



LAPIN YLIOPISTO  
UNIVERSITY OF LAPLAND

# Algoritmit avuksi sosiaalityöhön

Tekoälystä ratkaisu sosiaalityön arkeen?

Sosiaalityö  
Pro gradu -tutkielma

Elina Vikman

3.3.2026  
Lapin yliopisto



Lapin yliopisto

Tiedekunta: Yhteiskuntatieteiden tiedekunta

Työn nimi: Algoritmit avuksi sosiaalityöhön. Tekoälystä ratkaisu sosiaalityön arkeen?

Tekijä: Elina Vikman

Koulutusohjelma / opetuskokonaisuus / oppiaine: Sosiaalityö

Ohjaaja: Yliopistonlehtori Tarja Orjasniemi

Työn laji: Pro gradu -tutkielma

Sivumäärä: 60 sivua

Vuosi: 2026

Tiivistelmä:

Pro gradu -tutkielmani tavoitteena oli kuvata ja analysoida, miten tekoälyä voidaan hyödyntää sosiaalityössä ja millaisia mahdollisuuksia sekä rajoitteita tekoälyn käyttö tuo asiakastyön tiedonmuodostukseen, päätöksenteon tukeen ja ammatilliseen harkintaan. Tarkastelun kohteena oli erityisesti se, millaista tietoa tekoäly tuottaa sosiaalityön prosesseissa ja miten ammattilaiset voivat arvioida tämän tiedon laatua, läpinäkyvyyttä ja käyttökelpoisuutta.

Tutkimus oli laadullinen kirjallisuuskatsaus. Aineisto koottiin vuosina 2020–2025 julkaistuista vertaisarvioituista tutkimusartikkeleista, joissa on kuvattu tekoälyn käyttöönottoa erilaisissa sosiaalityön konteksteissa. Aineisto analysoitiin aineistolähtöisellä sisällönanalyysillä ja temaattisella analyysillä, joiden avulla tunnistettiin tekoälyn tukemia sosiaalityön tiedonmuodostuksen tehtäviä, tekoälypohjaisen tiedon vahvuuksia ja rajoitteita sekä vaikutuksia ammatilliseen harkintaan.

Tulosten mukaan tekoäly voi tukea sosiaalityötä erityisesti laajojen tietomassojen jäsentämisessä, tiedon poiminnassa ja tiivistämisessä sekä päätöksenteon valmistelussa, tukien esimerkiksi riskien ja vaihtoehtojen tarkastelua. Hyödyt kytkeytyvät työn sujuvoittamiseen, uudenlaisten näkökulmien tuottamiseen ja arviointien yhdenmukaistumiseen. Keskeisiä rajoitteita ovat kuitenkin tiedon mahdolliset asiavirheet, vinoumat ja datan laatuun liittyvät riskit, tekoälymenetelmien puutteellinen läpinäkyvyys sekä luottamuksellisuuteen liittyvät haasteet. Tekoäly ei voi korvata sosiaalityöntekijän eettistä ja kontekstisidonnaista harkintaa, vaan sen käyttö edellyttää osaamista, selkeitä käytäntöjä ja ihmisen tekemää varmistamista.

Avainsanat: sosiaalityö, tekoäly, tiedonmuodostus, kriittinen realismi, päätöksenteko

## Sisällys

<b>1</b>	<b>Johdanto</b>	<b>1</b>
<b>2</b>	<b>Sosiaalityöntekijät uuden äärellä – tekoäly pelottaa ja houkuttaa</b>	<b>4</b>
	2.1 Tietojärjestelmät haastavat sosiaalityön tiedonmuodostuksen	4
	2.2 Tekoäly rantautuu sosiaalityöhön	7
	2.3 Tekoäly työkaluna asiantuntijatyössä	10
<b>3</b>	<b>Sosiaalityön tiedon ja todellisuuden ulottuvuudet</b>	<b>13</b>
	3.1 Asiantuntijuus paineessa: liikaa tietoa, liian vähän varmuutta	13
	3.2 Kriittinen realismi todellisuuden jäsentäjänä	15
<b>4</b>	<b>Tutkimuksen toteutus</b>	<b>20</b>
	4.1 Tutkimuksen tavoite ja tehtävä	20
	4.2 Aineiston kerääminen	21
	4.3 Aineiston analyysi ja metodologia	30
	4.4 Tutkimuseettiset kysymykset	32
<b>5</b>	<b>Tekoälyn tuottama tieto ja tuotokset sosiaalityössä</b>	<b>34</b>
	5.1 Lausunnot, arviot ja ehdotukset tekoälykollegan tuottamana	34
	5.2 Koneoppiminen sosiaalisten ilmiöiden selittämisessä ja ennustamisessa	36
	5.3 Tiedonhaku ja keskustelu tekoälyn kanssa	40
<b>6</b>	<b>Tekoälypohjainen tiedonmuodostusprosessi: hyödyt ja rajoitteet</b>	<b>42</b>
	6.1 Tekoäly vakuuttaa ja harhauttaa	42
	6.2 Tekoälypohjainen tiedonmuodostusprosessi edellyttää läpinäkyvyyttä	45
<b>7</b>	<b>Ammatillinen harkinta ja vastuu tekoälyn aikakaudella</b>	<b>48</b>
	7.1 Päätöksenteon delegoinnin rajat	48
	7.2 Tekoäly ei ole moraalinen toimija	51
<b>8</b>	<b>Yhteenveto ja pohdinta</b>	<b>53</b>
<b>9</b>	<b>Lähteet</b>	<b>56</b>

## 1 Johdanto

Sitran, eli Suomen itsenäisyyden juhlarahaston, julkaisemia katsauksia seurataan tiiviisti suomalaisessa mediassa, liike-elämässä ja julkishallinnossa. Sitran vuonna 2026 julkaisemat megatrendit kuvaavat hyvin toimintaympäristöä, jonka puristuksessa julkisen palveluiden organisoinnista, taloudesta ja kehittämisestä vastaavat tahot joutuvat tasapainoilemaan. Sitran tunnistamia megatrendejä ovat suomalaisen väestön ikääntyminen, demokraattisen maailmanjärjestyksen horjumisen, ympäristökriisi sekä tekoälymurros (Sitra 2026).

Sosiaali- ja terveydenhuollon palvelutarpeet kasvavat ja hyvinvointialueilta edellytetään aiempaa vahvempaa tehokkuutta ja tuottavuutta nykyisillä tai jopa vähenevillä resursseilla. Väestön ikääntyessä työvoiman saatavuus heikkenee ja hyvinvointivaltion ylläpito muuttuu entistä kalliimmaksi (Valtiovarainministeriö 2025). Julkista sektoria kritisoidaan hitaudesta ottaa käyttöön uusia teknologioita tuottavuuden parantamiseksi. Uusien ratkaisujen käyttöönottoon suhtaudutaan varovaisesti, osin tiukan tietosuojasääntelyn vuoksi. Pilottien ja kokeilujen käynnistäminen on myös vaikeaa, koska niihin ei aina löydy organisaatioista tekijöitä valmiiksi kireän resurssitilanteen takia.

Olen työskennellyt sosiaali- ja terveydenhuollon ICT-kehittämistehtävissä usean vuoden ajan. Toimiessani projektipäällikkönä asiakas- ja potilastietojärjestelmien käyttöönotoissa olen saanut mahdollisuuden seurata paraatipaikalta, kuinka rakenteinen kirjaaminen tulee osaksi asiakas- ja potilastyön arkea järjestelmämuudistusten myötä. Olemme myös suunnitelleet muutosjohtamisen toimenpiteitä yhdessä palveluiden kanssa heidän ottaessaan käyttöön asiakkaille suunnattuna digitaalisia ratkaisuja. Näiden avulla asiakkaille on voitu laajemmin tarjota omatoimisen asioinnin mahdollisuuksia ja sitä kautta vahvistaa osallisuuden kokemusta heidän asioidessaan sosiaali- ja terveydenhuollon palveluissa.

Uuden teknologian käyttöönotto ei koskaan ole saumatonta, vaan se haastaa asiakkaat, ammattilaiset ja koko organisaation muuttamaan tuttuja käytäntöjä. Se edellyttää uusien toimintamallien omaksumista ja jopa kokonaisten palveluprosessien uudelleenorganisointia. Uuden teknologian opettelu saattaa turhauttaa ja nostaa epävarmuuden tunteita pintaan. Tietotekniset yksityiskohdat voidaan kokea vieraina ja pelottavina. Myös esihenkilöt ja johto joutuvat sietämään epävarmuutta; käyttöönoton hyödyt eivät usein realisoidu heti, vaan vasta sen jälkeen, kun oppimiselle – ja usein myös erehtymiselle - on annettu riittävästi aikaa.

Sosiaalityön digitalisaatio ei näyttäydä enää pelkkänä tulevaisuudentutkijoiden visiointina vaan se on jo tullut luontevaksi osaksi jokaisen ammattilaisen arkea (mm. Salovaara 2024a; Kääriäinen 2024). Paineet teknologian hyödyntämiseen ja laajentamiseen kohdistuvat niin yksittäisiin työntekijöihin osana laadukasta kirjaamista ja palvelua, kuin sosiaalipalveluiden johtajiin tiukentuneiden tuottavuusvaatimusten ja tiedolla johtamisen odotusten kautta.

Tekoälyratkaisut ovat viime vuosina tulleet laajasti osaksi monia yhteiskunnan sektoreita sosiaali- ja terveydenhuollon ulkopuolella. Asiakaspalvelussa tekoälypohjaiset chatbotit ovat yleistyneet tarjoamaan automatisoitua neuvontaa ja ohjausta. Myös viihteen ja kuluttajapalvelujen puolella tekoäly on arkipäivää: yksityiset toimijat hyödyntävät digitaalisten alustojen kautta kerättyä käyttäjätietoa esimerkiksi kohdennetun markkinoinnin ja sosiaalisen median sisältöjen optimoinnin välineinä. Suoratoistopalvelut, kuten Netflix ja Spotify, hyödyntävät algoritmeja yksilöllisten katselu- ja kuuntelusuositusten muodostamisessa.

Tekoälyn hyödyntämiseen liitetään kuitenkin myös merkittäviä riskejä. Autoritäärisissä valtioissa algoritmit on valjastettu yksilöiden liikkeiden valvontaan ja yhteiskunnallisen kontrollin välineeksi. Huolta herättää myös kysymys suurten teknologiajättien keräämien henkilötietojen käyttötarkoituksesta: käyttötarkoituksiin liittyvät epäselvyydet nähdään yksityisyyden suojaa heikentävinä käytäntöinä sekä kaupallisia ja poliittisia valtasuhteita muokkaavina toimintatapoina (ks. Coeckelbergh 2026). Näiden mahdollisuuksien ja riskien rinnakkaiselo selittää, miksi tekoälyyn suhtaudutaan monin paikoin yhtä aikaa sekä toiveikkaasti että varauksellisesti ja miksi sen hyödyntäminen herättää perusteltuja eettisiä ja yhteiskunnallisia kysymyksiä.

Tekoälyä on esitetty julkisessa keskustelussa ratkaisuksi julkisen sektorin tuottavuushaasteisiin (DigiFinland 2025). Perinteisiä internetin hakukoneita haastavan keskustelelevan chatbotin, ChatGPT:n, ensimmäisen julkaisun myötä marraskuussa 2022 tekoäly tuli aivan tavallisten ihmisten saataville ja tarjosi mahdollisuuden kokeilla tekoälyn kykyä luonnollisen kielen ymmärtämisessä ja tekstin tuottamisessa.

Pro gradu -tutkielmani käsittelee tekoälyn hyödyntämisen mahdollisuuksia sosiaalihuollossa. Haluan tutkimuksessani tarkastella erityisesti tekoälypohjaisen tiedonmuodostuksen prosessia osana sosiaalityön tehtäviä. Tarkastelen tiedonmuodostusta sekä teknisenä prosessina että tiedon epistemologisena jäsennyksenä eli minkälaisena ammattilaiset näkevät tekoälyn

tuottaman tiedon osana sosiaalityön tehtävää. Toteutan tutkimukseni kirjallisuuskatsauksena kansainväliseen tutkimukseen. Koska teknologian hyödyntämisestä sosiaalihuollon palveluissa löytyy jo monipuolisesti tutkimuksia, rajaan tutkimustehtävässäni tarkastelun nimenomaan tekoälyratkaisuihin. Tarkastelen pro gradu -tutkielmassani erilaisissa sosiaalityön ympäristöissä toteutettuja tekoälykokeiluja ja analysoin niiden roolia osana sosiaalihuollon tiedonmuodostusta.

Artikkeleissa kuvatut esimerkit toimivat sekä empiirisinä havainnollistuksina että keskustelun avaajina. Niiden kautta analysoin millaista tietoa tekoälyratkaisut tuottavat sosiaalityön tueksi ja millaisia kokemuksia niihin on liittynyt. Näiden esimerkkien tarkastelu mahdollistaa sekä teoreettisen että empiirisen tason yhdistämisen. Artikkeleiden pohjalta voidaan eritellä, millaisena tekoälypohjainen tiedonmuodostus näyttäytyy todellisissa työtilanteissa ja missä määrin se haastaa syvällisemmin sosiaalityön ammatillista identiteettiä, arvoja, päätöksenteon periaatteita ja tiedonmuodostuksen tapoja.

Pro gradu -tutkielmani etenee siten, että luvussa kaksi taustoitetaan digitalisaation vaikutuksia sosiaalityöhön, ammattietiikkaan ja tiedonmuodostukseen. Luvussa kaksi tarkastelen myös, minkälaisia merkityksiä tekoälyyn sosiaalityössä liitetään ja miten sen hyödyllisyyttä arvioidaan suhteessa asiakastyöhön, ammattietiikkaan ja päätöksenteon autonomiaan. Tässä luvussa esittelen myös tekoälyyn liittyvät peruskäsitteet ja selvennän tutkimukseni teknologista ulottuvuutta niiltä osin, kun se on tämän aihealueen osalta tarpeellista. Koska kyseessä on sosiaalityön pro gradu -tutkielma, joudun tekemään tiukkoja rajauksia sen suhteen, kuinka perusteellisesti avaan tekoälyn toimintalogiikkaa, astumatta kuitenkaan liian syväälle tietojenkäsittelytieteen maailmaan.

Kolmas luku muodostuu teoreettisena viitekehyksenäni toimivien käsitteiden tarkastelusta. Luvun alussa tarkastelen sosiaalityön asiantuntijuutta. Selvitän, mitkä ovat asiantuntijuutta määrittävät keskeiset tiedon lajit sosiaalityössä. Sosiaalityön tiedon ulottuvuuksien tunnistaminen avaa mahdollisuuden analysoida teknologian hyötyjä ja riskejä tiedonmuodostuksen tukena. Luvun toisessa osiossa käyn läpi kriittisen realismin viitekehystä, joka tarjoaa yhden näkökulman ymmärtää ja jäsentää sosiaalista todellisuutta ja sosiaalisia ilmiöitä. Luvussa syvennän tiedonmuodostuksen teoreettista viitekehystä kriittisen realismin avulla.

Luvussa neljä esittelen tutkimuksen toteutusprosessin. Toteutan tutkimukseni laadullisen kirjallisuuskatsauksena. Luku pitää sisällään tutkimukseni metodologisten menetelmien kuvauksen. Tässä luvussa esittelen myös kirjallisuuskatsauksessa käyttämäni aineiston. Hyvän tutkimuskäytännön mukaisesti itsearvioin myös tutkimukseni luotettavuutta ja rajoitteita.

Kappaleet viisi, kuusi ja seitsemän muodostavat työni analyysiosion, jossa tarkastelen tutkimusartikkeleihin nojaten tekoälypohjaisen tiedonmuodostuksen konkreettisia jäsenyyksiä sosiaalityössä. Syvennyn yksittäisiin käyttökohteisiin kuvatakseni tekoälyn toimintalogiikkaa sekä tarkastellakseni tekoälypohjaisen tiedon tuottamisen etiikkaa ja oikeutusta. Lisäksi erittelen, millaisia hyötyjä ja mahdollisuuksia tekoälyn hyödyntämiseen liitetään, ja missä kohdin on perusteltua suhtautua sen käyttöön varauksella. Luvussa kahdeksan käyn läpi aineistoni tulokset ja esittelen johtopäätökseni, jossa yhdistän aineistoni tulokset laajempaan teoreettiseen keskusteluun tutkimastani ilmiöstä.

Toivon, että tämä tutkimus tuo osaltaan näkyvyyttä tekoälyn hyödyntämiselle ja lisää ymmärrystä tekoälyteknologioiden ja -menetelmien käytöstä. Toivon myös, että se antaa sosiaalityön ammattilaisille ja opiskelijoille sekä tietojärjestelmien kehittäjille uusia ajatuksia siitä, miten tekoälyä voisi turvallisesti ja hallitusti kehittää, hyödyntää ja arvioida, sosiaalityön erityispiirteet huomioiden.

## **2 Sosiaalityöntekijät uuden äärellä – tekoäly pelottaa ja houkuttaa**

### **2.1 Tietojärjestelmät haastavat sosiaalityön tiedonmuodostuksen**

Digitalisaatio itsessään ei ole uusi ilmiö sosiaalityössä. Tietojärjestelmien käyttö on jo viime vuosina vakiintunut osaksi sosiaalityön ammattilaisen arkea ja järjestelmillä on merkittävä rooli sosiaalihuollon tiedonmuodostusprosessissa (Salovaara 2024a; Kääriäinen 2024). Parhaimmillaan tietojärjestelmät selkeyttävät kirjaamista, auttavat tiedonhaussa ja asiakkaan kokonaiskuvan muodostamisessa sekä ohjaavat asiakasprosessin etenemistä ja siten tukevat sosiaalityöntekijän työskentelyä erilaisissa asiakastilanteissa. Toisaalta tietoteknisten järjestelmien nähdään myös kaventavan ammattilaisen omaa harkintavaltaa (Meilvang & Dahler 2024, 32).

Salovaara määrittelee (2024a, 100) tietojärjestelmän datan keruun, tallennuksen, käsittelyn, jakamisen ja monipuolisen hyödyntämisen mahdollistavaksi tekniseksi artefaktiksi. Yhtä lailla se

myös toimii tiedonmuodostusta ohjaavana ja määrittävänä teknologiana. Rakenteet ja valinnat rajaavat ja ohjaavat sitä, mitä dataa ja tietoa tietojärjestelmillä kerätään. Salovaara muistuttaa (mt., 90) tietojärjestelmien olevan ihmisten konstruktion tuloksia, sillä niiden suunnittelussa ja toteutuksessa tehdään valintoja ja kompromisseja erilaisten tiedollisten intressien välillä. Tietojärjestelmien kehittämisvaiheessa on tyypillisesti mukana ihmisiä erilaisista organisaatioista ja ammattiryhmistä. Heidän arvomaailmansa voivat olla keskenään hyvinkin erilaisia. Näistä erilaisista lähtökohdista tulevat ihmiset määrittelevät, mitä tietoa järjestelmien avulla halutaan kerätä ja mikä tieto on olennaisempaa kuin toinen (mm. Salovaara 2024a, 100; Alam 2025, 4-5). Tekoälyjärjestelmät eivät tee tähän poikkeusta, sillä ne ovat tietojärjestelmiä siinä missä perinteiset asiakas- ja potilastietojärjestelmätkin.

Tietojärjestelmien kehittämisessä tiedollisten intressien yhteensovittamista vaikeuttaa lisäksi se, että yhteiskunnalliset tietotarpeet kasvavat ja monimutkaistuvat ajan kuluessa. Järjestelmiä joudutaan jatkuvasti kehittämään tietotarpeiden tai lainsäädännön muuttuessa. Tietoa asiakkaista ja sosiaalisista ilmiöistä tarvitaan yhä useampiin tarkoituksiin ja yhteiskunnan eri tasoille: tiedon tulee palvella sekä yksilöllisten asiakastilanteiden arviointia ja työskentelyä että organisaation johtamista, palveluiden suunnittelua ja järjestämistä. Lisäksi alueelliset ja kansalliset poliittiset päätöksentekijät hyödyntävät tietoa yhteiskunnallisten kysymysten jäsentämiseksi ja päätöksenteon perustaksi. (ks. Jokiranta 2024.)

Salovaara (2024a; 2024b) esittelee sosioteknisen viitekehyksen yhtenä tapana jäsentää tietojärjestelmien ja ihmisten välistä suhdetta tiedonmuodostuksessa. Sosioteknisessä viitekehyyksessä sosiaalinen ja tekninen ovat systeemin eri osia, jotka kuitenkin toimivat vuorovaikutuksessa toistensa kanssa. Vuorovaikutus voi tarkoittaa sitä, kuinka ihminen käyttää tietojärjestelmää, mutta se myös tarkoittaa sitä, miten tietojärjestelmät rakennetaan ja mitä tietoa valitaan kerättäväksi järjestelmiin.

Kääriäinen tarkastelee (2024) sosiaalityön asiakastyön dokumentointia tietojärjestelmiin sekä teknisenä että ei-teknisenä tiedonmuodostusprosessina, joissa tietojärjestelmät ja niissä käsiteltävä tieto kytkeytyvät vahvasti yhteen. Tietoarkkitehtuurilla tarkoitetaan organisaation käyttämiä tietoja, niiden rakenteita ja suhteita toisiinsa. Tietoarkkitehtuuria ja tietojärjestelmiä tulee kehittää käsi kädessä. Asiakasasiakirjat ovat tärkeä muistamisen väline, mutta samalla ne pakottavat tekemään valintoja siitä, mitä kirjataan ja mikä jätetään kirjaamatta tietojärjestelmään. Asiakasasiakirjoihin tallennettujen tietojen jatkokäytössä on olennaista huomioida tulkinvaraisuus: samat kirjaukset voivat avautua eri ammattilaisille ja asiakkaille

eri merkityksissä. Erilaiset tulkinnat asiakirjojen teksteistä voivat vaikuttaa jopa palveluiden saamiseen tai niiden eväämiseen. Tästä syystä onkin erittäin tärkeää yhdenmukaistaa sosiaalihuollon tietoa ja yhtenäistää käsitteistöä, mikä edistää palvelujen tasalaatuisuutta ja niiden arvioitavuutta. (Mt., 2024, luku I.)

Suomessa on toteutettu ja toteutetaan merkittäviä kansallisia tietoarkkitehtuurihankkeita, joissa pyritään varmistamaan, että sosiaali- ja terveydenhuollon palveluissa muodostuneet tiedot vastaavat niille asetettuja vaatimuksia ja tietotarpeita. Sosiaalihuollon asiakasvarannon kehittäminen osana Kansaneläkelaitoksen (KELA) valtakunnallisia Kanta-palveluita on yksi keskeisistä, parhaillaan käynnissä olevista hankkeista. Kehittämisen tueksi laaditussa, sosiaalihuollon toimijoille suunnatussa Kanta-palvelujen käsikirjassa (2025) määritellään sosiaalihuollon tietoarkkitehtuurin yhdenmukaistamisen periaatteet, jotta asiakastietojen siirtyminen Kantaan voidaan toteuttaa. Tavoitteena on, että asiakastiedot ovat jatkossa sekä asiakkaiden että viranomaisten saatavilla lakien sallimassa laajuudessa.

Toinen keskeinen tapa jäsentää sosiaalityön digitaalista tiedonmuodostusta ja dokumentointia on erottaa toisistaan kaksi tietomuotoa: narratiivinen, vapaamuotoinen teksti ja rakenteinen, määrämuotoinen tieto. Asiakastietojärjestelmiä koskeva kritiikki kohdistuu usein juuri näiden tietomuotojen väliseen jännitteeseen – siihen, miten ne painottuvat, miten ne ohjaavat kirjaamista ja miten ne asetetaan toistensa vaihtoehtoiksi. Wartainen ja kumppanit (2020) tarkastelevat narratiiviseen ja rakenteiseen muotoon tallennetun tiedon tarkoituksenmukaista suhdetta ja toteavat, että sosiaalityössä tarvitaan molempia.

Sosiaalihuollon Kanta -kehitystyön yhteydessä määrämuotoisen tiedon vaatimukset ja laajuus kasvavat. Määrämuotoisen kirjaamisen riskinä nähdään näkökulman kapeneminen, järjestelmän ohjatessa ammattilaista keskittymään vain etukäteen määriteltyihin tietoihin. Tämän pelätään johtavan sosiaalityön asiantuntijuuden ja asiakkaan kanssa käytävän vuorovaikutuksen rajoittamiseen, koska tietojärjestelmä ohjaa työntekijää liian vahvasti tuottamaan tietoa ennalta määritellyn rakenteen pohjalta. Äärimmäisessä tilanteessa ammattilainen havaitsee asiakkaan tilanteesta keskeisiä seikkoja, joita ei ole mahdollista dokumentoida tietojärjestelmään, koska niille ei ole varattuna asianmukaista kirjaamispaikkaa. Narratiivisen asiakaskertomuksen vaarana puolestaan on se, että vapaana tekstinä tallennetusta tiedosta voi olla jälkikäteen mahdotonta poimia asiakkaan tilanteen kannalta oleellisia asioita, niiden hävitessä pitkien kirjoitusten sekaan. (Wartainen ym. 2020, 104-106.)

Sosiaalityöntekijöiden ajoittainen kriittinen suhtautuminen tietojärjestelmiin on ymmärrettävää. Työskentely yhtäältä kasvavien kirjaamisvaatimusten, tietojärjestelmien käytettävyysongelmien ja asiakkaan kanssa tehdyn välittömän työn tarpeiden ristitulella voi asettaa sosiaalityöntekijät ahtaalle. Vaikka asiakastyön kirjaaminen ja dokumentointi on tarpeellista, ammattilaiset ovat verranneet kirjaamisvaatimusten kiristymistä jopa pedon ruokkimiseen (*“feeding the beast”*) (Wastell & White, 2014, 214).

## 2.2 Tekoäly rantautuu sosiaalityöhön

Sosiaalityön käytännöissä ja koulutuksessa tekoälyn nopea kehittyminen herättää sekä toiveita että huolia. Sosiaali- ja terveyspalveluiden digitalisaatio, resurssipaineet ja tiedolla johtamisen vahvistuminen muodostavat kontekstin, jossa algoritmiset järjestelmät, kielimallit ja päätöksenteon tukityökalut halutaan vahvemmin integroida ammatillisiin prosesseihin.

Wartiainen ja kumppanit tunnistavat (2020) tekoälypohjaisten ratkaisujen sovelluskohteita sosiaalihuollossa. Näitä ovat esimerkiksi palveluun hakeutumisen keskustelut digitaalisen asiakaspalvelijan kanssa tai rakenteisen tiedon automaattinen muodostaminen kirjoitetusta tai puhutusta narratiivista. Toisaalta tekoälyratkaisut voivat tuottaa asiakastiedon perusteella työnkulkujen ehdotuksia tai analysoida riskien ja eri interventioiden vaikuttavuutta sekä todennäköisiä hyötyjä. Tekoäly voitaneen jatkossa myös valjastaa myös väestötason palvelutarpeiden ennustamiseen. (Mt., 110.)

Tätä tutkimusta kirjoittaessani suomalaisella sote-sektorilla on käynnissä erilaisia tekoälykokeiluja. Muutamia rajattuja käyttöönottoja on toteutettukin, mutta kansallisestikin on edelleen vielä hyvin vähän kokemusta siitä, miten tekoälyä voisi laajemmin hyödyntää sote-palveluissa. Sosiaali- ja terveysministeriön vuonna 2024 käynnistämä ja DigiFinlandin koordinoima kansallinen sote-tekoälyn ekosysteemi pyrkii omalta osaltaan kannustamaan julkisia organisaatioita ja yksityistä sektoria yhteistyöhön tekoälyn käyttöönottojen kasvattamiseksi. Ekosysteemiin osallistuminen on vapaaehtoista, mutta kehittämistä on tuettu taloudellisilla kannustimilla (Sosiaali- ja terveysministeriö, 2024).

Jos kansalliset sote-sektorin tekoälykokeilut ovat vielä hyvin maltillisia niin alkutaipaleella on myös lainsäädäntö, jopa kansainvälisellä mittapuulla tarkasteltuna. Euroopan Unionin vuonna 2024 hyväksymä, tekoälyä koskeva asetusta (Euroopan parlamentin ja neuvoston asetusta (EU) 2024/1689), on maailman ensimmäinen tekoälyä koskeva oikeudellinen kokonaisuus, joka pyrkii

sääntelemään sekä tekoälyn kehitystä, että sen käyttöä. Asetuksen keskeisimmät kohdat kohdistuvat erityisesti tekoälyn vastuullisen käytön varmistamiseen ja tekoälyratkaisujen riskienhallintaan. Suomessa asetuksen säädösvalmistelu on Työ- ja elinkeinoministeriön käynnistämän hankkeen (TEM044:00/2024) vastuulla. Asetus pannaan täytäntöön vaiheittain niin, että se on kokonaisuudessaan voimassa viimeistään syyskuussa 2026. Siirtymäkauden aikana sekä julkisen että yksityisen sektorin toimijat valmistautuvat asetuksen tuomiin velvoitteisiin.

Aivan uudesta teknologiasta ei kuitenkaan ole kyse, edes sosiaalityössä. Tekoälyn hyödyntämisen yleisyyttä sosiaalitieteissä ovat tutkineet muun muassa Robila ja Robila (2020). Tutkimus tarjoaa katsauksen viime vuosikymmenen puolella toteutettuihin tekoälyteknologian käyttökohteisiin tai kokeiluihin. He analysoivat yhteensä 36 vertaisarvioitua tutkimusta, jotka toteutettiin vuosien 2010–2019 välillä sosiaali- ja käyttäytymistieteiden alalla ja joiden kohteena oli tekoälyn tai koneoppimismallien soveltuvuus sosiaalisten ilmiöiden selittämisessä tai ennustamisessa.

Robila ja Robila osoittavat (2020), että koneoppimismallien potentiaali on tunnistettu jo viime vuosikymmenellä sekä terveydenhuollon että sosiaalisen hyvinvoinnin kysymyksissä. Tekoälymenetelmiä on hyödynnetty erilaisten ihmiselämään liittyvien olotilojen tai olosuhteiden (*”conditions”*) diagnostiikan ja ennustamisen tehostamisessa. Tutkimusaineistossa useimmin esiin nousseet teemat liittyivät nuorten alkoholinkäytön malleihin ja itsemurhariskiin. Ennustemalleja on kehitetty myös masennuksen tunnistamiseen, autismin diagnostiikkaan sekä hoitoon sitoutumisen arviointiin. Erilaisia riskimalleja on pilotoitu myös rikollisen käyttäytymisen ja sosiaalisten riskien arvioinnissa. Niitä on hyödynnetty esimerkiksi uusintapidätyksen ja asunnottomuuden riskin ennustamiseen sekä lapsiin kohdistuvan kaltoinkohtelun riskien tunnistamiseen. (Mt., 2957, 2958, 2962.)

Li ynnä muut (2025) ovat tarkastelleet tekoälyratkaisujen yleisyyttä ja tekoälymenetelmien laatua sosiaalityössä. Heidän tutkimuksensa osoittavat, että tekoälyavusteisia ratkaisuja hyödynnetään jonkin verran sosiaalityön tehtävissä. Tällaisia ovat asiakasryhmien tunnistaminen, päätöksenteko, toimenpiteiden luokittelu, palvelujen seuranta sekä riskien ennustaminen (mt., 6).

Luonnollisen kielen käsittely ja robotiikka kasvattavat jatkuvasti rooliaan ihmisen kehityksen ja toiminnan ymmärtämisessä. Robilan ja kumppaneiden mukaan (2019, 2960) sosiaalisen median

ja avointen keskustelupalustojen yleistymisen on lisännyt mahdollisuuksia käyttää luonnollisen kielen käsittelyä ja koneoppimista näiden aineistojen hyödyntämiseen. Käytännön sovelluskohteet liittyivät erilaisten sosiaalisten ilmiöiden tai toistuvien narratiivien tunnistamiseen keskustelupalustojen laajoista sisällöistä. Vastaavanlaisia havaintoja tekevät Al Maamari kumppaneineen (2025). He tarkastelevat tekoälyratkaisujen hyödyntämispotentiaalia laajoja ihmisjoukkoja koskevissa onnettomuus- ja katastrofitilanteissa. Heidän kirjallisuuskatsauksensa kattaa 29 tutkimusta vuosilta 2010–2024. Tutkimukset painottuvat tarkastelujakson loppupuolelle, mikä viittaa siihen, että kiinnostus tekoälyn ja katastrofinhallinnan rajapintaan on voimistunut erityisesti viime vuosina (mt., 124).

Tutkimukset osoittavat, että koneoppiminen mahdollistaa sisällöllisesti laajojen aineistojen käsittelyn ja analysoinnin sosiaalisen median alustoilta, minkä avulla voidaan tehokkaasti tavoittaa laaja joukko ihmisiä ja tarjota heille käytännön opastusta ja psykologista tukea kriisitilanteessa. Alustoilta on mahdollista seuloa ja analysoida viestejä tai sisältöjä, jotka koskevat avun tai palveluiden tarvetta ja niiden avulla muodostaa ajantasaista tilannekuvaa tilanteesta (Al Maamari 2025, 125-127). Toisaalta tekoälyn avulla voidaan tarjota emotionaalista tukea katastrofien uhreille erilaisten skaalattavien ja nopeaan reaktiokykyyn perustuvien tekoälyassistenttien avulla. Tällaisten ratkaisujen tarjoaminen ei rajoitu ainoastaan katastrofitilanteisiin vaan sen avulla voidaan tavoittaa erilaisia kohderyhmiä, joiden kynnyksiä asioida viranomaisten kanssa on suuri. (Mt., 118.)

Useiden tutkimusten painopiste vaikuttaa olevan sosiaalityön ammattilaisten, opiskelijoiden tai tutkijoiden tekoälyvalmiuksien ja asenteiden arviointi suhteessa tekoälyn käyttöön (ks. Báez ym. 2025; Garkisch ym. 2025; Buğra ym. 2024; Kapur ym. 2025), kun taas tekoälyn soveltuvuutta konkreettisiin sosiaalityön työtehtäviin arvioidaan selvästi harvemmin. Sosiaalityöntekijät omaksuvat uutta teknologiaa usein varsin hitaasti ja varauksella. Tekoälyratkaisuja otetaan kuitenkin käyttöön kiihtyvää tahtia yhteiskunnan eri sektoreilla, mikä lisää painetta niiden käyttöönottoon myös sosiaalityössä. On kuitenkin havaittu, ettei sosiaalityöntekijöiden ymmärrys tekoälyratkaisujen toimintalogiikasta ja sen vaikutuksista päätöksentekoon ole riittävällä tasolla (Kapur ym. 2025, 550). Mikäli tekoälyn tuloon ei valmistauduta huolella, on riski, että tekoälyn käyttöönotto sosiaalityön kentälle toteutetaan hallitsemattomasti. Tavoitteena tulisi olla sellaisten toimintatapojen rakentaminen, jotka mahdollistavat tekoälyn hallitun käyttöönoton. Tämä edellyttää sitä, että tekoälyyn liittyvä osaaminen on tuotava osaksi sekä sosiaalityön käytännön työtä että sosiaalityön koulutusohjelmia. (ks. Alam 2025; James ym. 2024; Ahn ym. 2025; Kapur ym. 2025.)

Tekoälyn toimintalogiikka perustuu kyvykkyyteen käsitellä valtavia tietomassoja sekä järjestelmän kehitysvaiheessa että sen käytön aikana. Tekoälyn hyödyntämiseen liittyvät huolenaiheet painottuvat erityisesti eettisyyteen ja vastuullisuuteen. Tekoälyn pelätään vahvistavan epätasa-arvoisia ja jopa syrjiviä käytäntöjä. Alam hahmottelee (2025) suosituksia tekoälyn osallistavaan ja läpinäkyvään käyttöön. Hän muistuttaa, ettei eettinen arviointi saa rajoittua pelkkiin arvo- ja periaatekysymyksiin, vaan edellyttää riittävää teknologista ymmärrystä tekoälyratkaisun toiminnasta ja rajoitteista. Ilman perustietoa esimerkiksi mallien oppimisperiaatteista ja vinoumariskeistä ratkaisujen kriittinen arviointi jää pinnalliseksi. Tähän osaamiskokonaisuuteen hän viittaa käsitteellä eettinen tekoälylukutaito (*”Ethical AI literacy”*) (mt., 5), joka kytkee yhteen teknisen lukutaidon ja eettisen harkinnan käytännön päätöksenteossa. Hän peräänkuuluttaa myös tekoälyn läpinäkyvää hallintamallia, jossa korostuu sidosryhmien, kuten ammattilaisten ja asiakkaiden, mahdollisuus vaikuttaa järjestelmän käyttöönoton ja käytön ehtoihin. Kaiken kaikkiaan sosiaalityöntekijöiden tulisivat siirtyä ennakoivasta riskipuheesta aktiivisempaan rooliin tekoälyn käytön todellisten vaikutusten seurannassa ja arvioinnissa. (Mt., 8-9.)

Myös James ja kumppanit (2024) sekä Ahn kumppaneineen (2025) käyttävät termiä tekoälylukutaito. He vahvistaisivat algoritmien osaamis pohjaa jo sosiaalityön opiskeluiden aikana. Vahvan teknologiaosaaminen ja sosiaalityön teoriapohjan yhdistelmä tuottaisi valmiuksia arvioida uusia tekoälyteknologioita kriittisesti. Tämä mahdollistaisi sen, ettei järjestelmien käyttöönotto perustuisi yksistään teknologialähtöisiin oletuksiin vaan etenisi hallitusti, sosiaalityön keskeisten arvojen ja periaatteiden ohjaamana. (James ym. 2024, 1555, 1561-1565.)

Teknologiaymmärrys tarjoaa sosiaalityöntekijöille arvokkaan lähtökohdan tekoälyn yhteiskunnallisten vaikutusten harkittuun ja eettiseen käsittelyyn. Ahn ja kumppanit (2025, 14) haastavatkin sosiaalityöntekijät ottamaan vahvaa roolia rakenteellisesta sosiaalityöstä ja kasvattamaan tietoisuutta oikeudenmukaisesta ja vastuullisesta tekoälyn hyödyntämisestä laajemmin yhteiskunnan eri tasoilla.

### **2.3 Tekoäly työkaluna asiantuntijatyössä**

Tekoälystä kirjoitettaessa on syytä tarkastella muutamia siihen liittyviä keskeisiä määritelmiä ja termejä. Koneoppiminen, tekoäly ja algoritmit määritellään usein rinnakkaisiksi termeiksi tai

sekoitetaan keskenään, mutta ne eivät täsmällisesti tarkoita samaa asiaa. Tekoöly-termiä käytetään arkikielessä väljästi ja jopa huolimattomasti kuvaamaan ratkaisuja, joiden kyvykkyys jäljittelee ihmismielen ajattelua tai suoriutumista. Termi vaatisikin yleensä rinnalleen tarkennuksen siitä, mitä tekoölymenetelmää tarkoitetaan. Tekoöly on käsitteenä laaja ja hieman epätarkka yläkäsite, jonka alle kuuluu laaja joukko erilaisia menetelmiä.

Valtioneuvoston kanslia on julkaissut ohjeistuksen (VN/6190/2025) generatiivisen tekoölyn hyödyntämisestä työ tukena ja apuvälineenä julkisessa hallinnossa. Ohjeistuksessa määritellään tekoöly seuraavasti: ”Tekoöly (AI) tarkoittaa tietojärjestelmien kykyä suorittaa tehtäviä, jotka normaalisti vaativat inhimillistä älykkyyttä. Näihin tehtäviin kuuluvat esimerkiksi oppiminen, päättely, ongelmanratkaisu, kielen ymmärtäminen ja visuaalinen havaitseminen. Tekoölyjärjestelmät hyödyntävät algoritmeja ja suuria tietomääriä parantaakseen suorituskykyään. Tekoöly soveltaa edistyneitä analyysi- ja logiikkapohjaisia tekniikoita, mukaan lukien koneoppiminen, tulkitukseen tapahtumia, tukeakseen ja automatisoidakseen päätöksiä sekä suorittaakseen toimia”. Euroopan unionin tekoölyasetuksen (2024/1689) 3. artiklan mukaan tekoölyjärjestelmällä tarkoitetaan ”konepohjaista järjestelmää, joka on suunniteltu toimimaan käyttöönoton jälkeen vaihtelevilla autonomian tasoilla ja jossa voi ilmetä mukautuvuutta käyttöönoton jälkeen ja joka päättelee vastaanottamastaan syötteestä eksplisiittisiä tai implisiittisiä tavoitteita varten, miten tuottaa tuotoksia, kuten ennusteita, sisältöä, suosituksia tai päätöksiä, jotka voivat vaikuttaa fyysisiin tai virtuaalisiin ympäristöihin”.

Tekoöly perustuu tilastolliseen päättelyyn ja algoritmeihin, joiden kautta se pystyy prosessoimaan suuriakin määriä tietoa. Algoritmillä tarkoitetaan ennalta määritettyä, laskennallista menettelyä, jonka avulla tietoa käsitellään ja jonka seurauksena tuotetaan jokin tulos (Cormen ym. 2009, 5), kuten luokitus, pisteytys, suositus tai ennuste. Tekoöly rakentuu algoritmeista, eli eräänlaisista laskentaohjeista tai sääntökokoelmista, joiden avulla malli käsittelee ja muokkaa tietoa. Algoritmit ohjaavat esimerkiksi sitä, miten tekoöly löytää suurista tietomääristä olennaisia asioita, tekee luokitteluja tai ehdottaa ratkaisuja. Toisin sanoen ilman algoritmeja tekoöly ei pystyisi oppimaan tai tekemään päätelmiä, vaan ne muodostavat koko järjestelmän toiminnan perustan.

Koneoppiminen (Machine Learning, ML) on yksi tekoölyn keskeisistä osa-alueista, ja erilaisia koneoppimisen tekniikoita hyödynnetään laajasti tekoölyratkaisuissa. Kaikki koneoppimisen muodot ovat tekoölyä, mutta kaikki tekoölyratkaisut eivät sovelle koneoppimisen tekniikoita.

Koneoppimismallit analysoivat menneestä datasta löytyviä rakenteita ja säännönmukaisuuksia arvioidakseen tulevien tapahtumien todennäköisyyttä. Koneoppimismallit voidaan jakaa erilaisiin alalajeihin sen perusteella, mitä oppimistyyppiä tekniikassa sovelletaan, eli miten algoritmia opetetaan. Tällaisia alalajeja ovat muun muassa ohjattu ("supervised") ja ohjaamaton ("unsupervised") oppiminen sekä vahvistusoppiminen ("reinforcement learning") (VN/6190/2025).

Ohjattu koneoppimismalli hyödyntää ennalta merkittyjä lopputuloksia ja oppii niiden perusteella tunnistamaan aineistosta ne tapaukset, joilla on samankaltainen todennäköinen lopputulos. Ohjaamaton koneoppiminen puolestaan ei käytä ennalta merkittyjä lopputuloksia, vaan ryhmittelee tapaukset datassa ilmenevien yhteisten piirteiden ja mallien perusteella (Ahn ym. 2025, 10). Toisaalta koneoppimisen menetelmiä voidaan jaotella myös niiden tehtävän mukaan. Leslie ja kumppanit jaottelevat koneoppimisen kuvaileviin, ennustaviin ja analysoiviin (2020, 9-18).

Neuroverkot ovat tekoälymalleja, jotka oppivat merkityksellisiä yhteyksiä useiden datakerrosten läpi siten, että ne muodostavat monimutkaisia hierarkioita ihmisaivojen tapaan (Ahn ym. 2025, 10). Niihin perustuvat laajat kielimallit (Large language models, LLM) ja generatiivinen tekoäly. Generatiivinen tekoäly (generative AI) on tekoäly, joka pystyy luomaan uutta sisältöä, kuten tekstiä, kuvia, musiikkia tai videoita, perustuen sille annettuun dataan; se käyttää koneoppimismalleja, kuten syviä neuroverkkoja, tuottaakseen luovia ja monimutkaisia tuotoksia, jotka muistuttavat ihmisten luomaa sisältöä (VN/6190/2025). Generatiivinen tekoäly erotellaan usein omaksi teknologiakseen korostaakseen sen tehtävää, joka liittyy uuden luomiseen. Laajat kielimallit puolestaan viittaavat tekoälyjärjestelmiin, jotka käsittelevät ja analysoivat luonnollista kieltä sekä siihen liittyvää dataa, kuten tekstiä, puhetta ja muita viestinnän muotoja (Patton ym. 2023) ja ne pystyvät tuottamaan kieltä edistyksellisellä tavalla.

Tekoälymenetelmien ja koneoppimisen teknologioiden laajempi arviointi ei ole tämän tutkimuksen laajuudessa, joten rajaan tarkoituksella pois koneoppimisen alalajien toimintalogiikan yksityiskohtaisen kuvauksen. Kaikissa artikkeleissa ei ole myöskään avattu tutkimuksen toteutukseen käytettyä tekoälymenetelmää, joten menetelmien yksityiskohtainen tarkastelu tai vertailu ei saatavissa olevan tiedon perusteella ole mahdollista.

### 3 Sosiaalityön tiedon ja todellisuuden ulottuvuudet

#### 3.1 Asiantuntijuus paineessa: liikaa tietoa, liian vähän varmuutta

Jotta tekoälyn hyödyntämistä voidaan tarkastella analyttisesti ja tarkoituksenmukaisesti osana sosiaalityötä, on ensin jäsennettävä sosiaalityön tiedonmuodostuksen periaatteet sekä siihen kytkeytyvä keskustelu sosiaalityön asiantuntijuudesta. Keskeistä on hahmottaa, miten sosiaalityön tieto tuotetaan ja millaisista elementeistä se rakentuu, sekä millaisia haasteita ja reunaehdoja tiedonmuodostukseen liittyy. Näin voidaan kytkeä tarkastelu tutkielman ydinkysymykseen: miten tekoälyn hyödyntäminen liittyy sosiaalityön prosesseihin ja tiedonmuodostukseen. Tähän kokonaisuuteen kuuluu myös arvio siitä, millaisin edellytyksin sosiaalityön tietoa voidaan käsitteellistää ja systematisoida siten, että algoritmiset menetelmät kykenevät sitä analysoimaan ja tuottamaan sosiaalityön tarpeisiin relevanttia tietoa.

Lindh ynnä muut kuvaavat (2018) sosiaalityön asiantuntijuutta moniulotteiseksi ja osin vaikeasti määriteltäväksi kokonaisuudeksi, joka rakentuu sekä erilaisten asiakasryhmien kanssa tapahtuvasta yksilöllisestä ja yhteisöllisestä työstä että rakenteellisesta yhteiskunnallisesta vaikuttamisesta. Sosiaalityöntekijän asiantuntijuuden ja osaamisen tulisi muuttuvissa asiakastilanteissa rakentua vahvasti tiedon hyödyntämiselle ja tilanteiselle soveltamiselle. Työskentelyn perustaksi tarvitaan laaja ja monilähteinen tietopohja, jota sosiaalityöntekijä kykenee tulkitsemaan ja kytkemään asiakkaan tilanteen kokonaisuuteen. Sosiaalityöntekijöiden on kuvattu soveltavan käytännön työtä koskevassa reflektiossa ja tiedonmuodostuksessa eri tiedon lajeja: tutkimus- ja teorian tietoa, asiantuntijaa tietoa, asiakkaan tietoa, formaalia tietoa ja ammattieettistä tietoa (Karttunen & Hietamäki 2014, 320, 332-333).

Asiantuntijuutta muovaavat ja kuormittavat hyvinvointivaltion tehostamis- ja tuottavuuspaineet, jotka ohjaavat työn tavoitteita ja resursointia. Kansallisen palvelujärjestelmän uudistukset ja lainsäädännön muutokset tuottavat jatkuvia rakenteellisia muutospaineita. Globalisaatio ja siihen kytkeytyvät väestölliset ja sosiaaliset muutokset haastavat sosiaalityön toimintaympäristöä ja asiakaskunnan tarpeita. Sosiaaliset ongelmat monimutkaistuvat. Näihin kehityskulkuihin vastaaminen edellyttää sosiaalityön asiantuntijuuden vahvistumista, mikä tarkoittaa sosiaalityön menetelmien, tiedon ja identiteetin uudistamista ja tarkentamista. (Lindh ym. 2018, 9-11.)

Eri tiedonlajeihin liittyy keskinäisiä riippuvuussuhteita ja jännitteitä, joiden merkitys sosiaalityön tiedonmuodostuksen prosessissa ja päätöksenteossa on olennaista tunnistaa. Tekoälyn

käyttöönotto lisää kokonaisuuteen uuden ulottuvuuden, joka voi muuttaa tiedon painoarvoja, tiedon hierarkioita sekä päätöksenteon dynamiikkaa. Samalla nousee kysymys siitä, mihin sosiaalityössä jo tunnistettuun tiedonlajiin tekoälypohjainen tieto mahdollisesti asemoituu – vai muodostuuko siitä oma, erillinen tiedon muotonsa tässä kontekstissa.

Karttunen ja Hietämäki (2014) määrittelevät sosiaalityöntekijän asiakastyössä tarvitsemat tieto- ja taitoalueet seuraavasti: sosiaaliset ongelmat, työmenetelmät, vuorovaikutustaidot sekä arviointi- ja reflektointitaidot. Heidän mukaansa kokemustietoa on perinteisesti arvostettu ja hyödynnetty sosiaalityön asiakaskonteksteissa tyypillisesti enemmän kuin tutkimustietoa (mt., 327). Pohjalainen on tarkastellut (2012) Michael Polanyin jo vuonna 1966 tunnetuksi tekemää käsitettä ”tacit knowledge”, joka suomeksi käännettynä tarkoittaa hiljaista tietoa. Pohjalainen kiteyttää hiljaisen tiedon seuraavasti: ”Hiljainen tieto on uskomuksia, mielikuvia, ajatusrakennelmia ja näkemyksiä, jotka ovat ihmisen toiminnan taustalla. Hiljainen tieto on toisaalta myös ammattitaitoa, osaamista (mt., 2).

Hiljaisen tiedon hyödyntämistä kohtaan esiintyy kritiikkiä. Ongelmallista hiljaiselle tiedolle on se, että sitä on usein haastava tunnistaa ja pukea sanoiksi. Kuusisto-Niemi ja Kääriäinen (2005) näkevät, että sosiaalityön tiedon hiljaisuus on liitetty laajasti koko toimintatapakulttuuriin ja julkiseen keskusteluun sosiaalityön menettelytavoista. He kokevat hiljaisen tiedon käyttämisen ongelmallisena esimerkiksi asiakasta koskevien päätösten perusteena. Hiljaisen tiedon näkyväksi tekeminen hälventäisi sosiaalityön epämääräisyyttä. (Kuusisto-Niemi & Kääriäinen 2005, 454–457; Kääriäinen 2016.)

Viranhaltijoina sosiaalityöntekijöiden on kyettävä perustelemaan ja dokumentoimaan, mihin asiakastyössä tehdyt ratkaisut pohjautuvat. On tunnistettu kansallinen tarve kehittää ja vahvistaa sosiaalityön tietoperustaa ja tutkittuun tietoon perustuvia työskentelymenetelmiä (Pohjola 2025). Karttunen ja Hietämäki (2014) sisällyttävät tutkimustietoon muun muassa sosiaalityön substanssitiedon käsitteineen ja malleineen, menetelmätiedon sekä empiirisen tiedon sosiaalisista ilmiöistä ja ongelmista (mt., 2014, 333).

Sosiaalityö on laaja ja monialainen toimintakenttä, joka ulottuu kaikkiin ihmisen elämänvaiheisiin. Tämä monimuotoisuus tekee tutkimusperustaisen tiedon tuottamisesta erityisen vaativaa. Sosiaalityötä tehdään useilla tasoilla – yksilöiden, ryhmien, yhteiskunnan sekä kansainvälisen yhteistyön tasoilla – ja kaikille näille tasoille korostuu tarve tuottaa systemaattista tietoa (ks. Jokiranta 2024). Systemaattista tiedontuotantoa tarvitaan yhtäältä vaikuttavuuden

vahvistamiseksi, eli sosiaalisen muutoksen aikaansaamiseksi, ja toisaalta kehittämisen tueksi. Vain näin sosiaalityö kykenee vastaamaan siihen kohdistuviin uudistuspaineisiin ja säilyttämään asemansa sosiaalisen hyvinvoinnin edistämässä terveydenhuollon rinnalla tasa-arvoisena osaamisalueena.

Huolestuttavaa on kuitenkin se, että sosiaalityöntekijöille tehtyjen kyselytutkimusten mukaan (Kemppainen ym. 2018; Hiltula ym. 2024) tutkimusperustaisen tiedon hyödyntämiseen arkityössä ei ole riittävästi mahdollisuuksia tai aikaa. Pohjola ja Satka peräänkuuluttavat (2022) sosiaalisen osaamisen, tietoperustan ja työkäytäntöjen päämäärätietoista edistämistä. Heidän mukaansa on välttämätöntä vahvistaa sosiaalityön tieteellistä perustaa siten, että se rakentuu vankalle teoreettiselle kehikolle, hyödyntää systemaattiseen arviointiin pohjautuvia toimintamalleja ja perustuu vaikuttavuuden seurantaan.

### **3.2 Kriittinen realismi todellisuuden jäsentäjänä**

Tiedonmuodostuksen eri ulottuvuuksia voidaan tarkastella realistisen tieteenfilosofian näkökulmasta. Realismin suuntauksista erityisesti kriittinen realismi tarjoaa käyttökelpoisen viitekehyksen sosiaalityön tiedonmuodostuksen ymmärtämiseen ja jäsentämiseen. Tässä luvussa esittelen kriittisen realismin keskeiset lähtökohdat, taustoitan sen historiallista kehitystä sekä tarkastelen sen vaikutusta ja sovelluksia sosiaalityön tutkimuksessa.

Kriittinen realismi muotoutui 1970-luvulla filosofi Roy Bhaskarin (1944–2014) työn pohjalta vastauksena aikakauden keskeisiin tieteenfilosofisiin painotuksiin, joissa sosiaalisen todellisuuden nähtiin rakentuvan ensisijaisesti tulkintojen kautta ja joissa tiedon subjektiivisuutta korostettiin voimakkaasti. Bhaskar (2015) korostaa tieteen selitysvoimaa sosiaalisten ilmiöiden tutkimuksessa: hänen mukaansa sosiaalista todellisuutta on mahdollista tarkastella ja selittää systemaattisesti samankaltaisin tieteellisin pyrkimyksin kuin luonnontieteissä. Kriittinen realismi ei sinänsä kiistä sitä, että ihmisen tietoisuus ja tulkinnat muovaavat tapaa, jolla todellisuutta ymmärretään ja kuvataan. Tämän tieteenfilosofian kannattajat kuitenkin korostavat, että tämän lisäksi on olemassa myös ihmisestä ja hänen tietoisuudestaan riippumaton todellisuus (Danermark ym. 2019, 7).

Kriittisessä realismissa uskotaan, että empiirisesti havaittavien ilmiöiden lisäksi myös havaintojemme ulkopuolella olevaa maailmaa voidaan tutkia ja tehdä se näkyväksi. Ihmistieteiden ilmiöitä voidaan siis selittää paljastamalla niiden taustalla piileviä mekanismeja.

Sosiaalityön kontekstissa tämä merkitsee, että työntekijän tehtävä ei rajaudu ilmiöiden kuvaamiseen tai asiakkaan tilanteen luokitteluun, vaan siihen sisältyy sosiaalisten ongelmien juurisyiden ja niitä ylläpitävien rakenteellisten tekijöiden tarkastelu. Tavoitteena on ymmärtää, miten ongelmat syntyvät ja toistuvat, sekä arvioida, millaisin toimin ja millä tasoilla niihin on mahdollista vaikuttaa. (Pekkarinen & Tapola-Haapala 2009, Johdanto, Kriittinen realismi ja sosiaalityön tiedontuotanto.) Kriittisen realismin näkökulmasta arvoja ja sosiaalisia tulkintoja tarkastellaan lähtökohtaisesti refleksiivisesti ja arvioiden. Sosiaalisia ilmiöitä selittävä kriittinen analyysi voi parhaimmillaan paljastaa, miten sosiaaliin rakenteisiin on kietoutunut kyseenalaisia, virheellisiä tai harhaanjohtavia oletuksia, uskomuksia ja arvolatauksia, jotka voivat ylläpitää tai vahvistaa vinoutuneita käytäntöjä.

Sosiaaliset ongelmat ja ilmiöt näyttävät usein monimutkaisina, monisyisinä ja tilanteisesti vaihtelevina, mikä vaikeuttaa niiden tutkimista ja selittämistä. Sosiaalitieteille tunnusomaista onkin ilmiöiden monikerroksisuus. Toisin kuin luonnontieteissä, sosiaaliseen kontekstiin kietoutuvia ilmiöitä voidaan selittää vain osittain teoreettisten mallien avulla ja siitäkin huolimatta malleihin liittyy rajoitteita ja epävarmuustekijöitä. Sosiaalitieteille tyypillistä on, että teorian pohjalta voidaan kylläkin ennustaa tai odottaa tiettyjä faktoja, mutta teorioilla ei voi määrittää lopputulemaa. Sosiaaliset ilmiöt altistuvat jatkuvalla muutokselle, johtuen ihmisten ja yhteisöjen kyvystä muuttua. Luonnontieteissä vastaava muutos ei ole mahdollinen. (Danermark ym. 2019, 19-20, 29).

Kriittisen realismin mukaan sosiaalitieteissä todellisuutta ei voida kuvata tai selittää täysin kattavasti, mutta tutkimuksen tehtävänä on pyrkiä mahdollisimman hyvin perusteltuihin selityksiin. Yksilön toimintaa ja vuorovaikutusta sosiaalisessa ympäristössä ei yleensä voi palauttaa yksinkertaisiin syy-seurausketjuihin, koska ilmiöihin vaikuttaa samanaikaisesti useita tekijöitä ja rakenteita. Siksi tavoitteena on jäsentää ilmiöitä ja tunnistaa niiden taustalla vaikuttavia mekanismeja niin pitkälle kuin käytettävissä oleva aineisto ja teoreettinen ymmärrys kulloinkin mahdollistavat.

Sosiaalityön tiedonmuodostuksen tarkastelu nojautuu vahvaan konstruktionismin perinteeseen, jossa todellisuuden nähdään muodostuvan ihmisten välisistä tulkinnoista. Pekkarinen ja Tapola-Haapala (2009, Sosiaalityö, tieto ja totuus) kannustavatkin pohtimaan, millaisia mahdollisuuksia kriittinen realismi tarjoaa tietämisen ja totuuden kysymysten käsittelyyn sosiaalihuollon kontekstissa. Tiedonmuodostuksen kannalta tämä tarkoittaa sitä, että sosiaalityöntekijän henkilökohtaisen tiedon, intuition ja omakohtaisten arvioinnin rooli kapenee ja

sosiaalityöntekijän odotetaan perustavan työskentelynsä yhä vankemmin teoriapohjaiseen tai ennalta määriteltyyn ja strukturoituun tietoon.

Kriittinen realismi tukee näkemystä siitä, että sosiaalihuollon ammattilaisten ja organisaatioiden toimintamalleja tulisi systematisoida ja sitä kautta toimenpiteiden ja interventioiden tulisi perustua tutkittuun tietoon. Kriittisen realismin viitekehyksessä sosiaalityössä ei ole tilaa niin kutsutulle hiljaiselle tiedolle, vaan kaikki päätöksentekoa ohjaava tieto on käsitteellistettävä ja päätökset on perusteltava rationaalisesti. Tästä näkökulmasta rakenteinen kirjaus, erilaiset lomakkeet ja listat tukevat ammattilaisen työtä ja vähentävät muistinvaraisuuden aiheuttamaa vaihtelevuutta tiedontuotannossa.

Devlieghere ja kumppanit (2022) käsittelevät kahden lähestymistavan, dataismin ja relationshipismin, välisiä eroja sosiaalityön kontekstissa. Dataismi korostaa suurten tietomassojen merkitystä esimerkiksi päätöksenteossa tai palvelutarpeen ennakoinnissa. Relationshipismi puolestaan nojaa perinteiseen ajatukseen vuorovaikutteisesta asiakassuhteesta ja yksilöllisten tilanteiden erityislaatuisuudesta. Tekoälyyn soveltuu pitkälle samat vaatimukset kuin mihin tahansa tilastollisen päättelyn menetelmiin. Abstraktien ilmiöiden, joita sosiaalitieteissä tyypillisesti tarkastellaan, tutkiminen ja mittaaminen edellyttävät sitä, että käsitteet pystytään operationalisoimaan. Operationalisoinnissa käsitteet puretaan pienempiin osiin ja muokataan siten, että eri osien vaihtelua voidaan mitata. Ihmistieteissä haasteeksi muodostuu käsitteiden moniulotteisuus. Laajimmillaan ilmiöt koostuvat useista ylä- ja alakäsitteistä. Yhden yläkäsitteen alla saattaa olla kymmeniä uusia alakäsitteitä, jotka yhdessä muodostavat mitattavan kokonaisuuden. (Vehkalahti 2019, 18-19.)

Kriittisen realismin mukaan tutkimuskohteita voidaan jäsentää avoimina ja suljettuina systeeminä sen perusteella, millaisissa olosuhteissa tapahtumien välisiä säännönmukaisuuksia havaitaan (Danermark ym. 2019). Suljetussa systeemissä keskeiset vaikuttavat tekijät voidaan pitää vakioina. Tällöin tapahtumissa ilmenee usein lainkaltaisia säännönmukaisuuksia, jotka tekevät ilmiöiden ennustamisesta helpompaa. Suljetulle systeemille on myös tunnusomaista, ettei se ole vuorovaikutuksessa muiden ulkopuolisten systeemien kanssa. Avoimessa systeemissä sen sijaan useat mekanismit vaikuttavat samanaikaisesti ja voivat vahvistaa, heikentää tai kumota toistensa vaikutuksia. Tämän vuoksi samat lähtöehdot eivät johda joka kerta samanlaisiin tapahtumiin, eikä kausaalisuutta voida päätellä ja testata luotettavasti pelkkien toistuvien tapahtumien avulla (ks. Bhaskar 2025, 45).

Regressioanalyysiä on hyödynnetty menestyksekkäästi avoimien systeemien tutkimisessa (Danermark ym. 2019, 56), mutta tuloksia kannattaa tulkita varovasti: tilastolliset mallit kertovat ennen kaikkea, millaisia keskimääräisiä yhteyksiä aineistosta löytyy, ei sitä, että X aiheuttaa varmasti Y:n. Myös Bhaskar (2015, 45-46) näkee sosiaalitieteiden teorioiden arvon selittämisessä, ei niinkään ennustamisessa. Tulokset on hyvä yhdistää taustateoriaan ja kontekstiin, ja tarvittaessa tukea niitä muulla aineistolla tai lisäanalyysillä.

Sosiaalitieteissä on pitkään pohdittu avoimien ja suljettujen systeemien problematiikkaa. Useimmiten sosiaalista todellisuutta kuvataan avoimena systeeminä, vaikka on esitetty myös hybridimalleja, joissa yhdistyvät molempien systeemityyppien piirteet. Toisaalta nähdään, ettei ihmisten vuorovaikutus olisi ylipäättään mahdollista ilman tiettyjä säännönmukaisuuksia. Tällöin sosiaalinen vuorovaikutus nojaa osin myös suljettujen systeemien periaatteisiin, vaikkakaan nämä säännönmukaisuudet eivät ole pysyviä tai muuttumattomia. (Danermark ym. 2019, 54-58.)

Kriittisessä realismissa operationalisoinnin haasteeseen viittaa termi ”avoimen järjestelmän epävarmuus”. Se tarkoittaa sitä, että luonnontieteisiin verrattuna ihmistiede on moniulotteisempaa, eikä syy-seuraussuhdetta useinkaan pystytä tyhjentävästi ennustamaan tai selittämään. Sosiaalityössä avoimen järjestelmän epävarmuus tarkoittaa käytännön työssä sitä, että asiakkaan tilanteen kokonaiskäsityksen muodostamiseen sekä päätöksentekoon vaikuttavat muun muassa asiakkaan ja ammattilaisen tunnetaso, heidän kulttuuriset ja sosiaaliset kontekstinsa, molempien sosio-ekonominen tausta ja asema, ikä, subjektiivinen kokemus tilanteesta tai molempien henkilökohtainen historia ja kokemukset sekä arvomaailma, vain muutamia mainitakseni.

Tiedonmuodostuksen ja päätöksenteon kontekstissa nämä taustamuuttujat lisäävät päätöksenteon epävarmuutta. Kaikkien taustamuuttujien operationalisointi on haastavaa, jopa mahdotonta. Vaikka tekoälyalgoritmit pystyvät analysoimaan suuria määriä muuttujia, nekin pelkistävät sosiaalityön tapaukset erillisiksi muuttujiksi, jolloin monet kontekstuaaliset tekijät jäävät ulkopuolelle, koska niitä on liian haastavaa operationalisoida riittävän kattavasti (Kapur ym. 2025, 550). Tiedon tuottamisessa ja rajaamisessa on tehtävä valintoja, koska näkökulmien ja tiedon eri tasojen määrä on loputon (Pohjola, 2025).

Sosiaalisten ilmiöiden moniulotteisuus ja operationalisoinnin haaste voidaan todeta esimerkin kautta. Jos haluamme tarkastella huono-osaisuutta, jossa jo yksistään huono-osaisuuden määritelmä riippuu sovellettavasta teoriasta tai näkökulmasta. Saaren mukaan (2024, Huono-

osaisten hyvinvointi) huono-osaisuutta voidaan tarkastella esimerkiksi elämänlaadun, toimintakyvyn, osallisuuden, taloudellisten tai käytettävissä olevien resurssien näkökulmista. Toisaalta huono-osaisuutta voidaan tarkastella objektiivisena tai subjektiivisena kokemuksena. Kouluttamattomuus, toimeentulo-ongelmat, köyhyys ja mielenterveysongelmat kietoutuvat usein yhteen huono-osaisuuden taustalla (Kataja 2014). Ilmiön ymmärtämiseksi olisi kaikkia muuttujia tarpeen tarkastella vielä erikseen ja pilkkoa kukin yhä pienempiin taustamuuttujiin.

Kriittinen realismi kannustaa kuitenkin käsitteellistämään sosiaalisia ilmiöitä, kehittämään niiden operationalisointia ja kartuttamaan tutkimustietoa. Tässä viitekehyksessä myös tekoälyä ja tilastollisia menetelmiä voidaan hyödyntää entistä systemaattisemmin ymmärryksemme syventämiseksi ihmisistä, sosiaalisista ilmiöistä ja sosiaalityön tarpeista. Vaikka tutkimustiedon kehittyminen ja suurien tietomäärien käsittely teknologialla avaavatkin uusia mahdollisuuksia tutkia ja selittää sosiaalisia ilmiöitä, on arvioitava huolellisesti menetelmien ja teknologioiden luotettavuutta ja soveltuvuutta.

## 4 Tutkimuksen toteutus

### 4.1 Tutkimuksen tavoite ja tehtävä

Tutkimuksen tavoitteena on kuvata ja analysoida, miten tekoälyn avulla tuotetaan ja muokataan tietoa sosiaalityön prosesseissa. Tekoälyn hyödyntämiseen liitetään merkittäviä mahdollisuuksia, mutta myös eettisiä, ammatillisia ja yhteiskunnallisia haasteita. Ne voivat liittyä esimerkiksi datan vinoumiin ja sitä kautta syrjintää tuottaviin käytäntöihin. Huolta on myös esitetty tekoälyn läpinäkyvyyden puutetta kohtaan. Tieto tekoälyn toimintaperiaatteista on usein puutteellista, ja taustalla olevien monimutkaisten algoritmien toimintaa on maallikon vaikea – ellei jopa mahdoton – ymmärtää. Myös tietosuojaan liittyvät kysymykset ovat nousseet keskusteluun. Lisäksi pelkona nähdään inhimillisen kohtaamisen kaventuminen, tiedon, prosessien ja palveluiden siirtyessä yhä enemmän digitaaliseen muotoon. Toisaalta automaation tarjoamat mahdollisuudet työn uudelleenjärjestelyyn ovat lupaavia: työntekijöiden kuormitusta voidaan vähentää, kun osa tehtävistä voidaan siirtää järjestelmien hoidettavaksi.

Tutkimukseni on laadullinen kirjallisuuskatsaus. Laadullisen tutkimuksen aineistokeruumenetelmät ovat haastattelu, kysely, havainnointi ja erilaisiin dokumentteihin perustuva tieto (Tuomi & Sarajärvi 2018, luku 3). Tutkimuskirjallisuutta hyödyntämällä aion jäsentää, miten tekoälyn käyttö kietoutuu sosiaalityön arvoihin, tietoperustaan sekä ammattikäytäntöihin ja millaisia muutoksia tai jännitteitä tämä yhteen kietoutuminen tuottaa. Tutkimuskysymykseni ovat:

1. Minkälaisia sosiaalityön tiedonmuodostuksen tehtäviä tekoäly on tukenut?
2. Mitkä ovat tekoälypohjaisen tiedon muodostusprosessin vahvuudet ja rajoitteet?
3. Miten tekoäly vaikuttaa ammatilliseen harkintaan sosiaalityössä?

Tekoälyä on hyödynnetty sosiaalityön ulkopuolella hyvin monipuolisesti. Sitä voidaan hyödyntää tiedonhakuun, arviointiin, dokumentoinnin osana tai päätöksenteon tukeen – jopa automaattiseen päätöksentekoon. Sillä voidaan automatisoida muun muassa logistiikan, asiakaspalvelun tai toiminnanohjauksen prosesseja. Tutkimukseni pyrkii jäsentämään sitä, miten tietoa sosiaalihuollon asiakkaiden todellisuudesta ja erilaisista sosiaalisista ongelmista tuotetaan tekoälyn avulla sosiaalityön tehtävien toteuttamiseksi.

Tarkastelen erityisesti sosiaalityön kontekstissa tuotetun tekoälypohjaisen tiedon läpinäkyvyyttä ja laatua. Läpinäkyvydellä tarkoitan esimerkiksi sitä, kuinka hyvin tekoälymenetelmien

hyödyntämät lähtötiedot on kuvattu tai kuinka kattavasti tiedon muodostukseen hyödynnettyjä menetelmiä on selitetty. Lisäksi tarkastelen, onko osana prosessia tuotu esiin mahdollisia epävarmuustekijöitä tai riskejä sekä tiedontuotannon ja päätöksenteon vastuita.

Tutkimukseeni valikoituneet artikkelit kuvaavat tekoälyn käyttöä hyvin erityyppisissä sosiaalityön tilanteissa. Pyrin selvittämään, vaihtelee tekoälyn avulla tuotetun tiedon oikeutus sen mukaan, missä sosiaalityön tehtävissä tekoälyä hyödynnetään ja minkälaisien asiakasryhmien kanssa työskennellään. Tavoitteeni on tuoda yhteen erilaisia näkökulmia ja pohtia tekoälyn hyödyntämistä sosiaalityön ammatillisen identiteetin kautta.

## 4.2 Aineiston kerääminen

Olen hyödyntänyt tutkimuksessani aineistoa suomalaisista ja kansainvälisistä tutkimuksista. Hain tutkimusaineiston LUC Finnan SCOPUS-viitetietokannasta. Tutkimusartikkelit etsin aluksi englanninkielisillä hakusanoilla ”social work” ja ”artificial intelligence”. Uskoin, että nämä hakusanat mahdollistavat riittävän laajan otannan erilaisten tekoälyä koskevien näkökulmien hakemiseen. ”Artificial intelligence” toimii erilaisten tekoälyratkaisujen ja tekoälymenetelmien yläkäsitteenä. Lisäksi halusin tarkastella tekoälyä koskevaa tutkimusta yleisesti sosiaalityön näkökulmasta, jolloin haullla on mahdollista saada monipuolinen ymmärrys tekoälystä koko sosiaalityön alalta. SCOPUS-viitetietokannan tarjoamat artikkelit ovat vertaisarvioituja, joten erillistä hakua vertaisarvioinnin perusteella ei ollut tarpeen tehdä. Artikkeleiden määrä tarkentui tämän haun perusteella 205 tulokseen.

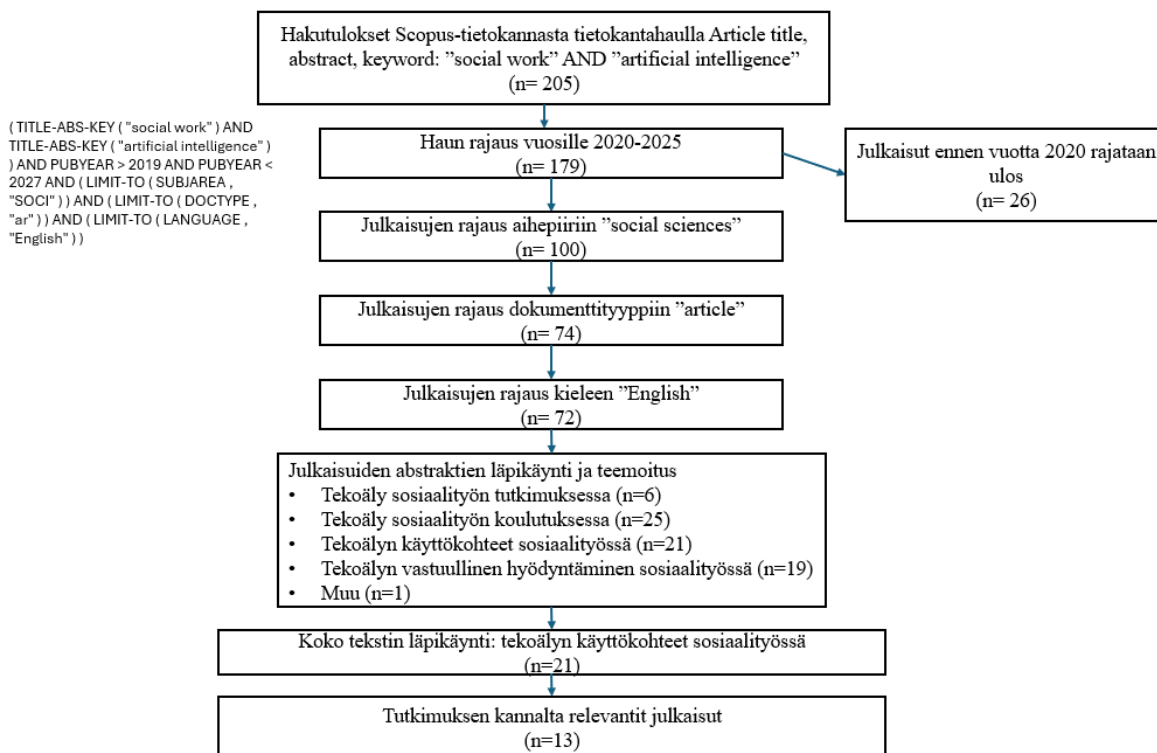
Rajasin aineiston niin, että otin mukaan hakuun ainoastaan vuoden 2020 jälkeen tehdyt julkaisut. Tämän rajauksen toteutin sen vuoksi, että teknologia on kehittynyt huimin harppauksin viimeisen muutaman vuoden ajan ja halusin nimenomaan keskittyä ajankohtaisimpiin tekoälyn ympärillä tehtyihin tutkimuksiin. Tällä rajauksella sain 179 tulosta. Vaikka hakusanana käytin termiä ”social work”, nousi tuloksiin artikkeleita, jotka eivät suoranaisesti kohdistuneet sosiaalityön käyttötapauksiin vaan olivat täysin muilta tieteenaloilta. Tein vielä tämän jälkeen haun, jossa oli mukana rajaus tieteenalan ”social sciences” mukaan. Tämä rajaus tuotti 100 tulosta. Tämän jälkeen rajasin mukaan vain tieteelliset artikkelit ja englanninkieliset julkaisut. Hakutuloksia oli nyt 72 ja tiivistelmien selailu antoi minulle varmuuden siitä, että toteutettu hakupolku vei minua oikeaan suuntaan.

Tämän jälkeen kävin läpi tarkemmin tutkimusartikkeleiden tiivistelmät. Tiivistelmien perusteella tunnistin tutkimusartikkeleista neljä teemaa. Nämä teemat olivat tekoäly sosiaalityön tutkimuksessa, tekoäly sosiaalityön koulutuksessa, tekoälyn käyttökohteet sosiaalityössä sekä tekoälyn vastuullinen hyödyntäminen sosiaalityössä. Joukossa oli yksi artikkeli, joka ei käsitellyt mitään näistä teemoista, joten rajasin kyseisen artikkelin ulos tarkastelun piiristä. Tekoälyn käyttökohteet sosiaalityössä -teeman alle sisällytin vain sellaiset artikkelit, joissa oli kuvattu tarkemmin sekä tekoälyratkaisun käyttöönottoprosessia että käyttöönoton lopputuloksia, ei ainoastaan teoreettista keskustelua tekoälyn käyttökohteista.

Pro gradu -tutkielmani tavoite on tarkastella, miten tekoälyratkaisuja on hyödynnetty sosiaalityössä ja mitä tietoa tekoäly tuottaa sosiaalityön prosesseissa. Tekoälyn käyttökohteet sosiaalityössä -teeman alle nousi koko aineistosta yhteensä 21 artikkelia. Näistä lopulliseen tarkasteluun valitsin sellaiset artikkelit, joissa oli tarkasteltu monipuolisesti sekä sosiaalityön käytötapausta että tekoälyn tuottamaa tietoa. Vastatakseni tutkimuskysymyksiin, keskityin artikkeleiden valinnassa kolmeen pääteemaan, jotka tulivat käydä ilmi aineistosta. Tutkimuksen tuli paikantua johonkin sosiaalityön kontekstiin, tutkimuksesta tuli käydä ilmi käytössä oleva teknologia sekä tiedonmuodostuksen logiikkaa tuli olla kuvattu.

Viisi artikkelia rajautui pois, koska ne olivat kirjallisuuksia aiemmista tutkimuksista ja kartoittivat tekoälyn käytön yleisyyttä sosiaalityössä. Lisäksi yksi artikkeleista rajautui pois, koska se kohdistui kolmannella sektorilla tehtävään vapaaehtoistyöhön, eikä siten käytötapauksena soveltunut suoraan tutkimukseni kohteeksi. Kahta artikkelia ei ollut sähköisesti saatavilla koko tekstinä LUC-FINNAN sähköisen haun kautta. Tämän haun perusteella tutkimuskysymyksiini vastasi parhaiten 13 artikkelia.

Kuvaus hakuprosessista (haku toteutettu 15.10.2025):

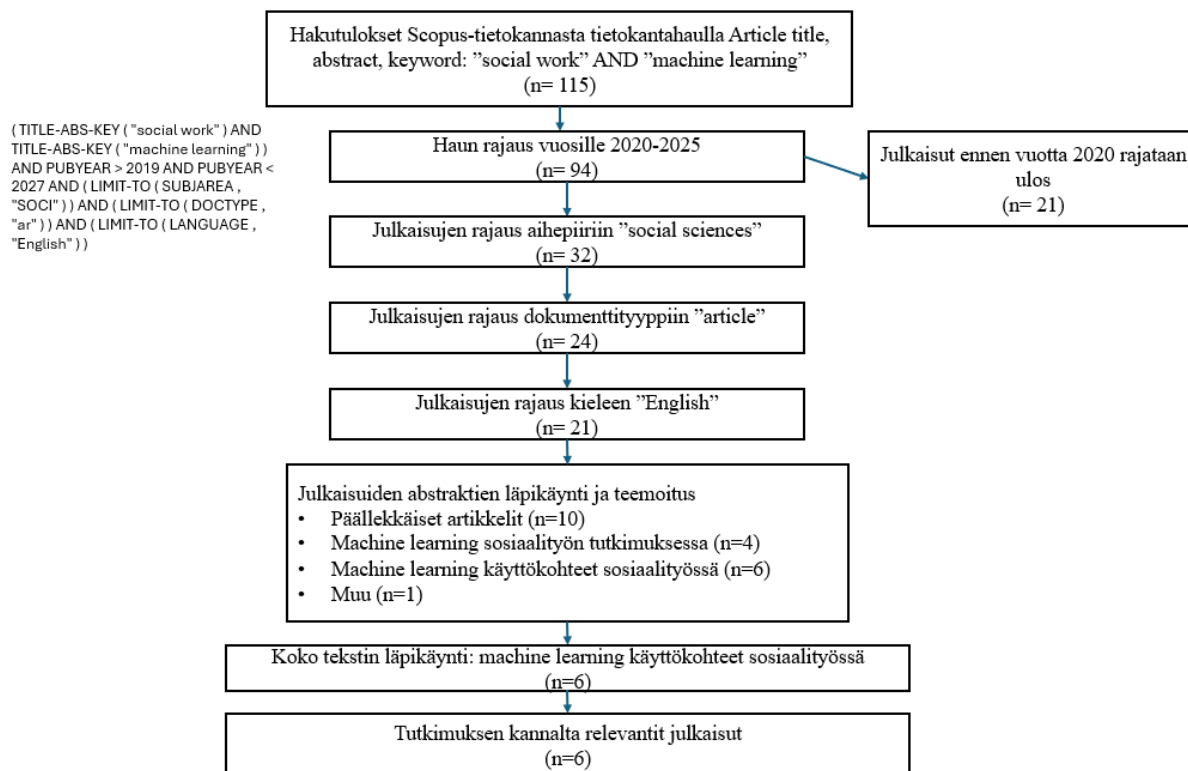


Perehtyessäni artikkeleihin tarkemmin, huomasin kuitenkin aineiston painottuvan hyvin vahvasti yhdenlaiseen tekoälymenetelmään: viime vuosien nopea teknologinen kehitys ja erityisesti laajojen kielimallien (Large Language Models, LLM) yleistyminen ja saatavuuden helpottuminen näkyvät julkaisuissa merkittävästi: kolmestatoista artikkelista yhdeksän keskittyy näiden mallien soveltuvuuden, mahdollisuuksien ja rajoitteiden arviointiin. Tutkimuksissa tarkastellaan LLM-teknologioiden potentiaalia tiedonhallinnan, dokumentaation ja ammatillisen päätöksenteon tukena.

Vain loput neljä käsittelee muita tilastollisia ja laskennallisia menetelmiä, kuten regressioanalyseja, ennakoivia malleja ja muita datalähtöisiä ratkaisuja, jotka tukevat sosiaalityön eri tehtäviä ja prosesseja. Vaikka aineisto kokonaisuudessaan heijasteli sosiaalityön kentän kasvavaa kiinnostusta tiedolla johtamista ja dataperusteista ammatillista kehittämistä kohtaan, antoi se yksipuolisen kuvan tekoälymenetelmien hyödyntämisestä.

Laajensin hakua koskemaan myös koneoppimista käsittelevät artikkelit ("machine learning"), jotta saisin monipuolisemman otoksen käyttötapauksista. Tein vastaavat rajaukset ja teemoituksen kuin edellisessä haussa ja sain tulokseksi yhteensä 6 artikkelia.

Kuvaus hakuprosessista (haku toteutettu 10.1.2026):



Molempien hakupolkujen lopputuloksena lopulliseen tarkasteluun valikoitui yhteensä 19 artikkelia. Artikkeleissa on edustettuna menetelmiltään sekä laadullisia että määrällisiä tutkimuksia. Niissä kuvataan tämän tutkimuksen tavoitteiden näkökulmasta riittävällä tarkkuudella tekoälyratkaisun kehittämistä, käyttöönottoa ja integrointia osaksi sosiaalityön tehtävää.

Alla olevissa taulukoissa olen kuvannut aineiston keruupolun SCOPUS-tietokannasta kahdella eri hakusanayhdistelmällä toteutettuna.

Taulukko 1: Aineistonkeruupolku hakusanalla "social work" ja "artificial intelligence"

Tietokanta	#	Hakurajaus	Tietokantakysely	Hakutulos
SCOPUS	1	Haku asiasanapareilla tekstistä ("Title", "Abstract" tai "Keyword"), rajausta vain vuodesta 2020 julkaistuihin aineistoihin, artikkeleihin, sosiaalialaan ja vain englannin kielellä julkaistuihin aineistoihin.	( TITLE-ABS-KEY ( "social work" ) AND TITLE-ABS-KEY ( "artificial intelligence" ) ) AND PUBYEAR > 2019 AND PUBYEAR < 2027 AND ( LIMIT-TO ( SUBJAREA , "SOC" ) ) AND ( LIMIT-TO ( DOCTYPE , "ar" ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) )	72 artikkelia

Ensimmäisen hakuvaiheen jälkeen toteutin haun toisella aineistonkeruupolulla, johon sisällytin koneoppimista ("machine learning") käsittelevät artikkelit.

Taulukko 2: Aineistonkeruupolku hakusanalla "social work" ja "machine learning"

Tietokanta	#	Hakurajaus	Tietokantakysely	Hakutulos
SCOPUS	2	Haku asiansanapareilla tekstistä ("Title", "Abstract" tai "Keyword"), rajaus vain vuodesta 2020 julkaistuihin aineistoihin, artikkeleihin, sosiaalialaan ja vain englannin kielellä julkaistuihin aineistoihin.	( TITLE-ABS-KEY ( "social work" ) AND TITLE-ABS-KEY ( "machine learning" ) ) AND PUBYEAR > 2019 AND PUBYEAR < 2027 AND ( LIMIT-TO ( SUBJAREA , "SOCI" ) ) AND ( LIMIT-TO ( DOCTYPE , "ar" ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) )	21 artikkelia

Useissa tekoälyä käsittelevissä kansainvälisissä tutkimuksissa nostetaan haasteeksi tutkimusnäytön vähäisyys (mm. Li ym. 2025 ja Garkisch ym. 2025). Tämä vastaa myös omaa havaintoani käydessäni aineistoa läpi: empiiristä tutkimusta tekoälyratkaisujen konkreettisista käyttökohteista sosiaalityössä on toistaiseksi melko niukasti saatavilla. Huolimatta kasvaneesta kiinnostuksesta tekoälyä kohtaan ja teeman ajankohtaisuudesta, on vielä vaikeampi löytää kotimaista tutkittua tietoa tekoälyn hyödyntämisestä sosiaalityössä. Lehtiniemen (2024) tutkimusartikkeli suomalaisessa lastensuojelun kontekstissa toteutetusta tekoälypohjaisesta riskien ennakoimista on ainoa hakutuloksista noussut artikkeli.

Taulukko 3: Lopulliset artikkelit

1. Perron Brian Edward, Hiltz Barbara S., Khang Erin M. & Savas Sue Ann (2025) AI-Enhanced Social Work: Developing and Evaluating Retrieval-Augmented Generation (RAG) Support Systems. Journal of Social Work Education, 61 (1), p. 3-13. 10.1080/10437797.2024.2411172	Artikkelissa tarkastellaan tekoälypohjaista retrieval-augmented generation (RAG) -teknologiaa, jonka tavoitteena on parantaa päätöksentekoa ja asiakastukea. Tapauksittain tutkimuksen avulla tarkastellaan RAG-pohjaisen työkalun kehittämistä ja arviointia ja käsitellään sen sovellettavuutta erilaisiin sosiaalityön haasteisiin.	Tulokset vahvistavat RAG-tekniikan potentiaalia uudistaa sosiaalityön käytäntöjä, sillä se tarjoaa tarkkaa ja luotettavaa tietoa ja parantaa tekoälyratkaisun tarkkuutta.
2. Karataş Zeki (2025) Can a Large Language Model Judge a Child's Statement?: A Comparative Analysis of ChatGPT and Human Experts in	Tutkimuksen tavoitteena on tarkastella tekoälyn käytön potentiaalia, rajoituksia ja ohjelmointia lasten seksuaaliväkivallan havaitsemiseksi ja hyväksikäyttöä koskevien lausuntojen	ihmisasiantuntijat arvioivat lausuntoja yhdenmukaisesti ja luotettavasti, mutta jyrkkä ja merkittävä luotettavuuden lasku havaittiin, kun tekoälymallin arvioita

<p>Credibility Assessment. Journal of Evidence-Based Social Work (United States). 10.1080/26408066.2025.2547211</p>	<p>arvioinnissa. Tämän lisäksi vertaillaan ihmisasiantuntijoiden sekä suuren kielimallin (LLM) tekemien arviointien luotettavuutta.</p>	<p>verrattiin ihmisten arvioihin. Tekoäly osoitti järjestelmällisiä erimielisyyksiä kriteereissä, jotka edellyttävät hienovaraista, kontekstisidonnaista arviointia.</p>
<p>3. Taskiran Tepe Hacer &amp; Aslantürk Hüsnünür (2025) Comparative Efficacy of AI LLMs in Clinical Social Work: ChatGPT-4, Gemini, Copilot. Research on Social Work Practice. 10.1177/10497315241313071</p>	<p>Tutkimuksessa tarkastellaan kolmen suuren kielimallin (LLM) – ChatGPT-4:n, Geminin ja Microsoft Copilotin – tehokkuutta kliinisessä sosiaalityössä. Mallien suorituskykyä arvioitiin esittämällä niille eritasoisia skenaarioita joihin arvioimalla vastauksia Ateşmanintarjoavien luottavuusindeksin sekä Likert-asteikolle perustuvan tarkkuusmittarin avulla.</p>	<p>Tulokset osoittivat, että Gemini saavutti korkeimman tarkkuuden, kun taas Microsoft Copilot erottui edukseen luottavuudessa. Tulokset viittaavat siihen, että vaikka suuret kielimallit tarjoavat lupaavia mahdollisuuksia sosiaalityössä, ne edellyttävät edelleen hienosäätöä, jotta ne voisivat vastata paremmin alan tarpeisiin.</p>
<p>4. Segal Michal (2025) Confronting and managing ethical dilemmas in social work using ChatGPT. European Journal of Social Work, 28 (1), p. 155-167. 10.1080/13691457.2024.2377786</p>	<p>Tässä tutkimuksessa ChatGPT:tä verrattiin Israelin eettisen komitean vastauksiin sen hyödyllisyyden arvioimiseksi sosiaalityöntekijöiden eettisen päätöksenteon tukena. Israelin eettiselle komitealle aiemmin lähetetyt kysymykset käännettiin ja esitettiin ChatGPT-4:lle.</p>	<p>ChatGPT huomioi monenlaisia näkökulmia ja toi johdonmukaisesti esiin ristiriitaisia velvollisuuksia ja ammatillisia rajoja. Eettisen komitean vastaukset puolestaan tarjosivat selkeämpää ja käytännönläheisempää ohjausta sekä painottivat enemmän juridisia kysymyksiä. ChatGPT:ssä on potentiaalia ammatillaisen tukityökaluksi, mutta se ei vielä kykene itsenäisesti ratkaisemaan eettisiä dilemmaa.</p>
<p>5. Lehtiniemi Tuukka (2024) Contextual social valences for artificial intelligence: anticipation that matters in social work. Information Communication and Society, 27 (6), p. 1110-1125. 10.1080/1369118X.2023.2234987</p>	<p>Lastensuojelun sosiaalityöntekijät pilotoivat tekoälytyökalua riskien ennakointiin. Artikkelissa tarkastellaan odotuksia tekoälyn arvosta ja suorituskyvystä sosiaalityössä ja laajemmin. Tekoälyn arvon oletetaan syntyvän sen kyvystä tuottaa ennakoivaa tietoa.</p>	<p>Tekoälyä ei tulisi nähdä ensisijaisesti tietoa tuottavana välineenä, vaan asiiantuntijoiden tiedonmuodostusprosesseja tukevana työkaluna.</p>
<p>6. Meilvang Marie Leth, Dahler Anne Marie (2024) Decision support and algorithmic support: the construction of algorithms and professional discretion in social work: Beslutningsstøtte og algoritmisk støtte: Konstruktionen af algoritmer og det professionelle skøn i socialt arbejde. European Journal of</p>	<p>Artikkelissa tarkastellaan kolmen päätöksenteon tukena käytettävän algoritmin taustalla olevia ajatuksia, niiden kehittämistä ja käyttöä. Algoritmit on kehitetty Tanskankuntien käyttöön haavoittuvassa asemassa olevien lasten ja perheiden sosiaalityöhön. Ne on tarkoitettu ammattilaisten päätöksenteon tueksi lähetettä koskeissa ratkaisuissa.</p>	<p>Analyysin perusteella havaitaan, että algoritmit on tarkoitettu tukemaan päätöksentekoa vakiomalla asiakastyötä, tekemällä päätösten perustasta objektiivisempi ja poistamalla vinoumia. Lisäksi tuomme esiin algoritmien tavoitteiden ja ammatillisen harkintavallan käsitteen läheisen suhteen.</p>

Social Work, 27 (1), p. 30-42. 10.1080/13691457.2022.2063806	Analysoimme päätöksenteon tukialgoritmien tavoitteita, suunnittelua ja toimintaa sosiaalityössä.	
7. Hao Zhichao, Yang Beichen, Ruggiano Nicole, Ma Yunying, Guo Yuqi, Pan Xiaofu (2022) Depression Prediction amongst Chinese Older Adults with Neurodegenerative Diseases: A Performance Comparison between Decision Tree Model and Logistic Regression Analysis. British Journal of Social Work, 52 (1), p. 274-290. 10.1093/bjsw/bcaa237	Tekoälyyn (AI) perustuvat data-analyysit eivät ole yleistyneet sosiaalityössä yhtä laajalti kuin muilla tieteenaloilla. Tutkimus keskittyi kiinalaisiin neurodegeneratiivisia sairauksia (ND) sairastaviin ikääntyneisiin ja pyrki (i) kehittämään päätöspuomenetelmään perustuvan ennustemallin masennukseen yhteydessä olevien tekijöiden tunnistamiseksi sekä (ii) vertaamaan päätöspuomallin ennustesuoriutumista logistiseen regressioanalyysiin.	Logistiseen regressioon verrattuna päätöspuomalli suoriutui masennuksen ennustamisessa selvästi paremmin. Tutkijoiden, ammattilaisten ja päättäjien tulisi kiinnittää huomiota keinoihin, joilla vähennetään masennusalttiutta kiinalaisilla ND-sairauksia sairastavilla ikääntyneillä. Lisäksi päätöspuomallia voidaan hyödyntää viitekehyksenä muiden fyysisten tai mielenterveys-sairauksien ennustamisessa ja analysoinnissa.
8. Wilkins David & Benett Verity (2025) Making Accurate Judgements in Child Welfare: Comparing ChatGPT With Qualified Social Workers. Child and Family Social Work. 10.1111/cfs.13304	Tutkimuksessa verrattiin lasten ja perheiden parissa työskentelevien sosiaalityöntekijöiden arviointitarkkuutta ChatGPT:hen. Osallistujat arvioivat 12 anonymisoitua asiakaslähetettä tulevien tapahtumien todennäköisyyttä. ChatGPT suoriutui hieman keskimääräistä sosiaalityöntekijää paremmin arviointitehtävistä.	Tulokset korostavat tekoälyn potentiaalia mutta myös sen rajoituksia sosiaalityön päätöksenteon tukemisessa, esimerkiksi liittyen monimutkaisten inhimillisten tarpeiden ja sosiaalisten tilanteiden ymmärtämiseen.
9. Perron Brian Edward, Luan Hui, Victor Bryan G., Hiltz-Perron Oliver, Ryan Joseph P. (2024) Moving Beyond ChatGPT: Local Large Language Models (LLMs) and the Secure Analysis of Confidential Unstructured Text Data in Social Work Research. Research on Social Work Practice. 10.1177/10497315241280686	Laajojen kielimallien käyttöä sosiaalityön tutkimuksessa rajoittavat luottamuksellisuuden ja tietoturvan kysymykset liittyen arkaluonteiseen dataan. Tutkimuksessa arvioitiin paikallisesti käytettävien kielimallien suorituskykyä pähteisiin liittyvien ongelmien luokittelussa ja poiminnassa jäsentämättömistä lastensuojelun tutkintayhteenvedoista.	Paikalliset kielimallit saavuttivat poikkeuksellisen hyvät tulokset, jotka olivat verrattavissa ihmisasiantuntijoiden suoritukseen. Tutkimus korostaa paikallisten kielimallien käyttökelpoisuutta suurten tekstimäärien tehokkaaseen analysointiin luottamuksellisesta ja rakenteettomasta aineistosta.
10. Chan Chitat, Zhao Jiahui, Zhao Yunmeng (2025) Our journey with an AI assistant offering narrative therapy on WhatsApp. Journal of Social Work Practice, 39 (2), p. 165-189. 10.1080/02650533.2024.2420314	Tutkimus selvittää, kuinka toteuttamiskelpoista on käyttää räätälöitäviä GPT-alustoja narratiivisen terapian istuntoihin suunnitellun chatbotin kehittämiseen ja arvioisen suoriutumista.	Chatbot voidaan konfiguroida ohjaamaan käyttäjiä terapiavaiheiden läpi ja mukautumaan vuorovaikutukseen. Käyttöönotto rajoittaa vaatimus perustason koodaus- ja teknisestä osaamisesta.
11. Segal Michal (2025) Social workers' evaluation of ChatGPT for solving ethical dilemmas within the limits of	Tutkimuksessa tarkastellaan, miten sosiaalityöntekijät arvioivat ChatGPT:n käyttöä luottamuksellisuuteen liittyvien eettisten	Analyysi tunnisti kolme päähaastetta: erot monimutkaisten ammatillisten kysymysten jäsentämisessä,

<p>confidentiality. Journal of Social Work Practice.</p> <p>10.1080/02650533.2025.2480092</p>	<p>kysymysten ratkaisemisessa. Sosiaalityön maisteriopiskelijat vertasivat omia eettisiin kysymyksiin pohtimaan ratkaisuja ChatGPT:ntuottamiin ratkaisuihin.</p>	<p>Viittauksissa eettisiin koodeihin ja lainsäädäntöön sekä tunneulottuvuudessa ja empatiassa. ChatGPT voi tukea työtä, mutta ei korvaa sosiaalityöntekijöiden eettistä harkintaa ja empatiaa.</p>
<p>12. Gökçe Arslan Elif, Tikaş Çelik Sevgi, Akdeniz Emel Tuba &amp; Öztürk Emel (2025) The Touch of Artificial Intelligence in Social Work: Analysis of Social Investigation Reports in Child Welfare with ChatGPT. Journal of Evidence-Based Social Work (United States), 22 (4), p. 607-625.</p> <p>10.1080/26408066.2025.2480396</p>	<p>Tutkimuksen tavoitteena oli tarkastella ChatGPT:n antamia vastauksia seitsemään lapsiin liittyvään anonymisoituun sosiaalitutkimusraporttiin.</p>	<p>Tekoälysovellus ottaa merkittävässä määrin hoitaakseen ihmisten aiemmin tekemiksi miellettyjä tehtäviä. Tulevat tutkimukset voisivat tarkastella tekoälyn käytön vaikutuksia sosiaalityön käytännöissä alan asiantuntijoiden näkökulmasta.</p>
<p>13. Raya Díez Esther, Trujillo-Carmona Manuel &amp; Carbonero Muñoz Domingo (2021) Using Big Data to Manage Social Inclusion Programs. Journal of Sociology and Social Welfare, 48 (3), 6, p. 74-98.</p>	<p>Tekoäly nojaa Big Data -hallintajärjestelmiin, jotka puolestaan mahdollistavat kuvailevan päättelyn, ennaltaehkäisevät toimenpiteet sekä dataan perustuvan päätöksenteon. Tässä artikkelissa kuvataan Big Datan keskeisiä piirteitä ja esitellään prosessi, jossa suunniteltiin sosiaalisen syrjäytymisen diagnosointiin tarkoitettu työkalu, SiSo-asteikko.</p>	<p>Tulokset vahvistavat työkalun soveltuvuuden ja arvon sosiaalisen syrjäytymisen riskitason mittaamisessa sekä osoittavat datalähtöiseen sosiaalisen osallisuuden järjestelmään perustuvan toteutuksen olevan mahdollista. Artikkelin tuo myös esiin Big Datan tarjoaman mahdollisuuden tuottaa tietoa sosiaalityön käyttöön.</p>
<p>14. Portela Manuel, Castillo Carlos, Tolan Songul, Karimi-Haghighi Marzieh &amp; Pueyo Antonio Andrés (2025), A comparative user study of human predictions in algorithm-supported recidivism risk assessment. Artificial Intelligence and Law, 33 (2), p. 471-517.</p>	<p>Artikkelissa tarkastellaan algoritmipohjaisen riskinarviointivälineen (RAI) käytön vaikutuksia väkivaltaisen uusintarikollisuuden riskinennustamisen tukena. Tutkimuksessa käytettiin koneoppimiseen perustuvaa versiota RiskCanvi-järjestelmästä.</p>	<p>RAI-välineen avulla tehty ennusteen tarkistaminen parantaa kaikkien ryhmien suoriutumista. Alan ammattilaiset suoriutuvat tehtävästä paremmin kuin ei-asiantuntijat. Lisäksi RAI-välinettä voidaan pitää keinona laajentaa ja tukea ammattilaisten osaamista ja kyvykkyyksiä.</p>
<p>15. Kim Jinyi, Hsu Julia Hsin Ping, Sohn Gahyun, Lee Giemyung &amp; Lee Myeongs (2025), Leveling Socioeconomic Disparities: The Role of Service Availability in School Dropout Rates. Research on Social Work Practice, 10497315251377009.</p>	<p>Tutkimuksen tavoitteena on tarkastella sosioekonomisen aseman (SES), mielen- ja päihdehoitopalvelujen saatavuuden sekä koulupudokkuusasteiden välisiä yhteyksiä kolmessa osavaltiossa. Paikkatietoon perustuvia koneoppimismenetelmiä hyödynnettiin aineistojen yhdistämiseen ja puuttuvien arvojen imputointiin.</p>	<p>Tulokset viittaavat siihen, että laskennalliset menetelmät voivat tukea moniosavaltioista koulutusaineistojen analyysia. Tutkimus osoittaa paikkatietojen ja koneoppimismenetelmien potentiaalisen monitasoisessa analyysissa julkisten palvelujen tukena.</p>

<p>16. Bogenschutz Matthew D., Broda Michael D., Dinora Parthenia A., Prohn Seb M., Lineberry Sarah Nelson &amp; West Angela (2025), Applying machine learning to understand factors predicting pharmacotherapy for mental health support among adults with intellectual and developmental disabilities. <i>Social Work in Mental Health</i>.</p>	<p>Yhdysvalloissa kehitysvammaisten henkilöiden mielenterveyteen liittyviä ongelmia hoidetaan muuta väestöä useammin lääkkeellisesti. Tässä tutkimuksessa hyödynsimme koneoppimistavanalysoidaksemme, mitkä tekijät ennustavat parhaiten lääkehoidon käyttöä mielenterveyden tukemisessa kehitysvammaisilla henkilöillä.</p>	<p>Käyttämämme algoritmi suoriutui erinomaisesti lääkehoidon käytön ennustamisesta. Tutkimus nostaa tietoisuutta kehitysvammaisten lääkehoidon juurisyistä sekä kannustaa hyödyntämään koneoppimisen menetelmiä vastaavissa tutkimuksissa.</p>
<p>17. Villumsen Anne Marie Anker, Rosholm Michael, Bodilsen Simone Tranberg, Toft Sanne Dalgaard, Berg Line Svolvegaard &amp; Nirmalarajan Liesanth Yde (2024), Ethical considerations in research when building predictive risk modelling in child and family welfare. <i>Journal of Comparative Social Work</i>, 19(1), p. 102-126.</p>	<p>Tässä artikkelissa esitellään ja tarkastellaan eettisiä kysymyksiä ja niiden vaikutuksia tutkimuksessa, jossa rakennetaan ennakoivaa riskimallia mahdolliseen käyttöön Tanskan lasten- ja perhepalveluissa. Tavoitteena tutkia, voisiko ennakoiva riskimalli tuottaa lisäarvoa lasten- ja perheiden palveluille riskiarvioinnissa – erityisesti ilmoituksiin liittyvän päätöksenteon tukena.</p>	<p>Riskienarviointimallin kehityksessä korostuvat eettiset näkökohdat, korostunut merkitys, laadullisen että määrällisen tutkimuksen kasvattaminen sekä yhteistyön merkitys mallien kehittäjien, sosiaalityön käytännön toimijoiden sekä riskissä olevien lasten ja perheiden välillä.</p>
<p>18. Victor Bryan G., Perron Brian Edward, Sokol Rebeccah Lyn, Fedina Lisa &amp; Ryan Joseph P. (2021), Automated identification of domestic violence in written child welfare records: Leveraging text mining and machine learning to enhance social work research and evaluation. <i>Journal of the Society for Social Work and Research</i>, 12(4), p. 631-655.</p>	<p>Tässä tutkimuksessa tarkastellaan tekstinlouhinnan ja koneoppimismenetelmien toteutettavuutta perheväkivaltaan liittyvien ongelmien tunnistamisessa lastensuojelun selvitysyhteenvetoihin dokumentoidusta tekstistä.</p>	<p>Koneoppimismallit saavuttivat yli 90 prosentin tarkkuuden verrattuna ihmiskoodaajiin. Tekstinlouhinta ja koneoppimismenetelmät voivat olla kustannustehokas ratkaisu merkityksellisen tiedon poimimiseen tekstiaineistoista. Vaikka menetelmät eivät sovellu tapauskohtaisiin ennustaviin analyyseihin, niiden tuottamat havainnot voivat olla erityisen hyödyllisiä sosiaalisten ilmiöiden tarkastelussa.</p>
<p>19. Qin Fangyu, Lv Zheqi, Wang Dan Ni, Hu Bo &amp; Wu Chao (2020), Health status prediction for the elderly based on machine learning. <i>Archives of Gerontology and Geriatrics</i>, 90, 104121.</p>	<p>Hoivaa tarvitsevien ikääntyneiden palvelujen järjestäminen edellyttää entistä tarkempia ennusteita ikääntyneiden terveydentilasta, jotta resurssit voidaan kohdentaa tarkoituksenmukaisesti. Testasimme koneoppimismenetelmien suorituskkyä ikääntyneiden hoivan tarpeen ennustamisessa</p>	<p>Koneoppimismenetelmät pystyvät perinteisiä analyysimenetelmiä tarkemmin mallintamaan muuttujien välisiä epälineaarisia suhteita.</p>

Vaikka tekoälyn mahdollisuuksista keskustellaan laajasti, valtaosa löydetystä kansainvälisestä tutkimuksesta keskittyy sosiaalityöntekijöiden ja opiskelijoiden asenteisiin tekoälyä kohtaan, tekoälyn käyttöön liittyviin eettisiin kysymyksiin sekä koulutuksellisiin ja osaamisen

kehittämiseen liittyviin näkökulmiin. Varsinainen musta aukko tutkimuksessa näyttää paikantuvan erityisesti tekoälypohjaisten ratkaisujen vaikuttavuuden arviointiin (mm. Li ym. 2025, 7).

Tämä tutkimuksellinen epäsuhta osoittaa, että tekoälyn käytännön sovellukset ovat sosiaalityön kentällä vielä alkuvaiheessa, ja siksi niitä koskevaa empiiristä tietoa on rajallisesti. Suunta kuitenkin on muuttumassa, sillä erityisesti vuoden 2025 aikana on julkaistu useampia tutkimuksia, jotka tarkastelevat asenteiden ja eettisten kysymysten ohella konkreettisia kokeiluja tai sovelluksia. Täten voin hyödyntää tutkimuksessani aivan ajankohtaisinta tietoa. Toivottavasti kansalliset ponnistelut tekoälypilottien tukemiseksi tuottavat myös osaltaan enemmän tämän suuntaista tieteellistä tutkimusta. Vain saattamalla tekoäly teoriasta konkretiaan voidaan muodostaa monipuolisempi kuva tekoälyn ja sosiaalityön välisestä suhteesta.

Tutkimusaineiston niukkuudesta huolimatta tutkimukseeni valikoidut artikkelit tarjoavat laaja-alaisen kuvan sosiaalityön tämänhetkisistä tehtävistä ja asiakastyön eri ulottuvuuksista suhteessa tekoälyyn. Aineiston perusteella on mahdollista muodostaa monipuolinen käsitys tekoälyn hyödyntämisen ja tiedonmuodostuksen välisestä yhteydestä sosiaalityön käytännöissä.

### **4.3 Aineiston analyysi ja metodologia**

Tutkimusstrategiana pro gradu -tutkielmassani käytän laadullista menetelmää. Tutkimukseni on laadullinen kirjallisuuskatsaus tekoälyn hyödyntämisestä sosiaalihuollossa. Näkökulmana on erityisesti tekoälyn hyödyntäminen tiedonmuodostuksessa ammattilaisen työn tukena. Koska tarkasteltava kokonaisuus on melko uusi eikä valitsemastani näkökulmasta ei löydy laajasti tutkimusaineistoa, käytän menetelmänä aineistolähtöistä sisällönanalyysia.

Aineistolähtöisessä analyysissä pyritään rakentamaan tutkimusaineistosta teoreettinen kokonaisuus siten, etteivät aiemmat havainnot tai teoriat ohjaa analyysin lopputulosta. Sisällönanalyysillä pyritään järjestämään aineisto tiiviiseen ja selkeään muotoon kadottamatta sen sisältämää informaatiota. Aineistolähtöisessä analyysissä on erityisen tärkeää arvioida analyysin objektiivisuutta sekä reflektoida tutkijan omaa positiotaan ja omia ennakkokäsityksiään suhteessa tutkimuskohteeseen ja aineistoon. (Tuomi & Sarajärvi 2018, Luku 4.4)

Analysoin aineiston temaattisella analyysillä, koska tavoitteena on tunnistaa aineistosta merkityksiä, teemoja ja rakenteita ilman ennalta määrättyä teoreettista kehikkoa. Tuomen ja Sarajärven (2018) mukaan aineistolähtöisessä analyysissä ja temaattisessa analyysissä on selviä yhtymäkohtia. Molemmista lähestymistavoissa analyysi etenee aineistosta käsin: aineistoa luetaan, koodataan ja jäsennetään siten, että siitä voidaan tunnistaa toistuvia teemoja sekä muodostaa niitä kuvaavia kategorioita tai teemoja. Menetelmien tavoitteena on tuoda esiin aineistossa ilmeneviä havaintoja ja jäsennyksiä ja rakentaa niiden pohjalta ymmärrettävä kokonaisuus, vaikka käsitteet ja analyysin painotukset voivat vaihdella tutkimustradition mukaan.

Tekoälypohjaisen tiedonmuodostuksen prosessin kuvaus, tekoälyn tuottama lopputulos sekä käytetty tekoälymenetelmä muodostavat tämän tutkimuksen aineistosta nousevat pääteemat. Syvennän näitä teemoja tarkentamalla niihin liittyviä alateemoja kunkin artikkelin sisällöstä.

Taulukko 4: Artikkeleista löytyvät teemat

	Menetelmä	Läpinäkyvyys	Tietopohja tekoälyn tukena	Lopputuotos
1	LLM + RAG	Ei	Yleiskäyttöinen kielimalli + toimialasidonnainen tietämyskanta	Narratiivinen keskustelu (yleinen ohjaus ja neuvonta)
2	LLM	Ei	Yleiskäyttöinen kielimalli	Sanallinen arvio/lausunto (lastensuojelu)
3	LLM	Ei	Yleiskäyttöinen kielimalli	Sanallinen arvio/lausunto (sosiaalityö)
4	LLM	Ei	Yleiskäyttöinen kielimalli	Sanallinen arvio/lausunto (eettiset kysymykset)
5	ML (oletus)	Ei	Aiempi yksilötason asiakas- ja potilastieto, muuttujia ei kuvattu	Riskin ennustaminen (lastensuojelu)
6	ML (oletus)	Ei	Aiempi yksilötason potilas- ja asiakastieto, aikaisemmat tukitoimet, tietoa vanhemmista, muuttujia ei kuvattu	Luokittelu (lastensuojeluilmoitus) Riskin ennustaminen (lastensuojelu)
7	ML	Kyllä	Hyvinvointiin ja terveyteen liittyvät muuttujat kuvattu	Riskin ennustaminen (ikäntyneet)
8	LLM	Ei	Yleiskäyttöinen kielimalli	Sanallinen arviointi/lausunto (lastensuojelu)
9	LLM	Ei	Yleiskäyttöinen kielimalli + lastensuojeluraportit	Tiedon haku ja jäsentäminen (väestö- ja aluetason hyvinvointitieto lastensuojelussa)
10	LLM	Ei	Yleiskäyttöinen kielimalli täydennettynä narratiivisen terapian perusrakenteen ja vuorovaikutuksen ohjeilla	Narratiivinen keskustelu (terapia)
11	LLM	Ei	Yleiskäyttöinen kielimalli	Sanallinen arviointi/lausunto (eettiset kysymykset)
12	LLM	Ei	Yleiskäyttöinen kielimalli + erityisalakohdainen lainsäädäntö	Sanallinen arviointi/lausunto (lastensuojelu)
13	Tilastollinen analytiikka	Ei	Sosiaalisen syrjäytymiseen liittyvät muuttujat kuvattu	Alueelliset muuttujat ja riippuvuudet: syrjäytyminen
14	ML	Kyllä	Rikoksen uusimiseen liittyvät muuttujat kuvattu	Riskin ennustaminen (rikollisuus)
15	ML	Kyllä	Alueelliset ja sosioekonomiset muuttujat kuvattu	Alueelliset muuttujat ja riippuvuudet: koulupudokkuus
16	ML	Kyllä	Aiemmat asiakas- ja potilaskirjaukset, tukitarpeet ja strukturoitu haastattelu, muuttujat kuvattu	Riskin ennustaminen (vammaset, mielenterveyden ongelmat)
17	ML	Kyllä	Aiemmat asiakas- ja potilaskirjaukset, muuttujat kuvattu	Riskin ennustaminen (lastensuojelu)
18	ML	Kyllä	Kirjalliset vapaamuotoiset yhteenvedot lasten kaltoinkohtelu epäilyjen selvityksistä, muuttujat kuvattu	Alueelliset palvelutarpeet: lastensuojelu
19	ML	Kyllä	Aiemmat asiakas- ja potilaskirjaukset, muuttujat kuvattu	Riskiennustaminen (ikäntyneiden terveydentila)

Ensimmäisessä vaiheessa ryhmittelen artikkelit sen mukaan, mitä tekoälymenetelmää ne käsittelevät. Artikkeleista nousee kaksi keskeistä menetelmää: laajoihin kielimalleihin perustuvat (LLM) sekä muut koneoppimisen (ML) alle asettuvat menetelmät. Tämä ryhmittely antaa mahdollisuuden arvioida sitä, kuinka monipuolisesti erityyppisiä menetelmiä on hyödynnetty sosiaalityössä. Koska kyseessä on sosiaalityön tutkimus, avaan tekoälyratkaisujen

tekniisiä yksityiskohtia vain sen verran kuin koen tarpeelliseksi niiden toimintalogiikan ymmärtämiseksi ja sosiaalityön asiayhteyteen kiinnittämiseksi.

Artikkelien jaottelusta mainittakoon vielä, että kahdessa artikkelissa tekoälyn toimintalogiikkaa esiteltiin teoreettisessa osiossa, mutta varsinaisesti tekoälymenetelmää ja sen käyttöönottoa ei täysin kuvattu. Tällöin taulukon menetelmää kuvaavaan sarakkeeseen lisäsin sivuhuomautuksen ”oletus”. Sosiaalista syrjäytymistä käsittelevässä artikkelissa puolestaan käytiin läpi tekoälypohjaisen ratkaisun tietopohjaan liittyviä vaatimuksia, mutta tekoälyratkaisun soveltamiseen liittyvä analyysivaihe jäi ohueksi, joten analyysissä käsittelen käyttötapausta puhtaasti tilastollisen analytiikan ratkaisuna, huomioiden sen edellytykset toimia tekoälypohjaisen ratkaisun taustalla.

Toisessa vaiheessa syvennän analyysiä tarkastelemalla tiedonmuodostuksen prosessia ja ulottuvuutta. Ryhmittelyperusteena arvioin, onko tiedonmuodostuksen prosessi läpinäkyvä, eli onko aineistossa kuvattu sekä tekoälymenetelmä että tietopohja, jota tekoälymenetelmä hyödyntää. Lopuksi jaottelen artikkelit myös tekoälyratkaisun tavoitellun lopputuotoksen mukaan, eli minkälaisen lopputuotoksen tekoälymenetelmä tuottaa sosiaalityöntekijän tiedonmuodostusprosessissa.

Tarkastelen artikkeleita myös sen perusteella, mitä sosiaalityön tehtäviä tutkimuksen kohteena olevat tekoälyratkaisut tukevat. Sosiaalityön tehtäviä voi tarkastella joko asiakasprosessin eri vaiheiden näkökulmasta, joissa korostuu vireilletulon käsittely, palvelutarpeen arviointi, palveluiden suunnittelu ja järjestäminen sekä palvelun toteutus (STM 2024). Asiakasprosessin eri kohdissa korostuvat erilaiset tietotarpeet. Vaihtoehtoinen tapa tarkastella sosiaalityön palveluita Suomessa on jakaa ne palvelutehtävittäin, jolloin määrittäväksi tekijäksi nousee asiakasryhmä, jota ne palvelevat. Palvelutehtävät eivät kuitenkaan sovellu tähän tutkimukseen, sillä kansainvälisten artikkeleiden toimintaympäristö ja palvelurakenne eroavat suomalaisesta verrokistaan. Päädyn tässä tutkimuksessa jäsentämään tuotoksia asiakasryhmien näkökulmasta, koska sitä kautta voin tarkastella, minkälaisia erityispiirteitä erilaisten asiakasryhmien kanssa tehtävään työhön ja päätöksentekoon liittyy.

#### **4.4 Tutkimuseettiset kysymykset**

Tutkimuseettinen neuvottelukunta (Asetus tutkimuseettisestä neuvottelukunnasta 1347/1991) on opetus- ja kulttuuriministeriön asettama asiantuntijaelin. Se tuottaa hyvien tieteellisten

käytäntöjen ohjeistuksia suomalaiselle tutkimuskentälle. Tutkimuseettinen neuvosto (TENK 2023) määrittelee hyvän ja laadukkaan tutkimustyön arvoiksi luotettavuuden, rehellisyyden, arvostuksen ja vastuunkannon. Kirjoitusprosessin aikana olen reflektoinut omaa työskentelyäni aktiivisesti näitä arvoja vasten ja pyrkinyt noudattamaan niitä omassa työssäni.

Tutkimusartikkeleihin pohjautuva aineistolähtöinen analyysi nojautuu muiden tutkijoiden tuottamiin tutkimuksiin ja analyyseihin tutkittavasti ilmiöstä. Aineistolähtöisessä analyysissä tutkijan on tarkasteltava kriittisesti omaa positiotaan suhteessa tutkittavaan ilmiöön ja sitä, miten omat tulkinnat rakentuvat suhteessa aineistoon ja tutkimuskohteeseen. Samalla on tärkeää arvioida myös lähdeaineiston luonnetta: millaisista lähtökohdista käytetyt artikkelit on tuotettu ja millaiset oletukset, rajaukset tai intressit voivat vaikuttaa niiden esitystapaan. Keskeiseksi nousee kysymys siitä, missä määrin täydellistä neutraaliutta tai objektiivisuutta on ylipäätään mahdollista saavuttaa tutkimusprosessissa. (Tuomi & Sarajärvi 2018.)

Kirjoitusprosessin aikana arvioin, millaisia ennakkokäsityksiä minulla on suhteessa tutkittavaan ilmiöön ja miten ne voivat vaikuttaa tulkintoihin. Ennakkokäsitykset vaikuttavat koko tutkimusprosessin aikana, aina tutkimuskohteen valinnasta johtopäätösten laatimiseen saakka. Pro gradu -tutkielmaani valitsemani aihe on oman työkokemukseni perusteella tuttu, koska olen työskennellyt teknologiaratkaisujen parissa pitkään.

Tunnistan olevani teknologiamyönteinen. Tämä lähtökohta voi vaikuttaa ajatusmaailmaani siten, että tarkastelen uuden teknologian hyödyntämisen mahdollisuuksia positiivisesti ja toiveikkasti, mikä voi johtaa näkökulmani vinoutumiseen ja sitä kautta myös tutkimustulosten vääristymiseen. Pyrin tunnistamaan aineistosta mahdollisimman puolueettomasti sekä myönteiset että kielteiset näkökulmat tekoälyn hyödyntämiseen siinä laajuudessa kuin ne aineistossa ilmenevät, ilman että omat oletukseni ohjaavat tulkintoja tai painotuksia. Aineistoa analysoidessani keskityin esittämään havainnot tarkasti ja totuudenmukaisesti.

Tutkimusprosessin aikana olen myös tarkastellut omaa tapaaani hyödyntää tekoälyä osana kirjoitusprosessia. Olen käyttänyt tämän tutkielman toteuttamisen tukena ChatGPT 5.2 -versiota. Lapin yliopisto on antanut linjaukset Tekoälypohjaisten työkalujen käytöstä (2023). Kielimalleja saa lähtökohtaisesti hyödyntää, mutta niiden käyttö tulee rajoittaa kielenhuollollisiin tehtäviin tai ideoiden hakemiseen. Tekoälyä ei saa hyödyntää varsinaisen tutkielmatekstin tuottamiseen, koska tekstin tulee perustua opiskelijan omaan ajatteluun ja itsenäiseen työskentelyyn. Olen noudattanut annettuja linjauksia tarkasti ja huolellisesti. Kielenhuollollisten tehtävien lisäksi

olen hyödyntänyt kielimalleja termien käännöksiin, mutta olen aina itse myös pyrkinyt varmistamaan käännöksen oikeellisuuden ja alkuperäisen merkityksen säilymisen.

## 5 Tekoälyn tuottama tieto ja tuotokset sosiaalityössä

### 5.1 Lausunnot, arviot ja ehdotukset tekoälykollegan tuottamana

Peräti yhdeksän, eli lähes puolet kaikista pro gradu -tutkielmassani käsitellyistä tutkimusartikkeleista tarkastelee laajoihin kielimalleihin perustuvien menetelmien hyödyntämistä sosiaalityön eri konteksteissa. Laajat kielimallit nousivat erityisesti esiin tietokannasta hakusanalla ”artificial intelligence” tehdyissä hauissa. Tämä on sinänsä kiinnostava yksityiskohta ja voi viitata siihen, että käsitteet ”laajat kielimallit” ja ”tekoäly” mielletään puhekielessä monissa yhteyksissä osittain päällekkäisiksi tai jopa synonyymeiksi. Sama käsitteellinen sekoittuminen näkyy myös akateemisessa tutkimuksessa. Laajoihin kielimalleihin liittyvien tutkimusten määrän kasvu erityisesti vuonna 2025 kertonee myös siitä, että tämän teknologian kehitys on ottanut huimia harppauksia ja se on nyt helposti tavallisten ihmisen ja ammattilaisten ulottuvilla.

Akateeminen kiinnostus näyttää kohdistuvan erityisesti laajojen kielimallien kyvykkyyteen laatia erityyppisiä sanallisia tilannearvioita tai kannanottoja. Sanallisten tuotosten pituus ja sisältö vaihtelevat kulloisenkin tehtävänannon mukaisesti. Tekoäly tuottaa joko sanallisen lausunnon, arvion tai toimenpide-ehdotuksen asiakkaasta annettujen taustatietojen perusteella. Tehtävänsä suorittamiseksi tekoälylle annetaan taustatietoina muun muassa ammattilaisten tekemiä lausuntoja, kuvitteellisia tai anonymisoituja asiakastapauksia ja lähetteitä (Karataş ym. 2025; Taskiran & Aslantürk 2025; Gökçearsan ym. 2025; Wilkins & Benett 2025). Arviointitehtävät ulottuvat myös sosiaalityöhön liittyvien eettisten kysymysten pohdintaan, joita tarkastellaan kahdessa erillisessä tutkimuksessa (Segal 2025a; 2025b). Tekoälyä pyydetään ottamaan kantaa konkreettisiin sosiaalityön kenttään liittyviin tilanteisiin, joita leimaavat eettiset ristiriidat tai joissa sosiaalityöntekijän käsillä on erityisen haastava eettinen päätöksentekotilanne. Tekoälyn eettistä arviointikykyä tarkastelen luvussa 7.2.

Keskeisiä tutkimuksia laajojen kielimallien hyödyntämisestä sanallisissa arvioinneissa löytyy erityisesti turkkilaisesta sosiaalityön toimintaympäristöstä. Karataş kumppaneineen (2025) tarkastelee tekoälyn käytön luotettavuutta lasten seksuaalista hyväksikäyttöä koskevien lausuntojen arvioinnissa suhteessa ammattilaskollegan arviointiin. Taskiran ja Aslantürk (2025)

vertailevat kolmen yleisimmän saatavilla olevan kielimallin – ChatGPT-4:n, Geminin ja Microsoft Copilotin – suoriutumista asiakkaan tilannekuvan arvioinnissa ja asiakaskohtaisten ohjeiden laatimisessa. Tekoälyn arviointitehtävää varten laaditaan viiteen eri vaikeustasoon asettuvaa kuvitteellista, mutta realistista sosiaalityön asiakastapausta, jotka kattavat laajan kirjon eritasoisia sosiaalisia tukitarpeita (mt., 3).

Turkissa selvitetään myös laajojen kielimallien soveltuvuutta lastensuojeluun liittyvien raporttien arvioinnissa: ChatGPT valjastetaan seitsemän lastensuojelun anonymisoidun selvitysraportin (*”social investigation reports in child welfare”*) laadulliseen ja sisällölliseen arviointiin. Lisäksi tekoälyä pyydetään arvioimaan asiakkaisiin liittyviä riskejä sekä laatimaan asiakaskohtaiset ja paikalliseen palveluverkoston soveltuvat toimenpidesuunnitelmat jatkotyöskentelyn tueksi (Gökçearslan ym. 2025, 610-611). Edellisistä tutkimuksista poiketen tekoälylle annetaan taustamateriaaliksi myös lastensuojelun erityisalaan liittyvää paikallista lainsäädäntöä.

Kun tekoälyn kykyä tuottaa sanallisia arvioita verrataan ihmiskollegan suoriutumiseen, tutkimukset osoittavat, että ihmisasiantuntijat suoriutuvat arviointitehtävästä kokonaisuutena paremmin ja heidän harkintansa tuottaa useammin luotettavampia arvioita. Lisäksi havaitaan, että ihmisasiantuntijan tekemät arviot ovat useimmiten tekoälyn tuottamia arvioita täsmällisempiä: ne huomioivat paremmin kontekstuaaliset tekijät ja mahdollistavat hienovaraisen ja tilannekohtaisen tulkinnan (Karataş 2025; Taskiran & Aslantürk 2025; Gökçearslan ym. 2025). Edellä mainittujen puutteiden lisäksi tekoälyn vastauksissa havaitaan myös epätarkkuuksia siitäkin huolimatta, että taustamateriaaliksi annetaan erityisalaan liittyvää lainsäädäntöä (Gökçearslan ym. 2025, 619-620).

Tutkimukset vahvistavat osittain aiempia huolia siitä, että tekoälyn arviointilogiikka poikkeaa edelleen olennaisesti asiantuntijoiden päättelystä. Arviot jäävät usein liian yleiselle tasolle, eikä tekoäly tavoita kaikkia niitä keskeisiä näkökulmia, joita monimutkaisten asiakastilanteiden ymmärtäminen edellyttää. Lisäksi kielimallit antavat toisinaan ristiriitaisia ohjeita tai tuottavat samaan kysymykseen toisistaan poikkeavia vastauksia eri aikaan esitettyinä. Toisaalta ne ovat kyvykkäitä suoriutumaan joistakin tehtävistä - kuten esimerkiksi kielellisten rakenteiden tarkastelusta - varsin hyvin (Karataş 2025, 11). Laajat kielimallit kykenevät tuottamaan ihmisajattelua jäljitteleviä vastauksia, mutta ne eivät sellaisenaan ole täysin luotettavia tarjoamaan tilannekohtaisia toimintaohjeita tai tuottamaan tarkempaa palveluihin ohjausta (Taskiran & Aslantürk 2025, 9-10).

Wilkins ja Benett päätyvät (2025) hieman yllättäen päinvastaiseen lopputulokseen tutkiessaan tekoälyn suoriutumista lastensuojeluilmoitusten arvioinnissa suhteessa sosiaalityöntekijöiden suorittamaan arviointiin. ChatGPT:tä (v3.5) pyydetään analysoimaan 12 anonymisoitua lastensuojeluilmoitusta ja arvioimaan niiden perusteella myöhempien toimenpiteiden, tapahtumien ja lopputulosten todennäköisyyttä. Vastauksia verrataan 581 sosiaalityöntekijän käsittelemiin ilmoituksiin. Tulokset osoittavat, että tekoäly suoriutui arvioinnista hieman paremmin kuin keskiverto sosiaalityöntekijä. Wilkins ja Benett esittävät (2025, 7), että ammattilaisten objektiivisena pidettyyn päätöksentekoon saattaa liittyä yksilötasolla tapahtuvia, tiedostamattomia vääristymiä, kun taas tekoäly käsittelee kaikki kysymykset samalla menettelyllä ja arviointiperiaatteilla.

Muut tutkimuksissa esitetyt positiiviset havainnot laajojen kielimallien kyvykkyydestä liittyvät tekoälyn kykyyn nostaa esiin monipuolisesti erilaisia näkökulmia, mikä voi edesauttaa arjen työssä vähemmälle huomiolle jäävien asiakkaiden ja asiakasryhmien huomioimista. Laajojen kielimallien kyvykkyys tuottaa uusia näkökulmia asiakastyöhön voi tukea erityisesti sellaisia sosiaalityön asiantuntijoita, joilla on vain vähän käytännön kokemusta sosiaalityön alueesta. Tekoäly suoriutui hyvin muun muassa riskitekijöiden ja suojaavien tekijöiden jäsentämisestä ja tulkitsi niiden vaikutuksia kattavasti, mikä voi lisätä arviointien läpinäkyvyyttä ja kokonaisvaltaisuutta. (Gökçearslan ym. 2025, 615, 620.)

Tutkimusten perusteella laajoihin kielimalleihin perustuvia tekoälyratkaisuita voitaisiin hyödyntää sosiaalityön tehtävissä erityisesti sosiaalityöntekijän mentorina tai asiantuntijakollegana. Laajojen kielimallien tuottamaan tietoon kannattaa kuitenkin suhtautua kriittisesti. Osassa tutkimuksia varoitetaan, etteivät laajojen kielimallien kyvykkyydet kaikilta osin vielä riitä ihmismielen tasoiseen ajatteluun. On kuitenkin myös huomioitava teknologian nopea edistyminen ja päättelykyvyn huima kehitys. Vaikka tekoäly ei sovellu itsenäiseksi arvioijaksi, se voisi silti tukea asiantuntijoita työprosessissa eräänlaisena ”kognitiivisena apulaisena” (Karataş 2025, 12). Tekoäly tarjoaa myös uusia näkökulmia ja tukee sosiaalityöntekijää tilanteissa, joissa on tarpeellista kuulla toinen kollegiaalinen arvio. Viime kädessä vastuu tiedon oikeellisuudesta on ammattilaisella.

## **5.2 Koneoppiminen sosiaalisten ilmiöiden selittämisessä ja ennustamisessa**

Koneoppimiseen perustuvat menetelmät esiintyvät tämän tutkimuksen lähdeaineistossa yhtä yleisesti kuin laajoihin kielimalleihin perustuvat järjestelmät. Kuten aiemmin todettiin,

koneoppiminen on yksi tekoälyn alalaji. Hakusanalla ”tekoäly” koneoppimiseen liitettyjä artikkeleita löytyi kuitenkin vain muutama, mikä viittaa jälleen tekoäly-termiin liittyviin epätarkkuuksiin. Vaikka algoritmiset päätöksentekoprosessit ovat olleet jo pitkään nykypäivää, mahdollistavat uudet tekoälypohjaiset ratkaisut paljon laajempien tietomassojen käsittelyn ja analysoinnin. Koneoppimisen mahdollisuuksia suurten datasettien analyysissä on hiljattain alettu hyödyntää sosiaalitieteissä. Koneoppimiseen kohdistuu suuria odotuksia, jotka perustuvat sen mahdollisuuksiin louhia tietoa (”*data mining*”) suurista tietomassoista ja kykyyn tunnistaa datasta säännönmukaisuuksia ja malleja (mm. Bogenschutz ym. 2025 ja Hao ym. 2022).

Ero perinteisten ja tekoälyn perustuvien algoritmisten tai koneoppimis pohjaisten menetelmien välillä on suuri. ”Perinteinen” algoritmeihin perustuva järjestelmä tuottaa suositellun tai määrätyn toimintatavan ennakkoon tunnettujen muuttujien perusteella, kun taas monimutkaiset tekoälyjärjestelmät, jotka koulutetaan suurilla tietoaaineistoilla, tekevät useimmiten todennäköisyyteen pohjautuvia päätelmiä useiden muuttujien perusteella (James ym. 2024, 1553). ”Koneoppimisen menetelmissä analyysia ohjaa aineisto - eivät tutkijan ennako-oletukset - ja juuri data vie kohti löydöksiä” (Bogenschutz ym. 2025, 121).

Seitsemän yhdeksästä koneoppimiseen keskittyvästä tutkimusartikkelista käsittelee sosiaalisten riskien ennustamista. Riskiarviointia voidaan kohdistaa esimerkiksi masennusriskin havainnointiin (Hao ym. 2022), rikoksen uusimisen arviointiin (Portela ym. 2025) ja psykotrooppisten lääkkeiden käytön todennäköisyyden ennustamiseen (Bogenschutz ym. 2025). Lisäksi ennakoivia malleja on hyödynnetty ikäihmisten terveydentilan muutosten havaitsemisessa ja seurannassa (Fang-Yu ym. 2020).

Eryyisesti Tanskassa on toteutettu tutkimuksia ennakoivien riskimallien hyödyntämiseksi lastensuojelun toimintaympäristössä (Meilvang & Dahler 2024; Villumsen ym. 2024). Suomalaisessa sosiaalityön kontekstissa toteutettua tekoälytutkimusta on tehty hyvin vähän. Lehtiniemi (2024) tarkastelee vuonna 2021 lastensuojelussa pilotoitua ennakoivaa riskimallia ja pureutuu tekoälyn rooliin tulevaisuustiedon tuottajana. Lastensuojelutyötä tehdään epävarmuudessa sekä kovan paineen alla. Lastensuojelun konteksti on päätöksenteon näkökulmasta monimutkaista ja siihen voivat vaikuttaa ammattilaisen tiedostamattomat uskomukset tai käytösmallit (Leslie 2020, 14). Kehitystyön tavoitteita ovat haavoittuvassa asemassa olevien lasten ja perheiden tilanteen arvioinnin tasalaatuisuuden parantaminen ja työntekijäkohtaisen subjektiivisen arvion minimointi. Tanskassa tekoälykehitykselle ovat luoneet paineita julkisuuteen nousseet tapahtumat, joissa työntekijät ovat kahdessa erillisessä

tapauksessa tehneet virheellisen arvion lapsen tilanteesta, ja kaltoinkohtelu on jatkunut useita vuosia (Meilvang & Dahler 2024, 35). Vaikka yllä kuvatut tekoälyprojektit ovat vielä alkuvaiheessa ja ne on toteutettu rajatussa ympäristössä, alustavia tuloksia on mahdollista tarkastella.

Tekoälystä odotetaan apua muun muassa lastensuojeluilmoitusten käsittelyyn, jolloin tekoäly luokittelisi saapuvat lastensuojeluilmoitukset kiireellisiin ja ei-kiireellisiin (Meilvang & Dahler 2024, 33). Tämän lisäksi uusista teknologioista odotetaan helpotusta asiakastilanteiden nopeampaan arviointiin. Tilannearviointin ja interventioiden tueksi kehitetyt riskien ennakointimallit tuottavat yksilökohtaisen riskiennusteen sekä joissain tapauksissa myös arvion tarvittavista tukitoimenpiteistä. Tyypillisesti ennakointimalli voisi toimia siten, että viranomaisen vastaanottaa ilmoituksen mahdollisesta tuen tarpeesta, jonka jälkeen malli tuottaa asiakaskohtaisen numeerisen riskiarvion hyödyntäen lapsesta ja hänen perheestään kerättyä asiakas- ja potilasrekisterin historiatietoa. Tällaisia riskin ennustamisen kohteita ovat esimerkiksi lapsen kaltoinkohtelu, kiireellinen sijoitus tai huostaanotto sekä lastensuojelun tukitoimien tarpeen ennakointi (Villumsen ym. 2024; Meilvang & Dahler 2024; Lehtiniemi 2024, 1115). Väkivaltarikosten uusimista puolestaan tarkastellaan usean sosiaalisen ja kliinisen taustamuuttujan avulla, joita ovat muun muassa rikos- ja vankeushistoria, elämäkerralliset tekijät, perhe- ja sosiaaliset tekijät, terveydelliset tekijät sekä asenne- ja persoonallisuustekijät (Portela ym. 2025, 479), joiden pohjalta algoritmeihin perustuva malli tuottaa rikoksen uusimista ennustavan riskiarvion.

Tutkimustulokset puoltavat uusien koneoppimisen menetelmien omaksumista perinteisten tilastollisten menetelmien rinnalle (Hao ym. 2022, 286–287). Eryteisesti tiedonlouhinta ja koneoppiminen laajentavat mahdollisuuksia rakentaa ja testata sosiaalisiin ilmiöihin kytkeytyviä ennustemalleja sekä tarkastella ilmiöiden taustalla vaikuttavia tekijöitä aiempaa monipuolisemmin. Hao ja kumppanit (2022) havainnollistavat tätä kehitystä tuottamalla tiedonlouhintaan perustuvan mallin, jolla ennustetaan neurodegeneratiivisia sairauksia sairastavien ikääntyneiden masennusriskiä. Mallin taustamuuttujiksi koottiin laaja kokonaisuus, johon sisältyivät kognitiivinen toimintakyky, masennusoireet ja demografiset tiedot, mutta myös sairauden vakavuus, sosiaalinen tuki, itsearvioitu terveydentila sekä päivittäinen aktiivisuus (mt., 278–279). Esimerkki osoittaa, että koneoppiminen mahdollistaa useiden samanaikaisten tekijöiden hyödyntämisen ennusteissa tavalla, joka voi olla perinteisille menetelmille raskaampi toteuttaa ja tulkita.

Samaa metodologista potentiaalia kuvastaa Yhdysvalloissa kehitysvammaisten henkilöiden mielenterveyttä tukevan lääkehoidon käyttöä tarkasteleva tutkimus, jossa kehitysvammaisten henkilöiden mielenterveysongelmia hoidetaan lääkkeillä suhteellisesti useammin kuin muussa väestössä. Bogenschutz ja kumppanit (2025) vertaavat kolmea erilaista koneoppimismallia selvittääkseen, mitkä tekijät selittävät mielenterveyttä tukevan lääkehoidon käyttöä kehitysvammaisilla henkilöillä. Analyysissa yhdistetään yksilötasoisista kyselyistä johdettuja terveydellisiä ja sosiaalisia muuttujia useista lähdeaineistoista koottuun kokonaisuuteen. Parhaiten suoriutunut malli saavuttaa 90 prosentin ennustetarkkuuden, mikä tukee näkemystä koneoppimisen käyttökelpoisuudesta sosiaalisten ilmiöiden tutkimisessa ja selittämisessä. Selitysovoimaisimmiksi taustatekijöiksi nousevat diagnosoitu ahdistuneisuushäiriö, mielialahäiriödiagnoosi sekä diagnosoitu psykoottinen häiriö (Mt., 131). Yhdessä nämä esimerkit havainnollistavat, että koneoppimisen vahvuudet eivät rajoitu vain ennustetarkkuuteen, vaan se tarjoaa myös välineitä tunnistaa ja priorisoida ilmiöihin kytkeytyviä tekijöitä – kuitenkin niin, että tulkinta ja merkityksellistäminen edellyttävät edelleen kontekstisidonnaista analyysia.

Myös Fang-Yu kumppaneineen (2020) arvostelevat perinteisiä analyysimenetelmiä riittämättömiksi sosiaalisten ilmiöiden tutkimisessa. Selvittäessään ikäihmisten terveydentilan muutoksiin vaikuttavia tekijöitä he osoittavat, että koneoppimismenetelmät pystyvät perinteisiä menetelmiä tarkemmin mallintamaan muuttujien välisiä epälineaarisia (*non-linear*) suhteita. Epälineaarisen muuttujan suhde tutkittavaan ilmiöön voi olla piilevä tai niin monimutkainen, ettei sitä ole mahdollista laskea perinteisin tilastollisin menetelmin. Koneoppimismenetelmin tuotettuja riskimalleja voidaan hyödyntää esimerkiksi resurssien kohdentamisen suunnittelussa ja niiden avulla voi löytyä aiemmin huomiotta jääneitä, piileviä tekijöitä. (Mt., 2, 5-6.).

Ennakoivat riskimallit eivät siis rajoitu ainoastaan yksilötason ennustamiseen, vaan mallit voivat tuottaa tietoa ja ymmärrystä rakenteellisiin ja yhteiskunnallisiin kysymyksiin, jolloin huomion kohteena ovat väestötason muuttajat. Koneoppimisen menetelmät voivat myös täydentää perinteisiä tilastollisia menetelmiä tutkimuksessa ja palvelutarpeiden jäsentämisessä. Kim kumppaneineen (2025) tarkastelevat paikkatieto- ja koneoppimistekniikoiden yhdistämistä monitasoisen analyysin välineinä. Tutkimuksessa selvitetään 13 Yhdysvaltain osavaltion koulupiiritason aineistolla sosioekonomisen aseman, mielenterveys- ja päihdepalvelujen saatavuuden sekä koulupudokkuusasteiden välisiä yhteyksiä. Koulupudokkuus osoittautuu monitasoiseksi ilmiöksi, jossa taustatekijät ulottuvat yksilö-, perhe-, koulu- ja yhteisötasolle (Mt., 2–3). Koneoppimista hyödynnetään erityisesti puuttuvien koulupudokkuustietojen täydentämisessä, minkä jälkeen yhteyksiä muuttujien välillä analysoidaan tilastollisilla malleilla.

### 5.3 Tiedonhaku ja keskustelu tekoälyn kanssa

Mielenkiintoisia ja lupaavia tekoälyratkaisujen sovellusalueita sosiaalityössä näyttäisivät olevan tiedonhaku sekä keskustelemaan tekoälyyn perustuvat ratkaisut. Perron ja kumppanit (2024; 2025) korostavat laajojen kielimallien potentiaalia suurten tietomassojen läpikäynnissä ja asiakastiedon jäsentämisessä: mallit voivat tukea olennaisen tiedon poimintaa, luokittelua ja tiivistämistä sekä auttaa rakentamaan kokonaiskuvaa tapauksesta silloin, kun aineisto on laaja ja pirstaleinen. Nämä odotukset liittyvät suoraan sosiaalityön tiedontuotantoa koskevaan keskusteluun narratiivisen tiedon haasteista. Asiakasdata koostuu usein pitkistä sanallisista kirjauksista, joita on kertynyt useiden vuosien ajalta ja useista kohtaamisista. Tällöin keskeinen riski on, että olennainen tieto – esimerkiksi riskitekijät, aiemmat interventiot, käännekohtat ja palvelutarpeen muutokset – jää hajanaisen ja laajan tekstimassan alle, mikä vaikeuttaa sekä ajantasaisen tilannearvion tekemistä että päätöksenteon perustelujen läpinäkyvää dokumentointia.

Laajat kielimallit osoittautuvat tehokkaaksi avuksi strukturoimattoman asiakastiedon jäsentämisessä ja luokittelussa. Perron ja kumppanit (2024) arvioivat tutkimuksessaan, missä määrin erilaisia laajoja kielimalleja voidaan hyödyntää päihteisiin liittyvien ongelmien luokittelussa ja päihteisiin liittyvän olennaisen tiedon poiminnassa lastensuojelun tutkintayhteenvedoista. Tutkijat kiinnittävät huomiota erityisesti siihen, että tekoäly kykenee järjestelmällisesti tunnistamaan ja poimimaan olennaista sisältöä aineistosta lähes ihmiskoodarin tasoisesti, mutta työskentelynopeus osoittautuu tekoälyllä selvästi ihmistä paremmaksi (mt., 702). Tiedonhakuun liittyvät käytötapaukset voivat vaihdella tutkimuksesta palveluiden suunnitteluun ja kohdentamiseen sekä asiakastyöhön. Tekoäly tarjoaa kustannustehokkaan tavan strukturoimattoman tiedon käsittelyyn erilaisissa sosiaalityön toimintaympäristöissä silloin, kun suuria tietomääriä on hidasta käydä läpi manuaalisesti.

Laajojen kielimallien ohella tarkastellaan myös koneoppimismenetelmien mahdollisuuksia tukea käytännön tiedonhallintaa. Bryan ja kumppanit (2021, 635) selvittävät, voidaanko lastensuojelun asiakkuudessa olevien perheiden strukturoimattomasta tekstiaineistosta tunnistaa perheväkivaltaan kytkeytyviä ajankohtaisia ongelmia tekstinlouhinnan ja ohjatun koneoppimisen menetelmillä. Tekoälyä varten luodaan sääntöpohjainen vertailumalli, jossa perheväkivaltaan liittyvät teemat, riskit tai tapahtumat operationalisoidaan muuttujiksi. Käytännössä ne ovat teksteissä ilmeneviä perheväkivaltaan liittyviä sanoja tai ilmauksia (Mt., 637-639).

Tulokset puoltavat sekä laajojen kielimallien että koneoppimismenetelmien käyttökelpoisuutta merkityksellisen tiedon poimimiseen tekstiaineistoista. Koneoppimismenetelmien ja laajojen kielimallien kustannustehokkuutta tai luotettavuutta tiedon jäsennyksen tukena ei ole lähdeaineiston perusteella mahdollista vertailla keskenään tässä tutkimuksessa. Laajojen kielimallien ja muiden koneoppimiseen perustuvien menetelmien mahdollisuudet vahvistaa hallinnollista tehokkuutta tulee ottaa vakavasti, koska suurin osa sosiaaliryöy organisaatioiden datasta on jäsentymätöntä (Perron ym. 2024, 704; Bryan ym. 2021, 650).

Huomionarvoista on kuitenkin se, että siinä missä laajojen kielimallien avulla toteutettu tiedonhaku ei käsitellyissä tutkimuksissa edellytä laajaa taustavalmistelua, on koneoppimismenetelmin toteutetun tutkimuksen koulutusdatan manuaalinen valmistelu erittäin työlästä. Manuaalisen operationalisoinnin lisäksi taustavalmisteluihin sisältyy koulutusdatana hyödynnetyn noin 1400 dokumentin manuaalinen luokittelu. Jokaisen dokumentin luokitteluun kuluu noin 6 minuuttia, jolloin yhteenlaskettu työmäärä nousee jopa 140 työtuntiin (Bryan ym. 2021, 648). Vasta tämän jälkeen tekoälymallin koulutus on mahdollista aloittaa.

Laajojen tietomassojen läpikäynnin lisäksi Perron ja kumppanit näkevät (2025) tekoälyn potentiaalin asiakaspalveluun liittyvissä tehtävissä. Tekoälypohjaisia asiakaspalvelubotteja on jo laajasti käytössä erityisesti yksityisellä sektorilla ja ne ovat yleistymässä myös julkisissa palveluissa. Muutamat suomalaisetkin hyvinvointialueet ovat jo ottaneet tekoälypohjaisia chatbot-ratkaisuja käyttöönsä asiakasohjauksen ja neuvonnan palveluihin. Tällaisista uusista palveluista ovat uutisoineet muun muassa Vantaan ja Keravan hyvinvointialue (2025) sekä Päijät-Hämeen hyvinvointialue (2025). Tekoälypohjaiset chatbotit perustuvat laajoihin kielimalleihin. Ne pystyvät keskustelemaan asiakkaan kanssa, hakemaan tietoa laajoista tietolähteistä ja tiivistämään tiedon kysyjälle joko toimintaohjeiksi tai yleisen neuvonnan tueksi. Laajoihin kielimalleihin nojaavat chatbotit kykenevät myös palvelemaan usealla eri kielellä, jolloin niiden hyödynnettävyys erilaisiin asiakaspalvelutehtäviin kasvaa.

Kokemuksia on saatu myös keskustelevan chatbotin soveltuvuudesta terapiatyöhön (Chan ym. 2025). Tutkimuksessa tekoälylle laadittiin vaiheittainen ohjeistus, jonka mukaisesti se ohjasi asiakkaan strukturoitua keskustelurakennetta pitkin ja läpi narratiivisen terapian eri vaiheiden. Tekoälypohjaista chatbotia hyödynnetään terapia- ja keskustelutukeen myös Suomessa MIELI ry:n tuottamassa Sekasin-chat-palvelussa (2025), jossa nuori voi halutessaan keskustella tekoälypurin kanssa anonymisti erilaisista mielenterveyteen ja arjen haasteisiin liittyvistä

asioista. Tulosten mukaan chatbot kykenee ohjaamaan terapiakeskustelua suunnitellusti ja mukautumaan käyttäjän reaktioihin keskusteluiden edetessä: keskustelujen eteneminen vastaa hyvin terapian vaiheistusta, ja käyttäjä kokee sekä keskustelun, että tekoälyn tuottaman yhteenvedon omasta tilanteestaan osuvaksi ja hyödylliseksi (Chan ym. 2025, 180).

Tutkimuksessa myös havaitaan, että terapiabotin kehitystyö vaatii paljon teknistä osaamista, kehitystyötä ja ohjeistuksia, jotta se saadaan toimimaan toivotulla tavalla (mt., 172-173). Tämä havainto vastaa myös omaa kokemustani ohjauksen ja neuvonnan botin kehittämisprojekteista. Mitä yksityiskohtaisempia ohjeita botille tuotetaan, sitä paremmin ja tarkemmin se pystyy palvelemaan sille annetuissa tehtävissä. Siitäkin huolimatta, että tekoälypohjainen botti toimii itsenäisesti, on sen jatkokehittämiseen ja monitorointiin panostettava, sillä asiakasrajapinnassa tapahtuva ohjaus ja neuvonta edellyttävät luotettavuutta ja korkeaa laatua. Ohjeita ja tekoälyn hyödyntämiä tietolähteitä tulee päivittää, jotta tekoälyn käyttämä tieto pysyy ajan tasalla.

## **6 Tekoälypohjainen tiedonmuodostusprosessi: hyödyt ja rajoitteet**

### **6.1 Tekoäly vakuuttaa ja harhauttaa**

Tekoälyn kyvykkyydet jäljitellä inhimillistä ajattelua ovat jo vakuuttaneet suuren yleisön. Sen kielelliset lahjakkuudet ja ilmaisukyky vastaavat ihmisten kielellisiä taitoja. Samaan aikaan kun laajojen kielimallien hyödyntäminen laajenee sosiaalityön tehtäviin, koulutukseen tai tutkimukseen, tulee sosiaalityön tiedeyhteisön olla valmis vastaamaan kriittisiin kysymyksiin, jotka liittyvät asiakkaiden hyvinvointiin, misinformaatioon, uskottavuuteen, eettisiin käytäntöihin sekä lähteettömän tiedon aiheuttamiin vinoumiin ja haittoihin (Patton ym. 2023, 553).

Kriittinen ajattelu ohjaa meitä tarkastelemaan tekoälyn toimintalogiikkaa. Laajat kielimallit, kuten ChatGPT, rakentuvat laajoihin verkkolähteistä kerättyihin aineistoihin, joita ei ole mahdollista yksiselitteisesti varmentaa (Patton ym. 2023, 557). Mallien koulutusdata painottuu usein sisältöihin, jotka ovat verkossa laajimmin saatavilla, näkyvimmin esillä ja eniten käytettyjä, joten niiden tuottamat vastaukset voivat heijastaa olemassa olevia tiedon tuotannon ja näkyvyyden valtarakenteita. Laajoja kielimalleja on kritisoitu länsimaisen kulttuurin ylivoimasta, koska useimmat ensimmäiset kielimallit (ChatGPT, Gemini, Claude, Llama) on kehitetty yhdysvaltalaisen teknologiyhtiöiden johdolla ja siten ne on koulutettu länsimaisia arvoja ja maailmankuvaa heijastelevalla tiedolla (mt., 559). Sitten rinnalle on noussut myös kiinalaisia ja myöhemmin eurooppalaisia toimijoita. Kritiikkiä malleille esitetään muun muassa kehittyvissä

maissa. Omorogiuwan ja kumppaneiden mukaan (2025, 101) ”generatiivinen tekoäly heijastaa usein koulutusdataansa sisältyviä kolonialistisia, rodullistavia ja sukupuoleen liittyviä vinoumia”.

Pahimmassa tapauksessa lähdeaineistojen vinoumat eivät rajoitu yksittäisiin asiavirheisiin, vaan voivat vaikuttaa siihen, millaiset näkökulmat ylipäättään normalisoituvat ”yleistiedoksi” järjestelmän tuottamissa vastauksissa. Käytännössä tämä voi tarkoittaa sitä, että laajoihin kielimalleihin pohjautuvat järjestelmät tuottavat suhteettoman usein länsimaista tulkintaa ”totuutena”, samalla kun globaalin etelän tutkijoiden ja asiantuntijoiden tuottama tieto jää vähäisemmälle huomiolle. Sosiaalityön näkökulmasta tämä on olennainen eettinen kysymys, koska ammattialan ydintehtävään kuuluu haavoittuvassa asemassa olevien ryhmien äänen vahvistaminen sekä rakenteellisen eriarvoisuuden tunnistaminen ja purkaminen. Jos tekoälyjärjestelmät toistavat tiedonvalinnan vinoumia, ne voivat huomaamatta vahvistaa juuri niitä valtasuhteita ja näkymättömyyden mekanismeja, joita sosiaalityössä pyritään kriittisesti tarkastelemaan ja uudistamaan. Tämän vuoksi generatiivisen tekoälyn käyttö sosiaalityön konteksteissa edellyttää tietoista ja kriittistä otetta, jotta ihmisoikeuksien toteutuminen ja sosiaalisen oikeudenmukaisuuden periaatteita voidaan vahvistaa sosiaalityön toimialueella (Patton ym. 2023, 561).

Tiedon valtarakenteiden epävarmuuden lisäksi tekoälyratkaisujen hyödyntämistä varjostaa hallusinoinnin riski. Hallusinaatio tarkoittaa tekoälyratkaisun tuottamaa virheellistä tietoa. Tekoälyn tuottama tieto saattaa itsessään kuulostaa vakuuttavalta ja kielellisesti oikeanlaiselta, mutta sisältö voi koostua asiavirheistä. Maallikkokäyttäjän on usein haastavaa päästä kiinni hallusinoinnin syihin tiedonmuodostusprosessin taustalla. On haastavaa, ellei jopa mahdotonta selvittää, johtuuko virheellinen tieto algoritmin toiminnasta vai virheellisestä tietolähteestä.

Perron ynnä muut (2025) sekä Gökçearsan kumppaneineen (2025, 621) tunnistavat hallusinoinnin riskin sosiaalityön tekoälyratkaisuissa. Siinä missä yleiset laajat kielimallit yksistään käytettynä saattavat generoida epäjohtonmukaista tai epätasaista laatua tuotoksissaan, voidaan hallusinaation riskiä pienentää sosiaalityön sensitiivisellä kentällä siten, että tekoälyratkaisulle tarjotaan valmiita tietolähteitä, joista se hakee tarvittavat aihealuekohtaiset tiedot. Tällainen tekoälyratkaisu hyödyntää laajoihin kielimalleihin perustuvaa teknologiaa, jolloin sillä säilyy kyvykkyys ymmärtää ja tuottaa kieltä. Tekoäly kuitenkin rajataan hyödyntämään ainoastaan tai pääsääntöisesti ennalta määritellyjä tietolähteitä. Tällöin pystytään parhaiten varmistumaan siitä, että tekoälyn tuottama tieto on täsmällistä.

Ylläkuvattua toteutustapaa kutsutaan lyhenteellä RAG (Retrieval-augmented generation). Perron kumppaneineen (2025) havainnollistavat RAG:in hyötyjä esittämällä yleisesti saatavilla olevalle ChatGPT-versiolle sosiaalityötä koskevan kysymyksen. Vastaus voi vaikuttaa kielellisesti kaikin puolin sujuvalta, mutta tarkemmassa tarkastelussa käy ilmi, että tekoäly antaa väärän tai epätarkan tiedon. Tämän jälkeen he esittävät saman kysymyksen uudelleen, mutta liittävät mukaan täsmällisen tietolähteen ja rajaavat mallin hakemaan tiedon nimenomaan siitä. Tällä kertaa tekoäly tuottaa selvästi täsmällisemmän ja tarkemman vastauksen.

Esimerkki tukee ajatusta, että vaikka laajan kielimalli tuottama kielellinen tuotos vaikuttaa hyvältä, yleiskäyttöiset laajat kielimallit eivät pääsääntöisesti sellaisenaan vastaa sosiaalityön spesifiin tarpeeseen, vaan niiden tarkkuutta on parannettava tarjoamalla toimialalle tyypillisiä ja tarpeellisia lähdemateriaaleja. Mitä tarkemmin tietolähteitä määritellään ja mitä selkeämpiä toimintaohjeista sille tuotetaan, sitä paremmin se voi suoriutua myös sosiaalityön substanssikysymyksistä.

Myös koneoppimista hyödyntävissä menetelmissä kohdataan samankaltaisia haasteita, sillä suurien tietomassojen käsittelyssä ja analyysissä piilee sudenkuoppia. Algoritmisiin ratkaisuihin kytkeytyy myös riski tuottaa ja vahvistaa vinoumia. Vinoumat heikentävät tekoälymallin luotettavuutta. Vinoumat johtuvat usein mallin pohjana käytetystä datasta. Vanhentunut, puutteellinen tai epätarkka data, kiistanalaiset kirjaukset tai suoranaiset virheet lähtödatassa tuottavat epätasällisiä tai vääristyneitä ennusteita ja tuloksia (Leslie ym. 2020, 9). Meilvang ja Dahler muistuttavat (2024), että päätöksenteon tukialgoritmien läpinäkyvyyttä ja luotettavuutta voi heikentää se, että algoritmeissa käytetty taustatieto on usein sosiaalityöntekijöiden itsensä tuottamaa, eikä sitä kerätä tai tuoteta yhdenmukaisin tavoin. Vaikka datan laatu tunnistetaan heikoksi, suuri datamäärä saa uskomaan, että algoritmin tuottamat tulokset olisivat objektiivisempia kuin yksittäisen sosiaalityöntekijän tekemä arvio. (Mt., 38-39.)

Tekoälypohjaisten ratkaisujen perusedellytys on, että niiden käyttö perustuu riittävän kattavaan, yhdenmukaisesti tuotettuun ja laadultaan varmistettuun taustadataan – muuten mallit skaalaavat myös puutteet, virheet ja vinoumat. Díezin ja kumppaneiden (2021) pääajatuksena on se, että heikkolaatuinen data voi johtaa väärin johtopäätöksiin ja päätöksiin sekä tuottaa vinoumia ja syrjintää, minkä vuoksi datan tulee olla validia ja luotettavaa, jotta se myös tuottaa luotettavia lopputuloksia (Díez ym. 2021, 80). Samassa tutkimuksessa kuvataan sosiaalista syrjäytymistä mittaavan SiSo-työkalun kehittämistä. Tutkimus havainnollistaa, miten datan ja mittaamisen laatua voidaan myös testata ja kehittää ennen laajempaa hyödyntämistä. Tekoälyn

hyöty sosiaalityön ja sosiaalitieteiden tiedontuotannossa ei siis synny pelkästään algoritmeista, vaan ennen kaikkea siitä, miten laadukas ja tarkoituksenmukaisesti operationalisoitu taustadata on, ja miten sen laatua varmennetaan ennen todellista tuotantokäyttöä.

## **6.2 Tekoälypohjainen tiedonmuodostusprosessi edellyttää läpinäkyvyyttä**

Lastensuojelu on laajasti edustettuna tekoälyratkaisuiden käyttökohteena. Kahdeksan eli lähes puolet kaikista aineistona hyödynnetyistä tutkimusartikkeleista käsittelee lastensuojelun toimintaympäristöä, kattaen hyvin erilaisia tehtäviä lastensuojeluilmoitusten vastaanottamisesta ja arvioinnista aina toimenpiteiden suunnitteluun ja huostaanottoriskin ennustamiseen. Lastensuojelu on erittäin sensitiivinen toimintakenttä, ja lastensuojelulliset toimenpiteet vaikuttavat heikommassa asemassa oleviin lapsiin ja perheisiin. Olipa kyseessä sitten ammattilaisen tiedonmuodostusprosessi tai tekoälyn tuottama tieto, viranomaistoimintaa koskeva vaatimusta toimenpiteiden ja päätösten perustelusta tulee noudattaa kaikissa sosiaalihuollon palveluissa.

Lastensuojelussa ennakoivat tekoälyratkaisut voivat esimerkiksi tunnistaa lapsen turvallisuutta uhkaavia tekijöitä ja arvioida, kuinka todennäköisiä lapsen kasvuun tai kasvuympäristöön liittyvät riskit ovat. Koneoppimismenetelmiin liittyy kuitenkin jo aiemmin kuvattu vinoumariski. Lastensuojelun toimintaympäristössä tämä voisi tarkoittaa, että tietyt taustatekijät – kuten köyhyys – alkavat mallissa ylikorostua lapsen kasvuympäristön vaarantumisen riskin indikaattoreina. Koneoppimismallien ennusteet alkavat vaikuttaa tulevaan tiedonkeruuseen, ja mallien tekemät suositukset tallentuvat osaksi tulevaa dataa, jota käytetään uusien mallien kouluttamiseen – mikä vahvistaa olemassa olevia vinoutumia edelleen. Pahimmillaan tällaiset mallit tuottavat ja toisintavat epätasa-arvoisia käytänteitä ja sosiaalisten ilmiöiden vääristynyttä tulkintaa. (Leslie ym. 2020, 9, 17.)

Laajojen kielimallien osalta puutteet näyttävät liittyvän tiedonmuodostusprosessin läpinäkyvyyteen. Yleiskäyttöisten tekoälymallien toimintalogiikkaa ei tutkimuksissa avata. Laajoja kielimalleja hyödynnetään matalalla kynnyksellä, mutta mahdollisuus tarkastella kriittisesti mallin tiedonmuodostusprosessia ei ole. Tutkimuksissa ei kuvata miten tekoälymallia on opetettu, mitä tietoa se hyödyntää tietolähteenään ja millä tavalla se päätyy annettuun vastaukseen. Toinen laajoihin kielimalleihin liitetty vakava huoli liittyy kysymyksiin yksityisyyden suojasta (mm. Patton 2023, 557; Perron 2024). Läpinäkyvyys ja ymmärrys järjestelmässä käytettävien tietojen käsittelystä tulee liittää osaksi tekoälyratkaisujen arviointia. Perronin

mukaan (mt., 2) ei ole koskaan täysin mahdollista varmistua siitä, että tiedot poistetaan kolmansien osapuolten taustajärjestelmistä tai siitä, ettei niitä hyödynnetä alkuperäisen käyttötarkoituksen vastaisesti. Tämä epätietoisuus asettaa hänen mukaansa rajoitteita hyödyntää laajoihin kielimalleihin pohjautuvia ratkaisuja täysimääräisesti.

Algoritmipohjaisia koneoppimismalleja koskevissa tutkimuksissa on yleisemmin jäsennetty, millä perusteilla malli tuottaa tuloksiaan. Samalla hyvin nopeasti käy kuitenkin ilmi, että mallien arviointi edellyttää vähintään perustason tilastollista ja menetelmällistä osaamista. Seitsemässä artikkelissa yhdeksästätoista tekoälymallin kehitystyötä kuvataan riittävän yksityiskohtaisesti, jotta niistä on mahdollista muodostaa jonkinlainen yleiskuva. Tutkimuksissa painotetaan kattavan testauksen ja mallin arvioinnin merkitystä kehitysprosessin aikana. Villumsen ja kumppanit (2024, 7) kehittävät ja vertailevat neljää erilaista koneoppimismallia riskiennusteiden tuottamiseen. Kaikki mallit perustuvat todennäköisyyksien laskentaan ja painottavat eri muuttujia ennusteen muodostamisessa. Tutkimus osoittaa, että koneoppimisen malleilla tuotetut riskiennusteet tuottavat erilaisia lopputuloksia riippuen siitä, minkälainen matemaattinen tekoälymalli otetaan käyttöön. Valitun mallin ennustetta arvioidaan lopuksi rinnakkain muiden kaltoinkohtelua indikoivien havaintojen sekä sosiaalityöntekijöiden tekemien ratkaisujen kanssa. Tutkimus tarjoaa mahdollisuuden arvioida tekoälymallin kehitysprosessia ja lopputulosta.

Niin ikään ikäihmisen masennusriskin ennustemallin kehitysprosessissa keskitytään laadukkaaseen taustavalmistautumiseen sekä datan laadun ja muuttujien arviointiin (Hao ym. 2022, 285-286). Erilaisten muuttujien painoarvoa tarkastellaan ja verrataan suhteessa masennusriskin ennustamiseen, mikä tarjoaa mahdollisuuden arvioida ja validoida mallin toimivuutta (mt., 280-281) ja siten tuottaa läpinäkyvyyttä mallin toimintaperiaatteisiin.

Tekoälyratkaisun käyttöönotto siis edellyttää, että mallin luotettavuutta voidaan arvioida riittäväällä tarkkuudella ja mallin toimintamekanismit ymmärretään riittäväällä tasolla. Villumsen kumppaneineen (2024, 118) painottaa, että päätöksentekoprosessit, mukaan lukien tekoälyn toimintaperiaatteet ja rajoitteet, on tehtävä ymmärrettäviksi - ei ainoastaan ammattilaisille, mutta myös asiakkaille, jotka ovat päätöksenteon kohteena. Lehtiniemi (2024, 10) havainnollistaa esimerkin, jossa asiakkaalla ei ollut riittävä ymmärrystä tekoälymallin käytöstä: kohonneen huostaanottoriskin käsittely mallin tuottaman arvion pohjalta ei tukenut asiakasta kyseisessä tilanteessa, vaan saattoi pikemminkin aiheuttaa tarpeetonta huolta ja mielipahaa.

Tutkimusartikkeleista nousee selvästi esiin, että tekoälyteknologian toimintaperiaatteiden ymmärtäminen tuottaa monille sosiaalityön kentällä toimivalle työntekijälle, tutkijalle tai opiskelijalle vaikeuksia. Sosiaalityön opinnoissa tunnistetaan olevan teorian ja teknologian välinen syvä kuilu. Tämä osaltaan heikentää ammattilaisten mahdollisuuksia tarkastella tekoälypohjaista tiedonmuodostusta kriittisesti.

## 7 Ammatillinen harkinta ja vastuu tekoälyn aikakaudella

### 7.1 Päätöksenteon delegoinnin rajat

Tekoälyn käyttöönottoa sosiaalityössä ajaa tarve kehittää (ainakin) kolmea asiakastyön osa-aluetta: käytäntöjen yhtenäistämistä, objektiivisuuteen pyrkimistä ja päätöksenteon vinoumien poistamista (Meilvang & Dahler 2024, 36). Näiden ohella taloudellisen tehokkuuden ja tuottavuuden tavoittelu toimivat tekoälykehityksen yhteiskunnallisen tason ajureina. Tarkastellessani lähdeaineistoa jäsensin kokonaisuutta pohtimalla, missä määrin tekoälyn avulla voidaan tavoitella tiedon ja päätöksenteon objektiivisuutta, kykeneekö tekoäly tuottamaan sellaista tietoa, jota voidaan sellaisenaan siirtää asiakastyöhön, sekä miten ammattilaiset hahmottavat oman harkintavaltansa suhteessa tekoälyn tuottamaan tietoon.

Meilvang ja Dahler tarkastelevat (2024, 31-32) algoritmisten päätöksentekoratkaisujen vaikutuksia sosiaalityön ammatillisen harkinnan rooliin ja merkitykseen. Tutkimuksessa selvitetään (mt., 31), miten ammattilaisten jäsentävät omaa ammatillista harkintavaltansa suhteessa päätöksentukea tarjoaviin algoritmeihin ja miten tekoäly muotoutuu osaksi ammatillisia käytäntöjä. Edellisessä luvussa on käsitelty tiedonmuodostuksen läpinäkyvyyden haasteita sekä sosiaalityössä ilmenevää osaamisvajetta tekoälypohjaisten menetelmien hyödyntämisessä. Jotta tekoälypohjaista tiedonmuodostusta voidaan käytännön työssä jäsentää suhteessa ammattilaisen harkintaan, menetelmien toimintalogiikan on oltava riittävän läpinäkyvä. Tällöin ammattilaisilla on mahdollisuus arvioida mallin tuottamia tuloksia ja tunnistaa menetelmiin liittyvät riskit, kuten vinoumat ja epätarkkuudet.

Tutkimuksissa on raportoitu myönteisiä tuloksia siitä, että algoritmisen tuki voi parantaa ammattilaisten päätöksenteon tarkkuutta sekä rikosten uusimisriskin arvioinnissa että vakavan kaltoinkohtelun riskien tunnistamisessa (Villumsen ym. 2024; Portela ym. 2025). Tutkimuksissa nähdään, että vuorovaikutus tutkitun arviointityökalun kanssa parantaa ammattilaisen tekemää ennustetarkkuutta sekä vaikuttaa myönteisesti ammattilaisten halukkuuteen tukeutua työkaluun arviointitehtävässä. Tutkimukset vahvistavat aiempia tuloksia, joiden mukaan päätöksenteon tarkkuus paranee hieman algoritmisen tuen avulla, osallistujat muuttavat omia ennusteitaan algoritmin ennusteiden suuntaan, ja riskien ilmaisemiseen käytetyt eri asteikot tuottavat erilaisia tarkkuustasoja. (Mt., 473.)

Tutkimuksissa suhtaudutaan tekoälypohjaisiin menetelmiin pääasiassa myönteisesti ja ennakkoluulottomasti. Lehtiniemi havaitsee (2024) työntekijöiden tiedostavan tekoälytyökalun

rajoitteet, mutta ennusteiden tarkkuus ei kuitenkaan noussut heidän arvioissaan keskeiseksi huolenaiheeksi. Työkalun tuottamaa tietoa pidetään hyödyllisenä ennen kaikkea siksi, että se näyttäytyy ulkopuolisena ja siten keskusteluun tuotavana "faktana". Samalla työntekijät asemoivat asiakastiedon ensisijaisiksi tuottajiksi ensisijaisesti itsensä: tekoäly nähdään korkeintaan prosessia tukevana välineenä, ei asiakkaasta tehtävän tiedon tai arvioiden korvaajana. (Mt., 1117.)

Ennakoivat riskimallit pureutuvat syvälle sosiaalityön päätöksenteon ytimeen. Siksi niiden kohdalla tulee käyttää erityisen tarkkaa eettistä harkintaa ja arvioida käyttöönoton hyötyjä ja riskejä perusteellisesti (Villumsen ym. 2024, 5). Algoritmisten järjestelmien tarkoituksena on lisätä menettelyllistä objektiivisuutta ja standardoida asiakastyötä, koska sosiaalityöntekijät käyttävät keskenään poikkeavaa ja subjektiivista harkintaa. Koska yksittäisten tapausten konteksti on kuitenkin sosiaalityössä tärkeä, ammatillinen harkinta ja tapauskohtainen tieto muodostavat välttämättömän tukirakenteen toimivalle algoritmille (Meilvang & Dahler 2024, 39).

Päätöksenteon täysimääräinen automatisointi koneoppimisen keinoin nähdään kaiken kaikkiaan ongelmallisena: tutkijoilla on vankka yhteinen näkemys siitä, ettei tiedontuotantoa ja päätöksentekoa voi ulkoistaa pois sosiaalityöntekijän ammatillisesta osaamisesta (mm. Lehtiniemi 2024, 11; Villumsen ym. 2024, 7). Myös lainsäätäjät ovat nähneet automaattisen päätöksenteon korkean riskin teknologiaksi (Portela ym. 2025, 472), johon tulee suhtautua äärimmäisellä varovaisuudella ja harkinnalla. EU:n tekoälyasetus velvoittaa organisaatioita tekemään tekoälyn riskien arviointi korkean riskin käyttötapauksissa. Tämän lisäksi tekoälyasetuksessa suoranaisiin kiellettyihin käyttötapauksiin kuuluu sosiaalisen pisteytys.

Joissain tutkimuksissa sen on osoitettu heikentävän ammattilaisen kontekstisidonnaista harkintaa ja estävän yksilöllisen ja tilannekohtaisen työskentelyn (James ym. 2024, 1558). Tekoälypohjaiset riskiennusteet tuottavat kertaluotoisen otoksen asiakkaan sen hetkisestä tilanteesta. On huomioitava, että oikeassa asiakastyössä tulevaisuuden ennakointi tai ennustaminen ei ole kertaluonteinen lopputuotos, vaan ennakointi tulee nähdä asiakkaan kanssa vuorovaikutuksessa tapahtuvana dynaamisena prosessina (Lehtiniemi 2024, 3-4, 7).

Ammattilaisen rooli tekoälyn tuottaman tiedon validoijana korostuu. Lupaavissa tukirooleissa tulkinnan, kontekstualisoinnin ja lopullisen päätöksenteon vastuun on pysyttävä selkeästi ihmisasiantuntijoilla (Karataş 2025, 12). Asiakastilanteet edellyttävät edelleen tilannekohtaista ja kontekstisidonnaista arviointia, eikä kaikkia päätöksentekoon vaikuttavia muuttujia pystytä

koneoppimisen menetelmistä huolimatta edelleenkin operationalisoimaan. Tekoälyn avulla ei ole esimerkiksi täysin mahdollista päätellä asiakkaan yksilöllisiä vahvuuksia, selviytymiskeinoja ja stressitekijöitä (Lehtiniemi 2024, 8-9). Edelleen ratkaisemattomaksi kysymykseksi jää intuition, hiljaisen tiedon, kokemustiedon tai organisatorisen tiedon rooli sosiaalityön päätöksenteossa ja algoritmisten ratkaisujen täydentäjänä. Tekoäly ei myöskään pysty ”lukemaan” rivien välistä.

Algoritmien tuki koetaan kuitenkin hyödyllisenä, koska se auttaa ammattilaista tarkistamaan ja arvioimaan omia johtopäätöksiään, tukee uusien työntekijöiden perehdytystä ja tuottaa yhdenmukaisia tuloksia samalla kun se vähentää päätöksenteon tunnepohjaisuutta. (Portela ym. 2025, 492). Tekoälypohjainen laajojen ja eri lähteistä koottujen aineistojen käsittely voi olla hyödyksi monimutkaisissa toimintaympäristöissä, joista sosiaalityöntekijöillä ei ole riittävästi kokemusta (James ym. 2024, 1557). Tekoäly voi toimia ammattilaisen tukityökaluna. Asiakastyössä tämä voisi esimerkiksi tarkoittaa tekoälyn tuottaman tiedon nostamista yhteiseen tarkasteluun asiakastapaamisessa ja sen tuottamaan tiedon havainnointia yhdessä asiakkaan kanssa. Ajatus on se, ettei tekoälypohjainen tieto itsessään oikeuta mahdollista interventiota (Lehtiniemi 2024, 1119).

Yleiskäyttöisiä laajoja kielimalleja hyödynnettäessä ei voida varmuudella sanoa, nojaako tekoälyn osaaminen ja päättely akateemiseen tutkimuskirjallisuuteen, erityisalakohotaisiin tietoihin vai pohjaako se pahimmassa tapauksessa osaamisensa yleiseen verkossa saatavilla olevaan sisältöön, jonka oikeellisuutta ei ole todennettu. Laajojen kielimallien uskottavalta tuntuvat vastaukset voivat saada sen kuulostamaan kaikkien alojen asiantuntijana, joka on kyvykäs tuottamaan ”absoluuttista totuutta”. Pelkkien lopputuotosten arvioinnin lisäksi tulisikin kasvattaa ymmärrystä yleiskäyttöisten laajojen kielimallien taustalla olevasta datasta, prosesseista tai algoritmeista. Laajoihin kielimalleihin pohjautuviin ratkaisuihin liittyy sama läpinäkyvyyden vaatimus kuin muihinkin koneoppimisen malleihin.

Mallien tuottamiin vastauksiin on suhtauduttava kriittisesti, mutta tietolähteiden rajauksen, datan laadun, laadukkaan operatoinalisoinnin, testauksen ja tarkan ohjeistuksen avulla voidaan mallien luotettavuutta kasvattaa ja siten toteuttaa käyttökelpoisia ammattilaisten tukityökaluja. Tekoäly voisi auttaa ”järjestämällä, käsittelemällä ja tekemällä näkyväksi asiakastietoja tai hoitohistoriaa, jotka muuten olisivat työläitä löytää” (Lehtiniemi 2024, 8). Mikään tekoälymalli ei kuitenkaan pysty poistamaan ihmistieteisiin liittyvää epävarmuutta - niiden tuottama hyöty perustuu erityisesti suurten massojen ja populaation analyysiin, mutta epävarmuudet kasvavat yksilöntason ennustamisessa (mm. Leslie ym. 2020; Villumsen ym. 2024).

Kaiken kaikkiaan automatisoitu päätöksenteko näyttäytyy sosiaalityön käytännöissä samanaikaisesti sekä lupaavana että merkittävien eettisten ja ammatillisten riskien lähteenä. Parhaimmillaan laajoihin ja yhdistettyihin tietomassoihin perustuvat riskimallit voivat tukea tapausten tunnistamista ja kokonaiskuvan muodostamista tuottamalla sosiaalityöntekijälle lisäinformaatiota asiakkaan tilanteesta, mutta algoritmien ylikäyttö ja asiakkaan ominaisuuksien tiukka kvantifiointi yhdistetään ihmisoikeus- ja oikeudenmukaisuuskysymyksiin sekä vinoumien vahvistumiseen. Mallien heikko selitettävyyttä voi heikentää asiakkaiden ja työntekijöiden mahdollisuuksia ymmärtää, haastaa ja oikaista algoritmin ohjaamia ratkaisuja.

## 7.2 Tekoäly ei ole moraalinen toimija

Tekoälyn hyödyntämiseen liittyy useita erilaisia eettisiä näkökulmia. Ne voivat liittyä joko tekoälyn tuottamiin vinoumiin ja sitä kautta vääristyneisiin tulkintoihin sosiaalisesta todellisuudesta. Toisaalta eettiset kysymykset voivat kytkeytyä kysymyksiin tekoälyjättien vastuullisuudesta. Ylen Ulkolinjan dokumentissa ”Tekoälyn ytimessä” (Yleisradio 2024) kuvataan, kuinka tekoälymallit pohjautuvat äärimmäisen laajaan määrään dataa, jonka kouluttamiseen tarvitaan valtava määrä ihmistyövoimaa ja luonnonvaroja – pääsääntöisesti kolmansista maista. Kouluttaessaan tekoälymalleja ihmiset altistuvat myös äärimmäisen vahingolliselle materiaalille, joka aiheuttaa heille inhimillistä kärsimystä.

Kansainvälisellä tasolla sosiaalityön eettistä harkintaa jäsentää ennen kaikkea kahden kansainvälisen sosiaalityön katto-organisaation IFSW:n (International Federation of Social Workers) ja IASSW:n (International Association of Schools of Social Work) tuottama Global Social Work Statement of Ethical Principles (2018), joka toimii yleisenä kehikkona eri maiden eettisille käytännöille. Ammattietiikan ytimessä olevat periaatteet ovat ihmisarvon tunnustaminen, ihmisoikeuksien edistäminen sekä sosiaalisen oikeudenmukaisuuden toteuttaminen. Näitä periaatteita vahvistetaan toimintavelvoitteilla, jotka edistävät itsemääräämisoikeutta ja osallisuutta sekä kunnioittavat luottamuksellisuutta ja yksityisyyttä. Nämä periaatteet ulottuvat myös teknologian käyttöön.

Ennakoivien riskimallien kehittämisessä ja hyödyntämisessä korostuvat erityisesti näkökulmat, jotka koskevat päätöksenteon oikeudenmukaisuutta ja laadukkuutta. The Alan Turing instituutin ja Oxfordin yliopiston yhteisesti laatimassa julkaisussa ”Lapsiperheiden sosiaalipalveluissa hyödynnettävien koneoppimisen ratkaisujen eettinen arviointi” (engl. ”Ethics review of Machine

Learning in Children's social Care") (Leslie ym. 2020) tarkastellaan koneoppimismallien hyödyntämisen eettisyyttä sosiaalityön olemassa olevan eettisen viitekehyksen kautta, jotta niihin liittyviä riskejä pystytään hallitsemaan. Leslie ynnä muut (2020) näkevät koneoppimisen ja datatieteen alueella paljon mahdollisuuksia, joiden valjastaminen lasten ja perheiden hyvinvoinnin parhaaksi tulee tehdä tietoisella otteella ja eettisesti harkiten.

Tekoälyasetuksen kansallista toimeenpanoa koskevan hallituksen esityksen luonnoksen lausuntokierroksella Amnesty International Suomen Osasto (2024) nosti esiin huolta siitä, että esimerkiksi muualta Suomeen tulevien henkilöiden ”turvallisuusuhan arviointi automatisoitujen riskinarviointi- ja profiloitijärjestelmien avulla voi [...] johtaa yksityisyyden ja tietosuojan loukkauksiin ja jopa vapauden ja turvallisuuden perusteettomiin rajoituksiin”. Toisaalta voidaan myös perustellusti kysyä, missä määriin ennakoivien riskimallien toimeenpano vahvistaa vinoumia sosiaalisten ongelmien tulkinnassa niiden ohjatesa tarkastelemaan sosiaalisia ilmiöitä riskilähtöisesti ja unohtamaan yksilön ja yhteisön käytössä olevat voimavarat (Lehtiniemi 2024, 1118).

Laajoihin kielimalleihin pohjautuvat menetelmät, kuten ChatGPT, jäsentävät ajattelua ilman myötätuntoa, järkeilyä, pohdintaa tai mielikuvitusta (Patton ym. 2023, 556). Sillä ei ole omaatuntoa, uskontoa tai oikeudenmukaisuuden tunnetta, vaan sen vastaukset heijastelevat internetissä yleisesti saatavilla olevaa tietoa. Segal arvioi (2025a & 2025b) laajoihin kielimalleihin perustuvan ChatGPT-tekoälyn kyvykkyyttä arvioida ja ratkaista yksilöllisten sosiaalityön tilanteiden eettisiä ristiriitoja. Kahdessa erillisessä tutkimuksessa tekoälyn kykyä ratkaista eettisiä kysymyksiä vertaillaan ammattilaisten tai sosiaalityö opiskelijoiden tekemään eettiseen harkintaan.

Ensimmäisessä tutkimuksessa tekoäly arvioi 14 kysymystä, jotka oli jo aiemmin osoitettu Israelin eettiselle komitealle. Eettisen komitean tehtävänä on käsitellä sosiaalityön alan yleisiä kysymyksiä sekä Israelissa työskentelevien sosiaalityöntekijöiden esittämiä kysymyksiä ja antaa niihin suosituksia. Kysymykset liittyvät esimerkiksi konkreettiseen asiakastilanteeseen tai sosiaalityön kentällä nouseviin organisatorisiin tai yhteiskunnallisiin kysymyksiin, jotka ovat erityisen haastavia eettisestä näkökulmasta. Segalin mukaan (2025a, 163) tekoäly suoriutuu erinomaisesti ristiriitaisten eettisten kysymysten tunnistamisessa ja jokseenkin hyvin lakinäkökulmien huomioimisessa. Ihmisammattilaisista koostuvaan eettiseen komiteaan verrattuna tekoäly ei kuitenkaan pysty antamaan ammattilaisen työn tueksi selkeitä,

yksiselitteisiä toimintaohjeita. Sen sijaan se tuottaa yleisluontoista ohjausta ja näkökulmia, joista sosiaalityöntekijän tulee itse lopulta arvioida oikea menettelytapa kyseiseen tilanteeseen.

Toisessa tutkimuksessa verrataan tekoälyn ja sosiaalityön opiskelijoiden eettistä päätöksentekoa. Tutkimuksessa tarkastellaan 20 erilaista sosiaalityön tapausesimerkkiä, joihin tekoäly ja opiskelijaryhmät laativat oman arvionsa. Analyysissa tunnistetaan kolme keskeistä ongelmakohtaa (Segal 2025b, 5): tekoälyn tapa jäsentää vaativia ammatillisia tilanteita vaihtelee, viittaukset eettisiin ohjeistuksiin ja lainsäädäntöön näyttäytyvät epäjohdonmukaisina ja vastauksissa ilmenee tunneherkkyyden sekä myötätunnon puutetta. Opiskelijoiden vastaukset koetaan laajemmiksi, oikeudellisesti vahvemmin perustelluiksi ja empaattisemmiksi verrattuna tekoälyn tuottamiin vastauksiin. Lopputulokset vahvistavat ajatusta siitä, että tekoälystä voi olla apua työssä, mutta se ei voi korvata sosiaalityöntekijän eettistä harkintaa eikä empatiaa. (Mt., 6.)

## 8 Yhteenveto ja pohdinta

Sosiaalityön tiedonmuodostus ja siihen kytkeytyvä päätöksenteko on usein monimutkaista ja perustuu hajanaiseen ja epävarmaan tietoon. Lisäksi sosiaalityöntekijät joutuvat käsittelemään valtavia määriä tietoa asiakasasiakirjoista, palvelujärjestelmistä ja muista hallinnollisista lähteistä, mikä vie huomattavasti aikaa asiakastyöltä, hidastaa käsittelyprosesseja ja lisää epävarmuutta ja muistinvaraisuutta ammattilaisten työssä. Asiakkaiden tilanteiden arviointi on vaativa tehtävä, sillä se edellyttää sekä yksilöllisten elämäntilanteiden syvällistä ymmärrystä että kykyä suhteuttaa ne lainsäädännön sekä palvelujärjestelmän ja resurssien asettamaan kehykseen. Käytännön työssä nämä haasteet näyttäytyvät kiireenä, kognitiivisena kuormituksena sekä päätöksenteon epätasalaatuisuutena.

Tekoälyn mahdollisuudet prosessien keventämisen ja systematisoinnin avulla on tunnistettu. Tekoäly nähdään esimerkiksi tiedon analysoinnin, vaihtoehtojen vertailun ja arviointiprosessien tukena. Samalla on kuitenkin olennaista huolehtia siitä, että tekoälyn tuottama lisäarvo täydentää sosiaalityöntekijän asiantuntijuutta ilman että se kaventaa inhimillisen kohtaamisen ja ammatillisen harkinnan keskeistä roolia. Tämä tapahtuu vain panostamalla laadukkaaseen arviointiin, käyttöönottoon sekä tekoälyn toimintalogiikan läpinäkyvyyteen ja tekoälyn taustalla olevan tiedon laatuun.

Tutkimukset osoittavat, että teknologiaosaamista tulee vahvistaa sosiaalityön koulutuksessa. Sosiaalityön koulutuksen tulee vahvemmin kytkeytyä teknologiamuutoksen hallitsemiseen.

James kumppaneineen (2024, 1554) kirjoittaa: ”Sosiaalityön koulutuksen on väitteemme mukaan yhä enemmän pystyttävä käsittelemään jännitteitä, jotka syntyvät toisaalta tulevien sosiaalityöntekijöiden valmentamisesta teknologialähtöiseen ammatilliseen toimintaympäristöön ja toisaalta opiskelijoiden perehdyttämisestä kriittisiin teorioihin, jotka tunnistavat sosiaalisen vallan epätasaisen jakautumisen. Teknologinen muutos sekä nousee tästä epätasa-arvosta että samalla uusintaa sitä abstraktimmalla tasolla.”

Tutkimus tekoälyn roolista sosiaalityössä on ajankohtaista ja tarpeellista. Ensinnäkin se tuottaa tietoa ammattilaisille ja opiskelijoille uudesta teknologiasta, joka on vasta hiljattain alkanut kytkeytyä sosiaalipalveluiden rakenteisiin. Tämä tieto on keskeistä, jotta voidaan ennakoida tekoälyn käyttöönoton edellyttämiä osaamistarpeita, ohjeistuksia ja eettisiä linjauksia. Toiseksi tutkimus auttaa paikantamaan niitä työelämän haasteita, joihin tekoäly voi tuoda arvokasta tukea. Tutkimus auttaa myös tunnistamaan ne tilanteet, joissa tekoälyn hyödyntäminen uhkaa ammatillisia arvoja, asiakkaan oikeuksia tai päätöksenteon läpinäkyvyyttä. Kolmanneksi tutkimuksen kautta on mahdollista tukea sosiaalityön kehittämistä siten, että tekoälyn käyttö rakentuu kriittisesti arvioiden ja ammattietiikan pohjalta, ei pelkästään teknologisen innovaation ehdoilla. Tutkimus tuottaa tietoa, joka palvelee sekä sosiaalityön koulutusta että käytännön työtä, ja tarjoaa välineitä laajempaan yhteiskunnalliseen kriittiseen tarkasteluun teknologian ja hyvinvointipalvelujen suhteesta.

Teknologinen kehitys tarjoaa uudenlaisia mahdollisuuksia tiedon hyödyntämiseen: Tekoälypohjaisten teknologioiden yleistyessä ja kehittyessä voidaan jopa päästä kokonaan eroon määrämuotoisen ja narratiivisen lähestymistavan vastakkainasettelusta. Koneoppiminen mahdollistaa suurien datamassojen hyödyntämisen rakenteisesta tiedosta. Se pystyy analysoimaan nopeasti ja tarkasti järjestelmään syötetyt määrämuotoiset tiedot. Toisaalta aivan viime vuosina harppauksina kehittyneet laajat kielimallit mahdollistavat tiedon haun ja analysoinnin myös silloin, kun se on pääasiassa narratiivisessa muodossa.

Koneoppimisen kyky tunnistaa ja analysoida ennalta tuntemattomia taustamuuttujia vauhdittaa keskustelua siitä, miten sosiaalisia ilmiöitä ylipäättään selitetään ja millaisin käsittein niitä jäsennetään. Samalla se haastaa tarkastelemaan uusia mahdollisuuksia sosiaalisten ilmiöiden operationalisoinnissa. Koneoppiminen tarjoaa uusia keinoja tarkastella sosiaalista todellisuutta ja tuottaa uudenlaisia menetelmiä aineistojen käsittelyyn ja mittarointiin. Menetelmät avaavat sosiaali- ja ihmistieteille uusia mielekkäitä tutkimuskohteita ja kytkevät sosiaalisten ilmiöiden tarkastelun ja sosiaalityön tiedonmuodostuksen lähemmäksi kriittisen realismin lähtökohtia.

Tavoitteeni oli selvittää, miten tekoälyratkaisut soveltuvat sosiaalihuollon tiedontuotantoon. Analyysin perusteella tekoälyn käyttökohteissa toistuu useita yhdenmukaisuuksia. Ensinnäkin tekoälyn käyttökohteet linkittyvät säännönmukaisesti tiettyihin teknologiatyyppeihin: generatiiviset kielimallit ja chatbotit sijoittuvat useimmiten sosiaalityöntekijän kollegan tai asiakkaan tuen rooleihin, kun taas ennustavat mallit ja algoritmeihin pohjautuvat päätöksenteon ratkaisut näyttäytyvät valvojina ja portinvartijoina, jotka luokittelevat, pisteyttävät ja priorisoivat asiakkaita. Toiseksi tiedon luonne eriytyy roolien mukaan: kollegaroolissa tekoäly tuottaa dialogista, narratiivista ja ehdottavaa tietoa, kun taas valvojaroolissa tieto tiivistyy numeeriseksi, jotka saatetaan tulkita faktana. Kolmanneksi useissa artikkeleissa tekoälyn asemaa sosiaalityön tehtävässä määritellään päätöksenteon tukena eikä korvaajana, mutta erityisesti valvoja- ja portinvartijaroolissa sen tuottama tieto kytkeytyy usein resurssien kohdentamiseen ja oikeudellisesti merkittäviin ratkaisuihin.

Toivon, että tämä pro gradu -tutkielma osaltaan rohkaisee käynnistämään lisää tekoälyyn liittyvää tutkimusta. Jatkossa olisi perusteltua kohdentaa tutkimusta erityisesti siihen, millaisin periaattein tekoälyratkaisuja kehitetään ja otetaan käyttöön sosiaalihuollossa. Samalla sosiaalityöhön tarvitaan selkeät arvioinnin ja valvonnan käytännöt, jotta käyttöönotto on vastuullista ja tiedonmuodostuksen läpinäkyvyys sekä selitettävyyden voidaan varmistaa. Tutkimus voi lisäksi tehdä näkyväksi tekoälyosaamiseen liittyviä puutteita ja siten tukea sosiaalityön koulutussisältöjen uudistamista. Kaikki nämä kokonaisuudet kytkeytyvät tiiviisti tekoälylukutaidon vahvistamiseen.

Teknologia kehittyi nopeasti ja edellisvuonna kehitetyt ratkaisut voivat jo tänään olla vanhentuneita. Uusia, entistä kehittyneempiä ratkaisuja on jo saatavilla. Juuri teknologian nopea kehittyminen ja yksityisten toimijoiden välinen kilpailu vaikuttavista ratkaisuista tekevät tekoälyn tutkimisesta mielenkiintoisen aiheen. Liikkuvaan junaan hyppääminen voidaan kokea vaikeaksi, mutta loikka kannattaa tehdä, sillä teknologian mahdollisuuksien ymmärtäminen avaa ovia uusien käytötapausten tunnistamiseen. Tekoälylukutaito ei ole vain mahdollisuus, se on vaatimus pysyä ajan hengessä mukana. Tärkeätä on hallita teknologiaa, jottei teknologia hallitse meitä.

## 9 Lähteet

- Ahn, Eunhye & Choi, Moon & Fowler, Patrick & Song, in Han (2025) Artificial Intelligence (AI) Literacy for Social Work: Implications for Core Competencies. *Journal of the Society for Social Work and Research* 16(1), 9-26. <https://doi.org/10.1086/735187>
- Al Maamari, Raya Hamed Hilal & Elsherbiny, Mohamed Mohamed Kamel (2025) Artificial Intelligence in Providing Psychosocial Support in Natural Hazards: A Semi- Systematic Literature Review. *Journal of Technology in Human Services* 43(2), 117-141. <https://doi.org/10.1080/15228835.2025.2470150>
- Alam, Nafees (2025) Social Work in the Age of Artificial Intelligence: A rights-Based Framework for evidence-Based Practice Through Social Psychology, Group Dynamics, and Institutional Analysis. *Journal of Evidence-Based Social Work United States* 23(1), 123-134. <https://doi.org/10.1080/26408066.2025.2547219>
- Amnesty International Suomen osasto (2024) Tekoälyasetuksen kansallinen toimeenpano. Lausunto 75/ 2024. [https://www.amnesty.fi/uploads/2024/12/lausunto\\_ai\\_fi\\_75\\_2024\\_tem\\_he\\_tekoalyasetuksen\\_toimeenpano.pdf](https://www.amnesty.fi/uploads/2024/12/lausunto_ai_fi_75_2024_tem_he_tekoalyasetuksen_toimeenpano.pdf) Viitattu 25.2.2026.
- Báez, Johanna Creswell & Ahn, Eunhye & Tamietti, Aubrey & Victor, Bryan G. & Goldkind, Lauri (2025) Clinical Social Workers' Perceptions of Large Language Models in Practice: Resistance to Automation and Prospects for Integration. *Journal of Evidence-Based Social Work United States* 23(1), 42-63. <https://doi.org/10.1080/26408066.2025.2542450>
- Bhaskar, Roy (2015) The possibility of Naturalism. A philosophical critique of the contemporary human sciences. 4. painos. 1. painos 1979. Oxon ja New York: Routledge.
- Chan, Chitai & Zhao, Jiahui & Zhao, Yunmeng (2025) Our journey with an AI assistant offering narrative therapy on WhatsApp. *Journal of Social Work Practice* 39(2), 165-189. <https://doi.org/10.1080/02650533.2024.2420314>
- Coeckelbergh, Mark (2026) Technofascism: AI, Big Tech, and the new authoritarianism. *AI and Society*. <https://doi.org/10.1007/s00146-026-02862-9>
- Cormen, Thomas H. & Leiserson Charles E. & Rivest, Ronald & Stein Clifford (2009) Introduction to Algorithms. 3. painos. Cambridge ja London: The MIT Press.
- Danermark Berth, Ekström Mats & Karlsson Jan Ch. (2019) Explaining society. *Critical Realism in the Social Sciences*. 2.painos. 1. painos 2001. Oxon ja New York: Routledge.
- 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>
- Díez, Esther Raya & Carmona, Manuel Trujillo & Muñoz, Domingo Carbonero (2021) Using Big Data to Manage Social Inclusion Programs. *Journal of Sociology and Social Welfare* 48 (3), 74-98. <https://doi.org/10.1177/10497315241280686>
- DigiFinland (2025) SOTE-tekoälyn ekosysteemin tekoälyvisio 2035. [https://digifinland.fi/wp-content/uploads/2025/06/SOTE-tekoalyn\\_ekosysteemin\\_tekoalyvisio\\_2035\\_final.pdf](https://digifinland.fi/wp-content/uploads/2025/06/SOTE-tekoalyn_ekosysteemin_tekoalyvisio_2035_final.pdf) Viitattu 3.3.2026.
- Euroopan parlamentti ja Euroopan unionin neuvosto (2024) Euroopan parlamentin ja neuvoston asetukset (EU) 2024/1689, annettu 13 päivänä kesäkuuta 2024, tekoälystä annettavista yhdenmukaistetuista säännöistä ja asetusten (EY) N:o 300/2008, (EU) N:o 167/2013, (EU) N:o 168/2013, (EU) 2018/858, (EU) 2018/1139 ja (EU) 2019/2144 sekä direktiivien 2014/90/EU, (EU) 2016/797 ja (EU) 2020/1828 muuttamisesta (tekoälysäädös), OJ L 2024/1689. <http://data.europa.eu/eli/reg/2024/1689/oj/fin> Viitattu: 27.2.2026.
- Garkisch, Michael & Goldkind, Lauri (2025) Considering a Unified Model of Artificial Intelligence Enhanced Social Work: A Systematic Review. *Journal of Human Rights and Social Work* 10 (1), 23-42. <https://doi.org/10.1007/s41134-024-00326-y>

- Gökçearslan, Elif & Tiktaş, Çelik Sevgi & Akdeniz, Emel Tuba & Öztürk, Emel (2025) The Touch of Artificial Intelligence in Social Work: Analysis of Social Investigation Reports in Child Welfare with ChatGPT. *Journal of Evidence-Based Social Work (United States)* 22(4), 607-625. <https://doi.org/10.1080/26408066.2025.2480396>
- Hao, Zhichao & Yang, Beichen & Ruggiano, Nicole & Ma, Yunying & Guo, Yuqi & Pan, Xiaofu (2022) Depression Prediction amongst Chinese Older Adults with Neurodegenerative Diseases: A Performance Comparison between Decision Tree Model and Logistic Regression Analysis. *British Journal of Social Work* 52(1), 274-290. <https://doi.org/10.1093/bjsw/bcaa237>
- Hiltula, Maija & Kemppainen, Tarja & Kostamo-Pääkkö, Kaisa & Liikanen, Sanna-Liisa & Sotaniemi, Ulla (2024) Sosiaalihuollon ammattilaiset pohjoisella yhteistyöalueella. Pohjois-Suomen sosiaalialan osaamiskeskuksen julkaisusarja, Nro 54, Pohjois-Suomen sosiaalialan osaamiskeskus, Rovaniemi. <https://lapitoy.sharepoint.com/:b/s/Virtu-tiedostomateriaali/EfMKZdF7GiBMtcVjleg-kcBRzPwAL2e158aCIU9dvlvFg?e=hRdNQO> Viitattu 24.11.2025.
- International Association of Schools of Social Work and International Federation of Social Workers (2018) Global Social Work Statement of Ethical Principles. 27.4.2018. <https://www.iasw-aiets.org/wp-content/uploads/2018/04/Global-Social-Work-Statement-of-Ethical-Principles-IASSW-27-April-2018-1.pdf> Viitattu 25.2.2026
- James, Paul & Lal, Jason & Liao, Ashley & Magee, Liam & Soldatic Karen (2024) Algorithmic decision-making in social work practice and pedagogy: confronting the competency/critique dilemma. *Social Work Education* 43(6), 1552-1569. <https://doi.org/10.1080/02615479.2023.2195425>
- Jokiranta, Harri (2024) Tietokäytäntöjen tiedontuotannon poliittisuus ja yhteydet yhteiskunnalliseen päätöksentekoon. SSOS1344V24-3001 Sosiaalityön tutkimusperustaisen työn kehittäminen. Luento 14.3.2025. Lapin yliopisto.
- Juvonen Tarja, Lindh Jari, Pohjola Anneli & Romakkaniemi Marjo (toim.) Sosiaalityön tutkimuksen vuosikirja 2018. Sosiaalityön muuttuva asiantuntijuus (2018)
- Kapur, Ishita & Kennedy, Reeve & Hickman, Christy (2025) Artificial Intelligence Algorithms, Bias, and Innovation: Implications for Social Work. *Journal of Evidence-Based Social Work* 22(4), 548-570. <https://doi.org/10.1080/26408066.2025.2470903>
- Karataş, Zeki (2025) Can a Large Language Model Judge a Child's Statement?: A Comparative Analysis of ChatGPT and Human Experts in Credibility Assessment. *Journal of Evidence-Based Social Work (United States)* 23(1), 78-93. <https://doi.org/10.1080/26408066.2025.2547211>
- Karttunen, Teija & Hietamäki, Johanna (2014) Tiedon käytön kysymyksiä ja haasteita sosiaalityön asiakastyössä. Teoksessa Riitta Haverinen, Marjo Kuronen & Tarja Pösö (toim.) Sosiaalihuollon tila ja tulevaisuus. Tampere: Vastapaino, 320–336.
- Kataja, Kati & Ristikari, Tiina & Paananen, Reija & Heino, Tarja & Gissler, Mika (2014) Hyvinvointiongelmiin ylisukupolviset jatkumot kodin ulkopuolelle sijoitettujen lasten elämässä. *Yhteiskuntapolitiikka* 79(1), 38–54. <https://urn.fi/URN:NBN:fi-fe201402131471> Viitattu 6.2.2026.
- Kim, Jinyi & Hsu, Julia Hsin Ping & Sohn, Gahyun & Lee, Giemyung & Lee, Myeong (2025) Leveling Socioeconomic Disparities: The Role of Service Availability in School Dropout Rates. *Research on Social Work Practice*. <https://doi.org/10497315251377009>
- Kääriäinen, Aino (2016) Hiljainen tieto ei kuulu sosiaalityöhön. *Talentia-lehti*. 20.7.2016. <https://www.talentia.fi/talentia-lehti/hiljainen-tieto-ei-kuulu-sosiaalityohon/> Viitattu 30.12.2025.
- Kääriäinen Aino (2024) Tekstit sosiaalityössä. Asiakirjat, kirjoittaminen ja tiedonmuodostus. Gaudeamus. E-kirja. Viitattu 12.1.2026.
- Kuusisto-Niemi, Sirpa & Kääriäinen, Aino (2005) Hiljaisen tiedon vallasta näkyvän tiedon valtaan - puheenvuoro tiedon näkyväksi tekemisen puolesta. *Janus: sosiaalipolitiikan ja*

- sosiaalityön tutkimuksen aikakauslehti 13(4), 452–460.  
<https://journal.fi/janus/article/view/50340/15196> Viitattu 2.12.2025.
- Lapin Yliopisto (2023) Tekoälypohjaisten työkalujen käyttö Lapin yliopistossa. Opetuksen johto- ja kehittämissyhmän kokous 2/2023 ja Strategisen johtoryhmän kokous 3/2023.
- Lehmuskoski, Antero & Palm, Niina & Korhonen, Katja & Suhonen, Marko (2025) Kanta-palvelujen käsikirja sosiaalihuollon toimijoille. THL.  
<https://yhteistyotilat.fi/wiki08/spaces/JULSOSK/overview> Viitattu 27.2.2026.
- Lehtiniemi, Tuukka (2024) Contextual social valences for artificial intelligence: anticipation that matters in social work. *Information Communication and Society* 27(6), 1110-1125.  
<https://doi.org/10.1080/1369118X.2023.2234987>
- Leslie, David & Holmes, Lisa & Hitrova, Christina & Ott, Ellie (2020) Ethics review of machine learning in children’s social care. *What Works for Children’s Social Care*. The Alan Turing Institute, Rees Centre, University of Oxford.  
[https://www.turing.ac.uk/sites/default/files/2020-02/wwcsc\\_ethics\\_of\\_machine\\_learning\\_in\\_csc\\_jan2020.pdf](https://www.turing.ac.uk/sites/default/files/2020-02/wwcsc_ethics_of_machine_learning_in_csc_jan2020.pdf) Viitattu 10.10.2025.
- Li, Lifei & Wang Miao & Jian Miao (2025) Artificial Intelligence-Assisted Case Management in Social Work Services: A Systematic Review. *Research on Social Work Practice*.  
<https://doi.org/10.1177/10497315251329531>
- Meilvang, Marie Leth & Dahler, Anne Marie (2024) Decision support and algorithmic support: the construction of algorithms and professional discretion in social work: Beslutningsstøtte og algoritmisk støtte: Konstruktionen af algoritmer og det professionelle skøn i socialt arbejde. *European Journal of Social Work*, 27 (1), 30-42.  
<https://doi.org/10.1080/13691457.2022.2063806>
- Mieli Ry (2024) Sekasin-chat hakee tekoälystä ratkaisuja nuorten mielenterveyskriisiin. *Verkkouutinen* 15.1.2025. <https://mieli.fi/uutiset/sekasin-chat-hakee-tekoalysta-ratkaisuja-nuorten-mielenterveyskriisiin/> Viitattu 15.1.2026.
- Omorogiwa, Tracy B. E. & Mugumbate, Rugare & Harms-Smith, Linda & Naami, Augustina & Diraditsile, Kabo (2025) Ethical and transparent use of generative Artificial Intelligence (AI): Ethics letter three (3) from the African Independent Ethics Committee (AIEC). *African Journal of Social Work* 15(10), 100-103. <https://dx.doi.org/10.4314/ajsw.v15i1.11>
- Opetus- ja kulttuuriministeriö (1991) Asetus tutkimuseettisestä neuvottelukunnasta 1347/1991. (15.11.1991). Suomen säädöskokoelma. Finlex. Viitattu 10.1.2026.
- Patton, Desmond Upton & Landau, Aviv Y. & Mathiyazhagan, Siva (2023) ChatGPT for Social Work Science: Ethical Challenges and Opportunities. *Journal of the Society for Social Work and Research* 14 (3), 553-562. <https://doi.org/10.1086/726042>
- Pekkarinen, Elina & Tapola-Haapala, Maria (2009) Kriittinen realismi sosiaalityössä: tiedontuotannosta emansipaatioon. Teoksessa Mäntysaari, Mikko & Pohjola, Anneli & Pösö, Tarja (toim.) *Sosiaalityö ja teoria*, 183-205. PS-kustannus. E-kirja.
- Perron, Brian Edward & Luan, Hui & Victor, Bryan G. & Hiltz-Perron, Oliver & Ryan, Joseph P. (2024) Moving Beyond ChatGPT: Local Large Language Models (LLMs) and the Secure Analysis of Confidential Unstructured Text Data in Social Work Research. *Research on Social Work Practice* 35(6), 695-710. <https://doi.org/10.1177/10497315241280686>
- Perron, Brian Edward & Hiltz, Barbara S. & Khang, Erin M. & Savas, Sue Ann (2025) AI-Enhanced Social Work: Developing and Evaluating Retrieval-Augmented Generation (RAG) Support Systems. *Journal of Social Work Education*, 61(1), 3-13.  
<https://doi.org/10.1080/10437797.2024.2411172>
- Pohjalainen, Marjut (2012) Hiljaisen tiedon käsite ja hiljaisen tiedon tutkimus: katsaus viime aikaiseen kehitykseen. *Informaatiotutkimus*, 31(3). <https://journal.fi/inf/article/view/7079> Viitattu 20.1.2026.

- Pohjola, Anneli & Satka, Mirja (2022) Sosiaalialan tietoperusta sosiaali- ja terveydenhuollon uudessa palvelujärjestelmässä. *Janus Sosiaalipolitiikan ja sosiaalityön tutkimuksen aikakauslehti*, 30(3), 289–295. <https://doi.org/10.30668/janus.121374>
- Pohjola, Anneli (2025) Sosiaalityön tutkimusperusteinen osaaminen ja vaikuttaminen palvelujärjestelmässä. SSOS1344V24-3001 Sosiaalityön tutkimusperustaisen työn kehittäminen. Luento 21.1.2025. Lapin yliopisto.
- Portela, Manuel & Castillo, Carlos & Tolan, Songul & Karimi-Haghighi, Marzieh & Pueyo, Antonio Andrés (2025) A comparative user study of human predictions in algorithm-supported recidivism risk assessment. *Artificial Intelligence and Law* 33 (2), 471-517. <https://doi.org/10.1007/s10506-024-09393-y>
- Päijät-Hämeen Hyvinvointialue (2025) Uutinen 18.6.2025. <https://vakehyva.fi/fi/ajankohtaista/ajankohtaista/tekoalypohjainen-chatbot-palvelu-testikayttoon-vaken-verkkosivuilla-5112025-alkaen-lue-usein-kysytyt-kysymykset-botin-kaytosta> Viitattu 12.12.2025.
- Robila, Mihaela & Robila, Stefan A. (2020) Applications of Artificial Intelligence Methodologies to Behavioral and Social Sciences. *Journal of Child and Family Studies* 29(10), 2954-2966. <https://doi.org/10.1007/s10826-019-01689-x>
- Saari Juho (2015 ) Elämän edellytykset. Huono-osaiset yhteiskunnan pohjalla. Helsinki: Gaudeamus.
- Salovaara, Samuel (2024a) Tietojärjestelmät osana sosiaalityön tiedonmuodostusta. Väitöskirja. Lapin yliopisto, Rovaniemi.
- Salovaara, Samuel (2024b) Sosiaalipalveluiden tiedolla johtaminen - Sosiotekninen ymmärrys tietoperusteisen toiminnan vahvistajana. SSOS1344V24-3001 Sosiaalityön tutkimusperustaisen työn kehittäminen. Luento 21.12.2024. Lapin yliopisto.
- Segal Michal (2025a) Confronting and managing ethical dilemmas in social work using ChatGPT. *European Journal of Social Work* 28(1), 155-167. <https://doi.org/10.1080/13691457.2024.2377786>
- Segal Michal (2025b) Social workers' evaluation of ChatGPT for solving ethical dilemmas within the limits of confidentiality. *Journal of Social Work Practice* 40(1), 5-18. <https://doi.org/10.1080/02650533.2025.2480092>
- Sitra (2026) Megatrendit. Kohti uutta yhteiskuntasopimusta. Sitran selvityksiä 251. <https://www.sitra.fi/julkaisut/megatrendit-2026/> Viitattu 7.1.2026.
- Sosiaali- ja terveysministeriö (2024) Sosiaalihuoltolain soveltamisopas 2024. Sosiaali- ja terveysministeriön julkaisuja 2024:1. <https://urn.fi/URN:ISBN:978-952-00-7155-4> Viitattu 2.3.2026.
- Sosiaali- ja terveysministeriö (2024) Yhdessä SOTE-tekoälyä kehittämään. Uutinen 20.6.2024. <https://stm.fi/-/yhdedssa-sote-tekoalya-kehittamaan> Viitattu 10.11.2025.
- Taskiran Tepe, Hacer & Aslantürk, Hüsnünur (2025) Comparative Efficacy of AI LLMs in Clinical Social Work: ChatGPT-4, Gemini, Copilot. *Research on Social Work Practice* 35(8), 968-980. <https://doi.org/10.1177/10497315241313071>
- Tulğan, Buğra & Pak Güre, Merve Deniz (2024) In the Echoes of Tomorrow: The Intersection of Social Work and Artificial Intelligence Through the Eyes of Turkish Students. *Journal of Social Service Research* 50 (4), 620-636. <https://doi.org/10.1080/01488376.2024.2354930>
- Tuomi, Jouni & Sarajärvi, Anneli (2018) Laadullinen tutkimus ja sisällönanalyysi. Uudistettu laitos. 1. painos 2002. Helsinki: Tammi. E-kirja.
- Tutkimuseettinen neuvottelukunta (2023) Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsitteleminen Suomessa. Tutkimuseettisen neuvottelukunnan julkaisuja 2/2023. [https://tenk.fi/sites/default/files/2023-03/HTK-ohje\\_2023.pdf](https://tenk.fi/sites/default/files/2023-03/HTK-ohje_2023.pdf) Viitattu 5.1.2026.
- Työ- ja elinkeinoministeriö (2024) EU:n tekoälyasetuksen kansallisen toimeenpanon työryhmä (hanke TEM044:00/2024; VN/11007/2024). Valtioneuvosto. Saatavilla: <http://valtioneuvosto.fi/hanke?tunnus=TEM044:00/2024> Viitattu 25.2.2026.

- Valtioneuvoston kanslia (2025) Ohjeistus generatiivisen tekoälyn hyödyntämisestä työn tukena ja apuvälineenä julkisessa hallinnossa VN/6190/2025. [https://valtioneuvosto.fi/documents/10623/1464506/Ohjeistus+generatiivisen+teko%C3%A4lyn+hy%C3%B6dynt%C3%A4misest%C3%A4+ty%C3%B6n+tukena+ja+apuv%C3%A4lineen%C3%A4+julkisessa+hallinnossa\\_saavutettava.pdf/e1f25f87-68fd-6994-f3ee-1c832988798f/Ohjeistus+generatiivisen+teko%C3%A4lyn+hy%C3%B6dynt%C3%A4mist%C3%A4+ty%C3%B6n+tukena+ja+apuv%C3%A4lineen%C3%A4+julkisessa+hallinnossa\\_saavutettava.pdf?t=1740975832781](https://valtioneuvosto.fi/documents/10623/1464506/Ohjeistus+generatiivisen+teko%C3%A4lyn+hy%C3%B6dynt%C3%A4misest%C3%A4+ty%C3%B6n+tukena+ja+apuv%C3%A4lineen%C3%A4+julkisessa+hallinnossa_saavutettava.pdf/e1f25f87-68fd-6994-f3ee-1c832988798f/Ohjeistus+generatiivisen+teko%C3%A4lyn+hy%C3%B6dynt%C3%A4mist%C3%A4+ty%C3%B6n+tukena+ja+apuv%C3%A4lineen%C3%A4+julkisessa+hallinnossa_saavutettava.pdf?t=1740975832781). Viitattu 12.12.2025.
- Valtiovarainministeriö (2025) Kokonaistaloudellisen ennusteen tausta-analyysjä 2025. Talousnäkyvät. Valtiovarainministeriön julkaisuja 2025:57. <https://urn.fi/URN:ISBN:978-952-367-184-3> Viitattu 20.2.2026.
- Vantaan ja Keravan hyvinvointialue (2025) Tekoälypohjainen chatbot-palvelu testikäyttöön VAKEn verkkosivuilla 5.11.2025 alkaen – lue usein kysytyt kysymykset botin käytöstä. Ajankohtaista 5.11.2025. <https://vakehyva.fi/fi/ajankohtaista/ajankohtaista/tekoalypohjainen-chatbot-palvelu-testikayttoon-vaken-verkkosivuilla-5112025-alkaen-lue-usein-kysytyt-kysymykset-botin-kaytosta> Viitattu 12.12.2025.
- Vehkalahti Kimmo (2019) Kyselytutkimuksen mittarit ja menetelmät. Helsingin yliopisto. 1. painos 2014. Finn Lectura. <https://doi.org/10.31885/9789515149817>
- Wartiainen, Piritta & Vuori, Anna & Tyllinen, Mari (2020) Digitalisaation mahdollisuudet sosiaalihuollossa asiakastietojärjestelmän näkökulmasta. Teoksessa Suoninen-Erhiö, Lea & Pohjola, Anneli & Satka, Mirja & Simola, Jenni (toim.) Sosiaaliala uudistuu. Tietopohjan ja vuorovaikutuksen kysymyksiä. Helsinki: Huoltaja-säätiön julkaisuja, 99-113.
- Wastell, David & White, Sue (2014) Beyond bureaucracy: Emerging trends in social care informatics. Health informatics journal, 20(3), 213-219. <https://doi.org/10.1177/1460458213487535>
- Wilkins, David & Benett, Verity (2025) Making Accurate Judgements in Child Welfare: Comparing ChatGPT With Qualified Social Workers. Child and Family Social Work. <https://doi.org/10.1111/cfs.13304>
- Yleisradio (2024) Tekoälyn ytimessä, Ulkolinja. TV1. Ensiesitys 20.2.2025; uusinta 23.2.2025.