T. & De Koninck, L. (2010). As time goes by: Constraint handling rules. *Theory and Practice of Logic Programming*, 10 (1), 1-47.
- Song, K.; Liu, Q. & Wang, Q. (2011). Olfaction and Hearing Based Mobile Robot Navigation for Odor/Sound Source Search. *Sensors*, 11 (2), 2129-2154.
- Spong, M. W.; Hutchinson, S. & Vidyasagar, M. (2004). *Robot Dynamics and Control* (2. painos). New York: John Wiley & Sons.
- Tekniikka&Talous (2017, 5. kesäkuuta). Näin Rolls-Royce kehittää robottilaivoja Suomessa - edelläkävijyydestä 50 000 € Teknologiapalkinto. Haettu 22.1.2019 osoitteesta [https://www.tekniikkatalous.fi/kaikki\\_uutiset/nain-rolls-royce-kehittaa-robottilaivoja-suomessa-edellakavijyydesta-50-000-teknologiapalkinto-6654759](https://www.tekniikkatalous.fi/kaikki_uutiset/nain-rolls-royce-kehittaa-robottilaivoja-suomessa-edellakavijyydesta-50-000-teknologiapalkinto-6654759)
- Teknologian tutkimuskeskus VTT Oy (2017, 8. maaliskuuta). VTT ja Rolls-Royce vievät autonomisten laivojen kehityksen uudelle vuosikymmenelle. Haettu 19.10.2017 osoitteesta <http://www.vtt.fi/medialle/uutiset/vtt-ja-rolls-royce-viev%C3%A4t-autonomisten-laivojen-kehityksen-uudelle-vuosikymmenelle>
- Tenorth, M. & Beetz, M. (2009). KnowRob – Knowledge Processing for Autonomous Personal Robots. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4261-4266.
- Trafton, J. G.; Hiatt, L. M.; Harrison, A. M.; Tamborello II, F. P.; Khemlani, S. S. & Schultz, A. C. (2013). ACT-R/E: An Embodied Cognitive Architecture for Human-Robot Interaction. *Journal of Human-Robot Interaction*, 2 (1), 30-55.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59 (236), 433-460.

- Turner, J. R. & Cochrane, R. A. (1993). Goals-and-methods matrix: coping with projects with ill defined goals and/or methods of achieving them. *International Journal of Project Management*, 11 (2), 93–102.
- Veres, S. M.; Molnar, L.; Lincoln, N. K. & Morice, C. P. (2011). Autonomous vehicle control systems – a review of decision making. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 225 (2), 155–195.
- Walters, M. L.; Dautenhahn, K.; Te Boekhorst, R.; Koay, K. L.; Syrdal, D. S. & Nehaniv, C. L. (2009). An empirical framework for human-robot proxemics. *Procs of new frontiers in human-robot interaction*.
- Wegner, P. (1998). Interactive foundations of computing. *Theoretical computer science*, 192 (2), 315–351.
- Weiss, A.; Wurhofer, D.; Buchner, R.; Tscheligi, M.; Blasi, L. & Plebani, M. (2009). Development of a Teleoperator Interface for Humanoid Robots by Means of Heuristic Evaluation Technique. *Proceedings of Towards Autonomous Robotic Systems (TAROS)*, 236–241.
- Yachting & Boating World, YBW (2017, 21. kesäkuuta). Rolls-Royce demonstrates world's first remotely operated commercial vessel. Haettu 18.5.2019 osoitteesta <https://www.ybw.com/news-from-yachting-boating-world/rolls-royce-demonstrates-worlds-first-remotely-operated-commercial-vessel-55113>
- Yanco, H. A. & Drury, J. (2004a). Classifying Human-Robot Interaction: An Updated Taxonomy. *IEEE International Conference on Systems, Man and Cybernetics*, 3, 2841–2846.
- Yanco, H. A. & Drury, J. (2004b). "Where am I?" Acquiring situation awareness using a remote robot platform. *IEEE International Conference on Systems, Man and Cybernetics*, 3, 2835–2840.
- Yanco, H. A.; Drury, J. L. & Scholtz, J. (2004). Beyond Usability Evaluation: Analysis of Human-Robot Interaction at a Major Robotics Competition. *Human-Computer Interaction*, 19 (1), 117–149.
- Yhdysvaltain puolustusministeriö (2018). *Summary of the 2018 Department of Defense Artificial Intelligence Strategy*. Yhdysvallat. Haettu 11.6.2018 osoitteesta <https://media.defense.gov/2019/Feb/12/2002088963/-1/-1/1/SUMMARY-OF-DOD-AI-STRATEGY.PDF>
- Yle (2017, 30. lokakuuta). Juuso Pekkinen: Tulevaisuuden aavelaivat. Haettu 18.5.2019 osoitteesta <https://areena.yle.fi/1-4261581>



Zhang, Q.; Wu, Y. N. & Zhu, S. C. (2018). Interpretable Convolutional Neural Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 8827-8836.