

Juha Koivisto, Teemu Keski-Kuha,  
Jaakko Lähteenmäki, Heba Sourkatti,  
Juha Pajula, Emmi Antikainen, Petri Kettunen,  
Juha Vesala, Viivi Rankka ja Juhana Riekkinen

# Monialaisen palvelukäytön ennakointi tekoälyn avulla

Kehittämisen perusteita ja suuntaviivoja

OHJAUS 15/2023



Ohjaus 15/2023

**Monialaisen palvelukäytön ennakointi  
tekoälyn avulla**  
Kehittämisen perusteita ja suuntaviivoja

© Kirjoittajat ja Terveyden ja hyvinvoinnin laitos

Kannen kuva: Tiina Kuoppala, Bigstock

Taitto: Tiina Kuoppala, Graforma

ISBN 978-952-408-145-0 (verkkojulkaisu)

ISSN 2323-4172 (verkkojulkaisu)

<http://urn.fi/URN:ISBN:978-952-408-145-0>

# Sisällys

<b>Kirjoittajat</b> .....	<b>6</b>
<b>Saatteeksi</b> .....	<b>7</b>
<b>Keskeiset käsitteet</b> .....	<b>9</b>
 <b>1 Mitä on monialainen palvelujen käyttö ja sen ennakointi?</b> .....	<b>11</b>
1.1 Monialaisten palvelujen käyttö ja tarve .....	13
1.2 Monialaisten palvelujen käytön ennakointi .....	13
 <b>2 Tekoälypohjaisen ennakkoinnin käytäntöjen kehittäminen</b> .....	<b>15</b>
2.1 Kehittämisen organisointi .....	17
2.2 Monialaisten palvelujen käytön analyysin totuttaminen .....	19
2.3 Vältettävissä olevan palvelujen käytön määrittely .....	20
2.4 Ennakkoinnin kehittämistarpeiden ja -tavoitteiden määrittely .....	21
2.5 Ennakkoinnin käytäntöjen määrittely .....	23
2.6 Ennakkoinnin käytäntöjen toimeenpano .....	28
2.7 Palvelujen käytön jatkuva seuranta ja arviointi .....	28
 <b>3 Tekoälypohjaisten ennustemallien kehittäminen</b> .....	<b>30</b>
3.1 Tavoitteen määrittely .....	33
3.2 Tietoaineiston määrittely .....	34
3.3 Aineiston hankinta .....	37
3.4 Aineiston poiminta ja esikäsittely .....	39
3.5 Mallin toteutus ja testaus .....	41
3.6 Operatiivisen käytön huomioiminen .....	45
3.7 Käyttäjäkokeilu .....	48
 <b>4 Tekoälypohjaisten ennustemallien tekninen toteuttaminen</b> .....	<b>49</b>
4.1 Tietojärjestelmien valmiuksien ja kyvykkyyksien suunnittelu .....	52
4.1.1 Operatiivisten tietojärjestelmien ominaisuuksien varmistaminen .....	54
4.1.2 Tekoälyohjelmistojärjestelmien kehittäminen .....	55
4.2 Tekoälymallien sovittaminen tietojärjestelmiin .....	58
4.3 Tekoälyratkaisujen teknologinen käyttöönotto .....	61
4.4 Tekoälyteknologian organisatorisen osaamisen kehittäminen .....	66





<b>5 Tietosuojalainsäädännön ottaminen huomioon tekoälypohjaisten ennustemallien kehittämisessä ja hyödyntämisessä</b> .....	<b>69</b>
5.1 Lainsäädäntö ja periaatteet .....	71
5.2 Ennustemallien kehittäminen toisilain perusteella .....	72
5.2.1 Tieteellinen tutkimus .....	72
5.2.2 Tietojohdaminen .....	73
5.3 Ennustemallien hyödyntäminen .....	74
5.3.1 Tietojohdaminen .....	74
5.3.2 Suostumus ja sen edellytykset .....	75
5.4 Automaattinen päätöksenteko .....	77
5.5 Vaikutustenarvioinnin laatiminen .....	80
5.5.1 Milloin vaikutustenarviointi on tehtävä? .....	81
5.5.2 Mitä vaikutustenarvioinnin tulee sisältää? .....	81
<b>Lähteet</b> .....	<b>84</b>



# Kirjoittajat

**Juha Koivisto**, VTT, dosentti, johtava asiantuntija, Tiedolla johtaminen ja vertaiskehittäminen -yksikkö, Terveyden ja hyvinvoinnin laitos

**Teemu Keski-Kuha**, DI, kehittäispäällikkö, Tiedolla johtaminen ja vertaiskehittäminen -yksikkö, Terveyden ja hyvinvoinnin laitos

**Jaakko Lähteenmäki**, TkL, johtava tutkija, Teknologian tutkimuskeskus VTT Oy

**Heba Sourkatti**, DI, tutkija, Teknologian tutkimuskeskus VTT Oy

**Juha Pajula**, Tkt, Erikoistutkija, Teknologian tutkimuskeskus VTT Oy

**Emmi Antikainen**, DI, Tutkija, Teknologian tutkimuskeskus VTT Oy

**Petri Kettunen**, Tkt, dosentti, yliopistotutkija, Tietojenkäsittelytieteen osasto, Helsingin yliopisto

**Juha Vesala**, OTT, yliopistonlehtori, oikeustieteiden tiedekunta, Lapin yliopisto

**Viivi Rankka**, LL.M., tutkija, oikeustieteiden tiedekunta, Lapin yliopisto

**Juhana Riekkinen**, OTT, VT, yliopistonlehtori, oikeustieteiden tiedekunta, Lapin yliopisto



# Saatteeksi

Tässä käsikirjassa tarkastellaan ja jäsennetään monialaisen palvelukäytön ennustamisen ja ennakkoinnin kehittämistä ja toteuttamista sosiaali- ja terveydenhuollon organisaatioissa. Käsikirjan tarkastelu kohdentuu erityisesti moderniin tekoälyyn tukeutuvaan ennustamiseen ja ennakkointiin. Käsikirja on suunnattu ensisijaisesti hyvinvointialueilla tapahtuvan kehittämistyön tueksi.

Monialaisella palvelujen käytöllä tarkoitetaan tässä käsikirjassa palvelukäyttöä, jossa asiakas käyttää erilaisten palveluryhmien palveluja, esimerkiksi perusterveydenhuollon, sosiaalihuollon ja päihdehuollon palveluja. Monialaisen palvelukäytön ennustamisella tarkoitetaan niiden riskitekijöiden ja ennusmerkkien tunnistamista, jotka voivat johtaa monialaiseen palvelukäyttöön. Ennustamisen analyysin yksikköä on yksittäinen henkilö, joita koskevista analyyseistä voidaan aggregoida asiakasryhmätasoisia ennusteita. Ennakkoinnilla tarkoitetaan niitä käytäntöjä ja toimenpiteitä, joita toteutetaan ennustetiedon perusteella asiakastyössä yksittäisen asiakkaan kohdalla ja palvelutoiminnan johtamisessa asiakasryhmätasolla.

Käsikirjassa tarkastellaan lisäksi ennustamisen ja ennakkoinnin kehittämiseen ja toteuttamiseen sekä yksittäistä asiakasta ja asiakasryhmiä koskevan datan hyödyntämiseen liittyviä lainsäädännöllisiä kysymyksiä.

Käsikirjassa tarkastellaan ennustamista ja ennakkointia monialaisesti palveluja käyttävien asiakasryhmien kehityksessä. Käsikirjan sisältää esimerkkejä Päijät-Hämeen hyvinvointialueella toteutetusta ikääntyneiden asiakasryhmään rajatusta monialaisen palvelukäytön ennustamisen ja ennakkoinnin kehittämisestä. Päijät-Hämeen kanssa kehitetystä tekoälypohjaisesta ennustemallista saatiin tämän käsikirjan kirjoitusaikana (vuonna 2023) valmiiksi demonstraatioversio.

Tämän käsikirjan jäsennyksiä on helppo soveltaa myös muiden rajatumpien ja tarkemmin määriteltyjen asiakasryhmien, esimerkiksi maahanmuuttajien monialaisen palvelukäytön ennustamisen ja ennakkoinnin tarpeisiin. Käsikirjan perusajatukset ja -välineet ovat sovellettavissa myös muun tyyppisten tekoälyyn perustuvien palvelujen käytön ennusteiden ja ennakkointien tekemiseen, esimerkiksi uusintakäyntien ennustamiseen ja ennakkointiin.

Tekoälyyn tukeutuvan ennustamisen ja ennakkoinnin kehittäminen on toimintaa, joka edellyttää onnistuakseen organisaation useiden eri toimijaryhmien osallistumista kehittämistoimiin. Kehittämisen kohteena on tällöin moninainen sosiotekninen systeemi, joka koostuu muun muassa erilaisista sote-toimijoista, tehtävistä, laeista, säännöistä ja periaatteista, malleista, tietojärjestelmistä ja tietoteknisestä infrastruktuurista. Tällaisten elementtien tulisi toimia mahdollisimman saumattomasti yhteen. Tämän käsikirjan lukijakunta on näin ollen moninainen, ja sen eri luvut on suunnattu eri kohderyhmille erilaisten rajattujen kehittämistasojen ja -kohteiden perusteella, vaikka käytännössä kyse on yhtenäisestä kokonaisuudesta. Käsikirjan luvuissa on jonkin verran päällekkäisyyttä juuri siksi, koska kyse on saumattomasta kokonaisuudesta.



Luku 2 on suunnattu erityisesti asiakasryhmätasoisia ennakoiteja tekeville johtajille ja esihenkilöille sekä yksittäisten asiakkaiden ja näiden palvelujen suunnittelun kanssa työskenteleville sote-ammattilaisille. Luku 3 on suunnattu erityisesti ennustemallien kehittämisen parissa työskenteleville data-analyytikoille. Luku 4 on suunnattu organisaation tietotekniikasta ja tietojärjestelmistä vastaaville tietotekniikan ja tietojohdamisen ammattilaisille. Luku 5 on suunnattu asiakas- ja potilastietojen hyödyntämiseen liittyvien oikeudellisten kysymysten parissa työskenteleville ammattilaisille. (Kuvio 1.)



**Kuvio 1.** Monialaisen palvelukäytön ennakoinnin kehittämisen kokonaisuudet

Tämä käsikirja on syntynyt sosiaali- ja terveysministeriön rahoittaman Monialaisen palvelutarpeen ennakointi tekoälyn avulla – kansallinen kokeilu ikäihmisten asiakasryhmää koskien -hankkeen tuotoksena (MAITE-hanke). Rahoitus oli osa kansallista Kotona asumisen teknologiat ikäihmisille -ohjelmaa (KATI). Hankkeen tutkimus- ja kehittämiskonsortiota johti ja koordinoi Terveyden ja hyvinvoinnin laitos. Muut toteuttajat olivat Teknologian tutkimuskeskus VTT Oy, Helsingin yliopisto (tietojenkäsittelytieteen osasto), Lapin yliopisto (oikeustieteiden tiedekunta) sekä kokeilualueena Päijät-Hämeen hyvinvointialue. THL vastaa luvuista 1 ja 2, VTT Oy luvusta 3, Helsingin yliopisto luvusta 4 ja Lapin yliopisto luvusta 5.

Tämä käsikirja toimii samalla hankkeen loppuraporttina.



# Keskeiset käsitteet

**Anonyymit tiedot:** Anonyymit tiedot ovat tietoja, jotka ”eivät liity tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, tai [...] joiden tunnistettavuus on poistettu siten, ettei rekisteröidyn tunnistaminen ole tai ei ole enää mahdollista” (tietosuoja-asetuksen johdanto-osan 26 kappale). Tietosuoja-asetusta ei sovelleta anonymisointeihin tietoihin.

**Aggregoidut tilastotiedot:** Aggregoidulla tilastotiedolla tarkoitetaan toisilaisissa ”tilastomuotoista, luotettavasti anonymisoitua tietoa” (3 §:n 18 kohta). Anonymisoidun tiedon ja aggregoidun tilastotiedon ero on se, että jälkimmäisessä henkilöä ei voi tunnistaa edes yhdistämällä tietoa muuhun anonymisoituun tietoon – aggregoidusta tilastotiedosta on poistettu anonyymien tiedon ns. jäännösriski, eli riski henkilön tunnistamisesta.

**Automaattinen päätöksenteko:** Kyse on tietosuoja-asetuksen 22 artiklassa tarkoitetusta automatisoidusta yksittäispäätöksestä, jos päätös perustuu pelkästään automaattiseen käsittelyyn, kuten profilointiin, ja jos päätöksellä on rekisteröityä koskevia oikeusvaikutuksia tai se vaikuttaa rekisteröityyn vastaavalla tavalla merkittävästi.

**Erityisten henkilötietojen ryhmä:** Erityisiin henkilötietoryhmiin kuuluvia henkilötietoja ovat tiedot, joista käy ilmi poliittisia mielipiteitä, uskonnollinen tai filosofinen vakaumus, rotu tai etninen alkuperä, tai ammattiliiton jäsenyys. Erityisiä henkilötietoryhmiä koskeva käsittely kattaa lisäksi geneettisten tai biometristen tietojen käsittelyn yksiselitteistä tunnistamista varten, sekä terveyttä tai seksuaalista käyttäytymistä ja suuntautumista koskevien tietojen käsittelyn (tietosuoja-asetuksen 9 artiklan 1 kohta). Potilastiedot ovat erityisiin henkilötietoryhmiin kuuluvia tietoja, ja sosiaalihuollon asiakastiedot rinnastuvat erityisiin henkilötietoryhmiin.

**Henkilötieto:** Henkilötiedoilla tarkoitetaan ”kaikkia tunnistettuun tai tunnistettavissa olevaan luonnolliseen henkilöön, jäljempänä ’rekisteröity’, liittyviä tietoja” (tietosuoja-asetuksen 4 artiklan 1 kohta). Tunnistettavissa olevalla tarkoitetaan luonnollista henkilöä, joka voidaan suoraan tai epäsuoraan tunnistaa erityisesti tunnistetietojen perusteella. Tunnistetietoja ovat esimerkiksi nimi, henkilötunnus ja sijaintitieto, sekä henkilölle tunnusomainen fyysinen, fysiologinen, geneettinen, psyykinen, taloudellinen, kulttuurillinen tai sosiaalinen tekijä.

**Henkilötietojen käsittelijä:** Henkilötietojen käsittelijällä tarkoitetaan luonnollista henkilöä tai oikeushenkilöä, viranomaista, virastoa taikka muuta elintä, joka ”käsittelee henkilötietoja rekisterinpitäjän lukuun” (tietosuoja-asetuksen 4 artiklan 8 kohta).

**Monialainen palvelujen käyttö:** Monialaisella palvelujen käytöllä tarkoitetaan tässä käsikirjassa palvelukäyttöä, jossa asiakas käyttää erilaisten palveluryhmien palveluja, esimerkiksi perusterveydenhuollon, sosiaalihuollon ja päihdehuollon palveluja.

**Monialaisten palvelujen käytön ennakointi:** Monialaisten palvelujen käytön ja monialaisen palvelukäytön ennakointi tarkoittaa tässä käsikirjassa toimintaan, jossa ennustemallin tuottaman tiedon avulla toteutetaan ennakoivia ja ennaltaehkäiseviä toimenpiteitä, jotta monialaista palvelujen käyttöä ei syntyisi. Ennakoinnin perustana on yksittäisiä asiakkaita koskevat ennusteet. Yksittäistä asiakasta koskevan ennusteen tukemana voidaan tehdä tätä koskevia ennaltaehkäiseviä toimenpiteitä. Yksittäisiä asiakkaita koskevien ennusteiden pohjalta voidaan aggregoida asiakasryhmätasoisia ennusteita ja ryhtyä asiakasryhmätasoihin ennakoiviin ja ennaltaehkäiseviin toimenpiteisiin.

**Monialaisten palvelujen käytön ennustaminen:** Monialaisten palvelujen käytön ennustamisella tarkoitetaan tässä käsikirjassa riskin arviointia sille, että henkilö tulee lähitulevaisuudessa käyttämään runsaasti erilaisia sosiaali- ja terveystalv palveluja. Ennustaminen perustuu asiakas- ja potilastietojärjestelmiin kertyneiden tietojen pohjalta luotuun ennustemalliin. Riskin lisäksi ennustemallin avulla arvioidaan monialaiseen palvelukäyttöön johtavia riskitekijöitä ja ennusmerkkejä.

**Pseudonyymit tiedot:** Pseudonymisoinnilla tarkoitetaan henkilötietojen käsittelyä niin, ettei henkilötietoja voi yhdistää tiettyyn henkilöön lisätietoja käyttämättä (tietosuoja-asetuksen 4 artiklan 5 kohta). Pseudonymisoidut tiedot ovat asetuksen johdanto-osan 26 kappaleen mukaan tunnistettavissa olevaa luonnollista henkilöä koskevia tietoja. Näin ollen pseudonymisoidut henkilötiedot ovat tietosuoja-asetuksen soveltamisen piirissä.

**Rekisterinpitäjä:** Rekisterinpitäjällä tarkoitetaan luonnollista henkilöä tai oikeushenkilöä, viranomaista, virastoa tai muuta elintä, joka yksin tai yhdessä toisten kanssa määrittelee henkilötietojen käsittelyn tarkoitukset ja keinot (tietosuoja-asetuksen 4 artiklan 7 kohta). Rekisterinpitäjä on se osapuoli, jonka tarkoituksia varten henkilötietoja käsitellään ja joka tekee päätökset henkilötietojen käsittelyä koskien. Vaikka asetuksessa käytetään termejä 'rekisteröity' ja 'rekisterinpitäjä', ei tietojen tarvitse olla yksinomaan henkilörekisterissä.

**Tekoälyn määritelmä:** Tekoäly voidaan yleisellä tasolla määritellä tietokoneohjelmaksi, joka kykenee tekemään älykkäinä pidettäviä toimintoja<sup>1</sup>. Koneoppiminen on tekoälyn osa-alue, jonka tarkoituksena on saada ohjelmisto toimimaan entistä paremmin pohjatiedon ja mahdollisen käyttäjän toiminnan perusteella<sup>2</sup>. Tässä käsikirjassa käsitellään asiakas- ja potilastietoja hyödyntävää koneoppimismallia.

1 <https://fi.wikipedia.org/wiki/Teko%C3%A4ly>

2 <https://fi.wikipedia.org/wiki/Koneoppiminen>



# MITÄ ON MONIALAINEN PALVELUJEN KÄYTTÖ JA SEN ENNAKOINTI?



Sosiaali- ja terveyspalvelujen yhteentoimivuuden ja integraation kehittämiseen on Suomessa panostettu viime vuosina erityisen paljon resursseja erilaisissa kehittämissuunnitelmissa ja -hankkeissa (esim. STM 2021). Taustalla on niin Suomessa kuin Euroopassa yleisemminkin terveydellisten ja sosiaalisten ongelmien ja haasteiden moninaistuminen ja yhtäaikaisuus sekä väestön vanheneminen, jolloin tarve monialaisille ja yhteentoimiville palveluille on kasvanut (European Commission 2017; Goodwin ym. 2017; Gilbert 2016; Koivisto & Liukko 2017; Melchiorre ym. 2018). Tarpeeseen on Suomessa yritetty vastata esimerkiksi kehittämällä asiakas/palveluohjausta, oma/vastuutyöntekijämallia, monialaista tiimityöskentelyä ja asiakassuunnitelmaa, asiakkuuksien segmentointia sekä monialaisia palvelukokonaisuuksia ja -ketjuja (Koivisto & Tiirinki 2020; Koivisto ym. 2020).

Monialaisten palvelujen tarjoamisen keskeinen haaste on löytää ja tunnistaa henkilöt ja asiakkaat, joiden terveys- ja/tai sosiaaliset ongelmat ovat kasaantuneet ja jotka mitä ilmeisemmin hyötyisivät monialaisista palveluista. Laki sosiaali- ja terveydenhuollon järjestämisestä (612/2021) velvoittaa hyvinvointialueet huolehtimaan yhteen sovitettuja palveluja laaja-alaisesti tarvitsevien asiakasryhmien ja asiakkaiden tunnistamisesta. Tällaisia tunnistamisen välineitä on kehitetty Suomessa paljon, esimerkiksi riskikartoitukset, kyselylomakkeet, alkuhaastattelut ja herätteet tietojärjestelmissä, mutta niiden käyttö on toistaiseksi ollut epäsystemaattista ja organisaatiokohtaista (Hietapakka ym. 2019).

Vielä tärkeämpää kuin tunnistaa ja löytää monialaisista palveluista mahdollisesti hyötyvät henkilöt olisi ennustaa ja ennakoida monialaista palvelujen käyttöä. Jos pystyisimme ennustamaan henkilön olevan vaarassa ajautua haasteelliseen elämäntilanteeseen ja monialaisten palvelujen käyttäjäksi, voisimme tarjota ennakoivaa ja ennaltaehkäisevää apua ja tukea ajoissa ja silloin monien eri palvelujen käyttöä ei välttämättä syntyisi tai sitä voitaisiin lykätä. Tällöin vältettäisiin inhimillistä kärsimystä, parannettaisiin henkilöiden hyvinvointia ja mahdollisesti säästettäisiin sosiaali- ja terveydenhuollon resursseja. Osa monialaisesta palvelujen käytöstä voi kuitenkin olla välttämätöntä esimerkiksi synnynnäisen vamman tai sairauden takia. Raja välttämättömän ja vältettävissä olevan monialaisen palvelujen käytön välillä lienee kuitenkin häilyvä.

Monialaisten palvelujen käyttö ei välttämättä ole kalleinta palvelujen käyttöä, mutta se voi olla kallista, jos siihen sisältyy runsaasti erikoissairaanhoidon, vanhuspalvelujen, vammaispalvelujen tai lastensuojelun palvelujen käyttöä (Haula & Peltola 2020; Leskelä ym. 2013).

Ennustamiseen ja ennakointiin voidaan käyttää erilaisia välineitä. Sosiaali- ja terveydenhuollon ammattilainen voi esimerkiksi asiakkaan tuntemukseen, asiakkaan läheisen tekemään huoli-ilmoitukseen tai ennakoivaan kotikäyntiin (Kempele 2021) perustuen ennustaa asiakkaan tilan kehittymistä ja hyödyntää esimerkiksi riskiarviointiin kehitettyjä kysymyspatteristoja. Tarvitsemme kuitenkin ammattilaisten tekemän asiakastyön ja palvelutuotannon johtamisen tueksi systemaattisempia, yksittäisiä asiakkaita koskevaa dataa analysoivien ennustamisen välineiden kehittämistä ja niiden soveltamista suuriin asiakasmääriin (esim. Pentti ym. 2019).

Tässä käsikirjassa jäsennetään ja tarkastellaan, miten monialaisen palvelukäytön ennakoointia voidaan kehittää ja toteuttaa hyödyntämällä modernin tekoälyteknologian tarjoamia mahdollisuuksia.



## 1.1 Monialaisten palvelujen käyttö ja tarve

Tämä käsikirja keskittyy monialaisten palvelujen käytön ennustamiseen ja ennakointiin. Ei kuitenkaan ole olemassa absoluuttista tai edes kansallista määritelmää sille, kuinka monen eri toimialan tai palveluryhmän palveluja asiakkaan tulee käyttää, jotta käyttö olisi monialaista palvelujen käyttöä. Lisäksi tulee huomioida, että määrän laskeminen perustuu palvelujen ja toimialojen luokituksiin tai ryhmittelyihin, jotka voivat olla erilaisia eri organisaatioissa. Se, mikä katsotaan monialaiseksi palvelujen käytöksi, määrittyy prosesseissa, joissa esimerkiksi kehitetään välineitä monialaisen palvelukäytön ennustamiseen ja ennakointiin tai monialaisista palveluista mahdollisesti hyötyvien tunnistamiseen. Paljon palveluja käyttävillä viitataan taas tavallisesti henkilöihin, jotka käyttävät määrällisesti paljon palveluja, mutta käyttö ei välttämättä ole monialaista palvelujen käyttöä. Palvelun pitkäaikaisessa käytössä asiakas käyttää tiettyä palvelua tai palveluja säännöllisesti jopa vuosia. (Koivisto & Tiirinki 2020.)

Monialaisesta palvelujen tarpeesta ja monialaisesta palvelujen käytöstä puhutaan usein epämääräisesti. Käytön käsite on selkeämpi. Asiakas käyttää tällöin eri toimialojen tai palveluryhmien palveluja. Palvelutarpeen käsite on haastavampi. Henkilön moninaiset sosiaaliset ja terveydelliset haasteet ei ole sama asia kuin monialainen palvelutarve. Palvelutarpeita on hyödyllistä jäsentää asioina, jotka määrittyvät asiakkaan ja palvelutarjonnan välisessä vuorovaikutuksessa. Asiakkaan palvelutarpeet määritellään tai ainakin tulisi nykypäivänä määritellä vuorovaikutteisesti asiakkaan, mahdollisesti tämän läheisen ja sosiaali- ja terveydenhuollon ammattilaisten välisessä vuorovaikutuksessa. Tavallisesti puhutaan palvelutarpeen arvioinnista. Palvelutarpeet eivät ikään kuin asusta asiakkaassa, vaan ne ovat hänen tavoitteidensa perusteella ja vuorovaikutuksessa määriteltäviä asioita. Asiakkaan palvelutarpeen määrittelyä muovaa siis osaltaan se, millaisia palveluja asiakkaalle on tarjottavissa (palveluvalikoima). Eri osapuolet tuovat oman asiantuntemuksensa ja tietämyksensä vuorovaikutustilanteeseen, jossa asiakkaan palvelutarpeet yhteistoiminnallisesti määritellään.

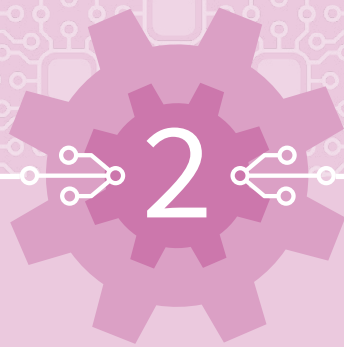
## 1.2 Monialaisten palvelujen käytön ennakointi

Monialaisten palvelujen käytön ennakointi tarkoittaa tässä käsikirjassa toimintaan, jossa tekoälypohjaisen ennustemallin tuottaman tiedon avulla ryhdytään ennakoiviin ja ennaltaehkäiseviin toimenpiteisiin, jotta monialaista palvelujen käyttöä ei syntyisi. Ennakoinnin perustana on yksittäisiä asiakkaita koskevat ennusteet. Yksittäistä asiakasta koskevan ennusteen tukemana voidaan tehdä tätä koskevia ennaltaehkäiseviä toimenpiteitä. Yksittäisiä asiakkaita koskevien ennusteiden pohjalta voidaan aggregoida asiakasryhmätasoisia ennusteita ja ryhtyä asiakasryhmätasoisin ennakoiviin ja ennaltaehkäiseviin toimenpiteisiin.

Kun ennakointia tehdään yksittäisen asiakkaan kohdalla, sosiaali- ja terveydenhuollon ammattilaiset ovat kiinnostuneita siitä, mitkä ennusmerkit viittaavat siihen, että henkilö on vaarassa ajautua monialaisten palvelujen käyttäjäksi, ja tekevät ennusteen tuella päätöksiä asiakkaan ennaltaehkäisevää palvelutarjontaa ja tukea koskien. Yksittäisten henkilöiden tietojen käyttöön liittyviä tietosuojakysymyksiä ja ennakointiin liittyviä lainsäädännön sallimia toimintamahdollisuuksia tarkastellaan tarkemmin luvussa 5.

Kun ennakointia tehdään asiakasryhmän tasolla, palvelujen järjestämistahon johtajat tai palvelujen tuottamisesta vastaavat esihenkilöt ovat kiinnostuneita siitä, millaisia ja kuinka suuria riskissä olevia asiakasryhmiä ilmenee yksittäisiä asiakkaita koskevien ennusteiden perusteella. Tämän ennustetiedon perusteella johto voi tehdä ennaltaehkäisyyn tähtääviä ratkaisuja palvelujen kehittämiseen, organisointiin ja resurssointiin liittyen.

Monialaisten palvelujen käytön ennustaminen erilaisten suurien datamääriä käsittelevien tekoälytekniikoiden avulla perustuu aina joihinkin valittuihin kriteereihin, arvoihin, periaatteisiin ja valikoituun dataan. Tällaiset tekniikat eivät tiedä henkilöiden reaaliaikaisesta elämäntilanteesta ja siihen vaikuttavista tekijöistä mitään. Tämän vuoksi ennustemallien avulla tehdyt ennusteet ovat aina epävarmoja varsinkin yksittäisen henkilön kohdalla. Kun ennustemallin määrityksiä ja valintoja muunnetaan, voi malli tuottaa toisenlaisia ennustetuloksia. Ennusteet perustuvat sananmukaisesti niihin valintoihin, joita mallia kehitettäessä tehdään. Ne voivat kuitenkin toimia sote-ammattilaisten hyödyllisenä tukena näiden tehdessä asiakkaan palveluja koskevia päätöksiä tai johdon tukena näiden tehdessä palvelujen kehittämistä, organisointia ja resurssointia koskevia päätöksiä.



# TEKOÄLYPOHJAISEN ENNAKOINNIN KÄYTÄNTÖJEN KEHITTÄMINEN



Sosiaali- ja terveydenhuollon uudistukseen sisältyvä yksi keskeinen tavoite on Suomessa siirtää sosiaali- ja terveydenhuollon painopistettä raskaista palveluista ennakoivaan ja ehkäisevään työhön (STM 2021). Terveydenhuollossa tavoitteena on siirtää painopistettä erikoissairaanhoidosta perustasolle. Sosiaalihuollossa tavoitteena on siirtää painopistettä varhaiseen tukeen ja matalan kynnyksen palveluihin. Erilaisten ennustemallien ja ennakoinnin toimintamallien kehittäminen, kokeilu ja hyödyntäminen palvelee näitä tavoitteita.

Tässä käsikirjassa tarkastellaan ja pohdintaa, miten monialaisten palvelujen käytön ennakointia voitaisiin hyvinvointialueilla ja sote-organisaatioissa kehittää asiakastyössä ja palvelutoiminnan johtamisessa hyödyntämällä modernia tekoälyä. Kuvio 2 kuvaa tekoälypohjaisen ennakoinnin käytäntöjen kehittämisen keskeisiä tehtäviä, joita seuraavassa jäsennetään tarkemmin. Luvuissa 3–5 jäsennetään monialaisen palvelukäytön tekoälypohjaisen ennustemallin kehittämisen ja toteuttamisen tehtäviä sekä tekoälypohjaisten ennustemallien kehittämisen ja hyödyntämisen lainsäädännöllisiä reunaehtoja.



**Kuvio 2.** Tekoälypohjaisen ennakoinnin käytäntöjen kehittäminen



## 2.1 Kehittämisen organisointi

Ennakointitoiminnan kehittäminen organisaatiossa edellyttää valmistelua, organisointumista ja suunnittelua. Ennakoinnin kehittämiseen pätee pääasiallisesti samat nykyaikaiset yhteiskehittämisen periaatteet kuin minkä tahansa asian kehittämiseen organisaatiossa. Ennakoinnin kehittämisen tulee olla jatkuvasti käynnissä olevaa toimintaa. Se, kuinka mittavasta kehittämisestä kulloinkin on kyse, määrittää paljolti sitä, miten ja millaisilla resursseilla kehittämistä tehdään. Pienempiä korjauksia voidaan tehdä osana arkista perustyötä. Mittavammat toiminnan uudistukset edellyttävät suunnitelmallisempaa järjestäytymistä.

Tässä käsikirjassa tarkastellaan ja jäsennetään monialaisten palvelujen käytön ennakointia yksittäisen asiakkaan kohdalla asiakastyössä ja asiakasryhmätasolla johtamisen käytäntöinä. Ennakoinnin välineenä tarkastellaan tekoälypohjaisen ennustemallin kehittämistä (Luku 3), jonka hyödyntäminen edellyttää tietojärjestelmien kehitystyötä (Luku 4) ja tietosuojalainsäädännön reunaehtojen huomioimista (Luku 5).

Tällainen ennakoinnin, ennustemallin ja tietojärjestelmien kehittäminen edellyttää omanlaistaan osallistamisen strategiaa kussakin kehittämisen toiminnossa. Kehittämiseen tulisi osallistua organisaatiossa niiden keskeisten toimijaryhmien edustajat, joiden työn kannalta kehitettävä asia on merkityksellinen. Eri toimijoilla on oma näkökulmansa ja asiantuntijuutensa kehitettävään asiaan ja heidän kuulemisensa on perusedellytys kehittämisen hyvän lopputuloksen kannalta.

Asiakastyön osalta tällaisia toimijoita voivat olla esimerkiksi vastaanottotyöntekijät, lääkärin, hoitajat, sosiaalityöntekijät, palvelu/asiakasohjaajat, sekä asiakaspäälliköt. Palvelutoiminnan johtamisen osalta toimijoita voivat olla esimerkiksi lähiesihenkilöt, muut esihenkilöt, palvelutuotannon johto sekä palvelujen järjestämisestä vastaava johto. Lisäksi kehittämiseen tulisi osallistua data-analytikoita ja sosiaali- ja terveydenhuollon asiakkaita. Osallistamisen menetelmät ja keinot voivat olla monenlaisia eikä niitä tarkastella tässä käsikirjassa tarkemmin.

Ennakoinnin kehittämistoiminnalle on tarpeellista laatia tiivis suunnitelma, jota päivitetään tarvittaessa kehittämisen edetessä. Kehittäminen ei etene lineaarisesti, vaan paremminkin vuorovaikutteisesti ja iteratiivisesti eri toimijoiden, tehtävien ja kehittämisen kohteiden tai toimintojen kesken. Kehittämissuunnitelman tulisi sisältää vastaukset mahdollisuuksien mukaan seuraaviin kysymyksiin:





- Kenellä on kehittämisen vetovastuu?
- Ketkä kuuluvat kehittämisen ydinryhmään ja mikä on eri toimijoiden työnjako ja vastuut?
- Mitkä ovat kehittämistarpeet?
- Mitkä ovat kehittämisen prosessi- ja tulostavoitteet?
- Mitkä ovat kehittämisen vaiheet ja kehittämistehtävät?
- Miten eri toimijaryhmien näkökulman huomioiminen ja osallistuminen varmistetaan kehittämistyön eri vaiheissa ja tehtävissä?
- Millaisia yhteiskehittämisen ja tiedon keruun menetelmiä käytetään kehittämisen eri vaiheissa?
- Millaisia erityisiä resursseja tarvitaan ja kuinka paljon?
- Ostetaanko kehittämistyötä ja millaista organisaation ulkopuolelta?
- Mikä on kehittämisen eri vaiheiden ja tehtävien aikataulu?
- Millaisilla mittareilla kehittämistä seurataan ja arvioidaan?



## PÄIJÄT-HÄME

### Osallistuminen kehittämiseen Päijät-Hämeessä

Päijät-Hämeen hyvinvointialueen hankkeessa kehitettiin tekoälyn tukemaa monialaisen palvelukäytön ennustamista ja ennakointia ikääntyneiden asiakasryhmää koskien. Kehittämisen vetovastuu oli jaettu ikääntyneiden palvelujen tulosaluejohtajan ja hyvinvointialueen tietojohdajan kesken. Kehittämisen operatiivisesta johtamisesta vastasi projektipäällikkö.

Päijät-Hämeen laajaan kehittäjäryhmään kuului monipuolinen joukko sote-ammattilaisia, kuten asiakasohjaajia, lääkäreitä, sairaanhoitajia, sosiaalityöntekijöitä ja -ohjaajia ja päihde- ja mielenterveyshuollon ammattilaisia, sekä data-analyttikoita, tietosuojan asiantuntijoita, tietotekniikan asiantuntijoita ja esihenkilöitä. Erilaisia ennakkoinnin ja ennustemallin kehittämisen vaiheita ja tehtäviä toteutettiin pienemmissä työryhmissä kulloistenkin tarpeiden pohjalta.



## 2.2 Monialaisten palvelujen käytön analyysin totuttaminen

Monialaisten palvelujen käytön ennustamisen ja ennakkoinnin kehittämisen perustan muodostaa analyysi siitä, millaista on tarkasteltavan asiakasryhmän, esimerkiksi yli 65-vuotiaiden, sosiaali- ja terveydenhuollon palvelujen käyttö ja erityisesti monialainen palvelujen käyttö tarkasteltavalla hyvinvointialueella. Tällaisen analyysin tarkoituksena on luoda kuva siitä, millaisiin segmentteihin/asiakasryhmiin kohderyhmä jakaantuu palvelukäytön osalta ja millaisiin mahdollisiin alasegmentteihin/ala-asiakasryhmiin monialaisesti palveluja käyttävät jakaantuvat. Suuri osa tarkasteltavasta asiakasryhmästä ei välttämättä käytä ollenkaan sosiaali- ja terveydenhuollon palveluja, osa käyttää palveluja epäsäännöllisesti tai hyvin vähän, osa voi käyttää esimerkiksi pitkäaikaissairauksen vuoksi säännöllisesti yhtä tiettyä palvelua ja vain pieni osa, arviolta alle kymmenen prosenttia (Haula & Peltola 2020) käyttää monialaisia palveluja.

Monialaisen palvelukäytön analyysin tekeminen edellyttää palvelujen luokittelua palveluryhmiin, jonka avulla voidaan määritellä, millaista on monialainen palvelujen käyttö. Palvelukäytön analyysin yksikkö on yksittäinen palveluja käyttävä asiakas. Yksittäisten asiakkaiden samantyyppisestä palvelukäytöstä voidaan luoda asiakasprofiileja, jotka kuvaavat kyseisten asiakkaiden tyypillistä palvelujen käyttöä. Kun samanlaista palvelukäyttöä on useammalla asiakkaalla, muodostavat nämä asiakassegmentin/ryhmän.

Palvelukäytön analyysi tehdään palvelujen ryhmittelyn pohjalta tietyltä aikajaksolta ja tiettyä asiakasryhmää koskien hyvinvointialueen sosiaali- ja terveydenhuollon asiakas- ja potilastietoja sekä muita tietoja hyödyntäen. Tietoja voi joutua kokoamaan useammasta tietojärjestelmästä, ellei tietoja ole saatavilla esimerkiksi erityisestä tietoaaltaasta tai järjestelmästä, joka kokoaa kaikki hyvinvointialueen sosiaali- ja terveydenhuollon potilas- ja asiakastiedot. Lisäksi hyödynnetään tarpeen mukaan muiden palveluntuottajien, esim. kuntien ja valtion tuottamaa tietoa. Palvelukäytön analyysiä voidaan lisäksi täydentää ja syventää muiden tiedonkeruun menetelmien avulla, esimerkiksi haastatteleamalla asiakkaita ja ammattilaisia.

Palvelukäytön analyysin tuottamaa määrällistä ja muuta mahdollista tietoa sekä sote-ammattilaisten tietämystä ja kokemusta hyödyntämällä voidaan määritellä ja kuvata monialaisesti palveluja käyttävien asiakasprofiileja. Kuvaamisen tulisi osallistua monialainen joukko sote-ammattilaisia ja esihenkilöitä. Asiakasprofiilissa voidaan kuvata muun muassa monialaisia palveluja käyttävän henkilön tyypillisiä piirteitä, palvelujen käyttötapoja ja palvelujen käyttöön liittyviä haasteita.





## PÄIJÄT-HÄME

### Ikääntyneiden monialaisten palvelujen käytön analyysi Päijät-Hämeessä

Päijät-Hämeessä tehtiin kehittämisen perustaksi asiakasdatoihin ja ammattilaisten asiantuntemukseen perustuva lähtötilanteen analyysi ikääntyneiden monialaisesta palvelujen käytöstä. Ensin tehtiin tilastollinen yleiskatsaus palvelujen käyttöön palvelukohtaisesti. Sen jälkeen määriteltiin monialaisessa työpajassa sote-ammattilaisten ja esimiesten tietämykseen ja kokemuksiin perustuen monialaisesti palveluja käyttävien ikääntyneiden asiakasprofieileja. Työpajassa määrittyi neljä keskeistä asiakasprofieilia:

1. Monialaisia palveluja käyttävä ikääntynyt muistisairas (paljon kotiin vietäviä ja muita palveluja; sairauden tunnottomat, psyykkiset ongelmat, jne.)
2. Ylipainosta ja siihen liittyvistä sairauksista kärsivä ikääntynyt (sydän- ja verisuonitaudit, 2 tyypin diabetes, tukielinvaivat, yms.)
3. Ikääntynyt päihde- ja mielenterveysasiakas (yksinäisyys, sosiaaliset ongelmat, masennus, alkoholin liikkakäyttöä, talousongelmat, yms.)
4. Diagnosoimaton ikääntynyt asiakas, joilla paljon yhteydenottoja ja asiointeja eri palveluissa; hoidosta ja palveluista ei vastaa kukaan, kokeilee palveluja

## 2.3 Vältettävissä olevan palvelujen käytön määrittely

Palvelujen käytön analyysi ja asiakasprofiilien kuvaaminen luo perustan vältettävissä olevan palvelujen käytön määrittelylle. Tällöin määritellään yhteistoiminnallisesti sellainen monialainen palvelujen käyttö, johon asiakkaiden ei toivota ajautuvan ja jota voitaisiin mahdollisesti ennaltaehkäistä tai lykätä. Osa monialaisesta palvelujen käytöstä voi kuitenkin olla välttämätöntä esimerkiksi synnynnäisen vamman tai sairauden takia. Raja välttämättömän ja vältettävissä olevan palvelukäytön välillä on kuitenkin häilyvä.

Vältettävissä olevan monialaisen palvelukäytön kombinaatioita voidaan määritellä yksi tai useampi riippuen palvelujen käytön analyysin tuloksista ja määrittelyistä monialaisesti palveluja käyttävien asiakasprofieileista. Vältettävissä olevan monialaisen palvelukäytön määrittelyjen perusteella voidaan aloittaa ennakoinnin toimintamallien ja ennustemallin kehittäminen. Vältettävissä olevan palvelukäytön määritykset voivat kehittyä ja muuttua ennustemallin kehittämisen edistyessä (ks. Luku 3).



## 2.4 Ennakoinnin kehittämistarpeiden ja -tavoitteiden määrittely

Palvelujen monialaisen käytön analyysi kertoo, millaisia monialaisia palveluja käyttäviä asiakasryhmiä alueella on ja kuinka paljon. Tämä tieto antaa pohjaa ennakoinnin kehittämistarpeiden määrittelylle. Lisäksi perustana on käsitys ja ymmärrys siitä, miten ja millaista ennakointia organisaatiossa jo tehdään ja millaisia käytäntöjä monialaisesti palveluja käyttävien osalta organisaatiossa on jo kehitetty ja käytössä.

Kehittämistarpeet määritellään yhteistoiminnallisesti eri toimijaryhmien näkemykset huomioiden. Kehittämistarpeita määritellään ainakin palvelutoiminnan johtamisen näkökulmasta ja asiakastyön näkökulmasta. Johtamisen näkökulmasta voi tarve olla esimerkiksi luoda monialaisen palvelujen käytön ennakoinnin toimintamalli valitsevan korjaavaan sosiaali- ja terveydenhuoltoon painottuvan toimintamallin rinnalle ja tätä tukevat palvelutoiminnan ohjauksen käytännöt. Asiakastyön näkökulmasta voi olla tarve esimerkiksi luoda ennaltaehkäisevä toimintatapa, jossa riskiasiakkaille pyritään tarjoamaan ehkäisevää tukea ja apua varhaisessa vaiheessa ja joka rakennettaisiin osaksi asiakasohjausta ja monialaisen palveluprosessin toimintokokonaisuutta.

Ennakointitoiminnan kehittämistä ei vielä aloiteta kehittämistarpeiden pohjalta, vaan tarpeet tulee ensin kääntää kehittämisen tavoitteiksi. Laveamman raamin kehittämistavoitteille antavat alueen sosiaali- ja terveydenhuollon strategiset tavoitteet ja muut ylätason tavoitteet. Ennakoinnin kehittämiseksi voidaan määritellä prosessitavoitteita ja tulostavoitteita. Prosessitavoitteet kertovat, millaisia toimintamalleja, palveluja, toimintoja ja tekniikoita on tavoitteena kehittää. Tulostavoitteet kertovat, millaista muutosta kehitettävillä asioilla tavoitellaan esimerkiksi monialaisesti palveluja käyttävien määrässä, asiakkaiden terveydessä ja hyvinvoinnissa sekä kustannuksissa.



## PÄIJÄT-HÄME

### Monialaisen palvelutarpeen ennakoinnin kehittämistarpeet ja -tavoitteet Päijät-Hämeessä

Päijät-Hämeessä määriteltiin monialaisen palvelukäytön ennakoinnin kehittämistarpeita ja -tavoitteita palvelutoiminnan johtamisen ja asiakasohjauksen näkökulmista.

#### Tarpeet:

Palvelutoiminnan johtamisen näkökulmasta kehittämisen tarpeiksi määriteltiin

- tarve suunnitella ja kehittää palveluja ja palvelutoimintaa osaltaan ennustetietoon ja ennakointiin tukeutuen,
- tarve löytää ja kohdentaa ennaltaehkäiseviä toimenpiteitä (yhteistyö kuntien kanssa) ennustetiedon tukemana (oikeiden asiakkaiden löytäminen, väliinputoaja-asiakkaiden huomioiminen, oikea-aikaisuus, jne.), ja
- tarve seurata ja arvioida (monialaista) palvelujen käyttöä johtamisen tukena (monialaisen palvelukäytön muuttuminen, ennustetiedon muuttuminen).

Asiakasohjauksen näkökulmasta kehittämisen tarpeiksi määriteltiin

- tarve hyödyntää asiakasprofiilien mukaisia ennuste- ja muita tietoja asiakkaiden ohjaamisessa ja neuvonnassa,
- tarve suunnitella ja yhteensovittaa palveluja asiakkaan kanssa ennustetietoa hyödyntäen, ja
- tarve kiinnittää erityisesti huomio ennaltaehkäiseviin toimenpiteisiin.

#### Tavoitteet:

Palvelutoiminnan johtamisen näkökulmasta

- prosessitavoitteeksi määriteltiin systemaattinen ennustetiedon hyödyntämisen toimintamalli osaksi johtamisen toimintoja, kuten palvelujen suunnittelu ja kehittäminen, ennaltaehkäisevien toimenpiteiden kohdentaminen, ja seuranta ja arviointi, ja
- tulostavoitteiksi määriteltiin:
  - 1) monialaisesti palveluja käyttävien ikäihmisten määrä ja kustannukset vähenevät ja
  - 2) asiakkaiden palveluiden käyttö on suunnitelmallista ja yhteensovitettua.

Asiakasohjauksen näkökulmasta

- prosessitavoitteiksi määriteltiin:
  - 1) ammattilaisten yhtenäiset toimintatavat ja kirjaamisohjeet sekä niiden noudattaminen, ja
  - 2) ammattilaisten yhtenäinen systemaattinen ennustetietoa hyödyntävä toimintatapa, joka huomioi asiakkaiden yksilölliset tarpeet.
- tulostavoitteiksi määriteltiin:
  - 1) asiakkaiden palvelut ovat oikea-aikaisia ja oikein kohdennettuja,
  - 2) asiakkaat pärjäävät pidempään ilman sote-palveluita ja kevyemmällä sote-palveluilla sekä
  - 3) asiakkaat ovat terveitä ja hyvinvoivia.



## 2.5 Ennakoinnin käytäntöjen määrittely

Ennakoinnin käytäntöjen kehittäminen tapahtuu asetettujen prosessi- ja tulostavoitteiden perusteella, mutta tavoitteita voidaan tarvittaessa tarkentaa ja määritellä uudelleen. Kehittämisen kohteena on tällöin monialaisten palvelujen käytön ennakointi asiakasryhmätasolla johdon toimintana ja ennakointi yksittäistä asiakasta koskien asiakastyössä.

Kehitettävää toimintaa voidaan jäsentää tavoitteellisina ennakoinnin käytäntöinä, joissa toimijat hyödyntävät tiettyä teknologiaa, tämän käsikirjan tarkastelussa tekoälypohjaista ennustemallia. Keskeisiä kysymyksiä tällöin ovat esimerkiksi seuraavat:



- *Millaista ennustetietoa ennustemallin tulisi tuottaa johdolle ja asiakastyöhön?*
- *Ketkä hyödyntävät ennustemallin tuottamaa tietoa ja tekevät päätöksiä sen tukemana?*
- *Mitä toimijoita tarvitaan ennustetiedon tuottamiseen ja kokoamiseen?*
- *Millaisia asioita koskien päätöksiä tehdään ennustetietoa hyödyntäen, esim. palvelujen kehittäminen, organisointi ja resursointi, sekä yksittäisen asiakkaan ennaltaehkäiseviin palveluihin, tukeen ja apuun liittyvät päätökset?*
- *Millaisista erityisistä tehtävistä ennakointi koostuu?*
- *Miten ennakoinnin käytännöt ja tehtävät kytkeytyvät ja linkittyvät muihin johtamisen ja asiakastyön käytäntöihin?*
- *Millaista osaamista ennakointiin osallistuvilta toimijoilta edellytetään?*
- *Millaisia uudenlaisten työroolien määrittelyä ja kuvaamista ennakointi edellyttää?*



Kehittämisen perustana on oltava jonkinlainen kuvaus ja ymmärrys siitä, miten monialaisia palveluja käyttävien kanssa toimitaan nykytilanteessa. Mikäli organisaatiossa toteutetaan jo jollakin tavalla monialaisen palvelukäytön ennakointia, on nämä käytännöt tarpeellista kuvata riittävällä tarkkuudella (toimijat, tehtävät, tieto ja välineet) lähtötilanteeksi, jota lähdetään kehittämään ja muuntamaan. Tämän jälkeen määritellään ja kuvataan yhteistoiminnallisesti uudet ennakoinnin käytännöt. Sekä lähtötilanteen että uusien käytäntöjen kuvaaminen tekee näkyväksi muutoksen. Uudet ennakoinnin käytännöt voivat korvata aiempia ennakoinnin käytäntöjä, mutta samalla ne linkittyvät muihin johtamisen ja asiakastyön käytäntöihin. Vanhojen ja uusien käytäntöjen kuvauksia voidaan hyödyntää myöhemmin, kun henkilökunta perehdytetään ja koulutetaan uusiin käytäntöihin.

Uusia käytäntöjä on tarpeellista lähteä kokeilemaan jo varhaisessa vaiheessa ja tehdä korjausliikkeitä, mutta tämä on mahdollista vasta, kun ennustemalli ja sen tuottamaa tietoa on käytettävissä. Käytännössä käytäntöjä ja ennustemallia kehitetään iteratiivisesti ja vuorovaikutteisesti askel askeleelta eteenpäin. Luvussa 3 jäsennetään ja tarkastellaan tarkemmin tekoälypohjaisen ennustemallin kehittämistä. Silloin kun hyödynnetään ja otetaan käyttöön jossakin muualla jo kehitetty ennustemalli, tulee malli sovittaa oman organisaation tarpeisiin sekä sen toiminnallisiin ja teknisiin vaatimuksiin.





## Ennakoinnin kehittäminen Päijät-Hämeessä

Päijät-Hämeen hankkeessa ennakointitoimintaa kehitettiin ensisijaisesti ikääntyneiden palvelutoimintaan johdon toiminnoksi ja asiakasohjauksen toiminnoksi. Työkokouksissa, joihin osallistui tulosalueen johto ja asiakasohjaaja, määriteltiin ensin, millaista tietoa ennustemallin tulisi tuottaa palvelutoiminnan johdolle ja asiakasohjaukseen. Tämän jälkeen mallinnettiin ensin johdon ja asiakasohjauksen nykyisiä toimintoja ja käytäntöjä ja sen jälkeen uusia, tekoälymallin tuottamaan ennustetietoon pohjaavia toimintoja ja käytäntöjä.

### Ennakointi palvelutoiminnan johdon toimintana:

Millaista tietoa monialaisen palvelutarpeen ennustemallin tulisi tuottaa Päijät-Hämeen sote-johdolle ja erityisesti ikääntyneiden palvelujen johdolle? (Tässä ja uusien tekoälyn tukemien käytäntöjen suunnittelussa ei huomioitu tietosuojarajoitteita. Luvun 5 lopussa tarkastellaan Päijät-Hämeessä hankkeen näkökulmasta tietosuojakysymyksiä.)

- kuinka suuri osuus Päijät-Hämeen soten palveluissa asiakkaana olevista ikäihmisistä kuuluu riskiryhmään, joka on vaarassa ajautua käyttämään monialaisesti palveluja; millaisiin alaryhmiin riskiryhmä jakaantuu
- kuinka suuri osuus Päijät-Hämeen ikääntyneiden palveluissa asiakkaana olevista ikäihmisistä kuuluu riskiryhmään, joka on vaarassa ajautua käyttämään monialaisesti palveluja; millaisiin alariskiryhmiin ko. asiakkaat jakaantuvat
- kuinka suuri osuus ikäihmisistä, jotka eivät ole ikäihmisten palveluissa (ja jotka halutaan pitää poissa sieltä), ovat riskiryhmään kuuluvia;
- millaiset riskitekijät (ennusmerkit) ennustavat riskiryhmään ja sen johonkin alaryhmään ajautumista

### Ikääntyneiden palvelujen johtamisen toiminnot ja uudet tekoälypohjaista ennustemallia hyödyntävät käytännöt Päijät-Hämeen hyvinvointialueella.

Yksikkö	Toiminto	Uusi käytäntö	Hyödynnettävä tieto	Toimijat
Ikääntyneiden palvelujen johto	Ennustaminen	Ennustetaan Päijät-Hämeen soten asiakkaina olevien ikääntyneiden riskiä ajautua monialaisesti palveluja käyttävien asiakkaiden ryhmään ja sen eri alaryhmiin. Ennustetaan, kuinka suuri osuus ikääntyneiden palveluissa asiakkaana olevista on riskissä ajautua ko. ryhmään ja sen eri alaryhmiin	Ennustemallin tuottama tieto riskiryhmistä ja riskitekijöistä	Toimialajohtaja, tulosaluejohtaja, johtoryhmä (sis. kontrolleri)





Yksikkö	Toiminto	Uusi käytäntö	Hyödynnettävä tieto	Toimijat
Ikääntyneiden palvelujen johto	Segmentointi	<p>Segmentoidaan Päijät-Hämeen soten asiakkaina olevat ikääntyneet kahteen pääsegmenttiin: 1) ei ole vaarassa ajautua käyttämään monialaisia palveluja - riskiryhmään kuulumattomat ja 2) vaarassa ajautua käyttämään monialaisesti palveluja - riskiryhmä ja sen alaryhmiin kuuluvat.</p> <p>Segmentoidaan ikääntyneiden palveluissa asiakkaana olevat ikääntyneet em. segmentoinnin mukaisesti.</p>	Ennustemallin tuottama tieto riskiryhmistä ja riskitekijöistä	<p>Toimialajohtaja, tulosaluejohtaja, johtoryhmä (sis. kontrolleri)</p> <p>Asiakasohjaajat</p>
Ikääntyneiden palvelujen johto	Ennakointi	<p>Ennakoiva toiminta riskiryhmien osalta, esim. turvallisten asumismuotojen organisointi, yhteisöllinen asuminen</p> <p>Toteutetaan ennustetiedon perusteella tarvittava ennakoiva yhteistyö hyvinvointialueen muiden sote-palvelujen kanssa</p> <p>Toteutetaan ennustetiedon perusteella tarvittava ennakoiva yhteistyö kuntien ja järjestöjen kanssa</p>	Ennustemallin tuottama tieto riskiryhmistä ja riskitekijöistä	Toimialajohtaja, tulosaluejohtaja, johtoryhmä (sis. kontrolleri)
Ikääntyneiden palvelujen johto	Suunnittelu ja kehittäminen	Ks. ennakointi. Kehitetään tarvittaessa uusia palveluja ja yhteistyön muotoja	Ennustemallin tuottama tieto riskiryhmistä ja riskitekijöistä	Asiakasohjaus, tulosaluejohtaja



Yksikkö	Toiminto	Uusi käytäntö	Hyödynnettävä tieto	Toimijat
Ikääntyneiden palvelujen johto	Seuranta ja arviointi	<p>Seurataan ikääntyneiden palveluissa asiakkaina olevien, monialaisesti palveluja käyttävien määrää ja kustannuksia; 1 kerta/kk</p> <p>Ikääntyneiden palvelujen asiakkaiden osalta monialaisen tuen asiakkuuksien kesto; 1 kerta/kk</p> <p>Riskiryhmään ja sen alaryhmiin kuuluvien ikääntyneiden palvelujen asiakkaiden osuuksien seuranta; 1 kerta/kk</p>	Ennustemallin tuottama tieto riskiryhmistä ja riskitekijöistä; palvelukäyttöä ja kustannuksia koskeva tieto asiakas- ja potilastietojärjestelmissä	Toimialajohtaja, tulosaluejohtaja, johtoryhmä (sis. kontrolleri)

### Ennakointi asiakasohjauksessa:

Millaista tietoa monialaisen palvelutarpeen ennustemallin tulisi tuottaa asiakasohjaukselle yksittäistä asiakasta koskien? (Tässä ja uusien tekoälyn tukemien käytäntöjen suunnittelussa ei huomioitu tietosuojarajoitteita. Luvun 5 lopussa tarkastellaan Päijät-Hämeessä hankkeen näkökulmasta tietosuojakysymyksiä.)

- tieto siitä, kuuluuko asiakas riskiryhmään (vaarassa ajautua käyttämään monialaisia palveluja) vai ei, ja jos kuuluu, niin millaiseen alariskiryhmään
- keskeisimmät asiat asiakasprofiilissa; asioinnit, yhteydenotot, käynnit, diagnoosit, peruutetut käynnit tai menemättä jättämättömät käynnit; mitkä asiat tai tekijät selittävät riskiryhmään kuulumista
- riskin vaaran asteet; esim. lievä, keskivaikea, kriittinen
- puheeksi ottamisen aiheet
- hälytykset asetettujen rajojen perusteella



**Ikääntyneiden palvelujen asiakasohjauksen toiminnot ja uudet tekoälypohjaista ennustemallia hyödyntävät käytännöt Päijät-Hämeen hyvinvointialueella.**

Yksikkö	Toiminto	Uusi käytäntö	Hyödynnettävä tieto	Toimijat
Ikääntyneiden asiakasohjaus	Asiakkaan ennakoiiva kontaktointi	Asiakasohjaaja tai ”robotti” ottaa asiakkaaseen yhteyttä ennakoivasti herätteen perusteella. Asiakasta suositellaan ottamaan yhteyttä asiakasohjauksen neuvontaan.  Ohjataan asiakas tekemään oma-arviointi.	Ennustemallin tuottama tieto asiakasta koskien	Asiakasohjaajat, neuvonnan työntekijät
Ikääntyneiden asiakasohjaus	Asiakkaan ongelman ratkaisu neuvonnassa tai asia siirtyy viranhaltija-puolelle	Neuvonnan työntekijä tarkastaa/huomioi asiakkaan ennusteen/herätteet	Ennustemallin tuottama tieto asiakasta koskien	Asiakas, neuvonnan työntekijät, asiakasohjaajat
Ikääntyneiden asiakasohjaus	Ratkaistaan asia kevyemmällä ohjauksella tai tehdään palvelutarpeen arviointi	Puheeksi ottaminen, mikä heräte oli ja minkä asteinen, mihin riskiryhmään kuulu	Ennustemallin tuottama tieto asiakasta koskien	Asiakas ja omainen, asiakasohjaaja, palvelun tuottaja tarvittaessa
Ikääntyneiden asiakasohjaus	Asiakkaalle laaditaan asiakas-suunnitelma ja järjestetään tarvittavat palvelut	Riskin asteen ja riskitekijöiden huomioiminen ja toimenpiteiden suosittelu ja suunnittelu	Ennustemallin tuottama tieto asiakasta koskien	Asiakas ja omainen, asiakasohjaaja
Ikääntyneiden asiakasohjaus	Asiakkaan palvelukäytön, tilanteen ja ennusteen seuranta	Asiakkaan tilanteen seuranta riskiryhmien suhteen; tilanteen ja palvelukäytön muuttuminen	Ennustemallin tuottama tieto asiakasta koskien	Omatyöntekijä, asiakas, omainen, palveluntuottaja
Ikääntyneiden asiakasohjaus	Palvelujen päättäminen	Tarkistetaan ennustetieto vielä palvelun päättyessä	Ennustemallin tuottama tieto asiakasta koskien	Omatyöntekijä, asiakasohjaaja

## 2.6 Ennakoinnin käytäntöjen toimeenpano

Monialaisen palvelukäytön ennakoinnin käytäntöjen toimeenpano ja toteuttaminen edellyttää siihen liittyvien toimenpiteiden suunnittelua ja sopimista. Toimeenpano on aina jonkinlainen organisatorinen muutos- ja oppimisprosessi (Nurminen ym. 2002), ei pelkästään esimerkiksi teknisen ennustemallin implementointi. Toimeenpano edellyttää vaiheistusta ja aikataulutusta sekä vastuuhenkilöiden nimeämistä:



- *Kuka vastaa koko ennakointitoiminnan toimeenpanosta?*
- *Kuka vastaa johdon ennakointitoiminnan toimeenpanosta?*
- *Kuka vastaa asiakastyössä tapahtuvan ennakoinnin toimeenpanosta?*
- *Millaisista vaiheista ja tehtävistä toimeenpano koostuu?*
- *Miten henkilöstön koulutus ja perehdytys toteutetaan?*



Ennakoinnin toimeenpano edellyttää ennakointia tekevien ammattilaisten ja johdon sekä muun tarvittavan henkilöstön kouluttamista ja perehdyttämistä. Toimeenpanon ja koulutuksen tärkein materiaali on ennakointitoimintaan liittyvän lähtötilanteen käytäntöjen ja uusien ennakoinnin käytäntöjen kuvaukset. Koulutuksen tulee sisältää tietoa toiminnan tavoitteisiin, sääntöihin ja toimintakulttuurin muutokseen liittyen sekä tietoa toimintaa sääntelevästä tietosuojalainsäädännöstä. Lisäksi koulutuksen tulee perehdyttää henkilöstöä ennustamisen problematiikkaan, datanlukutaitoon ja ennustemallin tekniseen käyttöön. Erityisesti tulee kiinnittää huomiota tekoälypohjaisten ennustemallien luonteen (epävarmuus, datariippuvuudet, mahdollinen mallien heikkeneminen ajan myötä) ymmärtämisen ja tulosten tulkintaan. Ennustemallin käytettävyyttä tarkastellaan luvussa 3 ja teknistä implementaatiota luvussa 4.

## 2.7 Palvelujen käytön jatkuva seuranta ja arviointi

Palvelujen tuottamisen ja toteuttamisen johtaminen ja ohjaus sekä palvelujen kehittäminen edellyttävät tarkasteltavan asiakasryhmän palvelujen käytön säännöllistä seuranta. Seurantatiedon perusteella johto voi ryhtyä tarvittaviin ohjauksen ja kehittämisen toimenpiteisiin. Yksittäisen asiakkaan kohdalla voidaan ryhtyä asiakasta koskeviin toimenpiteisiin.

Ennakoinnin ja ennustamisen kehittämisen lähtötilanteessa huolellisesti määritelty ja toteutettu palvelujen käytön analyysi tarjoaa kehyksen säännöllisesti toteutettavalle seurannalle ja arvioinnille. Seurannan ja arvioinnin tarkoitus, tarkkuus sekä toteuttamisvastuu ja -frekvenssi määritellään organisaatio- ja asiakasryhmäkohtaisesti. Seurannan perusanalyysiyksikkö on yksittäinen asiakas ja tämän palvelukäyttö. Tavallisesti seuranta on tehty esimerkiksi palveluryhmäperustaisesti tai palveluperustaisesti seuraamalla



palvelujen käytön kokonaismäärää. Kehittyneet tekoälyteknologiat tarjoavat kuitenkin mahdollisuuksia toteuttaa yksittäisen asiakkaan palvelukäyttöön perustuvaa asiakasryhmätasoisista seuranta.

Tällaisen seurantakehyksen avulla voidaan säännöllisesti seurata, millaisiin segmentteihin/asiakasryhmiin kohderyhmä jakaantuu palvelukäytön osalta, millaisiin alasegmentteihin/asiakasryhmiin segmentit jakaantuvat, esimerkiksi monialaisesti palveluja käyttävien alasegmentit, ja miten nämä muuttuvat ajan kuluessa. Yksittäisen asiakkaan kohdalla voidaan seurata palvelujen käytön, riskiryhmään kuulumisen ja ennusteen muutosta.



**PÄIJÄT-HÄME**

**Tekoälypohjaiseen ennustemalliin perustuva seuranta Päijät-Hämeestä**

Yksikkö	Toiminto	Uusi käytäntö	Hyödynnettävä tieto	Toimijat
Ikääntyneiden palvelujen johto	Ikääntyneiden asiakasryhmien seuranta ja arviointi	Seurataan ikääntyneiden palveluissa asiakkaina olevien, monialaisesti palveluja käyttävien määrää ja kustannuksia; 1 kerta/kk  Ikääntyneiden palvelujen asiakkaiden osalta monialaisen tuen asiakkuuksien kesto; 1 kerta/kk  Riskiryhmään ja sen alaryhmiin kuuluvien ikääntyneiden palvelujen asiakkaiden osuuksien seuranta; 1 kerta/kk	Ennustemallin tuottama tieto palvelukäytöstä, riskiryhmistä ja riskitekijöistä; kustannuksia koskeva tieto asiakas- ja potilastietojärjestelmissä	Toimialajohtaja, tulosaluejohtaja, johtoryhmä (sis. kontrolleri)
Ikääntyneiden asiakasohjaus	Asiakkaan palvelukäytön, tilanteen ja ennusteen seuranta ja arviointi	Asiakkaan tilanteen seuranta riskiryhmien suhteen; tilanteen ja palvelukäytön muuttuminen	Ennustemallin tuottama tieto asiakasta koskien	Omatyöntekijä, asiakas, omainen, palveluntuottaja





# TEKOÄLYPOHJAISTEN ENNUSTEMALLIEN KEHITTÄMINEN



Tässä käsikirjassa tarkastellaan tekoälyn hyödyntämistä sosiaali- ja terveydenhuollon palvelujen suunnittelussa ja toteutuksessa. Keskitymme koneoppimiseen perustuvaan ennustemalliin ja sen kehittämisen keskeisiin tehtäviin. Tekoäly voidaan yleisellä tasolla määritellä tietokoneohjelmaksi, joka kykenee tekemään älykkäinä pidettäviä toimintoja<sup>3</sup>. Koneoppiminen (engl. machine learning, ML) on tekoälyn osa-alue, jonka tarkoituksena on saada ohjelmisto toimimaan entistä paremmin pohjatiedon ja mahdollisen käyttäjän toiminnan perusteella<sup>4</sup>.

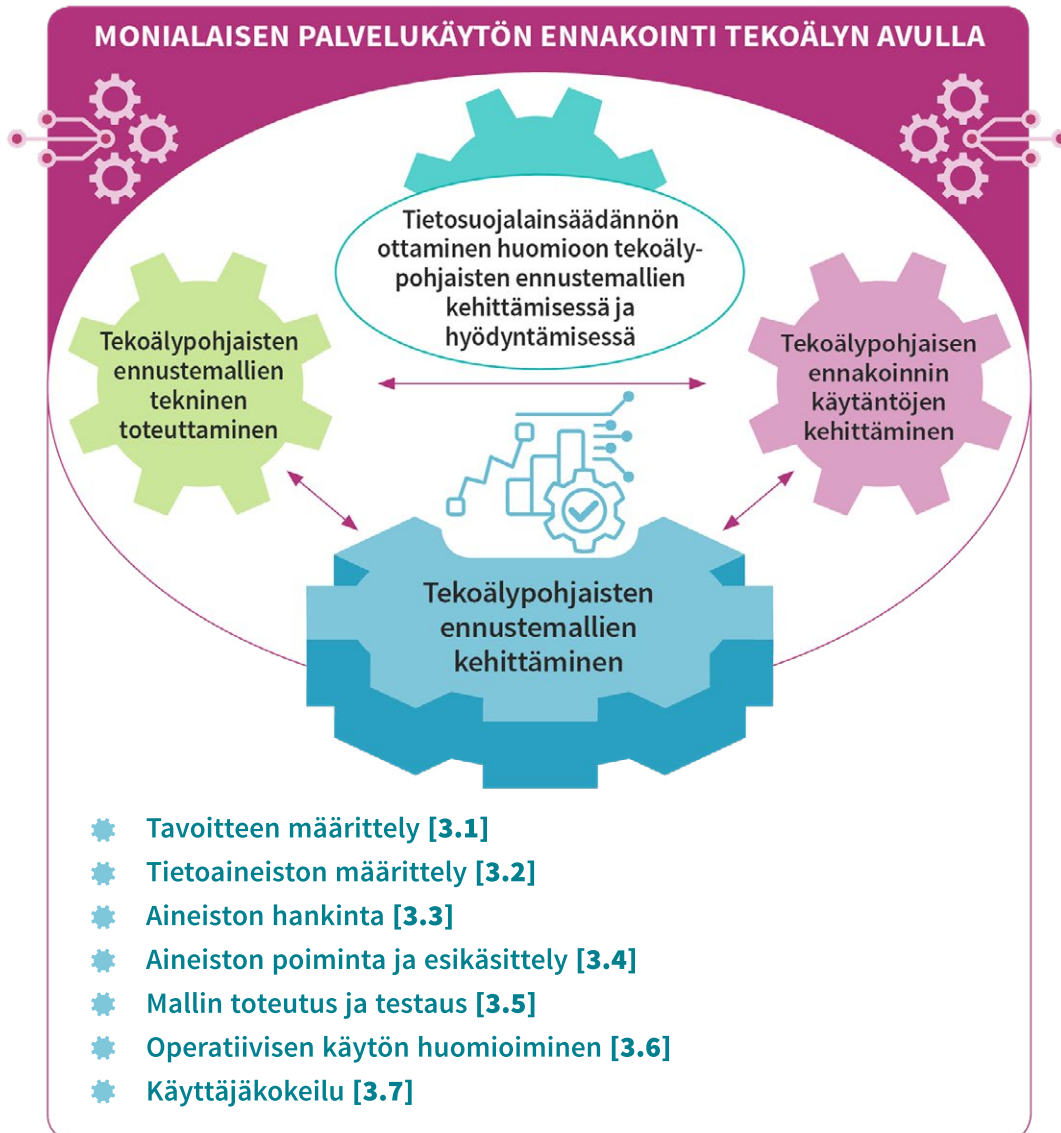
Tässä luvussa painopiste on tekoälypohjaisen ennustemallin kehittämisessä – mallin vientiä operatiiviseen tietojärjestelmäympäristöön käsitellään luvussa 4. Kutakin kehittämistehtävää käsitellään yleisesti ennustemallien kehittämisen näkökulmasta sekä kuvataan, miten kehittämistehtävä toteutui Päijät-Hämeen kehittämishankkeessa.

Sosiaali- ja terveydenhuollon palvelujen toteuttamista tukevan ennustemallin kehittämisessä tarvitaan asiakkaan terveyttä ja hyvinvointia kuvaavia tietoja. Tässä käsikirjassa lähtökohtana on ns. tosielämän datan (realworld data, RWD) hyödyntäminen, jolloin mallin kehittämiseen käytetään palvelujen käytön yhteydessä asiakas- ja potilasrekisteriin kertyneitä tietoja.

---

3 <https://fi.wikipedia.org/wiki/Teko%C3%A4ly>

4 <https://fi.wikipedia.org/wiki/Koneoppiminen>



**Kuvio 3.** Tekoälypohjaisten ennustemallien kehittämisen keskeiset tehtävät

Tekoälypohjaisia ennustemalleja on aiemmin kehitetty tyypillisesti sovelluksiin, joissa pyritään ennustamaan jonkin tietyn sairauden ilmaantumista, uusiutumista tai pahenemista (esim. Antikainen 2023, Liu 2020). Huomattavan vähän on raportoitu yrityksistä ennustaa ikääntyneiden riskiä monialaiseen palvelujen käyttöön (Bardsley 2011, Nakubulwa 2022).



## 3.1 Tavoitteen määrittely

Ennustemallin kehittämisen alkuvaiheessa määritellään tavoite – mihin ongelmaan tai ongelmiin mallin on tarkoitus tuoda ratkaisu. Tällöin tulee huomioida toisaalta tavoitteet johtamisen näkökulmasta sekä toisaalta asiakastyön näkökulmasta (katso tarkemmin luku 2). Tavoitetta määriteltäessä on lisäksi huomioitava operatiiviseen käyttöön liittyvät rajoitteet, mm. tietosuojalainsäädännöstä johtuen (ks. luku 5).

Ennustemallin lähtötietona ovat sote-organisaation asiakas- ja potilastiedot sekä mahdolliset muut yksilön hyvinvointiin liittyvät tiedot. Näiden tietojen pohjalta ennustemallin tulee tuottaa ennusteita, jotka auttavat tunnistamaan yksilöitä ja ryhmiä, joilla on kohonnut todennäköisyys kasvavaan ja monialaiseen sote-palvelun käyttöön. Johtamisen tasolla tavoitteena on tunnistaa erilaisia asiakasryhmiä siten, että kullekin ryhmälle sopivia terveyttä ja hyvinvointia ennakoivasti ylläpitäviä toimenpiteitä voidaan kohdennetusti toteuttaa. Asiakastyön näkökulmasta tavoitteena voi olla yksittäisen asiakkaan hoidon ja palvelujen suunnittelun tukeminen siten, että saadaan paras mahdollinen vaikuttavuus käytettävissä olevilla resursseilla.

Ennustemallin toteuttamista varten ylätasoon tavoitteet tulisi tarkentaa täsmällisiksi toiminnallisiksi vaatimuksiksi. Lisäksi olisi hyvä asettaa tavoite mallin ennustetarkkuudelle esimerkiksi tavoiteltava minimi AUC-arvo (area under curve). Täsmällisten vaatimusten ja tarkkuustavoitteen määrittely ei kuitenkaan aina ole mahdollista kehittämistyön alkuvaiheessa. Koneoppimiseen perustuvan mallin kehittäminen on tyypillisesti eksploraatiivinen prosessi, jossa yksityiskohdat, mukaan lukien syöttötiedot ja tulokset, täsmenntyvät tyypillisesti vasta kehityshankkeen kuluessa.

Kehittämisen alkuvaihe voi olla luonteeltaan tieteellinen tutkimushanke, jonka myötä selvitetään mitä tietoaineistoja ja mallinnusmenetelmiä ao. tapauksessa kannattaa hyödyntää ja millaiseen ennustetarkkuuteen on mahdollista päästä. Tutkimushankkeen myötä saadaan tietoa käytettävissä olevan datan laadusta, mikä vaikuttaa suoraan mallinnuksen tarkkuuteen sekä ylipäätään datan soveltuvuuteen aiottuun tarkoitukseen.



## PÄIJÄT-HÄME

### Tavoitteen määrittely Päijät-Hämeessä

Päijät-Hämeessä ennustemallin tavoitteeksi määriteltiin tunnistaa ikääntyneet henkilöt, joilla on korkea riski ajautua monialaiseen palvelukäyttöön, vaikka tämä tilanne olisi estettävissä puuttamalla tilanteeseen oikeilla palveluilla riittävän ajoissa. Ennustemallin tulisi palvella toisaalta yksittäisen asiakkaan palveluohjausta sekä toisaalta tukea tiedolla johtamista. Mallin antaman ennusteen perusteella on tarkoitus ohjata yksittäisiä asiakkaita sopivien palveluiden piiriin entistä aiemmin, ehkäistä heidän toimintakykynsä ja elämänlaatunsa heikkenemistä, sekä auttaa palvelujen suunnittelijoita ennakoimaan resurssien tarve eri palveluissa. Ennustemallissa hyödynnetään monipuolista ja pitkäaikaista asiakas- ja potilastietoa. Päijät-Hämeen kehittämishankkeeseen sisältyi ennustemallin demonstraatioversion (”Proof-of-concept”, PoC) toteutus sekä suppea kokeilu, jotka toteutettiin testikäyttöliittymän avulla.

Kehittäminen toteutettiin tieteellisenä tutkimushankkeena, jonka tavoitteena oli tutkia monialaisen palvelujen käytön ennustettavuutta. Mallin toiminnallisuudelle tai suorituskyvyille ei asetettu täsmällisiä vaatimuksia. Työpajatyöskentelyn tuloksena tunnistettiin tutkimuksen päätepiteitä, ts. tilanteita, joissa asiakkaalla on runsasta ja monialaista, ennaltaehkäistävässä olevaa palvelujen käyttöä. Lopullisia päätepiteitä ei kuitenkaan lyöty lukkoon vielä tutkimussuunnitelmassa, vaan päätepiteiden lopullinen määrittely jätettiin tehtäväksi tutkimuksen kuluessa.

## 3.2 Tietoaineiston määrittely

Tietoaineiston määrittely on keskeinen, ennustemallin kehittämisen alkuvaiheeseen sijoittuva tehtävä. Tekoälypohjaiset mallit tyypillisesti pyrkivät tunnistamaan ryhmien välisiä eroja päätepiteiden saavuttamisessa (esim. runsas palvelujen käyttö). Ryhmät voivat olla jo ennalta tunnistettuja (supervised learning) tai tunnistaminen voi tapahtua osana mallinnusta (unsupervised machine learning). Yhteistyötä sote-ammattilaisten kanssa tarvitaan sekä ryhmien ennalta tunnistamisessa että koneoppimisen tuottamien tulosten arvioinnissa. Oleellista on tunnistaa tavoitteen kannalta keskeiset tietoaineistot. Niiden määrittelyyn on tärkeää osallistaa riittävä määrä sote-ammattilaisia siten, että tarvittavat osaamisalueet tulevat katetuiksi.

Aineistojen hyödyntämistä rajoittavat usein käytännön seikat – erityisesti datan saatavuus ja hyödynnettävyys tarvittavassa aikataulussa. Siten tietoaineistojen määrittelyssä tarvitaan myös tietojärjestelmä- ja data-asiantuntijoita, jotka pystyvät ottamaan kantaa siihen, miten data on saatavissa asiakas- ja potilasrekistereistä kehitystyön käyttöön.

Sote-organisaatioiden lisäksi tietoaineistojen saatavuudesta saa tietoja Findatalta<sup>5</sup>, joka myös pääsääntöisesti vastaa aineistojen luvituksesta toisilain (Laki sosiaali- ja terveys-tietojen toissijaisesta käytöstä) mukaisesti. Biopankkiaineistojen saatavuustietoja saa Fingenious-palvelusta<sup>6</sup>.

5 <https://findata.fi/aineistot>

6 <https://site.fingenious.fi/en/>



Hyödynnettävien tietojen määrittelyssä on tärkeää arvioida tarkasti, mitkä tiedot ovat oleellisia kehitettävän ratkaisun kannalta. Huomioon otettavia seikkoja ovat mm.:

- **Datan minimointiperiaate (GDPR:n mukaisesti<sup>7</sup>).** Henkilötietoja ei saa ottaa käsiteltävään aineistoon vain siksi, että tiedot ovat saatavilla – tulee olla perusteltu syy sille, että tiedot voivat olla hyödyllisiä mallin kehityksen kannalta ja siksi niitä on tarpeen käsitellä.
- **Aineiston laatu.** Jos jonkin aineiston laatu on puutteellinen (esimerkiksi ei kata tarkasteltavaa aikajaksoa), on huolellisesti arvioitava, onko aineisto hyödynnettävissä mallin kehityksessä ollenkaan. Tällainen aineiston käyttö mallin kehityksessä voi vain huonontaa lopullisen mallin suorituskykyä (”garbage in garbage out”<sup>8</sup>).
- **Aineiston esikäsittelytarve.** Työ aineiston saattamiseksi kehitystyössä hyödynnettävään muotoon voi olla merkittävä esimerkiksi, jos aineisto ei ole valmiiksi rakenteisessa muodossa, vaan vapaana tekstinä.

On suositeltavaa rajata käytettävät data-aineistot sellaisiin, jotka ovat riittävän laadukkaita (esim. eivät sisällä puutteita tai katkoksia) ja jotka ehditään hyödyntämään niillä resursseilla, joita mallin kehitystyöhön on käytettävissä. Data-aineistoa voidaan rajata rajaamalla tarkasteltavia muuttujia, kohdejoukkoa ja tarkasteltavaa ajanjaksoa.

---

7 <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

8 <https://direct.mit.edu/qss/article/2/3/795/102771/Garbage-in-garbage-out-revisited-What-do-machine>



## PÄIJÄT-HÄME

### Päijät-Hämeessä hyödynnetyt tiedot

Päijät-Hämeen kehittämishankkeessa hyödynnetyt tietoaineistot määriteltiin alustavasti VTT:n tutkijoiden toimesta ja kirjattiin tutkimussuunnitelmaan. Tietoaineistoja päätettiin hyödyntää neljän vuoden ajalta (2018–2021). Tältä ajanjaksolta oli käytettävissä yhtenäinen ja kattava aineisto – aiemmat tiedot eivät olleet kaikilta osin yhtenäisesti saatavissa tutkimusta aloitettaessa. Tutkimussuunnitelmaan kirjattiin tutkimuksen inklusiokriteerit, jotka rajaavat poimittavaa data-aineistoa:

- asuinpaikka Päijät-Hämeen hyvinvointialueella ajalla 2018–2021
- syntymävuosi: 1932–1952 (ikä v. 2022: 70–90)
- Päijät-Hämeen hyvinvointialueen palvelujen käyttöä vähintään kerran vuonna 2018
- henkilö on elossa vuoden 2021 lopussa

Tietoaineistojen määrittelyssä hyödynnettiin Päijät-Hämeessä aikaisemmin toteutettua mallinnustyötä (”Ikäihmisten palvelutarpeen ennakointi”) sekä aiempia tutkimuksia (ks. [tutkimussuunnitelma](#)). Tutkimussuunnitelma kierrätettiin tiedoksi ja kommentoitavaksi valituilla erikoisasiantuntijoilla (geriatria, asiakasohjaus, tehohoito). Lisäksi tietoaineistomäärittelyä käsiteltiin useissa kokouksissa Päijät-Hämeen tietojärjestelmä- ja data-asiantuntijoiden kanssa. Näin pyrittiin varmistamaan, että aineistot ovat varmasti saatavissa. Päijät-Hämeessä on hiljattain otettu käyttöön tietoallasratkaisu, joka parantaa mahdollisuuksia tietoaineistojen käyttöön tiedolla johtamisessa, tilastoinnissa ja tutkimuksessa. Tutkimussuunnitelmaan sisällytettiin seuraavat aineistot:

#### Demografiatiedot:

ikä, sukupuoli, asuinpaikka

#### Keskeiset tiedot sote-palvelujen käytöstä, mm.:

- perus- ja erikoissairaanhoidon käynnit ja sairaalajaksot sisältäen diagnoosi-, toimenpide-, lääkitys- ja laboratoriotietoja
- päihdepalvelujen käyttö
- ikääntyneiden palvelujen käyttö
- suun terveydenhoidon käynnit

#### Keskeiset tiedot palvelupäätöksistä, mm.:

- kuntoutus
- asumispalvelut
- omaishoidon tuki
- kotihoito
- kuljetuspalvelut

#### Keskeiset tiedot palvelutarpeen ja toimintakyvyn arvioinneista:

- keskeiset RAI-mittaritiedot

### 3.3 Aineiston hankinta

Aineistoon hankintaan liittyvät järjestelyt riippuvat datan käsittelyn luonteesta suhteessa tietosuojalainsäädäntöön (General Data Protection Regulation, GDPR) ja kansalliseen lakiin sosiaali- ja terveystietojen toissijaisesta käytöstä (”toisiolaki”). Lisäksi on huomattava, että EU-tason lainsäädäntö ja sitä tukeva tekninen infrastruktuuri (European Health Data Space) tulevat valmistuttuaan vaikuttamaan oleellisesti sosiaali- ja terveysdatan hyödyntämiseen<sup>9</sup>.

Voimassa olevaan lainsäädäntöön perustuen kehittämishankkeet jakautuvat karkeasti neljään ryhmään riippuen siitä, luokitellaanko ne tietojohdamiseksi vai tieteelliseksi tutkimukseksi, ja siitä hyödynnetäänkö niissä yhden vai useamman rekisterinpitäjän tietoja (Taulukko 1).

Useissa aiemmissa hyvinvointialueiden hankkeissa dataa on hyödynnetty osana organisaation toiminnan ja palvelujen kehittämistä<sup>10</sup>. Tällöin hyvinvointialue määrää datan hyödyntämisen tavoitteet ja voi itse antaa luvan tietojen käsittelemiseen, jos käsittelyn kohteena on vain ao. hyvinvointialueen omien rekisterien tiedot. Käytännön työn toteuttaa tyypillisesti yritys tai tutkimuslaitos hyvinvointialueen toimeksiannosta. Tutkimuksen toteuttaja on tällöin GDPR:n näkökulmasta tiedon käsitelijän (”data processor”) roolissa. Järjestely on suosittu, koska pääsy dataan voidaan tyypillisesti järjestää nopeasti ilman Findatan lupahakemuskäsittelyä, eikä tietoja tarvitse viedä erilliseen sertifioituun käyttöympäristöön<sup>11</sup>.

Mallin kehitystyö on tarkoituksenmukaista luokitella tieteelliseksi tutkimukseksi, jos kyseessä on uudenlainen ennustemalli ja kehitystyön tulokset halutaan julkaista esim. tieteellisenä artikkelina. Tämä on tärkeää, jotta esitettyä mallinnustapaa voidaan vuoro-vaikutteisesti edelleen kehittää tiedeyhteisössä. Luokitteluun voi vaikuttaa myös hankkeen rahoitusmalli. Esimerkiksi yhteisrahoitteiset tutkimushankkeet luokitellaan tieteelliseksi tutkimukseksi. Myös tieteellisen tutkimuksen tapauksessa hyvinvointialue voi rekisterinpitäjänä myöntää luvan tietojen käyttöön, mutta tiedot on toisiolain mukaisesti vietävä sertifioituun käyttöympäristöön tutkimuksen toteuttamista varten.

9 [https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space\\_fi](https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_fi)

10 [Esiselvitys toisiolain vaikutuksista tiedolla johtamiseen](#)

11 <https://www.valvira.fi/terveydenhuolto/toisiolain-mukaiset-tietoturvalliset-kayttoymparistot/toisiokayttoymparistojen-rekisteri>

**Taulukko 1.** Henkilötietojen luvitus tietojohdantamiseen ja tieteelliseen tutkimukseen

	Tietojohdantaminen	Tieteellinen tutkimus
Rekisterinpitäjän omien tietojen käyttö	Sosiaali- ja terveydenhuollon palvelunantajalla on oikeus käyttää henkilötietoja omalla vastuullaan olevan palvelutoiminnan toteuttamiseksi. Tietoja voidaan käsitellä rekisterinpitäjän omassa käyttöympäristössä.	Rekisterinpitäjä voi itse myöntää tietoluvan tutkimuksen toteuttajalle.  Tietoja tulee käsitellä Findatan määräyksen mukaisesti sertifioitussa käyttöympäristössä
Sote-tietojen laajempi käyttö (useita rekisterinpitäjiä)	Findata voi tietopyynnön perusteella tuottaa aggregoidut tiedot vertailuaineistoksi sosiaali- ja terveydenhuollon palvelunantajalle.	Findata myöntää tietoluvan.  Tietoja tulee käsitellä Findatan määräyksen mukaisesti sertifioitussa käyttöympäristössä.

Kun datan käsittelyn luonne ja laajuus on taulukon 1 mukaisesti valittu, edetään luvan hakemisen ja sopimusten tekemiseen.

Lupaa haetaan rekisterinpitäjältä tai Findatalta ao. organisaation käytössä olevalla hakemuskaavakkeella. Tieteellisen tutkimuksen tapauksessa oleellinen osa tietolupahakemusta on hyvän tieteellisen käytännön mukainen tutkimussuunnitelma. Tähän sisältyy mm. käytettävien tutkimusmenetelmien ja tietoaaineistojen määrittely.

Hyväksytyt tietoluvan lisäksi tarvitaan osapuolten välisiä sopimuksia ja riskien hallintaan liittyviä dokumentteja, esim.:

- rekisterinpitäjän ja henkilötietojen käsittelijän välinen sopimus (Data Processing Agreement, DPA), esim. jos yritys/tutkimuslaitos kehittää mallia hyvinvointialueen toimeksiannosta
- yhteisrekisterinpitäjien välinen sopimus, esim. yhteisrahoitteisissa tutkimushankkeissa, joissa tutkimusrekisteri on toimijoiden yhteinen
- vaikutusten arviointi (Data Protection Impact Assessment, DPIA)
- tietosuojaseloste
- seloste henkilötietojen käsittelytoimista



## PÄIJÄT-HÄME

### Datan käyttöön liittyvät järjestelyt Päijät-Hämeessä

Ennustemallin kehittämistyön todettiin olevan luonteeltaan tieteellinen tutkimus, jossa hyödynnetään yhden rekisterinpitäjän tietoaainestoa. Tietojen käsittelyä varten perustettiin Päijät-Hämeen hyvinvointialueen, THL:n ja VTT:n yhteinen tutkimusrekisteri, johon liittyvät roolit ja vastuut määriteltiin osapuolten allekirjoittamassa yhteisrekisterinpitäjäyysopimuksessa, sisältäen liitteinä vaikutusten arvioinnin, tietosuojaselosteen ja selosteen käsittelytoimista.

Tietojen käsittelyn tuli tapahtua sertifioidussa tietoturvalisessa käyttöympäristössä – Päijät-Hämeen kehittämishankkeessa päätettiin käyttää oletusarvoista vaihtoehtoa: Findatan Kapseli-ympäristöä (kaikki toisilain mukaisesti sertifioidut käyttöympäristöt on kuvattu Valviran sivustossa)<sup>12</sup>. Tällöin yhteinen tutkimusrekisteri muodostui Kapseliin.

Tietolupa haettiin suoraan Päijät-Hämeen hyvinvointialueelta. Lupa saatiin nopeasti (alle kahdessa viikossa). Nopeaan käsittelyyn todennäköisesti vaikutti se, että hyvinvointialueen edustajat olivat olleet kiinteästi mukana tutkimuksen valmistelussa. Näin tutkimus oli hyvin Päijät-Hämeen tiedossa jo ennalta ja hyvinvointialueen asiantuntijat olivat päässeet vaikuttamaan sen sisältöön.

12 <https://www.valvira.fi/terveydenhuolto/toisilain-mukaiset-tietoturvaliset-kayttoymparistot/toisiokayttoymparistojen-rekisteri>

## 3.4 Aineiston poiminta ja esikäsittely

Aineiston poiminnalla viitataan datan hakuun lähdejärjestelmistä sen jälkeen, kun lupa aineiston käyttöön on saatu. Viime vuosien aikana sote-organisaatiot ovat toteuttaneet datan hyödyntämistä tukevia tietoaaltaita, joihin operatiivisten järjestelmien tiedot siirtyvät automaattisesti. Tietoaaltaiin on myös kytketty datan käsittely-ympäristöjä, joista on suora pääsy tietoaaltaan tietoihin, mm. tietojohdamiseen liittyvässä kehittämisessä. Tällöin ei välttämättä tarvita erillistä poimintaa, vaan data pysyy tietoaaltaassa käsittelyn aikana. Tutkimushankkeissa tiedot viedään toisilain mukaisesti tietoturvaliseen käsittely-ympäristöön (esim. Findata/Kapseli), jolloin tietojen poiminta tarvitaan. Tässäkin tapauksessa tietoaallas voi merkittävästi helpottaa tietojen löytämistä ja tutkimusaineiston muodostamista.

Esikäsittely sisältää työvaiheet, joissa aineiston laatu tarkastetaan sekä saatetaan sopivaan muotoon analyysiä varten. Tarkastuksella pyritään varmistamaan, että kohdejoukosta on käytettävissä yhdenmukainen kokoelma muuttujia kattaen tarkasteltavan ajanjakso (tai kohdejoukko jakautuu yhdenmukaisesti alijoukkoihin, jotka analysoidaan erikseen). Lisäksi pyritään löytämään aineistossa olevia virheitä, esim. yksittäisiä puuttuvia arvoja tai saman tiedon esiintymistä eri yksiköissä.



Puutteita voi aiheutua, mm. siitä, että tietoja joudutaan usein poimimaan eri lähteistä, jolloin on mahdollista, että kaikki tarvittavat lähteet eivät ole olleet mukana poiminnassa. Em. puutteiden löytämiseksi aineistoa kannattaa tarkastella ajallisesti ja varmistaa, että keskimääräiset vuosittaiset arvot (esim. palvelujen käyttöön liittyen) ovat loogisia. Myös vertailut julkisesti saatavilla olevaan tilastotietoon voivat paljastaa aineiston puutteita. Jos aineistossa on puutteita, ne pyritään korjaamaan hankkimalla puuttuvat tiedot, tai jos se ei ole mahdollista käsittelemällä puuttuvat tiedot asianmukaisella tavalla analyysissä.

Lisäksi aineistolle voi olla tarpeen tehdä muunnoksia tietoaineistojen yhdistämiseksi ja saattamiseksi sopivaan muotoon mallinnusta varten. Erityisesti, jos hankkeessa hyödynnetään eri sote-alueiden tietoja, voidaan joutua toteuttamaan erilaisia muunnoksia aineistojen saattamiseksi yhdenmukaisiksi.



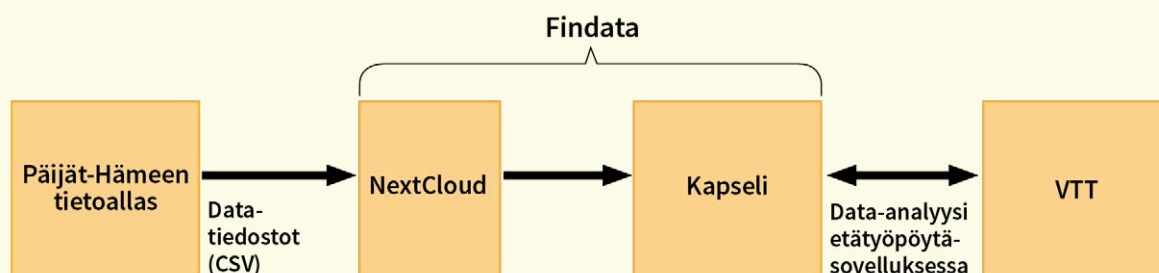
## PÄIJÄT-HÄME

### Aineiston poiminta ja esikäsittely Päijät-Hämeessä

Päijät-Hämeen hyvinvointialue toteutti tietojen poiminnan tietoaltaasta sekä siirron Kapseliin hyödynnäen Findatan NextCloud -palvelua. Vaihe oli suunniteltua työläämpi, koska ennakoitua suurempi osa tiedoista ei ollut valmiina mallinnettuna Päijät-Hämeen tietoaltaassa tai tiedot eivät olleet siellä tutkimuskäyttöön soveltuvassa muodossa. Tutkimussuunnitelmaan kootun tietokuvauksen pohjalta VTT laati ensin ehdotuksen tietorakenteista tiedon poimintaa varten. Ehdotus käsiteltiin ja sitä täydennettiin yhdessä Päijät-Hämeen data-asiantuntijoiden kanssa. Tämän jälkeen Päijät-Hämeen hyvinvointialue toteutti puuttuvat tietoallasintegraatiot ja niihin liittyvän mallinnustyön sekä laati tietorakennemäärittelyä vastaavat näkymät tietoaltaaseen. Mm. sote-uudistukseen liittyvien kiireiden vuoksi edellä mainittuihin tehtäviin kului aikaa n. 3 kk. Tämän jälkeen varsinainen tietojen haku ja siirto Kapseliin onnistui nopeasti Findatan ohjeistuksen mukaan.

Esikäsittelyn yhteydessä havaittiin joitakin puutteita. Alkuperäiseen tutkimussuunnitelmaan ei ollut sisällytetty kuolintietoa, eikä tietoa case manager -asiakkuudesta. Nämä tiedot päivitettiin suunnitelmaan. Päijät-Hämeen hyväksyttyä tutkimussuunnitelman muutoksen aineisto täydennettiin em. tiedoilla.

Kapseli on suljettu tietoturvallinen käyttöympäristö. Findata kontrolloi ympäristöön vietyjä ja sieltä pois tuotuja tietoja. Oheisen kuvion mukaisesti datan omistaja (Päijät-Häme) vie tutkittavan datan NextCloudin kautta Findatan tarkistusprosessiin, josta data siirtyy Kapseliin. Vastaavan prosessin kautta vietiin Kapseliin tutkimuksen tekijän tarvitsemat analyysiohjelmat ja aputiedostot.





## 3.5 Mallin toteutus ja testaus

Koneoppiminen tarjoaa menetelmiä datalähtöisten eli aiempiin havaintoihin perustuvien ennustemallien rakentamiseen. Koneoppimisalgoritmit pyrkivät oppimaan niille esitetyn data-aineiston perusteella aineiston ja ennustettavan muuttujan välistä suhdetta kuvaavan mallin. Oppimismenetelmät voidaan jakaa esimerkiksi ohjattuun (engl. supervised) ja ohjaamattomaan (engl. unsupervised) oppimiseen sekä vahvistusoppimiseen (engl. reinforcement learning). Tarkasti määritellyn ennustetavoitteen tapauksessa on usein kyse ohjatusta oppimisesta. Ohjatussa oppimisessa algoritmi vaatii, että koulutusta varten syötettyihin esimerkkeihin liittyy tunnettu tulos (esim. luokka tai jokin jatkuva muuttuja). Ohjaamattomassa oppimisessä ja vahvistusoppimisessä ei sen sijaan tarvita tunnettua tulosta, vaan algoritmit pyrkivät löytämään esimerkeistä sisäsyntyisiä riippuvuuksia tai oppimista ohjataan palkkio- tai rangaistusfunktioilla.

Ennustetavoite tai tutkimuskysymyksen asettelu (yllä mainitut vaihtoehdot huomioiden) määrittelee usein oppimisen lähtökohdan ja siihen sopivat mallit. Koneoppimismalleja itsessään voidaan ryhmitellä esimerkiksi lineaarisiin, tilastollisiin, yhdistelmä- (engl. ensemble) tai neuroverkko-malleihin. Mallin valintaan vaikuttaa päätepisteen tyyppin lisäksi koulutusaineiston koko, ennustettavan päätepisteen esiintyvyys aineistossa, laskentaresurssien saatavuus, syöteaineiston laajuus ja tyyppi. Esimerkiksi tietyille aineistotyypeille, kuten kuville tai tekstille, löytyy erityisesti niille kehiteltyjä neuroverkko-malleja.

Neuroverkot koostuvat ihmisen biologisia hermosoluja jäljittelevistä neuroneista, jotka aktivoituvat riittävästä ”ärsykkeestä” opitun aktivointifunktion mukaisesti. Syvät neuroverkot koostuvat useista kerroksista tällaisia neuronijoukkoja, joita optimoidaan iteroivasti: koulutusaineisto syötetään verkkoon toistuvasti useassa erässä ja verkon virhettä pyritään korjaamaan jokaisen erän jälkeen muokkaamalla neuronien parametreja. Syvien neuroverkkojen kerrokset pystyvät oppimaan syöteaineistosta piirteitä itseohjautuvasti (representation learning), mikä on usein hyödyllistä, jos aineistosta ei voida tieteellisin perustein valita riittävän hyvää joukkoa piirteitä.

Syvät neuroverkot ja syväoppiminen (engl. deep learning) ovat osa tekoälyn laajempaa konseptia (engl. artificial intelligence, AI). Toisaalta syvät neuroverkot voivat kerrosten lukumäärästä ja syöteaineiston laajuudesta riippuen vaatia paitsi merkittävän suurta aineistoa, myös hyvinkin tehokkaita laskentaresursseja. Lisäksi syviä neuroverkkoja on kritisoitu niiden vaikeasta tulkittavuudesta, vaikkakin viime vuosien aikana attention-menetelmät ovat tuoneet tulkittavuutta ja myös valmiiden mallien tulkitsemiseen on kehitetty uusia menetelmiä (esim. SHAP ja LIME).

Yksinkertaisemmat koneoppimismallit, jotka eivät käytä neuroneja, ovat laskennallisesti kevyempiä ja helposti ymmärrettäviä. Yksinkertaisuutensa vuoksi niitä pidetään tyyppillisesti neuroverkkoja luotettavampina. Ne vaativat kuitenkin tyyppillisesti enemmän resursseja manuaaliseen piirteiden valintaan ja piirrevalinnan onnistuminen vaikuttaa vahvasti näiden mallien tarkkuuteen. Mallin valinta vaikuttaakin vahvasti myös aineiston esikäsittelyyn.

Mallin selitettävyys ja ymmärrettävyys ovat tärkeitä ominaisuuksia erityisesti lääketieteellisessä kontekstissa, jossa malleja voidaan käyttää päätöksenteon tukena ja

vaikuttaa siten potilaiden terveyteen. Ennustemallin käytettävyyteen vaikuttaa myös ennusteen käyttökelpoisuus: mallin tulisi olla helposti saatavilla ja ennuste saada riittävän ajoissa, jotta sen perusteella voidaan toimia (actionable AI). Mallia kehitettäessä tulisi kiinnittää erityistä huomiota koulutusaineistoon: onko aineisto edustava ja riittävän monipuolinen ja laaja otos, joka tarjoaa mallille riittävän pohjan ja onko todellista käyttötilannetta heijastava koulutusasetelma mahdollinen. Lisäksi mallin validoinnissa tulisi erityisesti arvioida mallin toimivuutta uudella datalla ja sen toimintavarmuuden tasapuolisuutta eri potilasryhmissä (kuten sukupuolien välillä ja eri etnisissä ryhmissä) (Van Smeden 2022). Tyypillisesti sosiaalihuollon ja sairaaloiden tietokannat Suomessa ovat pitkän aikavälin otoksia, mutta rajatulta alueelta ja suhteellisen yksipuolisesta populaatiosta, jolloin malli ei välttämättä ole yhtä toimiva vähemmistöille.

Aineisto on luonnollisesti suuressa roolissa kaikissa datalähtöisissä menetelmissä. Koneoppimisessa aineisto jaetaan tyypillisesti koulutusvaiheen aineistoon (koulutus- ja validointisetit) sekä testiaineistoon, jota ei käytetä koulutusvaiheessa lainkaan. Kehitysvaiheessa voidaan tarvittaessa käyttää ristivalidointia, jossa mallin koulutus ja validointi suoritetaan iteratiivisesti siten, että validointiin käytettävä osajoukko aineistosta vaihtuu kullakin toistolla. Kunkin aineistoryhmän tulisi koostua monipuolisista havainnoista. Mikäli aineisto koostuu esimerkiksi useasta samaan yksilöön liittyvästä havainnosta, tulisi kaikki samaan yksilöön liittyvät havainnot olla vain yhdessä aineistoryhmässä, jotta validointi- tai testitulokset ei vääristyisi. Aineistoissa, joiden ennustetavoite on luokitteleva ja ennustettavia (positiivisia) tapauksia ei ole yhtä monta kuin muita (negatiivisia) tapauksia, tulee luokkien välinen epätasapaino ottaa huomioon. Luokkien osuudet voidaan tarvittaessa tasapainottaa (esimerkiksi laajentamalla aineistoa keino-tekoisesti tai valitsemalla vain sattumanvarainen osajoukko negatiivisia tapauksia) tai joissain tapauksissa muokata mallin sisäisiä painotuksia. Lisäksi mallin arvioimiseen on hyvä käyttää monipuolisia mittareita, jotka mittaavat ennustemallin suoriutumista kaikkien luokkien näkökulmasta.

Kun syötetiedot muodostavat aikasarjan, tuottaa se omat haasteensa mallin toteutukseen. Valitusta koneoppimismallista riippuen aineisto voidaan muun muassa syöttää numeerisena aikasarjana tietyille neuroverkoille tai esikäsitellä siten, että aikariippuvat muuttujat esitetään tiivistettyinä muutaman piirteen avulla (tilastollisten ja taajuustason muuttujien kautta). Aikasarjojen tapauksessa mallin ennustepiste tulisi lisäksi valita huolellisesti siten, että opetusaineiston ennustepiste vastaa ennustepistettä todellisessa käytössä. Jos mallia voidaan käyttää missä tahansa vaiheessa todellisen asiakkaan aikasarjaa, tulisi koulutusaineistossakin ennustepisteen olla ”missä tahansa” vaiheessa, eikä esimerkiksi aina tietyllä hetkellä päätetapahtuman ajankohtaan nähden.

Sosiaali- ja terveydenhuollon tietokannoissa lisähaasteen tuo myös tietojen heterogeenisyys: aineisto koostuu lukuarvoista, koodeista, kategorisista muuttujista, tekstistä ja kuvista. Tekstimuotoinen aineisto on kuitenkin mahdollista muokata numeeriseksi vektorimuotoon (ks. one-hot encoding, word embedding) tai tokenisoida etukäteen toisen algoritmin avulla muodostetun ”sanakirjan” mukaan, jossa toisiaan lähellä olevat numeeriset esitysmuodot viittaavat saman tyyppisiin tekstimuotoisiin sanoihin. Toisinaan voi olla tehokkainta kouluttaa mallit erikseen erityyppisille datalähteille (esimerkiksi kuville ja muulle datalle) ja muodostaa näistä yhdistelmämalli.



## PÄIJÄT-HÄME

### Ennustemallin toteutus Päijät-Hämeessä

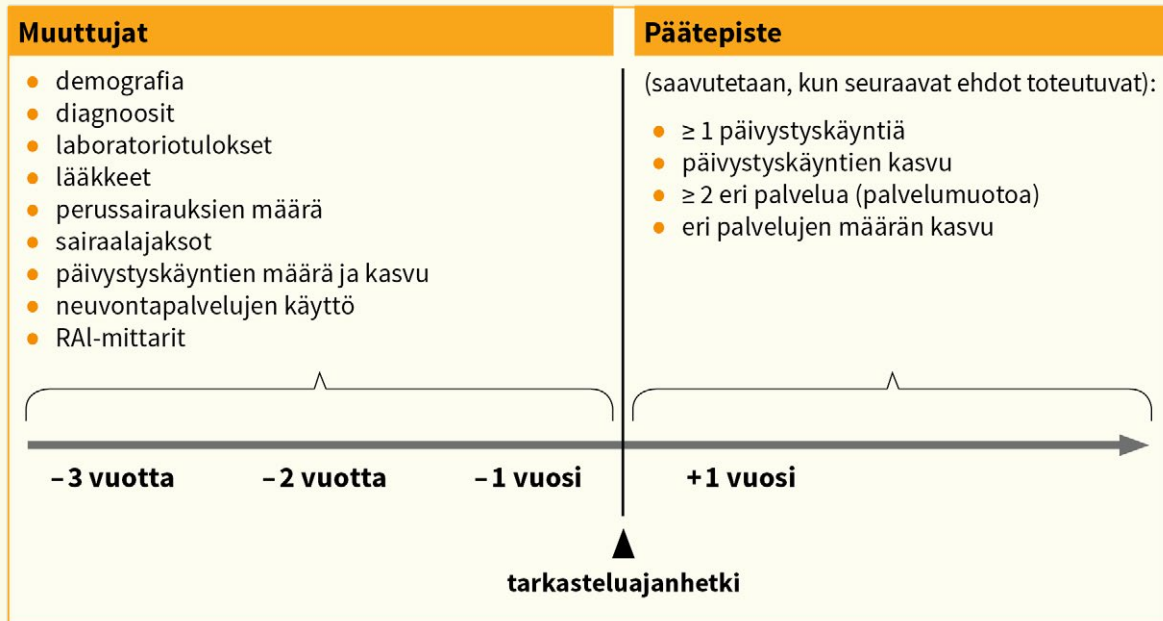
Ennustemalli hyödyntää monipuolista tietoa asiakkaiden terveydestä sekä aiemmasta sosiaali- ja terveydenhuollon palvelujen käytöstä. Mallinnuksen haasteena on se, että ikääntyneiden palvelutarve vaihtelee runsaasti ja riippuu mm. henkilön perussairauksista. Jo aikaisessa vaiheessa todettiin, että mallinnus tulee toteuttaa ryhmittäin siten, että voidaan tunnistaa kussakin ryhmässä ne tekijät, jotka ennustavat riskiä runsaaseen ja monialaiseen palvelukäyttöön. Tutkimusasetelmaa käsiteltiin Päijät-Hämeen sote-ammattilaisten kanssa työpajamuotoisesti. Tällöin tunnistettiin asiakasryhmiä (ks. Luku 2), joille on ominaista runsas monialaisten sote-palvelujen käyttö. Näiden perusteella pyrittiin määrittelemään tarkasteluryhmiä, jotka ovat yksikäsitteisesti muodostettavissa aineiston perusteella. Ryhmät muodostettiin ensisijaisesti diagnoositietojen perusteella. Ensimmäisten tulosten esittelyn yhteydessä nousi esille myös tarve tarkastella erillisenä ryhmänä ikääntyneiden palvelujen asiakkaita sekä kotihoitopalvelujen asiakkaita edellisen alaryhmänä. Nämä määriteltiin palvelukäyntien palvelumuototiedon perusteella. Näin saatiin kuusi osittain päällekkäistä tarkasteluryhmää:

- Muistisairaus
- Diabetes
- Mielenterveys- tai päihdeongelma
- Ei perussairautta
- Ikääntyneiden palvelujen asiakkuus
- Kotihoitopalvelujen asiakkuus

Päätepiste määriteltiin perustuen päivystyskäyntien määrään ja kasvuun sekä käytössä olevien eri palvelujen määrään ja kasvuun. Päätepiste laskettiin hyödyntämällä käyntitietoihin sisältyvää palvelumuototietoa. Käytettävissä olevan aineiston perusteella ei ollut mahdollista suoraan huomioida päätepuoleen määrittelyssä sitä, onko ao. palvelukäyttö ollut vältettävissä (ks. alaluku 2.3).

Toteutettu malli ennustaa tiettyyn tarkasteluryhmään kuuluvalle asiakkaalle riskin päätepuoleen saavuttamiseen, ts. runsaaseen ja kasvavaan erilaisten sote-palvelujen käyttöön, tarkasteluajanhetkeen nähden seuraavana vuonna. Syöttötietona malli käyttää henkilön taustatietoja ja sote-palvelujen käyttötietoja edellisen kolmen vuoden ajalta (kuva seuraavalla sivulla). Ennustemallissa hyödynnettiin 26 muuttujaa edustaen ala olevassa kuvassa mainittuja ryhmiä. Koko muuttujajoukkoa ei kuitenkaan ollut mahdollista hyödyntää kaikkien tarkasteluryhmien kohdalla. Esimerkiksi RAI-mittaritieto oli kattavasti saatavilla vain kotihoitopalvelujen asiakkailta.





### Ennustemallin muuttujat ja (ei toivottu) päätepiste.

Mallinnus toteutettiin erikseen kullekin ryhmälle hyödyntäen kahta eri menetelmää: logistista regressiota ja gradienttitehostettua päätöspuuta (XGBoost algoritmi). Mallin opetukseen käytettiin 70 % käytettävissä olevasta aineistosta. Loput 30 % käytettiin mallin testaukseen.

Alustavien analyysitulosten perusteella voidaan todeta, että mallin luokittelukyky jää matalalle tasolle. Luokittelukykyä kuvaava AUC-arvo on suuruusluokassa 0.6 molemmilla menetelmillä. Erityisesti tarkasteluryhmissä, joissa käytettävissä oleva aineiston koko oli pieni (mm. muistisairaat), AUC-arvo jäi alhaiseksi. Kaikkien tarkasteluryhmien kohdalla henkilön ikä sekä perussairauksien ja aiempien päivystyskäyntien määrä erottuvat odotetusti päätepestettä ennustavina tekijöinä. Muut merkittävät tekijät liittyivät mm. lääkkeiden käyttöön ja asuinpaikkaan ja niiden vaikutukset riskiin vaihtelivat tarkasteluryhmittäin.

Tuloksia arvioitaessa on huomattava, että vastaavia ennustemalleja on tutkittu vasta vähän ja myös aikaisemmissa tutkimuksissa luokittelutarkkuus on ollut rajallinen (Bardsley 2011, Nakubulwa 2022). Hyödynnettäessä ennustemallia ennakoivien palvelujen kohdentamiseen jo alhainenkin luokittelutarkkuus voi kuitenkin olla hyväksyttävä ja tuoda parannusta nykytilanteeseen.

Yksityiskohtainen ennustemallin kuvaus sekä tulokset mallin luokittelukykyvystä ja riskiin vaikuttavista tekijöistä julkaistaan erikseen.

## 3.6 Operatiivisen käytön huomioiminen

Tekoälyyn pohjautuvan ratkaisun kehittäminen on usein kokeiluluontoista ja tapahtuu esimerkiksi tutkimushankkeen yhteydessä. Tällaisen hankkeen tavoitteena on tyypillisesti tuottaa proof-of-concept (PoC) ohjelmisto, jonka pohjalta operatiivinen ratkaisu voidaan toteuttaa myöhemmin. Tutkimus- tai kokeiluhankkeessa on huomioitava tämä lopullinen tavoite niin, että PoC versiosta voidaan mahdollisimman suoraviivaisesti edetä tuotantokäyttöön.

Tulevaa operatiivista käyttöä ajatellen tulee jo hankkeen suunnitteluvaiheessa huomioida mm. seuraavat kysymykset:

*Onko ratkaisu voimassa olevan lainsäädännön mukainen? Esimerkiksi ovatko GDPR:n vaatimukset koskien rekisteröidyn oikeuksia täytettävissä? Tekoälyratkaisujen kannalta oleellinen on mm. rekisteröidyn oikeus olla joutumatta automaattisen päätöksenteon kohteeksi ilman nimenomaista suostumusta.*

*Onko lopullinen ratkaisu lääkintälaitelainsäädännön mukainen lääkinnällinen laite, jolloin sitä koskevat ao. sertifiointivaatimukset?*

*Onko ratkaisu asiakastietolain tarkoittama tietojärjestelmä tai sellaisen osa, jolloin sen on täytettävä THL:n sosiaali- ja terveydenhuollon tietojärjestelmiä koskevat olennaiset vaatimukset?*

*(Lähteenmäki ym. 2022)*

Tällä hetkellä voimassa olevan lainsäädännön lisäksi on tärkeää huomioida käynnissä oleva kansallinen ja EU-tason lainsäädäntötyö (mm. EU:n tekoälylainsäädäntö), johon sisältyy uusia vaatimuksia erityisesti koskien datan ja tekoälyn hyödyntämistä.

Lääkintälaitelainsäädännön näkökulmasta on oleellista, että kehitystyön aikana syntyy tarvittava dokumentaatio, jonka avulla mallin kehittämisen vaiheet voidaan tarvittaessa toistaa ja tarkistaa. Kerättäviin jäljitettävyysetietoihin sisältyy mm. ohjelmiston versiohistoria, koneoppimiseen käytetty menetelmä ja siihen liittyvät parametrit sekä tiedot mallin opettamiseen ja testaukseen käytetyistä tietoaineistoista. Regulaation lisäksi tulosten toistettavuus kuuluu myös hyvään tutkimuskäytäntöön.

Asiakastietolakiin liittyvät olennaiset vaatimukset koskevat erityisesti ratkaisun tietoturvasuutta sekä yhteensopivuuteen Kanta-palvelun kanssa. Lainsäädäntöön liittyviä kysymyksiä tarkastellaan yksityiskohtaisemmin luvussa 5.

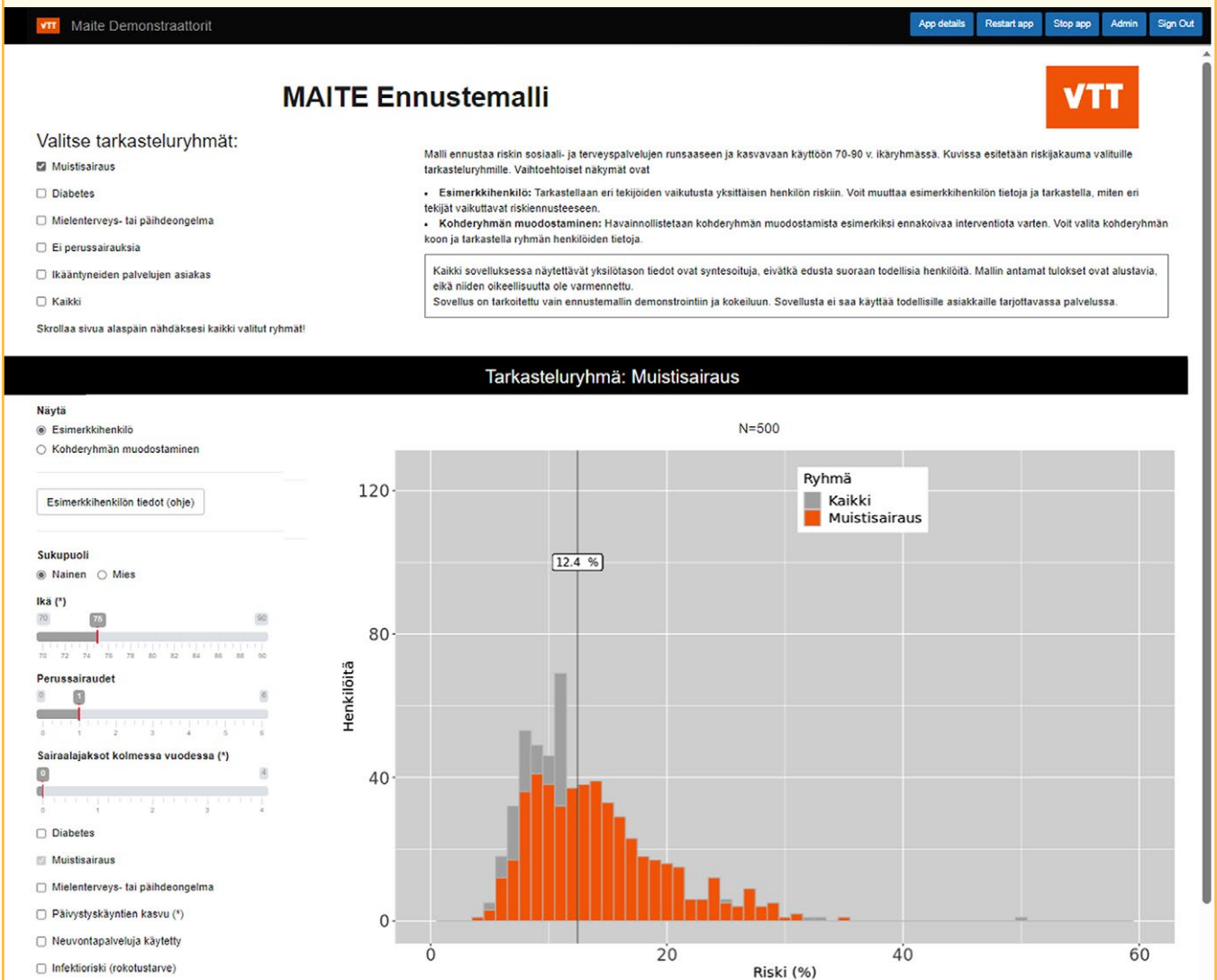


**PÄIJÄT-HÄME**

## Operatiivisen käytön huomiointi Päijät-Hämeessä

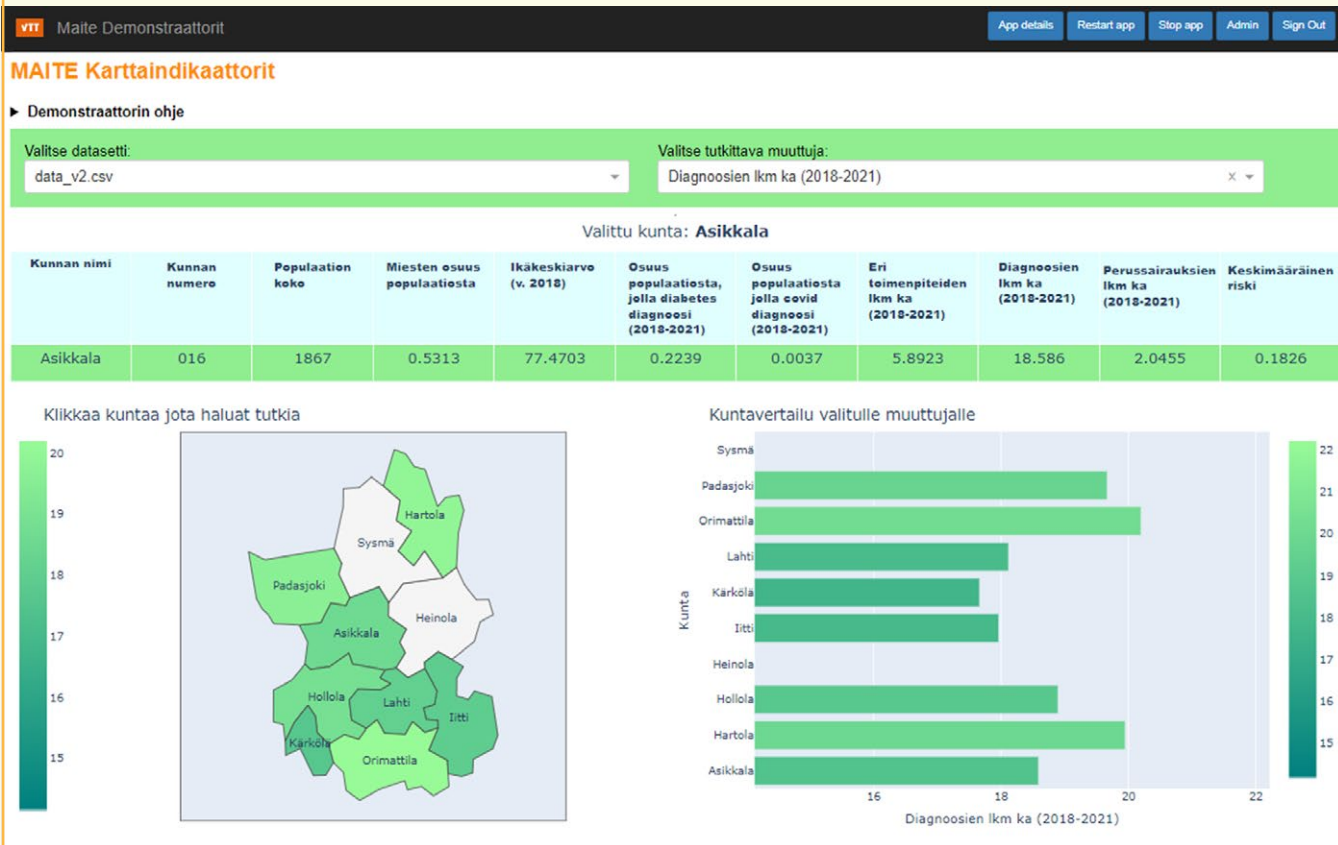
Tuleva operatiivinen käyttö huomioitiin erityisesti kolmesta näkökulmasta: (1) hyödynnetyt data-aineistot, (2) data-analyysien toistettavuus, (3) aikaisen vaiheen demonstrointi sote-organisaation potentiaalisille käyttäjille. Data-aineistojen osalta pitäydettiin niihin tietoihin, jotka olivat käytävissä hyvinvointialueen tietoaltaassa. Näin myös operatiivisen palvelun toteuttaminen olisi mahdollista hyödyntämällä Päijät-Hämeen omien rekisterien tietoja, eikä olisi tarvetta saada käyttöön esimerkiksi kansallisissa rekistereissä olevia tietoja. Data-analyysien toistettavuudesta huolehdittiin tallentamalla kattavat lokitiedot koko tiedonkäsittelyprosessista kattaen tietoaltaasta poimitun tiedon esikäsittelyn, viennin tutkimustietokantaan, ennustemallin kehittämisen ja testauksen.

Tutkimushankkeeseen sisällytettiin PoC-ohjelmiston toteutus ("MAITE demo"). MAITE demo mahdollistaa ennustemallin esittelyn ja aikaisen käyttäjäpalautteen, mm. liittyen siihen, miten malli parhaiten palvelisi sote-ammattilaisia käytännön työssä (kuva alla).



MAITE demo on selainpohjainen sovellus, jonka avulla voi tarkastella asiakasryhmien riskijakaumaa ja riskeihin vaikuttavia tekijöitä sekä havainnollistaa kohderyhmien valintaa esimerkiksi ennakoivia interventiota varten.

Lisäksi ohjelmiston avulla voidaan tutkia mallinnuksessa käytetyn aineiston muuttujia kunnittain karttapohjalla (kuva alla).



## 3.7 Käyttäjäkokeilu

Positiivinen käyttäjäkokemus ja koettu hyöty ovat keskeisiä tekijöitä uusien innovaatioiden käyttöönoton kannalta. Käyttäjät tuleekin saada mukaan kehittämisprosessiin jo kehitystyön alusta lähtien. Suunnittelu- ja määrittelyvaiheessa käyttäjien osallistaminen tapahtuu esimerkiksi työpajoissa ja haastattelujen avulla. Mahdollisimman aikaisessa vaiheessa tulisi lisäksi esitellä alustavia käyttöliittymädemonstraatiota, joilla havainnollistetaan ratkaisun hyödyntämistä sote-organisaatiossa.

Käyttäjäkokeilulla viitataan tässä laajasti erilaisiin järjestelyihin, joilla ratkaisun käyttäjät (ts. sosiaali- ja terveydenhuollon työntekijät) pääsevät kokeilemaan ennustemallia. Käyttäjäkokeilun laajuus riippuu projektin resursseista ja tavoitteista. Pienimuotoinen kokeilu toteutetaan tyypillisesti tilaisuutena, jossa ennustemalli demonstroidaan PoC sovelluksen avulla ja koekäyttäjät saavat lyhyen opastuksen sen käyttöön. Käyttäjät kokeilevat sitten ennustemallin toimintoja sekä lopuksi antavat palautetta ratkaisun toiminnallisuudesta ja toimivuudesta. Laajemmissa käyttäjäkokeiluissa käyttäjät saavat perusteellisemmän opastuksen ja saavat sitten sovelluksen itsenäiseen käyttöön pidemmäksi aikaa. Käyttäjäpalaute voidaan antaa suullisesti pienimuotoisen kokeilun yhteydessä. Varsinkin laajemmissa käyttäjäkokeiluissa palaute kerätään lomakekyselyllä. Jos käyttäjäkokeilun yhteydessä tallennetaan käyttäjien henkilötietoja, tulee huomioida tietosuojalainsäädännön (GDPR) edellyttämät toimenpiteet.



### PÄIJÄT-HÄME

#### Käyttäjäkokeilu Päijät-Hämeessä

Päijät-Hämeen pienimuotoinen demonstraatio/käyttäjäkokeilu toteutettiin MAITE demon avulla (ks. alaluku 3.6). MAITE demoon liittyvä taustainformaatio ja käytön opastus Päijät-Hämeen henkilöstön edustajille annettiin kahdessa verkkokokouksessa. Ensimmäiseen osallistui Päijät-Hämeen tietohallinnon edustajia ja toiseen asiakaspalvelujen edustajia. Kokouksissa käytiin aluksi läpi ennustemallin periaate. Tavoitteena oli antaa realistinen kuva siitä, mitä ennustemalli ennustaa ja mihin ennusteet perustuvat. Sen jälkeen esiteltiin MAITE demo jakautuen kahteen osaan: muuttujien tutkiminen (karttademo) ja ennustetulosten tutkiminen (ennustemallidemo). Kokousten päätteeksi annettiin käytännön ohjeet MAITE demoon kirjautumiseen, käyttötuen saamiseen sekä palautteen antamiseen sähköisellä lomakkeella. Päijät-Hämeen puolelta todettiin, että hyvinvointialueen kiireiden vuoksi, demonstraation itsenäiseen kokeiluun ei kaikilla välttämättä ole mahdollisuutta. Tunnukset toimitettiin kymmenelle henkilölle.

MAITE demon esittelyn ja itsenäisen kokeilun yhteydessä saatujen ensimmäisten palautteiden perusteella muuttujien tarkastelu karttapohjalla on havainnollinen ja ymmärrettävä esitystapa. Samoin kohderyhmän muodostaminen, jossa näytetään lista (virtuaalisista) riskihenkilöistä. Sen sijaan populaation riskiä kuvaava jakauma koetaan vain rajallisesti ymmärrettävänä ja sen korvaamista vaihtoehtoisella esitystavalla tulisi pohtia mahdollisen jatkokehittämisen yhteydessä. Palautteen perusteella ennustemalli voisi olla hyödyllinen ikääntyneiden palveluohjauksessa ja seulonnassa ennakoivia interventioita varten edellyttäen, että sovellus on todettu lainsäädännön mukaiseksi ja integroitu olemaan olevaan sovellusympäristöön. MAITE demo on Päijät-Hämeen käytettävissä MAITE-hankkeen loppuun saakka.



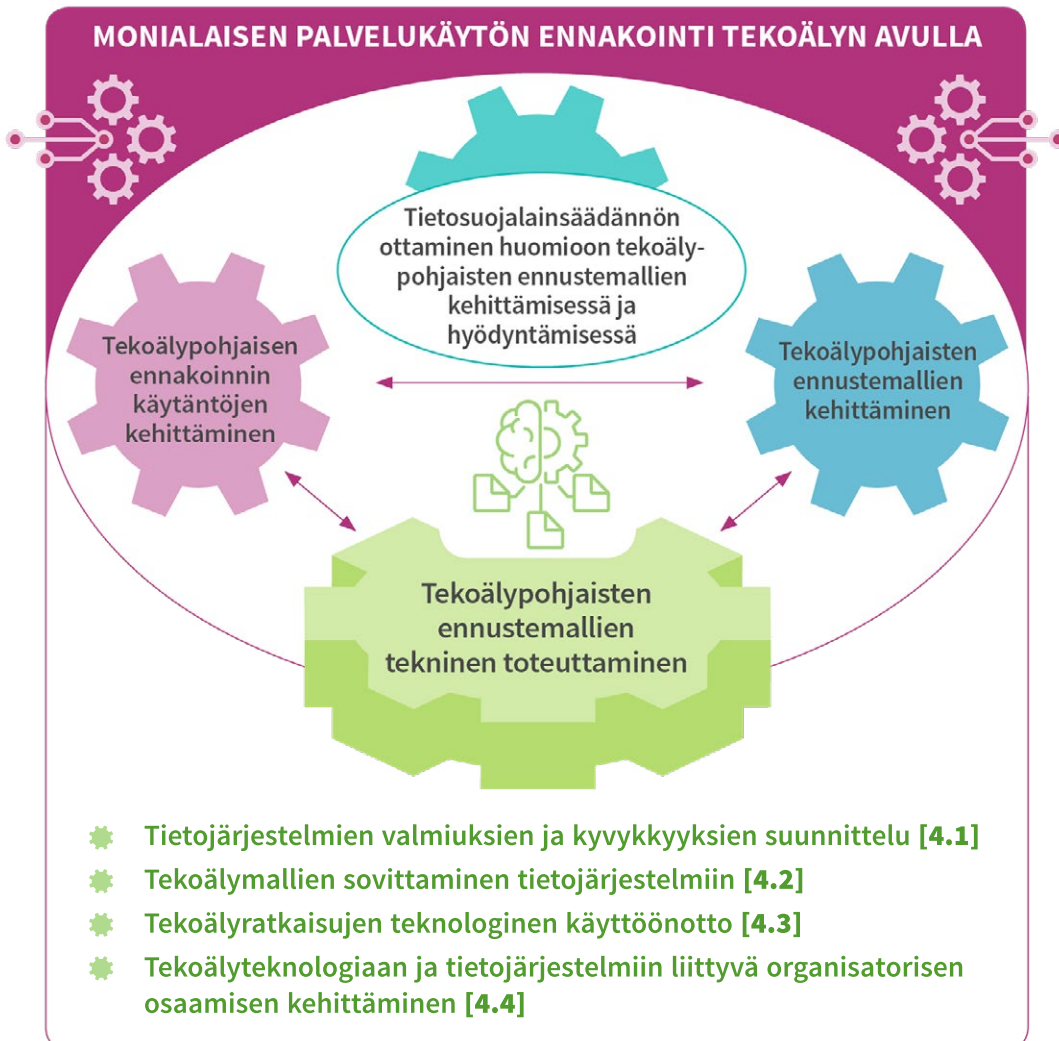


# TEKOÄLYPOHJAISTEN ENNUSTEMALLIEN TEKNINEN TOTEUTTAMINEN



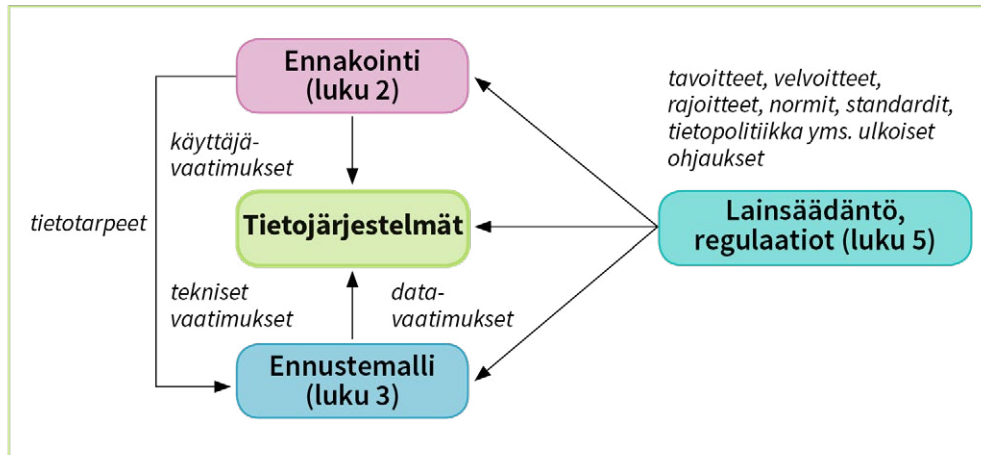
Tässä luvussa käsitellään informaatiojärjestelmien ja tietotekniikan (IT) näkökulmista sitä, mitä tekoälyteknologian tukeman toiminnan (tässä tapauksessa monialaisen palvelutarpeen ennakoinnin) kehittämiseen ja toteuttamiseen sosiaali- ja terveydenhuollon organisaatioissa liittyy ja on otettava huomioon. Tässä tekoälymallin realisointia käsitellään neljästä eri näkökulmasta omissa alaluvuissaan (Kuvio 4).

Palvelutuotantokäytössä tekoälymalli on suoritettava komponentti operatiivisessa tietojärjestelmässä. Se saa syötetietonsa valituista tietokannoista tai muista järjestelmistä ja sen laskemat tulokset välitetään niiden käyttäjille tietojärjestelmissä. Mallin käyttö tapahtuu organisatorisissa prosesseissa, jolloin siitä tulee osa palvelutuotannon työkaluja. Tällainen tekoälyn käyttöönotto vaikuttaa monialaisesti vaatiensa onnistuakseen suunnitelmallista muutosprosessia monin edellytyksin ja mahdollistavin tekijöin. Myös mahdollisia esteitä on käsiteltävä. Kaikki edellä oleva vaatii monia, uusiakin osaamisia (kompetensseja) ja osaamisyhdistelmiä, joita organisaatioissa on järjestelmällisesti kehitettävä. Käyttöönotto voi edellyttää muutoksia jopa työntekijöiden toimenkuviin ja organisaatorakenteisiin.



**Kuvio 4.** Tekoälypohjaisten ennustemallien teknisen toteuttamisen keskeiset tehtävät

Tämä luku liittyy käsikirjan muihin lukuihin kuviossa 5 esitetyllä tavalla: Ennakointi (luku 2) perustuu tietojärjestelmiin, jotka suorittavat kehitetyn tekoälymallin (luku 3) laskennan, sisältävät laskennassa käytetyt tietolähteet ja välittävät laskennan tulokset organisaatiossa käyttäjille; Lainsäädäntö ohjaa ja asettaa reunaehdoja koko toiminnalle ja järjestelmille (luku 5).



**Kuvio 5.** Tekoälymallin realisoinnin riippuvuudet

On huomattava myös, että edellä kerrotut eri osa-alueet kehittyvät rinnakkain iteraatiivisesti. Tekoälymallin realisointi ei siis ole kertaluontoinen tapahtuma, vaan pikemminkin organisatorinen muutospolku (“journey”) ja oppimisprosessi. Se ei oikeastaan pääty staattiseen tavoitetilään, vaan kehitettyä mallia on ylläpidettävä ja vastaavasti muita osa-alueita sovitettava jatkuvasti.

Huomattakoon, että tässä oletuksena on, että lähtötilanteena on kokonaan uuden tekoälymallin kehittäminen ja käyttöönotto olemassa olevassa, toimivassa organisaatiossa. Silloin kun hyödynnetään ja otetaan käyttöön jossakin muualla jo kehitetty ennustemalli, tulee malli sovitaa oman organisaation tarpeisiin sekä sen toiminnallisiin ja teknisiin vaatimuksiin.

Ennustemallin kehittäminen ja operatiivinen käyttö voidaan nähdä osana laajempaa tiedolla johtamisen prosessikokonaisuutta. Tietojärjestelmien näkökulmasta oleellista on se, että kaikissa vaiheissa tietoja kerätään, prosessoidaan ja jaetaan tyypillisesti useiden eri järjestelmien avulla. Jotta tiedolla johtamista voidaan toteuttaa täysipainoisesti, eri järjestelmien ja tietojen integrointi tulee olla mahdollisimman saumatonta tukien niiden hyödyntäjien todellisia tarpeita ja loppukäyttäjien käytännön työskentelyä.

## 4.1 Tietojärjestelmien valmiuksien ja kyvykkyyksien suunnittelu

Tarkoituksenmukaiset ja riittävän tehokkaat ICT-perustaiset tietojärjestelmät ovat nykyään välttämättömiä ja jopa kriittisiä sosiaali- ja terveydenhuollon organisaatioiden toiminnan mahdollistajia ja palveluiden toteutusvälineitä. Tietojärjestelmiin liittyvät tietojohdaminen, tiedolla johtaminen ja tietoperustainen johtaminen (Rytilä 2011). Koska erityisesti tekoälymallit pohjautuvat dataan ja laskentaan, tietojärjestelmät ovat keskeinen elementti tekoälyteknologian soveltamisessa.

Julkisissa organisaatioissa on raportoitu monia esteitä tietojärjestelmien tuottaman tiedon hyödyntämisessä. Niitä liittyy sekä tiedon laatuun – kattavuuteen, vertailukelpoisuuteen, ajantasaisuuteen ja luotettavuuteen että tietojen yhdistelyyn, muotoihin, rajapintoihin ja ylipäätään tarpeellisen tiedon puuttumiseen. (Salovaara ym. 2021.) Koska tiedon laatu on oleellista tekoälymallien kehittämisessä ja operatiivisessa käytössä, se onkin otettava erityisesti huomioon jatkuvasti. Mallin kehittämisvaiheessa määrittäytyy se, mitä dataa eri tietojärjestelmistä on oltava saatavissa ja millaista hyödynnettävää informaatiota malli muodostaa suoritusympäristössään. Organisaatioissa tulee olla kyky tunnistaa mallilla tietojärjestelmissä tuotetun tiedon ja siinä käytetyn datan luotettavuutta. Tekoälymallien hyödyntämisessä ongelmana voi olla myös se, että ne saattavat vaatia laajat asiakastietorekisterit ja tehokkaat tietojärjestelmät (Koivisto & Tiirinki 2020). On oltava valmius ja halukkuus tehdä investointipäätöksiä tietojärjestelmiin, jotka tallentavat ja tuottavat haluttua tietoa soveltuvasti ja käyttökelpoisesti.

Kehitettäviin ja operatiivisesti sovellettaviin tekoälymalleihin liittyviä eri tietojärjestelmiä voidaan kartoittaa esimerkiksi ns. koneoppimisen kehittyneen analytiikan kanvaasin avulla (Pentti ym. 2019). Kun ennakoitua tehdään asiakas- ja potilastietojen avulla, tarvitaan sopivat tietojärjestelmät ja -rekisterit, joista tietoja voidaan poimia. Tällöin ongelmana voi olla se, että erilaisia tietoja tallennetaan eri paikkoihin erillisiin järjestelmiin, jotka eivät ole vuorovaikutuksessa keskenään. Algoritmien, tiedon louhinnan ja tekoälyn hyödyntäminen vaatiikin usein myös tietojärjestelmien uudelleensuunnittelua – ottaen huomioon myös tietosuojaan liittyvät mahdolliset esteet tietojärjestelmien laajemmalle hyödyntämiselle (ks. luku 5). (Koivisto & Tiirinki 2020.)

*Onko organisaatiolla tietojohdamista yhdistävä alusta tai järjestelmä?*

*(Leskelä ym. 2019)*

Tietojärjestelmät ovat osa sote-organisaation kokonaisarkkitehtuuria. Yleisenä tavoitteena Suomessa on se, että ”kansallisia sähköisiä palveluja ja tietojärjestelmiä hyödynnetään mahdollisimman laajasti, linjassa kansallisen sote-kokonaisarkkitehtuurikehyksen kanssa” (Anttila ym. 2020). Sosiaalialan tietojärjestelmissä on pyritty muodostamaan kansallisesti yhtenäistä rakennetta tietoarkkitehtuurille, ja sosiaali- ja terveydenhuollon digitalisaatiossa tietojärjestelmien kokonaisarkkitehtuuria on määritetty THL:n johdolla jo pitkällä aikavälillä tavoitteena yhteinen kokonaisarkkitehtuuri, joka mahdollistaa yhtenäisen ja vertailukelpoisen tiedon (Salovaara ym. 2021).

*Mahdollistavatko organisaation tietoarkkitehtuuri ja teknologia-arkkitehtuuri tietosisältöjen ja tietotuotannon yhteensovittamisen (hyvinvointialue, kansalliset toimijat)? (Välikangas 2020)*

Kokonaisarkkitehtuurimalli käsittää neljä arkkitehtonista osa-aluetta (Lohtander ym. 2021): toiminta-arkkitehtuuri, tietoarkkitehtuuri, tietojärjestelmäarkkitehtuuri ja teknologia-arkkitehtuuri (organisaation teknologinen infrastruktuuri ja järjestelmäarkkitehtuurin teknologiavalinnat). Tekoälymalleilla on kytköksiä kaikkiin em. kokonaisarkkitehtonisiin osa-alueisiin. Ennustemallien soveltamisen kokonaisarkkitehtonisessa systeemisuunnittelussa onkin tarkasteltava järjestelmällisesti ammattilaiskäyttäjryhmien kannalta keskeisiä tietojärjestelmiä, integraatiotarpeiden jaottelua ja mallituotteen liittämistä sote-järjestelmiin (Lähtenmäki ym. 2020). Silloin kun otetaan käyttöön josakin muualla jo kehitetty ennustemalli, tulee malli sovittaa oman organisaation tarpeisiin sekä sen toiminnallisiin ja teknisiin vaatimuksiin.

*Mahdollistavatko organisaation ohjelmistointegraatioarkkitehtuurit ”älykkäät” sote-järjestelmät, joissa data-analytiikka ja tekoäly integroituvat saumattomasti? (Nasr ym. 2021)*

*Jos tekoälyratkaisut ovat erillisiä sovelluksia, integroituvatko ne saumattomasti organisaation muihin tietojärjestelmiin (esim. potilastietojärjestelmät)? (Nummela 2022)*

Kansallinen Virta-arkkitehtuuri linjaa sote-tietojärjestelmien kehittämistä valtakunnallisella tasolla ottaen huomioon hyvinvointialueiden tietojohdamisen tarpeet kokonaisuutena. Se on viitearkkitehtuuri, jonka avainsuunnitteluperiaatteena on ”kansalliseen käsittemallinnukseen pohjautuva tietoturvallinen toisiokäytön lain vaatimukset huomioiva automaattinen model-driven tietovarasto” (DigiFinland 2021). Virta-arkkitehtuurissa tarkoituksena on yleisesti tiedolla johtamisen tukeminen luotettavalla ja ajantasaisella tiedolla (DigiFinland, n.d.). Virta-arkkitehtuurin mukaisissa automaattisissa mallipohjaisissa tietovarastoissa pyritään siihen, että tiedot linkitetään lähdejärjestelmiin ja latausohjelmat muodostetaan automaattisesti (ETL (Extract, Transform and Load) / DW-kerros). Viitearkkitehtuuri ei ota kantaa siihen, mitä komponentteja tiedolla johtamisen tietojärjestelmissä pitää olla. (DigiFinland 2021.) Käytännössä saatetaan tarvita esimerkiksi jalostavia lähdedatan prosessointivaiheita, koska alkuperäinen tallennettu data saattaa olla puutteellista, epäyhtenäistä tai biasoitunutta (Ghassemi ym. 2019).

Kun sosiaali- ja terveydenhuollon organisaatioiden tietojärjestelmiä kehitetään mahdollistamaan ja tukemaan tekoälyavusteisuutta tai jopa -pohjaisuutta osana tiedolla johtamista, toimitaan monialaisesti sosiaali- ja terveyshuolto-osaamisten ja asiantuntemusten leikkauksissa (esim. sähköiset potilasrekisterit) (Aerts & Bogdan-Martin 2021). On otettava systemaattisesti huomioon organisaation sote-tietojärjestelmiä koskevat kansalliset määrittelyt ja linjaukset ja erityisesti rajapintoja koskevat kansainväliset standardit ja spesifikaatiot (mahdollisesti tulevatkin) (Lähtenmäki ym. 2020).

*Voidaanko tekoälyjärjestelmiä toteutettaessa hyödyntää komponentteja, jotka ovat jo käytössä sote-organisaatioissa, jolloin järjestelmien käyttöönotto ei vaadi suuria laite- tai ohjelmistoinvestointeja eikä myöskään muutoksia kokonaisarkkitehtuuriin? (Klemola ym. 2014)*

#### 4.1.1 Operatiivisten tietojärjestelmien ominaisuuksien varmistaminen

Kun tekoälymalli operationalisoidaan, tietojärjestelmäkokonaisuudessa on tarkasteltava seuraavia asioita:

1. Voidaanko annettu ennustemalli automatisoida (tehokkaasti)? Käytettävien tietojärjestelmien teknistä suorituskykyä voidaan arvioida ennakoivasti pilotointivaiheessa mittauksin ja tarvittaessa investoida lisäkapasiteettiin (ks. esim. Pentti ym. 2019).
2. Ennustemallin informaatio/datatarpeiden ja -lähteiden identifiointi ja analyysi (saatavuus, datamuodot, ohjelmalliset rajapinnat): Tietoja voi joutua kokoamaan useammasta järjestelmästä, ellei tietoja ole saatavilla esimerkiksi erityisestä tietotaltaasta tai järjestelmästä, joka kokoaa kaikki alueen/organisaation sosiaali- ja terveydenhuollon potilas- ja asiakastiedot.
3. Ennustemallin laskennallisten tulostietojen välittäminen käyttäjille (informaatiojärjestelmien väliset rajapinnat, käyttöliittymät)
4. Ennustemallin laskennan toteuttavan sovelluksen ohjelmistotekninen arkkitehtuuri: Tekoälyteknologiaan perustuvat ratkaisut ovat nykyisin nopeassa kehitysvaiheessa, mutta yleisperiaatteena teknologian käytön tulisi perustua yhtenäisiin prosesseihin ja hyödyntää sekä sote-alueen omia että kansallisia tietojärjestelmäpalveluita (Lähtenmäki ym. 2020).

Tekoälymallien kehitykseen, tuotantoon viemiseen ja operatiiviseen käyttöön liittyy useita eri tietojärjestelmiä ja niiden kytköksiä. Kun sote-organisaation tekoälymalleihin haetaan dataa ulkoisista rekistereistä (esim. Kela), on tarve määritellä yhtenäisiä käytäntöjä (Lähtenmäki ym. 2021). Kun kehitetty tekoälymalli julkaistaan operationaaliseen tuotantokäyttöön, se on integroitava olemassa olevien sovellusten ja mahdollisten muiden mallien kanssa. Jo tekoälyratkaisujen pilotointivaiheessa tulisi olla tiedossa tiekartta niiden tulevaan tuotteistamiseen ottaen huomioon organisaation muut jo käytössä olevat ja mahdollisesti tulevat kehitettävät ennustemallit. Koska mallien kehitys ja ylläpito on käytännössä toistuva prosessi, tarvitaan jatkuvan integroinnin ja julkaisun (CI/CD, ”continuous integration/deployment”) käytänteitä ja niitä mahdollistavia työkaluketjuja. Tekoälymalleissa tarvitaan lisäksi jatkuvan (uudelleen)opettamisen järjestelmällistä tukea (CT, ”continuous training”), jotta aikaansaadaan niiden turvallinen, jäljitettävä, testattava ja toistettava kehittäminen, kouluttaminen, julkaiseminen ja päivittäminen (John ym. 2021). Operatiivisessa tuotantokäytössä tekoälymalleja on monitoroitava, jotta voidaan luottaa siihen, että malli toimii tuotannossa halutulla tavalla ja tuottaa aiottua hyötyarvoa (Sallinen 2022). Tekoälymallit saattavat olla huomattavan laskentaintensiivisiä (tietokonekuormitus).

*Miten järjestelmiä seurataan (monitoroidaan) koneoppimismallien tuotantoonviennin jälkeen? (Mylyyaho ym. 2022)*

IT-infrastruktuuria tuleekin kehittää ja hallita jatkuvasti, jottei se muodostu tekoälymallien operatiivisen hyödyntämisen esteeksi (Pote ym. 2021). IT-infrastruktuuripohjan tulee olla vankka, verkko- ja tietoliikenneyhteyksin tulee olla riittävät, tietoturvasta tulee huolehtia ja ICT-ratkaisujen on oltava kaiken kaikkiaan toimintavarmoja.

- *Kuinka yhteensopivaa organisaation olemassa oleva informaatioteknologia (IT) on uuden tekoälyteknologian kanssa? (Madan & Ashok 2023)*
- *Onko käytetyissä mallinusteknologioissa ja tuotantoympäristön teknologioissa mallien integrointia hankaloittavia eroja? (Riungu-Kalliosaari ym. 2017)*
- *Miten tulevaisuuden infrastruktuuriarkkitehtuuria suunnitellaan tekoälytuetun sote-järjestelmän kannalta? (Tyrväinen ym. 2018)*

#### 4.1.2 Tekoälyohjelmistojärjestelmien kehittäminen

Tekoälyohjelmistoissa (“AI software”) integroidaan tekoälykomponentteja aikaansaamaan “älykkäitä” ohjelmistojärjestelmiä (Tamburri 2020). Sellaisten koneoppimiskomponentteja sisältävien järjestelmien konstruointiin ja muokkaamiseen tarvitaan uusia, strukturoituja tuotantomenetelmiä, jotka laajentavat perinteisiä ohjelmistotuotannon (“software engineering”) prosesseja ja teknologioita ottaen mukaan myös datatieteen ja tekoälyn metodeja ja teknologioita yleisemminkin (“AI engineering”) (Bosch ym., 2021). Koneoppimismallit ovat vain yksi osa ohjelmistojärjestelmäkokonaisuutta, jossa tarvitaan paljon muutakin toiminnallisuutta. Ohjelmistojärjestelmien arkkitehtuurisuunnittelussa on otettava huomioon sekä tavanomaiset että tekoälykomponentit.

- *Millaista järjestystä tekoäly/koneoppimisohjelmistotuotteiden kehityksessä organisaatiossa tyypillisesti noudatetaan: mallilähtöinen / tuotelähtöinen (laajentaminen koneoppimistuetulla toiminnallisuudella) / mallin ja tuotteen kehittäminen rinnakkain? (Nahar ym. 2022)*
- *Miten ja missä tehdään koneoppimissovelluskehityksen eri työvaiheet: vaatimusten analyysi, dataorientoituneet tehtävät, malliorientoituneet tehtävät, DevOps-tehtävät? (Kumeno 2019)*

Koneoppimiselementtejä sisältävien tekoälyohjelmistojärjestelmien kehittämiseen liittyy monia tunnistettuja haasteita. Ne liittyvät yleisesti datan hallintaan ja datan laatuun, mallien suunnittelumenetelmiin ja -prosesseihin sekä tuotantoonvientiin ottaen huomioon vaatimusten sääntöjenmukaisuus (“compliance”, ”security”, ”privacy”) (Bosch ym. 2021).

- *Miten kehityskäytäntein ja infrastruktuuriratkaisuin aikaansaadaan saumaton siirtymä tekoälyratkaisujen prototypoinnista tuotantoonvientiin? (Muiruri ym. 2022)*
- *Miten tekoälymallien/ratkaisujen kehitysvaiheen tuotokset (esim. mallin toteutuskoodi) paketoitaan ja dokumentoidaan (esim. ”model card”)? (Stirbu ym. 2022)*

Tekoälyohjelmistoteknologia on tällä hetkellä nopeasti kehittyvää, joten yleisesti hyväksyttyjä tekoälyohjelmistojen kehitysmetodologioita eikä vakiintuneita, kypsiä prosessimalleja ei juurikaan ole. Erityisesti se, miten tehdään tekoälyohjelmistojen (tekoälyominaisuuksien) testausta ja laadun validointia on osittain avoin tutkimusongelmakin. Tekoälyohjelmistoversioiden testaaminen poikkeaa merkittävästi perinteisestä ohjelmistotestaamisesta. Perinteiset ohjelmistotestausmenetelmät ja -työkalut eivät yleensä ole riittäviä. (Tao ym. 2019.) Luonnollisesti on testattava myös yleiset järjestelmäpalveluiden laatuparametrit (suorituskyky, luotettavuus, tietoturva, jne). Periaatteessa testausprosessissa ja sen automatisoinnissa saattaa olla mahdollista soveltaa itsessään tekoälyteknologiaa (esim. testidatan analysoinnissa), mutta tällaisia menetelmiä ja ratkaisuja ei vielä ole yleisesti käytössä.

••••• *Miten varmistutaan koko tekoälyjärjestelmän (ei vain koneoppimismallin) oikeellisuudesta ja vankkuudesta? (Myllyaho ym. 2021)*



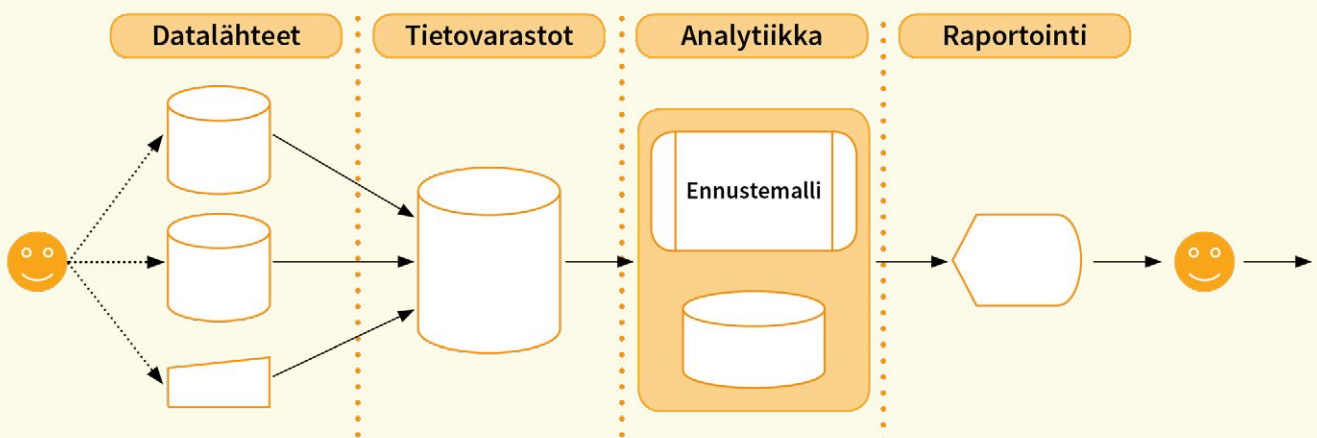
## PÄIJÄT-HÄME

### Tietojärjestelmien valmiuksien ja kyvykkyyksien suunnittelu Päijät-Hämeessä

Päijät-Hämeen hankkeessa kehitetyn ennustemallin tietojärjestelmätekniinen kehitysympäristö perustui Findatan tietokapseliin. Se ei ollut suoraan yhteydessä Päijät-Hämeen hyvinvointialueen tietovarastoon, josta mallin kehittämisessä käytetyt lähdetiedot toimitettiin (ks. kuva alaluvussa 3.4 Aineiston poiminta ja esikäsittely Päijät-Hämeessä).

Hankesuunnitelma oli budjettirajoituksista johtuen rajattu siten, että tavoitteena oli vain kehitettävän tekoälymallin demonstraatioversio (nk. Proof-of-Concept, PoC). Sen tarkoituksena oli mahdollistaa mallikonseptin esittely ja käyttäjäkokeilu rajatuin toiminnallisuuksin. Tekoälymallin tuotantoonvienti hyvinvointialueen operatiiviseen ympäristöön ja integrointi päätöstukisovelluksiin eivät siten tapahtuneet hankkeen aikana.

Hankkeen alussa ennustemallin tuotantoympäristöä hahmotettiin alla olevan yleisen kaavakuvan pohjalta. Sote-asiakkaiden/potilaiden (kuvassa vasemmalla) tiedot tietovarastoissa ovat syöteinä ennustemallin laskentaan analytiikkatietojärjestelmässä, josta tulokset välitetään käytettäväksi ammattilaisille (kuvassa oikealla).

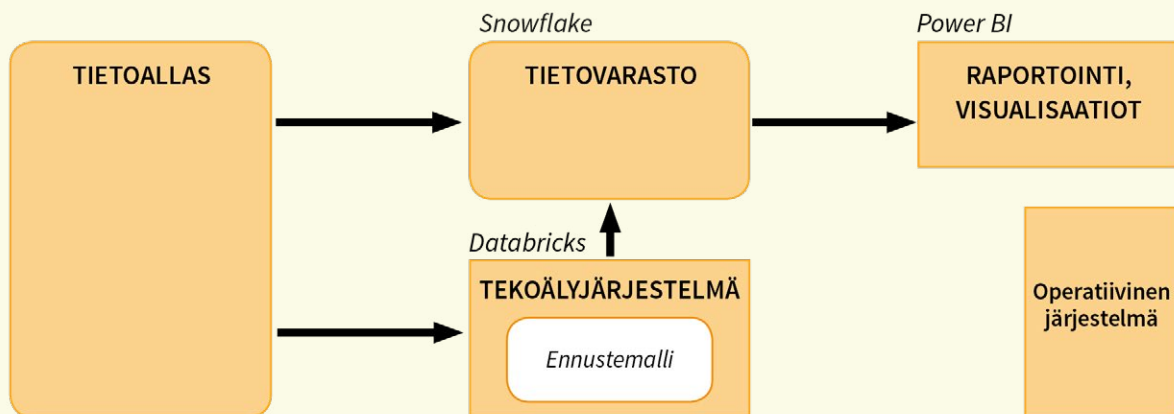




Hankkeen aikana mahdollinen tuotantoympäristön toteutus tarkentui ja konkretisoitui seuraavasti:

- 11.5.2022: Ennustemallin käyttöönotosta ja hyödyntämisestä jatkossa ei vielä ole konkreettista käsitystä.
- 19.12.2022: Toteutusajatuksena on se, että mallia ajetaan Databricks-ympäristössä Snowflake-tietovaraston kautta. Mallin tuloksia voidaan julkaista tietovarastosta mm. Power BI-analytiikka-sovelluksella.
- 12.1.2023: Aikomuksena on nimenomaan toteuttaa Power BI:lla raportointia.

Tämän tuloksena ennustemallin realisointi tuotantoympäristössä voisi siten olla tietojärjestelmä-arkkitehtonisesti alla olevan kaavakuvan mukainen:



Tämän hankkeen ulkopuolelle jäi siis se, miten kehitetty ennustemalli vietäisiin kehitysympäristöstään (kuva alaluvussa 3.4 Aineiston poiminta ja esikäsittely Päijät-Hämeessä) hyvinvointialueen tuotantoympäristöön (vrt. alaluku 3.6 Operatiivisen käytön huomiointi). Päijät-Hämeen hyvinvointialueella jo aikaisemmin kehitetty kansallisen Virta-arkkitehtuurin mukainen data-alusta kuitenkin mahdollistaa tietotuotteiden (kuten raporttien) teon. Siinä voidaan tehdä edistyksellistä data-analytiikkaa koneoppimisteknologioin yhdistäen ja hyödyntäen tietoja lainsäädännöllisten vaatimusten rajoitteiden mukaisesti (ks. luku 5).

Päijät-Soten Digistrategiassa (2021–2023) nk. digitalustalla on keskeinen asema. Se on jatkuvasti kehitettävä järjestelmäkokonaisuus, joka kokoaa yhteen digitaalisten palveluiden toiminnallisuudet ammattilaisille ja asiakkaille. Alustasta saadaan jatkuvasti kertyvää dataa tiedolla johtamiseen ja palvelujen jatkuvaan kehitykseen. Alustan ja sillä olevien palveluiden mittaaminen on tärkeä osa tietojärjestelmien mahdollistamaa tiedolla johtamista.

## 4.2 Tekoälymallien sovittaminen tietojärjestelmiin

Tietojärjestelmät ja tietojen hallinta ovat kiinteä osa nykyaikaisen sosiaali- ja terveydenhuolto-organisaation prosesseja, toiminnallisia rakenteita ja johtamista (Salovaara ym., 2021). Tietojärjestelmiä tuleekin kehittää niiden organisatorisessa kontekstissa ottaen huomioon niitä käyttävien ja tietoja hyödyntävien ihmisten toiminnan, työskentelytilanteet ja työtavat. Tunnettuja yleisiä ongelmia tietojärjestelmien käyttöönotoissa ovat järjestelmien toiminnallisuuden yhteensopimattomuudet organisaation työntekijöiden kanssa (Ekholm & Kinnunen 2016). Informaatiojärjestelmät ovat monimutkaisessa vuorovaikutuksessa organisaation sosioteknologisissa verkostoissa, joissa niiden tulisi olla sopuun sopuun työntekijöiden ominaisuuksien, preferenssien ja heidän työtehtäviensä suhteen (Ammenwerth ym. 2006).

*Miten tekoälyjärjestelmien toiminta selitetään eri asiantuntemustason sidosryhmille (ekspertit / noviisit (sote-ammattilaiset, maallikot))?*  
(Szymanski 2022)

Palvelutuotannon johtamisessa ja operatiivisissa työprosesseissa keskeisenä koordinoimien mahdollistajana on tiedon saanti koko palvelujärjestelmätason potilas- ja asiakastiedoista (Rautiainen 2021). Ennakointia tukevien tekoälymallien analytiikka ja laskentatulokset tuleekin integroida palvelu- ja hoitopolkuihin (National Health Service 2019).

*Miten tekoälymallit liittyvät organisaation tiedolla johtamisen prosesseihin?*  
(Helander ym. 2020)

*Miten tekoälyä hyödynnetään asiakkaan tilannekuvassa, ammattilaisen päätöksenteon tukena ja asiakkaan päätöksenteon tukena?*  
(Achte & Hiltunen-Toura 2020)

*Onko ammattilaisilla käytössä monia eri järjestelmiä, vai onko heidän mahdollista käsitellä (kaikkia) eri tietoja yhden järjestelmän kautta?*  
(Ruotsalainen & Tötterman 2022)

Tekoälyn soveltamisessa voidaan nähdä sellaisia kehitysmahdollisuuksia, että tekoälyratkaisusta tulee työntekoa tukevia virtuaaliassistentteja (”tukiäly”). Tällöin organisaation työntekijöitä ja toimintoja voidaan suunnitella kysyen, miten tekoälyä (koneoppimista) voidaan parhaiten hyödyntää ammattilaisten ja ”koneiden” yhteistyönä. (Akava ym. 2022.) Tällaisessa prosesseihin ja työntekijöihin vaikuttavassa digitalisaatiossa on kuitenkin myös potentiaalisia riskejä ja vaikeuttavia tekijöitä (Aerts & Bogdan-Martin 2021).

*Aiheuttaako tekoälyn käyttö päätöksenteon tukena merkittävän muutoksen verrattuna tavanomaisiin tietopohjaisiin päätöksentekotukijärjestelmiin?*  
(Cresswell ym., 2020)

Erilaisia ennakkoinnin malleja ja työkaluja on käytössä tai kehitteillä (Koivisto & Tiirinki 2020). Kun uusi tekoälymalli kehitetään ja otetaan käyttöön organisaatiossa, jonkun tai jonkin on suoritettava useita eri tehtäviä liittyen tietojärjestelmiin ja operatiivisiin prosesseihin. Eri tehtävistä vastaavat eri roolit organisatorisissa asemissa (mahdollisesti jopa eri organisaatioissa). Taulukossa 2 on esitetty tästä kokonaisuudesta luonnehtiva yhteenveto tietojärjestelmien ja IT:n näkökulmista (vrt. luku 3).

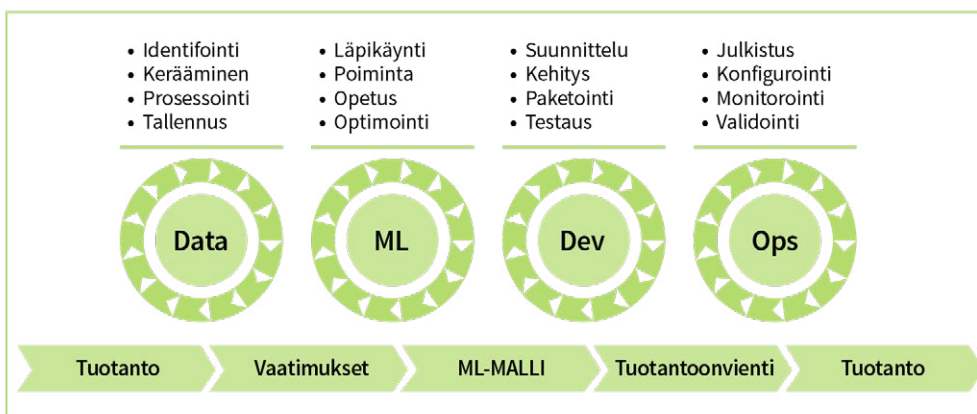
**Taulukko 2.** Tekoälymallin realisoinnin tehtävät ja toimijat

Vaihe / Tehtävät	Toimijat (tietojärjestelmät, IT)
Tekoälymallin tavoitteenasetanta prosesseissa ja työnkuluissa (tietojärjestelmien vaatimusmäärittely)	Tietojärjestelmien suunnittelijat IT (arkkitehdit)
Datan poiminta, prosessointi (opetus-, validointi-, testidata)	IT (tuki)
Mallin operatiivinen toteutus	IT (arkkitehdit, ohjelmistokehittäjät)
Mallin käyttöönotto (pilotointi, tuotantoon vienti)	IT (ohjelmistokehittäjät, MLOps)
Mallin operatiivinen tuotantokäyttö	IT (MLOps)
Mallin ylläpito (monitorointi)	IT (MLOps)

MLOps (Machine Learning Operations) on teknologinen ja organisatorinen konsepti, jossa tekoälymallien dataoperaatiot, mallien jatkuva kehittäminen ja niiden operatiivinen käyttö ja ylläpito mukaan lukien uudelleenopettaminen pyritään integroimaan ja automatisoimaan saumattomaksi, koko elinkaaren päästä-päähän kattavaksi kokonaisuudeksi (John ym. 2021). Motivaationa on yhtenäistää tekoälyjärjestelmien kehittämistä ja operatiivista käyttöä laentaen DevOps-periaatteita.

MLOps-toimintapojen menestykselliseen soveltamiseen liittyy useita ratkaistavia kysymyksiä ja mahdollisia esteitä, esimerkiksi (John ym. 2021) eksperimenttien seuraaminen ja vertaileminen, versionhallinnan puutteet, (monimutkaisten) mallien tuotantoon vienti ja rajoittavat regulaatiot. Tekoälyelementtejä sisältävien tekoälyohjelmistojärjestelmien jatkuva operointi (DataOps-ketju) poikkeaa radikaalisti tavanomaisista ohjelmistoista (DevOps) (Tamburri 2020). Koneoppimiskomponentteja on orkestroitava osana tekoälyohjelmistokokonaisratkaisua.

Yhteenvetona kuviossa 6 on esitetty keskeiset prosessit, työnkulut ja tehtävät ML-malleja sisältävien tekoälyohjelmistojen kehittämisessä ja operoinnissa (DataOps, MLOps, DevOps). Ne on kaikki jonkun/jonkin tehtävä organisaatiossa työkalujen tukemana ko. data- ja IT-ympäristössä. Huomattakoon lisäksi, että ne ovat iteratiivisia ja toistuvat mallien elinkaaren aikana.



**Kuvio 6.** Tekoälymalleja toteuttavan MLOps-ohjelmistotuotannon prosessivaiheet (mukailtu (Tamburri 2020))



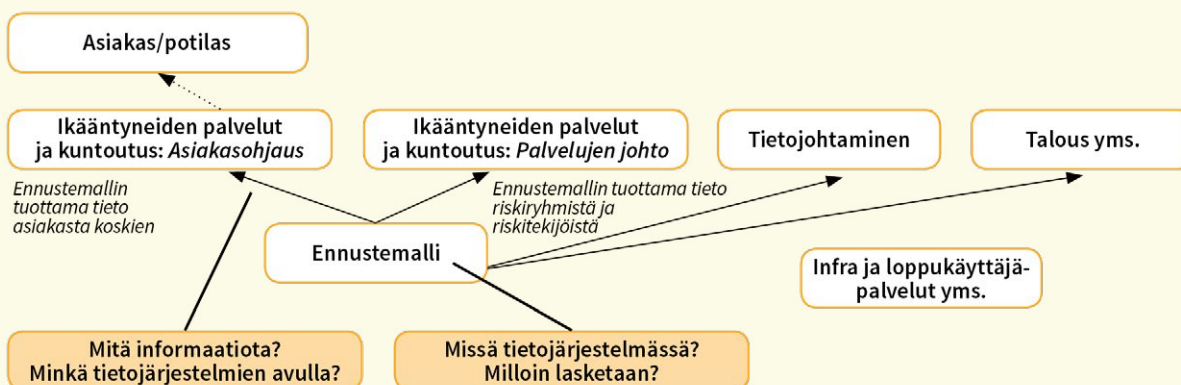
**PÄIJÄT-HÄME**

### Tekoälymallien sovittaminen prosesseihin ja työnkulkuihin tietojärjestelmissä Päijät-Hämeessä

Tässä hankkeessa kehitettyä ennustemallia ei projektin puitteissa viety pilotoinnista operatiiviseen tuotantokäyttöön. Sitä kuitenkin otettiin ennakoivasti huomioon mallin kehityksessä ja kokeilussa (ks. alaluku 3.6 Operatiivisen käytön huomiointi).

Mallilta haluttuja tietotarpeita ja mallin käyttöä eri ammattilaisten ja muiden organisatoristen hyödyntäjien työnteoissä ja työskentelyprosesseissa tunnistettiin ja määriteltiin järjestelmällisesti (ks. luku 2, Ennakoinnin kehittämistarpeiden tunnistaminen ja kehittämistavoitteiden määrittely, Ennakointi johdon toimintana ja Ennakointi asiakasohjauksessa, Ennustemalliin pohjautuva seuranta). Näin voitiin identifoida eri työaktiviteetit ja niissä uutta tekoälypohjaista ennustemallia hyödyntävät työkäytännöt perustuen sen tuottamaan informaatioon.

Em. pohjustusten ja kaavailujen perusteella on mahdollista projisoida ennustemallin ja sitä hyödyntävän ennakoinnin prosessien ja työnteokujen tarpeita ja vaatimuksia tietojärjestelmille alla olevan kaavakuvan esittämällä tavalla. Kehitetty ennustemalli on toteutettava tuotantoympäristön tietojärjestelmissä. Mallin käyttäjät ja hyödyntäjät saavat niistä mallin tuottamat tulostiedot raporteina ja mahdollisesti visualisointeina. Mallin laskenta on operationalisoitava mahdollisesti automaattiseksi. Kaikki tietovuot on tunnistettava ja määriteltävä ja ne on tietojärjestelmissä vastaavasti luotava ja kytkettävä ko. datalähteisiin.



Päijät-Hämeessä aikomuksena on toteuttaa Power BI:lla raportointia, jota ennustemallin hyödyntäjät voivat käyttää eri aktiviteeteissä. Aikaisemmin ennen tätä hanketta toteutetun ennustemallin tapauksessa (ikäihmisten palvelutarpeen ennakointi) mallia ajetaan säännöllisin väliajoin automatisoidusti.

Tuotantokäytössä ennustemallia on myös jatkuvasti seurattava ja tarvittaessa päivitettävä. Tällaiset tehtävät kuuluvat MLOps-tyyppisiin prosesseihin ja työnkulkuihin. Päijät-Hämeessä tämänkaltaisia toimintatapoja noudatetaan osittain jo tällä hetkellä.

Kaiken kaikkiaan on otettava huomioon lisäksi se, että koska tekoälymallit perustuvat dataan, niissä käytettävän datan järjestelmällinen kerääminen ja kirjaaminen jatkuvasti kaikissa ao. työprosesseissa on oleellinen mallien tuottaminen tulosten hyödynnettävyyden kannalta (tarkkuus, ajantasaisuus, luotettavuus). Päijät-Hämeessä tällaista tieto-orientoitumista onkin korostettu kaikkien ammattilaisten yhteisenä vastuuna tietojohdamisessa<sup>13</sup>.

Päijät-Hämeessä on määritelty organisaation tietotuotteille sisäiset kehitys- ja hallintaprosessit. Jos tekoälymalleja pidetään tietotuotteina, niiden käsittely on otettava huomioon ko. prosesseissa. Kuitenkaan kaikista tekoälykokeilusta ei välttämättä synny organisaation virallisia tietotuotteita.

13 <https://paijat-sote.fi/sote-uudistus/laadukkaasta-tiedosta-syntyy-parempia-palveluja%E2%80%AF/>

## 4.3 Tekoälyratkaisujen teknologinen käyttöönotto

Tekoälyratkaisujen käyttöönottoon liittyvät samankaltaiset periaatteet ja tekijät kuin yleisesti uusien teknologioiden tuomisessa työorganisaatioihin. Yksittäiset tekoälymallit eivät välttämättä liity organisaation kaikkiin tietojärjestelmiin eivätkä prosesseihin ja työnkulkuihin, mutta mitä useampia ja upotetumpia malleja organisaatioissa on, sitä enemmän syntyy riippuvuuksia ja kytköksiä. Laajimmillaan kyse voi olla koko organisaatiota kattavasta tekoälytransformaatiosta. Yleisesti teknologian vaikutusarvioinnissa tulee ottaa huomioon sekä henkilöstövaikutukset että varsinaiset teknologiaratkaisut (Anttila ym. 2020).

- *Onko organisaatiossa tekoälytiekartta? (Plastino & Purdy 2018)*
- *Millä organisatorisilla toimialoilla, funktioilla (tai vast.) tekoälypohjaisia ratkaisuja selvitetään, millaisin hyötytavoittein? (de Sousa ym. 2019)*
- *Millä tavoin organisaatiossa ymmärretään se, mitä tekoälyteknologiassa on kyse (teknologian luonne), miksi sitä otetaan käyttöön (teknologiastrategia) ja miten sitä käytetään päivittäin? (Sun & Medaglia 2019)*
- *Onko tekoäly sellainen teknologinen innovaatio, joka muokkaa merkittävästi sote-järjestelmää? (Zahid ym. 2021)*

Miten siis päästä alkuun ja eteenpäin tekoälyn/koneoppimisen hankinnassa ja hyödyntämisessä (Akava ym. 2022)? Kaiken kaikkiaan tekoälyn efektiivinen organisatorinen implementointi edellyttää sekä teknologisten että sosiaalisten systeemien integratiivista kehittämistä (Yu ym. 2023). Ylipäätään on luotava sellainen ympäristö ja olot, joissa tekoälyratkaisujen kehittäminen on mahdollista turvallisesti ja vaikuttavasti (Panch ym. 2019). Organisaatiossa on oltava selkeä visio siitä, miten se voi hyötyä tekoälyn soveltamisesta (Sira 2021). Usein olemassa olevissa, toimivissa organisaatioissa on luontevinta upottaa asteittain kehittyvää tekoälyanalytiikka (”advanced analytics”) niihin prosesseihin ja työkaluihin, joita työntekijät jo käyttävät muutenkin toiminnassaan ja sen ohjaamisessa (Finland’s AI Accelerator 2020).

*Miten varmistetaan se, onko tekoälyratkaisujen toteuttaminen mahdollista (”feasible”) organisaatiossa (esim. datan saatavuus, osaaminen, partnerointi, ulkoistettu työ)? (Desouza ym. 2020)*

*Voidaanko tekoälymallien käyttämän datan jatkuva saatavuus varmistaa niiden kehittämiseksi ja parantamiseksi käyttöönoton jälkeen? (Jiang ym. 2017)*

Kun tekoälyteknologiaa tuodaan olemassa olevaan organisaatioon, on tärkeää ymmärtää, miten eri työntekijät suhtautuvat siihen, jotta määritellyt prosessit ja koulutukset tukevat sen hallintaa, käyttöä ja soveltamista. Eri sidosryhmillä voi olla eri käsityksiä ja odotuksia. Monialaisella yhteistyöllä voidaan varmistaa, että tekoälyteknologiakehityksintressit ja ammattilaisten odotukset tulevat huomioonotetuiksi tasapainoisesti. (Shinners ym. 2021.)

*Missä tiedolla johtamisen alueissa tekoälyllä on eniten hyötyä tai potentiaalia? (Tujula 2021)*

Tekoälyteknologian menestyksellinen integrointi riippuu useista pääelementeistä ja tekijöistä, ml. teknologia-organisaatio-suunnastaminen (”alignment”) (Alami ym. 2021). Tekoälyn käyttöönottoon ja sen käyttöön liittyy useita organisatorisia alijärjestelmiä, ml. tekniset järjestelmät (esim. IT-infrastruktuuri kokeilujen mahdollistajana, tekniset kyvykkyydet tekoälyteknologioiden ja liiketoimintaprosessien integroimiseksi) (Yu ym. 2023).

*Liittyvätkö tekoälymallit organisaation digitaalisiin palveluihin? (Maier ym. 2021)*

Tekoälyn menestykselliselle käyttöönotolle ja hyödyntämisellä on tunnistettu useita tyypillisiä esteitä ja hidasteita, jotka liittyvät organisaatioympäristöön, rakenteisiin ja teknisiin elementteihin. Usein ongelmat liittyvät strategiaan: Organisaatiolla ei ole selvää tekoälystrategiaa. (Sira 2021.) Tekoälyjärjestelmiin liittyviä erityisiä esteitä, hidasteita ja riskejä ovat datan laatuongelmat ja informaationhallinta organisaatiossa (datan ”provenance, curation, integration” ja ”governance”) (National Health Service 2019).

●●●●● *Millaista tietopolitiikkaa organisaatioissa noudatetaan datan hallinnassa ja hyödyntämisessä? (Rousku ym. 2019)*

●●●●● *Jääkö organisaatioon kerääntyvä sote-tieto vähälle käytölle? (Varjola 2022)*

●●●●● *Käyttävätkö tekoälymallit yritysten omistamaa dataa ja/tai alustoja? (Lazer ym. 2020)*

Tekoälyn käyttöönotolla ja soveltamisella voi olla monitahoisia seurauksia ja vaikutuksia organisaatioissa työntekijöihin ihmisinä, organisaation toimintaan ja henkilöstön toimenkuviin (Yu ym. 2023). Onkin tärkeää ymmärtää organisaatiota kokonaisuutena työssä käytettävine tekoälyteknologisine järjestelmineen, joita työntekijät ihmisinä soveltavat omissa tehtävissään. Mitä kaikkea tapahtuu, kun ”tekoäly tulee työpaikalle” (Akava ym. 2022)? Ammattilaiset saattavat olla vähemmän halukkaita sen käyttöönottoon, elleivät he luota uuteen teknologiaan tai ymmärrä selvästi, miten se parantaa heidän palvelunsa tuloksellisuutta (Shinners ym. 2020).

●●●●● *Onko esitetty vaatimuksia, ettei tekoälyalgoritmeja saisi käyttää, jos niiden toimintaa ei pystytä tyhjentävästi selittämään, vaikka ne voidaan testeillä osoittaa toimiviksi? (Kärkkäinen 2020)*

Julkisen sektorin järjestelmissä kertyy jatkuvasti suuria määriä dataa, joka voi olla kiinnostavaa teknologia-alan yrityksille. Arvonmuodostuksen on oltava yhteishyödyllistä ja vastavuoroista. On kuitenkin ymmärrettävä se, ettei julkisen sektorin organisaatio pysty yksin kehittämään kaikkia potentiaalisia algoritmeja, vaan tarvitaan jaettuja yhteistyöhankkeita yrityssektorin kanssa. (Zeitoun & Ravaud 2020.) Kun tekoälyteknologiaa otetaan laajamittaisesti ja syvällisesti käyttöön, sen elementeistä tulee osa organisaation IT-strategiaa (Kitsios & Kamariotou 2021).

Tekoälyteknologian käyttöönotossa ja ML-mallien sisällyttämisestä olemassa oleviin organisaatioihin ja tietojärjestelmiin on käytännössä kyse evoluutiosta. Siinä voidaan nähdä yleisiä kehitys- ja kypsyysvaiheita. Ensimmäisellä tasolla tehdään vain kokeiluja ja prototyyppointia muttei varsinaista tuotantokäyttöä. Toisessa vaiheessa malleja viedään tuotantokäyttöön, mutta vain sellaisissa järjestelmissä, jotka eivät ole organisaation toiminnan kannalta kriittisiä. Kolmannella kypsyydystasolla edetään kriittisiin tuotantojärjestelmiin. Seuraavalla tasolla yksittäisistä malleista ja järjestelmistä muodostuu useiden elementtien verkkoja, joissa useat koneoppimismallit ovat toiminnallisissa ketjuissa. (Bosch ym. 2021.)

Useilla toimialoilla on nykyään tehty ja käynnissä tekoälykokeiluja, mutta kynnys viedä tekoälyratkaisut niistä suunnitellusti eteenpäin varsinaiseen tuotantoon ja hyötykäyttöön on edelleen suuri (Nortio 2020). Tekoälyratkaisujen jatkuva kehittäminen ja menestyksellinen hyödyntäminen organisaatioissa riippuukin kokonaisuudessaan useista tekijöistä, joissa organisoituminen ja vuorovaikutus ovat kriittisinä yhdistävinä elementteinä: Mitä ollaan ratkaisemassa ja mitä hyötyjä tavoitellaan? Mitkä ovat tekoälyratkaisujen tavoitteet ja mahdolliset haasteet ja kytkökset organisaation muuhun toimintaan? Varsinkin isoissa organisaatioissa on varottava ajautumasta koordinoimattomiin, keskenään jopa päällekkäisiin tai epäyhteensopiviin erillisiin kokeiluihin eri teknologioin ja eri ulkoisin yhteistyöpartnerein. (Finland’s AI Accelerator 2020.)

Onko tekoälyprototyypikehitys organisaatiossa ongelma-, teknologia- vai datalähtöistä? (Tyrväinen ym. 2018)

Tekoälyteknologian tarjoamien monien potentiaalisten hyötyjen ja mahdollisuuksien vastapainoksi on otettava kriittisesti huomioon myös siihen liittyvät riskit, epävarmuudet ja mahdolliset negatiiviset sivuvaikutukset. Yksi keskeinen tekoälypohjaisten sovellusten perusluonteeseen liittyvä ominaisuus on niiden mahdollisesti hankala selitettävyyttä ("explainability"). Onkin kaiken kaikkiaan arvioitava kriittisesti sitä, ovatko koneoppimis pohjaiset ratkaisut ylipäättään sopivia juuri niihin organisatorisiin tehtäviin, joita halutaan kehittää ja automatisoida (Storey ym. 2022). Nykyinen tekoäly/ML-teknologia ei välttämättä ole vielä riittävän kypsää kaikkiin tehtäviin.

Mikä merkitys on sillä, missä määrin tekoälyyn voi luottaa? (DeCamp & Tilburt 2019)

Yleisesti tarkastellen tekoälyn käyttöönotto voidaan nähdä osana julkishallinnon digitaalista transformaatiota. Kokonaisuudessaan se voi johtaa perustavanlaisiin uudelleensuunnitteluihin informaatiojärjestelmien arkkitehtuureissa ja niiden teknologisissa komponenteissa, kun dataa pyritään keräämään tehokkaasti ja hyödyntämään saumattomasti läpi organisatoristen prosessien. (Eom & Lee 2022.)

Nykyisin julkisen sektorin sote-organisaatiot tuottavat palveluitaan entistä enemmän yhteistyössä yksityisen sektorin toimijoiden kanssa. Monissa yrityksissä tekoälyteknologian käyttöönotossa on edetty jo pidemmän aikaa. Pk-sektorin yrityksissä uusien tekoälypohjaisten ohjelmistoratkaisujen soveltaminen on kuitenkin usein suuria yrityksiä vaikeampaa esimerkiksi puutteellisen osaamisen takia. Esimerkiksi Päijät-Hämeessä tehdyssä yrityskyselytutkimuksessa (vuonna 2020) kävi ilmi, että tekoälyn käyttö on vielä minimaalista, mutta yritykset näkevät tarvetta ja mahdollisuuksia sen hyödyntämiseen toiminnassaan. Tärkein syy tekoälyteknologiaratkaisujen käyttöön oli ajan säästö. (Sirviö 2021.)





## PÄIJÄT-HÄME

### Tekoälyratkaisujen teknologinen käyttöönotto organisaatiossa Päijät-Hämeessä

Päijät-Hämeessä on tehty tekoälypohjaisten ennusteratkaisujen kehitystä ennen tätä hanketta mm. ikäihmisten (yli 80-vuotiaat) palvelutarpeen ennakkoinnissa. Siinä lähtökohtana oli nk. koneoppimisen kanvaasi, jonka avulla määriteltiin tuotettavan ennustemallin perustietoja ja tavoitteita. Ennustemalli rakennettiin organisaation yhteiseen tietoaalustaan. Ennustemalli kehitettiin hyödyntämällä XGBoost-algoritmia.

Päijät-Hämeessä käynnistyi vuonna 2020 Tulevaisuuden sote-keskus -hanke osana silloista hyvinvointialueen (vuoden 2023 alusta PHHVA) valmistelua<sup>14</sup>. Siinä digitaalisuuteen liittyvät kehitystoimet ovat osa kaikkien tavoitteiden toteuttamista. Asiakastietojen hallinnan ja tiedolla johtamisen kehittämällä tavoitellaan parempaa kustannustietoisuutta ja uusia mahdollisuuksia kustannustehokkuuteen<sup>15</sup>.

Päijät-Soten Digistrategiassa (2021–2023) linjataan niin, että tekoäly ja tietojohdaminen ovat yhä useammin keskeisiä elementtejä ammattilaisten työssä. Lisäämällä automatisoituja prosesseja parannetaan nopeutta, tehokkuutta ja palvelujen tasalaatuisuutta.

Strateginen toimintakulttuurin muutos vaatii kuitenkin resursseja, minkä ymmärretään uusien ratkaisujen käyttöönottojen yhteydessä mahdollisesti aiheuttavan lyhytaikaisia prosessien hidastumisia. Käyttöönotot edellyttävätkin toimintaprosessien päivittämistä ja palveluiden jäsentämistä. Jatkuvaan kehitykseen osallistuvien yhdistämiseksi rakennetaan nk. digimentorien johtamia verkostoja.

Data on keskeinen elementti kaikissa tekoälyratkaisuisissa. Päijät-Sotessa on laadittu datastrategia ohjaamaan yhteistä tietojohdamista haastattelemalla työntekijöitä eri puolilta organisaatiota<sup>16</sup>. Keskeistä on järjestelmällistä datan keruuta ja kerättävän datan laatua, jotta sitä voidaan hyödyntää tehokkaasti tiedolla johtamisessa. Data-alusta kytkee ja yhdistää datavirrat tietopohjaksi, jonka päälle voidaan tehdä edistysellistä analytiikkaa mm. koneoppimisen keinoin. Tällaisen datastrategian toteuttaminen vaatii myös organisaation toimintakulttuurin kehittämistä, sillä kaikilta työntekijöiltä edellytetään uudenlaista tapaa nähdä tiedon arvo<sup>17</sup>. Data-asiantuntijoita tarvitaan erityisesti huolehtimaan datan laadun hallinnasta ja monitoroinnista sekä datan hallintaan liittyvistä velvoitteista, käytännöistä ja seurannasta noudattaen koko organisaation tietojohdamisen yhteistyömallia yhdenmukaisen tietopolitiikan mukaisesti<sup>18</sup>. Ylipäätään Päijät-Hämeessä pyritään kehittämään ja ylläpitämään data-alustaa siten, että sen avulla on mahdollista jalostaa datamassoista tietoa ymmärrettävään muotoon toiminnan hyödynnettäväksi tiedolla johtamisessa, erityisesti sote-toiminnanohjausjärjestelmissä (ERP)<sup>19</sup>.

Tekoälyratkaisujen käyttöönottoonkin liittyen Päijät-Hämeen organisaatiossa ollaan kehittämässä saatontamaprosessia työn tekemiseen ja kehittämiseen tarvittavien raporttien sisäiseen tilaamiseen tietojohdamisen tiimiltä<sup>20</sup>. Kustannustehokkuuden parantamiseksi tietotuotteiden käyttöä halutaan seurata tosiaikaisesti siten, että tarpeellimmat tietotuotteet pidetään jatkuvasti käytössä ja sellaiset, joita ei enää aktiivisesti hyödynnetä, poistetaan tarpeettomina.

14 <https://soteuudistus.fi/tulevaisuuden-sosiaali-ja-terveyskeskus-ohjelma/#ennaltaehk%C3%A4isy%20ja%20ennakointi>

15 <https://www.paijatha.fi/hankkeet/tulevaisuudensote/>

16 <https://paijat-sote.fi/sote-uudistus/laadukkaasta-tiedosta-syntyy-parempia-palveluja%E2%80%AF/>

17 <https://paijat-sote.fi/tietohallinnon-tyopaikat/tietojohdamisen-tiimi/>

18 <https://paijat-sote.fi/tyopaikat/> (Data-asiantuntija (2023))

19 <https://paijat-sote.fi/tietohallinnon-tyopaikat/tietojohdamisen-tiimi/>

20 <https://paijat-sote.fi/tietohallinnon-tyopaikat/tietojohdamisen-tiimi/>



## 4.4 Tekoälyteknologian organisatorisen osaamisen kehittäminen

Kun tekoälymalleja kehitetään ja otetaan käytännössä operatiiviseen sovelluskäyttöön organisaatioissa, tarvitaan monia, mahdollisesti uusiakin osaamisia (kompetensseja) ja niiden monialaisia yhdistelmiä. Jotta mallien toteuttaminen ja integrointi toiminnan tietojärjestelmiin olisi mahdollista ja tehokasta, on tärkeää tunnistaa kaikki tarvittavat osaamiset ja organisaation mahdolliset puutteet niissä. Osaamiset liittyvät monialaisesti sekä IT-teknoologiaan että varsinaiseen tekoälytietämykseen.

*Millaisia osaamiskehikoita organisaatiossa käytetään data-analytiikka-tehtäviin tarvittavien spesifisten taitojen arvioimiseen ja kehittämiseen? (Weiser ym. 2022)*

Sosiaalialan organisaatioissa on havaittu osaamispuutteiden olevan tyypillinen este tietojärjestelmien tuottaman tiedon hyödyntämisessä (Salovaara ym. 2021). Tekoälymallien tuottaman uudenlaisen ja -tasoisen tiedon tehokas hyödyntäminen vaatii entistä enemmän ja monipuolisempaa osaamista läpi organisaation. Yleisesti on tarve kehittää nykyisen työvoiman digitaalisia taitoja (Pote ym. 2021).

*Onko organisaatiossa sopiva yhdistelmä datatieteen ja ohjelmistokehityksen taitoja? (Myllyaho ym. 2022)*

Sote-alan työntekijöiden riittävä teknologian käyttöosaaminen ja tiedon hyödyntämisen taidot ovat olennaisia. On huolehdittava henkilökunnan jatkuvasta ajantasaisesta osaamisesta ja tietämyksestä tekoälyteknologiasta ja eri sovelluksista tekoälyavusteisessa ja -pohjaisessa toiminnassa. (Lähteenmäki ym. 2020.) Toistaiseksi useimmat nykyisistä työntekijöistä eivät liene saaneet perusteellista tekoälykoulutusta osana muodollisia opintojaan, joten henkilökunnan valmius tekoälyjärjestelmien soveltamiseen voi olla epävarmaa ja vaihdella suuresti (Shinners ym. 2021). Julkisen sektorin organisaatioissa on nykyään usein paineita vähentää kustannuksia ja mahdollisuus kokeneiden tekoälyasiantuntijoiden rekrytointiin voi olla käytännössä rajoitettua (Panch ym. 2019).

*Vaatiiko uusien tekoälyratkaisujen/järjestelmien käyttö ammattilaisilta muuta kuin tavanomaista ”digiosaamista”? (Nummela 2022)*

Tekoälymalleja kehitettäessä ja hyödynnettäessä ICT-sidonnaisten avainosaamisien identifiointia tulee tehdä läpi mallien linkkaaren ja palvelutuotannon arvoketjun: mallinnusvaiheessa, mallin toteuttamisessa ja integroimisessa operatiiviseen IT-ympäristöön, mallipohjaisen palvelun käyttämisessä ja hyödyntämisessä. Nykyään tyypillisenä puutteena voi olla analytiikkaosaaminen, jota tulisikin kehittää (Lähteenmäki ym. 2020).

*Miten organisaatiossa otetaan huomioon koneoppimisovelluskehityksen perinteisestä ohjelmistokehityksestä poikkeavat osaamisvaatimukset? (Kumeno 2019)*

Yleisesti tekoälyjärjestelmien käyttöönotto ja soveltaminen voi johtaa palvelutuotannon työskentelymallien uudelleenkonfigurointiin ja -suunnitteluun, jolloin saatetaan tarvita uusia osaamisia ja kapasiteetteja. On tärkeää, että tekoälypohjaisten järjestelmiä työssään käytännössä hyödyntävät kykenevät integroimaan ja sovittamaan niitä työskentelyrutiineihinsa. (Alami ym. 2021.)

Osaamisten kehittämistä voidaan tehdä järjestelmällisesti kokonaisarkkitehtuuri-kehityksen ohjaamana (Lohtander ym. 2021): MILLÄ: fyysiset tietomallit, tietovarannot; MITEN: tietojärjestelmäsalkku, teknologiaresurssit; MITÄ: tietojärjestelmäkartta, teknologiavalinnat; MIKSI: strategiset tavoitteet, arkkitehtuuriperiaatteet, standardisalkku, ohjaavat lait ja säädökset. Viimeisen kohtaan liittyen voidaan ja tuleekin laatia organisaation kyvykkyysskartta ottaen systemaattisesti huomioon asetetut kehittämisvaatimukset ja -tavoitteet. Osaamisten kehittämistä on hallittava tietoisesti, kun tekoälyteknologia ja sen soveltaminen ovat voimakkaasti kehittyvässä ja laajenevassa vaiheessa. Tarvitaan ajantasainen tilannekuva, tilannetietoisuus ja tilanneymmärrys (Mykkänen ym. 2021).

Kaiken kaikkiaan uusien teknologisten ratkaisujen – erityisesti tekoälypohjaisten – kehittäminen ja käyttöönotto sosiaali- ja terveydenhuollon organisaatioissa vaatii osaamista ja taitoja kaikilla organisaation tasoilla ja henkilöstöryhmissä sekä varsinaisten palvelujen operatiivisessa toimituksessa että niiden hallinnassa ja organisaation johdossa (National Health Service 2019). Osaamisen kehittämisen tulisikin olla organisaatiossa kokonaisvaltaista. Tekoälyteknologian soveltamisen laajentuminen organisaatiossa johtaa siihen, että data-, tekoäly- ja analytiikkaosaamiset ovat lopulta kaikille ainakin osittain relevantteja osaamisia, koska kaikkien organisaatiossa tulisi ymmärtää datan rooli ja merkitys omassa työssään ja kontekstissaan kyetäkseen käyttämään sitä hyödyllisesti. (Finland's AI Accelerator 2020.)



## PÄIJÄT-HÄME

### Tekoälyteknologian ja tietojärjestelmien organisatoristen osaamisten kehittäminen Päijät-Hämeessä

Päijät-Soten Digistrategian (2021–2023) visio linjasi koko organisaation digitaalisten kyvykkyyksien ja osaamisten pääsuuntaviivat ja lähiaikojen kehittämistavoitteet. Digitaalisten palveluiden kehittämisen teknisten kyvykkyyksien tasoa pyritään nostamaan.

Tietojohdamisen tiimissä kehitetään osaamisia toteuttaa tekoälypohjaisia ennustemalleja hyödyntää matemaattisia malleja. Osaamiskehityksessä tunnistettuja pääalueita on sekä tiedon johtamisessa eli datan hallinnasta että tiedolla johtamisesta eli tiedon hyödyntämisessä jalostamalla datamassoista tietoa ymmärrettävään muotoon<sup>21</sup>. Muutosohjelmassa 2020–2024 tietojohdamisen kehittäminen on uuden johtamisjärjestelmän peruspilareita, ja esimerkiksi uusilta datainsinööreiltä odotetaan osaamista automaattioratkaisuista ja koneoppimisesta sekä analytiikasta ja raportoinnista<sup>22</sup>. Datan ja tiedonhallinnan periaatteiden ja käytänteiden lisäksi data-asiantuntijoilta edellytetään myös ymmärrystä sääntelyn vaikutuksista ja sen huomioonottamista datan hallinnan kehittämistyössä<sup>23</sup>.

Tietojohdamisen datakyvykkyyksien vahvistamista tehdään pilvipohjaisen data-alustan ja tietovarastoinnin kehittämisessä. Teknologiavalintojen mukaisesti haettavilta datainsinööreiltä odotetaan osaamista mm. MS Azuresta, Databricksistä ja Snowflakesta<sup>24</sup>. Data-analytiikoilta odotetaan osaamista analytiikka- ja raportointiratkaisuista, kuten Power BI<sup>25</sup>.

Myös sote-järjestelmien ICT-kehityksen puolelta raportointijärjestelmien (Power BI) tuntemusta pidetään tarpeellisenä<sup>26</sup>. ICT-hankepäälliköillä tulee olla osaamista ja näkemystä sote-palvelutuotantoon liittyvästä teknologiasta sekä teknologia- ja kokonaisarkkitehtuurin hallinnasta ja kehittämisestä – myös monitoimittajaympäristössä<sup>27</sup>.

Palvelutuotantoon liittyen ICT-kehittäjien vastuulla ovat organisaation palvelutuotantoprosesseja tukevien sovellusten suunnittelu ja kehittäminen systemaattisesti ohjatun palvelutuotannon mahdollistamiseksi yhdessä sote-ammattilaisten ja palvelumuotoilun kanssa<sup>28</sup>. Tavoitteena on luoda kyvykkyyksiä ammattilaisille hallita tehtäviä koko ajan tietoisena tilannekuvasta.

Yleisesti organisaation digitaalisten kyvykkyyksien kehittäminen vaikuttaa koko organisaation toimintakulttuuriin. Työkulttuurissa edistetään digitaalisten työkalujen ja työtapojen käyttöönottoa<sup>29</sup>. Uutena organisatorisena roolina onkin nk. digimentori, joka on ”osaamisen edelläkävijä, tiedon välittäjä ja muutokseen kannustaja”.

21 <https://paijat-sote.fi/tietohallinnon-tyopaikat/tietojohdamisen-tiimi/>

22 <https://paijat-sote.fi/tyopaikat/> (Data engineer, tietojohdaminen (2022) • Data-asiantuntija (2023) • Data engineer, tietojohdaminen (2022) • Data analyst, tietojohdaminen (2022) • ICT-asiantuntija, sosiaalihuollon sovellukset (2022))

23 <https://paijat-sote.fi/tyopaikat/> (Data engineer, tietojohdaminen (2022) • Data-asiantuntija (2023) • Data engineer, tietojohdaminen (2022) • Data analyst, tietojohdaminen (2022) • ICT-asiantuntija, sosiaalihuollon sovellukset (2022))

24 <https://paijat-sote.fi/tyopaikat/> (Data engineer, tietojohdaminen (2022) • Data-asiantuntija (2023) • Data engineer, tietojohdaminen (2022) • Data analyst, tietojohdaminen (2022) • ICT-asiantuntija, sosiaalihuollon sovellukset (2022))

25 <https://paijat-sote.fi/tyopaikat/> (Data engineer, tietojohdaminen (2022) • Data-asiantuntija (2023) • Data engineer, tietojohdaminen (2022) • Data analyst, tietojohdaminen (2022) • ICT-asiantuntija, sosiaalihuollon sovellukset (2022))

26 <https://paijat-sote.fi/tyopaikat/> (Data engineer, tietojohdaminen (2022) • Data-asiantuntija (2023) • Data engineer, tietojohdaminen (2022) • Data analyst, tietojohdaminen (2022) • ICT-asiantuntija, sosiaalihuollon sovellukset (2022))

27 <https://paijat-sote.fi/tietohallinnon-tyopaikat/> (ICT-projektipäällikkö, rakentamishankkeet (2022) • ICT-kehittäjä, tuotannonohjaus (2022) • Digimentori (2022))

28 <https://paijat-sote.fi/tietohallinnon-tyopaikat/> (ICT-projektipäällikkö, rakentamishankkeet (2022) • ICT-kehittäjä, tuotannonohjaus (2022) • Digimentori (2022))

29 <https://paijat-sote.fi/tietohallinnon-tyopaikat/> (ICT-projektipäällikkö, rakentamishankkeet (2022) • ICT-kehittäjä, tuotannonohjaus (2022) • Digimentori (2022))

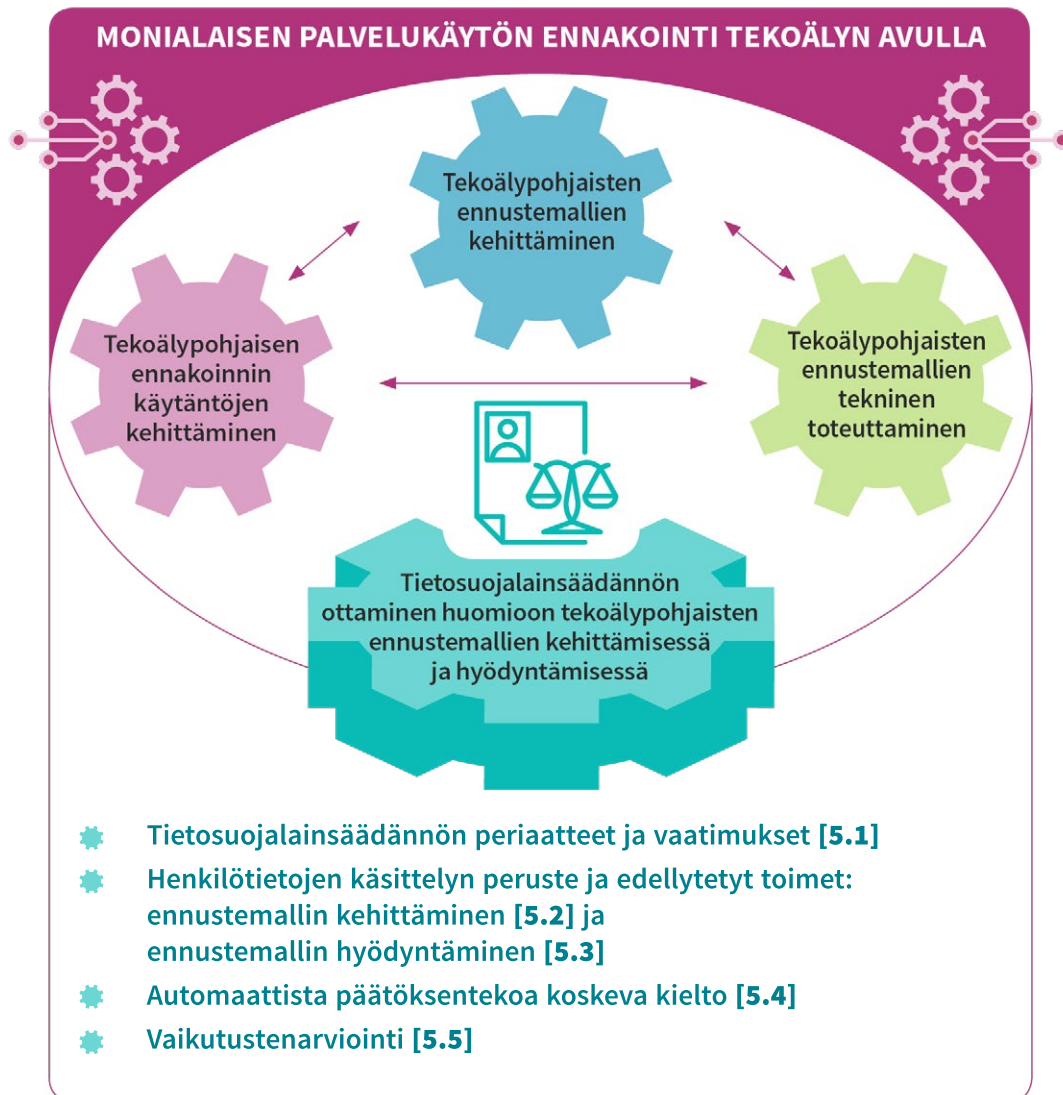


# TIETOSUOJA- LAINSÄÄDÄNNÖN OTTAMINEN HUOMIOON TEKOÄLYPOHJAISTEN ENNUSTEMALLIEN KEHITTÄMISESSÄ JA HYÖDYNTÄMISESSÄ



Sosiaali- ja terveysalan ennustemallien kehittäminen ja hyödyntäminen edellyttää tyypillisesti sosiaali- ja terveyspalvelujen asiakkaiden henkilötietojen käsittelyä. Jotta ennustemallien kehittäminen ja hyödyntäminen on tietosuojalainsäädännön kannalta mahdollista, on henkilötietojen käsittelylle oltava tietosuojalainsäädännössä hyväksytyt peruste ja sen tulee muuten täyttää lainsäädännön vaatimukset.

Kuviossa 7 esitetään keskeisiä tietosuojalainsäädännön vaatimuksia, jotka tulee ottaa huomioon tekoälypohjaisten ennustemallien kehittämisessä ja hyödyntämisessä.



**Kuvio 7.** Tietosuojalainsäädännön vaatimusten ottaminen huomioon tekoälypohjaisten ennustemallien kehittämisessä ja hyödyntämisessä

## 5.1 Lainsäädäntö ja periaatteet

Euroopan parlamentin ja neuvoston asetus (EU) 2016/679 luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta (’tietosuoja-asetus’, ’TSA’) tuli sovellettavaksi 25.5.2018. Asetusta sovelletaan laajasti henkilötietojen käsittelyyn, ja se määrittää henkilötietojen käsittelyn yleiset edellytykset EU:ssa.

Henkilötietojen käsittely on käsitteenä laaja, ja kattaa lähes kaikki toimenpiteet, jotka kohdistuvat henkilötietoihin. Henkilötietojen käsittelyllä tarkoitetaan ”toimintoa tai toimintoja, joita kohdistetaan henkilötietoihin tai henkilötietoja sisältäviin tietojoukkoihin joko automaattista tietojenkäsittelyä käyttäen tai manuaalisesti, kuten tietojen keräämistä, tallentamista, järjestämistä, jäsentämistä, säilyttämistä, muokkaamista tai muuttamista, hakua, kyselyä, käyttöä, tietojen luovuttamista siirtämällä, levittämällä tai asettamalla ne muutoin saataville, tietojen yhteensovittamista tai yhdistämistä, rajoittamista, poistamista tai tuhoamista” (TSA:n 4 artiklan 2 kohta). Asetuksella suojellaan luonnollisten henkilöiden oikeutta henkilötietojen suojaan sekä vahvistetaan säännöt, jotka koskevat henkilötietojen vapaata liikkuvuutta ja luonnollisten henkilöiden suojelua henkilötietojen käsittelyssä. Suomessa tietosuoja laki (1050/2018) täsmentää ja täydentää tietosuoja-asetusta ja käsittelee sen kansallista soveltamista. Siinä säädetään muun muassa erityisten henkilötietoryhmien käsittelystä, tietosuojavaltuutetun toiminnasta ja tietojenkäsittelyn erityistilanteista, kuten esimerkiksi henkilötietojen käsittelystä tieteellisiä tutkimustarkoituksia varten.

Tietosuojavelvoitteiden kohteena tietosuoja-asetuksessa ovat kaikki henkilötietojen käsittelystä vastaavat ja käsittelytoimia suorittavat tahot, jotka on tietosuoja-asetuksessa jaoteltu rekisterinpitäjiin ja henkilötietojen käsittelijöihin. Eräs asetuksen keskeisistä ominaispiirteistä on asetuksen 5 artiklaan kirjattujen tietosuojaperiaatteiden merkitys. Tietosuoja-asetuksessa määritellyt henkilötietojen käsittelyä koskevat periaatteet ovat: (a) lainmukaisuus, kohtuullisuus ja läpinäkyvyys; (b) käyttötarkoitussidonnaisuus; (c) tietojen minimointi; (d) täsmällisyys; (e) säilytyksen rajoittaminen; ja (f) eheys ja luottamuksellisuus. Näitä henkilötietojen käsittelyä koskevia periaatteita on noudatettava aina henkilötietoja käsiteltäessä. Periaatteiden noudattamisen ohella rekisterinpitäjän on nimenomaisesti pystyttävä osoittamaan se, että periaatteita on noudatettu, mikä korostaa etukäteisen suunnittelun ja henkilötietojen käsittelyn huolellisen dokumentoinnin merkitystä.

Laki sosiaali- ja terveystietojen toissijaisesta käytöstä (552/2019, ’toisilaki’), tuli voimaan 1.5.2019. Toisilaisissa annetaan tietosuoja-asetusta täydentävät säännökset, kun henkilötietoja käsitellään seuraaviin käyttötarkoituksiin, vaikka tietoja ei olisi alun perin tallennettu näissä tarkoituksissa: tilastointi, tieteellinen tutkimus, kehittämis- ja innovaatio toiminta, opetus, tietojohdaminen, sosiaali- ja terveydenhuollon viranomaisohjaus ja -valvonta, sekä viranomaisen suunnittelu- ja selvitystehtävä. Näistä toisilain mahdollistamista käyttötarkoituksista olennaisimpia tämän käsikirjan kannalta ovat tieteellinen tutkimus (38 §) ja tietojohdaminen (41 §). Toisilaki pyrkii lupakäsittelyn ja tietojen yhdistämisen sujuvoittamiseen. Sosiaali- ja terveysalan tietolupaviranomainen (Findata) vastaa tietojen kokoamisesta, yhdistämisestä, esikäsittelystä ja luovuttamisesta toissijaiseen käyttöön, sekä tekee päätökset koskien tietolupia ja tietopyyntöjä. Myös toisilaisissa määritetyt rekisterinpitäjät voivat tehdä muun muassa tietolupapäätöksiä koskien niiden omia aineistojaan. Findata myös ylläpitää tietoturvallista käyttöympäristöä ja vastaa julkaistavien tulosten anonymisoinnista.

## 5.2 Ennustemallien kehittäminen toisiolain perusteella

Kun potilas- ja asiakastietoja käytetään ennustemallien kehittämiseen (esim. koneoppimismallin opettamiseen), tietojen käsittelylle tulee olla lainmukainen peruste. Vaikka tietosuoja-asetus voi eräissä tilanteissa mahdollistaa ennustemallien kehittämisen (esim. suostumuksen perusteella), tarjoaa toisiolaki usein oikeudellisesti selkeämmät perusteet ennustemallien kehittämiseen. Toisiolain mahdollistamista potilas- ja asiakastietojen käyttötarkoituksista yleisimmin ennustemallien kehittämiseen voivat soveltua tieteellinen tutkimus (38 §) ja tietojohdaminen (41 §). Alla tarkastellaan niiden soveltumisen edellytyksiä ja haasteita ennustemallien kehittämisen näkökulmasta tarkemmin.

### 5.2.1 Tieteellinen tutkimus

Toisiolaki mahdollistaa asiakastietojen ja muiden toisiolain 6 §:ssä tarkoitettujen organisaatioiden henkilötietojen käytön tietoluvan nojalla tieteellistä tutkimusta varten (38 §).

Toisiolaissa ei määritellä tieteellisen tutkimuksen käsitettä, mutta lain esitöissä (HE 159/2017 vp, s. 123) viitataan TSA:n johdanto-osan 159 kappaleeseen, jonka mukaan tieteellisen tutkimuksen käsitettä tulee tulkita laajasti, niin että se sisältää muun muassa teknologian kehittämisen, soveltavan tutkimuksen, perustutkimuksen ja yksityisin varoin rahoitettavan tutkimuksen. Toisiolain osalta ei vielä ole oikeuskäytännössä otettu kantaa tieteellisen tutkimuksen käsitteeseen, mutta muun henkilötietoja ja salassa pidettäviä tietoja koskevan lainsäädännön osalta oikeuskirjallisuudessa ja oikeuskäytännössä tieteellisen tutkimuksen pääpiirteitä on katsottu olevan tutkimuksen tekijän riittävä tieteellinen pätevyys, tutkimuksen perustuminen asianmukaiseen tutkimussuunnitelmaan, sekä tutkimuksen tieteelliset tavoitteet ja tutkimuksen autonomisuus ja kriittisyys (Mäenpää 2016, KHO:2013:181). Tieteellisen tutkimuksen ei ole katsottu olevan riittävällä tasolla ennen maisterintutkintoa (Turun HO 15.1.2019).

Ennustemallien kehittämistä varten voi täyttää tieteellistä tutkimusta koskevat edellytykset. Ei ole kuitenkaan selvää, milloin kyseessä on lainsäädännössä tarkoitettu tieteellisestä tutkimuksesta ja milloin sellaisesta muusta toiminnasta, johon tietolupaa ei ole mahdollista saada. Tietoluvan myöntämistä koskevan päätöksen tekee joko tietolupaviranomainen (Findata) tai toisiolain 6 §:ssä mainittu yksittäinen rekisterinpitäjä, jotka siten arvioivat sitä, onko kyseessä toisiolaissa tarkoitettu tieteellinen tutkimus vai ei. Findata on verkkosivuillaan julkaissut ohjeita koskien tieteellisen tutkimuksen pääkriteerejä (ks. *Palvelut ja ohjeet*, 22.6.2023) ja antaa neuvontaa tietoluvan hakijoille.

Tietolupaan perustuva yksilötasoinen tietoaineisto luovutetaan luvansaajalle tietoturvallisessa käyttöympäristössä (toisiolain 51 §:n 3 mom.). Tietoturvallisella käyttöympäristöllä tarkoitetaan toisiolaissa ”teknistä, organisatorista ja fyysistä tietojen käsittelyn toimintaympäristöä, jossa tietoturvaluus on varmistettu asianmukaisin hallinnollisin ja teknisin toimin” (3 §:n 11 kohta). Ensisijainen käsittely-ympäristö lainsäädännön mukaan on Findatan oma tietoturvallinen käyttöympäristö (Kapseli). Tietoturvallinen käyttöympäristö on suljettu ympäristö, joten haasteeksi ennustemallien tuotantokäytön ja jakokehityksen kannalta voi osoittautua tietoluvan nojalla saadun aineiston pohjalta



tuotettujen tulosten (esimerkiksi koneoppimismalli) ulos tuominen käyttöympäristöstä. Muuhun tietoturvalliseen käyttöympäristöön aineistoja voidaan luovuttaa vain, jos se on välttämätöntä (toisiolain 20:n § 3 mom. ja 51 §). Findata määrittelee tietoturvallisia käyttöympäristöjä koskevat tarkemmat vaatimukset (toisiolain 24 §:n 2 mom; ks. myös Määräys: Muiden palveluntarjoajien tietoturvalisille käyttöympäristöille asetettavat vaatimukset (findata.fi)).

**Tieteellinen tutkimus henkilötietojen toissijaisena käyttötarkoituksena voi soveltua ennustemallien kehittämiseen, jos tutkimuksen tekijällä on riittävä tieteellinen pätevyys, tutkimus perustuu asianmukaiseen tutkimussuunnitelmaan ja kehittämisen tavoite on tieteellinen. Kun henkilötietojen käyttötarkoitus on tieteellinen tutkimus, on otettava huomioon tietoturvalisien käyttöympäristön vaatimukset koskien esimerkiksi aineiston ulos tuomista käyttöympäristöstä.**

## 5.2.2 Tietojohtaminen

Toisiolaki mahdollistaa sosiaali- ja terveydenhuollon palvelunantajan<sup>30</sup> sen omassa toiminnassa syntyneiden ja sen omiin rekistereihin tallennettujen tietojen käsittelyn tietojohtamista varten ilman tietolupaa (41 §:n 1 mom.). Tietojohtamisella tarkoitetaan toisiolaissa ”tiedon käsittelemistä palvelunantajan asiakas-, palvelu- ja tuotantoprosesseissa toiminnan, tuotannon ja talouden ohjauksen, johtamisen ja päätöksenteon tukena” (3 §:n 5 kohta).

Tietojen käsittelyn tulee lisäksi olla välttämätöntä palvelunantajan palvelutoiminnan ”tuottamista, seuranta, arviointia, suunnittelua, kehittämistä, johtamista ja valvontaa varten” (toisiolain 41 §:n 1 mom.). Toisiolain 41 §:n tulkinnassa on myös huomiotava TSA:n 9 artiklan 2 kohdan h alakohda, koska 41 § on kyseisessä alakohdassa tarkoitettua sellaista kansallista sääntelyä, joka on sallittua tietyin edellytyksin. Kohdan mukaan tällaisen käsittelyn tulee olla tarpeen ”ennalta ehkäisevää tai työterveydenhuoltoa koskevia tarkoituksia varten, työntekijän työkyvyn arvioimiseksi, lääketieteellisiä diagnooseja varten, terveys- tai sosiaalihuollollisen hoidon tai käsittelyn suorittamiseksi taikka terveys- tai sosiaalihuollon palvelujen ja järjestelmien hallintoa varten [...] noudattaen 3 kohdassa esitettyjä edellytyksiä ja suojatoimia”. Nämä suojatoimet edellyttävät, että tietoja käsittelee tai niiden käsittelystä vastaa henkilö, jota sitoo lakisääteinen salassapitovelvollisuus.

Ennustemallin kehittäminen voi olla mahdollista palvelunantajan toimesta toisiolaissa tarkoitettuna tietojohtamisena yllä mainittujen edellytysten täytyessä. Tietojohtamisen käsite on laaja, mutta se ei välttämättä kata kaikkia mahdollisia ennustemalleihin liittyviä pyrkimyksiä. Esimerkiksi toisiolain esitöiden mukaan tietojohtaminen ei sovellu yksilöä koskevaan päätöksentekoon (HE 159/2017 vp, s. 81). Kun henkilötietojen

30 ’Palvelunantajalla’ tarkoitetaan toisiolain mukaan viranomaista, joka järjestää, tuottaa tai toteuttaa sosiaali- tai terveydenhuoltoa taikka sosiaali- tai terveyspalveluja. Palvelunantajalla tarkoitetaan myös Yksityisistä sosiaalipalveluista annetun lain (922/2011) tai yksityisistä terveydenhuollosta annetun lain (152/1990) mukaista yksityistä palveluntuottajaa (3 §:n 13 kohta).

toissijaisena käyttötarkoituksena on tietojohdaminen, on tapauskohtaisesti varmistettava, että tietojen käyttö on toisilain määritelmän mukaista tietojohdamista.

Tietojohdaminen soveltuu vain palvelunantajan omassa toiminnassa syntyneisiin tietoihin tai sen omiin rekistereihin tallennettuihin tietoihin (toisilain 41 §:n 1 mom.). Uudenmaan hyvinvointialueita ja Helsingin kaupunkia koskevan erillisratkaisun johdosta Uudenmaan hyvinvointialueet ja Helsingin kaupunki voivat käsitellä ja yhdistellä tunnisteellisesti kunkin omassa toiminnassa syntyneitä tai kunkin omiin rekistereihin tallennettuja potilastietoja, sekä HUS-yhtymän hyvinvointialueiden lukuun tuottamia tietoja (41 §:n 3 mom.). Tietojohdaminen ei sovellu, mikäli on tarpeen käsitellä muita kuin edellä mainittuja tietoja. Muiden palvelunantajien tietoja on tietolupaviranomaiselle (Findata) tehtävän tietopyynnön perusteella mahdollista saada ja käyttää vertailutarkoituksia varten aggregoituna tilastotietona (41 §:n 2 mom.).

**Tietojohdaminen voi soveltua henkilötietojen toissijaiseksi käyttötarkoitukseksi ennustemallien kehittämiseen, jos ennustemallien kehittämisessä on kyse palvelunantajan palvelutoimintaa varten välttämättömästä henkilötietojen käsittelystä, käsittely tapahtuu TSA:n 9 artiklan 2 kohdan h alakohdan nojalla, ja tietoja käytetään palvelunantajan prosesseissa toiminnan ohjauksen ja johtamisen tukena. Kun käyttötarkoitus on tietojohdaminen, saa palvelunantaja käsitellä ja yhdistellä vain omassa toiminnassa syntyneitä tai omiin rekistereihin tallennettuja tietoja.**

## 5.3 Ennustemallien hyödyntäminen

Kun ennustemallia hyödynnetään sote-toiminnassa, potilas- ja asiakastietojen käsittely (mm. syötteenä koneoppimismalliin) edellyttää lainmukaista perustetta. Alla tarkastellaan toisilain tietojohdamisen käyttötarkoituksen soveltumista ennustemallien hyödyntämiseen sekä niitä edellytyksiä, joilla asiakkaan/potilaan suostumuksen perusteella henkilötietoja voidaan käsitellä ennustemallien hyödyntämiseen.

### 5.3.1 Tietojohdaminen

Toisiolaki mahdollistaa sosiaali- ja terveydenhuollon palvelunantajan tietojen käsittelyn tietojohdamista varten.

Kuten on esitetty edellä luvussa 5.2.2, tietojohdaminen soveltuu vain palvelunantajan omassa toiminnassa syntyneisiin tai sen omiin rekistereihin tallennettuihin tietoihin (toisilain 41 §:n 1 mom.). Mikäli on tarpeen käsitellä muita tietoja, tietojohdaminen ei sovellu. Tällöin tietojen käsittelylle on oltava jokin toinen peruste. Tietojen käytön on lisäksi oltava välttämätöntä palvelunantajan palvelutoiminnan kannalta (41 §:n 1 mom.). Ennustemallien hyödyntämisen yhteydessä tapahtuvalla tietojen käsittelyllä on näin ollen oltava yhteys palvelunantajan vastuulla toteutettavaan palvelutoimintaan. Toisilain 41 §:n tulkinnassa on myös huomioitava edellä mainittu TSA:n 9 artiklan 2 kohdan h alakohta.

Kyseen on lisäksi oltava edellä mainitun toisiolain määritelmän mukaisesta tietojohdamisesta. Erityisesti tulee huomata, kuten yllä todettiin, tietojohdaminen ei sovellu yksilöä koskevaan päätöksentekoon (HE 159/2017 vp, s. 81). Tietojohdamisessa on sen sijaan kyse palvelunantajan sisäisestä johtamisesta ja ohjauksesta, kuten esimerkiksi tuotantoprosessien kehittämisestä, potilaan hoidon yhteensovittamisesta ja asiakassegmentoinnista (HE 159/2017 vp, s. 92, 132). On kuitenkin huomattava, että tietojohdaminen ei sovellu sellaiseen asiakas- ja potilastietojen käsittelyyn, jossa käsittelyn tarkoituksena on tehdä yksilöitä koskevia päätöksiä. Tarkka rajanveto voi kuitenkin olla vaikeaa, koska toisiolaissa ei määritellä tarkemmin tietojohdamisen sisältöä tässä suhteessa.

**Tietojohdaminen voi soveltua henkilötietojen toissijaisena käyttötarkoituksena ennustemallien hyödyntämiseen, jos ennustemallien hyödyntämisessä on kyse palvelunantajan palvelutoimintaa varten välttämättömästä henkilötietojen käsittelystä, käsittely tapahtuu TSA:n 9 artiklan 2 kohdan h alakohdan nojalla, ja tietoja käytetään palvelunantajan prosesseissa toiminnan ohjauksen ja johtamisen tukena. Kun käyttötarkoitus on tietojohdaminen, saa palvelunantaja käsitellä ja yhdistellä vain omassa toiminnassa syntyneitä tai omiin rekistereihin tallennettuja tietoja. Tietojohdaminen ei sovellu yksilöä koskevaan päätöksentekoon.**

### 5.3.2 Suostumus ja sen edellytykset

Aina kun käsitellään henkilötietoja, on käsittelylle oltava lainmukainen peruste. TSA:n määrittämistä käsittelyn oikeusperusteista laajimmin ja todennäköisimmin ennustemallien hyödyntämiseen soveltuu asiakkaan/potilaan antama suostumus (TSA:n 6 artiklan 1 kohdan a alakohta). Suostumuksen tulee täyttää kaikki sitä koskevat TSA:ssa asetetut edellytykset.

Suostumus on annettava ennakkoon, ennen henkilötietojen käsittelyn aloittamista. TSA:ssa ei kategorisesti kielletä jälkipästä suostumusta, mutta TSA:n 6 artiklasta ja johdanto-osan 40 kappaleesta ”[...] seuraa loogisesti, että pätevä laillinen peruste on oltava olemassa ennen tietojen käsittelyn aloittamista. Suostumus olisi siksi annettava ennen käsittelytoimintaa” (Euroopan tietosuojaneuvoston, ‘ETN’, suuntaviivat 05/2020, s. 21). Mikäli suostumus annetaan kirjallisessa ilmoituksessa, joka koskee myös muita asioita, on suostumuksen antamista koskeva pyyntö esitettävä helposti ymmärrettävässä muodossa ja selkeästi muista asioista erillään (TSA:n 7 artiklan 2 kohta). Suostumus on myös voitava peruuttaa milloin tahansa, ja peruuttamisen on oltava yhtä helppoa kuin suostumuksen antamisen (7 artiklan 3 kohta).

Pätevä suostumus on yksilöity, tietoinen, aidosti vapaaehtoinen ja yksiselitteinen tahdonilmaisu (TSA:n 4 artiklan 11 kohta). Jokaisen edellä mainitun osatekijän on toteuduttava.

- **Yksilöity suostumus** vaatii, että jos henkilötietojen käsittelyllä on useita eri tarkoituksia, rekisteröidyn on annettava suostumuksensa jokaista käsittelyn tarkoitusta varten. Suostumuksen tulee siis kattaa kaikki samaa tarkoitusta varten toteutettavat käsittelytoimet (TSA:n johdanto-osan 32 kappale).

- **Tietoinen suostumus** edellyttää, että rekisteröity tietää vähintäänkin rekisterinpitäjän henkilöllisyyden ja henkilötietojensa käsittelyn tarkoitukset (TSA:n johdanto-osan 42 kappale). ETN:n suuntaviivoissa (05/2020) mainitaan lisäksi seuraavat vaatimukset: tieto siitä, mitä ja minkä tyyppisiä tietoja kerätään; tieto oikeudesta peruuttaa suostumus; tarvittaessa selvitys siitä, käytetäänkö tietoja TSA:n 22 artiklan 2 kohdan c alakohdan mukaisesti automatisoituihin päätöksiin; sekä tieto mahdollisista tiedonsiirron riskeistä, johtuen tietosuojan tason riittävyttä koskevan päätöksen ja 46 artiklan asianmukaisen suojatoimien puuttumisesta. Edellä mainitut vaatimukset koskevat suostumuksen vähimmäisisältöä, eli tapauskohtaisesti katsottuna saattaa olla tarpeellista tarjota rekisteröidylle enemmänkin tietoa. Olennaista ETN:n mukaan on, että rekisteröity ymmärtää henkilötietojensa käsittelyn ja käsittelyn tarkoituksen. Tiedot voi antaa kirjallisesti tai suullisesti. Rekisterinpitäjän osoitusvelvollisuudesta johtuen kirjallinen, dokumentoitava muoto on suositeltava (Korpisaari ym. 2022, 153–154).
- **Aidosti vapaaehtoinen suostumus** edellyttää, että rekisteröity voi peruuttaa suostumuksensa tai kieltäytyä suostumuksen antamisesta ilman, että peruuttamisella tai kieltäytymisellä on rekisteröidylle haitallisia vaikutuksia (TSA:n johdanto-osan 42 kappale). Esimerkkeinä haitallisista vaikutuksista ETN mainitsee suuntaviivoissaan (05/2020) lisäkustannukset, pakottamisen, harhaanjohtamisen ja pelottelun. Myös rekisteröidyn ja rekisterinpitäjän välillä oleva mahdollinen epäsuhta vaikuttaa suostumuksen vapaaehtoisuuteen, esimerkiksi jos rekisterinpitäjänä on viranomainen (TSA:n johdanto-osan 43 kappale). TSA:ssa ei kategorisesti kielletä viranomaisilta suostumukseen perustuvaa henkilötietojen käsittelyä, mutta käsittelylle pitäisi pääsääntöisesti löytyä jokin muu peruste (ETN suuntaviivat 05/2020, s. 9). Pääsääntönä vallan tasapainon suhteen voidaan pitää sitä, että suostumus on pätevä vain silloin, jos suostumuksen antamiseen ei liity ”kyvyttömyyttä harjoittaa vapaata tahtoa” (s. 10).
- **Yksiselitteinen tahdonilmaisu** tarkoittaa yksiselitteistä ilmaisuja, jolla rekisteröity antaa lausuman tai toteuttaa selkeästi suostumusta ilmaisevan toimen (TSA:n 4 artiklan 11 kohta). Suostumusta ei voi antaa passiivisella toimella, eli esimerkiksi vaikenemalla tai valmiiksi rastitetulla ruudulla. Rekisteröidyn on siis toimittava tavalla, joka selvästi osoittaa rekisteröidyn hyväksyvän henkilötietojen käsittelyn (TSA:n johdanto-osan 32 kappale). Esimerkiksi tapauksessa Planet49 Euroopan unionin tuomioistuimien katsoi, että suostumus ei ollut pätevä, koska lupa oli saatu jo valmiiksi rastitetulla ruudulla (asia C-673/17, Planet49). Myös oikeuskirjallisuudessa on katsottu, että passiivinen suostumus ei ole pätevä (Bensoussan ym. 2017). Suostumuksessa pitää olla kyse aktiivisesta toiminnasta (ETN suuntaviivat 05/2020, s. 20).

Ennustemallien hyödyntämisen yhteydessä tapahtuvassa henkilötietojen käsittelyssä on kyse erityisiin henkilötietoryhmiin kuuluvien tietojen (mm. terveyttä koskevat tiedot) käsittelystä. Erityisiin henkilötietoryhmiin kuuluvien tietojen käsittely vaatii yllä mainittujen suostumuksen edellytysten lisäksi sitä, että suostumus on nimenomainen (TSA:n 9 artiklan 2 kohdan h alakohta). Nimenomaisuus viittaa tapaan, jolla rekisteröity ilmaisee suostumuksensa. Käytännössä nimenomaisella suostumuksella tarkoitetaan sitä, että rekisteröity antaa yksiselitteisen, suostumusta ilmaisevan lausuman (ETN suuntaviivat 05/2020, s. 22).

Pätevälläkään suostumuksella ei voi syrjäyttää TSA:n 5 artiklan tietosuojaperiaatteita (Korpisaari ym. 2022, 117). Lainmukaisuus, kohtuullisuus ja läpinäkyvyys vaatii, että henkilötietojen käsittelylle on lainmukainen peruste, henkilötietoja ei käytetä väärin, ja että rekisteröity tietää, miten hänen tietojaan käsitellään. Käyttötarkoitussidonnaisuus edellyttää, että henkilötietoja kerätään vain tiettyä, nimenomaista ja laillista tarkoitusta varten, eikä tietoja käsitellä myöhemmin näiden tarkoitusten kanssa yhteensopimattomalla tavalla. Tietojen minimointi tarkoittaa, että vain käsittelyn tarkoituksen kannalta oleellisia ja tarpeellisia henkilötietoja käsitellään. Täsmällisyys vaatii, että henkilötiedot ovat täsmällisiä, ja virheelliset henkilötiedot oikaistaan tai poistetaan. Säilytyksen rajoittaminen edellyttää, että henkilötietoja säilytetään vain niin kauan kuin on tarpeen käsittelyn tarkoitusten toteuttamista varten. Eheys ja luottamuksellisuus tarkoittaa, että henkilötietoja on käsiteltävä niin, ettei tietojen eheys tai luottamuksellisuus vaarannu. (TSA:n 5 artiklan 1 kohdan a–f alakohdat). Lisäksi on huomioitava myös muut tietosuojalainsäädännön vaatimukset (kuten esimerkiksi vaikutustenarvioinnin tekeminen – tästä ks. tämän luvun kappale 5.5.).

On tärkeää muistaa, että TSA:n suostumus eroaa potilastietojen ja sosiaalihuollon asiakastietojen luovutusta koskevasta suostumuksesta. Potilaslain (785/1992) mukaan potilasasiakirjoihin sisältyvät tiedot ovat salassa pidettäviä (13 §:n 1 mom.), eikä niitä tule luovuttaa sivulliselle ilman potilaan kirjallista suostumusta (13 §:n 2 mom.). Sosiaalihuollon asiakaslain (812/2000) mukaan sosiaalihuollon asiakirjat ovat salassa pidettäviä (14 §:n 1 mom.), eikä asiakirjasta saa antaa tietoja ilman asiakkaan nimenomaista suostumusta (16 §). Näin ollen **TSA:n mukainen nimenomainen suostumus käsittelyn oikeusperusteena ei välttämättä yksin ratkaise sitä, saadaanko potilas- tai asiakastietoja käyttöön ennustemallien hyödyntämistä varten.**

Henkilötietojen käsittely ennustemallia hyödynnettäessä voi perustua rekisteröidyn eli asiakkaan tai potilaan nimenomaiseen suostumukseen, jos kaikki suostumuksen osatekijät (yksilöity, tietoinen, aidosti vapaaehtoinen ja yksiselitteinen tahdonilmaisu) ja edellytykset (suostumus on pyydettyä ennakkoon helposti ymmärrettävässä muodossa, ja se on voitava pystyä peruuttamaan) täyttyvät. Rekisterinpitäjä vastaa suostumuksen edellytysten toteutumisesta, joten rekisterinpitäjyyden määräytymiseen on tässä yhteydessä kiinnitettävä erityistä huomiota. Myös muut tietosuojalainsäädännön vaatimukset on huomioitava.

## 5.4 Automaattinen päätöksenteko

Eräs mahdollinen oikeudellinen ongelma tekoälypohjaisten ennustemallien käyttöön-otossa ja hyödyntämisessä sote-alalla liittyy tietosuojasetuksen 22 artiklaan, jonka 1 kohdan on tulkittu sisältävän automatisoitujen yksittäispäätösten yleisen kiellon (esim. WP 251 rev.01, s. 21 ja julkisasiamies Pikamäen ratkaisuehdotus 16.3.2023 Euroopan unionin tuomioistuimen asiassa C-634/21, SCHUFA Holding AG ym., kohta 31). Kielto ei koske kaikkea tekoälyyn perustuvaa tai tekoälyavusteista ratkaisutoimintaa, vaan ainoastaan tietyt edellytykset täyttävää kokonaan automatisoitua päätöksentekoa.

Tietosuoja-asetuksen 22 artiklan mukainen automatisoitu yksittäispäätös on päätös,

- joka perustuu henkilötietojen pelkästään automaattiseen käsittelyyn, kuten profilointiin, ja
- jolla on tietosuoja-asetuksessa tarkoitettua rekisteröityä koskevia oikeusvaikutuksia tai joka muuten vaikuttaa vastaavalla tavalla merkittävästi rekisteröityyn.

Jos nämä edellytykset eivät täyty, 22 artikla ei sovellu. Kääntäen ilmaisten artikla ei siis estä tekoälypohjaisen ennustemallin hyödyntämistä, jos järjestelmä on mahdollista suunnitella ja toteuttaa siten, ettei automatisoituja yksittäispäätöksiä muodostu.

Päätös ei perustu pelkästään automaattiseen käsittelyyn, jos joku luonnollinen henkilö – esimerkiksi lääkäri tai muu terveydenhuollon ammattihenkilö – osallistuu ennustemallin hyödyntämisen yhteydessä erilaisiin toimenpiteisiin liittyvien ratkaisujen tekemiseen merkityksellisellä tavalla niin, ettei esimerkiksi potilaan kontaktointia tai muuta toimenpidettä koskevan päätöksen voida katsoa syntyneen pelkästään automaattisen käsittelyn perusteella. Ammattihenkilöllä on oltava aidosti valtuudet ja edellytykset tarvittaessa muuttaa järjestelmän esittämää suositusta. Mikä tahansa muodollinen osallistuminen prosessiin tai järjestelmän esittämien toimenpiteiden rutiininluontoinen hyväksyminen ei siis riitä.

Toiseksi 22 artiklan mukainen kielto ei sovellu, jos ennustemallia käytettäessä ei synny lainkaan yksilöihin vaikuttavia päätöksiä tai näiden päätösten vaikutukset eivät ole merkittäviä. Jos ennustemallia käytetään esimerkiksi ainoastaan tulevan resurssien tarpeen arviointiin tai muihin vastaaviin toimenpiteisiin, joiden vaikutukset eivät kohdistu yksittäisiin rekisteröityihin (asiakkaisiin), 22 artikla ei muodostu esteeksi. Ennustemalleja hyödynnettäessä ei tyypillisesti synny suoranaisia oikeusvaikutuksia saavia päätöksiä. Muiden vastaavien vaikutusten osalta merkittävyyskynnyksen voivat sen sijaan ylittää esimerkiksi päätökset, jotka vaikuttavat tietyn henkilön mahdollisuuksiin saada terveydenhuoltopalveluja (WP 251 rev.01, s. 23). Jos ennustemallin käyttö vaikuttaa siihen, minkälaisia terveydenhuolto- tai sosiaalipalveluita rekisteröidyille tarjotaan tai ei tarjota, tämä edellytys voi täytyä, vaikka mitään muodollista hallintopäätöstä hoidon tai muun palvelun epäämisestä ei tehtäisikään.

Apulaistietosuojavaikuttettu on kahdessa vuonna 2022 antamassaan ratkaisussa (dnrot 3895/83/2022 ja 6482/186/2020) tulkinnut, että eräissä ennaltaehkäisyyn ja ennakointiin tähtäävissä potilastietojen käsittelytilanteissa todennäköisesti muodostuisi automatisoituja yksittäispäätöksiä. Arvioitavina olleilla järjestelmillä oli tarkoitus tunnistaa terveytensä puolesta riskissä olevia potilaita tietojen automatisoidun analyysin avulla ja saattaa näitä hoidon piiriin ottamalla heihin yhteyttä. Varsinaisia hoitopäätöksiä ei ollut tarkoitus automatisoida. Automatisoituja yksittäispäätöksiä katsottiin todennäköisesti syntyvän niiden rekisteröityjen osalta, joita ei poimittu lähempään tarkasteluun tai ammattihenkilöiden suorittamien jatkotoimenpiteiden kohteeksi. Näiden henkilöiden osalta järjestelmän ilman luonnollisen henkilön osallistumista antama riskiarvio muodostui lopulliseksi päätökseksi, joka apulaistietosuojavaikuttetun näkemyksen mukaan saattoi vaikuttaa heidän mahdollisuuksiinsa saada hoitoa ja oli siten mahdollisesti vaikutuksiltaan merkittävyyskynnyksen ylittävä.

**Tekoälypohjaisten ennustemallien hyödyntämisen yhteydessä voi muodostua tietosuoja-asetuksen 22 artiklassa tarkoitettuja ja lähtökohtaisesti kiellettyjä automatisoituja yksittäispäätöksiä, jos 1) päätöksen tekoprosessiin ei merkityksellisellä tavalla osallistu terveydenhuollon ammattihenkilö tai muu luonnollinen henkilö, ja 2) jos tällaisella päätöksellä on oikeusvaikutuksia tai muita merkityksellisiä vaikutuksia yksittäiselle potilaalle/asiakkaalle. Tällainen vaikutus voi olla esimerkiksi se, että henkilön mahdollisuudet saada hoitoa tai sosiaalihuollon palveluita heikentyvät. Automatisoituja yksittäispäätöksiä voi muodostua erityisesti sellaisten henkilöiden osalta, jotka ennustemallin mukaan eivät olisi jatkotoimenpiteiden tarpeessa, koska näiden henkilöiden osalta automaattinen profilointi voi jäädä viimeiseksi käsittelyvaiheeksi ja siten muodostua lopulliseksi päätökseksi.**

Jos ennustemallia käytettäessä syntyy automatisoituja yksittäispäätöksiä, ennustemallin hyödyntäminen voi kuitenkin olla sallittua jonkin 22 artiklan 2 kohdassa määritellyn poikkeusperusteiden nojalla. Näitä perusteita on kolme. Kiellosta voidaan poiketa, jos automatisoitu päätös on

- a. välttämätön rekisteröidyn ja rekisterinpitäjän välisen sopimuksen tekemistä tai täytäntöönpanoa varten;
- b. hyväksytty rekisterinpitäjään sovellettavassa unionin oikeudessa tai jäsenvaltion lainsäädännössä; tai
- c. perustuu rekisteröidyn nimenomaiseen suostumukseen.

Jos poikkeusperusteena on sopimuksen tekeminen tai täytäntöönpano tai rekisteröidyn nimenomainen suostumus, rekisterinpitäjän on toteutettava erityisiä suojaotoimia rekisteröidyn oikeuksien ja vapauksien sekä oikeutettujen etujen suojaamiseksi. Suojaotoimiin kuuluvat näissä tapauksissa vähintäänkin oikeus vaatia, että tiedot käsittelee rekisterinpitäjän puolesta luonnollinen henkilö, sekä oikeus esittää kantansa ja riitauttaa päätös. Jos automatisoitujen yksittäispäätösten hyväksyttävyyys taas perustuu erilliseen lainsäädäntöön, suojaotoimet on vahvistettava lainsäädännössä.

Sote-alan ennustemalleja hyödynnettäessä käsitellään tyypillisesti ns. erityisiin henkilötietoryhmiin kuuluvia terveyttä koskevia tietoja. Automatisoitujen yksittäispäätösten tekemiseksi näiden tietojen käsittely on mahdollista vain, kun se perustuu rekisteröidyn nimenomaiseen suostumukseen (TSA:n 9 artiklan 2 kohdan a alakohta) tai kun näiden tietojen käsittely on tarpeen tärkeää yleistä etua koskevasta syystä unionin oikeuden tai jäsenvaltion lainsäädännön nojalla (TSA:n 9 artiklan 2 kohdan g alakohta).

Koska sote-alan ennustemallien hyödyntämisestä ja automaattisesta päätöksenteosta ei nykyisin ole erityislainsäädäntöä, poikkeusperusteista voi tulla kyseeseen käytännössä vain nimenomainen suostumus, jonka olisi siis katettava myös se, että automatisoidun päätöksenteon yhteydessä käsitellään terveystietoja. Nimenomaisen suostumuksen tulee täyttää tietosuoja-asetuksessa asetetut edellytykset, joita on käsitelty edellä alaluvussa 5.3.2.

Tyypillinen riski viranomaisen ja yksityishenkilön välisessä suhteessa on se, ettei esimerkiksi julkisia palveluita tarvitsevalla henkilöllä välttämättä ole aitoa mahdollisuutta kieltäytyä henkilötietojensa käsittelystä. Tästä syystä viranomaisten suorittaman henkilötietojen käsittelyn tulisi pääsääntöisesti perustua muihin käsittelyperusteisiin kuin rekisteröidyn suostumukseen. Jos jonkin julkisen palvelun saamisen edellytyksenä on suostumus henkilötietojen käsittelyyn, joka ei liity välittömästi kyseisen palvelun tarjoamiseen, kyse ei lähtökohtaisesti ole aidosti vapaaehtoisesta suostumuksesta.

Kun ennustemallien käyttöön liittyvästä henkilötietojen käsittelystä ei ole erikseen säädetty laissa, ennustemallin käytöstä kieltäytyminen ei saisi johtaa siihen, että joltakulta evättäisiin pääsy palveluihin, joihin hänellä on oikeus ja joita hän tarvitsee. Ennustemallin avulla tapahtuva poiminta ei voi olla ainoa reitti palvelujen piiriin pääsemiseksi. Suostumuksen vapaaehtoisuuden ja siten henkilötietojen käsittelyn laillisuuden kannalta ongelmallista voi olla jo se, jos ennustemallin käytöstä kieltäytyminen tosiasiallisesti heikentää henkilön mahdollisuuksia saada julkisia palveluita merkittävästi.

On syytä huomata, että tekoälypohjaisen ennustemallin käyttöönotto sote-alalla lähtökohtaisesti edellyttää tietosuoja-asetuksen 35 artiklassa tarkoitettujen vaikutustenarvioinnin toteuttamista riippumatta siitä, onko kyse automatisoiduista yksittäispäätöksistä vai muunlaisesta (osittain) automaattisesta päätöksenteosta. Tätä velvollisuutta käsitellään tarkemmin alaluvussa 5.5.

**Vaikka automatisoituja yksittäispäätöksiä syntyisi, ennustemallin hyödyntäminen voi olla laillista, jos päätöksenteon kohteena olevat asiakkaat tai potilaat (rekisteröidyt) siihen nimenomaisesti suostuvat. Suostumuksen tulee kattaa terveystietojen automaattinen käsittely ja siihen perustuva päätöksenteko. Rekisterinpitäjän on tällöin myös huolehdittava erityisistä suojatoimista rekisteröityjen oikeuksien, vapauksien ja oikeutettujen etujen suojaamiseksi. Automatisoidut yksittäispäätökset voitaisiin mahdollistaa myös erityisellä lainsäädännöllä, jossa samalla määriteltäisiin riittävät suojatoimenpiteet.**

## 5.5 Vaikutustenarvioinnin laatiminen

Ennustemallien kehittämisen ja hyödyntämisen osalta tulee ottaa huomioon erityisesti edellä käsitellyt kysymykset eli se, mihin lainmukaiseen perusteeseen henkilötietojen käsittely perustuu, ja automaattisen päätöksenteon kieltö. Näiden lisäksi tietosuojalainsäädännön muut vaatimukset tulee täyttää. Näihin lukeutuu erityisesti se, että sote-alan ennustemalleja kehitettäessä ja hyödynnettäessä usein on tarpeen tehdä TSA:n 35 artiklassa säädetty tietosuojaa koskeva vaikutustenarviointi (Data Protection Impact Assessment). Vaikutustenarviointia edellytetään, koska rekisterinpitäjän on etukäteisesti arvioitava henkilötietojen käsittelystä rekisteröidylle potentiaalisesti aiheutuvia riskejä. Riskien arviointi liittyy TSA:n riskiperusteiseen lähestymistapaan ja rekisterinpitäjän osoitusvelvollisuuteen (Korpisaari ym. 2022, 400).



### 5.5.1 Milloin vaikutustenarviointi on tehtävä?

Vaikutustenarviointi on tehtävä silloin, kun henkilötietojen käsittely todennäköisesti aiheuttaa korkean riskin luonnollisen henkilön oikeuksien ja vapauksien kannalta. Riskien suhteen rekisterinpitäjän on huomioitava henkilötietojen käsittelyn luonne, laajuus ja asiayhteys, sekä käsittelyn tarkoitukset (TSA:n 35 artiklan 1 kohta). Henkilön oikeudet ja vapaudet on ymmärrettävä laajasti, kattaen tietosuojaan liittyvien oikeuksien lisäksi myös esimerkiksi liikkumisvapauden ja sananvapauden (Korpisaari ym. 2022, 400). Vaikutusten vakavuus voi olla vähäistä, kohtalaista, merkittävää tai kriittistä (Tietosuojan vaikutustenarvioinnin ohje). Vaikutustenarviointi on tehtävä ennen henkilötietojen käsittelyn aloittamista (TSA:n 35 artiklan 1 kohta). Vaikutustenarvio tulee eräissä tilanteissa liittää Findatan tietolupahakemukseen.

Vaikutustenarviointia vaaditaan erityisesti muun muassa silloin, kun rekisteröityjen henkilökohtaisia ominaisuuksia arvioidaan automaattiseen päätöksentekoon perustuen niin, että käsittelyllä on rekisteröityä koskevia oikeusvaikutuksia tai muita vastaavalla tavalla merkittäviä vaikutuksia (TSA:n 35 artiklan 3 kohdan a alakohta). Tämä muistuttaa TSA:n 22 artiklassa säänneltyä automaattisen päätöksenteon kieltoa. Vaikutustenarviointia koskeva 35 artikla on kuitenkin laajempi, ja koskee automaattisten päätösten lisäksi myös sellaisia päätöksiä, jotka vain perustuvat henkilötietojen automaattiseen käsittelyyn (Korpisaari ym. 2022, 402). Vaikutustenarviointia vaaditaan myös silloin, jos kyseessä on erityisiin henkilötietoryhmiin kohdistuva, laajamittainen käsittely (TSA:n 35 artiklan 3 kohdan b alakohta). Laajamittaisuutta arvioidessa on huomioitava rekisteröityjen lukumäärä, käsiteltävien henkilötietojen määrä, henkilötietojen käsittelytoimen kesto tai pysyvyys, ja käsittelyn toimen maantieteellinen ulottuvuus (WP 248 rev.01, s. 12).

Vaikutustenarviointi on usein tarpeen tehdä, koska sote-alan ennustemallien kehittämisen ja hyödyntämisen yhteydessä usein käsitellään terveystietoja ja muita erityisten henkilötietojen ryhmään kuuluvia tietoja ja varsinkin kehitysvaiheessa kyse voi olla hyvin laajamittaisesta henkilötietojen käsittelystä, ja koska käytettäessä ennustemalleja sote-alalla kyse saattaa olla automaattisesta päätöksenteosta tai muuten potilaiden/asiakkaiden etuihin merkittävästi vaikuttavasta toiminnasta (mm. palveluiden saaminen). On huomattava, että koska vaikutustenarviointi auttaa rekisterinpitäjää noudattamaan tietosuojalainsäädäntöä, tietosuojaryhmä suosittelee sen laatimista silloinkin, kun vaikutustenarvioinnin laatimisen tarpeellisuus on epäselvää (WP 248 rev.01, s. 9).

### 5.5.2 Mitä vaikutustenarvioinnin tulee sisältää?

Vaikutustenarvioinnin tulee sisältää vähintäänkin seuraavat seikat: kuvaus suunnitelluista käsittelytoimista ja käsittelyn tarkoituksista; arvio käsittelytoimien tarpeellisuudesta ja käsittelyn oikeasuhtaisuudesta käsittelyn tarkoituksiin nähden; arvio rekisteröityjen oikeuksia ja vapauksia koskevista riskeistä; sekä kuvaus suunnitelluista toimenpiteistä, joilla puututaan riskeihin ja joilla varmistetaan henkilötietojen suoja, sekä osoitetaan, että TSA:ta on noudatettu (TSA:n 35 artiklan 7 kohta). Vaikutustenarvioinnissa on tarkasteltava suunniteltuja toimenpiteitä, suojatoimia ja mekanismeja, joiden avulla on tarkoitus lieventää korkeaa riskiä ja varmistaa henkilötietojen suoja (TSA:n johdanto-osan 90 kappale). Muutoin TSA:ssa ei määritellä vaikutustenarvioinnin tarkkaa sisältöä.

Tietosuojavaltuutetun toimiston ohjeessa (Tietosuojan vaikutustenarvioinnin ohje) listataan vaikutustenarvioinnin sisällön määrittämisen avuksi seuraavat vaikutustenarviointia koskevat ydinkysymykset:



- Miksi ja miten henkilötietoja aiotaan käsitellä?
- Mitä henkilötietoja aiotaan käsitellä?
- Miten suunniteltu henkilötietojen käsittely vaikuttaa rekisteröityihin, kun käsittely sujuu suunnitelmien mukaisesti?
- Mikä voi mennä pieleen suunnitellussa henkilötietojen käsittelyssä?
- Miten todennäköistä on, että jokin menee pieleen?
- Miten yllä mainitun todennäköisyyttä voi vähentää?



Vaikutustenarviointia tulee tarkastella jatkuvana prosessina kertaluontoisen tehtävän sijaan (WP 428 rev.01, s. 16–17). Mikäli henkilötietojen käsittelytoimien sisältämät riskit muuttuvat, rekisterinpitäjän on tarkasteltava uudelleen, tapahtuuko tietojen käsittely vaikutustenarvioinnin mukaisesti (TSA:n 35 artiklan 11 kohta).



## PÄIJÄT-HÄME

### Päijät-Hämeessä kehitetyn ennustemallin suhtautuminen tietosuojalainsäädäntöön

Toisiolaki mahdollistaa sosiaali- ja terveydenhuollon palvelunantajan sen omassa toiminnassa syntyneiden ja sen omiin rekistereihin tallennettujen tietojen käsittelyn tietojohdamista varten (41 §:n 1 momentti). Tietojohdamisella tarkoitetaan toisio-laissa (3 §:n 5 kohta) ”tiedon käsittelemistä palvelunantajan asiakas-, palvelu- ja tuotantoprosesseissa toiminnan, tuotannon ja talouden ohjauksen, johtamisen ja päätöksenteon tukena.”

### Ennustemallien hyödyntäminen johtamisessa

Päijät-Hämeen pilotissa osa pohdituista ja luvussa 2 esitetyistä ennustemallien hyödyntämistavoista (mm. johtamiseen selvästi tarkoitettut, ei yksittäisiä asiakkaita koskevat toimintatavat) ovat sellaisia, että asiakas- ja potilastietojen käsittely voi olla mahdollista toisio-lain 41 §:ssä tarkoitettuna tietojohdamisena. Tulee kuitenkin varmistaa, että tietojen käyttö on yllä mainitun toisio-lain määritelmän mukaista tietojohdamista ja että tietojen käsittely on ”välttämätöntä palvelunantajan vastuulla toteutettavan palvelutoiminnan tuottamista, seuranta, arviointia, suunnittelua, kehittämistä, johtamista ja valvontaa varten.” (41 § 1 mom). On myös otettava huomioon, että toisio-lain 41 §:n tulkinnassa on huomioitava tietosuoja-asetuksen 9 artiklan 2 kohdan h alakohta, koska 41 § on kyseisessä alakohdassa tarkoitettua sellaista kansallista sääntelyä, joka on sallittua tietyin edellytyksin.

Vaikka ennustemallien hyödyntämistavoissa olisi kyse sellaisten ennusteiden tekemisestä, jotka koskevat asiakasryhmiä (ei lainkaan yksilöjä) eivätkä toimet muutenkaan johda yksilöjä koskeviin toimenpiteisiin, on tarpeen huomata, että toisio-laissa tarkoitettun tietojohdamisen ei ole tarkoitettu koskevan yksilöjä koskevaa päätöksentekoa. Tällaisiin käyttötarkoituksiin on oltava muu käsittelyperuste, ja toimintaa on lisäksi arvioitava automaattista päätöksentekoa koskevan kiellon kannalta (ks. seuraavalla sivulla).



Lisäksi on huomattava, että tietojohtaminen (41 §:n 1 mom.) soveltuu vain palvelunantajan omassa toiminnassa syntyneisiin tietoihin tai sen rekistereihin tallennettuihin tietoihin. Muiden palveluntarjoajien tietoja on kuitenkin tietolupaviranomaiselle tehtävän tietopyynnön perusteella mahdollista saada ja käyttää vertailutarkoituksia varten aggregoituna tilastotietona (41 §:n 2 mom.).

### Ennustemallien hyödyntäminen palveluohjauksessa

Osa pohdituista ja luvussa 2 esitetyistä ennustemallien hyödyntämistavoista koskisi yksilöjä, joihin kohdistettaisiin erilaisia toimenpiteitä ennustetiedon perusteella tai ainakin ennustetiedon avulla. Toisilain tietojohtaminen (41 §) ei soveltune asiakas- ja potilastietojen käsittelyyn tällaisia yksilöjä koskevia hyödyntämistarkoituksia varten, koska kyse olisi yksilöjä koskevista ennusteista ja toimista (ei tietojohtamisesta). Tarkka rajanveto on kuitenkin vaikeaa, koska toisilaisissa ei määritellä tarkemmin tietojohtamisen sisältöä tässä suhteessa, vaan sen valmisteluaineistosta ilmenee, ettei tietojohtamisen ole tarkoitus ulottua yksilöä koskeviin päätöksiin.

Henkilötietojen käsittelyyn ennustemallien hyödyntämisen yhteydessä voi kuitenkin soveltua jokin toinen käsittelyperuste. TSA:n määrittämistä käsittelyperusteista laajimmin ja todennäköisimmin ennustemallien hyödyntämiseen sote-alalla soveltuva henkilötietojen käsittelyn oikeusperuste on sellainen asiakkaan/potilaan antama suostumus, joka täyttää kaikki sitä koskevat TSA:ssa asetetut edellytykset. Suostumuksen on oltava mm. aidosti vapaaehtoinen. Tämä tarkoittaa muun ohella sitä, ettei suostumuksen antamisesta kieltäytyminen saa johtaa siihen, että rekisteröity jää ilman hänelle kuuluvia julkisia palveluita. Suostumus on annettava ennakkoon, ennen käsittelyn aloittamista. Monialaisen palvelukäytön ennakoinnissa kyse on sellaisista tiedoista (mm. terveyttä koskevat tiedot), että suostumuksen on tavanomaisten edellytysten (TSA:n 7 artikla ja 4 artiklan 11 kohta) lisäksi oltava nimenomainen (TSA:n 9 artikla). Lisäksi tulee ottaa huomioon TSA:n tietosuojaperiaatteiden noudattaminen ja TSA:n asettamat muut vaatimukset (kuten tarvittaessa vaikutustenarvioinnin tekeminen).<sup>31</sup>

Osa pohdituista ennustemallien hyödyntämistavoista on lisäksi sellaisia, että niitä on tarpeen arvioida automaattista päätöksentekoa koskevan kiellon kannalta (TSA:n 22 artikla). Vaikka osassa pohdituista hyödyntämistavoista ilmeisesti olisi kyse toiminnasta, jossa ammattihenkilöt aidosti arvioivat toimien tarpeellisuutta kunkin asiakkaan/potilaan kohdalla, on huomattava että täysin automaattisia päätöksiä on Suomessa katsottu syntyvän siltä osin, kuin luonnollisen henkilön suorittamaa harkintaa ei tehdä joidenkin yksilöiden osalta.<sup>32</sup> Tällaiset täysin automaattisesti tehdyt päätökset voivat vaikuttaa sillä tavoin merkittävästi yksilöihin (esim. palvelujen saamiseen), että niihin soveltuu 22 artiklan kielto. Tästä syystä ennustemallien hyödyntäminen on tarpeen suunnitella ja toteuttaa siten, ettei mitään osin synny tällaisia täysin automaattisesti tehtäviä päätöksiä (ilman luonnollisen henkilön aitoa harkintaa), joilla voi olla yksilöihin merkittäviä vaikutuksia. Vaihtoehtoisesti on mahdollista asiakkaan/potilaan nimenomaisen suostumuksen perusteella tehdä häntä koskevia päätöksiä täysin automaattisestikin, mutta tälle asetetaan TSA:ssa eräitä lisäedellytyksiä (mm. yksilöiden oikeuksia ja etuja turvaavia toimenpiteitä).

31 Lisäksi on huomioitava, että asiakas- ja potilastiedot ovat salassa pidettäviä ja että niiden luovuttaminen ennustemallien hyödyntämistä varten on mahdollista vain tietyillä perusteilla (esim. asiakkaan/potilaan tietojen luovuttamista koskeva suostumus/lupa).

32 Apulaistietosuojavaltuutetun päätös 3895/83/2022 ja 6482/186/2020.

# Lähteet

Achte, P., & Hiltunen-Toura, M. (2020) Tulevaisuuden Sote-keskus ja digitaaliset palvelut (Sote-uudistus). STM.

Aerts, A., & Bogdan-Martin, D. (2021) Leveraging data and AI to deliver on the promise of digital health. *International Journal of Medical Informatics*, 150, 104456.  
<https://doi.org/10.1016/j.ijmedinf.2021.104456>

Akava, Elinkeinoelämän keskusliitto EK, Kirkon työmarkkinalaitos, Kunta- ja hyvinvointialueyönantajat KT, Suomen Ammattiliittojen Keskusjärjestö SAK, Suomen Yrittäjät, STTK ry, & Valtion työmarkkinalaitos VTML (2022) Vuoropuhelu digitalisaation hyödyntämisestä työpaikoilla ja toimialoilla.

Alami, H., Lehoux, P. Denis, J.-L., Motulsky, A., Petitgand, C., Savoldelli, M., Rouquet, R., Gagnon, M.-P., Roy, D., & Fortin, J.-P. (2021) Organizational readiness for artificial intelligence in health care: insights for decision-making and practice. *Journal of Health Organization and Management*, 35(1), 106–114. DOI 10.1108/JHOM-03-2020-0074

Ammenwerth, E., Iller C., & Mahler, C. (2006) IT adoption and the interaction of task, technology and individuals: a fit framework and a case study. *BMC Med Inform Decis Mak*, 6(3).  
<https://doi.org/10.1186/1472-6947-6-3>

Antikainen, E., Linnosmaa, J., Umer, A., Oksala, N., Eskola, M., van Gils, M., Hernesniemi, J., & Gabbouj, M. (2023) Transformers for cardiac patient mortality risk prediction from heterogeneous electronic health records. *Scientific Reports* 2023 13:1, 13(1), 1–10. <https://doi.org/10.1038/s41598-023-30657-1>

Anttila, H., Pulli, K., Alarotu, E., Noro, A., Kehusmaa, S., Sinervo, T., Luoma, M.-L., Niemelä, M., & Lähteenmäki, J. (2020) Kotona asumisen teknologiat ikäihmisille (KATI) 2021-2023: Ohjelma ja hankeopas. Terveystyön ja hyvinvoinnin laitos.

Apulaistietosuojavaltuutetun päätös 23.6.2022, dnro 3895/83/2022.

Apulaistietosuojavaltuutetun päätös 23.8.2022, dnro 6482/186/2020.

29 artiklan mukainen tietosuojatyöryhmä (2017) Ohjeet tietosuojaa koskevasta vaikutustenarvioinnista ja keinoista selvittää ”liittykö käsittelyyn todennäköisesti” asetuksessa (EU) 2016/679 tarkoitettu ”korkea riski”. WP 248 rev.01 17/FI. Annettu 4. huhtikuuta 2017, viimeksi tarkistettu ja hyväksytty 4. lokakuuta 2017. (WP 248 rev.01)

29 artiklan mukainen tietosuojatyöryhmä (2018) Suuntaviivat automatisoiduista yksittäispäätöksistä ja profiloinnista asetuksen (EU) 2016/679 täytäntöön panemiseksi. WP 251 rev.01 17/FI. Annettu 3. lokakuuta 2017, viimeksi tarkistettu ja hyväksytty 6. helmikuuta 2018. (WP 251 rev.01)

Bardsley, M., Billings, J., Dixon, J., Georghiou, T., Lewis, G. H., & Steventon, A. (2011) Predicting who will use intensive social care: case finding tools based on linked health and social care data. *Age and Ageing*, 40(2), 265–270. <https://doi.org/10.1093/ageing/afq181>

Bensoussan, A., Henrotte, J., Gallardo, M., & Fanti, S. (2017) *General Data Protection Regulation: Text, commentaries and practical guidelines*. Wolters Kluwer.

Bosch, J., Holmström Olsson, H., & Crnkovic, I. (2021) Engineering AI Systems: A Research Agenda. In: A. K. Luhach & A. Elçi (Eds.), *Artificial Intelligence Paradigms for Smart Cyber-Physical Systems* (pp. 1–19). IGI Global. <https://doi.org/10.4018/978-1-7998-5101-1>

Cresswell, K., Callaghan, M., Khan, S., Sheikh, Z., Mozaffar, H., & Sheikh, A. (2020) Investigating the use of data-driven artificial intelligence in computerised decision support systems for health and social care: A systematic review. *Health Informatics Journal*, 26(3), 2138–2147.  
<https://doi.org/10.1177/1460458219900452>



DeCamp, M, & Tilburt, J. C. (2019) Why we cannot trust artificial intelligence in medicine. *The Lancet Digital Health*, 1(8), E390. [https://doi.org/10.1016/S2589-7500\(19\)30197-9](https://doi.org/10.1016/S2589-7500(19)30197-9)

Desouza, K. C., Dawson, G. S., & Chenok, D. (2020) Designing, developing, and deploying artificial intelligence systems: Lessons from and for the public sector. *Business Horizons*, 63(2), 205–213. <https://doi.org/10.1016/j.bushor.2019.11.004>

DigiFinland. (2021) Virta-arkkitehtuuri. [https://digifinland.fi/wp-content/uploads/2021/04/Virta-arkkitehtuuri\\_11.3.2021.pdf](https://digifinland.fi/wp-content/uploads/2021/04/Virta-arkkitehtuuri_11.3.2021.pdf)

DigiFinland (12/3/21) Lisätietoa Virta-hankkeesta. <https://digifinland.fi/toimintamme/virta-hanke/virta-hanke-lisatietoa/>

Ekholm, S., & Kinnunen, U-M. (2016) Tietojärjestelmän käyttöönottoa tukevat teoreettiset mallit terveydenhuollossa. *Finnish Journal of EHealth and EWelfare*, 8(2-3), 63–73.

Eom, S.-J., & Lee, J. (2022) Digital government transformation in turbulent times: Responses, challenges, and future direction. *Government Information Quarterly*, 39(2), 101690. <https://doi.org/10.1016/j.giq.2022.101690>

Euroopan tietosuojaneuvosto (2020) Asetuksen 2016/679 mukaista suostumusta koskevat suuntaviivat 05/2020. Versio 1.1, hyväksytty 4. toukokuuta 2020. (ETN suuntaviivat 05/2020)

Euroopan parlamentin ja neuvoston asetukset (EU) 2016/679 luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta.

European Commission (2017) Tools and methodologies to assess integrated care in Europe. Report by Expert Group on health Systems Performance Assessment. European Commission, Luxembourg. 13.4.2017.

Euroopan unionin tuomioistuimien asia C-673/17, Planet49, tuomio (suuri jaosto) 1.10.2019.

Euroopan unionin tuomioistuimien asia C-634/21, SCHUFA Holding AG ym., julkisasiamies Pikamäen ratkaisuehdotus 16.3.2023.

Findatan määräys 1/2022, Dnro THL/214/14.00.07/2022, liite 1, 'Tietoturvallisen käyttöympäristön vaatimukset' (19.1.2022). Liite-1-Määräys-1-2022-Tietoturvallisen-kayttoympariston-vaatimukset.pdf (findata.fi).

Finland's AI Accelerator (2020) Opas tekoälyn jatkuvaan hyödyntämiseen. Suomen Tekoälykiihdyttämö • Teknologiateollisuus.

Ghassemi, M., Naumann, T., Schulam, P., Beam, A. L., Chen, I. Y. & Ranganath, R. (2019) Practical guidance on artificial intelligence for health-care data. *The Lancet Digital Health*, 1(4), E157–E159. [https://doi.org/10.1016/S2589-7500\(19\)30084-6](https://doi.org/10.1016/S2589-7500(19)30084-6)

Gilbert, H. (2016) Supporting integration through new roles and working across boundaries. *TheKingsFund*, June 2016.

Goodwin, N., Nolte, E. & Stein, V. (2017) What is Integrated Care? In: Amelung, V., Stein, V., Goodwin, N., Balicer, R., Nolte, E. & Suter, E. (Eds.), *Handbook Integrated care*. Springer International publishing AG, 3-24.

Haula, T & Peltola, M. (2020) Monialainen tai kallis palvelujen käyttö Suomessa. Teoksessa: Koivisto J, Tiirinki H, toim. Monialaisen palvelutarpeen tunnistaminen sosiaali-, terveys- ja työvoimapalveluissa. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 2020:38, 16-40.

Hallituksen esitys 159/2017 vp eduskunnalle laiksi sosiaali- ja terveystietojen toissijaisesta käytöstä sekä eräiksi siihen liittyviksi laeiksi.

Helander, N., Ahonen, O., Houhala, K., & Jääskeläinen, A. (2020) Tiedolla johtaminen julkisella sektorilla: käytännön tapauksia eri hallinnon aloilta. *Focus Localis*, 48(3), 22–42.

Hietapakka L., Karjalainen P., Liukko E., Sinervo T. Monialaista palvelujen yhteensovittamista tarvitsevien asiakkaiden sekä työ- ja toimintakyvyn tukitarpeiden tunnistaminen Suomessa. Tutkimuksesta tiiviisti 52, 2019. Terveyden ja hyvinvoinnin laitos, Helsinki.



Kempele (2021) Kempeleen vanhuspoliittinen ohjelma vuosille 2021-2025 (2030). Kempeleen kunta, 2021.

KHO:2013:181., 583/2/12.

Jiang F., Jiang Y., Zhi H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017) Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. DOI: 10.1136/svn-2017-000101

John, M. M., Holmström Olsson, H., & Bosch, J. (2021) Towards MLOps: A Framework and Maturity Model. In M. T. Baldassarre, G. Scanniello & A. Skavhaug (Eds.), *Proceedings 2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA 2021)* (pp. 334–341). IEEE. DOI: 10.1109/SEAA53835.2021.00050

Kitsios, F., & Kamariotou, M. (2021) Artificial Intelligence and Business Strategy towards Digital Transformation: A Research Agenda. *Sustainability*, 13(4), 2025. <https://doi.org/10.3390/su13042025>

Klemola, K., Uusi-Illikainen, J. & Askola, T. (2014) *Sosiaali- ja terveystalouden tietojohdantamisen käsikirja*. Sitra.

Koivisto, J. & Tiirinki, H. (toim.) (2020) *Monialaisten palvelutarpeiden tunnistamisen ja ennakkoinnin toimintamallit ja työkalut – väliraportti (Työpäpaperi 4/2020)*. Terveyden ja hyvinvoinnin laitos. <http://urn.fi/URN:ISBN:978-952-343-459-2>

Koivisto, J., Liukko, E., Tiirinki, H. & Lyytikäinen, M. (2020) *Palvelukokonaisuuksien ja -ketjujen määrittely, ohjauksen ja seurannan käsikirja*. Terveyden ja hyvinvoinnin laitos, Ohjaus 11/2020.

Koivisto, J. & Liukko, E. (2017) *Palvelujen yhteensovittamisen haaste*. Katsaus Baskimaan, Skotlannin ja Pohjois-Irlannin sosiaali- ja terveydenhuollon uudistukseen. Työpäpaperi 35/2017, Terveyden ja hyvinvoinnin laitos.

Korpisaari, P., Pitkänen, O., & Warma-Lehtinen, E. (2022) *Tietosuoja (2. painos)*. Alma Talent.

Kumeno, F. (2019) Software engineering challenges for machine learning applications: A literature review. *Intelligent Decision Technologies*, 13(4), 463–476. DOI: 10.3233/IDT-190160

Kärkkäinen, L. (2020) *Tekoälyä terveydenhoitoon*. Aikakauskirja Duodecim, 136(17), 1980–1981.

Lazer, D. M. J., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., Freelon, D., Gonzalez-Bailon, S., King, G., Margetts, H., Nelson, A., Salganik, M. J., Strohmaier, M., Vespignani, A., & Wagner, C. (2020) Computational social science: Obstacles and opportunities. *Science*, 369, 1060–1062. DOI:10.1126/science.aaz8170

Leskelä, R.-L., Haavisto, I., Jääskeläinen, A., Helander, N., Sillanpää, V., Laasonen, V., Ranta, T. & Torkki, P. (2019). *Tietojohdantaminen ja sen kehittäminen: tietojohdantamisen arviointimalli ja suosituksia maakuntavalmistelun pohjalta (Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 2019:42)*. Valtioneuvoston kanslia. <http://urn.fi/URN:ISBN:978-952-287-754-3>

Leskelä RL., Komssi, V., Sandström S., Pikkujämsä S., Haverinen A, Olli, SL. & Ylitalo-Katajisto K. (2019) *Paljon sosiaali- ja terveystaloudellisia käyttäviä asiakkaat Oulussa*. Suomen Lääkärilehti. 2013;68:3163-3169c.

Liu, W., Stansbury, C., Singh, K., Ryan, A. M., Sukul, D., Mahmoudi, E., Waljee, A., Zhu, J. & Nallamothu, B. K. (2020) Predicting 30-day hospital readmissions using artificial neural networks with medical code embedding. *PLoS ONE*, 15(4). <https://doi.org/10.1371/JOURNAL.PONE.0221606>

Lohtander, M., Hätönen, H., Katainen, J., Lundkvist, T., Siira, T. & Hovi, A-M. (2021) *Hyvinvoinnin ja terveyden edistämisen kokonaisarkkitehtuuri (1.0)*. STM.

Lähteenmäki, J., Niemelä, M., Hammar, T., Alastalo, H., Noro, A., Pylsy, A., Arajärvi, M., Forsius, P., Pulli, K. & Anttila, H. (2020) *Kotona asumista tukeva teknologia - kansallinen toimintamalli ja tietojärjestelmät (KATI-malli) (VTT Technology No. 373)*. VTT Technical Research Centre of Finland. <https://doi.org/10.32040/2242-122X.2020.T373>

Lähteenmäki, J., Pajula, J., Antikainen E. (2022) Development of medical applications based on AI models and register data – regulatory considerations. *Proc of the 18th Scandinavian Conference on Health Informatics*, August 22-24, Tromsø, Norway, pp. 141-146. <https://doi.org/10.3384/9789179293444>



- Lähteenmäki, J., Vuorinen, A.-L., Pajula, J., Harno, K., Lehto, M., Niemi, M. & Van Gils, M. (2021) Integrating data from multiple Finnish biobanks and national health-care registers for retrospective studies: Practical experiences. *Scandinavian Journal of Public Health*, 50, 482–489. <https://doi.org/10.1177/14034948211004421>
- Madan, R. & Ashok, M. (2023) AI adoption and diffusion in public administration: A systematic literature review and future research agenda. *Government Information Quarterly*, 40(1), 101774. <https://doi.org/10.1016/j.giq.2022.101774>
- Maier, E., Reimer, U. & Wickramasinghe, N. (2021) Digital healthcare services. *Electronic Markets*, 31, 743–746. <https://doi.org/10.1007/s12525-021-00513-z>
- Melchiorre, MG., Papa, R., Rijken, M., van Ginneken, E., Hujala, A. & Barbabella, F. (2018) eHealth in Integrated Care Programs for people with multimorbidity in Europe; Insights from the ICARE4EU project. *Health Policy*, 122(1).
- Muiruri, D., Lwakatare, L. E., Nurminen, J. K. & Mikkonen, T. (2022) Practices and Infrastructures for Machine Learning Systems: An Interview Study in Finnish Organizations. *IEEE Computer*, 55(6), 18–29. DOI: 10.1109/MC.2022.3161161
- Mykkänen, M., Miettinen, M., Siponen, T. & Saranto, K. (2021) Sairaalan reaaliaikainen tilannekuva päivittäisessä johtamisessa. *Finnish Journal of EHealth and EWelfare*, 13(4), 425–441. <https://doi.org/10.23996/fjhw.109940>
- Myllyaho, L., Raatikainen, M., Männistö, T., Mikkonen, T. & Nurminen, J. K. (2021) Systematic literature review of validation methods for AI systems. *Journal of Systems & Software*, 181, 111050. <https://doi.org/10.1016/j.jss.2021.111050>
- Myllyaho, L., Raatikainen, M., Männistö, T., Nurminen, J. K. & Mikkonen, T. (2022) On misbehaviour and fault tolerance in machine learning systems. *Journal of Systems & Software*, 183, 111096. <https://doi.org/10.1016/j.jss.2021.111096>
- Mäenpää, O. (2016) *Julkisuusperiaate* (3. painos). Talentum Pro.
- Nahar, N., Lewis, G., Zhou, S. & Kästner, C. (2022) Collaboration Challenges in Building ML-Enabled Systems: Communication, Documentation, Engineering, and Process. In *ICSE '22: Proceedings of the 44th International Conference on Software Engineering* (pp. 413–425). ACM. <https://doi.org/10.1145/3510003.3510209>
- Nakubulwa, M., Junghans, C., Novov, V., Lyons-Amos, C., Lovett, D., Majeed, A., Aylin, P. & Woodcock, T. (2022) Factors associated with accessing long-term adult social care in people aged 75 and over: a retrospective cohort study. *Age and Ageing*, 51(3), 1–9. <https://doi.org/10.1093/ageing/afac038>
- Nasr, M., Islam, M. M., Shehata, S., Karray, F. & Quintana, Y. (2021) Smart Healthcare in the Age of AI: Recent Advances, Challenges, and Future Prospects. *IEEE Access*, 9, 145248–145270. DOI: 10.1109/ACCESS.2021.3118960
- National Health Service (2019) *The Topol Review: Preparing the healthcare workforce to deliver the digital future*. Health Education England.
- Nortio, J. (2020) Tekoäly tositoimissa / Kynnys korkealla. *Automaatioväylä*, 36(4), 13.
- Nummela, A. (2022) Tekoälyn käyttö kotihoidossa: hoitajien kokemuksia Gillie.AI alustan käytöstä [YAMK-opinnäytetyö, LAB-ammattikorkeakoulu]. <http://www.theseus.fi/handle/10024/756392>
- Nurminen, M., Reijonen, P. & Vuoreneimo, J. (2002) Tietojärjestelmän organisatorinen käyttöönotto: kokemuksia ja suuntaviivoja. *Turun kaupungin terveystoimen julkaisuja*, Sarja A, Nro 1/2002.
- Panch, T., Pearson-Stuttard, J., Greaves, F. & Atun, R. (2019) Artificial intelligence: opportunities and risks for public health. *The Lancet Digital Health*, 1(1), E13–E14. [https://doi.org/10.1016/S2589-7500\(19\)30002-0](https://doi.org/10.1016/S2589-7500(19)30002-0)
- Pentikäinen, M., Kärkkäinen, A., Mykkänen, J., Penttinen, J., Hyppönen, K., Siira, T. & Jalonen, M. (2019) Sosiaali- ja terveydenhuollon asiakas- ja potilastietojen kansallinen kokonaisarkkitehtuuri (2.1). *Terveyden ja hyvinvoinnin laitos*. <http://www.julkari.fi/handle/10024/138840>

Pentti, J., Välikangas, E., Uotinen, P., Alanne, P.-M., Kinanen, T.-M., Natunen, K., Pekkanen, J., Sarpola, E., Valkonen, M., Nylander, O., Pikkariainen, M. & Stigell, P. (2019) Tietojohdamisen pilotti SoteDigi Oy – Kainuun sote loppuraportti (v1.0). SoteDigi Oy.

Plastino, E. & Purdy, M. (2018) Game changing value from Artificial Intelligence: eight strategies. *Strategy & Leadership*, 46(1), 16–22. <https://doi.org/10.1108/SL-11-2017-0106>

Pote, H., Rees, A., Holloway-Biddle, C. & Griffith, E. (2021) Workforce challenges in digital health implementation: How are clinical psychology training programmes developing digital competences? *Digital Health*, 7, 1–11. DOI: 10.1177/2055207620985396

Rautiainen, I. (2021) Monialaisen tuen asiakkaan asiakasohjauksen kriittiset menestystekijät Päijät-Hämeen hyvinvointiyhtymässä [eMBA-lopputyö, LUT-yliopisto].

Riungu-Kalliosaari, L., Kauppinen, M. & Männistö, T. (2017) What Can Be Learnt from Experienced Data Scientists? A Case Study. In M. Felderer, D. Méndez Fernández, B. Turhan, M. Kalinowski, F. Sarro & D. Winkler (Eds.), *Product-Focused Software Process Improvement. PROFES 2017. Lecture Notes in Computer Science*, vol 10611 (pp. 55–70). Springer, Cham. [https://doi.org/10.1007/978-3-319-69926-4\\_5](https://doi.org/10.1007/978-3-319-69926-4_5)

Rousku, K. (Ed.), Andersson, C., Stenfors, S., Lähteenmäki, I., Limnell, J., Mäkinen, K., Kopponen, A., Kuivalainen, M. & Rissanen, O.-P. (2019) Glimpses of the future – Data policy, artificial intelligence and robotisation as enablers of wellbeing and economic success in Finland (Publications of the Ministry of Finance 2019:39). Ministry of Finance, Finland. <http://julkaisut.valtioneuvosto.fi/handle/10024/161675>

Ruotsalainen, R. & Tötterman, H. (2022) Rage against the machine: How bad IT systems undermine nurses' work – and what to do about it. *Reaktor Health*.

Rytilä, M. (2011) Tietoperustainen johtaminen palvelutoiminnan suunnittelussa julkisella terveydenhuoltoalalla. [Väitöskirja, Lapin yliopisto]. Lapin yliopistokustannus.

Sallinen, A. (21.3.2022) MLOps – mitä ja miksi? Koodia Suomesta. <https://koodiasuomesta.fi/mlops-mita-ja-miksi/>

Salovaara, S., Leinonen, J. & Silén, M. (2021) Barriers to the utilisation of data gathered through information systems in the management of social welfare organisations. *Finnish Journal of EHealth and EWelfare*, 13(4), 372–387. <https://doi.org/10.23996/fjhw.109930>

Shinners, L., Aggar, C., Grace, S. & Smith, S. (2020) Exploring healthcare professionals' understanding and experiences of artificial intelligence technology use in the delivery of healthcare: An integrative review. *Health Informatics Journal*, 26(2), 1225–1236. DOI: 10.1177/1460458219874641

Shinners, L., Aggar, C., Grace, S. & Smith, S. (2021) Exploring healthcare professionals' perceptions of artificial intelligence: Validating a questionnaire using the e-Delphi method. *Digital Health*, 7, 1–9. DOI: 10.1177/20552076211003433

Sira, M. (2021) AI Adoption Transformation Trends. *Business Ethics and Sustainable Development: Interdisciplinary Theoretical and Empirical Studies*, 2, 16–22.

Sirviö, J. (2021) Artificial intelligence in social care and healthcare SMEs in the Päijät-Häme region. In: Peltonen, K. & Hartikainen, A. (Eds.), *LAB Health Annual Review 2021 (The Publication Series of LAB University of Applied Sciences, part 30, pp. 235–245)*. LAB University of Applied Sciences.

de Sousa, W. G., de Melo, E. R. P., De Souza Bermejo, P. H., Sousa Farias, R. A. & Gomes, A. O. (2019) How and where is artificial intelligence in the public sector going? A literature review and research agenda. *Government Information Quarterly*, 36(4), 101392. <https://doi.org/10.1016/j.giq.2019.07.004>

Stirbu, V., Granlund, T. & Mikkonen, T. (2022) Continuous design control for machine learning in certified medical systems. *Software Qual J*. <https://doi.org/10.1007/s11219-022-09601-5>

STM (2021) Tulevaisuuden sosiaali- ja terveyskeskus 2020–2023. Ohjelma ja hankeopas. Sosiaali- ja terveysministeriön julkaisuja 2021:27. Sosiaali- ja terveysministeriö, Helsinki 2021.

Storey, V. C., Lukyanenko, R., Maass, W. & Parsons, J. (2022) Explainable AI. *Communications of the ACM*, 65(4), 27–29. 10.1145/3490699





Sun, T. Q. & Medaglia, R. (2019) Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, 36(2), 368–383. <https://doi.org/10.1016/j.giq.2018.09.008>

Szymanski, M. (2022) Designing and evaluating explainable AI for non-AI experts: challenges and opportunities. In J. Golbeck, F. M. Harper, V. Murdock, M. Ekstrand, B. Shapira, J. Basilico, K. Lundgaard, E. Oldridge (Eds.), *RecSys '22 Proceedings of the 16th ACM Conference on Recommender Systems* (pp. 735–736). ACM. <https://doi.org/10.1145/3523227.3547427>

Tamburri, D. A. (2020) Sustainable MLOps: Trends and Challenges. In *Proceedings 2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)* (pp. 17–23). IEEE. DOI 10.1109/SYNASC51798.2020.00015

Tao, C., Gao, J. & Wang, T. (2019) Testing and Quality Validation for AI Software—Perspectives, Issues, and Practices. *IEEE Access*, 7, 120164–120175. DOI: 10.1109/ACCESS.2019.2937107

Tietosuojaavaltuutetun toimisto: Tietosuojan vaikutustenarvioinnin ohje (12/2021) Tietosuojan vaikutustenarvioinnin ohje.

Toisiolain tulkintojen selkeyttäminen ja viestintä: yhteishankkeen loppuraportti (2022). Koostanut Deloitte Oy.

THO:2019:3., H 18/1882.

Tujula, P. (2021) Tekoälyavusteinen tiedolla johtaminen [YAMK-opinnäytetyö, LAB-ammattikorkeakoulu]. <http://www.theseus.fi/handle/10024/449677>

Tyrväinen, P., Silvennoinen, M., Talvitie-Lamberg, K., Ala-Kitula, A. & Kuoremäki, R. (2018) Identifying opportunities for AI applications in healthcare — Renewing the national healthcare and social services. In J. L. Vilaça, T. Grechenig, D. Duque, N. Rodrigues & N. Dias (Eds.), *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)* (pp. 1-7). IEEE. DOI: 10.1109/SeGAH.2018.8401381

Varjola, V. (2022) Tiedolla johtaminen sosiaali- ja terveydenhuollon kompleksisessä toimintaympäristössä [Pro gradu -tutkielma, Vaasan yliopisto]. <https://urn.fi/URN:NBN:fi-fe2022042530411>

Väläkangas, E. (2020) Virta-hankkeen tilanne ja kehitystyö SoteDigissä – Sote-järjestämisen tietomalli ja Sote-tietopaketti. SoteDigi Oy.

Weiser, O., Kalman, Y. M., Kent, C. & Ravid, G. (2022) 65 Competencies: Which Ones Should Your Data Analytics Experts Have? *Communications of the ACM*, 65(3), 58–66. DOI: 10.1145/3467018

Yu, X., Xu, S. & Ashton, M. (2023) Antecedents and outcomes of artificial intelligence adoption and application in the workplace: the socio-technical system theory perspective. *Information Technology & People*, 36(1), 454–474. DOI 10.1108/ITP-04-2021-0254

Zahid, A., Poulsen, J. K., Sharma, R. & Wingreen, S. C. (2021) A systematic review of emerging information technologies for sustainable data-centric health-care. *International Journal of Medical Informatics*, 149, 10442. <https://doi.org/10.1016/j.ijmedinf.2021.104420>

Zeitoun, J.-D. & Ravaud, P. (2020) Artificial intelligence in health care: value for whom? *The Lancet Digital Health*, 2(7), E338–E339. [https://doi.org/10.1016/S2589-7500\(20\)30141-2](https://doi.org/10.1016/S2589-7500(20)30141-2)

