

SAVONIA

ammattikorkeakoulu

OPINNÄYTETYÖ - YLEMPI AMMATTIKORKEAKOULUTUTKINTO
TEKNIIKAN JA LIIKENTEEN ALA

TEKOÄLY JA SOSIAALIALA

Blogimuotoisen opinnäytteen raportti

TEKIJÄ Katariina Hemming

Koulutusala Tekniikan ja liikenteen ala	
Tutkinto-ohjelma Master's Degree Programme in Digital Health	
Työn tekijä(t) Katariina Hemming	
Työn nimi Tekoäly ja sosiaaliala – Blogimuotoisen opinnäytteen raportti	
Päiväys 4.4.2024	Sivumäärä/Liitteet 62/24
Toimeksiantaja/Yhteistyökumppani(t)	
Tiivistelmä	
<p>Tekoälyteknologiat muokkaavat yhteiskuntaamme ja samalla sosiaalialan toimintaympäristöä. Tekoälyltä ja digitalisaatiolta odotetaan tukea yhteiskunnalliselle kehitykselle ja taloudelliselle kasvuille. Sosiaali- ja terveysalalla digitaalisten palveluiden strateginen kehittäminen liittyy tiedonhallinnan infrastruktuurin parantamiseen. Tavoitteena on, että syntyvää tietoa voidaan käyttää päätöksen teon tukena ja mahdollistetaan tiedon liikkuminen sujuvasti eri järjestelmien ja palveluiden välillä. Kehitys ei kuitenkaan voi tapahtua täysin kriittikittävästi, sillä etenkin sosiaalialalla sekä tiedon lähteenä että kohteena ovat yhteiskunnan heikommassa asemassa olevat jäsenet.</p> <p>Tämän opinnäytetyön tavoitteena on yhdistää keskustelu tekoälyteknologioista sosiaalialan toimintaympäristöön. Tutkimustehtävänä on tarkastella tekoälyn hyödyntämisen nykytilaa sosiaalialalla sekä pohtia tekoälyn ja sosiaalialan suhteen eettisyyttä. Tutkimustehtävä jäsenyy kahden tutkimuskysymyksen kautta, jotka ovat: 1.Miten tekoälyä jo hyödynnetään sosiaalialalla? 2. Miten eettisyys on huomioitu tekoälypohjaisissa järjestelmissä? Tutkimuksen viitekehityksessä yhdistyy sosiaalitiiteellinen, teknologinen ja filosofinen teoriaperusta, joiden yhteinen tekijä löytyy tiedosta. Tutkimus toteutettiin integroivana kirjallisuuskatsauksena ja sen aineiston muodostavat kansainväliset vertaisarvioituid tiedartikkelit. Analyysi toteutettiin sisällönanalysina. Tutkimusraportin ohessa opinnäytteestä julkaistu blogi on luettavissa osoitteessa:</p> <p>https://blogi.savonia.fi/peoplechangetheworld/tekoaly-sosiaalialalla/</p> <p>Tekoälyä hyödynnetään sosiaalialalla päätöksenteon tukena, mutta toistaiseksi ei juurikaan itsenäisesti vaan ensisijaisesti vertaillen koneoppimisalgoritmeja ja niiden tekemien päätösten osuvuutta suhteessa ihmisen tekemiin päätöksiin. Toinen tuloksissa painottuva näkökulma on tekoälyn tarjoamat mahdollisuudet ennaltaehkäisevään ja vaikuttavaan sosiaalityöhön, ei pelkästään asiakastyössä vaan myös alan ammattilaisten työssäjaksamisen tukena. Tekoälypohjaisten järjestelmien eettisten erityiskysymysten tarkastelussa on aineiston perusteella tunnistettavissa niin ikään kaksi pääluokkaa, jotka ovat tekoälykehitykseen liittyvän aineistoharhan tunnistaminen, sekä olemassa olevien sovellusten ja tekoälyjärjestelmien eettinen käyttö.</p> <p>Tekoälyllä on potentiaalia muuttaa sosiaalityötä ja tehostaa asiakkaille, organisaatioille ja yhteisöille tarjottavia palveluita. Eettinen ja oikeudenmukaisuuteen tähtäävä tekoälykehitys vaatii yhteistyötä lainsäätäjien, sovelluskehittäjien, loppukäyttäjien ja esimerkiksi sote-alan ammattilaisen välillä, jotka toimivat sekä tekoälyratkaisuiden käyttäjinä että asiantuntijoina asiakkaiden suuntaan.</p>	
Avainsanat Tekoäly, sosiaaliala, eettisyys, kirjallisuuskatsaus	

Field of Study Technology, Communication and Transport	
Degree Programme Master's Degree Programme in Digital Health	
Author(s) Katariina Hemming	
Title of Thesis Artificial Intelligence and Social Work – A report of the thesis published as a blog	
Date 4.4.2024	Pages/Appendices 62/24
Client Organisation /Partners	
<p>Abstract</p> <p>Artificial Intelligence (AI) technologies are shaping our society and, at the same time, the environment in which the social sector operates. AI and digitalisation are expected to support social development and economic growth. In the social and health sector, the strategic development of digital services is linked to improving the infrastructure for data management. The aim is to ensure that the information generated can be used to support decision-making and to enable the smooth flow of information between different systems and services. However, this development cannot be entirely uncritical, since, particularly in the social field, the disadvantaged members of society are both the source and the target of information.</p> <p>The aim of this thesis is to link the debate on AI technologies to the knowledge base of the social sector. The research task is to examine the current state of AI in the social sector and to reflect on the ethics of the relationship between AI and the social sector. The research task is structured through two research questions, which are: 1. How is AI already used in the social sector? 2. How is ethics considered in AI-based systems? The theoretical framework combines a social science, technological and philosophical knowledge base, with knowledge as the common element. The study was carried out as an integrative literature review and is based on international peer-reviewed scientific articles. The analysis was carried out using content analysis. The blog texts related to this research report are found from the link below: https://blogi.savonia.fi/peoplechangetheworld/tekoaly-sosiaalialalla/</p> <p>AI is being used in the social sector to support decision making, but so far not so much independently but primarily by comparing machine learning algorithms and the accuracy of their decisions with those of humans. Another aspect that is emphasised in the results is the potential of AI for preventive and effective social work, not only in client work but also as a support for the working life of professionals. In the examination of the specific ethical issues of AI-based systems, two main categories can also be identified on the basis of the data, namely the identification of data bias related to AI development, and the ethical use of existing applications and AI systems.</p> <p>AI has the potential to transform social work and improve the efficiency of services provided to clients, organisations, and communities. Ethical and fair AI development requires collaboration between legislators, application developers, end-users and, for example, social care professionals who act both as users of AI solutions and as experts towards clients.</p>	
<p>Keywords Artificial intelligence, AI, social sector, ethics, literature review</p>	

SISÄLTÖ

1	JOHDANTO	6
2	TUTKIELMAN TARKOITUS	7
3	MONITIETEINEN NÄKÖKULMA TEKÖÄLYTUTKIMUKSESSA	8
3.1	Vaikuttava sosiaalityö	8
3.2	Tekoäly ja etiikka	9
3.2.1	Moraaliteoriat tekoälykehityksessä	10
3.2.2	Big data ja siihen liittyvät eettiset kysymykset sosiaalialalla	11
3.2.3	Periaatteellinen ja metodinen eettisyys tekoälykehityksessä	12
4	INTEGROIVAN KIRJALLISUUSKATSAUKSEN TOTEUTUS	14
4.1	Integroivan kirjallisuuskatsauksen vaiheet	14
4.2	Aineistonhaku	15
4.3	Aineiston arviointi	17
4.4	Aineiston kuvaus	18
4.5	Aineiston analyysi	18
5	TULOKSET	21
5.1	Tekoälyjärjestelmien hyödyntäminen sosiaalialalla	21
5.1.1	Päätöksenteon tukeminen	21
5.1.2	Ennaltaehkäisy	22
5.1.3	Sosiaalisten ongelmien varhainen tunnistaminen ja niihin reagointi	22
5.1.4	Sosiaalialan ammattilaisten työssäjaksaminen ja muutoksenkyvykyys	23
5.2	Tekoälyjärjestelmien eettisyys	24
5.2.1	Aineistoharhan tunnistaminen tekoälykehityksessä	24
5.2.2	Sosiaalialalle suunnattujen tekoälyjärjestelmien ja sovellusten eettinen käyttö	24
6	JOHTOPÄÄTÖKSET JA POHDINTA	26
6.1	Tutkielman luotettavuus ja eettiset näkökulmat	27
6.2	Pohdinta	29
	LÄHTEET	32
	LIITE 1. ARTIKKELILUETTELO	38
	LIITE 2. BLOGI-TEKSTIT	45

KUVALUETTELO

KUVA 1 Yhteisönäkökulman tekijät teknologisen kehityksen eettisessä arvioinnissa (Leikas 2008, 77.)	13
KUVA 2 Laadullisen kirjallisuuskatsauksen tutkimusprosessi Fink 2005, 54, ref. Salminen 2011, 11, mukail- len.	15
KUVA 3 Esimerkki aineiston pelkistämisestä ja samankaltaisten ilmaisujen koodaamisesta	19
KUVA 4 Esimerkki aineiston analyysistä	19

TAULUKOT

TAULUKKO 1 Englanninkielisen aineiston sisäänotto- ja poissulkukriteerit	16
--	----

1 JOHDANTO

Tekoälyteknologiat muokkaavat yhteiskuntaamme ja samalla sosiaalialan toimintaympäristöä. Sosiaalialalla digitalisaatio on nähty pitkään uusien sosiaalisten ongelmien, kuten digitaalisen eriarvoistumisen, nettiriippuvuuksien tai yleistyvän, anonyymien vihapuheiden aiheuttajana (vrt. Kivistö 2017, 22). Tekoälyn suosion kasvun ja käytön arkipäiväistymisen myötä on kuitenkin herännyt kiinnostus myös siihen, miten teknologiaa voidaan hyödyntää sosiaalisten haasteiden ratkaisemiseen.

Tekoälyn ja digitalisaation mahdollisuudet yhteiskunnallisen hyvinvoinnin tukemiseksi on otettu tarkasteluun julkishallinnossa. Sosiaali- ja terveysministeriö sekä Terveyden ja hyvinvoinnin laitos ovat koonneet vuodesta 2016 lähtien tekoälyn ja robotisaation hyödyntämistä sosiaali- ja terveysalalla Hyteairo- ohjelman alle (Terveyden ja hyvinvoinnin laitos julkaisuaika tuntematon). Tekoälyä hyödynnetään virtuaalisina avustajina eli ns. chatbotteina sekä apuna lupaprosesseissa ja päätöksenteossa. Esimerkiksi Kansaneläkelaitoksen (KELA) hakemuksista yli puolet saapuu sähköisessä muodossa ja vuosittain järjestelmä tekee automaattisesti kymmeniä miljoonia hakijan kannalta myönteisiä suorakorvauspäätöksiä. (Koivisto, Leikas, Auvinen, Vakkuri, Saariluoma, Hakkarainen & Koulu 2019, 22.)

Pääministeri Orpon (2023) hallituksen ohjelmassa tekoälylle ja digitalisaatiolle annetaan painoarvo koko yhteiskunnallisen kehityksen ja taloudellisen kasvun mahdollistajana. Sosiaali- ja terveysalalla (myöhemmin sote) digitaalisten palveluiden strateginen kehittäminen liittyy tiedonhallinnan infrastruktuurin parantamiseen. Tavoitteena on, että syntyvää tietoa voidaan käyttää päätöksen teon tukena ja mahdollistetaan tiedon liikkuminen sujuvasti eri järjestelmien ja palveluiden välillä. Lisäksi ihmisillä pitää olla paremmat mahdollisuudet oman sote-tiedon hallintaan. Hallitusohjelmassa painotetaan myös mahdollistavaa lainsäädäntöä, jolla halutaan edistää tekoälyn hyödyntämistä esimerkiksi sosiaali- ja terveydenhuollon työvuorosuunnittelussa, ennaltaehkäisyssä sekä omahoidossa ja arvioinnissa. (Valtioneuvosto 2023.)

Tekoälyteknologioiden omaksuminen osaksi julkishallintoa ja julkisia sote-palveluita ei voi tapahtua täysin kriittittömästi. Euroopan komissio (2019) on julkaissut ylätasoinen ohjeet tekoälyn kehitykselle. Niissä otetaan kantaa tekoälyn kehittämiseen ihmisten perusoikeuksien, itsemääräämisoikeuden kunnioittamisen, yksilön vapauden ja tasa-arvon näkökulmasta. Ohjeistus ei kuitenkaan ulotu yksittäisiin, esimerkiksi ihmisten asiakkuuksiin tai työkuviin vaikuttaviin muutoksiin, joita tekoälykehitys ja sen laajamittainen soveltaminen yhteiskunnassa aiheuttavat. Parviainen & Rantala (2020) nostavat kriittiseen tarkasteluun erityisesti tekoälypohjaiseen, ennakoivaan päätöksentekoon liittyvät ongelmat, jotka voivat johtaa vääristymiin ja itseään toteuttavien, negatiivisten ennusteiden toteutumiseen. Esimerkiksi syrjäytymisen tarkasteluun ja ennaltaehkäisyyn kehitetyn koneoppimisalgoritmin avulla tehtyjen profiloitien vuoksi riskiryhmiin kuuluville voidaan kohdentaa tarpeettomia tukitoimia, ja siten leimata heidät epäonnistujiksi jo etukäteen. (emt. 5–6.)

2 TUTKIELMAN TARKOITUS

Opinnäytetyö on opiskelijan itsenäinen, mutta ohjattu tuotos, jolla pyritään osoittamaan opintojen aikana hankitun kokonaisvaltaisen tietoperustan analyttinen yhdistäminen rajattuun ongelmaan. Lisäksi opinnäytetyö valmentaa tekijäänsä tiedeviestintään ja tieteelliseen kirjoittamiseen, joka on paljon muutakin kuin kirjoittamista ja tekstin tuottamista. Tieteellinen kirjoittaminen on ajattelua, luetun uudelleen jäsentämistä ja akateemista toimintaa. Se yhdistää parhaimmillaan ajankohtaisen ammatillisen keskustelun tieteenalan tuottamaan tutkimustietoon. (Vilka 2020.)

Tämän opinnäytetyön tavoitteena on yhdistää keskustelu tekoälyteknologioista sosiaalialan tietoperustaan. Tutkimustehtävänä on tarkastella tekoälyn hyödyntämisen nykytilaa sosiaalialalla sekä pohdita tekoälyn ja sosiaalialan suhteen eettisyyttä.

Tutkimustehtävä jäsentyy kahden tutkimuskysymyksen kautta, jotka ovat:

1. Miten tekoälyä jo hyödynnetään sosiaalialalla?
2. Miten eettisyys on huomioitu tekoälypohjaisissa järjestelmissä?

”Sosiaaliala pyrkii edistämään yhteiskunnallista hyvinvointia ja auttamaan niitä, jotka tarvitsevat apua ja tukea. Tekoäly (AI) puolestaan edustaa teknologiaa, joka lupaa mullistaa useita eri aloja. Kuinka nämä kaksi aluetta liittyvät yhteen ja millaisia vaikutuksia niillä voi olla?”, kirjoittaa ChatGPT pyydettyäessä sitä ideoimaan johdanto tekoälyn ja sosiaalialan yhteyttä tarkastelevalle blogikirjoitukselle. Opinnäytetyön tulosten raportointi toteutetaan blogimuotoisena, sillä henkilökohtaisen ammatillisen kasvun ja asiantuntijuuden syventämisen lisäksi, blogi tarjoaa välineen herättää yhteiskunnallista keskustelua. Blogin avulla on mahdollista parantaa tieteelliseen tutkimukseen saavutettavuutta sekä kansantajuistaa tutkimusaiheita. (Jarreau 2015, 21–22.) Lisäksi blogi mahdollistaa tutkitun ja tärkeänä koetun asian näkyväksi tekemistä myös visualisoinnin keinoin (Helsingin yliopisto julkaisu-aika tuntematon). Blogissa hyödynnetään OpenArtAI tekoälyä kuvituksessa ja tekstien ideointiin käytetään ChatGPT:tä. Tekoälyn hyödyntäminen opinnäytteessä ja sen tuottama aineisto kerrotaan avoimesti ja selvästi.

Opinnäytteeseen liittyvät blogijulkaisut ovat luettavissa osoitteessa:

<https://blogi.savonia.fi/peoplechangetheworld/tekoaly-sosiaalialalla/>

3 MONITIETEINEN NÄKÖKULMA TEKÖÄLYTUTKIMUKSESSA

Tekoäly on yleistermi teknologioista, joiden avulla tietokoneet saadaan tekemään älykkäitä ratkaisuja esimerkiksi havaintotiedon perusteella. Tekoälylle ei ole yhteisesti sovittua määritelmää, mutta yleensä kyky suorittaa tehtäviä ilman valvontaa (autonomisuus) sekä suorituskyvyn parantaminen oppimalla (adaptiivisuus) ovat keskeisiä tekoälyyn liitettäviä ominaisuuksia. (Euroopan parlamentti 2023)

Tekoälyteknologiat luokitellaan kolmeen sen perusteella, kuinka hyvin ne jäljittelevät ihmisen ominaisuuksia ja millaista tekniikkaa tähän jäljittelyyn käytetään. Nykyiset laajalle levinneet tekoälyratkaisut edustavat kapeaa tekoälyä. Kapea tekoäly (Artificial Narrow Intelligence, ANI) suoriutuu hyvin sille opetetusta erityisestä tehtävästä, joka voi olla esimerkiksi kuvantunnistussovellus, shakkipeli tai työtehtävää suorittava robotti. (Hallamaa, Haikarainen & Kalliokoski 2021, 12.) Itsenäisesti toimiva, vahva tekoäly (Artificial General Intelligence, AGI) vastaa jo ihmisen kykyjä kokonaisuuksien hallinnassa ja pystyy soveltamaan älyään minkä tahansa ongelman ratkaisussa. Kun tähän liitetään myös inhimillisiä ominaisuuksia, kuten esimerkiksi minätietoisuus, voidaan jo puhua supertekoälystä (Artificial Super Intelligence, ASI), joka on monessa suhteessa kykenevämpi kuin ihminen. Kaksi jälkimmäistä tekoälyn muotoa ovat vasta hypoteettisia. (Boucher 2020, 13.)

Tutkielman teoreettinen lähestymistapa on interdisiplinaarinen eli monitieteinen ja siinä yhdistyy sosiaalitieteellinen, teknologinen ja filosofinen tietoperusta. Viitekehyksessä sovelletun kolmen tieteenalan yhteinen tekijä löytyy tiedosta. Ansaitakseen yhteiskunnallisen verovaroin kustannettavan toiminnan legitimitetin, sosiaalityön on pohjauduttava tietoon (Kivipelto & Saikkonen 2013, 313). Filosofia tarjoaa puitteet tiedon ja tietämisen ymmärtämiselle ja samalla auttaa arvioimaan kriittisesti erilaisia tietoon liittyviä kysymyksiä, kun taas tekoäly perustuu nykyisellään suurelta osin datan analysointiin, käsittelyyn ja mallintamiseen. Uuden tekniikan kehitys ja soveltaminen on prosessi. Tähän prosessiin liittyy useita virstanpylväitä, joiden aikana on pystyttävä arviomaan teknologiseen ratkaisuun liittyviä eettisiä periaatteita sekä punnittava ratkaisun vaikutuksia parhaan olemassa olevan tiedon perusteella. Tarkastelun kohteena voi olla käyttäjän näkökulma eli esimerkiksi ihmisarvoon, itsemääräämisoikeuteen ja valinnanvapauteen liittyvät periaatteet. (Leikas 2008, 77.) Tällaista pohdintaa on esimerkiksi se, loukkaako palveluasumisen tai sairaalan turvallisuutta parantamaan rakennettu videovalvonta yksityisyyttä ja muuttuu kontrollin välineeksi. (Vrt. EOAK 31.12.2019.)

3.1 Vaikuttava sosiaalityö

Sosiaalityö määritetään ammatti- ja tieteenalaksi, jonka tehtävä on edistää muutosta kohti oikeudenmukaisempaa yhteiskuntaa sekä auttaa ihmisiä erilaisten hyvinvointia heikentävien ongelmatilanteiden ratkaisemisessa (Sosnet julkaisuaika tuntematon). Määritelmän perusteella sosiaalityössä voidaan erottaa kaksi erilaista työorientaatiota, jotka ovat emansipatorinen ja yhteiskunnallista muutosta edistävä yhteisösosiaalityö sekä materiaalista tukea tarjoava ja hyvinvointia edistävä yksilötyö.

Molemmissa suuntauksissa keskeistä on asiakkaan itsensä aktivointi ongelmien ratkaisemiseksi ja elämänlaadun parantamiseksi. (Esim. Karjalainen 2016, 119–120.)

Vaikuttavan sosiaalityön tavoitteena on saavuttaa myönteisiä muutoksia asiakkaiden elämässä ja yhteiskunnassa sekä arvioida ja kehittää menetelmiä, joilla suotuisia vaikutuksia edistetään. Käsitteeseen liittyy läheisesti vaikuttavuustieto, jonka avulla pyritään resurssien mahdollisimman tehokkaaseen kohdentamiseen. Kivipellon ja Saikkosen mukaan (2013, 314) ”*Vaikuttavuuden arvioinnissa tarvitaan tietoa siitä, miten palvelut tukevat hyvinvointia, vähentävät eriarvoisuutta, ehkäisevät sosiaalisten ongelmien syvenemistä tai tuottavat tietoa yhteiskunnan epäoikeudenmukaisista rakenteista.*” Syntyvän tiedon hyödyntäminen resurssien kohdentamisessa sekä hyvinvointia edistävien menetelmien kehittäminen tekevät vaikuttavasta sosiaalityöstä sopivan kontekstin myös tekoälypohjaisten ratkaisuiden tarkastelulle.

Opinnäytetyössä tutkimuskontekstia ei kuitenkaan sidota pelkästään professionaaliseen sosiaalityöhön, vaan laajempaan sosiaalialan käsitteeseen, jolloin tarkasteluun voidaan ottaa mukaan koko sosiaalihuollon prosessi palvelutarpeen arvioinnista päätöksen tekoon sekä myös yhteiskunnalliset rakenteet, jotka vaikuttavat yksilön hyvinvointiin. (Tieteen termipankki julkaisuaika tuntematon.) Esimerkiksi sosiaalihuollon palvelutarpeen arviointeja ja päätöksiä tekee myös muut ammattihenkilöt kuin sosiaalityöntekijät, muun muassa toimeentulotuki siirtyi Kelan vastuulle vuoden 2017 alusta lähtien. (Esim. Saikkonen 2016, 33).

3.2 Tekoäly ja etiikka

Etiikka on tieteenala, joka vastaa kysymyksiin ”Miten pitäisi toimia?” ja ”Millainen on hyvä elämä?”. Tekoälyn etiikka tutkii eettisiä kysymyksiä, jotka liittyvät tekoälyn suunnitteluun, kehittämiseen, toteuttamiseen ja käyttöön. Se on soveltavan etiikan osa-alue, jonka tarkastelee ”hyvää” tekoälykehityksen taustalla. Soveltavan etiikan ohessa tekoälyn kehitystyön eettisyyttä ohjaa insinöörin ammatitietikka, tutkimusetiikka, sekä nykyisin yhä enenevässä määrin myös vastuullisen tekniikan kehittämisen periaatteet. Vastuullisen tekniikan kehityksen näkökulmasta uuden tekniikan kehitystyö ei saa olla itseisarvo. Sen sijaan tekniikka on väline, jonka tarkoitus on edistää ihmiselämän arvokkaita päämääriä, kuten demokratiaa, ihmisoikeuksia ja suvaitsevaisuutta. (Niiniluoto 2020.)

Tekoälyllä ei ole eettistä harkintaa eikä se pysty monitulkintaisuuteen, joten sen ohjelmoinnissa täytyy rajoittaa epäeettistä käyttäytymistä antamalla tekoälylle joko ehdottomia sääntöjä (top down) tai kuvaamalla lopputulosta (bottom up). Top down- lähestymistavassa erilaisia eettisiä periaatteita ja teorioita voidaan koodata järjestelmään ja antaa selkeät valinnan mahdollisuudet. Bottom up- lähestymistavassa tekoäly oppii ja kehittyy kokemusten ja datan perusteella ilman tiukkoja ennalta määriteltyjä sääntöjä. Usein monimutkaisissa, ja etenkin henkeen ja terveyteen vaikuttavissa, järjestelmissä käytetään hybridimallia. Tällöin järjestelmä tunnistaa tilanteet, joissa sen täytyy ehdottomasti noudattaa ennalta määrättyjä sääntöjä ja vastaavasti tilanteet, joissa oppiminen ja adaptoituminen on mahdollista. (Charisi, Dennis, Fisher, Lieck, Matthias, Slavkovik, Sombetzki, Winfield & Yampolskiy 2017, 4–5.)

Euroopan komission (2019) ohjeistuksessa luotettavalle tekoälykehitykselle edellytetään tekoälyjärjestelmän täyttävän koko elinkaarensa ajan lainmukaisuuden, teknisen ja sosiaalisen luotettavuuden sekä eettisyyden vaatimukset. Lainmukaisuuden täyttymiseksi tekoälykehityksessä tulee noudattaa kaikkia sovellettavia lakeja ja asetuksia. Järjestelmien sosiaalinen ja tekninen luotettavuus tarkoittaa esimerkiksi järjestelmän suorittaman päätöksenteon toistettavuutta, tahattoman haitan välttämistä ja tietoturvallisuutta. Eettisyyden vaatimus täydentää teknologiakehityksessä lainsäädäntöä, joka ei aina pysy nopean teknologisen kehityksen perässä. Mitä enemmän tekoälyjärjestelmiä otetaan käyttöön ja niille annetaan vastuuta päätöksenteosta sitä tärkeämpää on varmistaa, että ne toimivat ehdottomien arvojen mukaisesti, oikeudenmukaisesti ja järjestelmien elinkaaren aikaiset vastuut ja velvollisuudet ovat selkeät. (Euroopan komissio 2019, 11)

3.2.1 Moraaliteoriat tekoälykehityksessä

Normatiivisena etiikkana tunnettu tutkimusala pyrkii löytämään ratkaisuja siihen, mikä on oikeaa ja väärää toimintaa ja antaa mielenkiintoisen lähtökohdan pohtia tekoälyratkaisuiden eettistä ja oikeudenmukaista sovellettavuutta sosiaalialalla. Moraaliteorioista - jollaisia ovat esimerkiksi velvollisuusetiikka, utilitarismi ja sopimusteoriat - voidaan myös johtaa muun muassa Euroopan komission peräänkuuluttamia ehdottomia arvoja kehitystyön tueksi. Velvollisuusetiikan näkökulmasta tekoälykehityksessä arvioidaan aina tekoa, ei sen seurausta. (Esim. Spaulding 2020, 380.) Jos ihmisen yksityisyyden suoja on tärkein arvo yhteiskunnassa, tekoälyä ei voida esimerkiksi palveluasumisessa käyttää etävalvontaan, vaikka sen avulla olisi mahdollista parantaa asukkaiden turvallisuutta tai terveyttä.

Sen sijaan utilitaristisen eli seurausetiikan näkökulmasta on tärkeä punnita aiheuttaako toiminta enemmän hyötyä kuin vahinkoa. Tekoälyrobotteja voidaan kouluttaa tekemään tuottavia arvioita, joilla maksimoidaan hyöty suurimmalle joukolle ihmisiä. Klassinen esimerkki on tilanne, jossa autonomisesti ajava auto täynnä matkustajia joutuu tekemään ratkaisun, jossa se ajaa jalankulkijan yli, mutta samalla säästetään bussilastillinen ihmisiä. Teoreettisesti näin on säästetty useampi ihminen, mutta pohdittavaksi jää, onko ratkaisu silti oikea. (Kumar & Choudhury 2023.)

Uudemmissa moraaliteorioissa huomioidaan myös toiminnan konteksti sekä korostetaan empatiaa eettisen harkinnan taustalla. Esimerkiksi partikularismissa moraalinen päätös perustuu tarkkaan harkintaa yksittäisten tilanteiden erityispiirteistä. Sen mukaan tiukat eettiset säännöt tai periaatteet voivat johtaa epäoikeudenmukaisiin tai epärealistisiin päätöksiin. Relationismissa ja feministisessä etiikassa korostuu ihmisten väliset suhteet moraalisen päätöksenteon taustalla. Arvot rakentuvat suhteissa muihin ihmisiin ja yhteisöihin. (Esim. Gunkel 2020, 552.) Uudemmat moraaliteoriat eivät tarjoa ehdottomia ohjenuoria sosiaalialalla käytettävien tekoälypohjaisten ratkaisujen eettisen arvioinnin tueksi, vaan pikemminkin korostavat asiakassuhteen ainutkertaisuutta. Tekoälyjärjestelmä pystyy auttamaan ammattilaista päätöksenteossa, mutta lopullinen harkinta perustuu aina ihmisen toi-

mintaan. Teorioita kohtaan on myös esitetty kritiikkiä, sillä niiden avulla voidaan päätyä epäjohdonmukaisuuteen ja vaarannetaan yhdenvertaisuus, mikäli samankaltaisissa tilanteissa ei sovelleta samankaltaisia eettisiä periaatteita. (Leikas 2008, 63.)

Samankaltainen kohtelu samankaltaisissa tilanteissa viittaa vahvasti oikeudenmukaisuuteen. Oikeudenmukaisuus on yhteiskuntafilosofiassa ja etiikassa periaate, jonka mukaan jokainen saa osansa (Tieteen termipankki julkaisuaika tuntematon). Käsitteen määritelmä ei ole yksiselitteinen ja voi vaihdella eri kulttuurien, yhteiskuntien ja yksilöiden välillä. Lisäksi sen merkitys voi riippua kontekstista ja tilanteesta. John Rawlsin oikeudenmukaisuusteoria on yksi vaikutusvaltaisimmista sopimus-teorioista nykyaikaisessa moraalifilosofiassa. Se korvasi auktoriteettiin perustuvan ja ulkoa ohjautuvan moraalisen toiminnan ja korvasi sen tasa-arvoisten ihmisten vapaaseen tahtoon perustuvalla, yhteiskunnalliseen oikeudenmukaisuuteen ja tasapuolisuuteen tähtäävällä toiminnalla. (Herne 2015, 43–52; Lahtinen 2022.)

Oikeudenmukaisuus tekoälykehityksen taustalla tarkoittaa, että tekoälyn kehityksessä huomioidaan tasapuolisesti kaikkien ihmisten tarpeet ja varmistetaan, että kaikilla on yhteneväiset mahdollisuuden menestyä tekoälyyn pohjautuvassa maailmassa. Sosiaalialalla tulee huolehtia, että erityisesti heikommassa asemassa olevien ihmisten, kuten lasten, vammaisten ja syrjäytymisvaarassa olevien ihmisten tarpeet ja näkökulma tulee huomioiduksi. (Euroopan komissio 2019, 11, 16.) Nämä ihmiset ovat usein kuitenkin heitä, joiden data puuttuu tekoälyn koulutukseen käytettävästä datasta (esim. Kak, 2020) tai se on puolueellista (esim. Bankins, Formosa, Griep, & Richards 2022).

3.2.2 Big data ja siihen liittyvät eettiset kysymykset sosiaalialalla

Big data eli massadata on sähköisessä muodossa olevaa dataa, jonka käsittelyssä hyödynnetään koneoppimista, tekoälyä ja tilastollisia malleja. Massadataa kertyy useasta eri lähteestä, kuten asiakastietojärjestelmistä, rekistereistä, asiakastapaamisista ja -kyselyistä, älypuhelimista, sovelluksista, valvontalaitteista, antureista jne. (Zetino & Mendoza 2019, 409). Data vaihtelee muodoltaan, määrältään ja laadultaan, jolloin sekä sen prosessoimiseksi että yhteensovittamiseksi eri järjestelmien välillä tarvitaan uudenlaisia, tyypillisesti tekoälypohjaisia ratkaisuita. (Günther, Rezazade, Huysman, & Feldberg 2017, 191).

Tekoälyä hyödynnetään etenkin tunnistamaan sellaisia säännönmukaisuuksia monimutkaisesta tietomassasta, joihin ihminen ei pysty. Tällä pyritään palveluiden parempaan ennustettavuuteen ja sitä kautta resurssien oikea-aikaiseen kohdentamiseen. Niukkenevien resurssien myötä myös sosiaalipalveluiden tuottamaa arvoa suhteessa ihmisten toimintakykyyn, terveyteen ja hyvinvointiin tullaan arvioimaan kokonaisvaltaisemmin ja dataperusteisesti. (Zetino & Mendoza 2019, 410.) Suomessa sekä Jyväskylän yliopisto (Talvitie-Lamberg, Silvennoinen, Moilanen & Korpela 2018) että Espoon kaupunki (STT info 2018) ovat toteuttaneet hankkeita, joissa asiakasdatasta on pyritty tunnistamaan lasten ja nuorten syrjäytymistekijöitä, joita voidaan käyttää edelleen ennaltaehkäisyyn ja palvelutarpeiden ennakoinnin välineinä.

Massadatan käyttöön sosiaalisten ongelmien ratkaisussa liittyy kuitenkin useampi eettinen kysymys. Ensimmäinen on edellisessä luvussa esitetty datavinouma, joka johtuu heikommassa asemassa olevien ihmisten näkymättömyydestä lähdeaineistossa eli aliedustavuudesta. Toinen merkittävä eettinen kysymys liittyy erityisesti henkilökohtaisten tietojen käyttöön ja siihen vaadittavaan suostumukseen. Suostumuksen ehtona on, että sen antaja ymmärtää riittävällä tavalla keillä on pääsy dataan ja miten sitä käytetään. Datan on oltava yksilöitävissä myös eri järjestelmien kesken, jolloin sen on sisällettävä riittävästi yksilöllistä tunnistetietoa. Lisäksi esimerkiksi tekoälymallien kehitykseen valitaan datasta usein tietty otanta, jolloin tuon otannan satunnaistaminen vaikuttaa mallien luotettavuuteen ja ennustuskykyyn. Kehittäjien vastuulla on myös ohjelmoida tekoäly tunnistamaan sosiaaliset ongelmat, joihin on hankala löytää yleispätevää määritelmää ja joihin tyypillisesti liittyy vahva subjektiivisuus. Tällöin esimerkiksi tekoälypohjaisen järjestelmän tekemään päätöksentekoon voi liittyä osittain tahatontakin puolueellisuutta, epäjohdonmukaisuutta ja satunnaisuutta. (Gillingham & Graham 2017, 138–141.)

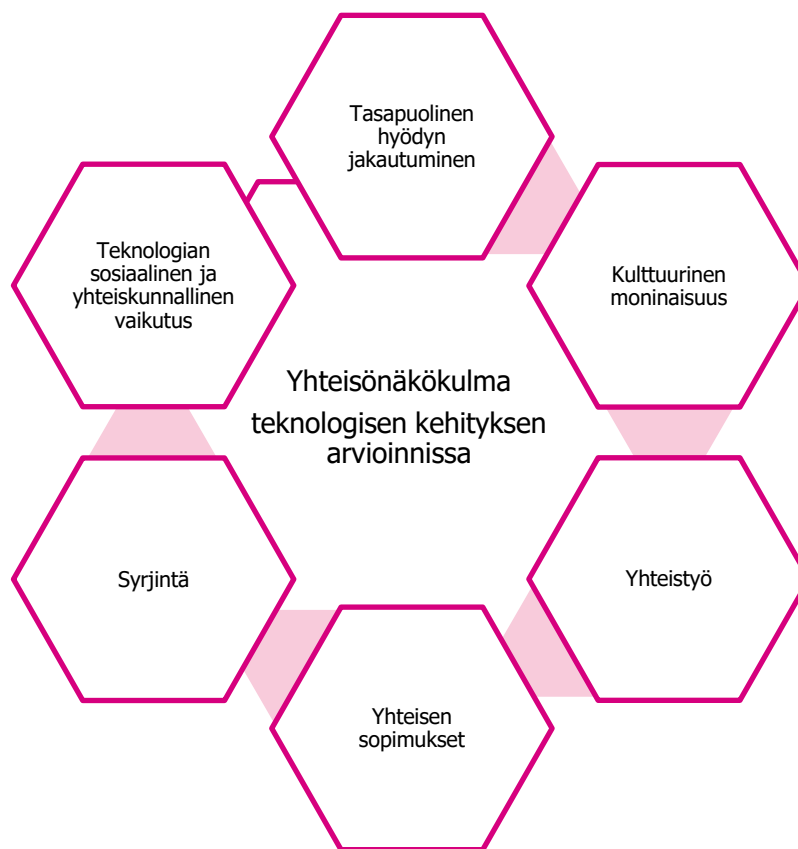
3.2.3 Periaatteellinen ja metodinen eettisyys tekoälykehityksessä

Tekoälykehitykseen liittyvää eettistä pohdintaa voidaan lähestyä myös periaatteellisesta tai metodisesta näkökulmasta. (Koivisto, Leikas, Auvinen, Vakkuri, Saariluoma, Hakkarainen & Koulu 2019, 15.) Periaatteellinen näkökulma on lähellä velvollisuusetiikkaa ja huomioi kehityksessä globaalit ihmisoikeussopimukset ja normistot, jollaisia ovat esimerkiksi EU:n tekoälykehitystä koskevat vaatimukset (Euroopan komissio 2019) tai Virginia Dignummin (2017) kehittämä ART (accountability, responsibility, transparency) menetelmä. Menetelmä perustuu yhteiskunnan valmiuteen ottaa vastuu tekoälyn vaikutuksista yhteiskunnallisen muutoksen agenttina. Sekä lainsäätäjien, tutkijoiden että tekoälykehittäjien tulee kouluttautua tiedostamaan oma vastuunsa tekoälyjärjestelmien kehittämistä. Vastuullisuus ei ulotu pelkästään kehitystyön ajalle, vaan vastuu on kannettava myös teknologisten ratkaisuiden jälkeen syntyvistä seuraamuksista, jotka eivät mahdollisesti ole vielä kehityshetkellä tiedossa. Läpinäkyvyys tarkoittaa sekä algoritmien toiminnan ymmärrettävyyttä että ihmisten oikeutta saada tietoa sekä heitä koskevan datan käytöstä että arkisessa toiminnassa kohtaamistaan tekoälyratkaisuista. ART periaatteita noudattamalla voidaan vähentää tekoälykehitykseen vaikuttavia satunnaisuuksia sekä hallitsemattomia vaikutuksia. (Dignum 2017, 4700)

Metodisessa näkökulmassa huomio on suunnitteluprossissa ja yksi sovelletuimmista, eettistä pohdintaa tukevista menetelmistä on Value Sensitive Design VSD (arvoherkkä suunnittelu). Yhteisöllä ja kulttuureilla on omat sosiaalisesti jaetut arvot ja normit, jotka ohjaavat myös teknologiakehitystä. VSD on periaatteellinen lähestymistapa, jonka tavoitteena on ottaa uuden teknologian kehityksessä huomioon eettisesti tärkeät arvot paitsi käyttäjien näkökulmasta myös kaikkien muiden, joihin teknologiat vaikuttavat riippumatta siitä, käyttävätkö kyseiset henkilöt koskaan tekniikoita. (Friedman 1996; Umbrello & van de Poel 2021, 284.)

Methodinen ja periaatteellinen lähestymistapa kytkeytyvät yhteen Leikaksen (2008) mallissa, joka on yhteisönäkökulma uuden teknologisen kehityksen arviointiin (kuva 1.). Mallin avulla on mahdollista

tarkastella tekniikan ja yhteisön välisen suhteen arvoperustaa, ja siksi se soveltuu myös tekoälykehitykseen liittyvän eettisyystarkastelun perustaksi.



KUVA 1 Yhteisönäkökulman tekijät teknologisen kehityksen eettisessä arvioinnissa (Leikas 2008, 77.)

Mallissa kuvatuilla yhteisillä sopimuksilla tarkoitetaan kansallisia ja kansainvälisiä lakeja ja standardeja, jotka ohjaavat tekoälyyn liittyvää kehitystä. Euroopan parlamentti hyväksyi kesällä 2023 tekoälyasetuksen, jota pidetään ensimmäisenä askeleena kohti sitovaa, yleiseurooppalaista tekoälykehitykseen liittyvää lainsäädäntöä (Euroopan parlamentti 2023). Sote-alalla uusien tekniikoiden käyttöönoton arvioinnissa on sovellettava lisäksi esimerkiksi ammattialalla yhteisesti sovittuja eettisiä ohjeita.

Teknologiat vaikuttavat aina ihmisen tapaan toimia ja ajatella maailmassa. Ne muuttavat työmenetelmiä, heijastuvat ihmissuhteisiin ja tapaamme kommunikoida, ja lopulta muokkaavat myös esimerkiksi sanavarastoa. Vaikutukset ulottuvat aina yksilöstä yhteisötasolle. Tällöin arvioinnissa on huomioitava, kuinka uusi teknologia palvelee eri taustoista tulevia ihmisryhmiä. (Hallamaa, Leikas, Malkavaara & Vesterinen 2020, 84.) Tekoälykehityksessä on tunnistettu riski, joka liittyy datan aiheuttamiin vääristymiin ja johtaa pahimmillaan vähemmistöryhmien entistä rajumpaan syrjintään. Syrjimättömyyden ja tasapuolisesti jakautuvaan hyödyn näkökulmaan liittyy yhteisön jäsenten yhtäläiset mahdollisuudet päästä käyttämään tekoälypohjaisia sovelluksia iästä, varallisuudesta ja digitaidoista riippumatta. (Koivisto, Leikas, Auvinen, Vakkuri, Saariluoma, Hakkarainen & Koulu 2019, 8.)

4 INTEGROIVAN KIRJALLISUUSKATSAUKSEN TOTEUTUS

Tutkielma toteutettiin laadullisena, integroivana kirjallisuuskatsauksena, ja tulosten julkaisussa hyödynnettiin perinteisen tutkimusraportin lisäksi myös Savonia Ammattikorkeakoulun blogi- alustaa. Integroiva kirjallisuuskatsaus sisältää piirteitä sekä systemaattisesta että kuvailevasta kirjallisuuskatsauksesta, mutta se soveltuu systemaattista kirjallisuuskatsausta paremmin kartoittavien, laajojen tutkimusongelmien ratkaisuun. (Salminen 2011, 11.) Lisäksi integroiva kirjallisuuskatsaus sopii hyvin kokonaiskuvan muodostamiseen ilmiöstä, kuten esimerkiksi tekoäly sosiaalialan kontekstissa, josta on vielä vähän tutkittua tietoa olemassa (vrt. Torraco 2005, 357). Integroiva, narratiivinen kirjallisuuskatsaus palvelee myös blogimuotoista julkaisemista, sillä sen tavoitteena on muodostaa epäyhtenäisestäkin lähdejoukosta tarinallinen, helppolukuinen kokonaisuus (Salminen 2011, 8). Aineistosta muodostuva synteesi välittyy lukijoille blogiteksteinä.

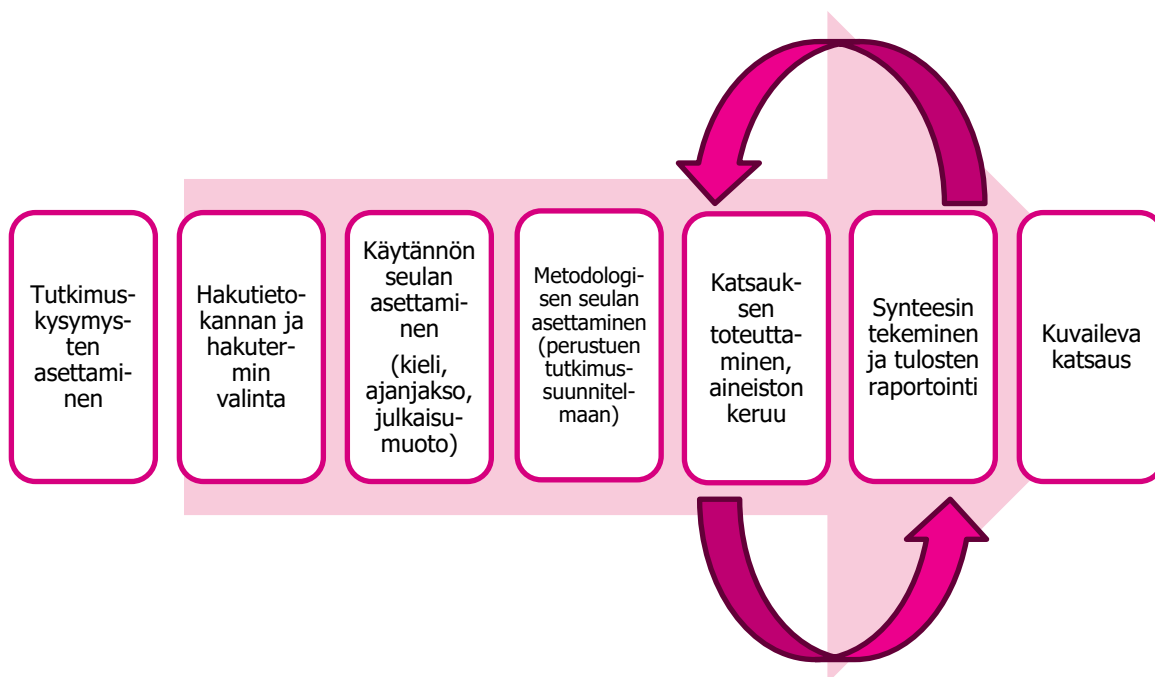
Tiedeblogi voidaan määritellä blogiksi, joka levittää, selittää, kommentoi, tutkii, kokoaa tai muuten käsittelee tiedettä ja tieteellistä tutkimusta (Jarreau 2015, 10). Se on paitsi tiedeviestintää nykyisin myös yhä enemmän tiedepolitiikkaa, jolla pyritään vaikuttamaan yhteiskunnallisiin mielipiteisiin ja päätöksentekoon (Lauhanen & Laasasenaho 2021).

Blogikirjoituksessa kiteytetään sanoma ja käytetään yleistajuista kieltä. Lisäksi se tarjoaa asiantuntijalle mahdollisuuden tuoda oma persoona ja näkemys esiin kielellisen ja visuaalisen ilmaisun keinoin. (Wallin 2021.) Tästä syystä opinnäytteen blogijulkaisuissa ei noudateta Savonia ammattikorkeakoulun opinnäytetyön ohjeistettua raportointia, vaan opinnäyteraportti pidetään erillään itse blogijulkaisuista. Raportissa kerrotaan kattavammin menetelmävalinnoista ja aineiston tuottamisesta sekä si-
dotaan tutkittava aihe tutkimuskontekstin käsitteisiin ja tutkimusperinteisiin. Blogikirjoitukset on liitetty osaksi raporttia ja ennen julkaisua ne on tarkastettu plagiaatintunnistusohjelmalla.

4.1 Integroivan kirjallisuuskatsauksen vaiheet

Kirjallisuuskatsauksen vaiheet voidaan jäsentää ylätasolla neljään vaiheeseen, jotka ovat: 1. tutkimuskysymyksen muodostaminen, 2. aineiston valitseminen, 3. kuvailun rakentaminen ja 4. tulosten tarkasteleminen. (Kangasniemi, Utriainen, Ahonen, Pietilä, Jääskeläinen & Liikanen 2013, 292). Myös tässä opinnäytteessä noudatettiin kyseisiä vaiheita eli tutkimustyö käynnistyi tutkimusongelman asettamisella, jota seurasi aineiston hankinta, arviointi ja tulkinta.

Nykyään myös laadullisilta kirjallisuuskatsauksilta vaaditaan systemaattisuutta vähintäänkin suunnitelmallisuuden, läpinäkyvyyden ja tutkimusprosessin kuvaamisen osalta (Aveyard 2014, 119–120) ja integroiva kirjallisuuskatsaus noudattaa systemaattista kirjallisuuskatsausta etenkin tutkimuksen raportoinnissa. Aineiston tuottaminen ja analyysi kuvataan mahdollisimman tarkasti, jotta tutkimuksen luotettavuutta ja toistettavuutta on mahdollista arvioida. (Cooper 1998, 15.) Tämän opinnäytteen tutkimusprosessi on kuvattu Finkia (2005, 54, ref. Salminen 2011, 11) mukailleen kuviossa 2.



KUVA 2 Laadullisen kirjallisuuskatsauksen tutkimusprosessi Fink 2005, 54, ref. Salminen 2011, 11, mukaillen.

Alkuperäisestä kuvaajasta poiketen prosessi on kuvattu horisontaalisesti ja siihen on lisätty nuolet aineiston keruun ja synteesin muodostamisen välille. Integroivassa kirjallisuuskatsauksessa on systemaattista kirjallisuuskatsausta enemmän vapauksia lähdeaineiston valinnassa ja se voi sisältää sekä laadullisia että määrällisiä aineistoja. Lisäksi aineistoa voidaan tuottaa lumipallomenetelmällä (Saaranen-Kauppinen & Puusniikka 2009, 49), joka kirjallisuuskatsauksessa tarkoittaa, että aineistoa täydennetään harkinnanvaraisesti esimerkiksi aineistoon jo valitun artikkelin lähdeluettelo hyväksikäyttäen. Tältä osin tutkimusprosessissa on mahdollista täydentää aineistoa vielä analyysivaiheen aikana, ja palata synteesistä edelliseen vaiheeseen.

4.2 Aineistonhaku

Tutkielman aineiston muodostavat integroivan, narratiivisin maustein täydennetyt kirjallisuuskatsauksen periaatteita noudattaen kerätyt kansainväliset ja kotimaiset tiedejulkaisut sekä vertaisarvioitut artikkelit. Opinnäytetyön aineistonhaku toteutettiin eksplisiittisesti eli se on kuvattu tutkimusraportissa mahdollisimman tarkasti. Valitut tietokannat ja kriteeristö on kirjoitettu auki lukijalle. (Snyder 2019, 334; Kangasniemi, Utriainen, Ahonen, Pietilä, Jääskeläinen & Liikanen 2013, 295.)

Kirjallisuuskatsauksen hakuprosessi määrittelee tutkimuksen onnistumisen, sillä virheellinen haku johtaa myös aineiston vääristymään ja harhaisuuteen. Aineiston harhaisuutta voi estää tekemällä koehakuja useammasta eri tietokannoista, testaamalla hakufraaseja etukäteen sekä laatimalla näiden perusteella hakustrategian. (Kraus, Breier, Lim, Dabić, Kumar, Kanbach, Mukherjee, Corvello, Piñeiro-Chousa, Liguori, Palacios-Marqués, Schiavone, Ferraris, Alberto, Fernandes & Ferreira, 2022, 2584). Opinnäytteen koehaut suoritettiin UEF Primo hakupalvelusta, jonka tulosten perusteella englanninkielinen tutkimushaku päätettiin toteuttaa EBSCOhost Academic Search Premier tietokannasta ja suomenkielisen tutkimuskirjallisuuden osalta ARTO artikkeliviitetietokannasta.

Varsinainen tutkielman aineistohaku käynnistyi tietokantahauulla EBSCOhost Academic Search Premier tietokannasta. Ensimmäisen kierroksen sisäänottokriteereiksi muodostui englanninkieliset, vertaisarvioidut, tieteelliset artikkelit, joiden kokoteksti on saatavissa maksuttomana verkkojulkaisuna. Ajallisesti kirjallisuuskatsaukseen valikoitui vuoden 2013 tammikuusta joulukuulle 2023 tehdyt julkaisut. Haku termeillä ”artificial intelligence or ai or a.i. or machine learning or deep learning” AND ”social work or social workers or social work practice or social services” tuotti yhteensä 198 tulosta. Hakuä tasmennettiin rajaamalla haun ulkopuolelle Amerikan alkuperäisväestöön viittaava käsite NOT ”American Indians”, joka lyhennetään myös AI. Näillä ehdoilla artikkeleiden määräksi tuli 117 kappaletta. Hakuehdoissa huomioitiin laajennus ”Apply equivalent subjects”, jolloin hakukone tunnistaa haussa käytetyille käsitteille rinnakkaiset termit tietokannan asiansastoissa. Kaksoiskappaleiden poistamisen ja tiivistelmien lukemisen jälkeen aineistoon valikoitui tässä vaiheessa 21 artikkelia.

Uudella hakukierroksella hakutermiin lisättiin AND ”ethics or ethical issue or ethical concerns or ethical dilemmas”, jolloin artikkelimäärä tippui yhdeksään. Otsikoiden, asiansanojen ja abstraktien perusteella näistä yhdeksästä artikkelista aineistoon mukaan valikoitui neljä. Aineiston sisäänotto- ja poissulkukriteerit on esitetty seuraavassa taulukossa:

TAULUKKO 1 Englanninkielisen aineiston sisäänotto- ja poissulkukriteerit

Sisäänottokriteerit	Poissulkukriteerit
Vertaisarvioitu, akateeminen artikkeli	Tutkimushankkeiden hakukuulutukset, ns. call for papers sekä pääkirjoitukset
Kokoteksti saatavilla verkossa	Maksulliset ja ns. predatorilehtien artikkelit
Englannin kieli	Muun kieliset julkaisut
Tekoälyn etiikkaa sosiaalialalla käsittelevät vertaisarvioidut akateemiset artikkelit	Robotikkaan ja erityisesti kyborgeihin liittyvät julkaisut
Koneoppimista ja kielimalleja käsittelevät artikkelit, joiden tutkimuskontekstina on jokin yhteiskunnallinen ilmiö (katastrofi, vähemmistötyö)	Tekoälyn soveltamista terveydenhuollossa käsittelevät sekä tekoalan soveltamista terveys- ja sosiaalialan koulutuksessa käsittelevät artikkelit
Tekoälyn soveltamista sosiaalialalla käsittelevät artikkelit	Tekoälyn hyödyntämistä tutkimustyössä ja kielin huollossa käsittelevät artikkelit

Vaikka katsauksesta jätettiin ulkopuolelle hakukuulutukset, erään tieteellisen konferenssin hakukuulutuksen lähdeluettelosta aineistoon valikoitui mukaan artikkeli, jossa sosiaalityöntekijöille annettiin suosituksia generatiivisen tekoälyn ja kielimallien käytössä omassa työssään. Aineistoon tässä vaiheessa valikoituneet englanninkieliset artikkelit tallennettiin omiin kansioihin odottamaan tarkempaa luentaa ja analyysia.

Suomenkielinen haku ARTO artikkeliviitetietokannasta toteutettiin hakukomennoilla ”tekoäly” OR ”AI” and ”sosiaaliala” OR ”sosiaalityö”. Aineistoa haettiin samalta ajanjaksolta kuin englanninkielistä aineistoa eli vuosien 2013–2023 väliltä. Haku rajattiin tieteellisiin artikkeleihin ja e-artikkeleihin, jolloin haku tuotti tuloksena yhden artikkelin. Artikkelin ei soveltunut aiheensa puolesta tutkimusaineistoon, mutta jälleen sen lähdeluettelosta löytyi tutkimusongelman kannalta kiinnostavia suomenkielisiä lähteitä, joita voi hyödyntää tutkimusaiheen taustoittamisessa ja teoreettisessa viitekehyksessä. Hakuehtojen laajentaminen lisäämällä yksi kerrallaan aineistotyyppisiin väitöskirjat, tutkimusraportit ja tutkimusjulkaisut ei tuottanut lisää tulosta, joten suomenkielinen aineisto jätettiin tässä vaiheessa pois tiedonhausta.

4.3 Aineiston arviointi

Koska kuvaileva kirjallisuuskatsaus mahdollistaa aineistoon sisällytettävän sekä laadullista että määrällistä tutkimusta - ja laadullinen tutkimus on aina jossain määrin ainutlaatuista - tutkijalla ei ole käytettävissä samanlaista mittaristoa aineiston arvioinnissa kuin systemaattista kirjallisuuskatsausta tekevällä tutkijalla (Saini & Shlonsky 2012, 9). Integroivan kirjallisuuskatsauksen aineistonvalintaprosessiin liittyy kuitenkin sisäänkirjoitettu kriittisyys (Salminen 2011, 8), joten kirjallisuuskatsauksen valikoituvia tiedonlähteitä on pystyttävä arvioimaan sekä aineiston laadun että relevanssin näkökulmasta.

Aineiston arvioinnissa voidaan soveltaa Aveyardia (2018, 108) mukailten seuraavia kysymyksiä:

1. Mistä tieto on peräisin? Löytyikö se sattumalta vai systemaattisen haun tuloksena?
2. Onko lähde riittävän laadukas sisällytettäväksi aineistoon? (esim. kieliasu, toteutus, johdonmukaisuus)
3. Milloin, miksi ja missä tieto on julkaistu?
4. Kuka tiedon on julkaissut? Voiko se olla puolueellista? (esim. poliittiset tai uskonnolliset julkaisut)
5. Onko aineisto oleellinen tutkimusongelman kannalta?

Ensimmäisiin apukysymyksiin vastasi riittävällä tavalla sisäänottokriteereissä huomioitu vertaisarvioitu, akateeminen artikkeli, joten seuraavalla karsintakierroksella luenta keskittyi viidenteen apukysymykseen eli aineiston oleellisuuden kirjallisuuskatsauksen ja tutkimusongelman näkökulmasta. Tarkemman luennan jälkeen aineiston muodostivat 11 englanninkielistä tutkimusartikkelia, jotka täyttivät sisäänottokriteerit ja vastasivat tutkimusongelmaan. Kuvailevan kirjallisuuskatsauksen luonteen mukaisesti lopullinen päätös aineistosta ei kuitenkaan perustu tutkimuksen täydelliseen toistettavuuteen, vaan sen argumentointivoimaan suhteessa tutkimuskysymykseen. (Kangasniemi, Utraiainen, Ahonen, Pietilä, Jääskeläinen & Liikanen 2013, 296).

4.4 Aineiston kuvaus

Aineistosta rajautui ensimmäisen luennan jälkeen pois pääkirjoitukset (n=2), hakukuulutus (n=1), kommentaari (n=1) sekä isoimpana joukkona artikkelit, jotka eivät lainkaan käsitelleet tekoälyn soveltamista sosiaalialan kontekstissa (n=9). Myös aikaisemmin tarkempaan tarkasteluun päätyneet artikkelit, jossa sosiaalityöntekijöille annettiin suosituksia generatiivisen tekoälyn ja kielimallien käytössä omassa työssään, rajautui kriittisen tarkastelun jälkeen ulos aineistosta, sillä se ei vastannut tutkimuskysymykseen, vaan oli pikemmin ohjeistus tekoälyn käytöstä. Analyysivaiheessa aineisto täydentyi kirjallisuushaun perusteella mukaan valitun artikkelin lähteenä käytetystä artikkelista, jossa käsiteltiin koneoppimista työssä uupumisen tunnistamiseksi. Myös tämä artikkeli täytti sisäänottokriteerit ja laatuvaatimukset ollen vertaisarvioitu, englanninkielinen tutkimusartikkeli. Lopullisen tutkimusaineiston 12 artikkelia on kuvattu liitteessä 1.

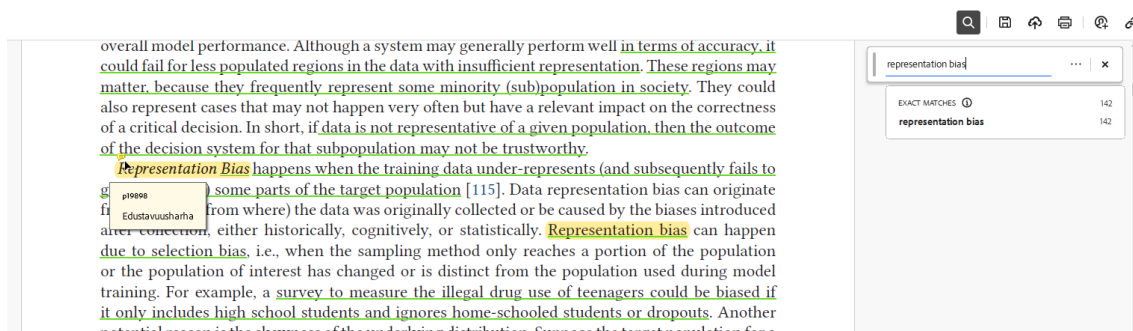
Aineistossa on mukana sekä kvalitatiivisella että kvantitatiivisella otteella toteutettuja tutkimusartikkeleita. Ajallisesti aineistoon valituista artikkeleista suurin osa on julkaistu 2020- luvulla lukuun ottamatta yhtä raskauden aikaisten riskitekijöiden ja sosiaalityön interventioiden välistä yhteyttä tutkivaa tutkimusta. Aineisto jakautuu tutkimuskysymyksiin siten, että kahdeksan artikkelia käsittelee tekoälyn käyttöä sosiaalialalla ja neljä tekoälyn käytön eettisyyttä. Jälkimmäisistä artikkeleista kaksi ei käsittele puhtaasti tekoälykehityksen eettisyyttä sosiaalialalla, vaan aineistovääritysten syntymistä ja niiden korjaamista yhteiskunnallisen ja historiallisen datan perusteella. Toisessa datavinoumiin keskittyvässä artikkelissa huomioitiin lisäksi kehittyvät maat ja maantieteellisestä köyhyydestä johdettu ihmisten näkymättömyys koneoppimismallien ja tekoälyn kehityksessä, joten aihepiiriensä puolesta molemmat artikkelit ovat relevantteja, kun tarkastellaan tekoälyjärjestelmien eettisyyttä.

4.5 Aineiston analyysi

Laadulliset kirjallisuuskatsaukset ovat aineistolähtöisiä ja tämän opinnäytteen osalta analyysi käynnistyi artikkeleiden temaattisella luennalla (Vrt. Raitakari & Takala 2023, 311). Luennan aikana aineistoa pelkistetään, eli siitä karsitaan tutkimuskysymysten kannalta merkityksetöntä sisältöä ja samalla lisätään aineiston informaatioarvoa. Analyysin alkuvaiheessa pitäydytään aineiston alkuperäisilmaisuuksissa. (Juuti & Puusa 2020.) Opinnäytteeseen valittujen artikkeleiden temaattisen luennan aikana aineistoon tehtiin alleviivauksia ja lisättiin huomiota. Artikkelit on tallennettu paikallisesti pdf-muodossa, joten myös merkinnät tehtiin sähköisesti.

Aineistolähtöisen eli induktiivisen sisällönanalyysin tavoitteena on muodostaa aineistosta tiivis ja selkeä kokonaisuus, joka vastaa tutkimusongelmaan. Pelkistämisen jälkeen aineistoa pilkotaan yhä pienempiin osiin ja niitä ryhmitellään samankaltaisten tai eroavien ilmaisujen joukoiksi. Tätä voidaan kutsua myös aineiston koodaamiseksi, jolloin samaa tarkoittavia ilmaisuja merkitään yksinkertaisin koodein. Koodien avulla on mahdollista tarkastella samaa tarkoittavien ilmaisujen esiintyvyyttä, joka auttaa analyysin myöhemmässä vaiheessa havaintojen luokittelussa. (Puusa 2020.) Koska tämän

opinnytteen aineistoon valitut artikkelit oli tallennettu sähköisessä muodossa, koodaamisessa pystyi hyödyntämään ilmaisujen etsimisessä CTRL + F hakutoimintoa. Löydetty tekstipätkät alleviivattiin, mikäli näin ei ollut tehty vielä temaattisen luennan aikana. Täsmälleen samoihin sanamuotoihin lisättiin myös korostusväriä.



KUVA 3 Esimerkki aineiston pelkistämisestä ja samankaltaisten ilmaisujen koodaamisesta

Koodaamisen jälkeen samankaltaisista tai toisistaan selkeästi poikkeavista ilmaisuista muodostetaan kategoria, joista voidaan tutkimusperinteestä riippuen käyttää myös nimeä tema. Kuitenkin temaattinen analyysi ja sisällönanalyysi eivät ole identtisiä tutkimusperinteitä, vaikka käsitteitä käytetäänkin ristiin. Sisällönanalyysissä kategorioiden muodostamisessa pyritään hierarkkisesti etenemään pelkistetyistä ilmaisuista ryhmittelyn kautta kohti aineiston abstrahointia eli käsitteellistämistä. Kategorioiden tulee vastata tutkimuskysymyksiin, mutta induktiivisessa sisällönanalyysissä yhdistäviä luokkia muodostetaan aineiston perusteella, eikä esimerkiksi teoriasta nousevien teemojen mukaan. (Tuomi & Sarajärvi 2018.)

Alkuperäinen ilmaus	Pelkistetty ilmaus	Alaluokka	Yläluokka	Yhdistävä luokka
Automated Decision Systems (ADS) are widely used in society, ranging from fire prevention by predicting high-risk buildings to recruiting automation by screening for competitive candidates. However, historical data used for decision-making might not be objective; it could inherit historical biases in the algorithm design.	inherit historical biases in the algorithm design	Algoritmiin periytyvät historialliset aineistovääritykset	historialliset aineistovääritykset	Aineistoharha tekoälykehityksessä
AI contributes to what Ash et al. (2018) term geographies produced by the digital, where "the digital is mediating and augmenting the production of space and transforming spatial relations" (2018, p. 29), including notions of a "digital divide."	digital divide	Digitaalinen eriarvoisuus, joka perustuu asuinpaikkaan	maantieteellisen eriarvoisuuden aiheuttamat aineistovääritykset	Aineistoharha tekoälykehityksessä
This datafication is creating new forms of spatial comparison	creating new forms of spatial comparison	Digitaalinen eriarvoisuus, joka perustuu yhteiskunnalliseen asemaan.	asemaan tai asuinpaikkaan perustuva epätasa-arvoisuus aineistossa	Aineistoharha tekoälykehityksessä
AI is likely to become part of intensification and disruption that both reinforces uneven geographies of development.	uneven geographical development	Digitaalinen eriarvoisuus, joka johtuu elinympäristön kehittymättömyydestä	asemaan tai asuinpaikkaan perustuva epätasa-arvoisuus aineistossa	Aineistoharha tekoälykehityksessä
Representation Bias happens when the training data under-represents (and subsequently fails to generalize well) some parts of the target population	data under-represents some parts of the target population	Tietyn osan aliedustavuus kohdeväestön datassa.	väestöryhmän aliedustus otoksessa	Aineistoharha tekoälykehityksessä

KUVA 4 Esimerkki aineiston analyysistä

Kuvassa on esimerkki opinnäytteen induktiivisen sisällönanalyysin etenemisestä aineiston pelkistämisestä ryhmittelyn kautta abstrahointiin. Vaikka induktiivinen sisällönanalyysi on aineisto-ohjaavaa, ei sitä pysty täysin pitämään erillään tutkijan esiyymmärryksestä tutkittavasta aiheesta, joka on muodostunut tutkimusaiheen teoreettisen tarkastelun ja käsitelmäärittelyiden myötä. Laadullisen aineiston käsittely on kuitenkin aina myös aineiston tulkintaa ja samalla tulosten kannalta kriittistä, sillä lopulta tutkija päättää mitkä ilmaisut luokitellaan mihinkin kategoriaan. (Puusa 2020.) Analyysivaihetta voidaan tehdä lukijalle läpinäkyväksi kuvaamalla sen vaiheet mahdollisimman tarkasti ja aineistoesimerkein, kuten edellä on pyritty tekemään. Integroivassa kirjallisuuskatsauksessa aineiston valintaa, tulkintaa ja raportointia ohjaa sisäänkirjoitettu kriittisyys (Salminen 2011, 8). Analyysivaiheessa tämä tarkoittaa, että aineistoa voidaan edelleen karsia tai täydentää, mikäli valitulla lähteellä ei ole argumentointivoimaa suhteessa tutkimuskysymykseen. Tämän opinnäytteen kohdalla pitäydettiin yhtä artikkelia lukuun ottamatta tietokantahakujen tuottamassa aineistossa, sillä niistä muodostui riittävän monipuolinen kuva tutkittavasta ilmiöstä.

5 TULOKSET

5.1 Tekoälyjärjestelmien hyödyntäminen sosiaalialalla

Tässä luvussa vastataan ensimmäiseen tutkimuskysymykseen: Miten tekoälyä jo hyödynnetään sosiaalialalla? Aineiston perusteella tekoälyn käytöstä sosiaalialalla tunnistettiin kaksi pääsuuntausta. Ensimmäinen sitä käytetään päätöksenteon tukena, ja näissä tutkimuksissa on ensisijaisesti vertailtu koneoppimisalgoritmeja ja niiden tekemien päätösten osuvuutta suhteessa ihmisen tekemiin päätöksiin. Toinen pääkategoria muodostuu tekoälyn tarjoamista mahdollisuuksista ennaltaehkäisevään ja vaikuttavaan sosiaalityöhön. Kiinnostavin tulos on tekoälypohjaisten järjestelmien hyödyntäminen asiakkaan tukemisen lisäksi myös sosiaalityöntekijöiden työssäjaksamisen tukemisessa ja työuupumuksen ennakoivassa tunnistamisessa.

5.1.1 Päätöksenteon tukeminen

Tekoälymalleja käytetään sosiaalialan päätöksentekoprosessissa palvelutarpeen arvioinnin tukena, riskiarvioinnissa ja resurssoinnissa. Tutkimuksen mukaan koneoppimisalgoritmit suoriutuivat samankaltaisesti tai ylittivät noin joka kolmannessa tapauksessa lomakepohjaisen päätöksenteon sosiaalialan palvelutarpeen arvioinnissa. Tulos vaikuttaa heikolta, mutta sitä selittää puutteellinen tai virheellinen asiakasdata. Koneoppimismallien hyödyt tulevatkin parhaiten esiin silloin, kun aineistomäärä kasvaa. Esimerkiksi moniulotteisesta sosiaalialan ja terveydenhuollon yhteisestä asiakasdatasta pystyttiin tunnistamaan selvästi nykyistä käytäntöä paremmin synnytyskomplikaatioita ennustavia sosiaalisia riskejä (esimerkiksi mielenterveysongelmat, perheväkivalta tai vankeus), vaikka näille tekijöille annettiin koneoppimisalgoritmeissa terveysdataa pienempi painotus. (Pan, Nolan, Brown, Khan, Van Der Boor, Harris, & Ghani 2017.)

Riskiarvioinnin ja resurssointipäätösten tukena on kokeiltu myös kotiin asennettavaa teknologiaa. Sen avulla pyrittiin tunnistamaan ja ennakoimaan mahdolliset vaaratilanteet etenkin aikuissosiaalityössä poikkeavan käytöksen perusteella ja siten kohdentamaan uusia tukitoimia ja tehostamaan palvelua. Kotianturoinnista syntyvän datan mallintamisessa ja säännönmukaisuuksien etsimisessä hyödynnettiin tekoälyä. Anturoinnilla ei kuitenkaan ole vielä pystytty saavuttamaan haluttuja tuloksia, sillä niiden käyttöönottoon ja datan tulkitsemiseen liittyi paljon epävarmuuksia, jolloin järjestelmän tekemisiin suosituksiin tai päätösehdotuksiin ei voitu luottaa ja tilanteiden todentamiseksi tarvittiin aina ihmisen tekemä interventio. (Glasby, Litchfield, Parkinson, Hocking, Tanner, Roe & Bousfield 2023, 40–41.)

Teknisesti tarkasteltuna tekoälypohjaisten päätöksentekomallien kehityksessä suurimmaksi hidasteeksi on muodostunut lähdedataan liittyvät haasteet. Data on usein vaillinaista, jolloin siitä puuttuu esimerkiksi sellaisia merkittäviä sosiodemografisia muuttujia, jotka vaikuttavat suoraan mallien ennustevoimaan. Tällaisia ovat esimerkiksi ikään, siviilisäätyn ja sosioekonomiseen asemaan liittyvät taustatekijät. (Cresswell, Callaghan, Khan, Sheikh, Mozaffar, Sheikh 2020, 941). Toinen dataan liit-

tyvä huoli on sen sensitiivisyys, koska dataan kirjataan paljon yksityiskohtaista tietoa asiakkaan lisäksi myös hänen perhetilanteestaan ja läheisistään. Vaarana on myös järjestelmien systemaattinen puolueellisuus etenkin riskien arviointimalleissa, joissa on käytetty opetusmateriaalina pelkästään asiakasdataa. (Gillingham 2021, 44.)

Päätöksentekomallien yleistymistä sosiaalialalla hidastaa sekä virkamiesten että päättäjien ennakkoluulot järjestelmiä kohtaan, joka johtuu useimmiten käytetyn teknologian tuntemattomuudesta. Etenkin tekoälyjärjestelmät, jotka vaativat toimiakseen erikseen asennettavaa teknologiaa koetaan vieraiksi ja vaivalloisiksi. Toisaalta toimialan resurssien niukkuuden takia aikaa perehtymiselle ja koulutukselle ei myöskään ole. (Glasby, Litchfield, Parkinson, Hocking, Tanner, Roe & Bousfield 2023.)

5.1.2 Ennaltaehkäisy

Suomessa on pyritty muuttamaan sosiaalityön työorientaatiota reaktiivisesta ja korjaavasta kohti ehkäisevää ja vaikuttavaa sosiaalityötä (Karjalainen & Sarvimäki 2005, 79). Ennaltaehkäisevä työorientaatio on myös globaali trendi, (esim. Banks 1999), joten on luontevaa, että sen tukemiseksi etsitään ratkaisuja uudesta teknologiasta. Ennaltaehkäisevän sosiaalityön tueksi on kokeiltu ennakkoluulottomasti kotiin asennettavaa teknologiaa, esimerkiksi anturointia, jonka avulla pyritään tunnistamaan asiakkaan poikkeava käytös. Anturit keräävät tietoa ympäristöstään ja tekoälyn avulla pyritään tunnistamaan tiettyjä käyttäytymismalleja sekä erityisesti poikkeamia. (Glasby, Litchfield, Parkinson, Hocking, Tanner, Roe et al. 2023.) Ennaltaehkäisevän tutkimusorientaation alta pystyi tunnistamaan kaksi alakategoriaa, joista toinen kohdistuu asiakkaan ongelmien varhaiseen tunnistamiseen ja toinen sosiaalialan ammattilaisten työssäjaksamiseen.

5.1.3 Sosiaalisten ongelmien varhainen tunnistaminen ja niihin reagointi

Kotona annettavan sekä perhehoidon hoivan määrä kasvaa globaalisti, joten samalla myös hoitohenkilökunnan tarve kasvaa. Digitaalisista teknologioista haetaan ratkaisua, jolla voidaan parantaa hoitoa sekä vastata sosiaalihuoltojärjestelmään kohdistuviin paineisiin niukkenevista resursseista huolimatta. Tutkimuksissa oli tarkasteltu kotiin asennettavan teknologian toimivuutta sekä asiakkaan näkökulmasta että sosiaalialan ammattilaisten näkökulmasta. Kotiin ja palveluasumiseen asennettavien tekoälyjärjestelmien tarkoitus oli tunnistaa sellaisia muutoksia asiakkaisen käyttäytymisessä, jotka viittaavat mahdollisesti vakavampiin ongelmiin. Kyseessä ei ollut valvontakamerat vaan liiketunnistimiin perustuva anturointi. (Litchfield, Glasby, Parkinson, Hocking, Tanner, Roe, & Bousfield 2023) Jatkotutkimuksessa tarkasteltiin ammattilaisten ja päättäjien näkemystä teknologian hyödynnettävyydestä asiakkaan tuen tarpeen ennakoimisessa, mutta haasteeksi koettiin teknologian hyödyntämisen strategian puuttuminen sekä teknologian käyttöönoton asettamat vaatimukset henkilökunnalle. Vaikka ammattilaiset kokivat, että teknologialla on mahdollisuus tukea asiakkaiden itsenäistä

ja aktiivista elämää, testattu teknologia ei sellaiseen palveluun sosiaalialan toimintaympäristössä. Henkilökunta raportoi kokevansa tilanteen siten, että ihmisten tulee sopeutua teknologian vaateisiin eikä teknologian ihmisten tarpeisiin. (Emt; Glasby J, Litchfield I, Parkinson S, Hocking L, Tanner D, Roe B, et al. 2023.)

Kotiin asennettavan teknologian lisäksi tutkimuksissa tarkasteltiin sosiaalityöntekijöiden tekemiä vapaaehtoisia asiakaskirjauksia luonnollisen kielen analyysillä (NLP). Tutkimuksen tarkoitus oli tukea päätöksentekijöitä resurssoinnissa ja toiminnan suunnittelussa. Koneoppimista hyödyntämällä kirjauksista pyrittiin tunnistamaan sairaalassosiaalityöntekijöiden asiakkaan tukemiseksi tekemiä interventioita, sekä niiden laatua ja määrää suhteessa muuhun potilaskertomukseen. Yleisimmät kirjauksista tunnistetut sosiaalityöntekijöiden palvelut olivat: hoidon koordinointiin liittyvät toimenpiteet (21,5 %), asiakkaan koulutus (21,0 %), taloussuunnittelu (18,5 %) ja asiakkaan ohjaaminen kunnan tai järjestöjen palveluihin (17,1 %). Tutkimuksessa havaittiin, että sairaalassa vietetyt päivät lyhenivät merkittävästi, kun sosiaalityöntekijä osallistui sairaalassa hoidon suunnitteluun ja kotiutuksen koordinointiin. (Bako, Taylor, Wiley, Zheng, Walter-McCabe, Kasthurirathne & Vest 2021.)

5.1.4 Sosiaalialan ammattilaisten työssäjaksaminen ja muutuskävykyys

Sosiaalityön digitalisaatiolla on kahdenlaisia vaikutuksia sosiaalialan ammattilaisten ammattitaidon ylläpitämiseen ja uusien ammattilaisten kouluttamiseen. Ensinnäkin toimialan digitalisaatio asettaa uudelleen koulutusvaatimuksia ammattilaisille, jotta he oppivat käyttämään ja tulkitsemaan uutta tekniikkaa eettisesti ja asianmukaisesti. Toisaalta myös opetus on muuttunut yhä enemmän teknologiavälitteiseksi. Perinteisesti kasvokkain ja harjoitteluissa opittu toisten ihmisten kunnioittava kohtelu, eettinen arviointi ja ihmisoikeuskysymykset voivatkin muuttua kustannuspaineiden alla yhä enemmän itsenäisiksi verkko-opinnoiksi. (Hodgson, Goldingay, Boddy, Nipperess & Watts 2022)

Sosiaalialan ammattilaisille on tarjolla lukuisa määrä tekoälyä hyödyntäviä sovelluksia, joita voidaan käyttää tukena asiakastyössä ja harjoittaa niiden avulla asiakkaiden sosiaalisia taitoja, vihan hallintaa sekä tarjota tukea yksinäisyyteen (Reamer 2023, 56). Tutkimuksista ilmeni, että näiden suoraan asiakasrajapinnassa toimivien sovellusten lisäksi tarjolla on myös teknologiaa, jonka avulla voidaan arvioida sosiaalityöntekijöiden työkuormaa ja tunnistaa työuupumusriskiä. Koneoppimisalgoritmit tunnistavat hyvin työssä uupumiseen ja loppuun palaamiseen viittaavat muuttujat kompleksisista aineistoista. Menetelmä mahdollistaa siirtymisen yksilöllisestä työssä jaksamisen tukemisesta kohti tiimityöskentelyä ja auttaa tunnistamaan työyhteisöjä, joissa työkuorma on kova ja tukea tarvitsevia työntekijöitä on paljon. Ennaltaehkäisevänä toimenpiteenä pystytään esimerkiksi tehostamaan esihenkilötyöskentelyä, parantamaan työsuunnittelua sekä edesauttamaan mahdollisia lainsäädännöllisiä muutoksia vaativia toimenpiteitä (esimerkiksi henkilöstömitoitus), joiden avulla työntekijöiden loppuun palamisen riskiä voidaan pienentää. (Grządzielewska 2021.)

5.2 Tekoälyjärjestelmien eettisyys

Tämä luku vastaa toiseen tutkimuskysymykseen: Miten eettisyys on huomioitu tekoälypohjaisissa järjestelmissä? Aineistosta on tunnistettavissa kaksi pääluokkaa, jotka ovat tekoälykehitykseen liittyvä aineistoharhan tunnistaminen, sekä olemassa olevien sovellusten ja tekoälyjärjestelmien eettinen käyttö.

5.2.1 Aineistoharhan tunnistaminen tekoälykehityksessä

Yksi keskeisin eettinen kysymys tekoälykehityksen taustalla on tekoälyn koulutuksessa käytetyn datan laatu ja siitä johtuva edustavuusharha (Shahbazi, Yin, Asudeh, & Jagadish 2023, 2). Edustavuusharha johtuu tekoälykoulutuksessa käytetyn datan yksipuolisuudesta, algoritmin virheellisyydestä tai käyttöympäristön muutoksista. Tällöin tekoäly soveltaa oppimaansa väärin, koska sen koulutusaineisto ei kata kaikkia mahdollisia tilanteita tai ihmisryhmiä, joita se kohtaa käytännössä. Tekoälystä haetaan ratkaisuja useisiin globaaleihin huono-osaisuuteen johtaviin tekijöihin, mutta kääntöpuolena sen ennustetaan vahvistavan eriarvoistumista etenkin kehittyvissä yhteiskunnissa, joissa tuloerot ja koulutuserot ovat suuret. Köyhyydessä ja syrjäisillä seuduilla asuvilla ihmisillä ei ole mahdollisuutta käyttää teknologiaa, jolloin heistä ei jää digitaalista jälkeä, jota käytetään tekoälyjärjestelmän koulutusaineistossa. Tämä voi johtaa siihen, että vinoumien tunnistamisen sijaan tekoäly ruokkii - osin tahattomastikin - eriarvoistumista. (McDuie-Ra & Gulson 2019.)

Tekoälyalgoritmin virheellisyys voi johtua myös aineiston valikoivuusharhasta. Tällöin koulutusdatan satunnaistaminen ei ole onnistunut ja käytetty otos ei ole suhteessa populaatioon. Tämä riski on tunnistettavissa etenkin vapaamuotoisissa aineistoissa: kuva-aineistoissa, elokuvissa, audiolähteissä, sosiaalisen median lähteissä sekä historiallisissa aineistolähteissä, joista esimerkiksi vähemmistöryhmät puuttuvat kokonaan tai ne esitetään stereotyyppisesti valtaväestön näkökulmista. Myöskään niin sanottu rakenteellinen data, eli tietokantoihin määrättyssä muodossa ja järjestyksessä tallennettava data, ei ole vapaa aineistoon liittyvistä harhoista. Lisäksi aineistojen tallentaminen ja säilyttäminen on liiketoimintaa, joka mahdollisesti rajaa ihmisten ja ihmisryhmien näkyvyyttä tekoälykoulutuksessa käytettävissä olevasta aineistosta. (Shahbazi, Yin, Asudeh, & Jagadish 2023, 5–8.)

5.2.2 Sosiaalialalle suunnattujen tekoälyjärjestelmien ja sovellusten eettinen käyttö

Tekoälyn kehittyessä yhä itsenäisemmäksi, tutkijoiden eettinen huoli liittyy sen vaikutuksiin ihmisten autonomisuuteen ja itsemääräämisoikeuteen. Tekoälykehityksessä on tärkeä huomioida, että ihmisiä ei kohdella pelkästään tietolähteenä. Ihmisten itsemääräämisoikeus ymmärretään useimmissa demokratioissa moraalisenä periaatteena ja sitä tulee kunnioittaa kaikissa tilanteissa. Tekoälyä ei saa missään tilanteessa käyttää ilman sen vaikutuspiirissä olevien ihmisten tietoisuutta olemassa olevasta

teknologiasta, sillä tällaisissa tilanteissa tekoälyn käyttö voidaan tulkita laittomaksi valvonnaksi ja kontrolloinniksi. (Formosa 2021, 600.)

Autonomisuuteen ja itsemääräämisoikeuteen liittyvien huolien lisäksi tekoölyyn nivoutuva eettinen keskustelu kulminoituu sen mahdolliseen manipulatiiviseen ja sortavaan käyttöön etenkin heikommassa asemassa olevien väestöryhmiä kohtaan. Tutkimuksissa on tunnistettu tekoölyä hyödyntävien virtuaaliavustajien seksistinen suhteutuminen naisiin, rasistisen suhtautuminen vähemmistöihin (Formosa 2021, 609.) sekä syrjäytymisuhan alla oleviin nuoriin kohdistuvaa epäeettistä markkinointia (Formosa 2021, 609; Reamer 2023, 62).

Tutkimuksissa on tunnistettu myös ihmisten taipumus liittää inhimillisiä piirteitä elottomiin olentoihin. Tämä antropomorfismina tunnettu ilmiö on sitä todennäköisempi mitä enemmän esimerkiksi tekoölyavusteinen robotti muistuttaa ihmistä olemukseltaan ja ääneltään. Ilmiöistä voi olla hyötyä yksinäisyyden ehkäisemisessä, toimintakyvyn vahvistamisessa ja ihmisten autonomisuuden tukemisessa, etenkin jos robottien avulla pystytään tukemaan kotona pärjäämistä. Lisäksi vuorovaikutteisilla roboteilla voidaan vahvistaa ihmisten oikeudenmukaisuutta ja käyttää niitä rikollisuuden ehkäisemisessä esimerkiksi vahvistamalla empatian kehittymistä ja eettistä harkintaa vanheinhoidossa. (Formosa 2021, 606.)

6 JOHTOPÄÄTÖKSET JA POHDINTA

Tekoälyllä on potentiaalia muuttaa sosiaalityötä ja tehostaa asiakkaille, organisaatioille ja yhteisöille tarjottavia palveluita. Tekoäly mahdollistaa nopean datan analysoinnin tavalla, jota voidaan sosiaalityössä käyttää riskiarvioinnissa, interventioiden arvioinnin tukena sekä palvelujärjestelmän systemaattisten vinoumien tunnistamiseksi ja korjaamiseksi (vrt. Reamer 2023, 53). Päätöksenteon ja palveluiden parantamisen lisäksi tekoälyn avulla voidaan edistää sosiaalialan ammattilaisten työssä jaksamista ja tunnistaa loppuun palaamista ennakoivia riskitekijöitä. Lisäksi keskustelevala tekoäly sekä vuorovaikutteiset robotit voivat toimia asiakkaan kuntoutusprosessin tukena.

Uuden teknologian kehittämiseen liittyvät eettiset kysymykset tulevat selvimmin esille teknologioissa, jotka mahdollistavat käyttäjien tunnistamisen ja valvonnan. Esimerkiksi mobiililaitteet, langattomat verkot ja hajautetut tietojärjestelmät pystyvät aktiivisesti seuraamaan yksilön toimintaa ja reagoimaan siihen ilman suoraa käyttäjän vuorovaikutusta. Verkkojen ja palveluiden käytöstä kertyy runsaasti arkaluonteisia tunnistetietoja ja henkilökohtaisia tietoja, mikä korostaa yksityisyyden suojan tärkeyttä. Tekoälyjärjestelmiä voidaankin käyttää tehokkaasti syrjinnän ja kulttuurisen sarron välineenä, mikäli kehityksessä ei huomioida monipuolisesti eri ihmisryhmien tarpeita ja lähtökohtia. (Leikas 2008, 76.)

Eettinen ja oikeudenmukaisuuteen tähtäävä tekoälykehitys vaatii yhteistyötä lainsäätäjien, sovelluskehittäjien, loppukäyttäjien ja esimerkiksi sote-alan ammattilaisen välillä, jotka toimivat sekä tekoälyratkaisuiden käyttäjinä että asiantuntijoina asiakkaiden suuntaan. Jopa hyvin suunnitellut toimet voivat aiheuttaa laajalle ulottuvia, ei-toivottuja ekologisia, sosiaalisia, poliittisia ja kulttuurisia vaikutuksia. (Hallamaa & Kalliokoski 2022, 9.) Vaikka tekoälyratkaisut voidaan toteuttaa teknisesti eri tavoin, tekoälyn toiminta perustuu tietoon, ja tuon tiedon lähde on aina ihminen. Tekoälyn käyttämän datan vääristymät sekä algoritmin vinoumat johtuvat siis inhimillisestä toiminnasta ja näitä vääristymiä voidaan ehkäistä lisäämällä uuden teknologian kehitykseen ja käyttöönottoon eettistä pohdintaa. Sosiaalisten ongelmien ratkaisuun tähtäävässä tekoälyn koulutuksessa ja sovelluskehityksessä on välttämätöntä, että alan ammattilaiset toimivat tiiviissä yhteistyössä kehittäjien kanssa. Samalla pitää pystyä varmistamaan, että tekoälyjärjestelmät noudattavat alan lakeja, standardeja ja eettisiä normeja. Tämä on myös uusi osaamisvaade sosiaalialan ammattilaisille, sillä onnistunut yhteistyö edellyttää ainakin jossain määrin ymmärrystä kehitysmalleista, algoritmeista ja tekoälyn koulutuksessa käytettävän datan muodostumisesta. (Formosa 2021, 611.)

Kokonaisuudessaan tekoälypohjaisista järjestelmistä sosiaalialalla on julkaistu vain vähän tutkimusta. Tutkimuskohteen kypsymättömyydestä kertoo se, että julkaistussa tutkimuksissa on suurta laadullista vaihtelevuutta ja tutkimuksia julkaistaan puutteellisin tiedoin. Tämä aiheuttaa vaikeuksia tekoälyjärjestelmien jatkokehittämiseksi sekä järjestelmien toiminnallisuuden arvioimiselle. Jos esimerkiksi sovellusten kehittämisessä ja päätöksenteossa käytettyä lähestymistapaa ei ole kerrottu, päätöksentekojärjestelmän luotettavuutta ja läpinäkyvyyttä on vaikea arvioida. Tekoälykehityksessä kaksi yleisimmin käytettyä lähestymistapaa ovat tietoon perustuva (knowledge-driven) ja dataan perustuva (data-driven) kehitysmalli. Näistä ensimmäinen nojaa vahvasti alan asiantuntijoiden tietoon ja toimialaosamiseen ja jälkimmäinen asiakkailta kerättyyn suureen tietomäärään ja sitä vasten rakennettuihin koneoppimismalleihin. (Cresswell, Callaghan, Khan, Sheikh, Mozaffar, Sheikh

2020.) Lähestymistavasta riippumatta toimialatuntemus sekä sosiaalialan ammattilaisten mukaan ottaminen päätöksentekomallien ja sovellusten kehitykseen on välttämätöntä, jotta tekoälykehityksessä tulee riittävästi huomioitu sosiaalialan toimintaympäristön erityispiirteet.

6.1 Tutkielman luotettavuus ja eettiset näkökulmat

Laadullisen tutkimuksen luotettavuuden arvioinnista on moninaisten tutkimusperinteitten takia useampia näkemyksiä, mutta toisaalta tiedekentällä vallitsee hyvinkin yksimielinen näkemys siitä, että laadullisen tutkimuksen luotettavuuteen on kiinnitettävä huomiota tutkimuksissa. Yleisimmin tutkimuksen luotettavuuden arviointiin kuuluu tutkimuksen eettinen arvio, reliabiliteetin eli toistettavuuden ja validiteetin eli pätevyyden arviointi. (Saaranen-Kauppinen & Puusniekka 2009, 24–25.) Laadullinen tutkimus on aina jossain määrin tekijänsä näköinen, joten toistettavuuteen ei laadullisissa tutkimuksissa aina pyritä. Toistettavuuteen vaikuttaa eri tutkimusperinteiden käsitykset tiedosta ja siitä juontuvat totuustulkinnat, jotka ovat aina jossain määrin riippuvaisia myös tutkijan kokemuksesta ja ennakkokäsityksistä. Vaatimus kohdistuu sen sijaan esimerkiksi aineiston hankinnan läpinäkyvyyteen, analyysivaiheen kuvaamiseen sekä tutkimuksen teoriaperinteestä ja tutkijasta johtuvien riippuvuuksien tunnistamiseen. Tutkija on lukijoille velkaa sen, että he pystyvät seuraamaan tutkimuksen etenemistä tutkimusongelman asettamisesta, tiedonhankinnan ja analyysin kautta johdopäätöksiin saakka. Kuvaamalla nämä vaiheet mahdollisimman tarkasti, niin tutkija kuin lukija voi arvioida tutkimuksen pätevyyttä suhteessa tutkimusongelmaan. (Tuomi & Sarajärvi 2018.) Aineiston valinnan kriteerit on tämän opinnäytteen tutkimusraportissa kuvattu tarkasti mukaan lukien hakufraasit, sisäänottokriteerit sekä poikkeamat. Lisäksi myös analyysivaihetta on avattu lukijalle aineistonäytteiden avulla. Kategoriat ja niihin mahdollisesti liittyvät alaluokat on lisäksi kerrottu liitteessä 1. kunkin aineistoartikkelin kohdalla.

Vaikka kirjallisuuskatsaus perustuu kokonaan aikaisemmalle kirjallisuudelle, se ei kuitenkaan ole vapaa eettisestä arvioinnista. Menetelmän väljyys päinvastoin pakottaa tutkijan perustelemaan ja pohtimaan kriittisesti tekemiään valintoja. (Kangasniemi, Utriainen, Ahonen, Pietilä, Jääskeläinen & Liikanen 2013, 297.) Kirjallisuuskatsauksessa luotettavuus ja eettisyys kulkevat rinnakkain, jolloin tutkimusprosessin mahdollisimman tarkka kuvaaminen parantavat samalla molempia. Erityisesti kuvailvan kirjallisuuskatsauksen eettiset kysymykset ja luotettavuuden arviointi liittyvät tutkimuskysymyksen muotoiluun sekä itse tutkimusprosessin eettisyyteen ja läpinäkyvyyteen. Tutkimuskysymyksen tulee olla riittävän täsmällinen, jotta siihen voi löytää vastauksen, mutta kuitenkin niin väljä, että se mahdollistaa ilmiön monipuolisen tarkastelun. Läpinäkyvyys auttaa tunnistamaa mahdollisia subjektiivisia vinoumia. (Kangasniemi, Utriainen, Ahonen, Pietilä, Jääskeläinen & Liikanen 2013, 297.)

Tutkimuskysymys ohjaa aineiston valintaa ja aineiston valinnassa ja käsittelyssä korostuu tutkimusetiikan noudattaminen. Aineistolla pyritään mahdollisimman suureen kattavuuteen, mutta tuloksiin voi silti vaikuttaa esimerkiksi julkaisuharha tai kieliharha. (Snyder 2019, 334.) Kieliharha liittyy esimerkiksi pelkästään englanninkielisen julkaisun sisällyttämistä aineistoon, mutta käytettävissä olevat

resurssit kuitenkin yleensä pakottavat tekemään rajauksia. Julkaisuharhaan tutkija voi vaikuttaa pyrkimällä valitsemaan aineistoon mukaan vielä julkaisematonta tutkimusta. Tutkimusprosessin kriittinen arviointi tuo tietoisuuteen tutkimusprosessin eri vaiheisiin liittyvät valinnat sekä parantaa tutkimuksella tuotetun tiedon luotettavuutta. (Kangasniemi, Utriainen, Ahonen, Pietilä, Jääskeläinen & Liikanen 2013, 297.) Tämän opinnäytteen aineisto muodostui pelkästään englanninkielisistä, jo julkaistuista tutkimusartikkeleista. Aineiston artikkeleista ainoastaan yksi oli julkaistu maassa, jossa virallinen kieli on muu kuin englanti, joten tätä voidaan pitää selvänä heikkoutena aineiston monipuolisuuden ja edustavuuden näkökulmasta. Toisaalta rajaamalla aineisto tieteellisiin, vertaisarvioituihin artikkeleihin, pyrittiin varmistamaan aineiston laadukkuus vähän tutkitun aiheen ympärillä.

Eettinen reflektio jatkuu tutkimuksen analyysivaiheessa sekä synteesin kirjoittamisessa. Tutkija vastaa, että aineisto tulee huomioitua kattavasti ja sen koko potentiaali hyödynnetään. Kirjallisuuskatsaukselle on kuitenkin tyypillistä, että analyysin edetessä aineisto kehittyy yleisluontoisesta täsmällisempään. Lisäksi analyysin subjektiivisuuden takia eri tutkijat voivat päätyä samasta aineistosta hyvin erilaisiin johtopäätöksiin. (Kangasniemi, Utriainen, Ahonen, Pietilä, Jääskeläinen & Liikanen 2013, 298.) Tutkija on aina vastuussa muulle tutkimusyhteisölle hyvän tieteellisen käytännön edistämisestä, plagioinnin estämisestä sekä tutkimustyön laadun parantamisesta pitkällä aikavälillä. (Snyder 2019, 339). Myös tässä opinnäytteessä noudatettiin koko tutkimusprosessin ajan eettisiä ohjeita ja hyvän tieteellisen käytännön periaatteita, joihin kuuluu rehellisyys, avoimuus, huolellisuus sekä yleinen tarkkuus niin tutkimustyössä, lähdeviittauksissa kuin tuloksien raportoinnissa (Tutkimuseettinen neuvottelukunta 2023).

Laadullisen tutkimuksen moninaisten tutkimusperinteiden takia myös valittavissa olevien tutkimusmenetelmien kirjo on laaja (Tuomi & Sarajärvi 2018). Tutkimusongelmassa pyrittiin kartoittamaan vasta vähän tunnettua ilmiötä, eli kuinka tekoälyä sovelletaan sosiaalialalla ja kuinka eettisyys huomioidaan kehitystyössä. Tähän kartoittavaan tiedontarpeeseen kirjallisuuskatsaus vastaa hyvin. Molempiin tutkimuskysymyksiin olisi voinut lähteä etsimään vastausta myös esimerkiksi erilaisilla haastattelumenetelmillä. Erityisesti tekoälyä hyödyntävien järjestelmien eettisessä kehitystyössä olisi ollut mielenkiintoista haastatella asiantuntijoita sekä sosiaalialan ammattilaisten että sovelluskehittäjien joukosta ja jakaa alan parhaita käytäntöjä asiantuntijoiden kesken tutkimusraportin ja blogijulkaisujen muodossa. Tutkimusta voisi jatkaa edelleen haastatteleamalla valittua joukkoa ammattilaisia, muodostaa näin saadusta aineistoista yksityiskohtainen kuva esimerkiksi tietyn sosiaalialan organisaation tavasta soveltaa tekoälyä eettisesti ja lähteä kartoittamaan tavan yleisyyttä valtakunnallisella kyselylomaketutkimuksella. Koska tekoäly on sosiaalialalla varsin tuore ilmiö, tutkimuskohteita muodostuu jatkossa yhä enemmän. Tutkimuksen avulla voidaan kehittää alan käytäntöjä ja ammattilaisten tietoisuutta tekoälyn eettisestä soveltamisesta ja näin pyrkiä myös ennaltaehkäisemään uuden tekniikan mahdollisia ennakoimattomia vaikutuksia, mikä on myös jokaisen uuden tekniikan kehittäjän ja käyttöönottajän moraalinen vastuu. (Heikkerö 2009, 65–66.)

6.2 Pohdinta

Teoreettisessa viitekehyksessä yhdistyy sosiaalitieteellinen, filosofinen ja teknologinen näkökulma tekoälyn sovellettavuuteen sosiaalialalla, joiden yhdistäminen ei ole ongelmattonta. Sosiaalityön tietoperusta nojaa osallisuuden vahvistamiselle, kun taas tekniikan tarkoitus on ollut lähtökohtaisesti vähentää ihmisen osallisuutta. Esimerkiksi robotteja käytetään tehtaissa korvaamaan ja tehostamaan ihmisen suorittamia työtehtäviä. Teknologinen kehitys onkin pyritty pitämään erillään moraaliiin ja etiikkaan liittyvistä tarkasteluista. (Leikas 2008, 74.) Posthumanistisen aatesuuntauksen mukaan ihmisen kehitys kulkee kuitenkin vääjäämättä kohti tilannetta, jossa yhä enemmän ihmisen orgaanisia osia on korvattavissa keinotekoisilla varaosilla kuten hammasimplanteilla, kuulokojeilla ja proteeseilla. Apuvälineisiin liitettävä älytekniikka liudentaa entisestään koneen ja ihmisen välistä rajaa. Kehitys kulkee myös toiseen suuntaan: ihmisistä koneisiin, jolloin robotteihin yritetään tuoda ihmisen tietoisuutta muistuttavia ominaisuuksia ja erityisesti tämä on tavoitteena supertekoälykehityksessä. (Niiniluoto 2020.) Tekoälymallit ovat lisäksi tulleet yhä kiinteämmäksi osaksi ihmisten arkitodellisuutta, joten sen sijaan että tutkimuksella vahvistetaan ihmistieteiden ja teknologian vastakainasettelua, pitäisi pystyä löytämään keinoja hyödyntää teknologiaa sen alkuperäisessä tarkoituksessa eli ihmisten apuvälineenä. (Hallamaa, Haikarainen & Kalliokoski 2021, 17.)

Tekoälykehityksen näkökulmasta ei kuitenkaan ole enää ongelmattonta palauttaa tarkastelua sen alkujuurille eli tekniikaksi, joka on luotu palvelemaan ihmistä - etenkin, jos tarkastellaan globaalia kehitystä. Tekoälyteknologian arkipäiväistyminen mobiilisovellusten kautta on tuonut sovellukset lähes kaikkien saavutettaviksi, mutta edelleen maailman köyhempien ja samalla väestörikkaimpien alueiden asukkaiden poissaolo lähdedatasta vääristää esimerkiksi kuvantunnistusalgoritmien ope- tusta. (Shahbazi, Yin, Asudeh, & Jagadish 2023; Hallamaa & Kalliokoski 2021, 9.) Ajankohtainen kysymys on, syventääkö tekoäly polarisaatiota edelleen vai tarjoaa meille työkalun eriarvoistumiskehi- tyksen katkaisemiseksi.

Sosiaalialan toimintaympäristö on vahvasti sidoksissa ympäröivään yhteiskunnalliseen todellisuuteen, joten myös uuden tekniikan suunnittelussa ja käyttöönotossa tulee huomioida ympäristön asettamat reunaehdot. Toisaalta tekoälyn yleistyminen muuttaa sosiaalialan toimintaympäristöä myös asiakas- työn näkökulmasta ja vaatii alaa uudistumaan. On tiedossa, että tekoäly on vasta tekemässä tuloaan sosiaalialalle, jonka takia siihen liittyvää tieteellistä tutkimusta on julkaistu vähän. Lisäksi julkaistujen tutkimuksien laadussa on suurta vaihtelua, joka johtuu tekoälyn käsitteen epätarkkuudesta sekä eri käsitteiden, kuten esimerkiksi koneoppiminen ja tekoäly, päällekkäisestä ja ristikkäisestä käytöstä tutkimuksissa. (Cresswell, Callaghan, Khan, Sheikh, Mozaffar & Sheikh 2020, 2145.) Tästä syystä sosiaalialalla ollaan osin tuntemattoman edessä, joten eettisen arvioinnin merkitys on yhä tärkeäm- pää.

Pohjimmiltaan kaiken uuden tekniikan kehitystyössä on kyse arvoista ja arvoalinnoista, jolloin eetti- sessä arvioinnissa pätee samat filosofiset peruskysymykset kuin missä tahansa muussa eettisessä arvioinnissa: mitä ja kenen etua kehityksellä tavoitellaan? Sovellettavien eettisten periaatteiden tai moraaliteorian valitseminen uuden älykkään teknologian kehityksessä ei ole yksinkertaista, mutta

välttämätöntä. Uuden tekniikan kehittäjät tulevat harvoin ajatelleeksi millaisia kauaskantoisia vaikutuksia teknologialla on. Yhteisesti sovittujen kriteereiden tai periaatteiden avulla voidaan tarkastella saavutettujen tulosten ja asetettujen tavoitteiden suhdetta. (vrt. Leikas 2008, 77.)

Toisaalta tekoäly on jo integroitunut kattavasti moniin yhteiskunnallisesti merkittäviin järjestelmiä, kuten Apotti ja Omavero, jolloin sitä on mahdoton irrottaa muusta systeemistä (Hallamaa & Kallio-koski 2022, 7.) Jääkö tällöin sosiaalialan tehtäväksi huolehtia, että myös kaikkein heikommassa asemassa olevat yhteiskunnan jäsenet saavat riittävät tiedot ja taidot tunnistaa ja käyttää tekoälyjärjestelmiä? Tekoälyn vahva integroituminen osaksi arkisessa käytössä olevia laitteita kasvattaa myös sen koulutuksessa käytettävän aineiston määrää, jolloin massadatan merkitys yhteiskunnallisten ongelmien ratkaisemisessa kasvaa edelleen. Tällöin on tärkeä tunnistaa datan vinoumat ja pyrkiä estämään niitä sovelluskehityksessä ja tarkastella datan käyttöä eettiseltä kannalta. Eettisessä arvioinnissa palataan jälleen peruskysymyksiin ääreen, eli mitä hyvää ja kenen etua tekoälykehityksellä tavoitellaan. Sosiaalialan näkökulmasta peruskysymys kuuluu: mikä on hyväksyttävä tapa hankkia tietoa palvelujärjestelmän kehittämiseksi? Onko oikein syöttää yksittäisen ihmisen asiakastietoja koneoppimisjärjestelmään ja siten auttaa tunnistamaan mahdollisia syrjäytymistä ennakoivia riskitekijöitä, tai voiko massadatan avulla luotujen tekoälymallien avulla ylipäätään tehdä yksilötasolla ulottuvaa riskiarviointia ja mahdollisia ennaltaehkäiseviä interventioita. Ratkaisuksi on esitetty toimintamallia, jossa tekoälymalleja sovelletaan yhteisötasolla ja esimerkiksi nuorten syrjäytymistä edistäviä tekijöitä pyritään tunnistamaan koulukohtaisesti. (Talvitie-Lamberg, Silvennoinen, Moilanen & Korpela 2018, 24; myös Leikas 2008, 77.)

Terveydenhuollossa teknologia-avusteisella työllä on pitkä historia. Alan tutkimuksissa on tunnistettu muun muassa robotiikan hyödyntämisestä johtuvia positiivisia vaikutuksia työn tehokkuuden, laadun ja asiakastyytyväisyyden kannalta (Hänninen 2021, 3). Sosiaalialalla teknologia-avusteinen työ ei yhtä tuttua kuin terveydenhuollossa. Virtuaalista työvoimaa hyödynnetään vielä vähän, joten sosiaalialalle ei ole syntynyt kovin laajaa käyttökokemusta teknologian soveltamisesta työn tukena. Uuden teknologian laajamittaisen käytön esteeksi onkin tunnistettu tiedon puute sekä tekniikkaan liittyvät ennakkoluulot. (Emt. 42.) Toisaalta sosiaaliala pystyy myös oppimaan terveydenhuollon digitalisatiosta ja sitä kautta ennakoimaan käyttöönottoon liittyviä riskejä ja suunnittelemaan toimintatapoihin vaadittavia muutoksia. Muutosvastarintaa voi loiventaa järjestämällä ammattilaisille riittävästi koulutusta, osallistamalla heitä käyttöönottoon ja varmistamalla riittävästi lähitukea paitsi uuden teknologian käyttöönoton ajalla myös koko järjestelmän elinkaaren ajalle. (Vehko, Hyppönen, Ryhänen-Tompuri, Heponiemi, 2019, 12–13.)

Uuden teknologian käyttöönotossa sosiaalialalla on myös muistettava asiakkaat, joiden on tärkeä tietää missä kaikessa teknologiaa käytetään. Etenkin tekoälyn käytön leviäminen sosiaalialalle on samalla myös tärkeä muistutus sosiaalialan ammattietiikan tarpeellisuudesta digitaalisten työkalujen kehitystyössä ja tämä työ vaatii alan ammattilaisilta ennakkoluulotonta ja ennakoivaa työtettä, jotta voidaan varmistaa asiakkaiden eettinen kohtelu joka tilanteessa. (vrt. Reamer 2023, 66.)

Tämän opinnäytteen tarkoitus oli syventää asiantuntemusta sosiaalialalla käytettävistä tekoälypohjaisista ratkaisuista sekä niiden käyttöön liittyvistä eettisistä erityiskysymyksistä. Tekoälykehitys on tällä hetkellä varsin vauhdikasta, joten myös uutta tutkimusta ja samalla jatkotutkimusaihetta syntyy

lähes viikoittain. Esimerkiksi tekoälyn soveltamiseen liittyy aina merkittäviä tietoturvakysymyksiä, joita nopea kehitys jatkuvasti myös haastaa. Ne jäivät tässä työssä tarkastelematta, mutta olisivat merkittävä jatkotutkimusaihe. Viimeisin suomalaissaavutus tekoälykehityksessä on Turun yliopiston ja Silo AI:n yhteistyössä kehittämä ”Poro” tekoäly. Se on avoimen lähdekoodin kielimalli pienille kielille, joille on saatavilla vähän koulutusdataa. Kehitystyötä on ohjannut eurooppalaiset monimuotoiset ja demokraattiset arvot ja niiden ylläpitäminen. (Turun yliopisto 2024.) Arvojen kirjaaminen osaksi teknologista kehitystä tuo näkyvämmäksi myös insinööritieteisiin kohdistuvaa vaatimusta yksilöllisen hyvän edistämisestä kohti yhteisöllistä hyvää (Heikkerö 2009, 96) ja samalla se puoltaa myös monitieteistä tutkimusotetta sosiaalisten ongelmien ratkaisemiseksi viimeisimpiä teknologisia innovaatioita hyväksikäyttäen.

Monitieteisyys oli kuitenkin myös tämän opinnäytteen heikkous. Sosiaalityön, tekniikan ja filosofian tietoperusta on laaja, jotta siitä saisi puristettua opinnäytteeseen soveltuvan, mutta riittävän perusteellisen teoreettisen perustan. Eri tieteenalojen yhteinen perusta löytyi tiedosta, mutta samalla valinta ohjasi painotusta sosiaalityön asiakassuhteessa syntyvän sensitiivisen tiedon muodostumiseen ja sen hyödyntämiseen liittyviin filosofisiin tarkasteluihin. Tekoälykehitykseen käytettyjen teknisten ratkaisuiden tarkastelu jäi vähäisemmäksi, eikä niitä käsitelty yksityiskohtaisesti myöskään aineistoon valituissa artikkeleissa. Tutkimusongelmaan tutkielma antoi riittävästi ja kattavasti vastausta ja lisäksi laajensi käsitystä asiakastyössä hyödynnettävistä tekoälypohjaisista ratkaisuista myös sosiaalialan ammattilaisten työssäjaksamisen tukemisessa hyödynnettäviin sovelluksiin. Tämä opinnäyteprosessi alkoi vuoropuhelulla ChatGPT:n kanssa. Vaikka vuoropuheluista jäi vähän jäljelle lopulliseen toteutukseen, valkoisen paperin ongelmaan ja ideointiin niistä oli apua. Tästä syystä on perusteltua päättää tämä tutkielmaprosessi ChatGPT:n muotoilemaan sitaattiin, joka sopii muistisäänöksi myös sosiaalialalla hyödynnettävien tekoälypohjaisten innovaatioiden kehitykselle:

"Tasapainossa tekoälyn voima, sosiaalityön sydän ja eettinen kompassi."

LÄHTEET

Työssä on käytetty seuraavasti tekoälyä:

ChatGPT 2023. OpenAI. GPT-3.5. Käytetty ideoimaan blogitekstejä, lokakuu 2023. <https://chat.openai.com>

OpenArtAI 2023. OpenAI. Käytetty blogin kuvitukseen, lokakuu-maaliskuu 2023. <https://openart.ai/home>

Aveyard, Helen. 2018. *Doing a Literature Review in Health and Social Care: A Practical Guide*. 4th ed. Maidenhead: McGraw-Hill Education.

Bako, Abdulaziz Tijjani, Taylor, Heather L., Wiley, Kevin, Zheng, Jiaping, Walter-McCabe, Heather Kasthurirathne, Suranga N & Vest, Joshua R. 2021. Using Natural Language Processing to Classify Social Work Interventions. *The American Journal of Managed Care* 27 (1): E24–E31. <https://doi.org/10.37765/ajmc.2021.88580>.

Bankins, Sarah, Formosa, Paul, Griep, Yannick, & Richards, Deborah 2022. AI Decision Making with Dignity? Contrasting Workers' Justice Perceptions of Human and AI Decision Making in a Human Resource Management Context. *Information Systems Frontiers*, 24(3), 857–875. <https://doi.org/10.1007/s10796-021-10223-8>

Boucher, Philip 2020. Artificial intelligence: How does it work, why does it matter, and what can we do about it? European Union. [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641547/EPRS_STU\(2020\)641547_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641547/EPRS_STU(2020)641547_EN.pdf) Viitattu 25.10.2023

Cain, Geoffrey 2022. *Totaalinen poliisivaltio*. Jyväskylä: Atena.

Charisi, Vicky, Dennis, Loïuse, Fisher, Michael, Lieck, Robert, Matthias, Andreas, Slavkovik, Marija, Sombetzki, Janina, Winfield, Alan F. T., & Yampolskiy, Roman 2017. Towards Moral Autonomous Systems. arXiv.Org. <https://doi.org/10.48550/arxiv.1703.04741>

Cooper, Harris 1998. *Synthesizing Research: A Guide for Literature Reviews*, 3rd ed. Thousand Oaks, CA: Sage.

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

Dignum, Virginia 2017. Responsible Autonomy. ArXiv.Org. <https://doi.org/10.48550/arxiv.1706.02513>. Viitattu 12.12.2023

EOAK 31.12.2019 nro 4110/2018 Kuvaaminen ja äänittäminen lastensairaalassa. <https://www.oikeusiamies.fi/r/fi/ratkaisut/-/eoar/4110/2018> Viitattu 6.12.2023

Euroopan komissio 2019. Viestintäverkkojen, sisältöjen ja teknologian pääosasto. Luotettavaa tekoälyä koskevat eettiset ohjeet. Julkaisutoimisto. <https://data.europa.eu/doi/10.2759/068692>

Euroopan parlamentti 2023. Verkkojulkaisu. MEPs ready to negotiate first-ever rules for safe and transparent AI. <https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai> Viitattu 6.12.2023

Euroopan parlamentti 2023. Mitä tekoäly on ja mihin sitä käytetään? https://www.europarl.europa.eu/pdfs/news/expert/2020/9/story/20200827STO85804/20200827STO85804_fi.pdf Viitattu 23.10.2023

Formosa, Paul 2021. Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy. *Minds and Machines (Dordrecht)* 31 (4): 595–616. <https://doi.org/10.1007/s11023-021-09579-2>.

Friedman, Batya 1996. Value-Sensitive Design. <https://dl.acm.org/doi/pdf/10.1145/242485.242493>

Gillingham, Philip & Graham, Timothy 2017. Big Data in Social Welfare: The Development of a Critical Perspective on Social Work's Latest Electronic Turn. *Australian Social Work* 70 (2): 135–47. <https://doi.org/10.1080/0312407X.2015.1134606>.

Gillingham, Philip 2021. Algorithmically Based Decision Support Tools: Skeptical Thinking about the Inclusion of Previous Involvement. *Practice (Birmingham, England)* 33 (1): 37–50. <https://doi.org/10.1080/09503153.2020.1749584>.

Grant Maria J & Booth Andrew 2009. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J.*26(2), 91-108. doi: 10.1111/j.1471-1842.2009.00848.x. PMID: 19490148. Viitattu 8.12.2023

Grzadzielewska, Malgorzata. 2021. Using Machine Learning in Burnout Prediction: A Survey. *Child & Adolescent Social Work Journal* 38 (2): 175–80. <https://doi.org/10.1007/s10560-020-00733-w>.

Gunkel, David J. 2020. Perspectives on Ethics of AI: Philosophy. Teoksessa: Markus D. Dubber, Frank Pasquale, and Sunit Das (toim.). *The Oxford Handbook of Ethics of AI*. Oxford Academic, <https://doi-org.ezproxy.uef.fi:2443/10.1093/oxfordhb/9780190067397.013.35>.

Günther, Wendy Arianne, Rezazade, Mohammad H., Huysman, Mehrizi Marleen & Feldberg, Frans 2017. Debating Big Data: A Literature Review on Realizing Value from Big Data. *The Journal of Strategic Information Systems* 26 (3), 191–209. <https://doi.org/10.1016/j.jsis.2017.07.003>.

Hallamaa, Jaana, Haikarainen, Raine & Kalliokoski, Taina 2021. Tekoälyteknologiat ja turvallisuus. *Tieteessä tapahtuu*, 39(4), 11–17. <https://journal.fi/tt/article/view/111214> Viitattu 24.10.2023

Hallamaa, Jaana & Kalliokoski Taina 2022. AI Ethics as Applied Ethics. *Frontiers in Computer Science (Lausanne)* 4. <https://doi.org/10.3389/fcomp.2022.776837>.

Halsband, Aurélie 2022. Sustainable AI and Intergenerational Justice. *Sustainability (Basel, Switzerland)* 14(7), 3922-.

- Heikkerö, Topi 2009. Tekniikka ja etiikka: Johdatus teoriaan ja käytäntöön. Tekniikan akateemisten liitto.
- Herne, Kaisa 2015. Mitä oikeudenmukaisuus on? Helsinki: Gaudeamus.
- Hodgson, David, Goldingay, Sophie, Boddy, Jennifer, Nipperess, Sharlene & Watts, Lynelle 2022. Problematising Artificial Intelligence in Social Work Education: Challenges, Issues and Possibilities, *The British Journal of Social Work*, Volume 52, Issue 4, June 2022, Pages 1878–1895, <https://doi.org/10.1093/bjsw/bcab168>
- Helsingin yliopisto, julkaisuaika tuntematon. Blogikirjoitus. Tiedeviestintä. Tutkijablogin pitäminen. <https://blogs.helsinki.fi/tiedeviestinta/verkko/tutkijablogin-pitaminen/> Viitattu 26.10.2023
- Hänninen, Pasi 2021. Robotiikka sosiaali- ja terveydenhuollon tukena. Informaatioteknologian tiedekunnan julkaisuja no. 90/2021. Jyväskylän yliopisto. https://jyx.jyu.fi/bitstream/handle/123456789/77700/1/90-2021_Robotiikka%2520sosiaali-%2520ja%2520terveydenhoi-%2520don%2520tukena_VERKKO.pdf Viitattu 2.3.2024
- Jarreau, Paige Brown 2015. All the Science That Is Fit to Blog: An Analysis of Science Blogging Practices. LSU Doctoral Dissertations. 1051. https://repository.lsu.edu/gradschool_dissertations/1051 Viitattu 26.10.2023
- Kak, Amba 2020. the Global South is everywhere, but also always somewhere. National policy narratives & AI Justice. AIES 2020 - Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 307–312. <https://doi.org/10.1145/3375627.3375859>
- Karjalainen, Pekka 2016. Sosiaalityön uusi asento: paikka, rooli ja orientaatio. Teoksessa: Blomgren, Sanna, Karjalainen, Jouko, Karjalainen, Pekka, Kivipelto, Minna, Saikkonen, Paula & Saikku, Peppi (toim.) Sosiaalityö, palvelut ja etuudet muutoksessa. Tampere: Juvenes Print – Suomen yliopistopaino.
- Karjalainen, Pekka & Sarvimäki, Pirjo 2005. Sosiaalityö hyvinvointipolitiikan välineenä 2015-toimintapideohjelma. Sosiaali- ja terveysministeriön julkaisuja 2005:13. Yliopistopaino, Helsinki 2005. Osoitteessa https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/72019/julkaisu_2005_13_sosiaalityo_2015.pdf?sequence=1&isAllowed=y Viitattu 29.2.2024
- Kivipelto Minna & Saikkonen Paula, 2013. Tiedon tuotanto ja vaikuttavuustieto sosiaalityössä. *Yhteiskuntapolitiikka* 78:3, 313–321
- Kivistö, Mari 2017. Sosiaalityö digitalisaatiossa ja eettisyyden vaade. Teoksessa: Kivistö, Mari & Päykkönen, Kirsi (toim.). Sosiaalityö digitalisaatiossa. Lapin yliopiston yhteiskuntatieteellisiä julkaisuja C. Työpapereita 58. Rovaniemi: Lapin yliopistopaino. 21–42
- Kraus, Sascha, Breier, Matthias, Lim, Weng Marc, Dabić, Marina, Kumar, Satish, Kanbach, Dominik, Mukherjee, Debmalya, Corvello, Vincenzo, Piñeiro-Chousa, Juan, Liguori, Eric; Palacios-Marqués, Daniel, Schiavone, Francesco, Ferraris, Alberto, Fernandes, Cristina, Ferreira, João J 2022. Literature Reviews as Independent Studies: Guidelines for Academic Practice. *Review of Managerial Science* 16 (8), 2577–95. <https://doi.org/10.1007/s11846-022-00588-8>. Viitattu 7.12.2023

- Kumar, Shailendra & Choudhury, Sanghamitra 2023. Normative ethics, human rights, and artificial intelligence. *AI Ethics* 3, 441–450. <https://doi.org/10.1007/s43681-022-00170-8> Viitattu 13.11.2023
- Laakkonen, Mika & Ahvenainen, Sakari 2021. Informaatioteknologian filosofia, etiikka ja digitalisoitunut yhteiskunta. Jyväskylän yliopisto.
- Lauhanen, Risto & Laasasenaho, Kari 2021. Tiedettä, politiikkaa vai tiedepolitiikkaa? *SeAMK asiantuntijablogi* 17.5.2021. <https://lehti.seamk.fi/asiantuntijablogi/tiedetta-politiikkaa-vai-tiedepolitiikkaa/> Viitattu 12.11.2023
- Litchfield, Ian, Glasby, Jon, Parkinson, Sarah, Hocking, Lucy, Tanner, Denise, Roe, Bridget & r Bousfield, Jennife. 2023. 'Trying to Find People to Fit the Tech...': A Qualitative Exploration of the Lessons Learnt Introducing Artificial Intelligence-Based Technology into English Social Care. *Health & Social Care in the Community* 2023: 1–11. <https://doi.org/10.1155/2023/9174873>.
- Niiniluoto, Ilkka 2020. Tekniikan filosofia. Helsinki: Gaudeamus.
- Pan, Ian, Nolan, Laura B, Brown, Rashida R, Khan, Romana, van der Boor, Paul, Harris, Daniel G & Ghani, Rayid. 2017. Machine Learning for Social Services: A Study of Prenatal Case Management in Illinois. *American Journal of Public Health* 107 (6): 938–44. <https://doi.org/10.2105/AJPH.2017.303711>.
- Parviainen, Jaana & Rantala, Juho 2020. Ennakoiva analytiikka ja tekoälyn etiikka: Miten ennakoivat teknologiat taipuvat hallintajärjestelmäksi? *Futura* 39(1), 61–70.
- Puusa, Anu 2020. Näkökulmia laadullisen aineiston analysointiin. Puusa, Anu & Juuti, Pauli (toim.). *Laadullisen tutkimuksen näkökulmat ja menetelmät*. Helsinki: Gaudeamus.
- Puusa, Anu & Juuti, Pauli 2020. Laadullisen aineiston analysointi. Teoksessa Puusa, Anu & Juuti, Pauli (toim.). *Laadullisen tutkimuksen näkökulmat ja menetelmät*. Helsinki: Gaudeamus.
- Raitakari, Suvi & Takala, Henna 2023. Asunnottomuutta kokeneet aikuiset ja asunnottomuustyö COVID-19-tutkimuksen kohteena. Teoksessa Raitakari, Suvi, Harrikari, Timo, Hekkala, Marjo & Kivipelto, Minna (toim.) *Sosiaalityö kriiseissä ja poikkeuksellisissa olosuhteissa. Sosiaalityön tutkimuksen vuosikirja 2023*. Jyväskylä: Sophi. 301–332.
- Reamer, Frederic G. 2023. Artificial Intelligence in Social Work: Emerging Ethical Issues. *International Journal of Social Work Values and Ethics* 20 (2): 52–71. <https://doi.org/10.55521/10-020-205>.
- Saaranen-Kauppinen, Anita & Puusniekka, Anna 2009. Menetelmäopetuksen tietovaranto. Kvali-MOTV kvalitatiivisten menetelmien verkko-oppikirja. Yhteiskuntatieteellinen tietoarkisto. Tampereen yliopisto. <https://www.fsd.tuni.fi/fi/tietoarkisto/julkaisut/kvalimotv.pdf> Viitattu 26.10.2023
- Saikkonen, Paula 2016. Kuntakentän kirjavat käytännöt. Teoksessa: Blomgren, Sanna, Karjalainen, Jouko, Karjalainen, Pekka, Kivipelto, Minna, Saikkonen, Paula & Saikku, Peppi (toim.) *Sosiaalityö, palvelut ja etuudet muutoksessa*. Tampere: Juvenes Print – Suomen yliopistopaino. 29–40
- Saini, Michael & Shlonsky, Aron 2012. *Systematic synthesis of qualitative research*. New York, NY: Oxford University Press.

- Salminen, Ari 2011. Mikä kirjallisuuskatsaus? Johdatus kirjallisuuskatsauksen tyyppeihin ja hallintotieteellisiin sovelluksiin. Vaasan yliopiston julkaisuja. Opetusjulkaisuja 62. Julkisohtaminen 4. https://www.uwasa.fi/materiaali/pdf/isbn_978-952-476-349-3.pdf Viitattu 26.10.2023
- Snyder, Hannah 2019. Literature Review as a Research Methodology: An Overview and Guidelines. *Journal of Business Research* 104, 333–39. <https://doi.org/10.1016/j.jbusres.2019.07.039>. Viitattu 7.12.2023
- Sosnet julkaisuaika tuntematon. Sosiaalityön määritelmä. Verkkojulkaisu. <https://www.sosnet.fi/FI/Sosnet/Sosiaalityon-maaritelma> Viitattu 25.10.2023
- Spaulding, Norman W. 2020. Is Human Judgment Necessary? Artificial Intelligence, Algorithmic Governance, and the Law. Teoksessa: Markus D. Dubber, Frank Pasquale, and Sunit Das (toim.), *The Oxford Handbook of Ethics of AI*, Oxford Academic. <https://doi-org.ezproxy.uef.fi:2443/10.1093/oxfordhb/9780190067397.013.25>. Viitattu 19.2.2024.
- STT Info 7.6.2018. Espoon kokeilu todisti: Tekoäly tunnistaa tukea tarvitsevia. Verkkojulkaisu. <https://www.sttinfo.fi/tiedote/68401520/espoon-kokeilu-todisti-tekoaly-tunnistaa-tukea-tarvitsevia?publisherId=3385> Viitattu 7.12.2023
- Talvitie-Lamberg, Karoliina, Silvennoinen, Minna, Moilanen, Hannu & Korpela, Jari 2018. Tekoäly sosiaalisen syrjäytymisen ennakoivassa tunnistamisessa - Case THL kouluterveyskysely. Jyväskylän yliopisto. Informaatioteknologian tiedekunnan julkaisuja. No. 62/2018.
- Terveyden ja hyvinvoinnin laitos julkaisuaika tuntematon. Verkkojulkaisu. Hyvinvoinnin tekoäly ja robotiikka -ohjelma (Hyteairo). <https://thl.fi/fi/tutkimus-ja-kehittaminen/tutkimukset-ja-hankkeet/hyvinvoinnin-tekoaly-ja-robotiikka-ohjelma-hyteairo-> Viitattu 12.11.2023
- Tieteen termipankki julkaisuaika tuntematon. Verkkojulkaisu. <https://tieteentermipankki.fi/wiki/Nimitys: oikeudenmukaisuus>. Viitattu 24.2.2024
- Torraco, Richard J. 2005. Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review* 4 (3): 356–67. <https://doi.org/10.1177/1534484305278283>. Viitattu 7.12.2023
- Tuomi, Jouni & Sarajärvi, Anneli 2018. Laadullinen tutkimus ja sisällönanalyysi. Uudistettu laitos. Helsinki: Kustannusosakeyhtiö Tammi.
- Tutkimuseettinen neuvottelukunta 2023. Hyvä tieteellinen käytäntö (HTK). Verkkojulkaisu. <https://tenk.fi/fi/tiedevilppi/hyva-tieteellinen-kaytanta-htk> Viitattu 8.12.2023
- Turun yliopisto 2024. Euroopan avoin kielimalli Poro: eurooppalaisen tekoälyn ja kielten monimuotoisuuden virstanpylväs. Verkkojulkaisu. <https://www.utu.fi/fi/ajankohtaista/uutinen/euroopan-avoin-kielimalli-poro-eurooppalaisen-tekoalyn-ja-kielten> Viitattu 2.3.2024
- Umbrello, Steven & van de Poel, Ibo 2021. Mapping value sensitive design onto AI for social good principles. *AI and Ethics* 1, 283–296. <https://doi.org/10.1007/s43681-021-00038-3> Viitattu 12.12.2023

Valtioneuvosto 2023. Vahva ja välittävä Suomi. Pääministeri Petteri Orpon hallituksen ohjelma. Verkkojulkaisu. <https://valtioneuvosto.fi/hallitukset/hallitusohjelma#/> Viitattu 12.11.2023

Vehko, Tuulikki, Hyppönen, Hannele, Ryhänen-Tompuri, Miia & Heponiemi, Tarja 2019. Miten tietojärjestelmät palvelevat terveydenhuollon ammattilaisten työtä? Vaikutukset työhön ja työhyvinvointiin. Digitsyö ja stressi -hanke, loppuraportti. <http://urn.fi/URN:ISBN:978-952-343-279-6> Viitattu 2.3.2024

Vilkka, Hanna 2020. Akateemisen lukemisen ja kirjoittamisen opas. PS-kustannus.

Wallin, Riikka 2021. Hiiltä ja timanttia. Opittajat pedagogiikan rapapinnoilla. 7 askelta onnistuneeseen asiantuntijablogikirjoitukseen. Metropolia-blogi 17.9.2021. <https://blogit.metropolia.fi/hiilta-ja-timanttia/2021/09/17/7-askelta-onnistuneeseen-asiatuntijablogikirjoitukseen/> Viitattu 12.11.2023

Zetino, Jeffrey & Mendoza, Natasha 2019. Big Data and Its Utility in Social Work: Learning from the Big Data Revolution in Business and Healthcare. *Social Work in Public Health* 34 (5): 409–17. <https://doi.org/10.1080/19371918.2019.1614508>.

LIITE 1. ARTIKKELILUETTELO

Artikkelit 1. tutkimuskysymykseen *Miten tekoälyä jo hyödynnetään sosiaaliilla?*

Tutkimuksen tekijä(t) Julkaisuvuosi Julkaisu Julkaisumaa	Tutkimuksen nimi	Tutkimusteh- tävä/ tutkimusongel- mat	Aineisto ja menetelmät	Keskeiset tulok- set	Yhdistävä luokka ja mahdollinen täydentävä pää- luokka
<p>Pan, I., Nolan, L. B., Brown, R. R., Khan, R., Van Der Boor, P., Harris, D. G., & Ghani, R.</p> <p>2017</p> <p><i>American Journal of Public Health (1971), 107(6), 938–944.</i> https://doi.org/10.2105/AJPH.2017.303711</p> <p>Yhdysvallat</p>	<p>Machine learning for social services: A study of prenatal case management in Illinois.</p>	<p>Arvioida koneoppimisalgoritmien hyödynnettävyyttä raskauden aikaisten riskien tunnistamiseksi käytössä olevan paperipohjaisen riskiarvioinnin tehostamiseksi.</p> <p>Varhaisen tunnistamisen avulla pyritään sosiaalipalveluiden parempaan kohdentamiseen synnytyksen jälkeen.</p>	<p>Tilastollinen analyysi.</p> <p>Koneoppimismallin kehittämiseen käytettiin Illinoisin osavaltion hallinnollista dataa 6457 naisesta heinäkuusta 2014 toukokuuhun 2015. Data sisälsi tietoja äidin raskaudesta, tapaamisista sosiaalityöntekijöiden kanssa sekä synnytyksestä. Dataan yhdistettiin ammattilaisten arvioita ja päätöksiä etuuksista sekä erilaisiin tukiohjelmiin osallistumisesta. Lopulliseen tietojoukkoon pääsemiseksi yksittäinen synnytys oli yhdistettävä äidin tunnistamiseen, vastasyntyneen tunnistamiseen, raskausajan seurantataulukon, syntymärekisteriin sekä synnytyksen jälkeisiin seurantatietoihin. Koneoppimismallin suorituskyky arviointiin Pythonin k-kertaisella (k-Fold) ristiin validointimallilla.</p>	<p>Koneoppimisalgoritmit suoriutuivat samankaltaisesti tai ylittivät paperipohjaisen riskinarvioinnin 36 %:ssa tutkimustapauksissa.</p>	<p>Ennaltaehkäisy</p> <p>Sosiaalisten ongelmien tunnistaminen ja niihin puuttuminen</p>
<p>Cresswell K, Callaghan M, Khan S, Sheikh Z, Mozaffar H, Sheikh A.</p>	<p>Investigating the use of</p>	<p>Tarkastella dataohjattua tekoälypohjaisiin algoritmeihin</p>	<p>Systemaattinen kirjallisuuskatsaus, jonka lopullisen aineiston</p>	<p>Tutkimuksissa suurta laadullista</p>	<p>Päätöksenteon tukeminen</p>

<p>2020</p> <p>Health Informatics J. 2020 Sep;26(3):2138-2147. doi: 10.1177/1460458219900452. Epub 2020 Jan 22. PMID: 31964204.</p> <p>Iso-Britannia</p>	<p>data-driven artificial intelligence</p> <p>in computerised decision support systems for health and social care:</p> <p>A systematic review</p>	<p>perustuvaa päätöksentekojärjestelmää</p> <p>terveydenhuollon ja sosiaalihuollon kontekstissa.</p> <p>Tutkimuskysymykset:</p> <p>Millaisia datapohjaiseen tekoälyyn yhdistettyjä päätöksentekojärjestelmiä käytetään terveydenhuollon ja sosiaalihuollon ympäristöissä?</p> <p>Mitä todisteita päätöksentekojärjestelmien tehokkuudesta ja hoidon laadun ja turvallisuuden parantamisesta on todistettu?</p>	<p>muodosti 5 vertaisarvioitua englannin kielistä artikkelia, jotka julkaistu 1/2013–9/2018 välillä.</p>	<p>vaihtelua, eikä aineistossa ollut yhtään sosiaalialaan liittyvää tutkimusta, joissa olisi hyödynnetty tekoälypohjaista päätöksen tekoa.</p>	
<p>Bako AT, Taylor HL, Wiley K Jr, Zheng J, Walter-McCabe H, Kasthurirathne SN, Vest JR.</p> <p>2021</p>	<p>Using Natural Language Processing to Classify Social Work Interventions</p>	<p>Tarkoitus poimia ja luokitella potilastietojärjestelmään kirjoitetuista sairaalassosiaalityöntekijöiden muistiinpanoista</p>	<p>Luonnollisen kielen analyysi (NLP), jossa aineistona 815 sosiaalityöntekijän tekemää asiakasmuistiinpanoa potilastietojärjestelmästä. Analyysissä hyödynnettiin kirjallisuuden perusteella tehtyä luokittelua ja itse analyysi toteutettiin Pythonin RegEx:llä,</p>	<p>Koneoppimista voidaan hyödyntää, kun halutaan tunnistaa sosiaalityöntekijöiden tekemistä vapaa-muotoisista muistiinpanoista tehtyjä toimenpiteitä</p>	<p>Ennaltaehkäisy</p> <p>Sosiaalisten ongelmien tunnistaminen ja niihin puuttuminen</p>

<p>Am J Manag Care. 2021 Jan 1;27(1): e24-e31. DOI: 10.37765/ajmc.2021.88580. PMID: 33471465; PMCID: PMC8005360.</p> <p>Yhdysvallat</p>		<p>toimenpiteitä, joilla on vastattu asiakkaan tuen tarpeisiin.</p>	<p>jolla säännöllisiä ilmaisuja etsittiin tekstistä.</p>	<p>ja interventioita. Tämä taas auttaa päätöksentekijöitä resurssoinnissa ja toiminnan suunnittelussa.</p>	
<p>Gillingham, Philip</p> <p>2021</p> <p>Practice, 33:1, 37-50, DOI: 10.1080/09503153.2020.1749584</p> <p>Iso-Britannia</p>	<p>Algorithmically Based Decision Support Tools: Skeptical Thinking about the Inclusion of Previous Involvement</p>	<p>Tarkastella algoritmeihin perustuvien menetelmien kehittämistä ja käyttöä sosiaalityön päätöksentekovälineenä.</p>	<p>Kirjallisuuskatsaus</p>	<p>Sosiaalialan ammattilaisten pitää ymmärtää massadatan muodostuminen sekä myös periaatteet, joten sitä käytetään algoritmin mallintamisessa. Lisäksi sosiaalityöntekijöiden tulee olla mukana päätöksentekomallien kehittämisessä.</p>	<p>Päätöksenteon tukeminen</p>
<p>Grządzielewska, Małgorzata</p> <p>2021</p> <p>Child and Adolescent Social Work Journal 38:175–180</p> <p>https://doi.org/10.1007/s10560-020-00733-w</p> <p>Puola</p>	<p>Using Machine Learning in Burnout Prediction: A Survey</p>	<p>Tutkimuksessa verrattiin koneoppimisalgoritmia tilastolliseen monimuuttujamenetelmään sosiaalityöntekijöiden työssä uupumisen varhaisen tunnistamisen tukena.</p>	<p>Survey- tutkimus</p>	<p>Artikkelissa esitetään näkemys koneoppimis-malleista, jotka liittyvät pääasiassa työssä uupumisen tunnistamiseen. Koneoppimismalli tunnistaa hyvin työssä uupumiseen viittaavat muuttujat kompleksista aineistoista ja algoritmit voivat luoda uusia</p>	<p>Ennaltaehkäisy</p> <p>Sosiaalialan ammattilaisten työssä jaksaminen ja muutoskyvykyys</p>

				teoreettisia malleja loppuun palamisesta.	
<p>Hodgson, D., Goldingay, S., Boddy, J., Nipperess, S. & Watts, L.</p> <p>2022</p> <p>The British Journal of Social Work, Volume 52, Issue 4, June 2022, Pages 1878–1895, https://doi.org/10.1093/bjsw/bcab168</p> <p>Iso-Britannia</p>	<p>Problematising Artificial Intelligence in Social Work Education: Challenges, Issues and Possibilities</p>	<p>Tarkastella tekoälyn mahdollisia tulevaisuuden vaikutuksia sosiaalityön koulutukseen.</p>	<p>Kirjallisuuskatsaus (review)</p>	<p>COVID-19 pandemia ja tekoälykehityksen nopea harppaus muuttavat pysyvästi sosiaalityön koulutusta, jolloin koko alan akateeminen käytäntö on uudelleenarvioinnin edessä.</p>	<p>Ennaltaehkäisy</p> <p>Sosiaalialan ammattilaisten työssäjaksaminen ja muutostyövykyys</p>
<p>Litchfield, I., Glasby, J., Parkinson, S., Hocking, L., Tanner, D., Roe, B., & Bousfield, J.</p> <p>2023</p> <p><i>Health & Social Care in the Community</i>, 2023, 1–11. https://doi.org/10.1155/2023/9174873</p> <p>Iso-Britannia</p>	<p>Trying to Find People to Fit the Tech. . .": A Qualitative Exploration of the Lessons Learnt Introducing Artificial Intelligence-Based Technology into English Social Care</p>	<p>Tarkastella tekoälyominaisuuksin varustettujen sensorijärjestelmien käyttöönottoa englantilaisessa sosiaalihuollossa. Järjestelmien tarkoitus oli tunnistaa sellaisia muutoksia asiakkaisten käyttäytymisessä, jotka viittaavat mahdollisesti vakavampiin ongelmiin.</p>	<p>Haastattelututkimus Tutkimuksessa toteutettiin 18 puolistrukturoitua haastattelua kolmessa sosiaalihuollon toimipisteessä, joiden asiakkailta on ollut käytössä kotiin asennettuja sensorivusteisia tekoälyjärjestelmiä. Haastateltavina oli päättäjiä, operatiivisia johtajia ja hoitohenkilökuntaa.</p>	<p>Teknologialla on mahdollisuus tukea itsenäistä ja aktiivista elämää ja auttaa sekä tuen tarpeessa olevia henkilöitä että auttajia. Testattu teknologia ei sellaiseen palvelut sosiaalialan toimintaympäristöä, ja lisäksi sen käyttöohjeissa oli epäselvyyttä. Hoitohenkilökunta raportoi kokevansa tilanteen siten,</p>	<p>Ennaltaehkäisy</p> <p>Sosiaalisten ongelmien tunnistaminen ja niihin puuttuminen</p>

				että ihmisten tulee sopeutua teknologian vaateisiin eikä teknologian ihmisten tarpeisiin.	
<p>Glasby J, Litchfield I, Parkinson S, Hocking L, Tanner D, Roe B, et al.</p> <p>2023</p> <p>Health Soc Care Deliv Res 2023;11(9). https://doi.org/10.3310/HRYW4281</p> <p>Iso-Britannia</p>	<p>New and emerging technology for adult social care – the example of home sensors with artificial intelligence (AI) technology</p>	<p>Arvioida, kuinka kotiin asennettuja tekoälyominaisuuksilla varustettuja antureita osataan hyödyntää tuen tarpeen arvioinnissa ja järjestämisessä.</p>	<p>Haastattelututkimus</p> <p>Kaksivaiheinen toteutus, jossa ensimmäisenä toteutettiin kirjallisuuskatsaus, yhdeksän haastattelua ja kolme suunnittelutyöryhmäkokousta.</p> <p>Toisessa vaiheessa haastateltiin 20 sosiaalialan ammattilaista kolmessa eri toimipaikassa sekä kolme sensoriteknologian valmistajan ja säätelyviranomaisen edustajaa.</p>	<p>Organisaatiot odottivat teknologialta paljon tukea kotihoidon järjestämiseksi ja kustannusten säästämiseksi. Käyttökoulutuksesta huolimatta ammattilaisten oli vaikea hyödyntää tekniikkaa, jolloin se ei vastannut odotuksia.</p> <p>Sosiaalityö on perinteeltä pikemminkin reaktiivista kuin proaktiivista, jolloin myös päättäjiltä puuttuu strateginen ymmärrys teknologian käytöstä ja rajoituksista päätöksenteon tukena.</p>	<p>Ennaltaehkäisy</p> <p>Sosiaalisten ongelmien tunnistaminen ja niihin puuttuminen</p>

<p>McDuié-Ra D, & Gulson K.</p> <p>2020</p> <p>Area 52: 626–633. https://doi.org/10.1111/area.12602</p> <p>Asutralia</p>	<p>The backroads of AI: The uneven geographies of artificial intelligence and development</p>	<p>Tarkastella tekoäly-tekniikoiden kehityksen taustalla olevia maantieteellisesti epätasaisesti jakautuvia datalähteitä.</p>	<p>Käsiteanalyysi, jossa aineistona käytetty Maailmanpankin ja Intian valtion populistisena pidettyä #AIforAll hankkeen dataa.</p>	<p>Big data, algoritmit ja koneoppiminen poistavat vastuun kehityksestä yhteisöiltä ja yksilöiltä ja siirtää sen tekoälylle. Kehitys on huolestuttavaa, mikäli esimerkiksi globaali etelä puuttuu kokonaan aineistosta.</p>	<p>Aineistoharha tekoälykehityksessä</p>
<p>Formosa, P.</p> <p>2021</p> <p><i>Minds and Machines (Dordrecht)</i>, 31(4), 595–616. https://doi.org/10.1007/s11023-021-09579-2</p> <p>Australia</p>	<p>Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy</p>	<p>Tarkastella vuorovaikutteisten ja tekoälyperustaisten robottien vaikutusta ihmisten autonomiaan.</p> <p>Tarkastelun keskiössä on robottien muuttuminen työvälineestä toimijaksi.</p>	<p>Kuvaileva kirjallisuuskatsaus.</p>	<p>Artikkeli antaa suosituksia, kuinka AI ja robottekehityksessä voidaan huomioida ja kunnioittaa käyttäjiensä autonomiaa esimerkiksi sääntelyllä, standardoinnilla, käyttäjätutkimuksilla ja tietoisuuden lisäämisellä.</p> <p>Koska artikkeli tarkastelee ihmisten autonomisuuden tukemista tekoäly- ja robottekehityksessä, samalla tarkastelun ulkopuolelle rajautuu muut eettiset näkökulmat,</p>	<p>Sosiaalialalle suunnattujen tekoälyjärjestelmien ja sovellusten eettinen käyttö</p>

				kuten yhteinen etu ja oikeudenmukaisuus.	
Reamer, F. 2023 International Journal of Social Work Values and Ethics, 20(2), 52-71. https://doi.org/10.55521/10-020-205 Yhdysvallat	Artificial Intelligence in Social Work: Emerging Ethical Issues	Tarkastella tekoälyn käyttöön liittyviä eettisiä kysymyksiä sosiaalityössä.	Kirjallisuuskatsaus, jonka lisäksi arvioitu Yhdysvalloissa käytössä olevia ammattikäyttöön suunniteltuja tekoälyavusteisia sovelluksia.	Artikkeli antaa suositukset tekoälyn käytön eettiseksi standardeiksi sosiaalityössä ja hahmottelee sosiaalityöntekijöiden tekoälyn eettisen käytön strategiaa.	Sosiaalialalle suunnattujen tekoälyjärjestelmien ja sovellusten eettinen käyttö
Shahbazi, N., Lin, Y., Asudeh, A., & Jagdish, H. V. 2023 ACM Computing Surveys, 55(13s), 1–39. https://doi.org/10.1145/3588433 Yhdysvallat	Representation Bias in Data: A Survey on Identification and Resolution Techniques	Tutkia tekniikoita esitysharhan ja vääristymien tunnistamiseksi ja ratkaisemiseksi tutkimusdatassa (esim. luonnollinen kieli), jota käytetään koneoppimismalleissa ja tekoälyn opetuksessa.	Survey-tutkimus, jossa on tarkasteltu aineistoharhaan liittyviä tutkimuksia, joissa data on jäsentelemätöntä ja esitetty sekä taulukkomuodossa että muilla tavoin kuten kuvina, kaavioina, tekstinä tai puheena.	Nykyiset tekniikat, joilla datavääristymiä tunnistetaan, eivät ole suoraan sovellettavissa uudenlaisiin datatyyppeihin kuten esimerkiksi suoratoistodataan. Muutenkin etenkin edustusharhasta johtuvien datavääristymien korjaamiseksi tarvitaan uudenlaisia kattavuuden ja edustavuuden mittareita.	Aineistoharha tekoälykehityksessä

LIITE 2. BLOGI-TEKSTIT

TEKOÄLY SOSIAALIALALLA

<https://blogi.savonia.fi/peoplechangetheworld/tekoaly-sosiaalialalla/>

Tekoälyteknologiat muokkaavat yhteiskuntaamme ja samalla sosiaalialan toimintaympäristöä. Sosiaalialalla digitalisaatio on nähty pitkään uusien sosiaalisten ongelmien, kuten digitaalisen eriarvoistumisen, nettiriippuvuuksien tai yleistyvän, anonyymin vihapuheen aiheuttajana (vrt. Kivistö 2017, 22). Tekoälyn suosion kasvun ja käytön arkipäiväistymisen myötä on kuitenkin herännyt kiinnostus myös siihen, miten teknologiaa voidaan hyödyntää sosiaalisten haasteiden ratkaisemiseen.

Tämä osa blogista toimii tieteellisen viestinnän kanavana YAMK opinnäytteelle, jonka tutkimustehtävänä on tarkastella tekoälyn hyödyntämisen nykytilaa sekä eettistä soveltamista sosiaalialalla.



"Social exclusion" by OpenART AI

Blogin avulla on mahdollisuus parantaa tieteelliseen tutkimukseen saavutettavuutta, sekä tehdä tutkittua ja tärkeänä koettua asiaa näkyväksi myös visualisoinnin keinoin. Blogin lisäksi opinnäytetyöstä julkaistaan Savonia AMK ohjeistuksen mukaisesti tutkimusraportti, joka noudattaa hyvän tieteellisen käytännön periaatteita. Sekä blogipostaukset että raportti tarkastetaan plagioinnin varalta.

Lähteet

Sekä opinnäytteessä että blogissa hyödynnetään seuraavasti tekoälyä:

ChatGPT 2023. OpenAI. GPT-3.5, tekstien ideointiin. <https://chat.openai.com>

OpenArtAI 2023. OpenAI, blogin kuvitukseen. <https://openart.ai/home>

Kivistö, Mari 2017. Sosiaalityö digitalisaatiossa ja eettisyyden vaade. Teok-
sessa: Kivistö, Mari & Pääkkönen, Kirsi (toim.). Sosiaalityö digitalisaatiossa. La-
pin yliopiston yhteiskuntatieteellisiä julkaisuja C. Työpapereita 58. Rovaniemi:
Lapin yliopistopaino. 21–42

Opinnäytetyö on päätös!

<https://blogi.savonia.fi/peoplechangetheworld/opinnaytetyo-on-paatos/>

Miksi lattiakaivojen pudistaminen hammasharjalla on parasta tekemistä mitä voi tietää, sähköpostin lukeminen tuntuu vastenmieliseltä ja kalenterissa vilisevät viikot nostavat hikikarpalot otsalle?

Vastaus löytyy hupenevasta opiskeluajasta, joka pakottaa tekemään opinnäytetyön tutkimussuunnitelmaa ja etenemään valmistumisprosessissa.



"Student who is struggling to graduate and instead of studying she is doing something else" by OpenART AI

Opinnäytetyö on opiskelijan itsenäinen, mutta ohjattu tuotos, jolla pyritään osoittamaan opintojen aikana hankitun kokonaisvaltaisen tietoperustan analyttinen yhdistäminen rajattuun ongelmaan. Lisäksi opinnäytetyö valmentaa tekijäänsä tiedeviestintään ja tieteelliseen kirjoittamiseen, joka on paljon muutakin kuin kirjoittamista ja tekstin tuottamista. Tieteellinen kirjoittaminen on ajattelua, luetun uudelleen jäsentämistä ja akateemista toimintaa. Se yhdistää parhaimmillaan ajankohtaisen ammatillisen keskustelun tieteenalan tuottamaan tutkimustietoon.

Opinnäytetyöprosessiin kuuluu ideointia, suunnittelua ja vuoropuhelua ohjaajien kanssa. Prosessi vaatii päättäväisyyttä, sitkeyttä ja toisinaan myös puuduttavaa istuskelua tietokoneen ruutua tuijottaen ja inspiraatiota etsien. Digitaaliset työkalut voivat kuitenkin tuoda viihteellisyyttä ajoittain yksinäiseen puurtamiseen. Tekoälyn hyödyntämisestä opinnäytteessä on julkaistu oppilaitoskohtaisia ohjeita, joissa käsitellään muun muassa tekoälyyn liittyvää tietosuojaa sekä käytön avoimuutta. ChatGPT:n kanssa käytävä vuoropuhelu voi myös keventää tunnelmaa ja kenties auttamaan pääsemään eteenpäin jokaisen opinnäytetyötekijän kohtaamasta jumituksesta. Ohessa ChatGPT:n kanssa käyty keskustelu ideoista blogin otsikoiksi.

YOU

Hei! Keksitkö minulle nimiehdotuksia YAMK insinööriyötä käsittelevälle blogille?

ChatGPT

Tietysti! Tässä muutamia nimiehdotuksia YAMK insinööriyötä käsittelevälle blogille:

"Insinöörin Insight"

"Tekniikan Taival"

"GraduGuru - Insinöörin Ääni"

"Innovaatiot ja Insinööriyö"

"Tekniikan Timantit: YAMK Insinööriyö"

"Insinööritiede: Maisteriksi Askel Askeleelta"

"Insinööri-Inspiraatio"

"Innovaatioteknologia YAMK: Insinööriyön Maailma"

"Teknologian Tähtihetket"

"Insinöörin Älykkyys: YAMK Tutkimusmatka"

Toivottavasti näistä löytyy sopiva tai ainakin antavat inspiraatiota oman bloginimesi!

Opinnäytetyö on ensisijaisesti harjoitustyö, joka on parhaimmillaan valmistuessaan. YAMK opinnoista haetaan uutta suuntaa ja etenemismahdollisuuksia urapolulle sekä syvennetään asiantuntijuutta. Kelpoisuus pätevöittää myös johtotehtäviin, ja johtajan pitää pystyä tekemään päätöksiä.

Myös opinnäytetyön tekeminen on lopulta päätös.



"Surrealistic painting of child protection" by OpenART AI

Blogissa on käytetty seuraavasti tekoälyä:

ChatGPT 2023. OpenAI. GPT-3.5. Käytetty ideoimaan blogien nimiä, joulukuu 2023. <https://chat.openai.com>

OpenArtAI 2023. OpenAI. Käytetty blogin kuvitukseen, joulukuu 2023. <https://openart.ai/home>

Lähteet

Hakala, Juha T. 2017. Tulevan maisterin graduopas. Helsinki: Gaudeamus.

Vilkkä, Hanna 2020. Akateemisen lukemisen ja kirjoittamisen opas. PS-kustannus.

Helsingin yliopisto, julkaisuaika tuntematon. Blogikirjoitus. Tiedeviestintä. Tutkijablogin pitäminen. <https://blogs.helsinki.fi/tiedeviestinta/verkko/tutkijablogin-pitaminen/>

Miten tekoälylle opetetaan etiikkaa?

<https://blogi.savonia.fi/peoplechangetheworld/miten-tekoalylle-opetetaan-etiikkaa/>



Tekoälyä, jota nykyään löytyy lähes jokaisesta älypuhelimesta, kuvantunnistus-sovelluksesta ja virtuaaliavustajasta, kutsutaan kapeaksi tekoälyksi. Kapean tekoälyn kyvykkyys rajoittuu yhteen tehtävään, joten se ei pysty monitulkintaisuuteen eikä moraalisten arvioiden tekemiseen. Moraali ja eettinen harkinta ovatkin ihmisyyteen keskeisesti liitettäviä kognitiivisia toimintoja, joihin tekoäly ei kykene.

Koska tekoälyllä ei ole eettistä harkintaa, tulee sen ohjelmoinnissa rajoittaa epäeettistä käyttäytymistä antamalla tekoälylle joko ehdottomia sääntöjä, joita sen tulee noudattaa tai kuvaamalla haluttua lopputulosta. Tällöin tekoäly oppii ja kehittyy kokemusten ja datan perusteella ilman ennalta määriteltyjä sääntöjä. Usein monimutkaisissa järjestelmissä, ja erityisesti niissä, jotka vaikuttavat ihmisten henkeen ja terveyteen, käytetään hybridimallia. Klassinen esimerkki tällaisesta monimutkaisesta järjestelmästä on itseohjautuva auto. Sen tekoälyjärjestelmä tunnistaa tilanteet, joissa täytyy ehdottomasti noudattaa ennalta määrättyjä sääntöjä ja vastaavasti tilanteet, joissa oppiminen ja soveltaminen on mahdollista.

Koulutusaineiston eettisyys

Tekoälyn kehitykseen käytettävä data syntyy inhimillisen toiminnan seurauksena ja inhimillisen toiminnan seurauksena se myös päättyy tekoälyn tekemään päätöksenteko- tai analysointimalliin. Tekoälyn etiikka rakentuuakin pitkälti sen käyttämän koulutus- ja lähdeaineiston laatuun, joten onnistumisen edellytys on, että tuo aineisto on luotettavaa, edustavaa ja kattavaa. Mitä enemmän tekoälyä hyödynnetään sosiaali- ja terveyspalveluiden analysoinnissa ja päätöksenteon tukena, sitä tärkeämpää on ymmärtää tekoälykehityksessä käytettävän datan merkitys.

Tekoälyn koulutuksessa käytettävä data ja siihen liittyvät virheet ovat yksi keskeisimmistä eettisistä haasteista tekoälykehityksessä. Dataan liittyvät haasteet voidaan tiivistää kahteen pääluokkaan, jotka ovat datavinoumat sekä sensitiivisten tietojen käyttö koulutusdatassa ja erityisesti siihen tarvittavat suostumukset. Datan liittyvät vinoumat ja stereotypiat on helppo havainnollistaa, kun pyytää tekoälypohjaista kuvantunnistussovellusta piirtämään henkilökuvia. Alla olevissa kuvissa OpenArt AI:ta on pyydetty piirtämään kuvat etenemisvaikeuksissa olevista opiskelijoista ilman, että kehotteessa mainittiin opiskelijan sukupuolta. Oikeanpuoleisen kuvan kehotteessa lisättiin opiskelun sijaistoiminnoksi siivoaminen, josta voinee tulkita, että viittaaminen siivoamiseen sai tekoälyn piirtämään kuvaan tummaihoisen naisen siivoamassa.



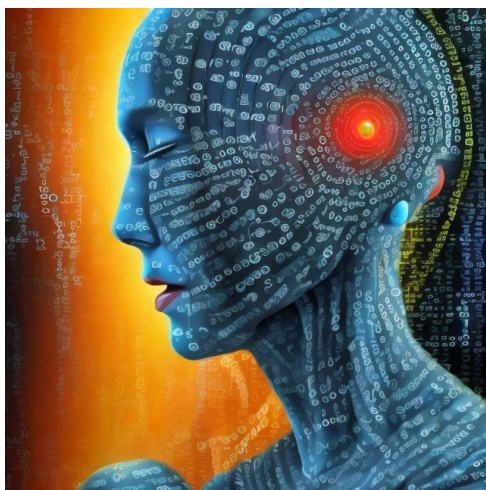
Sosiaalialan koulutusdatan erityispiirteet

Tekoälyn koulutukseen käytettävää aineistoa voidaan kerätä asiakaskirjauksista, kuvista, taulukoista tai vaikkapa äänilähteistä. Sosiaalialalla aineistoa kertyy paitsi asiakkaista myös heidän läheisistään ja se sisältää runsaasti henkilökohtaista tietoa. Sensitiivisen datan käyttöön liittyviin eettisiin kysymyksiin on pyritty löytämään vastauksia datan anonymisoinnista, mutta silloin tekoälyn ennustevoima kärsii. Sosiaalialalla datan on oltava myös yksilöitävissä eri järjestelmien kesken, jolloin sen on sisällettävä riittävästi tunnistetietoa.

Tekoälyä ei saa missään tilanteessa käyttää ilman sen vaikutuspiirissä olevien ihmisten tietoisuutta olemassa olevasta teknologiasta. Henkilökohtaisten tietojen käyttäminen ja tallentaminen edellyttää lisäksi aina suostumusta ja tuon suostumuksen ehtona on, että sen antaja ymmärtää riittävällä tavalla keillä on pääsy dataan ja miten sitä käytetään.

Sosiaalialalla käytetään paljon käsitteitä (esimerkiksi syrjäytyminen, oikeudenmukaisuus), joiden yksiselitteinen määrittäminen on hankalaa. Kehittäjien vastuulla on ohjelmoida tekoäly tunnistamaan sosiaaliset ongelmat, joihin on vaikea löytää yleispätevää määritelmää ja joihin tyypillisesti liittyy vahva subjektiivisuus. Tällöin tekoälypohjaisen järjestelmän tekemään päätöksentekoon voi liittyä osittain tahatontakin puolueellisuutta, epäjohdonmukaisuutta ja satunnaisuutta.

Tekoäly tarvitsee ihmisälyä eettiseksi tuekseen sosiaalialalla



Tekoälylle etiikan opettaminen on verrattavissa siihen, kun lapselle opetetaan moraalisia periaatteita ja sosiaalisia normeja. Tekoäly tarvitsee ohjausta kehittäjiltä ja käyttäjiltään pystyäkseen tunnistamaan hyväksyttäviä ja sopivia toimintatapoja erilaisissa tilanteissa.

Kansalliset ja kansainväliset lait ja standardit pyrkivät varmistamaan oikeudenmukaisen tekoälykehityksen. Eettisyyden vaatimus täydentää lainsäädäntöä, joka ei aina pysy nopean teknologisen kehityksen perässä. Mitä enemmän tekoälyjärjestelmiä otetaan käyttöön ja niille annetaan vastuuta päätöksenteosta sitä tärkeämpää on varmistaa, että ne toimivat ehdottomien arvojen mukaisesti, oikeudenmukaisesti ja järjestelmien elinkaaren aikaiset vastuut ja velvollisuudet ovat selkeät. Tässä tarvitaan sosiaalialan ammattilaisen osallisuutta kehittäjien rinnalla.

Lähteet

Charisi, Vicky, Dennis, Loïuse, Fisher, Michael, Lieck, Robert, Matthias, Andreas, Slavkovik, Marija, Sombetzki, Janina, Winfield, Alan F. T., & Yampolskiy, Roman 2017. Towards Moral Autonomous Systems. arXiv.Org.
<https://doi.org/10.48550/arxiv.1703.04741>

Euroopan komissio 2019. Viestintäverkkojen, sisältöjen ja teknologian pääosasto. Luotettavaa tekoälyä koskevat eettiset ohjeet. Julkaisutoimisto. <https://data.europa.eu/doi/10.2759/068692>

Gillingham, Philip & Graham, Timothy 2017. Big Data in Social Welfare: The Development of a Critical Perspective on Social Work's Latest Electronic Turn. Australian Social Work 70 (2): 135–47. <https://doi.org/10.1080/0312407X.2015.1134606>.

Hallamaa, Jaana & Kalliokoski Taina 2022. AI Ethics as Applied Ethics. Frontiers in Computer Science (Lausanne) 4. <https://doi.org/10.3389/fcomp.2022.776837>.

Litchfield, Ian, Glasby, Jon, Parkinson, Sarah, Hocking, Lucy, Tanner, Denise, Roe, Bridget & r Bousfield, Jennife. 2023. 'Trying to Find People to Fit the Tech...': A Qualitative Exploration of the Lessons Learnt Introducing Artificial

Intelligence-Based Technology into English Social Care. Health & Social Care in the Community 2023: 1–11. <https://doi.org/10.1155/2023/9174873>.

Talvitie-Lamberg, Karoliina, Silvennoinen, Minna, Moilanen, Hannu & Korpela, Jari 2018. Tekoäly sosiaalisen syrjäytymisen ennakoivassa tunnistamisessa – Case THL kouluterveyskysely. Jyväskylän yliopisto. Informaatioteknologian tiedekunnan julkaisuja. No. 62/2018.

Blogin kuvitus: OpenART AI.

Kehotteet:

“automated vehicle controlled by AI

“a painting of a student who is struggling to graduate.”

“a painting of a student who is struggling to graduate and instead of studying the student is cleaning.”

“painting of Artificial intelligence and social work”

Tekoälyn soveltaminen sosiaalialalla

<https://blogi.savonia.fi/peoplechangetheworld/tekoalyn-soveltaminen-sosiaalialalla/>



Englantilaistutkimuksessa testattiin aikuissosiaalityön tukena kotiin asennettavaa monianturointitekniikkaa, josta syntyvästä datasta pyrittiin tekoälyn avulla tunnistamaan ongelmista ennakoivia poikkeamia asiakkaiden käyttäytymisessä. Kyseessä ei ollut kameroihin perustuva valvonta, vaan erilaisten antureiden keräämää havaintodataa ihmisten arjesta kotiloissa, johon voitiin lisäksi yhdistää

muuta tietoa asiakkaan terveydestä ja hyvinvoinnista. Perienglantilaiseen tapaan yksi keskeinen seurattava mittari oli vedenkeittimen päälle laittaminen.

Millaisia sosiaalisia interventioita voimme perustaa siihen, että asiakas jättää väliin kello viiden teen? Tuskin pelkästään teen väliin jättäminen johtaa väliintuloihin, mutta yhdistämällä tieto elintoimintoihin liittyvään dataan, liikkeen havainnoimiseen ja ennalta tunnistettuihin riskeihin, saadaan jo huomattavasti enemmän arvokasta tietoa mahdollisesta hyvinvointia uhkaavasta tilanteesta. Tekoäly havaitseekin ihmistä paremmin laajoista, monipuolisista aineistoista säännönmukaisuuksia, jolloin työskentelytapaa on mahdollista muuttaa reaktiivisesta proaktiiviseen.

Tekoälyllä on potentiaalia muuttaa sosiaalityötä ja tehostaa asiakkaille, organisaatioille ja yhteisöille tarjottavia palveluita. Tekoäly mahdollistaa nopean datan analysoinnin tavalla, jota voidaan sosiaalityössä käyttää riskiarvioinnissa, interventioiden arvioinnin tukena sekä palvelujärjestelmän systemaattisten virheiden tunnistamiseksi ja korjaamiseksi. Päätöksenteon ja palveluiden parantamisen lisäksi, tekoälyn avulla voidaan edistää sosiaalialan ammattilaisten työssä jaksamista ja tunnistaa loppuun palaamista ennakoivia riskitekijöitä. Li-

säksi keskustelevala tekoäly sekä vuorovaikutteiset robotit voivat toimia asiakkaan kuntoutusprosessin tukena. Hyviä kokemuksia on saatu esimerkiksi vankeinhoidosta.

Sosiaalialan uudet vaateet

Tekoälypohjaiset järjestelmät tekevät vasta tuloaan sosiaalialalle, joten julkaistua tutkimusta on niukasti saatavilla. Ylipäättään teknologiaa ja virtuaalista työvoimaa hyödynnetään sosiaalialalla vasta vähän, joten alalle ei ole myöskään syntynyt kovin laajaa käyttökokemusta tekniikasta työn tukena. Uuden teknologian laajamittaisen käytön esteeksi onkin tunnistettu tiedon puute sekä tekniikkaan liittyvät ennakkoluulot. Terveystuollossa teknologia-avusteisella työllä on sen sijaan pitkä historia. Sosiaaliala voisikin ottaa oppia lähitoimialalta, jossa teknologia-avusteisesta työstä on saatu runsaasti käyttökokemusta ja tunnistettuja hyötyjä.

Tekoälyjärjestelmien kehityksessä on huomioitava, että saman teknologian käyttö eri tilanteissa voi tehdä siitä eettisesti kyseenalaista. Usein eettiset ongelmat erottuvat selvimmin järjestelmissä, joissa on mukana valvonnan ja kontrollin elementtejä. Kasvojen tunnistaminen on perusteltu toiminnallisuus esimerkiksi, kun kirjaudutaan omalle puhelimelle, mutta saman teknologian käyttö läsnäolon seurantaan työpaikalla on kontrollointia. Alun esimerkissä kuvattu kotiin asennettava teknologia voi olla jossain tilanteissa perusteltua ja hyväksyttävissä, mutta siinäkin on tarkkaan harkittava, missä menee asiakkaan yksityisyyden raja suhteessa tekniikasta saavutettaviin hyötyihin.

Tekoälyn käytön yleistymisellä on monenlaisia vaikutuksia sosiaalialan ammattilaisten ammattitaidon ylläpitämiseen ja uusien ammattilaisten kouluttamiseen. Ensinnäkin toimialan digitalisaatio asettaa uudelleen koulutusvaatimuksia ammattilaisille, jotta he oppivat käyttämään ja tulkitsemaan uutta tekniikkaa eettisesti ja asianmukaisesti. Sosiaalialan ammattilaiset toimivat sekä tekoälyratkaisuiden käyttäjinä että asiantuntijoina asiakkaiden suuntaan. Toisaalta myös opetus on muuttunut yhä enemmän teknologiavälitteiseksi. Perinteisesti kasvokkain ja harjoitteluissa opetettu toisten ihmisten kunnioittava kohtelu, eettinen arviointi ja ihmisoikeuskysymykset voivatkin muuttua kustannuspaineiden alla yhä enemmän itsenäisiksi verkko-opinnoiksi.

Sosiaaliala on vahvasti sidoksissa yhteiskunnalliseen todellisuuteen. Tekoälyn käytön arkipäiväistyminen vaikuttaa monin tavoin ihmisiin, yhteiskunnan rakenteisiin ja palveluihin ja vaikuttaa siten väistämättä myös sosiaalialaan. Kokonaisuutena eettinen ja oikeudenmukaisuuteen tähtäävä tekoälykehitys vaatii yhteistyötä lainsäätäjien, sovelluskehittäjien, loppukäyttäjien ja sote-alan ammattilaisen välillä. Jopa hyvin suunnitellut toimet voivat aiheuttaa laajalle ulottuvia, ei-toivottuja ekologisia, sosiaalisia, poliittisia ja kulttuurisia vaikutuksia.



Lähteet

Formosa, Paul 2021. Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy

Hodgson, David, Goldingay, Sophie, Boddy, Jennifer, Nipperess, Sharlene & Watts, Lynelle 2022. Problematizing Artificial Intelligence in Social Work Education: Challenges, Issues and Possibilities, *The British Journal of Social Work*, Volume 52, Issue 4, June 2022, Pages 1878–1895, <https://doi.org/10.1093/bjsw/bcab168>

Litchfield, Ian, Glasby, Jon, Parkinson, Sarah, Hocking, Lucy, Tanner, Denise, Roe, Bridget & Bousfield, Jennife. 2023. 'Trying to Find People to Fit the Tech...': A Qualitative Exploration of the Lessons Learnt Introducing Artificial Intelligence-Based Technology into English Social Care. *Health & Social Care in the Community* 2023: 1–11. <https://doi.org/10.1155/2023/9174873>.

Reamer, Frederic G. 2023. Artificial Intelligence in Social Work: Emerging Ethical Issues. *International Journal of Social Work Values and Ethics* 20 (2): 52–71. <https://doi.org/10.55521/10-020-205>.

Blogin kuvitus: OpenArtAI

Kehotteet:

5 o'clock tea, elegant tea set on a vintage lace tablecloth, porcelain teapot and teacups, dainty floral teacups, delicate china saucers, vintage silver tea strainer, warm and inviting lighting, cozy atmosphere, high quality, vintage, elegant, delicate, warm lighting, dainty tea set, floral design

impact of AI to social work

10 käskyä tekoälyn eettiselle käytölle sosiaalialalla

<https://blogi.savonia.fi/peoplechangetheworld/10-kaskya-tekoalyn-eettiselle-kaytolle-sosiaalialalla/>

Seuraavat kuvakortit on laadittu yhteistyössä ChatGPT:n kanssa, joka muodosti tekstit kortteihin ja OpenART AI kuvitti ne. Osaa kuvista on muokattu Photoshop Express sovelluksella ja tekstejä on paranneltu.



Inhimillinen ohjaus ja päätöksenteko

Säilytä päätösvalta ja vastuu tärkeissä päätöksissä ihmisillä. Tekoälyn tulee toimia ihmisten hyväksi eikä korvata inhimillistä harkintaa.



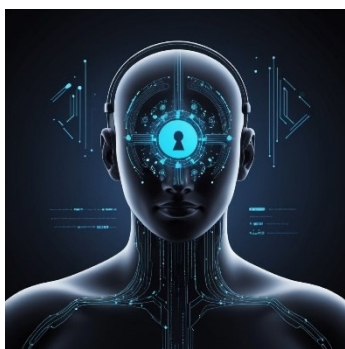
Läpinäkyvyys

Tee tekoälyn toimintaperiaatteet ja päätöksentekoprosessit läpinäkyviksi. Käyttäjille ja sosiaalityöntekijöille tulee kertoa, miten tekoäly tekee päätöksiä ja mitkä tiedot siihen vaikuttavat



Oikeudenmukaisuus

Varmista, että tekoäly ei vahvista olemassa olevia epäoikeudenmukaisuuksia. Tarkkaile algoritmeja ja dataa, jotta vältetään syrjiviä tai vinoutuneita tuloksia.



Yksityisyys

Suojaa asiakkaiden ja käyttäjien yksityisyyttä. Varmista, että tekoäly käsittelee henkilötietoja asianmukaisesti ja noudattaa tietosuoja-asetuksia.



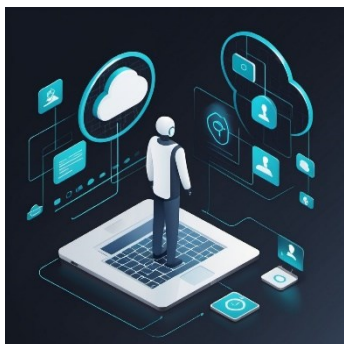
Koulutus ja tietämys

Kouluta sosiaalityöntekijöitä ja muita sidosryhmiä tekoälyn käytöstä, sen rajoituksista ja eettisistä haasteista. Tietämys auttaa varmistamaan vastuullisen ja eettisen käytön.



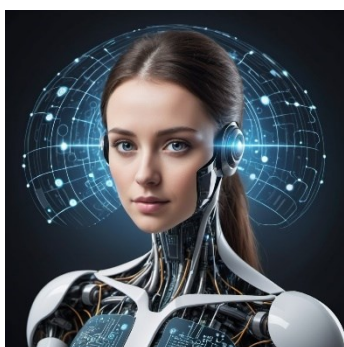
Osallistava suunnittelu

Ota sosiaalityöntekijät ja asiakkaat mukaan tekoälyn suunnitteluun ja kehitykseen. Näin varmistat, että teknologia palvelee erilaisia ja monimuotoisia tarpeita.



Vastuullisuus ja vastuun selkeys

Määritä selkeästi, kuka on vastuussa tekoälyn toiminnasta ja mahdollisista virheistä. Huolehdi siitä, että vastuu on asianmukaisesti jaettu eri osapuolille.



Palaute ja jatkuva parantaminen

Kerää palautetta tekoälyn käytöstä ja käytä sitä järjestelmän jatkuvassa parantamisessa.



Etiikan ja sääntelyn noudattaminen

Tunnista ja noudata soveltuvia eettisiä periaatteita ja sääntelyä. Huomioi paikalliset ja kansainväliset ohjeet tekoälyn käytössä sosiaalialalla.



Riskien hallinta

Arvioi mahdolliset riskit, kuten väärinkäytön mahdollisuudet ja tekniset haasteet. Kehitä strategioita niiden hallitsemiseksi.

Blogin kuvitus OpenART AI

ChatGPT Komento: "Laadi ohjeet tekoälyn eettiselle käytölle sosiaalialalla."