

Vastuun kantaminen sosioteknisen järjestelmän epäonnistumisessa

JAANA HALLAMAA & TAINA KALLIOKOSKI

TIIVISTELMÄ Artikkeliki käsittelee vahingon ja vastuun välisiä epäjatkuvuuksia. Mikä estää ja luo edellytyksiä yhteistoiminnan jatkamiselle sosioteknisten järjestelmien toiminnan epäonnistuttua virheen vuoksi? Sosioteknisiä systeemejä ovat esimerkiksi modernit organisaatiot, joiden ydintehtävien hoitamiseksi tarvitaan ihmistoimijoiden ja monien teknologisten järjestelmien sujuvaa yhteistyötä ja joiden vakaa toiminta on olennaista turvallisuuden takaamiseksi. Yhteiskunta edellyttää, että vakavan virheen tehnyt tai vahingon aiheuttanut kantaa vastuun toiminnastaan alistumalla oikeusjärjestelmän hänelle langettamiin sanktioihin. Kun virheen tekee teknologinen pseudotoimija tai sosiotekninen järjestelmä, perinteiset filosofiset vastuullisen toimijan intentionaalisuuteen, kausaalisuuteen sekä kontrolliin perustuvat ehdot synnyttävät vastuuaukon. Aineistona ovat epäonnistumisen tutkimukset sekä filosofiset yhteistoiminnan ja vastuun teoriat. Niitä kriittisesti soveltaen muotoillaan yhteistoiminnan onnistumisen ja epäonnistumisen sekä yhteisen vastuun kantamisen käsitteellisiä, toiminnallisia ja moraalipsykologisia ehtoja. Tekoälyteknologiat ja systeemiajattelu haastavat käsityksiä vastuusta ja syyllisyydestä. Tarkastelemalla vastuunkantamista relationaalisesti sosiaalisena toimintana osa menneeseen ja tulevaan suuntautuvan vastuun jakamisen vaikeuksista voidaan ratkaista. Vastuuaukon ylittämiseksi esitetään osallisuuden käsitettä ja erilaisia vahinkoon osallistumisen tapojen tunnistamista.

ASIASANAT Vastuu, jaettu vastuu, yhteistoiminta, epäonnistuminen, sosiotekninen systeemi, virhe.

JOHDANTO

Ihmisten yhteistoiminta on 2000-luvulla raken-
tunut yhä kiinteämmin digitaalisten ja tekoöly-
avusteisten järjestelmien sujuvan toiminnan
varaana. Teknologisten ratkaisujen keskeinen
tavoite on korvata inhimillisen toiminnan puut-
teita. Vaikka tavoite usein toteutuu, ihmisten ja
teknologioiden muodostamat sosiotekniset jär-
jestelmät eivät toimi virheettömästi, vaan luovat
uudenlaisia epäonnistumisen lähteitä.¹

Epäonnistumisia on perinteisesti käsitelty
muun muassa sovittelemalla niiden aiheut-
tamia konflikteja mutta myös pyytämällä ja
antamalla anteeksi. Jos teknologinen järjestelmä
on osa epäonnistumiseen johtaneita tekijöitä,
tavanomaiset tavat selittää tapahtuman syitä ja
korjata aiheutuneita vahinkoja eivät kuitenkaan
riitä. Myös kysymys siitä, kuka kantaa vastuun
epäonnistumisen vaikutuksista, mutkistuu
sosioteknisissä järjestelmissä.

Sosiotekniset järjestelmät voivat olla pieniä,
kuten ihmisen ja jonkin laitteen joissakin sosiaa-
lisisissa olosuhteissa muodostama toiminnallinen
kokonaisuus, tai suuria, kuten iso teollisuuslaitos,
maa-, meri- tai ilmailukenneväline logistisine
taustajärjestelmineen tai tieto-, tekoöly- tai
automaatiojärjestelmiin perustoinnoissaan
tukeutuva yhteiskunnallinen organisaatio. Teko-
ölyllä tarkoitamme tässä niin sääntöpohjaisia
kuin koneoppimisperustaisia, suuria datamää-
riä prosessoivia järjestelmiä, joita luonnehtii
autonomisuus ja adaptiivisuus. Autonomisuus
tarkoittaa kykyä tehdä päätöksiä ilman ihmisen
välitöntä osuutta ja ohjausta ja adaptiivisuus
tekoölyjärjestelmän kykyä mukauttaa omaa toi-
mintaa digitaalisesta ja fyysisestä ympäristöstä
kerättyjen syötteiden perusteella.²

Vastuunalaisuutta määrittävät filosofiset ja
oikeudelliset käsitykset rakentuvat oletukselle,
että vahingon on saanut aikaan tunnistettava
ihmistoimija, jonka vallassa siihen johtaneet
toimet ovat olleet. Tekoölyjärjestelmien vastuul-
lisuuden määrittämistä tutkineet Mehdi Dastani
ja Vahid Yazdanpanah kutsuvat tätä toimijakes-
keiseksi suhtautumiseksi vastuuseen erotuksena
tapahtumakeskeisestä kausaalisuhteiden selvit-
tämiseen keskittyneestä vastuukeskustelusta.³

Mittavia, onnettomuuksiksi luokiteltuja epäon-
nistumisia on tavattu tarkastella kausaalisina
tapahtumakulkuina, joiden toistumista pyritään
estämään välttämällä moraalisia kannanottoja
tapahtumiin osallistuneiden syyllisyydestä. Mo-
raalisesti neutraalin lähestymistavan on todettu
tarjoavan parhaan perustan toiminnan kehittä-
miselle onnettomuustapahtumista oppimalla.⁴

Jos epäonnistuminen johtuu kollektiivisista
tekijöistä tai sen saa aikaan kollektiivinen tai
teknologinen toimija, vastuullista tahoja ei voida
nimetä paikantamalla yksittäinen kausaalinen
aiheuttaja.⁵ Vastuusuhteiden muuttuminen epä-
selviksi vaikeuttaa toiminnan jatkamista tai es-
tää sen, ja toimintaa sääteleviä ehtoja on muutet-
tava, jotta päästäisiin eteenpäin. Sosioteknisissä
järjestelmissä, joissa ihmiset yksilöinä ja ryhmi-
nä toimivat yhdessä tekoölyavusteisten tekno-
logioiden kanssa, vastuunkantamiseen näyttää
muodostuvan jonkinlainen epäjatkuvuus.

Tässä artikkelissa tarkastelemme, millaisia
toiminnan epäjatkuvuuskohtia syntyy, kun
epäonnistuminen tapahtuu sosioteknisessä
järjestelmässä, ja millaisten vastuukäsitysten
avulla niitä voidaan silloittaa, jotta toimintaa
voidaan jatkaa. Toiminta tarkoittaa tässä yhtey-
dessä aktiivisuutta, jota toimija käyttää keinona
toteuttaakseen päämääräksi asettamansa tavoit-
teen tai edistääkseen sitä. Yksinkertainen määri-
telmä vastaa praktisena syllogismina tunnettua
tavoitteellisen toiminnan määritelmää.⁶ Sen

- 1 Käytämme termejä ”sosiotekninen järjestelmä” ja ”sosiotekninen systeemi” toistensa synonyymeina. Sosioteknis-systeemisestä ajattelusta ja käsitteen historiasta ks. Davis et al. 2014.
- 2 Elements of AI s.a.
- 3 Dastani & Yazdanpanah 2022.
- 4 Ks. Raine Haikaraisen kirjoitus tässä julkaisussa.
- 5 List 2021.
- 6 Tuomela 1977, 133.

varassa voidaan paitsi tarkastella ihmisagenttien pyrkimyksiä myös analysoida sosioteknisten järjestelmien toimintaa, jossa koneet ja laitteet ohjelmoidaan huolehtimaan erilaisista tehtävistä.

Tekoälysystemejäkin voidaan tarkastella toimijoina esimerkiksi tunnetun BDI-mallin (belief, desire, intention) mukaisesti, vaikka ne eivät täytä moraalille toimijalle tyypillisesti asetettuja kriteerejä.⁷ BDI-mallissa uskomukset ja halut on luotu järjestelmälle opetetun datan ja koodattujen toimintaperiaatteiden varaan, ja intentio tarkoittaa toimintaa ohjaavia periaatteita edistää järjestelmään ohjelmoitujen uskomusten ja halujen perusteella luotuja tavoitteita.⁸

Valittu lähtökohta sopii yhteen myös erityisesti Bruno Latourin tutkimuksiin yhdistetyn toimijaverkostoteorian (Actor Network Theory, ANT) kanssa. Siinä luonnollisen, sosiaalisen ja teknisen maailman ilmiöitä, niiden vuoro- ja yhteisvaikutuksia tarkastellaan samalla tasolla ilman, että niitä yritetään samalla analysoida sosiaalisten lainalaisuuksien avulla. Sosiaaliset lainalaisuudet ymmärretään ihmisten luomiksi, todellisuutta selittäviksi konstruktioiksi, joita niitäkin voidaan tarkastella ilmiömaailman osina.⁹

Toiminnan epäjatkuvuuskohdalla tarkoitamme tilannetta, jonka vuoksi toiminnan jatkaminen suunnitellulla tavalla on vaikeutunut tai estynyt kokonaan, minkä vuoksi toimintaa sääteleviä ehtoja on muutettava. Tarkastelemme ensin toiminnan keskeyttävien epäonnistumisten käsitteellistämistä sekä epäonnistumiseen johtavien syiden tutkimis- ja selittämistapoja. Tukeudumme tarkastelussa toiminnan filosofiaan ja virheiden ja systeemisten epäonnistumisten tutkimisessa kehitettyihin malleihin. Kysymme sitten, millaisia uusia elementtejä vastuun jakamiseen tarvitaan, kun epäonnistumiseen on osallistunut muitakin kuin ihmistoi-mijoita. Keskustelemme aiheesta soveltamalla kriittisesti filosofisia jaetun vastuun teorioita ja pohtimalla ihmistoi-mijoiden ja pseudotoimijoiden yhteistoiminnan onnistumisen ja yhteisen vastuunkannon käsitteellisiä, toiminnallisia ja moraalipsykologisia ehtoja.

TOIMINTA, VIRHEET JA EPÄONNISTUMINEN

Virheet ovat erottamaton osa inhimillistä toimintaa. Mitä vaativammista ja useamman toimijan yhteisistä hankkeista on kysymys, sitä todennäköisemmin toteuttajat epäonnistuvat jossakin vaiheessa jollakin tavalla. Virheiden systemaattisen tutkimuksen uranuurtajan James Reasonin mukaan epäonnistumisiin johtavat virheet edustavat saman asian eri puolia. Esimerkiksi digitaalisen laskentaohjelman toiminnan opetteleminen tutkimushakemuksen talousarvion laatimiseksi voi osoittautua virheeksi, jos siihen kuluu niin paljon aikaa, että hakemusta ei saada jätetyksi määräaikaan mennessä. Epäonnistuminen koskee kuitenkin vain keinoa: tavoitteessa oppia käyttämään ohjelmaa ei sinänsä ole mitään vikaa.¹⁰

Toimimalla toimija toteuttaa tärkeänä pitämänsä päämäärän tai edistää toivottuina, suotuisina ja hyvinä pitämiään tavoitteita. Tavoitteena voi olla tulevan toiminnan edellytysten luominen, jonkin haitallisen tapahtuman estäminen tai olemassa olevien hyvien ylläpitäminen, tukeminen tai suojeleminen.¹¹ Sosioteknisen järjestelmän toiminta, jossa ihmistoi-mijat ja teknologiset järjestelmät muodostavat yhteistoimintakokonaisuuden, voi koostua useista toimintatyypeistä. Toimijoilla voi olla jaettujen tavoitteiden lisäksi henkilökohtaisia päämääriä. Jos jaetut päämäärät ja henkilökohtaiset päämäärät ovat keskenään ristiriidassa, epäonnistumisen riski kasvaa.¹²

7 Hallamaa & Kalliokoski 2020.

8 Bosse, Memon & Treur 2011; Norling 2004. Mallin soveltamisesta ks. Hallamaa et al. 2023.

9 Latour 2005.

10 Reason 1990, 2.

11 Hallamaa 2017, 24–26.

12 Kalliokoski 2020, 158–168.

Epäonnistuminen on kaikille tuttua: suunniteltu tavoite ei toteudukaan valitulla keinolla odotetussa ajassa. Ajatus onnistumisesta on kuitenkin osa toiminnan käsitettä: toiminta on tavoitteellista tekemistä, jonka ansiosta toimija tietää tai uskoo haluamansa asian toteutuvan tai edistyvän.¹³ Käsitteellisesti onnistuminen on toiminnan normaalitapaus, joka ei edellytä lisätoimia toisin kuin epäonnistuminen, joka saa etsimään syytä epäonnistumiselle.

Teoista tai erilaisista vaiheista koostuvan toiminnan onnistuminen ja onnistuminen suhteessa tavoitteen saavuttamiseen ovat kaksi eri asiaa. Arkipäivän sattumusta voi käyttää kuvaamaan, kuinka suunniteltu toiminta kyllä onnistuu, mutta tavoite jää silti saavuttamatta, esimerkiksi kun tietoteknisen sovelluksen käyttäminen ohjeiden mukaisesti ei saakaan aikaan sovelluksen hoidettavaksi annetun asian toteutumista.

Onnistumisen kaksi aspektia – teon suorittamisen onnistuminen ja tavoitteen onnistunut saavuttaminen – menevät helposti kielenkäytössä sekaisin. Sosioteknisten systeemien toiminnan tarkastelussa onnistuminen keinon käyttämisenä ja tavoitteen saavuttamisena on tärkeää erottaa toisistaan. Esimerkiksi aseteknologian käytön onnistumista arvioidaan suhteessa kohteelle aiheutettuun vahinkoon. Ohjuksen laukaisu voi onnistua suunnitellusti ja oikeaan aikaan, mutta ohjus voi silti ohittaa maalin tai kohde voi jäädä tuhoutumatta. Vaikka yhteistoiminta onnistuu toimintana, siihen osallistuneiden näkökulmat toiminnan aikaansaamista vaikutuksista ja niiden arvosta voivat vaihdella.¹⁴

Toiminnan onnistuminen ei tarkoita, että toiminta olisi moraalisesti arvostettavaa tai edes suotuisaa sen kohteiden kannalta. Yhteistoiminnan onnistuminen edellyttää suunnitelman mukaisen tekosarjan suorittamista riittävällä tavalla ja tavoitteen saavuttamista eli onnistumista toiminnan faktuaalisen intension suhteen. Lisäksi on onnistuttava toteuttamaan toiminnan aksiologinen intentio eli tuottamaan ne toimijoiden arvokkaina, suotuisina tai hyvinä pitämät asiat, joiden vuoksi toimintaan on ryhdytty. Yksittäisen toimijan onnistuminen voi olla hyvä

hänen omalta kannaltaan, mutta vahingollista joidenkuiden muiden näkökulmasta, mikä antaa heille perusteen toivoa, ettei toimija onnistu.¹⁵

Ansion ja vastuun käsitteitä käytetään yhdistämään toisiinsa toimijan teot ja niiden kausaaliset seuraukset ja (kerrannais)vaikutukset. Jos toimijan itselleen asettaman tavoitteen saavuttaminen on suotuisaa myös muiden kannalta, toiminnan seurauksia voidaan pitää toimijan ansioina ja häntä voidaan kiittää niiden aikaansaamisesta. Tilanne muuttuu, jos toiminta epäonnistuu ja se vaikuttaa epäedullisesti muihin osapuoliin, vahingoittaa heitä tai aiheuttaa heille haittaa. Toiminta keskeytyy tai estyy ja syntyy tarve selvittää syitä ja syyllisiä, moittia ja vaatia vastuuseen. Lisäksi vahingot on korjattava ja korvattava, jotta tilanne voidaan palauttaa sellaiseksi, että yhteistyötä – tai elämää – voidaan jatkaa.¹⁶

Sosioteknisen systeemin epäonnistumisia voidaan jäsentää kysymällä, mitä epäonnistuminen rikkoo. Hyvin toimivassa systeemissä kukin osa toimii tarkoituksenmukaisesti suhteessa kokonaisuuteen ja niveltyy saumattomasti siihen, mitä muut osatoimijat tekevät. Jos systeemissä tapahtuu yllättävä virhe eli jokin keinoiksi omaksutuista toimista epäonnistuu, perusteet jatkaa aiempaa systeemistä toimintaa raukeavat: systeemi on osoittautunut epäluotettavaksi. Luottamus toisiin toimijoihin on yhteistoiminnan liima ja yhteistoimijoiden luotettavuus moraalinen resurssi, jota epäonnistumiset horjuttavat.¹⁷ Jos joku toimii tietoisesti systeemin tavoitetta vastaan ja systeemin toiminta epäonnistuu sen vuoksi, ihmistoimijoiden luottamus toisiinsa ja tavoitetta vastaan toimineen hyvántahtoisuuteen rakoilee.

13 Hallamaa 2017, 41–44.

14 Kalliokoski 2020, 152 ja 156–158.

15 Hallamaa 2017, 65; Kalliokoski 2020, 151–152.

16 Hallamaa & Kalliokoski 2022.

17 Luottamuksen merkityksestä yhteistoiminnalle Elster 2007, 345–352; Hallamaa 2017, 125–149.

Jos epäonnistuminen johtuu toiminnasta, joka rikkoo myös lain säättämiä velvoitteita, tapahtuu rikos, josta tuomittu on lain sääätäneen yhteiskunnan silmissä epäilyttävä yhteistyökumppani. Sattumanvaraisista tekijöistä johtuvan onnettomuuden tapahduttua siitä kärsineiden tai siihen osallisten luottamus todellisuuteen ja maailman turvallisuuteen ja hallittavuuteen on uhattuna. Se, minkä suhteen epäonnistutaan, vaikuttaa siihen, millaisia korjaustoimia tarvitaan, jotta yhteistoimintaa voidaan jatkaa.

Esittelemme seuraavaksi epäonnistumisen syistä tehtyä tutkimusta ja palaamme sen jälkeen erilaisiin tapoihin käsitellä epäonnistumisia sosioteknisissä systeemeissä moraalisen ja juridisen vastuun kannalta.

KUINKA EPÄONNISTUMISTA ON SELITETTY

Epäonnistumisia on tutkittu etenkin osana teollisen tuotannon kehittämistä. Koneellistetussa valmistamisessa jokainen prosessin osa voidaan suunnitella ja toteuttaa niin, että kaikkia askeleita voidaan kontrolloida palvelemaan asetettua tavoitetta. Tarkoin säädelyssä tuotantoketjussa ihminen käyttää hyväkseen luonnon lainalaisuuksia omien tavoitteidensa edistämiseen.

Epäonnistumiseen johtavia tekijöitä on voitu tarkastella järjestelmällisesti ja järkipäisesti vasta sen jälkeen, kun omaksuttiin käsitys, että epäonnistumiset johtuvat ainakin osaksi ihmisen toiminnasta, eivätkä ole kohtalon tai yliluonnollisten voimien aiheuttamia.¹⁸ Toiminnan epäjatkuvuuskohtien paikantaminen epäonnistumiseen johtaviin tekijöihin vakiintui tutkimusalaksi 1910-luvun loppupuolella Yhdysvalloissa. Konetekniikan kehittymisen ansiosta laitteistot voitiin rakentaa toimintavarmiksi, joten epäonnistumisten syitä alettiin etsiä yksilötoimijan virheistä.¹⁹ Vaikka tutkijat ovat 1900-luvun lopulta alkaen havahtuneet siihen, ettei yksilötoimijan virheitä etsimällä voida ymmärtää systeemisten epäonnistumisten luonnetta, yksilötoimijan toiminnan arviointia tarvitaan edelleen, kun epäonnistumisia pyritään ymmärtämään.²⁰

Sosioteknisten systeemien epäonnistumisiin voidaan puuttua, kun tunnistetaan virheiden

perustyyppit, joita voidaan arvioida toiminnan käsitteen peruselementtien – tavoitteen ja keinon – avulla. Teknologiset järjestelmät saavat aikaan kausaalisia vaikutuksia maailmassa, minkä nojalla niitä voidaan pitää kohtuullisen itsenäisinä pseudotoimijoina²¹ tai oikeintoimijoina,²² jotka edistävät jotakin yleensä ihmistoimijoiden määrittämää tavoitetta annettujen normien mukaan. Sovellamme niiden toiminnan ja yhteistoiminnan epäonnistumiseen virheluokitteluita, joita James Reason on eritellyt urauurtavasti ihmistoimijoiden tekemistä virheistä.²³

Tavoitteen saavuttaminen edellyttää, että toimijalla on riittävän oikea käsitys todellisuuden rakenteesta ja sen osien välillä vallitsevista kausaalisuhteista. Sen varassa toimija voi tehdä

- 18 Guarnieri 1992, 152; Kölbl-Ebert 2005, 83–85. Ks. myös Neiman (2002, 14–18) joka kuvaa, kuinka Lissabonin 1755 tuhonnut maanjäristys käynnisti keskustelun luonnononnettomuuksien syistä. Tapahtumat Jumalan rangaistuksesi tulkitsevien selitysten sijaan ryhdyttiin pohtimaan ihmisten mahdollisuuksia estää ja lievittää onnettomuuksia. Lissabonissa suuri osa tuhoista ei johtunut itse maanjäristyksestä vaan lyhytnäköisestä maankäytöstä ja surkeasta paloturvallisuudesta.
- 19 Reiman & Oedewald 2008, 39–42.
- 20 Inhimillisiä tekijöitä (engl. human factors) ja niiden osuutta turvallisuuden ylläpitämisessä ja turvallisuuden heikentymisessä tutkitaan edelleen monitieteisessä turvallisuustutkimuksessa. Esim. Teperi et al. 2023 ovat kiinnittäneet huomiota läheiseen yhteyteen, joka organisaatiolla, ryhmällä, työtehtävillä ja työn luonteella on inhimillisiin tekijöihin.
- 21 Hallamaa & Kalliokoski 2020.
- 22 Antti Kauppinen (2021) ehdottaa autonomisille tekoälyjärjestelmille *oikeintoimijan* kategoriaa: ne kykenevät noudattamaan niille asetettuja sääntöjä, mutta eivät välttämättä kykene muotoilemaan moraalisia arvostelmia tai kantamaan moraalista vastuuta.
- 23 Reasonin klassikkoteoksen *Human Error* ensimmäinen painos julkaistiin 1990.

suunnitelman muuttaa todellisuutta niin, että se hänen toimintansa vaikutuksesta vastaa paremmin sitä, millainen hän toivoo sen olevan.²⁴ Sosioteknisen systeemin osana toimivan ihmisen ymmärrys syy- ja seuraussuhteista perustuu usein teknisten järjestelmien välittämään informaatioon. Se koostuu mittauksiin perustuvista tietosyötteistä, jotka yhä autonomisemmiksi suunnitelluissa systeemeissä ohjaavat koko järjestelmää ilman, että ihminen joutuu puuttumaan siihen. Näin systeemi muuttuu ihmisapuolten kannalta osaksi olosuhteita, joihin reagoimisesta tulee toiminnan keskeinen tavoite.

Jos käsitys todellisuudesta ja siinä vallitsevista vaikutussuhteista on virheellinen esimerkiksi puutteellisen tilannetiedon tai teknistä systeemiä ohjaavien, mutta käyttäjälle tuntemattomien tai väärin tulkittujen toimintaperiaatteiden vuoksi, toimijan on vaikea tai mahdotonta laatia sellaista keinoja koskevaa suunnitelmaa, jonka avulla hän voisi saavuttaa tavoitteensa.²⁵ Vääränlaiseen suunnitelmaan pohjautuvan epäonnistumisen välttäminen edellyttää ajatustyötä: toimijan on keskityttävä havainnoimaan ja pohtimaan, miltä osin ja miksi hänen käsityksensä todellisuudesta ja siinä vallitsevista vaikutussuhteista on väärä.²⁶

Erinomaisenkin suunnitelman toteutus voi kuitenkin epäonnistua toimijoiden taitamattomuuden, oikeasta tekosarjasta lipeämisen tai erehtymisen vuoksi. Taitamaton toimija ei hallitse tekoja teknisinä suorituksina niin, että hän saisi aikaan sen, mitä tavoittelee (skill-based failure).²⁷ Teknologisten toimijoiden taitopohjaiset virheet jäävät usein kehitysvaiheeseen ja laboratorioon, sillä laitteita ja niiden luotettavaa käyttöä varten on säädetty tarkat tuoteturvallisuus- ja laatumääräykset erityisesti turvallisuus-kriittisillä aloilla.²⁸ Ennakoimattomissa tilanteissa, esimerkiksi luonnonilmion yllättäessä, teknologisella toimijalla ei aina ole riittävästi taitoa tulkita tilannetta ja toimia sen vaatimalla, luovaa ongelmanratkaisukykyä edellyttävällä tavalla.²⁹

Hyvät taidotkaan eivät takaa onnistumista, jos toimija ei onnistu soveltamaan niitä oikein.

Epäonnistuminen voi johtua siitä, että toimija lipeää (slip) tavoitteeseen johtavalta polulta, kun päämäärän kannalta tarkoituksenmukaiseen tekojen sarjaan sekaantuu käsillä olevan tehtävän kannalta epätarkoituksenmukainen tekoyksikkö jostain toisesta suorituskokonaisuudesta. Lipeämisestä on kyse, kun autoilija näkee tiellä esteen ja aikoo jarruttaa, mutta jalka osuu jarrun sijaan kaasulle, eikä auto pysähdykään, vaan ampaisee kovempaan vauhtiin.³⁰

Virhe voi olla myös *erehdys* (lapse): henkilö suorittaa toimintakokonaisuuteen kuuluvia tekoyksiköitä, mutta väärässä järjestyksessä.³¹ Uuden älypuhelimien hankkinut turvautuu tekosarjaan, johon on tottunut käyttäessään aiempaa, toisentyypistä laitetta. Jos älylaitteiden toimintaperiaatteet poikkeavat toisistaan, aiemmin omaksutun toimintamallin toteuttaminen ei tuota haluttua tulosta.

Sekä lipeämiset että erehdykset ovat yleisiä ja luonteeltaan triviaaleja virheitä, mutta ne voivat olla kohtalokkaita, jos ne tapahtuvat ympäristössä, jossa moitteettomasti etenevä, standardoitu

- 24 Hallamaa 2017, 29–39; Reason 1990, 5.
 25 Esimerkiksi oppivat tekoälyjärjestelmät voivat luoda nk. mustan laatikon ongelman, jos ihmistoimijat eivät ymmärrä järjestelmän tekemien päätösten tuottamismekanismia, päätösten perusteita eivätkä osaa siten mukauttaa omaa toimintaansa teknologisen systeemin toimintaan järkevällä tavalla. Mustan laatikon ongelmaa on yrittänyt ratkaista mm. Laitinen 2021.
 26 Reason 1990, 9–10.
 27 Reason 1990, 55–56.
 28 Jatkuvasti päivitettävä laitedirektiivi (*Machine Directive*) on keskeinen EU-tason säätelyväline. Directive 2006/42/EC (2006).
 29 Haikarainen 2022, 274.
 30 Reason kiinnostui virheistä, kun hän havaitsi lusikoivansa teelehtiä kissanruokakuppiin: teenkeiton lomassa mieleen muistunut kehotus ruokkia kissa johti virhesuoritukseen. Reason 2013, 1–5.
 31 Reason 1990, 11–13.

toiminta on tärkeää, esimerkiksi leikkaussalis-
sa.³² Reasonin mukaan tällaisia virheitä tekevät
erityisesti tehtävissään rutinoituneet osaajat.
Kun automatisoituneeseen suoritukseen ei enää
tarvitse keskittyä siitä selvittääkseen, toimijan
ajatukset eksyvät tehtävästä johonkin muuhun,
jolloin eri taidot tai kontekstit voivat sekaantua
keskenään. Aloittelijat puolestaan tekevät enem-
män puutteelliseen suunnitteluun pohjautuvia
virheitä kuin kokeneet toimijat.³³

Reasonin yleinen virheiden luokittelujärjes-
telmä (Generic Error-Modelling System) kuvaa
vain toiminnan suunnitteluun ja toteuttamiseen
kuuluvia seikkoja, eikä ota kantaa moraalisiin
kysymyksiin.³⁴ Reasonin asiantuntemus ei kui-
tenkaan ole vain tieteellistä, vaan hän on myös
erittäin kokenut onnettomuustutkija.³⁵ Faktojen
ja erilaisten syy-yhteyksien selvittäminen irral-
laan moraalista näkökulmasta ja vastuukysy-
mysten arvioinnista on onnettomuustutkinnan
keskeinen lähtökohta ja toimintaperiaate.³⁶
Näin pyritään takaamaan, ettei tarve arvioida
syyllisyyskysymystä vääristä kausaalisten
tapahtumakulkujen selvittämistä ja johda fakta-
aineiston valikoimiseen. Todennettavat tosiasiat
tarjoavat perustan myös oikeudelliselle ja
moraaliselle arvioinnille, kun niiden merkitystä
punnitaan päätettäessä, onko perusteita aloittaa
rikosprosessi ja määrätä vahingonkorvauksia.

SOSIOTEKNISISSÄ SYSTEEMEISSÄ EPÄONNISTUMISET EMERGOITUVAT

Teollisten prosessien turvallisuustutkimuksessa
on havaittu, etteivät epäonnistumiset suinkaan
aina selity yhden ihmisen virheillä, vaan monien
ihmisten ja teknisten järjestelmien yhteistoi-
minnan systeemiset tekijät ovat aiheuttaneet
ne. Kasvattamalla tuotantojärjestelmää ja pilk-
komalla sitä eri tehtävistä huolehtiviin osiin
toimintoja voidaan optimoida, mutta samalla
enennetään riskejä: toiminnan epäonnistumi-
nen systeemin yhdessä osassa voi levitä hallitse-
mattomasti sen muihin osiin. Järjestelmätason
epäonnistuminen voi aiheuttaa valtavia vahin-
koja myös ulkopuolisille.³⁷

Suurten teollisuusonnettomuuksien³⁸ tut-
kinta on osoittanut, että selityksen löytämiseksi

on tarkasteltava kokonaisuutta, joka koostuu
ihmistoimijoiden yhteistyöstä suhteessa järjes-
telmäteknikkaan.³⁹ Keskittyminen heikoimman
lenkin etsimiseen perustuu jälkikäteen hah-
moteltuun kuvaukseen tapahtuman kannalta
merkityksellisistä seikoista, mutta ei ota huomi-
oon tilannetta sellaisena, jollaisena se on siinä
mukana olleille näyttäytynyt.⁴⁰

Kasvaneita epäonnistumisten riskejä on
yritetty hallita luomalla paitsi parempia tapoja
suunnitella sosioteknisten systeemien osasten
yhteistoimintaa myös luomalla malleja sys-
teemisen turvallisuuden ylläpitämiseen. Osa-
tekijöiden toimintaa pyritään kontrolloimaan
suoraviivaistamalla komentoketju. Esimerkiksi
Nancy Levesonin turvallisuusmallissa systeemin
eri osien välillä kulkevat viestintävirrat järjes-

32 Haikarainen 2022, 210–222.

33 Reason 1990, 58–59.

34 Reason 1990, 61–63.

35 Reason 2013.

36 Ks. Haikarainen tässä julkaisussa.

37 Ks. esim. Pallardy (2022) Deepwater Hori-
zon öljynporauslautan uppoamisen aiheut-
taman massiivisen ympäristö- ja terveyska-
tastrofin tutkinta.

38 Englanninkielisessä turvallisuustutkimuk-
sessa epäonnistumisista, joilla on vakavia
seurauksia, on ryhdytty käyttämään acci-
dent-sanana asemesta vauriota ja vahingol-
lista tapahtumaa tarkoittavaa injury-sanana.
Näin vältetään latinan *accidens*-sanasta
juontuva merkitysvivahde tapahtuman
ennakoimattomuudesta ja vääjäämättömyy-
destä. Guarnieri 1992, 153–154. Myös suomen
sanoilla onnettomuus ja vahinko on har-
haanjohtavia sivumerkityksiä: puhe onnesta
ja onnettomuudesta tuottaa mielleyhtymän
siitä, että tapahtumia säätelevät (ainakin
osittain) voimat, joita ihminen ei voi hallita
tai ennakoida. Sanan vahinko konnotaatiot
saavat ajattelemaan, että epäonnistumiset
tulevat sattumalta kohdalle. Kielitoimiston
sanakirja s.a.

39 Taylor et al. 2016; Taylor et al. 2017.

40 Guarnieri 1992, 152–153.

tetään tarkan hierarkian mukaan. Toisenlaisen lähestymistavan systeemisen turvallisuuden ylläpitämiseen on esittänyt Karl Weick tutkimusryhmineen. Siinä systeemin vakaa ja turvallinen toiminta perustuu ihmistoimijoiden jatkuvaan tietoisuuteen meneillään olevista tapahtumista, kun toimintaan osallistuvat havainnoivat organisaation ja laitteiden lähettämiä heikkoja signaaleja. Vähäisistä muutoksista ja poikkeamista raportoimalla ja niitä analysoimalla pyritään vahvistamaan organisaation kykyä reagoida yllättäviin ja mahdollisesti kohtalokkaisiin muutoksiin.⁴¹

Sosiotekninen epäonnistuminen voidaan usein paikantaa ihmistoimijoiden ja ei-inhimillisten teknologisten toimijoiden välisen yhteistyön puutteisiin. Ihmiset prosessoivat tietoa ja havainnoivat ympäristöään eri tavalla kuin teknologiset systeemit, eivätkä ihmiset siksi aina osaa tulkita niiden tuottamia viestejä oikein, mikä voi johtaa vaaratilanteita aiheuttaviin reaktioihin. Psykologi J. J. Cañas huomauttaa, että yhteistoiminnalla voi olla eettisesti väärää seurausta (ethically incorrect consequences), vaikka yhteistoiminnan ihmis- ja teknologiaosapuolten oma toiminta olisikin eettisesti hyväksyttävää. Näin sosioteknisen systeemin epäonnistuminen voi emergoitua ihmis- ja pseudotoimijoiden oikein toimimisesta huolimatta.⁴²

Kausaaliyhteyksiä rekonstruovat reduktionistiset selitykset etsivät yksilötoimijoiden virheitä vahinkoa edeltäneestä toiminnasta, kun taas systeemistä näkökulmaa hyödyntävät emergenttiset mallit perustuvat havaintoon, että toiminta voi epäonnistua katastrofaalisella tavalla, vaikkei kukaan siihen osallistunut olisi tehnyt – ainakaan erityisen suurta – virhettä. Virheen etsiminen on ymmärrettävää, koska sen avulla vastuukysymysten määrittäminen helpottuu. Systeeminen tarkastelu ja emergenttiset mallit puolestaan näyttävät vaikeuttavan vastuullisuuden arvioimista tai tekevän sen mahdottomaksi etenkin suurissa sosioteknisissä järjestelmissä, kun niissä on luovuttu mahdollisuudesta osoittaa, mikä yksittäisten toimijoiden osuus – ja vastuu – on ollut epäonnistumiseen johtaneessa toimintakokonaisuudessa.⁴³

Tarkastelemme seuraavaksi tarkemmin vastuun käsitettä.

VASTUUNKANTAMINEN ON SUHDETOIMINTAA

Vahinkoon johtaneen virheen tehnyttä pidetään lähtökohtaisesti moraalisesti ja rikosoikeudellisesti vastuunalaisena, mikä tarkoittaa sitä, että hänen on vastattava toimintansa seurauksista ja vaikutuksista. Taaksepäin, menneeseen suuntautuva vastuu voidaan erottaa eteenpäin, tulevaisuuteen suuntautuvasta vastuusta eli velvollisuuksista hoitaa lain, roolien, sopimusten ja lupausten varassa toteutettavat tehtävät. Menneeseen suuntautuvassa vastuussa maallikoita kiinnostavat syyllisten löytäminen ja rankaiseminen. Filosofit taas ovat muotoilleet välttämättömiä ja riittäviä ehtoja sille, että toimija voidaan vaatia vastuuseen. Kummankin aiheen keskiössä on vastuun *locus*.⁴⁴

Tässä artikkelissa vastuuta tarkastellaan suhteessa toteutuvana vastuunkantamisena.⁴⁵ Arkikielen ilmaiset ”jakaa vastuuta”, ”antaa vastuuta” ja ”saattaa vastuuseen” osoittavat, että kyseessä on vuorovaikutteinen toiminta, joka luo

41 Kahta systeemiturvallisuuden mallia vertailevat Lintern & Kugler 2017.

42 Ihmisen ja tekoälyteknologian yhteistoimintavirheiden vaikutuksista ks. Elish 2019.

43 Reiman et al. 2015; Dekker & Nyce 2011.

44 Emmanuel Lévinas edustaa filosofiassaan paljon laajempaa näkemystä vastuusta, jonka lähtökohta ovat toisen kasvat, kielto ”älä tapa” ja vaatimus tehdä lähimmäiselle hyvää. Vastuu siis edeltää toimintaa eikä määrity toiminnan kausaalisten seurausten mukaan, ks. esim. Burggraeve 1999, 32–34. Näin ymmärretyn vastuukäsitteen ottaminen epäonnistumisen tarkastelun pohjaksi ei kuitenkaan ole mielekäästä.

45 Yazdanpanah et al. (2021, 16) puhuvat monen toimijan systeemin vastuusta suhdekäsitteenä, mutta käsittelevät vain vastuullisen toimijan ja vastuunalan välistä suhdetta.

suhteen. Vastuussa olemisen moraalista merkitystä kuvastaa se, että vastuu painaa taakkana, jota joutuu kantamaan. Vastuuseen joudutaan, kun tehtävien hoitaminen epäonnistuu eivätkä aiotut tavoitteet toteudu. Vastuuseen saattaminen on osa prosessia, jossa epäonnistumisen syytä, seurauksia ja seuraamuksia tarkastellaan.

Vastuusuhteessa kietoutuvat toisiinsa vastuunalainen taho, vastuun ala sekä vastuun antava taho. Niiden välille muodostuu kolmi-paikkainen relaatio:

R(a, c, b)⁴⁶

Lausekkeen R tarkoittaa vastuusuhdetta (responsibility), a on *vastuunkantaja*, c *vastuun ala* ja b on *vastuuttaja*, taho, jolle a on vastuussa c:stä. Luonnollisella kielellä lauseke ilmaisee, että ”a on vastuussa c:stä b:lle”. Vastuusuhteen muodostuminen edellyttää, että vastuunkantaja ja vastuuttaja tunnistavat oman ja suhteen toisen osapuolen aseman. Vastuunkantaja ja vastuuttaja voivat olla sama toimija: itsensä vastuuttaminen ja vastuun kantaminen omasta toiminnastaan on moraalisen toimijuuden edellytys. Retrospektiivisesti vastuunkantotoimien määrääminen ja niiden oikeasuhtaisuuden hyväksyminen edellyttävät, että b ottaa ne vastaan.

Mikä tekee toimijasta kelvollisen vastuunkantajan niin, että hänelle voidaan sälyttää vastuuta tulevista toimista tai hänet voidaan saattaa vastuuseen menneistä tapahtumista? Vastaukseksi on tarjottu kriteerejä, joilla punnitaan vastuulliseksi nimettävän toimijan episteemisiä kykyjä, mentaalaisia tiloja tai tietoa moraalinormeista ja sitoutumista niihin. Episteemisillä kyvyillä tarkoitetaan esimerkiksi tietoa ympäristöstä, omasta osaamisesta ja päätöksentekostrategioista. Mentaalisiin tiloihin sisällytetään toimijan kyky toimia intentionaalisesti ja teon tai tekosarjan tahallisuus jonkin spesifin vastuunalan suhteen. Jotta toimijaa voidaan pitää moraalisesti vastuullisena, häneltä yleensä edellytetään myös tietoa jaetuista moraalikäsitteistä ja normeista.⁴⁷

Jaana Hallamaan mukaan vastuullista toimijaa luonnehtii kolme ehtoa. Toimijan on pystyttävä saamaan aikaan tavoittelemiaan

kausaalisia vaikutuksia (kyvykkyysehto), kyettävä ymmärtämään tekemistensä ja niiden vaikutusten välinen yhteys (rationaalisuusehto) ja toimijalla on lisäksi oltava kyky asettaa tavoitteita (intentionaalisuusehto). Etenkin viimeinen ehto aiheuttaa pulmia silloin, kun on ratkaistava, missä määrin toimijat ovat vastuussa niistä tekojensa vahingollisista seurauksista, joita he eivät ole aikoneet.⁴⁸

Kun tarkastellaan spesifiä epäonnistumistapahtumaa jälkikäteen ja pohditaan, onko perusteita pitää jotakin toimijaa moraalisesti tai juridisesti vastuunalaisena, on kyettävä osoittamaan toimijan ja tapahtuman välinen kausaalinen yhteys. Lisäksi on kyettävä todentamaan, että toimija olisi tosiasiallisesti voinut toimia myös toisin ja siten välttää epäsuotuisan tapahtuman.⁴⁹ Tarvitaan faktoja kausaalisesta tapahtumainkulusta, jota onnettomuus- ja rikostutkintakin pyrkivät selvittämään. Vastuunalaisuuden osoittaminen edellyttää perustetta sanoa, että toimija on tehnyt teon, joka on aiheuttanut vastuusuhteen kannalta merkityksellisen vaikutuksen.

Pelkkä teon kausaalinen aikaansaaminen ei aina riitä perusteeksi edellyttää vastuun kantamista. On otettava huomioon tapahtumiin vaikuttaneet tilannetekijät ja toiminnan tavoitteel-

46 Pelkonen 2012; Hallamaa & Kalliokoski 2022.

47 Dastani & Yazdanpanah (2022) erottelevat toisistaan toimijakeskeisen ja tapahtumakeskeisen vastuukeskustelun, joita voidaan pitää myös saman vastuupulman eri puolina.

48 Hallamaa 2017, 83–84. Tekoöllysystemit sosioteknisten systeemien osana haastavat kaikkia näitä ihmistoimijuuden ehtoja. Hallamaa & Kalliokoski 2020; Hallamaa & Kalliokoski 2022, 416–417.

49 Braham & van Hees (2018) nimittävät näitä kausaalisuusehdoksi ja välttämisehdoksi. Välttämisehdolla kiteytetään vuosisatainen filosofinen ja teologinen keskustelu ihmisen vapaudesta valita toimintansa. Talbert 2019.

lisuus. Jos toimijalle on annettu mahdollisuus lukea allekirjoittamansa sopimus, mutta hän epäonnistuu siihen kirjattujen velvollisuuksien hoitamisessa, on perusteltua sanoa hänen olleen tietoinen vastuunalastaan sopimuksen toiselle osapuolelle. Jos toimijan epäonnistuminen on vahinko, hän voi kiistää vastuunalaisuutensa.

Kahvia juodessaan Anna läikyttää sitä Beatalta lainaamalleen älylaitteelle, joka rikkoutuu. Hän on rikkoutumisen syy, *causa*, mutta voi torjua vastuunkantamisen edellyttämät korvausvaatimukset sanomalla, ettei hänen tarkoituksensa ollut särkeä laitetta. Laitteen lainannut Beata voi perustellusti sanoa Annan valinneen sopimattoman paikan kahvihetkelle. Vastuunkantamisen toteutuminen edellyttää vastuunalaisuuden tunnistamista.

Tarkastelussa olemme keskittyneet toiminnallisiin epäonnistumisiin, mutta epäonnistuminen voi olla myös juridista ja moraalista. Toiminnan juridinen epäonnistuminen on sitä, että lakia on rikottu joko tekemällä jotain lainvastaista tai jättämällä tekemättä jotain, mihin laki velvoittaa. Juridisen vastuusuhteen syntyminen edellyttää, että toimijaa voidaan kohdella oikeushenkilönä, jonka vastuu suhteessa lakiin on yleinen: vastuun ala on toimia lain velvoittamalla tavalla ja olla rikkomatta lakia pidättymällä teoista, jotka ovat lainvastaisia. Juridisen vastuun asettaja on yhteiskunta kokonaisuudessaan, ja oikeuskoneiston tehtävä on valvoa, että vastuutaan laistavat saatetaan siitä edesvastuuseen. Rikkomus voi aiheuttaa vahinkoa, jota kärsineistä tulee asianomistajia, ja vastuuseen saattamiseen kuuluu silloin korvauksen tai hyvityksen suorittaminen heille.

Toiminnan moraalinen epäonnistuminen on moraalinormin vastaista toimintaa. Vastuunalaisuuden arvioimisen kannalta on tärkeää selvittää, voidaanko toimijan sanoa toimineen tahallisesti vai tahattomasti. Moraali-instituutio perustuu sosiaaliin sitoumuksiin, ja moraalinormien vastaisesti toimivia pidetään vastuullisina muille ja yhteisölle mutta myös itselleen.⁵⁰

Systemisiä epäonnistumisia tarkastelevassa tutkimuksessa ei yleensä käsitellä juridisia ja moraalisia epäonnistumisia, jotka perustuvat

usein toimijan tietoiseen pyrkimykseen tehdä asioista omalta kannaltaan yksinkertaisempia, hankkia itselleen jotain oikeudetonta etua tai oikoa mutkia. Juridisesti ja moraalisesti moitittavan epäonnistumisen syynä voi olla myös omien epäonnistumista estävien tai turvallisuutta ylläpitävien tehtävien laiminlyöminen⁵¹ tai liiallinen tai väärin kohdistunut luottamus toisten ihmisten hyvántahtoisuuteen tai riskien todennäköisyyteen. Tällaiset epäonnistumiset voi niin halutessaan välttää, eikä toimijan väitettä, että hän toimi rikollisesti tai moraalin vastaisesti vahingossa, voida pitää uskottavana.

Yhteiskunnassa erilaiset vastuut on sidottu eri organisaatioissa hoidettaviin virkoihin ja rooleihin, ja vastuun ala määräytyy niiden mukaan. Organisaation toimintaa voidaan arvioida tutkimalla, kuinka hyvin vastuista huolehditaan.⁵² Vastuu on näin keskeinen käsite myös sosioteknisten järjestelmien toiminnan tarkastelussa.

Vastuiden hahmottaminen vaikeutuu, kun joudutaan pohtimaan, kuinka vastuut tosiasiasa jakautuvat. Johtoasemassa olevien sanotaan usein kantavan muita suurempaa vastuuta. Epäonnistumisten sattuessa hierarkian mukaiset vastuusuhteet eivät kuitenkaan toteudu, jos epäonnistumisten syytä etsitään suorittavan portaan yksilötoimijoista, kuten vakavien onnettomuuksien kohdalla usein tehdään.⁵³ Systemisestä näkökulmasta työntekijöiden panos perustuu organisaation sääntöihin, kulttuuriin ja resursseihin, joista ennen muuta johto on vastuussa. Työntekijät ovat voineet ennen onnettomuuden tapahtumista tuoda esiin puutteita

50 Haklin & Mäkelän (2019) mukaan vastuullisen toimijuuden keskeinen kriteeri on moraaliyhteisön jäsenyys, minkä vuoksi tekoälytoimijoita ei voida pitää vastuullisina agenteina.

51 Vakava tietoturvan laiminlyöminen teki mahdolliseksi Psykoterapiakeskus Vastaamon tietomurron. Valviran tiedote 19.5.2021.

52 Bovens 1998, 11.

53 Elish 2019.

ja turvallisuusriskejä, joihin johto on kuitenkin ollut haluton puuttumaan.⁵⁴

MITEN VASTUUTA EPÄONNISTUMISESTA JAETAAN

Sosioteknisissä systeemeissä epäonnistumisen syy on usein se, ettei toimijoiden ja erityisesti ihmisen ja teknologisen systeemin yhteistoiminta suju suunnitellusti.⁵⁵ Jos osapuolina on ollut tekoälytoimijoita, kausaaliketjuja ei aina ole mahdollista rekonstruoida. Tekoälysovellukset hoitavat monia tehtäviä melko itsenäisesti etenkin, jos ne ohjelmoidaan oppimaan omasta toiminnastaan.⁵⁶ Silloin voidaan tarkastella tyypikkauselektioita eli tilastollisia yhteyksiä.⁵⁷ Vastuunalaisuuden edellytyksinä edellä esitellyt kausaalisuus- ja intentionaalisuus- eivätkä näytä riittävän vastuunalaisuuden määrittämiseksi tällaisissa tapauksissa. Toimintaan osallistuvat saavat aikaan erilaisia muutoksia ja vaikutuksia toisissaan ja ympäristössään, ja vaikka jonkun mukana olevan vaikuttamistoimet eivät jossain tilanteessa muuttaisi fyysikaalista todellisuutta, hänkin on osallinen lopputuloksen aikaansaamiseen tai vaikutusten muodostumiseen.

Christopher Kutzin mukaan jaettu suunnitelma, vaikka toinen osapuoli ei edistäisikään sitä, tekee suunnittelijat vastuunalaisiksi lopputuloksesta ja sen seurauksista.⁵⁸ Valaisemme asiaa esimerkillä:

Ove ehdottaa Avan sähköauton kyydissä istuessaan, että tämä näyttäisi, kuinka nopeasti auton ajojärjestelmä kiihdyttää sen kovaan vauhtiin. Ava innostuu, antaa autolle komennon, mutta poliisi pysäyttää heidät. Onko Ove vastuussa liikenneturvallisuuden vaarantamisesta ja sen seuraamuksena annetuista sakoista? Ilman Ovea Ava ei olisi tullut ajaneeksi ylinopeutta. Ei ehkä ole perusteita sanoa, että Ove ja Ava olisivat *toimineet yhdessä*, mutta voidaan sanoa, että heillä oli yhteinen intentio, jonka Ava toteutti kiihdyttämällä autoaan, minkä vuoksi häntä sakotettiin. Ove osallistui ajatuksen esittämiseen ja yhteisen intention perusteella lopputuloksesta aiheutuneisiin vaikutuksiin, vaikka ei itse kiihdyttänyt autoa. Ove oli toiminnassa osallinen.

Vaikka ei-biologiset toimijat kuten tekoälyohjelmistot ja muut kehittyneet tekniset systeemit eivät pysty toimimaan suunnittelijoiden asettamien tavoitteiden toteutumista, nekin voivat silti olla keskeisiä toimeenpanevia tekijöitä ja osallisia epäonnistumiseen.⁵⁹ Millä perusteella teknologisia pseudotoimijoita voidaan pitää vastuusuhteen osapuolina? Filosofit Christian List rinnastaa tekoälyn vastuullisena pitämisen pulmallisuuden ongelmiin, joista on keskusteltu kollektiivin vastuuta pohdittaessa.

Kollektiivin vastuuta määriteltäessä filosofit ovat tunnistanee epäjatkuvuuskohdan, jota on ruvettu kutsumaan vastuuaukoksi (responsibility gap). Vastuuaukon käsite kuvaa menneisyyteen suuntautuvan vastuun määrittämisen vaikeutta tilanteessa, jossa kollektiivin jäsenet saavat yhdessä aikaan vahingotapahtuman, jota kukaan yksittäinen jäsen ei ole aikunut tuottaa, eikä kukaan siksi tunnusta – tai ehkä edes tiedosta – vastuunalaisuuttaan. Jos vastaava tapahtuma olisi ollut yhden ihmistoimijan aikaansaama, häntä pidettäisiin vastuunalaisena.⁶⁰

Vastuuaukko voi ilmetä niin ihmiskollektiivin kuin tekoälyohjelmistonkin aikaansaamien vaikutusten jälkiselvittelyssä.⁶¹ Vaikutusten

54 Dekker & Nyce 2012. Ks. myös Final Committee Report 2020.

55 Cañas 2022.

56 Automatisoitujen toimintojen tuottamista erityyppisistä ongelmista ks. Haikarainen 2022, 251–263.

57 Dastani & Yazdanpanah 2022.

58 Kutz (ks. Lawon 2013) edustaa käsitystä heikosta kollektiivisesta yhteistoiminnasta. Käsitystä vahvasta kollektiivisesta yhteistoiminnasta kannattavan Raimo Tuomelan (2007, 233–251) mukaan tiiviin ryhmän jäsenyys tekee ryhmään kuuluvat vastuunalaisiksi ryhmän tai sen yksittäisen jäsenen väärästä toiminnasta.

59 List 2021; Dastani & Yazdanpanad 2022.

60 Vastuuaukon ontologiasta Himmelreich 2019, 734. Ilmiöstä puhutaan myös *vastuutyhjiönä* (responsibility void). Braham & van Hees 2018.

emergoituminen sosioteknisessä systeemissä johtaa siihen, ettei kukaan tunnista aiheuttaneensa niitä eikä ole valmis ottamaan niistä moitteita. Vaikka teknologiset sovellukset toimisivat autonomisesti, niillä ei ole eikä niille voida ohjelmoida kykyä tunnistaa tai tuntea moraalista vastuuta. Myöskään ihmistoimijat eivät tunnista aiheuttaneensa vahinkoja tai ainakaan saaneensa niitä yksin aikaan. Perinteisen vastuukäsitteistön ja vastuunalaisuuden kriteerien perusteella ihmiset voivat kiistää, että heidän toimintansa täyttää vastuunalaisuuden edellyttämän kausaalisuusehdon. Käytännössä ilmenevää pulmaa on kuvattu monien käsien ongelmaksi, joka voi tosin ilmaista myös ihmistoimijoiden tietoista vastuunalaisuuden pakoilua.⁶² Ryhmien ja sosioteknisen systeemin epäonnistuuksessa vastuu ei näytä olevan kenelläkään.

Vastuuaukon täyttämiseen on ehdotettu *ryhmävastuun käsitettä*, jonka varassa järjestäytyneet organisaatio voidaan saattaa vastuuseen kollektiivina eikä vain yksittäisinä henkilöinä. Organisaatioita voidaan juridisesti myös kohdella kuin persoonia (oikeushenkilö), vaikka niitä ei ontologisesti pidettäisi sellaisina.⁶³ Ryhmävastuussa olevien ihmistoimijoiden ja teknologisten pseudo toimijoidenkin kohdalla jää ratkaisematta, miten vastuu kannetaan ja jaetaan. Siksi on tarpeen eritellä ja näin auttaa toimijoita tunnistamaan erilaiset yhteistoimintaroolit ja vaikutussuhteet, joissa he ovat osallistuneet systeemiseen toimintaan.

Eteenpäin suuntautuva jaettu vastuu perustuu siihen, että useammalla toimijalla, ihmisillä tai teknologisilla toimijoilla, on ryhmässä jokin tehtävä. Sen merkittävyys suhteessa koko ryhmän tehtävään määrittää yksittäisten jäsenten vastuita, kun yhteinen toiminta tai yhteisen tavoitteen saavuttaminen epäonnistuvat. Vastuiden tosiasiallista jakautumista on vaikea selvittää, jos ei ole ensin määritely, mikä ryhmän tehtävä on, kuinka sitä hoidetaan ja mikä kunkin osuus siinä on.

Jonkin tehtävän hoitamiseksi määritellyssä roolissa toimivien ihmisten vastuu säilyy silloinkin, kun epäonnistumisen aiheuttaa tekoäly tai kollektiivi. Vastuun niiden epäonnistumi-

sesta jakavat Listin mukaan ihmiset, joilla on toiminnan mahdollistava rooli, kuten manageri tai systeemin operaattori, mikäli he toimivat ilman pakkoa, tietoisesti ja omasta halustaan. Vastuussa ovat myös auktoriteettiasemassa toimivat omistajat, tuottajat ja tekoälyn säätelijät. Esimerkiksi hallituksen jäsen, omistaja tai osakkeenomistaja vastaavat kollektiivin ja tekoälyn toiminnan normeista. Kolmas vastuunkantajaryhmä ovat organisaation tai systeemin suunnittelijat ja tuotantoon osallistuneet. Näissä rooleissa toimivilla on tehtäviensä mukaiset ja laakisäteiset velvoitteet lopputuloksen aikaansaamisessa. Myös hyvään ja moraalisesti kestävään toimintaan ohjaava ammattietikka määrittää siihen sitoutuneiden ihmisten vastuunalaa. Vastuuaukkoja voi silti ilmaantua: ihmistoimijoiden joukkoa ei aina pidetä samalla tavalla vastuunalaisena systeemin aikaansaamasta vahingosta kuin yksittäistä ihmistoimijaa.⁶⁴

Vastuurelaation osapuoleksi voi päätyä myös passiivisen vastuun⁶⁵ perusteella eli osallistumalla toimintaan epäsuorasti. Filosofi Gregory Mellema⁶⁶ esittää yhdeksän tapaa osallistua väärintekemiseen — ja tulla vastuunalaiseksi siitä. Mellema tukeutuu Tuomas Akvinolaisen *Summa theologiae* -teoksessa erittelemiin perusteisiin, joiden mukaan muiden kuin varkaan olisi osallistuttava anastetun omaisuuden palauttamiseen tai korvaamiseen. Mellema soveltaa Akvinolaisen perusteita moderniin maailmaan.

Jaottelimme Melleman esittämät osalliseksi tulemisen tavat kolmeen luokkaan myötävaikuttavaksi, kannustavaksi ja reaktiiviseksi osallisuudeksi sen mukaan, millaisen vaikutussuhteen ne luovat toimijoiden välille. Pohdimme, mitä nämä osallisuuden tavat voisivat tarkoittaa teknologisen systeemin jaetun vastuun kohdalla.

61 List 2021.

62 Elish 2019; Noorman 2021; Hallamaa & Kalliokoski 2022.

63 List 2021; Bovens 1998, 58–59.

64 List 2021.

65 Termi Bovensilta 1998.

66 Mellema 2016, 18–26.

Toimijasta voi tulla vastuunalainen toimintaan myötävaikuttava osapuoli sillä perusteella, että hänellä on tieto toisen suunnitelmasta toimia väärin, eikä hän pyri estämään sen toteuttamista, vaikka hänellä olisi siihen moraalinen velvollisuus, kuten vanhemmalla suhteessa lapseensa. Myös avustava tai toimintaa tukeva rooli merkitsee epäsuorasti osallistumista toimintaan.⁶⁷ Onko esimerkiksi teknologisen systeemin kehittänyt ja markkinoille saattanut taho osallinen teknologisen sovelluksen käytön aiheuttamista vahingoista? Jos kehittäjä ja markkinoija tietävät, että systeemillä voidaan aiheuttaa haittaa tai että sitä voidaan helposti käyttää väärin, mutta eivät ryhdy varotoimiin tällaisen käytön estämiseksi, he ovat osallistuneet epäsuorasti vahingon aiheuttamiseen.

Kannustava vastuurelaatio syntyy, kun toimija käskee muiden tehdä jotain tai ohjaa heitä siihen neuvoillaan. Osallistuminen voi olla myös toiminnan kehumista tai rohkaisemista toimintaan.⁶⁸ Voitaisiko teknisten sistemien markkinointia ja kritiikitöntä kehumista pitää tällaisen osallisuussuhteen perustana? Kesken-eräisen tai osittain toimimattoman teknologian markkinoiminen ja sen toiminnasta annetut ylimitoitettut lupaukset voivat kannustaa ja rohkaista ihmistoimijoita käyttämään viallista, vahingollista tai huonosti toimivaa teknologiaa tai antamaan toimintoja sellaisen hoidettavaksi.

Reaktiiviseen vastuurelaatioon toimija asettuu silloin, kun hän reagoi tehtyyn virheeseen tai väärintekoon suojelemalla tai auttamalla tekijää tai olemalla ilmoittamatta tapahtuneesta viranomaisille. Reaktiivinen vastuurelaatio syntyy myös silloin, jos toimija ei ilmaise tekijälle, että pitää tämän tekoa vääränä.⁶⁹ Voitaisiko tällä perusteella moraalisesti kyseenalaisen tai mahdollista vahinkoa aiheuttavan teknologisen systeemin tukijaa, aktiivista käyttäjää tai hiljaista hyväksyjää pitää osallisena sen aiheuttamasta vahingosta?

Melleman esittelemiä osallisuuden tapoja yhdistää toimijan intentio edistää väärintekemistä tai olla puuttumatta siihen. Toimijan osallisuus väärään tai vahingolliseen tekoon on moraalisesti moitittavaa silloinkin, kun hän ei ole kausaali-

sesti osallistunut sen aikaansaamiseen eikä itse hyödy siitä.

Edellä eriteltyt osallisuuden tavat tarjoavat välineen jakaa vastuuta sosioteknisissä systeemeissä, joissa pseudotoimijat toimivat kausaalina voimina ihmisten ohella ja joissa heidän yhteistoimintansa voi aiheuttaa epätoivottuja seurauksia, joita kukaan ei ole tahtonut saada aikaan. Tilanne haastaa arvioimaan uudelleen paitsi ihmisten mahdollisuuksia toimia tiukasti määriteltyinä moraalitoimijoina⁷⁰ myös vakiintunutta käsitystä vastuusta.

LOPUKSI

Olemme edellä tarkastelleet sosioteknisen järjestelmän epäonnistumisen aiheuttamia epäjatkuvuuskohtia ja niiden silloittamisyriytyksiä. Mitä kompleksisemmista ihmisen ja teknologian yhteistyöhön perustuvista organisaatioista on kysymys, sitä vaikeammaksi toimintaan osallistuvien erilaisten tekijöiden yhteisvaikutusten hallitseminen käy. Epäonnistumisia ja niiden aiheuttamia onnettomuustapahtumia on siksi pidetty tavanomaisena osana modernien sosioteknisiksi systeemeiksi miellettyjen organisaatioiden toimintaa.⁷¹ Onnettomuuksia voidaan estää hyvällä suunnittelulla ja luomalla kestävää toimintakulttuuria, mutta suuri merkitys on myös sillä, miten toimintaa jatketaan epäonnistumisen jälkeen.⁷²

Systemisen turvallisuuden kohentaminen ei poista tarvetta selkeyttää ja perustella vastuunkannon tai vastuun jakamisen perusteita. Syste-

67 Mellema 2016, 20.

68 Mellema 2016, 19.

69 Mellema 2016, 23–24.

70 Hallamaa & Kalliokoski 2020.

71 Ajatus perustuu Charles Perrow'n Normal Accident Theory -nimellä tunnetulle analyysille kompleksisten sosioteknisten sistemien toiminnasta, ks. Perrow 1984.

72 Huomion siirtymisestä epäonnistumisten estämisestä niiden käsittelyyn ks. Haikarainen 2022, 129–131.

miajattelu kuitenkin haastaa perinteisiä moraalikäsityksiä vastuusta ja syyllisyydestä varsinkin, jos epäonnistumisia tarkastellaan emergenttisten mallien avulla. Tässä artikkelissa esitettyjen jaetun vastuun ja osallisuuden käsitteiden avulla voidaan kenties muuttaa asennoitumista siihen, miten menneeseen ja tulevaan suuntautuvan moraalisen vastuun virheiden aiheuttamista epäonnistumisista ajatellaan jakautuvan.

Jotta ihmistoimijat kykenevät jakamaan ja kantamaan vastuuta myös epäonnistumisen aiheuttamassa kriisissä, heillä on ensinnäkin oltava selvä käsitys itsestään vastuusuhteen osapuolina. Pseudotoimijoihin voidaan ohjelmoida toimintoja, jotka vastaavat tietoa vastuunalaisuudesta tai jonkin toiminnan moitittavuudesta tai velvollisuudesta hoitaa se, mutta niillä ei ole kykyä tunnistaa moraalisen vastuun taakkaa ja seurauksia. Käsitykseen itsestä vastuusuhteen osapuolena kuuluu oman moraalitoimijuuden tunnustaminen. Toiseksi moraalitoimijan on tunnistettava niin menneeseen kuin tulevaankin suuntautuva vastuunalansa. Kolmanneksi moraalitoimijalla on oltava ymmärrys siitä, että osallistumalla omalla toiminnallaan organisaation perustehtävän toteuttamiseen hän on vastuusuhteessa paitsi itseensä ja omaan organisaatioonsa myös laajempaan yhteisöön, jota varten organisaatio on olemassa.

Yritykset silloittaa epäonnistumisen aiheuttamia toiminnan epäjatkuvuuskohtia ovat osa oikeudenmukaisuuden toteuttamista. Rikosoikeudelliset vastuukysymykset ja niiden ratkaiseminen rikosprosessissa toteuttavat retributiivista oikeudenmukaisuutta ja niiden ratkaisu sovittelumenettelyllä restoratiivista oikeudenmukaisuutta. On selvittävää, onko vahinkoon johtanut toiminta ollut normien mukaista ja minkä tasoisia normeja on mahdollisesti rikottu. Vaikka lakia ei olisi rikottukaan, epäonnistumisen aiheuttamia vahinkoja voidaan korvata esimerkiksi liikenne- tai potilasvahinkovakuutusjärjestelmän kautta, jotta päästäisiin lähemmäksi ennen epäonnistumista vallinnutta tilannetta.

On tärkeää tunnustaa myös, että kaikkia sosioteknisen yhteistoiminnan epäjatkuvuus-

kohtia ei voida ylittää: vahinkoja ei voida aina korvata. Etenkään kärsimykselle ei voida määrittellä hintaa, vaan on löydettävä muita keinoja sopeutua menetyksiin ja jatkaa elämää. Epäonnistumiset ja niiden selvittämisessä esille tulevat vastuunalaisuuden määrittämisen ongelmat ilmentävät elämään kuuluvaa traagista sattumanvaraisuutta, jota parhainkaan varautuminen ja vastuunkantamismekanismit eivät kykene täysin poistamaan.

TT Jaana Hallamaa on sosiaalietiikan professori Helsingin yliopistossa. Hän on Strategisen tutkimuksen neuvoston rahoittaman tekoälyn eettistä ohjaamista yhteiskunnan palveluksessa tutkivan ETAIROS-hankkeen (327352) varajohtaja.

- jaana.hallamaa@helsinki.fi

TT Taina Kalliokoski on sosiaalietiikan tutkijatohtori Helsingin yliopistossa. Hän työskentelee Strategisen tutkimuksen neuvoston rahoittamassa tekoälyn eettistä ohjaamista yhteiskunnan palveluksessa tutkivassa ETAIROS-hankkeessa (327352).

- taina.kalliokoski@helsinki.fi

KIRJALLISUUS

- Ames, Jonathan (2021). Sarah Everard Murder Leaves Women's Trust in the Police Almost Halved. *The Times*. <https://www.thetimes.co.uk/article/sarah-everard-murder-leaves-womens-trust-in-the-police-almost-halved-kh6pwr5gc> (luettu 23.9.2022).
- Bosse, T., Z. A. Memon & J. Treur, (2011). A Recursive BDI Agent Model for Theory of Mind and its Applications. *Applied Artificial Intelligence* 25:1, 1–44.
- Bovens, Mark (1998). *The Quest for Responsibility: Accountability and Citizenship in Complex Organizations*. Cambridge: Cambridge University Press.

- Braham, Matthew & Martin van Hees (2018). Voids or Fragmentation: Moral Responsibility for Collective Outcomes. *The Economic Journal* 128:612, F95–F113. DOI:10.1111/eoj.12507.
- Burggraave, Roger (1999). Violence and the Vulnerable Face of the Other: The Vision of Emmanuel Levinas on Moral Evil and Our Responsibility. *Journal of Social Philosophy* 30:1, 29–45.
- Cañas, José J. (2022). AI and Ethics When Human Beings Collaborate With AI Agents. *Frontiers in Psychology, Human-Media Interaction*. 04 March 2022. DOI: 10.3389/fpsyg.2022.836650.
- Dastani, Mehdi & Vahid Yazdanpanah (2022). Responsibility of AI Systems. *AI & Society*. DOI:10.1007/s00146-022-01481-4.
- Davis, Matthew C, Rose Challenger, Dharshana N.W. Jayewardene & Chirs W. Clegg (2014). Advancing Socio-technical Systems Thinking: A Call for Bravery. *Applied Ergonomics* 45:2, 171–180. DOI: 10.1016/j.apergo.2013.02.009.
- Dekker, Sidney & James Nyce (2012). Cognitive Engineering and the Moral Theology and Witchcraft of Cause. *Computational Science and Engineering* 14:3, 207–212. DOI: 10.1007/s10111-011-0203-6.
- Directive 2006/42/EC - New Machinery Directive (2006). <https://osha.europa.eu/en/legislation/directives/directive-2006-42-ec-of-the-european-parliament-and-of-the-council>.
- Elements of AI (s.a.). <https://course.elementsofai.com/ai/1>. (luettu 30.3.2023).
- Elish, Madelene Clare (2019). Moral Crumple Zones: Cautionary Tales in Human-robot Interaction. *Engaging Science, Technology, and Society* 5, 40–60. DOI:10.17351/ests2019.260.
- Elster, Jon (2013). *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. New York: Cambridge University Press.
- Guarnieri, Michael (1992), Landmarks in the History of Safety. *Journal of Safety Research* 23, 151–158. DOI: 10.1016/0022-4375(92)90018-5.
- Haikarainen, Raine (2023). Ilmailun turvallisuustutkimta ja kriminalisointipaineet. *Teologinen Aikakauskirja* tässä numerossa.
- Hallamaa, Jaana (2017). *Yhdessä toimimisen etiikka*. Helsinki: Gaudeamus.
- Hallamaa, Jaana & Taina Kalliokoski (2020). How AI Systems Challenge the Conditions of Moral Agency? *Culture and Computing, HCII 2020, Proceedings*. Toim. Matthias Rauterberg. Cham: Springer Nature Publishing AG, 54–64. DOI: 10.1007/978-3-030-50267-6_5.
- Hallamaa, Jaana & Taina Kalliokoski (2022). Placing Blame in Multi-Agent Systems. *Culture and Computing, HCII 2022, Proceedings*. Toim. Matthias Rauterberg. Cham: Springer Nature Publishing AG, 413–425. DOI: 10.1007/978-3-031-05434-1_28.
- Hallamaa, Jaana, Tomi Janhunen, Jyrki Nummenmaa, Timo Nummenmaa, Pertti Saariluoma & Elizaveta Zimina (2023 tulossa). Towards Designing Ethically Acceptable AI Security Systems Through Agent Modeling.
- Hakli, Raul & Pekka Mäkelä (2019). Moral Responsibility of Robots and Hybrid Agents. *The Monist* 102:2, 259–275.
- Himmelreich, Johannes (2019). Responsibility for Killer Robots. *Ethical Theory and Moral Practice* 22:3, 731–747.
- Kalliokoski, Taina (2020). *Yhteisöllisyyden rajat yhteistoiminnan ja ihmisen perushyviiden näkökulmasta*. Väitöskirja. Helsinki: Helsingin yliopisto.
- Kauppinen, Antti (2021). Osaammeko rakentaa moraalisia toimijoita? *Tekoäly, ihminen ja yhteiskunta: filosofisia näkökulmia*. Toim. Panu Raatikainen. Helsinki: Gaudeamus, 131–156.
- Kielitoimiston sanakirja (s.a.). <https://www.sanakirja-fi.libproxy.helsinki.fi/kotus/finnish-finnish/vahinko> (luettu 18.10.2022).
- Kölbl-Ebert, Martina (2005). Lissabon 1755: Anatomie einer Erderschütterung. *Archeopteryx* 23:83–98.
- Laitinen, Arto (2021). Mustan laatikon ongelma ja oikeus selityksen saamiseen. *Tekoäly, ihminen ja yhteiskunta: filosofisia näkökulmia*. Toim. Panu Raatikainen. Helsinki: Gaudeamus, 181–196.
- Latour, Bruno (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.
- Lintern, Gavan & Peter N. Kugler (2017). Sociotechnical System Safety: Hierarchical Control versus Mindfulness. *System Engineering* 20:4, 307–317. DOI: 10.1002/sys.21396.

- List, Christian (2021). Group Agency and Artificial Intelligence. *Philosophy and Technology* 34:4, 1213–1242. DOI: 10.1007/s13347-021-00454-7.
- Mellema, Gregory (2016). *Complicity and Moral Accountability*. Notre Dame, IN: University of Notre Dame Press.
- Neiman, Susan (2002). *Evil in Modern Thought: An Alternative History of Philosophy*. Princeton, NJ: Princeton University Press.
- Norling, E. (2004). Folk Psychology for Human Modelling: Extending the BDI paradigm. *Proceedings of AAMAS 2004*. New York: IEEE Computer Society, 202–209.
- Noorman, Merel (2017). Computing and Moral Responsibility. *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition). Toim. Edward N. Zalta <https://plato.stanford.edu/archives/sum2017/entries/collective-responsibility>.
- Pallardy, Richard (2022). Deepwater Horizon Oil Spill: Environmental Disaster, Gulf of Mexico [2010]. *Encyclopedia Britannica*. <https://www.britannica.com/event/Deepwater-Horizon-oil-spill> (luettu 13.10.2022).
- Pelkonen, Taina (2012). *Kouluumpumiset moraalisena toimintana Helsingin Sanomien kantaa ottavissa kirjoituksissa*. Pro gradu -tutkielma. Helsingin yliopisto. <http://urn.fi/URN:NBN:fi-fe201204103209>.
- Perrow, Charles (1984). *Normal Accidents: Living with High Risk Technologies*. Princeton, NJ: Princeton University Press.
- Reason, James (1990). *Human Error*. Cambridge: Cambridge University Press.
- Reason, James (2013). *Life in Error: From Little Slips to Big Disasters*. Farnham: Ashgate.
- Reiman, Teemu & Pia Oedewald (2008). *Turvallisuuskriittiset organisaatiot: Onnettomuudet, kulttuuri ja johtaminen*. Helsinki: Edita.
- Reiman, Teemu, Carl Rollenhagen, Elina Pietikäinen, & Jouko Heikkilä (2015). Principles of Adaptive Management in Complex Safety-critical Organizations. *Safety Science* 71 (part B), 80–92. DOI: 10.1016/j.ssci.2014.07.021.
- Talbert, Matthew (2019). Moral Responsibility. *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition). Toim. Edward N. Zalta <https://plato.stanford.edu/archives/win2019/entries/moral-responsibility/>.
- Taylor, Richard, Lorenzo van Wijk, John May & Neil Carhart (2016). A Study of the Precursors Leading to "Organisational" Accidents in Complex Industrial Settings. *Process Safety and Environmental Protection* (PSEP) 93, 50–67. DOI: 10.1016/j.psep.2014.06.010.
- Taylor, Richard, John May, Andrew Weyman & Neil Carhart (2017). Understanding Organisational and Cultural Precursors to Events. *Forensic Engineering (ICE)* 170:3, 1–10. DOI: 10.1680/jfoen.17.00006.
- Teperi, Anna-Maria, Teemu Paaajanen, Ilkka Asikainen & Eero Lantto (2023). From Must to Mindset: Outcomes of Human Factor Practices in Aviation and Railway Companies. *Safety Science* 158. Online 1.11.2022. DOI: 10.1016/j.ssci.2022.105968.
- Tuomela, Raimo (1977). *Human Action and its Explanation: A Study on the Philosophical Foundations of Psychology*. Synthese Library 166. Dordrecht: Reidel.
- Valviran tiedote 19.5.2021 (2021). Psykoterapiakeskus Vastaamo Oy laiminlöi useita velvollisuuksiaan. *Valvira*. <https://www.valvira.fi/-/psykoterapiakeskus-vastaamo-oy-laiminloi-useita-velvollisuuksiaan> (luettu 19.4.2023).
- Yazdanpanah, Vahid, Enrico H. Gerding, Sebastian Stein, Corina Cirstea, M. C. Schraefel & Timothy J. Norman (2021). Different Forms of Responsibility in Multiagent Systems: Sociotechnical Characteristics and Requirements. Theme Article: Sociotechnical Perspectives. *IEEE Computer Society* 25: 6, 15–22. Online December 2021. DOI: 10.1109/MIC.2021.3107334.