Suvi Rantanen

Tekoälyn käytön eettiset ulottuvuudet sosiaali- ja terveydenhuollossa

Dokumenttianalyysi sosioteknisestä näkökulmasta

Johtamisen akateeminen yksikkö
Sosiaali- ja terveyshallintotieteet
Pro gradu tutkielma
Hallintotieteet

Vaasa 2025

VAASAN YLIOPISTO
Johtamisen akateeminen yksikkö
Tekijä: Suvi Rantanen
Tutkielman nimi: Tekoälyn käytön eettiset ulottuvuudet sosiaali- ja terveydenhuollossa: Dokumenttianalyysi sosioteknisestä näkökulmasta
Tutkinto: Hallintotieteiden maisteri
Oppiaine: Sosiaali- ja terveyshallintotiede
Työn ohjaaja: Kaisa Kurkela
Valmistumisvuosi: 2025 **Sivumäärä:** 95

TIIVISTELMÄ:
Yhteiskunnassa on käynnissä nopea teknologinen murros, jossa tekoälyratkaisut integroituvat yhä laajemmin sosiaali- ja terveydenhuollon toimintaan. Tämä kehitys luo merkittäviä mahdollisuuksia hoidon laadun ja tehokkuuden parantamiseen, mutta vaatii kokonaisvaltaista tarkastelua, jossa huomioidaan niin teknologiset, sosiaaliset kuin eettisetkin ulottuvuudet.

Tämän pro gradu -tutkielman tavoitteena on syventää ymmärrystä tekoälyn käyttöön liittyvistä eettisistä ulottuvuuksista sosiaali- ja terveydenhuollossa sekä tarkastella, miten nämä eettiset näkökulmat kytkeytyvät työn ja teknologian vuorovaikutukseen. Tutkimus tuo esiin, millaisia eettisiä haasteita tekoälyratkaisujen käyttöönotto luo alan organisaatioille ja ammattilaisille. Teoreettisena viitekehyksenä hyödynnetään sosioteknistä teoriaa, joka korostaa teknologian ja sosiaalisten rakenteiden vuorovaikutusta, sekä ihmiskeskeisen tekoälyn periaatteita ja tekoälyn eettisen arvioinnin viitekehystä. Näiden teoreettisten näkökulmien avulla jäsennetään tekoälyn käyttöönoton vaikutuksia organisaatioiden toimintaan, työn sisältöön ja työntekijöiden rooleihin.

Tutkimus toteutetaan laadullisena dokumenttianalyysinä teoriaohjaavan sisällönanalyysin menetelmällä. Aineistona käytettään kahdeksaa ajankohtaista suomalaista ja kansainvälistä asiantuntija- ja viranomaisjulkaisua, joissa käsitellään tekoälyn hyödyntämistä sosiaali- ja terveydenhuollossa, sen eettisiä kysymyksiä ja sääntelyä.

Tulosten perusteella yksityisyyden ja tietosuojan turvaaminen nähdään eettisen hyväksyttävyyden edellytyksenä, ja läpinäkyvyys sekä selitettävyys korostuvat tekoälyratkaisujen luotettavuuden ja käyttäjien luottamuksen rakentamisessa. Vastuullisuus liittyy erityisesti oikeudellisten velvoitteiden noudattamiseen ja päätöksenteon selkeään vastuunjakoon. Datan laatu ja algoritminen oikeudenmukaisuus nousevat keskeisiksi edellytyksiksi vinoumien ehkäisemiseksi ja yhdenvertaisuuden varmistamiseksi. Teknologian ja työn vuorovaikutuksen tarkastelu osoittaa, että tekoälyn käyttöönotto muuttaa työprosesseja ja rooleja, mikä edellyttää teknologian ja sosiaalisten rakenteiden jatkuvaa yhteensovittamista.

Johtopäätöksenä voidaan todeta, että tekoälyn vastuullinen käyttöönotto sosiaali- ja terveydenhuollossa edellyttää kokonaisvaltaista lähestymistapaa, jossa tekninen kehitys, lainsäädäntö ja eettiset periaatteet kytkeytyvät organisaatiotason muutoksiin, kuten työnjaon, johtamisen ja toimintakulttuurin uudistamiseen. Tutkimus tuo esiin tarpeen jatkuvalle arvioinnille, avoimuudelle ja yhteistyölle, jotta tekoäly voi toimia sosiaali- ja terveydenhuollon tukena turvallisesti ja eettisesti kestävällä tavalla.

AVAINSANAT: tekoäly, etiikka, sosiaali- ja terveydenhuolto, vastuullisuus, tietosuoja, läpinäkyvyys, sosiotekninen teoria

Sisällys

1 Johdanto 6

1.1 Johdatus tutkielman teemaan 6

1.2 Tutkimuskysymykset sekä tutkimuksen rakenne 9

2 Tekoäly sosiaali- ja terveysalalla 11

3 Sosiotekninen teoria 15

4 Generatiivinen tekoäly ja sen vaikutukset 20

5 Tekoälyn käytön eettiset näkökulmat sosiaali- ja terveydenhuollossa 24

5.1 Tietosuoja, yksityisyys ja korkean riskin tiedot 26

5.2 Lait ja sääntely 28

5.3 Avoimuus ja läpinäkyvyys 32

5.4 Vastuullisuus 34

5.5 Datan käyttö ja algoritmiset vinoumat 36

6 Tutkimuksen toteutus 40

6.1 Tutkimusmenetelmä 40

6.2 Aineisto 42

6.3 Teoriaohjaava sisällönanalyysi 45

7 Tulokset 51

7.1 Tietosuoja ja yksityisyys 51

7.2 Läpinäkyvyys ja selitettävyys 54

7.3 Vastuullisuus 56

7.4 Data ja algoritminen vinouma 58

7.5 Teknologian ja työn vuorovaikutus 61

8 Johtopäätökset 65

8.1 Eettiset ulottuvuudet 67

8.2 Teknologian ja työn vuorovaikutus sekä sosiotekninen näkökulma 72

8.3 Tutkimuksen luotettavuus 78

8.4 Jatkotutkimusaiheet 81

Lähteet 83

Liitteet 93

Kuviot

Kuvio 1. Tutkielman eteneminen. 10

Kuvio 2. Sosioteknisen teorian neljä pääulottuvuutta mukaillen Leavitt (1975). 17

Kuvio 3. Aineiston valintakriteerit. Kuvio laadittu hyödyntäen tekoälyä (ChatGPT, OpenAI, 2024). 44

Kuvio 4. Tekoälyn eettiset pääulottuvuudet sosiaali- ja terveydenhuollossa. 65

Kuvio 5. Pääluokkien suhteellinen painotus kansainvälisissä tuloksissa. 77

Kuvio 6. Pääluokkien suhteellinen painotus kansallisissa tuloksissa. 77

Taulukot

Taulukko 1. Keskeiset säädökset tekoälyn käytössä sosiaali- ja terveydenhuollossa. 31

Taulukko 2. Esimerkki pelkistämisestä. 49

Taulukko 3. Esimerkki sisällönanalyysista. 50

Taulukko 4. Eettisten ulottuvuuksien vaikutukset työhön ja käytäntöihin sekä niiden tulkinta sosioteknisen teorian näkökulmasta. 64

1 Johdanto

1.1 Johdatus tutkielman teemaan

Tekoäly on noussut viime vuosina laajamittaisen yhteiskunnallisen keskustelun kohteeksi, ja sen sovellukset näkyvät yhä selvemmin myös sosiaali- ja terveydenhuollossa. Teknologinen kehitys etenee nopeasti, ja erityisesti generatiivisen tekoälyn odotetaan vaikuttavan merkittävästi alan työn tuloksellisuuteen, tiedolla johtamiseen ja ammattilaisten työnkuviin tulevina vuosina (Heinäsenaho ja muut, 2023). Tekoälyn mahdollisuuksia on verrattu antibioottien ja lääketieteellisen kuvantamisen kaltaisiin innovaatioihin, jotka ovat mullistaneet terveydenhuollon toimintaa (STM, 2018, s. 11). Vertaus havainnollistaa tekoälyyn liitettävää yhteiskunnallista merkitystä.

Tekoälyn käyttöönotto ei kuitenkaan merkitse pelkästään teknologista uudistusta, vaan se käynnistää myös laajemman sosiaalisen ja organisatorisen muutoksen, joka ulottuu työn rakenteisiin, asiakaskohtaamisiin ja ammattilaisten rooleihin (Salovaara, 2024, s. 5). Sosioteknisen teorian näkökulmasta teknologia ei ole neutraali väline, vaan sen merkitykset ja vaikutukset rakentuvat vuorovaikutuksessa organisaatioiden sosiaalisten rakenteiden, työkäytäntöjen ja toimijoiden kanssa (Orlikowski, 1992). Onnistunut käyttöönotto edellyttää teknisen toimivuuden lisäksi sitä, että teknologia sovitetaan työympäristön olosuhteisiin ja ihmisten tarpeisiin, mikä korostaa teknisten ja sosiaalisten tekijöiden yhteensovittamisen merkitystä (Salovaara, 2024, s. 5).

Ihmiskeskeisen tekoälyn ja tukiälyn (augmented intelligence) näkökulmat korostavat, että teknologia toimii välineenä, joka tukee ja vahvistaa inhimillistä päätöksentekoa sen sijaan, että korvaisi ihmisen. Sosiaali- ja terveydenhuollossa tämä konkretisoituu muun muassa tekoälypohjaisissa diagnostiikkatyökaluissa, hoitoprosessien optimoinnissa ja resurssien hallinnan kehittämisessä (Johnson ja muut, 2020; Larsio, 2023). Tekoäly kytkeytyy siten osaksi sosioteknistä järjestelmää, jossa teknologiset ja sosiaaliset elementit muodostavat erottamattoman kokonaisuuden.

Samanaikaisesti tekoälyn hyödyntäminen nostaa esiin keskeisiä eettisiä kysymyksiä, jotka liittyvät muun muassa tietosuojaan, potilasturvallisuuteen ja algoritmiseen vinoumaan. Erityisesti algoritmisten vinoumien ehkäisy on ratkaisevaa, kun tekoälyjärjestelmät tukevat diagnoosien tekemistä tai hoitosuositusten muodostamista (Hoffman, 2020). Myös tekoälyratkaisujen oikeudenmukaisuus on keskeinen tekijä järjestelmien eettisessä hyväksyttävyydessä (Kumar ja muut, 2023). Luottamuksen rakentaminen edellyttää, että tekoäly on läpinäkyvää, vastuullista ja selitettävää (Al-Ansi ja muut, 2024). Näiden ehtojen toteutuminen vaatii sekä huolellista teknistä suunnittelua että onnistunutta integrointia osaksi sosiaali- ja terveydenhuollon arjen käytäntöjä.

Tutkimusaihe on yhteiskunnallisesti ja hallintotieteellisesti merkittävä, sillä tekoälyn käyttöönotto muuttaa sosiaali- ja terveydenhuollon toimintaympäristöä monitasoisesti. Se edellyttää teknologian ja työn yhteensovittamista sekä eettisten periaatteiden konkretisoimista arjen käytännöissä. Eettisten kysymysten jäsentäminen ja sosioteknisen näkökulman huomioiminen ovat keskeisiä, jotta tekoälyratkaisut voidaan integroida osaksi palvelujärjestelmiä tavalla, joka tukee sekä ammattilaisia että kansalaisten oikeuksia ja turvallisuutta (Salovaara, 2024; Sartor & Lagioia, 2020).

Tutkielma perustuu tuoreeseen suomalaiseen ja kansainväliseen dokumenttiaineistoon, joka koostuu politiikka- ja ohjausdokumenteista, asiantuntijaraporteista sekä suosituksista. Aineistoksi on valittu julkisesti saatavilla olevia julkaisuja, jotka käsittelevät tekoälyn eettisiä ulottuvuuksia sosiaali- ja terveydenhuollossa. Valinnassa on painotettu ajankohtaisuutta, asiantuntijalähtöisyyttä ja tutkimuskysymysten kannalta olennaista sisältöä, jotta analyysi kohdistuu sosiaali- ja terveydenhuollon kannalta keskeisiin eettisiin teemoihin.

Tämän tutkielman tavoitteena on syventää ymmärrystä tekoälyn käyttöön liittyvistä eettisistä ulottuvuuksista sosiaali- ja terveydenhuollossa sekä tarkastella niitä sosioteknisen teorian näkökulmasta. Tutkimuksessa analysoidaan, millaisia eettisiä

näkökulmia tekoälyyn liitetään alan keskeisissä politiikka- ja ohjausdokumenteissa ja miten nämä näkökulmat heijastuvat käytännön toimintaympäristöihin. Erityistä huomiota kiinnitetään vastuullisuuden, läpinäkyvyyden ja selitettävyyden teemoihin, joita korostetaan erityisesti kansainvälisissä eettisissä linjauksissa ja suosituksissa tekoälyn hyväksyttävyyden edellytyksinä (WHO, 2021). Yksityisyyden suojan ja datan käytön kysymykset liittyvät läheisesti tietosuojaan ja algoritmisten järjestelmien oikeudenmukaisuuteen, joita Sartor ja Lagioia (2020) tarkastelevat yleisen tietosuoja-asetuksen (GDPR) ja tekoälyn rajapinnan näkökulmasta. Algoritmisten vinoumien ja työn ja teknologian vuorovaikutuksen merkitystä puolestaan painotetaan tekoälyn yhteiskunnallisia vaikutuksia ja valtasuhteita analysoivassa tutkimuksessa (Roberts ja muut, 2020), jossa tuodaan esiin, miten eettiset periaatteet ja teknologiset ratkaisut kietoutuvat organisatorisiin ja sosiaalisiin rakenteisiin. Lisäksi tarkastellaan, miten nämä eettiset ulottuvuudet kietoutuvat sosiaalisiin ja organisatorisiin rakenteisiin ja millaisin mekanismein ne muovaavat tekoälyn käyttöönoton ehtoja sekä työn arkea.

Tekoälyn eettisiä kysymyksiä on aiemmassa kirjallisuudessa tarkasteltu useista näkökulmista. Osa tutkimuksesta painottuu teknisiin ratkaisuihin ja oikeudellisiin puitteisiin (Sartor & Lagioia, 2020), osa eettisiin periaatteisiin ja hallinnollisiin kysymyksiin (WHO, 2021) ja osa yhteiskunnallisiin vaikutuksiin ja vallan jakautumiseen (Roberts ja muut, 2020). Sosiotekninen näkökulma syventää keskustelua tarkastelemalla, miten teknologiset ja sosiaaliset tekijät kietoutuvat toisiinsa organisatorisissa toimintaympäristöissä. Tutkielma sijoittuu näiden lähestymistapojen väliin ja tarkastelee tekoälyn eettisiä ulottuvuuksia osana sosiaali- ja terveydenhuollon sosioteknistä järjestelmää. Tutkielma tuottaa jäsenneltyä tietoa eettisten periaatteiden ja käytännön toimintaympäristöjen välisestä dynamiikasta. Tuloksissa tarkastelen havaintoja suhteessa teoreettiseen viitekehykseen ja osoitan, miten ne tukevat tai täydentävät sosioteknisen teorian näkökulmia.

1.2 Tutkimuskysymykset sekä tutkimuksen rakenne

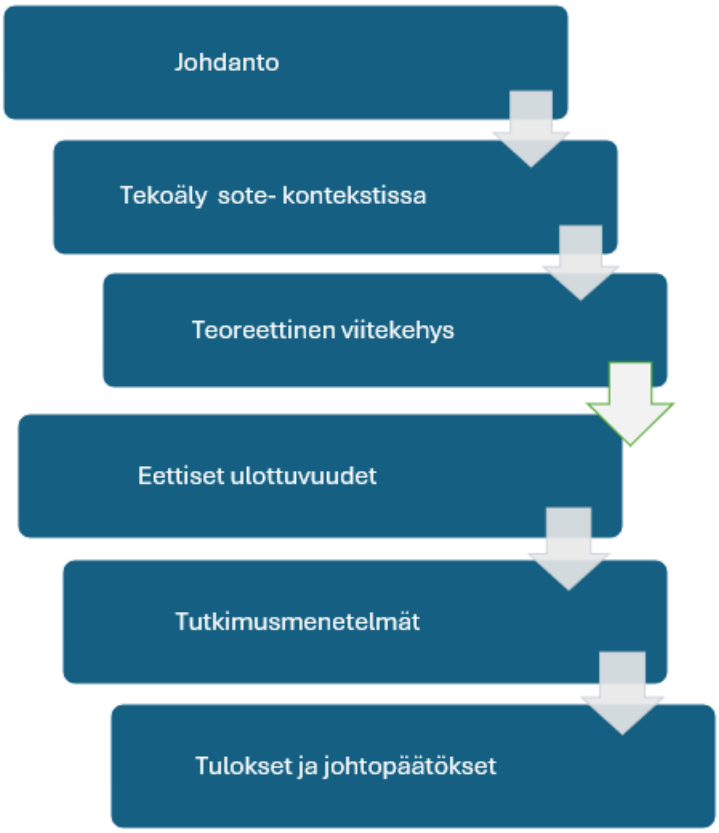
Tämän tutkielman tavoitteena on syventää ymmärrystä tekoälyn käyttöön liittyvistä eettisistä näkökulmista sosiaali- ja terveydenhuollossa. Tekoälyn käyttöönotto sosiaali- ja terveysalan kontekstissa tuo mukanaan monia mahdollisuuksia, mutta myös haasteita, jotka liittyvät muun muassa vastuullisuuteen, läpinäkyvyyteen, selitettävyyteen, yksityisyydensuojaan, datan vinoumiin sekä tekniikan integroimiseen arjen työhön. Näiden eettisten ulottuvuuksien tarkastelu on keskeistä, kun tekoälyä sovelletaan arkaluonteisia tietoja sisältäviin palveluihin, joissa päätöksenteolla voi olla merkittäviä vaikutuksia yksilöiden hyvinvointiin ja oikeuksiin.

Tutkimuksessa tarkastellaan lisäksi, miten nämä eettiset näkökulmat kytkeytyvät organisaatioiden arkeen ja teknologian käyttöönottoon sosioteknisestä näkökulmasta. Sosiotekninen teoria korostaa, että teknologian vaikutukset eivät ole ennalta määrättyjä, vaan ne rakentuvat sosiaalisten ja teknisten elementtien vuorovaikutuksessa. Tekoälyn käyttöönottoa ei voida ymmärtää pelkästään teknisenä muutoksena, vaan se edellyttää myös sosiaalisten rakenteiden, toimintatapojen ja arvojen yhteensovittamista.
Tutkielman päätutkimuskysymyksiä ovat:

1. Millaisia eettisiä näkökulmia tekoälyyn liitetään sosiaali- ja terveysalalla?
2. Miten eettiset näkökulmat kytkeytyvät sosioteknisiin kysymyksiin?

Tutkielma käynnistyy johdantoluvulla, jossa esitellään tutkimuksen tausta, tavoitteet ja tutkimuskysymykset. Toisessa luvussa tarkastellaan tekoälyn hyödyntämistä sosiaali- ja terveydenhuollon kontekstissa, mikä luo perustan aiheen ymmärtämiselle ja tutkimuksen rajaukselle. Kolmannessa luvussa syvennytään tutkimuksen teoreettiseen viitekehykseen, joka rakentuu sosioteknisen näkökulman ohella ihmiskeskeisen tekoälyn periaatteille sekä tekoälyn eettisen tarkastelun viitekehykselle (Johnson ja muut, 2020; Ojanperä, 2023). Näiden lähestymistapojen kautta tarkastellaan, miten tekoäly voidaan integroida vastuullisesti ja tarkoituksenmukaisesti osaksi sosiaali- ja terveydenhuollon prosesseja.

Neljännessä luvussa tarkastellaan generatiivista tekoälyä ja sen vaikutuksia erityisesti palvelurakenteisiin ja ammatilliseen työhön. Viidennessä luvussa käsitellään tekoälyn käyttöön liittyviä keskeisiä eettisiä ulottuvuuksia, joita tarkastellaan tietosuojan, sääntelyn, läpinäkyvyyden, vastuullisuuden sekä datan ja algoritmisten vinoumien näkökulmista. Kuudennessa luvussa kuvataan tutkimuksen metodologinen toteutus. Luvussa esitellään valittu tutkimusmenetelmä, aineiston muodostuminen sekä teoriaohjaavan sisällönanalyysin käytännön toteutus. Seitsemäs luku kokoaa analyysin tulokset teemoittain ja tarkastelee, millaisia merkityksiä tekoälyn eettiseen käyttöön liitetään valituissa dokumenteissa. Tutkielma päättyy johtopäätöksiin, jossa esitetään keskeiset havainnot, arvioidaan tutkimuksen luotettavuutta sekä pohditaan mahdollisia jatkotutkimuksen suuntaviivoja. Kuviossa 1 esitetään yleiskuva tutkielman rakenteesta ja etenemisestä.



Kuvio 1. Tutkielman eteneminen.

2 Tekoäly sosiaali- ja terveysalalla

Tekoälyä pidetään yhtenä keskeisimmistä teknologisista murroksista, jonka arvioidaan muuttavan perusteellisesti terveydenhuollon toimintaa (Larsio, 2023, s. 4). Autioniemi (2020, s. 5) luonnehtii tekoälyä innovaation lähteeksi, joka tuo mukanaan merkittäviä muutoksia erityisesti julkisen sektorin työelämään. Johnson ja muut (2020) puolestaan korostavat, että tekoälyä hyödyntävät järjestelmät osoittautuvat tehokkaammiksi kuin ilman tekoälyä toimivat ratkaisut. Tekoäly voi parantaa terveydenhuollon saatavuutta, tukea ammattilaisten päätöksentekoa ja edistää kansalaisten kokonaisvaltaista hyvinvointia tarjoamalla älykkäitä ratkaisuja diagnostiikkaan, hoitoprosessien optimointiin ja resurssien hallintaan (Larsio, 2023, s. 20).

Heinäsenaho ja muut (2023) kuvailevat generatiivisella tekoälyllä olevan potentiaalisia ja hyvin mittavia hyötyjä terveydenhuollossa. Tekoäly pystyy esimerkiksi analysoimaan suuria määriä potilastietoja, ennakoimaan sairauksien kehittymistä sekä suosittelemaan yksilöllisiä hoitomuotoja. Tekoälyn avulla on mahdollista yhdistellä monimutkaisia tietoja, kuten potilaan geneettisiä ominaisuuksia, sukupuolta ja ympäristötekijöitä, joiden vuorovaikutuksen avulla voidaan tunnistaa sairauksien riskitekijöitä ja ennaltaehkäistä vakavia sairauksia ennen niiden puhkeamista. Heinäsenaho ja muut (2023) enteilevät generatiivisen tekoälyn olevan pian osa jokapäiväisiä ja toistuvia hoitotilanteita. Salo (2024, s. 235) ennustaa, että vuoden 2024 jälkeen ei juuri olisi tietotyöläisiä, jotka eivät olisi kuulleet luovan tekoälyn palveluista ja että seuraavan vuoden aikana jokainen kyseisen alan ammattilainen tulee käyttämään näitä palveluita jo työssään.

Tiedämme, että tekoälyjärjestelmällä on kyky toimia kognitiivisesti. Johnsonin ja muiden (2020) mukaan tekoälyjärjestelmän neljä pääominaisuutta ovat päättely, oppiminen, ymmärtäminen ja vahvistaminen. Näiden ominaisuuksien ansiosta tekoäly kykenee lukemaan, käsittelemään ja tulkitsemaan sekä strukturoitua että jäsentämätöntä dataa laajassa mittakaavassa. Tämä mahdollistaa monimutkaisten sairauksien syy-

seuraussuhteiden tunnistamisen, hoitoprosessien optimoinnin ja yksilöllisten hoitosuositusten tarkentamisen, mikä tukee kliinistä päätöksentekoa entistä tarkemmin potilaskohtaisten tarpeiden perusteella. (Johnson ja muut, 2020.)

Tekoäly tarjoaa erityisen merkittäviä mahdollisuuksia muun muassa diagnostiikassa, kuvantamisessa ja hoitosuunnitelmien laatimisessa. Tekoälystä käytetyt erilaiset muodot, kuten syväoppiminen ja koneoppiminen, ovat mahdollistaneet päätöksenteon parantamisen sekä ihmisen arviointikyvyn rajojen ylittämisen. Esimerkiksi perusterveydenhuollossa tekoälyn avulla voidaan ennakoida potilaiden terveydentilan muutoksia ja tarjota heille kohdennettua hoitoa. Ojanperä (2023, s. 134–135) tuo esille, että tekoälyä on jo hyödynnetty laajasti potilastietojen analysoinnissa, ja sen rooli tässä tulee vain kasvamaan. Johnson ja muut (2020) toteavat, että tekoälyn käyttö radiologiassa on jo vakiintunut ja tekoäly pystyy analysoimaan röntgenkuvia tehokkaammin ja luotettavammin kuin perinteiset menetelmät. Larsio (2023, s. 20) korostaa, että tekoäly ei ole korvaamassa terveydenhuollon ammattilaisia, vaan toimii ennemmin tukiälynä, joka vahvistaa päätöksentekoa ja lisää diagnostiikan tarkkuutta.

Tekoälyn hyödyntäminen edellyttää myös ihmiskeskeistä lähestymistapaa, jossa huomioidaan terveydenhuollon ammattilaisten ja potilaiden tarpeet. Tyrväinen ja muut (2018) korostavat tekoälyhankkeiden onnistumisessa käyttäjälähtöisyyttä ja yhteistoiminnallista kehitystyötä. Esimerkiksi tekoälyyn pohjautuvat chatbotit voivat tarjota kansalaisille ohjausta ja neuvontaa terveyteen liittyvissä kysymyksissä, mutta samalla on huolehdittava siitä, että tekoäly toimii eettisesti ja vastuullisesti (Heinäsenaho ja muut, 2023).

Yleisesti tekoälyn hyötyjä ovat sen kyky analysoida suuria tietomääriä nopeasti, tukea terveydenhuollon ammattilaisten työtä sekä parantaa potilaiden hoidon laatua. Kuitenkin tekoälyn käyttöön liittyy myös eettisiä ja teknisiä haasteita, kuten algoritmisen syrjinnän mahdollisuus ja tietosuoja (Kumar ja muut, 2023), sekä vastuullisuus, läpinäkyvyys ja oikeudenmukaisuus, jotka vaikuttavat tekoälyn hyväksyttävyyteen ja

luotettavuuteen (Khosravi ja muut, 2024). Hoffman (2020) korostaa, että tekoälyn vinoumat voivat aiheuttaa vakavia oikeudenmukaisuuskysymyksiä, erityisesti terveydenhuollossa, missä väärät päätökset voivat vaikuttaa ihmisten elämänlaatuun ja selviytymiseen. Lisäksi tekoälyn selitettävyys ja läpinäkyvyys ovat keskeisiä haasteita, sillä päätösten perusteiden ymmärrettävyys vaikuttaa potilaiden ja ammattilaisten luottamukseen tekoälyä kohtaan. Tekoälyn käyttöönotto edellyttää kansalaisten oikeutta ymmärtää tekoälyn tekemien päätösten perusteita, mikä tukee luottamusta teknologiaa kohtaan (Al-Ansi ja muut, 2024.)

Tekoälyn yleistyminen sosiaali- ja terveysalalla edellyttää rinnalleen vahvaa lainsäädäntöä ja sääntelykehystä. Koska tekoälyjärjestelmät käsittelevät arkaluonteisia potilastietoja ja tukevat hoitopäätöksiä, on olennaista varmistaa niiden luotettavuus, läpinäkyvyys ja oikeudenmukaisuus. Euroopan unionin sääntely, erityisesti tekoälyasetus ja yleinen tietosuoja-asetus (GDPR), muodostavat keskeisen kehyksen, jonka tarkoituksena on suojata yksilöiden oikeuksia ja varmistaa järjestelmien eettinen käyttö (Ojanperä, 2023, s. 163–165). Samansuuntaisesti Al-Ansi ja muut (2024) korostavat, että sääntelyllä luodaan puitteet tekoälyn vastuulliselle soveltamiselle. Lainsäädäntö toimii ennaltaehkäisevänä mekanismina algoritmisen syrjinnän ja tietosuojaloukkauksien kaltaisia riskejä vastaan, ja samalla se ohjaa teknologista kehitystä vastuullisuuden ja ihmisoikeuksien näkökulmasta (Ferede, 2025). Tulevaisuudessa korostuvat kansainvälisesti yhdenmukaisten standardien ja ohjeistusten kehittäminen, jotta tekoälyratkaisujen käyttöönotto terveydenhuollossa olisi yhdenvertaista, turvallista ja eettisesti kestävää kaikkialla maailmassa (Kumar ja muut, 2023). Ferede (2025) puolestaan painottaa, että oikeudellisten rakenteiden kehittäminen on keskeistä globaalien eriarvoisuuksien kaventamiseksi tekoälyn soveltamisessa.

Vaikka tekoälystä on keskusteltu jo vuosia, sen käyttöönotto sosiaali- ja terveysalalla on yhä alkuvaiheessa. Viime vuosina tekoälyn soveltamisesta on julkaistu yhä enemmän tutkimuksia, mutta sosiaali- ja terveysalan kontekstissa julkaisut ovat edelleen verrattain

uusia. Autioniemi (2020, s. 5) korostaa, että julkishallinnossa tekoälyn tutkimus on vasta kehittymässä. Tekoälyn hyödyntämismahdollisuuksista saadaan jatkuvasti lisää tietoa, mutta samalla tunnistetaan myös keskeisiä haasteita, kuten tietosuoja, eettiset kysymykset ja sääntelyn esteet (DigiFinland, 2024a). Larsio (2023, s. 4–5) nostaa esiin datan liikkuvuuden ja saatavuuden ongelmat merkittävinä esteinä tekoälyn laajemmalle käyttöönotolle sosiaali- ja terveysalalla. Jotta tekoälyä voidaan integroida tehokkaasti terveydenhuoltoon, tarvitaan jatkuvaa tutkimusta, kehitystyötä ja lainsäädännön sopeuttamista, jotta sen hyödyt voidaan maksimoida ja samalla minimoida mahdolliset riskit.

3 Sosiotekninen teoria

Tekoälyn integrointi sosiaali- ja terveysalalle ei ole pelkästään tekninen, vaan myös sosiaalinen muutos, joka vaatii uusien työtapojen ja organisaatiokulttuurin mukauttamista. Teknologian onnistunut käyttöönotto edellyttää sen sovittamista organisaation sosiaalisiin rakenteisiin, työprosesseihin ja työntekijöiden tarpeisiin (Valta, 2013, s. 22–23).

Sosiotekninen teoria auttaa analysoimaan, miten ihmiskeskeinen tekoäly ja tukiäly voivat edistää terveydenhuollon tehokkuutta ja turvallisuutta, sekä millaisia haasteita ja mahdollisuuksia niiden käyttöön liittyy. Teorian avulla voidaan jäsentää tekoälyn vaikutuksia sosiaali- ja terveydenhuollon työn rakenteisiin sekä tarkastella, miten teknologian ja inhimillisten tekijöiden välinen tasapaino voidaan säilyttää digitalisoituvassa työympäristössä (Trist, 1981.)

Sosiotekninen lähestymistapa korostaa, että teknologiaa ei tulisi arvioida vain teknisin perustein, vaan myös sen mukaan, miten se vaikuttaa työhön sosiaalisessa kontekstissaan (Salovaara, 2024, s. 4–5). Tietojärjestelmät eivät ole neutraaleja työkaluja, vaan ne vaikuttavat siihen, millaista tietoa tuotetaan, miten sitä jäsennetään ja millaisia käytäntöjä työssä muodostuu (Salovaara, 2024, s. 5, 51). Salovaaran näkemys liittyy suoraan Orlikowskin (1992) esiin nostamaan ajatukseen teknologian ja sosiaalisen toiminnan yhteen kietoutumisesta, jossa teknologiat eivät ainoastaan tue työtä, vaan osallistuvat aktiivisesti sen muovaamiseen.

Sosiotekninen lähestymistapa on saanut laajaa huomiota erityisesti sellaisilla toiminta-alueilla, joissa yhdistyvät inhimillinen toiminta ja teknologia. Sosiotekninen systeemi voidaan ymmärtää sosiaalisten ja teknisten osien muodostamaksi kokonaisuudeksi, jolla on selkeä tavoitteellinen toiminta (Walker ja muut, 2008, s. 480). Salovaara (2024, s. 42–43) korostaa, että sosiotekninen tarkastelu auttaa ymmärtämään teknologian ja sosiaalisten rakenteiden tiivistä yhteen kietoutumista. Walker ja muut (2008, s. 480–481)

painottavat, että järjestelmän suorituskyky rakentuu näiden osien vuorovaikutuksesta, joka sisältää sekä suunniteltuja että odottamattomia, usein monimutkaisia ja ei-lineaarisia suhteita. Tällöin esimerkiksi tekoälyyn perustuvat päätöksenteon tukijärjestelmät voivat muuttaa työn kulkua tavoilla, joita ei voida täysin ennakoida. Vaikutukset voivat olla sekä myönteisiä että kielteisiä riippuen siitä, miten ratkaisut kytkeytyvät työn käytännön rakenteisiin.

Berg (1999) huomauttaa, että tietojärjestelmät eivät yksiselitteisesti tehosta tai häiritse toimintaa, vaan niiden vaikutukset riippuvat aina käytännön tilanteesta ja siitä, kuinka hyvin teknologia linkittyy olemassa olevaan työhön. Teknologian ja työn välinen suhde on dynaaminen: järjestelmät voivat uudistaa työnjaon logiikkaa, vastuuketjuja ja jopa työn merkityksiä. Myös Walker ja muut (2008) nostavat esille tämän avoimen järjestelmän näkemyksen, jossa painottuvat jatkuva sopeutuminen ja kontekstuaalinen muutos. Tekoälyn käyttöä ei tulisi arvioida vain sen ennakoitavien hyötyjen perusteella, vaan myös sen mukaan, miten se muovaa työn käytäntöjä, valtasuhteita ja vuorovaikutusta esimerkiksi potilaan ja ammattilaisen välillä. Sosiaali- ja terveysalalla tämä edellyttää herkkyyttä arvioida teknologisten ratkaisujen vaikutuksia paitsi tehokkuuteen ja tiedonkulkuun myös hoitotyön eettisiin ja inhimillisiin ulottuvuuksiin.

Koska sosiotekninen teoria nojaa avoimien järjestelmien periaatteisiin, tekoälyn käyttöönottoa ei voida nähdä yksittäisenä, staattisena muutoksena. Sen sijaan se vaatii jatkuvaa sopeutumista ja vuorovaikutusta työntekijöiden, asiakkaiden ja ympäristön kanssa (Walker ja muut, 2008, s. 487–488). Teknologian kehittäminen ja käyttöönotto ei ole irrallinen tekninen prosessi, vaan siihen sisältyy organisaation roolien, vastuiden ja palautemekanismien uudelleenarviointia. Valtan (2013, s. 31) nostaman huomion mukaisesti pelkkä tekninen kehittäminen ei riitä, vaan myös sosiaalista järjestelmää on kehitettävä rinnakkain. On olennaista, että teknologiaratkaisujen suunnittelussa kuullaan niiden käyttäjiä – ilman henkilöstön ääntä järjestelmäkehitys jää yksipuoliseksi ja altistaa epäonnistumiselle (Valta, 2013, s. 31).

Sosioteknisen teorian tulkinnoissa on esitetty erilaisia näkemyksiä, sillä osa tutkijoista korostaa sosiaalisen ja teknisen järjestelmän kaksijakoa, kun taas toiset jäsentävät teoriaa kolmen ulottuvuuden kautta, joita ovat ihminen, teknologia ja työtehtävät (Valta, 2013, s. 31–34). Valta (2013, s. 31) kuvaa organisaatiota Leavittia (1975) mukaillen neljän keskenään vuorovaikutuksessa olevan elementin avulla: ihmisen, organisaation, teknologian ja työtehtävien (ks kuvio 2). Nämä yhdessä muodostavat sosioteknisen kokonaisuuden. Lähestymistapa konkretisoi tutkimukseni teoreettisen viitekehyksen perusajatuksen, jonka mukaan tekoälyn käyttöönottoa sosiaali- ja terveydenhuollossa on tarkasteltava näiden neljän ulottuvuuden kautta samanaikaisesti. Malli korostaa sosioteknisen teorian keskeistä viestiä siitä, että yksittäinen osa-alue ei voi kehittyä erillään muista, vaan kokonaisuuden toimivuus rakentuu niiden välisestä vuorovaikutuksesta ja tasapainosta.

Kuvio 2. Sosioteknisen teorian neljä pääulottuvuutta mukaillen Leavitt (1975).

Orlikowski (1992, s. 403–405) muistuttaa, että teknologia ei ole neutraali työkalu, vaan se rakentuu ihmisten toiminnassa ja muokkaa samalla sitä, miten työtä tehdään. Sosiaali-

ja terveysalalla tämä näkyy esimerkiksi tekoälyn käytössä asiakastietojen tulkinnassa, hoidon tarpeen arvioinnissa tai palveluiden kohdentamisessa. Vaikka ammattilainen tekee lopulliset päätökset, tekoälyn antamat suositukset voivat vaikuttaa siihen, mitä vaihtoehtoja pidetään realistisina tai perusteltuina. Tekoäly voi ohjata päätöksentekoa hienovaraisesti ja tuoda mukanaan uusia vastuusuhteita sekä riippuvuutta teknologisista rakenteista.

Tekoälyn eettisesti, toiminnallisesti ja ammatillisesti kestävä käyttö edellyttää sen tarkastelua osana laajempaa sosioteknistä järjestelmää. Tällöin huomio kiinnittyy paitsi tekoälyn tekniseen toimivuuteen myös siihen, miten se vaikuttaa työntekijöiden rooleihin, vuorovaikutukseen asiakkaiden kanssa ja tiedon muodostumiseen arjen käytännöissä. Salovaara (2024, s. 40–42) korostaa, että tekoälyn käyttö muokkaa organisaatioiden toimintaa ja työn rakenteita jatkuvassa vuorovaikutuksessa sosiaalisten tekijöiden kanssa. Walker ja muut (2008) puolestaan tarkastelevat sosioteknisiä järjestelmiä avoimina ja dynaamisina kokonaisuuksina, joiden vaikutuksia ei voida palauttaa yksittäiseen tekijään. Orlikowski (1992, s. 406–407) vie tätä ajatusta pidemmälle toteamalla, että teknologia ja sosiaaliset käytännöt muovaavat toisiaan jatkuvasti. Näin sosiotekninen näkökulma haastaa perinteiset käsitykset teknologiasta neutraalina välineenä ja nostaa esiin sen, että tekoäly on osa ihmisten ja koneiden yhteistoimintaa, jossa merkitys ja vaikutukset syntyvät käytännön konteksteissa.

Tässä yhteydessä on tärkeää tarkastella, miten tekoälyn käyttöönotto muuttaa työn luonnetta. Siirtääkö se esimerkiksi painopistettä asiakastyöstä järjestelmätyöhön, vähentääkö se ammatillista harkintaa vai tukee sitä, ja millaisia uusia osaamistarpeita se synnyttää? Tekoäly voi parhaimmillaan toimia tukena, joka lisää työn vaikuttavuutta ja sujuvuutta. Ilman riittävää reflektiota ja osallistavaa kehittämistä tekoäly voi kuitenkin myös vahvistaa kontrollin ja seurannan elementtejä, jotka kaventavat työn autonomiaa ja asiakaslähtöisyyttä. Sosiotekninen näkökulma tuo esiin tarpeen jatkuvalle dialogille teknologiakehittäjien, ammattilaisten ja asiakkaiden välillä, jotta tekoälyn muovaama tulevaisuus olisi yhteistoiminnallinen ja inhimillisesti kestävä.

Sosiotekninen teoria mahdollistaa tekoälyn tarkastelun kokonaisvaltaisena muutosilmiönä, jossa teknologia ja sosiaalinen konteksti muokkaavat toisiaan. Teoria auttaa jäsentämään, miten tekoäly vaikuttaa työn sisältöön, vastuisiin ja vuorovaikutukseen sosiaali- ja terveysalalla. Kyse ei ole vain teknisestä käyttöönotosta, vaan laajemmasta yhteistoiminnallisesta järjestelmästä, jossa teknologian eettinen, ammatillinen ja organisatorinen kestävyys rakentuvat vuorovaikutuksessa työn arjen kanssa.

4 Generatiivinen tekoäly ja sen vaikutukset

Älykkyydestä on tutkijoiden keskuudessa erilaisia määritelmiä. Yleisesti älykkyys liitetään kykyihin oppia ja ymmärtää asioita (Ala-Korpela ja muut, 2007, s. 51). Älykkyyden tarkemman määrittelyn vaihtelevuus tuo oman haasteensa ja monimuotoisuutensa myös tekoälyn määrittelemiseen.

Tekoäly (engl. Artificial Intelligence, AI) käsitteenä on ollut olemassa jo vuosikymmeniä, saaden alkunsa Yhdysvaltojen New Hampshiressa sijaitsevasta Darthmouth Collegessa pidetystä kesätutkimushankkeesta vuonna 1956. Käsitteen luojana pidetään matematiikan professori John McCarthya, joka yhdessä muiden tutkijoiden kanssa olettivat, että on olemassa mahdollisuus saada tietokone toimimaan ihmismielen mukaisesti. Tuolloin vuonna 1956 tutkijat lähtivät oletuksesta, jonka mukaan oppiminen ja älykäs toiminta voidaan ohjelmoida tietokoneeseen. Se voisi toteuttaa erilaisia toimintoja, kuten käyttää kieltä, muodostaa käsitteitä, ratkoa ongelmia inhimilliseen tapaan sekä parantaa omaa toimintaansa. (Moor, 2006, s. 87; Pesonen, 2021). Näistä varhaisista ajatuksista tekoälyn kehitys on edennyt merkittävästi, erityisesti neuroverkkojen ja koneoppimisen avulla (Ojanperä, 2023, s. 24–26).

Erilaiset nykyajan määritelmät tekoälystä ovat kuitenkin pääpiirteissään samankaltaisia. Tekoälyllä tarkoitetaan tietokoneohjelmia, jotka pystyvät oppimaan ihmisen kognitiivisia toimintoja, kuten oppimista tai ongelmanratkaisua (Perez & Grande, 2020). Tietokoneohjelmien mukaisesti tekoäly voi olla laite, ohjelmisto tai järjestelmä, jolla on inhimillinen kyvykkyys oppia ja tehdä päätöksiä (Työ- ja elinkeinoministeriö, 2019, s. 15–16). Digi Finland (2024a) määrittelee tätä koneen kykynä käyttää ihmisen älyyn liitettäviä taitoja, joita ovat mm. päättely, oppiminen, suunnittelu tai luominen. Tekoälyn avulla koneet, laitteet, järjestelmät, ohjelmat ja palvelut ovat kykeneväisiä toimimaan tilanteeseen sopivalla tavalla järkevästi (TEM, 2019, s. 15–16).

Jo 1950- luvulla tutkijoilla oli erilaisia näkemyksiä tekoälyn mahdollisuuksista ja tulevaisuuden muodoista, ja samankaltaiset erimielisyydet tekoälyn mahdollisuuksista ovat jatkuneet vuosikymmeniä (Gaskin, 2008, s. 30–32; Moor, 2006, s. 87–91). Tekoälyn syntymisen taustalla ei ole yksimielistä teoriaa, eikä myöskään yksimielistä tulevaisuuden suuntaa (Moor, 2006). Elinkeinoelämän tutkimuslaitoksen hankkeessa (Kääriäinen ja muut, 2018, s. 9) tekoälyn korostetaan olevan yksittäisen teknologian sijaan useita erilaisia menetelmiä, teknologioita, sovelluksia sekä tutkimussuuntia.

Tekoälyn korostetaan olevan muutakin, kuin teknologiaa, ja nämä kaikki ovat yhdessä osa digitalisaation laajempaa kehitystä (Työ- ja elinkeinoministeriö, 2019, s. 15–16). Sen myötä tekoälyä voidaan hyödyntää niin sanotun monimutkaisten luovien toimintojen automatisointiin, jolloin ihminen vapautuu korkeamman tason ajatteluun. (Ojanperä, 2023, s. 102).

Generatiivinen tekoäly on herättänyt paljon keskustelua sen vaikutuksista työelämään ja yhteiskuntaan. Ojanperä (2023, s. 130) kuvaa tekoälyn luovan uusia mahdollisuuksia ja innovaatioita, sekä muuttavan tapaa, jolla työtä tehdään. Johnson ja muut (2020) kuvaavat tämän johtavan työvoiman muutokseen. Valtioneuvoston julkaisussa (Heinäsenaho ja muut, 2023) viitataan investointipankin Goldman Sachs:in arvioon, jonka mukaan generatiivinen tekoäly saattaa korvata jopa neljäsosan läntisen maailman työpaikoista, ja nostaa tuottavuutta kokonaisuudessaan 1,5 prosenttia. Brynjolfsson ja muut (2023, s. 2–3) tarkentavat tekoälyn käytön hyödyn kohdistuvan työntekijätasolla erityisesti vähän koulutettuihin työntekijöihin, vaikutuksen ollessa minimaalinen verratessa kokeneisiin sekä korkeasti koulutettuihin työntekijöihin.

Tekoälyn hyödynnettävyyden korostetaan kuitenkin kohdistuvan työntekijöiden työn tehostamiseen (Ojanperä, 2023, s. 122) sekä tukemiseen (Johnson, 2020; Ojanperä, 2023, s. 122; Salo, 2024, s. 228). Tekoälystä käytetäänkin sanoja *tukiäly* (Ojanperä, 2023, s. 126) tai *lisätty älykkyys* (Johnson ja muut, 2020). Ojanperä (2023, s. 122) korostaa työntekijöille kohdistetun koulutuksen ja tuen merkitystä tekoälyn hyödynnettävyydessä.

Samoin Ojanperä (2023, s. 142) nostaa esiin, miten generatiivisen tekoälyn kehitys tuo mukanaan uudenlaisia työnteon muotoja, joissa tekoäly toimii enemmän ihmisen tukena kuin tämän korvaajana. Tekoälyn ja koneoppimisen peruskäsitteiden, nykyisen tekniikan tason ja tulevaisuuden visioiden ymmärtämisen merkitystä korostetaan sekä yleisesti että ammatillisesta näkökulmasta kuten terveydenhuollossa (Matheny ja muut, 2019).

Tekoälyn kehityksen myötä erityisen huomion kohteeksi on noussut generatiivinen tekoäly, joka perustuu suurten kielimallien ja neuroverkkojen hyödyntämiseen. Salo (2023, s. 14–15) kuvaa neuroverkkoja ihmisen hermoston toiminnasta inspiroituneiksi tietojenkäsittelymalleiksi, jotka simuloivat aivojen toimintaa yhdistämällä suuren määrän laskentayksiköitä eli neuroneja. Näin muodostuvat verkostot kykenevät tunnistamaan monimutkaisia elementtejä ja datajoukkoja sekä tuottamaan niiden perusteella erilaisia ennusteita. Ojanperä (2023, s. 27–29) jatkaa, että generatiivinen tekoäly, kuten GPT-mallit ja niiden pohjalta kehitetyt sovellukset kuten ChatGPT, kykenee luomaan uutta sisältöä, kuten tekstiä, kuvia ja ohjelmakoodia. Näiden järjestelmien taustalla ovat niin sanotut massiiviset perusmallit, jotka rakentuvat laajoihin tietoaineistoihin ja hyvin monimutkaisiin neuroverkkorakenteisiin.

Nykyiset generatiiviset tekoälymallit kuten OpenAI:n kehittämä GPT-4 perustuvat malleihin, jotka mahdollistavat laajamittaisen kielianalyysin ja kontekstuaalisen ymmärryksen. Näiden mallien oppimisprosessia on verrattu ihmisen kielen omaksumiseen. Ensin tunnistetaan yksittäiset sanat, sen jälkeen lauseet ja lauserakenteet ja vähitellen järjestelmä oppii ymmärtämään ja ennustamaan merkityksiä (Ojanperä, 2023, s. 38-39.)

Generatiivisen tekoälyn kehitys on avannut uusia mahdollisuuksia esimerkiksi työelämässä, koulutuksessa ja luovilla aloilla. Se on tuonut mukanaan uusia mahdollisuuksia ja haasteita, jotka vaikuttavat laajasti yhteiskuntaan ja työelämään. Sen myötä on myös noussut keskustelua tekoälyn eettisistä vaikutuksista, kuten tiedon väärinkäytöstä, tekijänoikeuksista ja tekoälyn läpinäkyvyydestä (Ojanperä, 2023, s. 142).

Keskustelua käydään erityisesti siitä, miten generatiivista tekoälyä tulisi säädellä ja hyödyntää vastuullisesti yhteiskunnassa. Tekoälyn hyödyntämiseen liittyvät eettiset kysymykset sekä vaikutukset vaativat jatkuvaa tarkastelua ja sääntelyä.

5 Tekoälyn käytön eettiset näkökulmat sosiaali- ja terveydenhuollossa

Tekoälyn on todettu tekevän sosiaali- ja terveysalalla nopeaa muutosta ja kehitystä, ja tähän muutokseen liitettävät tehokkuuden elementit tulevat useiden ennusteiden mukaan tehostumaan entisestään tulevaisuudessa. Strange ja Tucker (2024) pitävät COVID-19-pandemiaa merkittävänä käännekohtana tekoälyn käyttöönotossa sosiaali- ja terveydenhuollossa. Ennen pandemiaa tekoälyn käyttöä tarkasteltiin usein kriittisesti erityisesti tietosuojan, valvonnan ja suurten teknologiayritysten vallan näkökulmasta. Pandemian myötä tekoälystä tuli kuitenkin keskeinen ratkaisu terveyskriisin hallinnassa, kuten tartunnanjäljityksessä, diagnooseissa ja terveydenhuollon resurssien hallinnassa (Strange & Tucker, 2024.)

Keinotekoisen älykkyyden käyttö terveydenhuollossa aiheuttaa kuitenkin väistämättä useita eettisiä ongelmia, jotka hidastavat sen laajamittaista käyttöönottoa Kumar ja muut (2023) esittävät, että potilasturvallisuuden varmistaminen edellyttää monien keskeisten haasteiden, kuten yksityisyydensuojan, luottamuksen, oikeudenmukaisuuden ja sääntelyn, ratkaisemista. Strange ja Tucker (2024) tarkastelevat eettisiä haasteita erityisesti algoritmisen vinouman näkökulmasta ja pohtivat sen seurauksia. Heidän mukaansa tekoälypohjaisten järjestelmien hyödyntämä potilastietojen kerääminen ja analysointi voi lisätä terveydenhuollon tehokkuutta ja parantaa hoidon saavutettavuutta, mutta samalla se voi vaarantaa henkilötietojen suojan ja johtaa epäoikeudenmukaiseen päätöksentekoon. Erityisesti algoritmien käyttö diagnoosien ja hoitosuositusten tukena voi sisältää vinoumia, jotka syrjivät tiettyjä väestöryhmiä, jos mallit on koulutettu epätasapainoisella tai harhauttavalla datalla. Lisäksi Strange ja Tucker (2024) nostavat esiin tekoälyn kehitykseen liittyvät kaupalliset intressit. Suurten teknologiayritysten keskeinen rooli terveydenhuollon digitalisaatiossa herättää huolta siitä, missä määrin kehitystä ohjaavat ensisijaisesti taloudelliset tavoitteet julkisen terveyden edun sijaan.

Käsiteltäessä potilaiden henkilö- ja terveystietoja ollaan tekemisissä aina korkean riskin tietojen kanssa. Tässä yhteydessä tietosuoja ja potilaiden yksityisyyden suojaaminen nousevat ensisijaisiksi asioiksi. Terveydenhuollossa tietosuojan varmistaminen on yksi tekoälyjärjestelmiin liittyvä prioriteetti (Khosravi ja muut, 2024). Wang ja muut (2023) korostavat, että terveydenhuollossa tekoälyn tulee noudattaa tiukimpia eettisiä standardeja. Eettisiä ohjeistuksia noudattamalla voidaan varmistua tekoälyn vastuullisesta käytöstä, mikä käsittää myös potilaiden yksityisyyden suojaamisen luotettavalla tiedonvaihdolla.

Khosravi ja muut (2024) käsittelevät vuoteen 2023 mennessä tehtyjen tutkimusten perusteella näitä tekoälyn eettisiä haasteita, jotka osaltaan vaikuttavat tekoälyn suunnitteluun, kehitykseen, käyttöönottoon sekä vaikuttavuuteen terveydenhuollossa. Nämä kaikki osa-alueet ovat osa vastuullista terveydenhuoltoa. Khosravi ja muut (2024) painottavat, että tekoälyn käyttö terveydenhuollossa edellyttää monitieteistä ja osallistavaa lähestymistapaa, sillä tekoälyteknologian käytössä tulee huomioida juuri terveydenhuollon tavoitteet ja arvot. Tekoälyn käyttöön liittyvien eettisten haasteiden osalta on myös todennettu, että yksinkertaisuuden tai universaalisuuden sijaan ongelmat näyttävät olevan usein hyvin kontekstisidonnaisia ja dynaamisia. (Khosravi ja muut, 2024.)

Al-Ansin ja muiden (2024) mukaan keskeisimpiä esille nousevia haasteita tekoälyn käytössä sosiaali- ja terveydenhuollossa ovat läpinäkyvyys ja selitettävyys, kun pyritään välttämään tekoälyyn liittyvää epäluottamusta, syrjintää tai oikeudellisia ongelmia. Näiden kehittäminen edesauttaa tekoälyn vastuullista ja oikeudenmukaisuutta yleisesti julkishallinnossa. (Al-Ansi ja muut, 2024.)

Vaikka nämä eettiset kysymykset ovat edelleen läsnä, Strange ja Tucker (2024) näkevät, että suhtautuminen tekoälyyn on muuttunut pandemian myötä. He osoittavat, että globaalin hallinnan instituutiot, kuten WHO, UNESCO ja OHCHR, ovat alkaneet painottaa tekoälyn myönteisiä vaikutuksia, erityisesti sen merkitystä terveydenhuollon kriisien

hallinnassa ja palveluiden tehostamisessa. COVID-19-pandemia vahvisti käsitystä tekoälystä keskeisenä ratkaisuna. Tämä muutos heijastaa laajempaa kehitystä, jossa tekoäly nähdään yhä tärkeämpänä osana modernia terveydenhuoltoa, samalla kun siihen liittyviä ongelmia, kuten algoritmista syrjintää ja tietoturvariskejä, käsitellään vähemmän kriittisesti. Tämä kehitys herättää kysymyksiä siitä, missä määrin pandemian myötä syntynyt ajattelutapa on vaikuttanut tekoälyä koskevaan sääntelyyn ja sen hyväksyttävyyteen sosiaali- ja terveydenhuollossa (Strange & Tucker, 2024).

5.1 Tietosuoja, yksityisyys ja korkean riskin tiedot

Tekoälymallit kykenevät keräämään ja käsittelemään valtavia tietomääriä. Sosiaali- ja terveydenhuollossa käytettävien arkaluontoisten, korkean riskin tietojen käsittely herättää odotetusti tietosuoja- aiheisia kysymyksiä sekä huolta yksityisyyden säilymisestä. Jo itsessään suuri tietomäärä, joka käsittää arkaluontoisia potilastietoja, aiheuttaa huolta yksityisyyden, tietojen luvattoman käytön sekä tietovuotojen osalta (Kumar ja muut, 2023.) Strange ja Tucker (2024) muistuttavat, että nämä riskit ovat olennaisia terveydenhuollon siirtyessä digitaalisiin järjestelmiin. WHO:n raportissa on myös tunnistettu riski, että laajamittainen terveysdatan kerääminen, esimerkiksi kontaktien jäljittäminen ja tautiseuranta, voi heikentää yksityisyydensuojaa, jos tietoja käytetään ilman riittävää suojaa ja läpinäkyvyyttä (Strange & Tucker, 2024).

Yksityisyys nähdään perusoikeutena ja -arvona, joka koetaan uhatuksi käytettäessä tekoälyteknologiaa terveydenhuollossa. Yksityisyyden säilyttämisellä tarkoitetaan yksilön oikeutta hallita omia henkilökohtaisia ja arkaluonteisia tietojaan. (Cohen ja muut, 2020.) Tekoäly käsittelee näitä yksityisiä tietoja keräämällä, tallentamalla ja jakamalla. Yksityisyyden suojan uhan tunnetta vahvistavat uudenlaiset laskennalliset tekniikat, joiden avulla yksilöiden tunnistaminen on mahdollista, vaikka tiedot olisivat esimerkiksi terveystiedoista anonymisoitu ja tunnisteet olisi poistettu. Tämänkaltaiset uudet teknologian kehittymisen myötä ilmenneet uhkakuvat näkyvät viimeaikaisissa tutkimustuloksissa. (Khosravi ja muut, 2024.)

Tekoälyyn liittyvä läpinäkyvyyden elementti korostuu myös tietosuojaan liittyvissä kysymyksissä. Läpinäkyvyys ja vastuullisuus tietojen käsittelyssä vahvistavat tekoälyn hyötyjen käyttöä suojaamalla potilaiden yksityisyyttä ja luottamusta tekoälyä kohtaan. (Wang ja muut, 2023.) Oleellista on myös potilailta saatu suostumus tietojen käyttöön (Kumar ja muut, 2023). Samaan aikaan tietosuoja asettaa ristiriitoja läpinäkyvyyden osalta, sillä avoimuutta arkaluontoisten tietojen osalta ei voida toteuttaa. (Al-Ansi ja muut, 2024).

Strange ja Tucker (2024) osoittavat, että yksityisyydensuojan kysymykset olivat pitkään keskeinen osa tekoälyn sääntelyä ja globaalien hallintaelinten, kuten WHO:n ja Unescon, asiakirjoja. Ennen COVID-19-pandemiaa keskustelussa painottuivat huoli valvonnan lisääntymisestä, teknologiayritysten vallan kasvusta ja potilaiden oikeudesta hallita omia tietojaan. Pandemian aikana ja sen jälkeen painopiste kuitenkin muuttui. Tekoälyn nopea käyttöönotto kriisitilanteissa johti siihen, ettei yksityisyyden suojaa aina arvioitu yhtä perusteellisesti kuin aiemmin (Strange & Tucker, 2024). WHO (2021) korostaa, että erityisesti tartunnanjäljitysjärjestelmien ja diagnostiikkatyökalujen kaltaiset teknologiat laajensivat terveysdatan käyttöä tavoilla, joiden eettisiä ja oikeudellisia riskejä ei ollut täysin ennakoitu. Tämä kehitys viittaa siihen, että yksityisyydensuojan merkitys voi jäädä taka-alalle, jos teknologioita otetaan käyttöön kiireellisesti ilman riittävää harkintaa. (Strange & Tucker, 2024).

Tekoälyn käytön yhteydessä on todettu useita tietosuojaloukkauksia. Esimerkiksi Siala ja Wang (2022) nostavat esiin tapauksen, jossa Iso-Britanniassa potilaiden suostumuspyynnöt ohitettiin, minkä seurauksena arkaluonteisiin potilastietoihin päästiin käsiksi luvattomasti, vaarantaen jopa 1,6 miljoonan potilaan yksityisyyden. Tällaiset tietosuojarikkomukset lisäävät terveysdataan liittyviä riskejä ja herättävät huolta sekä potilaiden että terveydenhuollon ammattilaisten keskuudessa (Siala & Wang, 2022). Kehittyvissä maissa nämä riskit voivat olla vielä suurempia, sillä

tietosuojalainsäädäntö ja tekninen infrastruktuuri ovat usein puutteellisia, mikä tekee potilastiedoista alttiimpia väärinkäytöksille (Ferede, 2025).

Tietosuoja, yksityisyys ja korkean riskin tietojen käsittely muodostavat keskeisen eettisen perustan tekoälyn hyödyntämiselle sosiaali- ja terveydenhuollossa. Ne eivät ole vain teknisiä tai juridisia vaatimuksia, vaan heijastavat yksilön oikeuksia, itsemääräämisoikeutta ja luottamusta hoitojärjestelmiin. Aiemmassa tutkimuksessa on tarkasteltu tietosuojaa ja yksityisyyttä eri näkökulmista, joista osa tutkimuksesta keskittyy datan käsittelyn turvallisuuteen ja sääntelyyn, kuten GDPR:n ja tekoälyn rajapintaan (Sartor & Lagioia, 2020), osa korostaa yksityisyyttä globaalina periaatteena ja nostaa esiin valvontaan liittyviä riskejä (Strange & Tucker, 2024; WHO, 2021), ja osa tarkastelee tietosuojaa laajemmin osana luottamuksen ja hyväksyttävyyden rakentumista terveydenhuollon kontekstissa (Roberts ja muut, 2020).

5.2 Lait ja sääntely

Tekoälyn nopea yleistyminen sosiaali- ja terveydenhuollossa luo painetta lainsäädännölle ja valvonnalle. Kasvava käyttö ja uudet sovellukset haastavat sääntelyviranomaisia ja lainsäätäjiä kehittämään joustavia mutta luotettavia kehyksiä, jotka turvaavat potilaiden oikeudet ja luovat edellytykset turvalliselle, läpinäkyvälle ja oikeudenmukaiselle teknologian käytölle (Bouncer ja muut, 2022, s. 26). Ojanperä (2023, s. 158) korostaa, että sääntelyn on pystyttävä reagoimaan teknologian kehitykseen ilman, että se menettää ennakoitavuuttaan ja selkeyttään. Strange ja Tucker (2024) painottavat, että keskeisenä tavoitteena on tasapainottaa innovaation edistäminen ja riskien hallinta niin, että tekoälyratkaisut tukevat hoidon laatua, mutta eivät vaaranna yksityisyyttä tai perusoikeuksia.

Sosiaali- ja terveydenhuollon arkaluontoisten henkilötietojen käsittely edellyttää, että tietosuojamekanismeja ja lakisääntelyä noudatetaan systemaattisesti. Al-Ansi ja muut (2024) korostavat, että viranomaisten ja lainsäätäjien on asetettava

vähimmäisvaatimukset tekoälypohjaisten päätösten perusteltavuudelle ja raportoinnille. GDPR ja muu tietosuojalainsäädäntö takaavat kansalaisten oikeuden tietää, miten heihin liittyviä päätöksiä tehdään, ja mahdollisuuden tarkastella ja kyseenalaistaa niitä (Al-Ansi ja muut, 2024). Kumar ja muut (2023) painottavat, että sääntelyn rinnalle tarvitaan kansainvälisesti hyväksyttyjä eettisiä periaatteita erityisesti terveydenhuollon sovelluksissa.

Tekoälyjärjestelmien käyttö edellyttää myös jatkuvaa valvontaa ja selkeitä standardeja. Siala ja Wang (2022) esittävät, että järjestelmien luotettavuus varmistetaan säännöllisillä tarkastuksilla ja arvioinneilla sekä eettisten ja oikeudenmukaisten toimintaperiaatteiden noudattamisella niin kehittämisessä kuin käytössä. Heidän mukaansa tekoälyn on toimittava hoitoprosessissa ihmisen päätöksenteon tukena, ei sen korvaajana. Ferede (2025) täydentää näkemystä toteamalla, että erityisesti kehittyvissä maissa sääntelyn ja ohjeistusten puute voi lisätä tietosuojaloukkauksia ja hoidon eriarvoisuutta. Kansainvälisesti yhtenäiset standardit ja eettinen hallinta ovat siksi välttämättömiä.

Sääntelyn haasteena on teknologian nopea kehitys. Ojanperä (2023, s. 161) huomauttaa, että generatiivisen tekoälyn, kuten ChatGPT:n, nopea kehitys on tehnyt osasta sääntelykehyksiä vanhentuneita jo ennen niiden käyttöönottoa. Myös sääntelyn taloudelliset vaikutukset ovat herättäneet keskustelua. Strange ja Tucker (2024) korostavat, että vaikka COVID-19-pandemia loi perusteen tekoälyn nopealle käyttöönotolle, yksityisyyden suojan ja ihmisoikeuksien on pysyttävä sääntelyn keskiössä. Muutoin vaarana on pysyvä muutos siinä, miten tietosuojaa ja yksityisyyttä kunnioitetaan terveydenhuollossa ja yhteiskunnassa laajemmin.

Keskeisin henkilötietojen käsittelyä ohjaava säädös on EU:n yleinen tietosuoja-asetus (GDPR, EU 2016/679), joka määrittää periaatteet, kuten käyttötarkoitussidonnaisuuden, tietojen minimoinnin ja rekisteröidyn oikeudet. GDPR varmistaa, että tekoälyratkaisut ovat tietosuojaa kunnioittavia ja läpinäkyviä (Euroopan unioni, 2016). EU:n tekoälyasetus (AI Act, EU 2024/1689) puolestaan luo riskiperusteisen sääntelykehyksen,

joka asettaa tiukimmat vaatimukset korkean riskin tekoälyjärjestelmille, kuten kliinisille sovelluksille. Asetus velvoittaa riskienhallintaan, tekniseen dokumentointiin, inhimilliseen valvontaan ja läpinäkyvyyteen sekä vahvistaa kansalaisten oikeutta selkeisiin ja jäljitettävissä oleviin päätöksiin (Al-Ansi ja muut, 2024; Euroopan parlamentti ja neuvosto, 2024).

Lainsäädäntö muodostaa perustan tekoälyratkaisujen eettisesti hyväksyttävälle käytölle. GDPR:n joustava tulkinta mahdollistaa henkilötietojen hyödyntämisen esimerkiksi koneoppimisen harjoitusdatan luomiseen, kunhan käyttö on yhteensopivaa alkuperäisen tarkoituksen kanssa ja sisältää riittävät suojatoimet (Sartor & Lagioia, 2020, s. 76–81). EU:n tekoälyasetus täydentää GDPR:ää ja vahvistaa kansalaisten luottamusta asettamalla velvoitteita, jotka lisäävät selitettävyyttä, jäljitettävyyttä ja turvallisuutta koko tekoälyn elinkaaren ajan (Euroopan parlamentti ja neuvosto, 2024).

Kansallisella tasolla asiakastietolaki, potilaslaki ja asiakaslaki täydentävät EU-sääntelyä turvaamalla asiakkaiden ja potilaiden oikeuden hyvään hoitoon, tietosuojaan ja päätösten perusteltavuuteen (Finlex, 1992/785; 2000/812; 2023/703). Tämä kokonaisuus muodostaa puitteet, joissa tekoälyratkaisujen on toimittava tukien potilaiden oikeusturvaa ja eettisesti kestävää hoitotyötä. Taulukossa 1 on esitetty keskeiset tekoälyn käyttöä ohjaavat säädökset ja niiden merkitys sosiaali- ja terveydenhuollon kontekstissa.

Säädös	Tarkoitus ja merkitys tekoälyn käytölle
EU:n yleinen tietosuoja-asetus (GDPR, EU 2016/679)	Määrittää periaatteet henkilötietojen käsittelylle (käyttötarkoitussidonnaisuus, minimointi, rekisteröidyn oikeudet). Varmistaa, että tekoälyratkaisut ovat tietosuojaa kunnioittavia ja läpinäkyviä.
EU:n tekoälyasetus (AI Act, EU 2024/1689)	Ensimmäinen koko EU:n kattava sitova sääntelykehys tekoälylle. Luokittelee tekoälyjärjestelmät riskitasoihin ja asettaa korkean riskin sovelluksille tiukat vaatimukset mm. riskienhallinnasta, dokumentaatiosta, läpinäkyvyydestä, inhimillisestä valvonnasta ja kyberturvallisuudesta.
EU:n lääkinnällisiä laitteita koskeva asetus (MD-asetus, EU 2017/745)	Velvoittaa lääkinnällisinä laitteina toimivat tekoälyratkaisut CE-merkintään ja tiukkoihin turvallisuus- ja suorituskykyvaatimuksiin ennen käyttöönottoa.
Asiakastietolaki (703/2023)	Ohjaa sosiaali- ja terveydenhuollon asiakas- ja potilastietojen käsittelyä ja tietojärjestelmien vaatimuksia.
Potilaslaki (785/1992) ja asiakaslaki (812/2000)	Turvaavat potilaiden ja asiakkaiden oikeuden hyvään hoitoon, tietosuojaan ja tiedonsaantiin.

Taulukko 1. Keskeiset säädökset tekoälyn käytössä sosiaali- ja terveydenhuollossa.

Lait ja sääntely muodostavat perustan tekoälyn eettisesti hyväksyttävälle käytölle sosiaali- ja terveydenhuollossa. Ne määrittävät yksityisyyden suojan, vastuunjaon ja oikeudenmukaisuuden vähimmäistason, jonka puitteissa eettisiä periaatteita tulkitaan

muuttuvassa teknologisessa ympäristössä (Euroopan unioni, 2016; Euroopan parlamentti ja neuvosto, 2024). Aiemmassa tutkimuksessa on tarkasteltu sääntelyä monipuolisesti eri näkökulmista. Säädösten soveltamista ja oikeudellisia reunaehtoja on analysoitu erityisesti GDPR:n ja tekoälyn rajapinnan yhteydessä (Sartor & Lagioia, 2020) sekä vastuullisuutta ja läpinäkyvyyttä korostavassa tutkimuksessa (Al-Ansi ja muut, 2024). Hyväksyttävyyden ja luottamuksen rakentamista on painotettu kansainvälisissä linjauksissa (WHO, 2021; Kumar ja muut, 2023). Lisäksi valvonnan ja vallan kysymyksiä on tarkasteltu kriittisesti yhteiskunnallisissa analyyseissa, jotka korostavat teknologian käytön laajempia vaikutuksia (Ferede, 2025; Strange & Tucker, 2024).

5.3 Avoimuus ja läpinäkyvyys

Terveydenhuollossa tekoälyn käytössä pyritään eettisesti kestävään ja tehokkaaseen toimintaan. Wang ja muut (2023) toteavat tätä tuettavan erityisesti toiminnan läpinäkyvyydellä, jolla voidaan tarkoittaa sekä algoritmista että yleistä läpinäkyvyyttä ja selitettävyyttä tekoälyn käyttöön liittyen. Läpinäkyvyydellä tarkoitetaan erityisesti sitä, että tekoälyjärjestelmien toimintaperiaatteet, päätöksentekoprosessit ja mahdolliset rajoitukset esitetään selkeästi ja ymmärrettävästi sekä ammattilaisille, potilaille että sidosryhmille. (Khosravi ja muut, 2024.) Al-Ansi ja muut (2024) korostavat, että puutteet tekoälyn pohjalla vaikuttavien algoritmien ymmärtämisessä tai tarkastelussa vaikuttavat sekä hallinnon vastuullisuuteen, että kansalaisten luottamukseen hallintoa ja heidän ohjaamaa digitalisaatiota kohtaan. Selitettävyydellä, eli selittävän tekoälyn (Explainable AI, XAI) käytöllä ja siihen painottamisella tuetaan näiden tekoälyn algoritmien ymmärtämistä. (Al-Ansi ja muut, 2024.)

Avoimuuteen liitetään erityisesti tiedon jakamisen ja saatavuuden varmistaminen. Tekoälyn toimintalogiikan, sen käyttämän datan ja algoritmien käytön tulee olla esitettävissä mahdollisine virheineen ja rajoituksineen. (Khosravi ja muut, 2024.) Niin potilaiden kuin ammattihenkilöiden on tiedettävä, kun he ovat tekemisissä tekoälyn kanssa (Wang ja muut, 2023).

Khosravin ja muiden (2024) tuloksissa näkyy myös selkeästi luottamustekijä, joka pohjautuu erityisesti lääkäreiden ammattitaitoon tekoälyn arviointikyvyssä. On nähtävissä, että lääkäreiden ja terveydenhuollon ammattilaisten ymmärrys ja arviointikyky tekoälyn toiminnasta ja käytöstä vaikuttaa suoraan potilaiden luottamukseen tekoälyä kohtaan. Tämä vaatii ammattilaisten selkeää ja avointa kommunikointia tekoälyn käytöstä, hyödyistä ja riskeistä. Lääkärin rooli nähdään keskeisenä tekijänä siinä, miten potilas kykenee luottamaan tekoälyyn. (Khosravi ja muut, 2024.)

On huomioitavaa, että kehittyneet AI- algoritmit, kuten syväoppiminen ja koneoppimismallit, ovat hyvin monimutkaisia ja vaikeasti ymmärrettävissä myös asiantuntijoille. Tämä vaikeuttaa osaltaan selitettävyyden piirrettä. (Al-Ansi ja muut, 2024.) Kun tähän liitetään terveydenhuollon konteksti, jossa käsitellään arkaluonteisia potilastietoja, on tekoälyjärjestelmien toimintaprosessien selittäminen haastavaa myös tekoälyn kanssa työskenteleville ammattilaisille. Strange ja Tucker (2024) kuvaavatkin tekoälyn päätöksentekoprosessin olevan usein läpinäkymätön. Terveydenhuollossa tämänkaltainen läpinäkymätön prosessi ei mahdollista tarkastelua ja vastuullisuuden varmistamista, mikä on ongelmallista, sillä seuraukset kantautuvat potilaiden hoitoon ja oikeuksiin saakka. Tämä aiheuttaa selkeitä ristiriitoja läpinäkyvyyden ja tietosuojan välille. (Strange ja Tucker, 2024.) Läpinäkyvyyden puuttuminen altistaa tilanteelle, jossa ammattilaisten sekä potilaiden luottamus tekoälyä kohtaan heikkenee, mikä altistaa osaltaan vaikeasti tunnistettaville ja korjattaville virheille. Tällä voi olla vaikutuksia myös hoidon laadun heikkenemiseen, kun annettavaan hoitoon ei uskalleta luottaa. (Wang ja muut, 2023.)

Läpinäkyvyyden ja avoimuuden merkitys ei rajoitu pelkästään riskien hallintaan. Esimerkkinä käytännön keskustelusta aiheen ympärillä voidaan mainita Backlundin (2024) juontama podcast, jossa Haataja ja Puolakka tuovat esiin, että tietoisuus toisten toiminnasta, joka voidaan nähdä osaksi läpinäkyvyyttä, voi edistää innovaatiota. Kun alan toimijat näkevät, mitä muut tekevät, tämä voi lisätä vuorovaikutusta, rohkaista

kokeilemaan uusia ratkaisuja ja tuottaa uusia näkökulmia tekoälyn soveltamiseen. Cohen ja muut (2020) muistuttavat, että myös tietoturvan varmistaminen on olennainen osa luottamuksen rakentamista, sillä ilman sitä avoimuus ei johda hyväksyttävyyteen.

Luottamuksen puute on yhdistetty läpinäkyvyyden puuttumiseen yleisesti monissa yhteyksissä. Tekoälyn hyväksyminen osaksi terveydenhuoltoa edellyttää läpinäkyvyyden ja avoimuuden periaatteiden toteutumista. Näillä voidaan taata järjestelmien eettinen, luotettava ja oikeudenmukainen toiminta ja integraatio terveydenhuoltoon. (Khosravi ja muut, 2024.) Tekoälyjärjestelmien tietoturvan varmistaminen on olennainen osa rakentaessa luottamusta ja tätä kautta hyväksyntää järjestelmien käyttöön terveydenhuollossa (Cohen ja muut, 2020).

Avoimuus ja läpinäkyvyys muodostavat keskeisen eettisen ja käytännöllisen perustan tekoälyn hyväksyttävälle käytölle terveydenhuollossa. Ne tukevat vastuullisuutta, virheiden tunnistamista ja luottamuksen rakentumista sekä mahdollistavat eettisesti kestävän teknologian kehittämisen ja käytön. Aiemmassa tutkimuksessa on korostettu läpinäkyvyyden merkitystä muun muassa järjestelmien toimintalogiikan ja datan avoimuuden varmistamisessa (Khosravi ja muut, 2024; Wang ja muut, 2023), selitettävyyden ja vastuullisuuden vahvistamisessa luottamuksen ja oikeuksien turvaamiseksi (Al-Ansi ja muut, 2024; Strange & Tucker, 2024) sekä innovoinnin ja yhteistyön edistämisessä (Backlund, 2024; Cohen ja muut, 2020).

5.4 Vastuullisuus

Vastuullisuus nousee terveydenhuollossa tärkeäksi teemaksi kietoen yhteen eettisten ja oikeudellisten standardien ja normien noudattamisen. Siala ja Wang (2022) korostavat, että terveydenhuollossa vastuullisuus näyttäytyy vielä haasteena. Khosravi ja muut (2024) pohjaavat vastuullisuuden elementtien merkitystä kirjallisuuden kautta. Sen mukaan tekoälyjärjestelmien vaatimat luottamuksen, oikeudenmukaisuuden, turvallisuuden ja laadun merkitykset korostuvat, kun ollaan tekemisissä

terveydenhuoltoalalla, jossa on kyse ihmiselämästä ja terveydestä. Bouncer ja muut (2022, s. 5) korostavat, että muidenkin terveydenhuoltoon liitettävien asioiden, kuten lääkkeiden tai rokotusten, osalta käyttöön ja kehitykseen osallistuvien tahojen tulee kantaa oma vastuunsa osana tätä käyttö- ja kehittämisprosessia. Tällä tavoin pyritään minimoimaan potilaille koituvien vahinkojen syntymistä. Samalla tavalla vastuun jakamisen osalta odotetaan meneteltävän myös tekoälyä käytettäessä. Yhteistä vastuunjakamista tulee tapahtua niin valmistajien kuin muidenkin tekoälyn käyttöön osallistuvien tahojen kesken.

Bouncer ja muut (2022, s. 5) nostavat esille myös viranomaistahon valvonnan tarpeen. Säännöllisillä tarkastuksilla ja riskinarvioinneilla pystytään osaltaan vaikuttamaan vastuullisuuteen. Wang ja muut (2023) toteavat, että vastuuvelvollisuudet tulee suunnitella selkeästi, ja virheen sattuessa haitat ja virheet tulee pystyä jäljittämään ja edelleen käsittelemään asianmukaisesti. Tässä voidaan käyttää raportointimenetelmiä tai valvontajärjestelmiä, joiden avulla tekoälyn toimintaa voidaan jatkuvasti seurata ja arvioida. (Wang ja muut, 2023.) Kyky virheiden selitettävyyteen tai korjaamiseen perustellaan myös sillä, että kansalaisilla tulee olla mahdollisuus osoittaa valitus tai kyseenalaistaminen virheestä, joka on löydettävissä (Al-Ansi ja muut, 2024.) Kumar ja muut (2023) muistuttavat, että tekoälyn tulee toimia päätöksenteon tukena nojautuen selkeisiin eettisiin ja oikeudellisiin kehyksiin, eikä se saa korvata inhimillistä arviointia (Kumar ja muut, 2023).

Tekoälyn käyttö terveydenhuollossa aiheuttaa merkittäviä puutteita vastuullisuuden jakautumisessa erityisesti silloin, kun vahinkoa on tapahtunut ja yritetään jäljittää tahoa, joka olisi vastuussa tilanteesta. Bouncer ja muut (2022, s. 26-27) toteavat vastuun määrittelyn olevan erittäin epäselvää. Tämä johtuu siitä, että vaikka tekoälyn valmistajan todettaisiin olevan vastuussa virheistä, niin yhteen tahoon katseita ei voida kohdistaa. Tekoälyprosessien taustalla vaikuttavat sekä monimutkainen tekoälyvälinein tapahtuva toiminta, että useat eri toimijat. Toimijoiksi voidaan luetella potilas, kliinikko ja esimerkiksi terveyskeskus, sekä lisäksi tekoälyjärjestelmän kehittäjät, valmistajat sekä

tutkijat. Bouncer ja muut (2022, s. 26-27) nimeävät tällaisen vastuun tunnistamisen ongelman lääketieteelliseksi tekoälyn monitoimijaongelmaksi.

Vastuullisuus muodostaa tekoälyn käytön eettisen ja oikeudellisen perustan terveydenhuollossa. Se liittyy potilasturvallisuuden ja oikeusturvan turvaamiseen sekä järjestelmien oikeudenmukaisuuteen. Aiemmassa tutkimuksessa on tarkasteltu vastuullisuutta muun muassa sääntelyn, valvonnan ja jäljitettävyyden näkökulmista (Al-Ansi ja muut, 2024; Kumar ja muut, 2023; Wang ja muut, 2023) sekä korostettu vastuun selkeää määrittelyä ja jaettua vastuuta monitoimijaympäristössä (Bouncer ja muut, 2022). Lisäksi tutkimus on tuonut esiin, että vastuun käsitys voi vaihdella esimerkiksi kansallisten arvojen ja hallintotapojen mukaan.

5.5 Datan käyttö ja algoritmiset vinoumat

Tekoälyn hyödyntäminen edellyttää laajaa ja laadukasta dataa, ja terveydenhuollon digitalisaatio on keskeinen tekijä tekoälyn käyttöönoton mahdollistamisessa (Johnson ja muut, 2020). Tekoälyn integrointi terveydenhuoltoon tuo mukanaan sekä mahdollisuuksia että haasteita, joista merkittävimpiä ovat datan laatu, turvallisuus ja analytiikan haasteet (Johnson ja muut, 2020).

Algoritmeilla tarkoitetaan tarkkoja toimintaohjeita (Toivonen, 2023, s. 35). Yksi tekoälyn merkittävimmistä ongelmista on algoritminen vinouma, joka voi syntyä heikon ohjelmoinnin tai virheellisen datan seurauksena (Autioniemi, 2020, s. 12). Vaikka algoritmeja testataan laajasti ennen käyttöönottoa, vinoumia ilmenee edelleen usein, mikä voi johtaa epäoikeudenmukaisiin tai virheellisiin lopputuloksiin.

Erityisesti terveydenhuollossa algoritmiset vinoumat voivat vaikuttaa hoitopäätöksiin ja potilaiden kohteluun. Siala ja Wang (2022) tuovat esille tapauksen, jossa algoritmi arvioi terveydenhoitotarpeita eriarvoisesti esimerkiksi ihonvärin perusteella. Samankaltaisia vinoumia on havaittu myös liittyen esimerkiksi sosioekonomiseen taustaan,

asuinalueeseen ja koulutustasoon. Hoffman (2020) huomauttaa, että vääränlaisten luokitusten seuraukset kohdistuvat usein vähemmistöryhmiin, mikä on merkittävä epäoikeudenmukaisuuden lähde. Algoritmista vinoumaa nähdäänkin esiintyvän usein terveydenhuollon kentällä (Hoffman, 2020; Johnson ja muut, 2020; Matheny ja muut, 2019). Lisäksi Ferede (2025) korostaa, että kehittyvissä maissa tekoälyn hyödyntäminen voi pahentaa eriarvoisuutta, sillä algoritmit perustuvat usein kehittyneiden maiden dataan, mikä voi johtaa vinoutuneisiin ja epäluotettaviin päätöksiin paikallisessa kontekstissa.

Vinoumat voivat vaikuttaa paitsi terveydenhuollon yhdenvertaisuuteen myös potilasturvallisuuteen. Esimerkiksi Siala ja Wang (2022) kuvaavat tapauksen, jossa syövän hoitosuositusten kehityksessä käytettiin heikkolaatuista dataa, mikä johti algoritmisen vinouman syntymiseen. Tekoälyjärjestelmä tuotti tämän seurauksena virheellisiä ja jopa potilasturvallisuudelle vaarallisia hoitosuosituksia. Strange ja Tucker (2024) korostavat, että juuri tällaisten riskien vuoksi tekoälyä ei tulisi käyttää yksinään diagnoosien ja hoitosuositusten tekemisessä, vaan inhimillinen arviointi on välttämätöntä.

Algoritmiset vinoumat voivat johtaa myös algoritmiseen syrjintään, jossa tietyt potilasryhmät jäävät järjestelmän päätöksenteossa huomiotta tai joutuvat epäedulliseen asemaan (Siala & Wang, 2022). Tämä voi tarkoittaa esimerkiksi sitä, että algoritmi aliarvioi potilaan terveydelliset riskit, jolloin vakavan sairauden varhainen tunnistaminen jää tekemättä ja potilas altistuu vakaville seurauksille (Ojanen ja muut, 2022). Algoritminen syrjintä kohdistuu usein tiettyihin väestöryhmiin, ja Hoffman (2020) pitää tätä yhtenä merkittävimmistä tekoälyn riskeistä. Siala ja Wang (2022) esittävät useita esimerkkejä, joissa algoritmiset vinoumat aiheuttavat syrjintää ja korostavat kehitystyön sekä koulutuksen merkitystä riskien vähentämisessä.

Matheny ja muut (2019) sekä Johnson ja muut (2020) huomauttavat, että algoritmien vinoumat voivat johtaa potilaan hoidon eriarvoisuuteen ja epäoikeudenmukaisuuteen.

Ferede (2025) täydentää tätä näkökulmaa tuomalla esille, että kehittyvissä maissa tekoälyyn liittyvät resurssivajeet ja lainsäädännölliset puutteet voivat lisätä vinoumien vaikutusta entisestään, jolloin tietyt potilasryhmät voivat jäädä täysin tekoälypohjaisten järjestelmien ulkopuolelle. Tiedostamattomat vinoumat voivat heikentää kliinistä päätöksentekoa ja vaikuttaa terveyden laatuun (Johnson ja muut, 2020).

Kaiken kaikkiaan tekoälyn suurimpia haasteita terveydenhuollossa ovat puolueellisuus ja epäoikeudenmukaisuus, jotka voivat heikentää potilasturvallisuutta ja hoidon laatua (Alhosani & Alhashmi, 2024). Algoritmisten vinoumien ja syrjinnän tunnistaminen, ennaltaehkäisy ja korjaaminen ovat keskeisiä kehitysalueita tekoälyn luotettavuuden ja oikeudenmukaisuuden varmistamisessa. Ferede (2025) painottaa, että erityisesti kehittyvien maiden kohdalla on välttämätöntä varmistaa, että tekoälyjärjestelmät suunnitellaan paikalliset tarpeet huomioiden, jotta ne eivät syvennä jo olemassa olevia terveyseroja vaan edistävät tasa-arvoisempaa hoitoa.

Datan käyttö ja algoritmiset vinoumat muodostavat yhden tekoälyn eettisen tarkastelun ydinalueista. Aiemmassa tutkimuksessa on tarkasteltu erityisesti datan laatua, anonymisointia ja sääntelyä vinoumien hallinnan välineinä (Johnson ja muut, 2020; Sartor & Lagioia, 2020; Toivonen, 2023) sekä vinoumien vaikutuksia oikeudenmukaisuuteen ja yhdenvertaisuuteen ja tarvetta kehittää toimintaperiaatteita, jotka ehkäisevät syrjiviä käytäntöjä (Hoffman, 2020; Siala & Wang, 2022). Lisäksi aikaisempi tutkimus on tuonut esiin, miten datan ja algoritmien vinoumat voivat kietoutua rakenteellisiin eriarvoisuuksiin ja vaikuttaa hoidon saatavuuteen ja laatuun (Ferede, 2025; Roberts ja muut, 2020).

Tämä tutkielma tarkastelee kokonaisuutena sitä, miten tekoälyn käyttöön liittyvät keskeiset eettiset periaatteet, kuten tietosuoja ja yksityisyys, sääntelyyn liittyvät eettiset periaatteet, avoimuus ja läpinäkyvyys, vastuullisuus sekä datan käyttö ja algoritmiset vinoumat, ilmenevät valituissa ohjausdokumenteissa sekä miten ne kytkeytyvät sosiaali- ja terveydenhuollon toimintaympäristöihin ja muovaavat tekoälyn hyväksyttävyyttä

osana sosioteknistä järjestelmää. Teoriaosuus luo näin perustan analyysille, jossa näitä teemoja tarkastellaan systemaattisesti suhteessa aineistoon.

6 Tutkimuksen toteutus

6.1 Tutkimusmenetelmä

Tämän tutkielman metodologinen lähestymistapa perustuu laadulliseen tutkimukseen, jonka tavoitteena on syventää ymmärrystä tutkittavasta ilmiöstä sekä tuottaa merkityksellistä tietoa sen luonteesta ja rakenteista. Laadullista tutkimusta kutsutaan usein ymmärtäväksi tutkimukseksi, sillä se ei pyri mittaamaan ilmiöitä määrällisesti, vaan tulkitsemaan ja kuvaamaan niitä syvällisesti. (Tuomi & Sarajärvi, 2018, s. 25.) Laadullisessa tutkimuksessa pyritään kuvaamaan aineiston merkityksiä systemaattisesti mutta joustavasti, jolloin tutkijan rooli tulkitsijana korostuu (Schreier, 2024).

Tutkielman päätutkimuskysymyksiä ovat: Millaisia eettisiä näkökulmia tekoälyyn liitetään sosiaali- ja terveysalalla? Miten eettiset näkökulmat kytkeytyvät sosioteknisiin kysymyksiin? Ensimmäinen kysymys suuntautuu empiiriseen tarkasteluun ja ohjaa tulosten systemaattista läpikäyntiä aineiston pohjalta. Toinen kysymys on laajempi ja teoreettisempi, ja tarkastelee eettisten näkökulmien ja sosioteknisten kysymysten välistä yhteyttä yleisemmällä tasolla. Tähän liittyviä havaintoja käsitellään tuloksissa erityisesti teknologian ja työn vuorovaikutuksen kautta, mutta sosiotekninen näkökulma nostetaan johtopäätöksissä erikseen esiin laajempana teoreettisena tarkasteluna ja kokoavana huomiona.

Keskeisenä menetelmänä tässä tutkielmassa toimii dokumenttianalyysi. Menetelmä tarkoittaa kirjallisten aineistojen systemaattista tarkastelua, jonka avulla pyritään tunnistamaan niistä olennaisia sisältöjä, merkityksiä ja yhteyksiä (Bowen, 2009). Dokumenteilla viitataan tässä yhteydessä laajasti kaikenlaiseen kirjalliseen materiaaliin, joka voi liittyä esimerkiksi organisaatioiden, viranomaisten tai yhteisöjen toimintaan (Tuomi & Sarajärvi, 2018, s. 117). Prior (2008) korostaa, että dokumentit voivat olla sekä virallisia että epävirallisia tekstejä, ja niiden merkitys rakentuu aina osana sosiaalista ja institutionaalista kontekstia. Lisäksi dokumentteja voidaan tarkastella paitsi niiden sisältämän tiedon myös niiden roolien ja vaikutusten näkökulmasta tutkimuskohteessa.

Prior (2008, s. 832–833) muistuttaa, että dokumentit eivät ole passiivisia tallenteita, vaan niillä on aktiivinen suhde ympäröivään todellisuuteen.

Vaikka dokumenttianalyysiä on perinteisesti käytetty täydentävänä tutkimusmenetelmänä, sitä voidaan hyödyntää myös päämenetelmänä silloin, kun tutkimuksen fokus kohdistuu nimenomaan kirjallisiin aineistoihin (Anttila, 2000, s. 277–278; Bowen, 2009, s. 29; Morgan, 2022, s. 66). Dokumenttianalyysin etuina pidetään muun muassa ajallista ja taloudellista tehokkuutta, hyvää saatavuutta sekä sitä, että analysoitava aineisto on usein tutkijasta riippumatonta (Bowen, 2009, s. 31; Morgan, 2022, s. 65–66). Lisäksi dokumentit sisältävät usein tarkkoja merkintöjä esimerkiksi niiden julkaisuajankohdasta, alkuperästä ja laatijasta, mikä tukee aineiston läpinäkyvää käyttöä tutkimuksessa (Morgan, 2022, s. 65).

Toisaalta dokumenttianalyysin heikkouksina voidaan pitää sitä, että dokumentteja ei ole alun perin tuotettu tutkimustarkoitukseen (Anttila, 2000, s. 278), jolloin ne voivat sisältää puutteita tutkijan kannalta olennaisten kysymysten osalta. Lisäksi erityisesti organisaatiokohtaiset dokumentit voivat heijastaa julkaisijan näkökulmaa, arvoja tai tavoitteita, mikä edellyttää tutkijalta kriittistä otetta ja kontekstin ymmärrystä (Bowen, 2009, s. 31–32; O'Leary, 2021, s. 283). O´Learyn (2021, s. 280–283) mukaan dokumenttiaineiston käyttö vaatiikin tutkijalta tietynlaista joustavuutta ja kykyä tulkita aineistoa avoimesti, mahdollisesti myös tavoilla, joita ei alun perin osattu ennakoida.

Morganin (2022, s. 65–66) mukaan dokumenttianalyysi tarjoaa erityisen käyttökelpoisen lähestymistavan silloin, kun kenttätutkimus ei ole mahdollista esimerkiksi resurssi- tai turvallisuussyistä. Lisäksi dokumenttien analyysi voi vähentää eettisiä huolia, joita liittyy osallistujien suostumukseen, anonymiteettiin tai aliedustettujen ryhmien tutkimiseen (Morgan, 2022, s. 65–66). Tämä on merkityksellistä tutkimuksessani, jossa tarkastellaan tekoälyn käytön eettisiä ulottuvuuksia alalla, jossa tutkimuskohteet voivat olla sensitiivisiä ja tietosuojaa koskevat vaatimukset tiukkoja (WHO, 2021).

Morgan (2022, s. 71–72) korostaa myös, että dokumenttianalyysi ei ole metodologisesti passiivinen toimi, vaan se vaatii tutkijalta kykyä arvioida aineiston autenttisuutta, uskottavuutta, edustavuutta ja merkityksellisyyttä. Näiden neljän arviointiperusteen avulla voidaan varmistaa, että tutkimuksessa käytetty aineisto palvelee tutkimusongelman kannalta tarkoituksenmukaisesti (Morgan, 2022, s. 71–72). Tekoälyn eettisten vaikutusten tarkastelussa sosiaali- ja terveyspalveluissa on erityisen tärkeää perehtyä muun muassa kansallisiin strategioihin, lainsäädäntöluonnoksiin sekä kansainvälisten järjestöjen ohjeistuksiin. Näiden moniulotteisten dokumenttien kriittinen analyysi mahdollistaa sen, että tutkimuksessa voidaan tuottaa syvällistä ymmärrystä niistä vastuullisuuden, läpinäkyvyyden, selitettävyyden ja yksityisyydensuojan vaatimuksista, joita tekoälyn käyttöön tällä sektorilla kohdistuu (Schreier, 2024; WHO, 2021).

6.2 Aineisto

Tutkielman myötä perehdyin laajasti kotimaisiin julkaisuihin, jotka liittyvät tutkielman teoreettiseen viitekehykseen ja käsiteltävään ilmiöön. Aineiston valintaa ohjasivat tutkielman teoreettisena viitekehyksenä toimineet tekoälyn eettiset periaatteet sosiaali- ja terveydenhuollon kontekstissa. Tämän lisäksi sosiotekninen teoria on toiminut taustakehyksenä, vaikuttaen siihen, millaisia aineistoja pidin tutkielman kannalta merkityksellisinä. Tarkastelin julkaisuja erityisesti siitä näkökulmasta, miten niissä käsitellään tekoälyn integrointia osaksi sosiaali- ja terveydenhuollon prosesseja, työn arkea ja inhimillistä vuorovaikutusta. Voidaan todeta, että tekoälyn käytön eettiset ulottuvuudet toimivat analyysin ensisijaisena jäsentäjänä, kun taas sosioteknisen näkökulman kautta tarkastellaan teknologian ja sosiaalisen järjestelmän välistä vuorovaikutusta ja sen merkitystä käytännön sovelluksissa. Tuomen ja Sarajärven (2018, s. 90) kuvailun mukaisesti analyysiä on toteutettu koko tutkimusprosessin ajan, eikä ainoastaan analyysivaiheessa.

Laadullisessa tutkimuksessa käytettävä aineisto voi olla tekstiä, puhetta, ääntä tai visuaalista materiaalia (Eriksson & Kovalainen, 2016, s. 83). Schreierin (2024) mukaan dokumenttianalyysin aineisto voi koostua hyvin moninaisista valmiista materiaalista, kuten virallisista asiakirjoista, sopimuksista, verkkosivujen sisällöistä, arkistomateriaaleista tai mediateksteistä. Tutkielmaan alkukartoituksessa valikoidut julkaisut ovat luonteeltaan selvityksiä, viranomaisraportteja, katsauksia, ohjeistuksia sekä toimintasuosituksia. Ne tarjoavat sisällöllisesti monipuolisen ja ajankohtaisen materiaalin ja soveltuivat dokumenttianalyysimenetelmän aineistoksi.

Käytin aineistona tunnistettavissa olevien toimijoiden tuottamia julkaisuja, jotka ovat julkisesti saatavilla internetissä. Olen ladannut ja säilyttänyt dokumenttien PDF-versiot, jotta aineisto on käsiteltävissä analyysivaiheessa, ja sen alkuperä on tarvittaessa todennettavissa. Nämä asiat tukevat tutkimuksen läpinäkyvyyttä ja vahvistavat sen luotettavuutta. Valitsemieni dokumenttien (ks Liite 1) pituus vaihtelee 9–183 sivun välillä. Käyttämieni dokumenttien etsimisen, arvioinnin ja tallentamisen olen tehnyt maaliskuun 2025- toukokuun 2025 välisenä aikana.

Tutkielman aineistona toimii kahdeksan huolellisesti valittua julkaisua, jotka edustavat sekä kansallista että kansainvälistä asiantuntija- ja viranomaistason keskustelua tekoälyn eettisestä käytöstä sosiaali- ja terveydenhuollossa. Julkaisut käsittelevät tekoälyn käyttöä konkreettisessa sote-kontekstissa ja nostavat esiin eettisiä, toiminnallisia ja tietopoliittisia näkökulmia, jotka ovat olennaisia tutkittavan ilmiön ymmärtämisessä. Aineiston valinnassa on painotettu ajantasaisuutta, luotettavuutta ja sisällöllistä yhteyttä tutkimuskysymyksiin, erityisesti tekoälyn eettisiin reunaehtoihin ja sen vaikutuksiin sosiaali- ja terveysalan toimintaympäristössä. Kuviossa 3 havainnollistetaan aineiston keskeiset valintakriteerit ja niiden merkitykset analyysin kannalta.

Aineisto on koottu tukemaan sekä eettisten periaatteiden analysointia että sosioteknisen viitekehyksen soveltamista. Osa julkaisuista painottaa järjestelmätason ratkaisuja ja tietopolitiikkaa, ja toiset korostavat ihmisen ja teknologian välistä

vuorovaikutusta, asiakaslähtöisyyttä sekä ammatillista toimijuutta. Tämä mahdollistaa tekoälyn eettisten ulottuvuuksien tarkastelun sekä periaatteellisella että käytännön tasolla. Aineisto auttaa hahmottamaan, miten teknologian käyttöönotto jäsentyy osaksi organisaatioiden toimintaa ja rakenteita.

Kuvio 3. Aineiston valintakriteerit. Kuvio laadittu hyödyntäen tekoälyä (ChatGPT, OpenAI, 2024).

Kotimaiset julkaisut tarjoavat ajankohtaista ja syventävää näkökulmaa muun muassa tekoälyn käytännön sovelluksiin, käyttöönoton haasteisiin, datan hallintaan ja teknologian eettisiin reunaehtoihin. Mukana on asiantuntijaraportteja, hallinnollisia

selvityksiä, kyselyaineistoja ja toimintasuosituksia, jotka tuovat esiin sekä strategisen tason linjauksia että konkreettisia havaintoja kentältä. Osassa aineistoa korostuu sosiotekninen näkökulma, jossa tekoäly nähdään osana laajempaa organisatorista ja kulttuurista muutosta.

Kansainväliset julkaisut täydentävät kokonaisuutta tarjoamalla laajemman viitekehyksen tekoälyn eettiseen tarkasteluun. Ne konkretisoivat esimerkiksi EU:n tekoälysääntelyyn liittyviä keskeisiä periaatteita, kuten turvallisuutta, syrjimättömyyttä ja läpinäkyvyyttä, sekä korostavat teknologian eettisesti kestävää käyttöä globaalissa terveydenhuollon kontekstissa. Kansainvälinen aineisto vahvistaa analyysin vertailtavuutta ja auttaa jäsentämään, miten kansalliset linjaukset suhteutuvat osaksi laajempaa eettistä ja oikeudellista keskustelua.

Aineiston merkitys ei rajaudu eksplisiittisiin eettisiin viittauksiin, vaan se tarjoaa empiiristä tukea ymmärrykselle siitä, miten teknologiset ja eettiset kysymykset kietoutuvat osaksi arjen käytäntöjä. Aineisto mahdollistaa sen arvioimisen, kuinka tekoälyn käyttöä sote-sektorilla jäsennetään suomalaisessa keskustelussa ja kuinka nämä näkökulmat asettuvat suhteessa kansainvälisiin periaatteisiin.

6.3 Teoriaohjaava sisällönanalyysi

Puusa (2020, s. 149) kuvaa sisällönanalyysin monivaiheiseksi prosessiksi, jossa tulkinta on jatkuvaa analyysin eri vaiheissa. Menetelmän keskeisiä vaiheita ovat aineiston pelkistäminen, luokittelu ja tulkinta, joiden avulla jäsennetään aineistosta esiin nousevia teemoja ja muodostetaan tutkimuksen tavoitteiden kannalta relevantteja johtopäätöksiä. Laadullisen aineiston monimuotoisuus ja runsaus tekevät analyysista kiinnostavan mutta samalla vaativan (Puusa, 2020, s. 148). Tuomi ja Sarajärvi (2018, s. 104–105) huomauttavat, että analyysia voivat haastaa useat mahdolliset tulkintavaihtoehdot, minkä vuoksi on tärkeää pysyä valittujen teemojen ja näkökulmien piirissä. Puusa (2020, s. 148) korostaa, että analyysin tavoitteena on tuottaa ennen

kaikkea perusteltuja ja oivaltavia tulkintoja. Luotettavuuden ja tieteellisyyden takaamiseksi analyysin tulee olla systemaattinen, huolellisesti dokumentoitu ja jäljitettävissä (Puusa, 2020, s. 145–149).

Tässä tutkielmassa valitsin analyysimenetelmäksi teoriaohjaavan eli abduktiivisen sisällönanalyysin. Menetelmä sijoittuu induktiivisen ja deduktiivisen lähestymistavan välimaastoon ja perustuu sekä aiempaan teoriaan että aineistosta nouseviin uusiin merkityksiin (Tuomi & Sarajärvi, 2018, s. 110–111). Schreier (2024) korostaa, että teoriaohjaava analyysi soveltuu erityisen hyvin monitulkintaisten asiakirjojen jäsentämiseen ja mahdollistaa aineistosta nousevien havaintojen ja teorian välisen vuoropuhelun. Analyysia ohjasivat aiemmasta tutkimuksesta johdetut yläluokat, kun taas alaluokat muodostuivat aineistolähtöisesti. Analyysin logiikka rakentuu abduktiiviseen päättelyyn, jossa tutkija liikkuu joustavasti teorian ja aineiston välillä (Holopainen ja muut, 2020, s. 239; Tuomi & Sarajärvi, 2018, s. 107–111), ja jossa teorian roolina on aineistohavaintojen reflektoiva tulkinta sekä sopivien käsitteellisten linssien etsiminen analyysin tueksi (Kork ja muut, s. 278-280). Tässä työssä abduktiivinen lähestymistapa näkyy erityisesti siinä, että analyysi eteni vuorottelevasti aineistolähtöisten havaintojen ja teoreettisen viitekehyksen välillä, jolloin aineistosta nousseita uusia merkityksiä suhteutettiin jatkuvasti aiempaan tutkimukseen ja käsitteelliseen jäsennykseen.

Käytännössä teoriaohjaavassa sisällönanalyysissä prosessi alkaa aineistoon perehtymisellä, minkä myötä analyysin yläluokat luodaan tutkielmassa käytetyn tutkimustiedon ja kirjallisuuden perusteella. Alaluokkien muodostumiseen vaikuttaa aineistosta nouseva empiirinen sisältö. (Tuomi & Sarajärvi, 2018, s. 133.) Tutkimusprosessissa analyysi käynnistyi laajalla aineistoon perehtymisellä. Julkaisut valikoituivat niiden tarjoamien eettisten, toiminnallisten tai sosioteknisten näkökulmien perusteella. Dokumentoin ja perustelin valinnat systemaattisesti, jotta aineisto muodostaisi tasapainoisen ja tutkimuskysymyksiin vastaamisen kannalta tarkoituksenmukaisen kokonaisuuden.

Teoriaohjaava sisällönanalyysi tarjoaa tarkoituksenmukaisen menetelmän dokumenttiaineiston jäsentämiseen. Se mahdollistaa ilmiöiden tarkastelun monitulkintaisissa asiakirjoissa sekä niiden merkitysten suhteuttamisen tutkimuksen teoreettiseen viitekehykseen Teoriaohjaava sisällönanalyysi tarjoaa tarkoituksenmukaisen menetelmän dokumenttiaineiston jäsentämiseen. Se mahdollistaa ilmiöiden tarkastelun monitulkintaisissa asiakirjoissa sekä niiden merkitysten suhteuttamisen tutkimuksen teoreettiseen viitekehykseen (Tuomi & Sarajärvi, 2018, s. 110–111). Menetelmä sallii myös uusien, teorian ulkopuolelta nousevien havaintojen huomioimisen. Tämä on erityisen perusteltua eettisten ja yhteiskunnallisten jännitteiden, kuten tekoälyn vastuullisuuden, selitettävyyden ja oikeudenmukaisuuden tarkastelussa sosiaali- ja terveydenhuollon kontekstissa.

Teoriaohjaava analyysi tukee myös tutkielman sosioteknistä lähestymistapaa (Orlikowski, 1992; Salovaara, 2024), jonka ytimessä on teknologian ja sosiaalisten rakenteiden välinen vuorovaikutus Tämä näkökulma auttaa jäsentämään, miten tekoälyn eettiset haasteet liittyvät paitsi teknisiin ratkaisuihin myös organisatorisiin käytäntöihin, päätöksenteon prosesseihin ja normatiivisiin odotuksiin.

Yläluokat muodostin aikaisemman kirjallisuuden ja tutkimuksen pohjalta (Tuomi & Sarajärvi, 2018, s. 133). Teoriaan perustuvat viisi pääluokkaa, eli tietosuoja ja yksityisyys, läpinäkyvyys ja selitettävyys, vastuullisuus, data ja algoritminen vinouma sekä teknologian ja työn vuorovaikutus, jäsensivät analyysia. Alaluokat syntyivät aineistosta nousseiden merkitysten perusteella. Huomioin sosioteknisen näkökulman yhtenä pääluokkana, jossa tarkastellaan työn ja teknologian vuorovaikutusta erityisesti tekoälyn käyttöönoton yhteydessä.

Pääluokkien määrittely perustui sekä teoriakirjallisuuteen että tutkijan ennakko-oletuksiin siitä, millaisia teemoja dokumenttiaineistosta todennäköisesti nousee. Jo ennen analyysia oli havaittavissa, että eettisessä keskustelussa korostuvat yksityisyyden

suoja, läpinäkyvyys, vastuullisuus ja algoritmien oikeudenmukaisuus, mikä perustelee niiden valinnan pääluokiksi. Luokka data ja algoritminen vinouma osoittautui analyysissa erityisen kiinnostavaksi, sillä sen esiintyminen oli vähemmän eksplisiittistä ja vaati tarkkaa tulkintaa. Vastaavasti teknologian ja työn vuorovaikutus laajensi analyysin näkökulmaa yksittäisistä eettisistä periaatteista kohti organisatorisia ja rakenteellisia ulottuvuuksia, mikä on johdonmukaista tutkielman sosioteknisen viitekehyksen kanssa.

Analyysi eteni aineiston systemaattisella koodauksella, jonka tavoitteena oli jäsentää aineisto merkityksellisiin kokonaisuuksiin mahdollisimman objektiivisesti ja toistettavasti. Kuten Eriksson ja Kovalainen (2016, s. 119–121) toteavat, koodaus on laadullisen analyysin ydinvaihe, ja teoriaohjaavassa analyysissa se voidaan toteuttaa ennalta määriteltyjen luokkien avulla. Tässä tutkimuksessa rakensin analyysirungon kirjallisuuteen perustuvien pääluokkien varaan. Pyrin toteuttamaan aineiston tulkintaa johdonmukaisesti, jolloin tutkijan omat arvot ja näkemykset pidettiin tietoisesti analyysin ulkopuolella. Pyrin tekemään kaikki analyysivaiheen ratkaisut läpinäkyvästi ja perustellusti, mikä vahvistaa tutkimuksen luotettavuutta ja tieteellistä uskottavuutta.

Toteutin koodauksen Excel-ohjelman avulla, ja kävin aineiston systemaattisesti läpi dokumentti kerrallaan. Kirjasin taulukkoon suorat lainaukset, pelkistetyt ilmaisut, alaluokat, yläluokat sekä pääluokat, ja lisäksi merkitsin sivunumerot jäljitettävyyden varmistamiseksi. Koodaus tarkoittaa tässä yhteydessä lauseiden ja ilmauksien luokittelua niitä kuvaavilla käsitteillä, minkä avulla aineiston sisältö jäsennetään tutkimuskysymysten kannalta merkitykselliseksi kokonaisuudeksi (Eriksson & Kovalainen, 2016, s. 120).

Pelkistämisvaiheessa poistin aineistosta tutkimuksen kannalta epäolennaisen sisällön, ja pyrin säilyttämään ydinsanoman analyysin perustaksi (Tuomi & Sarajärvi, 2018, s. 123). Taulukossa 2 havainnollistan pelkistämisen logiikkaa esimerkkien avulla. Siinä alkuperäiset englanninkieliset lainaukset on tiivistetty tutkimuksen analyysirunkoon

sopiviksi pelkistyksiksi, mikä kuvaa ensimmäistä vaihetta aineiston jäsentämisessä ja teoreettisen viitekehyksen suuntaamisessa.

Suora lainaus	**Pelkistys**
AI technologies should not encode biases to the disadvantage of identifiable groups, especially groups that are already marginalized.	Algoritminen vinouma ei saa syrjiä
Use of limited, low-quality, non-representative data in AI could perpetuate and deepen prejudices and disparities in health care.	Vinoutunut data voi lisätä terveydenhuollon epätasa-arvoa
Tests and evaluations should be regular, transparent and of sufficient breadth to cover differences in the performance of the algorithm according to race, ethnicity, gender, age and other relevant human characteristics.	Algoritmit on testattava säännöllisesti eri väestöryhmissä
Societal bias and discrimination are often replicated by AI technologies...	Yhteiskunnalliset vinoumat heijastuvat tekoälyyn

Taulukko 2. Esimerkki pelkistämisestä.

Tämän vaiheen pohjalta muodostuivat analyysin alaluokat, jotka tarkensivat pääluokkien sisältöä ja rakensivat siltaa yksittäisten ilmaisujen ja teoreettisen viitekehyksen välille. Tämän pohjalta muodostuivat analyysin alaluokat, jotka tarkensivat pääluokkien sisältöä ja rakensivat siltaa yksittäisten ilmausten ja teoreettisen viitekehyksen välille. Taulukossa 3 esitetään, miten yksittäinen ilmaus etenee analyysin eri tasojen läpi, edeten alkuperäisestä lainauksesta pelkistyksen ja alaluokan kautta yläluokkaan ja lopulta pääluokkaan. Taulukko osoittaa, miten aineistohavainnoista johdettiin käsitteellisiä kokonaisuuksia.

Suora lainaus	Pelkistys	Alaluokka	Yläluokka	Kokoava pääluokka
"Tekoälyratkaisun tulee olla ymmärrettävä käyttäjälle."	Käytettävyyden selitettävyys tärkeää	Käytettävyyden vaatimus	Selitettävyys	Läpinäkyvyys ja selitettävyys

Taulukko 3. Esimerkki sisällönanalyysista.

Analyysin selkeyden ja visuaalisen hahmotettavuuden tueksi hyödynsin värikoodausta. Jokaiselle pääluokalle määriteltiin oma väri, jota käytettiin koodatuissa ilmaisuissa. Tämä helpotti teemojen esiintyvyyden hahmottamista ja tuki analyysin eheää etenemistä sekä vertailtavuutta eri dokumenttien välillä. Kokonaisuutena analyysiprosessi eteni vaiheittain aineiston merkitysten tunnistamisesta niiden luokitteluun, pelkistämiseen ja teemoitteluun, eli laadullisen sisällönanalyysin ydinprosessin mukaisesti (Puusa, 2020, s. 152).

Tutkimusprosessin aikana tarkastelin alkuperäisaineistoa toistuvasti suhteessa analyysirunkoon, ja päivitin koodausrunkoa aineistosta nousseiden uusien merkitysten mukaisesti. Esimerkiksi yksittäisiä koodauksia siirrettiin analyysin edetessä toiseen alaluokkaan, mikäli niiden merkitys tarkentui toiseen teemaan sopivammaksi. Kuten Puusa ja Juuti (2020, s. 189) korostavat, laadulliseen analyysiin sisältyy väistämättä tutkijan subjektiivisia valintoja. Tämän vuoksi pyrin tekemään tutkijana oman roolini, tulkintani ja päätökseni mahdollisimman näkyviksi sekä tarkastelemaan niitä reflektiivisesti.

7 Tulokset

Ensimmäisessä tutkimuskysymyksessä pyrittiin selvittämään, millaisia eettisiä näkökulmia tekoälyyn liitetään sosiaali- ja terveysalalla ja miten nämä näkökulmat ilmenevät tarkastelluissa aineistoissa. Kappaleessa on pyritty nostamaan esiin aineistosta nousseita keskeisiä ilmiöitä, jotka syventävät ymmärrystä tekoälyn eettisistä ulottuvuuksista ja niiden merkityksestä sosiaali- ja terveydenhuollon kontekstissa. Toisen tutkimuskysymyksen kautta tarkastellaan, miten nämä eettiset näkökulmat kytkeytyvät sosioteknisiin kysymyksiin ja mitä tämä tarkoittaa teknologian käyttöönoton kannalta.

Tulosten esittely on jaettu alalukuihin analyysissa muodostuneiden pääluokkien mukaisesti. Pääluokat, eli tietosuoja ja yksityisyys, läpinäkyvyys ja selitettävyys, vastuullisuus, data ja algoritminen vinouma sekä teknologian ja työn vuorovaikutus, pohjautuvat tutkimuksen teoreettiseen viitekehykseen ja aineistosta tehtyihin havaintoihin. Tällainen jaottelu tarjoaa systemaattisen ja selkeän rakenteen tulosten esittämiselle sekä mahdollistaa eettisten teemojen tarkastelun suhteessa tutkimuskysymyksiin ja sosiotekniseen viitekehykseen.

7.1 Tietosuoja ja yksityisyys

Tulosten perusteella yksityisyyden ja tietosuojan kysymykset nousevat merkittäviksi ja läpileikkaaviksi teemoiksi kaikissa tarkastelluissa aineistoissa. Ne kytkeytyvät yhtä lailla juridisiin velvoitteisiin, teknologisiin ratkaisuihin kuin eettiseen hyväksyttävyyteen ja luottamukseen sosiaali- ja terveydenhuollon tekoälyratkaisuja kohtaan. Keskeisiksi nousevat rekisteröidyn oikeudet, suostumuksen vapaaehtoisuus, tietosuoja-arvioinnit, sääntelyn tulkintaerot sekä monipuoliset tekniset ja organisatoriset keinot yksityisyyden suojaamiseksi.

Esiin nousee vahva näkemys siitä, että henkilötietojen käsittelyn on perustuttava lakiin tai yksiselitteiseen, aidosti vapaaehtoiseen suostumukseen. Palveluiden käyttöä ei tule

tehdä riippuvaiseksi siitä, että käyttäjä luovuttaa tietonsa tekoälyn käsiteltäväksi. Yksilöllä tulee olla mahdollisuus kieltäytyä ilman haitallisia seurauksia, mikä liittyy kiinteästi niin sanottuun omadata-periaatteeseen, jossa yksilöllä on oikeus hallita omaa dataansa ja päättää sen käytöstä (Ailisto ja muut, 2022). Samalla tunnistetaan automaattiseen profilointiin ja luokitteluun liittyvät riskit, jotka voivat vaikuttaa yksilön oikeuksiin, ja korostetaan tarvetta tuoda nämä riskit näkyviksi tietosuoja-arviointien avulla.

> (1) Yksilön tulee voida luottaa dataa koskevaan tietosuojaan ja hänellä tulee olla oikeus päättää siitä. (Ailisto ja muut, 2022, s. 9)

Tulokset osoittavat, että sääntely ja tekniset ratkaisut on nähtävä toisiaan täydentävinä kokonaisuuksina. GDPR:n ja muun tietosuojasääntelyn tulkintaerot aiheuttavat epävarmuutta, ja yksittäiset oikeustapaukset tai viranomaislinjaukset voivat muodostaa ennakkotapauksia, jotka vaikuttavat laajemmin koko alan käytäntöihin. Selkeän ohjeistuksen ja yhdenmukaisten tulkintojen katsotaan tukevan sekä tietosuojan toteutumista että uusien ratkaisujen käyttöönottoa.

Teknisistä ja organisatorisista ratkaisuista korostuvat anonymisointi, pseudonymisointi (eli tunnistetietojen korvaaminen toisilla tiedoilla siten, ettei henkilö ole suoraan tunnistettavissa), turvalliset käyttöympäristöt, EU-tasoisesti hyväksytyt tietoturvaratkaisut sekä differentiaalinen tietosuoja, joka voi vähentää yksilön tunnistettavuutta myös suurissa aineistoissa. Arkaluontoisten tietojen, kuten lääkitystietojen ja sijaintitietojen, käsittelyssä edellytetään erityistä huolellisuutta. Lisäksi painotetaan tietojen eheyden ja muodon säilymistä järjestelmien välillä osana tietosuojan ja luotettavuuden varmistamista.

Organisatorisella tasolla nähdään tärkeänä kehittää tietoturvaosaamista kaikilla toimijatasoilla, yhtenäistää kansallisia käytäntöjä ja varmistaa viranomaisvalvonnan sekä auditointien riittävyys. Asiakasdatan hallinta, selkeät suostumuskäytännöt ja

tietoturvalliset infrastruktuurit koetaan paitsi oikeudellisina velvoitteina myös luottamusta rakentavina tekijöinä palvelujen käyttäjien ja järjestelmien tarjoajien välillä.

Tuloksissa nousee esiin tasapainon tarve yksityisyyden suojan ja tiedon hyödynnettävyyden välillä. Tämä tulee esille esimerkiksi tilanteissa, joissa terveystietojen tai sijaintitietojen käyttö voi edistää yhteiskunnallisesti merkittäviä tavoitteita, kuten tutkimusta, päätöksentekoa tai ilmastonmuutoksen torjuntaa, mutta samalla luoda yksityisyyteen liittyviä riskejä. Näissä tilanteissa korostetaan läpinäkyvyyden ja luotettavan oikeudellisen kehyksen merkitystä.

Kansainvälisessä kontekstissa korostuu tarve vahvistaa tietosuojaa globaalilla tasolla. Terveyteen liittyvät henkilötiedot ovat erityisen arkaluontoisia ja edellyttävät vahvoja oikeudellisia ja organisatorisia puitteita. Läpinäkyvyys, suostumus ja yksilön informointi nähdään keskeisinä periaatteina, joiden toteutumista tulee arvioida tekoälyjärjestelmien koko elinkaaren ajan.

> (2) It also requires protection of privacy and confidentiality and obtaining valid informed consent through appropriate legal frameworks for data protection. (WHO, 2021, s. 25)

Huomio kiinnittyy myös siihen, että puutteelliset tai kulttuurisesti epäsensitiiviset tietosuojakäytännöt voivat lisätä eriarvoisuutta ja heikentää tekoälyn hyväksyttävyyttä. Tämä näkyy erityisesti kehittyvissä maissa, joissa lainsäädäntö ja tekninen infrastruktuuri voivat olla heikompia.

Tuloksissa painottuu lisäksi se, että tietosuoja ei ole irrallinen osa teknologian kehitystä, vaan sen tulee olla integroitu tekoälyjärjestelmän suunnittelusta käyttöönottoon ja ylläpitoon asti. Tämä kattaa käytettävien datakokonaisuuksien laadun ja tarkoituksenmukaisuuden, rekisteröityjen oikeudet saada tietoa automatisoidun päätöksenteon perusteista sekä mahdollisuuden inhimilliseen valvontaan. Näin

tietosuoja kytkeytyy myös muihin pääluokkiin, kuten läpinäkyvyyteen ja vastuullisuuteen.

Konkreettisella tasolla esiin nousee esimerkkejä teknologisista ratkaisuista ja sääntelyhankkeista, jotka tukevat yksilön oikeuksia. EU-tasolla on kehitteillä sääntelyä, jonka tavoitteena on vahvistaa yksilöiden mahdollisuuksia hallita omia sähköisiä terveystietojaan sekä edistää tietojen turvallista käyttöä tutkimuksessa ja päätöksenteossa. Myös kansalliset mallit, kuten valvotut käyttöympäristöt ja anonymisointipalvelut, nähdään tärkeinä tietosuojan toteutumisen varmistamisessa.

Aineistossa näkyy jännite yksilön tietosuojan ja datan laajan hyödyntämisen välillä. Kaikki dokumentit eivät erittele, miten vapaaehtoisuus ja sekundäärikäyttö sovitetaan yhteen ilman niin sanottua pakotettua suostumusta. Kansalliset ohjeet painottavat käytännön toteutusta, kun taas osa kansainvälisistä linjauksista jää periaatetasolle. Tämä täydentää WHO:n (2021) ja Sartor & Lagioian (2020) esittämiä näkemyksiä osoittamalla, että GDPR:n joustavat tulkintamahdollisuudet voivat käytännössä johtaa tietosuojan toimeenpanon vaihtelevuuteen eri toimijoiden ja tilanteiden välillä.

Kaiken kaikkiaan tulokset osoittavat, että tietosuoja ja yksityisyys ovat paitsi juridisia ja teknisiä kysymyksiä myös keskeisiä luottamuksen rakentajia. Niiden toteutuminen edellyttää johdonmukaista sääntelyä, tehokkaita teknisiä ratkaisuja, yksilön oikeuksien turvaamista sekä eettisesti kestävää tiedonhallintaa. Ilman näitä tekijöitä tekoälyn hyväksyttävä ja vastuullinen käyttö sosiaali- ja terveydenhuollossa on vaarassa, mikä voi heikentää sekä potilaiden että ammattilaisten luottamusta järjestelmiin.

7.2 Läpinäkyvyys ja selitettävyys

Aineiston perusteella läpinäkyvyys ja selitettävyys nousevat keskeisiksi edellytyksiksi tekoälyn eettiselle ja turvalliselle hyödyntämiselle sosiaali- ja terveydenhuollossa. Läpinäkyvyydellä tarkoitetaan sitä, että tekoälyjärjestelmien toimintaperiaatteet,

päätöksentekoprosessit ja mahdolliset rajoitukset esitetään selkeästi ja ymmärrettävästi eri toimijoille, kuten ammattilaisille, potilaille ja sidosryhmille (Khosravi ja muut, 2024). Selitettävyydellä puolestaan viitataan selittävän tekoälyn käyttöön ja siihen liittyviin menetelmiin, joilla pyritään avaamaan järjestelmän algoritmien toimintaa ja päätöksenteon logiikkaa käyttäjille (Al-Ansi ja muut, 2024). Tämä on erityisen tärkeää terveydenhuollon kontekstissa, jossa monimutkaiset algoritmit vaikuttavat potilaiden hoitoon ja oikeuksiin. Erityisesti korkean riskin sovellusten kohdalla käyttäjälle on ilmoitettava selkeästi järjestelmän käytöstä ja sen roolista päätöksenteossa.

Aineiston mukaan merkittävä haaste liittyy siihen, että monet tekoälyratkaisut jäävät läpinäkymättömiksi niin sanottuina “mustina laatikkoina”, jolloin käyttäjät eivät ymmärrä algoritmien toimintaa tai päätöksenteon perusteita. Tämä koskee sekä teknistä läpinäkyvyyttä, kuten datan lähteitä, algoritmien toimintaperiaatteita ja kehityksessä tehtyjä oletuksia, että prosessien läpinäkyvyyttä, kuten käyttöoikeuksia, vastuunjakoa ja hallinnollisia päätöksiä. Erityisesti monimutkaiset mallit, kuten syvät neuroverkot, ovat vaikeasti selitettäviä, mikä rajoittaa käyttäjien mahdollisuuksia arvioida niiden toimintaa ja luotettavuutta. Aineistossa korostuu näkökulma, jonka mukaan eettinen suunnittelu ei voi rajoittua vain teknisiin ratkaisuihin, vaan siihen sisältyy myös kansalaisten tietoisuuden vahvistaminen ja selitettävyyden edistäminen.

> (1) Kykyä tunnistaa ne avaintekijät, jotka vaikuttavat algoritmin päätöksentekoprosessiin, kutsutaan selitettävyydeksi. (Leikas, 2022, s. 118)

Läpinäkyvyyttä voidaan edistää useilla käytännön mekanismeilla. Näitä ovat esimerkiksi selkeät ilmoitukset käyttäjälle automaation käytöstä, tekoälytuotosten merkintä ja tunnistettavuus, järjestelmien jäljitettävyys sekä avoimen lähdekoodin hyödyntäminen osallistumisen ja tarkastettavuuden lisäämiseksi. Myös käyttäjäviestinnän ymmärrettävyys ja saavutettavuus korostuvat, jotta tekninen tieto ja käyttöperiaatteet välittyvät selkeästi eri kohderyhmille.

Aineistossa nousi esille myös se, miten eri tavoin läpinäkyvyyttä kuvataan, sillä osassa dokumentteja painottuu tekninen, ja toisissa taas hallinnollinen tai viestinnällinen ulottuvuus. Tämä täydentää Khosravin ja muiden (2024) sekä Al-Ansin ja muiden (2024) näkemyksiä osoittamalla, että selitettävyys edellyttää myös prosessin jäljitettävyyttä ja vastuusuhteiden avaamista.

Aineistossa läpinäkyvyys ja selitettävyys jäsentyvät neljään toisiinsa kietoutuvaan kokonaisuuteen. Käytetyn datan laatu, alkuperä, käsittely ja algoritmien toimintaperiaatteet tulee avata ja selittää ymmärrettävästi. Lisäksi käyttäjälle on tarjottava selkeää viestintää siitä, milloin tekoälyä käytetään ja millainen rooli sillä on päätöksenteossa. Järjestelmien toiminnan ja päätöksenteon tulee olla jäljitettävissä ja todennettavissa. Viimeisenä kokonaisuutena ilmenee, että läpinäkyvyyden ja selitettävyyden toteutumista tukevat oikeudelliset ja eettiset velvoitteet, jotka ohjaavat teknologian kehittämistä ja käyttöä.

Analyysin havainnot osoittavat, että läpinäkyvyyden ja selitettävyyden vahvistaminen on keskeistä paitsi teknisen toimivuuden ja luotettavuuden, myös käyttäjien luottamuksen ja yhteiskunnallisen hyväksyttävyyden varmistamiseksi.

7.3 Vastuullisuus

Havaintojen perusteella vastuullisuus tekoälyn kehittämisessä ja käytössä rakentuu useista toisiinsa liittyvistä ulottuvuuksista, joissa korostuvat sekä oikeudelliset että eettiset velvoitteet. Ydinasemassa on palvelun käyttäjien ja kansalaisten oikeuksien turvaaminen. Tekoälyn ei tule missään tilanteessa johtaa oikeuksien heikentymiseen, vaan sen käytön on vahvistettava oikeusturvaa ja edistettävä yhdenvertaisuutta. Tämä edellyttää, että järjestelmät suunnitellaan lainsäädäntöä noudattaen, perusoikeusvaikutukset arvioidaan systemaattisesti ja että oikeudelliset vastuut säilyvät selkeästi määriteltyinä myös tekoälyn käytön lisääntyessä.

> (1) Avainasemaan nousee tasapainon löytäminen uuden teknologian hyödyntämisen ja etiikan välillä. Tekoälyn etiikkaa ei tule ymmärtää vain tekemistä rajoittavana tekijänä, vaan myös uutta luovana, synergioita ja mahdollisuuksia lisäävänä. (Työ- ja elinkeinoministeriö, 2019, s. 100)

Eettinen arviointi nousee keskeiseksi osaksi vastuullisuutta. Tekoälyhankkeiden suunnitteluun ja käyttöönottoon tulee aina sisällyttää arvio siitä, miten ratkaisut huomioivat ihmisten tarpeet, turvallisuuden ja perusoikeudet. Vastuullinen tekoäly ei voi perustua pelkästään teknisiin ratkaisuihin, vaan se edellyttää jatkuvaa eettistä harkintaa, keskustelua ja kollektiivista vastuunkantoa. Käyttäjillä, kehittäjillä, viranomaisilla ja palveluntuottajilla on kullakin oma roolinsa. Käyttäjät vastaavat hyväksyttävyydestä ja tarkoituksenmukaisesta käytöstä, kehittäjät järjestelmien luotettavuudesta ja viranomaiset sääntelyn toteutumisesta. Erityisesti terveydenhuollossa korostuvat ammatillinen vastuu, tietosuojan varmistaminen sekä potilasturvallisuuden takaaminen.

Aineistossa nousee esiin myös vastuun jakautumisen ongelmallisuus. Useiden toimijoiden osallistuminen tekoälyjärjestelmien kehittämiseen ja käyttöön voi johtaa niin sanottuun “useiden käsien ongelmaan”, jossa vastuun kohdentaminen hämärtyy. Tämä riski korostaa läpinäkyvyyden ja selkeiden vastuusuhteiden merkitystä. Kollektiivisen vastuun mallit voivat toisaalta vahvistaa vastuunkantoa, mutta ne eivät saa johtaa vastuun hajautumiseen tavalla, joka jättää eettiset velvoitteet toteutumatta. Vastuullisuus ilmenee myös riskienhallinnan näkökulmasta. Tekoälyn käytössä tulee tunnistaa ja minimoida riskit, jotka voivat liittyä esimerkiksi potilasturvallisuuteen, syrjiviin algoritmeihin, tietoturvaan tai resurssien väärään kohdentamiseen. Riskienhallinta edellyttää sekä ennakoivia vaikutusten arviointeja että jatkuvaa seurantaa tekoälyn koko elinkaaren ajan. Tämä koskee erityisesti korkean riskin järjestelmiä, joissa vaikutukset terveyteen, turvallisuuteen ja perusoikeuksiin voivat olla merkittäviä.

Samalla vastuullisuus edellyttää rohkeaa hyötyjen korostamista ja tekoälyn käyttämistä yhteiskunnallisen hyvän edistämiseksi. Analyysin havaintojen perusteella vastuullinen

tekoälyn hyödyntäminen voi tukea ihmiskeskeistä ja oikeudenmukaista kehitystä, luoda hyvinvointia ja talouskasvua sekä vahvistaa luottamusta julkisiin palveluihin. Tämä edellyttää kuitenkin, että tekoälyä koskeva sääntely, eettiset ohjeistukset ja kansainvälinen yhteistyö luovat selkeät pelisäännöt.

Aineistossa on tunnistettavissamyös vastuunjaon haaste, kun useiden toimijoiden osallistuminen hämärtää lopullista vastuuta. Ammattilaiselle jää päätösvalta, mutta päätösperustat voivat olla algoritmisia. Tämä konkretisoi Kumarin ja muiden (2023) sekä Wangin ja muiden (2023) esittämiä vaatimuksia vastuunjaon selkeyttämisestä ja täydentää myös Al-Ansin ja muiden (2024) hallintomallia osoittamalla, että pelkkä periaatteiden taso ei riitä ilman käytännön vastuunjaon toimintatapoja.

Vastuullisuutta ei voida rajata yksittäiseen näkökulmaan, vaan se on monitasoinen kokonaisuus, jossa yhdistyvät juridiset vastuut, eettiset periaatteet, organisatoriset käytännöt ja yksittäisten käyttäjien toiminta. Olennaista on vastuullisuuden jatkuva arviointi, läpinäkyvyys ja yhteisesti jaettu näkemys siitä, mitä vastuullinen tekoäly käytännössä tarkoittaa. Näin voidaan varmistaa, että tekoälyn kehitys ja käyttö tapahtuvat tavalla, joka sekä tukee yksilön oikeuksia että edistää laajemmin sosiaalista oikeudenmukaisuutta ja kestävää yhteiskunnallista hyötyä.

7.4 Data ja algoritminen vinouma

Aineistosta nousee esiin laaja kirjo havaintoja liittyen datan laatuun, saatavuuteen, hyödyntämiseen sekä algoritmiseen vinoumaan. Analyysin havainnot osoittavat, että datan merkitys tekoälyratkaisujen kehittämisessä ja soveltamisessa sosiaali- ja terveydenhuollossa on ratkaiseva, mutta samalla siihen liittyy monia haasteita, jotka voivat johtaa vinoutuneisiin lopputuloksiin ja eettisiin ongelmiin.

Ensinnäkin aineistossa korostuu datan laadun keskeinen merkitys. Laadukas ja edustava data nähdään edellytyksenä tekoälyn hyödyntämiselle tutkimuksessa, hoitoprosesseissa

ja päätöksenteossa. Heikkolaatuinen tai epäedustava data puolestaan vaarantaa tekoälyn tuottamien tulosten luotettavuuden ja voi johtaa vinoumien toistumiseen, syrjiviin päätöksiin sekä yhteiskunnallisesti kestämättömään käyttöön. Myös datan jatkuva laadunvalvonta ja siihen kohdennettava rahoitus nousivat tärkeinä tekijöinä.

Toiseksi aineistossa korostuvat datan saatavuuteen, yhteentoimivuuteen ja omistajuuteen liittyvät haasteet. Tiedon liikkuminen, eri järjestelmien yhteensopivuus sekä avoimen datan hyödyntäminen nähdään keskeisinä edellytyksinä tekoälyratkaisujen kehittämiselle. Datan saavutettavuuden esteiksi tunnistetaan tekniset rajapinnat, lainsäädännölliset rajoitteet ja tietosuojaan liittyvät kysymykset. Samalla keskusteluun nousee datan omistajuus ja suvereniteetti, esimerkiksi tarve turvata eurooppalaisen datan hallinta ja käyttö eurooppalaisissa käsissä.

Kolmas keskeinen havainto on, että algoritmit oppivat jatkuvasti datasta, mutta oppimisen laatu riippuu vahvasti datan kattavuudesta ja monimuotoisuudesta. Syrjivät algoritmit, kontekstisidonnaiset vinoumat sekä koulutusdatan puutteellinen monimuotoisuus tunnistetaan suurimmiksi riskeiksi. Erityisesti vähemmistöjen ja heikommassa asemassa olevien ryhmien osalta vinoutunut koulutusdata voi johtaa syrjiviin ja epätasa-arvoisiin hoitopäätöksiin. Myös kysymysten asettelulla, harjoitusaineiston valinnoilla ja datan keruumenetelmillä on merkittävä vaikutus tekoälyn toimintaan.

Neljänneksi esiin nousee tekoälyn soveltamisen monimuotoisuus ja laaja käyttö potilasdatassa, sensoritiedossa sekä kliinisessä diagnostiikassa. Tekoälyä hyödynnetään esimerkiksi kuvantamisessa, riskiryhmien tunnistamisessa ja sairauksien ennustamisessa. Samalla kuitenkin havaitaan, että mikäli tekoälyn käyttämä data on vinoutunutta, tämä heijastuu suoraan sen päätöksiin ja voi vääristää hoitokäytäntöjä.

Lisäksi aineistosta tulee esiin tarve säännölliselle arvioinnille, valvonnalle ja vinoumien ehkäisylle. Algoritmien suorituskykyä on testattava eri väestöryhmien osalta, ja

harjoitusdatan monipuolisuus on varmistettava. Vinoumien hallintaa voidaan edistää lainsäädännön avulla, mutta myös tekniset ratkaisut, kuten siirto- ja vahvistusoppiminen sekä federatiiviset datan jakamisen mallit, tarjoavat keinoja haasteiden lieventämiseen. Aineistossa korostuu, että datan vinoutuminen ei ole vain teoreettinen riski, vaan se voi heijastua suoraan tekoälyn toimintaan ja vaikuttaa sen lopputuloksiin.

(1) Jos tekoälyn kehittämiseen tai päivittämiseen käytetty data on vinoutunutta tai haitallista, se voi peilautua tekoälyn toimintaan kuten luokitteluun ja päätöksentekoon (Työ- ja elinkeinoministeriö, 2019, s. 109).

Lisäksi esiin nousee huoli datan vinoutumisen seurauksista laajemmassa yhteiskunnallisessa kontekstissa. Tekoälyjärjestelmät voivat vahvistaa jo olemassa olevia yhteiskunnallisia vinoumia ja syrjiviä rakenteita, mikäli datan laatu ja monimuotoisuus eivät ole riittävällä tasolla. Tämä voi näkyä paitsi yksittäisten potilaiden hoitopäätöksissä myös resurssien kohdentamisessa, palveluiden saatavuudessa ja yleisessä luottamuksessa teknologiaa kohtaan. Vinoumien ehkäisy ei siten ole vain tekninen kysymys, vaan myös oikeudenmukaisuuteen, tasa-arvoon ja sosiaaliseen kestävyyteen liittyvä haaste.

Aineistosta nousee myös esiin kielellisen, kulttuurisen ja käsitteellisen vinouman riski. Terveys, riski tai hoitotarve eivät ole universaaleja käsitteitä, vaan niiden tulkinta voi vaihdella kontekstista toiseen. Jos tekoälyratkaisuja kehitetään vain yhden kulttuuriympäristön lähtökohdista, ne voivat menettää käyttökelpoisuutensa ja oikeudenmukaisuutensa toisenlaisissa olosuhteissa. Kansainvälisessä aineistossa korostuu lisäksi, että globaalilla tasolla tekoälyratkaisut perustuvat usein korkean tulotason maiden dataan, jolloin matalan ja keskitulotason maiden väestöt jäävät heikosti edustetuiksi. Tämä voi johtaa virheellisiin johtopäätöksiin, väärään priorisointiin ja jopa hoidon tarpeen ohittamiseen.

Lisäksi osa dokumenteista korostaa datan laatua, mutta puolueellisuuden järjestelmällinen tarkastaminen ja eri väestöryhmien huomioiminen jää usein mainitsematta. Lisäksi tavoite lisätä datan liikkuvuutta on ristiriidassa sen kanssa, että sen omistajuus ja hallinta halutaan pitää tarkasti rajattuina. Tämä vastaa Hoffmanin (2020) sekä Johnsonin ja muiden (2020) havaintoja, mutta tuo esiin tarpeen selkeille ja käytännönläheisille arviointitavoille, joita nykyiset ohjeet eivät aina tarjoa.

Kokonaisuutena analyysin havainnot osoittavat, että data ja algoritminen vinouma muodostavat monitasoisen kysymyksen, jossa yhdistyvät tekniset, oikeudelliset ja eettiset ulottuvuudet. Luotettava ja monimuotoinen data on edellytys oikeudenmukaiselle ja kestävälle tekoälyn hyödyntämiselle sosiaali- ja terveydenhuollossa. Samalla on välttämätöntä varmistaa, että vinoumien riskit tunnistetaan ja niihin puututaan järjestelmällisesti, jotta tekoälyn käytön hyödyt voidaan saavuttaa ilman, että ne syventävät olemassa olevia eriarvoisuuksia.

7.5 Teknologian ja työn vuorovaikutus

Teknologian ja työn vuorovaikutus on keskeinen tekoälyn eettisen tarkastelun näkökulma sosiaali- ja terveydenhuollossa. Sosioteknisen teorian mukaan teknologia ei ole neutraali väline, vaan sen merkitys rakentuu vuorovaikutuksessa työn käytäntöjen ja organisaation rakenteiden kanssa (Orlikowski, 1992; Salovaara, 2024). Tämän vuoksi tässä luvussa tarkastellaan erikseen, miten tekoälyn käyttöönotto muuttaa työn sisältöä, työnjakoa ja vuorovaikutusta sekä millaisia mahdollisuuksia ja haasteita nämä muutokset tuovat mukanaan.

Tekoälyn käyttöönotto sosiaali- ja terveydenhuollossa sekä laajemmin yhteiskunnassa muokkaa työn ja teknologian välistä suhdetta monella tasolla. Keskeiseksi nousee ajatus siitä, että teknologia ei itsessään ratkaise ongelmia, vaan sen toimivuus rakentuu vuorovaikutuksessa ihmisen työn, osaamisen ja organisatoristen prosessien kanssa.

Tekoäly näyttäytyy yhtä aikaa työn tehostajana, kuormituksen aiheuttajana sekä muutosvoimana, joka edellyttää jatkuvaa oppimista ja sopeutumista.

Tekoäly voi tukea ammattilaisten työtä esimerkiksi tuottamalla luonnoksia, kokoamalla tietoa ja helpottamalla rutiininomaisia tehtäviä, jolloin ihmiselle jää enemmän aikaa keskittyä asiantuntijuutta vaativiin ja inhimillistä vuorovaikutusta edellyttäviin tehtäviin. Ammattilaisen rooli ei kuitenkaan katoa, vaan korostuu hyväksyntä- ja tarkistusvastuun kautta. Tämä luo uudenlaisen työnjaon, jossa tekoäly toimii ehdotusten ja ratkaisujen tuottajana, ja ihminen tekee lopulliset päätökset sekä kantaa eettisen ja ammatillisen vastuun.

> (1) Suuririskisyyden poissulkemiseksi tekoälyratkaisun käyttöönotossa onkin suositeltavaa toteuttaa tarvittavat hallinnolliset ja tekniset toimet sen varmistamiseksi, että sote-ammattilainen todella tarkistaa tekoälyn luonnostelemat kirjaukset ja tekee asiakasta tai potilasta koskevat päätökset itse kirjausten pohjalta. (DigiFinland, 2024b, s.9)

Vuorovaikutuksessa korostuu myös käytettävyyden ja sujuvuuden merkitys. Jos järjestelmä koetaan hankalaksi tai vuorovaikutus kömpelöksi, se estää teknologian omaksumisen. Käytettävyyttä parantavat ratkaisut, kuten puheohjattavat palvelut ja reaaliaikainen yhteistyö tekoälyn kanssa, voivat sen sijaan lisätä työn mielekkyyttä ja vähentää hallinnollista kuormitusta. Tämä heijastuu myös työhyvinvointiin, kun tekoäly vapauttaa työntekijöitä rutiineista ja mahdollistaa mielekkäämpien työtehtävien tekemisen.

Teknologian ja työn vuorovaikutuksessa nousee esiin moniammatillisen yhteistyön haasteet ja mahdollisuudet. Hajanaiset tietojärjestelmät ja standardoinnin puute voivat hidastaa yhteistyötä ja lisätä manuaalista työtä. Toisaalta avoimiin rajapintoihin ja yhteisiin tietomalleihin perustuvat ratkaisut tukevat tiedon sujuvaa liikkuvuutta ja helpottavat työn koordinointia. Tämä edellyttää myös kansallista koordinaatiota ja yhtenäisiä käytäntöjä, jotta tekoälyratkaisut voidaan viedä laajempaan käyttöön.

Tekoälyn vaikutus työn sisältöön on kaksijakoinen. Se voi parantaa tehokkuutta ja tukea päätöksentekoa, mutta samalla lisätä työkuormaa, erityisesti silloin kun uudet järjestelmät vaativat lisää valvontaa, datan hallintaa ja teknisiä perustaitoja. Tämä korostaa koulutuksen, osaamisen kehittämisen ja muutosjohtamisen merkitystä. Työntekijöiden digiosaamisesta tulee välttämätön edellytys, ja samalla tarvitaan uusia organisatorisia rakenteita, jotka tukevat työn ja teknologian yhteensovittamista.

Vuorovaikutus ei koske vain työntekijän ja teknologian suhdetta, vaan ulottuu myös asiakassuhteisiin. Tekoäly voi tukea ammattilaisen ja asiakkaan välistä vuorovaikutusta tai muuttaa sitä, esimerkiksi mahdollistamalla suoraa asiointia ilman ammattilaisen väliintuloa. Tämä voi lisätä saavutettavuutta, mutta samalla herättää kysymyksiä valvonnasta, luottamuksesta ja työn sisällön muuttumisesta. Potilaan ja ammattilaisen välinen suhde voi muuttua, jos tekoälylle annetaan merkittävä rooli päätöksenteossa.

Työn ja teknologian vuorovaikutusta määrittää myös se, että tekoäly ei ole irrallinen työkalu vaan osa laajempaa teknologista ekosysteemiä. Sen hyödyntäminen vaatii datan lisäksi muita teknologioita ja järjestelmiä. Tämä moniteknologinen ympäristö lisää työn monimutkaisuutta ja edellyttää kykyä hallita eri teknologioiden keskinäistä riippuvuutta. Samalla se avaa mahdollisuuksia kehittää palveluja, jotka helpottavat kansalaisten arkea, lisäävät saavutettavuutta ja parantavat yhteiskunnan resilienssiä.

Kokonaisuutena työn ja teknologian vuorovaikutus rakentuu tasapainosta hyötyjen ja kuormituksen välillä. Tekoäly voi merkittävästi tehostaa työtä, lisätä työn mielekkyyttä ja tukea yhteiskunnallisia tavoitteita, mutta vain silloin kun sen käyttöönotto huomioi työntekijöiden osaamisen, vuorovaikutuksen laatuun liittyvät tekijät sekä eettisen vastuun rajat. Näitä tuloksia voidaan syventää tarkastelemalla, miten eri eettiset ulottuvuudet kytkeytyvät työn käytäntöihin ja sosioteknisen näkökulmaan. Taulukkoon 4 on koottu pääluokat, niiden vaikutukset työhön sekä tulkinta sosioteknisen teorian valossa. Tämä auttaa hahmottamaan, miten teknologiset ja sosiaaliset tekijät lomittuvat tekoälyn käyttöönotossa ja millaisia muutoksia ne tuovat työn arkeen. Taulukko

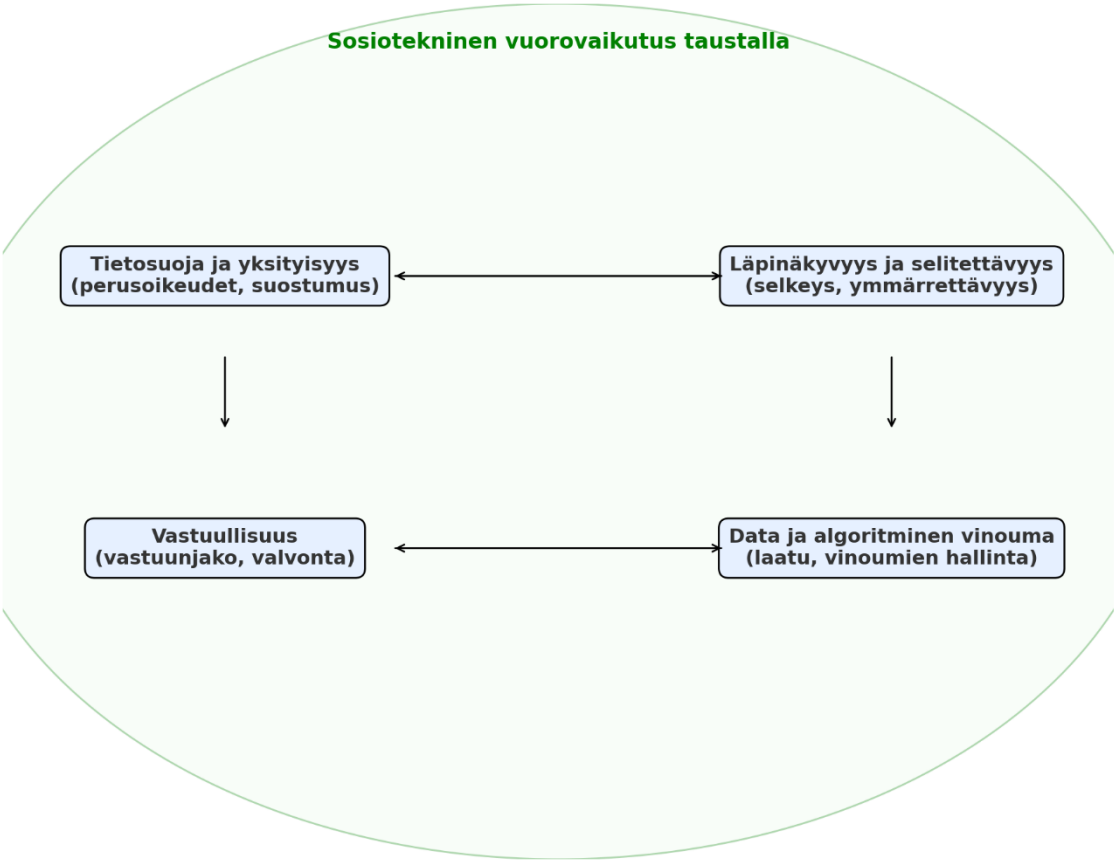
havainnollistaa, että eettiset ulottuvuudet eivät näyttäydy erillisinä ilmiöinä, vaan ne kietoutuvat toisiinsa ja muokkaavat samanaikaisesti työn rakenteita, käytäntöjä ja vuorovaikutussuhteita. Lisäksi se konkretisoi, miten sosiotekninen näkökulma tarjoaa välineitä näiden ilmiöiden kokonaisvaltaiseen ymmärtämiseen.

Eettinen ulottuvuus	**Vaikutukset työhön ja käytäntöihin**	**Sosiotekninen tulkinta**
Tietosuoja ja yksityisyys	Edellyttää tietosuoja-arviointeja, teknisten ratkaisujen hallintaa ja selkeitä suostumuskäytäntöjä.	Tietosuoja ei ole vain tekninen kysymys, vaan vaatii organisaation roolien ja vastuiden uudelleenmäärittelyä. Teknologia toimii työn muokkaajana.
Läpinäkyvyys ja selitettävyys	Selittämättömät prosessit vaikeuttavat ammattilaisten työtä ja horjuttavat luottamusta tekoälyratkaisuihin.	Läpinäkyvyys on edellytys luottamukselle ja yhteiselle tiedonmuodostukselle. Selitettävyys mahdollistaa työntekijöiden ja teknologian yhteistoiminnan.
Vastuullisuus	Vastuunjaon epäselvyys lisää painetta ammattilaisille, joille jää lopullinen vastuu.	Teknologia muuttaa työnjaon logiikkaa ja tuo mukanaan uusia valtasuhteita. Vastuu syntyy vain, kun tekniset ja sosiaaliset rakenteet tukevat toisiaan.
Data ja algoritminen vinouma	Datan puutteet voivat vääristää päätöksiä ja lisätä työkuormaa.	Vinoumat heijastavat myös sosiaalisia rakenteita ja päätöksenteon logiikkaa. Data kytkeytyy aina kontekstiin ja muokkaa työn käytäntöjä sekä resurssien jakautumista.
Työn ja teknologian vuorovaikutus	Tekoäly helpottaa rutiineja, mutta lisää tarvetta tarkistamiseen ja koulutukseen. Vaikuttaa myös ammattilaisen ja potilaan suhteeseen.	Teknologia muokkaa työnjakoa ja vuorovaikutusta, edellyttäen jatkuvaa sopeutumista.

Taulukko 4. Eettisten ulottuvuuksien vaikutukset työhön ja käytäntöihin sekä niiden tulkinta sosioteknisen teorian näkökulmasta.

8 Johtopäätökset

Tämän tutkielman tavoitteena oli syventää ymmärrystä tekoälyn käyttöön liittyvistä eettisistä ulottuvuuksista sosiaali- ja terveydenhuollon kontekstissa sekä tarkastella, miten nämä ulottuvuudet kytkeytyvät työn ja teknologian vuorovaikutukseen sosioteknisestä näkökulmasta. Tutkielma tarkasteli aihetta dokumenttianalyysin ja teoriaohjaavan sisällönanalyysin avulla, hyödyntäen kansallisia ja kansainvälisiä ohjaus, sääntely- ja suositusdokumentteja viimeisten kuuden vuoden ajalta. Tavoitteena oli tunnistaa keskeiset eettiset teemat sekä jäsentää niiden merkitys tekoälyn käyttöönoton ja hyväksyttävyyden kannalta. Kuviossa 4 on esitetty tiivistetysti tutkielman keskeiset eettiset ulottuvuudet ja niiden suhteet sosiaali- ja terveydenhuollon kontekstissa.



Kuvio 4. Tekoälyn eettiset pääulottuvuudet sosiaali- ja terveydenhuollossa.

Tulosten valossa keskeiset eettiset ulottuvuudet jäsentyvät viiden teeman ympärille, joita ovat tietosuoja ja yksityisyys, läpinäkyvyys ja selitettävyys, vastuullisuus, data ja algoritminen vinouma sekä teknologian ja työn vuorovaikutus. Näiden perusteella tekoälyn eettinen hyväksyttävyys sosiaali- ja terveydenhuollossa rakentuu tasapainosta, jossa juridiset velvoitteet, tekniset ratkaisut ja eettiset periaatteet limittyvät käytännön soveltamiseen. Lisäksi analyysi toi esiin yksittäisiä aineistolähtöisiä havaintoja, kuten jännitteen läpinäkyvyyden ja tietosuojan välillä, jotka täydentävät pääluokkien tarkastelua.

Tulokset osoittavat, että eettiset kysymykset eivät ole pelkästään teknisiä tai sääntelyllisiä, vaan ne ilmenevät kaikilla organisaation ja työn tasoilla sekä niiden välisessä vuorovaikutuksessa. Tekoäly vaikuttaa työnjakoon, osaamisvaatimuksiin ja prosesseihin sekä tuo esiin datan ja algoritmien merkityksen työn konkreettisena toimintaympäristönä. Lisäksi tuloksista käy ilmi ristiriita tekoälyn hyötyjen ja työkuormituksen välillä. Vaikka tekoäly voi tehostaa toimintaa ja vapauttaa aikaa asiakastyöhön, se voi samanaikaisesti lisätä valvonta- ja tarkistusvastuuta sekä kasvattaa ammattilaisten kuormitusta, jos järjestelmät eivät ole saumattomasti integroituneet työhön. Tämä havainto on linjassa sosioteknisen teorian näkemyksen kanssa siitä, että teknologiset ratkaisut voivat yhtä aikaa helpottaa ja kuormittaa työn käytäntöjä riippuen niiden kontekstuaalisesta sovittamisesta (Berg, 1999). Orlikowski (1992) puolestaan korostaa, että teknologian vaikutukset rakentuvat aina vuorovaikutuksessa sosiaalisten rakenteiden ja työn käytäntöjen kanssa, mikä selittää havaittua jännitettä.

Eettiset ulottuvuudet toimivat sosioteknisen järjestelmän keskeisinä kytkentäpisteinä, joissa teknologiset, organisatoriset ja inhimilliset tekijät kohtaavat. Ne vaikuttavat tekoälyn käyttöönottoon sekä edistävien että rajoittavien mekanismien kautta. Edistäviä tekijöitä ovat esimerkiksi luottamuksen rakentaminen, datan laatu ja läpinäkyvyys, kun taas rajoittavia tekijöitä ovat sääntelyn tulkintaerot, vinoumat ja vastuun hajautuminen. Näin ollen eettisyys ei näyttäydy irrallisena lisänä, vaan

rakenteellisena osana tekoälyn ja työn yhteismuodostusta sosiaali- ja terveydenhuollossa.

8.1 Eettiset ulottuvuudet

Ensimmäinen tutkimuskysymys tarkasteli, mitkä eettiset ulottuvuudet nousevat esiin sosiaali- ja terveydenhuollon tekoälyä koskevissa asiakirjoissa. Tulosten perusteella tunnistettiin viisi keskeistä eettistä ulottuvuutta, jotka jäsentävät tekoälyn hyväksyttävää käyttöä sosiaali- ja terveydenhuollossa. Nämä ovat tietosuoja ja yksityisyys, läpinäkyvyys ja selitettävyys, vastuullisuus, data ja algoritminen vinouma sekä teknologian ja työn vuorovaikutus. Tietosuoja ja yksityisyys muodostavat sääntelyllisen ja eettisen perustan tekoälyn hyväksyttävyydelle. Ne kytkeytyvät sekä lainsäädännöllisiin puitteisiin että teknisiin ratkaisuihin, ja niiden toteutuminen on keskeinen luottamuksen edellytys. Läpinäkyvyys ja selitettävyys toimivat vastavoimana tekoälyn mustan laatikon, eli päätöksenteon läpinäkymättömyyden ongelmille. Järjestelmien toimintaperiaatteiden ja päätöksenteon logiikan avaaminen on välttämätöntä sekä ammattilaisten että potilaiden näkökulmasta. Vastuullisuus rakentuu oikeudellisten velvoitteiden, eettisen harkinnan ja kollektiivisen vastuunjaon varaan, mutta siihen liittyy myös haasteita, kuten vastuun jakautumisen epäselvyys useiden toimijoiden välillä. Data ja algoritminen vinouma vaikuttavat suoraan tekoälyn oikeudenmukaisuuteen, tehokkuuteen ja luotettavuuteen. Datan laatu, monimuotoisuus, omistajuus ja saavutettavuus määrittävät järjestelmien toimintaa ja vinoumariskejä. Teknologian ja työn vuorovaikutus kuvaa tekoälyn käyttöönoton vaikutuksia työnjakoon, osaamisvaatimuksiin ja vuorovaikutussuhteisiin ja tuo esiin sen, että teknologian toimivuus rakentuu aina sosiaalisissa ja organisatorisissa rakenteissa.

Tulosten perusteella tekoälyn eettinen hyväksyttävyys sosiaali- ja terveydenhuollossa rakentuu tasapainosta, jossa juridiset velvoitteet, tekniset ratkaisut ja eettiset periaatteet kietoutuvat yhteen käytännön soveltamisessa. Yksityisyyden ja tietosuojan keskeinen rooli eettisen hyväksyttävyyden perustana korostuu tuloksissa.

Henkilötietojen käsittely perustuu laissa säädettyihin perusteisiin ja aidosti vapaaehtoiseen suostumukseen, mikä on linjassa aikaisemmassa tieteellisessä keskustelussa korostetun perusoikeuslähtöisyyden kanssa (Khosravi ja muut, 2024; Wang ja muut, 2023). Tulokset kuitenkin tarkentavat, että käytännössä suostumuksen vapaaehtoisuus ei aina toteudu, jos palvelun käyttö edellyttää datan luovuttamista. Tämä havainto tuo esiin jännitteen periaatteiden ja käytännön välillä: teoriassa yksilön oikeus hallita omia tietojaan on kiistaton, mutta operatiivisesti se horjuu erityisesti palveluissa, joissa vaihtoehdot ovat rajallisia.

Tulokset korostavat, että sääntelyn tulkintaerot synnyttävät epävarmuutta, joka voi estää uusien ratkaisujen käyttöönottoa. Tämä osoittaa, että sääntely ei toimi ainoastaan ennaltaehkäisevänä riskienhallinnan välineenä, kuten aikaisemmassa tieteellisessä keskustelussa on esitetty, vaan myös luottamuksen infrastruktuurina, joka mahdollistaa innovoinnin silloin, kun tulkinnat ja ohjeet ovat selkeitä ja ennakoitavia. Tekniset ratkaisut, kuten pseudonymisointi, anonymisointi ja differentiaalinen tietosuoja, jossa yksilön tunnistettavuutta vähennetään lisäämällä satunnaista vaihtelua suuriin aineistoihin, näyttäytyvät käytännön mekanismeina, joiden avulla teoriassa määritellyt periaatteet muuttuvat toimiviksi käytännöiksi. Näin tietosuoja muodostaa tekoälyn hyväksyttävän käytön perustan, jota ilman muut eettiset vaatimukset jäävät haavoittuviksi.

Läpinäkyvyyden ja selitettävyyden osalta aikaisempi tieteellinen keskustelu korostaa niiden merkitystä luottamuksen rakentajina (Al-Ansi ja muut, 2024), ja tulokset vahvistavat tämän näkemyksen. Käytännön tasolla läpinäkyvyys ilmenee datan alkuperän avaamisena, algoritmien toimintaperiaatteiden ja kehitysoletusten dokumentointina, prosessien jäljitettävyytenä sekä käyttäjäviestinnän selkeytenä. Tämä laajentaa aiemmin esitettyä painotusta teknisestä selitettävyydestä sosiaaliseen selitettävyyteen, jossa ymmärrettävä kieli ja saavutettavuus ovat yhtä tärkeitä kuin algoritmien tekninen avaaminen (Johnson ja muut, 2020).

Tulokset tuovat esiin aikaisemmassa tieteellisessä keskustelussa tunnistetun rakenteellisen jännitteen. Mitä enemmän mallien toimintaa ja dataa avataan, sitä suurempi on riski paljastaa yksityisyyteen liittyviä arkaluonteisia tietoja. Erityisesti syvien neuroverkkojen kohdalla, joita on kuvattu ihmisen hermoston toiminnasta inspiroituneiksi laskennallisiksi malleiksi (Salo, 2023), täydellinen selitettävyys ei ole realistista. Sen sijaan toiminnallinen läpinäkyvyys, joka tarkoittaa sitä, että käyttäjät tietävät milloin tekoälyä käytetään, millaisessa roolissa se toimii ja millaisia valitusmahdollisuuksia heillä on, on aina mahdollista toteuttaa. Tulokset vahvistavat aikaisemman tieteellisen keskustelun näkemyksen, jonka mukaan läpinäkyvyys ei ole pelkästään teknologinen ominaisuus vaan myös hallinnollinen ja viestinnällinen ratkaisu, jonka avulla voidaan vahvistaa luottamusta tekoälyyn (Al-Ansi ja muut, 2024; Johnson ja muut, 2020).

Vastuullisuuden osalta tulokset vahvistavat aikaisemmassa tieteellisessä keskustelussa esitetyn näkemyksen, jonka mukaan tekoäly ei voi korvata ammatillista harkintaa vaan sen tulee toimia päätöksenteon tukena selkein vastuusuhtein. Larsio (2023) näkee, että juuri ammattilaisen roolin säilyttäminen päätöksenteossa on keskeinen edellytys hyväksyttävälle tekoälyn käytölle. Tulokset tukevat tätä näkemystä ja tuovat esiin niin kutsutun useiden käsien ongelman, jossa monitoimijaiset tekoälyprosessit hämärtävät vastuun kohdentamista. Tämä havainnollistaa sitä, mitä Valta (2013) on korostanut jäljitettävyyden ja auditointien merkityksenä. Ilman selkeitä prosesseja eettinen vastuunkanto jää helposti abstraktiksi.

Tulokset konkretisoivat myös inhimillisen valvonnan vaatimuksen hallinnolliseksi ja tekniseksi kysymykseksi. Ammattilaisen hyväksyntä- ja tarkistusvastuu ei toteudu yksinomaan sääntelyllä, vaan se edellyttää resursointia ja työn organisointia. Lisäksi tulokset osoittavat, että vastuullisuus ei ole pelkkä rajoite, vaan se voi toimia myös kehityksen mahdollistajana. Tämä on linjassa Larsion (2023) esittämän näkemyksen kanssa, jonka mukaan selkeät vastuut, valvonta ja riskienhallinta eivät ainoastaan rajoita tekoälyn käyttöä, vaan vahvistavat sen hyväksyttävyyttä ja tukevat uusien ratkaisujen

käyttöönottoa. Valta (2013) puolestaan korostaa, että eettiset reunaehdot voivat toimia innovaatioita suuntaavina rakenteina, jotka tukevat kestävää kehitystä teknologian käyttöönotossa.

Datan ja algoritmisen vinouman osalta aikaisempi tieteellinen keskustelu painottaa, että laadukas ja edustava data on oikeudenmukaisuuden ja luotettavuuden edellytys, kun taas vinoutunut aineisto voi johtaa syrjiviin ja vaarallisiin päätöksiin. Kumar ja muut (2023) korostavat, että datan edustavuus on ratkaisevaa, jotta tekoälyjärjestelmien päätökset olisivat eettisesti kestäviä. Tulokset vahvistavat tämän näkemyksen ja konkretisoivat sen käytännön ilmentymiä. Ne osoittavat, että vinoumien ehkäisy ei ole kertaluontoinen toimi vaan jatkuvaa laadunvalvontaa, suorituskykytestausta ja ryhmäkohtaista arviointia edellyttävä prosessi.

Tulokset nostavat esiin myös datan yhteentoimivuuden ja saatavuuden eettisen ulottuvuuden. Aikaisemmassa keskustelussa nämä on kuvattu usein teknisinä vaatimuksina, mutta tulosten perusteella ne muodostuvat oikeudenmukaisuuden edellytyksiksi. Ilman yhteentoimivaa dataa on vaarana, että tekoälyn tuottamat suositukset kasaavat virheitä niihin ryhmiin, jotka ovat muutenkin aliedustettuja. Hoffman (2020) muistuttaa, että oikeudenmukaisuuden toteutuminen ei ole mahdollista ilman jatkuvaa valvontaa ja vinoumien arviointia kaikissa tekoälyn käytön vaiheissa. Tulokset tukevat tätä näkemystä ja tuovat esiin, että pelkät tekniset ratkaisut eivät riitä, vaan ilman hallinnollisia käytäntöjä ja systemaattista seurantaa ne voivat jopa vahvistaa vinoumia. Tulokset täydentävät aikaisempaa tieteellistä keskustelua osoittamalla, että oikeudenmukaisuuden vaatimus ei jää pelkäksi periaatteelliseksi lähtökohdaksi, vaan se konkretisoituu vasta silloin, kun vinoumien hallinta on systemaattisesti sisällytetty tekoälyn kehittämiseen, käyttöönottoon ja jatkuvaan arviointiin.

Tulokset vahvistavat aikaisemmassa tieteellisessä keskustelussa esitetyn näkemyksen, jonka mukaan tekoälyn eettinen hyväksyttävyys sosiaali- ja terveydenhuollossa rakentuu useiden toisiinsa kytkeytyvien tekijöiden kokonaisuudesta. Kuten Khosravi ja muut (2024)

sekä Wang ja muut (2023) korostavat, tietosuoja ja yksityisyys muodostavat perustan, jonka varaan muut eettiset periaatteet rakentuvat. Tulokset tukevat tätä näkemystä ja osoittavat, että ilman vahvaa tietosuojan ja yksityisyyden toteutumista muut eettiset vaatimukset jäävät haavoittuviksi.

Johnson ja muut (2020) sekä Al-Ansi ja muut (2024) näkevät, että läpinäkyvyys ja selitettävyys ovat keskeisiä tekijöitä luottamuksen rakentamisessa tekoälyyn. Tulokset tuovat tämän näkemyksen käytännön tasolle osoittamalla, että läpinäkyvyys ja selitettävyys eivät rajoitu teknisiin ratkaisuihin vaan ilmenevät myös hallinnollisina ja viestinnällisinä käytäntöinä, jotka tekevät tekoälyn toiminnasta ymmärrettävää niin ammattilaisille kuin käyttäjille. Larsio (2023) puolestaan korostaa, että tekoäly ei voi korvata ammatillista harkintaa, vaan sen tulee toimia päätöksenteon tukena selkein vastuusuhtein. Tulokset täydentävät keskustelua nostamalla esiin käytännön haasteen, niin kutsutun useiden käsien ongelman, jossa vastuun kohdentaminen monitoimijaisissa tekoälyprosesseissa hämärtyy. Tämä havainto vahvistaa Valtan (2013) esille tuoman ajatuksen jäljitettävyyden ja auditointien keskeisestä merkityksestä eettisen vastuunkannon varmistamisessa.

Luotettavuutta ei ole nostettu tässä tutkielmassa omaksi pääluokakseen, mutta se kietoutuu kaikkiin eettisiin teemoihin. Läpinäkyvyys, vastuullisuus ja datan hallinta muodostavat keskeiset edellytykset sille, että tekoälyratkaisuja voidaan pitää luotettavina. WHO:n (2021) tulokset korostavat, että luottamus ylittää teknisen selitettävyyden. Kyse ei ole vain yksittäisten teknisten vaatimusten, kuten läpinäkyvyyden tai tietosuojan, toteutumisesta, vaan kokonaisvaltaisesta hyväksyttävyydestä, joka perustuu eettiseen suunnitteluun, ihmiskeskeiseen toteutukseen ja julkiseen legitimiteettiin. Wang ja muut (2023) korostavat, että luottamus rakentuu yksilön oikeudesta hallita omia tietojaan. Johnson ja muut (2020) puolestaan painottavat, että selitettävyyden on oltava saavutettavaa ja ymmärrettävää, jotta se voi todella vahvistaa luottamusta. Tulokset täydentävät, että luotettavuus syntyy vasta, kun nämä tekijät yhdistyvät käytännön soveltamisessa.

Luottamus tekoälyyn näyttäytyy näin paitsi teknologisena myös sosiaalisena ja institutionaalisena prosessina. Sen rakentaminen edellyttää laajaa osallistumista, eettistä sitoutumista ja toimivaa sääntely-ympäristöä, kuten myös Valta (2013) on todennut sosioteknisestä näkökulmasta. Erillisenä teemana luotettavuus olisi voinut korostaa tätä näkökulmaa entistä selkeämmin, mutta tässä tutkielmassa se ilmenee läpileikkaavasti muiden eettisten ulottuvuuksien kautta.

Eettiset periaatteet ja käytännön toteutus ovat erottamattomia. Ilman juridista ja organisatorista selkänojaa tekniset ratkaisut eivät kykene rakentamaan luottamusta, ja ilman teknisiä ja operatiivisia mekanismeja eettiset periaatteet jäävät pelkiksi kirjatuiksi tavoitteiksi vailla käytännön vaikutusta. Johnson ja muut (2020) korostavat, että luottamus edellyttää yhtä aikaa teknisiä, hallinnollisia ja sosiaalisia ratkaisuja, mikä näkyy myös tuloksissa. Voidaan todeta, että eettisyys ei ole erillinen lisä tekoälyn käyttöönotossa, vaan sen hyväksyttävyyden ja toimivuuden keskeinen edellytys sosiaali- ja terveydenhuollossa (Al-Ansi ja muut, 2024). Wang ja muut (2023) täydentävät tätä näkökulmaa korostamalla, että eettisten periaatteiden toteutuminen vaatii jatkuvaa seurantaa ja selkeitä toimintamalleja.

Tulokset osoittavat lisäksi, että taustalla kaikissa eettisissä ulottuvuuksissa vaikuttaa työn ja teknologian vuorovaikutus, jonka merkitys näkyy läpileikkaavasti sosioteknisen teorian mukaisesti. Orlikowski (1992) kuvaa tätä vuorovaikutusta teknologian ja sosiaalisten rakenteiden yhteismuodostuksena. Valta (2013) puolestaan tuo esiin, että teknologian käyttöönotto rakentuu organisaation, ihmisten, teknologian ja työtehtävien välisessä dynamiikassa.

8.2 Teknologian ja työn vuorovaikutus sekä sosiotekninen näkökulma

Tämän tutkielman toinen tutkimuskysymys tarkasteli, miten eettiset näkökulmat kytkeytyvät sosioteknisiin kysymyksiin. Kysymys suuntaa tarkastelun yksittäisiä toimialoja laajemmalle ja korostaa eettisten ja sosioteknisten ulottuvuuksien välistä

yleistä vuorovaikutusta. Tulokset vahvistavat teoriaosuudessa esiin nostetun sosioteknisen teorian keskeisen näkemyksen siitä, että teknologia ei ole neutraali väline, vaan sen merkitys rakentuu aina vuorovaikutuksessa sosiaalisten rakenteiden, työn käytäntöjen ja inhimillisen toiminnan kanssa (Orlikowski, 1992; Salovaara, 2024). Tämä merkitsee sitä, että myös tekoälyn eettiset ulottuvuudet saavat konkreettisen sisältönsä näissä sosioteknisissä risteyskohdissa, joissa teknologiset ratkaisut, organisatoriset rakenteet ja ammatilliset käytännöt kietoutuvat yhteen.

Tulokset osoittavat, että teknologian ja työn vuorovaikutus on keskeinen sosiotekninen kohtaamispiste, jossa eettiset kysymykset tulevat konkreettisiksi. Tekoäly vaikuttaa työn rakenteisiin kaksijakoisesti. Se voi helpottaa työtä vähentämällä rutiinitehtäviä ja tukemalla päätöksentekoa, mutta samalla se voi lisätä kuormitusta, jos järjestelmät ovat vaikeakäyttöisiä, tai jos valvonta- ja tarkistusvastuu siirtyy liiallisesti ammattilaisille. Tämä havainto tukee Walkerin ja muiden (2008) esittämää sosioteknisen teorian periaatetta, jonka mukaan järjestelmien suorituskyky ja vaikutukset riippuvat sosiaalisten ja teknisten elementtien yhteen kietoutumisesta eikä teknologiasta itsestään. Tekoäly voi esimerkiksi vapauttaa aikaa asiakastyöhön, mutta jos hyväksyntä- ja tarkistusvastuu jää epäselväksi, työmäärä voi jopa kasvaa. Näin ollen tekoälyn käyttöönotto on väistämättä myös eettinen ja organisatorinen kysymys, joka edellyttää jatkuvaa tasapainottelua hyötyjen ja kuormituksen välillä.

Käytettävyyden ja työn sujuvuuden merkitys nousevat aineistossa vahvasti esiin. Berg (1999) on todennut, että tietojärjestelmät eivät yksiselitteisesti tehosta tai hidasta toimintaa, vaan niiden vaikutukset määräytyvät aina kontekstin mukaan. Tulokset osoittavat, että jos järjestelmät on suunniteltu ja integroitu huolellisesti osaksi työprosesseja, ne voivat lisätä työn mielekkyyttä ja vähentää hallinnollista kuormaa. Kömpelöt ja heikosti sovitetut ratkaisut sen sijaan estävät järjestelmien omaksumista ja voivat jopa lisätä epävarmuutta ja kuormitusta. Nämä havainnot tukevat Valtan (2013) esittämää näkemystä siitä, että onnistunut teknologian käyttöönotto edellyttää

sosiaalisen ja teknisen järjestelmän rinnakkaista kehittämistä sekä tiivistä vuoropuhelua käyttäjien ja kehittäjien välillä.

Sosiotekninen näkökulma näkyy myös moniammatillisen yhteistyön ja tiedon liikkumisen kysymyksissä. Hajanaiset tietojärjestelmät ja standardoinnin puute voivat hidastaa yhteistyötä ja luoda organisatorisia siiloja, kun taas avoimiin rajapintoihin ja yhteisiin tietomalleihin perustuvat ratkaisut sujuvoittavat tiedonvaihtoa. Tämä vahvistaa sosioteknisen lähestymistavan ajatusta avoimesta järjestelmästä, jossa teknologian vaikutukset ovat monisuuntaisia ja osin ennakoimattomia (Walker ja muut, 2008). Tekoäly ei siis ainoastaan tue olemassa olevia käytäntöjä, vaan myös muokkaa työnjaon logiikkaa, vastuuketjuja sekä ammattilaisen ja asiakkaan välistä vuorovaikutusta.

Tulokset tuovat esiin myös sen, että tekoälyn käyttöönotto muuttaa väistämättä ihmisten ja teknologian välisiä vuorovaikutussuhteita. Suora asiointi tekoälyjärjestelmien kanssa voi lisätä saavutettavuutta ja sujuvoittaa palveluja, mutta samalla se herättää kysymyksiä valvonnasta, luottamuksesta ja työn sisällön muutoksesta. Tämä vastaa Orlikowskin (1992) näkemystä siitä, että teknologia ei ainoastaan rakenna uusia työprosesseja, vaan myös muovaa niitä toimintavaihtoehtoja, joita ihmiset pitävät realistisina ja perusteltuina. Tulokset osoittavat, että tekoäly voi ohjata päätöksentekoa hienovaraisesti ja tuoda mukanaan uusia vastuusuhteita, jotka eivät ole yksinomaan teknisiä, vaan myös sosiaalisia ja eettisiä.

Useat aineiston dokumentit tunnistavat jännitteen tietosuojan ja läpinäkyvyyden välillä, mikä on myös aiemmassa kirjallisuudessa keskeinen eettinen ja sosiotekninen kysymys (Al-Ansi ja muut, 2024 ; Wang ja muut, 2023). Tietosuoja perustuu yksilön oikeuteen hallita omia tietojaan ja siihen liittyviin oikeudellisiin velvoitteisiin, mutta samanaikaisesti tekoälyjärjestelmien luotettavuus ja yhteiskunnallinen hyväksyttävyys rakentuvat läpinäkyvyydelle ja selitettävyydelle. Potilaan tai käyttäjän oikeus tietää, mihin ja miten hänen tietojaan käytetään, on luottamuksen kannalta keskeistä (Wang ja muut, 2023), mutta tietojen avoin käsittely voi toisaalta lisätä yksityisyyden

vaarantumisen riskiä, mikä erityisesti terveysdatan kontekstissa voi olla merkittävää (Al-Ansi ja muut, 2024). Tämä jännite heijastaa sosioteknisen teorian väitettä siitä, että teknologiset ja sosiaaliset tavoitteet eivät ole aina ongelmattomasti yhteensovitettavissa, vaan niiden yhteys rakentuu käytännöissä, joissa eettiset, juridiset ja tekniset näkökohdat limittyvät (Orlikowski, 1992; Walker ja muut, 2008).

Tulosten perusteella tämä jännite korostuu erityisesti kansanterveydellisissä käyttötapauksissa, kuten sijaintitietojen hyödyntämisessä, joissa yksilön tietosuoja ja kollektiiviset terveyshyödyt voivat olla jännitteisessä suhteessa. Myös sääntelyn ja sen tulkinnan epäselvyydet, erityisesti GDPR:n osalta, voivat aiheuttaa joko varovaisuutta tekoälyn käytössä tai liian löyhästi tulkittuina heikentää yksilön oikeuksien suojaa (Strange & Tucker, 2024). Tämä havainto vastaa teoreettista käsitystä siitä, että normatiiviset ja institutionaaliset puitteet määrittävät teknologian käyttöönoton ehtoja ja vaikuttavat siihen, millaisia ratkaisuja organisaatiot ja ammattilaiset pitävät hyväksyttävinä (Berg, 1999). Valta (2013) puolestaan tuo esiin, että teknologian ja työn välinen dynamiikka vaikuttaa keskeisesti siihen, miten eettiset periaatteet ja sääntely konkretisoituvat käytännössä.

Laajamittaiset tietosuojaongelmat, kuten Iso-Britanniassa esiin nousseet tapaukset, joissa potilastietoja käsiteltiin ilman asianmukaista suostumusta ja valvontaa (Siala & Wang, 2022), konkretisoivat näiden jännitteiden riskipotentiaalin. Vaikka vastaavia tapauksia ei tarkastellussa aineistossa Suomessa tai kansainvälisesti systemaattisesti raportoitu, niiden esiintyminen laajassa kansainvälisessä keskustelussa korostaa, että eettiset ja juridiset puitteet ovat ratkaisevia tekoälyn hyväksyttävän käytön kannalta. Yksityisyys jäsentyy näin sekä juridisena että eettisenä oikeutena, joka liittyy yksilön autonomiaan ja päätösvaltaan omista tiedoistaan (Cohen ja muut, 2020). Tätä periaatetta konkretisoi niin sanottu omadata, joka tarkoittaa yksilön oikeutta hallita omaa dataansa ja päättää sen käytöstä (Ailisto ja muut, 2022). Näin aineiston ja kirjallisuuden näkökulmat limittyvät siten, että tietosuoja ja yksityisyys muodostavat keskeisen perustan, jonka päälle muut eettiset periaatteet, kuten selitettävyys,

oikeudenmukaisuus ja vastuullisuus, rakentuvat. Khosravi ja muut (2024) korostavat tietosuojan merkitystä luottamuksen ja hyväksyttävyyden edellytyksenä, Wang ja muut (2023) puolestaan painottavat läpinäkyvyyden ja selitettävyyden roolia tietosuojan rinnalla. Johnson ja muut (2020) tuovat esiin sen, että tietosuoja on myös keskeinen osa järjestelmien luotettavuuden rakentamista.

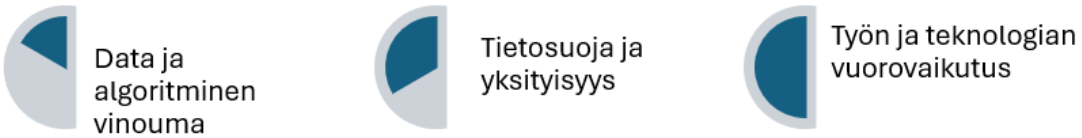
Eettiset näkökulmat kytkeytyvät tiiviisti sosioteknisiin kysymyksiin, sillä ne konkretisoituvat niissä risteyskohdissa, joissa teknologia, työn käytännöt ja organisatoriset rakenteet kohtaavat (Orlikowski, 1992; Walker ja muut, 2008). Tietosuoja ei ole pelkästään juridinen velvoite, vaan se ohjaa työn organisointia ja teknologisten ratkaisujen muotoa (Khosravi ja muut, 2024; Valta, 2013). Läpinäkyvyys ja selitettävyys vaikuttavat sekä teknologian toiminnallisuuteen että vuorovaikutukseen ja luottamuksen rakentumiseen käyttäjien ja ammattilaisten välillä (Johnson ja muut, 2020; Al-Ansi ja muut, 2024). Vastuullisuus määrittyy työnjaon ja ammatillisen harkinnan kautta, ja epäselvät vastuusuhteet voivat johtaa useiden käsien ongelmaan, jossa vastuu tekoälyratkaisujen päätöksistä ja vaikutuksista hämärtyy (Larsio, 2023; Valta, 2013). Algoritmisen vinouman hallinta on sekä tekninen että sosiaalinen prosessi, joka edellyttää jatkuvaa valvontaa ja kontekstin huomioimista. Tämä on linjassa sekä aineistosta nousseiden havaintojen että kirjallisuuden painotusten kanssa oikeudenmukaisuuden ja luotettavuuden edellytyksistä. Kumar ja muut (2023) korostavat datan ja algoritmien edustavuuden merkitystä eettisesti kestävien ratkaisujen varmistamisessa. Hoffman (2020) puolestaan painottaa jatkuvan valvonnan ja arvioinnin roolia vinoumien hallinnassa. Näin eettisyys näyttäytyy sosioteknisen kokonaisuuden sisäisenä ominaisuutena, ei erillisenä lisänä tai ulkoisena vaatimuksena.

On myös huomionarvoista, että teknologian ja työn vuorovaikutus esiintyy aineistossa eri tavoin eri konteksteissa. Kotimaisissa julkaisuissa korostuvat työn arkeen, ammattilaisten rooleihin ja vastuunjakoon liittyvät sosiotekniset jännitteet, kun taas kansainvälisissä lähteissä painottuvat rakenteelliset ja institutionaaliset kysymykset, kuten sääntelykehykset ja hallinnolliset järjestelyt (Al-Ansi ja muut, 2024; Johnson ja

muut, 2020; Wang ja muut, 2023). Tämä ero heijastaa sosioteknisen teorian näkemystä siitä, että periaatteet teknologian ja sosiaalisten järjestelmien yhteen kietoutumisesta ovat yleisiä, mutta niiden ilmenemismuodot vaihtelevat kontekstin, infrastruktuurin ja normatiivisten rakenteiden mukaan (Berg, 1999; Walker ja muut, 2008). Tekoälyn käyttöönotto nostaa esiin yhteisiä eettisiä ja organisatorisia kysymyksiä, mutta niiden painotukset ja ratkaisut muotoutuvat ympäristön erityispiirteiden mukaan.

Kuvioissa 5 ja 6 havainnollistetaan näitä kontekstuaalisia eroja vertaamalla pääluokkien suhteellisia painotuksia kansainvälisissä ja kansallisissa aineistoissa. Kuvioista käy ilmi erityisesti se, että teknologian ja työn vuorovaikutukseen liittyvät kysymykset painottuvat kotimaisissa lähteissä selvästi enemmän kuin kansainvälisissä, joissa rakenteelliset ja institutionaaliset näkökulmat korostuvat. Tietosuoja ja yksityisyys on sisällytetty molempiin kuvioihin osoittamaan, että tämän teeman esiintyvyydessä ei havaittu merkittäviä eroja. Myös vastuullisuus sekä läpinäkyvyyden ja selitettävyyden kysymykset nousivat vahvasti esille molemmissa konteksteissa, vaikka niitä ei ole tässä vertailussa havainnollistettu erillisillä kuvioilla. Vaikka painotukset vaihtelevat eri kontekstien välillä, niin sosiotekniset ja eettiset kysymykset läpäisevät aineiston eri tasoilla ja muodostavat yhteisen viitepohjan tekoälyn hyväksyttävyyden tarkastelulle.

Kuvio 5. Pääluokkien suhteellinen painotus kansainvälisissä tuloksissa.



Kuvio 6. Pääluokkien suhteellinen painotus kansallisissa tuloksissa.

Tutkielman keskeinen johtopäätös on, että tekoälyn hyväksyttävä ja kestävä käyttö rakentuu sosioteknisestä kokonaisuudesta, jossa eettiset periaatteet, teknologiset ratkaisut ja sosiaaliset sekä organisatoriset käytännöt limittyvät toisiinsa (Orlikowski, 1992; Valta, 2013; Walker ja muut, 2008). Tietosuoja, läpinäkyvyys, vastuullisuus ja algoritmisen vinouman hallinta eivät toimi irrallisina eettisinä vaatimuksina, vaan ne ovat kiinteä osa sosioteknistä järjestelmää, joka määrittää tekoälyn vaikutukset ja hyväksyttävyyden eri ympäristöissä. Vasta näiden ulottuvuuksien yhteensovittaminen mahdollistaa tekoälyn, joka tukee työn laatua, vahvistaa luottamusta ja edistää kestävää kehitystä sosiaali- ja terveydenhuollossa sekä laajemmin yhteiskunnassa.

Kuvioon 4 palaten voidaan todeta, että eettiset ulottuvuudet muodostavat kokonaisuuden, joka yhdistää juridiset, tekniset ja organisatoriset näkökulmat tekoälyn käyttöönotossa. Kuvio toimii hyödyllisenä jäsennyksenä myös tulosten synteesissä, sillä se osoittaa, miten eri periaatteet ja käytännöt limittyvät hyväksyttävän tekoälyn käytön perustaksi sosiaali- ja terveydenhuollossa.

Tulosten perusteella keskeisiä kehittämiskohteita ovat tietosuojan ja läpinäkyvyyden vahvistaminen, vastuunjaon selkeyttäminen monitoimijaympäristöissä sekä datan laadun ja monimuotoisuuden systemaattinen parantaminen. Näiden toteuttaminen edellyttää viranomaisten, organisaatioiden ja teknologian kehittäjien tiivistä yhteistyötä sekä selkeitä kansallisia linjauksia ja yhteisiä toimintamalleja. Lisäksi eettiset periaatteet tulee kytkeä konkreettisesti teknologian suunnittelu- ja käyttöönottoprosesseihin, jotta tekoälyn hyväksyttävä ja kestävä käyttö sosiaali- ja terveydenhuollossa voidaan varmistaa.

8.3 Tutkimuksen luotettavuus

Tämän tutkielman luotettavuutta ja rajoitteita on tärkeää tarkastella osana kokonaisarviointia. Aineisto koostuu asiakirjoista ja raporteista, mikä rajaa näkökulmaa ja jättää ulkopuolelle esimerkiksi ammattilaisten tai potilaiden kokemuksellisen tiedon.

Näin ollen tuloksia ei voida yleistää suoraan kaikkiin käytännön konteksteihin, vaan ne tarjoavat ennen kaikkea tulkintoja dokumentoiduista näkemyksistä ja ohjeistuksista. Laadullisen tutkimuksen luonteeseen kuuluu, ettei tutkimustuloksia voida pitää irrallisina tutkijasta tai aineiston tulkintatavasta; puhdasta objektiivista tietoa ei ole olemassa (Tuomi & Sarajärvi, 2018, s. 19). Tämä asettaa tutkijalle velvoitteen kuvata ja perustella tutkimuksellisia valintoja sekä tiedostaa oma roolinsa osana tiedonmuodostusta (Juuti & Puusa, 2020, s. 175).

Tutkimuksen uskottavuutta on vahvistettu useilla keinoilla. Aineisto on monipuolinen ja sisältää sekä kansallisia että kansainvälisiä dokumentteja, kuten ohjaus-, sääntely- ja suositusasiakirjoja. Aineiston valinta, analyysi ja tulosten raportointi on toteutettu läpinäkyvästi ja systemaattisesti. Analyysissa on esitetty runsaasti suoria esimerkkejä ja sivunumeroin varustettuja lainauksia, mikä tukee laadullisen tutkimuksen uskottavuutta (Aaltio & Puusa, 2020, s. 170–176). Analyysiprosessi eteni vaiheistetusti esiluvusta alustavaan koodaukseen, ryhmittelyyn ja teemojen kokoamiseen, ja jokainen havainto on jäljitettävissä alkuperäiseen dokumenttiin ja sivunumeroon. Prosessista on ylläpidetty aineistolokia, koodikirjan versiohistoriaa ja analyysimuistioita, mikä mahdollistaa auditoinnin jälkikäteen ja vahvistaa tutkimuksen läpinäkyvyyttä.

Tekoälyteknologian ja siihen liittyvien sovellusten nopea kehitys asettaa lisäksi erityisiä haasteita aineiston valinnalle ja analyysin ajankohtaisuudelle. Koska uutta tutkimusta, sääntelyä ja käytännön sovelluksia julkaistaan tiheästi, aineiston rajaus on jouduttu tekemään kriittisesti sen hetkisen saatavilla olevan tiedon perusteella. Tämä tarkoittaa, että osa keskeisistä dokumenteista voi vanhentua nopeasti tai jäädä lähinnä historiallisiksi kuvauksiksi teknologian kehityksestä ja sen soveltamisesta. Teknologian nopea kehitys muokkaa jatkuvasti myös eettisten kysymysten kenttää. Kysymykset, jotka aiemmin olivat marginaalisia, voivat nousta keskeisiksi, kun taas jotkin huolenaiheet voivat teknisten ratkaisujen tai institutionaalisen ohjauksen myötä menettää merkitystään. Tässä tutkielmassa tähän haasteeseen on vastattu valitsemalla aineisto

siten, että se ei ainoastaan kuvaa tekoälyn nykytilaa, vaan auttaa myös jäsentämään eettisiä ja sosiaalisia ilmiöitä, joiden merkitys säilyy teknologian kehittyessä.

Tutkijan oma esiymmärrys ja näkökulma ovat myös vaikuttaneet tutkimuksen kulkuun. Aihepiirin tuntemus, kiinnostus tekoälyn eettisiin kysymyksiin sekä aiempi perehtyneisyys sosiaali- ja terveydenhuollon toimintaympäristöön ovat ohjanneet niin aineiston valintaa kuin analyysin painotuksiakin. Laadulliselle tutkimukselle ominaisesti subjektiivisuus ei ole välttämättä heikkous, vaan se voi lisätä tutkimuksen uskottavuutta, kun tutkija tunnistaa ja tekee näkyväksi oman asemansa ja vaikutuksensa tutkimusprosessiin (Aaltio & Puusa, 2020, s. 170). Tässä työssä tämä on näkynyt erityisesti pyrkimyksenä systematisoida aineiston analyysi teoriaohjaavan sisällönanalyysin avulla, jotta henkilökohtainen esiymmärrys ei hallitsisi tulkintoja yksinään.

Tutkimuksen siirrettävyyttä on vahvistettu kuvaamalla tarkasti tutkimuksen konteksti ja rajaukset. Aineiston ajallinen rajaus, maantieteellinen painotus sekä dokumenttityyppien valinta on esitetty selkeästi. Näin lukija voi arvioida tulosten sovellettavuutta muihin sosiaali- ja terveydenhuollon ympäristöihin tai maihin. Tutkielmassa on noudatettu hyvän tieteellisen käytännön (HTK) periaatteita Tutkimuseettisen neuvottelukunnan (TENK) ohjeiden mukaisesti. Koko prosessin ajan työssä on pyritty luotettavuuteen, rehellisyyteen, avoimuuteen ja vastuullisuuteen (Tutkimuseettinen neuvottelukunta, 2023, s. 11–12). Aineiston valinta, analyysi ja tulosten raportointi on toteutettu läpinäkyvästi ja systemaattisesti, eikä tutkimus ole sisältänyt sellaista aineistonkeruuta, joka olisi edellyttänyt erillistä tutkimuslupaa tai eettistä ennakkoarviointia. Tutkielman luotettavuutta on vahvistettu kuvaamalla aineistonkeruun rajauksia, analyysimenetelmää sekä esittämällä aineistosta suoria esimerkkejä, mikä tukee laadullisen tutkimuksen uskottavuutta (Aaltio & Puusa, 2020, s. 170–176).

Tutkimusetiikka edellyttää tutkijalta myös tietoisuutta tutkimusaiheen vaikutuksista ja sen käsittelyyn liittyvistä riskeistä. Tutkimusaihe on eettisesti perusteltu, sillä tekoälyyn liittyvät ratkaisut voivat vaikuttaa yksilöiden oikeuksiin, turvallisuuteen ja tasa-arvoon, mikä tekee näiden kysymysten tarkastelusta yhteiskunnallisesti merkittävää. Tutkimus ei kohdistunut yksittäisiin henkilöihin, vaan asiakirjoihin, jotka ovat julkisesti saatavilla. Henkilötietoja ei kerätty, eikä tutkimus edellyttänyt tutkimuslupaa tai eettistä ennakkoarviointia. Dokumentteja on käsitelty alkuperäistä kontekstia ja alkuperäislähteitä kunnioittaen, ja kaikki aineisto on asianmukaisesti arkistoitu koko tutkimusprosessin ajaksi.

Tässä tutkielmassa on hyödynnetty OpenAI:n kehittämää GPT-4o -kielimallia ChatGPT-palvelun kautta Vaasan yliopiston ohjeistusta noudattaen. Tekoälyä on käytetty kuvioiden laatimisessa, sekä kirjoittajan oman tekstin jäsentämisen ja englanninkielisen aineiston käännösten tukena. Kaikesta sisällöstä ja sen lopullisesta muotoilusta vastaa tutkielman tekijä. Tekoälyn ehdotuksia on hyödynnetty harkinnanvaraisesti, ja lopulliset ratkaisut on tehty itsenäisesti.

8.4 Jatkotutkimusaiheet

Tekoälykehityksen nopea eteneminen tekee tutkimuskentästä hyvin dynaamisen, sillä uusia ilmiöitä, haasteita ja eettisiä kysymyksiä nousee esiin jatkuvasti, mikä luo runsaasti jatkotutkimuksen mahdollisuuksia. Tämä tutkielma keskittyi erityisesti tekoälyn eettisiin ulottuvuuksiin sosiaali- ja terveydenhuollossa, mutta tarkastelun ulkopuolelle jäivät useat teemat, joiden tutkiminen olisi perusteltua täydentämään kokonaiskuvaa.

Tässä työssä rajattiin pois asiakirjat, jotka käsittelevät ensisijaisesti sääntelyä ja oikeudellisia velvoitteita, vaikka sääntely muodostaa keskeisen kontekstin tekoälyratkaisujen kehittämiselle ja käyttöönotolle (Euroopan parlamentti ja neuvosto, 2024; Ojanperä, 2023, s. 158). Jatkossa olisi tärkeää tutkia tarkemmin, miten eettiset periaatteet ja oikeudelliset velvoitteet kytkeytyvät toisiinsa käytännön sovelluksissa.

EU:n tekoälyasetuksen ja kansallisen lainsäädännön tarkastelu voisi syventää ymmärrystä siitä, kuinka normatiiviset rakenteet tukevat tai rajoittavat tekoälyn vastuullista hallintaa sosiaali- ja terveydenhuollossa (Euroopan parlamentti ja neuvosto, 2024; Finlex, 1992/785; 2000/812; 2023/703). Vaikka tekoälyn eettisiä jännitteitä yksityisyyden, läpinäkyvyyden ja oikeudenmukaisuuden näkökulmista on tutkittu melko paljon (Cohen ja muut, 2020; Khosravi ja muut, 2024), tutkimusta eettisten periaatteiden ja sääntelyn käytännön yhteen kietoutumisesta on vielä vähän. Tämä tutkimusaukko korostuu erityisesti sosiaali- ja terveydenhuollon kontekstissa (European Commission, 2024; Sartor & Lagioia, 2020).

Toinen jatkotutkimusehdotus liittyy asenteisiin, pelkoihin ja ennakkoluuloihin tekoälyä kohtaan. Ammattilaisten, asiakkaiden ja kansalaisten suhtautuminen vaikuttaa ratkaisevasti siihen, kuinka hyväksyttävänä tekoälyn käyttö nähdään ja millaisia eettisiä jännitteitä sen käyttöönottoon liittyy. Näiden tutkiminen täydentäisi juridisia ja teknisiä tarkasteluja ja auttaisi tunnistamaan luottamusta vahvistavia tai vastarintaa synnyttäviä tekijöitä sosiaali- ja terveydenhuollossa. Luottamuksen merkitys on aiemmissa tutkimuksissa tunnistettu keskeiseksi käyttöönoton edellytykseksi (Ojanperä, 2023, s. 160; WHO, 2021), mutta suomalaiseen kontekstiin kohdistuva empiirinen tutkimus on edelleen niukkaa.

Lisäksi kansainvälinen vertailu olisi kiinnostava jatkotutkimusaihe, sillä se voisi tuoda esiin yhtäläisyyksiä ja eroja tekoälyn eettisissä kysymyksissä ja niiden hallinnassa eri maiden sosiaali- ja terveydenhuollossa. Tätä yhteyttä ei tässä tutkielmassa tarkasteltu, mutta aiemmat tutkimukset viittaavat siihen, että sääntely- ja hallintakehykset sekä resurssien ja teknologisen kyvykkyyden tasot vaihtelevat merkittävästi (Sartor & Lagioia, 2020; WHO, 2021). Tässä luvussa esitetyt jatkotutkimusaiheet vastaavat tunnistettuihin aukkoihin ja täydentäisivät olemassa olevaa tutkimuskenttää.

Lähteet

Aaltio, I., & Puusa, A. (2020). Laadullisen tutkimuksen luonne ja luotettavuus. Teoksessa A. Puusa & I. Juuti (toim.), Laadullisen tutkimuksen näkökulmat ja menetelmät (s. 167–188). Gaudeamus.

Ala-Korpela, M., Inkinen, S. & Suna, T. (2007). Kyborgin käsikirja. Havaintoja informaatiosta, ihmisestä ja koneesta, elämästä ja älykkyydestä. Otava.

Al-Ansi, Ab., Garad, A., Jacoob, M. & Al-Ansi, A. (2024). Elevating e-government: Unleashing the power of AI and IoT for enhanced public services. Heliyon, 10 (23). https://doi.org/10.1016/j.heliyon.2024.e40591

Alhosani, K. & Alhashmi, S. (2024). Opportunities, challenges, and benefits of AI innovation in government services: a review. Discover artificial intelligence, 4, 18 (2024). https://doi.org/10.1007/s44163-024-00111-w

Ailisto, H., Myllymäki, P., Tarkoma, S., Kämäräinen, J.-K., Röning, J., Salakoski, T., Solin, A., Saariluoma, P., Mikkonen, T., van Gils, M., Väänänen, K., Puolamäki, K., Ylén, P., Roos, T., Leikas, J., Honkela, A., Kutila, M., Ruotsalainen, L., Ylikoski, P., & Linturi, R. (2022). Tekoälyratkaisut tänään ja tulevaisuudessa (Eduskunnan tulevaisuusvaliokunnan julkaisu 1/2022). Eduskunta. https://www.parliament.fi/FI/naineduskuntatoimii/julkaisut/Documents/tuvj_1+2022.pdf

Anttila, P. (2000). Tutkimuksen taito ja tiedon hankinta: Taito- taide- ja muotoilualojen tutkimuksen työvälineet (3. painos). Akatiimi Oy.

Autioniemi, Jari. (2020). Tekoälyn yhteiskehittäminen julkisella sektorilla. Hallinnon Tutkimus 39(1), s. 5–20. https://doi.org/10.37450/ht.98075

Backlund, L. (Juont.). (2024). Jakso 7: Miten tekoäly muuttaa työelämää seuraavien vuosien aikana? [Podcast-jakso]. Backlund & AI – Tekoälyn tulevaisuus [Podcast]. Tekoälyn tulevaisuus. https://www.cgi.com/fi/fi/podcast/alykasta-tekoa

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Brynjolfsson, E., Li, D. & Ray, L. (2023). Generative AI at work. NBER. https://arxiv.org/pdf/2304.11771

Cohen, G., Evgeniou, T., Gerke, S. & Minssen, T. (2020). The European artificial intelligence strategy: implications and challenges for digital health. The Lancet digital health. https://doi.org/10.1016/S2589-7500(20)30112-6

DigiFinland. (2024a). Tekoäly hyvinvointialueilla: Sosiaali- ja terveydenhuollon käyttötapaukset ja lainsäädäntö. Selvitys ja suositukset yhteisistä tulkinnoista hyvinvointialueille, joulukuu 2024. Sosiaali- ja terveydenhuollon tekoälyn käyttötapaukset ja lainsäädäntö

DigiFinland. (2024b). Tekoälyn oikeudelliset reunaehdot sosiaali- ja terveydenhuollossa. DigiFinland Oy. https://digifinland.fi/wp-content/uploads/2024/12/Liite-2.-Tekoalyn-oikeudelliset-reunaehdot-sosiaali-ja-terveydenhuollossa.pdf

Eriksson, P., & Kovalainen, A. (2016). Qualitative research in business studies (2nd ed.). SAGE Publications.

Euroopan parlamentti ja neuvosto. (2024). Euroopan parlamentin ja neuvoston asetus (EU) 2024/1689 tekoälyä koskevista yhdenmukaistetuista säännöistä ja eräiden säädösten muuttamisesta (tekoälysäädös). Euroopan unionin virallinen lehti, L

202, 12.7.2024, 1–144. https://eur-lex.europa.eu/legal-content/FI/TXT/PDF/?uri=OJ:L_202401689

Euroopan unioni. (2017). Euroopan parlamentin ja neuvoston asetus (EU) 2017/745 lääkinnällisistä laitteista, adi-direktiiviä 2001/83/EY, asetuksia (EY) N:o 178/2002 ja (EC) N:o 1223/2009 muuttavista ja neuvoston direktiivien 90/385/ETY ja 93/42/ETY kumoavasta säädöksestä (Medical Devices Regulation). Euroopan unionin virallinen lehti, L 117, 1–175. https://eur-lex.europa.eu/legal-content/FI/TXT/PDF/?uri=CELEX:32017R0745

Euroopan unioni. (2016). Euroopan parlamentin ja neuvoston asetus (EU) 2016/679 luonnollisten henkilöiden suojelusta henkilötietojen käsittelyssä sekä näiden tietojen vapaasta liikkuvuudesta ja direktiivin 95/46/EY kumoamisesta (yleinen tietosuoja-asetus). Euroopan unionin virallinen lehti, L 119, 1–88. https://eur-lex.europa.eu/legal-content/FI/TXT/PDF/?uri=CELEX:32016R0679

European Commission. (2024). Artificial Intelligence – Questions and answers. Press release. European Union. https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1683

Ferede, D. (2025). Artificial Intelligence (AI) and Healthcare Capabilities: A Systematic Literature Review. F1000Research, 14:20. https://doi.org/10.12688/f1000research.158477.1

Gaskin, J. (2008). Whatever happened to artificial intelligence? Network World, 25, no 25, 2008, s. 30–32. https://www.proquest.com/trade-journals/whatever-happened-artificial-intelligence/docview/215995707/se-2?accountid=14797

Heinäsenaho, M., Äyräs-Blumberg, O. & Lähesmaa, J. (2023). Tekoäly mullistaa terveydenhuoltoa- mahdollisuudet hyödynnettävä viipymättä. Valtioneuvoston

kanslia. https://valtioneuvosto.fi/-//1271139/tekoaly-mullistaa-terveydenhuoltoa-mahdollisuudet-hyodynnettava-viipymatta

Hoffman, S. (2020). The Emerging Hazard of AI- Related health care discrimination. The Hasting center report. Vol. 51 (1), p.8-9. https://onlinelibrary-wiley-com.proxy.uwasa.fi/doi/full/10.1002/hast.1203

Holopainen, A., Puusa, A. & Juuti, P. (2020) Grouded theory: Aineistolähtöinen tutkimustapa. Teoksessa A. Puusa & P. Juuti (toim.). Laadullisen tutkimuksen näkökulmat ja perusteet. (s. 239–255). Gaudeamus.

Johnson, K., Wei, W., Weeraratne, D., Frisse, M., Misulis, K., Zhao, J & Snowdon, J. (2020). Cin Transl Sci 14, 86–93. https://doi.org/10.1111/cts.12884

Juuti, A., & Puusa, A. (2020). Laadullisen tutkimuksen toteuttaminen. Teoksessa A. Puusa & I. Juuti (toim.), Laadullisen tutkimuksen näkökulmat ja menetelmät (s. 173–192). Gaudeamus.

Khosravi, M., Zare, Z., Mojtabaeian, S. & Izadi, R. (2024). Ethical challenges of using articifial intelligence in healthcare delivery: a thematic analysis of a systematic review of reviews. Journal of Public Health. Advance online publication. https://doi.org/10.1007/s10389-024-02219-w

Kork, A.-A., Turja, T., Kurkela, K., Jones, M., & Jäntti, A. (2023). Informaatiolukutaito osallisuuden ja hyvinvoinnin politiikkana. *Politiikka*, 65(4), 269–297. https://doi.org/10.37452/politiikka.127226

Kumar, P., Chauhan, S. & Awasthi, L.K. (2023). Artificial intelligence in healthcare: Review, ethics, trust challenges & future research directions. Engineering applications of articifial intelligence. 120. https://doi.org/10.1016/j.engappai.2023.105894

Kääriäinen, J. (toim.), Aihkisalo, T., Halén, M., Holmström, H., Jurmu, P., Matinmikko, T., Seppälä, T., Tihinen, M., & Tirronen, J. (2018). Ohjelmistorobotiikka ja tekoäly – soveltamisen askelmerkkejä (Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 65/2018). Valtioneuvoston kanslia. https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161123/65-2018-Ohjelmistorobotiikka%20ja%20tekoaly.pdf?sequence=1&isAllowed=y

Laki sosiaali- ja terveydenhuollon asiakastietojen käsittelystä (2023). 30.6.2023/703. Suomen säädöskokoelma. Helsinki. http://www.finlex.fi

Laki sosiaalihuollon asiakkaan asemasta ja oikeuksista (2000). 22.9.2000/812. Suomen säädöskokoelma. Helsinki. http://www.finlex.fi

Laki potilaan asemasta ja oikeuksista (1992). 17.8.1992/785. Suomen säädöskokoelma. Helsinki. http://www.finlex.fi

Larsio, A. (2023). Datasta voimaa sote- järjestelmään. Sote- dataa hyödyntämällä parempaa hoitoa ja kustannussäästöjä. Sitra. https://www.sitra.fi/wp/wp-content/uploads/2023/05/datasta-voimaa-sote-jarjestelmaan.pdf

Leikas, J. (2022). Tekoälystä vastuullista. Teoksessa H. Ailisto, P. Myllymäki, S. Tarkoma, J.-K. Kämäräinen, J. Röning, T. Salakoski, A. Solin, P. Saariluoma, T. Mikkonen, M. van Gils, K. Väänänen, K. Puolamäki, P. Ylén, T. Roos, J. Leikas, A. Honkela, M. Kutila, L. Ruotsalainen, P. Ylikoski, & R. Linturi, *Tekoälyratkaisut tänään ja tulevaisuudessa* (s. 114-120). Eduskunta, Tulevaisuusvaliokunta (Julkaisu 1/2022). https://www.parliament.fi/FI/naineduskuntatoimii/julkaisut/Documents/tuvj_1+2022.pdf

Matheny, M., Israni, S. T., Ahmed, M. & Whicher, D. (2019) Artificial intelligence in health care. The hope, the hype, the promise, the peril. National Academy of medicine. https://ehealthresearch.no/files/documents/Rapporter/Andre/2019-12-AI-in-Health-Care.pdf

Moor, J. (2006). The Darthmouth College Artificial Intelligence Conference: The Next Fifty Years. AI Magazine, 27, no. 4, 2006, s. 87–91. https://doi.org/10.1609/aimag.v27i4.1911

Morgan, H. (2022). Conducting a qualitative document analysis. *The Qualitative Report, 27*(1), 64–77. https://doi.org/10.46743/2160-3715/2022.5044

Ojanen, A., Sahlgren, O., Vaiste, J., Björk, A., Mikkonen, J., Kimppa, K., Laitinen, A. & Oljakka, N. (2022). Algoritminen syrjintä ja yhdenvertaisuuden edistäminen. Arviointikehikko syrjimättömälle tekoälylle. (Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 2022:54). Valtioneuvoston kanslia. https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/164290/2022_VNTEAS_54.pdf?sequence=4&isAllowed=y

Ojanperä, T. (2023). Tekoälyn vallankumous. Käsikirja. Alma Talent.

O'Leary, Z. (2021). The essential guide to doing your research project (4. painos). SAGE Publications Ltd.

OpenAI. (2024). GPT-4o (13.5.2024) [laaja kielimalli]. https://chat.openai.com/chat

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. Organization Science, 3(3), 398–427. https://doi.org/10.1287/orsc.3.3.398

Perez, M. & Grande, R. (2020). Application of artifical intelligence in the diagnosis and treatment of hepatocellular carcinoma: A review. Baishideng publishing group, 26(37): 5617–5628. https://doi.org/10.3748/wjg.v26.i37.5617

Pesonen, R. (2021). Tekoäly. https://filosofia.fi/fi/ensyklopedia/tekoaly

Prior, L. (2008). Repositioning documents in social research. *Sociology, 42*(5), 821–836. https://doi.org/10.1177/0038038508094564

Pulkkinen, J. (2025). Tekoälyn hyödyntämisen haasteet ja ratkaisut julkisen sektorin sote-palveluissa: Selvitys 6.2.2025. WebFellows Oy. https://www.webfellows.fi/wp-content/uploads/2025/02/Sote-ai-haasteet-Pulkkinen.pdf

Roberts, H., Cowls, J., Morley, J., Taddeo, M., Wang, V. & Floridi, L. (2020). The chinese approach of artificial intelligence: an analysis of policy, ethics ans regulation. https://doi.org/10.1007/s00146-020-00992-2

Salo, I. (2023). Luova tekoäly mullistaa kaiken. ChatGPT näyttää tietä. Kauppakamari.

Salo, I. (2024). Luova tekoäly työn supervoimana. Kauppakamari.

Salovaara, S. (2024). Tietojärjestelmät osana sosiaalityön tiedonmuodostusta (Väitöskirja, Lapin yliopisto). Lapin yliopisto. https://urn.fi/URN:ISBN:978-952-337-436-2

Sartor, G., & Lagioia, F. (2020). The impact of the General Data Protection Regulation (GDPR) on artificial intelligence. European Parliamentary Research Service (EPRS), Scientific Foresight Unit (STOA). European Parliament.

https://op.europa.eu/en/publication-detail/-/publication/dc544697-19b8-11ec-b4fe-01aa75ed71a1/language-en

Schreier, M. (2024). *Qualitative content analysis in practice*. SAGE Publications Ltd. https://doi.org/10.4135/9781529682571

Siala, H. & Wang, Y. (2022). SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. Socian Science & Medicine 296. https://doi.org/10.1016/j.socscimed.2022.114782

Sosiaali- ja terveysministeriö. (2018). Eheä yhteiskunta ja kestävä hyvinvointi: Sosiaali- ja terveysministeriön tulevaisuuskatsaus. Valtioneuvoston julkaisusarja 22/2018. https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160904/22_TUKA_STM_WEB.pdf?sequence=4&isAllowed=y

Strange, M., & Tucker, J. (2024). Global governance and the normalization of artificial intelligence as 'good' for human health. *AI & Society, 39*, 2667–2676. https://doi.org/10.1007/s00146-023-01774-2

Toivonen, H. (2023). Mitä tekoäly on? Teos.

Trist, E. L. (1981). The evolution of socio-technical systems. *Ontario Quality of Working Life Centre*. https://www.lmmiller.com/blog/wp-content/uploads/2013/06/The-Evolution-of-Socio-Technical-Systems-Trist.pdf

Tutkimuseettinen neuvottelukunta. (2023). Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsitteleminen Suomessa. Tutkimuseettisen neuvottelukunnan julkaisuja 2/2023. https://tenk.fi/sites/default/files/2023-03/HTK-ohje_2023.pdf

Tyrväinen, P., Silvennoinen, M., Talvitie-Lamberg, K., Ala-Kitula, A., & Kuoremaki, R. (2018). Identifying opportunities for AI applications in healthcare - Renewing the national healthcare and social services. 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH), 1–8. https://doi.org/10.1109/SeGAH.2018.8401381

Työ- ja elinkeinoministeriö. (2019). Edelläkävijänä tekoälyaikaan: Tekoälyohjelman loppuraportti (Työ- ja elinkeinoministeriön julkaisuja 23/2019). Työ- ja elinkeinoministeriö. https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161447/23_19_Tekoalyraportti.pdf

Tuomi, J. & Sarajärvi, A. (2018). Laadullinen tutkimus ja sisällönanalyysi (uudistettu laitos). Kustannusosakeyhtiö Tammi.

Valta, M. (2013). *Sähköisen potilastietojärjestelmän sosiotekninen käyttöönotto: Seitsemän vuoden seurantatutkimus odotuksista omaksumiseen* [Väitöskirja, Itä-Suomen yliopisto]. Publications of the University of Eastern Finland. Dissertations in Social Sciences and Business Studies, 62. https://erepo.uef.fi/server/api/core/bitstreams/e45fb3a7-53c2-47e1-a849-c27afebfce0a/content

Walker, G. H., Stanton, N. A., Salmon, P. M., & Jenkins, D. P. (2008). A review of sociotechnical systems theory: A classic concept for new command and control paradigms. *Theoretical Issues in Ergonomics Science, 9*(6), 479–499. https://doi.org/10.1080/14639220701635470

Wang, C., Liu, S., Yang, H., Guo, J., Wu, Y & Liu, J. (2023). Ethical considerations of using chatGPT in health care. Journal of Medicial internet research. 25, e48009. https://doi.org/10.2196/48009

WHO. (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. World Health Organization. https://iris.who.int/bitstream/handle/1065/341996/9789240029200-eng.pdf

Liitteet

Liite 1. Tutkimusaineistona käytetyt dokumentit.

Dokumentti ja lähde	Julkaisija	Vuosi	Perustelu aineistoon valinnalle
Sote-AI-haasteet. Saatavilla: https://www.webfellows.fi/wp-content/uploads/2025/02/Sote-ai-haasteet-Pulkkinen.pdf	WebFellows Oy / STM	2025	Kotimainen katsaus, joka analysoi tekoälyn käyttöönoton keskeisiä haasteita ja nostaa esiin sääntelyn ja käytännön välisiä kuiluja.
Sosiaali- ja terveydenhuollon tekoälyn käyttötapaukset ja lainsäädäntö. Saatavilla: https://digifinland.fi/wp-content/uploads/2024/12/DigiFinland_Sosiaali-ja-terveydenhuollon-tekoalyn-kayttotapaukset-ja-lainsaadanto_04122024.pdf	DigiFinland	2024	Ajantasainen asiantuntijaraportti, joka kuvaa konkreettisia tekoälyn käyttötapauksia sote-sektorilla sekä sääntelyn vaatimuksia.
Tekoälyn oikeudelliset reunaehdot sosiaali- ja terveydenhuollossa. Saatavilla: https://digifinland.fi/wp-content/uploads/2024/12/Liite-2.-Tekoalyn-oikeudelliset-	DigiFinland Oy	2024	Oikeudellinen analyysi, joka tarkastelee tekoälyn käyttöä sote-sektorilla tietosuojan, tekoälysäädöksen ja hallinnon näkökulmista.

reunaehdot-sosiaali-ja-terveydenhuollossa.pdf			
Datasta voimaa sote-järjestelmään. Saatavilla: https://www.sitra.fi/wp/wp-content/uploads/2023/05/datasta-voimaa-sote-jarjestelmaan.pdf	Sitra	2023	Politiikkaraportti, joka esittää tietopolitiikan ja tekoälyn kehittämisen ratkaisuja sekä datan hallinnan ja hyödyntämisen näkökulmia.
Tekoälyratkaisut tänään ja tulevaisuudessa. Saatavilla: https://www.eduskunta.fi/FI/naineduskuntatoimii/julkaisut/Documents/tuvj_1%2B2022.pdf	Eduskunnan tulevaisuusvaliokunta	2022	Valiokuntajulkaisu, joka tarkastelee tekoälyn tulevaisuusnäkymiä, eettisiä kysymyksiä ja yhteiskunnallisia vaikutuksia.
EU AI Act: Komission Q&A. Saatavilla: https://ec.europa.eu/commission/presscorner/api/files/document/print/en/qanda_21_1683/QANDA_21_1683_EN.pdf	Euroopan komissio	2021	Selkeä yhteenveto EU:n tekoälysäädöksestä kysymys–vastaus -muodossa; tiivistää keskeiset sääntelysisällöt.
Ethics & Governance of Artificial intelligence for health. Saatavilla:	World health organization (WHO)	2021	Kansainvälinen eettinen ohjeistus, joka esittää suosituksia tekoälyn

https://iris.who.int/bitstream/handle/10665/341996/9789240029200-eng.pdf			vastuulliseen käyttöön terveydenhuollossa.
Edelläkävijänä tekoälyaikaan. Tekoälyohjelman loppuraportti. Saatavilla: https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161447/23_19_Tekoalyraportti.pdf	Työ- ja elinkeinoministeriö (TEM)	2019	Hallinnollinen selvitys, joka raportoi tekoälyn mahdollisuuksista ja riskeistä sosiaali- ja terveysalalla sekä antaa alan suosituksia.