

VALTIONEUVOSTON  
SELVITYS- JA TUTKIMUSTOIMINTA

Raija Koivisto, Jaana Leikas, Heidi Auvinen, Ville Vakuri, Pertti Saariluoma, Jenni Hakkarainen ja Riikka Koutu

## **Tekoäly viranomaistoiminnassa - eettiset kysymykset ja yhteiskunnallinen hyväksyttävyyys**

**Helmikuu 2019**

Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 14/2019

# KUVAILULEHTI

<b>Julkaisija ja julkaisuaika</b>	Valtioneuvoston kanslia, 1.2.2019
<b>Tekijät</b>	Raija Koivisto, Jaana Leikas, Heidi Auvinen, Ville Vakkuri, Pertti Saari- luoma, Jenni Hakkarainen ja Riikka Koulu
<b>Julkaisun nimi</b>	Tekoäly viranomaistoiminnassa - eettiset kysymykset ja yhteiskunnalli- nen hyväksyttävyyys
<b>Julkaisusarjan nimi ja numero</b>	Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 14/2019
<b>Asiasanat</b>	Etiikka, tekoäly, ohjelmistorobotiikka, hyväksyttävyyys, viranomaistoi- minta

## Julkaisun osat/ muut tuotetut versiot

<b>Julkaisuaika</b>	Helmikuu, 2019	<b>Sivuja</b> 62	<b>Kieli</b> Suomi
---------------------	----------------	------------------	--------------------

## Tiivistelmä

Tekoälyn ja ohjelmistorobottien käyttö lisääntyy - myös viranomaistoiminnassa - muuttaen prosesseja ja toimintatapoja. Uusien teknologioiden ja järjestelmien käyttöönotolla on moninaisia vaikutuksia niin kansalaisten, virkamiesten kuin koko yhteiskunnankin näkökulmasta. Yhteiskunnan arvojen mukainen eettinen toiminta ja siitä kumpuava luottamus tulisi säilyttää myös tekoälyn aikakaudella.

Tässä selvityksessä on aluksi tarkasteltu käsityksiä etiikasta ja hyväksyttävyydestä yleisesti ja erityisesti teknologian kehittämisen ja käyttöönoton yhteydessä. Tuloksena on esitetty selvitystyön aikana kehitetty eettinen toimintamalli. Koska eettisyys on vahvasti kontekstisidonnaista, on tekoälyn eettisyyttä viranomaistoiminnassa tarkasteltu käyttötapausten avulla, joita selvitettiin haastatteleamalla 13 valtionhallinnon toimijaa.

Käyttötapausten eettisiä ja hyväksyttävyyssymptomiksi tarkasteltiin työpajoissa. Hyväksyttävyyttä selvitettiin lisäksi tekemällä kansalaiskysely ja analysoimalla internetissä käytyä keskustelua tekoälystä ja sen käytöstä viranomaistoiminnassa. Raportissa on esitetty vahvimmin esille tulleet eettisyyteen ja hyväksyttävyyteen liittyvät kysymykset sekä kysymyksiin liittyviä ratkaisuehdotuksia.

Tekoäly viranomaistoiminnassa saatetaan nähdä eettisten ongelmien aiheuttajana, mutta se avaa paljon myös positiivisia mahdollisuuksia, kuten monenlaisia laadun parantamisen keinoja. Siten tekoäly viisaasti käytettynä on eettisyyden ja yhteiskunnallisen hyväksyttävyyden kannalta positiivisen muutoksen mahdollistaja ja voi parhaimmillaan jopa ratkaista eettisyyteen ja yhteiskunnalliseen hyväksyttävyyteen liittyviä ongelmia.

Tämä julkaisu on toteutettu osana valtioneuvoston vuoden 2018 selvitys- ja tutkimussuunnitelman toimeenpanoa (tietokayttoon.fi).

Julkaisun sisällöstä vastaavat tiedon tuottajat, eikä tekstisisältö välttämättä edusta valtioneuvoston näkemystä.

# PRESENTATIONSBLAD

<b>Utgivare &amp; utgivningsdatum</b>	Statsrådets kansli, 1.2.2019
<b>Författare</b>	Raija Koivisto, Jaana Leikas, Heidi Auvinen, Ville Vakkuri, Pertti Saari- luoma, Jenni Hakkarainen och Riikka Koulu
<b>Publikationens namn</b>	Artificiell intelligens inom myndighetsutövande - etiska och acceptans aspekter
<b>Publikationsseriens namn och nummer</b>	Publikationsserie för statsrådets utrednings- och forskningsverksamhet 14/2019
<b>Nyckelord</b>	Etik, artificiell intelligens, AI, mjukvarurobotik, acceptans, myndighets- utövande

## Publikationens delar /andra producerade versioner

<b>Utgivningsdatum</b>	Februar, 2019	<b>Sidantal</b> 62	<b>Språk</b> Finska
------------------------	---------------	--------------------	---------------------

## Sammandrag

Artificiell intelligens (AI) och mjukvarurobotar används allt mera – också av offentliga myndigheter – och förändrar därmed processer och verksamhetsformer. Ibrukttagandet av nya teknologier och system har en mångskiftande inverkan på medborgarna, tjänstemännen och också på hela samhället. En etisk verksamhet i enlighet med samhällets normer, samt det förtroende som följer av detta, bör bibehållas också under AI eran.

Den här rapporten granskar först begrepp inom etik och acceptans på ett generellt plan, och därefter mera specifikt i samband med utvecklingen och ibrukttagandet av teknologi. Som resultat presenteras en verksamhetsmodell, som utvecklats inom utredningen. Eftersom etik är starkt beroende av kontexten, granskar rapporten med hjälp av användningsfall etiken rörande AI inom myndighetsutövande. Användningsfallen identifierades genom att intervjua 13 aktörer inom statsförvaltningen.

Användningsfallens etiska och acceptans aspekter granskades i workshoppar. Acceptans utreddes dessutom via en medborgarenkät samt genom att analysera Internet diskussionerna om AI och om hur myndigheterna använder AI. Rapporten tar fram de mest uppenbara frågorna rörande etik och acceptans, samt förslag till lösningar av frågorna.

AI inom myndighetsutövande kan anses förorsaka etiska problem. Å andra sidan öppnar AI många positiva möjligheter; den erbjuder bl.a. många medel att förbättra kvaliteten. Därför möjliggör en klokt tillämpad AI etiska och acceptabla samhälleliga förändringar och kan i bästa fall t.o.m. lösa problem rörande etik och samhällelig acceptans.

Den här publikationen är en del i genomförandet av statsrådets utrednings- och forskningsplan för 2018 ([tietokayttoon.fi/sv](http://tietokayttoon.fi/sv)).

De som producerar informationen ansvarar för innehållet i publikationen. Textinnehållet återspeglar inte nödvändigtvis statsrådets ståndpunkt

## DESCRIPTION

<b>Publisher and release date</b>	Prime Minister's Office, 1.2.2019
<b>Authors</b>	Raija Koivisto, Jaana Leikas, Heidi Auvinen, Ville Vakkuri, Pertti Saariluoma, Jenni Hakkarainen and Riikka Koulu
<b>Title of publication</b>	Artificial intelligence in authority use - ethical and societal acceptance issues
<b>Name of series and number of publication</b>	Publications of the Government's analysis, assessment and research activities 14/2019
<b>Keywords</b>	Ethics, artificial intelligence, software robotics, acceptability, authority use

### Other parts of publication/ other produced versions

<b>Release date</b>	February, 2019	<b>Pages</b> 62	<b>Language</b> Finnish
---------------------	----------------	-----------------	-------------------------

### Abstract

Artificial intelligence and software robots are used more often also by authorities thus changing the operations and procedures. The use of new technologies and systems has manifold impacts from the point of view of citizens, civil servants and the whole society. The ethical functioning reflecting the prevailing values and the consequent trust in the society should remain also in the era of AI.

This report first studies the concepts of ethics and acceptability in general but especially in the context of technology development and application. As a result, an ethical model for technology development generated during the study is presented. Because ethics is strongly context dependent, the ethics of AI in the authority use is studied by using use cases. These were identified by interviewing 13 civil service organizations.

Ethical and acceptability issues in use cases were discussed in workshops. In addition, acceptability was studied also with the help of an interview among citizens and by analyzing the internet discussions about AI and its use by authorities. The report introduces the most obvious issues concerning ethics and acceptability as well as ideas on how to solve these issues.

AI in the authority use may be seen as the origin for ethical issues. However, AI may open many opportunities as well such as offering different solutions for quality improvement. Therefore, AI, if applied in a wise way, can be an enabler for an ethical and acceptable societal change and at its best even solve the ethical and acceptability problems in the society.

This publication is part of the implementation of the Government Plan for Analysis, Assessment and Research for 2018 ([tietokayttoon.fi/en](http://tietokayttoon.fi/en)).

The content is the responsibility of the producers of the information and does not necessarily represent the view of the Government.



# SISÄLLYS

<b>Tiivistelmä</b> .....	<b>6</b>
<b>1. Johdanto</b> .....	<b>10</b>
1.1. Tausta .....	10
1.2. Tutkimuksen tavoitteet.....	11
<b>2. Tekoälyn etiikka ja yhteiskunnallinen hyväksyttävyys</b> .....	<b>13</b>
2.1. Eettiset periaatteet .....	13
2.2. Tekoälyn etiikan erityiskysymyksiä .....	14
2.3. Eettisen suunnittelun lähestymistapoja.....	16
2.4. Yhteiskunnallisen hyväksyttävyyden käsite .....	17
2.5. Viranomaistoiminta, oikeudelliset reunaehdot ja etiikka .....	18
<b>3. Tekoälyn käyttötapaukset viranomaistoiminnassa</b> .....	<b>21</b>
3.1. Muutos perinteisestä automaatiosta tekoälyyn.....	21
3.2. Tekoälyn käyttöönoton nykytila.....	22
3.3. Eettisyyden ja hyväksyttävyyden tarkastelu valituissa käyttötapauksissa.....	27
<b>4. Kansalaisnäkökulma tekoälyyn viranomaiskäytössä</b> .....	<b>40</b>
4.1. Kansalaiskysely yhteiskunnallisesta hyväksyttävyydestä.....	40
4.2. Kansalaiskeskustelu internetissä .....	45
<b>5. Yhteenveto ja pohdinta</b> .....	<b>48</b>
5.1. Johtopäätöksiä käyttötapauksista .....	48
5.2. Johtopäätöksiä kansalaisnäkökulmasta .....	50
5.3. Eettisyyden ja hyväksyttävyyden kysymyksiä .....	52
5.4. Eettinen toimintamalli.....	54
<b>Lähteet</b> .....	<b>58</b>

# TIIVISTELMÄ

Tietojärjestelmien ja tekoälyn kehitys on näihin päiviin saakka ollut teknologiavetoista, ja kiinnostavaa on ollut, mihin kaikkeen teknologia pystyy. Nyt tässä kehityksessä ollaan murroskohdassa, jossa tietotekniikka ja -talous muuttavat kaikkia yhteiskunnan rakenteita. Tekninen kehitys on otettava haltuun myös muista näkökulmista niin, että se palvelee yhteiskuntaa eikä päinvastoin. Kansalaisten perusoikeuksien tulee toteutua myös tietotalouden aikana, jotta luottamusyhteiskuntamme hyvät peruspilarit eivät sorru.

**Eettisyys ja hyväksyttävyy**s ovat luottamuksen syntymisen edellytyksiä. Etiikka on filosofinen oppiala, joka tutkii ja argumentoi moraalia. Moraali taas tarkoittaa ihmisten käsityksiä hyvästä ja pahasta, oikeasta ja väärästä, sekä siitä, miten näitä tulisi käytännön toimissa ilmentää. Teknologian sosiaalisella ja yhteiskunnallisella hyväksyttävyydellä tarkoitetaan sitä, miten ihmiset käytännössä hyväksyvät teknologian. Tämä viittaa ensinnäkin hyvään käytettävyyteen, käyttäjäkokemukseen ja käyttökelpoisuuteen ihmisten kokemana asiana. Toiseksi se viittaa käyttötarkoituksen objektiiviseen oikeuttamiseen eettisessä, juridisessä ja käytännöllisessä mielessä, toisin sanoen siihen, että toiminnat objektiivisesti ottaen täyttävät yhteiskunnan toiminnan arvopäämäärät. Näin eettisyys ja yhteiskunnallinen hyväksyttävyy tyypillisesti kulkevat käsi kädessä.

Tässä selvityksessä tarkasteltiin tekoälyn ja ohjelmistorobotiikan (myöhemmin tekoäly) eettisyyttä ja yhteiskunnallista hyväksyttävyyttä viranomaiskäytössä. Koska eettisyyttä ja hyväksyttävyyttä on aina tarkasteltava kontekstissaan, kuvasimme viranomaiskäyttöä tekoälyn ja ohjelmistorobotiikan käyttötapausten muodossa. Haastatelluissa 13 eri ministeriöiden hallinnonalan organisaatioissa tekoälyn käyttöönotto oli eri vaiheissa; kokeiluja, suunnitelmia ja ideoita tekoälyn käytöstä oli kaikilla useita.

Perinteistä automaatiota käytetään jo laajasti viranomaistoiminnassa. Automaattinen käsittely perustuu selkeisiin lainsäädännöstä lähteviin laskentamalleihin, joihin ei liity harkintaa. **Tekoäly voi suorittaa myös harkintaa sisältäviä toimintoja ja sen käyttö voi muuttaa koko viranomaisen palveluprosessia.** Suurimmiksi muutoksiksi tunnistettiin asiakkaan kontaktointitapojen monipuolistuminen, aika- ja paikkariippumattomuus, palveluprosessin nopeutuminen, työn ja eri virastojen välisen yhteistyön muuttuminen, asiakkaan parempi huomiointi päätöksenteossa, sekä tarkistusprosessin ja perustelujen esittämisen ja viranomaisten koulutuksen muuttuminen. Näiden lisäksi sekä asiakkaan että viranomaisen kyky kyseenalaistaa asioita voi heikentyä ja asiakkaan suhtautuminen viranomaisiin voi muuttua.

Eettisyyttä ja hyväksyttävyyttä tarkasteltiin sekä käyttötapausten avulla että yleisestä kansalaisnäkökulmasta. Käyttötapaustarkastelut toteutettiin haastattelujen ja kolmen työpajan avulla.

## **Esimerkkejä toistuvista, useille käyttötapauksille yhteisistä eettisyyden ja hyväksyttävyyden haasteista**

Seuraavassa esitetään kootusti selvityksen perusteella tunnistettuja yksittäisiä eettisyyden ja hyväksyttävyyden kysymyksiä ja pohditaan niihin ratkaisujen ensiaskeleita. Tässä esitetyt kysymykset ovat esimerkkejä toistuvista, useille käyttötapauksille yhteisistä eettisyyden ja hyväksyttävyyden haasteista, mutta ne eivät korvaa tapauskohtaista tarkastelua, johon ehdotamme luvussa 5.4 esitettyä toimintamallia.

Miten huomioidaan algoritmien katvealueet? Huolena esitetään se, miten inhimillinen näkökulma voidaan ottaa päätöksenteossa huomioon, kun käsittelijänä on ihmisen sijaan kone. Henkilökohtaiset tilanteet, harkinnanvaraisuus, asiaan vaikuttavat sivuseikat (näiden lukeminen ns. rivien välistä), yms. jäävät tekoälyjärjestelmään perustuvan päätöksenteon ulkopuolelle.

- Kansalaisella olisi oltava mahdollisuus päästä asioimaan myös ihmisvirkailijan kanssa.
- Kansalaisen oikeudet tekoälypalvelujen käytössä olisi kommunikoitava selvästi.

Rapautuuko luottamus virkamiesosaamiseen? Sekä valtionhallinnon sisällä että yksilön tai yhteiskunnan kannalta on merkittävää säilyttää luottamus virkamiesten osaamiseen siitakin huolimatta, että prosesseja pyritetään älykkäiden järjestelmien avulla. Vikatilanteiden ratkaiseminen, erikoistapausten käsittely, valitukset, yms. vaativat myös tulevaisuudessa ihmisosaamista.

- Toiminnan tulee perustua selkeään lainsäädäntöön.
- Järjestelmien valvonta ja esimerkiksi virheiden tunnistaminen tulee varmistaa virkamiesten osaamisen kautta.
- Tekoälypalvelun käytön perustelut erityisesti harkinnanvaraisissa tilanteissa on kommunikoitava: läpinäkyvyys ja asian ymmärtäminen lisäävät luottamusta.
- Hyvien kokemusten - ns. voitonnarratiivien - kerryttäminen käynnistyy aloittamalla yksinkertaisemmista sovelluksista, joissa tekoäly helpoimmin tuottaa positiivisia tuloksia. Selkeä kommunikointi ja vastuunkanto epäonnistuttaessa auttaa säilyttämään luottamuksen.
- Kansalaisilla tulisi olla saatavilla tieto siitä, missä omat tiedot ovat ja miten ja mihin niitä hyödynnetään.
- Järjestelmien ja sovellusten helppokäyttöisyydellä ja positiivisten käyttäjäkokemusten avulla lisätään luottamusta käytännön tasolla.

Onko julkishallinnossa riittävää hankinta-, ylläpito- ja kehittämisosaamista? Tekoälyjärjestelmien hankintaan liittyvät resurssi- ja laatuksymykset ovat keskeisiä ratkaisujen toteuttamisen, käyttöönoton ja ylläpidon onnistumisessa. Puutteellisen osaamisen riskejä ovat etiikan ja hyväksyttävyyden näkökulmista vajavaiset järjestelmät, yksityisyyden vaarantuminen, toimittajaloukku ja järjestelmän ylläpidon ja uudistamisen organisoinnin haasteet.

- Roolijaot tulee määritellä selkeästi (ml. yhteistyö ja vuorovaikutus) ja niihin liittyvät tekniset ja muut osaamistarpeet: tilaaja, omistaja, käyttäjä, työkalut, vastuut, jne.
- Hankinnat tulee avata (tavoitteet, vaiheet, toimijat, muutokset, hyödyt, kustannukset) arkikielelle ymmärryksen lisäämiseksi ja keskustelun mahdollistamiseksi; myös kansalaiset käyttäjinä voisivat osallistua tarvemäärittelyyn.
- Resurssien yhdistäminen esimerkiksi hallinnonalojen tai tieteenalojen kesken on tärkeää osaajien ja osaamisen hankkimiseksi.
- Valtionhallinto voisi täsmentää yleistä ohjeistusta tekoälyhankintoja varten.
- Viestintään on kiinnitettävä erityistä huomiota.

Kuka vaatii tekoälyn toimintaa? Virheiden tunnistaminen tekoälyn toiminnassa vaatii jatkuvaa seuranta ja mahdollisesti sekä ihmis- että konetyötä. Järjestelmien opetusaineistot ja algoritmit ovat monimutkaisia ja niissä esiintyvät virheet voivat jäädä pitkäksi aikaa piiloon. Toisaalta myös selvien virhetoimintojen korjaaminen ja väärin päätösten oikaisu voi olla haasteellista ja kallista. Myös esimerkiksi vastuun määrittämät viranomais toiminnassa ovat nykyisellään epäselviä tai toimimattomia tekoälyjärjestelmiä sovellettaessa.

- Algoritmien ja datan avaaminen kansalaisille voisi olla yksi ratkaisu, mutta toisaalta se voisi aiheuttaa väärinkäytön riskin. Tasapainottavia ratkaisuperiaatteita tarvitaan.
- Tekoälyjärjestelmän oppimiseen ja kehittämiseen on luotava selvät, avoimesti kommunikoidut periaatteet ja prosessit.
- Vastuukysymysten määrittelyt ihmisen ja järjestelmien suhteen tulee laatia tekoälykontekstiin sopiviksi.

Miten käydään rajanvetoa tekoälyn käyttämisestä ennakoivassa viranomaistyössä? Älykkäät järjestelmät mahdollistavat tiedon analyysien avulla tapahtuvan ennakoinnin, jota voidaan hyödyntää viranomaistyössä monella tapaa, esimerkkeinä varhainen puuttuminen riskikäyttäytymiseen tai tukien myöntäminen ilman hakemusta. Avoin kysymys on, missä määrin tällaisia toimintoja voidaan ylipäänsä sisällyttää tulevaisuuden viranomaistyöhön.

- Ennakoivan toiminnan periaatteena tulisi olla aina kansalaisen ja yhteiskunnan etu. Tulisi keskustella, määritellä ja kommunikoida missä tilanteissa, sovelluskohdissa ja järjestelmissä ennakoiva työ on hyväksyttävää.

Hämärtääkö tekoäly viranomaisprosessien ja -päätösten ymmärrettävyyttä? Riskinä pidetään sitä, että tekoälyjärjestelmien käyttöönoton myötä viranomaistyön perusteet ja prosessit jäävät pimentoon. Poikkihallinnolliset järjestelmät ja datamassojen ristiin käyttö haastavat tilanteita edelleen. Kansalaisen mahdollisuus ja kyky nähdä, mihin omia tietoja käytetään, miten niitä voi hallita ja voiko niitä poistaa, hämärtyy. Läpinäkymättömyys ja ymmärrettävyyden puutteet haittaavat sekä kansalaista että virkamiestä.

- Jotta viranomaistyö säilyy ymmärrettävänä, tekoälyratkaisun tulee olla hyvin toteutettu, avoimesti periaatteiltaan dokumentoitu ja ymmärrettävästi viestitty.
- Osoptimoinnin sijaan viranomaispalvelun kokonaisprosessi on arvioitava uudelleen, kun tekoäly tai ohjelmistorobotti on osa sitä.
- Tekoälyjärjestelmän ja siihen liittyvän viranomaisprosessin kannalta kaikilla osapuolilla tulee olla selvyyttä siitä, kuka päättää ja miten, mikä on hyväksyttävää. Mitä enemmän harkinnanvaraisuutta koneen tehtäväksi annetaan, sitä paremmin sen käyttö on perusteltava ja viestittävä.

Ovatko tekoälyyn liittyvät väärinkäytös- ja turvallisuusriskit hallittavissa? Tekoälyjärjestelmiin liitetään uhkakuvia tietomurroista, ja toisaalta pelätään myös käyttäjien voivan manipuloida järjestelmän toimintaa omaksi edukseen.

- Laajamittaiset riskit vaativat varautumista, mutta myös varautumisesta tiedottamista.
- Toteutuneisiin uhkisiin reagoimisen tulee olla luottamuksen säilyttämisen vuoksi mahdollisimman ajantasaista ja avointa.

Mitä kansalaisen ylipäänsä tulee tietää tekoälystä? Yleinen tietoisuus tekoälyn mahdollisuuksista ja uhkista on vielä puutteellista eikä välttämättä pohjautu faktatietoon. Kaikkien käyttäjä- ja erityisryhmien pääsy palveluun on taattava. Riittävän tietotason määrittely, ja tätä tukeva informaatio-ohjaus ovat siten tärkeitä. Keskeinen tekijä on myös kansalaisen käsitys omista digitaidoistaan. Ylipäänsä rajanvetoa tarvitaan myös siinä, tuleeko kansalaisen tietää, liittyykö viranomaispalvelun toteuttamiseen mahdollisesti tekoälyn hyödyntämistä.

- Kaikilla tulisi olla yhtäläinen pääsy ja mahdollisuus käyttää sovellusta taustasta, toimintakyvystä, digitaidoista tai kielitaidosta riippumatta.
- Käytön ja käyttöliittymän tarkoituksenmukaisuuteen, helppouteen ja selkeyteen tulisi kiinnittää huomiota järjestelmää suunniteltaessa niin, että tekoäly ei monimutkaista



käyttäjäkokemusta tai muodostu sen esteeksi. Päinvastoin, tekoälyjärjestelmän avulla tulisi pyrkiä parantamaan käyttäjäkokemusta.

### **Eettinen toimintamalli**

Seuraavassa on kuvattu toimintamalli eettisten kysymysten pohdintaan viranomaistoiminnassa, jota esitämme sovellettavaksi tekoälyn käyttöönoton suunnittelussa ja toteutuksessa. Toimintamalli on kuvattu kokonaisuudessaan luvussa 5.4. Toimintamallin keskeiset vaiheet ovat:

1. *Ottakaa eettinen näkökulma.*
2. *Hankkikaa tarkka käsitys arvioitavasta tapauksesta.*
3. *Yrittäkää nähdä sovellukseen liittyvät eettiset kysymykset ja mahdolliset perinteiset ratkaisut, jotka tapaukseen liittyvät.*
4. *Verratkaa tapausta muihin vastaavanlaisiin tapauksiin.*
5. *Hyödyntäkää erilaisia systemaattisia analyysitekniikoita*
  - Eri ammattiryhmien eettiset koodistot
  - Roolien ja vastuiden arviointi
  - Eri sidosryhmien arviointi
  - Järjestelmällinen toimintaohjeiden arviointi
  - Eettisten teorioiden arviointi
  - Tekoälyyn liittyvien eettisten periaatteiden huomiointi.
6. *Tehkää asiaankuuluvat eettiset johtopäätökset tapauksesta.*
7. *Miettikää tapausta tulevaisuuden kannalta.*

Tekoäly viranomaistoiminnassa saatetaan nähdä eettisten ongelmien aiheuttajana, mutta se avaa paljon myös positiivisia mahdollisuuksia, kuten monenlaisia laadun parantamisen keinoja. Näitä ovat esimerkiksi nopeus, aikariippumattomuus, yhdenmukaisuus, reilu kohtelu, säännönmukaisuus, isojen volyymien käsittelykyky, resurssitehokkuus, uusien palvelujen mahdollistuminen ja datan maksimaalinen hyödyntäminen. Siten tekoäly viisaasti käytettynä on eettisyyden ja yhteiskunnallisen hyväksyttävyyden kannalta positiivisen muutoksen mahdollistaja ja voi parhaimmillaan jopa ratkaista eettisyyteen ja yhteiskunnalliseen hyväksyttävyyteen liittyviä ongelmia.

# 1. JOHDANTO

## 1.1. Tausta

Eri ministeriöissä on meneillään tai suunnitteilla useita tekoälyyn liittyviä kehityshankkeita omien prosessiensa tueksi (Klemettinen ja Peltola 2018). Yleisesti ottaen tietojärjestelmien ja tekoälyn kehitys on näihin päiviin saakka ollut teknologiavetoista, ja kiinnostavaa on ollut, mihin kaikkeen teknologia pystyy. Nyt tässä kehityksessä ollaan murroskohdassa, jossa tietotekniikka ja -talous muuttavat kaikkia yhteiskunnan rakenteita. Tekninen kehitys on otettava haltuun myös muista näkökulmista niin, että se palvelee yhteiskuntaa eikä päinvastoin. Kansalaisten perusoikeuksien tulee toteutua myös tietotalouden aikana.

Yksi tietotalouteen liittyvistä keskustelluimmista aiheista onkin tekoälyn eettisyys ja ennen kaikkea tiedon omistajuus. USA:ssa tiedon omistavat yritykset, kun taas Kiinassa sitä hallinnoi valtio. Eurooppa pyrkii globaaliin kilpailuun eettisyyden saralla (European Commission 2018a; European Commission 2018b; Rabesandratana 2018; Goodman 2018) ja merkittävien kansainvälien tekniikan alan järjestö IEEE - Institute of Electrical and Electronics Engineers (Bryson ja Winfield 2018) on ehdottanut jopa tekoälyn ja autonomisten järjestelmien eettisen suunnittelun standardointia. Useat Euroopan unionin jäsenmaat ovat julkaisseet tekoälyyn liittyviä strategiapapereita, joissa eettisyys on vahvasti esillä (House of Lords 2018; Villani 2018; Die Bundesregierung 2018; The Expert Group on data ethics 2018).

Suomessa tekoälyn käyttöönottoa valtionhallinnossa tukee mm. valtiovarainministeriön käynnistämä kansallisen tekoälyohjelma Auroran esiselvityshanke toimikaudelle 15.9.2018-28.2.2019 (Valtiovarainministeriö 2018). Auroran tavoitteena on vauhdittaa julkisen hallinnon siirtymistä tekoälyaikaan turvallisesti ja eettisesti, ja siten mahdollistaa kansalaisille monien palveluntuottajien palveluista koostuvat saumattomat ja sujuvat kokonaisuudet erilaisiin elämäntapahtumiin. Käytännön ohjeet siitä, mitä eettinen tekoäly on tai miten sitä tuotetaan ovat kuitenkin vielä puutteellisia. Tutkimusta ja keskustelua asiasta tarvitaan. Mistä sitten puhutaan, kun puhutaan tekoälyn etiikasta? Etiikan rooli on kuvata sitä, miten ihmiset kohtelevat toisiaan, ja miten heidän tulisi elää. Se tarkastelee mikä on oikein ja mikä väärin, ja miten hyvä elämä voisi toteutua. Tällaiset kysymykset ovat luonteeltaan käytännöllisiä, ja siksi etiikka nähdään yleensä käytännöllisen filosofian osana. Etiikka jakautuu moniin koulukuntiin, kysymyksenasetteluihin ja ajattelutapoihin. Olennaista etiikan tutkimuksessa on määritellä ihmisten toiminnan pohjana olevat normit ja arvot, tehdä ne tietoisiksi kaikille ja alistaa ne näin sosiaalisen tarkastelun, keskustelun ja analyysin kohteeksi.

Koska teknologia on yksi nykyisen ihmisen tapa kohdata monia elämään ja toimeentuloon liittyviä ongelmia, on luonnollista ajatella, että teknologian käyttö ja suunnittelu ovat monin tavoin kytköksissä eettisiin periaatteisiin. Koska teknologia vaikuttaa ihmisten elämään ja arkeen, myös sen käyttöä ja suunnittelua säätelevät relevantit eettiset kysymykset on syytä ottaa tarkastelun kohteeksi.

Eettinen keskustelu toimii teknologiakehityksen yhteydessä usealla eri tasolla. Korkeimman tason muodostavat yleiset ihmisten toimintaa säätelevät periaatteet. Esimerkkinä voidaan mainita ”Älä tee vahinkoa toiselle” -periaate. Tällaisen yleisen eettisen periaatteen tulisi säädellä kaikkea ihmisten yhteistoimintaa. Tältä tasolta on kuitenkin syytä laskeutua alemmas tarkastelemaan teknologioiden hyödyntämiseen liittyviä konkreettisempia kysymyksiä. Tällöin yleiset periaatteet saavat uusia muotoja.

”Älä tee vahinkoa toiselle” -periaate muuttuu konkreettisella tasolla moniksi uusiksi kysymyksiksi. Tällaisia ovat esimerkiksi oikeus yksityisyyteen, fyysiseen koskemattomuuteen tai viranomaistoimintojen läpinäkyvyyteen. Kaikkiin näihin teknologian käyttökysymyksiin liittyy uhkia, jotka voivat toteutuessaan johtaa ihmisen henkiseen tai fyysiseen vahingoittamiseen. Suurella osalla eettisiä peruskysymyksiä on vastaavia erityismuotoja uusien teknologioiden sosiaalisen hyväksymisen ja käyttöönoton osana. Käytännössä käytötapaukset liittyvät toiminnan prosesseihin.

## 1.2. Tutkimuksen tavoitteet

Tämä raportti on loppuraportti tutkimuksesta, jonka VNK on tilannut aiheesta Robotiikan ja tekoälyn sääntelyn ja hyväksikäytön etiikka sekä yhteiskunnallinen hyväksyttävyyden (5.2 D). Hankkeen tavoitteena oli selvittää tekoälyn ja ohjelmistorobotiikan käyttöön viranomaistoinnassa liittyviä eettisiä ja yhteiskunnallisen hyväksyttävyyden kysymyksiä. Osatavoitteet olivat:

1. robotiikan ja tekoälyn viranomaiskäyttöä kuvaavien edustavien käytötapauksien selvittäminen
2. eettisten kysymysten ja juridisten reunaehtojen kuvaaminen yleensä
3. tunnistettujen käytötapauksien eettisyyskysymysten selvittäminen
4. tunnistettujen käytötapauksien yhteiskunnallisen hyväksyttävyyden selvittäminen
5. kansalaisten ja viranomaisten toimintamallien uudelleen organisointi uuden teknologian vaatimuksia vastaaviksi
6. ratkaisuvaihtoehtojen kehittäminen ongelmallisiin tilanteisiin.

Tutkimuksen lähtöoletuksena oli, että tekoälyteknologioiden, ja laajemmin digitalisaation järjestelmällinen hyödyntäminen sekä julkisella sektorilla että yrityksissä on avaintekijä tuottavuuden kasvulle, uusille liiketoiminnoille ja sitä kautta kansalliselle menestykselle. Sen tarkoituksena oli siten rakentaa ymmärrystä ja tietopohjaa eri hallinnonalojen strategiselle suunnittelulle ja sovelluskohtaisille hankkeille.

Hankkeen tulosten toivotaan auttavan päättäjiä ymmärtämään tekoälyn mahdollisuudet ja lähtemään hyödyntämään niitä omalla (hallinnon) alallaan. Vastaavasti ymmärrys auttaa päättäjiä havaitsemaan tekoälyteknologioiden tuomat haasteet ja riskit sekä löytämään niihin uusia ratkaisuja. Selvityksen tuloksista hyötyvät myös tekoälyyn pohjaavia ratkaisuja kehittävät yritykset. Kolmas tuloksista hyötyvä taho julkisen hallinnon ja yritysten ohella ovat kansalaiset ja media.

Tässä selvityksessä tarkennamme Ailiston ym. (2018) määritelmää tekoälystä Pöntisen (2018) esittämällä tavalla tarkoittamaan ”algoritmin ja sen hyödyntämän datan yhteisvaikutusta, joka mahdollistaa laitteen tai muun vastaavan ei-elollisen toimivan tehtävän ja tilanteen mukaisesti järkevällä tavalla”. Vastaavasti ohjelmistorobotiikalla tarkoitamme tässä ”ohjelmia, joita voidaan konfiguroida käyttämään organisaation tietojärjestelmiä niin kuin ihmisenkin niitä käyttäisi” (Kääriäinen et al. 2018). Tässä raportissa puhuttaessa tekoälystä viranomaistoinnassa viittaamme sekä tekoälyn että ohjelmistorobotiikan soveltamiseen viranomaistyössä.

Raportin luku 2 esittelee etiikan ja yhteiskunnallisen hyväksyttävyyden käsitteitä ja periaatteita. Luvussa edetään yleisestä tasosta tekoälyn erityiskysymyksiin sekä sen hyödyntämiseen viranomaistoinnassa. Luku 3 esittelee hankkeen empiiristä tutkimusta, jossa haas-

tattelujen ja työpajojen avulla kartoitettiin ja analysoitiin tekoälyn nykyisiä ja tulevia käyttötapauksia viranomaistoiminnassa. Empirian toinen osa kuvataan luvussa 4, jossa kyselyaineistoon ja internetin kansalaiskeskusteluun perustuen pureudutaan tekoälyn hyväksyttävyyteen viranomaispalveluja käyttävän kansalaisen näkökulmasta. Luku 5 vetää yhteen hankkeen johtopäätökset, koosteen eettisyyden ja hyväksyttävyyden keskeisimmistä kysymyksistä ratkaisusuositukseen sekä tulosten perusteella esitettävän eettisen toimintamallin.

## 2. TEKOÄLYN ETIIKKA JA YHTEISKUNNALLINEN HYVÄKSYTTÄVYYS

### 2.1. Eettiset periaatteet

Etiikalla (kreik. ethos = vakiintuneet tavat) tarkoitetaan moraaliin liittyvien kysymysten, kuten oikean ja väärän, arvojen ja hyvän elämän pohdintaa. Eettiset periaatteet ovat sääntöjä, normeja, oikeuksia ja velvoitteita, joiden tulisi ohjata meitä päätöksenteossamme. Ne sisältävät ihanteita yksilön arvostamisesta ja kunnioittamisesta sekä kieltoja (esimerkiksi ”älä vahingoita”). Eettisesti ongelmallisissa tilanteissa pyrimme etsimään ratkaisua näistä yleisesti hyväksytyistä periaatteista. Eettinen dilemma syntyy silloin, kun käsissämme on kaksi erilaista, mutta mahdollista ongelman ratkaisuvaihtoehtoa. Eettiset periaatteet eivät kuitenkaan ole absoluuttisia tai riippumattomia. Samoin eettiset kysymykset ovat luonteeltaan monimerkityksisiä, eikä niihin voi soveltaa standardiratkaisuja. Etiikka ei ole eksakti tiede, eikä ongelmiin useimmiten voida tarjota yhtä ainoaa oikeaa ratkaisua. Silti selkeiden, tarkkojen tai yksioikoisten vastausten puuttuminen ei oikeuta suhtautumaan eettisiin kysymyksiin leväperäisesti. (Leikas 2008)

*Itsemääräämisoikeuden kunnioittaminen (autonomy)* on keskeisin yksilön ihmisarvoa ja vapautta ilmaiseva periaate. On tärkeää, että ihminen voi ilmaista aidosti omia tunteitaan, arvojaan ja pyrkimyksiään sekä päättää itse tärkeistä valinnoistaan. Itsemääräämisoikeuden mukaan emme saa kumota yksilön tahtoa itsemääräämisoikeuteen pakottamalla, uhkaamalla tai millään muulla vapautta rajoittavalla tavalla, vaikka hänen päätöksensä tuntuisikin meistä miten huonolta tahansa eikä edistäisi mielestämme ihmisen hyvää. Tähän sopii myös älä vahingoita toisia (do no harm) -periaate sekä toiminnan vapauden ja valinnanvapauden kunnioittaminen.

Itsemääräämisoikeuden rinnalla *yhteisen hyvän, yhteisöllisyyden ja oikeudenmukaisuuden (justice)* periaatteet vaativat osakseen huomiota. Oikeudenmukaisuuden periaate pyrkii tarkastelemaan sitä, miten tietyssä yhteisössä yhteisön jäsenten tulisi toimia ja miten haittojen ja etujen tulisi yhteisössä jakaantua. Oikeudenmukaisuudessa on siten kyse myös siitä, mitä oikeuksia ja velvollisuuksia yhteisön jäsenillä on. Oikeudenmukaisuus voidaan ymmärtää hyvin eri tavoilla sen mukaan, millaisen näkökulman yhteisö on valinnut.

*Tasa-arvoisuusperiaatteen* mukaisesti etiikka ei salli kenenkään suosivan rikkaita köyhien kustannuksella, mahtavia heikkojen kustannuksella, miehiä naisten kustannuksella, valkoihoisia värillisten kustannuksella, nuoria vanhojen sijaan, terveitä vammaisten sijaan jne. Huolimatta ihmisen asemasta tai taustasta jokaista henkilöä tulee kohdella samalla tavalla, kun etuja ja haittoja arvioidaan.

Arvioitaessa teknologian suhdetta yksilöön ja yhteisöön sekä näiden suhdetta toisiinsa esiin nousevat esimerkiksi seuraavat, keskenään limittyvät periaatteet (Leikas 2008):

*Teknologian käyttäjän näkökulmasta:* ihmisarvo (dignity), loukkaamattomuus (integrity), oikeuksien kunnioittaminen (respect for rights), itsemääräämisoikeus (autonomy), tietoinen suostumus (informed consent), oikeus kieltäytyä (right to decline), luottamus (trust), pätevyys (competence), yhdenvertaiset mahdollisuudet kaikille, tasa-arvoisuus (democracy), identiteetti (identity), käyttäjän osallistuminen (participation), käyttäjän suojaaminen (protec-

tion of user), valvonta (surveillance), turvallisuus (safety), saavutettavuus (access), vahingoittaminen (do no harm), valinnanvapaus (choice), vapaaehtoisuus (voluntariness) ja yksityisyyden suoja (privacy).

*Yhteisön näkökulmasta:* tasapuolinen hyödyn jakautuminen (equal benefit), kulttuurinen moninaisuus, yhteistyö (cooperation), yhteiset sopimukset (conventions), syrjintä (freedom from bias) sekä teknologian sosiaalinen ja yhteiskunnallinen vaikutus (social impact of technology and role in the society).

*Teknologian kehittäjän näkökulmasta:* luotettavuus (reliability), valvonta (surveillance), turvallisuus (security), yksimielisyys (agreement), pätevyys (competence), vastuullisuus (accountability), tekijänoikeuksien kunnioittaminen (respect for intellectual property rights) ja ymmärrys (comprehension). Tekoäly teknologiana nostaa esiin uusia, tekoälyn toimintaan ja hyödynnettävyyteen liittyviä eettisiä kysymyksiä. Näitä tarkastellaan seuraavassa luvussa.

## 2.2. Tekoälyn etiikan erityiskysymyksiä

Viime vuosien teknologiset edistysaskeleet sekä uudet sovelluskohteet ovat tuoneet tekoälyn ihmisten pariin. Enää sovellukset eivät toimi vain suljetussa ympäristössä, vaan ovat verkon ja autonomisten systeemien, kuten autojen tekoälysovellusten kautta, osa ihmisen arkea. Tekoälyn hyödyntäminen ja sen tulevaisuuden käyttöpotentiaali eri sovelluskohteissa ovat tehneet tekoälystä uudella tavalla ihmisten elämään vaikuttavaa teknologiaa. Tekoälyn hyödyntäminen osana ihmisten arkielämää on tehnyt myös tekoälyyn liittyvät eettiset kysymykset ajankohtaisiksi ja relevanteiksi. (Charisi 2017)

Tekoälyn etiikan määrittely itsessään on haaste tutkimukselle: Teknisen tekoälyn ja filosofisen etiikan yhdistäminen muodostaa luonteeltaan monitieteisen ja metodologisesti rikkaan tutkimuskentän. Monitieteisen taustan vuoksi tekoäly ja etiikka liittyvät toisiinsa usealla eri tasolla. Tekoälyn etiikan aihealueen määrittelyyn on tarjottu useampia malleja – tai pikemminkin aihealueiden kategorisointeja – mutta alalle ei ole vielä vakiintunut yksittäistä käsitystä siitä, mitä tekoälyn etiikka pitää sisällään tai mitä sen tulisi pitää sisällään. (Leikas et al. 2018; Vakuri & Abrahamsson 2018).

Tekoälyn etiikka voidaan käsitteellisesti hahmottaa koostuvan ainakin kolmesta kentästä (Dignum 2018). Yksi kenttä on etiikka, jossa on kyse luonteeltaan eettisen päättelyn tai päätöksenteon integroimisesta osaksi autonomisia systeemejä (engl. ethics for design). Toinen on etiikka, joka koskee erilaisten autonomisten ja älykkäiden järjestelmien suunnittelu (engl. ethics in design). Etiikka liittyy tällöin tekoälyyn sellaisten sääntelyn ja tekniikan sekä suunnittelun metodien muodossa, jotka tukevat tekoälyjärjestelmien eettisten kysymysten analysointia ja arviointia näiden järjestelmien korvataessa ja muovatessa perinteisiä sosiaalisia rakenteita. Kolmas kenttä on etiikka, joka ymmärretään sellaisina menettelytapoina, standardeina ja sertifiointeina, joiden tarkoituksena on turvata käyttäjien ja kehittäjien toiminta hyväksytyjen moraaliperiaatteiden pohjalta (engl. ethics for design(ers)). Tämä koskee niin käyttöä, suunnittelua, toteutusta ja hallintaa kuin tutkimustakin.

Dignummin (2018) kolmijakoa mukaillen tekoälyn etiikka voidaan pelkistellysti jakaa kolmelle eri tasolle:

- Eettisten tekijöiden suunnittelu (Ethics by design):

- Luonteeltaan eettisen päättelyn tai päätöksenteon kyvykkyyden toteuttaminen osaksi tekoälyjärjestelmää hyödyntämällä teknisiä ja algoritmisia ratkaisuja.
- Suunnitteluetiikka (Ethics in design):
  - Autonomisten ja älykkäiden järjestelmien suunnitteluajattelu
  - Tekoälyjärjestelmien eettisten vaikutusten suunnittelu ja arviointi, sekä tähän tarvittavien taitojen ja menetelmien kehittäminen järjestelmien sulautuessa osaksi yhteiskunnan rakenteita
  - Esimerkiksi sääntely ja metodien kehittäminen.
- Ammattietiikka (Ethics for design(ers)):
  - Tekoälyn kehittäjien ja käyttäjien eettisen ja moraalisen toiminnan tukeminen järjestelmien tutkimuksen, suunnittelun, toteutuksen, käyttöönoton ja hallinnan näkökulmista.
  - Esimerkiksi toimintamallien kehittäminen, codes of conduct, standardit ja sertifiointi.

Etiikan huomioimiseksi osana tekoälyn ja tekoälysovellusten kehittämistä on olemassa kaksi pääasiallista lähestymistapaa: periaatteellinen ja metodinen. *Periaatteellisissa lähestymistavoissa* etiikkaa lähestytään ohjaavien periaatteiden ja periaatekokoelmien kautta, kuten esimerkiksi AI at Google: our principles, IEEE Ethically aligned design, ART-Design principles (Dignum 2017) ja Euroopan komission Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems. *Metodisissa lähestymistavoissa* taas etiikkaa lähestytään suunnitteluajattelun ja vaiheittain etenevien menetelmien kautta. Näitä ovat esimerkiksi Value-Sensitive Design (VSD) (Friedman et al. 2006), Care-Centered Value Sensitive Design (CCVSD) (Wynsberghe 2013) ja Life-Based Design (LBD) (Saariluoma et al. 2016). Metodiset lähestymistavat ponnistavat suunnitteluajattelusta ja suunnittelua ohjaavista lähestymistavoista, prosesseista ja menetelmistä, kun taas ohjaavat periaatteet pyrkivät tarkastelemaan tekoälyn eettisiä periaatteita.

Edellisessä luvussa esitettyjen periaatteiden lisäksi tekoälykeskustelussa on esitetty muun muassa seuraavia tekoälyn tai tekoälyjärjestelmien kehitystä ohjaavia periaatteita:

<i>Läpinäkyvyys (transparency)</i>	Tekoälyjärjestelmien toiminnan sekä kehittämisen läpinäkyvyys, algoritmien läpinäkyvyys
<i>Vastuullisuus (accountability/responsibility)</i>	Ihmisten ja tekoälyjärjestelmien kyky vastata päätöksistä sekä tunnistaa virheitä ja odottamattomia seurauksia.
<i>Reiluus (fairness/justice/equality)</i>	Järjestelmien kyky toimia tasapuolisesti
<i>Selitettävyyys (explainability/accountability)</i>	Kyky selittää ja oikeuttaa päätökset järjestelmää käyttävälle
<i>Jäljitettävyyys (traceability/explainability)</i>	Kyky jäljittää ja ymmärtää järjestelmän päätöksiä sekä niiden lähtökohtia.

## 2.3. Eettisen suunnittelun lähestymistapoja

### Arvosensitiivinen suunnittelu

Yhtenä lähestymistapana tekoälyn etiikan kysymyksiin on esitetty arvosensitiivisen suunnittelun (Value-Sensitive Design) -lähestymistapaa. Friedman, Kahn ja Borning (2003; 2006) ovat käsitelleet moraalisten arvojen merkitystä teknologian suunnittelussa. Heidän esittämänsä lähestymistapa, Value-Sensitive Design (VSD), nostaa esille moraaliset kysymykset tuotesuunnittelussa. VSD:n mukaan suunnittelija liittää tuotteeseen inhimillisiä arvoja, joko tietoisesti tai tiedostamattaan, niin käyttämiensä suunnittelumenetelmien kuin toteuttamiensa suunnitteluratkaisujen kautta. VSD painottaa eettisyyttä tuotesuunnittelun lähtökohdaksi ja korostaa sellaisia moraalisia arvoja kuin yksityisyyden suojaa (privacy), oikeutta tulla kohdelluksi tasa-arvoisesti (non-discriminatory) sekä yksilön tarvetta riippumattomuuteen (autonomy) ja oikeuteen päättää omaa toimintaansa koskevista asioista (control).

VSD:n mukainen suunnitteluprosessi voidaan aloittaa käsitteellistämisvaiheesta, jossa tunnistetaan teknologian tai palvelun käytön eri sidosryhmät, ts. sekä suoraan että epäsuorasti suunniteltavan teknologian kanssa tekemisissä olevat tahot sekä arvot, joita nämä tahot noudattavat. Keskeiset arvot määritellään käsitteellisesti: mitä tarkoitetaan esimerkiksi luottamuksella, mistä luottamus syntyy ja miten se määritellään mahdollisimman konkreettisesti. Empiirisessä vaiheessa selvitetään, miten nämä keskeiset arvot toteutuvat käytännössä ihmisten ja sosiaalisten järjestelmien toiminnassa ja vuorovaikutuksessa. Empiiristä tutkimusta tarvitaan myös suunnittelun tulosten menestyksellisyyden arvioimiseen. Teknologian kehittämisen vaiheessa selvitetään, miten teknologian ominaisuudet ja sen taustamekanismit tukevat tai ehkäisevät tiettyjen arvojen toteutumista. Tähän vaiheeseen kuuluu myös tekninen ennakoiva suunnittelu siten, että suunnittelun lopputulos tukee käsitteellisessä vaiheessa tunnistettuja ja määriteltyjä arvoja.

Arvosensitiivisen suunnittelun tukena voidaan käyttää arvolistoja heuristiseen tapaan suunnitteluohjeina (Friedman ym. 2006). Tyypillisiä arvoja ovat esimerkiksi ihmisen hyvinvointi, vapaus, vastuu omista teoista sekä huomaavaisuus. Useat Friedmanin ja kollegoiden niimeämistä arvoista liittyvät yksityisyyteen ja tiedon käsittelyyn. VSD:tä on kritisoitu siitä, ettei se huomioi useiden vaihtoehtoisten arvojen yhtäaikaista olemassaoloa

### Teknologiakeskeisestä suunnittelusta elämäkeskeiseen suunnitteluun

Teknologioiden nopea kehitys on johtanut uusien suunnitteluajattelun mallien kehittymiseen. Perinteisesti teknologiasuunnittelu on ollut tekniikan suunnittelua. On pyritty tuottamaan koneita, laitteita ja ohjelmistoja, ja sen jälkeen etsimään niille sovellutuksia. Perinteiseen ajattelutapaan liittyy kuitenkin kapea-alaisuutta, mistä syystä on ryhdytty kehittämään kokonaisvaltaisempia teknologian suunnittelumalleja, joissa tekniikka pyritään suunnittelemaan osana ihmisten toiminnan ja arjen suunnittelua. Tällainen ihmisten tarpeiden ja elämäntilanteiden ymmärtämisestä ponnistava, elämäkeskeiseksi kutsuttu suunnittelutapa (Life-Based Design; Saariluoma et al. 2016) korostaa ihmisten elämänlaadun parantamista teknologian kehittämisen lähtökohdaksi. Siksi elämäkeskeisessä suunnittelussa kysytäänkin millainen rooli teknologialla voi olla elämänlaadun parantajana ja miten teknologia voi kyseisessä kontekstissa tukea kohteena olevaa toimintaa tärkeiden päämäärien saavuttamiseksi. Suunnittelijan on pohdittava missä muodossa ja millä (ja kenen) ehdoilla teknologia toivotetaan tervetulleeksi ja halutaan käyttöön.



Elämäkeskeisessä suunnittelussa tarkastellaan yhtäältä tosiasioita, kuten esimerkiksi tietyn viraston tehtäviä ja toimintatapoja, ja toisaalta arvoja ja normeja, joita yhteiskunta ja kansalaiset noudattavat. Faktujen kertoessa miten asiat ovat tällä hetkellä tai mihin suuntaan tapahtumat ovat johtamassa, arvot kertovat siitä, miten asioiden tulisi olla. Mikä on mahdollista, mikä on sallittua ja mikä on tavoiteltavaa, ovat tyypillisiä arvojen pohjalta avautuvia suunnittelunäkökuja. Elämäkeskeistä suunnitteluajattelua hyödyntäen on mahdollista määrittellä teknologian vaatimukset vastaamaan hallinnon toimintatapoja ja tehtäviä, ja vastaavasti hahmottaa uusia hallinnon toimintatapoja tukemaan mahdollisimman hyvin kansalaisten arkea ja hyvää elämää.

Elämäkeskeinen suunnittelu avaa ihmistieteiden kirjon suunnittelijoiden käyttöön tuomalla esimerkiksi psykologian ja sosiologian kysymykset osaksi teknologiaratkaisujen suunnittelua. Tältä pohjalta on mahdollista toteuttaa eettisesti kestäviä ja hallinnon käytännön toimintamalleja tukevia teknologiaratkaisuja. Ihmistieteiden kysymysten sisällyttäminen teknologia-suunnitteluun lisää suunnitteludiskurssien syvyyttä, ja auttaa hahmottamaan teknologian hyödyntämiseen, haltuunottoon ja vaikuttavuuteen liittyviä eettisiä kysymyksiä niin yksilöiden kuin yhteiskunnan näkökulmasta.

## Vastuullinen innovaatiotoiminta

Vastuullinen tutkimus- ja innovaatiotoiminta (Responsible Research and Innovation, RRI) on viime vuosina noussut esiin erityisesti Euroopan Komission toimesta. RRI-ajattelun tausta on tieteen ja teknologian tuottamien yhteiskunnallisten riskien kasvussa ja tutkimustulosten yhteiskunnallisessa hyväksyttävyydessä. RRI on määritelty läpinäkyväksi, vuorovaikutteiseksi prosessiksi, jonka kautta yhteiskunnalliset toimijat ja innovoijat tulevat avoimiksi toistensa näkemyksille teknologian (eettisestä) hyväksyttävyydestä, kestäväyydestä ja yhteiskunnallisesta toivottavuudesta (von Schomberg 2011). Keskeistä RRI-toiminnalle on siis eri toimijoiden yhteiskunnallinen vuorovaikutus sekä tutkimus- ja innovaatiotoiminnan laajasti ymmärretty yhteiskunnallinen kestävyys ja hyväksyttävyys.

Eettisyys ja perusoikeuksien kunnioittaminen on yksi RRI-toiminnan kulmakivistä. Muut ovat toimijoiden osallistaminen laajasti, sukupuolten tasa-arvo, sekä tieteellisen koulutuksen, tulosten ja prosessien avoimuus. RRI korostaa innovaatioiden yhteiskehittämistä, minkä taustalla on ajatus siitä, että osallistamalla käyttäjät ja muut yhteiskunnan toimijat voidaan kehittää yhteiskunnallisesti vaikuttavia innovaatioita. Tämä tarkoittaa esimerkiksi sitä, että teknologisia innovaatioita kehitettäessä yksi merkittävistä tekijöistä on kestävyysajattelu. Monet teknologiapohjaiset yhteiskunnalliset innovaatiot pyrkivät yhteiskunnalliseen muutokseen, tai yhteiskunnallisten haasteiden purkamiseen, jolloin on tärkeää ymmärtää, että vastuullisuudesta tulee tällöin myös jaettava: kehittämisessä mukana olleista toimijoista tulee yhdessä vastuullisia tulosten vaikutusten ja seurausten suhteen (Nieminen et al. 2018).

## 2.4. Yhteiskunnallisen hyväksyttävyyden käsite

Eettisistä periaatteista ei tule todellisia elämää ohjaavia käytäntöjä, elleivät ihmiset sisäistä niitä ja sitoudu niihin. Tekoälyn etiikkaa kehitettäessä on tärkeää kiinnittää erityistä huomiota uuden teknologian edellyttämien eettisten käytänteiden hyväksyttävyyden tarkastelemiseen. Käytännössä tämä merkitsee sitä, että etiikkaa on syytä tarkastella aluksi sekä normatiivisena että reaalisena ilmiönä.

Normatiivinen etiikka tarkastelee eettisiä periaatteita oikeutettavissa olevina ideaaleina. Normien ajatellaan olevan oikein tai väärin (oikeutettuja tai virheeksi osoitettuja) erilaisten perusteiden kautta. Mitkä ja millä perusteella vallitsevat säännöt, arvot, normit ja lait ovat oikeutettuja ja miten ne voidaan oikeuttaa? Tämä on olennaista, koska yhteiskunnallisesti hyväksyttäviä eettisiä periaatteita voivat olla vain ne, jotka ovat oikeutettua.

Oikeuttaminen on itse asiassa yksi demokratian keskeisiä toimintaperiaatteita. Yhteisen keskustelun ja säätämistoiminnan kautta syntyvät oikeutetut normit ja lait sekä niiden taustalla olevat arvo-oletukset. Arvot eivät kuitenkaan ole ikuisia, vaan ne elävät.

Demokratiassa arvojen muutos perustuu vapaaseen keskusteluun ja vallitsevien arvojen analyysiin. Tätä kutsutaan diskurssietiikaksi. Sen edellytyksenä ovat demokraattisten perusarvojen, kuten sananvapauden, yhdistymisvapauden ja vaaleilla valittujen päättäjiin perustuvien toimintamallien tunnustaminen. Ne tekevät mahdolliseksi sosiaalisen arvojen harkinnan ja mahdollisuuden keskustella arvoista ja analysoida niitä mielekkäästi yhteiskunnan kontekstissa.

Teknologian tehtävänä on ihmisen elämänlaadun parantaminen. Tämä koskee myös tekoälyä. Elämänlaatu tarkasteltuna ihmisten arjen sekä yleisen henkisen ja fyysisen hyvinvoinnin pohjalta luo hyvän mittarin sille, olemmeko me onnistuneet ottamaan uuden teknologian osaksi yhteiskunnallista toimintaamme mielekkäällä tavalla. Voidaan puhua yhteiskunnallisen arjen mielekkyydestä. Ravinto, asuminen, sivistys, vapaus, terveys, oikeudenmukaisuus, turvallisuus ja onnellisuus ovat tällöin elämänlaadun arvioinnin osatekijöitä.

Määreiksi teknologiaeettisen elämänlaadun tarkastelemista varten sopivat myös jo edellä esitetyt : ihmisarvo, yksityisyys, fyysinen koskemattomuus, teknologioiden tarkoituksenmukaisuus, ihmisten kohtelun reiluus, luottamus ja tasa-arvo, avoimuus ja vastuullisuus, läpinäkyvyys ja jäljitettävyyys. Lisäksi voisi vielä mainita: ennakoitavuus, oikeus kasvatukseen, osallistaminen ja kuulluksi tuleminen, sananvapaus ja oikeus tietoon, sekä oikeus osallistua uusien yhteiskunnallisten toimintakulttuurien luomiseen. Kaikki edellä esitetyt näkökulmat ovat käsitteellisiä, kunnes ne tulevat todelliseksi osaksi ihmisten elämää.

Tässä selvityksessä teknologian sosiaalisella ja yhteiskunnallisella hyväksyttävyydellä tarkoitetaan sitä, miten ihmiset käytännössä hyväksyvät teknologian. Tämä viittaa ensinnäkin hyvään käytettävyyteen, käyttäjäkokemukseen ja käyttökelpoisuuteen ihmisten kokemana asiana. Toiseksi se viittaa käyttötarkoituksen objektiiviseen oikeuttamiseen eettisessä, juridisessä ja käytännöllisessä mielessä, ts. siihen, että toiminnot objektiivisesti ottaen täyttävät yhteiskunnan toiminnan arvopäämäärät.

## **2.5. Viranomaistoiminta, oikeudelliset reunaehdot ja etiikka**

### **Oikeudelliset reunaehdot yhteiskunnallisen hyväksyttävyyden edellytyksenä**

Robottiikan ja tekoälyn hyödyntämisen oikeudelliset reunaehdot ovat laaja, eritasoisia normistoja leikkaava kokonaisuus. Viranomaisen toimintaa jäsentävät yhtäältä kansainvälinen ihmisoikeusnormisto, perustuslaissa asetetut perusoikeudet sekä viranomaisen toimintaa ohjaava yleinen- ja erityinen lainsäädäntö.

Perustuslain funktio yksilöiden ja valtion välistä suhdetta määrittelevänä normistona on keskeinen, sillä perustuslaissa asetetaan viranomaistoiminnalle laadullisia vaatimuksia. Perus-

tuslaissa turvataan muun muassa viranomaistoiminnan lakisidonnaisuus ja yksilöiden oikeus oikeussuojakeinoihin suhteessa julkista valtaa käyttävään viranomaiseen. Sektorikoh- taisessa lainsäädännössä säännellään yksityiskohtaisemmin esimerkiksi viranomaisen tie- donsantioikeuksista, oikeusturvakeinojen käyttämisestä sekä viranomaistoiminnan julki- suudesta. Tätä normistoa tarkastellaan laajemmin Valtioneuvoston kanslian selvityksessä “Algoritmi päätöksentekijänä? Tekoälyn hyödyntämisen mahdollisuudet ja haasteet kansalli- sessa sääntely-ympäristössä”, jossa selvitetään tekoälyn hyödyntämisen esteitä ja mahdolli- suuksia lainsäädännön näkökulmasta.

## **Viranomaistoiminnan legitimitetti**

Oikeuden ja etiikan välinen suhde ei suinkaan ole yksiselitteinen, eikä viranomaistoiminnan eettisyyttä tai hyväksyttävyyttä kansalaisten silmissä voida yksiselitteisesti palauttaa kysy- mykseen toiminnan lainmukaisuudesta. Oikeudellisessa tutkimuksessa oikeusjärjestelmää ja sitä kautta myös julkisen vallan käyttämistä kohtaan tunnettuun hyväksyttävyyteen viita- taan usein legitimitetin käsitteellä. Normatiivisessa legitimitetissä legitiimin järjestelmän toiminnan edellytys on sen lakiperusteisuus. Viranomaistoiminnan legitimitettiin voidaan tarkastella kuitenkin myös normatiivisesta järjestelmästä eli lainsäädännöstä erillisenä, yh- teiskunnan tuntemaan luottamukseen kiteytyvänä ilmiönä (Kulla 2006).

Oikeuden ja etiikan näennäisestä erillisyydestä huolimatta voidaan keskeisten lakien, kuten rikoslain ja perustuslain, katsoa heijastavan yhteiskunnallisesti hyväksyttäväksi katsottuja arvoja ja ohjaavan viranomaisen toimintaa yleisesti hyväksyttävään suuntaan. Voimassaole- vat lait asettavat julkisen hallinnon toiminnalle reunaehdoja, joita sen tulee noudattaa. Digita- lisoitaessa viranomaistoimintaa tai lisäämällä automaatiota viranomaisessa, on keskeistä huomioida etenkin hyvän hallinnon periaatteiden, oikeussuojakeinojen ja kuulluksi tulemisen kokemuksen (Lind 1988) kiinteä suhde legitimitetikkysymykseen ja kansalaisten tuntemaan luottamukseen.

Perinteisen näkökulman mukaan sellaiset hallinnon periaatteet kuten yhdenvertaisuus, tar- koituksenmukaisuus ja luottamuksen suoja ovat turvanneet viranomaistoiminnan legitimi- teettiä (Koivisto 2007, Tuori 2002). Hyvän hallinnon periaatteet ja toimiva oikeusjärjestelmä konkretisoituu kansalaisille etenkin muodollisen ja tosiasiallisen oikeusturvan saatavuus- dessa (Lind 1988). Tehokkaat oikeussuojakeinot toimivat viimesijaisena mekanismina, jolla kansalainen voi kyseenalaistaa esimerkiksi väärinä tai epäoikeudenmukaisina koetut toi- menpiteet. Kehitettäessä eettisesti kestäviä sovelluksia julkishallinnon toiminnan tueksi kai- ken toiminnan lähtökohdaksi on siis otettava tekoälysovellusten hyödyntäminen tavalla, joka on sekä hyväksyttävää että voimassa olevan lainsäädännön mukaista.

## **Keskeisestä lainsäädännöstä: perustuslaki, yhdenvertaisuus ja tiedon käsit- tely**

Keskeisiksi lainsäädännöstä nouseviksi haasteiksi tekoälykehitykselle ja etenkin tekoälyn hyödyntämiselle viranomaistyössä on tässä selvityksessä, sekä tutkimushankkeen muissa osioissa tunnistettu muun muassa tiedon käsittelyyn liittyvät kysymykset. Kun tekoälyn käyt- töä ja automaatiota lisätään yhteiskunnan eri sektoreilla, nousee esille lisäksi kysymys virka- vastuun perusteiden määräytymisestä. Ongelmaksi nähdään usein se, voidaanko virkavas- tuun määräytymistä säänteleviä vahingonkorvaus- ja rikosoikeudellisia normeja soveltaa sellaisenaan myös tekoälyjärjestelmiin, ja mikä taho näissä järjestelmissä on perustuslain, vahingonkorvauslain tai rikoslain tarkoittama virkamies. Julkisyhteisön virkavastuulla on kes-

keinen funktio viranomaistoiminnan järjestelmätason toimivuuden, legitimitietin ja luottamuksen ylläpitäjänä (Koskinen ja Kulla 2016). Lisäksi olennaisiksi oikeudellisiksi kysymyksiksi tekoälyaikana nousevat muun muassa hyvän hallinnon toteutumisen varmistaminen ja se, kuinka yksilöille turvataan riittävät oikeussuojakeinot.

Tekoälyn ja viranomaistoiminnan välisen suhteen arvioinnin lähtökohdaksi voidaan ottaa esimerkiksi perustuslain 2 § (731/1999, PL). PL 2 §:ssä säädetään oikeusvaltioperiaatteesta, jossa oikeusvaltion toiminnan edellytykseksi asetetaan vallankäytön lakisidonnaisuus. Kaiken julkisen vallankäytön on perustuttava lakiin. Lisäksi perustuslain 2 §:n mukaan yksilöllä tulee olla oikeus osallistua yhteiskunnan kehittämiseen. Yhteiskunnan kehittäminen kansanvaltaisuuden mukaisesti edellyttää avointa julkista hallintoa ja riittävää demokraattista keskustelua etenkin merkittävien uudistusten yhteydessä.

Yhdenvertaisuus turvataan perustuslain 6 §:ssä ja periaatetta täydennetään useissa sektoriaeissa tarkemmin määritellyillä syrjinnän kielloilla. Kansallisessa ja kansainvälisessä keskustelussa keskeisimmäksi haasteeksi koneoppimismalleja hyödyntävien sovellusten ja palveluiden kehittämisessä on havaittu niihin liittyvä riski syrjivään tai epäyhdenvertaiseen kohteluun. Riski syrjimisestä tai syrjäytymisestä voi realisoitua esimerkiksi tilanteissa, jossa yhteiskunnan jäsenet asettuvat eriarvoiseen asemaan teknisen kyvykkyyden tai erilaisten tietoteknisten laitteiden saatavuuden seurauksena.

Perustuslain 21 §:ssä säädetään oikeusturvatakeista. Lailla turvataan jokaiselle kuuluva oikeus saada asiansa käsiteltäväksi toimivaltaisessa tuomioistuimessa tai viranomaisessa. Keskeinen teema digitalisaatiota ja viranomaistoiminnan automaatiota tarkastelevassa tutkimuksessa on se, kuinka kansalaisille varmistetaan riittävät oikeussuojakeinot, sekä turvataan muodollisen ja tosiasiallisen oikeusturvan saatavuus. Oikeussuojakeinojen käyttäminen edellyttää muun muassa sitä, että yksilöllä on riittävä ymmärrys häntä koskevien päätösten perusteista ja vaikutuksista, sekä mahdollisuus valittaa päätöksistä.

Riittävä avoimuus, julkisuus ja tiedonsaanti tilanteissa, joissa yksilö asioi viranomaisessa tai joissa tähän kohdistetaan julkista valtaa, on siis edellytys tehokkaiden oikeussuojakeinojen käyttämiselle. Esimerkiksi yksilön mahdollisuus saada selkeä tieto häntä koskevan päätöksen perusteista on edellytys sekä virheellisten päätösten havaitsemiseksi että valitusoikeuden käyttämiseksi. Jäljempänä esiteltävissä kyselyn tuloksissa tehokkaiden oikeussuojakeinojen ja viranomaistoiminnan läpinäkyvyyden merkitys osoittautui keskeiseksi edellytykseksi kansalaisten luottamukselle tekoälyavusteiseen viranomaistoimintaan.

Henkilötietojen suoja määritellään Euroopan unionin perusoikeuskirjassa (12.12.2007, 8 art.) perusoikeudeksi. Tietosuojalainsäädännön uudistamisen myötä ja yleisessä tietosuojasetuksessa (2016/679) säädettyjen henkilötiedon käyttöä ohjaavien normien vaikutuksesta henkilötietojen käytön lainmukaisuus ja vastuullisen käytön merkitys on korostunut julkisessa keskustelussa. Suurin osa tekoälyksi luettavia teknologioita on ns. dataintensiivisiä eli mallin kehittäminen ja operatiivinen hyödyntäminen edellyttävät merkittäviä datamääriä. Koska viranomaistoimintaan liittyy usein henkilötietojen käsittelyä, ovat tietojen tietoturallinen ja lainmukainen käyttö avainasemassa kehitettäessä tekoälyjärjestelmiä viranomaistoiminnan tueksi.

## 3. TEKOÄLYN KÄYTTÖTAPAUKSET VIRANOMAIS-TOIMINNASSA

### 3.1. Muutos perinteisestä automaatiosta tekoälyyn

Perinteistä automaatiota käytetään jo laajasti viranomaistoiminnassa. Automaattinen käsittely perustuu selkeisiin lainsäädännöstä lähteviin laskentamalleihin, joihin ei liity harkintaa. Ohjelmistorobotiikka ja tekoäly voivat suorittaa älykkäämpiä, myös harkintaa sisältäviä toimintoja, jotka voivat muuttaa monia asioita myös viranomaistoiminnassa. Asiaa pohdittiin selvityshankkeen järjestämässä työpajassa 21.8.2018 eri toimijoiden näkökulmasta. Työpajassa tunnistettiin seuraavat muutokset tai haasteet:

- *Asiakkaan kontaktointitavat monipuolistuvat.* Perinteisen virastossa käynnin ja kaavakkeen täytön tilalle tulee - ja on jo tullut - sähköinen asiointi: kaavakkeen voi täyttää internetissä vaikkapa mobiililaitteella. Asiointi voi myös tapahtua suullisesti chatbotin tai virkailijan kanssa, jotka tekevät dokumentoinnit, eikä kansalaisen tarvitse täyttää kaavakkeita.
- *Asioiden käsittely ja päätöksenteko nopeutuvat.* Asia tulee vireille heti, kun se on dokumentoitu järjestelmään. Tapaukseen tarvittava tiedon haku myös muista järjestelmistä ja tiedon analysointi tapahtuvat hetkessä koneellisesti. Päätös voidaan saada myös asiakkaan tietoon tosiaikaisesti. Vakuutustoiminnassa nopeutta on demonstroitu: vakuutusrahat voivat olla vakuutuksenottajan tilillä muutaman sekunnin kuluttua sen jälkeen, kun hakija aloitti korvaushakemuksen puhelimella.
- *Riippumattomuus vuorokauden ajasta.* Asian vireillepano, käsittelyn seuraaminen, mahdollisten lisätietojen toimittaminen ja päätöksen saaminen voi tapahtua mihin vuorokauden aikaan tahansa eikä vain virastoaikana, mikä niin ikään nopeuttaa asian käsittelyä asiakkaan näkökulmasta.
- *Tarkistusprosessi ja päätösten perustelujen esittäminen muuttuvat.* Jos järjestelmät ovat läpinäkyviä, perustelut ovat myös selviä. Mikäli käytetään oppivia algoritmeja, joiden toiminta ei ole läpinäkyvää, päätösten perustelut voivat olla haastavampia.
- *Yksilön parempi huomiointi päätöksenteossa.* Käytössä olevan kattavamman tiedon perusteella asiakkaan tilanne voidaan huomioida kokonaisvaltaisemmin. Se voi tarkoittaa sellaisen etuuden myöntämistä, mitä asiakas ei osaisi hakeakaan. Toisaalta asiakkaan voimaannuttaminen voi olla mahdollista ohjaamalla hänet ennakoivasti hänen tarvitsemansa palvelun äärelle ilman, että hän sitä itse osaisi kaivata. Äärimmillään tullaan kysymykseen, kuinka paljon viranomaiset voivat 'ohjailla' asiakasta.
- *Sekä asiakkaan että viranomaisen kyky kyseenalaistaa voi heikentyä,* kun kaikki tapahtuu nopeasti ja 'oikeiden' algoritmien tukemana.
- *Asiakkaan suhtautuminen viranomaiseen muuttuu,* kun keskustelukumppanina ei olekaan ihminen vaan kone. Algoritmin antamaa tulosta voi helpommin muuttella antamalla sille sellaisia tietoja, että tulos on halutun lainen ('laillinen huijaaminen')

- *Objektiivisuuskäsitys voi muuttua*, kun päätöksentekoon tulee enemmän dataa useammalta taholta.
- *Työn luonne muuttuu*. Järjestelmä tekee rutiinit ja helpot tehtävät ja jatkossa myös vaativia tehtäviä, mutta haasteelliset tapaukset jäävät asiantuntijan ratkaistavaksi. Asiantuntijalla on tukenaan suuri määrä dataa ja järjestelmän tekemiä ratkaisuehdotuksia. Asiantuntijan täytyy pystyä ymmärtämään isoja vaativia kokonaisuuksia päätöksiä tehdessään.
- *Eri virastojen välisen yhteistyön luonne muuttuu*. Tietojen yhteiskäyttö lisääntyy ja tiedolla voidaan ohjata asiakasta eri toimijoiden palveluihin.
- *Asiantuntijoiden koulutus muuttuu*. Vaaditaan suurta vastenopeutta niin ratkaisujen tekemisessä kuin järjestelmien ylläpidossa.
- *Sähköinen asiointi on paikasta riippumatonta asiakkaan kannalta*, mutta ratkaisuissa tarvitaan paikallista tietoa - tämä on otettava huomioon järjestelmien ja data-lähteiden suunnittelussa.
- *Uuden muuttujan huomiointi asian käsittelyssä*. Järjestelmän tulisi olla sellainen, että se sallisi uuden muuttujan lisäämisen.

### 3.2. Tekoälyn käyttöönoton nykytila

Tekoälyn ja ohjelmistorobotiikan käyttötapauksien kartoittamiseksi viranomaistoiminnassa toteutettiin haastattelututkimus eri ministeriöiden hallinnonalan organisaatioiden kuulemiseksi. Haastattelujen tarkoituksena oli selvittää mitä tekoäly- tai ohjelmistorobotiikkasovelluksia viranomaiset ovat jo ottaneet käyttöön tai mitä sovelluksia on tunnistettu tai viety suunnittelu- tai kokeiluvaiheeseen. Keskusteluissa käytiin myös läpi yhden tai useamman käyttötapauksen osalta mm. tekniseen ratkaisuun, vaikutuksiin, läpinäkyvyyteen, hankintaan ja kohdattuihin haasteisiin liittyviä seikkoja. Puolistrukturoidut haastattelut toteutettiin Liitteessä 1 esitettyjen kysymysten ohjaamana. Haastatteluja tehtiin heinä-lokakuussa 2018 yhteensä 13 kpl puhelimitse tai internetin välityksellä. Seuraavassa on esitetty yhteenveto eri viranomaisten käytössä tai suunnitteilla olevista tai ideoiduista tekoälyn ja ohjelmistorobotiikan käyttötapauksista.

**Kansaneläkelaitoksessa (KELA)** on siirrytty sähköiseen asiointiin, ja yli 60 % hakemuksista tulee sähköisesti. KELAssa on käytössä tai tunnistettu seuraavat kuusi peruskäyttötapauksia:

- Etuuksien ratkaisuprosessissa järjestelmä tekee kymmeniä miljoonia suorakorvauksia vuosittain, esim. opintotukipäätös opiskelupaikan perusteella tai lääkekorvaukset. Nämä päätökset ovat aina hakijan kannalta myönteisiä. Monimutkaisemmissa tapauksissa järjestelmä tarjoaa tietoa käsittelijälle.
- Asiakaspalveluun ja neuvontaan suunnitellaan chatbottia, joka neuvoo oikeaan osoitteeseen tai pystyy auttamaan hakemuksen teossa. Pidemmälle vietyinä Aurorapalvelu voisi tunnistaa asiakkaan tarpeita.
- Tiedolla ohjaaminen - ennakoitukykyä analytiikan avulla: tulevaisuudessa järjestelmä voisi tietoonsa perustuen ja asiakkaan elämäntilanteen huomioiden ohjata

asiakkaan hänen tarvitsemiensa palvelujen äärelle. Kaikki tapaukset eivät ole mahdollisia nykyisten säännösten puitteissa (asiakkaan pitää asioida, muuten ei voida antaa neuvoja).

- KELAn oman tutkimuksen apuna esimerkiksi aineiston keräämisessä ja analysoinnissa sekä aineiston luotettavuuden analysoinnissa.
- KANTA-palvelut. Kansalaisen omien terveystietojen tallentaminen ja jakaminen. Tulevaisuudessa terveydenhuoltotietoja sisältävä Kanta-palvelu voi laajeta Kansa-palveluksi, jossa on mukana myös sosiaalihuollon tiedot.
- Tulos- ja vaikuttavuusseuranta esim. kuntoutuspalveluista. Idea-asteella.

**Kyberturvallisuuskeskuksen** palveluissa on käytössä sääntöpohjaista ohjelmistorobotiikkaa muun muassa asiakaspalvelun automatisoituna tukena sekä tietojärjestelmien tiedonhaku automatisoituna taustamateriaalin ja lähteiden keräämisessä. Tarkoitus on siirtyä tekoälyn hyödyntämiseen esimerkiksi tietoturvaloukkauksen palvelun kehittämisessä (Havaro2 Datajärvi)

Havaro 2 -hanke on käynnissä oleva hanke, jonka osana tarkoitus on koota yhteen kansallinen tietovaranto, Datajärvi, johon tietoa tuottaisivat Kyberturvallisuuskeskuksen rinnalla useat vastaavan tyyppistä työtä tekevät organisaatiot, esimerkiksi teleoperaattorit ja tietoturva-alan yritykset. Tekoälyä on tarkoitus hyödyntää datajärveen kertyvän suuren tietomassan käsittelyyn, uuden tyyppisen analyysin toteuttamiseen sekä ilmiöiden ja mallien tunnistamiseen tietoturvariskeihin liittyen. Tarkoitus on hyödyntää koneoppimismenetelmiä asiantuntijoiden työn tukena. Tällä hetkellä datajärven taustajärjestelmiä ollaan perustamassa, eikä kuvattu tekoälyratkaisu ole vielä käytössä.

**Liikenneviraston** omassa selvityksessä on tunnistettu seuraavat käyttötapaukset tekoälyn hyödyntämiseen:

- väylänpidon päätöksenteon tuki
- liikennetilanteen hallinta, liikenteen ohjauspäätökset häiriöiden tunnistaminen ja niiden käsittely.

Ohjelmistorobotiikan tai tekoälyn osalta käytössä ei ole vielä sovellutuksia, lukuun ottamatta muutamia taloushallinnon ohjelmistorobotiikan kokeiluja. Liikenneviraston tavoitteena ei ole olla uuden teknologian kokeilukohde, vaan valita omat ratkaisut käyttöönoton toisessa aallossa valtiohallinnon kokeilujen tulosten perusteella.

Molempien käyttötapauksien taustalla on tavoite tukea kattavaa ja reaaliaikaista tilannekuvaa väylien tilasta ja kunnosta, liikenteen olosuhteista sekä liikenteen määrästä. Väylänpidon päätöksenteon tuen käyttötapauksessa on tarkoituksena tukea hitaan muutoksen tunnistamista, ennakoimista ja resurssien optimaalista käyttöä, esimerkkinä väylän kuluman arviointi. Liikennetilanteen hallinnan käyttötapauksessa tarkoituksena on reagoida nopeaan muutokseen, kuten säätilan myötä tehtäviin ohjaustoimenpiteisiin. Molemmissa käyttötapauksissa kuvantulkinta sekä runsaiden eri tyyppisten tietolähteiden käyttö on arvioitu olevan keskeisessä roolissa tulevaisuuden tekoälyratkaisuisissa.

**Maahanmuuttovirasto** on panostanut ohjelmistorobotiikan ratkaisuihin tukemaan asiakaspalvelua sekä virkamiestyötä:



- asiakaspalvelun chatbot
- automaattinen edellytysten tarkistus (käytössä jo joissakin asiaryhmissä)
- päätösten (täys)automatisaatio (suunnittelilla).

Sääntöpohjaisesti toimiva asiakaspalvelun chatbot, neuvontarobotti on otettu käyttöön keväällä 2018 Migrin verkkosivuilla. Chatbotin tarkoituksena on keventää puhelinneuvonnan kuormaa vastaamalla säännöllisesti toistuviin asiakaspalvelun kysymyksiin ja tarvittaessa ohjaamaan eteenpäin henkilövirkamiehelle. Vuodesta 2015 käytössä olleen ohjelmistorobottin tarkoituksena on kerätä ennen päätöksentekoa rekistereistä ja hakijan antamista tiedoista tapauksen käsittelijälle nähtäväksi päätöstä varten tarvittavat tiedot. Kullekin edellytykselle luodaan automaattisia sääntöjä, joiden avulla pyritään mekaanisesti selvittämään, täyttyykö kyseinen edellytys vai ei. Edellytysten tarkistus on sääntöpohjainen prosessi, jossa käsittelijä näkee, millaisia selvityksiä automatiikka on tehnyt ja millaisia tuloksia selvitykset ovat tuottaneet.

**Oikeusministeriön hallinnonalalla** on selvitetty ohjelmistorobotiikan sekä tekoälyn hyödyntämisen tulevaisuuden sovelluskohteita. Soveltuviksi käyttötapauksiksi on muun muassa selvitetty puheentunnistusta ja päätöksenteon tukea. Näiden lisäksi on tunnistettu seuraavat tekoälyn käyttötapaukset, jotka voisivat olla hyödyllisiä alalla: anonymisointi, tiivistelmien tuottaminen, automaattinen luokittelu ja riskiprofiilien laatiminen. Puheentunnistuksen kaavailtuna sovelluskohteena ovat litteroitavat pöytäkirjat. Päätöksenteon tuen kaavailtuja sovelluskohteita ovat lainopillisen päätöksenteon ja tuomarin työn tukeminen automaattisen tiedonkeruun ja vastaavien päätösten esittämisen avulla sekä olennaisten lähteiden kokoaminen. Ratkaisut edellyttävät luonnollisen kielen käsittelyä koneoppimismenetelmin. Kaikki tekoälyn käyttötapaukset ovat tällä hetkellä vielä suunnitteluasteella.

**Opetushallitus** on kartoittanut alalle soveltuvia ohjelmistorobotiikan ja tekoälyn käyttötapauksia erilaisten oppilaitosten ja koulutuksentarjoajaverkostojen kokeilujen kautta. Tällä hetkellä käynnissä on tekoälykokeiluja muun muassa opetukseen liittyvästä ennakoinnista, oppimisanalytiikan hyödyntämisestä, älykkäistä oppimispeleistä, opiskelijan edistymisen seurannasta sekä ohjauksen saatavuuden ja oikea-aikaisuuden kehittämisestä. Opetushallituksen sisällä on pohdittu tekoälyn hyödyntämistä (mm. chatbottien käyttöä) koulutukseen liittyvissä tiedotustehtävissä sekä asiakaspalvelussa. Opetushallitus, opetus- ja kulttuuriministeriö ja työ- ja elinkeinoministeriö vastaavat yhdessä AuroraAI:n Kiinni työelämässä osamisen kehittämisen avulla -työpaketista, jossa kehitetään ja kokeillaan tapoja tekoälyn hyödyntämiseen jatkuvan oppimisen ja työllisyyden kontekstissa huomioiden vastuullisuus ja eettiset näkökulmat.

**Rajavartiolaitos** on ohjelmistorobotiikan ja tekoälyratkaisujen osalta suunnittelu- ja testausvaiheessa. Ohjelmistorobotiikan ratkaisuja on välillisesti jo käytössä henkilöstöhallinnon prosesseissa Valtion talous- ja henkilöstöhallinnon palvelukeskuksen (Palkeet) tuottamina. Rajavartiolaitoksen sisäisessä valmistelussa on hallinnollisen työn automatisointiin soveltuvia ohjelmistorobotiikan ratkaisuja. Tulevaisuuden tehtävien toteuttamisessa tekoälyratkaisulla nähdään olevan keskeinen rooli. EU tasolla ollaan kulkemassa kohti tekoälyn hyödyntämistä rajavalvonnassa, mikä tulee asettamaan myös vaatimuksia kansalliselle rajavalvonnalle ja siinä käytettäville tietojärjestelmille. Lisäksi kasvavan informaation ja datamassojen myötä tekoälyn hyödyntäminen nähdään ainoaksi toteuttamiskelpoiseksi ratkaisuksi tulevaisuudessa. Rajavartiolaitoksella on kaksi keskeistä tekoälyn hyödyntämisen käyttötapauksia:

- sensori-informaation analysointi tekoälyä hyödyntäen



- matkustajatiedon ennakkokäsittelyn automatisointi ja analysointi.

Sensori-informaation, eli suurelta osin rajavartiolaitoksen kameroiden stillkuvan ja videokuvan analysoinnissa on tarkoitus ottaa käyttöön konepohjaista hahmontunnistamista tukemaan rajavalvontaa suorittavien viranomaisten työtä. Tekoälyratkaisulla haetaan ensivaiheessa tukea inhimillisten virheiden välttämiseksi sekä työn kuormittavuuden vähentämiseksi teknisten valvontalaitteiden kuvahavaintojen tarkistuksessa. Tulevaisuudessa tekoälyjärjestelmät voisivat tuottaa jalostuneempaa ja yksityiskohtaisempaa analyysitietoa kuvainformaatiosta. Lisäksi pidemmällä tähtäimellä tekoälyratkaisujen on visioitu kehittyvän johtamisen tueksi siten, että järjestelmät voisivat antaa toimintasuosituksia tai lähettää automaattisesti ilmoituksia kentälle. Matkustajatiedon ennakkokäsittelyyn, jossa rajavartiolaitos toimii poliisia tukevana viranomaisena, on suunnitteilla konepohjaista automatisointia ja analysointia. Tässä käyttötapauksessa tekoälyä voitaisiin hyödyntää muun muassa taustarekisterihakuihin ja niiden pohjalta tehtyyn analyysiin, joka tällä hetkellä toteutetaan manuaalisesti. Rajavartiolaitos toteuttaa kehitysratkaisuja tekoälyn hyödyntämiseen yhteistyössä EU-hankkeissa sekä toimittajien kanssa kahdenvälisissä kokeiluissa. Esimerkiksi EWISA Videovalvonnan kehittäminen EU:n määräysten valvonnassa -hanke on ollut käynnissä vuodesta 2014.

Yleisesti **sosiaali- ja terveydenhuollon** (sote) sektorilla on toistaiseksi käytössä mm.:

- potilastietokantoihin integroituja, esim. Duodecim -tietämyskantaa hyödyntäviä päätöksenteon tukijärjestelmiä lääkärintyössä
- asiakasrajapinnan chatbot -neuvontapalveluja sekä puhelimitse että verkossa.

Muita varhaisen tai testivaiheen käyttötapauksia ovat esim. röntgenkuvien ja syöpäsolusientöjen kuva-analyysit, potilasrajapinnan kansalaispalvelut (kuten hoidontarpeen arviointi tai laboratoriolähete ilman lääkärikäyntiä) sekä sosiaaliseen kanssakäymiseen ja kotona asumisen tukeen liittyvät robotit. Jatkossa voidaan edetä reseptin uusimisiin tai automaattisiin diagnooseihin, mutta ne vaativat vielä useampia kehitysaskelaita.

Toistaiseksi sovellukset ovat pääasiassa alhaisen tason tekoälyratkaisuja, esim. ammattikäyttäjän avustajan roolissa, sillä vakiintuneet käytännöt ja usein myös lainsäädäntö edellyttävät päätöksenteon tapahtuvan ihmisen toimesta.

**Kunnallinen sosiaali- ja terveydenhuollon esimerkki, Kerava.** Tekoäly ja erityisesti ohjelmistorobotiikka nähdään hyödyllisenä ja tarpeellisenä osana sosiaali- ja terveydenhuollon työtä. Kaikki rutiininomainen työ, kuten sosiaali- ja terveydenhuollon taloushallinto, on menossa kohti automatisaatiota. Ohjelmistorobotiikan hankkeita ollaan käynnistämässä lähivuosina. Tekoälyn hyödyntämisen osalta tunnistettuja käyttötapauksia ovat:

- Tuleva tietojärjestelmä Apotti. Järjestelmään liittyvät analytiikan palvelut, lääkärin avustaminen, potilastiedon pohjalta tapahtuva indikointi sekä proaktiivinen puuttuminen kuntalaisten sosiaali- ja terveydenhuollon riskeihin
- Kotihoito: sensorikka ja sensori-informaatiosta tunnistettavat poikkeamat.

Apotti-järjestelmä on sosiaali- ja terveydenhuollon toiminnanohjausjärjestelmä, jossa kaikki on samassa järjestelmässä. Uudella järjestelmällä haetaan digitaalisuutta, toimintavarmuutta, tehokkuutta ajankäyttöön ja yhdenmukaista toimintatapaa alalle. Tekoälyratkaisut, kuten analytiikan palvelut, osana uutta järjestelmään nähdään arvokkaana lisänä, mutta

edellä kuvatut ensisijaisina. Lisäksi järjestelmästä haetaan välineitä kuntalaisten hyvinvoinnin edistämiseen ja tasapuolisen kohtelun varmistamiseen. Apotti-järjestelmän osalta Kerava on tulossa mukaan vasta 2020, ja tällä hetkellä käynnissä on valmistautuminen uuteen digitaaliseen toimintatapaan.

Kotihoidossa on käytössä jo digitaalisia palveluita, kuten videokotikäyntejä. Sensoriteknologian myötä digitaalisia palveluita ollaan kehittämässä kohti koneen varassa tapahtuvaa analytiikkaa esimerkiksi poikkeamien tunnistamiseksi. Sensoriteknologian hyödyntämistä varten haetaan helposti mitattavia tekijöitä ja pilottikokeilujen aikataulu sekä teknologiset ratkaisut ovat vielä suunnitteluasteella.

**Suomen ympäristökeskuksen** selvityksessä nousi esille kaksi tekoälyn käyttötapaa, jotka molemmat liittyvät satelliittikuvan tulkintaan:

- Itämeren vedenlaatu
- maanpeitteen tulkinta.

Satelliittikuvan tulkinnalla on pitkä historia osana ympäristön seuranta. Tekoälyn - tarkemmin koneoppimismenetelmien - on tarkoitus kehittää edelleen satelliittikuvan prosessointia, jossa on edelleen ihmistyöhön perustuvia vaiheita. Koneoppimismenetelmien käyttöä puoltaa monimuotoiset tulkintakohteet, käsiteltävän datan määrän kasvu, koneen kyky käsitellä ja hallita taustamateriaalia sekä kansainvälinen kehitys menetelmien käyttöön. Osin kehitettävät menetelmät ovat jo kehittymässä ihmistä paremmaksi kuvan tulkinnassa. Tulkinnassa koneoppimissovellukset eivät tee itsenäisiä päätelmiä, vaan laativat suosituksia asiantuntijoiden työn tueksi, joiden tehtävä on tuottaa tietoarvoa satelliittikuvien pohjalta.

**Tilastokeskus** hakee tällä hetkellä tapoja hyödyntää automaattisia menetelmiä tilastotutannon tukena sekä uusien tietolähteiden käytössä. Tarkasteltavaksi käyttötapaukseksi valittiin tilastojen tuotantoa varten tarvittavan tietoaineiston ja datan tarkastaminen, validointi ja käsittely.

Käyttötapauksen tarkoituksena on nykyisten datankäsittelyprosessien pidemmälle viety automatisointi koneoppimisen ja ohjelmistorobotiikan keinoin. Uutta teknologiaa pyritään hyödyntämään tietotuotteiden raaka-aineen validoinnissa sekä käsittelyn uudistamisessa ja tehostamisessa. Järjestelmä tekee päätöksiä siitä, mikä osa sisään tulleesta datasta tarkastetaan ihmisavusteisesti ja mikä osa päästetään tuotantoon, toisin sanoen, miten ihmistyötä kohdennetaan.

**Valtion talous- ja henkilöstöhallinnon palvelukeskuksen (Palkeet)** osalta käyttötapauksina nostettiin tarkasteluun:

- talous- ja henkilöstöhallinnon ohjelmistorobotiikka ja tuotanto prosessien automatisointi
- Asiakaspalvelu AI -kehittämishanke (suunnitteilla).

Palkeet tuottaa valtion talous- ja henkilöstöhallinnon palveluita ja vastaa näihin toimintoihin liittyvien järjestelmien palveluiden kehittämisestä valtionhallinnossa. Käyttötapauksien on tarkoitus tukea tätä toimintaa. Jo käytössä olevaa ohjelmistorobotiikkaa käytetään tukemaan sääntöpohjaisia, aiemmin mekaanisesti ihmistyöllä tehtyjä, tarkastuksia. Ohjelmistorobotit muun muassa reitittävät menotositteita ja tarkastavat valtionhallinnon laskuja. Asiakaspalvelu AI -kehittämishanke on suunnitteilla oleva tekoälyn hyödyntämishanke, jonka tavoit-

teena on rakentaa tekninen alusta, jolle toteutetaan tekoälyratkaisuna chatbot tukemaan virkamieskunnalle tarjottavaa asiakaspalvelua ympäri vuorokauden. Tekoälyratkaisun ja sen mukanaan tuoman toimintakulttuurin muutoksen nähdään mahdollistavan asiakaslähtöinen toimintamalli, jossa voidaan yhdistellä paremmin olemassa olevia sisältöjä myös muiden valtionvirastojen palveluista.

**Verohallinnossa** automaatio on jo pitkään ollut isossa roolissa, ja esimerkiksi ohjelmistorobottiikka nähdään automaation tapaan uutena työkaverina. Kolme ohjelmistorobotiikan käyttökohdetta, jotka Proof of Concept (PoC) -vaiheen jälkeen on viety tuotantoon vaiheittain tammikuusta 2018 alkaen, ovat:

- varainverotus: massatallennuskohde tietojen tallentamiseen
- muutosverotus: asiakkaan kuulemiskirjeen lähettäminen
- verotarkastus: ajoneuvotietojen haku kolmannen osapuolen järjestelmästä (Trafin Ajoneuvoliikenteen tietojärjestelmä) internetin yli.

Nämä kolme käyttötapausta etenivät PoC-vaiheen jälkeen tuotantoon, ja toteutusteknologiat ja -kumppanit valittiin dynaamisessa hankintajärjestelmässä kilpailuttaen keväällä 2017. Käyttötapauksissa ohjelmistorobotiikka automatisoi tiettyjä prosessiosia, kuten tietojen poiminta eri järjestelmistä. Muita käyttötapauksia ovat esim. vakiintuneet ohjelmistorobotiikan sovellukset testausautomaatioissa. Tekoäly ja koneoppimista ei vielä hyödynnetä, mutta mahdollisia käyttötapauksia tutkitaan ja kokeillaan, esimerkkinä heikon tekoälyn chatbotit.

### 3.3. Eettisyyden ja hyväksyttävyyden tarkastelu valituissa käyttötapauksissa

Koska eettisyys on kontekstisidonnaista, täytyy konteksti kuvata eettisten kysymysten tarkastelua varten. Tieto- ja viestintäteknologian etiikan tarkasteluun on esitetty kontekstin kuvaamista käyttötapauksien avulla (esim. Lucivero 2016). Haastatteluissa tunnistettujen käyttötapauksien, työpajassa (21.8.2018) ideoitujen ja äänestettyjen käyttötapauksien sekä hankkeen ohjausryhmäkeskustelun perusteella esillä olleista käyttötapauksista valittiin kolme jatkokäsittelyyn. Tapauksien katsottiin tyypiltään edustavan sekä nykyisiä että tulevia sovelluskohteita, joihin liittyvät tilanteet ja haasteet koskevat useita organisaatioita. Työpajassa 8.10.2018 toteutettuun eettiseen arviointiin valitut käyttötapaukset olivat:

1. asiakaspalvelun chatbot
2. ennakoiva anomalian tunnistus
3. elinikäisen oppimisen järjestelmä.

Hankkeen viimeisen työpajan 27.11.2018 yhteiskunnallisen hyväksyttävyyden arviointiin nostettiin näistä käyttötapauksista pisimmälle käytäntöön levinnyt skenaario *asiakaspalvelun chatbot*. Sen lisäksi hyväksyttävyyden skenaariotyöpajassa käsiteltiin kaikkia käyttötapauksia leikkaavina aiheina *tekoälyjärjestelmien hankinta* sekä *tekoäly viranomaisen työkaverina*.

Työpajassa 8.10.2018 arviointi toteutettiin kolmiportaisena siten, että kunkin tapauksen kohdalla pohdittiin ensin **yleisiä eettisiä periaatteita**:

- **Ihmisarvo ja loukkaamattomuus:** Henkilöä tulee kunnioittaa. Teknologia ei saa loukata ihmisarvoa eikä ihmisyyttä.

- **Itsemääräämisoikeus:** Omistusoikeus omaan informaatioon. Henkilöllä tulee olla oikeus päättää itse mihin tarkoitukseen teknologiaa käytetään, jos käytetään ollenkaan.
- **Yksityisyyden kunnioitus ja suoja:** Henkilöllä tulee olla oikeus päättää häntä koskevan henkilökohtaisen tiedon käytöstä ja suojata fyysisistä, psyykkistä ja sosiaalista intymiteettiään.
- **Yhdenvertaisuus ja tasa-arvoisuus:** Oikeudenmukainen kohtelu; Palvelujen tulee olla kaikkien käyttäjäryhmien saatavilla riippumatta fyysisistä tai psyykkisistä ominaisuuksista.

Tämän jälkeen tarkasteltiin kunkin tapauksen erityisesti tekoälyn etiikkaan liittyviä **erityiskysymyksiä** seuraavien periaatteiden osalta:

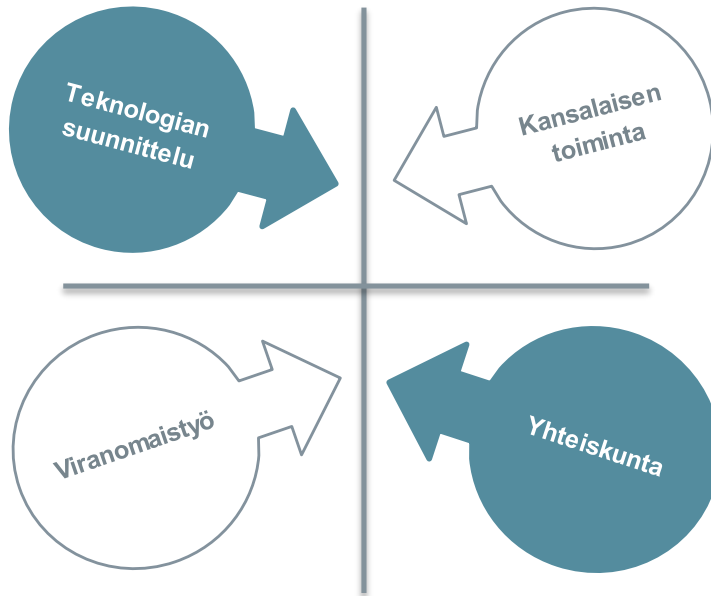
- **Tietoisuus:** Tietoisuus teknologian merkityksestä toiminnassa, teknologian toiminta-periaatteista ja sen mahdollisuuksista; Tasa-arvoisuus osaamisessa.
- **Avoimuus:** Informaation vapaa saatavuus, luotettavuus, oikeus tosiasiatietoon. Teknisen sovelluksen tulee olla luotettava siinä tehtävässä, mihin se on tarkoitettu. Se ei saa uhata käyttäjän fyysisistä tai henkistä terveyttä.
- **Kontrolloitavuus:** Teknologiaa tulee voida tarkastella ja varmistaa, että sovellukset eivätkä aiheuta vahinkoa kenellekään. Oikeus tarkistaa ja korjata omia tietoja.

Lopuksi käytiin läpi **viranomaistyön etiikan ja virkavastuun** kysymyksiä seuraavien periaatteiden valossa:

- **Teknologian rooli yhteiskunnassa:** Yhteiskunnan tulee hyödyntää teknologiaa siten, että sovellukset ja palvelut lisäävät elämänlaatua.
- **Viranomaisen vastuu:** Viranomaiset ovat vastuullisia järjestelmän toiminnasta, päätösvallan käyttämisen vastuu, mahdollisuus valittaa päätöksistä, mahdollisuus saada päätöksen perusteita, vaatimus työn oikeellisuudesta.

Kolmiportaisen arvioinnin avulla voitiin tunnistaa käyttötapauksissa keskeisiä eettisiä kysymyksiä huomioiden niin yleiset eettiset periaatteet kuin tekoälyn ja laajemmin teknologian mukanaan tuomat erityispiirteet sekä viranomaistyön luonne.

Jotta esitettyjen periaatteiden hahmottaminen tulisi konkreettisemmaksi, periaatteiden tarkastelussa käytettiin apuna nelikenttää, jonka muodostavat teknologian suunnittelu, kansalaisen toiminta, viranomaistyö ja yhteiskunta (Kuva 1).



Kuva 1. Eettisen tarkastelun nelikenttä.

"Teknologian suunnittelu" huomioi tietojärjestelmien ja niiden käyttöliittymien suunnittelun sisältäviä spesifejä eettisiä kysymyksiä ja vaatimuksia. "Kansalaisten toiminta" arvioi yksittäisen kansalaisen eli yksilön näkökulmasta viranomaistoimintaa sekä palvelujen hyödyntämistä ja käyttöä. "Viranomaistyö" arvioi toimintaa viranomaisten sekä viranomaistyön vaatimusten näkökulmasta. "Yhteiskunta" arvioi tekoälyn ja ohjelmistorobotiikan viranomaiskäyttöä yksilöä ja yksittäisiä toimintoja laajemmin koko yhteiskunnan näkökulmasta. Kaikki edellä esitetyt näkökulmat tulisi pyrkiä huomioimaan, kun arvioidaan tekoälyn ja ohjelmistorobotiikan käyttöä osana viranomaistyötä.

Työskentelytapana käytettiin hiljaista aivoriittä, parityöskentelyä ja yhteiskeskustelua. Keskusteluissa esiin nousseet ajatukset asetettiin osallistujien keskusteltavaksi eettisen tarkastelun nelikenttään.

Työpajan 27.11.2018 käyttötapausskenaarioiden yhteiskunnallisia hyväksyttävyysskysymyksiä arvioitiin hyödyntäen elämänlaadun mittaamiseen soveltuvia kriteerejä:

- ihmisarvo
- yksityisyys
- fyysinen koskemattomuus
- tarkoituksenmukaisuus
- ratkaisujen oikeutus
- reiluus
- luottamus
- saavutettavuus
- tasa-arvo
- avoimuus
- vastuullisuus
- läpinäkyvyys, jäljitettävyys
- ennakoitavuus
- oikeus kasvatukseen
- osallistaminen ja kuulluksi tuleminen
- sananvapaus ja oikeus tietoon
- oikeus osallistua uuden toimintakulttuurin luomiseen.

Seuraavissa alaluvuissa esitellään sekä eettisen arvioinnin työpajan käyttötapausten tulokset (asiakaspalvelun chatbot, ennakoiva anomalian tunnistus ja elinikäisen oppimisen ja osaamisen järjestelmä) että yhteiskunnallisen hyväksyttävyyden työpajan skenaariotulokset (asiakaspalvelun chatbot, tekoälyjärjestelmien hankinta ja tekoäly viranomaisen työkalu).

## Asiakaspalvelun chatbot

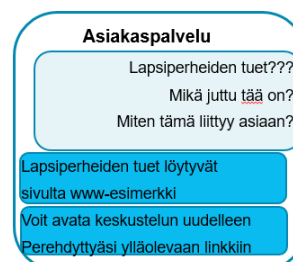
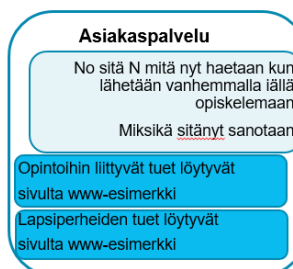
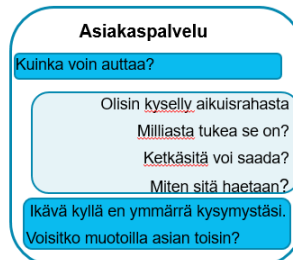
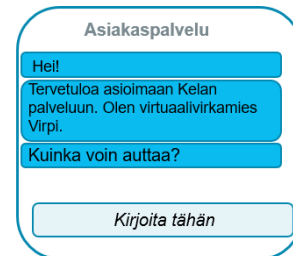
Chatbot-skenaarion avulla pyrittiin selvittämään tekoälyjärjestelmien ymmärrettävyyttä kansalaisen näkökulmasta. Chatbotteja on jo otettu käyttöön myös viranomaispalveluissa. Chatbot tarjoaa ensikäden opastusta ja voi antaa vastauksen moniin yksinkertaisiin kysymyksiin tai ohjata kysyjän oikeaan osoitteeseen säästämällä näin kaikkien osapuolten aikaa ja nopeuttamalla palvelua. Yksinkertaiselta palvelulta kuulostavan chatbotin käyttöön liittyy monia, kaikille käyttötapausten yhteisiä eettisyyden ja hyväksyttävyyden liittyviä kysymyksiä, joiden tapaus on hyödyllistä tarkastella.

Haja-asutusalueella asuva kansalainen Matti Meikäläinen, 52 v., on päättänyt hakea etuisuutta ja ottaa netin kautta yhteyttä viranomaiseen selvittääkseen, mihin etuuteen hän on oikeutettu. Palvelun etusivulla hän törmää chat-ikkunaan:

1. Matti aloittaa keskustelun tuesta hyvin rennosti, mutta asiakaspalvelu ei ymmärrä Mattin muotoilua.

2. Matti muotoilee kysymyksen uudelleen. Asiakaspalvelu tunnistaa sanat "opiskelu" ja "vanhemmalla". Asiakaspalvelu ohjeistaa Mattin perehtymään www-sivuihin tarkemmin tunnistaakseen tuen mistä halusi kysyä.

3. Matti hämmentyy vastauksesta. Matti ei saa asiaansa hoidettua eikä kysymykseen vastausta. Hänelle kuitenkin syntyy käsitys, että tämä on paras vastaus mitä asiakaspalvelusta voi saada. Matti pettyy palveluun ja sulkee keskustelun.



Seuraavaksi tarkastellaan asiakaspalvelun chatbottia yhtenä esimerkkinä chatbotteihin liittyvistä yhteiskunnallisista kysymyksistä.

**Toiminta ja päämäärä:** Toiminnan tavoitteena on asiakaspalvelun puhelinneuvonnan kuorman keventäminen ja asiakaspalvelun laadun parantaminen.

**Järjestelmä:** Asiakaspalvelun laatua pyritään parantamaan sääntöpohjaisesti toimivan asiakaspalvelun I. chatbot-järjestelmän kautta. Järjestelmä toimii verkkosivuilla ympäri vuorokauden. Se vastaa kysymyksiin ennalta määrättyjen sääntöjen pohjalta suomeksi ja englanniksi. Järjestelmä tarjoaa tietoja käsittelyajoista sekä yhteystiedoista, kuten sähköpostiosoitteista ja toimipisteiden osoitteista. Tarvittaessa järjestelmä ohjaa eteenpäin henkilövirkamiehelle. Järjestelmä on otettu käyttöön keväällä 2018.

**Toimijat:** Toimijat liittyvät järjestelmään ainakin kahdella tasolla: asiakkaat ja kehittäjä/ylläpitäjät. Ensimmäisessä järjestelmää käyttävät hyvin erilaisista kulttuuritaustoista tulevat asiakkaat, jotka hakevat tietoa palveluista ja yhteystiedoista. Usein tämä on asiakkaan ensikontakti viranomaisen toimintaan. Laajin näkymä järjestelmään on omilla kehittäjillä/ylläpitäjillä, jotka tarkkailevat botin toimintaa varmistuen sen toimivuuden asiakaspalvelussa sekä erityisesti annettujen neuvojen oikeellisuuden. Kehittäjä ja ylläpitäjät tuntevat prosessit sekä asiakaspalvelun tavoitteet ja ovat perehtyneet ohjelmistorobotiikkaan niin, että kykenevät toteuttamaan muutoksia järjestelmään.

**Kohde:** Toiminnan kohteena on asiakaspalvelutiedon antaminen sekä viranomaisen palveluissa ohjaaminen. Chatbot vastaa yleisimpiin ja yksinkertaisimpiin asiakaspalvelun kysymyksiin keventäen vastaamisen kuormaa puhelinpalvelusta. Toiminnalla tavoitellaan asiakkaan omia valmiuksia hakea ja saada tarvitsemaansa tietoa.

**Kontekstit:** Järjestelmä toimii useiden kontekstien rajapintana. Näitä ovat asiakaspalveluprosessit, viranomaisorganisaatio, verkko, sekä asiakkaiden laajasti poikkeavat kulttuuriset ja koulutukselliset taustat.

Taulukossa 1 on esitetty eettisyystarkastelun tulokset jaoteltuna neljään eri tarkastelunäkökulmaan.

Taulukko 1. Asiakaspalvelun chatbot: eettisyys- ja hyväksyttävyyksymyksiä eri toimijoiden näkökulmista.

Teknol. suunnittelu	Kansalaisen toiminta	Viranomaistyö	Yhteiskunta
<ul style="list-style-type: none"> <li>Saako kuvata itseään vapaamuotoisesti: luonnollinen kieli, metaforien käyttö, vai alusvetovalikot?</li> <li>Itsekorjaavuus: Poikkeamatilanteiden raportointi on oleellista.</li> <li>Millainen "luonne" botille sopii? Millaiseen bottiin luotettaisiin?</li> <li>Kuinka paljon on mahdollista saada taustatietoa selville heti keskustelun alussa? Miten? Tietääkö AI heti ihmisen taustat yhdistelemällä tietoja toisiinsa? Onko tämä hyvä asia?</li> <li>Miten chatbot saa käyttäjän äidinkielen tai kielitaidon selville?</li> </ul> <p>Tulisiko botin olla ääneltään "mies" vai "nainen"?</p>	<ul style="list-style-type: none"> <li>Onko oikeus hylätä chatbot-keskustelu ja vaatia keskustelukumppaniksi oikeaa ihmistä?</li> <li>Jos ei luota, ei pysty antamaan tunnistetietoja (esim. maahanmuuttajat).</li> <li>Chatbotissa ollaan anonyymejä ja kerrotaan helpommin erilaisia ja enemmän asioita. Kansalainen tarvitsee kykyä muotoilla asiansa eksaktisti.</li> <li>"Juhannusyön syndrooma": herkässä mielentilassa voidaan kertoa itsestä asioita, joita ei normaalisti kerrottaisi.</li> <li>Näkö- ja kuulovammaiset: miten käytetään bottia?</li> </ul>	<ul style="list-style-type: none"> <li>Kansalaisen tulee saada tietää, miksi hänet on ohjattu chatbotin keskustelukumppaniksi</li> <li>Kansalaisen tulee saada tietää, millä perusteella dataa kerätään. Läpinäkyvyys luo ja säilyttää luottamusta.</li> <li>Jos kansalaisille annetaan mahdollisuus kertoa itsestään vapaasti, saisiko viranomainen näistä "kipupisteistä" hyödyllistä tietoa? Millaisen tiedon kerääminen on oikein?</li> </ul>	<ul style="list-style-type: none"> <li>Luottamuksen säilyminen on olennaista. Chatbotin kanssa keskustelu voi olla vastenmielistä; miten silloin lisätä kansalaisten luottamusta?</li> <li>Miten luodaan kansalaisille luottamus siihen, että chatbot toimii yhtä luotettavasti kuin ihminen?</li> <li>Jos chatbot olisi tekoälyn varassa toimiva luonnollisen kielen käsitteijä, ja oppisi käyttyjen keskustelujen perusteella, tulisiko kansalainen paremmin kuulluksi?</li> </ul>

Chatbotin sisältämän palvelun laatu ja tarkoitus on käytävä ilmi käyttäjälle. Jokaisella kansalaisella on oikeus tulla toisen ihmisen kuulluksi. Näin ollen heti aluksi on oltava selvää, onko keskustelija ihminen vai robotti, ja käyttäjällä tulee olla mahdollisuus päästä myös juttelemaan ihmisen kanssa. Palvelun tarkoitus, kuten käyttäjän ja robotin roolit ja vastuut, pitäisi olla näkyvissä käyttäjälle kuten myös palvelun laajempi tarkoitus koko viranomaisprosessissa. Chatbotin toimintaperiaatteen ymmärtäminen lisää käyttäjän mahdollisuutta esittää oikeita kysymyksiä, joihin palvelusta on mahdollisuus saada vastauksia. Palvelun käytön perustelujen (nopeampi ja kattavampi palvelu, säästöt, jne.) esittäminen lisäävät läpinäkyvyyttä ja hyväksyttävyyttä.

Palvelun toiminnasta tulisi olla selkeät ohjeet käyttäjälle: miten palvelun kanssa kommunikoidaan, mitä chatbot ymmärtää tai ei ymmärrä (kieli, murteet, lausemuodot, sanojen sijamuodot, jne.). Ohjeiden perusteella tulisi käydä ilmi mm. perustuuko palvelun käyttämä tunnistus avainsanoihin vai voiko palvelu ymmärtää asioita ja asiayhteyksiä kokonaisvaltaisesti esimerkiksi hahmontunnistukseen perustuen.

Käyttöliittymän suhteen kansalaiset on hemmoteltu kaupallisten palveluiden hyvillä käyttöliittymillä ja ne ovat muodostuneet standardiksi. Käytön tarkoituksenmukaisuuteen, helppouteen ja selkeyteen tulisi siis kiinnittää huomiota.

Tietoturva on oleellinen asia. Käyttäjän tulisi saada tietää, mitä hänen antamillaan tiedoilla tehdään, miten niitä käytetään ja mihin ne mahdollisesti myös voivat päätyä. Pitäisi myös



olla selvää, missä vaiheessa keskustelu on vain keskustelua ja milloin tilanne muuttuu viralliseksi käyttäjän asian hoitamiseksi, mihin liittyy luottamuksellisen tiedon jakamista. Selvää tulisi myös olla, millaisia oikeuksia käyttäjällä on, ja että chatbotin kanssa käyty keskustelu ei vaikuta varsinaiseen asiointipäätökseen.

Koneeseen suhtaudutaan eri tavalla kuin ihmiseen ja käyttäjälle voi olla houkutus kokeilla chatbotin rajoja antamalla sille jopa väärää tietoa tai käyttää sitä keskustelukumppanina yksinäisyyden lieventämisessä.

Chatbot-palvelun kehittyminen käytön seurauksena pitää suunnitella ja tehdä läpinäkyväksi myös käyttäjälle. Palvelu ei voi kehittyä pelkästään sen perusteella, millaista kommunikointia käyttäjien kanssa tapahtuu, vaan tarkoituksenmukaisuus ja eettiset periaatteet on otettava huomioon.

Mikä on 'avuttoman' chatbotin oikeutus? Edut chatbotin käytöstä tulisi esittää käyttäjälle. Niitä voi olla nopeampi ja täsmällisempi palvelu ilman turhia kontakteja ja odotusta, palvelun saatavuus jopa kaikkina vuorokauden aikoina, paikkariippumattomuus, käytettävyyden parantuminen (ei turhauttavia kaavakkeiden täyttöjä), jne. Käyttäjälle on kuitenkin tarjottava myös vaihtoehtoa chat -palvelulle.

Viranomaispalvelun kokonaisprosessi on arvioitava uudelleen, kun chatbot on osa sitä, ja muotoiltava sellaiseksi, että chatbot yhtenä sen osana tuottaa optimaalisen hyödyn eikä ole lisävaihe kokonaisuudessa. Samalla on tarkasteltava myös sitä, kuka ja miten päättää siitä, mikä on hyväksyttävää.

## Ennakoiva anomalian tunnistus

**Toiminta ja päämäärä:** Toiminnan tavoitteena on muuttaa viranomaisten tapaa toimia kansalaisten aloitetta odottavasta toiminnasta proaktiiviseksi, palveluita tarjoavaksi ja resurssija kohdentavaksi toiminnaksi. Käyttökohde-esimerkkejä voivat olla muun muassa Kelan etuuskien tarjoaminen, ennakoiva sosiaalitoimi ja terveystalvet.

**Järjestelmä:** Tekoälyjärjestelmä (koneoppimiseen perustuva järjestelmä), joka kykenisi tunnistamaan huomionarvoisia tapauksia ja poikkeuksia viranomaisia tai kansalaisia paremmin. Anomalia on distanssi keskiarvosta, mutta anomalia syntyy vasta siitä, että tehdään sisältöanalyysi: 1) matemaattinen anomalia (esim. 1/1000), 2) sisällön analyysi, 3) synteesi - anomalian määrittely (viranomaisen kannalta merkittävä/huomionarvoinen tapaus).

**Toimijat:** Järjestelmän kanssa toimivat virkamiehet, valtionhallinnon järjestelmäasiantuntijat, palveluntuottajat sekä valvova taho.

**Kohde:** Resurssien kohdentaminen oleellisten kohteiden tunnistamisen kautta, kuten apua eniten tarvitsevien tai tehokkaan puuttumisen tarpeessa olevien tunnistaminen. Neuvonta, suositukset, seuraaminen.

**Kontekstit:** Laajasti dataa kokoava ja käsittelevä järjestelmä toimisi useiden viranomaiskontekstien rajapinnassa, jossa kohtaavat hallinnon eri organisaatiot, verkko- sekä palvelutuottajien moninaiset kontekstit, esim. kuka tuottaa ja missä.

Taulukossa 2 esitetään Ennakoiva anomalian tunnistus -käyttötapausten tarkastelussa tunnistetut eettisyys- ja hyväksyttävyyssymykset.

Taulukko 2. Ennakoiva anomaliatunnistus: eettisyys- ja hyväksyttävyyksymyksiä eri toimijoiden näkökulmista.

Teknol. suunnittelu	Kansalaisen toiminta	Viranomaistyö	Yhteiskunta
<ul style="list-style-type: none"> <li>• Vaatii erityisosaimista sekä domainista että tekoälyjärjestelmästä.</li> <li>• Poikkeamien analyysiosijat: ilmiöiden tunteminen on oleellista.</li> <li>• Pystytäänkö tunnistamaan myös tukijärjestelmän ja algoritmien anomaliat?</li> <li>• Ratkaisujen perusteltavuus? Koneen loogikka? Koneoppimisen luonne?</li> <li>• Mikä on oikea yleisyyden taso? Liian yleinen on ylimalkaista, liian yksilöity on tungettelua.</li> <li>• Ihmisarvo: Muuttuuko anomalioiden tunnistaminen profiloinniksi? Kuinka arviointeja tehdään?</li> <li>• Järjestelmän voi tunnistettua myös omat anomaliansa (Suunnittelija, viranomainen).</li> <li>• Tietoisuus, avoimuus, kontrolloitavuus. Kuinka profiloinnit tehdään? Onko ymmärrettäviä ja voidaanko perustella? (Suunnittelija, viranomainen).</li> <li>• Onko palvelu luotettava -&gt; luottamus (koskee kaikkia 4 tahoja).</li> </ul>	<ul style="list-style-type: none"> <li>• Voiko puhua valinnanvapaudesta?</li> <li>• Järjestelmällä pelaaminen.</li> <li>• Käyttäjän oikeus tietää, mihin tietoihin profilointi perustuu ja kenellä käyttöoikeus tietoihin?</li> <li>• Ihmisarvo: Tunteeko ihminen, että palvelu on apu ja tuki vai että se on kontrollia ja valvontaa.</li> <li>• Onko pakko olla tavoitettavissa ja mukana? Verotus, terveydenhuolto, sosiaaliturva / Onko systeemin ulkopuolella elämää?</li> <li>• Kohteena olevan kokemus: itsemääräämisoikeus, luotettavuus (miksi minä), luottamus siihen, että omia tietoja on käsitelty asianmukaisesti.</li> <li>• Oikeus omien tietojen korjaamiseen ja oikeus tulla poistetuksi. Miten tehdään? Oikeus kieltäytyä.</li> <li>• Muuttuvatko päätelmät itseään toteuttaviksi ennusteiksi? Miten ihminen ymmärtää ja käsittelee mahdollisesti yllättäen tulleen tiedon (luokittelun)?</li> </ul>	<ul style="list-style-type: none"> <li>• Tavoite: Säästää rahaa ja palvelu paremmin. Kumpi on ensisijaista?</li> <li>• Järjestelmien on palveltava perustehtävää.</li> <li>• Kuka saa hakea ja yhdistää tietoja?</li> <li>• Mitä viranomaisvastuu koskee? Tietojen hakua/Tietojen yhdistämistä/Järjestelmien toimivuutta.</li> <li>• Tietoisuus, avoimuus, kontrolloitavuus: Kaikki perustuu keskinäiseen luottamukseen ja viranomaisten ja kansalaisten sitoutumiseen toistensa hyvään osana omaa hyvää.</li> </ul>	<ul style="list-style-type: none"> <li>• Yhdenvertaisuusperiaate: Voiko järjestelmän ulkopuolella saada yhtä hyvän palvelun? (Viranomainen, yhteiskunta).</li> <li>• Yhdenmukaisuudesta pahe; leikkaa pois luovuuden, ihmisyhteen kuuluvat rodot, poikkeavuuden, joka tuottaa jotain uutta? (Viranomainen, yhteiskunta)</li> <li>• Kaikki universaalit periaatteet voivat vaarantua, jos systeemiin ei voi luottaa (Viranomainen, yhteiskunta)</li> <li>• Miten kriteeri poikkeamasta määritellään?</li> <li>• Tiedon käyttöoikeus?</li> <li>• Koostettu data; Mihin kaikkeen voidaan ja saadaan käyttää?</li> <li>• Kuinka järjestelmää valvotaan? Kuinka järjestelmää kehitetään?</li> <li>• Henkilö nostetaan esiin erilaisuuden perusteella.</li> <li>• Kuinka poikkeavuus määritellään?</li> <li>• Jo järjestelmän kutsuminen anomalioiden tunnistamiseksi on viesti, epäkunnioittava viesti.</li> </ul>

## Elinikäisen oppimisen ja osaamisen järjestelmä

**Toiminta ja päämäärä:** Toiminnan tavoitteena on innostaa ja aktivoida kansalaisia oman osaamisen kehittämiseen ja tätä kautta tulevaisuuden vaatimuksiin. Tämä tapahtuu kahden erillisen toiminnon yhdistämisen kautta: ensinnäkin kokoamalla kansalaisille oppimisprofiileja omista oppimisen tavoitteista sekä niiden oppilaitosten rekistereistä, joissa opintoja on suoritettu. Toiseksi kokoamalla yhteen kuratoidusti täydennyskoulutuksen tarjoajien kurssit.

**Järjestelmä:** Tekoälyjärjestelmä, joka kykenee kohdentamaan relevantteja sisältöjä luotettavasti kansalaisen yksilöityihin tarpeisiin.

**Toimijat:** Kansalaiset, kuratoivat virkamiehet, valtionhallinnon järjestelmäasiantuntijat, palveluntuottajat ja valvova taho. Kansalaiset ovat täysi-ikäisiä jo jonkin koulutustaustan omaavia (ammattillisista opinnoista korkeakoulutettuihin). Järjestelmää kuratoivat virkamiehet ovat koulutusalan ammattilaisia, jotka tuntevat koulutuksen yhteiskunnalliset tavoitteet ja työelämän vaatimukset. Järjestelmäasiantuntijat ovat valtiohallinnon työntekijöitä, jotka ovat perehtyneet teknisiin vaatimuksiin ja pystyvät osoittamaan vaatimuksia palveluntuottajille.

**Kohde:** Kansalaisten elinikäisen oppimisen toteuttaminen kohdentamalla henkilökohtaisten kiinnostusten sekä työelämän vaatimusten kannalta merkityksellisiä pieniä opintokokonaisuuksia. Järjestelmä ohjaa kansalaista opintoja kohti tarjoamalla opintoja valittaessa vaihtoehtoja personoidussa muodossa jokaisen omia lähtökohtia ymmärtäen. Järjestelmän ulkoisten puitteiden kautta oppiminen pyritään muuttamaan jatkuvaksi ja elinikäiseksi.

**Kontekstit:** Laajasti dataa kokoava ja käsittelevä järjestelmä toimisi useiden viranomaiskontekstien rajapinnassa, jossa kohtaavat hallinnon ja oppilaitosten organisaatiot sekä verkko- ja palveluntuottajien moninaiset kontekstit.

Taulukossa 3 esitetään Elinikäisen oppimisen ja osaamisen järjestelmä -käyttötapaukseen liittyvät työpajassa tunnistetut eettisyys- ja hyväksyttävyyssymykset.

Taulukko 3. Elinikäisen oppimisen ja osaamisen järjestelmä: eettisyys- ja hyväksyttävyyssymyksiä eri toimijoiden näkökulmista.

Teknol. suunnittelu	Kansalaisen toiminta	Viranomaistyö	Yhteiskunta
<ul style="list-style-type: none"> <li>• Miten huomioida työntekijöiden tuottama data? Onko mahdollista sisällyttää? Haaste eettisesti?</li> <li>• Miten kansalaisen tulisi reagoida järjestelmän suosituksiin?</li> <li>• Suunnittelu vs. ylläpito, säilyykö kokonaiskuva?</li> <li>• Yhä yksityisemmän datan kerääminen ja tietosuojan varmistaminen entistä jännitteisemmässä suhteessa? Kansalaisten järjestelmään syötämä data on turvattu. Se ei saa joutua kolmansille osapuolille.</li> <li>• Kontrolloitavuus. Kansalaisen syötämän tiedon väärinymmärryksen välttäminen.</li> <li>• Voidaanko varmistaa, että palvelu parantaa elämän laatua?</li> <li>• Miten toiveet kirjataan järjestelmään? Onko valmiita pohjia, liukuja vai vapaata tekstiä?</li> </ul>	<ul style="list-style-type: none"> <li>• Itsemääräämisoikeus.</li> <li>• Jääkö ihmisen oikeus omaan dataan vain unelmaksi?</li> <li>• Mydatan hyödyntäminen.</li> <li>• Itsemääräämisoikeus, henkilö omistaa itseään koskevan data ja päättää sen käytöstä.</li> <li>• Haavedata = kuva kansalaisen haaveista ja unelmista erittäin herkkää ja arkaluonteista.</li> <li>• Oikeus kulkea koko elämän läpi.</li> <li>• Itsemääräämisoikeus vs. yhteinen hyvä.</li> <li>• Teknologia mahdollistajana, käyttäjän roolin korostaminen käytössä.</li> </ul>	<ul style="list-style-type: none"> <li>• Vaihdetaan koulutusprofiili osaamisprofiiliksi, jonne voi tallentaa muutakin elämäkokemusta.</li> <li>• Koulutusta tarjoavat tahot voisivat saada tietoja koulutuksen tarjoamisen tarpeista järjestelmän kautta.</li> <li>• Muodostaako osaamisen omakannan?</li> <li>• Miten kielioppivirheet ja kansalaisten heikot kirjalliset kyvyt huomioidaan?</li> <li>• Elämäkokemus ja koko elämäkaari mukaan järjestelmään.</li> <li>• Vaatimus avoimuuteen ja vastuu systeemisistä vaikutuksista.</li> <li>• Kaksoisrooli koulutuksen suunnittelijana vs. työtarjonnan hallinta.</li> <li>• Virkavastuu tekoälyjärjestelmässä? Työparina tekoälyn kanssa?</li> <li>• Vastuu työelämää koskevan tiedon realiteetista.</li> <li>• Kyky toimia ajassa.</li> </ul>	<ul style="list-style-type: none"> <li>• Järjestelmän oikeus tappaa kansalaisten unelmia?</li> <li>• Työllisyys- ja osamistarpeiden ennakointi, järjestelmän tarjoaa oikeaa tietoa.</li> <li>• Tietoisuus jatkuvasta muutoksesta, tarve jatkuvalle kommunikoinnille yhteiskunnan kanssa.</li> <li>• Yksityisyys vs. valtion kontrolli.</li> <li>• Saako tietoihin koskea ilman lupaani?</li> <li>• Yksityisyyden kunnioitus ja suoja. Jalostettava henkilötieto oltaisiin anonyymia.</li> </ul>

## Tekoälyjärjestelmien hankinta

Tarkastelun lähtökohtana on kuvitteellinen skenaario tekoälyjärjestelmän hankintaprosessista viranomaistyön tueksi. Skenaarion avulla työapajassa pohdittiin, miksi skenaariossa esitetty kuvitteellinen hankinta epäonnistui, mitä olisi voitu tehdä paremmin ja miten tietojärjestelmän hankinnassa epäonnistuminen vaikuttaa laajemmin kansalaisten luottamukseen viranomaistyötä kohtaan. Työskentelyssä pyrittiin tunnistamaan kansalaisen näkökulmasta myös haasteita tai riskejä valtionhallinnon tekoälyjärjestelmien hankintaan liittyen sekä pohtimaan erityisesti luottamusta, hyväksyttävyyttä ja mahdollisia kehitysideoita.

Vaiheittain etenevä skenaario:

1. Esimerkkiministeriö tunnistaa tarpeen oman toiminnan tehostamiseen ja päättää hankkia käyttöönsä tekoälyjärjestelmän. Ministeriöllä ei ole omaa tekoälyosaamista

hankinnan suunnittelun, järjestelmätoteutuksen tai ylläpidon osalta. Hankinnan suunnittelussa turvaudutaan konsulttityöhön.

2. Tekoälyjärjestelmä tilataan yhdessä konsulttien asiantuntemuksen kanssa laadittujen vaatimusten perusteella. Konsultti edustaa tekoälyjärjestelmäosaamista ja hankinnan suunnittelussa mukana olevat virkamiehet oman hallinnonalansa osaamista. Kilpailutuksen kautta järjestelmätoimittajaksi valikoituu kansainvälinen toimittaja. Järjestelmätoimittaja räätälöi ministeriön tarpeisiin sopivan ratkaisun ja kouluttaa henkilökuntaa.
3. Ministeriön prosessiin tulee muutos ja järjestelmä ei ole kykenevä reagoimaan muutokseen yhtä nopeasti kuin virkamieskunta aiemmin. Käytön aikana selviää, ettei järjestelmäkumppani ole ymmärtänyt ministeriön omia tarpeita riittävän selvästi. Syntyy epäselvyyttä järjestelmätoimittajan kanssa siitä, miten muutokset toteutetaan. Järjestelmätoimittajaa joudutaan vaihtamaan. Ministeriö menettää järjestelmään kertyneen älyllisen pääoman, lähdekoodit, mallit, jne. Hankinta joudutaan aloittamaan puhtaalta pöydältä.

Työpajatyöskentelyässä päällimmäisenä kysymyksiä ja pohdintaa herätti kysymys siitä, poikkeako tekoälyjärjestelmän hankkiminen perinteisten tietojärjestelmien hankinnasta. Keskustelussa esille nousi vastauksena em. kysymykseen tekoälyjärjestelmien poikkeavuuksina seuraavat seikat: uudenlainen kompleksisuus, korkeammat osaamisvaatimukset sekä mahdolliset laajemmat vaikutukset tapaan tehdä työtä. Skenaarion kuvaaman hankintaosaamisen puutteen ja siitä seuraavien epäonnistumisien nähtiin syövän viranomaisyön uskottavuutta ja siten heikentävän luottamusta. Kilpailutujen tekoälyosaajien rekrytoiminen ja hankintaosaamisen kehittäminen ovat tahtokysymyksiä, joihin myös julkisella sektorilla on mahdollisuuksia, kun eri resursseja yhdistetään. Julkiselle sektorille on esimerkiksi muodostunut konsortioita, joiden avulla on paikattu rajallisia resursseja ja saatu siten käyttöön riittävää tekoälyosaamista.

Tekoälyjärjestelmien hankinnan lähtökohdat nousivat esille toiminnan hyväksyttävyyden ehdonä. Mitä hankinnalla tavoitellaan ja miksi hankinta tehdään? Lähtökohtana tulisi olla kansalaisten palveleminen ja palvelun parantaminen, kustannustehokkuus ei yksin riitä. Voisiko olla olemassa yleistä ohjeistusta valtionhallinnon tekoälyjärjestelmien hankintoja varten? Näin voitaisiin taata hankinnoissa selvä roolijako, jossa kuvataan esim. tilaajat, omistajat, käyttäjät sekä työkalut ja vastuut..

Hankinta tulisi avata arkikielelle. Kaikki hankinnan tavoitteet, vaiheet ja toimijat tulisi olla ymmärrettävissä kansalaiselle ilman erityisosaamista. Tämän kaltaisen toiminnan nähtiin paitsi tukevan kansalaisten ymmärrystä sekä luottamusta, mutta myös edesauttavan valtionhallinnon sisäistä ymmärrystä siitä, mitä hankinnalla tavoitellaan ja miten. Ymmärrettäväksi tekeminen helpottaisi laajempaa kansalaiskeskustelua sekä osallistumista viranomaispalveluiden tarvemäärittelyyn kansalaisten näkökulmasta. Ymmärrettävyyden rinnalla yksityisyyttä pidettiin hyväksynnän reunaehtona. Kysymyksenä nousi esille: Miten yksityisyys taataan uusissa järjestelmissä? Miten käydä riittävää keskustelua, ettei hankinnassa synny epäselvyyttä siitä, kuka lopulta omistaa ratkaisut?

Julkisuudessa esillä olleet tiedot epäonnistuneista tietojärjestelmähankinnoista sekä järjestelmiin liittyvät uhkakuvat luovat uusien hankintojen ympärille negatiivisen kierteen, joka syö luottamusta hankintaa, järjestelmää ja viimekädessä viranomaistoimintaa kohtaan. Tulevaisuudessa tarvitaan enemmän tietoa onnistumisista. Tarvitaan voitonarratiivi siitä, että hankinta voi onnistua ja sillä voidaan saavuttaa tavoiteltuja hyötyjä. Tämä tukisi kansalaisten

luottamusta siihen, että veronmaksajien rahoja käytetään oikein. Avoimuus hankintoihin liitetyistä kustannuksista, hyödyistä ja siitä, millainen muutos syntyy, auttaisi myös ymmärtämään suurempia investointeja. Negatiivisen kierteen ongelmana nähtiin lisäksi uhkakuvien pelko, vaikka todellisia uhkakuvia ei olisikaan. Huonot kokemukset ja pelot ohjaavat ajatuksia, jolloin ohjeet ja lainmukainen toiminta eivät riitä luottamuksen saamiseksi. Huolenaiheita ovat esim.: Voinko asioida suomeksi/ruotsiksi? Mihin tietoni päätyvät?

Hyväksyttävä toiminta edellyttää kommunikaatiota sekä vastuunkantoa, kun toiminnassa epäonnistutaan. Viranomaistyöllä tulee olla oikeus oppia virheistä ja korjata toimintaansa. Epäonnistumisia pitää pystyä käsittelemään ja niistä tiedottamaan. Myös vastuukysymykset tulee olla selvät varsinkin silloin, kun hankinnoissa on mukana ulkopuolisia osapuolia. Epäonnistumisien pohdinta tiivistä hyväksynnän kysymykseen: Jos tarve on perusteltu, saako hankinnassa mennä ”metsään”?

### **Tekoäly viranomaisen työkaverina**

Tarkastelun lähtökohtana esitettiin kuvitteellinen skenaario tulevaisuuden viranomaistyöstä, jossa tekoälyjärjestelmä toimii virkamiehen ”työparina”. Jo tällä hetkellä valtionhallinnon alaisuudessa on useampia yksittäisiin tehtäviin keskittyviä ohjelmistorobotiikan ratkaisuita, joita nimitetään ”digityökaveriksi” tai ”robottityökaveriksi”. Skenaarion avulla työpajassa pohdittiin, miten virkamiehen työ muuttuu, jos rinnalla toimisi yhä laaja-alaisempaan ja moniulotteisempaan hahmottamiseen kykenevä tekoälyjärjestelmä, jolla olisi myös kyky tehdä aiemmin puhtaasti ihmisille kuulunutta työtä. Työskentelyssä pyrittiin tunnistamaan kansalaisen näkökulmasta myös haasteita tai riskejä esitettyyn viranomaistoimintaan liittyen sekä pohtimaan erityisesti luottamusta, hyväksyttävyyttä ja mahdollisia kehitysideoita.

Vaiheittain etenevä skenaario:

1. Ministeriön alaiseen virastoon hankitaan omaa prosessia tukeva tekoälyjärjestelmä, joka tukee virkamiesten työtä, oppii olemassa olevista prosesseista ja toteuttaa niitä itsenäisesti. Viraston työntekijöiden työ helpottuu ja aikaa vapautuu.
2. Välikaudella prosessi kulkee sekä ihmisen että tekoälyjärjestelmän kautta, ja näiden välinen roolijako hämärtyy synnyttäen epäselviä vastuutilanteita. Ihmistyö painottuu asiantuntijuuteen, ja työtavat muuttuvat merkittävästi.
3. Käytön myötä tekoälyjärjestelmä hioutuu viraston prosessin tarpeisiin, mutta mahdollisesti myös kangistuu sen kaavoihin. Henkilöstöä on voitu vähentää tai kohdistaa ja kouluttaa uudella tavalla. Virastosta on tullut järjestelmästä riippuvainen, ja kokonaisnäkyminen prosessiin on kaventunut.

Työpajakeskusteluissa kiinnitettiin huomiota siihen, onko hyödyllistä vai vaarallista puhua tekoälyjärjestelmistä työkaverina. Varsinkin alkuvaiheen sovelluksissa voitaisiin ennemminkin puhua tukiälystä ja työntekoa avustavista järjestelmistä. Tekoälyn inhimillistämiseen ja koneen ja ihmisen vastakkainasetteluun liittyy riskejä, ja ylipäänsä tulisi tehdä tietoon pohjautuvia päätöksiä siitä, onko tekoälyltä syytä odottaa ja vaatia samaa kuin ihmiseltä. Aihepiiristä riippuen (kuten päätöksenteon läpinäkyvyys, virheiden sietäminen tai vastuukysymykset) yhdessä tapauksessa koneelta vaaditaan enemmän kuin ihmiseltä, toisessa vähemmän ja kolmannessa tapauksessa konetta käsitellään samoin kuin ihmistä.

Hyväksyttävyyteen liittyvinä avoimina kysymyksinä esitettiin mm. millä tavalla uskomme ja luotamme tekoälyjärjestelmään tai koneeseen. Kokemukset ja viestintä muokkaavat osaltaan tätä luottamusta, ja esimerkiksi tilanne, jossa kone oppii valehtelemaan, voi murentaa hyväksyttävyyttä nopeasti. Rakentuuko järjestelmää kohtaan nykyisiin virkamiehiin verrannollinen luottamus vai onko luottamus vähäisempää vai suurempaa? Riskien ohella on myös tunnistettavissa suurta parannuspotentiaalia, mikäli tekoälyllä onnistutaan toteuttamaan ja viestimään yhä parempi tasapuolisuus ja oikeudenmukaisuus viranomaispäätöksissä.

Avoimuus ja läpinäkyvyys, esimerkiksi viranomaispäätöksen perustelut ja siihen liittyvä valitusoikeus, ovat korostuneen tärkeitä käytettäessä tekoälyä viranomaisprosesseissa. Tarkoituksenmukaisen tason löytäminen on kuitenkin haaste. Esimerkiksi algoritmien avaamisessa ei tavallisen kansalaisen näkökulmasta ole hyötyä, mutta se aiheuttaisi suuren väärinkäytösten riskin. Tarkoituksenmukaista läpinäkyvyyden tasoa tulee etsiä myös järjestelmän opettamiseen käytettyjen aineistojen ja datan suhteen. Keskusteluissa tunnistettiin tilanteita, joissa kasvava toimintavarmuus ja luottamus järjestelmään voi tarkoittaa sen läpinäkyvyyden ja toimintojen perusteltavuuden heikentymistä. Tällaisiin tilanteisiin tulisi löytää eri näkökulmia tasapainottavia ratkaisuperiaatteita.

Hyväksyttävyyden ja luottamuksen kannalta tekoälyjärjestelmien käyttöönotosta viestimisellä on suuri merkitys. Tämä koskee paitsi viraston sisäistä viestintää myös ulkoista kansalaisviestintää ja tiedotusta. Keskustelun politisoituminen ja vaikkapa järjestelmään pettyneiden virkamiesten yksittäiset avautumiset sosiaalisessa mediassa voivat aiheuttaa suurta haittaa luottamukselle, vaikka itse järjestelmä toimisi moitteetta. Huomiota tulee kiinnittää myös käyttöönotettavan organisaation muutosten hyväksyttävyyteen, ja siihen, miten esimerkiksi työtavat, -kulttuuri ja suhde työnantajaan kehittyvät.

Työpajatyöskentelyssä todettiin, että varsinkin tekoälyjärjestelmien vaiheittaista kehitystä ajatellen on tärkeää tunnistaa sovelluskohteet ja prosessit, joissa tekoäly ensimmäisenä ja helpoimmin tuottaa positiivisia tuloksia. Sääntöpohjaisuuteen taipuvat toiminnot ovat esimerkiksi nopeastikin toteutuskelpoisista kohteista, joissa myös luotettavuutta ja hyväksyttävyyttä on helpompi rakentaa. Monimutkaisemmat ja esimerkiksi harkintaa edellyttävät toiminnot taas vaativat enemmän, ja varsinkin aikaisen vaiheen käyttökohteissa tekoäly lienee näissä hyvin rajatussa roolissa oleva, ihmistyötä täydentävä apu.

Tekoälyn hoitaessa omaa osaansa viranomaistyöstä ihmistyön määrä ja etenkin merkitys eivät välttämättä suinkaan laske. Vapautuville ihmisresursseille löytyy paljon tarvetta ihmiskohtaamista vaativissa tehtävissä ja esimerkiksi päätöksiin kohdistuvien valitusten käsitteilyssä. Myös uudet tehtävät, kuten järjestelmien opettaminen, valvonta ja seuranta, koneen avustaminen pulmatilanteissa tai järjestelmien uudistaminen viranomaistehtävien muutostilanteissa vaativat ihmisresursseja.

Riippuvuus tekoälyjärjestelmistä nähtiin nykyjärjestelmiin liittyvien riippuvuuksien kanssa analogisena. Hankintaosaaminen sekä hinnan ja muiden kriteerien tasapainottaminen ovat ratkaisun osia näihin ongelmiin.



## 4. KANSALAINÄKÖKULMA TEKÖÄLYYN VIRANOMAISKÄYTÖSSÄ

### 4.1. Kansalaiskysely yhteiskunnallisesta hyväksyttävyydestä

Suomessa luottamus viranomaiseen ja viranomaistoimintaan, kuten poliisiin, on perinteisesti ollut korkea. Luottamusta ylläpidetään muun muassa selkeän sääntelyn avulla, kuten määrittelemällä viranomaistoiminnan laadulliset edellytykset, viranomaisen toiminnan rajat ja lakisidonnaisuus julkisen vallan käytön edellytykseksi. Viime sijassa luottamus kuitenkin palautuu kansalaisten kokemukseen.

Selvityksen tässä osiossa tekoälyn viranomaiskäytön hyväksyttävyyttä tarkastellaan kyselytutkimuksen avulla. Kyselyllä selvitettiin kansalaisten suhtautumista viranomaistoiminnan digitalisointiin, sekä erilaisten ohjelmistorobotiikka-, automaatio- ja tekoälysovellusten käyttöön liittyviä toiveita ja pelkoja. Lisäksi kyselyllä kartoitettiin sitä, millaisia ehtoja kansalaiset asettavat erilaisten tekoälysovellusten käytölle, millaisilla edellytyksillä kansalaiset ovat valmiita hyväksymään tekoälyn käytön viranomaistyössä, sekä luottamuksen murentumiseen vaikuttavia tekijöitä.

Kyselyn tulokset heijastavat yksilöiden kokemusta oikeudenmukaisuudesta ja digitalisaatiosta, eikä niinkään sitä, ovatko järjestelmät oikeuden- tai lainmukaisia jonkin ulkoisen mittapuun mukaan. Tulokset eivät siten mittaa tekoälyn käytön normatiivisia edellytyksiä, vaan ennemminkin kansalaisten intuitiivisesti tuntemaa luottamusta tai kokemusta. Kyselytutkimus on laadittu heinä- ja elokuussa 2018 yhteistyössä Pentagon Insight Oy:n kanssa. Tiedonkeruu toteutettiin Taloustutkimuksen verkkopaneelissa elokuussa 2018 ja kyselyyn vastasi 1031 18-79 -vuotiasta suomalaista. Vastaaajajoukko painotettiin edustamaan Suomen 18-79 -vuotiasta väestöä iän, sukupuolen ja asuinalueen mukaan. Kyselytutkimuksen tulokset on esitetty [Tekoälyn hyväksyttävyyden viranomaistyössä ja julkisissa palveluissa](#) -raportissa.

#### Asenteet ja suhtautuminen erilaisiin sovelluskohteisiin

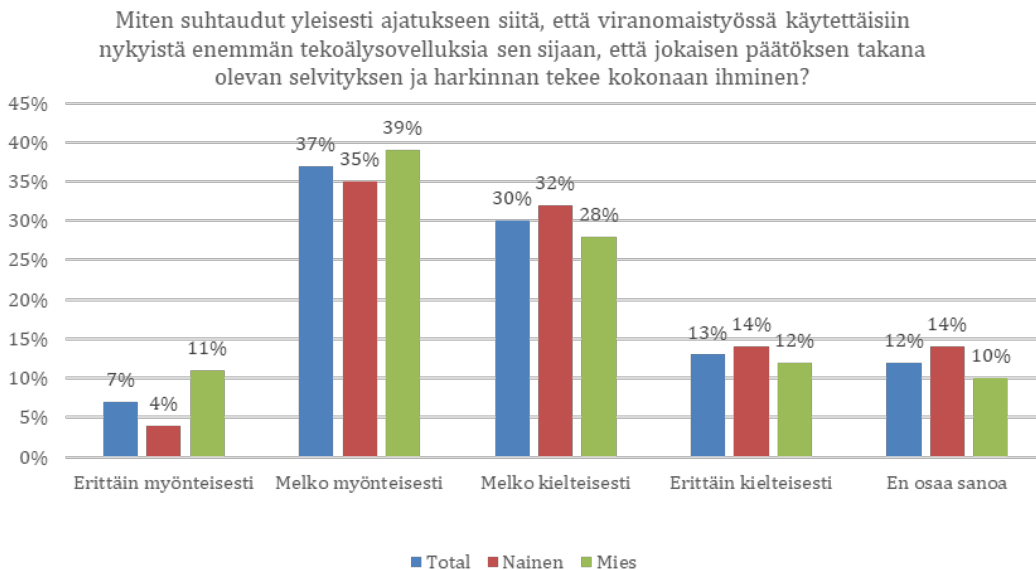
Tekoälyjärjestelmiä ja niiden käyttöä on kritisoitu voimakkaasti siitä, että niiden toimintalogiikan selittäminen ei aina ole mahdollista. Toisin sanoen, järjestelmää käyttävä ihminen tai viranomainen ei voi järjestelmän toimintaa tarkastelemalla päätellä, millä tavalla se toimii tai miten tuottaa päätöksiä. Lisäksi tekoälyteknologiaan liittyy voimakkaita myyttejä, jotka voivat olla omiaan herättämään epäluuloja kyseisen teknologian käyttöä kohtaan.

Vastaaajista kaksi kolmasosaa (65 %) koki, että he eivät tiedä riittävästi tekoälyn hyödyntämisestä viranomaisessa. Arvioidessaan oman tietämisen tasoa ikäluokkien välillä on merkittäviäkin eroja. 18-24 -vuotiaat uskovat tietotasoonsa eniten: 54 % uskoo tietävänsä varmasti tai melko varmasti riittävästi tekoälyn käytöstä viranomaisessa. Koko kyselyn keskiarvon oli 28 %. 65-79 -vuotiaat kokevat eniten epätietoisuutta: 75 % sanoo ettei tiedä melko varmasti tai varmasti riittävästi tekoälyn käytöstä viranomaistoiminnassa.

Huolimatta siitä, että kansalaiset kokivat tekoälyn käytön itselle vieraaksi asiaksi, suuri osa vastanneista ei silti tyrmännyt ajatusta suoralta kädeltä. Sen sijaan tekoälyn käyttöön viranomaistoiminnan tukena suhtauduttiin maltillisesti (Kuva 2). 44 % suhtautuu erittäin (7 %) tai melko (37 %) myönteisesti, 43 % suhtautuu erittäin (13 %) tai melko (30 %) kielteisesti. Kun



jakaumaa tarkastellaan ikäjakauman mukaan, nuoret (18-24 -vuotiaat) suhtautuvat tekoälyn hyödyntämiseen kaikkein myönteisemmin. Tässä ikäryhmässä erittäin kielteisesti tekoälyn käytön lisäämiseen suhtautui 0 % vastaajista.



Kuva 2. Suhtautuminen tekoälyn käytön lisäämiseen yleisesti.

Suuri osa vastaajista ei ollut omaksunut voimakkaan kielteistä tai myönteistä kantaa. Tämä voi tarkoittaa sitä, että vastaisuudessa suoritettavilla toimenpiteillä voi olla merkittäväkin vaikutus siihen, millaiseksi kansalaisten suhtautuminen digitaalisiin palveluihin yleensä tai tekoälykehitykseen muuttuu.

## Yleistä viranomaistoiminnan kehitystarpeista

Yleisellä tasolla etenkin nuoret kokivat kansalaisia koskevien viranomaispäätösten nopeuttamisen (63 %) ja päätösten tekemisen läpinäkyväksi (51 %) tärkeiksi kehityskohteiksi viranomaistoiminnassa. 65-79 -vuotiaat nostivat keskeisimmiksi kehityskohteiksi asiointin ja omien tietojen tarkastelun helpottamisen. Kysyttäessä, mitkä olisivat hyväksyttäviä syitä lisätä nimenomaan tekoälyn käyttöä olettaen, että hyödyt todella toteutuisivat, kansalaiset kokivat etenkin viranomaisessa asiointin helpottamisen ja päätösten tasapuolistumisen hyväksyttäväksi syiksi. Tärkeimmät viisi syytä on kyselyn perusteella:

1. Työntekijöiden ajan vapauttaminen niiden asiakkaiden palvelemiseen, jotka tarvitsevat apua eniten (62 % vastaajista)
2. Kansalaisia koskevien viranomaispäätösten teon nopeutuminen ja odotusaikojen lyhentyminen (55 % vastaajista)
3. Kustannusten vähentäminen (47 % vastaajista)
4. Kansalaisia koskevien viranomaispäätösten tarkkuuden parantaminen ja tasalaatuisuus (43 % vastaajista)
5. Kansalaisia koskevien viranomaispäätösten oikeudenmukaisuuden ja puolueettomuuden lisääminen (42 % vastaajista).

Kun jakaumia tarkastellaan ikäluokittain, nuoret erottuvat muista ikäluokista tekoälyn käytön motivaatiossa. 18-24-vuotiailla korostuu etenkin kustannusten vähentämisen merkitys (67 % vastaajista, koko kyselyn keskiarvon ollessa 47 %), sekä työn tehokkuuden kasvattaminen (51 % vastaajista, koko kyselyn keskiarvon ollessa 41 %).

Naiset korostavat miehiä selvästi enemmän eniten apua tarvitsevien palvelemisen tehostamista. 69 % naisista pitää sitä hyvänä syynä lisätä tekoälyn käyttöä, miehistä 56 %. Miehiä sen sijaan motivoi tekoälyn käyttöönottoon työn tehokkuuden kasvattaminen. 50 % miehistä pitää tätä hyväksyttävänä syynä lisätä tekoälyn käyttöä, naisista 32 %. Miehet nimeävät lisäksi kustannusten vähentäminen hyväksyttäväksi syyksi lisätä tekoälyn käyttöä. 53 % miehistä nimeää tämän hyväksyttäväksi syyksi, naisista 41 %.

## Tekoälyn hyödyntäminen eri sovelluskohteissa

Kyselytutkimuksen keskeinen tavoite oli selvittää sitä, kuinka vastaajat suhtautuivat tekoälyn käyttöön tietyissä prosesseissa ja sovelluskohteissa. Yleisenä huomiona voidaan sanoa, että tutuimpien käyttökohteiden automatisoimiseen tai hoitamiseen osin- tai kokonaan tekoälyn avulla suhtaudutaan myönteisemmin. Tarkastellut käyttökohteet olivat esitäytetty veroehdotus, tiedon etsiminen viranomaisen verkkosivuilta, rutiininomaisten päätösten tekeminen (kuten opintotuen jatkohakemuksen käsittely ja ratkaisu), yhteiskunnallisesti toivottavien valintojen edistäminen valvonnalla ja ohjauksella (kuten liikenneturvallisuuden edistäminen tai tautiepidemioiden ennakointi), virtuaaliset assistentit kuten chatti-neuvonta, poliisin voimavarojen ja valvonnan ohjaaminen, uusintarikollisuuden todennäköisyyden arviointi ja lääketieteellisten diagnoosien tekeminen.

85 % vastaajista suhtautui erittäin tai melko myönteisesti tekoälyn hyödyntämiseen etsittäessä tietoa viranomaisten verkkosivuilta. Esitäytetyn veroilmoituksen automatisoimiseen suhtautui myönteisesti 87 % vastaajista. 59 % vastaajista suhtautui erittäin tai melko myönteisesti tekoälyavusteisten chatbottien eli viranomaisen verkkosivuilla avustavan virtuaalisten assistenttien käyttöön. Kaikissa edellä mainituissa esimerkeissä tekoälyn tai ohjelmistorobotiikan funktio on pääasiassa tiedon etsimisen ja prosessoimisen helpottaminen, ja sen rooli on ennen kaikkea avustava. Rutiiniluontoisten päätösten, kuten opintotuen jatkohakemuksen käsittelyn automatisointiin suhtauduttiin niin ikään myönteisesti (72 % vastaajista).

70 % vastaajista koki yhteiskunnallisesti tärkeiden asioiden valvomisen ja edistämisen tekoälyavusteisesti erittäin tai melko myönteisenä. Myös uusintarikollisuuden arviointiin tekoälyn avulla suhtauduttiin melko positiivisesti, jopa 49 % vastaajista koki tekoälyavusteisen arviointisovelluksen käytön erittäin tai melko myönteisenä. Toisaalta jopa 16 % vastaajista ilmoitti, että eivät osaa arvioida tällaisen sovelluksen hyväksyttävyyttä. On huomattava, että näissä sovelluskohteissa tekoälyn funktio ei enää ole rutiininomaisten prosessien tukeminen, vaan niillä ennemminkin ennustetaan tulevia käyttäytymismalleja ja jopa puututaan yksilöiden toimintaan tai vapauksiin.

Sen sijaan merkittävästi vastaajan oikeuksiin vaikuttavien henkilökohtaisten toimien hoitamiseen tekoälyllä suhtauduttiin kielteisemmin. Kaikkein kielteisintä suhtautuminen oli lääketieteellisten diagnoosien tekemiseen tekoälyavusteisesti. 35 % vastaajista suhtautui tällaisen sovelluksen käyttöön erittäin tai melko myönteisesti, erittäin tai melko kielteisesti siihen suhtautui 51 % vastaajista. Toisaalta myös uusintarikollisuuden arviointi jakoi kantoja, 35 % vastaajista suhtautui sovellukseen melko tai erittäin kielteisesti.

## Tekoälyn hyödyntämisen rajoista

Kun vastaajilta kysyttiin yleisellä tasolla sitä, millaisiin toimintoihin he ylipäättänsä olisivat valmiita hyväksymään tekoälyn käytön, ilmeni vastauksista, että hyväksyttävyyden kokemus laski sitä mukaa, mitä itsenäisempi asema tekoälysovelluksella on. 33 % oli valmis hyväksymään tekoälyn käytön tukitoiminnoissa, joissa tekoäly ainoastaan avustaa tietojen syöttämisessä ja käsittelyssä. Ihmisen tehtäväksi jää tiedon analysointi ja päätösten tekemisestä vastaaminen.

35 % vastaajista on valmis asettamaan tekoälyn hyödyntämisen ylärajan rutiiniluontoisten toimintojen ja päätösten tuottamiseen. Tekoälylle siis ollaan valmiita hyväksymään tietynlainen autonomia, mutta ainoastaan helpoissa tapauksissa. Hyväksyttävyyys lähtee laskuun puoli-itsenäisen tekoälysovelluksen kohdalla. Yksi viidesosa vastaajista on valmis hyväksymään sellaisen tekoälysovelluksen käytön, jossa järjestelmä itsenäisesti sekä tuottaa päätöksiä että pistää ne täytäntöön. Ihmisen rooli näissä järjestelmissä on puuttua ainoastaan vaikeisiin tapauksiin.

5 % vastaajista oli valmiita hyväksymään sellaisen tekoälyjärjestelmän, jossa ihminen vastaa ainoastaan järjestelmän kehittämisestä. 3 % vastaajista määritteli hyväksyttävyyden rajaksi korkeimman autonomian asteen, eli sellaisen järjestelmän jossa koneet toimisivat täysin itsenäisesti. 5 % vastaajista ei hyväksynyt tekoälyn käyttöä missään tilanteissa.

## Hyväksyttävyyden reunaehdot – pelot

Kyselyn viimeisessä osassa selvitettiin kansalaisten viranomaistoimintaan kohdistamia odotuksia ja tekoälyn käyttöön liittyviä pelkoja. Pelkojen kartoittamisen lisäksi vastaajilta kysyttiin, millaisia positiivisia puolia tekoälyn käytön lisäämiseen yhdistetään. Kuten edellä todettiin, kansalaiset eivät ole toistaiseksi omaksuneet erityisen vakiintuneita käsityksiä tekoälyn käytön hyväksyttävyydestä. Etenkin melko myönteiset ja melko kielteiset kannat ovat sellaisia, joihin voidaan vaikuttaa. Jotta varmistetaan myönteisten kantojen kasvu, tulee tekoälysovelluksia ottaa käyttöön siten, että ne vahvistavat kansalaisten myönteisiä kokemuksia ja hyödyttävät kansalaisia.

65 % ovat huolestuneita inhimillisen harkinnanvaran katoamisesta, kun tekoälyn käyttöä viranomaisen työn välineenä lisätään. Inhimillisen harkinnan poistuminen on selvästi suurin huoli. Etenkin 18-24 -vuotiaat vastaajat (80 %) kokivat tämän uhkaksi. Sen sijaan vanhempien ikäluokkien suhtautuminen inhimillisen harkinnan katoamiseen oli lievempää. 65-79 vuotiaista 56% piti inhimillisyyden katoamista uhkana. Avointen vastausten perusteella huolena on etenkin inhimillisyyden vaarantuminen päätöksenteon yhteydessä ja yksilöllisen harkinnanvaran katoaminen päätöksenteosta. Esiin nousi myös huoli sellaisten kansalaisten syrjäytymisestä, jotka eivät ymmärrä tekoälyä tai sen käyttöä.

Muut isoimpina pidetyt uhkakuvat liittyvät viranomaistoiminnan läpinäkyvyyteen ja prosessien selkeyteen: 39 % on huolissaan siitä, etteivät tiedä ketkä kaikki henkilötietoja käyttävät ja 36 % on huolissaan siitä, että eivät voi olla varmoja siitä, etteivät henkilötiedot leviä ulkopuolisten käsiin. Tietojen käsittelyyn liittyvissä uhkakuvissa eri ikäryhmien välillä on suurta vaihtelua. Nuorimmasta ikäryhmästä vain 15 % nimesi uhkaksi sen, että eivät tiedä kuka käyttää henkilötietoja. Sen sijaan 50-64- ja 65-79 -vuotiaista 44 % ja 45 % nimesivät tämän yhdeksi suurimmista riskeistä.

Naiset olivat miehiä enemmän huolissaan tekoälyn aiheuttaman syrjäytymisen lisääntymisestä. 51 % naisista ja 36 % miehistä pelkäsi, että tekoälyn käyttö lisää etenkin sellaisten

ihmisten tarpeiden syrjäyttämisen, jotka eivät ymmärrä tai voi ymmärtää tekoälyä teknologiana. Vanhemmissa ikäryhmissä syrjäytymisen lisääntyminen koettiin niin ikään merkittävämpänä uhkana kuin nuoremmissa ikäryhmissä. Sen sijaan yleistä eriarvoisuuden lisääntymistä ei koettu yhtä suurena uhkana kuin syrjäytymistä. Vain 18 % vastaajista nimesi tämän yhdeksi huolenaiheeksi.

Itse päätösten laatu ja oikeudenmukaisuus huolettaa vastaajia selvästi vähemmän. Vain alle viidesosa, 19 %, pelkää päätösten oikeudenmukaisuuden heikkenevän. Lisäksi vain 16 % esitti pelkonaan sen, että luottamus viranomaiseen voisi heiketä tekoälysovellusten käytön seurauksena. Kaikkein vähiten huolestuttaa se, että yhä suurempi osa palveluista tuotettaisiin tulevaisuudessa sähköisesti.

## Hyväksyttävyyden reunaehdot – odotukset

Perustuslain 2 § mukaan julkisen vallan käytön tulee perustua lakiin, ja kaikessa julkisessa toiminnassa on noudatettava lakia. Tämä on luonteva lähtökohta myös tekoälyavusteisen viranomaistoiminnan arvioinnille. Lisäksi kuten edellä on todettu, yksilöiden tuntema luottamus ja viranomaistoiminnan legitimizeetti voidaan joiltain osin palauttaa selkeään lainsäädäntöön ja lakia toiminnassaan noudattavaan viranomaiseen. Kyselyn viimeisessä osiossa selvitettiin kansalaisten kokemusta siitä, millaisia odotuksia tekoälyä hyödyntäville viranomaisille asetetaan. Vastaajilta kysyttiin, miten eri seikat vaikuttaisivat vastaajan suhtautumiseen tekoälyn käytöstä.

Vastauksissa korostui etenkin selkeän sääntelykehikon luomisen merkitys hyväksyntää ja luottamusta lisäävänä tekijänä. 81 % vastaajista sanoi suhtautuvansa myönteisemmin tekoälyn käytön lisäämiseen, mikäli käytölle laaditaan selkeät ohjeet ja rajoitukset. Ohjeet ja rajoitukset voivat olla sääntöjä, mutta ennen kaikkea velvoittavia lakeja. Toiseksi suurin hyväksyntää lisäävä tekijä kyselyn perusteella on tekoälyn tekemien päätösten toimintalogiikan avaaminen asianosaisille ja asianosaisten mahdollisuus saada tietoa siitä, mitä tietoa tekoäly on yksittäisessä päätöksessä tai yksittäisessä toiminnassa hyödyntänyt. 78 % vastaajista koki, että päätösten yksityiskohtainen selittäminen ja yksityiskohtainen tieto tekoälyn prosessoimasta tiedosta lisäävät hyväksyttävyyden kokemusta.

76 % vastaajista piti tärkeänä, että viranomainen tiedottaa selkeästi, mikäli päätöksenteossa on käytetty tekoälyä. Niin ikään 76 % vastaajista koki, että heidän myönteinen suhtautuminen tekoälyn käyttöön kasvaa, mikäli myös tekoälyn suorittamista toimista ja päätöksistä vastaa aina ihminen. Vastauksissa on selkeästi havaittavissa läpinäkyvyyden ja avoimuuden merkitys. Useissa tekoälyn eettisiä reunaehtoja käsittelevissä julistuksissa läpinäkyvyyden toteuttamiselle on asetettu erilaisia kriteereitä. Päätöksenteon perustelu asianosaisille muodostaa eräänlainen minimivaatimuksen. Osassa kansainvälisiä julistuksia on ehdotettu myös viranomaisen käyttämän tekoälyohjelmiston lähdekoodin julkaisemista. Oikeustila on tältä osin kuitenkin avoin.

Se, että viranomainen käyttää päätöksenteon tukena ainoastaan viranomaisen hallussa olevaa tietoa ja muiden tietojen hyödyntämiseen pyydetään aina erikseen lupa, vaikutti 63 % vastaajan kohdalla hyväksyntää lisäävänä tekijänä. 61 % vastaajista koki suhtautuvansa myönteisemmin tekoälyyn, mikäli lailla säädettäisiin selkeästi myös yksityisten toimijoiden mahdollisuudesta käyttää tekoälyä omassa toiminnassaan, kuten esimerkiksi luotonannossa. 56 % vastaajista koki, että julkisella arvokeskustelulla olisi myönteistä merkitystä heidän suhtautumiseensa tekoälyn käytön lisäämiseen.

Erityisen kielteisesti suhtauduttiin ajatukseen siitä, että tekoälyä käytettäisiin sellaisissa prosesseissa, joista ei ole mahdollista valittaa. Siinä missä vain 7 % mukaan tämä lisäisi myönteistä suhtautumista, 77 % sanoi suhtautumisensa muuttuvan negatiiviseksi, jos tekoälyä hyödynnettäisiin näissä prosesseissa. Viranomaisen toiminnan lakiperusteisuuden ja -sidonnaisuuden tärkeydestä ja läpinäkyvyyden vaatimuksesta viestii myös se, että 64 % vastaajista sanoi suhtautumisen muuttuvan negatiiviseksi, jos tekoälyä käytettäisiin tilanteissa, joissa viranomainen hyödyntää päätöksenteon perusteena tietoja joiden käyttöön vastaaja ei ole erikseen antanut lupaa. Tulokset saattavat jossain määrin heijastaa 2018 aikana käytyä keskustelua tietosuojan, tietoturvan ja henkilötietojen lainmukaisen käytön merkityksestä.

## Muita huomioita

Verrattaessa ristiin vastaajien tärkeiksi nostamia kehityskohteita ja hyviä syitä lisätä tekoälyn käyttöä, suurin korrelaatio löytyy sellaisista teemoista jotka koskevat kansalaisen oman asioinnin helpottamista. Etenkin kansalaisia koskevien viranomaispäätösten nopeutuminen ja jonotusajojen lyhentäminen on teema, jonka kansalaiset kokevat sekä tärkeäksi kehityskohteeksi että hyväksyttäväksi syyksi lisätä tekoälyn käyttöä. Tekoälyn sovelluskohteet on mahdollisesti aluksi paikannettava sellaisiin rutiininomaisiin sovelluskohteisiin, jotka ovat kansalaisille jo tuttuja, ja joista nämä kokevat hyötyvänsä.

Kansalaisten negatiivista suhtautumista tekoölyyn selittää muun muassa tekoälytuntemuksen vähäisyys ja omien valmiuksien arvioiminen heikoiksi suhteessa julkista valtaa käyttävään viranomaiseen. Suhtautuminen tekoälyn käyttöön muuttuu sitä kielteisemmäksi, mitä vieraampi sovelluskohde on kyseessä. Kielteinen suhtautuminen tekoälysovelluksiin kasvoi sitä korkeammaksi, mitä autonomisempi asema tekoälysovellukselle annettiin. Kun tekoälyn käyttöä lisätään, on olennaista huolehtia siitä, että käyttötapauksista viestitään avoimesti ja sovelluksia käytetään mahdollisimman avoimesti.

Kolme keskeistä teemaa, jotka nousivat esille hyväksyntää lisäävinä tekijöinä ovat selkeän sääntelyn laatiminen, päätöksenteon tekeminen mahdollisimman avoimeksi ja läpinäkyväksi, sekä riittävän ihmiskontrollin varmistaminen. Läpinäkyvien viranomaisprosessien tarve koskee sekä tekoälyjärjestelmän toimintaa, että järjestelmän prosessoimaa dataa. Legitiimiteetin ja luottamuksen näkökulmasta on tärkeää, että tekoälyn käytön lainsäädännöllistä puolta kehitetään, ja huolehditaan siitä, että viranomaisella on käytössään selkeä ohjeistus jonka noudattamiseen kansalaiset voivat luottaa. Järjestelmien kehittämisen rinnalla tulee siis huomioida myös olemassa olevan lainsäädännön päivittäminen. Lainsäädäntöä tulee päivittää siten, että tekoälyä käyttävän viranomaisen toimintaa ohjaa sellainen normisto, jossa huomioidaan riittävä oikeusturvan ja oikeussuojakeinojen toteutuminen.

## 4.2. Kansalaiskeskustelu internetissä

Tekoälyn yhteiskunnallisen hyväksyttävyyden kyselytutkimusta täydennettiin internetissä tapahtuvan kansalaiskeskustelun analyysillä. Suppean katsauksen tavoitteena oli tunnistaa ihmisten suhtautumistapoja, asenteita, odotuksia ja epäilyksiä tekoälyn suhteen. Aineistona läpikäytiin kotimaisten sanomalehtien verkkoartikkelien kommentoinnissa käytyä keskustelua sekä aikakauslehtien tai muiden verkkopalveluiden keskustelupalstoilla tapahtuvaa ajatustenvaihtoa. Näistä lähteistä etsittiin tekoölyyn ja erityisesti sen hyödyntämiseen viranomaispalveluissa liittyviä keskusteluja, ja niiden pohjalta tunnistettiin tyypillisiä puheenaiheita eettisyyden ja hyväksyttävyyden näkökulmia painottaen.

Aineistohaun ja mediaseurannan jälkeen tarkemman analyysin tietolähteinä läpikäytiin seuraavat keskustelut, perustuen niiden temaattiseen osuvuuteen sekä kommentoinnin ja keskustelun vilkkauteen (päivämäärät viittaavat keskustelujen aloitushetkeen, eli keskustelun avanneen artikkelin tai viestin julkaisupäivään):

- Helsingin Sanomat, MieliPide: Tekoäly voi pahimmillaan suositella potilaalle täysin tarpeettomia hoitoja (28.10.2018)
- Iltta-Sanomat: Espoo pestasi tekoälyn sote-työntekijäksi – millaisin tuloksin? (25.4.2018)
- Iltta-Sanomat: Tutkija kertoo, mikä tekoälyssä huolestuttaa – ”Hysteria tappajaroboteista on hölynpölyä” (27.09.2018)
- Iltta-Sanomat: ETK: Tekoäly ennustaa työkyvyttömyyden jopa kaksi vuotta ennakoon – tarkkuus 78 prosenttia (17.4.2018)
- Iltta-Sanomat: Näin tekoäly seuloo työnhakijoita – ”Robotit etsivät puutteitasi ja virheitäsi” (15.4.2018)
- Tiede, Keskustelu: Tekoäly tunnistamassa huostaanottoa ennustavia tekijöitä (25.7.2018)
- Tiede, Keskustelu: Viekö tekoäly työpaikat? (20.6.2018)
- Vauva, Keskustelu: Tekoäly oppi jo tappamaan, vaikka sitä ei oltu siihen ohjelmoitu. (7.11.2017)
- Vauva, Keskustelu: Miksi tekoäly ei korvaa lakimiehiä? Maailman yksinkertaisin asia! (20.12.2017)
- Suomi24.fi, Keskustelu: Tekoälyä lisää vanhustenhoitoon (23.10.2017)
- Suomi24.fi, Keskustelu: Työmarkkinatori tulee MOL:in tilalle? Chatbot jne. (11.10.2017).

Taulukko 4 vetää yhteen läpikäydyistä keskusteluista tunnistettuja puheenaiheita. Löydökset on jaoteltu SWOT-analyysin keinoin (strengths, weaknesses, opportunities and threats) tekoälyn ja ohjelmistorobotiikan vahvuuksiin, heikkouksiin, mahdollisuuksiin ja uhkiin. Lisäksi listataan joitakin sekalaisia puheenaiheita. On huomattavaa, että tietty keskustelunaihe tai teema saattoi jossakin keskustelussa nousta esiin positiivisena ja toisessa negatiivisena asiana (esimerkiksi koneen ja algoritmien kyky mahdollistaa tunteiden erkauttaminen päätöksenteosta).

*Taulukko 4. Kansalaiskeskustelusta tunnistettuja puheenaiheita: tekoälyn hyödyntämisen vahvuuksia, heikkouksia, mahdollisuuksia, uhkia ja muita huomioita.*

<p><b>Tekoälyn vahvuudet</b></p> <p>Luotan enemmän laajan data-aineiston analyysiin perustuvaan tekoölyyn kuin ihmisen rajalliseen osaamiseen.</p> <p>Tekoäly toimii juuri niin kuin se on ohjelmoitu, ja se on siten hyvä apuri.</p> <p>Varhainen puuttuminen (ennakoivat viranomaispalvelut) kuulostaa hyvältä ja resursseja tehokkaammin ohjaavalla ratkaisulla.</p> <p>Tekoälyn vahvuus mutta myös heikkous on kyky olla huomioimatta tunnepohjaisia asioita.</p> <p>Tekoäly käy läpi salamannopeasti suuretkin tietoa-aineistot ja voi verrata käsiteltävää tapausta aiempiin.</p> <p>Tietokone tekee jo nyt oikeampia diagnooseja ja ehdotuksia kuin ihminen (lääkäri) ja osaa löytää vertailuaineistosta myös harvinaiset tai historialliset tapaukset.</p> <p>Tekoäly voi helpottaa arkielämää monella tapaa ja avustaa myös päättäjiä.</p>	<p><b>Tekoälyn heikkoudet</b></p> <p>Tekoäly ei osaa tehdä päätelmiä, se osaa ainoastaan suorittaa algoritmeja ja tehdä tilastollisia tuloksia, jotka voivat hyvinkin mennä pieleen.</p> <p>Luottavatko virkamiehet kylmän pragmaattisesti robottien ja tekoälyn ennusteisiin?</p> <p>Tekoälyä hehketään turhaan, kyse on toistaiseksi vain ihmisen tekemistä yksinkertaisista laskukaavoista, eikä sen avulla nykyään voida ennustaa edes säätä.</p> <p>Ennustavat viranomaispalvelut kuulostavat kyllä hyvältä, mutta onko ennakoivan tiedon hyödyntämiseen resursseja ja osaamista (esim. sosiaalipalveluiden puuttumiset tai työkyvyttömyyden käsittely).</p> <p>Tekoälyn tuottama tieto on enemmän todennäköisyyksiä kuin totuuksia, ja tämä pitää osata ottaa huomioon.</p>
<p><b>Tekoälyn mahdollisuudet</b></p> <p>Robotti olisi parempi päätöksentekijä ja voisi korvata jopa eduskunnan tai koko demokratian.</p> <p>Tekoäly mahdollistaisi työajan lyhentämisen ja enemmän tilaa harrastuksille yms.</p> <p>Odotan innolla näiden sovellusten käyttöönottoa omassa ammatissani. Kunhan riskit on huomioitu, voidaan tekoälyn avulla parantaa elintasoja valtavasti.</p> <p>Odotakaahan kun kuulemme uutisen siitä, että tekoäly oppi rakastamaan.</p> <p>Miksi suotta oletamme tekoälyn olevan pahantahainen tai murhanhimoisen? Miksemme oleta sen haluavan kehittää asioita parempaan suuntaan?</p> <p>Tekoälyllä on paremmat mahdollisuudet toimia arvojen mukaisesti, moraalisesti ja oikeudenmukaisesti kuin ihmisillä (lobbaus, rahoitus, etujen ajaminen, lahjonta, yms.).</p> <p>Ei tekoälystä kehity rasistista tai syrjivää, jos huolehdimme, että lähtöaineisto on kunnossa ja opeamme sitä periaatteella ”tee kuten sanon, älä niin kuin teen”.</p> <p>Tekoällyhype on huvittava ilmiö; kehitystä ja sovelluksia on, mutta ne eivät vain näytä yleistyvän käytössä, varsinkaan julkisella puolella.</p> <p>Tekoällyssä voi edesauttaa ratkaisuja isoihinkin ongelmiin, kuten yltäkylläisyyden ja nälänhädän polarisoima maailma.</p>	<p><b>Tekoälyn uhat</b></p> <p>Tekoälyn käytön esteeksi voi nousta se, ettei siltä sallita lainkaan virheitä, vaikka ihmisen tekemiin virheisiin verrattuna olisivat marginaalisia.</p> <p>Mitä voikin tapahtua, kun kaikki tieto on verkoissa ja hakkerit iskevät? Miten käy yksityisyyden?</p> <p>En halua viranomaisen käyttävän ja ennustavan käytöstäni robottien tai tekoälyn avulla.</p> <p>Onko koneiden kanssa seurustelun motiivi ja myös lopputulos se, ettemme osaa enää kohdata ihmistä?</p> <p>Ennakoivat viranomaistoiminta (kuten huostaanotto-prosessit) on ongelmien epäilemistä ja voi edesauttaa ongelmien realisoitumista.</p> <p>Ongelma on se, että tekoäly voi kehittää uusia toimintamalleja, joita sille ei ole edes ohjelmoitu.</p> <p>Tekoälyä kehittävät älykkäät, mutta mekanistisesti ajattelevat ihmiset. Kehittäjillä pitäisi olla osaamista myös moraalista, etiikan ja sosiaalisuuden asioista; teknologiaa kehitetään ihmisille.</p> <p>Tekoäly ja chatbotit mahdollistavat työttömän työnhakijan kyykyttämisen ja sähköisen valvonnan, ja edellyttävät, että olet sähköisesti tavoitettavissa ja reagoimassa ympäri vuorokauden.</p> <p>Kone ja algoritmit ovat kylmiä ja tunteettomia, ja siksi koneeseulonta ja bigdatan käyttö on hirveää.</p> <p>Ahdistaa tunnetasolla ajatus, että tekoäly tietää minusta liikaa ja yhdistelee asioita tietämättäni.</p>



### Sekalaiset

Mikäli tekoäly toimii (paremmin kuin ihminen), ei ole edes väliä, miten se sen tekee.

Tekoälyn ja sen soveltamisen tulisi kuulua osana (lääkäriin) koulutukseen.

Tuskin maltan odottaa, että tekoäly korvaa lääkärin tai ainakin toimii hänen apunaan.

Robottiikkaa ja investointeja sotateollisuuteen kyllä riittää, mutta siviilipuolella sovelluksia ei näy.

Kuka päättää mitä tietoja koneelle syötetään ja miten?

On tekoälystä varmaan apua älykkäiden hommissa, mutta odotan sitä, että osaisivat selittää minullekin, jotta edes hieman ymmärtäisin.

Miksi piirre, jota tavallisessa ihmisessä pidettäisiin ennakkoluuloisuutena, onkin tekoälyn tapauksessa hieno ja hyödyllinen?

Miten jaamme robottien tuomat hyödykkeet ja hyödyt, jäävätkö työn tulokset robottien omistajille?

## 5. YHTEENVETO JA POHDINTA

### 5.1. Johtopäätöksiä käyttötapauksista

Tekoälyn ja ohjelmistorobotiikan soveltamista viranomaistoiminnassa selvitettiin aluksi haastattelututkimuksella. Käyttötapauksista valittiin viisi käyttötapaussovellusta ja -tilannetta (asiakaspalvelun chatbot, ennakoiva anomalian tunnistus, elinikäisen oppimisen ja osaamisen järjestelmä, tekoälyjärjestelmien hankinta ja tekoäly viranomaisen työkaverina), joita yleistäen käytettiin tarkemmassa eettisyys- ja hyväksyttävyydestarkastelussa. Näiden selvitysten tärkeimmät johtopäätökset esitetään tässä luvussa.

Läpikäydyissä valtionhallinnon organisaatioissa on selvitetty laajasti ohjelmistorobotiikan ja tekoälyn hyödyntämismahdollisuuksia viranomaistoiminnassa. Käyttötapauksia oli tunnistettu runsaasti, ja valtionhallinnossa voidaan tulkita olevan vahva tahtotila ja kiinnostus teknologian hyödyntämiseen. Tekoälyn ja ohjelmistorobotiikan hyödyt palvelun laadun, nopeuden ja saatavuuden paranemisena sekä kustannus- ja resurssisäästöinä oli myös laajasti tunnistettu.

Tyypillisiä käyttötapauksia olivat viranomaisten päätöksentekoprosessien tai niiden osien automatisointi ohjelmistorobotiikkaa hyödyntäen sekä chatbotit joko viranomaisen työssä tai asiakaspalvelussa. Esimerkkejä ihmistyöstä koneelle siirtyvistä tehtävistä ovat tietojen haku järjestelmistä ja rekistereistä, hakemustietojen käsittely ja analyysi, kuvantulkinta, laatutarkastukset ja neuvontapalvelut (chatbotit). Useimmissa käyttötapauksissa tekoäly- tai ohjelmistorobotiikka oli viranomaistyötä tukeva työkalu, apuväline tai jopa työkaveri, mutta joukossa oli myös asiakasrajapinnassa tai ns. omillaan toimivia sovelluksia.

Haastatellut organisaatiot olivat hyvin erilaisissa tilanteissa tunnistamiensa käyttötapauksen toteuttamisessa. Osa oli edennyt enintään suunnittelun ja varhaisen kokeilemisen vaiheeseen, kun taas osalla oli jo useita järjestelmiä rutiinikäytössä. Eräissä pitkälle edenneissä organisaatioissa (kuten Kansaneläkelaitos, Verohallinto ja Tilastokeskus) automaatio oli jo aiemminkin ollut merkittävässä roolissa, ja ohjelmistorobotiikan ja tekoälyn hyödyntäminen nähdään pitkälti tämän kehittämistyön jatkumona. Näillä tahoilla oma osaaminen oli myös vahvaa, ja sen edelleen kehittämistä ja osaamisen omavaraisuutta pidettiin tärkeänä, vaikka teknisiä ratkaisuja ja muuta apua hankittaisiinkin ulkopuolelta. Myös niissä organisaatioissa,



joissa tekoälyn ja ohjelmistorobotiikan hyödyntämisessä oltiin varhaisemmassa vaiheessa, mainittiin oman osaamisen kasvattaminen avainasiana tai jopa viitattiin osaamisen tai henkilöstöresurssien puutteisiin kehityksen yhtenä pullonkaulana.

Chatboteilla nähtiin olevan iso merkitys viranomaistyön avustajana asiakkaan neuvojana ja ohjaajana tai myös virkailijan työn apuna. Aluksi sen hoitamat tehtävät voivat olla yksinkertaisiin kysymyksiin vastaamisia ja asiakkaan ohjaamista oikeaan palveluun. Kehittyneempiä muotoja voisivat olla asiakkaan todellisen tarpeen selvittämistä tai avustamista hakuprosessissa.

Hankintojen merkitys tekoälyn ja ohjelmistorobotiikan käyttöönotossa oli keskeinen kysymys, joka nousi laajasti esille haastattelukierroksella (mm. Liikennevirasto ja Kyberturvallisuuskeskus). Tasapainottelu omien resurssien ja hankittavien resurssien välillä voi olla monimutkaista, ja eri organisaatioilla oli tässä erilaisia valintoja vaihdellen omavaraisuuden pyrkimyksistä hankkijaroolin painottamiseen. Hankintaosaaminen tulee joka tapauksessa olemaan tärkeä kehityskohde läpi valtionhallinnon, ja rajanveto oman ja hankittavan tekemisen välillä tulee oletettavasti olemaan erilainen eri organisaatioissa. Valinnat tämän suhteen tulisi kuitenkin tehdä tietoisesti ja aina tietyn organisaation viranomaistehtävien kannalta tarkoituksenmukaisesti. Huomioitavia osaamisresurssityyppejä ovat esim. viranomaistyön prosessituntemus, prosessien tekninen määrittely ja kuvaaminen, hankintaosaaminen, ohjelmointi- ja ohjelmisto-osaaminen, järjestelmien ylläpito ja päivitys sekä käyttötuki. Ulkopuolisiin hankintoihin liittyviä riskejä ovat mm. toimittajariippuvaisuus, toimitettujen järjestelmien läpinäkymättömyys ja mahdolliset päivittämisshaasteet sekä viranomaistyön riittämätön ymmärrys sovellusten toteutusratkaisuisissa. Hankintaa tarkasteltiin työpajassa yhtenä käyttötapauksena; tulokset siitä on esitetty luvussa 3.3.

Haastatteluissa korostuivat myös tekoälyn ja ohjelmistorobotiikan henkilöstövaikutukset ja tähän liittyvät epävarmuudet. Käyttötapauksilla arvioitiin monesti olevan isojakin muutoksia viranomaistyön tehtäviin, viranomaisen rooliin, ammatti-identiteettiin ja töiden määriin. Toisaalta teknologiaratkaisut tukevat ihmistyötä, mutta niiden vaikutukset ihmistyön korvaajana herättävät helposti uhkakuvia. Sisäinen viestintä ja avoimuus sekä positiivisista että negatiivisista vaikutuksista nähtiin siten tärkeänä muutosten eteenpäinviemiseksi hyväksyttävällä ja innostavalla tavalla. Pääasiallisesti kokonaisvaikutukset viranomaistyön kannalta nähtiin työnkuvan kannalta positiivisena, kun kone tukee asiantuntijatyötä sekä hoitaa rutiinitehtävät ja ihmisresurssit vapautuvat mielekkäiden ja haastavien osatehtävien pariin. Osin nähtiin myös ihmistyötarpeen vähenemistä, mutta toisaalta monilla alueilla viranomaispalvelujen kysyntä tai laajuus on kasvussa, eikä teknologiaratkaisujen käyttöönotosta huolimatta ennakoitu henkilöstövähennysten tarpeita. Työpajassa käsiteltiin skenaariota Tekoäly työkaverina; sen tulokset on esitetty luvussa 3.3.

Tekoälyn ja ohjelmistorobotiikan käyttötapauksissa oli joitakin eri organisaatioiden yhteisiä hankkeita, ja yhteistyökuvioita oli myös esimerkiksi tiedonhakuprosesseissa eri organisaatioiden järjestelmien välillä. Puheen, kielen ja murteiden tunnistus ja niihin perustuvien chatbottien kehittäminen olisi järkevää tehdä yhteistyönä. Älykkäiden järjestelmien käyttöön liittyvä vastuiden ja korvausmenettelyiden määrittely on niin ikään yhteinen haaste. Kun data on kaiken perusta, sen laatu, luotettavuus ja jäljitettävyyys ovat oleellisia ja yhteistyö sen takaamiseksi on oleellista.

Selvityksen kuluessa tunnistettiin kaksi tekoälyn soveltamisen aluetta, jotka kiinnostivat laajasti: ennakoivan anomalian tunnistaminen ja elinikäiseen oppimiseen liittyvät sovellukset. Niihin molempiin todettiin liittyvän eettisiä ja vastuukysymyksiä, jotka tulee ratkaista ennen

tekoälyn käyttöönottoa. Näitä kysymyksiä pohdittiin työpajassa, jonka tulokset on esitetty luvussa 3.3.

Laajempikin poikkihallinnollinen yhteistyö voisi olla hyödyksi kehitystyössä sekä yhteistyöhankkeiden kannustamiseksi että kokemusten vaihtamiseksi. Esimerkiksi hankintaosaamiseen liittyvä tiedonvaihto mainittiin erityisen hyödyllisenä yhteisponnistuksena. Myös kansainvälinen yhteistyö tai kansainvälisten esimerkkien vertailu mainittiin joissakin haastattelussa, mutta kansainvälistä tietotaitoa ja yhteistyötä voitaisiin hyödyntää laajemminkin.

## 5.2. Johtopäätöksiä kansalaisnäkökulmasta

Kansalaisten suhtautumista tekoälyyn selvitettiin laajan kyselytutkimuksen avulla sekä analysoimalla internetin kansalaiskeskustelua. Näiden selvitysten tulokset tärkeimmät johtopäätökset esitetään tässä luvussa.

Tietosuojaulottuvuus ja henkilötietojen suoja ovat keskeisiä seikkoja esimerkiksi profiloinnin ja automaattisen päätöksenteon yhteydessä, mutta myös kerätessä henkilöistä tietoa viranomaisissa asioitaessa tai muutoin osana julkisen vallan käyttöä. Kyselyn tuloksissa on selkeästi nähtävissä kansalaisten kasvanut tietämys henkilötietojen käyttöön liittyvissä kysymyksissä. Yhdeksi vastausten keskeisimmistä teemoista nousi nimenomaisesti huoli omien tietojen käyttöön liittyvistä ongelmista.

Kuten edellä on todettu, etiikan, kansalaisten tunteman luottamuksen ja lainsäädännön väliset riippuvuussuhteet ovat kaikkea muuta kuin selkeät. Toisin sanoen, pelkällä viranomais toiminnan lakisidonnaisuudella ei voida ylläpitää legitimiä järjestelmää, mikäli kansalaiset kokevat järjestelmän epäoikeudenmukaisena. Toisaalta on mahdollista löytää yhtymäkohtia viranomaisen toiminnan lakisidonnaisuuden, viranomaistoiminnan legitimitietin ja yksilöiden viranomaista kohtaan tunteman luottamuksen välillä. Kyselytutkimuksen tulokset viittaavat paikoin siihen, että selkeällä lainsäädännöllä voidaan tukea kansalaisten tuntemaa luottamusta. Tehty kyselytutkimus selvitti ensi sijassa kansalaisten oletuksia ja tuntemuksia automaatiota ja tekoälyä kohtaan.

Keskeisiksi normatiivisiksi haasteiksi on sekä tässä selvityksessä että muissa meneillään olevissa selvityksissä tunnistettu toimijuuteen, vastuuseen ja tietoon liittyvät kysymykset. Alla esitetty lista on esimerkinomainen, eivätkä teemat suinkaan tyhjenny näihin kysymyksiin. Ne kuvaavat osaa niistä kysymyksistä, joihin vastaamista normatiivisen legitimitietin säilyttäminen saattaa edellyttää.

Toimijuus ja vastuu:

- Mihin tai kehen sekä perustuslaillinen että rikosoikeudellinen virkavastuu kohdistuu?
- Kuinka turvata yksilöille riittävät oikeusturvatakeet ja riittävä tiedollinen perusta oikeusturvatakeiden käyttämiseksi?
- Perustuuko julkisen vallan käyttö tässä yksittäisessä tapauksessa lakiin?
- Millä tavalla päätösten perusteluvollisuus voidaan toteuttaa tapauksissa, joissa ei tarkalleen tiedetä, miten lopputulokseen on päädytty?
- Miten ratkaistaan kuka automatisoidun päätöksen on tehnyt?

Tieto:

- Onko viranomaisella oikeus käsitellä henkilötietoja yleisesti ja oikeus käyttää henkilötietoja tässä yksittäisessä tapauksessa?
- Voiko tiedon käyttö loukata yksilön muita perusoikeuksia kuin henkilötietojen suoja, esimerkiksi yhdenvertaisuutta?
- Kuinka turvataan kansalaisen tiedollinen asema suhteessa julkista valtaa käyttävään viranomaiseen?
- Millä edellytyksillä eri tietolähteitä on mahdollista yhdistellä, ja millä edellytyksillä tietoa voidaan käyttää sen alkuperäisestä käyttötarkoituksesta eroaviin tarkoituksiin?

**Laajan kyselytutkimuksen** päähavainnot voidaan tiivistää seuraavasti:

- 1) Suomalaiset suhtautuvat maltillisesti ajatukseen tekoälyn hyödyntämisestä julkisissa palveluissa

Kysyttäessä suhtautumisesta tekoälyn käyttöön viranomaistyössä myönteiset ja kielteiset kannat normaalisti jakavat vastaajia. Vastakkainasettelu ei kuitenkaan ole jyrkkää. Vahvasti puolesta tai vahvasti vastaan olevat ovat vähemmistössä, ja isoin osuus vastaajista valitsee lievemmän kannan tekoälyn käytön lisäämisen puolesta tai vastaan. Yli kymmenesosa ei ole muodostanut kantaansa lainkaan.

- 2) Tekoälyn käyttö viranomaistyössä ei ole suomalaisille erityisen tuttu tai tunnettu asia

Lähes kaksi kolmasosa kokee, että ei tiedä riittävästi mahdollisista tekoälyn hyödyntämisen tavoista julkisessa hallinnossa tai viranomaistoiminnassa.

- 3) Tekoälyn käyttöönotossa huolestuttaa selkeästi eniten inhimillisen harkinnanvaran poistuminen

Etenkin avoimissa vastauksissa ilmenee huoli siitä, voiko ohjelmisto koskaan aidosti ymmärtää inhimillisen elämän monimutkaisuutta, ja siten tuottaa oikeina ja oikeudenmukaisina pidettäviä päätöksiä. Myös henkilötietojen väärinkäyttö tai niiden luvaton käyttö huolestuttaa.

- 4) Päätöksenteon ja tekoälyn käytön läpinäkyvyyden takaaminen, ja selkeän säännösten luominen lisäävät tekoälyn käytön hyväksyttävyyttä

Viranomaisen tulee ylipäätensä olla avoin tekoälyn käytöstä. Kansalaiset haluavat saada tietää, millä perusteilla tekoälyjärjestelmä on tehnyt yksittäisen ratkaisuun. Julkisuuden ja avoimuuden takaamiseksi tekoälyn käytölle halutaan julkisissa palveluissa selkeät pelisäännöt, selkeä sääntely ja valvonta.

- 5) Tekoälyn käyttöönotto on kansalaisten kannalta hyväksyttävintä aloittaa yksinkertaisista, rutiiniluontoisista käyttökohteista

Kokemuksia ja käytäntöjä näistä kokeiluista tulee kommunikoida avoimesti, jotta kansalaisten tietotaso tekoälyn käytöstä lisääntyisi ja tulisi asiana tutummaksi.

**Internetin kansalaiskeskustelun** analyysin päälöydöksiä voidaan nostaa esiin positiiviset odotukset oikeudenmukaisesti ja luotettavan oikeellisesti toimivasta tekoälystä ja ohjelmistorobotiikasta, joka ottaa suuret datamassat hyötykäyttöön ja mahdollistaa myös ennakoivan viranomaistoiminnan kautta toteutuvat hyödyt ja kustannussäästöt. Heikkouksina tunnistettiin tilastomaiseen, tunnepuolen huomioimattoman koneen katvealueet sekä epäily siitä, osataanko tekoäly ohjelmoida ja opettaa oikein ja osataanko sen tuotoksia hyödyntää. Keskusteluissa mahdollisuuksina tuotiin esille tekoälyn potentiaalia parantaa elämää sekä arjessa että suurissa yhteiskunnallisissa haasteissa. Sovellusten kehittämiseen liittyvistä eettisistä ja hyväksyttävyyden haasteista huolimatta tekoäly ja ohjelmistorobotiikka nähtiin myös eettisen, moraalisen ja oikeudenmukaisen yhteiskunnan työkaluna ja mahdollistajana. Suurimpia uhkia nähtiin järjestelmien ja tietojen väärinkäyttösmahdollisuuksissa ja yksityisyyden menettämisessä. Erityisesti ennakoiviin ja seurannan tekoälysovelluksiin suhtauduttiin varauksella.

### 5.3. Eettisyyden ja hyväksyttävyyden kysymyksiä

Tässä luvussa esitetään kootusti selvityksen perusteella tunnistettuja yksittäisiä eettisyyden ja hyväksyttävyyden kysymyksiä ja pohditaan niihin ratkaisujen ensiaskeleita. Tässä esitetyt kysymykset ovat esimerkkejä toistuvista, useille käyttötapauksille yhteisistä eettisyyden ja hyväksyttävyyden haasteista, mutta ne eivät korvaa tapauskohtaista tarkastelua, johon ehdotamme luvussa 5.4 esitettyä toimintamallia.

**Miten huomioidaan algoritmien katvealueet?** Huolena esitetään se, miten inhimillinen näkökulma voidaan ottaa päätöksenteossa huomioon, kun käsittelijänä on ihmisen sijaan kone. Henkilökohtaiset tilanteet, harkinnanvaraisuus, asiaan vaikuttavat sivuseikat (näiden lukeminen ns. rivien välistä), yms. jäävät tekoälyjärjestelmään perustuvan päätöksenteon ulkopuolelle.

- Kansalaisella on oltava mahdollisuus päästä asioimaan myös ihmisvirkamiehen kanssa.
- Kansalaisen oikeudet tekoälypalvelujen käytössä on kommunikoidava selvästi.

**Rapautuuko luottamus virkamiesosaamiseen?** Sekä valtionhallinnon sisällä että yksilön tai yhteiskunnan kannalta on merkittävää säilyttää luottamus virkamiesten osaamiseen siitäkin huolimatta, että prosesseja pyöritetään älykkäillä järjestelmillä. Vikatilanteiden ratkaiseminen, erikoistapausten käsittely, valitukset, yms. vaativat myös tulevaisuudessa ihmisosaamista.

- Toiminnan tulee perustua selkeään lainsäädäntöön.
- Järjestelmien valvonta ja esimerkiksi virheiden tunnistaminen tulee varmistaa virkamiesten osaamisen kautta.
- Tekoälypalvelun käytön perustelut erityisesti harkinnanvaraisissa tilanteissa on kommunikoidava: läpinäkyvyys ja asian ymmärtäminen lisäävät luottamusta.
- Hyvien kokemusten - ns. voitonnarratiivien - kerryttäminen käynnistyy aloittamalla yksinkertaisemmista sovelluksista, joissa tekoäly helpoimmin tuottaa positiivisia tuloksia. Selkeä kommunikointia ja vastuunkanto epäonnistuttaessa auttaa säilyttämään luottamuksen.

- Kansalaisilla tulisi olla saatavilla tieto siitä, missä omat tiedot ovat ja miten ja mihin niitä hyödynnetään.
- Järjestelmien ja sovellusten helppokäyttöisyydellä ja positiivisten käyttäjäkokemusten kerryttämisen avulla lisätään käytännön tasolla luottamusta.

**Onko julkishallinnossa riittävää hankinta-, ylläpito- ja kehittämisosaamista?** Tekoälyjärjestelmien hankintaan liittyvät resurssi- ja laatuksymykset ovat keskeisiä ratkaisujen toteuttamisen, käyttöönoton ja ylläpidon onnistumisessa. Puutteellisen osaamisen riskejä ovat etiikan ja hyväksyttävyyden näkökulmista vajavaiset järjestelmät, yksityisyyden vaarantuminen, toimittajaloukku ja järjestelmän ylläpidon ja uudistamisen organisoimisen haasteet.

- Roolijaot tulee määritellä selkeästi (ml. yhteistyö ja vuorovaikutus) ja niihin liittyvät tekniset ja muut osaamistarpeet: tilaaja, omistaja, käyttäjä, työkalut, vastuut, jne.
- Hankinnat tulee avata (tavoitteet, vaiheet, toimijat, muutokset, hyödyt, kustannukset) arkikielelle ymmärryksen lisäämiseksi ja keskustelun mahdollistamiseksi; myös kansalaiset käyttäjinä voisivat osallistua tarvemäärittelyyn.
- Resurssien yhdistäminen esimerkiksi hallinnonalojen tai tieteenalojen kesken on tärkeää osaajien ja osaamisen hankkimiseksi.
- Valtionhallinto voisi täsmentää yleistä ohjeistusta tekoälyhankintoja varten.
- Viestintään on kiinnitettävä erityistä huomiota.

**Kuka vahtii tekoälyn toimintaa?** Virheiden tunnistaminen tekoälyn toiminnassa vaatii jatkuvaa seurantaa ja mahdollisesti sekä ihmis- että konetyötä. Järjestelmien opetusaineistot ja algoritmit ovat monimutkaisia, ja näissä esiintyvät virheet voivat jäädä pitkäksi aikaa pimentoon. Toisaalta myös selvien virhetoimintojen korjaaminen ja väärin päätösten oikaisu voi olla haasteellista ja käydä myös kalliiksi. Myös esimerkiksi vastuun määrittelyt viranomistoiminnassa ovat nykyisellään epäselviä tai toimimattomia tekoälyjärjestelmiä sovellettaessa.

- Algoritmien ja datan avaaminen kansalaisille voisi olla yksi ratkaisu, mutta se toisaalta voisi aiheuttaa väärinkäytön riskin. Tasapainottavia ratkaisuperiaatteita tarvitaan.
- Tekoälyjärjestelmän oppimiseen ja kehittämiseen on luotava selvät, avoimesti kommunikoidut periaatteet ja prosessit.
- Vastuukysymysten määrittelyt ihmisen ja järjestelmien suhteen tulee tarkistaa tekoälykontekstiin sopiviksi.

**Miten käydään rajanvetoa tekoälyn käytämisestä ennakoivassa viranomaistyössä?**

Älykkäät järjestelmät mahdollistavat tiedon analyysien avulla tapahtuvan ennakkoinnin, jota voidaan hyödyntää viranomaistyössä monella tapaa, esimerkkeinä varhainen puuttuminen riskikäyttäytymiseen tai tukien myöntäminen ilman hakemusta. Avoin kysymys on, missä määrin tällaisia toimintoja voidaan ylipäänsä sisällyttää tulevaisuuden viranomaistyöhön.

- Ennakoivan toiminnan periaatteena tulisi olla aina kansalaisen ja yhteiskunnan etu. Tulisi keskustella, määritellä ja kommunikoida missä tilanteissa, sovelluskohdissa ja järjestelmissä ennakoiva työ on hyväksyttävää.

**Hämärtääkö tekoäly viranomaisprosessien ja -päätösten ymmärrettävyyttä?** Riskinä pidetään sitä, että tekoälyjärjestelmien käyttöönoton myötä viranomaistyön perusteet ja prosessit jäävät pimentoon. Poikkihallinnolliset järjestelmät ja datamassojen ristiin käyttö haastavat tilanteita edelleen. Kansalaisen mahdollisuus ja kyky nähdä, mihin omat tiedot menevät, miten niitä voi hallita ja voiko niitä poistaa, hämärtyy. Läpinäkymättömyys ja ymmärrettävyyden puutteet haittaavat sekä kansalaista että virkamiestä.

- Jotta viranomaistyö säilyy ymmärrettävänä, tekoälyratkaisun tulee olla hyvin toteutettu, avoimesti periaatteiltaan dokumentoitu ja ymmärrettävästi viestitty.
- Osaoptimoinnin sijaan viranomaispalvelun kokonaisprosessi on arvioitava uudelleen, kun tekoäly tai ohjelmistorobotti on osa sitä.
- Tekoälyjärjestelmän ja siihen liittyvään viranomaisprosessin kannalta kaikilla osapuolilla tulee olla selvyys siitä, kuka päättää ja miten, mikä on hyväksyttävää. Mitä enemmän harkinnanvaraisuutta koneen tehtäväksi annetaan, sitä paremmin sen käyttö on perusteltava ja viestittävä.

**Ovatko tekoälyyn liittyvät väärinkäytösten ja turvallisuusriskit hallittavissa?** Tekoälyjärjestelmiin liitetään uhkakuvia tietomurroista, ja toisaalta pelätään myös käyttäjien voivan manipuloida järjestelmän toimintaa omaksi edukseen.

- Laajamittaiset riskit vaativat varautumista, mutta myös varautumisesta tiedottamista.
- Toteutuneisiin uhkiin reagoimisen tulee olla luottamuksen säilyttämisen vuoksi mahdollisimman ajantasaista ja avointa.

**Mitä kaikkea kansalaisen ylipäänsä tulee tietää tekoälystä?** Yleinen tietoisuus tekoälyn mahdollisuuksista ja uhkista on vielä puutteellista eikä välttämättä pohjaudu faktatietoon. Kaikkien erityisryhmien pääsy palveluun on taattava. Riittävän tietotason määrittely, ja tätä tukeva informaatio-ohjaus ovat siten tärkeitä, ja keskeinen rooli on myös kansalaisen käsityksellä omista digitaidoistaan. Ylipäänsä rajanvetoa tarvitaan myös siinä, tuleeko kansalaisen tietää, liittyykö viranomaispalvelun toteuttamiseen mahdollisesti tekoälyn hyödyntämistä.

- Kaikilla tulisi olla yhtäläinen pääsy ja mahdollisuus käyttää sovellusta taustasta, toimintakyvystä, digitaidoista tai kielitaidosta riippumatta.
- Käytön ja käyttöliittymän tarkoituksenmukaisuuteen, helppouteen ja selkeyteen tulisi kiinnittää huomiota järjestelmää suunniteltaessa, niin että tekoäly ei monimutkaista käyttäjäkokemusta tai jopa muodostu esteeksi. Päinvastoin tekoälyjärjestelmän avulla pyritään parantamaan käyttäjäkokemusta.

## 5.4. Eettinen toimintamalli

Seuraavassa on kuvattu toimintamalli eettisten kysymysten pohdintaan viranomaistoiminnassa, jota esitämme sovellettavaksi tekoälyn käyttöönoton suunnittelussa ja toteutuksessa viranomaistoimintaan. Eettisen toimintamallin perustana on yleinen eettinen lähtökohta: eettisyyttä tulisi tarkastella ihmisen hyvän elämän näkökulmasta. Yleisten eettisten periaatteiden mukaisesti viranomaistoiminnassakin tulisi siis nähdä asiakkaat ja työntekijät ensisijai-

sesti ihmisinä. Toiminnan ytimessä on oikeellisuuden vaatimus ja kansalaisten tuen ja turvallisuuden varmistaminen siten, että yksilön tarpeet pysyvät säädösten ja asioiden keskiössä. Eettisessä tarkastelussa voivat toimia kompassina edellä esitetyt yleiset eettiset periaatteet ja normit. Niiden tulisi toteutua palvelujen vastuullisen järjestämisen keskeisenä osana (Nieminen et al. 2018).

Vastuullisen toiminnan voidaan olettaa myös tuottavan kustannussäästöä, jolloin vastuullisuus kytkeytyy eettisten arvojen lisäksi moniin pragmaattisiin arvoihin ja hyveisiin, kuten tehokkuuteen. Universaalien arvojen lisäksi palvelujen suunnittelua ohjaavana tekijänä tulisi olla merkityksellisyyden ymmärtäminen: Millaista lisäarvoa teknologia tai palvelu tuo kansalaisten elämään? Välttömän käyttötilanteen lisäksi tulisi siis tarkastella myös palvelun vaikutusta ihmisten arkeen.

Eettisten tilanteiden tunnistaminen tai tietyn toiminnan arvioiminen eettisesti vaatii paneutumista tarkastelun kohteena olevaan tapaukseen. Useimmiten etiikkaan liittyvät todelliset ongelmat eivät ole valitsemista joko oikean tai väärän välillä, sillä moraalisesti väärin olevat toimintavaihtoehdot ovat usein selkeästi tunnistettavissa. Eettinen ongelma syntyy siitä, kun on tehtävä päätös kahden hyvältä ja oikealta tuntuvalta valinnan kesken. Tällöin eettiset kysymykset ovat monisyisiä ja -ulotteisia, ja niiden ratkaiseminen edellyttää laajaa erilaisten toimijoiden ja ryhmien osallistamista sekä, viranomaistoiminnan kyseessä ollessa, toiminnan avaamista laajemmin yhteiskunnalliseen tarkasteluun.

Eettisessä tarkastelussa voidaan edetä erilaisia ohjeita ja neuvoja apuna käyttäen. Eettisen arvioinnin voidaan ajatella käsittävän erilaisia askelia, joiden avulla voidaan arvioida suunniteltavan tuotteen tai sovelluksen eettisyyttä (ks. esim. Bynum 2004). Vastuullisuusajattelun mukaisesti askelten läpikäyntiin tulisi kutsua osallistujia mahdollisimman kattavasti sovelluksen tai palvelun käytön ja hyödyntämisen eri sidosryhmistä.

### **1. Ottakaa eettinen näkökulma**

Kun pohditaan, onko jokin toiminta eettisesti hyväksyttävää, ensimmäinen askel on eettinen näkökulma, joka kohtelee kaikkia ihmisiä tasa-arvoisina ja oikeudenmukaisesti, ja kunnioittaa jokaisen yksilön tarpeita ja oikeuksia.

### **2. Hankkikaa tarkka käsitys arvioitavasta tapauksesta.**

Ennen kuin eettistä arviointia voidaan tehdä, on tärkeää saada selkeä ja yksityiskohtainen kuvaus käsillä olevasta sovelluksesta, siihen liittyvistä toimijoista, heidän tehtävistään, rooleistaan ja yhteyksistään toisiin toimijoihin. Kohteena olevaa toimintaa on hyvä tarkastella annettujen faktojen valossa. Uusien, faktoihin perustumattomien kuvausten luomista on syytä välttää. Mikäli tietolähteet perustuvat artikkeleihin ja julkaisuihin, on hyvä varmistua siitä, että käytetyt avainsanat sekä monimerkityksiset ja epämääräiset termit ymmärretään varmasti oikein.

### **3. Yrittäkää nähdä sovellukseen liittyvät eettiset kysymykset ja mahdolliset perinteiset ratkaisut, jotka tapaukseen liittyvät**

Pyrkikää löytämään sovellukseen liittyvät eettiset avainkysymykset ja miettikää, toimivatko niiden käsittelyssä perinteiset eettiset ohjesäännöt. Ovatko projektin päämäärät tai tilaajan käyttöajatukset projektin lopputulokselle hyväksyttäviä? Suuri osa eettisistä päätöksistä voidaan tehdä tämän pohjalta. Jos sovellus tai koko tapaus on epätavallinen eikä sille löydy menettelyohjeita tai jos se vaatii syvempää eettistä tapauksen ymmärtämistä, on syytä jatkaa eettistä pohdintaa.



#### 4. Verratkaa tapausta muihin vastaavanlaisiin tapauksiin

Onko tapauksesta olemassa ennakkotapauksia tai vastaavanlaisia muita tapauksia, joihin kyseessä olevaa tapausta voidaan verrata? On hyvä miettiä myös vastakohtaisia tapauksia. Miten vastaavat tapaukset on aikaisemmin toteutettu? Mitä yhtäläisyyksiä tällä ja muilla vastaavilla tapauksilla on? Onko tämä tapaus niin erilainen, että edelliset aikaisemmin käytetyt ratkaisumallit eivät tähän päde? Mikä tekee tästä tapauksesta niin erilaisen? Ketkä vastustaisivat tätä ratkaisua ja miksi? Todennäköisesti vastustaja joko kokisi itse olevansa eettisesti uhattuna tai vastustaisi ratkaisua virkansa tai tehtäviensä puolesta. Pohtimalla tapausta kuvitellun vastustajan kautta on mahdollista päästä käsiksi ratkaisun kannalta oleellisiin asioihin. Joskus kannattaa myös vaihtaa eri rooleja ja miettiä, miltä eri toimijoista tuntuisi tässä ratkaisuvaihtoehdossa. Miten he toivoisivat tapauksen ratkaistavan?

#### 5. Hyödyntäkää erilaisia systemaattisia analyysitekniikoita

Edellä esitettyjen askelten avulla on mahdollista luoda käsitys tapaukseen liittyvistä erilaisista toimijoista, heidän rooleistaan, tehtävistään ja yhteistyöstään sekä tapaukseen liittyvistä keskeisistä eettisistä kysymyksistä. Tämän avulla voidaan jo hahmottaa, millaisia ratkaisuja tulisi tehdä ja millaisia ohjesääntöjä noudattaa eettisissä valinnoissa. Lisäapua voivat antaa:

- Eri ammattiryhmien eettiset koodistot  
Useat eri ammattijärjestöt ovat koonneet omat eettiset koodistonsa. Ratkaisumallia voidaan arvioida valitun eettisen koodiston jokaisen periaatteen valossa.
- Roolien ja vastuiden arviointi  
Jokaiseen rooliin liittyy vastuita ja useimpiin myös oikeuksia. On hyvä selvittää tapaukseen liittyvien roolien sisältämiä vastuita ja pohtia, onko roolit ja vastuut toteutettu ratkaisumallissa oikein ja arvostaen. Mikäli käytettävä teknologia luo uusia rooleja, mitkä olisivat näihin rooleihin sisältyvät oikeudet ja vastuut?
- Eri sidosryhmien arviointi  
Eettisiä kysymyksiä voidaan arvioida paremmin, kun jokainen tapaukseen liittyvä sidosryhmä käydään erikseen läpi ja pohditaan, mitkä ovat kuhunkin sidosryhmään liittyvät hyödyt, haitat ja oikeudet. Ovatko edut ja haitat jakautuneet tasapuolisesti? Onko ihmisten oikeudet nostettu esille ja niitä kunnioitettu vai onko niitä loukattu tai piiloteltu? Jos teknologia nostaa esiin täysin uusia etuja tai haittoja, miten ne tulisi jakaa tasapuolisesti ja miten asianosaisten oikeuksia tulisi oikein kunnioittaa?
- Järjestelmällinen toimintaohjeiden arviointi  
Jokaisessa sosiaalisessa kontekstissa on joukko erilaisia toimintaohjeita ja sääntöksiä, lakeja, periaatteita ja käytäntöjä. Niitä ovat kansainväliset sopimukset, kansalliset ja kunnalliset lait, valtiolliset ohjesäännöt, ammatilliset standardit ja menettelyohjeet, yrityksen toimintatavat sekä yhteisölliset ja yksilölliset arvot. Näitä voidaan hyödyntää käsillä olevassa tapauksessa miettimällä, mitkä olisivat soveliaita tapauksen kannalta (hyvän toimintatavan ohjeet vai eettiset koodistot) ja onko ohjeita noudatettu oikein. Mikä on eri ohjeiden painoarvo toisiinsa nähden? Jos teknologian käyttö tuo uusia toimintatapoja, millaisia ohjeistuksia se vaatii ja miksi?
- Eettisten teorioiden arviointi  
Moraaliteoriat voivat osaltaan auttaa tilanteen hahmottamisessa silloin kun haetaan eettistä ymmärrystä tietyille tapaukselle.



- Tekoälyyn liittyvien eettisten periaatteiden huomiointi (tähän käyttötapauksista nousevia juttuja omina kohtinaan (esimerkkeinä))

## **6. Tehkää asiaankuuluvat eettiset johtopäätökset tapauksesta**

Mieti, mitkä ovat tapauksen avainkysymykset. Kohdellaanko ketään epäeettisesti ja jos, niin miksi? Mitä toimenpiteitä tarvitaan, ja miten ne olisivat eettisesti oikeutettuja? Jos tapaukseen liittyy erilaisia keskenään kilpailevia arvoja, miten ne asetetaan tärkeysjärjestykseen ja miksi?

## **7. Miettikää tapausta tulevaisuuden kannalta**

Jos tapauksessa on jotakin epäeettistä, miten sellaisen voisi ehkäistä jatkossa?

Jos tämän kaltaiselle tapaukselle ei ole olemassa eettisiä toimintaohjeita tai -malleja, millaisia uusia toimintatapoja olisi mahdollista suositella ja miten ne oikeutettaisiin eettisesti?

# LÄHTEET

- Ailisto, H. (toim.) (2018). Tekoälyn kokonaiskuva ja osaamiskartoitus. Valtioneuvoston selvitys- ja tutkimustoiminta, Selvitys- ja tutkimustoiminnan julkaisusarja 46/2018, kesäkuu 2018.
- Boddington, P. (2017). Towards a Code of Ethics for Artificial Intelligence. Cham: Springer.
- Bryson, J and Winfield, A. (2018). Standardizing Ethical Design for Artificial Intelligence and Autonomous Systems. *Computing Edge* February 2018, Vol 4, Number 2, pp. 10-13. IEEE Computer Society.
- Bynum, T.W. & Rogerson, S. (toim.) (2004) Computer Ethics and Professional Responsibility. Malden: Blackwell Publishing.
- Bynum, T.W. (2018). A very short history of computer ethics. Southern Connecticut State University. [https://web.archive.org/web/20080418122849/http://www.southernct.edu/organizations/rccs/resources/research/introduction/bynum\\_shrt\\_hist.html](https://web.archive.org/web/20080418122849/http://www.southernct.edu/organizations/rccs/resources/research/introduction/bynum_shrt_hist.html) (Retrieved 15.8.2018).
- Cerqui D., Weber J., Weber Karsten, E. (eds.) Special Issue: Ethics in robotics. *International Review of Information Ethics*, Vol. 6 (12/2006).
- Die Bundesregierung (2018). Eckpunkte der Bundesregierung für ein Strategie Künstliche Intelligenz. Stand: 18. Juli 2018.
- Digital Single Market (2018). EU Member States sign up to cooperate on Artificial Intelligence. *Digibyte*, 10 April 2018.
- Dignum, V. (2017). Responsible autonomy. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17), pp. 4698-4704. <https://www.ijcai.org/proceedings/2017/0655.pdf>
- Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20:1, 1-3.
- EDPS Ethics Advisory Group. (2018). Towards a digital ethics. [https://edps.europa.eu/sites/edp/files/publication/18-01-25\\_eag\\_report\\_en.pdf](https://edps.europa.eu/sites/edp/files/publication/18-01-25_eag_report_en.pdf)
- European Commission (2018a). COM(2018) 237 final. Communication from the commission to the European parliament, the European council, the council, the European economic and social committee and the committee of the regions, Artificial Intelligence for Europe. Brussels, 25.4.2018.
- European Commission (2018b). Statement on Artificial intelligence, robotics and 'autonomous' systems. European Group on Ethics in Science and New Technologies. March 2018. Doi: 10.2777/531856
- European Group on Ethics in Science and New Technologies, (2018). Artificial Intelligence, Robotics and 'Autonomous' Systems. European Commission. [https://ec.europa.eu/research/ege/pdf/ege\\_ai\\_statement\\_2018.pdf](https://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf)
- Euroopan parlamentin päätöslauselma 16. helmikuuta 2017 suosituksista komissiolle robotiikkaa koskevista yksityisoikeudellisista säännöistä (2015/2103(INL)) <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//FI>
- The Expert Group on DATA ETHICS (2018). Data for the Benefit of the People. Recommendations from the Danish Expert Group on Data Ethics. November 2018. <https://cis-india.org/internet-governance/files/data-for-the-benefit-of-people>
- EWISA Videovalvonnan kehittäminen EU:n maarajojen valvonnassa hanke: <http://www.ewisa-project.eu/>
- Friedman, B. & Kahn, P.H. Jr. (2003) Human values, ethics, and design. Teoksessa: Jacko, J.A. & Sears, A. (toim.) *The Human-Computer Interaction Handbook. Fundamentals, evolving technologies and emerging applications*. Mahwah, NJ: Lawrence Erlbaum Associates Inc. pp. 1177–1201.

- Friedman, B., Kahn, P.H. Jr. & Borning, A. (2006) Value Sensitive Design and Information Systems. Teoksessa: Zhang, P. & Galletta, D. (toim.) Human-Computer Interaction in Management Information Systems: Applications. Volume 6/2006. New York: M.E.Sharpe Inc. pp. 348–372
- Goodman, J. (2018). The only way is ethics. The Law Society, Gazette 21 May 2018. <https://www.law-gazette.co.uk/features/the-only-way-is-ethics/5066162.article>
- Hallamaa, J. (2017). Yhdessä toimimisen etiikka. Helsinki: Gaudeamus.
- Heikkerö, T. (2012). Ethics in technology. Plymouth: Lexington Books.
- House of Lords (2018). AI in the UK: ready, willing and able? Select Committee on Artificial Intelligence, Report of Session 2017–19. HL Paper 100.
- Hoven, J. van den, & Weckert, J. (Eds.). (2008). Information technology and moral philosophy. Cambridge: Cambridge University Press.
- Kivisaari, S., Lovio, R., & Väyrynen, E. (2004). Managing experiments for transition: Examples of societal embedding in energy and health care sectors.
- Klemettinen ja Peltola (2018). Tekoälyaika, esitys Tietoyhteiskunta-Akatemiassa 15.02.2018.
- Koivisto, I. (2007). Hallinnon hyvyys kommunikatiivisina käytäntöinä? Oikeus 2007 (36), s. 164-182.
- Koskinen, S. (2016). Virkamiesoikeuden perusteet. 7. uudistettu painos. Helsinki: Talentum Pro.
- Kulla, H. (2006). Hallintolainkäytön legitimeetti. Oikeussosiologiaa ja kriminologiaa, Juhlajulkaisu Ahti Laitinen 1946-24/4-2006, Turun yliopisto, oikeustieteellinen tiedekunta 2006.
- Kääriäinen, J., (toim.), Aihkisalo, T., Halén, M., Holmström, H., Jurmu, P., Matinmikko, T., Seppälä, T., Tihinen, M. ja Tirronen, J. (2018). Ohjelmistorobotiikka ja tekoäly – soveltamisen askelmerkkejä. Valtioneuvoston selvitys- ja tutkimustoiminta, Selvitys- ja tutkimustoiminnan julkaisusarja 65/2018, loka-kuu 2018.
- Leikas, J., & Koivisto, R. (2015). Ethics Assessment in different fields: Social Gerontechnology. Report of SATORI - Stakeholders Acting Together on the Ethical Impact Assessment of Research and Innovation. June 2015. <http://satoriproject.eu/media/2.c.3-Social-gerontechnology.pdf>
- Leikas, J. (2008). Ikääntyvät, teknologia ja etiikka – Näkökulmia ihmisen ja teknologian vuorovaikutus-tutkimukseen ja –suunnitteluun. VTT Working Papers; 110, ISBN 978-951-38-7171-0. Espoo: VTT.
- Leikas, J., Koivisto, R., & Gotcheva, N. (2018). Ethics of autonomous systems. In Heikkilä, E. (Ed.) Effective autonomous systems. VTT Framework for developing effective autonomous systems. VTT White Paper, December 2018.
- Lin, P. (2017). Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence. Oxford: Oxford University Press.
- Lind, E. A. (1988). The social psychology of procedural justice. New York: Plenum Press.
- Lucivero, F. (2016). Ethical assessments of emerging technologies. Appraising the moral plausibility of technological visions. The International Library of Ethics, Law and Technology, Vol 15. Heidelberg: Springer.
- Nieminen, M., Rilla, N., Leikas, J., & Ikonen, V. (2018). Innovaatiotoiminnan yhteiskunnallinen vastuullisuus ja vaikuttavuus. Teoksessa P. Juuti & S. Viljanen (toim.) Yhteiskunnalliset yritykset eettisen innovatiivisuuden lähteinä. Edita & Arvo-liitto.
- Pichai, S. (2018). AI at Google: our principles. Published Jun 7, 2018 in: <https://www.blog.google/technology/ai/ai-principles/>
- Pöntinen, H. (2018). Tekoälyn etiikka. Etiikan rooli Britannian, Japanin, Ranskan, Suomen ja USA:n tekoälystrategioissa. Graduluonnos, Helsingin Yliopisto.

Rabesandratana, T. (2018). Europe moves to compete in global AI arms race - With €1.5 billion for artificial intelligence, European Commission pins hopes on ethics. *Science* 4 May 2018, Vol 360 Issue 6388, p. 475.

Saariluoma, P., Cañas, J.J., & Leikas, J. (2016). *Designing for Life - A human perspective on technology development*. London: Palgrave MacMillan.

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2*. IEEE, 2017. [http://standards.ieee.org/develop/indconn/ec/autonomous\\_systems.html](http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html). Noudettu 20.06.2018.

Tuori, K. (2002). Käsiteläinopin itsepuolustus. *Lakimies* 7-8/2002, s. 1295-1320.

Vakkuri, V. and Abrahamsson, P. "The Key Concepts of Ethics of Artificial Intelligence." 2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). IEEE, 2018.

Valtiovarainministeriö (2018). Esiselvitys kansallisesta tekoälyohjelma Aurorasta. Tiedote 25.9.2018.

Villani, C. (2018). For a meaningful artificial intelligence - towards a French and European strategy. Mission assigned by the Prime Minister Édouard Philippe, A parliamentary mission from 8th September 2017 to 8th March 2018. [https://www.aiforhumanity.fr/pdfs/MissionVillani\\_Report\\_ENG-VF.pdf](https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf)

von Schomberg, R. (2011). *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields. A Report from the European Commission Services: European Commission*.

Wynsberghe, A. (2013). *Designing Robots for Care: Care Centered Value-Sensitive Design. (Report)*. *Science and Engineering Ethics*, 19(2), p. 407.

## Liite 1: HAASTATTELUKYSYMYKSET

Arvoisa Vastaanottaja,

Teemme selvitystä valtioneuvoston kanslian toimesta aiheesta **Robottiikan ja tekoälyn sääntelyn ja hyväksikäytön etiikka sekä yhteiskunnallinen hyväksyttävyyden (5.2 D).**

*Hankkeen tavoitteena on selvittää tekoälyn ja ohjelmistorobotiikan käyttöön viranomaistoiminnassa liittyviä eettisiä ja yhteiskunnalliseen hyväksyttävyyteen liittyviä kysymyksiä. Osatavoitteet ovat:*

- 1. Robotiikan ja tekoälyn viranomaiskäyttöä kuvaavien edustavien käyttötapauksien selvittäminen*
- 2. Eettiset kysymykset ja juridiset reunaehdot yleensä*
- 3. Tunnistettujen käyttötapauksien eettisyyskysymysten selvittäminen*
- 4. Tunnistettujen käyttötapauksien yhteiskunnallisen hyväksyttävyyden selvittäminen*
- 5. Kansalaisten ja viranomaisten toimintamallien uudelleen organisointi uuden teknologian vaatimuksia vastaaviksi*
- 6. Ratkaisuvaihtoehtojen kehittäminen ongelmallisiin tilanteisiin.*

*Hankkeen tuloksena saadaan edustava otos käyttötapauksista sekä niiden eettinen ja yhteiskunnallisen hyväksyttävyyden analyysi ratkaisuvaihtoehtoihin.*

Selvityksen tekee VTT:n koordinoima konsortio, johon kuuluvat myös Jyväskylän Yliopisto ja Helsingin Yliopisto. Käyttötapauksien selvittämiseksi haastattelemme valtionhallinnon edustajia. Haastattelut voivat tapahtua joko sähköpostilla, tarvittaessa puhelin- tai skypekeskustelulla täydentäen, tai sovitussa tapaamisessa. Haastattelun voidaan nauhoittaa, jos ennen haastattelua yhteisesti niin sovitaan.

Nauhoitetut haastattelut ja haastattelutulokset käsitellään luottamuksellisesti: kaikki tutkimuksen aikana kerätyt tiedot ovat luottamuksellisia ja aineisto käsitellään niin, ettei yksittäisten henkilöiden tietoja voida tunnistaa tutkimustuloksista raportoitaessa. Aineisto luovutetaan ainoastaan tämän hankkeen tutkijoiden käyttöön. Tiedot kootaan henkilötiedot poistettuina hankkeen yhteiseen VTT:n hallinnoimaan työtilaan, josta ne tuhoetaan VTT:n tietosuojakäytäntöjen mukaan. Hankkeen tulokset julkaistaan VN TEAS -raporttisarjassa.

Haastatteluja varten on laadittu haastattelurunko, joka on esitetty seuraavalla sivulla. Haastattelun muodosta ja ajankohdasta sovitaan VTT:n, Jyväskylän Yliopiston tai Helsingin Yliopiston vastaavan yhteyttä ottaneen tutkijan kanssa. Haastatteluun osallistuminen on vapaaehtoista. Ennen haastattelua pyydämme vapaaehtoisen suostumuksen (free consent) haastatteluun osallistumisesta.

Lisätietoja hankkeesta, haastatteluista ja tietojen käytöstä antaa:

Raija Koivisto, Johtava tutkija, VTT  
[raija.koivisto@vtt.fi](mailto:raija.koivisto@vtt.fi), puh. 040 5313311.

Nimi ja asema:

Organisaatio:

Päivämäärä:

Allekirjoitus / haastattelusuostumus yllä esitettyjen ehtojen mukaisesti:

Haastattelukysymykset:

1. Mitä tekoälyn tai ohjelmistorobotiikan käyttötapauksia organisaatiossanne on käytössä tai suunnitellaan otettavaksi käyttöön?
2. Valittuun käyttötapaukseen liittyen:
  - Mihin viranomaistehtävään käytetään, miten tukee viranomaisen toimintaa
  - Mitä tekee (tekeekö päätöksiä tai päätöksentekoa, kerääkö tietoa, jne.)
  - Mikä tekninen ratkaisu on valittu
  - Käyttäjien roolit (viranomaiskäyttö, kansalainen, jne.)
  - Keihin käytöllä on vaikutusta (loppuasiakkaat)
  - Mitä vaikutukset ovat (palvelun laatu, nopeus, saatavuus, jne)
3. Käyttötapauksen hankinta, suunnittelu, kehittäminen ja päivittäminen?
4. Miten viranomaistyön läpinäkyvyys toteutuu tällä hetkellä yleisesti?
5. Käyttötapauksen läpinäkyvyys ja jäljitettävyyt? Viranomaisnäkökulma ja loppukäyttäjän näkökulma
  - Käytetyn datan suhteen
  - Käytettyjen algoritmien suhteen
  - Käyttäjän toiminnan suhteen (käyttölogit)
  - Viranomaisvalvonta
6. Käyttötapauksen eettisyyden ja hyväksyttävyyden arviointi:
  - Käytetyt menetelmät
  - Keiden näkökulmasta
7. Mitä viranomaistyössä on robotisaation myötä muuttumassa?
  - Viranomaisorganisaation kannalta
  - Kansalaisten kannalta
8. Välittömät ja pidemmän tähtäimen odotukset uuden tekniikan käyttöönoton myötä?
9. Minkälaisia haasteita on tullut tähän mennessä vastaan?
  - Tekniset
  - Eettiset
  - Juridiset
  - Käyttö
  - jne.
10. Viranomaistyön perusarvot ja robotisaatio?
  - Mitä arvot ovat
  - Mitä riskejä syntyy.

**VTT**



JYVÄSKYLÄN YLIOPISTO  
UNIVERSITY OF JYVÄSKYLÄ



VALTIONEUVOSTON  
SELVITYS- JA TUTKIMUSTOIMINTA

[tietokayttoon.fi](http://tietokayttoon.fi)

ISSN 2342-6799 (pdf)

ISBN 978-952-287-643-0 (pdf)

