LAPIN YLIOPISTO
UNIVERSITY OF LAPLAND

Tekoälyn eettisyys lastensuojelussa – kuvaileva kirjallisuuskatsaus

Emilia Alamäki, Y2402626
Pro gradu –tutkielma
Sosiaalityö
Lapin yliopisto 2026

LAPIN YLIOPISTO
UNIVERSITY OF LAPLAND

Lapin yliopisto
Yhteiskuntatieteiden tiedekunta
Tekoälyn eettisyys lastensuojelussa – Kuvaileva kirjallisuuskatsaus
Tekijä: Emilia Alamäki
Koulutusohjelma/oppiaine: Sosiaalityö
Ohjaaja: Professori Elina Virokannas
Työn laji: Pro gradu -tutkielma
Sivumäärä: 56, 3 liitettä
Vuosi: 2026

Tiivistelmä:

Tutkielmassa tarkastellaan tekoälyn käyttöön lastensuojelussa liittyviä eettisiä kysymyksiä. Se on toteutettu kuvailevana kirjallisuuskatsauksena. Tutkielman aineisto koostuu 12 vertaisarvioidusta tutkimusartikkelista vuosilta 2022–2026. Tutkimusartikkelit sisältävät viisi empiiristä alkuperäistutkimusta, kaksi teoreettista artikkelia sekä viisi kirjallisuuskatsausta.

Analyysissa tunnistetaan viisi keskeistä tekoälyn käyttöön lastensuojelussa liittyvää eettistä kysymystä. Ensimmäinen kysymys muodostuu epäoikeudenmukaisuuden ja eriarvoistumisen vahvistumisesta, sillä tekoälyjärjestelmät voivat sisältää ja vahvistaa olemassa olevia yhteiskunnallisia ja rakenteellisia vinoumia. Toiseksi tekoälyn käyttöön kytkeytyy riski lapsen ja perheen tilanteen virheellisestä tulkinnasta, koska algoritmit eivät kykene huomioimaan kaikkia tilanteisiin liittyviä yksilöllisiä ja kontekstuaalisia tekijöitä. Kolmanneksi tekoälyn käyttö voi kaventaa sosiaalityöntekijän ammatillista harkintaa ja hämärtää päätöksentekoon kohdistuvaa vastuuta. Neljänneksi tekoälyn kyky tunnistaa tuen tarpeita ei itsessään takaa lasten ja perheiden tarpeisiin vastaamista, vaan tuen toteutuminen edellyttää toimivaa palvelujärjestelmää ja riittäviä resursseja. Viidenneksi tekoälyn hyödyntämiseen sisältyy merkittäviä yksityisyyden suojaan, tietoturvaan ja valvontaan liittyviä kysymyksiä.

Eettisten kysymysten lisäksi tutkielmassa tarkastellaan keinoja vastata tunnistettuihin haasteisiin. Tutkielma osoittaa, että eettisesti kestävät toimintatavat edellyttävät teknisten ratkaisujen lisäksi ammatillisen harkinnan ja päätösvallan, oikeudenmukaisuuden, tietosuojan sekä lapsen oikeuksien huomioimista. Tekoälyn hyödyllisyyttä ei siten voida arvioida ainoastaan sen ennustetarkkuuden perusteella, vaan myös sen mukaan, miten se tukee sosiaalityöntekijöiden työtä ja lasten ja perheiden tarpeisiin vastaamista. Eettisesti kestävän tekoälyn lähtökohtana on ihmisen säilyminen päätöksenteon keskiössä sekä tekoälyn sovittaminen lastensuojelun tavoitteisiin ja sosiaalityön eettisiin periaatteisiin.

Avainsanat: eettisyys, lastensuojelu, tekoäly

Sisällys

1 JOHDANTO 4

2 LASTENSUOJELU JA TEKOÄLY 7

2.1 Lastensuojelun lähtökohdat 7

2.2 Tekoäly tiedon tuottamisen ja päätöksenteon välineenä 10

2.3 Teknologisoituva lastensuojelu 11

3. TUTKIMUKSEN TOTEUTUS 14

3.1 Tutkimustehtävä 14

3.2 Kuvaileva kirjallisuuskatsaus 15

3.3 Tutkimusaineiston haku ja valinta 16

3.4 Induktiivinen sisältöanalyysi 18

3.5 Tutkimuksen eettiset lähtökohdat 20

4. TEKOÄLYN EETTISET ULOTTUVUUDET LASTENSUOJELUSSA 22

4.1Aineiston kuvaus 22

4.2 Epäoikeudenmukaisuus ja eriarvoisuuden vahvistuminen 24

4.3 Lapsen ja perheen tilanteen virheellinen tulkinta 26

4.4 Ammatillisen harkinnan kaventuminen ja vastuun hämärtyminen 29

4.5 Lapsen tarpeisiin vastaamisen ja tuen toteutumisen vaarantuminen 33

4.6 Yksityisyyden suoja, tietoturva ja valvonta 35

4.7 Eettisesti kestävän tekoälyn edellytykset 37

5. JOHTOPÄÄTÖKSET 41

5.1Analyysista johtopäätöksiin 41

5.2 Tutkimusprosessin tarkastelu 46

LÄHTEET 50

LIITTEET 57

Liite 1. Aineiston hakusanat 58

Liite 2. Aineiston valintaprosessi: tunnistaminen, seulonta ja valittu aineisto 58

Liite 3. Lopullinen aineisto 59

1 JOHDANTO

Informaatioteknologian ja tekoälyn käyttö on lisääntynyt merkittävästi viime vuosikymmeninä yhteiskunnan eri osa-alueilla, kuten sosiaali- ja terveydenhuollossa, poliisissa ja koulutuksessa (Brandão 2025). Myös sosiaalityössä on alettu hyödyntämään tekoälyyn perustuvia ratkaisuja, ja kehitys linkittyy lastensuojeluun (Gillingham 2016, 1044; Reamer 2023, 52). Sosiaalityö on tekoälyn kehityksessä erityisen merkityksellinen alue, sillä siinä käsitellään arkaluonteisia henkilötietoja ja tehdään ihmisten elämään vaikuttavia päätöksiä (Reamer 2023). Tekoälyllä tarkoitetaan laajaa joukkoa teknologisia menetelmiä, joista koneoppiminen on yksi keskeinen lähestymistapa. Koneoppimisessa järjestelmät tunnistavat aineistoista malleja ja rakenteita, joiden perusteella niitä voidaan hyödyntää esimerkiksi ennusteiden ja luokittelujen tuottamisessa (Gillingham, 2016, 1044–1045; Kivipelto 2025, 3; Sengar ym. 2024, 23662). Tekoälyn hyödyntämistä sosiaalityössä on perusteltu muun muassa toiminnan tehokkuuden, vaikuttavuuden, yhdenmukaisuuden ja päätöksenteon systemaattisuuden lisäämisellä. Terveyden ja hyvinvoinnin laitoksen (THL) mukaan sosiaalihuollossa tekoälyn tämänhetkisiä hyödyntämiskohteita ovat esimerkiksi lausuntojen ja päätösten luonnostelu. Tulevaisuudessa tekoälylle asetetaan odotuksia myös esimerkiksi palvelutarpeen arvioinnissa ja ennakoivassa analytiikassa. (THL 2026a; THL 2026b.)

Tekoälyn käyttöönotto sosiaalityössä ei ole ainoastaan teknologinen tai toiminnallinen kysymys, vaan siihen liittyy olennaisesti kysymys sosiaalityön perustehtävästä, arvoista ja eettisistä lähtökohdista (Livingston & Risse 2019; Raso ym. 2018, 1–5). Sosiaalityötä ohjaavat ihmisoikeudet, ihmisarvon kunnioittaminen ja sosiaalisen oikeudenmukaisuuden edistäminen (International Federation of Social Workers ym., 1994). Teknologisen kehityksen myötä on tärkeää tarkastella, millä tavoin tekoälyn käyttö asettuu suhteeseen näiden periaatteiden kanssa. Vaikka tekoälyn hyödyntämiseen voi liittyä merkittäviä hyötyjä, sen käyttö sosiaalityössä herättää moniulotteisia eettisiä kysymyksiä (Reamer 2023, 53).

Tekoälyn käyttö voi muuttaa tapaa, jolla asiakkaiden tilanteita kuvataan, arvioidaan ja tulkitaan. Algoritmisessa päätöksenteossa asiakkaasta saatavilla oleva tieto voidaan muuntaa erilaisiksi muuttujiksi, luokituksiksi ja ennusteiksi. Tällöin vaarana on, että asiakkaan elämäntilanteen monimuotoisuus ja siihen liittyvät sosiaaliset, kulttuuriset ja yksilölliset tekijät jäävät vähäisemmälle huomiolle. Teknologisen tehokkuuden ja sosiaalityön ihmiskeskeisen

toimintatavan välille voi näin muodostua jännite. (Gillingham 2016, 1044; Landau ym. 2022, 580; Agbana 2025, 1648.) Äärimmillään ihmisen yksilöllisyyden ja inhimillisten ominaisuuksien sivuuttaminen voi kytkeytyä dehumanisaation riskiin. Dehumanisaatiolla tarkoitetaan prosessia, jossa yksilöiltä tai ihmisryhmiltä kielletään heidän ihmisarvoansa, inhimillisyytensä tai kyky tuntea (Haslam 2006, 252).

Lastensuojelussa aiemmin mainittujen kysymysten tarkastelu on erityisen merkityksellistä. Lastensuojelulle on ominaista julkisen vallan käyttö sekä interventiot, joilla voi olla voimakkaita ja pitkäkestoisia vaikutuksia lasten ja perheiden elämään (Parton & Kirk 2009, 23–26). Lastensuojelun päätöksenteossa arvioidaan siten tilanteita, joissa ammatillisella harkinnalla ja käytettävällä tiedolla voi olla huomattavia vaikutuksia lapsen ja perheen oikeuksiin, turvallisuuteen ja elämäntilanteeseen. Tekoälyn hyödyntäminen tällaisissa prosesseissa nostaa esiin kysymyksiä esimerkiksi päätöksenteon oikeudenmukaisuudesta, tietoaineistojen vinoumista, ennusteiden luotettavuudesta, läpinäkyvyydestä ja ammatillisen vastuun säilymisestä (Rodriguez ym. 2019; Landau ym. 2022). Lastensuojelun kontekstissa tekoälyn käyttöä onkin tarpeen tarkastella erityisesti suhteessa siihen, millä tavoin teknologiset ratkaisut vaikuttavat lasten ja perheiden tilanteiden arviointiin sekä heidän kohtaamiseensa palvelujärjestelmässä.

Tekoälyn käyttöön liittyy myös kysymys ammatillisen harkinnan asemasta. Sosiaalityössä asiakkaan tilanteen arviointi ei perustu ainoastaan määrälliseen tai rakenteiseen tietoon, vaan ammattilaisen tulee huomioida myös tilanteen konteksti, asiakkaan oma näkökulma sekä sellaiset yksilölliset tekijät, joita ei välttämättä voida muuttaa algoritmin käsittelemäksi dataksi. Jos tekoälyn tuottamia ennusteita, luokituksia tai suosituksia ryhdytään pitämään ammatillista harkintaa ohjaavana tai jopa ratkaisevana tietona, vaarana on yksilöllisen tilannearvion kaventuminen. Tällöin lapsen ja perheen ainutlaatuiset olosuhteet voivat jäädä päätöksenteossa vähäisemmälle huomiolle. (Agbana 2025.) Tekoälyn eettisyyttä sosiaalityössä ei tule tarkastella ainoastaan teknologian ominaisuuksien näkökulmasta, vaan myös suhteessa siihen, miten tekoälyn käyttö vaikuttaa ammatilliseen harkintaan, päätöksentekoon ja kokonaisuutena asiakkaan asemaan.

Tekoälyn nopea kehitys ja sen lisääntyvä hyödyntäminen lastensuojelussa tekevät aiheesta ajankohtaisen ja tärkeän sosiaalityön näkökulmasta. Tutkimusaiheen valintaan on vaikuttanut myös oma mielenkiintoni tekoälyn käyttöön lastensuojelussa. Kiinnostus aiheeseen on syntynyt

työkokemusteni kautta lastensuojelun kentällä sosionomina ja harjoittelijana. Työelämässä tekoälyn hyödyntämiseen liittyvät eettiset kysymykset ovat herättäneet uteliaisuutta, mutta samalla niiden sisältö ja erilaiset ulottuvuudet ovat näyttäytyneet jossain määrin epäselvinä. Tämä havainto on osaltaan ohjannut tarkastelemaan sitä, mitä tekoälyn eettisyydellä lastensuojelun kontekstissa tarkoitetaan ja millaisia kysymyksiä eettisyyteen liittyy.

Pro gradu -tutkielmani tavoitteena on tarkastella aikaisemman tutkimustiedon perusteella tekoälyn käyttöön lastensuojelussa liittyviä eettisistä kysymyksiä. Tutkielma toteutetaan kuvailevana kirjallisuuskatsauksena. Tekoälyn eettistä käyttöä koskeva suomalainen tutkimus sosiaalityön ja erityisesti lastensuojelun kontekstissa on toistaiseksi rajallista. Kansainvälistä tutkimusta aiheesta on kuitenkin saatavilla, ja tekoälyn käyttöön liittyviä eettisiä kysymyksiä tarkastellaan eri näkökulmista.

2 LASTENSUOJELU JA TEKOÄLY

2.1 Lastensuojelun lähtökohdat

Suomalaista lastensuojelua määrittelee lastensuojelulaki, joka turvaa lapsen oikeutta turvalliseen kasvuympäristöön, tasapainoiseen ja monipuoliseen kehitykseen sekä erityiseen suojeluun (THL 2025). Vanhemmilla ja muilla huoltajilla on ensisijainen vastuu lapsen hyvinvoinnista, mutta mikäli he eivät kykene turvaamaan lapsen hyvinvointia ja kasvua, viranomaisten tehtävänä on tukea perhettä sekä tarvittaessa ohjata lapsi ja perhe lastensuojelun piiriin. Lastensuojelun tehtävänä on tällöin turvata lapsen etu ja järjestää lapselle ja perheelle tarvittavia tukitoimia (Lastensuojelulaki 417/2007, 2 §). Lastensuojelu voidaan jakaa ehkäisevään sekä lapsi- ja perhekohtaiseen lastensuojeluun. Hyvinvointialueet vastaavat lapsi- ja perhekohtaisen sekä ehkäisevän lastensuojelun, jonka järjestämisen vastuu kuuluu myös kunnille, organisoimisesta (THL 2025).

Ehkäisevän lastensuojelun piiriin kuuluu esimerkiksi neuvolan, varhaiskasvatuksen sekä opetuksen ja nuorisotyön tarjoama tuki ja erityinen tuki (THL 2025). Jos ehkäisevän lastensuojelun keinot eivät ole riittävät lapsen kehityksen turvaamiseen, siirrytään lapsi- ja perhekohtaiseen lastensuojeluun (HE 237/2018 vp). Lapsi- ja perhekohtaista suojelua toteutetaan esimerkiksi avohuollon tukitoimien järjestämisellä. Avohuollon tukitoimet ovat ensisijaisia, pääasiassa vapaaehtoisia, lastensuojelutoimenpiteitä. Niiden tavoitteena on turvata lapsen hyvinvointi ja tukea vanhemmuutta, jotta lapsi voi asua kotona. Tukitoimet suunnitellaan yhdessä perheen kanssa. Avohuollon tukitoimia ovat esimerkiksi perhetyö, tukihenkilötyöskentely, taloudellinen tuki, perhekuntoutus ja lapsen kuntoutus (Räty 2023, 380, 383).

Myös sijaishuolto ja jälkihuolto kuuluvat lapsi- ja perhekohtaiseen lastensuojeluun (Lastensuojelulaki 417/2007, 3 §). Sijaishuolto tarkoittaa lapsen hoidon ja kasvatuksen järjestämistä kodin ulkopuolella, mikäli lapsi on kiireellisesti sijoitettu tai huostaanotettu (Räty 2023, 582). Kiireellisessä sijoituksessa lapsi otetaan tilapäisesti kotoaan pois ja lapsi voidaan sijoittaa esimerkiksi perhehoitoon tai laitokseen. Kiireellinen sijoitus toteutetaan, jos lapsi on välittömässä vaarassa kotiolojensa tai oman toimintansa takia. (Mt., 409.) Huostaanotto on viimesijainen keino lapsen hyvinvoinnin turvaamiseen muiden tukitoimien ollessa riittämättömiä. Huostaanotto puuttuu perustuslaissa säädettyihin oikeuksiin, kuten lapsen

henkilökohtaiseen vapauteen. Se kuitenkin turvaa myös lapsen oikeutta riittävään suojeluun ja lapsen edun mukaiseen elämään. (Mt., 435, 438.) Jälkihuolto turvaa lapsen ja nuoren valmiuksia ja tukea itsenäiseen elämään sijaishuollon jälkeen (mt., 871). Jälkihuollon tuki voi olla esimerkiksi asumiseen, toimeentuloon, koulutukseen, työllistymiseen ja arjenhallintaan liittyvää tukea. Tuen tulee perustua nuoren yksilölliseen tilanteeseen ja tarpeisiin, sillä jälkihuollossa olevien nuorten tuen tarpeet vaihtelevat huomattavasti. (Hirschovits-Gerz & Weckroth 2024, 1–2.)

Lapsen etu on keskeinen lähtökohta lastensuojelun toiminnassa. Lapsen osallisuus ja mahdollisuus kuulluksi tulemiseen ovat keskeisiä lastensuojelun periaatteita, mutta niiden toteutuminen käytännössä ei ole aina ongelmatonta. Lastensuojelussa lasten näkemykset ja kokemukset voivat jäädä aikuisten näkökulmien varjoon, mikä voi heikentää lapsen mahdollisuuksia vaikuttaa itseään koskeviin asioihin. (Bardy 2013, 7–9; Jahnukainen & Paasivirta 2022, 188–189.) Sosiaalityöntekijällä on keskeinen rooli lapsen edun arvioinnissa ja toteuttamisessa, minkä vuoksi hänen tulee samanaikaisesti huomioida lapsen näkemykset, vanhempien oikeudet ja lapsen suojelun tarve. Lastensuojelun tulee olla mahdollisimman hienovaraista ja avohuollon tukitoimet ovat ensisijaisia. Sijaishuolto tulee kuitenkin järjestää viipymättä, mikäli lapsen edun toteutuminen sen vaatii. Sijaishuolto järjestetään lapsen edun mukaisesti ja taustalla on aina tavoite perheen jälleenyhdistämisestä (Räty 2026, 8–9).

Lapsen etu ei ole yksiselitteinen tai yksinomaan teknisesti määriteltävissä oleva asia, vaan sen arviointi edellyttää lapsen tilanteen kokonaisvaltaista tarkastelua. Lapsen näkemyksillä ja osallisuudella on keskeinen merkitys lastensuojeluprosessissa, ja lapsen mielipide tulee ottaa huomioon hänen ikänsä ja kehitystasonsa mukaisesti. Lastensuojelun tutkimuksessa on kuitenkin havaittu, että lasten osallisuus ei aina toteudu käytännössä, vaan lapset voivat jäädä sivuun heitä koskevasta päätöksenteosta. (Schulz 2024.) Tämä korostaa lastensuojelutyössä tarvittavaa ammatillista harkintaa sekä kykyä tarkastella lapsen elämäntilannetta kokonaisuutena.

Erilaiset arviot ja päätökset ovat keskeisessä roolissa lastensuojelussa, ja niillä on merkittäviä vaikutuksia lasten ja perheiden elämään (Taylor & White 2006). Lastensuojelu on jatkuvan yhteiskunnallisen keskustelun kohteena ja sen tarkastelu tutkimuksen näkökulmasta on tärkeää (Parton & Kirk 2009, 23–26). Lastensuojelun sosiaalityöntekijät tasapainottelevat eettisten ja oikeudellisten näkökulmien välillä ja puntaroivat lapsen edun mukaisia ratkaisuja samalla, kun

lapsen ja huoltajien näkemykset tulee ottaa huomioon (Forrester 2024). Päätöksiä tehdään usein epävarmoissa olosuhteissa, jolloin sosiaalityöntekijä joutuu arvioimaan sekä menneitä tapahtumia että mahdollisia tulevia kehityskulkuja (Taylor & White 2006). Päätöksiin sisältyy aina myös tulkinta lapsen tilanteesta ja siitä, millaiset toimenpiteet ovat lapsen edun mukaisia. Esimerkiksi lastensuojeluilmoituksen sulkeminen sisältää oletuksen siitä, että lapsen kasvuolosuhteet ovat riittävät turvaamaan lapsen kehityksen. Näin lastensuojelun päätöksenteko ei näyttäydy yksittäisen tiedon tai muuttujan perusteella tapahtuvana ratkaisuna, vaan prosessina, jossa erilaista tietoa punnitaan ja tulkitaan suhteessa lapsen tilanteeseen.

Lastensuojelun päätöksenteko rakentuu sekä implisiittisten että eksplisiittisten arvioiden varaan (Taylor & White 2006; Clardy 2022). Suomalaisessa lastensuojelussa arviointi ja päätöksenteko näyttäytyvät tutkimuksen perusteella monivaiheisina ja sosiaalisissa suhteissa rakentuvina prosesseina. Lamposen (2022) tutkimuksessa sosiaalityöntekijöiden päätöksenteko perustui useanlaiseen tietoon, kuten faktatietoon, kokemukselliseen tietoon ja intuitiiviseen tietoon. Arviointia ja päätöksentekoa tehdään tilanteissa, joissa käytettävissä oleva tieto voi olla pirstaleista tai siinä voi olla merkittäviä aukkoja. Sosiaalityöntekijä rakentaa käsitystä lapsen tilanteesta suhteuttamalla menneisyyttä, nykyhetken tietoja ja tulevaisuutta koskevia arvioita toisiinsa (Lamponen 2022). Lastensuojelun päätöksenteossa on siten olennaista paitsi käytettävissä oleva tieto myös se, miten tietoa tulkitaan, suhteutetaan muuhun tietoon ja muutetaan ammatillisiksi arvioiksi ja päätöksiksi.

Lastensuojelun sosiaalityössä päätöksentekoon vaikuttavat myös työn rakenteelliset ja organisatoriset olosuhteet. Suomessa lastensuojelun toimivuutta on pyritty vahvistamaan muun muassa sosiaalityöntekijöiden asiakasmäärää koskevalla henkilöstömitoituksella. THL:n selvityksen mukaan lastensuojelun sosiaalityön organisoinnissa ja palvelujen toimivuudessa on kuitenkin edelleen kysymyksiä, jotka liittyvät muun muassa henkilöstöresursseihin, työn organisointiin ja palvelujen saatavuuteen (Yliruka ym. 2022). Lastensuojelun sosiaalityö tapahtuu siten toimintaympäristössä, jossa ammatilliselle harkinnalle asetetaan samanaikaisesti oikeudellisia, eettisiä, organisatorisia ja ajallisia vaatimuksia.

2.2 Tekoäly tiedon tuottamisen ja päätöksenteon välineenä

Tekoälyn (artificial intelligence, AI) käsitteen kehittäjänä pidetään Stanfordin yliopiston professoria John McCarthya, joka käytti termiä ensimmäisen kerran vuonna 1955 (Stanford HAI 2022). Tekoälyllä tarkoitetaan ihmisen tai eläimen kognitiivisiin kykyihin perustumatonta älyllistä toimintaa (El Naqa & Murphy 2015). Se koostuu heterogeenisistä kone- ja syväoppimisien malleista, joiden avulla suuria datamääriä voidaan analysoida nopeasti. Tekoäly voi koneoppimisen avulla löytää erilaisia kaavoja, käsitellä laajoja tietoaineistoja ja ennustaa sekä tehdä päätöksiä suurten ja keskenään erilaisten aineistojen pohjalta. (Kivipelto 2025, 3.)

Generatiivinen tekoäly on tekoälyn osa-alue, joka luo itsenäisesti inhimilliseltä vaikuttavaa aineistoa aiemman datan pohjalta (Sengar, Bin Hasan, Kumar & Carroll 2024). ChatGPT on esimerkki generatiivisesta tekoälystä. Se tuottaa vastauksia hyödyntämällä koulutusdataansa pohjautuvia kielellisiä rakenteita ja todennäköisyyksiä ennustaen, mikä sana, lause tai kokonaisuus on kontekstissa todennäköisin seuraava. (Wilkins & Benett 2025.) Generatiiviset mallit ovat uutta ja ihmismäistä sisältöä tuottamalla muokanneet useita toimialoja ja synnyttäneet uusia ympäristöjä tekoälyn hyödyntämiselle (Sengar, Bin Hasan, Kumar & Carroll 2024).

Bickin, Blandin & Demingin (2024) mukaan generatiivisen tekoälyn nopea leviäminen voidaan rinnastaa aiempien digitaaliteknologioiden, kuten tietokoneen ja internetin, omaksumiseen. Brandão (2025) mukaan tekoälyn voidaan nähdä tulleen perustavanlaatuiseksi osaksi yhteiskuntaa ja sen käytön jatkuva laajentuminen kertoo laajemmasta digitaalisesta transformaatiosta. Generatiivisen tekoälyn on luonnehdittu sopivan parhaiten tehtäviin, joissa ihminen käsittelee ja tarkastaa sen tuotokset eivätkä yksittäisten virheiden haitat ole merkittäviä. Kuten apuvälineeksi ideoiden ja tekstiluonnosten tuottamiseen sekä kielentarkistukseen. (Kauhanen, Pajarinen & Rouvinen 2023, 7; United Nations Economic Commission for Europe 2025, 13.)

Pääministeri Petteri Orpon (20.6.2023-) hallitusohjelma määrittelee Suomen teknologiseksi edelläkävijäksi, joka hyödyntää laajasti digitalisaation ja tekoälyn tarjoamia mahdollisuuksia (Valtioneuvosto 2023, 40). Tekoälyn nähdään tulevaisuudessa sisältyvän melkein kaikkeen sosiaali- ja terveydenhuollon palveluihin sekä toimintaan (Ministry of Social Affairs and Health

2024, 11). Tekoälyn avulla pyritään sosiaali- ja terveydenhuollossa parantamaan palvelujen laatua, helpottamaan ammattilaisten työskentelyä ja vapauttamaan aikaa asiakastyön kannalta keskeisiin tehtäviin, kuten vuorovaikutukseen. Lisäksi tekoälyn käytön tavoitteena on vahvistaa palvelujärjestelmän toimivuutta ja tuottavuutta sekä edistää uusien liiketoimintamahdollisuuksien ja työpaikkojen syntymistä myös kansainvälisesti. (Mt., 5.) Tekoälyä mahdollistavassa lainsäädännössä ilmaistaan, että tekoäly mahdollistetaan sosiaali- ja terveydenhuollossa perusoikeudet turvaten. Ne sosiaali- ja terveydenhuollon tehtävät, jotka ovat todennäköisimmin automatisoitavissa, selvitetään ja erilaiset riskit ja mahdollisuudet kartoitetaan. (Valtioneuvosto 2023, 40.) Sosiaali- ja terveysministeriön linjausten mukaan tekoälyn kehittämisen ja käytön sosiaali- ja terveydenhuollossa tulee perustua avoimuuteen, vastuullisuuteen, oikeudenmukaisuuteen ja eri toimijoiden osallistamiseen (Digital Health Finland, 2026). Valtioneuvoston mukaan on tärkeää, että tekoälyn käyttöä valvotaan ja vastuut määritellään, sillä tekoälymallin mukaisissa ohjelmistoissa on olemassa virheiden riski. Virheet voivat johtua monista tekijöistä, eikä niitä voida aina ennaltaehkäistä. Koneoppimiseen perustuvat järjestelmät voivat esimerkiksi muokkaantua käytössä ja mallien käyttöä ja käyttäjien osaamista tulee valvoa. (Valtioneuvosto 2023, 40.)

Sosiaalityössä tekoälyä voidaan käyttää esimerkiksi asiakastyössä, palveluohjauksessa, esihenkilötyössä, avustavassa työssä ja tutkimuksessa (Reamer 2023, 52). Tomiţăn ja kumppaneiden (2024) mukaan teknologian hyödyntämiseen sosiaalityössä liittyy mahdollisuuksia, mutta myös eettisiä kysymyksiä, jotka koskevat muun muassa asiakkaiden yksityisyyttä, ammatillisia rajoja ja yhdenvertaista teknologian saatavuutta (13–25). Generatiivinen tekoäly, kuten ChatGPT, on puolestaan tuonut sosiaalityöhön uusia mahdollisuuksia tekstin tuottamiseen ja ammatillisen työn tukemiseen (Goldkind ym. 2023, 103–105). Ennakoivaa tekoälyanalytiikkaa voidaan hyödyntää myös päätöksenteon ja palvelujen kehittämisen tukena (Goldkind 2021, 372–373).

2.3 Teknologisoituva lastensuojelu

Sosiaalityötä ja erityisesti lastensuojelua on kritisoitu julkisuudessa käsiteltyjen epäonnistumisten vuoksi jo 1970-luvulta lähtien. Vastauksena kritiikkiin poliittisilla ja hallinnollisilla muutoksilla on lisätty lastensuojelun tilivelvollisuutta ja korostettu tehokkuutta ja tuloksellisuutta. (Parton & Kirk 2009, 32; Devlieghere ym. 2022.) Vaatimusten ja käytännön

työn välistä jännitettä on pyritty ratkaisemaan esimerkiksi kehittämällä systemaattisempia työskentelytapoja sekä hyödyntämällä teknologiaa ja tietojärjestelmiä osana lastensuojelun prosesseja. Teknologian käyttöönotto on lisännyt työn kontrolloitavuutta ja standardointia sekä siirtänyt luottamusta ammatillisesta harkinnasta osittain järjestelmiin ja niiden tuottamaan tietoon (Parton &Kirk 2009, 718–719; Smith 2004). Samalla sosiaalityön tiedollinen painopiste on osin siirtynyt sosiaalisen ymmärtämisestä yhä enemmän tiedon keräämiseen, käsittelyyn ja hallintaan (Parton & Kirk 2009, 253–254).

Teknologisoituminen näkyy käytännössä lastensuojelussa esimerkiksi juuri tekoälyn käyttöönottona. Kansainvälisesti tekoälyä on hyödynnetty erityisesti ennakoivassa riskimallinnuksessa (predictive risk modelling), tietotyön tukena sekä palvelujen priorisoinnissa ja kohdentamisessa. Ennakoivassa riskimallinnuksessa kehitetään malleja, jotka arvioivat esimerkiksi lastensuojeluasiakkuuden todennäköisyyttä tai lisätuen tarvetta laajoihin hallinnollisiin tietokantoihin perustuvan datan avulla. Mallien tavoitteena on tunnistaa varhaisen tuen tarvetta sekä tukea resurssien kohdentamista. (Gillingham 2016, 1044–1045.) Lisäksi tekoälyä on kehitetty tukemaan lastensuojelun arjen tietotyötä. Sitä voidaan hyödyntää esimerkiksi tekstin tuottamisessa, kirjaamisen helpottamisessa sekä asiakasdokumenteista löytyvän tiedon jäsentelyssä ja poiminnassa. Tekoälyä voidaan käyttää myös päätöksenteon tukena esimerkiksi suosituksia tuottavien järjestelmien muodossa. Palvelujen kohdentamisessa ja priorisoinnissa tekoälyä on hyödynnetty esimerkiksi ilmoitusten käsittelyn priorisointiin, resurssien suuntaamiseen sekä ennaltaehkäisevien palvelujen kohderyhmien tunnistamiseen. (Roberts, O'Brien & Pecora 2018, 2–3.)

Suomalaisessa lastensuojelukontekstissa tekoälyn hyödyntämistä on toistaiseksi tarkasteltu pääasiassa erilaisten pilotointien ja tutkimushankkeiden kautta. Näissä kokeiluissa tekoälyä on käytetty esimerkiksi riskien ennakointiin, ennaltaehkäisevän työn tukemiseen sekä sosiaalityöntekijöiden päätöksenteon apuvälineenä (Aalto-yliopisto 2019; Lehtiniemi ym. 2024). Yksi esimerkki tekoälyn hyödyntämisestä lastensuojelussa on Espoon kaupungin vuonna 2018 toteuttama kokeilu, jossa tekoälyä käytettiin ennustamaan lastensuojelun asiakkuuden riskiä laajojen rekisteriaineistojen perusteella. Kokeilua ei kuitenkaan otettu suoraan käyttöön päätöksenteossa, sillä siihen liittyi merkittäviä eettisiä ja oikeudellisia haasteita, kuten yksityisyyden suojaan, datan käyttöön sekä algoritmisen päätöksenteon läpinäkyvyyteen liittyviä kysymyksiä (Aalto-yliopisto 2019; University of Helsinki 2019). Tekoälyn mahdollisuuksia sosiaalihuollossa tutkitaan parhaillaan käynnissä olevissa

kehittämishankkeissa. Vaasan ammattikorkeakoulussa käynnistyi vuonna 2025 hanke, jossa selvitetään tekoälyn potentiaalia tehostaa sosiaalihuollon palveluja, mukaan lukien lastensuojelu (Vaasan ammattikorkeakoulu 2025). Lapin yliopistossa käynnistyi vuonna 2025 tutkimushanke Eettinen ja vastuullinen tekoäly sosiaalityössä: muuttuvien tietokäytäntöjen seuraukset asiakkaille, sosiaalityöntekijöille ja organisaatioille, jossa selvitetään, miten tekoälyn hyödyntäminen muuttaa sosiaalityön tiedon muodostumista ja ammatillisia käytäntöjä sekä vaikuttaa asiakkaiden asemaan (Lapin yliopisto 2025). Pirkanmaan hyvinvointialueella aloitettiin lokakuussa 2025 kokeilu, jossa tekoälyn mahdollisuuksia testataan lastensuojeluilmoitusten käsittelyn ja lapsen palvelutarpeen arvioinnin tukena. Tekoälyratkaisu tuottaa työntekijöiden tekemistä kirjauksista yhteenvedon, jonka ammattilainen tarkistaa ja tarvittaessa korjaa. Lisäksi Pirkanmaan hyvinvointialueella on arvioitu ja testattu tekoälyavusteisen riskiarvioinnin mahdollisuuksia, jossa tavoitteena on ollut tunnistaa lapsen kasvuun ja kehitykseen vaikuttavia riskitekijöitä automaattisesti taustadatasta (Pirkanmaan hyvinvointialue 2025; 2026).

Teknologian lisääntyvä hyödyntäminen lastensuojelussa liittyy laajemmin sosiaalityön muutokseen. Sosiaalityö on lähtökohtaisesti eettisesti latautunutta toimintaa, jossa päätöksentekoa ohjaavat sekä lainsäädäntö että ammattieettiset periaatteet. Sosiaalialan ammattietiikan keskeisiä arvoja ovat ihmisarvon kunnioittaminen, ihmisoikeuksien edistäminen, sosiaalinen oikeudenmukaisuus, yhdenvertaisuus sekä asiakkaan osallisuuden ja itsemääräämisoikeuden tukeminen. Sosiaalityöntekijän tehtävänä on edistää asiakkaiden oikeuksien toteutumista sekä arvioida toimintansa vaikutuksia erityisesti haavoittuvassa asemassa olevien ihmisten näkökulmasta. (Talentia 2024; Raunio 2009.) Eettisesti kestävä sosiaalityö edellyttää myös oman toiminnan jatkuvaa arviointia ja reflektiota. Sosiaalityöntekijän tulee tunnistaa omien ratkaisujensa taustalla vaikuttavia arvoja, oletuksia ja valtasuhteita sekä tarkastella kriittisesti toimintansa vaikutuksia asiakkaisiin. Ammatillinen asiantuntijuus ei siten perustu ainoastaan säädösten tai menetelmien soveltamiseen, vaan myös kykyyn perustella ratkaisuja eettisesti sekä arvioida niiden seurauksia muuttuvissa tilanteissa. (Kananoja ym. 2022; Talentia 2024.)

Erityisesti lastensuojelussa ammatillinen toiminta perustuu jatkuvaan harkintaan ja reflektioon. Päätöksiä tehdään usein epävarmoissa tilanteissa, joissa työntekijän on arvioitava samanaikaisesti lapsen etua, perheen näkemyksiä, lainsäädännön velvoitteita sekä päätösten mahdollisia seurauksia. Asiakastyössä syntyvä tieto rakentuu vuorovaikutuksessa asiakkaiden

kanssa, minkä vuoksi tilanteiden ymmärtäminen edellyttää yksilöllisten olosuhteiden, kokemusten ja merkitysten huomioimista. Sosiaalityön asiantuntijuutta ei siten voida määritellä yksinomaan standardoituihin arviointimenetelmiin tai teknisiin ratkaisuihin, vaan keskeisessä asemassa säilyvät ammatillinen harkinta, eettinen arviointi ja vastuu päätösten seurauksista (Taylor & White 2006; Forrester 2024).

Tekoälyn hyödyntäminen lastensuojelussa herättää moniulotteisia eettisiä kysymyksiä. Algoritmeihin ja tietojärjestelmiin perustuvat arviointimenetelmät edellyttävät usein asiakkaiden luokittelua sekä tiedon systemaattista koodaamista, mikä voi olla ongelmallista sosiaalityön näkökulmasta. Lastensuojelutyölle on ominaista, että jokainen asiakastilanne on uniikki, eikä asiakaskohtaamisissa rakentuvia vuorovaikutustilanteita voida täysin ennakoida tai standardoida (Jokinen 2016, 145). Tämä korostaa yksilöllisen tilanteen huomioimisen merkitystä myös silloin, kun asiakkaiden tilanteita pyritään jäsentämään erilaisten luokitusten ja kategorioiden avulla. Luokittelu voi helpottaa monimutkaisten asiakastilanteiden jäsentämistä, mutta samalla se voi pelkistää yksilön tietyn ongelman tai kategorian edustajaksi ja jättää hänen yksilölliset kokemuksensa ja tilanteensa taka-alalle. (Pohjola 2010, 29–35.) Tämän vuoksi teknologisten ratkaisujen soveltaminen lastensuojeluun edellyttää huolellista eettistä arviointia sekä ymmärrystä niiden vaikutuksista asiakkaiden oikeuksiin, yhdenvertaisuuteen ja ammatilliseen päätöksentekoon.

3. TUTKIMUKSEN TOTEUTUS

3.1 Tutkimustehtävä

Tarkastelen pro gradu –tutkielmassani tekoälyn käyttöön lastensuojelussa liittyviä eettisiä kysymyksiä. Eettisyyden arvioinnissa keskeistä on lapsen edun sekä ihmisoikeuksien toteutuminen kaikissa tilanteissa, joissa tekoälyä hyödynnetään. Saatavilla olevan tiedon mukaan tekoälyä hyödynnetään suomalaisessa lastensuojelussa jonkin verran ja jatkuvasti enenevissä määrin. Toteutan tutkielmani kirjallisuuskatsauksena tutkimusartikkeleista, jotka ovat pääosin kansainvälisiä. Kansainväliset tutkimukset eivät tarjoa suoria ratkaisuja suomalaiseen lastensuojelujärjestelmään, mutta niitä ja niiden pohjalta syntyneitä eettisiä kysymyksiä voi soveltaa myös suomalaiseen kontekstiin. Tutkimusaineisto kostuu 12

vertaisarvioidusta tutkimusartikkelista vuosilta 2022–2026. Tutkimuskysymykseni on: Millaisia eettisiä kysymyksiä tekoälyn käyttöön lastensuojelussa liittyy tutkimuskirjallisuuden perusteella?

3.2 Kuvaileva kirjallisuuskatsaus

Kirjallisuuskatsauksen avulla voidaan tarkastella aiempaa tutkimusta kokonaisuutena ja muodostaa käsitys siitä, millaista tietoa tarkasteltavasta ilmiöstä on jo tuotettu. Kirjallisuuskatsausten välillä on eroja niiden tavoitteissa, aineiston valinnassa ja siinä, miten aiempaa tutkimustietoa analysoidaan ja yhdistetään (Snyder 2019, 333–334). Salminen (2011, 6) sijoittaa kirjallisuuskatsauksen pääluokiksi kuvailevan ja systemaattisen kirjallisuuskatsauksen sekä meta-analyysin. Vilkka (2013, 19) puolestaan jäsentää kirjallisuuskatsaukset narratiiviseen, integratiiviseen ja systemaattiseen kirjallisuuskatsaukseen sekä meta-analyysiin. Tutkielmassani kirjallisuuskatsaus toimii menetelmänä, jonka avulla kokoan aikaisempaa tutkimustietoa ja tulkitsen sitä suhteessa asetettuun tutkimuskysymykseen. Sosiaalityön tutkimuksessa kirjallisuuskatsauksia voidaan hyödyntää esimerkiksi aiemman tutkimustiedon kokoamiseen ja tutkimuskohteen kokonaiskuvan muodostamiseen (Laaksonen, Kuusisto-Niemi & Saranto 2010, 26–28).

Valitsin tutkimusmetodikseni kuvailevan kirjallisuuskatsauksen narratiivisella orientaatiolla. Kuvaileva kirjallisuuskatsaus mahdollistaa tutkittavan ilmiön kattavan tarkastelun ilman systemaattisen kirjallisuuskatsauksen tiukkoja menetelmällisiä rajauksia (Salminen 2011, 6). Narratiivisessa kirjallisuuskatsauksessa aikaisempaa tutkimustietoa voidaan tarkastella ja yhdistää siten, että tutkittavasta ilmiöstä muodostuu kokonaiskuva (Salminen 2011, 7–8; Vilkka 2013, 21–22). Katsauksen tarkoituksena ei tällöin ole tuottaa yksittäisten tutkimusten tuloksista tilastollista tai muuta systemaattista yhteenvetoa, vaan jäsentää hajanaista tutkimustietoa ja tunnistaa tutkimuksissa esiintyviä näkökulmia ja niiden välisiä yhteyksiä. Narratiivinen kirjallisuuskatsaus mahdollistaa aiemman tutkimustiedon kokoamisen ja eri tutkimuksissa esiin nousseiden näkökulmien sekä tulosten tarkastelun suhteessa toisiinsa (Snyder 2019, 334). Menetelmä soveltuu erityisesti tilanteisiin, joissa tutkimusaiheesta on saatavilla eri tavoin toteutettuja tutkimuksia ja tutkimustieto on hajanaista. Tällöin katsauksen avulla voidaan muodostaa kokonaiskuvaa ilmiöstä ja tunnistaa tutkimuksessa esiintyviä keskeisiä teemoja.

Tutkimusaineistoni muodostuu erilaisista näkökulmista tekoälyn käytöstä lastensuojelussa. Osissa artikkeleissa käsitellään suoraan tekoälyn käytön eettisiä kysymyksiä, osassa eettiset kysymykset linkittyvät muuten tutkimukseen. Aineiston valinnassa painotin tutkimuksia, joissa tekoälyä tarkastellaan lastensuojelun tai siihen läheisesti liittyvien palvelujen yhteydessä ja joiden avulla voin tarkastella tekoälyn käyttöön liittyviä eettisiä kysymyksiä.

Kuvailevan kirjallisuuskatsauksen etu onkin sen joustavuus aineiston suhteen. Katsaukseen voidaan sisällyttää erilaisia tutkimusasetelmia ja tutkimusmenetelmiä käyttäen tuotettua tutkimustietoa, jolloin tarkasteltavasta ilmiöstä voidaan saada monipuolinen kuva (Salminen 2011, 6–8; Vilkka 2013, 21–22). Tämä on tutkielmani kannalta tärkeää, koska tekoälyn käyttöä lastensuojelussa koskeva tutkimus on vielä uutta ja tutkimusta on tehty erilaisista lähtökohdista. Valikoituneissa tutkimuksissa tekoälyä on tarkasteltu esimerkiksi ennakoivan riskimallinnuksen sekä algoritmisen päätöksenteon näkökulmista. Lähestymistavoiltaan erilaisten tutkimuksen tarkastelu samassa katsauksessa mahdollistaa tekoälyn eettisten kysymysten tarkastelun riittävän laajasti.

3.3 Tutkimusaineiston haku ja valinta

Hankin tutkielmani aineiston vaiheittain etenevällä tiedonhaulla, jonka tavoitteena oli tunnistaa tutkimustehtävän kannalta olennaiset vertaisarvioidut tutkimusartikkelit. Tiedonhaku eteni alustavasta aiheeseen perehtymisestä varsinaiseen kirjallisuushakuun, aineiston seulontaan ja lopullisen tutkimusaineiston muodostamiseen. Hakuprosessia ohjasivat ennalta määritellyt sisäänotto- ja poissulkukriteerit, joiden avulla rajasin aineiston tutkimuskysymysten kannalta tarkoituksenmukaiseksi.

Alustavassa tiedonhaussa perehdyin tekoälyn käyttöä käsittelevään tutkimukseen lastensuojelun ja sosiaalityön kontekstissa. Tämän jälkeen tarkensin hakua koskemaan tekoälyn eettisyyttä. Jo alustavan tiedonhaun aikana kävi ilmi, että aiheesta oli saatavilla vain vähän suomalaista tutkimusta. Tämän vuoksi laajensin tiedonhaun kansainväliseen tutkimuskirjallisuuteen. Kansainvälisestä kirjallisuudesta löytyi tutkimuksia sekä tekoälyn käytöstä lastensuojelussa että tekoälyn eettisyydestä sosiaalityössä. Sen sijaan nimenomaan tekoälyn eettisyyttä lastensuojelussa käsitteleviä tutkimuksia oli vähän. Tämän vuoksi lopullinen aineisto muodostui sekä tutkimuksista, joissa tarkastellaan tekoälyn eettisyyttä

lastensuojelussa, että tutkimuksista, joissa kuvataan tekoälyn käyttöä lastensuojelussa eri näkökulmista.

Taulukko 1

Sisäänottokriteerit	**Poissulkukriteerit**
Eettisiä kysymyksiä ja näkökantoja tarkastellaan	Ei käsitellä eettisiä kysymyksiä
Käsittelee lastensuojelua	Ei käsittele keskeisesti lastensuojelua
Saatavilla luettavaksi ilmaiseksi tai LUC-finna	Maksullinen tai ei vapaasti saatavilla
Suomen- tai englanninkielinen	Muu kuin suomen- tai englanninkielinen
Vertaisarvioitu	Ei vertaisarvioitu

Taulukko 1 kuvaa aineiston sisäänotto- ja poissulkukriteereitä. Hyväksyin mukaan vertaisarvioidut suomen- ja englanninkieliset tutkimukset, jotka käsittelivät lastensuojelua ja tarkastelivat tekoälyyn liittyviä eettisiä kysymyksiä tai olivat sisällöltään sovellettavissa lastensuojelun kontekstiin. Lisäksi tutkimusten tuli olla saatavilla LUC-Finnan kautta tai ilmaiseksi.

Varsinaisen kirjallisuushaun toteutin Lapin korkeakoulukirjaston tarjoamissa kotimaisissa ja kansainvälisissä tietokannoissa. Käytettyjä tietokantoja olivat LUC-Finna, Duodecim, LUC-Finnan kansainväliset e-aineistot, AgeLine, Social Science Database, Scopus ja Web of Science. Lisäksi tein täydentäviä hakuja kotimaisten tieteellisten lehtien julkaisu- ja hakupalvelu Journal.fi:ssä ja ChatGPT:n avulla. Jotta hakulausekkeet kattaisivat mahdollisimman laajasti tutkimusaiheen keskeiset käsitteet, hyödynsin hakusanojen muodostamiseen asiasanasto- ja ontologiapalveluita. Suomenkielisissä tietokannoissa käytin hakulausekkeena lastensuojelu OR sosiaalipalvelut OR sijaishuolto AND generatiivinen tekoäly OR tekoäly OR koneoppiminen AND eettisyys OR etiikka OR ammattietiikka OR arvofilosofia OR moraali OR normit. Journal.fi:ssä suomenkielisiin tietokantoihin hyödyntämäni hakulauseke ei tuottanut tuloksia, joten täsmensin hakulausekkeeksi (lastensuojelu OR tekoäly). Kansainvälisissä tietokannoissa hakulausekkeeksi muodostui child welfare OR child-related work OR child protection AND artificial intelligence OR generative artificial intelligence OR knowledge technology OR machine learning AND ethics OR ethicality OR moral OR norms OR ethical. Liite 1 kuvaa hakusanoja.

Seulonnan ensimmäisessä vaiheessa arvioin artikkelien tiivistelmiä. 213 artikkelista valikoitui jatkoon 70 artikkelia tiivistelmän perusteella. Artikkeleista 25 oli LUC-Finnan kansainvälisistä e-aineistoista, 22 Journal.fi:stä, 15 Web of Sciencestä, neljä ChatGPT:n löytämää, kaksi Social Science Databasesta sekä kaksi Scopucsesta. Seulonnan toisessa vaiheessa tarkastelin näitä 70:ää artikkelia lukemalla niiden johdannot sekä arvioimalla kokotekstejä yleisellä tasolla. Johdantojen ja koko tekstin tarkastelun jälkeen jatkoon valikoitui 24 artikkelia.

Kokotekstiseulonnassa luin 24 artikkelia läpi kokonaisuudessaan. Kokotekstiseulonnassa jatkoon valitsemisen keskeisenä perusteena oli, että tutkimus tarkasteli tekoälyä lastensuojelun kontekstissa tai sen tulokset olivat perustellusti sovellettavissa lastensuojeluun. Lisäksi asetin etusijalle tutkimukset, joissa tekoälyä käsiteltiin nimenomaan eettisestä näkökulmasta. Tämän seulontavaiheen jälkeen aineistoon jäi 17 artikkelia.

Lopuksi tarkastelin vielä tutkimusten julkaisuvuosia suhteessa tutkimusaiheeseen. Koska nykyisen kaltaisen generatiivisen tekoälyn voidaan katsoa yleistyneen 2020-luvun alussa suurten kielimallien läpimurron myötä (Sengar ym. 2024), rajasin lopullisen aineiston koskemaan vuonna 2020 tai sen jälkeen julkaistuja tutkimuksia. Aikarajauksella pyrin tarkastelemaan mahdollisimman ajankohtaista tutkimustietoa. Lopulliseksi tutkimusaineistoksi valikoitui 12 vertaisarvioitua tutkimusartikkelia. Aineiston hakuprosessi on kuvattu liitteessä 2 ja lopulliseen aineistoon valikoituneet tutkimukset liitteessä 3.

3.4 Induktiivinen sisältöanalyysi

Kirjallisuuskatsauksen analyysi voidaan toteuttaa joko induktiivisesti tai deduktiivisesti. Induktiivisessa analyysissa teemat ja tulkinnat muodostetaan aineistolähtöisesti, kun taas deduktiivisessa analyysissa tarkastelu perustuu ennalta määriteltyyn teoreettiseen viitekehykseen (Vilkka 2023, 86). Kallisen ja Kinnusen (2021) mukaan laadullista tutkimusta voidaan aineistovetoisuuden sijaan kuvata analyysivetoiseksi, jolloin empiirisellä aineistolla on keskeinen asema, mutta analyysia ohjaa lisäksi jokin paradigma, menetelmä tai teoria. Toteutin tutkimuksessani analyysin induktiivisen sisältöanalyysin avulla.

Aineiston analyysi eteni induktiivisen sisältöanalyysin periaatteita noudattaen. Analyysiprosessin alkuvaiheessa tarkastelin tekoälyn käyttöä lastensuojelussa hyötyjen ja haittojen näkökulmasta. Aineistosta nousi esiin sekä tekoälyn käytön mahdollisia hyötyjä että

haittoja, mutta niiden tarkastelu sellaisenaan ei muodostanut eettisen näkökulman kannalta tarkoituksenmukaista kokonaisuutta. Tämän vuoksi kohdensin tarkastelua hyötyjen ja haittojen sijaan siihen, millaisia eettisiä kysymyksiä tekoälyn käyttöön lastensuojelussa liittyy. Hyötyjen ja haittojen tarkastelu toimi kuitenkin analyysiprosessin alkuvaiheessa hyödyllisenä lähtökohtana. Se auttoi hahmottamaan, millaisia vaikutuksia tekoälyn käytöllä lastensuojelussa voi olla ja millaisissa tilanteissa näihin vaikutuksiin voi liittyä eettisiä kysymyksiä. Näin alustava tarkastelu tuki aineiston myöhempää jäsentämistä tutkimuksen eettisestä näkökulmasta.

Sisältöanalyysi voidaan jakaa kolmeen vaiheeseen: analyysin valmisteluun, organisointiin ja raportointiin (Vilkka 2023, 86). Valmisteluvaiheessa aineisto järjestetään analysoitavaan muotoon, ja tutkimuksista tarkastellaan keskeisiä elementtejä, kuten aiheita, teoreettisia lähtökohtia, menetelmiä sekä mahdollisia muuttujia (mt., 86–87). Analyysini käynnistyi lukemalla valitut artikkelit uudelleen kokonaiskuvan muodostamiseksi. Samalla laadin yleiskuvan aineistosta kirjaamalla ylös artikkeleissa käsiteltyjä teemoja, keskeisiä ilmiöitä, toimijoita sekä mahdollisia poikkeavia näkökulmia.

Organisointivaiheessa aineistosta tunnistetaan tutkimuskysymysten kannalta olennaiset havainnot, joiden pohjalta muodostetaan tulokset ja johtopäätökset (Vilkka 2023, 87). Aineistoa tarkastellaan etsimällä esimerkiksi samankaltaisuuksia, toistuvia teemoja, säännönmukaisuuksia sekä poikkeuksia. Kävin läpi tutkimukset paikantaen niistä tekoälyn käyttöön liittyviä eettisiä kysymyksiä. Muodostin havaintojen pohjalta artikkelikohtaisia tekoälyn eettisyyteen liittyviä teemoja, joista osa oli keskenään päällekkäisiä. Tuloksia kirjoittaessani yhdistin osin päällekkäisiä teemoja niiden sisällöllisen yhteensopivuuden perusteella ja jäsensin havainnot osaksi sitä teemaa, johon ne sisällöllisesti parhaiten sopivat. Teemojen jäsentämisen tukena hyödynsin myös ChatGPTä, jonka avulla tarkastelin vaihtoehtoisia tapoja hahmottaa teemojen välisiä yhteyksiä. Lopulliset ratkaisut teemojen muodostamisesta ja niiden välisistä suhteista perustuivat kuitenkin omaan arviooni aineistosta.

Tulosten raportointivaiheessa tunnistetut teemat kootaan ja esitetään systemaattisesti (Vilkka 2023, 87). Raportointivaiheessa kokosin samaan teemaan liittyvät havainnot eri artikkeleista yhteen ja tarkastelin, millä tavoin eri tutkimukset kuvasivat samaa eettistä kysymystä. Esitin kunkin teeman yhteydessä artikkeleissa esiin nousseet keskeiset havainnot ja toin esille myös tutkimusten välisiä eroja ja yhtäläisyyksiä. Tämän perusteella muodostin viisi pääteemaa, joiden alle ryhmittelin artikkeleissa tunnistetut eettiset kysymykset. Kirjoitin tulokset näiden

viiden pääteeman mukaisesti siten, että kunkin teeman yhteydessä kuvasin eri artikkeleissa esiin nousseita havaintoja ja vertasin niiden sisältöjä toisiinsa. Kuudes teema muodostui eettisiin kysymyksiin artikkeleissa esitetyistä ratkaisuista. Johtopäätökset muodostin erikseen tulosten perusteella ja niitä tarkastellaan tutkielman johtopäätösluvussa.

3.5 Tutkimuksen eettiset lähtökohdat

Kirjallisuuskatsauksessa tutkimuseettiset kysymykset liittyvät erityisesti tiedonhankinnan, aineiston analysoinnin ja tutkimustulosten raportoinnin eettisyyteen. Rauhalan ja Virokannaksen (2011, 10) mukaan tieto ei ole koskaan täysin arvovapaata eikä tietoarvoa voida irrottaa toimijoistaan ja toimijoiden sosiaalisista kytkennöistä. Esimerkiksi koulutusalani, sosiaalityön tutkimuksen perinteet, aiemmat kokemukseni sekä yhteisöt, joihin kuulun, vaikuttavat lähestymistapaani tutkielmani aiheeseen. Pietarisen (2002, 59) mukaan toiminnan hyvin tai oikein suorittamisen lähtökohtana voidaan käyttää Aristoteleen ohjetta, jossa kysytään ensin toiminnan pääasiallista tehtävää ja tarkoitusta. Lähtökohta on merkityksellinen myös tekoälyn käyttöä sosiaalityössä tarkasteltaessa, sillä tekoälyn hyödyntämistä tulee arvioida suhteessa sosiaalityön perustehtävään ja eettisiin tavoitteisiin**.** Sosiaalityöllä on klassinen perinne ihmisoikeusalana, ja heikommassa asemassa olevien ihmisten oikeuksien puolustaminen on yksi sen päämääristä (Turtiainen, Jäppinen & Ingo 2021, 27). Sosiaalityön arvoperusta muodostaa keskeisen lähtökohdan myös arvioitaessa tekoälyn käytön eettisyyttä lastensuojelussa, jossa teknologian vaikutukset voivat kohdistua haavoittuvassa asemassa oleviin lapsiin ja perheisiin.

Sosiaalityön eettistä koodistoa ylläpitävän sosiaalialan korkeakoulutettujen ammattijärjestö Talentia ry:n ammattieettiset ohjeet (2024) määrittelevät sosiaalialan arvoiksi ihmisarvon ja ihmisoikeuksien kunnioittamisen, itsemääräämisoikeuden kunnioittamisen, osallisuuden edistämisen, henkilön kokonaisvaltaisen huomioimisen sekä sosiaalisen oikeudenmukaisuuden. Nämä periaatteet muodostavat myös tutkimuksessani keskeisen eettisen lähtökohdan tekoälyn käyttöä arvioitaessa ja lähestymistapani tukee ammattieettisten ohjeiden sisältöä.

Lapset ovat ikänsä vuoksi erityisen suojelun kohteena, ja lastensuojelussa korostuvat lapsen edun ensisijaisuus sekä lapsen mahdollisuus osallistua ja vaikuttaa omissa asioissaan. Lastensuojelutyöhön liittyy myös viranomaisvaltaa, sillä sosiaalityöntekijä toimii

virkavastuulla esimerkiksi lapsen ja perheen tilanteen arvioinnissa sekä päätöksenteossa. (Malja ym. 2019.) Tämä korostaa tarvetta tarkastella kriittisesti myös sitä, millaisia vaikutuksia tekoälyn hyödyntämisellä voi olla lapsen ja perheen asemaan sekä heidän mahdollisuuksiinsa tulla kuulluiksi. Lapsen edun ja ihmisoikeuksien näkökulma kulkee mukana systemaattisesti koko tutkimusprosessin ajan.

Lähestymistapani tutkimusprosessissa on analyyttinen, ja tarkastelen aineistosta esiin nousevia havaintoja erityisesti niiden eettisten ulottuvuuksien näkökulmasta. Tavoitteeni ei ole normalisoida tekoälyn käyttöä lastensuojelussa, mutta en myöskään pyri esittämään sitä tarpeettoman kielteisessä valossa. Tutkijapositiooni ja käsityksiini voi vaikuttaa se, että olen suorittanut harjoitteluni lastensuojelussa ja työskennellyt sosiaalialan tehtävissä lastensuojelun kentällä. Kuitenkaan tekoälyn käytöstä lastensuojelussa tai sosiaalityössä minulla ei ole erityisesti kokemusta. Oma suhtautumiseni tekoälyn käyttöön lastensuojelussa on kaksijakoinen. Ajattelen, että tekoälystä voi olla hyötyä esimerkiksi kirjaamisessa ja muunlaisessa tietotyössä erityisesti aloittelevan sosiaalityöntekijän kohdalla. Suhtaudun kuitenkin kriittisesti tekoälyn käyttöön esimerkiksi asiakkaiden luokittelua vaativissa toiminnoissa.

Tekoäly ei näyttäydy julkisen sektorin, kuten lastensuojelun, organisaatioissa vain teknisenä ratkaisuna, vaan myös erilaisten diskurssien kautta, joissa tehokkuus, palvelun parantaminen ja hallinnollinen modernisaatio kietoutuvat julkisen vallan tavoitteisiin. Tällaiset diskurssit heijastavat paitsi taloudellisia tehokkuusintressejä myös laajempia valtasuhteita ja poliittisia prioriteetteja. (Hjaltalín & Sigurdarson, 2024.) Tutkielmaa tehdessäni tiedostan taustalla olevan poliittisen keskustelun. Eettisen katsantokantani pohjalla vaikuttaa määritelmä, jonka mukaan yhteiskunnallisen hyödyn tuottaminen asiakasryhmän hyvinvoinnin kustannuksella ei saa olla toiminnan tavoitteena (Peled & Leichtentritt 2002, 149–150).

Foucault’n (1980) näkemyksen mukaan tieto tuottaa valtaa ja valta tietoa. Tieto ei ole siis neutraalia, vaan se määrittää mitä pidetään totena ja arvokkaana. Focaultia peilaten on tärkeä tarkastella mitä ja miten dokumentoidaan sekä millaisia seurauksia dokumentoinnilla on. Kuvaan tutkimuksessa aineiston hakuprosessin läpinäkyvästi sekä perustelen, millä perusteilla aineistoa on haettu ja lopullinen aineisto valittu. Prosessin aikana pohdin myös sitä, onko kaikki relevantti tieto aiheesta saatavilla. Tekoälyn käytön eettisyyttä lastensuojelun kontekstissa kuvaavat artikkelit ovat yhtä lukuun ottamatta ulkomaisia eivätkä tarjoa suoria johtopäätöksiä Suomen tilanteeseen, sillä lastensuojelujärjestelmät ovat maakohtaisia (Pösö ym. 2019, 440). Huomioin, että tuottamallani tiedolla voi olla vaikutusta siihen, miten tekoälyn käyttöä

suomalaisessa lastensuojelussa lähestytään. Tutkimusprosessin lisäksi dokumentoinnissa erityisen tärkeää on lähteisiin viittaaminen asianmukaisesti. Erotan omat tulkintani lähteiden sisällöstä ja pidän tärkeänä mahdollisten ristiriitaisten näkemysten esille tuontia.

4. TEKOÄLYN EETTISET ULOTTUVUUDET LASTENSUOJELUSSA

Esittelen tutkimuksen tulokset jäsentämällä ne teemoihin, joiden kautta tekoälyn käytön eettisiä haasteita lastensuojelussa voidaan tarkastella. Osa aineistoesimerkeistä liittyy useampaan teemaan, minkä vuoksi teemojen välinen luokittelu ei ole yksiselitteinen eikä ehdoton. Teemoittelu on kuitenkin toiminut analyysin välineenä ja mahdollistanut aineistossa esiin nousseiden eettisten kysymysten systemaattisen tarkastelun. Jäsensin aineiston analyysin perusteella tulokset viiteen pääluokkaan, joiden avulla tarkastelen tekoälyn käyttöön lastensuojelussa liittyviä eettisiä kysymyksiä. Pääluokkien sisällä teksti etenee niin, että aluksi muodostan yleiskuvan aineistosta esiin nousseista teemoista painottaen erityisesti usein toistuvia havaintoja. Tämän jälkeen syvennän tarkastelua nostamalla esiin artikkeleiden välisiä eroja sekä yksityiskohtaisempia, eettisestä näkökulmasta merkittäviä, huomioita. Aineistossa nousi esiin myös useita näkemyksiä siitä, miten tekoälyn käyttöä voidaan kehittää eettisesti kestävämmäksi lastensuojelussa. Näiden havaintojen pohjalta muodostin kuudennen pääluokan, jossa tarkastelen eettisesti kestävän tekoälyn edellytyksiä. Ennen varsinaiseen analyysiin ja tutkimuksen tuloksiin siirtymistä esittelen aineiston lyhyesti.

4.1Aineiston kuvaus

Analyysi rakentuu tutkimusartikkeleista johdettujen teemojen perusteella. Tutkimusartikkelit koostuvat viidestä empiirisestä alkuperäistutkimuksesta, kahdesta teoreettisesta artikkelista sekä viidestä kirjallisuuskatsauksesta. Artikkeleista suurin osa on pohjoisamerikkalaisia: neljä yhdysvaltalaista ja yksi Yhdysvaltoihin ja Kanadaan sijoittuva. Kaksi artikkelia on brittiläistä. Yksittäisiä artikkeleita oli Alankomaista, Australiasta, Etelä-Koreasta, Suomesta ja Tanskasta.

Empiirisissä tutkimuksissa tarkastellaan tekoälyn ja algoritmisten järjestelmien käyttöä lastensuojelussa erityisesti riskien arvioinnin, ennakoivan päätöksenteon ja ammatillisen harkinnan tukemisen näkökulmista. Tutkimuksissa tarkastellaan myös sosiaalityöntekijöiden

kokemuksia ja näkemyksiä tekoälyn käytöstä sekä siihen liittyvistä hyödyistä, rajoituksista ja eettisistä kysymyksistä. Tutkimusten kontekstit vaihtelevat eri maiden ja lastensuojelujärjestelmien välillä, mikä tuo aineistoon näkökulmia tekoälyn hyödyntämisestä erilaisissa toimintaympäristöissä.

Kimin ja Yoonin (2026) tutkimus tarkastelee Etelä-Korean lastensuojelun asiantuntijoiden näkemyksiä tekoälyä hyödyntävän big datan käytöstä ja siihen liittyvistä eettisistä kysymyksistä. Wilkinsin ja Benettin (2025) tutkimuksessa verrataan ChatGPT:n ja sosiaalityöntekijöiden arvioita lastensuojelun päätöksenteossa sekä tarkastellaan tekoälyn roolia ammatillisen arvioinnin tukena. Ratnerin ja Schrøderin (2024) tutkimus puolestaan tarkastelee algoritmisten mallien kehittämiseen ja käyttöön liittyviä julkisia kiistoja sekä eettisiä kysymyksiä Tanskan lastensuojelussa. Lehtiniemen (2024) tutkimuksessa tarkastellaan sosiaalityöntekijöiden kokemuksia ennakoivasta tekoälystä sekä siihen liitettyjä odotuksia ja merkityksiä. Kawakamin ja kumppaneiden (2022) tutkimuksessa tarkastellaan sosiaalityöntekijöiden kokemuksia tekoälyavusteisesta päätöksenteosta ja erityisesti Allegheny Family Screening Tool (AFST) -järjestelmän käyttöön liittyviä näkemyksiä.

Agbanan (2025) sekä Landaun ja kumppanien (2022) tutkimukset edustavat teoreettisia eettisiä analyyseja, joissa tarkastellaan tekoälyn ja koneoppimisen käyttöön liittyviä eettisiä haasteita lastensuojelun ja riskinarvioinnin konteksteissa. Agbanan (2025) katsausartikkeli käsittelee ennakoivan tekoälyn hyödyntämiseen liittyviä eettisiä haasteita lastensuojelun riskinarvioinnissa sekä korostaa ammatillisen harkinnan, läpinäkyvyyden ja vastuullisen hallinnan merkitystä. Landau ym. (2022) tarkastelevat artikkelissaan koneoppimiseen perustuvien lapsen pahoinpitelyn ja laiminlyönnin tunnistamismallien kehittämiseen liittyviä eettisiä haasteita sekä esittävät suosituksia niiden vastuulliseen suunnitteluun ja arviointiin.

Kirjallisuuskatsauksissa on yksi kartoittava katsaus ja neljä systemaattista katsausta. Salmisen (2011) ja Vilkan (2013) määritelmiin peilaten kartoittavassa kirjallisuuskatsauksessa on yhtäläisyyksiä integratiiviseen kirjallisuuskatsaukseen. Bullin, Kiselyn, Bettsin ja Hunin (2025) systemaattinen kirjallisuuskatsaus arvioi lastensuojelussa käytettävien koneoppimismallien suorituskykyä sekä niiden oikeudenmukaisuuteen, läpinäkyvyyteen ja harhariskeihin liittyviä kysymyksiä. Gabriels, Kuiper ja Harder (2025) toteuttivat kartoittavan kirjallisuuskatsauksen, jossa tarkastellaan algoritmipohjaisia päätöksentukityökaluja sekä niiden hyötyjä, rajoituksia ja eettisiä haasteita lastensuojelussa. Yun ja Rosen (2025)

systemaattinen kirjallisuuskatsaus tarkastelee algoritmiavusteisten päätöksenteon tukityökalujen käyttöönottoa lastensuojelussa sekä niiden eettisiä haasteita ja oikeudenmukaisuutta. Saxena ja Guha (2024) toteuttivat katsauksen, jossa tutkitaan algoritmisten järjestelmien vaikutuksia lastensuojelun käytäntöihin, organisaatioihin ja päätöksentekoon sekä niiden aiheuttamia eettisiä ja prosessuaalisia haittoja. Hallin, Sagen, Scottin ja Josephin (2024) systemaattinen kirjallisuuskatsaus arvioi koneoppimiseen perustuvien päätöksentekotyökalujen suorituskykyä sekä niiden eettisiä, yhdenvertaisuuteen ja vinoumiin liittyviä, haasteita lastensuojelussa.

4.2 Epäoikeudenmukaisuus ja eriarvoisuuden vahvistuminen

Tarkasteltavien artikkelien perusteella oikeudenmukaisuus ja yhdenvertaisuuden haasteet näyttäytyvät yhtenä tekoälyn käytön merkittävimmistä eettisistä kysymyksistä lastensuojelussa. Artikkeleissa korostettiin, että tekoälyjärjestelmät eivät ole arvoneutraaleja, vaan niiden toiminta perustuu dataan, joka voi sisältää historiallisia ja rakenteellisia vinoumia (Kim & Yoon 2026; Lehtiniemi 2010; Agbana 2025). Tämän seurauksena tekoäly voi ylläpitää tai vahvistaa olemassa olevia eriarvoisuuksia ja vaikuttaa epäoikeudenmukaisesti lasten ja perheiden arviointiin.

Artikkelit kuvasivat kontekstiin liittyviä vinoumia. Landaun ym. (2022) mukaan esimerkiksi rodullisiin vähemmistöihin kuuluvien perheiden kohdalla pahoinpitelyn ja laiminlyönnin riski voi korostua perusteettomasti. Kimin ja Yoonin (2026) artikkelissa osallistujat arvioivat, että laajoihin tietoaineistoihin perustuvat ennustemallit eivät välttämättä kykene tunnistamaan korealaiselle yhteiskunnalle ominaisia kulttuurisia erityispiirteitä. Sen mukaan kulttuuristen tekijöiden puutteellinen huomioiminen voi heikentää kaltoinkohtelun tunnistamista ja johtaa tilanteisiin, joissa lasten ja perheiden olosuhteita arvioidaan heidän kulttuurinen kontekstinsa sivuuttaen. Tällöin tekoälyyn perustuvat arviot voivat tuottaa epäoikeudenmukaisia lopputuloksia ja asettaa eri ryhmät eriarvoiseen asemaan.

Kimin ja Yoonin (2026) mukaan vinouma ilmeni myös valikoitumisongelmana: esimerkiksi yksityisissä oppilaitoksissa opiskelevat tai ulkomailla asuvat lapset saattoivat jäädä järjestelmän ulkopuolelle, mikä vääristää riskiväestön määrittelyä ja heikentää ennusteiden luotettavuutta.

Ratner ja Schrøder (2024) puolestaan tunnistivat ikään liittyvä vinouman, joka johti oletukseen laiminlyönnin vakavuuden kasvamisesta lapsen iän myötä.

Artikkeleissa korostui historiallisten tietojen käyttöön liittyvät eettiset haasteet. Tekoälymallit rakentavat ennusteensa aikaisemman datan perusteella, minkä vuoksi menneisyyden tapahtumat voivat saada arvioinnissa liian suuren merkityksen. Vaikka asiakkaan aiemmat tiedot voivat tarjota arvioinnin kannalta merkityksellistä taustatietoa, niiden perusteella ei voida tehdä suoria johtopäätöksiä tulevista tapahtumista tai riskeistä oikeudenmukaisesti. Lehtiniemen (2022, 1118) mukaan eräs sosiaalityöntekijä kuvasi tekoälyn vaikutusta omaan ajatteluunsa.

> It felt stigmatizing to sift through that client history. My method is to meet clients without looking at their files too closely because that sets up prejudices. I felt [AI] was feeding my prejudice.

Kommentissa korostuu näkemys siitä, että menneisyyden liiallinen painottaminen voi ohjata arviointia ennakkokäsitysten suuntaan ja heikentää asiakkaan nykyisen tilanteen huomioimista. Tämän vuoksi tekoälyn tuottamaa tietoa ei tule tarkastella irrallaan asiakkaan kokonaistilanteesta, vaan ammattilaisen on arvioitava kriittisesti, kuinka hyvin järjestelmän tunnistamat riskit vastaavat asiakkaan tämänhetkistä tilannetta.

Kirjallisuuskatsauksien mukaan vinoumat voivat olla sisäänrakennettuina käytettyyn dataan ja siten välittyä algoritmien toimintaan. Saxenan ja Guhan (2024) katsauksessa tuotiin esiin, että aikaisempaan dataan perustuvissa arvioinneissa voi olla vinoumia, jotka ovat muodostuneet aiempien työntekijöiden toiminnan seurauksena. Huonosti suunnitellut työkalut vaarantavat oikeudenmukaisen ja yhdenvertaisen eettisen päätöksenteon. Joissakin katsauksien esittelemissä tekoälymalleissa oikeudenmukaisuus oli huomioitu jo tekoälymallien suunnitteluvaiheessa tavoitteena kehittää algoritmeja, jotka tuottavat mahdollisimman yhdenvertaisia ennusteita ilman syrjiviä vaikutuksia. Esimerkiksi Gabrielsin ym. (2025) mukaan rodullista vinoumaa pyrittiin vähentämään poistamalla tekoälymallista etnisyyttä kuvaava muuttuja. Artikkelissa kuitenkin todettiin, ettei tällainen ratkaisu välttämättä poista vinoumaa, sillä myös muut muuttujat, kuten asuinalue tai perheen tulotaso, voivat välittää samoja rakenteellisia eroja ja johtaa edelleen epäoikeudenmukaisiin ennusteisiin.

Osan kirjallisuuskatsauksien perusteella oikeudenmukaisuutta lisäävät ratkaisut voivat joissakin tilanteissa heikentää mallin ennustetarkkuutta, vaikka ratkaisut samalla vähentäisivät syrjiviä vaikutuksia (Gabriels ym. 2025; Hall ym. 2023). Hallin ym. (2023) katsauksessa kuitenkin todettiin, ettei algoritmin hyvää suorituskykyä tule pitää osoituksena sen eettisyydestä tai yhdenvertaisuudesta. Jos koulutusaineisto sisältää rakenteellisia vinoumia, algoritmi voi oppia ennustamaan niitä tehokkaasti ja samalla ylläpitää tai vahvistaa institutionaalisia eriarvoisuuksia.

Muutamassa artikkelissa nostettiin esille myös se, ettei tekoälyn arvioita tule verrata virheettömään inhimilliseen päätöksentekoon (Hall ym. 2023; Wilkins & Benett 2025). Sosiaalityötekijöiden arviot eivät ole vapaita ennakkokäsityksistä, kokemukseen perustuvista vinoumista eikä inhimillisistä virheistä. Näin ollen sekä tekoälyn että ihmisen tekemään arviointiin liittyy eettisiä riskejä, minkä vuoksi niiden välistä suhdetta ei voida tarkastella yksinkertaisena vastakkainasetteluna. Hallin ym. (2024) tarkastelemassa tutkimuksessa havaittiin, että lastensuojelun ilmoituspuhelimissa työskentelevät työntekijät ohittivat algoritmin suosituksia samankaltaisesti eri riskitasoilla. Tulosten perusteella päätöksiin vaikuttivat usein enemmän työntekijöiden omat arviot kuin algoritmin tuottamat riskiluokitukset, mikä osoittaa, että myös inhimilliseen päätöksentekoon liittyy subjektiivisuutta ja mahdollisia vinoumia. Kawakamin ym. (2022) tutkimuksessa nousi esille mielenkiintoinen näkökulma siitä, että tekoäly voi myös auttaa tunnistamaan työntekijöiden omia tiedostamattomia ennakkokäsityksiä. Näin ollen tekoälyllä nähtiin potentiaalia juuri oikeudenmukaisuuden lisääjänä.

4.3 Lapsen ja perheen tilanteen virheellinen tulkinta

Lasten ja perheiden tilanteiden virheellinen tulkinta lastensuojelussa on merkittävä eettinen kysymys. Tekoäly ei artikkelien perusteella kykene huomioimaan kaikkia yksittäisiin asiakastilanteisiin liittyviä erityispiirteitä ja kontekstuaalisia tekijöitä, minkä vuoksi sen tuottamat arviot voivat jäädä puutteellisiksi tai virheellisiksi. Virheellisten ja puutteellisten arvioiden näkökulmasta korostui sosiaalityön inhimillisen harkinnan merkitys. Inhimillisen harkinnan puute voi johtaa tilanteisiin, jossa asiakkaan tilannetta tulkitaan väärin ja näin ollen jatkotoimet ovat vääriä. (Ratner & Schrøder 2024.) Tämän vuoksi sosiaalityöntekijän rooli tekoälyn tuottaman tiedon tulkitsijana ja kriittisenä arvioijana korostui.

Tekoäly perustuu oletukseen, jonka mukaan samankaltaisilla asiakkailla on myös samankaltaiset riskit. Sosiaalityössä asiakkaiden tilanteet ovat kuitenkin yksilöllisiä, eikä riskejä voida arvioida luotettavasti pelkästään aiempien tietojen tai tilastollisten yhtäläisyyksien perusteella. Virheellisillä arvioilla voi olla lastensuojelussa vakavia seurauksia. Lapsen kaltoinkohtelun arvioinnissa väärä positiivinen arvio voi johtaa tarpeettomiin tutkimuksiin ja aiheuttaa perheelle ylimääräistä kuormitusta, kun taas väärä negatiivinen arvio voi merkitä sitä, että kaltoinkohtelu jää tunnistamatta ja lapsi jää ilman tarvitsemaansa suojelua. Vastaavasti kodin ulkopuolista sijoitusta koskevissa arvioissa lapsi voidaan joko sijoittaa tarpeettomasti tai jättää vaarallisiin olosuhteisiin. (Agbana 2025; Lehtiniemi 2010; Landau ym. 2022.)

Erityisesti ennakoivien algoritmien rajoitteet korostuivat monimutkaisissa ja muuttuvissa elämäntilanteissa, joissa arviointi edellyttää tilanteen laaja-alaista ymmärrystä ja ammatillista harkintaa. Hallinnolliseen dataan ja yksittäisiin indikaattoreihin perustuvat mallit eivät välttämättä kykene tunnistamaan monimutkaisia ilmiöitä, kuten lapsiin kohdistuvaa kaltoinkohtelua. Wilkins ja Benett (2025, 709) kuvasivat ChatGPT-tekoälymallia

> ChatGPT does not engage in mathematical calculations (perhaps most social workers do not either) or deep understanding (which social workers can and do). Rather, it creates text responses by identifying patterns in the data it has been exposed to, much like a parrot mimics phrases without comprehending their meaning.

Tutkijat korostavat, ettei tekoälyn tuottama kielellisesti johdonmukainen vastaus merkitse tilanteen syvällistä ymmärtämistä. Lastensuojelussa arviointi edellyttää syvällistä tiedon tulkitsemista suhteessa lapsen ja perheen yksilölliseen tilanteeseen ja sen kontekstiin. Kimin ja Yoonin (2026) tutkimuksessa tuotiin esiin, että monet tekoälymallit perustuvat ensisijaisesti määrällisiin indikaattoreihin eivätkä kykene huomioimaan esimerkiksi parisuhdekonflikteja tai kulttuurisia erityispiirteitä, jotka voivat olla olennaisia lapsen tilanteen ymmärtämisessä. Datan edustavuus on tässä keskeistä, sillä rajoittuneisiin aineistoihin perustuvat mallit voivat toimia epätarkasti eri väestöryhmissä, mikä voi johtaa epäoikeudenmukaisiin lopputuloksiin (Agbana 2025). Wilkinsin ja Benettin (2025) artikkelin mukaan on vältettävä naiivia empirismiä, jonka perusteella arvioita voitaisiin tehdä yksinomaan havaittavissa olevan datan perusteella ilman teorian, kontekstin tai moraalisen päättelyn merkitystä. Sosiaalityöntekijät hyödyntävät

arvioinnissaan monenlaista tietoa, kuten asiakkaiden kertomuksia, aiempia tapahtumia, viranomaisten havaintoja sekä tilanteeseen liittyviä olosuhteita.

Artikkeleissa korostui tekoälyn ja sosiaalityön tiedonmuodostuksen välinen ero. Algoritminen järjestelmä voi käsitellä suuria tietomääriä ja tunnistaa niiden välisiä tilastollisia yhteyksiä, mutta tämä ei merkitse lapsen tai perheen tilanteen syvällistä ymmärtämistä. Sosiaalityöntekijän tekemä arvio perustuu määrällisen tiedon ohella asiakkaiden kertomuksiin, vuorovaikutuksessa muodostuvaan tietoon, tilanteen kontekstin huomioimiseen sekä sosiaalityöntekijän ammatilliseen harkintaan. Tämän vuoksi tekoälyn tuottamaa ennustetta tulee tarkastella yhtenä tietolähteenä muiden joukossa eikä lapsen tilanteen kokonaisarviona. Lamponen (2022, 107–109) tuo esiin, että lastensuojelun päätöksenteossa hyödynnetään useita erilaisia tiedon muotoja, kuten faktatietoa, intuitiivista tietoa ja kokemustietoa, joita sosiaalityöntekijä suhteuttaa toisiinsa päätöksenteossa. Tekoälyn tuottama tieto voidaan siten nähdä yhtenä tiedon muotona, mutta sen ei tulisi saada automaattisesti muita tiedon muotoja korkeampaa asemaa. Erityisen tärkeäksi muodostuu asiakkaan ja lapsen oman tiedon säilyminen osana arviointia. Lamponen (166–167) osoittaa, että lapsen tiedolla on lastensuojelun päätöksenteossa erityinen merkitys ja että erilaisilla tiedon lähteillä on päätöksenteossa erilainen painoarvo. Jos algoritminen tieto alkaa määrittää sitä, mitä pidetään merkityksellisenä tietona asiakkaan tilanteesta, teknologia ei enää ainoastaan tue tiedonmuodostusta vaan osallistuu aktiivisesti sen rajaamiseen.

Lehtiniemen (2022) artikkelissa sosiaalityöntekijät korostivat, että asiakkailla on yksilöllisiä vahvuuksia, selviytymiskeinoja ja kuormitustekijöitä, joita ei voida päätellä luotettavasti algoritmisten riskiluokitusten perusteella. Työntekijöiden mukaan merkityksellinen riski rakentuu aina suhteessa perheen tilanteeseen, eikä näennäisesti samanlaisissa olosuhteissa elävien lasten tai perheiden tilanteita voida arvioida samalla tavalla. Havainto kytkeytyy sosiaalityön person-in-situation-ajatteluun, jossa yksilön tilannetta tarkastellaan suhteessa hänen sosiaaliseen ympäristöönsä ja elämäntilanteeseensa sen sijaan, että sitä selitettäisiin yksinomaan yksilön henkilökohtaisilla ominaisuuksilla. Sosiaalityöntekijöiden mukaan tekoäly ei kykene huomioimaan tätä kontekstuaalista ulottuvuutta samalla tavalla kuin ammattilainen. Kawakamin ym. (2022) artikkelissa esitettiin, että AFST-järjestelmää käyttäneet työntekijät ja esihenkilöt kuvasivat tilanteita, joissa korkea riskipisteytys heijasti pikemminkin järjestelmän kyvyttömyyttä huomioida asiakkaan tilanteeseen liittyvää merkityksellistä kontekstia kuin

todellista lapsen turvallisuusriskiä. Tämän vuoksi algoritmin antamia arvioita jouduttiin usein täydentämään laadullisella tiedolla ja ammatillisella harkinnalla.

Kirjallisuuskatsauksissa tarkasteltiin, kuinka tarkasti tekoäly- ja ennustemallit pystyvät ennustamaan arviointien tai päätösten lopputuloksia sekä millaisia virheitä ennusteisiin voi liittyä. Bullin ym. (2025) mukaan virheiden riski säilyy myös silloin, kun mallien ennustetarkkuus on korkea. Gabrielsin ym. (2025) katsauksessa tarkasteltiin AFST-järjestelmän vaikutusta päätöksiin, joissa arvioidaan, mitkä tapaukset voidaan rajata pois jatkoarvioinnista. Tutkimuksessa havaittiin, että järjestelmän käyttö heikensi jonkin verran näiden screen-out-päätösten tarkkuutta. Wilkinsin ja Benettin (2025) katsauksessa puolestaan havaittiin, että ChatGPT oli arvioinneissaan selvästi epätarkempi kuin parhaiten suoriutunut sosiaalityöntekijä. Näiden katsausten perusteella siis tekoäly- ja ennustemallien ennustetarkkuus vaihteli, ja korkeastakin ennustetarkkuudesta huolimatta niiden käyttöön liittyi virheiden riski sekä päätösten tarkkuuteen liittyviä rajoituksia.

4.4 Ammatillisen harkinnan kaventuminen ja vastuun hämärtyminen

Tarkasteltavissa artikkeleissa ammatillisen harkinnan kaventuminen ja vastuun hämärtyminen kytkeytyivät erityisesti siihen, millaiseksi sosiaalityöntekijän ja algoritmisen päätöksenteon tukijärjestelmän välinen suhde muodostuu (Agbana 2025; Gabriels ym. 2025; Kawakami ym. 2022). Tekoälyn ja algoritmisten järjestelmien käytön ei katsottu itsessään merkitsevän ammatillisen harkinnan korvautumista, mutta niiden tuottamat arviot voivat alkaa ohjata päätöksentekoa tavalla, joka kaventaa sosiaalityöntekijän mahdollisuuksia tehdä itsenäistä ja kokonaisvaltaista arviota lapsen tilanteesta. Keskeiseksi riskiksi aineistossa nousi automaatioharha, jolla tarkoitetaan ammattilaisen taipumusta antaa algoritmin tuottamalle tiedolle ja suosituksille suurempi painoarvo kuin omalle ammatilliselle harkinnalleen. Tällöin algoritmin tuottama arvio voi saada päätöksenteossa aseman, joka ylittää sen tarkoituksen päätöksenteon tukena.

Algoritmisten arvioiden merkitys ei kuitenkaan rajoitu siihen, miten sosiaalityöntekijä käyttää niiden tuottamaa tietoa päätöksenteossa. Sosiaalityön arviot ovat myös keinoja rakentaa sosiaalista todellisuutta, sillä ne voivat vaikuttaa siihen, miten lapsen ja perheen tilanne ymmärretään ja millaisiin toimenpiteisiin tilanteessa päädytään. Arvioihin liittyy siten myös

moraalista harkintaa siitä, mikä nähdään vääränä ja milloin tilanteeseen tulisi puuttua. Tällaiset arviot eivät ole yksinkertaisia tai pysyviä kuvauksia todellisuudesta, vaan niillä voi olla vaikutusta ihmisiä koskeviin käsityksiin ja toimintaan. (Wilkins & Benett 2025.) Tekoälyn tuottama analyysi voi vahvistaa tätä vaikutusta, jos sen perusteella muodostetuille arvioille annetaan ammatillisessa päätöksenteossa suuri painoarvo.

Gabrielisin ym. (2025) katsauksessa koneoppimismallien heikko tulkittavuus tunnistettiin keskeiseksi haasteeksi, sillä erityisesti monimutkaisten mallien toimintaperiaatteita voi olla vaikea ymmärtää ja niiden tuottamia ennusteita arvioida kriittisesti. Jos sosiaalityöntekijä ei kykene ymmärtämään, millä perusteella järjestelmä on muodostanut tietyn riskiluokituksen tai ennusteen, algoritmin tuottamaa tulosta voi olla vaikea arvioida kriittisesti. Tällöin vaarana on, että algoritminen arvio näyttäytyy objektiivisena ja kiistattomana tietona sen sijaan, että sitä tarkasteltaisiin yhtenä päätöksenteossa hyödynnettävänä tietolähteenä. Agbana (2025) nosti esille sen, kuinka läpinäkyvyyden ja selitettävyyden puute siten välillisesti vahvistaa automaatioharhaa ja vaikeuttaa ammatillisen harkinnan käyttöä. Niin sanottu musta laatikko -mallin läpinäkymättömyys korostuu erityisesti korkean riskin tilanteissa (mt., 1653).

> Many advanced AI models, especially deep learning architectures, operate as “black boxes,” producing outputs without clear insight into the decision-making process. This opacity can undermine stakeholder confidence and complicate the validation of predictions, particularly in high-stakes contexts.

Selitettävyyden merkitys korostuu lastensuojelussa, jossa algoritmisten arvioiden perusteella tehdyt ratkaisut voivat vaikuttaa merkittävästi lapsen ja perheen elämään. Artikkeleissa läpinäkyvyyttä tarkasteltiin tästä syystä myös edellytyksenä sille, että algoritmin tuottamaa tietoa voidaan perustellusti käyttää ammatillisen päätöksenteon tukena. Tekoälymallin tulisi tarjota riittävästi tietoa siitä, miten tiettyyn arvioon on päädytty, jotta sosiaalityöntekijä voisi arvioida sen merkitystä, kyseenalaistaa sen tarvittaessa ja suhteuttaa sen muuhun asiakasta koskevaan tietoon. (Agbana 2025.) Kawakamin ym. (2022) tutkimuksessa työntekijät toivoivat AFST:n tarjoavan pelkän riskipistemäärän sijaan tietoa, joka auttaisi päätösten perustelemisessa niille tahoille, joihin päätökset vaikuttavat. Selitettävyyden puute ei näin ollen kosketa ainoastaan sosiaalityöntekijän mahdollisuutta ymmärtää algoritmia, vaan myös päätöksenteon perusteltavuutta ja sen avoimuutta asiakkaille sekä muille päätöksentekoon ja valvontaan osallistuville tahoille.

Kawakamin ym. (2022) tutkimuksessa ammatillisen harkinnan kaventuminen ilmeni myös konkreettisesti AFTS:n käyttöön liittyvinä organisatorisina paineina. Osa työntekijöistä koki, että tekoälyn pisteytyksestä poikkeamista seurattiin organisaatiossa ja että työntekijöihin kohdistui odotus järjestelmän suositusten noudattamisesta. Tämän seurauksena työntekijät saattoivat toimia algoritmin suosituksen mukaisesti myös silloin, kun se oli ristiriidassa heidän oman ammatillisen arvionsa kanssa. Työntekijät kuvasivat ristiriitaa oman päätöksentekonsa lähtökohtien ja AFST:n ennustetavoitteiden välillä: sosiaalityöntekijät painottivat lapsen välitöntä turvallisuutta, kun taas järjestelmä pyrki ennustamaan haitallisten lopputulosten riskiä pidemmällä aikavälillä. Osa työntekijöistä suhtautui tästä syystä kriittisesti algoritmisen ennusteen merkitykseen omassa työssään, kun taas toiset pitivät pitkän aikavälin riskin huomioimista periaatteessa hyödyllisenä, mutta epäselvänä sen suhteen, miten tietoa tulisi käytännössä hyödyntää.

Kawakamin ym. (2022) tutkimus toi esiin myös sen, ettei algoritmin vaikutus ammatilliseen harkintaan välttämättä ole työntekijälle itselleen tietoinen tai suora. Useat työntekijät kokivat tekoälyn roolin vähäiseksi eivätkä katsoneet sen muuttavan omia päätöksiään, mutta tutkimuksessa havaittiin järjestelmän vaikuttavan päätöksentekoon ainakin epäsuorasti organisaation asettamien odotusten kautta. Algoritmin vaikutus ei siten rajoitu tilanteisiin, joissa työntekijä pitää järjestelmää luotettavana, vaan sitä voivat vahvistaa myös organisaation seuranta- ja arviointikäytännöt. Aineisto osoittaa, että ammatillista harkintaa voivat kaventaa myös käytännölliset tekijät. AFST:n suosituksesta poikkeaminen edellytti kirjallisia perusteluja, mikä koettiin työlääksi ja hyödyttömäksi. Järjestelmän suositusten noudattaminen saattoi siten olla työntekijälle helpompi vaihtoehto kuin niiden kyseenalaistaminen.

Ammatillisen harkinnan ja algoritmisen tiedon välinen suhde kytkeytyi aineistossa ammatillisen harkinnan lisäksi laajemmin vastuukysymykseen. Tekoälyyn perustuvassa päätöksenteossa vastuu ei kohdistu yksiselitteisesti yhteen toimijaan, vaan päätöksentekoon voivat liittyä sosiaalityöntekijän lisäksi järjestelmän kehittäjät ja työnantajaorganisaatio. Jos algoritmin tuottama suositus johtaa haitalliseen lopputulokseen, voi olla vaikeaa määrittää, missä määrin vastuu kuuluu yksittäiselle ammattilaiselle, järjestelmän kehittäjälle tai organisaatiolle. Aineistossa vastuuta liitettiin esimerkiksi tilanteisiin, joissa haitallinen lopputulos johtuu virheellisestä algoritmista, puutteellisesta käyttäjien koulutuksesta tai siitä,

ettei järjestelmän rajoituksista ole annettu riittävästi tietoa. (Agbana 2025.) Lisäksi algoritmisten ennusteiden tulkinta voi synnyttää uudenlaisia eettisiä ja oikeudellisia kysymyksiä erityisesti tilanteissa, joissa ammattilainen päättää olla noudattamatta algoritmin ennustetta (Landau ym. 2025). Tällöin ongelmaksi muodostuu, kuka vastaa algoritmin virheellisestä arviosta sekä miten ammattilaisen omaa ratkaisua suhteessa algoritmin suositukseen tulisi perustella ja arvioida.

Sosiaalityöntekijän ammatillinen harkinta on keskeistä eettisessä arvioinnissa ja päätösten seurauksista kannettavan vastuun näkökulmasta. Lastensuojelussa arvioidaan samanaikaisesti lapsen etua, perheen näkemyksiä, lainsäädännön velvoitteita ja päätösten mahdollisia seurauksia. Asiakkaan tilanteen ymmärtäminen edellyttää yksilöllisten olosuhteiden, kokemusten ja merkitysten huomioimista, minkä vuoksi sosiaalityön asiantuntijuutta ei voida sitoa yksinomaan standardoituihin arviointimenetelmiin tai teknisiin ratkaisuihin. Keskeisiksi jäävät edelleen ammatillinen harkinta, eettinen arviointi ja vastuu päätösten seurauksista. Ammatillisen harkinnan merkitys korostuu erityisesti tilanteissa, joissa tekoälyn tuottama arvio ohjaa lastensuojelullisen huolen määrittymistä. Suomen lain mukaan lastensuojelun tulee olla mahdollisimman hienovaraista ja avohuollon tukitoimet ovat ensisijaisia. Sijaishuolto tulee kuitenkin järjestää viipymättä silloin, kun lapsen etu sitä edellyttää. (Räty 2026, 8–9.) Tekoälyn käytössä tähän liittyy erityinen jännite. Jos tekoäly tuottaa lapsesta tai perheestä korkean riskin arvion, voi tieto parhaimmillaan auttaa ammattilaista havaitsemaan sellaisen tuen tai suojelun tarpeen, joka muuten jäisi huomaamatta. Toisaalta korkea riskiluokitus voi myös johtaa siihen, että lapseen tai perheeseen kohdistetaan aiempaa voimakkaampia toimenpiteitä kuten sijaishuollon käynnistämistä, ennen kuin ammattilainen on arvioinut tilanteen kokonaisuudessaan.

Vaikka myös inhimilliseen päätöksentekoon liittyy mahdollisuus virheisiin, vastuu päätöksestä on tällöin selkeämmin ihmisellä. Tekoälyn hyödyntäminen päätöksenteossa puolestaan voi tehdä vastuun kohdentumisesta haastavampaa, mikä korostaa ammattilaisen tosiasiallisen päätösvallan ja vastuun säilymisen merkitystä. Artikkeleista muodostuu kokonaisuutena kuva tilanteesta, jossa tekoälyn eettinen ongelmallisuus ei liity yksinomaan algoritmin kykyyn tuottaa virheellisiä tai vinoutuneita arvioita, vaan myös siihen, miten sen tuottama tieto sijoittuu osaksi sosiaalityön ammatillisia ja organisatorisia käytäntöjä. Algoritmi voi kaventaa ammatillista harkintaa sekä suoran luottamuksen että epäsuorien organisatoristen paineiden, läpinäkymättömyyden ja epäselvien toimintatapojen kautta. Samalla vastuun säilyttäminen

ihmisellä edellyttää, että sosiaalityöntekijällä on riittävät tiedot, mahdollisuus kyseenalaistaa algoritmin tuottama arvio sekä tosiasiallinen mahdollisuus poiketa siitä. Näin ollen aineistossa korostuu tarve sellaiselle ihmisen ja tekoälyn väliselle työnjaolle, jossa algoritminen tieto tukee ammatillista harkintaa, mutta ei määritä sen lopputulosta.

4.5 Lapsen tarpeisiin vastaamisen ja tuen toteutumisen vaarantuminen

Osassa artikkeleista tuotiin esiin, että tekoälyn kyky tunnistaa palvelujen tai tuen tarpeessa olevia lapsia ei vielä itsessään takaa sitä, että tunnistettuihin tarpeisiin pystytään käytännössä vastaamaan (Gabriels ym. 2025; Kim & Yoon 2026; Saxena & Guha 2024). Tekoäly voi tuottaa ennusteita ja tunnistaa riskissä olevia lapsia, mutta ennusteiden hyödyntäminen edellyttää toimivaa palvelujärjestelmää, riittäviä resursseja ja mahdollisuutta tarjota lapselle hänen tilanteeseensa soveltuvaa tukea. Mikäli ennusteiden ja palvelujen välille jää katkos, vaarana on, että tekoälyllä tunnistettu tarve ei johda konkreettisiin toimenpiteisiin. Tällöin tekoälyn hyödyntäminen ei välttämättä vahvista lasten saamaa tukea, vaan voi päinvastoin johtaa tilanteeseen, jossa tunnistettuihin tarpeisiin ei pystytä vastaamaan. Artikkeleissa tämä tuli esiin erityisesti tekoälyjärjestelmän ja olemassa olevien palvelurakenteiden välisenä yhteensopimattomuutena.

Kimin ja Yoonin (2026, 6) tutkimuksessa osallistujat toivat esiin, ettei ennustejärjestelmä yksinään riitä, mikäli sitä ei ole integroitu olemassa olevaan palvelujärjestelmään.

> As Participant D points out, ‘The effectiveness of the e-System is being questioned due to the lack of integration with existing services,’ highlighting that despite the potential for better outcomes, the failure to link policies with services limits their impact.

Järjestelmän vaikuttavuutta rajoitti erityisesti se, ettei ennusteiden ja käytännön palvelujen välillä ollut toimivaa yhteyttä. Ilman toimivaa palvelurakennetta ennusteita ei voida siirtää tehokkaasti käytännön toimintaan, vaikka niiden tarkkuus olisi hyvä. Tämä osoittaa, että tekoälyn tuottama tieto ei itsessään takaa lapselle oikea-aikaista tai tarkoituksenmukaista tukea. Palvelujen ja ennusteiden välisen yhteyden puute korostui myös korkean riskin lapsille suunnattujen seuranta- ja ennaltaehkäisevien palvelujen kohdalla. Aineistossa kuvattiin tilanteita, joissa ennustemallit pystyivät tunnistamaan lapsia, joilla arvioitiin olevan kohonnut

riski päätyä myöhemmin lastensuojelun asiakkaiksi. Tälle tiedolle ei kuitenkaan ollut vakiintuneita ennaltaehkäiseviä toimintamalleja, joiden avulla tunnistettuun riskiin olisi voitu puuttua. Edellä mainitussa tilanteessa tekoäly kykenee tunnistamaan mahdollisen ongelman, mutta järjestelmä ei pysty vastaamaan siihen ennen kuin ongelma mahdollisesti toteutuu. Eettisesti ongelmalliseksi muodostuu siten tunnistetun tarpeen ja siihen vastaamisen välinen kuilu. Pelkkä riskin tunnistaminen ei palvele lasta, jos tunnistaminen ei johda oikea-aikaiseen tukeen tai palveluun.

Joissakin artikkeleissa nousi esiin myös palvelukapasiteetin merkitys tekoälyn tuottamien suositusten toteutumiselle. Saxenan ja Guhan (2024) katsauksessa tuotiin esille se, että algoritmisista suosituksista huolimatta sijoituspäätöksiin vaikuttivat merkittävästi käytettävissä olevat palvelut ja resurssit. CANS-algoritmin tuottamia sijaisvanhemmille suunnattuja rahoitus- ja hoitomuotosuosituksia toteutettiin harvoin laadukkaiden sijaiskotien puutteen vuoksi. Tämä osoittaa, ettei algoritminen päätöksenteon tuki pysty ratkaisemaan palvelujärjestelmän rakenteellisia puutteita. Jos lapsen tarpeisiin soveltuvia palveluja ei ole saatavilla, tekoälyn tuottama suositus ei välttämättä johda lapsen tilanteen kannalta tarkoituksenmukaiseen ratkaisuun.

Puutteellinen vastaaminen asiakkaan tarpeisiin näkyi aineistossa myös henkilöstöresurssien näkökulmasta. Vaikka tekoäly voisi auttaa tunnistamaan ja priorisoimaan tukea tarvitsevia lapsia, henkilöstöpula ja työntekijöiden rajallinen kapasiteetti rajoittavat mahdollisuuksia vastata tunnistettuihin tarpeisiin. Aineistossa kuvattiin, että ennusteiden tuottaminen voi päinvastoin kasvattaa jo ennestään kuormittunutta työmäärää. Tällöin tekoälyn avulla tunnistettujen tapausten määrän kasvu ei automaattisesti tarkoita lasten saaman tuen lisääntymistä. Päinvastoin suuri tapausmäärä suhteessa käytettävissä olevaan henkilöstöön voi vaikeuttaa lasten tarpeiden perusteellista arviointia ja niihin vastaamista. (Saxena & Guha 2024.)

Tarkasteltavissa artikkeleissa korostui ristiriita tekoälyn tarjoaman tunnistamiskyvyn ja palvelujärjestelmän tosiasiallisen toimintakyvyn välillä. Tekoäly voi lisätä tietoa siitä, ketkä lapset saattavat tarvita tukea, mutta se ei itsessään lisää palvelujen, henkilöstön tai sopivien sijoituspaikkojen saatavuutta. Eettisesti ongelmalliseksi muodostuu tilanne, jossa lapsen tuen tarve tunnistetaan, mutta siihen ei pystytä vastaamaan käytännössä. Tällöin tekoälyn

hyödyntäminen ei välttämättä vähennä lasten ja perheiden haavoittuvuutta, vaan tunnistetun tarpeen ja toteutuvan tuen välinen kuilu voi jäädä ennalleen.

4.6 Yksityisyyden suoja, tietoturva ja valvonta

Tekoälyn käyttöön lastensuojelussa liittyy yksityisyyden suojaan, tietoturvaan ja valvontaan kytkeytyviä eettisiä kysymyksiä. Esimerkiksi ennakoiva tekoäly perustuu usein laajoihin tietoaineistoihin, joissa voidaan yhdistää terveyteen, koulutukseen, sosiaalipalvelujen käyttöön ja rikosoikeudellisiin seikkoihin liittyviä tietoja. Tällainen tietojen laaja-alainen kerääminen ja yhdistäminen lisää riskiä siitä, että lasten ja perheiden yksityisyyttä loukataan tai arkaluonteisia tietoja käytetään tarkoituksiin, joita lapset ja perheet eivät ole voineet ennakoida. Kimin ja Yoonin (2026) tutkimuksessa osallistujat toivat esiin huolen siitä, että hallinnollista dataa voidaan lainsäädännön nojalla käyttää lastensuojelukontekstissa ilman asiakkaan suostumusta.

Yksityisyyden suojan näkökulmasta keskeiseksi kysymykseksi nousi myös se, kuinka paljon tietoa lapsesta ja perheestä on eettisesti perusteltua kerätä ja yhdistää. Ratner ja Schrøder (2024) kuvaavat RISK-malliin liittynyttä oikeudellista keskustelua, jossa tarkasteltiin erityisesti sitä, millä perusteella ja missä laajuudessa lapsia koskevia tietoja voidaan kerätä ja hyödyntää algoritmisessa päätöksenteossa. Tähän liittyen artikkelissa (mt., 52) tuodaan esiin tietojen minimoinnin periaatteeseen liittyvä keskeinen jännite: lapsesta kerättävien tietojen tarpeellisuutta tulisi arvioida yksilöllisesti, mutta algoritminen tiedonkeruu perustuu erilaisten tietojen keräämiseen ja arviointiin.

> A main objection concerns the legal requirement to individually consider which information about a child to collect from a principle of data minimisation; a requirement rendered impossible by the standardised algorithmic collection and assessment of various data points.

Tietojen minimoinnin mukaan henkilötietoja tulisi kerätä ja käsitellä vain siinä määrin kuin se on käyttötarkoituksen kannalta tarpeellista. Erilaiset standardoidut tekoälymällit hyötyvät mahdollisimman laajoista tietoaineistoista. (Agbana 2025, Ratner & Schrøder 2024.) Hallin ym. (2023) katsauksen tulokset ovat linjassa yleisesti hyväksytyn käsityksen kanssa siitä, että suuremmilla otoskoilla koulutetut mallit kykenevät paremmin optimoimaan suorituskykyä.

Ennakoivassa mallinnuksessa tähän liittyy kuitenkin ristiriita, sillä mahdollisimman laajan muuttujajoukon käyttäminen voi parantaa mallin ennustetarkkuutta. Tällöin tekoälyn suorituskyvyn tavoittelu voi kannustaa keräämään ja yhdistelemään yhä enemmän tietoa lapsista ja perheistä, vaikka kaikkien tietojen tarpeellisuutta yksittäisen asiakkaan tilanteen näkökulmasta ei pystyttäisi perustelemaan. Tämä muodostaa jännitteen tekoälyn ennustetarkkuuden ja asiakkaan yksityisyyden suojan välille. (Agbana 2025, Ratner & Schrøder 2024.)

Tietoturvaan liittyvä eettinen ongelma puolestaan liittyy siihen, että laajoihin ja arkaluonteisiin aineistoihin kohdistuva luvaton pääsy voi aiheuttaa lapsille ja perheille merkittävää haittaa. Artikkeleissa nostettiin esille sekä kyberhyökkäysten että organisaation sisäisen väärinkäytön mahdollisuus. (Landau ym. 2022; Saxena & Guha 2024.) Tämän vuoksi tietojen turvallinen säilyttäminen, salaaminen ja käyttöoikeuksien rajaaminen ovat keskeisiä edellytyksiä tekoälyn vastuulliselle hyödyntämiselle. Tietoturvan merkitys korostuu erityisesti lastensuojelussa, jossa käsiteltävät tiedot voivat olla hyvin henkilökohtaisia ja niiden paljastumisella voi olla pitkäkestoisia vaikutuksia lapsen ja perheen asemaan. Tiedonjakokanavien tulee olla turvallisia ja toimintatapojen yhtenäisiä. Datan väärinkäytön riskien ehkäiseminen edellyttää tietoturvan lisäksi avointa keskustelua datan asianmukaisesta käytöstä sekä lainsäädäntöä, joka turvaa yksityisyyden suojan myös teknologian kehittyessä ja huomioi julkisen edun ja yksilön oikeuksien välisen tasapainon (Agbana 2025).

Yksityisyyden suojaan liittyvät kysymykset kietoutuivat artikkeleissa myös valvonnan ja perusoikeuksien toteutumisen kysymyksiin (Agbana 2025; Kim & Yoon 2026; Ratner & Schrøder 2024). Eri tietolähteiden yhdistäminen voi mahdollistaa aiempaa kattavamman kuvan muodostamisen lapsen tilanteesta, mutta samalla se voi lisätä kohdennetun ja epäoikeudenmukaisen valvonnan riskiä. Siksi tietojen yhdistämistä ei voida tarkastella ainoastaan teknisenä kysymyksenä, vaan sitä on arvioitava suhteessa lasten ja perheiden oikeuksiin, yksityisyyteen ja yhdenvertaiseen kohteluun. (Agbana 2025.) Kimin ja Yoonin (2026) tutkimuksessa osallistuja kritisoi e-järjestelmää siitä, että se voi perustua yleiseen epäilyyn ja johtaa perheiden ennakolliseen syyllistämiseen. Eettisesti ongelmalliseksi muodostuu tällöin tilanne, jossa lapsesta tai perheestä kerättyä tietoa käytetään ennakoimaan mahdollista tulevaa ongelmaa ennen kuin mitään konkreettista tapahtumaa on välttämättä tapahtunut. Laajamittainen tiedonkeruu voi siten muuttaa lastensuojelun toimintaa kohti ennakoivaa valvontaa, jossa perheitä tarkastellaan mahdollisten tulevien riskien eikä ainoastaan

heidän tämänhetkisen tilanteensa perusteella. Tietojen yhdistäminen tulisi toteuttaa tavalla, joka suojaa yksityisyyttä eikä mahdollista esimerkiksi tiettyihin väestöryhmiin kohdistuvaa epäoikeudenmukaista valvontaa. Yksityisyyden suojan näkökulmasta keskeiseksi eettiseksi kysymykseksi muodostuu siten se, missä kulkee hyväksyttävän tiedonkeruun ja yksilöihin kohdistuvan liiallisen seurannan raja. Tämä korostuu erityisesti lastensuojelussa, jossa tekoälyn avulla tehtävät ennusteet voivat vaikuttaa siihen, millaisena lapsi ja hänen perheensä nähdään jo ennen konkreettisen huolen tai tuen tarpeen ilmenemistä.

Yksityisyyden suojaan liittyvät kysymykset vahvistavat tutkielmani ihmisoikeusperustaista tarkastelua. Sosiaalityössä oikeuksien edistäminen ja suojeleminen ovat keskeinen osa ammatillista toimintaa, mutta oikeusperustaiseen työskentelyyn liittyy myös tilanteita, joissa eri oikeudet ja tavoitteet voivat joutua keskenään jännitteiseen suhteeseen (Forrester 2024, 35–36). Lastensuojelussa tekoälyn käyttöä ei voida arvioida ainoastaan sen perusteella, kuinka tehokkaasti järjestelmä tuottaa ennusteita, vaan myös sen perusteella, millaisia vaikutuksia teknologialla on lapsen oikeuksiin, osallisuuteen, yksityisyyteen ja yhdenvertaisuuteen (Reamer 2023, 61–62). Tekoälyn hyödyntäminen voi näin ollen synnyttää uudenlaisen jännitteen lastensuojelun suojelutehtävän ja lapsen sekä perheen oikeuksien välille: laaja tietojen kerääminen voidaan perustella lapsen suojelemisella, mutta samalla sama tiedonkeruu voi vahvistaa valvontaa ja kaventaa yksityisyyttä (Gillingham 2016, 1046–1047; Reamer 2023, 80–81). Eettisesti kestävä ratkaisu edellyttää näiden tavoitteiden välistä tapauskohtaista ja oikeusperustaista punnintaa, jossa lapsen suojelemisen tarvetta arvioidaan suhteessa lapsen ja perheen muihin oikeuksiin.

4.7 Eettisesti kestävän tekoälyn edellytykset

Eettisten kysymyksien lisäksi artikkeleissa esitettiin useita näkemyksiä siitä, miten eettisiin kysymyksiin voidaan vastata. Ratkaisuissa korostuivat yleisesti tekoälyjärjestelmien huolellinen kehittäminen ja arviointi, eri sidosryhmien osallistaminen, ammatillisen harkinnan säilyttäminen, järjestelmien läpinäkyvyys sekä eettisten ja oikeudellisten näkökulmien huomioiminen koko tekoälyn elinkaaren ajan. Artikkeleissa korostui myös näkemys siitä, ettei tekoälyn eettisyyttä voida varmistaa ainoastaan teknisillä ratkaisuilla, vaan huomio tulee kiinnittää myös siihen organisatoriseen ja ammatilliseen toimintaympäristöön, jossa järjestelmiä käytetään. Kokonaisuutena eettisesti kestävä tekoälyn käyttö näyttäytyi prosessina,

joka alkaa järjestelmän suunnittelusta ja jatkuu sen käyttöönoton, käytön ja vaikutusten jatkuvana arviointina. (Agbana 2025; Yu & Roderick 2025.)

Epäoikeudenmukaisuuden ja eriarvoistumisen ehkäisemisen kannalta merkittäväksi nähtiin datan edustavuus, vinoumien tunnistaminen sekä eri sidosryhmien osallistaminen järjestelmien kehittämiseen. Rajoittuneisiin tai vinoutuneisiin aineistoihin perustuvat mallit voivat toimia eri väestöryhmissä epätarkasti ja johtaa epäoikeudenmukaisiin lopputuloksiin. Tämän vuoksi mallien kehittämisessä tulisi hyödyntää mahdollisimman monipuolisia aineistoja ja arvioida niiden toimintaa eri väestöryhmissä. Pelkkä monimuotoisen koulutusaineiston käyttö ei kuitenkaan poista vinoumien riskiä, minkä vuoksi malleja tulisi kehittää yhteistyössä erilaisia elettyjä kokemuksia omaavien henkilöiden ja keskeisten sidosryhmien, kuten sosiaalityöntekijöiden, kanssa. Yhteiskehittämisen avulla voidaan tuoda esiin erilaisia näkökulmia jo suunnitteluvaiheessa ja tunnistaa sellaisia oletuksia ja mahdollisia haittoja, joita teknisessä kehittämisessä ei välttämättä havaita. (Agbana 2025.) Kirjallisuuskatsauksissa esitettiin vinoumien vähentämiseksi vinouma-auditointeja sekä oikeudenmukaisuutta huomioivien mallinnusmenetelmien hyödyntämistä (Hall ym. 2023). Oikeudenmukaisuuden varmistaminen edellyttää kuitenkin teknisten ratkaisujen lisäksi jatkuvaa seurantaa ja eettisten tarkistuspisteiden sisällyttämistä koko järjestelmän elinkaareen (Agbana 2025).

Lapsen ja perheen tilanteen virheelliseen tulkintaan liittyviä riskejä vähennetään artikkelien perusteella kehittämällä tekoälyjärjestelmiä siten, että ne huomioivat paremmin lastensuojelun toimintaympäristön ja ammatillisen tiedon erityispiirteet. Tekoälytyökalujen tulisi olla näyttöön perustuvia, ja niiden kehittämisessä tulisi hyödyntää lastensuojelun ammattilaisten asiantuntemusta. (Agbana 2025.) Lisäksi korostui tarve kehittää sellaisia tietoaineistoja ja dokumentointikäytäntöjä, joiden avulla tekoälylle välittyvä tieto olisi mahdollisimman laadukasta ja yhdenmukaista. Landaun ym. (2022) mukaan datan standardoinnin vähäisyys ja organisaatioiden vaihtelevat dokumentointikäytännöt vaikeuttavat luotettavien vertailuaineistojen muodostamista ja mallien yleistettävyyttä. Tämän vuoksi artikkelissa ehdotettiin riskinarviointiin liittyvien tietoelementtien suurempaa standardointia sekä erilaisten tietolähteiden yhdistämistä. Artikkelit toivat kuitenkin esiin sen, ettei laadukkaampikaan data yksin ratkaise lapsen tilanteen tulkintaan liittyvää ongelmaa (Kim & Yoon 2026; Lehtiniemi 2022; Wilkins & Benett 2025). Sosiaalityöntekijöiden tulee edelleen arvioida algoritmin tuottamaa tietoa suhteessa lapsen ja perheen ajankohtaiseen tilanteeseen sekä muuhun saatavilla olevaan tietoon. Erityisesti ennusteiden tulisi kohdistua asiakkaan nykytilanteen kannalta

merkityksellisiin asioihin, jolloin tekoäly voisi toimia vuorovaikutusta ja ammatillista tiedonmuodostusta tukevana välineenä. (Lehtiniemi 2022.)

Ammatillisen harkinnan kaventumisen ja vastuun hämärtymisen ehkäisemiseksi korostui ihmisen säilyminen päätöksenteon keskiössä. Human-in-the-Loop-lähestymistavassa tekoälyn tuottamat riskipisteet ja hälytykset integroidaan osaksi olemassa olevia lastensuojeluprosesseja, mutta lopullinen päätösvalta säilyy pätevillä ammattilaisilla. (Agbana 2025, Lehtiniemi 2022.) Tällainen toimintatapa on erityisen tärkeä tilanteissa, joissa algoritminen arvio ei tavoita tapauksen yksilöllisiä tai kontekstuaalisia piirteitä. Sosiaalityöntekijän tehtävänä on tällöin suhteuttaa algoritmin tuottama tieto laadullisiin havaintoihin, vuorovaikutukseen ja muuhun kontekstuaaliseen tietoon, jota järjestelmä ei kykene täysin huomioimaan. Ihmisen suorittama arviointi ennen toimenpiteitä voi samalla vähentää liiallista riippuvuutta todennäköisyyspohjaisista tuloksista. Human-in-the-Loop-malli tukee myös vastuun selkeyttämistä, sillä ammattilainen voi dokumentoida perustelunsa sekä silloin, kun hän yhtyy tekoälyn suositukseen, että silloin, kun hän päättää poiketa siitä. (Mt.)

Ammatillisen harkinnan säilyttämiseen liittyi myös näkemys siitä, ettei algoritmien suositusten noudattamista tulisi tehdä pakolliseksi. Algoritmien tehtävänä on tukea sosiaalityöntekijöitä ja perheitä sekä täydentää päätöksentekoa, ei korvata sitä. (Gabriels ym. 2025.) Työkalujen kehittämisessä tulisi siksi hyödyntää sosiaalityöntekijöiden asiantuntemusta, tarjota työntekijöille riittävästi koulutusta järjestelmien mahdollisuuksista ja rajoituksista sekä ylläpitää avointa keskustelua niiden käytöstä. Sosiaalityöntekijöiden tulisi olla mukana jo järjestelmien kehittämisen alkuvaiheessa eikä vasta niiden loppukäyttäjinä. Selkeät ja ei-tekniset selitykset sekä jatkuva koulutus voivat vahvistaa työntekijöiden ymmärrystä algoritmisista työkaluista ja tukea niiden tarkoituksenmukaista käyttöä. (Kawakami ym. 2025.) Osassa tarkasteltavia artikkeleita korostettiin lisäksi päätösvallan selkeää jakamista ihmisten ja järjestelmien välillä sekä organisatoristen tekijöiden huomioimista järjestelmien käyttöönotossa (Agbana 2025; Kawakami ym. 2025). Vastuukysymysten selkeyttämiseksi viranomaisten ja teknologiatoimittajien välisten sopimusten tulisi puolestaan määritellä selkeästi eri toimijoiden roolit, vastuut ja mahdolliset korvausmekanismit (Agbana 2025).

Lapsen tarpeisiin vastaamisen ja tuen toteutumisen näkökulmasta artikkelit eivät tarjonneet yhtä selkeää yksittäistä ratkaisua, vaan painottivat tekoälyn käyttötarkoituksen ja toimintaperiaatteiden sovittamista lastensuojelun tavoitteisiin. Sosiaalityöntekijät kokivat, että

työkalujen tulisi palvella heidän ammatillista tiedonmuodostustaan eikä toimia välineenä, jolla oikeutetaan tiettyjä toimenpiteitä. (Lehtiniemi 2022.) Esiin nousi myös tarve sellaiselle järjestelmälle, jossa ennustaminen ja interventiot ovat tiiviisti yhteydessä toisiinsa (Kim & Yoon 2026). Tämän vuoksi tekoälyn hyödyllisyyttä ei tulisi arvioida ainoastaan sen ennustetarkkuuden perusteella, vaan myös sen mukaan, tukeeko se tosiasiallisesti sosiaalityöntekijöiden työtä ja edistääkö se lasten ja perheiden tarvitseman tuen toteutumista. Algoritmisten työkalujen vaikutuksia perheisiin tulisi myös arvioida pitkittäistutkimuksilla ja verrata niiden vaikutuksia muihin päätöksenteon välineisiin. (Gabriels ym. 2025.)

Yksityisyyden suojan, tietoturvan ja valvonnan näkökulmasta eettisesti kestävä tekoälyn käyttö edellyttää erityisesti tietojen asianmukaista suojaamista ja niiden käytön hallintaa. Keskeisiksi suojauskeinoiksi tunnistettiin tietojen salaaminen, turvallinen tallennus ja roolipohjaiset käyttöoikeudet. Myös viranomaisten välinen tietojenvaihto tulisi järjestää siten, että se täyttää lakisääteiset vaatimukset ja perustuu asianmukaiseen suostumukseen. (Agbana 2025.) Tietosuojan ja tutkimuksen välisen ristiriidan vuoksi esiin tuotiin myös anonymisoidun datan käyttö turvallisissa ympäristöissä, ohjelmien avoin toimintaperiaate sekä tapa, jossa tekoälymallia voidaan opettaa ilman, että varsinaista tietoa tarvitsee jakaa (Bull ym. 2025). Yksityisyyden suojaa voidaan lisäksi parantaa käyttämällä yksittäisten ihmisten tietojen sijaan esimerkiksi tietyn alueen ihmisistä koottua tietoa. Tällaisen lähestymistavan on katsottu vähentävän yksityisyyteen liittyviä riskejä ja samalla ehkäisevän väestötason mallien virheellistä soveltamista yksittäisiin tapauksiin. (Yu & Roderick 2025.)

Eettisesti kestävän tekoälyn edellytykset liittyvät artikkelien perusteella sekä mallien teknisiin ominaisuuksiin että siihen, miten tekoälyjärjestelmät suunnitellaan, otetaan käyttöön ja kytketään osaksi lastensuojelun toimintaa. Eettisesti kestävän käytön edellytyksenä on, että järjestelmien kehittämisessä huomioidaan oikeudenmukaisuus, läpinäkyvyys, vastuullisuus ja yksityisyyden suoja sekä lastensuojelun keskeiset arvot, kuten ihmisarvon ja yksilön arvon kunnioittaminen. (Agbana 2025.) Eri sidosryhmien, erityisesti sosiaalityöntekijöiden sekä lastensuojelun asiakkaiden, osallistaminen voi auttaa varmistamaan, että kehitetyt ratkaisut vastaavat todellisiin tarpeisiin ja että niiden mahdolliset haitat tunnistetaan jo ennen käyttöönottoa (Yu & Roderick 2025). Lisäksi tekoälyjärjestelmien käyttöä tulisi seurata ja arvioida jatkuvasti, eikä eettistä hyväksyttävyyttä nähdä kertaluonteisena arviointina. Näin ollen tekoäly näyttäytyy sosiaalityön ammatillista harkintaa täydentävänä välineenä. Eettisesti kestävä ratkaisu ei ole sellainen, joka siirtää päätöksentekoa teknologialle, vaan sellainen, joka

tukee ammattilaista lasten ja perheiden tilanteiden arvioinnissa samalla, kun päätöksenteko, vastuu ja eettinen harkinta säilyvät ihmisillä.

5. JOHTOPÄÄTÖKSET

5.1Analyysista johtopäätöksiin

Tarkastelin tutkielmassani tutkimuskirjallisuuden perusteella tekoälyn käyttöön lastensuojelussa liittyviä eettisiä kysymyksiä. Tarkasteluun valikoituneiden artikkelien perusteella tekoälyn käyttöön lastensuojelussa liittyvät eettiset haasteet jäsentyivät viiteen pääluokkaan: epäoikeudenmukaisuuteen ja eriarvoisuuden vahvistumiseen, lapsen ja perheen tilanteen virheelliseen tulkintaan, ammatillisen harkinnan kaventumiseen ja vastuun hämärtymiseen, lapsen tarpeisiin vastaamisen ja tuen toteutumisen vaarantumiseen sekä yksityisyyden suojaan, tietoturvaan ja valvontaan liittyviin kysymyksiin. Näiden lisäksi tarkastelin eettisesti kestävän tekoälyn edellytyksiä artikkelien pohjalta. Tutkimustulosteni pohjalta tekoälyn käyttöön lastensuojelussa liittyy useita eettisiä kysymyksiä. Eettiset kysymykset eivät kuitenkaan tee tekoälyn hyödyntämisestä lastensuojelussa mahdotonta. Sen sijaan tekoälyn käytössä on tärkeää huomioida eettisyys koko prosessin ajan, aina tekoälymallien kehittämisen alkuvaiheista niiden käyttöönottoon ja tiedon hyödyntämiseen asti.

Yksi tutkielmani keskeinen johtopäätös on, että tekoälyn käytön eettiset haasteet lastensuojelussa eivät liity yksittäisiin teknisiin ongelmiin, vaan laajemmin siihen, miten tekoälyn tuottama tieto muuttaa sosiaalityön tiedonmuodostusta, ammatillista harkintaa ja päätöksentekoa sekä sitä, millaisena lapsi ja perhe tulevat näissä prosesseissa nähdyiksi. Tekoäly voi vahvistaa käytetyssä datassa olevia vinoumia, tuottaa puutteellisia tulkintoja monimutkaisista elämäntilanteista ja vaikuttaa ammatilliseen harkintaan tavalla, jota työntekijä ei välttämättä itse tiedosta. Samalla tekoälyn vaikutukset eivät määräydy ainoastaan teknologisten ominaisuuksien perusteella, vaan myös sen perusteella, millaiseen organisatoriseen ja ammatilliseen toimintaympäristöön järjestelmä kytketään.

Sitä, miten lapsen ja perheen tilanne lastensuojelun prosesseissa ymmärretään ja nähdään, voidaan tarkastella sosiaalityön systeemiteoreettisen ja ekologisen näkökulman kautta (Healy 2014, 121–123). Ekologisessa näkökulmassa ihmisen tilannetta ei tarkastella irrallaan hänen

ympäristöstään, vaan huomio kohdistuu yksilöön ja hänen sosiaalisen ympäristönsä välisiin suhteisiin sekä siihen, miten erilaiset ympäröivät järjestelmät vaikuttavat hänen tilanteeseensa (mt., 124–127). Näkökulma vahvistaa havaintoa siitä, että lapsen ja perheen tilannetta ei voida ymmärtää laajasti yksittäisten muuttujien, riskitekijöiden tai algoritmisen riskiluokituksen avulla.

Lastensuojelun kontekstissa merkityksellisiä ovat myös lapsen ja perheen ihmissuhteet, elämäntilanne, kasvuympäristö sekä palvelujärjestelmän toiminta. Hietamäen (2015, 150) tutkimus tuo esiin yhteisen näkemyksen merkityksen perheen tilanteen ymmärtämisessä: sosiaalityöntekijän ja vanhempien samanlainen näkemys perheen tilanteesta oli yhteydessä hyvään asiakassuhteeseen ja vanhemman osallisuuteen. Tekoälyn käytössä tämä tarkoittaa, että algoritmin tuottama tieto voi tarjota yhden näkökulman lapsen tilanteeseen, mutta se ei yksinään kuvaa lapsen ja perheen tilannetta osana heidän sosiaalisia suhteitaan ja ympäristöään. Eettisesti kestävän tekoälyn näkökulmasta keskeiseksi kysymykseksi muodostuu, kaventaako teknologia lapsen tilanteen tarkastelua yksittäisiin mitattaviin ominaisuuksiin vai tukeeko se lapsen ja perheen tilanteen kokonaisvaltaista ymmärtämistä osana laajempaa sosiaalista ja yhteiskunnallista kokonaisuutta (Kim & Yoon 2026; Lehtiniemi 2022; Wilkins & Bennett 2025).

Tarkastelemieni artikkelien perusteella yksi keskeinen eettinen kysymys liittyy siihen, millaisen aseman tekoälyn tuottama tieto saa ammatillisessa päätöksenteossa. Tekoälyn tarkoituksena voidaan nähdä olevan sosiaalityöntekijän tiedonmuodostuksen ja päätöksenteon tukemisen, mutta sen tuottama arvio voi käytännössä saada suuremman painoarvon kuin on tarkoitettu. Erityisesti automaatioharha, algoritmien läpinäkymättömyys sekä organisaatioiden toimintatavat voivat vahvistaa algoritmisen tiedon asemaa. (Kawakami ym. 2022, Saxena & Guha 2024.) Tällöin vaarana on, että algoritmin tuottama riskiluokitus alkaa näyttäytyä objektiivisena tietona, jota ammattilaisen on vaikea kyseenalaistaa. On tärkeää tunnistaa, että tekoäly voi ohjata ammattilaisen ajattelua myös silloin, kun sen vaikutusta ajatteluun ei tiedosteta. Tekoälyn vaikutusten tunnistaminen on merkityksellistä, sillä vaikutusten huomioiminen mahdollistaa myös niiden hallinnan.

Ammatillisen harkinnan merkitystä voidaan syventää Michael Lipskyn (2010) katutason byrokratian teorialla. Lipskyn (mts., 13–16) mukaan julkisten palvelujen työntekijät toimivat tilanteissa, joissa asiakkaiden tarpeet, organisaatioiden tavoitteet ja käytettävissä olevat resurssit eivät aina ole keskenään tasapainossa. Tällöin työntekijöiden harkintavalta on

keskeinen osa palvelujen toteuttamista. Sosiaalityössä ammatillinen harkinta ei siten ole ainoastaan yksittäisen työntekijän ominaisuus, vaan olennainen osa sitä, miten julkista sosiaalipolitiikkaa käytännössä toteutetaan (Kawakami ym. 2022; Saxena & Guha 2024). Tekoälyn käyttöönotto voi tässä suhteessa muuttaa harkinnan ehtoja: algoritminen tieto voi yhtäältä tukea työntekijää epävarmoissa tilanteissa, mutta toisaalta kaventaa hänen mahdollisuuksiaan tehdä tapauskohtaista arviota. Tarkastelemani artikkelit osoittavat, että tekoälyn käyttöönotossa ei ole kyse vain uuden työvälineen ottamisesta osaksi sosiaalityötä, vaan myös siitä, miten sosiaalityöntekijän ammatillinen harkinta ja päätöksenteko muuttuvat.

Tekoälyn käytössä lastensuojelussa ongelmallista ei ole ainoastaan se, että järjestelmä voi ehdottaa epätarkoituksenmukaista tai liiallista toimenpidettä. Ongelmallista voi olla myös se, jos algoritminen riskitieto alkaa määrittää sitä, kuinka vakavana lapsen ja perheen tilanne nähdään ja millaiseen puuttumiseen tilanteessa päädytään. Lastensuojelun hienovaraisuuden periaate ja avohuollon tukitoimien ensisijaisuus edellyttävät, että toimenpiteiden tarpeellisuutta arvioidaan suhteessa lapsen yksilölliseen tilanteeseen eikä ainoastaan algoritmin tuottamaan riskiluokitukseen (Räty 2026, 8–9). Samalla on tärkeää tunnistaa, ettei voimakkaampi puuttuminen ole itsessään eettisesti ongelmallista silloin, kun lapsen etu sitä edellyttää. Keskeistä on, että päätös perustuu lapsen ja perheen tilanteen kokonaisarvioon ja sosiaalityöntekijän omaan harkintaan. Algoritmin antama ennuste ei saa yksin ratkaista päätöstä.

Ammatillista harkintaa voidaan tarkastella myös sosiaalityön ammatillisen autonomian näkökulmasta. Tekoälyn ei tulisi korvata ammatillista harkintaa, vaan sen tulisi täydentää sitä sellaisella tiedolla, jonka merkityksen ja rajoitukset sosiaalityöntekijä pystyy arvioimaan. Reamerin (2023, 63–65) mukaan tekoälyn eettinen käyttö sosiaalityössä edellyttää läpinäkyvyyttä, vastuullisuutta ja sitä, että teknologiaa käytetään asiakkaiden hyvinvointia ja oikeuksia tukevalla tavalla. Näin ollen ihmisen säilyminen päätöksenteon keskiössä ei tarkoita ainoastaan sitä, että ihminen tekee lopullisen päätöksen, vaan myös sitä, että ammattilaisella säilyy tosiasiallinen mahdollisuus ymmärtää, arvioida ja tarvittaessa kyseenalaistaa tekoälyn tuottama tieto.

Tärkeä eettinen havainto liittyi myös oikeudenmukaisuuteen. Tekoälyjärjestelmät perustuvat dataan, joka voi sisältää aikaisemman toiminnan seurauksena syntyneitä historiallisia ja rakenteellisia vinoumia. Algoritmi voi siten oppia ennustamaan menneisyydessä esiintyneitä eroja ja vahvistaa niitä edelleen. Toisaalta yhdessä artikkelissa (Kawakami ym. 2022)

kuvattiin, että tekoälyn hyödyntäminen saattoi vähentää työntekijöiden ennakkoluulojen vaikutusta arviointiin. Havainnon pohjalta voidaan ajatella, että on mahdollista kehittää jopa oikeudenmukaisuutta vahvistavia malleja. Tekoälyjärjestelmiä voidaan myös kehittää ja mukauttaa niiden käytöstä saatavan kokemuksen ja tiedon perusteella, jolloin kehittämisessä voidaan huomioida sosiaalityöntekijöiden käytännön kokemukset ja tarpeet (Lehtiniemi 2022; Kawakami ym. 2022). Oikeudenmukaisuuden kysymys kytkeytyy myös siihen, miten sosiaalityön arviot rakentavat kuvaa lapsesta ja perheestä. Menneisyyteen perustuva algoritminen tieto voi vahvistaa ennakkokäsityksiä asiakkaasta ja ohjata ammattilaisen huomiota pois tämänhetkisestä tilanteesta. Tällöin aikaisemmista tapahtumista tai tietokantaan tallentuneista ominaisuuksista voi tulla osa lapsen tai perheen pysyvämpää määrittelyä, vaikka tilanne olisi muuttunut. Lastensuojelun näkökulmasta tämä on merkityksellistä, koska arvio ei ainoastaan kuvaa asiakasta, vaan voi vaikuttaa siihen, millaisiin toimenpiteisiin hänen kohdallaan päädytään.

Oikeudenmukaisuutta voidaan tarkastella Value Sensitive Design -lähestymistavan kautta (Friedman & Hendry 2019, 35–38). Sen mukaan teknologian suunnittelussa tulisi tunnistaa ja huomioida teknologiaan liittyvät inhimilliset arvot jo suunnitteluprosessin aikana. Lähestymistapa on tutkielmani kannalta kiinnostava, koska se siirtää eettisen tarkastelun painopistettä teknologian käyttöönoton jälkeisestä arvioinnista jo kehittämisvaiheeseen. Lastensuojelussa tämä tarkoittaa esimerkiksi sitä, että yhdenvertaisuus, ihmisarvo, yksityisyys, lapsen osallisuus ja ammatillinen harkinta eivät ole tekoälyn käytön seurauksena myöhemmin arvioitavia ominaisuuksia, vaan niiden tulisi ohjata järjestelmän suunnittelua alusta lähtien. Tämä tukee tutkielmani johtopäätöstä siitä, että eettisesti kestävän tekoälyn rakentaminen edellyttää sosiaalityöntekijöiden ja muiden asianosaisten osallistumista jo teknologian kehittämisvaiheessa.

Tekoälyn tunnistamiskyvyn ja palvelujärjestelmän toimintakyvyn välille muodostui artikkelien perusteella selkeä jännite: tekoäly voi auttaa tunnistamaan tuen tarpeessa olevia lapsia ja perheitä, mutta se ei itsessään lisää palvelujen, henkilöstön tai muiden resurssien saatavuutta. Havainto on merkityksellinen myös suomalaisessa lastensuojelukontekstissa, jossa tekoälyyn kohdistuu kasvavia odotuksia palvelujen toimivuuden, tuottavuuden ja ammattilaisten työn helpottamisen näkökulmasta. Tekoälyn hyödyntämisen tavoitteena on vapauttaa ammattilaisten aikaa asiakastyöhön sekä tehostaa palvelujärjestelmän toimintaa (Ministry of Social Affairs and Health 2024, 5). Samalla suomalaisessa sosiaali- ja terveydenhuollossa tekoälyn hyödyntämiseen kohdistuu odotuksia palvelutarpeen arvioinnin ja ennakoivan analytiikan

suhteen (THL 2026). Tarkastelemieni artikkelien mukaan on tärkeää tunnistaa, ettei tekoälyn kyky tuottaa tietoa tai tunnistaa tuen tarvetta poista niitä rakenteellisia ongelmia, jotka liittyvät palvelujen ja henkilöstön riittävyyteen. Tekoälytyökalujen tuominen ennakoivaan analytiikkaan ja päätöksenteon tueksi voi parhaimmillaan tukea ammattilaisten työtä, mutta se ei itsessään ratkaise lastensuojelun resurssipulaa. Tämän vuoksi on perusteltua pohtia myös sitä, miten rajallisia resursseja tekoälyn kehittämiseen ja käyttöönottoon kohdennetaan. Jos vähäisiä resursseja suunnataan tekoälyjärjestelmien kehittämiseen tilanteessa, jossa järjestelmät eivät lopulta tuota tavoiteltua hyötyä, on mahdollista, että nämä resurssit ovat pois muusta asiakastyöstä ja palvelujen järjestämisestä. Tekoälyn käyttöönotossa tulisikin siten arvioida paitsi sen mahdollisia hyötyjä myös sitä, tuottaako siihen käytetty resurssi todellista lisäarvoa suhteessa muihin tapoihin vahvistaa lastensuojelun palveluja ja henkilöstön mahdollisuuksia vastata asiakkaiden tarpeisiin.

Yksityisyyden suojaan liittyvät kysymykset nousivat tarkastelussani esiin harvemmin kuin tutkimuksen alkuvaiheessa oletin, sillä niitä käsiteltiin vain osassa artikkeleista. Tästä huolimatta esiin nousseet näkökulmat osoittautuivat eettisesti merkittäviksi ja korostivat yksityisyyden suojan merkitystä tekoälyn hyödyntämisessä lastensuojelussa. Keskeiseksi kysymykseksi muodostui se, kuinka paljon tietoa lapsesta ja perheestä on perusteltua kerätä ja yhdistää. Esimerkiksi ennakoivassa tekoälyn käytössä pyritään hyödyntämään laajoja tietoaineistoja, mikä voi olla ristiriidassa tietojen minimoinnin periaatteen kanssa. Mitä enemmän lapsesta ja perheestä kerätään ja yhdistetään tietoa, sitä yksityiskohtaisemmaksi heistä kerätty data muodostuu. Samalla lisääntyy riski siitä, että tietoa hyödynnetään tavoilla, joita asiakas ei ole voinut ennakoida. Näin yksityisyyden suojaan liittyvä kysymys ei koske ainoastaan sitä, miten kerättyjä tietoja suojataan, vaan myös sitä, millaisten tietojen kerääminen ja yhdistäminen on ylipäätään perusteltua. Tämä herättää laajemman yhteiskunnallisen kysymyksen siitä, millaista lapsia ja perheitä koskevaa tietoa on hyväksyttävää kerätä tekoälyn käyttöä varten ja millä perusteella. Laajamittainen tiedonkeruu voi siten muuttaa myös lastensuojelun toimintaa kohti ennakoivaa valvontaa. Jos lapsia ja perheitä tarkastellaan mahdollisten tulevien riskien perusteella jo ennen konkreettisen huolen ilmenemistä, riskin ennustamisesta voi tulla osa sitä tietoa, joka määrittää asiakkaan asemaa. Tällöin yksityisyyden suoja ei rajoitu tietojen tekniseen suojaamiseen, vaan kysymys ulottuu myös siihen, millä oikeudella ja millä perusteella lapsesta tai perheestä muodostetaan ennusteita tulevasta.

Tutkielmani perusteella eettisesti kestävän tekoälyn käyttö sosiaalityössä edellyttää teknisten ratkaisujen lisäksi muutoksia ammatillisiin ja organisatorisiin käytäntöihin. Datan tulee olla

mahdollisimman edustavaa ja mahdollisia vinoumia tulee arvioida. Järjestelmien toimintaperiaatteiden tulee olla riittävän läpinäkyviä ja niiden käyttöön liittyvät vastuut selkeästi määriteltyjä. Sosiaalityöntekijöiden asiantuntemusta tulisi hyödyntää jo järjestelmien kehittämisvaiheessa, ja ammattilaisille tulee tarjota riittävästi tietoa järjestelmien mahdollisuuksista ja rajoituksista. Eettisesti kestävän tekoälyn kehittäminen edellyttää myös riittävää tietopohjaa siitä, millaisia vaikutuksia tekoälyn käytöllä on sosiaalityöhön ja sen asiakkaisiin.

Ratnerin ja Elmholdtin (2023) mukaan tekoälyn tulee toimia ammatillista harkintaa tukevana välineenä eikä sitä rajoittavana tekijänä. Tämä on erityisen tärkeää sosiaalityössä, jossa haavoittuvassa asemassa olevien ihmisten tilanteet ovat usein monimutkaisia ja edellyttävät tapauskohtaista arviointia. Tekoälyn ei siten tulisi ohjata sosiaalityöntekijää tarkastelemaan asiakkaan tilannetta vain järjestelmän tuottamien tietojen tai ennusteiden kautta, vaan sen tehtävänä tulisi olla ammatillisen harkinnan tukeminen. Sosiaalityön ihmisoikeuksiin ja sosiaaliseen oikeudenmukaisuuteen perustuva ammattietiikka asettaa tekoälyn käytölle erityisiä vaatimuksia, minkä vuoksi myös tekoälyn vaikutuksia sosiaalityöhön ja sen asiakkaisiin tulee seurata ja arvioida jatkuvasti. Value Sensitive Design -näkökulma syventää tätä johtopäätöstä siten, että eettinen arviointi ulotetaan koko teknologian elinkaareen. (Friedman & Hendry 2019, 35–39). Näkökulman mukaan teknologian suunnittelussa tulee tarkastella samanaikaisesti teknisiä ratkaisuja ja niihin liittyviä arvoja sekä huomioida myös teknologian välillisesti vaikuttamat tahot. Lastensuojelussa tämä tarkoittaa, että esimerkiksi lapsen oikeudet, yksityisyys, yhdenvertaisuus, osallisuus ja ammatillinen harkinta tulisi huomioida jo järjestelmän tavoitteiden määrittelyssä, datan valinnassa ja mallin suunnittelussa eikä vasta sen käyttöönoton jälkeen.

5.2 Tutkimusprosessin tarkastelu

Tutkielmani tuloksia tarkasteltaessa on huomioitava kirjallisuuskatsauksen aineiston muodostumiseen liittyvät rajaukset. Lopullinen aineisto koostui 12 artikkelista, jotka eivät yhtä suomalaista tutkimusta lukuun ottamatta käsitelleet suomalaista lastensuojelun toimintaympäristöä. Lastensuojelujärjestelmät, lainsäädäntö, palvelurakenteet ja tekoälyn käyttötavat vaihtelevat kuitenkin eri maiden välillä. Tämän vuoksi kansainvälisen tutkimuksen tuloksia ei voida suoraan siirtää suomalaiseen lastensuojeluun. Tuloksia voidaan pikemminkin

hyödyntää tunnistamaan sellaisia eettisiä kysymyksiä, jotka voivat olla merkityksellisiä myös suomalaisessa toimintaympäristössä.

Aineiston erityispiirteenä oli, että artikkeleissa tekoälyn käyttöä tarkasteltiin erityisesti päätöksenteon ja ennakoivan analytiikan yhteydessä. Tämän vuoksi aineisto ei tarjonnut yhtä kattavaa kuvaa niistä eettisistä kysymyksistä, joita tekoälyn hyödyntämiseen muissa sosiaalityön tehtävissä voi liittyä. Tämä tuli itselleni erityisesti näkyväksi tutkimusprosessin loppuvaiheessa, kun tutustuin uudessa työssäni käytännön lastensuojelutyössä tekoälyn hyödyntämiseen kirjaamisen tukena. Tutkimusprosessin aikana ymmärrykseni tekoälyn erilaisista käyttötavoista sosiaalityössä laajeni, mikä herätti myös kysymyksen siitä, olisiko tutkimuksessa ollut perusteltua tarkastella ennakoivan analytiikan ja päätöksenteon lisäksi laajemmin muita tekoälyn käyttötapoja. Tämä olisi voinut tuoda näkyviin myös sellaisia eettisiä kysymyksiä, jotka jäivät tämän tutkielman aineistossa vähäisemmälle huomiolle. Vaikka laajempi tarkastelu olisi voinut tuoda esiin myös muita tekoälyn käyttöön liittyviä eettisiä kysymyksiä, käytettävissä olevan tutkimustiedon perusteella näin laajan näkökulman toteuttaminen kirjallisuuskatsauksena olisi ollut haastavaa. Tästä huolimatta tutkielma tarjoaa yleiskuvan keskeisistä tekoälyn käyttöön liittyvistä eettisistä kysymyksistä lastensuojelussa. Lisäksi se nostaa esiin tarpeen jatkotutkimukselle, jossa tarkastellaan erityisesti suomalaisessa lastensuojelussa käyttöön otettavien uusien tekoälyratkaisujen eettisiä kysymyksiä ja niiden vaikutuksia käytännön sosiaalityöhön.

Tutkimusprosessin aikana tehtyjä rajauksia voidaan tarkastella myös suhteessa siihen, että tekoäly on sosiaalityössä osa muuttuvaa sosioteknistä toimintaympäristöä. Tekoälyn käyttötavat, organisaatioiden toimintamallit ja ammattilaisten roolit kehittyvät samanaikaisesti, minkä vuoksi tutkielmani tulokset kuvaavat tiettyä vaihetta nopeasti muuttuvassa tutkimus- ja toimintaympäristössä (Lehtiniemi 2023, 1122–1123). Tämä korostaa tarvetta seurata tekoälyn käyttöönottoa myös pitkittäisesti eikä arvioida sen eettisiä vaikutuksia ainoastaan ennen käyttöönottoa tai yksittäisen tutkimushetken perusteella.

Myös aineiston hakuprosessi ja kuvailevan kirjallisuuskatsauksen luonne ovat vaikuttaneet siihen, millaisena tekoälyn eettisyys tutkielmassani näyttäytyy (Salminen 2011). Muodostin hakusanat lastensuojeluun, tekoälyyn ja etiikkaan liittyvistä käsitteistä. Tämä on voinut osaltaan ohjata aineistoa kohti tutkimuksia, joissa tekoälyä tarkastellaan nimenomaan eettisten kysymysten näkökulmasta. Tutkielman perusteella ei siten voida tehdä kattavaa päätelmää kaikista eettisistä kysymyksistä, joita tekoälyn mahdollisiin käyttötapoihin lastensuojelussa

liittyy. Pikemminkin tutkielma kuvaa sitä, millaisia eettisiä kysymyksiä valitussa tutkimuskirjallisuudessa on tunnistettu ja millaisia ehtoja eettisesti kestävälle tekoälyn käytölle on asetettu. Tämä on tärkeää huomioida myös arvioitaessa sitä, kuinka yleisiä tai merkittäviä aineistossa esiin nousseet kysymykset ovat. Vaikka aineistossa esiintyi useita erilaisia näkökulmia ja tekoälyn käyttötapoja, tietyt eettiset kysymykset toistuivat tutkimuksissa ja muodostivat tämän tutkielman keskeiset analyysiluokat.

Tutkimusintressiäni on ohjannut laajempi kiinnostus tekoälyn eettisyyteen sosiaalityössä. Eettisyyttä ei kuitenkaan tutkielmassani ole sidottu yhteen tiettyyn eettiseen teoriaan tai ennalta määriteltyyn etiikan analyysikehikkoon, vaan eettiset kysymykset ovat muodostuneet aineistolähtöisesti. Tämän vuoksi tutkielman tuloksia voidaan pitää kuvauksena siitä, millaisissa tilanteissa tekoälyn käyttö näyttäytyy tutkimuskirjallisuudessa eettisesti ongelmallisena ja millaisia ehtoja eettisesti kestävälle käytölle asetetaan. Tuloksia ei tule tulkita yksiselitteisenä määritelmänä sille, mitä tekoälyn eettisyys lastensuojelussa tarkoittaa.

Tutkimusprosessin perusteella voidaan kuitenkin todeta, että vaikka eettisiä kysymyksiä ei ole sidottu yhteen ennalta määriteltyyn teoriaan, tutkielman tulokset kytkeytyvät vahvasti sosiaalityön keskeisiin arvoihin. Erityisesti ihmisarvon, ihmisoikeuksien, sosiaalisen oikeudenmukaisuuden, osallisuuden ja ammatillisen harkinnan merkitys korostuu tutkielman tuloksissa (Talentia 2024). Näin aineistolähtöinen analyysi mahdollistaa sosiaalityön keskeisten arvojen ja eettisten periaatteiden tarkastelun tekoälyn käyttöön liittyvien kysymysten yhteydessä.

Tutkimusprosessin aikana käsitykseni tutkielman aiheesta myös tarkentui. Tutkimuksen alkuvaiheessa huomioni kohdistui erityisesti siihen, millaisia eettisiä ongelmia tekoälyn käyttöön lastensuojelussa voi liittyä. Tutkimuksen edetessä näkökulma laajeni kuitenkin yksittäisistä eettisistä ongelmista myös siihen, millaisissa tehtävissä tekoälyä käytetään, millaiseen tietoon sen toiminta perustuu ja millaisessa asemassa sen tuottama tieto on suhteessa sosiaalityöntekijän ammatilliseen harkintaan. Tämä laajempi tarkastelu osoitti, ettei tekoälyn eettisyyttä voida arvioida irrallaan siitä toimintaympäristöstä, jossa sitä käytetään. Tutkielman tekeminen on myös vahvistanut näkemystä siitä, että tekoälyn hyödyntämistä sosiaalityössä ei voida tarkastella pelkästään teknologian näkökulmasta. Sen käyttöönotto vaikuttaa samalla siihen, miten ammatillista harkintaa käytetään, miten asiakkaiden oikeudet toteutuvat ja millä tavoin sosiaalityön perustehtävää toteutetaan osana palvelujärjestelmää.

Kokonaisuutena tutkielman tulokset osoittavat, että tekoälyn eettisyyttä lastensuojelussa tulee tarkastella suhteessa lastensuojelun perustehtävään ja sen keskeisiin arvoihin. Sosiaalityössä päätöksiä tehdään tilanteissa, joissa asiakkaiden elämäntilanteet ovat yksilöllisiä, muuttuvia ja usein epävarmoja. Sosiaalityöntekijän tehtävänä on arvioida päätösten vaikutuksia erityisesti haavoittuvassa asemassa olevien ihmisten näkökulmasta sekä perustella ratkaisujaan eettisesti. Tekoälyn käyttöönotto ei poista tätä vastuuta, vaan voi tehdä sen entistä merkityksellisemmäksi. Tämän vuoksi tekoälyn käyttöä tulisi arvioida aina suhteessa siihen, tukeeko se sosiaalityöntekijän ammatillista harkintaa ja asiakkaan etua vai vaarantaako se niiden toteutumisen.

LÄHTEET

Aalto-yliopisto (2019) Tekoäly tunnistaa harvinaiset sairaudet ja lastensuojeluasiakkuuden riskitekijät. https://www.aalto.fi/fi/uutiset/tekoaly-tunnistaa-harvinaiset-sairaudet-ja-lastensuojeluasiakkuuden-riskitekijat

Agbana, Tonbara Mike (2025) Ethical considerations in using predictive AI for risk assessment in child protection social work. International Journal of Research Publication and Reviews, 6(8), 1648–1663. https://doi.org/10.55248/gengpi.6.0825.2922

Bardy, Marjatta (toim.) (2013) Lastensuojelun ytimissä (4. uud. p.). Terveyden ja hyvinvoinnin laitos.

Bick, Alexander, Blandin, Adam, & Deming, David J. (2024) The rapid adoption of generative AI. *NBER Working Paper No. 32966*. National Bureau of Economic Research. https://www.nber.org/papers/w32966

Brandão, Pedro Ramos (2025) Artificial intelligence as a transformative technology: Societal implications and future trajectories. *Societies, 6*(8), Article 190. https://www.nber.org/system/files/working_papers/w32966/w32966.pdf

Bull, Claudia, Kisely, Steve, Betts, Kim & Hu, Yanan (2025) Understanding the development, performance, fairness, and transparency of machine learning models used in child protection prediction: A systematic review. Child Abuse & Neglect, 169, 107630. https://doi.org/10.1016/j.chiabu.2025.107630

Clardy, Alan (2022) What Can We Know About the Future? Epistemology and the Credibility of Claims About the World Ahead." Foresight 24: 1–18.

Devlieghere, Jochen, Gillingham, Philip & Roose, Rudi (2022) Dataism versus relationshipism: A social work perspective. Nordic Social Work Research, 12(3), 328–338. https://doi.org/10.1080/2156857X.2022.2052942

Digital Health Finland (2026) AI in healthcare and social services: Principles for development and use in Finland.
https://digifinland.fi/wp-content/uploads/2026/02/AI-in-Healthcare-and-Social-Services.pdf

El Naqa, Issam & Martin J. Murphy (2015) What Is Machine Learning? In Machine Learning in Radiation Oncology, edited by I. El Naqa, R. Li, and M. Murphy, Springer, Cham. https://doi-org.ezproxy.ulapland.fi/10.1007/978-3-319-18305-3_1

Forrester, Donald (2024) The Enlightened Social Worker: An Introduction to Rights-Focused Practice. Policy Press.

Foucault, Michel (1980) Power & knowledge. Selected interviews and other writings 1972–1977 by Michel Foucault. Toim. Colin Gordon. New York: Pantheon Books.

Friedman, Batya & Hendry, David G. (2019) Value Sensitive Design: Shaping Technology with Moral Imagination. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/7585.001.0001.

Gabriels, Anne Elise, Kuiper, Chris & Harder, Annemiek T. (2025) Algorithmic tools to help social workers in child protection service make well-informed decisions: A scoping review. Child Abuse & Neglect, 169, 107678. https://doi.org/10.1016/j.chiabu.2025.107678

Gillingham, Philip (2016) Predictive risk modelling to prevent child maltreatment and other adverse outcomes for service users: Inside the "black box" of machine learning. Journal of Public Child Welfare, 11(2), 1–23. https://doi.org/10.1080/15548732.2016.1255697

Goldkind, Lauri (2021) Social Work and Artificial Intelligence: Into the Matrix. Social Work, 66(4), 372–374. https://doi.org/10.1093/sw/swab028

Goldkind, Lauri, Wolf, Lea, Glennon, Alexis, Rios, Juan & Nissen, Laura (2023) The end of the world as we know it? ChatGPT and social work. Social Work, 69(1), 103–105. https://doi.org/10.1093/sw/swad044

Grant, Maria J. & Booth, Andrew. (2009) A typology of reviews: An analysis of 14 review types and associated methodologies. Health Information & Libraries Journal, 26(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Hall, Seventy F., Sage, Melanie, Scott, Carol F. & Joseph, Kenneth (2024) A systematic review of sophisticated predictive and prescriptive analytics in child welfare: Accuracy, equity, and bias. *Child and Adolescent Social Work Journal, 41*(6), 831–847. https://doi.org/10.1007/s10560-023-00931-2

Haslam, Nick (2006) Dehumanization: An integrative review. Personality and Social Psychology Review, 10(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

HE 237/2018 vp: Hallituksen esitys eduskunnalle laiksi lastensuojelulain muuttamisesta. Viitattu 13.3.2026. https://www.eduskunta.fi/FI/vaski/KasittelytiedotValtiopaivaasia/Sivut/HE_237+2018.aspx

Healy, Karen (2014) Social Work Theories in Context: Creating Frameworks for Practice. United Kindom: Palgrave Macmillan.

Hietamäki, Johanna (2015) Lastensuojelun alkuarvioinnin vaikutukset vanhempien näkökulmasta. Jyväskylän Yliopisto.

Hirschovits-Gerz, Tanja & Weckroth, Niina (2024) Lastensuojelun jälkihuollon tilannekuva 2024. Tutkimuksesta tiiviisti 47/2024. Terveyden ja hyvinvoinnin laitos. https://urn.fi/URN:ISBN:978-952-408-422-2

Hjaltalín, Illugi Torfason & Sigurdarson, Hallur Thor (2024) The strategic use of AI in the public sector: A public values analysis of national AI strategies. Government Information Quarterly, 41(1), 101914. https://doi.org/10.1016/j.giq.2024.101914

Hämeen ammattikorkeakoulu (n.d.) AI Tuki!Tekoäly sosiaalihuollon arjen tukena. https://www.hamk.fi/projektit/ai-tuki/

International Federation of Social Workers, International Association of Schools of Social Work, & United Nations Centre for Human Rights. (1994) Human rights and social work: A manual for schools of social work and the social work profession. United Nations.

Jahnukainen, Johannes & Paasivirta, Annukka. (2022) Tekoja ja tahtotilaa – Mitä lapsen osallisuuden vahvistaminen vaatii lastensuojelussa? Teoksessa Laura Yliruka, Pia Eriksson, Liisa Jokinen & Kaisa Pasanen (toim.), Kohti monitoimijaista lastensuojelua hyvinvointialueilla (THL Työpaperi 52/2022, 99–110). Terveyden ja hyvinvoinnin laitos. https://urn.fi/URN:ISBN:978-952-343-984-9

Jokinen, Arja (2016) Asiakkaan ja sosiaalityöntekijän välinen suhde. Teoksessa Maritta Törrönen (toim.), Vastavuoroinen sosiaalityö. Helsinki: Gaudeamus, 138–147.

Kallinen, Timo & Kinnunen, Taina (2021) Etnografia. Teoksessa J. Vuori (toim.) Laadullisen tutkimuksen verkkokäsikirja. Tampere: Yhteiskuntatieteellinen tietoarkisto. Saatavissa: https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/teoreettis-metodologiset-viitekehykset/etnografia/. [Viitattu 3.10.2026].

Kananoja, Aulikki, Pohjola, Anneli, Lähteinen Merja (toim.) (2022) Sosiaalityö muuttuvassa yhteiskunnassa. 3. uudistettu painos. Helsinki: Gaudeamus.

Kauhanen, Antti, Pajarinen, Mika & Rouvinen, Petri (2023) Generatiivisen tekoälyn vaikutuksista (ETLA Muistio Brief No. 128). Elinkeinoelämän tutkimuslaitos. https://www.etla.fi/wp-content/uploads/ETLA-Muistio-Brief-128.pdf

Kawakami, Anna, Sivaraman, Venkatesh, Cheng, Hao-Fei, Stapleton, Logan, Cheng, Yanghuidi, Qing, Diana, Perer, Adam, Wu, Zhiwei Steven, Zhu, Haiyi & Holstein, Kenneth (2022) Improving human-AI partnerships in child welfare: Understanding worker practices, challenges, and desires for algorithmic decision support. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, 1–18. https://doi.org/10.1145/3491102.3517439

Kim, Sook Hyun & Yoon, Yoewon (2026) Child protection in the digital age in South Korea: Professionals' perceptions on artificial intelligence (AI) and big data efficacy and ethical challenges. Child & Family Social Work.

Kivipelto, Minna (2025) Voisiko tekoäly parantaa tukea tarvitsevien lapsiperheiden palvelujen vaikuttavuutta? ITLA-blogi. https://itla.fi/blogi/voisiko-tekoaly-parantaa-tukea-tarvitsevien-lapsiperheiden-palvelujen-vaikuttavuutta/

Laaksonen, Maarit, Kuusisto-Niemi, Sirpa & Saranto, Kaija (2010) Sosiaalityön tiedonhallinnan tutkimus – kirjallisuuskatsaus tutkimusmetodina. Finnish Journal of eHealth and eWelfare, 2(1), 26–32.

Lamponen, Tuuli (2022) *Kiireellisen sijoituksen päätöksenteko lastensuojelun sosiaalityöntekijän työnä*. Tampere: Tampere University. Tampere University Dissertations 644. https://urn.fi/URN:ISBN:978-952-03-2509-1

Landau, Aviv Y. Ferrarello, Susi, Blanchard, Ashley, Cato, Kenrick, Atkins, Nia, Salazar, Stephanie, Patton, Desmond U & Topaz, Maxim (2022) Developing machine learning-based models to help identify child abuse and neglect: Key ethical challenges and recommended solutions. Journal of the American Medical Informatics Association, 29(3), 576–580. https://doi.org/10.1093/jamia/ocab286

Lapin yliopisto (2025) Eettinen ja vastuullinen tekoäly sosiaalityössä: muuttuvien tietokäytäntöjen seuraukset asiakkaille, sosiaalityöntekijöille ja organisaatioille. Lapin yliopisto.

Lastensuojelulaki 417/2007. https://www.finlex.fi/fi/laki/ajantasa/2007/20070417

Lehtiniemi, Tuukka (2023) Contextual social valences for artificial intelligence: Anticipation that matters in social work. Information, Communication & Society, 27(6), 1110–1125. https://doi.org/10.1080/1369118X.2023.2234987

Livingston, Steven & Risse, Mathias (2019) The Future Impact of Artificial Intelligence on Humans and Human Rights. Ethics & International Affairs 33(2), 141–158. https://doi.org/10.1017/S089267941900011X.

Ministry of Social Affairs and Health (2024) Tekoälyn hyödyntäminen sosiaali- ja terveydenhuollossa: Ohjeistus ja periaatteet. Digital Health Finland. https://digifinland.fi/wp-content/uploads/2024/06/stm-Tekoalyn-hyodyntaminen-sotessa_24052024.pdf

Lipsky, Michael (2010) Street-Level Bureaucracy: Dilemmas of the Individual in Public Services. 30th Anniversary Expanded Edition. New York: Russell Sage Foundation.

Malja, Marja, Puustinen-Korhonen, Aila, Petrelius, Päivi & Eriksson, Piia (toim.) 2019. Lastensuojelun laatusuositus. Sosiaali- ja terveysministeriön julkaisuja 2019:8. Helsinki: Sosiaali- ja terveysministeriö.

Nerg, Riitta (2024) Töissä tekoälyn kanssa: Tietotyötä tukevat tekoälysovellukset sosiaalityössä [Väitöskirja tai tutkimusraportti]. Helsingin yliopisto

Parton, Nigel, & Kirk, Stuart A. (2009) The nature and purposes of social work. Teoksessa I. Shaw, K. Briar-Lawson, J. Orme, & R. Ruckdeschel (toim.), The SAGE handbook of social work research (ss. 23–36). SAGE Publications.

Peled, Einat & Leichtentritt, Ronit (2002) The Ethics of Qualitative Social Work Research. Qualitative Social Work, 1(2), 145–169.

Pietarinen, Juhani (2002) Eettiset perusvaatimukset tutkimustyössä. Teoksessa Sakari Karjalainen, Veikko Launis, Risto Pelkonen & Juhani Pietarinen (toim.) Tutkijan eettiset valinnat. Helsinki: Gaudeamus, 58–68.

Pirkanmaan hyvinvointialue (2025) Tekoäly avuksi – uusi kokeilu lapsiperheiden tueksi. 16.10.2025, päivitetty 21.10.2025.

Pirkanmaan hyvinvointialue (2026) Tekoäly avuksi lasten palvelutarpeen arvioinnin yhteenvedon käsittelyyn. 5.5.2026.

Pohjola, Anneli (2010) Asiakkuus sosiaalityössä. Helsinki: Gaudeamus.

Pösö, Tarja & Lönnqvist, Tove & Petrelius, Päivi & Salminen, Matti & Väkeväinen, Niina (2019) Väkivallaton lapsuus – toimenpidesuunnitelma lapsiin kohdistuvan väkivallan ehkäisystä 2020–2025. Sosiaali- ja terveysministeriö. https://urn.fi/URN:ISBN:978-952-00-4123-6. Viitattu 29.1.2026.

Raso, Filippo A., Hilligoss, Hannah, Krishnamurthy, Vivek, Bavitz, Christopher & Kim, Levin Yerin (2018) Artificial Intelligence & Human Rights: Opportunities & Risks. Berkman Klein Center Research Publication No. 2018-6. Harvard University. https://ssrn.com/abstract=3259344.

Ratner, Helene Friis & Elmholdt, Kasper (2023) Algorithmic Constructions of Risk: Anticipating Uncertain Futures in Child Protection Services. Big Data & Society 10 (2). https://journals.sagepub.com/doi/10.1177/20539517231186120

Ratner, Friis & Schrøder, Helene Ida (2024). Ethical plateaus in Danish child protection services: The rise and demise of algorithmic models. Science and Technology Studies, 37(3), 44–61. https://doi.org/10.23987/sts.126011

Rauhala, Pirkko-Liisa & Virokannas, Elina (2011) Sosiaalityön tutkimuksen etiikka, opettaminen ja tietoarvo. Teoksessa Pehkonen Aini & Väänen-Fomin Marja (toim.): Arvot ja Etiikka. Jyväskylä: PS-kustannus, 235–255.

Raunio, Kyösti (2009) Olennainen sosiaalityössä. Helsinki: Gaudeamus.

Reamer, Frederic G. (2023) Artificial intelligence in social work: Emerging ethical issues. International Journal of Social Work Values and Ethics, 20(2), 52–71. https://doi.org/10.55521/10-020-205

Roberts, Yvonne H., O'Brien, Kirk, & Pecora, Peter J. (2018) Considerations for implementing predictive analytics in child welfare (pp. 2–3). Casey Family Programs. https://www.casey.org/media/Considerations-for-Applying-Predictive-Analytics-in-Child-Welfare.pdf

Rodriguez, M. Y., DePanfilis, D., & Lanier, P. (2019) Bridging the gap: Social work insights for ethical algorithmic decision-making in human services. IBM Journal of Research and Development, 63(4–5), 8:1–8:8. https://doi.org/10.1147/JRD.2019.2934047

Räty, Tapio (2023) Lastensuojelulaki: Käytäntö ja soveltaminen (5. uudistettu painos.). Edita.

Salminen, Ari (2011) Mikä on kirjallisuuskatsaus? Johdatus kirjallisuuskatsauksen tyyppeihin ja hallintotieteellisiin sovelluksiin. Vaasan yliopiston julkaisuja. Opetusjulkaisuja 62, julkisjohtaminen 4. Vaasa 2011.

Saxena, Devansh & Guha, Shion (2024) Algorithmic harms in child welfare: Uncertainties in practice, organization, and street-level decision-making. ACM Journal on Responsible Computing, 1(1), 1–32. https://doi.org/10.1145/3616473

Sengar, Sandeep Singh, Bin Hasan, Affan, Kumar, Sanjay, & Carroll, Fiona (2024) Generative artificial intelligence: A systematic review and applications. Multimedia Tools and Applications, 84, 23661–23700. https://doi.org/10.1007/s11042-024-20016-1

Smith, Carol A. (2004) Social work and the modernisation agenda: Confidence in systems or trust in moral agents? Social Work & Social Sciences Review, 11(2).

Snyder, Hannah (2019) Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Stanford Institute for Human-Centered Artificial Intelligence. (2022) Brief definitions of key terms in AI. Stanford University. Stanford HAI: Brief Definitions of Key Terms in AI

Talentia (2024) Arki, arvot ja etiikka. Sosiaalialan ammattihenkilöiden eettiset ohjeet. 11–20.

Taylor, Carolyn. & White, Sue (2006) Knowledge and reasoning in social work: Educating for humane judgement. British Journal of Social Work

THL (2025) Lastensuojelun käsikirja. Mitä on lastensuojelu? https://thl.fi/julkaisut/kasikirjat/lastensuojelun-kasikirja/tyoprosessi/mita-on-lastensuojelu Viitattu 16.1.2026

THL (2026a) Hyvinvointialueiden tietohallinnosta vastaavien tulevaisuuden visioita digitalisaatiokehityksestä https://www.julkari.fi/server/api/core/bitstreams/02d681d9-9b6f-41bd-8c6e-a2765ee847f7/content

THL (2026b) Tekoäly nousee hyvinvointialueiden digivisioiden ytimeen. https://thl.fi/-/tekoaly-nousee-hyvinvointialueiden-digivisioiden-ytimeen

Tomiţă, Mihaela, Biriş, Diana Laura & Feher, Claudia (2024) Ethical considerations on the integration of technology in social work practice for enhancing resilience and promoting transformative action. Social Work Review, 1, 13–25.

Tzirides, Anastasia Olga, Saini, Akash, Zapata, Gabriela, Searsmith, Duane, Cope, Bill, Kalantzis, Mary, Castro, Vania, Kourkoulou, Theodora, Jones, Jon, Abrantes da Silva, Rodrigo, Whiting, Jen, & Kastania, Nikoleta Polyxenia (2023) *Generative AI: Implications and applications for education*.

Turtiainen, Kati, Jäppinen, Maija & Stamm, Ingo (2021) Ihmisoikeudet aikuissosiaalityön lähtökohtana. Teoksessa Aila-Leena Matthies, Anu-Riina Svenlin & Kati Turtiainen (toim.) Aikuissosiaalityö – tieto, käytäntö ja vaikuttavuus. Helsinki: Gaudeamus, 27–41

United Nations Economic Commission for Europe (2025) Use of generative AI for communication in official statistics. UNECE. https://unece.org/sites/default/files/2025-01/HLG2024_Use%20of%20Gen%20AI%20for%20communication%20in%20official%20statistics_0.pdf

University of Helsinki (2019) Can artificial intelligence be sued? Who is responsible when AI makes decisions? https://www.helsinki.fi/en/news/economics/can-artificial-intelligence-be-sued-who-responsible-when-ai-makes-decisions

Vaasan ammattikorkeakoulu (2025) Uusi tutkimushanke tekoälyn mahdollisuuksista sosiaalihuollossa – tavoitteena sujuvampi työ ja parempi tuottavuus. 23.6.2025. https://www.vamk.fi/ajankohtaista/uusi-tutkimushanke-tekoalyn-mahdollisuuksista-sosiaalihuollossa

Valtioneuvosto (2023) Vahva ja välittävä Suomi: Pääministeri Petteri Orpon hallituksen ohjelma 20.6.2023. Valtioneuvoston julkaisuja 2023:58.

https://julkaisut.valtioneuvosto.fi/server/api/core/bitstreams/a89b334b-2f1e-4b41-bd97-ff91a7449a7a/content

Vilkka, Hanna (2023) Kirjallisuuskatsaus metodina, opinnäytetyön osana ja tekstilajina. Helsinki: Art House. E-kirja.

Wilkins, David & Benett, Verity (2026) Making accurate judgements in child welfare: Comparing ChatGPT with qualified social workers. Child & Family Social Work, 31(2), 703–712. https://doi.org/10.1111/cfs.13304

Yliruka, Laura, Harrikari, Timo, Forsell, Martta, Kuoppala, Tuula, Purhonen, Emmi & Salmela, Marianna (2022) Lapsiperheiden ja lastensuojelun sosiaalityön organisoiminen sekä lastensuojelulain mukaisen henkilöstömitoituksen toteutuminen: Kuntakyselyjen tulokset. Työpaperi 44/2022. Terveyden ja hyvinvoinnin laitos. https://urn.fi/URN:ISBN:978-952-343-928-3

Yu, Meng-Hsuan & Rose, Roderick A. (2026) Algorithmic-assisted decision-making tools in child welfare practice: A systematic review. Research on Social Work Practice, 36(4), 382–397. https://doi.org/10.1177/10497315251350933

LIITTEET

Liite 1. Aineiston hakusanat

Teemat	**Suomenkieliset hakusanat**	**Englanninkieliset hakusanat**
Lastensuojelu	lastensuojelu, sosiaalipalvelut, sijaishuolto	child welfare, child protection, child-related work
Tekoäly	tekoäly, generatiivinen tekoäly, koneoppiminen	artificial intelligence, generative artificial intelligence, machine learning, knowledge technology
Etiikka	etiikka, eettisyys, ammattietiikka, arvofilosofia, moraali, normit	ethics, ethicality, ethical, moral, norms

Liite 2. Aineiston valintaprosessi: tunnistaminen, seulonta ja valittu aineisto

Tutkimusten valinta tietokannoista		
Tunnistaminen	Tunnistettu aineisto: **LUC-Finna (n = 2)** **Duodecim (n=1)** **LUC-Finnan kansainväliset e-aineistot (n=25)** **AgeLine (n=2)** **Social Science Database (n=17)** **Scopus (n=2)** **Web of Science (n=134)** **Journal.fi (n=22)** **ChatGPT (n=8)** **N=213**	Kaksoiskappaleena poistettu aineisto **(n=0)**
Seulonta	Tiivistelmän perusteella seulottu aineisto **(n = 213)**	Poissuljettu aineisto **(n = 143)**
	Johdannon ja koko tekstin silmäilyn perusteella seulottu aineisto **(n = 70)**	Poissuljettu aineisto **(n =46)**
	Koko tekstin perusteella seulottu aineisto **(n=24)**	Poissuljettu aineisto **(n=7)**
	Aikarajauksen perusteella seulottu aineisto **(n = 17)**	Poissuljettu aineisto **(n=5)**
Valitut	Sisällytetty aineisto **(n = 12)**	

Liite 3. Lopullinen aineisto

1.	2026	**Kim, Sook Hyun & Yoon, Yoewon**: Child Protection in the Digital Age in South Korea: Professionals' Perceptions on Artificial Intelligence (AI) and Big Data Efficacy and Ethical Challenges.	Child & Family Social Work	Empiirinen alkuperäistutkimus
2.	2025	**Bull, Claudia; Kisely, Steve; Betts, Kim & Hu, Yanan:** Understanding the development, performance, fairness, and transparency of machine learning models used in child protection prediction: A systematic review	Child Abuse & Neglect	Systemaattinen kirjallisuuskatsaus
3.	2025	**Gabriels, Anne Elise; Kuiper, Chris & Harder, Annemiek T:** Algorithmic tools to help social workers in child protection service make well-informed	Child Abuse & Neglect	Kartoittava kirjallisuuskatsaus
4.	2025	**Meng-Hsuan Yu & Roderick A. Rose:** Algorithmic-Assisted Decision-Making Tools in Child Welfare Practice: A Systematic Review	Research on Social Work Practice	Systemaattinen kirjallisuuskatsaus
5.	2025	**Agbana Tonbara Mike:** Ethical Considerations in Using Predictive AI for Risk Assessment in Child Protection Social Work	International Journal of Research Publication and Reviews	Teoreettinen analyysi?
6.	2025	**Wilkins, David & Benett, Verity:** Making Accurate Judgements in Child Welfare: Comparing ChatGPT With Qualified Social Workers	Child & Family Social Work	Empiirinen alkuperäistutkimus
7.	2024	**Ratner, Friis & Schrøder, Helene Ida:** Ethical Plateaus in Danish Child Protection Services: The Rise and Demise of Algorithmic Models	Science and Technology Studies	Empiirinen alkuperäistutkimus
8.	2024	**Saxena, Devansh & Guha, Shion:** Algorithmic Harms in Child Welfare: Uncertainties in	ACM Journal on Responsible Computing	Kirjallisuuskatsaus

		Practice, Organization, and Street-level Decision making		
9.	2023	**Hall, Seventy F.; Sage, Melanie; Scott, Carol F. & Joseph Kenneth:** A Systematic Review of Sophisticated Predictive and Prescriptive Analytics in Child Welfare: Accuracy, Equity, and Bias	Child & Adolescent Social Work Journal	Systemaattinen kirjallisuuskatsaus
10.	2022	**Lehtiniemi, Tuukka:** Contextual social valences for artificial intelligence: anticipation that matters in social work	INFORMATION, COMMUNICATION & SOCIETY	Empiirinen alkuperäistutkimus
11.	2022	**Kawakami, Anna; Sivaraman, Venkatesh; Cheng, Hao-Fei; Stapleton, Logan; Cheng, Yanghuidi; Qing, Diana; Perer, Adam; Wu, Zhiwei Steven; Zhu, Haiyi & Holstein, Kenneth:** Improving Human-AI Partnerships in Child Welfare: Understanding Worker Practices, Challenges, and Desires for Algorithmic Decision Support	Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems	Empiirinen alkuperäistutkimus
12.	2022	**Landau, Aviv Y.; Ferrarello, Susi; Blanchard, Ashley; Cato, Kenrick; Atkins, Nia; Salazar, Stephanie; Patton, Desmond U & Topaz, Maxim:** Developing machine learning-based models to help identify child abuse and neglect: key ethical challenges and recommended solutions	Journal of the American Medical Informatics Association	Teoreettinen analyysi