

## ***Vaikuttavuuskertomus 2:***

### ***Yhteiskunnallinen turvallisuus ja vakaus - Tekoälyn käyttö ja rooli eri toimijoiden yhteistyössä, ennalta ehkäisevänä toimijana, sekä informaatio-ohjauksessa ja teknologian hyödyntämisessä***

Tekoäly muuttaa yhteiskunnan toimintatapoja pitkällä ja lyhyellä aikavälillä. On olennaista varautua teknologian murroksen mukanaan tuomiin muutoksiin edeltä käsin. Tekoäly on yhä keskeisempi tekijä myös kansallista ja yhteiskunnallista turvallisuutta ja vakautta tukevien järjestelmien kehittämisessä. Tekoälyn etiikka on siksi myös teollisuusmaiden hallitusten kiinnostusten kohteena. Kulttuurin ja lainsäädännön pohjana olevaa etiikkaa pidetään tärkeänä välineenä hallita muutosta ja lisätä yhteiskunnallista resilienssiä. Etiikalla on tärkeä asema sekä suomalaisen yhteiskunnan että kansainvälisen yhteisön toiminnan sääntelyssä, kun yritetään hallita tekoälyn laajamittaisen käyttöönoton aiheuttamia sosiaalisia muutoksia ja vahvistaa yhteiskunnallista resilienssiä. Eriarvoistuminen, huono-osaisuus, koulutuksen riittämättömyys ja epätarkoituksenmukainen suuntautuminen ovat tyypillisiä sosiaalisia ongelmia, joihin hallinnossa on syytä kiinnittää huomiota. Lisäksi on hyvä muistaa, että tekoäly teknologiana tarjoaa hyviä mahdollisuuksia hallinnolle reagoida erilaisiin resilienssihaasteisiin.

ETAIROS-hankkeessa tavoitellaan pysyvää vaikuttavuutta kehittämällä yleisiä turvallisuutta ja yhteiskunnallista vakautta tukevien tekoälysovellusten käyttöönottoa ja käyttämistapoja koskevia eettisiä ja sosiaalisia toimintamalleja, sekä pyritään tukemaan tekoälyn soveltamisen sääntöpohjaisuutta. Se, millaista tekoälyteknologiaa tarvitaan eri kohteissa ja miten sitä voidaan parhaalla mahdollisella tavalla käyttää, on kansallisen turvallisuuden kannalta ydinasia. Jokainen teknologinen kehitysaskel muuttaa työtapoja ja yhteiskunnan rakenteita. Syntyy uusia tarpeita muuttaa työtapoja, koulustusta, innovaatiojärjestelmiä ja sosiaalisia toimintamalleja.

Hankeessamme luodaan menetelmiä ja käytäntöjä, joilla tekoälyn käyttöä suunnitellaan ja analysoidaan eri toimijoiden yhteistyönä. Erityisesti huomioidaan eettiset näkökulmat ja vaikutukset, ja työhön integroidaan informaatio-ohjaus ja erilaisten teknologioiden ja datalähteiden hyödyntäminen. Kyse on konsortion kokonaisvaikuttavuudesta.

Tutkimuksemme kohdistuu yhteistyön erilaisiin muotoihin, ennalta ehkäisevän toiminnan merkitykseen sekä erilaisen informaatio-ohjauksen ja teknologian hyödyntämisen mahdollisuuksiin.

***Kansallisen turvallisuuden kysymykset***, joita pohditaan yhdessä puolustus-, ja ulkoministeriön sekä Valtioneuvoston kanslian kanssa, tarjoavat hyvän pohja-aineiston yleiselle tekoälyn etiikan suunnittelulle. Olemme aloittaneet työn kartoittamalla keskeisten avainpartnereiden (Puolustusministeriön, Ulkoministeriön ja Valtioneuvoston kanslian) kanssa tiedon tarpeita. Kansallisen turvallisuuden näkökulmasta keskeisintä on toistaiseksi ollut yhteistyö Puolustusministeriön kanssa, sillä sen vastuulla on suurin osa keskeisistä kansallisen turvallisuuden kysymyksistä.

PLM:n kanssa tehtävän yhteistyön oleellinen toimintafoorumi on maanpuolustuksen tieteellinen neuvottelukunta (MATINE), jonka kanssa olemme selvittäneet puolustusministeriön kansainvälistä tekoälyn etiikkaan ja kansainvälisiin sopimuksiin liittyvää toimintaa. Työ on olennaista erityisesti Geneven sopimusten uudistamista ja kehittämistä koskevien neuvottelujen kannalta. Aseteknologioilla on tärkeä asema tekoälyn soveltamisessa, ja niiden sopimuksellinen sääntely on olennaista kansainvälisen turvallisuuden kannalta. Eettisellä ajattelulla on olennainen rooli soveltamiskäytänteitä luotaessa. Kansainväliseen yhteistyöhön olemme osallistuneet yhdessä PLM:n kanssa myös MCDC (Multinational, concept development conference) kapasiteettien kehitys forumissa (Madrid 28.10-1.11)

Ulkoministeriön kiinnostuksen kohteena ovat niin ikään kansainvälisten sopimusten tekemiseen liittyvät kysymykset. Etiikan keskeinen asema kulttuurin, tekniikkaa ja sen käyttöä koskevien sopimusten ja lakien laatimisessa sekä sosiaalisessa sääntelyssä ylipäättään selittää kansainvälisen kiinnostuksen tekoälyn etiikkaa kohtaan. Aktiivisesti toimien suomalaiset voivat saada maamme kokoa vaikuttavamman aseman säädösten ja sopimusmallien kehittämisessä.

Toinen keskeinen painopistealue on tekstiaineiston käsittely. Sekä Valtioneuvoston kanslia että ulkoministeriö pitävät erilaisten tekstin analysointiin perustuvien tekoälysovellusten kehittämistä tärkeänä. Tätä problematiikkaa tukevia ratkaisuja tarkastellaan ETAIROS:issa jatkossa. VMK:n kanssa on pohdittu myös emotionaalisen ilmapiirin mittaamisen ja seuraamisen tärkeyttä.

### **Tavoiteltu vaikuttavuus ja konsortion saavutukset tällä hetkellä**

Vaikuttavuustavoitetta arvioidaan tapaustutkimuksen kautta. Tässä vaiheessa merkittävimmissä asemassa on tekoälyjärjestelmän suunnittelu ostoskeskusturvallisuuden parantamiseksi.

Turvallisuuden hallinta on aiempaa vaativampaa ja siinä käytetään entistä enemmän erilaisia datalähteitä, jotka välittävät informaatiota muun muassa rakennusten tekniikasta ja sen toiminnasta mutta myös ihmisten ja ihmismassojen seuraamisesta. Kokonaisuuden reaaliaikainen hallinta on erittäin vaativa tehtävä etenkin, kun turvallisuuteen vaikuttavia päätöksiä joudutaan tekemään yleensä tarjolla olevan, parhaan mahdollisen tilannearvion perusteella.

Suunnittelutyössä ovat tässä vaiheessa mukana SPEK, ostoskeskus Sello, ja kiinteistöalan kaupalliset toimijat Newsec ja Lassila-Tikanoja. Näiden edustajia on tavattu.

Konsortio on pyrkinyt tässä vaiheessa keräämään partnerien tiedontarpeita. Kahdenkeskisiä palavereja on ollut Keski-Suomen keskussairaalan, Rambollin, PLM:n, UM:n, VNK:n ja Spekin kanssa.

## Havainnot konkreettisista vaikutuksista

Hanke on vasta alkanut, ja nyt voi vain sanoa, että aihe herättää suurta kiinnostusta osapuolissa ja se on jo saanut osapuolia miettimään asioita uusista näkökulmista.

Puolustuskapasiteettien parantaminen on olennainen osa kansallisen turvallisuuden kehittämistä. Tässä yhteydessä olemme osallistuneet monikansalliseen kapasiteettien kehittämisseminaariin (MCDC) Madridissa. Eettisten käytänteiden suunnitteluun on osallistuttu myös tieteellisissä kokouksissa, joissa on ollut mahdollista tarkastella periaatteellisia tekoälyn etiikan ongelmia, kuten Humen giljotiinin problematiikkaa.

Ihmisen ja älykkään teknologian vuorovaikutuskysymykset ovat nousseet nopeasti kansainvälisen tiedeyhteisön agendalle. ETAIROSissa tehtyä tutkimustyötä on kuvattu neljässä tieteellisessä julkaisuissa, jotka on esitelty IHJET – Human Interaction and Emerging Technologies -konferenssissa Nizzassa 22-24.8.2019, IHSED – Human Systems Engineering and Design -konferenssissa Münchenissä 16-18.9.2019, ja tullaan esittämään IHSI III – Human Systems Interaction -konferenssissa Modenassa 18-20.2.2020 sekä IHJET II – Human Interaction and Emerging Technologies -konferenssissa Lausannessa 23-25.4.2020.

Lisäksi olemme toimineet tekoälysession puheenjohtajana MATINEn tutkimusseminaarissa 21.11.2019. Tällaisten seminaarien kautta on mahdollista saada elävä kuva siitä, miten tekoälykysymysten käsittely etenee suomalaisessa yhteiskunnassa.

Konsortio on vasta käynnistänyt työnsä, joten vielä ei voida esitellä varsinaisia konsortion luomia tuloksia. Tutkimustyö on kuitenkin jo käynnistynyt konkreettisesti, erilaisia skenaarioita on analysoitu ja sopivia menetelmiä on alettu suunnitella. Eettistä arviointikehikkoa rakennetaan.

Teemme tutkimusta suunnittelumenetelmistä, joilla sekä kuvataan alueen toimijoiden rooleja että mallinnetaan tekoälyjärjestelmää ja siihen liittyviä järjestelmiä ja datalähteitä siten, että myös eettiset näkökulmat kuvataan ja mallinnetaan, jotta niitä voidaan analysoida edelleen. Tutkimukseen kytketään myöhemmin viranomaistoiminnan, lainsäädännön ja julkisen hallinnon näkökulmat.