

Nea Auranen

KERTOMUKSIA TEKOÄLYLUKUTAIDOSTA

Tekoälyn tunnistaminen mediassa aikuisten sanoin

Kasvatustieteiden ja kulttuurin tiedekunta
Pro gradu -tutkielma
Joulukuu 2024

TIIVISTELMÄ

Nea Auranen: Kertomuksia tekoälylukutaidosta : Tekoälyn tunnistaminen mediassa aikuisten sanoin
Pro gradu -tutkielma
Tampereen yliopisto
Kasvatuksen ja yhteiskunnan tutkimuksen maisteriohjelma, elinikäinen oppiminen ja kasvatus
Joulukuu 2024

Tarkastelen tässä tutkielmassa aikuisten tekoälylukutaitoa eläytymismenetelmällä. Tutkielman tarkoituksena on luoda katsaus aikuisten ajatuksiin ja käsityksiin tekoälylukutaidosta tekoälyn tunnistamisen osalta. Tekoälylukutaito on suhteellisen uusi käsite, jolle ei ole vakiintunut selkeää määritelmää. Lisäksi generatiivisen tekoälyn viimeaikainen kehitys on tehnyt vain muutaman vuoden takaisista määritelmistä vajavaisia. Generatiivinen tekoäly on aivan viime vuosina auennut kaikille saavutettavaksi sovellusten, kuten ChatGPT:n, muodossa ja sen seurauksena aiheuttanut valtavan aallon generatiivisella tekoälyllä tuotettua sisältöä. Muutaman vuoden takaiset tekoälylukutaidon määritelmät eivät voineet vielä huomioida tätä kehitystä, joten niistä puuttuu tekoälyn tunnistamisen näkökulma, joka on nyt erityisen ajankohtainen. Tekoälylukutaitoa ja tekoälyn tunnistamista on siis syytä tutkia. Tässä tutkielmassa pyrin vastaamaan tutkimuskysymykseen: millaisia tekoälyn tunnistamisen keinoja aikuisten kertomuksissa ilmenee?

Tutkielman aineisto koostuu aikuisten kirjoittamista eläytymismenetelmäkertomuksista. Eläytymismenetelmän tarkoituksena on kerätä vastaajilta lyhyitä kertomuksia, jotka perustuvat tutkijan luomaan kehyskertomukseen. Kehyskertomus ohjaa vastaajia kirjoittamaan kertomuksen halutusta aiheesta rajaamatta kertomuksen sisältöä liikaa. Kehyskertomuksesta on tyypillisesti vähintään kaksi variaatiota, jotta on mahdollista tarkastella varioinnin tuottamia eroavaisuuksia aineistossa. Tutkimusjoukko koostuu tässä tutkielmassa eri ikäisistä aikuisista. Tein aineistolle aineistolähtöisen sisällönanalyysin, erikseen kehyskertomukseen A perustuville kertomuksille ja kehyskertomukseen B perustuville kertomuksille.

Tutkielman tulokseksi muodostui viisi pääluokkaa: valppaus, medialukutaito, tekoälyn tunnistamiseen harjaantuminen, tarkempi tilannekohtainen analyysi ja yhdessä oppiminen. Pääluokkia yhdistää yhdistävä luokka, tekoälyn tunnistamisen keinot, joka vastaa tutkimuskysymykseen. Pääluokkien lisäksi aineistosta nousi teema, joka ei sopinut yhdistävän luokan alle: huoli tekoälyn kehityksen suunnasta. Tarkastelin kuitenkin teemaa, sillä se on tutkielman aiheen kannalta olennainen ja jatkotutkimuksen kannalta mielenkiintoinen. Variointi ei tuottanut tässä tutkielmassa merkittäviä eroja, ja pienet eri variaatioihin perustuvien kertomusten välillä ilmenneet erot saattoivat johtua vastausten epätasaisesta jakautumisesta kahden kehyskertomuksen välillä.

Kokonaisuudessaan tutkielman tulokset antavat katsauksen tekoälyn tunnistamisen keinoihin, sellaisena kun ne ilmenevät aikuisten kertomuksissa. Tuloksia ei voi pitää yleistettävänä, sillä ne perustuvat vastaajien näkemyksiin, ajatuksiin ja kuvitteellisiin kertomuksiin. Ne tarjoavat kuitenkin näkökulman tekoälylukutaitoon, tekoälyn tunnistamiseen ja tekoälyyn liittyviin asenteisiin aikuisten silmin. Tutkielman tulosten perusteella olisi mielekästä tutkia laajemmin tekoälylukutaitoa sekä erityisesti tekoälyyn liittyviä asenteita, huolia ja epävarmuuksia.

Avainsanat: tekoälylukutaito, generatiivinen tekoäly, medialukutaito, eläytymismenetelmä

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -ohjelmalla.

ABSTRACT

Nea Auranen: Stories about AI literacy : Recognizing AI in media - told by adults
Master's Thesis
Tampere University
Master's programme in educational sciences, lifelong learning and education
December 2024

In this thesis I examine adults' artificial intelligence literacy (AI literacy) using the method of empathy-based stories. The purpose of this thesis is to create an outlook into adults' thoughts and perceptions of AI literacy regarding specifically the recognition of generative artificial intelligence in media. AI literacy is a relatively new concept that does not yet have a set definition. Also, the fast development of generative artificial intelligence in the past few years has outdated some of the existing definitions of AI literacy. In recent years generative artificial intelligence has become available to the public and therefore caused a wave of content created by generative artificial intelligence. Definitions of AI literacy from few years ago could not have taken this into account and are often missing the now very topical perspective of recognizing generative artificial intelligence in media. That is why researching recognizing artificial intelligence is important. I attempt to answer the following research question: What means of recognizing artificial intelligence surface from adults' stories?

The research data of this thesis consists of empathy-based stories written by adults. The purpose of the method of empathy-based stories is to collect short stories that are based on a prompt created by the researcher. The prompt guides the stories into a wanted direction without limiting the content too much. There are typically at least two variations of the prompt to explore differences caused by the variation. The population of this thesis consists of adults in different age groups. I analysed the data by performing two data-driven content analyses: one for stories based on prompt A and the second for stories based on prompt B.

As a result, I found five main categories from the data: vigilance, media literacy, learning to recognize AI by practice, thorough situational analysis, and learning together. These categories are tied together by a connecting category, means of recognizing artificial intelligence, which provides an answer to my research question. Also, another theme that did not fit among the other categories surfaced from the data: worries about the direction of the development of artificial intelligence. I decided to explore this theme anyway since it is relevant to the theme of this thesis as well as interesting. The variation used in data collection did not cause significant differences in the stories and the minor differences that appeared might have been caused by a difference in the amounts of stories per prompt.

Overall, the results of this thesis provide an outlook into the means of recognizing artificial intelligence as they appear in stories written by adults. The results cannot be considered objective since they are based on the views, thoughts, and imaginary stories written by adults. They do however provide a viewpoint into AI-literacy, recognizing artificial intelligence, and attitudes towards artificial intelligence through the eyes of adults. Based on the results of this thesis, it would be interesting to study AI literacy as a whole and especially the attitudes, worries, and uncertainties regarding artificial intelligence.

Keywords: AI literacy, generative artificial intelligence, media literacy, method of empathy-based stories

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

TEKOÄLYN KÄYTTÖ OPINNÄYTTEESSÄ

Opinnäytteessäni on käytetty tekoälysovelluksia:

- Ei
- Kyllä

Ilmoitukseni mukaan olen käyttänyt opinnäytteessäni tutkielmaprosessin aikana seuraavia tekoälysovelluksia:

Tekoälysovellusten nimet ja versiot: Scopus AI

Käyttötarkoitus: Käytin Scopus AI:ta tiedonhakuun.

Osiot, joissa tekoälyä on käytetty: Luku 2.

Olen tietoinen siitä, että olen täysin vastuussa koko opinnäytteeni sisällöstä, mukaan lukien osat, joissa on hyödynnetty tekoälyä, ja hyväksyn vastuun mahdollisista eettisten ohjeiden rikkomuksista.

SISÄLLYS

1	JOHDANTO	6
2	MEDIALUKUTAITO JA TEKOÄLY	8
2.1	Medialukutaito	8
2.2	Generatiivinen tekoäly	11
2.3	Tekoälylukutaito	13
3	ELÄYTYMISMENETELMÄ	17
3.1	Eläytymismenetelmä	17
3.1.1	<i>Eläytymismenetelmän tausta</i>	18
3.1.2	<i>Eläytymismenetelmällä saatujen tutkimustulosten luonne ja merkitys</i>	20
3.2	Eläytymismenetelmä tässä tutkielmassa	21
4	TUTKIELMAN TOTEUTUS	22
4.1	Kehyskertomukset	22
4.2	Aineiston kerääminen	23
4.3	Analyysi.....	25
5	TULOKSET	28
5.1	Millaisia tekoälyn tunnistamisen keinoja aikuisten kertomuksissa ilmenee?	28
5.1.1	<i>Valppaus</i>	31
5.1.2	<i>Medialukutaito</i>	31
5.1.3	<i>Tekoälyn tunnistamiseen harjaantuminen</i>	32
5.1.4	<i>Tarkempi tilannekohtainen analyysi</i>	33
5.1.5	<i>Yhdessä oppiminen</i>	34
5.2	Huoli tekoälyn kehityksen suunnasta.....	35
6	POHDINTA	36
6.1	Tutkielman tulokset teoreettisen viitekehyksen valossa	36
6.2	Tutkielman eettisyys ja luotettavuus	39
6.3	Johtopäätökset.....	42
	LÄHTEET	43
	LIITTEET	47
	Liite 1: Kehyskertomukset.....	47
	Liite 2: Tietosuojailmoitus.....	48

TAULUKOT

TAULUKKO 1.	ESIMERKKI PELKISTETYISTÄ ILMAUKSISTA JA NIISTÄ MUODOSTUNEESTA ALALUOKASTA ANALYYSISSÄ A	26
TAULUKKO 2.	ESIMERKKI PÄÄLUOKAN MUODOSTUMISESTA ANALYYSISSÄ A	27
TAULUKKO 3.	ANALYYSISSÄ A MUODOSTUNEET LUOKAT	30
TAULUKKO 4.	ANALYYSISSÄ B MUODOSTUNEET LUOKAT	30

1 JOHDANTO

Muuttuva maailma asettaa ihmisille jatkuvasti uudenlaisia vaatimuksia digitaalisen osaamisen osalta; digitaaliset ympäristöt, alustat ja media muuttuvat jatkuvasti ja digitaalisuus on levinnyt lähes kaikille elämänalueille. Digitaalisessa ulottuvuudessa navigoimiseen tarvitaan medialukutaitoa: kykyjä, jotka mahdollistavat monipuolisen ja turvallisen median käyttämisen, analysoinnin, arvioinnin ja tuottamisen sen eri muodoissa (ks. esim. Aufderheide, 1992; Livingstone, 2004; Potter, 2022). Kuitenkin professori Teemu Roosin mukaan ihmisten medialukutaito on nyt koetuksella, sillä tekoälyn viimeaikaisen kehittymisen myötä ”emme voi luottaa siihen, mitä kuulemme tai näemme” (Vaaherkumpu & Terävä, 27.5.2023); tekoälyn kehittyminen haastaa käsityksiämme ja toimintatapojamme (Tiernan ym., 2023). Viime vuosina tekoälyn kehityksessä on nähty merkittäviä loikkia ja huomiota ovat herättäneet muun muassa OpenAI:n julkaisema ChatGPT ja Googlen Bard (nykyään Google Gemini) (Vaaherkumpu & Terävä, 27.5.2023). Generatiivisen tekoälyn viimeaikainen kehitys mahdollistaa monenlaisia hyötyjä, mutta tuo mukanaan myös ongelmia, riskejä ja vaaroja. Esimerkiksi rikolliset hyödyntävät monia tekoälyn ominaisuuksia kyberhyökkäyksissä (Trafficom, 2022). Lisäksi generatiivisen tekoälyn avulla on entistä helpompaa tuottaa disinformaatiota (Lim ym., 2023). Myöskään esimerkiksi käyttäjien datan yksityisyyteen liittyvät seikat eivät ole aina läpinäkyviä (Long & Magerko, 2020). Tekoälyyn liittyy siis monenlaisia eettisiä kysymyksiä ja mahdollisia ongelmia, joista median käyttäjän olisi hyvä olla tietoinen.

Kasvatustieteen kentällä keskustelu tekoälystä on keskittynyt pitkälti koulu- maailmaan, arviointiin ja oppimiseen (Tiernan ym., 2023). Vaikuttaa siltä, että tekoälylukutaidon (ks. esim. Long & Magerko, 2020; Ng ym., 2021) näkökulmaa on tutkittu vähemmän ja siksi sitä on mielekästä tutkia, etenkin ottaen huomioon generatiivisen tekoälyn aivan viimeaikaisen nopean kehityksen, jonka seurauksena voitaisiin ajatella, että tekoälylukutaidon merkitys on kasvanut digitaalisen median täytyessä tekoälyllä generoidulla sisällöllä. Tekoälylukutaito voidaan kenties

mieltää osaksi medialukutaitoa, sillä se on nykyään erityisen olennainen taito mediaa navigoidessa. Pohdinkin tässä tutkielmassa hieman medialukutaidon ja tekoälylukutaidon suhdetta, mutta ennen kaikkea keskityn yhden tekoälylukutaidon osa-alueen, tekoälyn tunnistamisen, tarkasteluun.

Tarkastelen tässä tutkielmassa tekoälyn tunnistamista aikuisten silmin; kerään aineiston eläytymismenetelmällä, jonka tarkoituksena on kerätä osallistujilta lyhyitä tutkijan luomaan kehyskertomukseen perustuvia kertomuksia (esim. Eskola, 1997). Tutkimuksen tarkoituksena on valaista tekoälyn tunnistamisen keinoja ja mahdollisesti tekoälyyn liittyviä asenteita ja siten luoda pohjaa jatkotutkimukselle tekoälylukutaidon luonteesta ja määritelmästä. Tutkielman tutkimuskysymys on: *minkälaisia tekoälyn tunnistamisen keinoja aikuisten kertomuksissa ilmenee?*

2 MEDIALUKUTAITO JA TEKOÄLY

Käsittelen tässä luvussa tämän tutkielman kannalta keskeisiä teemoja: medialukutaitoa, generatiivista tekoälyä ja tekoälylukutaitoa. Käsittelen ensin medialukutaidon määritelmää ja merkitystä luvussa 2.1. Sen jälkeen, luvussa 2.2, kuvailen generatiivista tekoälyä ja tarkastelen siihen liittyvää keskustelua ja tutkimusta kasvatustieteiden kentällä. Lopuksi keskityn vielä tekoälylukutaitoon, sen määritelmään ja siihen liittyvään aiempaan tutkimukseen luvussa 2.3.

2.1 Medialukutaito

Medialukutaito on tämän tutkielman kannalta olennainen teema, sillä se liittyy läheisesti tekoälylukutaitoon ja tekoälylukutaito voidaan kenties mieltää jopa osaksi sitä. Keskitän näkökulmani aikuisten medialukutaitoon, sillä tutkielman näkökulma keskittyy aikuisiin. Aikuisten medialukutaidon erottaminen lasten ja nuorten medialukutaidosta on olennaista, sillä yksilöiden medialukutaitoon liittyvät tarpeet muuttuvat läpi elinkaaren (Rasi ym., 2019).

Medialukutaidolla tarkoitetaan kykyjen joukkoa, jota tarvitaan mediassa navigoimiseen (ks. esim. Aufderheide, 1992; Livingstone, 2004; Potter, 2022). Medialukutaidon käsitteen määritelmästä ja määritelmien yhteneväisyydestä tutkijoiden keskuudessa on monia mielipiteitä (Potter, 2022). Potterin (2022) mukaan vaihtelevuus medialukutaidon määritelmässä on valtavaa, myös sellaisten artikkeleiden välillä, joissa on viitattu samaan lähteeseen. Medialukutaidon määrittelyn yhteydessä mainitaan usein määritelmän vakiintumattomuus ja tiedostetaan monenlaisten variaatioiden olemassaolo (esim. Wuyckens ym., 2022). Sen lisäksi, että medialukutaidon käsitteelle ei ole selkeää yhteisesti hyväksyttyä määritelmää, sen suhde toisiin lukutaitokäsitteisiin, kuten informaatiolukutaitoon, datalukutaitoon, algoritmilukutaitoon tai digilukutaitoon, ei ole vakiintunut. Wuycken-

sin ja kollegojen (2022) mukaan käsitteet kuten medialukutaito, informaatiolukutaito ja digilukutaito vaativat vielä selvennystä niin määritelmien kun käsitteiden välisten suhteidenkin osalta.

Eräs medialukutaidon määritelmä, johon kuitenkin usein viitataan, on *National Leadership Conference on Media Literacy* -konferenssissa vuonna 1992 muodostettu määritelmä (Potter, 2022), jonka mukaan medialukutaitoinen ihminen kykenee tulkitsemaan, arvioimaan, analysoimaan ja tuottamaan sekä perinteistä että digitaalista mediaa (Aufderheide, 1992). Monet määritelmät sisältävät siis jonkinlaisen kokoonpanon näistä kyvyistä. Myös esimerkiksi Livingstone (2004) määrittelee medialukutaidon neljästä kyvystä koostuvaksi kokonaisuudeksi: mediaan käsiksi pääseminen, median analysointi, median arviointi ja viestien luominen erilaisissa konteksteissa. Aikaiset määritelmät ovat luoneet medialukutaidolle peruspilarit, joita uudemmissakin määritelmässä usein esiintyy. On kuitenkin syytä huomioida, että media on jatkuvassa muutoksessa, ja yli kahdenkymmenen vuoden takaiset määritelmät medialukutaidosta pohjautuvat monilta osin hyvin erilaiseen mediaan, kuin millaisena sen tänä päivänä tunnemme.

Terras ja kollegat (2011) korostavat psykologisen näkökulman merkitystä medialukutaidolle, sillä se auttaa ymmärtämään medialukutaitoa ja siihen liittyviä kyvykkyyksiä kokonaisvaltaisemmin. Myös Potter (2004) lähestyy medialukutaitoa psykologisesta, tarkemmin kognitiivisesta, näkökulmasta. Potter (2004) ehdottaa medialukutaidolle kognitiivista mallia, jonka mukaan medialukutaito koostuu neljästä pääfaktorista: tietorakenteista, päätösten motivoinnista, informaatioprosessoinnin työkaluista ja informaatioprosessoinnin tehtävien virtauksesta. Potterin (2004) mukaan nämä neljä faktoria toimivat vuorovaikutuksessa toistensa kanssa. Olennaista teoriassa on se, että informaation prosessointi voi tapahtua tietoisesti tai tiedostamattomasti ja medialukutaito vaatii tietoista prosessointia (Potter, 2004). Potterin kognitiivinen malli syventyy tarkemmin medialukutaitoon ja tarjoaa kokonaisvaltaisemman selityksen medialukutaidon määritelmään usein liitetystä informaationkäsittelytaitojen taustasta.

Potterin (2004) mukaan medialukutaito on erityisen tärkeää vallitsevan informaatiotulvan takia; media tarjoaa niin suuren määrän informaatiota, että ihmisten on pakko prosessoida osa siitä automaattisesti. Automaattisella prosessoinnilla hän tarkoittaa selaamista "autopilotilla", kunnes jokin olennainen kiinnittää huomion. Siinä missä informaation automaattinen prosessointi auttaa välttämään

ei-toivottuja viestejä, se myös nostaa väärinymmärrysten ja vääristyneiden kokonaiskuvien muodostumisen riskiä ja heikentää merkitysten muodostamisen taitoa (Potter, 2004). Hän huomauttaa, että medially on hyvin erilaiset motivaatiot informaation esittämiseen, kuin ihmisillä informaation vastaanottamiseen, ja automatisoimalla informaation prosessointia päädymme vastaamaan median tavoitteisiin omien tavoitteidemme kustannuksella. Heikentyneet taidot taas johtavat Potterin (2004) mukaan siihen, että medially on yhä suurempi vaikutusvalta tärkeiden asioiden määrittäjänä. Potterin kaksikymmentä vuotta vanha jäsennys median ja ihmisten suhteesta ja medialukutaidon merkityksestä on mielestäni edelleen erityisen ajankohtainen, etenkin lyhytvideoiden vallatessa alaa jokaisella alustalla ja generatiivisen tekoälyn mahdollistaessa entistä nopeamman informaatiomassan tuottamisen.

Sen lisäksi, että medialukutaito auttaa yksilöä vastaamaan median vyöryttämään informaatiotulvaan, se edistää digitaalista turvallisuutta (esim. Lee, 2018). Medialukutaito auttaa yksilöitä muun muassa tulkitsemaan ja analysoimaan mediaa, mikä edistää esimerkiksi valeuutisten, disinformaation ja huijausten tunnistamista (Lee, 2018). Lee (2018) huomauttaa, että tutkimus medialukutaidon kentällä keskittyy pääosin nuoriin ja lapsiin, mutta koska juuri aikuiset ovat usein erilaisten huijausten ja valeuutisten kohderyhmää, aikuisten medialukutaidon tutkiminen ja edistäminen on myös tärkeää.

Kuten Lewiskin (2021) mainitsee, elämme median kyllästävässä maailmassa, ja halusimme tai emme, se vaikuttaa meihin. Lewis (2021) huomauttaa, että medialukutaito ei auta ihmisiä huomioimaan median ja teknologiantäyteen maailman vaikutuksia itseensä, vaan keskittyy enemmänkin median käyttämisen, tulkitsemisen ja muiden medialukutaitoon liittyvien kyvykkyyksien kehittämiseen. Terras ja kollegat (2011) mainitsevat vuonna 2011, että "uuden median" vaikutusten tutkiminen on hyvin alkuvaiheissa, ja voitaisiin ajatella, että tämä ajatus pätee edelleen median muuttuessa jatkuvasti nopeaan tahtiin. Median vaikutuksia ihmisiin ei siis ehkä tunneta kovin kokonaisvaltaisesti, etenkin median nopean muutoksen takia, eikä medialukutaitokaan välttämättä auta suhtautumaan näihin vielä jokseenkin tuntemattomiin vaikutuksiin. Medialukutaitoa muiden lukutaitokäsitteiden ohella on siis myös kritisoitu, mutta se on kuitenkin erityisen tärkeä

taito muun muassa siksi, että se auttaa yksilöitä navigoimaan mediassa ja käsittelemään median tarjoamaa informaatiotulvaa sekä edistää digitaalista turvallisuutta.

2.2 Generatiivinen tekoäly

Kuten olen edellisessä luvussa todennut, media on jatkuvassa muutoksessa. Eräs merkittävä muutos on generatiivisen tekoälyn kehitys ja yleistyminen erilaisilla elämän osa-alueilla, muun muassa myös kasvatuksessa ja koulutuksessa (Lim ym., 2023). Käsittelen tässä luvussa pintapuolisesti generatiivista tekoälyä ja sen jälkeen tekoälyyn liittyvää keskustelua ja aiempaa tutkimusta kasvatustieteiden kentällä.

Tekoälyllä on pitkä historia, mutta viime vuosien polttava puheenaihe on ollut *generatiivinen tekoäly*. Limin ja kollegojen (2023) mukaan generatiivinen tekoäly voidaan määritellä teknologiaksi, joka luo annettujen ohjeiden perusteella ihmismäistä sisältöä, kuten tekstiä ja kuvia. Generatiivinen tekoäly on kokenut merkittäviä harppauksia viime vuosina uusien tekniikoiden ansiosta (Wach ym., 2023). Suuret kielimallit (*Large Language Models, LLM*), kuten ChatGPT ja Google BARD pystyvät muun muassa tuottamaan ja tiivistämään tekstiä (Tiernan ym., 2023). Erilaiset tekoälyohjelmat voivat generoida muun muassa kuvia, videoita ja musiikkia (Tiernan ym., 2023). Generatiivinen tekoäly kehittyy internetistä ja mediasta saadun valtavan datamäärän avulla ja luo sisältöä tunnistamalla ja toistamalla yleisiä, toistuvia kaavoja (Fengchun & Holmes, 2023). Generatiivinen tekoäly on monimutkainen tekoälyteknologia, jonka taustoihin ja komponentteihin ei kuitenkaan ole mahdollista syventyä sen tarkemmin tämän tutkielman puitteissa.

Generatiivisen tekoälyn kehitykseen liittyy monia etuja ja mahdollisuuksia, jotka voivat tehostaa ihmisten toimintaa ja helpottaa monenlaisia asioita (Wach ym., 2023), mutta toisaalta kehitys on kuitenkin tuonut mukanaan myös negatiivisia asioita, kuten valtavan aallon disinformaatiota (Wach ym., 2023; Fengchun & Holmes, 2023; Frau-Meigs, 2023). Wachin ja kollegojen (2023) mukaan esimerkiksi ChatGPT on hyvin altis väärinkäytölle. Disinformaation lisäksi muita generatiivisen tekoälyn tuomia ongelmia ovat esimerkiksi syväväärengökset (*deepfake*) ja vääristymät algoritmeissa (Wach ym. 2023). Lim ja kollegat (2023)

puhuvatkin generatiivisen tekoälyn paradokseista: se on samanaikaisesti sekä ystävällinen että vihollinen.

Tiernanin ja kollegojen (2023) mukaan aiempi tekoölyyn liittyvä tutkimus kasvatustieteiden kentällä on viime aikoina painottunut työkalujen, kuten ChatGPT:n, vaikutuksiin muun muassa arviointiin ja opetukseen. Vaikka näkökulmani on aikuisissa, käsittelen kuitenkin tässä luvussa joitain kouluun ja opetukseen liittyviä tutkimuksia havainnollistaakseni tekoälyn tutkimuksen tilannetta kasvatustieteiden kentällä. Tekoölyä on tarkasteltu viime aikoina koulutuksen kontekstissa esimerkiksi sen potentiaalinen ja riskien näkökulmista. ChatGPT:n julkaisun jälkeen kasvatustieteiden kentällä heräsi nopeasti keskustelua generatiivisella tekoölyllä huijaamisesta: generatiivinen tekoöly voi esimerkiksi kirjoittaa kirjoitustehtäviä oppilaiden puolesta ja sitä on hankala jäljittää (Cope & Kalantzis, 2024). Cope ja Kalantzis (2024, s. 5) nostavat esiin listan muitakin generatiiviseen tekoölyyn liittyviä ongelmia kasvatustieteellisestä näkökulmasta: 1. se ei tunnista totuutta valheesta, 2. se ei pysty kertomaan käyttämiään lähteitä luotettavasti, 3. se ei ymmärrä kuinka ihmistietämys rakentuu, 4. se sanoo loukkaavia ja vaarallisia asioita, jos ihmiset eivät aseta sille rajoitteita, 5. se esittää ihmistä ja 6. se varastaa sisältöä sekä loukkaa yksityisyyttä.

Ciampa ja kollegat (2023) pohtivat tekoälyn hyötyjä ja riskejä nuorten ja aikuisten koulutuksen kontekstissa. Tekoölyllä on potentiaalia tukea muun muassa luovuutta ja kriittistä ajattelua sekä avustaa tiedonhaussa (Ciampa ym., 2023). Toisaalta se herättää eettisiä kysymyksiä liittyen muun muassa turvallisuuteen ja luotettavuuteen (Ciampa ym., 2023). He toteavat, että on tärkeää, että tekoöly toimii ajattelua tukevana työkaluna, ei sen korvaajana. Mahdollisuuksia hyödyntää tekoölyä opetuksessa on siis varmasti monenlaisia, mutta ne ovat vielä pitkälti tutkimattomia (Jones, 2024). Łodzikowskin ja kollegojen (2024) mukaan generatiivinen tekoöly avaa mahdollisuuksia ainakin oppilaiden osaamisen syvällisemmälle arvioinnille ja uudelleinteraktiiviselle opiskelulle. He korostavat, että tekoälyn tarkoituksena ei kuitenkaan ole korvata opettajaa, vaan toimia opettajan työtä tehostavana työkaluna. Myös Jonesin (2024) mukaan tekoölyllä on potentiaalia helpottaa ylikuormitettujen opettajien työtaakkaa.

Tekoöly on herättänyt kasvatustieteissä keskustelua myös eettisten kysymysten osalta (esim. Lähdesmäki, 2024; Jones, 2024). Lähdesmäki (2024) ko-

rosta tekoälyyn liittyvän vankan eettisen perustan luomisen tarvetta kasvatuksen ja koulutuksen kontekstissa, jotta sitä voitaisiin hyödyntää ihmisten ajattelua tukevana työkaluna. Myös Łodzikowski ja kollegat (2024) huomauttavat, että koulutuksessa on tekoälyn suhteen nojaututtava ulkoisiin, esimerkiksi Euroopan unionin asettamiin, eettisiin ohjeistuksiin. Jones (2024) nostaa esille suhtautumistapojen merkityksen tekoälyn hyödyntämisen ja eettisyyden kannalta: jos esimerkiksi koulun suhtautuminen tekoälyyn on negatiivinen, he voivat jäädä paitsi tekoälyn hyödyistä, kun taas liian optimistinen suhtautuminen voi sokaista mahdollisille riskeille ja ongelmakohtille. Jonesin (2024) mukaan olennaista on, että eettisten kysymysten pohtiminen on erityisen tärkeää siinä vaiheessa, kun tekoälyohjelmia ja hyödyntämistapoja koulutukseen suunnitellaan. Lisäksi Frau-Meigs (2023) nostaa ajatuksen siitä, että generatiivisen tekoälyn eettinen hyödyntäminen vaatii media- ja informaatiolukutaitoa.

Esimerkiksi Lähdesmäki (2024), Jones (2024) ja Łodzikowski ja kollegat (2024) peräänkuuluttavat siis eettisyyden huomioimista, kun tekoälyä hyödynnetään koulutuksessa. Tekoäly ja siihen liittyvät eettiset kysymykset ovat varmasti olennaisia erilaisissa viitekehyksissä. Tiernan ja kollegat (2023) tarkastelevat tutkimuksessaan digilukutaidon viitekehyksiä ja sitä, kuinka niissä huomioidaan informaatiolukutaito, medialukutaito sekä tekoäly. Katsauksessa selvisi, että viitekehykset ovat reagoineet hitaasti tekoälyn tuomiin muutoksiin (Tiernan ym., 2023). Hidas reagointi tekoälyn tuomiin muutoksiin ei varmasti edistä myöskään eettisten näkökulmien huomiointia.

Generatiivisen tekoälyn viimeaikaisiin muutoksiin liittyy siis sekä mahdollisuuksia että uhkia, niin yksilöiden jokapäiväisessä median käytössä, kuin kasvatuksen ja koulutuksen kontekstissakin. Tekoälyn kohtaamiseen tarvitaan media-lukutaitoa, ja tekoälyn lukutaitoon onkin viitattu tekoälylukutaito -käsitteellä. Muun muassa Long ja Magerko (2020) ja Ng ja kollegat (2021) käyttävät käsitettä tekoälylukutaito (*AI-literacy*), mutta eivät varsinaisesti rinnasta sitä medialukutaitoon tai osaksi sitä. Kuvailen tekoälylukutaitoa tarkemmin seuraavassa luvussa.

2.3 *Tekoälylukutaito*

Voisi ajatella, että viimeaikaisiin generatiiviseen tekoälyyn liittyviin muutoksiin suhtautuminen vaatii erityisesti tekoälylukutaitoa. Tekoälylukutaito on kuitenkin

suhteellisen uusi käsite ja sen määritelmät vaihtelevia luultavasti muun muassa tekoälyn nopean ja jatkuvan kehityksen vuoksi. Okkonen ja Kotilainen (2019) kirjoittavat jo vuonna 2019 tarpeesta ottaa tekoäly huomioon medialukutaidon määritelmässä ja opetuksessa. Tiernan ja kollegat (2023) taas mainitsevat, että tekoälyn vaikutukset media- ja informaatiolukutaitoihin ovat vielä melko tuntemattomia, sillä tekoäly on tuore tutkimusalue. Frau-Meigs (2023) argumentoi, että usein eriteltyt lukutaidot, kuten algoritmilukutaito, datalukutaito ja tekoälylukutaito voidaan käsittää osaksi media- ja informaatiolukutaidon ”paradigmaa”. Käsitän myös tässä tutkielmassa tekoälylukutaidon osaksi laajempaa medialukutaitoa, mutta tarkastelen sitä kuitenkin tässä luvussa medialukutaidosta erillisenä.

Tekoälylukutaidolle ei medialukutaidon tapaan ole vakiintunut yhteisesti hyväksyttyä määritelmää (Laupichler ym., 2022; Laupichler ym., 2023). Kuitenkin eräs määritelmä, johon usein viitataan, on Longin ja Magerkon (2020) muodostama määritelmä (Laupichler ym., 2022; Laupichler ym., 2023). Longin ja Magerkon (2020, s. 2) määritelmä kuuluu vapaasti suomennettuna seuraavasti: joukko kyvykkyyksiä, joka antaa yksilölle mahdollisuudet arvioida kriittisesti tekoälyteknologioita; kommunikoida ja tehdä yhteistyötä tehokkaasti tekoälyn kanssa; käyttää tekoälyä työkaluna verkossa, kotona ja työpaikalla. Long ja Magerko (2020, s. 3) jäsentävät tekoälylukutaitoa viiden teeman avulla: Mikä on tekoäly? (*What is AI?*); Mitä tekoäly voi tehdä? (*What can AI do?*); Miten tekoäly toimii? (*How does AI work?*); Kuinka tekoälyä tulisi käyttää? (*How should AI be used?*); Kuinka ihmiset käsittävät tekoälyn? (*How do people perceive AI?*). Jokaiseen teemaan liittyy joukko tarkempia kykyjä: esimerkiksi ensimmäisen teeman *Mikä on tekoäly?* yksi kyvykkyyks on tekoälyn tunnistaminen, joka vaikuttaisi olevan tämän tutkielman kannalta keskeisin osa-alue.

Ng ja kollegat (2021) kokoavat yhteen tekoälylukutaidon määritelmiä kirjallisuuskatsauksessaan. He (s. 4) järjestävät tekoälylukutaidon osa-alueita hierarkisesti: alimmilla osaamisen tasoilla ovat tekoälyyn liittyvä tietämys ja ymmärrys (*know and understand AI*), keskellä tekoälyn käyttäminen ja hyödyntäminen (*use and apply AI*) ja ylimpänä tekoälyn arviointi ja luominen (*evaluate and create AI*). Vaikka Ng ja kollegat (2021) tarkastelevat nimenomaan kasvatustieteellistä tutkimusta, määritelmä, jonka he ovat koonneet keskittyy kuitenkin pitkälti tekoälyn käyttöön ja teknologiaan. Se sivuuttaa muita esimerkiksi medialukutaidolle tyypil-

lisiä tunnistamiseen, analysoimiseen ja tulkitsemiseen liittyviä lukutaitokyvykkyksiä. Syynä tähän lienee se, että generatiivinen tekoäly on kehittynyt merkittävästi ja sen myötä alkanut näkymään selkeämmin mediassa vasta aivan viime vuosina. Mielestäni voitaisiinkin ajatella, että viime vuosien kehitys on nostanut medialukutaitomaisen tunnistamisen, tulkitsemisen ja analysoimisen merkitystä osana tekoälylukutaitoa, sillä generatiivinen tekoäly on saatavilla kaikille ja sen myötä netti ja media ovat pullollaan generatiivisella tekoälyllä tuotettua sisältöä. Tekoäly on kehittynyt merkittävästi viime aikoina, joten tekoälylukutaidon määritelmiä muutaman vuoden takaa on syytä tarkastella viimeaikaiset muutokset mielessä pitäen, ja siksi aiheetta on myös mielekästä tutkia.

Frau-Meigs (2023, s. 13) ehdottaa media- ja informaatiolukutaidon kanssa linjassa olevaa viitekehystä tekoälykyvykkyyksille. Tämä viitekehys jakautuu neljään kategoriaan: tieto, taidot, asenteet ja arvot. Tieto-kategoria koostuu muun muassa tekoälyn potentiaalin ja riskien tunnistamisesta ja tekoälyn taustalla vaikuttavien toimijoiden, omistajuuksien ja motivaatioiden tunnistamisesta. Taidot-kategoria taas koostuu muun muassa tekoälytyökalujen vastuullisesta käytöstä, ei-inhimillisten toimijoiden ja objektien kanssa toimimisesta ja informaation lähteiden tarkistamisesta. Asenteet-kategoriaan kuuluu muun muassa datan ja sen käyttötarkoitusten luova ja kriittinen tulkitseminen, haitallisen sisällön tunnistaminen ja riskien välttäminen. Arvo-kategoriaan taas kuuluvat muun muassa mielipiteen- ja sananvapauden sekä tasa-arvon edistäminen, yksityisyydensuojan tukeminen ja massavalvonnan vastustaminen (Frau-Meigs, 2023). Frau-Meigsin (2023) viitekehys eroaa edellä mainituista tekoälylukutaidon jäsennyksistä muun muassa siten, että se ei painotu yhtä vahvasti tekoälyn käyttämiseen tai toimintaperiaatteiden ymmärtämiseen.

Lähdesmäki (2024) taas puhuu tekoälytoimijuudesta: kyvyistä ja taidoista käyttää tekoälyä, sekä eettisestä älykkyydestä. Lähdesmäen (2024) mukaan eettisen perustan luominen kasvatuksessa ja koulutuksessa on avain tekoälyn hyödyntämiseen ihmisiä ja yhteiskuntaa tukevalla tavalla. Lähdesmäen (2024) kuvailema tekoälytoimijuus on mielestäni rinnastettavissa tekoälylukutaitoon, ja se eroaa edellä mainituista jäsennyksistä siten, että se korostaa eettisen älykkyyden keskeistä roolia.

Laupichler ja kollegat (2022) ovat tarkastelleet tekoälylukutaitoa aikuisten näkökulmasta kirjallisuuskatsauksessaan. He huomauttavat, että esimerkiksi

Long ja Magerko (2020) sekä Ng ja kollegat (2021) ovat tekoälylukutaidon määritelmässään sitä mieltä, ettei siinä ole kyse esimerkiksi ohjelmoinnista vaan enemmänkin tekoälykonseptien ymmärtämisestä. Laupichlerin ja kollegojen (2022) mukaan tekoälylukutaidon käsitteen sisältö vaatii yhtenäistämistä, ja päätöksiä sen suhteen ovatko esimerkiksi ohjelmointitaidot osa sitä. Laupichler ja kollegat (2023) ovat pyrkineet myöhemmin muodostamaan viitekehystä tekoälylukutaidon arviointiin, vaikka toteavat edelleen, ettei tekoälylukutaidolle ole selkeää määritelmää.

Tekoälylukutaidolla tarkoitetaan siis useimmiten tekoälyn hyödyntämiseen liittyviä kyvykkyyksiä, mutta käsitteelle ei ole yhtä vakiintunutta määritelmää ja eri määritelmien välillä esiintyy sisällöllistä vaihtelua. Vaikuttaa siltä, että tekoälyn nopea kehitys lisää edelleen jatkuvaa uudelleenmäärittelyn tarvetta ja esimerkiksi generatiivisen tekoälyn medialukutaitomaisen tulkinnan ja arvioinnin korostamista. Monet keskeiset jäsenyykset ja määritelmät ovat useamman vuoden takaisia, ja viime vuosien nopeat muutokset muun muassa generatiivisen tekoälyn toiminnassa ja saavutettavuudessa ovat nostaneet esiin tarpeen uudentilaisille tekoälykyvykkyyksille. Tästä syystä tekoälyn tunnistamisen näkökulmaa on mielekästä tarkastella tässä tutkielmassa.

3 ELÄYTYMISMENETELMÄ

Tämän tutkielman tavoitteena on siis vastata tutkimuskysymykseen: *millaisia tekoälyn tunnistamisen keinoja aikuisten kertomuksissa ilmenee?* Eläytymismenetelmä sopii tekoälylukutaidon tutkimiseen hyvin, sillä se mahdollistaa ilmiöön syventymisen tutkittavien kokemusperustan ja ajatusmaailman muodostaman linssin läpi. Mönkkösen (2021) mukaan eläytymismenetelmä toimii myös hyvin erilaisten käännekohtien tutkimisessa, jollaiseksi myös viimeaikaisen generatiivisen tekoälyn kehityksen ja yleistymisen voisi mieltää. Käsittelen ensin yleisemmin eläytymismenetelmää sekä sen taustaa ja luonnetta alaluvussa 3.1. Sen jälkeen, alaluvussa 3.2, kerron eläytymismenetelmästä tässä tutkielmassa.

3.1 Eläytymismenetelmä

Eläytymismenetelmä (*the method of empathy-based stories, MEBS*) on laadullinen tutkimusmenetelmä, joka perustuu tutkittavien kirjoittamiin lyhyisiin kertomuksiin, jotka on kirjoitettu tutkijan antaman kehyskertomuksen tilanteeseen eläytyen (Eskola, 1997; Eskola ym., 2017b; Wallin ym., 2018). Menetelmänä erityisen eläytymismenetelmästä tekee variointi eli kehyskertomusten pieni keskinäinen erilaisuus; kehyskertomuksissa on jokin elementti, jota varioidaan, ja jonka kertomuksissa aiheuttamia muutoksia lopuksi tarkastellaan (Eskola, 1997). Kehyskertomuksia on siis tyypillisesti vähintään kaksi, ja ne eroavat toisistaan yhden muuttujan osalta; tarkoituksena on tarkastella sekä yksittäisten tarinoiden sisältöä että eroja eri kehyskertomuksiin pohjautuvien tarinoiden välillä (Eskola ym., 2017a). Yksi vastaaja kirjoittaa kertomuksen tyypillisesti yhden kehyskertomuksen perusteella, eikä ole tietoinen toisista kehyskertomuksista (Wallin ym., 2018). Menetelmän etu on sen joustavuus, sillä se sopii monenlaisiin aihepiireihin (Eskola ym., 2017a), sekä monenlaisille ikäryhmille (Wallin ym., 2018). Aineiston kerääminen eläytymismenetelmällä voi onnistua hyvinkin nopeasti: vastausaikaa annetaan tyypillisesti noin 15–30 minuuttia (Wallin ym., 2018).

Eläytymismenetelmä tarjoaa mahdollisuuden monenlaisille näkökulmille, sillä se antaa vastaajalle huomattavasti enemmän vapautta kuin esimerkiksi kyselylomake valmiine vaihtoehtoineen (Eskola, 1997; Eskola ym., 2017b). Eläytymismenetelmä tuottaa tietoa, joka ei välttämättä vastaa todellisuutta, mutta enemmänkin kertoo, miten asiat voisivat olla tai mikä olisi mahdollista (Eskola, 1997; Eskola ym., 2017b), ja kuinka vastaaja käsittää kehyskertomuksessa kuvatun tilanteen (Mönkkönen, 2021). Mönkkönen (2021) korostaa menetelmällä saatavien jäsennysten, oletusten, ja mahdollisten tulevaisuuskuvien mielenkiintoisuutta elämän erilaisten muutosten äärellä: yllättävät käännekohdat elämässä, kuten kriisitilanteet tai sota, synnyttävät erilaisia skenaarioita tulevaisuudesta. Eläytymismenetelmä ei siis tuota yleistettävää faktatietoa todellisuudesta vaan tietoa tiettyyn aiheeseen liittyvistä käsityksistä, odotuksista ja arvoista (Wallin ym., 2018).

Vaikka eläytymismenetelmälle ominainen kuvittelu ei tarjoa varsinaista tietoa siitä, kuinka asiat ovat, se avaa erilaisia näkökulmia. Mönkkönen (2021) huomauttaa, että “kun toimitaan niin sanotusti varman päälle, saatamme sulkea ulkopuolelle kiinnostavia seikkoja todellisuudesta”. Siis unohtamalla yleistettävien tulosten tuottamisen vaateen, se mahdollistaa uudenlaisia näkökulmia ilmiöihin ja ihmisten ajatuksiin ja kokemuksiin. Eläytymismenetelmä vaikuttaa siis tarjoavan mahdollisuuksia näkökulmiin, joita muilla menetelmillä ei välttämättä voida saavuttaa. Se antaa myös vastaajille erityisen mahdollisuuden käyttää mielikuvitustaan ja vastata ehkä vapaammin kuin esimerkiksi haastattelutilanteessa.

3.1.1 Eläytymismenetelmän tausta

Eläytymismenetelmä rantautui Suomeen 80-luvulla amerikkalaisesta *role playing* -menetelmästä vaikutteita saaneena (Mönkkönen, 2021). Eskolan (1997) mukaan eläytymismenetelmän alkuvaiheissa se voitiin jakaa kahteen vaihtoehtoon: aktiiviseen (*active role-playing*) ja passiiviseen (*passive role-playing*). Aktiivinen vaihtoehto tarkoittaa roolileikin esittämistä annetun tilanteen perusteella, kun taas passiivinen vaihtoehto tarkoittaa nykyään paremmin eläytymismenetelmänä tunnettua kertomuksen kirjoittamista tilanteeseen eläytyen (Eskola, 1997). Eläytymismenetelmää kehiteltäessä oli tarkoituksena löytää eettinen tapa käyttää koellisen tutkimuksen logiikkaa huijaamatta tai manipuloimatta tutkittavia (Eskola,

1997). Esimerkkinä tällaisesta huijaamisesta Eskola käyttää Milgramin kuuluisaa tutkimusta, jossa tutkimus tottelevaisuudesta naamioitiin oppimiskokeeksi ja jossa tutkittavia pyydettiin antamaan sähköiskuja rangaistuksena väärin menneistä vastauksista toiselle ihmiselle (Milgram, 1984; Eskola, 1997, s. 10, mukaan). Eläytymismenetelmä tarjoaa siis paljon eettisemmän vaihtoehdon kokeelliselle tutkittavien manipuloinnille (Eskola, 1997).

Mönkkönen (2021) korostaa menetelmän tulkitsevaa luonnetta vertaamalla eläytymismenetelmää psykologian projektiivisiin testeihin. Ero psykologisiin testeihin on kuitenkin Mönkkösen (2021) mukaan siinä, että eläytymismenetelmässä tarkoituksena ei ole tehdä tulkintoja tarinan taustalla vaikuttavista (psykologisista) tekijöistä vaan tarkastella itse tarinaa ja sen sisältöä. Toisaalta Heikkinen nostaa ajatuksen siitä, että narratiivisessa tutkimuksessa tarkastellaan pelkkien kertomusten sijaan koko kerronnallista prosessia (Bold, 2012, s. 17; Heikkinen, 2018, mukaan). Eläytymismenetelmän kohdalla voitaisiin kuitenkin kyseenalaistaa, mitä kerronnalliseen prosessiin oikeastaan itse kertomuksen lisäksi sisältyy, ja kuinka sitä voidaan tarkastella nimettöminä ja valmiina vastaanotetuissa kertomuksissa.

Mönkkösen (2021) mukaan eläytymismenetelmä tosiaan kuuluu narratiivisiin tutkimusmenetelmiin. Narratiivisuudella tarkoitetaan tutkimusaiheiden tarkastelemista kertomusten kautta (Eskola & Suoranta, 1998). Narratiivista tutkimusta pohjustavat epistemologiset ja ontologiset taustaoletukset eroavat selkeästi aiemmin vallinneesta objektiivisuutta ihannoivasta paradigmasta (Heikkinen, 2018). Eläytymismenetelmän taustalla vaikuttavat konstruktivistiset ja relativistiset epistemologiat (Wallin, ym., 2018); tietämisen suhteellisuutta ja subjektiivisuutta korostavat käsitykset (Heikkinen, 2018). Konstruktivismilla viitataan käsitykseen tiedon muodostamisesta, jonka mukaan ihminen rakentaa tiedon aieman tiedon ja kokemuserustan päälle (Heikkinen, 2018). Tieteenfilosofiset käsitykset eläytymismenetelmän taustalla vaikuttavat siis nojaavan vahvasti käsitykseen kokemuksen subjektiivisuudesta.

Eläytymismenetelmä itsessään on siis suhteellisen uusi tutkimusmenetelmä, jonka juuret ovat roolileikissä ja narratiivisessa tutkimuksessa. Taustalla vaikuttavien tekijöiden värittämä eläytymismenetelmätutkimus on tulkitsevaa ja konstruktionismiin, kokemuksen subjektiivisuuteen, eläytymiseen ja kerronnallisuuteen pohjautuvaa.

3.1.2 Eläytymismenetelmällä saatujen tutkimustulosten luonne ja merkitys

Eläytymismenetelmä ei siis tuota faktatietoa, vaan esimerkiksi tietoa siitä miten asiat voisivat olla, minkälaisia käsityksiä asioista tai ilmiöistä nousee kertomuksissa tai miten asioita jäsennetään kertomuksissa (ks. esim. Mönkkönen, 2021; Eskola, 1997). Eläytymismenetelmäaineiston luotettavuutta on siis usein turha mitata kertomusten todenperäisyydellä, sillä sen ei ole tarkoitus tuottaa objektiivisia, yleistettäviä tuloksia vaan ennemminkin katsauksia ilmiöihin tutkittavien ajatusten ja näkemysten kautta. Vaikka eläytymismenetelmä ei tuota objektiivista ja yleistettävää, perinteisin tieteellisen tutkimuksen kriteerein luotettavaa tietoa, se luo silti merkityksellisiä ikkunoita ilmiöihin erityislaatuaisella tavalla.

Eläytymismenetelmään liittyvästä kirjallisuudesta kumpuaa mielestäni erityisesti kolme eläytymismenetelmää ja sen tuottamaa tietoa kuvastavaa piirrettä: tulkinnallisuus, subjektiivisuus, erityislaatuisuus. Eläytymismenetelmä pohjautuu narratiiviseen tutkimukseen ja roolileikkiin, mikä viittaa menetelmän tulkinnalliseen luonteeseen. Eläytymismenetelmällä kerätty aineisto vaatii syvällisempää tulkintaa: tekemällä tulkintoja vastaajien kirjoittamista kertomuksista, voidaan löytää yhteyksiä, näkemyksiä ja oivalluksia tutkittaviin teemoihin liittyen.

Subjektiivisuudella viitataan siihen, ettei menetelmän tarkoituksena ole tuottaa objektiivista tai yleistettävää tietoa. Sen aineisto koostuu ihmisten yksilöllisistä kokemuksista, elämistä ja ajatusmaailmoista kumpuaviin kertomuksiin, jotka saattavat perustua todellisiin tapahtumiin tai saattavat olla täyttä mielikuvituksen tuotetta tai jotain siltä väliltä. Eläytymismenetelmällä kerätystä aineistosta ei siis voi vetää yleistettäviä johtopäätöksiä asioista tai ilmiöistä. Sen sijaan tarkoituksena on avata ikkunoita tutkittaviin teemoihin ihmisten kirjoittamien kertomusten kautta: herätellä uudenlaisia näkökulmia ja ajatuksia.

Erytislaatuaisuudella taas tarkoitan eläytymismenetelmän kokeellista luonetta, jonka kehyskertomusten variointi mahdollistaa. Kehyskertomusten variointi erottaa eläytymismenetelmän monista muista laadullisen tutkimuksen menetelmistä. Tarkastelemalla ja vertailemalla eri variaatioihin perustuvien kertomusten välisiä eroja ja yhteneväisyyksiä päästään yhä syvemmälle tutkittavaan ilmiöön. Kuvaillemiäni piirteiden yhdistelmä luo siis ainutlaatuisen tutkimusmenetelmän,

jolla päästään syventymään erilaisiin ilmiöihin ihmisten yksilöllisten ajatusmaailmojen ja jäsenystapojen muodostaman linssin läpi. Eläytymismenetelmä mahdollistaa siis itse ilmiön tarkastelemisen, mutta sen lisäksi myös kenties vallitsevan ajan, kulttuurin ja paradigmojen heijastumien tulkitsemisen vastaajien kertomuksissa. Voitaisiin siis ajatella, että eläytymismenetelmän avulla tuotettu tieto on yhdistelmä kertomusten kirjoittajien ja tulkintoja tekevän tutkijan näkemyksiä maailmasta.

3.2 Eläytymismenetelmä tässä tutkielmassa

Kuten edellä mainitsin, eläytymismenetelmä sopii hyvin elämän käännekohtien tarkasteluun. Generatiivisen tekoälyn nopea kehitys ja uudenlainen saavutettavuus kaikille ihmisille, sekä sen yleistynyt näkyvyys mediassa, ovat kokonaisuutena mielestäni ilmiö, jota voisi luonnehtia käännekohtaksi ainakin digitaalisessa elämässä. Pyrkimyksenäni on luoda katsaus aikuisten käsityksiin tekoälyn tunnistamisesta mediassa ja samalla mahdollisesti ilmeneviin asenteisiin ja käsityksiin tekoälystä. Eläytymismenetelmä mahdollistaa juuri tällaisen katsauksen tekemisen ihmisten ajatuksiin, käsityksiin ja jäsenyyksiin ilmiöistä. Kuvailen eläytymismenetelmän soveltamista tässä tutkielmassa tarkemmin seuraavassa luvussa.

4 TUTKIELMAN TOTEUTUS

Kuvailen tässä luvussa tutkielman toteutuksen etenemistä. Kerron ensin luvussa 4.1 kehyskertomusten muodostamisesta ja varioinnista. Luvussa 4.2 kuvailen aineistonkeruun vaiheita tässä tutkielmassa ja lopuksi kerron aineiston analyysistä luvussa 4.3.

4.1 Kehyskertomukset

Tutkijan luomat kehyskertomukset ovat avainasemassa eläytymismenetelmässä. Kehyskertomusten luomiselle ei ole yleistä ohjeistusta, sillä niiden sisältö riippuu pitkälti tutkimuskysymyksistä (Wallin ym., 2018). Tyypillisesti kehyskertomukset ovat kuitenkin lyhyitä, niitä on vähintään kaksi, ja ne eroavat toisistaan yleensä vain yhden muuttujan osalta ja ovat muuten keskenään samankaltaisia (Wallin ym., 2018). Kehyskertomuksen laatimisessa tulee pohtia muun muassa kohde-ryhmää sekä kehyskertomuksen ymmärrettävyyttä ja selkeyttä (Eskola ym., 2017b). On hyvä pohtia, onko kehyskertomukseen tarkoitus vastata esimerkiksi minä -muodossa vai onko kehyskertomuksessa esiintyvällä henkilöllä nimi (Eskola ym., 2017b). Jos henkilöllä on nimi, tulee pohtia, minkälaisia ajatuksia nimi vastaajassa herättää esimerkiksi sukupuolen tai etnisyyden osalta, sillä mahdollisimman ”neutraali nimi jättää tilaa vastaajan mielikuvitukselle” (Eskola ym., 2017b). Kehyskertomuksen on hyvä olla pelkistetty, mutta on tärkeää huomioida, ettei kehyskertomus tuota liian rajaamattomia vastauksia, ja siten jätä vastaamatta tutkimuskysymyksiin (Eskola ym., 2017b).

Tämän tutkielman tarkoituksena on selvittää, millaisia tekoälyn tunnistamisen keinoja aikuisten kertomuksissa ilmenee, joten olen pyrkinyt muodostamaan kehyskertomukset siten, että ne johdattelevat vastaajien kertomuksia haluttuun suuntaan:

(A) Tuttavasi kertoo sinulle, että törmäsi taannoin netissä erääseen julkaisuun, ja epäili, että tekoäly on tehnyt sen. Hän kertoo, millainen julkaisu oli

ja mitä hän oli ajatellut pohtiessaan, oliko se todella tekoälyn luoma. Käytä mielikuvitustasi ja kerro keskustelustanne.

(B) Tuttavasi kertoo sinulle, että törmäsi taannoin netissä erääseen julkaisuun, ja oli varma, että tekoäly on tehnyt sen. Hän kertoo, millainen julkaisu oli ja mitä hän oli ajatellut pohtiessaan, oliko se todella tekoälyn luoma. Käytä mielikuvitustasi ja kerro keskustelustanne.

Molemmat kehyskertomusvariaatiot keskittyvät tilanteeseen, jossa vastaaja keskustelee tuttavansa kanssa, joka kertoo törmänneensä netissä erääseen julkaisuun. Kehyskertomuksessa A tuttava *epäilee* julkaisun olevan tekoälyn tuotos, kun taas kehyskertomuksessa B hän *on varma*, että se on tekoälyn tuotos. Julkaisun alustaa tai sisältöä ei ole tarkennettu, jotta vastaukset eivät rajautuisi esimerkiksi pelkkää tekstiä tai kuvaa koskeviksi. Kehyskertomukset ohjaavat vastaajaa eläytymään keskusteluun tuttavalla tarinallisuuden korostamiseksi. Henkilö, joka kuvailee kertomuksissa julkaisua ja tekoälyn tunnistamista on tuttava, eikä vastaaja itse, kuvitteellisen tilanteen keksimisen sekä eläytymisen helpottamiseksi. Oletuksena siis on, että kehyskertomuksien ohjaamana vastaajien kertomuksissa ilmenee keinoja, joita tekoälyn tunnistamiseen käytetään. Keskusteluun eläytyminen tarjoaa kuitenkin vastausmahdollisuuden myös sellaisille vastaajille, joille tekoäly on tuntemattomampi teema. Variaation tarkoituksena taas on ilmentää mahdollisia eroja sen suhteen, minkälaisilla keinoilla voidaan kenties varmemmin tunnistaa tekoälyn tuottamaa sisältöä, tai minkälaiset julkaisut tai niiden piirteet ovat varmemmin tunnistettavissa tekoälyn tuotoksiksi.

4.2 Aineiston kerääminen

Alkuperäinen suunnitelmani oli kerätä aineistoa kirjaston tai kirjastojen digiopastustapahtumien yhteydessä, sillä digiopastuksiin kokoontuu joukko eri ikäisiä aikuisia, erilaisista taustoista. Toiveena oli tavoittaa eri ikäisiä aikuisia erilaisista taustoista, rajaamatta vastaajia muun kuin aikuisuuden ja suomenkielisyyden suhteen. Aineistonkeruun ensimmäinen vaihe tapahtuikin alkuperäisen suunnitelman mukaisesti erään kirjaston kahdessa digiopastustilaisuudessa. Kerroin vastaajille lyhyesti tutkielman tavoitteesta ja menetelmästä ja jaoin vastaajille A4-paperin, jossa oli ohjeistus sekä linkki ja QR-koodi, joiden avulla pääsi verkkolomakkeeseen omalla laitteellaan. Paperin kääntöpuolella oli kehyskertomus ja tilaa kirjoittaa kertomus. Osassa papereista oli kehyskertomus A ja osassa B, ja

vastaavasti linkit ja QR-koodit johtivat verkkolomakkeeseen, jossa oli joko kehyskertomus A tai B. Vastaajat saivat siis päättää kirjoittavatko vastauksensa paperille vai omalla laitteellaan digitaalisesti. Lomakkeessa olevien ohjeiden lisäksi annoin vastaajille lyhyen suullisen ohjeistuksen kertomuksen kirjoittamiseen, jossa mainitsin kertomuksen vapaamuotoisuuden, tutkimukseen osallistumisen vapaaehtoisuuden ja nimettömyyden. Kehotin myös vastaajia olemaan kirjoittamatta kertomukseen nimeään tai muita henkilötietoja. Itse kirjoitustehtävän lisäksi lomakkeessa kysyttiin taustatietoina ainoastaan vastaajan suurpiirteinen ikä noin kymmenen vuoden intervallein.

Opastustilaisuuksien yhteydessä keräämäni aineiston määrä jäi kuitenkin vähäiseksi, vähäisten osallistujamäärien, sekä vähäisen kiinnostuksen vuoksi. Digiopastustilaisuuksien yhteydessä vastauksia kertyi seitsemän kappaletta. Heikon menestyksen ja sopivien digiopastustilaisuuksien puutteen vuoksi oli tarpeen muuttaa aineistonkeruusuunnitelmaa. Seuraava suunnitelma oli kerätä aineistoa eräästä digiopastukseen liittyvästä avoimesta Facebook-ryhmästä, mutta tämä jäi suunnitelman tasolle alustaan liittyvien teknisten vaikeuksien vuoksi. Viimeinen suunnitelma oli kerätä aineisto korkeakoulu yhteisön yhteisen viestintäkanavan kautta. Tämä kolmas aineistonkeruuvaihtoehto ei ollut täysin ideaali, sillä siinä korkeakoulu yhteisö rajaa vastaajien taustaa vastoin alkuperäisen suunnitelman ajatusta taustan rajaamattomuudesta. Toteutin kuitenkin suunnitelman ja jaoin linkit molempiin lomakkeisiin viestin saattamana viestintäkanavalle. Tämä toinen onnistunut aineistonkeruuvaihe erosi ensimmäisestä siten, että se tapahtui kokonaan sähköisesti, eli vastaajat eivät saaneet suullista ohjeistusta. Sama ohjesisältö löytyi kuitenkin tutkimuslomakkeen johdannosta. Tässä vaiheessa vastauksia kertyi vielä 28 kappaletta lisää.

Kokonaisuudessaan kertomuksia kertyi siis 35 kappaletta, joista 22 vastasi ensimmäiseen kehyskertomukseen ja 13 toiseen kehyskertomukseen. Eläytymismenetelmässä sopiva määrä aineistoa on 10–20 kertomusta kehyskertomusta kohden (Eskola, 1997), eli tässä tilanteessa yhteensä noin 20–40 vastausta, joten vastauksia kertyi oikein sopiva määrä. Vastaajien ikä vaihteli ikähaarukasta 18–29 vuotta, ikähaarukkaan 80–89 vuotta, mutta 60–69-vuotiaita vastaajia ei ollut yhtäkään. Suurin osa vastaajista oli alle 40-vuotiaita. Kertomusten pituus vaihteli muutamasta sanasta muutama sataan sanaan ja kaikki kertomukset olivat käyttökelpoisia.

4.3 *Analyysi*

Eskolan ja Eskolan (1998) mukaan eläytymismenetelmäaineistoa voidaan analysoida eri tavoin, mutta olennaisinta analyysissä on huomioida kehyskertomuksen variaatiot. Kehyskertomusten variointi ja erot eri kehyskertomuksiin vastaavissa kertomuksissa ovat se, mikä erottaa eläytymismenetelmän monesta muusta menetelmästä (Eskola, 1997). Eläytymismenetelmän analyysin tarkoituksena on siis tarkastella kertomusten sisältöä sekä paljastaa, vertailla ja tulkita muutoksia ja samankaltaisuuksia, joita kehyskertomusten eri variaatioihin pohjautuvissa kertomuksissa ilmenee.

Tässä tutkielmassa perehdyn aineistoon aineistolähtöisen sisällönanalyysin avulla. Aineistolähtöinen sisällönanalyysi voidaan erottaa kolmeen vaiheeseen: aineiston pelkistäminen, aineiston ryhmittely ja käsitteellistäminen (Miles & Huberman, 1994; Tuomi & Sarajärvi, 2018, mukaan). Ensimmäisessä vaiheessa aineistosta poimitaan analyysin kannalta olennaiset osat (Tuomi & Sarajärvi, 2018). Toisessa vaiheessa näistä olennaisista osista etsitään samankaltaisuuksia ja eroavaisuuksia, ja muodostetaan niistä luokkia. Muodostuneista alaluokista yhdistellään yläluokkia ja edelleen pääluokkia. Kolmannessa vaiheessa muodostetaan teoreettisia käsitteitä tutkimukselle olennaiseksi valikoituneen tiedon perusteella ja pyritään siten vastaamaan tutkimustehtävään (Tuomi & Sarajärvi, 2018). Aineistolähtöinen sisällönanalyysi sopii hyvin tämän tutkielman aineistolle, sillä tarkoituksena on tarkastella, minkälaisia tekoälyn tunnistamisen keinoja aineistossa ilmenee, ja aineistolähtöisen analyysin avulla on mahdollista etsiä ja jäsentää niitä.

Toteutin aineistolähtöisen sisällönanalyysin kahteen kertaan: kehyskertomuksen A kertomuksille ja kehyskertomuksen B kertomuksille erikseen. Eri kehyskertomuksiin vastaavat kertomukset oli analysoitava erikseen, jotta varioinnin vaikutusta oli mahdollista tarkastella. Aloitin analyysin lukemalla kertomukset huolellisesti läpi, jonka jälkeen alleviivasin kertomuksista tutkielman kannalta olennaiset ilmaukset. Olennaiset ilmaukset tekstissä olivat yhden tai useamman virkkeen mittaisia pätkiä. Kirjasin ylös olennaisista ilmauksista pelkistämäni ilmaukset, joista lähdin muodostamaan luokkia. Taulukossa 1 on esimerkki kehyskertomukseen A perustuvien kertomusten analyysin (jatkossa analyysi A) yhdestä alaluokasta ja pelkistetyistä ilmauksista, joista se muodostui.

TAULUKKO 1. Esimerkki pelkistetyistä ilmauksista ja niistä muodostuneesta alaluokasta analyysissä A

Pelkistetyt ilmaukset	Alaluokka
Ohjeiden etsiminen netistä	Ohjeiden ja apuvälineiden hyödyntäminen
Tekniset apuvälineet tekoälyn tunnistamiseen	
Netistä löytyvät apuvälineet tekoälyn tunnistamiseen	
Tekoälyn tuottaman tekstin tunnistamiseen tarkoitetut sivustot	

Muodostettuani pelkistetyistä ilmauksista alaluokat, siirryin yhdistelemään alaluokkia yläluokiksi, joita yhdistin edelleen pääluokiksi. Taulukossa 2 on pääluokien muodostumista havainnollistava esimerkki analyysistä A. Siinä pelkistetyistä ilmauksista muodostetut alaluokat muodostavat yläluokkia, jotka taas muodostavat pääluokan.

TAULUKKO 2. Esimerkki pääluokan muodostumisesta analyysissä A

Alaluokat	Yläluokat	Pääluokka
Tekoälyn toimintaperiaatteiden tunteminen	Tekoälyn tunteminen	Tekoälyn tunnistamiseen harjaantuminen
Ohjeiden ja apuvälineiden hyödyntäminen		
Tekoälyn tunnistamiseen harjaantuminen	Tekoälyn tunnistamiseen harjaantuminen	
Tekoälyyn liittyvä hämmennys		
Tekoälyn tunnistamisen vaikeus		
Outous julkaisussa	Luonnottomuuden tunnistaminen	
Luonnottomuus tai epäinhimillisuus		

Analysoin siis keräämäni aineistoa huomioimatta tässä vaiheessa teoreettista viitekehystä. Kuvailen tuloksia ja vertaan kehyskertomusten varioinnin vaikutuksia keskenään luvussa 5. Tarkastelen sen jälkeen aineistosta nousseita tuloksia teoreettisen viitekehysten valossa luvussa 6.

5 TULOKSET

Tämän tutkielman tarkoituksena on siis vastata tutkimuskysymykseen: *millaisia tekoälyn tunnistamisen keinoja aikuisten kertomuksissa ilmenee?* Kuvailen tässä luvussa tutkielman aineistosta analyysillä saatuja tuloksia. Kehyskertomusten erilliset analyysit muotoutuivat molemmat samoiksi pääluokiksi. Pääluokkia muodostui viisi: valppaus, medialukutaito, tekoälyn tunnistamiseen harjaantuminen, tarkempi tilannekohtainen analyysi ja yhdessä oppiminen. Pääluokkia yhdistää yksi yhdistävä luokka: tekoälyn tunnistamisen keinot. Kuvailen luvussa 5.1 tarkemmin jokaista pääluokkaa ja luokkakohtaisia eroja kehyskertomusten välillä.

5.1 *Millaisia tekoälyn tunnistamisen keinoja aikuisten kertomuksissa ilmenee?*

Aineisto jakautui siis kahteen osaan, kertomuksiin, jotka perustuivat kehyskertomukseen A ja kertomuksiin, jotka perustuivat kehyskertomukseen B:

(A) Tuttavasi kertoo sinulle, että törmäsi taannoin netissä erääseen julkaisuun, ja epäili, että tekoäly on tehnyt sen. Hän kertoo, millainen julkaisu oli ja mitä hän oli ajatellut pohtiessaan, oliko se todella tekoälyn luoma. Käytä mielikuvitustasi ja kerro keskustelustanne.

(B) Tuttavasi kertoo sinulle, että törmäsi taannoin netissä erääseen julkaisuun, ja oli varma, että tekoäly on tehnyt sen. Hän kertoo, millainen julkaisu oli ja mitä hän oli ajatellut pohtiessaan, oliko se todella tekoälyn luoma. Käytä mielikuvitustasi ja kerro keskustelustanne.

Viitataan jatkossa kehyskertomukseen A perustuviin kertomuksiin kertomuksina A ja kehyskertomukseen B perustuviin kertomuksiin kertomuksina B luettavuuden helpottamiseksi. Sitaattien yhteydessä merkitsen kertomuksen koodilla, jossa kirjain viittaa analyysiin A tai B ja numero kertomukselle osoitettuun numeroon (esimerkiksi A1).

Kehyskertomusvariaatioiden välinen ero on siis siinä, kuinka varma kehyskertomuksessa esiintyvä tuttava on siitä, että julkaisu on tekoälyn tekemä. Kehyskertomusten variointi ei kuitenkaan tässä tutkielmassa tuottanut merkittävää eroa eri kehyskertomuksiin perustuvissa kertomuksissa. Joitain eroavaisuuksia on kuitenkin nähtävissä. Kertomuksissa A kuvailtiin julkaisuja, jotka olivat kuvia, tekstiä ja videoita tai näiden yhdistelmiä, kun taas kertomukset B kuvailivat pääosin tekstijulkaisuja, mutta myös muutamia kuvajulkaisuja. Kertomuksissa A oli myös selvästi enemmän mainintoja julkaisijan, julkaisijan profiilin ja sivuston arvioinnista, kuin kertomuksissa B. Kertomuksissa A mainittiin myös useammin netistä löytyvät sivustot, joilla voidaan pyrkiä selvittämään, onko jokin tekoälyn tuotosta. Nämä eroavaisuudet voivat johtua siitä, että kertomuksia A oli enemmän kuin kertomuksia B. Toisaalta voitaisiin tulkita, että kenties tekstijulkaisut koetaan selvemmin tunnistettaviksi ja että tilanteessa, jossa ei voi olla aivan varma, tulee vasta käännettyä julkaisukontekstin puoleen. Lisäksi tilanteissa, joissa voi olla heti varma tekoälyn osallisuudesta, ei todennäköisesti tarvitse ajatella ulkoisiin työkaluihin turvautumista.

Analyysit A ja B johtivat keskenään hyvin samankaltaisiin tuloksiin: samaan yhdistävään luokkaan ja samoihin viiteen pääluokkaan. Yläluokat, joista pääluokat muodostuvat eroavat kuitenkin eri kehyskertomusten analyysseissä toisistaan jonkin verran. Analyysissä A muodostuneet yhdistävä luokka, pääluokat ja yläluokat ovat nähtävissä taulukossa 3.

TAULUKKO 3. Analyysissä A muodostuneet luokat

Tekoälyn tunnistamisen keinot				
Valppaus	Medialukutaito	Tekoälyn tunnistamiseen harjaantuminen	Tarkempi tilannekohtainen analyysi	Yhdessä oppiminen
Tarkkaavaisuus Varauksella suhtautuminen	Medialukutaito Kriittinen ajattelu Lähdekriittisyys	Tekoälyn tunnistaminen Tekoälyn tunnistamiseen harjaantuminen Luonnottomuuden tunnistaminen	Tekstin tilannekohtainen analyysi Kontekstin arviointi Yksityiskohtien arviointi	Yhdessä oppiminen

Yhdistävä luokka, *tekoälyn tunnistamisen keinot*, koostuu siis viidestä pääluokasta: valppaus, medialukutaito, tekoälyn tunnistamiseen harjaantuminen, tarkempi tilannekohtainen analyysi ja yhdessä oppiminen. Analyysin B vastaavat luokat ovat nähtävissä taulukossa 4. Käsittelen pääluokkien sisältöä ja kehyskeromusten välisiä eroja tarkemmin luokkakohtaisissa alaluvuissa (5.1.1, 5.1.2, 5.1.3, 5.1.4 ja 5.1.5.).

TAULUKKO 4. Analyysissä B muodostuneet luokat

Tekoälyn tunnistamisen keinot				
Valppaus	Medialukutaito	Tekoälyn tunnistamiseen harjaantuminen	Tarkempi tilannekohtainen analyysi	Yhdessä oppiminen
Tarkkaavaisuus Varovaisuus	Medialukutaito ja kriittinen ajattelu Lähdekriittisyys	Tekoälyn tunnistamiseen harjaantuminen Luonnottomuuden tunnistaminen	Tekstin sisällön analysoiminen Tekstin kielellisten seikkojen analysoiminen	Yhdessä oppiminen

Taulukoissa 3 ja 4 näkyvien pääluokkien lisäksi molemmissa analyysissä nousi selkeästi esiin teema, *huoli tekoälyn kehityksen suunnasta*, jota käsittelen erillisenä luvussa 5.2, sillä se on mielestäni tutkielman kannalta olennainen, mutta ei varsinaisesti istu muiden pääluokkien kanssa yhdistävän luokan alle. Käytän jatkossa kursiivua luokkien nimissä luettavuuden helpottamiseksi.

5.1.1 Valppaus

Ensimmäinen pääluokka, *valppaus*, koostuu kehyskertomuksen A tapauksessa *tarkkaavaisuudesta* ja *varauksella suhtautumisesta*. Kehyskertomuksen B tapauksessa taas *tarkkaavaisuudesta* ja *varovaisuudesta*. Tämän pääluokan osalta kehyskertomusten variointi ei tuottanut selkeitä eroavaisuuksia kertomuksissa. Luokat koostuvat ilmauksista, joissa mainitaan muun muassa tarkkaavaisuus netissä liikkua ja varauksella suhtautuminen sellaisiin julkaisuihin, jotka saattavat olla tekoälyllä generoituja. Esimerkiksi eräässä kertomuksessa valppaus ilmenee seuraavasti:

Täytyy vaan ite olla hereillä sen suhteen, että mitä lukee ja minkä ottaa totena ja minkä ei. [A19]

Otteessa pohditaan, että on yksilön omalla vastuulla olla tarkkaavainen ja arvioida sellaisen sisällön luotettavuutta, joka on tai saattaa olla tekoälyn generoimaa. Kokonaisuudessaan *valppaus*-pääluokka vaikuttaa viittaavan siihen, että median käyttäjä on itse vastuussa siitä, kuinka ja minkälaisella valppaudella ja varauksella suhtautuu kohtaamaansa sisältöön.

5.1.2 Medialukutaito

Medialukutaito-pääluokka muodostui analyysissä A yläluokista *medialukutaito*, *kriittinen ajattelu* ja *lähdekriittisyys*, ja analyysissä B yläluokista *medialukutaito* ja *kriittinen ajattelu* sekä *lähdekriittisyys*. Tämänkään pääluokan tapauksessa kehyskertomusten variointi ei tuottanut merkittävää eroa. *Medialukutaitoon* liittyvät maininnat kertomuksissa kuvasivat medialukutaidon merkitystä tekoälyllä luodun sisällön tunnistamisessa:

Hän kertoo kuitenkin olleensa huolissaan siitä, että uutistoimisto oli julkaissut kuvan sosiaalisen median kanavalleen. Tuttavani piti tekoa vastuuttomana

ja oli sitä mieltä, että tapaus osoittaa huonoa medianlukutaitoa toimittajilta. [A18]

Sitaatin kertomuksessa esiintyvä tuttava on huolissaan siitä, että toimittajat olivat julkaisseet tekoälyllä luodun kuvan sosiaalisessa mediassa, ja pitää tätä merkkinä heikosta medialukutaidosta. *Kriittiseen ajatteluun* liittyvät maininnat taas liittyivät usein julkaisujen kontekstin kriittiseen arviointiin, kuten esimerkiksi seuraavassa otteessa:

Niin kyllähän käsityöhön jää yleensä käsityön jälki vaikka ois kuinka taitava. Tekoäly varmaan vetää mutkat suoriksi eikä tiedä miten oikeasti toteutetaan asioita. [A17]

Otteessa pohditaan kriittisesti kertomuksessa kuvailtua kuvaa epäilyttävän täydellisestä käsityöstä. *Lähdekriittisyyteen* liittyvät ilmaukset kertomuksissa taas liittyivät lähteiden puutteen huomaamiseen, epämääräisten lähteiden tunnistamiseen ja ulkopuolisten lähteiden hakemiseen tekoälyn luoman julkaisun tunnistamiseksi. Mielestäni nämä kaikki kolme luokkaa sopivat *medialukutaito*-pääluokkaan, sillä ne ovat rinnastettavissa tyypillisesti medialukutaidon määritelmään sisältyviin kyvykkyyksiin, kuten median tulkintaan ja arviointiin, kriittiseen ajatteluun ja lähdekriittisyyteen.

5.1.3 Tekoälyn tunnistamiseen harjaantuminen

Tekoälyn tunnistamiseen harjaantuminen -pääluokan sisällöissä on nähtävissä pieni ero analyyseissä A ja B. Molemmissa analyyseissä pääluokka muodostui *tekoälyn tunnistamiseen harjaantuminen*- ja *luonnottomuuden tunnistaminen* -yläluokista, mutta sen lisäksi analyysissä A muodostui selkeämmin *tekoälyn tunteminen* -yläluokka. *Tekoälyn tunnistamiseen harjaantuminen* näyttäytyi aineistossa esimerkiksi mainintoina generatiivisen tekoälyn luoman sisällön tyypillisistä ominaisuuksista:

Tuttu kertoo esimerkiksi toistuvista lauserakenteista joita on tottunut näkemään esimerkiksi ChatGPT:tä käyttäessä. [B9]

Otteessa mainitaan lauserakenteet, joita ChatGPT saattaa toistaa usein. Näistä muodostuu tunnistettavia piirteitä, kun generatiivisen tekoälyn kanssa on tekemisissä tarpeeksi. *Luonnottomuuden tunnistaminen* taas ilmeni aineistossa muun

muassa mainintoina luonnottoman täydellisyys ja liikkeen tunnistamisena tai esimerkiksi tekstin robottimaisuuden huomaamisena. Esimerkiksi seuraavassa otteessa luonnon täydellisyys kuvassa herättää epäilyksiä:

Kasvot oli sit mun mielestä taas oudon tasaiset ja symmetriset, niin se herätti epäilykset. [A19]

Tekoälyn tunteminen muodostui selkeäksi yläluokaksi analyysissä A. *Tekoälyn tunteminen* -yläluokka koostuu alaluokista *tekoälyn toimintaperiaatteiden tunteminen* ja *ohjeiden ja apuvälineiden hyödyntäminen*. *Tekoälyn toimintaperiaatteiden tunteminen* esiintyi aineistossa mainintoina tekoälyn toimintamekanismien pohtimisesta tekoälyn tunnistamisen apuna:

Muista, että kielimalli on kielimalli, joka arvaa miten sanat muodostuvat ja liittyvät toisiinsa. [A5]

Ohjeiden ja apuvälineiden hyödyntäminen taas ilmeni mainintoina netistä löytyvistä ohjeista ja sivustoista, joilla on mahdollista pyrkiä tarkistamaan, onko jokin sisältö tekoälyn tuottamaa. Kokonaisuudessaan siis pääluokka, *tekoälyn tunnistamiseen harjaantuminen*, vaikuttaa viittaavan siihen, että olemalla tekemisissä tekoälyn kanssa oppii tuntemaan sen toimintaperiaatteita, tyypillisiä piirteitä ja hyödyntämään siihen liittyviä ohjeita ja apuvälineitä, ja siten kenties myös tunnistamaan tekoälyn tuottamaa sisältöä tehokkaammin.

5.1.4 Tarkempi tilannekohtainen analyysi

Tarkempi tilannekohtainen analyysi -pääluokka on osin päällekkäinen *tekoälyn tunnistamiseen harjaantumisen* kanssa, sillä esimerkiksi yläluokka, *luonnottomuuden tunnistaminen*, sopisi kumpaan tahansa pääluokkaan. Tämä pääluokka eroaa sisällöltään hieman analyysien A ja B välillä. Analyysissä A *tarkempi tilannekohtainen analyysi* muodostui yläluokista *tekstin tilannekohtainen analyysi*, *kontekstin arviointi* ja *yksityiskohtien arviointi*. Näissä kertomuksissa pohdittiin teksti-, kuva- ja videomuotoisten julkaisujen analysointia. *Tekstin tilannekohtainen analyysi* -yläluokka koostui tässä analyysissä *tekstin sisällön tarkastelusta* ja *tekstin kieliasun tarkastelusta*. *Kontekstin arviointi* -yläluokka taas on osin päällekkäinen luvussa 5.1.2 mainitsemani *kriittinen arviointi* -yläluokan kanssa, johon liittyi myös kontekstin kriittistä arviointia. *Kontekstin arviointi* ilmeni aineistossa

muun muassa julkaisijan, julkaisijan profiilin tai sivuston tarkasteluna suhteessa julkaisuun sekä myös motiivien pohtimisena. *Yksityiskohtien arviointi* taas ilmeni kertomuksissa yksityiskohtien tarkasteluun liittyvinä mainintoina:

Hän kertoo, että tarkasteltuaan kuvaa tarkemmin, siinä näkyi outoja yksityiskohtia, kuten epämuodostuneita yleisön jäseniä ja kannustuskyllttejä, joissa ollut teksti ei tarkoittanut mitään. Päällisin puolin ja nopeasti katsottuna kuva näytti kuitenkin ihan oikealta uutiskuvalta. [A18]

Analyysissä B *tarkempi tilannekohtainen analyysi* taas koostui *tekstin sisällön analysoiminen-* ja *tekstin kielellisten seikkojen analysoiminen -yläluokista*. Tämän analyysin pääluokka vastasi siis sisällöltään pitkälti ensimmäisessä analyysissä muodostuneen pääluokkaan sisältyvää yläluokkaa *tekstin tilannekohtainen analyysi*. Kertomukset B painottuivat enemmän tekstimuotoisiin julkaisuihin ja siten myös niiden analysoimiseen. *Tekstin sisällön analysoiminen* ilmeni aineistossa muun muassa mainintoina tekstin geneerisyydestä, odottamattomista näkökulmista ja kirjoittajan oman äänen puuttumisesta:

julkaisun teksti oli kirjoitettu hyvin monotonisella tavalla, eikä siitä pystynyt erottamaan kirjoittajan omaa "ääntä". [B8]

Tekstin kielellisten seikkojen analysoiminen esiintyi aineistossa esimerkiksi mainintoina virheistä, sujuvuuden puutteesta ja oudon tuntuista ilmauksista tekstissä.

Analyysissä A muodostuneen pääluokan monipuolisuus verrattuna toisen kehyskertomuksen kertomusten analyysiin, saattoi johtua suuremmasta määrästä kertomuksia. Kehyskertomuksessa B kuvaillaan tilannetta, jossa kertomuksen "tuttava" on varma, että julkaisu on tekoälyn tuotos. Kenties voitaisiin myös pohtia, että koska tämän kehyskertomuksen kertomukset keskittyvät pääosin tekstimuotoisiin julkaisuihin, tekstin tunnistaminen varmasti tekoälyn tuotokseksi olisi helpompaa, kuin kuvien ja videoiden tunnistaminen.

5.1.5 Yhdessä oppiminen

Yhdessä oppiminen on kaikista pääluokista yksinkertaisin ja koostuu ainoastaan kertomuksissa ilmenneistä maininnoista, jotka viittaavat yhdessä miettimiseen, muilta ihmisiltä oppimiseen ja kysymiseen. *Yhdessä oppimisen* pääluokka oli hy-

vin samankaltainen molemmissa analyyseissä. Yhdessä oppiminen ilmeni kertomuksissa usein mainintoina siitä, että pyritään yhdessä pohtimaan, onko julkaisu tekoälyn tuotos tai opitaan muilta jotain siihen liittyvää:

Oltiin aikasemmin puhuttu vastaavasta kuvasta, josta olin osottanu tekoälyyn viittaavat tunnusmerkit, ja nyt hän sano epäilevänsä tätäkin kuvaa sellaiseksi. [B12]

Otteessa kuvaillaan tilannetta, jossa kirjoittaja on auttanut tuttavaansa tunnistamaan tekoälyyn viittaavia tunnusmerkkejä kuvasta, ja tunnusmerkit oppineena tuttava osaa epäillä toista julkaisua tekoälyn luomaksi. Pääluokka, *yhdessä oppiminen*, vaikuttaa siis kuvaavan vertaisten tuen hyötyä tekoälyn tunnistamisessa.

5.2 Huoli tekoälyn kehityksen suunnasta

Molempien kehyskertomusten tapauksissa analyysissä nousi vielä yksi teema: *huoli tekoälyn kehityksen suunnasta*. Käsittelen teemaa erillisenä muista, sillä se ei varsinaisesti sovi yhdistävän luokan, *tekoälyn tunnistamisen keinot*, alle mutta on mielestäni kuitenkin tutkielman kannalta olennainen. *Huoli tekoälyn kehityksen suunnasta* ilmeni aineistossa huolena tulevaisuudesta ja pohdintana siitä, voiko mihinkään enää luottaa. Kertomuksissa pohdittiin tekoälyllä tuotetun sisälön taustalla olevia toimijoita ja tahoja, jotka sisältöä luovat sekä näiden motiiveja. Lisäksi kertomuksissa pohdittiin, mitä tapahtuu, jos generatiivista tekoälyä ruokitaan ”pahuudella” ja käytetään pahaan. Kertomuksissa pohdittiin myös tekoälyn kehityksen suuntaa, ja että onko tulevaisuudessa enää mahdollista erottaa totuutta valheesta, ja voiko mihinkään enää luottaa. Seuraava sitaatti kuvastaa kertomusten pohdintoja hyvin:

Tuttavani kertoo pelkäävänsä, että tekoälyllä luodut kuvat lisääntyvät tulevaisuudessa ja voivat levitessään aiheuttaa vääristyneitä ajatuksia ja pahimmillaan johtaa johonkin epämieluiseseen tapahtumaan. Pahimmissa tapauksissa emme enää erota, mikä on totta ja mikä ei. [A18]

Nämä maininnat ovat tärkeitä, sillä ne valaisevat ihmisten asenteita ja tekoälyyn liittyviä epävarmuuksia ja huolia. Asenteet ja huolet ovat varmasti yhteydessä myös siihen, kuinka innokkaita ihmiset ovat oppimaan tekoälystä ja käyttämään erilaisia generatiivisen tekoälyn sovelluksia, joka taas tämän tutkielman tulosten perusteella vaikuttaisi olevan yhteydessä tekoälyn tunnistamiseen (luku 5.1.3).

6 POHDINTA

Tarkastelen tässä luvussa tarkemmin tutkielman tuloksia teoreettisen viitekehyksen valossa. Pohdin lisäksi tutkielmaan liittyviä eettisiä kysymyksiä ja tutkielman luotettavuutta ja rajoituksia. Lopuksi kuvailen tutkielman johtopäätöksiä ja pohdin mahdollisia jatkotutkimusaiheita luvussa 6.3.

6.1 Tutkielman tulokset teoreettisen viitekehyksen valossa

Tutkielman tarkoituksena on vastata tutkimuskysymykseen: *millaisia tekoälyn tunnistamisen keinoja aikuisten kertomuksissa ilmenee?* Pyrin tarkastelemaan tutkimuskysymyksen avulla tekoälylukutaidon osa-aluetta, tekoälyn tunnistamista, ja tarkastelemaan samalla kertomuksista mahdollisesti ilmeneviä asenteita tekoälyä kohtaan. Tutkielman tuloksena muodostui viisi luokkaa, joita yhdistää yhdistävä luokka: *tekoälyn tunnistamisen keinot*. Pääluokiksi muodostuivat *valppaus, medialukutaito, tekoälyn tunnistamiseen harjaantuminen, tarkempi tilannekohtainen analyysi ja yhdessä oppiminen*. Pääluokat antavat vastauksen tutkimuskysymykseen havainnollistaen millaisia tekoälyn tunnistamisen keinoja aikuisten kirjoittamissa kertomuksissa ilmenee. Pääluokkien lisäksi kertomuksista nousi tärkeä teema, joka ei kuitenkaan sovi muodostuneiden pääluokkien joukkoon: *huoli tekoälyn kehityksen suunnasta*.

Pääluokat *valppaus* ja *medialukutaito* ovat molemmat kykyjä, jotka voivat auttaa tunnistamaan tekoälyn tuottamaa sisältöä. Ne voivat myös auttaa muun muassa tunnistamaan ja välttämään uhkia. Lee (2018) kuvailee disinformaatiota, vale uutisia ja huijauksia, joihin yksilöt voivat mediassa törmätä, sekä medialukutaidon roolia näiden uhkien torjunnassa. Generatiivisen tekoälyn kehitys on epäilemättä vain lisännyt tämänkaltaisia uhkia mediassa: esimerkiksi Limin ja kollegojen (2023) mukaan generatiivinen tekoäly mahdollistaa entistä nopeamman ja helpomman disinformaation tuottamisen. Lisäksi generatiivinen tekoäly on tuonut mukanaan muita ongelmia, kuten syvävääreännöksiä ja vääristymiä algoritmeissa

(Wach ym., 2023). Kuten medialukutaito, myös kertomuksista noussut *valppaus* edistää varmasti tällaisten uhkien tunnistamista ja välttämistä. Pääluokat, *valppaus* ja *medialukutaito*, ovat myös osin yhteneväisiä Frau-Meigsin (2023) tekoälykyvykkyyksien viitekehyksen asenteet-kategorian kanssa, johon liittyy muun muassa datan ja sen käyttötarkoitusten kriittinen tulkitseminen, haitallisen sisällön tunnistaminen ja riskien välttäminen.

Tekoälyn tunnistamiseen liittyvät *valppaus*- ja *medialukutaito*-pääluokat ovat myös olennaisia Potterin (2004) kuvaileman median automaattisen prosessoinnin kannalta. Potterin (2004) kuvaama median ja ihmisen suhde, jossa ihminen vastaa suureen määrään informaatiota prosessoimalla osan siitä automaattisesti on hyvin ajankohtainen informaatiotulvan kasvaessa jatkuvasti – myös generatiivisen tekoälyn luoman sisällön seurauksena. Potterin (2004) mukaan jatkuva automaattinen prosessointi mediassa heikentää yksilön medialukutaitoa ja antaa medialle vaikutusvaltaa. *Valppaus* ja *medialukutaito* auttavat kenties taistelemaan automaattista prosessointia vastaan ja sen sijaan edistävät tietoista prosessointia, joka vaikuttaisi Potterin (2004) mukaan edistävän median käyttäjän toimijuutta mediassa ja siten ehkä myös tekoälyn tunnistamista. Medialukutaito voitaisiin ainakin tämän tutkielman tulosten perusteella nähdä osana tekoälyn tunnistamista, mutta toisaalta tekoälyn tunnistaminen ja tekoälylukutaito voitaisiin käsittää myös osaksi laajempaa medialukutaitoa, kuten esimerkiksi Frau-Meigs (2023) argumentoi.

Tämän tutkielman tulokset ovat osin linjassa aiempien tekoälylukutaidon määritelmien kanssa. Voisi ajatella, että *tekoälyn tunnistamiseen harjaantuminen* ja *tarkempi tilannekohtainen analyysi* vaativat molemmat tekoälyn kanssa toimimista ja ainakin pintapuolista tekoälyn toimintaperiaatteiden tuntemista. Ne ovat siis ainakin osin yhteneväisiä Longin ja Magerkon (2020) tekoälylukutaitoviitekehyksen osa-alueiden *Mikä on tekoäly?*, *Kuinka tekoäly toimii?* ja *Mitä tekoäly voi tehdä?* kanssa. Longin ja Magerkon (2020) mukaan yksi *Mikä on tekoäly?* -osa-alueeseen liittyvä kyvykkyys on tekoälyn tunnistaminen, jonka alle tämän tutkielman tulosten voisi ajatella sopivan. Tässä tutkielmassa ilmenneet tekoälyn tunnistamisen keinot vaikuttavat siis olevan linjassa Longin ja Magerkon (2020) viitekehyksen kanssa laajemmin, kuin ainoastaan tämän yhden osa-alueen yhden kyvykkyuden kanssa. Ng ja kollegat (2021) taas luonnehtivat tekoälylukutaitoa

hierarkkisena rakenteena, jonka pohjalla ovat tekoälyyn liittyvä tietämys ja ymmärrys, keskellä tekoälyn käyttäminen ja hyödyntäminen ja huipulla tekoälyn arviointi ja luominen. Tässäkin tilanteessa *tekoälyn tunnistamiseen harjaantuminen ja tarkempi tilannekohtainen analyysi* voitaisiin rinnastaa tekoälyyn liittyvään tietämykseen ja ymmärrykseen, sekä osin tekoälyn käyttämiseen ja arviointiin. Tässä tutkielmassa ilmenneet tekoälyn tunnistamisen keinot vaikuttavat siis olevan osin yhteneväisiä tämänkin tekoälylukutaidon määritelmän kanssa, vaikka siitä puuttuu kokonaan tekoälyn tunnistamisen näkökulma.

Kertomuksista nousi esiin huolia tekoälyn kehityksen suunnasta ja siitä, miten tekoälyn kehitys vaikuttaa mediaan tulevaisuudessa. Kertomuksissa pohdittiin, voiko mihinkään luottaa ja onko tulevaisuudessa mahdollista enää erottaa, mikä on totta. *Huoli tekoälyn kehityksen suunnasta* koostui muun muassa tekoälyn taustalla oleviin toimijoihin ja näiden motiiveihin liittyvistä maininnoista. Nämä maininnat ovat hyvin samankaltaisia Frau-Meigsin (2023) tekoälykyvykkyyksien viitekehysten tieto-kategorian kanssa, johon liittyy muun muassa tekoälyn taustalla vaikuttavien toimijoiden, omistajuuksien ja motivaatioiden tunnistaminen.

Huolestuttaviin tulevaisuuskuviin voitaisiin kenties vastata huolehtimalla tekoälyn eettisestä käytöstä sekä yksilöiden osalta, että laajempiin eettisten ohjeistuksiin nojaten (ks. esim. Lähdesmäki, 2024; Frau-Meigs, 2023; Jones, 2024; Łodzikowski ym., 2024). Lähdesmäen (2024) mukaan eettinen älykkyys on olennainen osa tekoälytoimijuutta. Frau-Meigs (2023) taas korostaa media- ja informaatiolukutaidon roolia generatiivisen tekoälyn eettisen käytön pohjana. Eettisyys on osa myös Longin ja Magerkon (2020) tekoälylukutaidon viitekehystä yhtenä kyvykkyytenä *Miten tekoälyä tulisi käyttää?* -osa-alueessa. Myös asenteilla on merkitystä eettisyyden kannalta: Jonesin (2024) mukaan liian negatiivisen suhtautumisen seurauksena voi jäädä paitsi tekoälyn hyödyistä, kun taas liian positiivinen suhtautuminen voi sokaista tekoälyyn liittyville riskeille ja ongelmille. Epäluottamus tai huolestunut suhtautuminen tekoälyyn voi kenties lisätä sen välttelemistä, ja siten estää esimerkiksi tekoälyn tunnistamiseen harjaantumista ja tekoälyn toimintaperiaatteiden tuntemista, ja siten heikentää tekoälyn tunnistamista tai tekoälylukutaitoa. Long ja Magerko (2020) huomauttavat, että tekoäly muuttuu nopeasti ja siksi on tärkeää pohtia mitä tekoäly voisi tehdä tulevaisuu-

dessa. Tulevaisuuden kuvittelu on osa heidän tekoälylukutaitoviitekehystään. Tämä voisi tukea ajatusta, että kertomuksista noussut teema *huoli tekoälyn kehityksen suunnasta* olisi kuitenkin voinut sopia tässä tutkielmassa ilmenneiden viiden pääluokan joukkoon.

Tämän tutkielman tulokset ovat siis monella tavalla yhteneväisiä eri teemoihin liittyvän teorian ja aiemman tutkimuksen kanssa. On kuitenkin tärkeää pitää mielessä eläytymismenetelmällä kerätyn aineiston luonne: se ei ole yleistettävää tai objektiivista tietoa, vaan perustuu vastaajien käsityksiin, ajatuksiin ja mielikuvitukseen. Tulokset tarjoavat kuitenkin katsauksen aikuisten ajatuksiin tekoälyn tunnistamisesta ja osin myös tekoälyyn liittyviin asenteisiin. Ne valaisevat tekoälyn tunnistamisen keinoja aikuisten silmin, mutta eivät muodosta käsitystä tekoälylukutaidosta. Tekoälyn tunnistaminen on vain yksi tekoälylukutaidon osa-alue, joka tosin on hyvin pintapuolinen tai kokonaan puuttuva tarkastelemisani suosituissa tekoälylukutaidon määritelmässä (esim. Long & Magerko, 2020).

6.2 Tutkielman eettisyys ja luotettavuus

Tutkielman tekemiseen liittyy suuri määrä eettisiä kysymyksiä ja käytäntöjä, joista tämän tutkielman kannalta olennaista osaa käsittelen tässä luvussa. Pohdin lisäksi tutkielman luotettavuutta. Kuulan (2011) mukaan kaikkiin eettisiin ongelmiin ja kysymyksiin ei ole olemassa valmiita vastauksia, vaan tutkijan on tehtävä ratkaisuja tapauskohtaisesti. Tutkimuksen eettisyys edellyttää yhteisten sääntöjen ja ohjeiden tuntemista ja noudattamista. Tutkimuseettinen neuvottelukunta tarjoaa ohjeistuksen hyvän tieteellisen käytännön peruseriaatteista ja menettelytapoista (Tutkimuseettinen neuvottelukunta [TENK], 2023). Hyvän tieteellisen käytännön peruseriaatteisiin kuuluvat ”luotettavuus, rehellisyys, arvostus ja vastuunkanto” (TENK, 2023). Hyöty, vahingon välttäminen, autonomian kunnioittaminen ja oikeudenmukaisuus ovat pääsääntöisesti ihmistieteiden normien perusta (Kuula, 2011). Olen toteuttanut tämän tutkielman hyvän tieteellisen käytännön periaatteita noudattaen.

Yleisten eettisten ohjeistuksien lisäksi on pohdittava tarkemmin eettisyyttä prosessin jokaisessa vaiheessa tapauskohtaisesti. Jo aiheen valinnassa on pohdittava, voiko aiheesta saada tietoa siten, ettei siitä koidu haittaa tutkittaville. Tekoälyn tunnistamisen keinojen tutkimisesta ei tässä tapauksessa pitäisi aiheutua

haittaa tutkittaville. On toki mahdollisuus, että vastaaja esimerkiksi huolestuu tai ahdistuu omasta osaamisen tasostaan, mutta mahdollinen haitta ei vaikuta kovin todennäköiseltä tai suurelta. Huomioin tämän kuitenkin korostamalla lomakkeen ohjeistuksessa, ettei ole syytä stressaantua vastaamisesta, ja että kaikki vastaukset ovat toivottuja riippumatta vastaajan tietotasosta.

Tutkimuksen toteutuksessa on pohdittava omaa osaamistaan eettisesti muun muassa metodin hallinnan kannalta. Opinnäytetyötä tehdessä oma asiantuntijuus tutkijana on vielä alkutekijöissään. Tutustumalla huolellisesti tutkimusmenetelmäkirjallisuuteen ja eettisiin ohjeisiin voi edistää tutkimuksen eettisyyttä ja minimoida omaan osaamiseen ja toimintaan liittyviä riskejä. Teoreettista viitekehystä kootessa on otettava huomioon aiempi tutkimus ja huolehdittava asianmukaisesta viittaamisesta muihin tutkimuksiin (TENK, 2023). Tutustuin huolellisesti eläytymismenetelmään ennen aineiston keräämistä. Huolehdin myös viitustekniikasta ja siitä, että kaikkiin lähteisiin on viitattu asianmukaisesti.

Aineiston keräämiseen liittyy monenlaisia eettisiä kysymyksiä, joita tulee pohtia. Esimerkiksi jos tutkimus voidaan tehdä olemassa olevalla aineistolla, uuden aineiston kerääminen olisi tutkittavien vaivaamista (Kuula, 2011). Jos kuitenkin on tarpeen kerätä aineisto, osallistujia on informoitava muun muassa oikeuksistaan, tietosuojasta ja kerättävien tietojen käytöstä sekä luottamuksellisuudesta (Cohen ym., 2018). Tietosuojasta huolehtiminen kuuluu aina ihmistieteiden normeihin (Kuula, 2011). Tietosuojasta huolehtimista voi edistää jo suunnitteluvaiheessa pohtimalla, mitä tietoja tutkittavista on ylipäättään tarpeen kerätä ja onko tutkimus mahdollista toteuttaa keräämättä henkilötietoja. Aineiston säilyttämisestä ja tuhoamisesta on myös huolehdittava ja tiedotettava osallistujille (Cohen ym., 2018). Tässä tutkielmassa ei ollut tarpeen kerätä henkilötietoja ja ohjeistin vastaajia, ettei kertomuksiin tule kirjoittaa henkilötietoja, sillä tutkielmaan vastattiin anonyymisti. Huomioin kuitenkin mahdollisten epäsuorien henkilötietojen ilmenemisen mahdollisuuden kertomuksissa ja tein tietosuojailmoituksen (Liite 2). Tietosuojailmoituksessa informoidaan vastaajia siitä, mitä henkilötietoja tutkielmassa kerätään, kuinka niitä käytetään, säilytetään ja kenellä on pääsy kerättyyn henkilötietorekisteriin. Kerroin myös vastaajille, että tutkimukseen osallistuminen on mahdollista keskeyttää kesken vastaamisen.

Tutkielman luotettavuutta voidaan arvioida eettisyyden pohtimisen lisäksi monenlaisista näkökulmista, muun muassa aineiston keruun ja raportoinnin näkökulmista (Tuomi & Sarajärvi, 2018). Tässä tutkielmassa on syytä pohtia joitain näihin teemoihin liittyviä seikkoja luotettavuuden näkökulmasta. Tutkielman aineistonkeruu osoittautui odotettua haastavammaksi, mikä muutti hieman tutkimusjoukon rakennetta. Alkuperäinen suunnitelma oli kerätä aineistoa tilanteessa, jossa mikään tekijä ei rajaa osallistujien taustaa. Kuitenkin aineistonkeruusuunnitelmaa oli muutettava ja sen seurauksena osaa vastaajista yhdistää korkeakoulu organisaationa. On todennäköistä, että yhdistävän taustan, korkeakoulun, takia tavoitin enemmän tekoälystä tietäviä ja sen kanssa enemmän tekemisissä olevia henkilöitä aineistonkeruun toisessa vaiheessa.

Analyysivaiheessa on hyvä pohtia analyysin subjektiivisuutta. Laadullisen tutkimuksessa ja etenkin aineistolähtöisessä tutkimuksessa analyysi on tutkijan tulkintoihin perustuvaa ja siten tutkijan oman taustan ja lähtökohtien värittämää (Tuomi & Sarajärvi, 2018). Tutkijan tulee pyrkiä itsetietoisuuteen ja tarkkaavaisuuteen analyysia tehdessään, vaikka täysi objektiivisuus ei ole mahdollista (Cohen ym., 2018). Tutkijalla on myös eettinen vastuu esittää tulokset rehellisesti, tarkasti ja oikein (Cohen ym., 2018). Tämän tutkielman analyysi perustui omiin tulkintoihini kertomuksista ja erilaisten asioiden yhteyksistä. Pyrin huolelliseen harkintaan ja tarkkuuteen analyysissä, ja analyysin läpinäkyvään kuvaamiseen luvussa 4.3. Tutkielman tuloksissa ei näkynyt merkittävää kehyskertomusten variaationista johtuvaa vaihtelua, joten eläytymismenetelmälle ominainen variointi ei tässä tutkielmassa onnistunut aivan toivotulla tavalla. Tämän tutkielman luotettavuutta pohdittaessa on myös syytä huomioida eläytymismenetelmällä saadun tiedon luonne. Kuten olen luvuissa 3 ja 6.1 todennut, eläytymismenetelmä ei tuota objektiivista tai yleistettävää faktatietoa, vaan katsauksia ilmiöihin, asenteisiin ja ajatuksiin kertomusten kautta. Tämän tutkielman tuloksia ei siis voida pitää yleistettävänä tietona tai valmiina kuvauksena tekoälylukutaidon osa-alueesta, vaan enemmänkin katsauksena siihen, kuinka tekoälyn tunnistaminen näyttäytyy vastaajien kertomuksissa.

6.3 Johtopäätökset

Tämän tutkielman tulokset tarjoavat siis katsauksen aikuisten ajatuksiin ja näkemuksiin tekoälyn tunnistamisesta. Tulokset eivät ole objektiivisia tai yleistettäviä, vaan tarjoavat ainoastaan näkökulmia tutkimuksen teemaan tutkittavien ajatusten kautta. Tutkimuksen tarkoituksena oli vastata tutkimuskysymykseen: *millaisia tekoälyn tunnistamisen keinoja aikuisten kertomuksissa ilmenee?* Aineiston analyysissä muodostui viisi pääluokkaa, jotka vastaavat tutkimuskysymykseen: valppaus, medialukutaito, tekoälyn tunnistamiseen harjaantuminen, tarkempi tilannekohtainen analyysi ja yhdessä oppiminen. Lisäksi aineistosta nousi erillinen, mutta tärkeä teema: huoli tekoälyn kehityksen suunnasta. Eläytymismenetelmälle tyypillinen variointi tuotti tässä tutkielmassa vain pieniä eroja, jotka saattoivat johtua siitä, että toiseen kehyskertomukseen perustuvia kertomuksia oli enemmän. Tutkielman tulokset olivat osin yhteneväisiä aiempien tekoälylukutaidon määritelmien kanssa, vaikka nämä määritelmät ovat jo muutaman vuoden vanhoja ja siten eivät ole huomioineet viimeaikaista generatiivisen tekoälyn kehitystä, joka on erityisesti kasvattanut tekoälyn tunnistamisen merkitystä. Tekoälyn tunnistamista olisikin mielekästä tutkia vielä tarkemmin tekoälylukutaidon määritelmän päivittämiseksi. Erityisen mielenkiintoista olisi tutkia jatkossa tässäkin tutkielmassa ilmenneitä tekoälyn kehityksen suuntaan liittyviä huolia ja epävarmuuksia. Lisäksi tekoälylukutaidon ja medialukutaidon suhteen tarkempi tutkiminen ja jäsentäminen olisi erilaisten lukutaitokäsitteiden päällekkäisyyden vähentämisen ja selkeyttämisen näkökulmasta mielekästä.

LÄHTEET

- Aufderheide, P. (1992). *Media Literacy. A Report of the National Leadership Conference on Media Literacy*. The Aspen Institute.
- Ciampa, K., Wolfe, Z. M. & Bronstein, B. (2023). ChatGPT in education: Transforming digital literacy practices. *Journal of Adolescent & Adult Literacy*, 67(3), 186–195. <https://doi.org/10.1002/jaal.1310>
- Cohen, L., Manion, L. & Morrison, K. (2018). *Research methods in education* (Eighth edition.). Routledge, Taylor & Francis Group.
- Cope, B. & Kalantzis, M. (2024). On Cyber-Social Learning: A Critique of Artificial Intelligence in Education. Teoksessa D. Kourkoulou, A.-O. Tzirides, B. Cope & M. Kalantzis (toim.), *Trust and Inclusion in AI-Mediated Education : Where Human Learning Meets Learning Machines* (s. 3–34). Springer.
- Eskola, J. (1997). *Eläytymismenetelmäopas*. Tampereen yliopisto.
- Eskola, K. & Eskola, J. (1998). Tuottaako eläytymismenetelmä ainoastaan stereotypioita? – esimerkkinä teatterissa käymisen tutkiminen. Teoksessa J. Eskola, *Eläytymismenetelmä sosiaalitutkimuksen tiedonhankintamenetelmänä* (s. 321–337). Taju.
- Eskola, J., Mäenpää, T. & Wallin, A. (2017a). Digitalisoituva ja monikielistyvä eläytymismenetelmä aikamme muutosten tutkimusvälineenä. Teoksessa J. Eskola & T. Mäenpää & A. Wallin (toim.), *Eläytymismenetelmä 2017: Perusteema ja 11 muunnelmaa* (s. 6–15). Tampere University Press.
- Eskola, J., Karayilan, S., Kaski, T., Lehtola, T., Mäenpää, T., Nishimura-Sahi, O., Oede, A.-M., Rantanen, M., Saarinen, S., Toivikko, P., Valtonen, M. & Wallin, A. (2017b). Ohjeita ja kokemuksia menetelmästä kiinnostuneille. Teoksessa J. Eskola & T. Mäenpää & A. Wallin (toim.), *Eläytymismenetelmä 2017: Perusteema ja 11 muunnelmaa* (s. 6–15). Tampere University Press.

- Eskola, J. & Suoranta, J. (1998). *Johdatus laadulliseen tutkimukseen*. Vastapaino.
- Fengchun, M. & Holmes, W. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization, UNESCO. <https://doi.org/10.54675/EWZM9535>
- Frau-Meigs, D. (2024). *User empowerment through media and information literacy responses to the evolution of generative artificial intelligence (GAI)*. United Nations Educational, Scientific and Cultural Organisation, UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000388547>
- Heikkinen, H. L. T. (2018). Kerronnallinen tutkimus. Teoksessa R. Valli (toim.), *Ikkunoita tutkimusmetodeihin 2. Näkökulmia aloittelevalle tutkijalle tutkimuksen teoreettisiin lähtökohtiin ja analyysimenetelmiin*. (5., uudistettu ja täydennetty painos.). PS-kustannus.
- Jones, J. W. (2024). AI in Education: Some Thoughts About Ethics, Equity, and Social Impact. Teoksessa D. Kourkoulou, A.-O. Tzirides, B. Cope & M. Kalantzis (toim.), *Trust and Inclusion in AI-Mediated Education : Where Human Learning Meets Learning Machines* (s. 59–72). Springer.
- Kuula, Arja. (2011). *Tutkimusetiikka : aineistojen hankinta, käyttö ja säilytys* (2. uud. p.). Vastapaino.
- Laupichler, M. C., Aster, A. & Raupach, T. (2023). Delphi study for the development and preliminary validation of an item set for the assessment of non-experts' AI literacy. *Computers and Education: Artificial Intelligence* 4, 1–10. <https://doi.org/10.1016/j.caeai.2023.100126>
- Laupichler, M. C., Aster, A., Schirch, J. & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. *Computers and Education: Artificial Intelligence*, 3, 1–15. <https://doi.org/10.1016/j.caeai.2022.100101>
- Lee, N. M. (2018). Fake news, phishing, and fraud: a call for research on digital media literacy education beyond the classroom. *Communication Education*, 67(4), 460–466. <https://doi.org/10.1080/03634523.2018.1503313>
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I. & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The*

- International Journal of Management Education*, 21(2), 1–13.
<https://doi.org/10.1016/j.ijme.2023.100790>
- Livingstone, S. (2004). What is media literacy? *Intermedia*, 32(3), 18–20.
- Łodzikowski, K., Foltz, P. W. & Behrens, J. T. (2024). Generative AI and Its Educational Implications. Teoksessa D. Kourkoulou, A.-O. Tzirides, B. Cope & M. Kalantzis (toim.), *Trust and Inclusion in AI-Mediated Education : Where Human Learning Meets Learning Machines* (s. 35–57). Springer.
- Long, D. & Magerko, B. (2020). *What is AI Literacy? Competencies and Design Considerations. 2020 CHI Conference on Human Factors in Computing Systems - Proceedings*. <https://doi.org/10.1145/3313831.3376727>
- Lähdesmäki, S. (2024). Eettinen perusta tekoälyn hyödyntämiseen koulutuksessa. *Kasvatus & Aika*, 18(1), 41–49.
<https://doi.org/10.33350/ka.142410>
- Mönkkönen, K. (2021). Eläytymismenetelmä. Teoksessa S. Ryyänen & A. Rannikko (toim.), *Tutkiva mielikuvitus. Luovat, osallistuvat ja toiminnalliset tutkimusmenetelmät yhteiskuntatieteissä*. Gaudeamus.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W. & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education. Artificial Intelligence* 2, 1–11.
<https://doi.org/10.1016/j.caeai.2021.100041>
- Okkonen, J. & Kotilainen, S. (2019). Minors and Artificial Intelligence : Implications to Media Literacy. *Information Technology and Systems: Proceeding of ICITS 2019*, 881–890. https://doi.org/10.1007/978-3-030-11890-7_82
- Potter, W. J. (2022). Analysis of definitions of media literacy. *The Journal of Media Literacy Education*, 14(2), 27–43. <https://doi.org/10.23860/JMLE-2022-14-2-3>
- Potter, W. J. (2004). *Theory of media literacy : a cognitive approach* (1st ed.). SAGE.
- Rasi, P., Vuojärvi, H. & Ruokamo, H. (2019). Media Literacy Education for All Ages. *The Journal of Media Literacy Education*, 11(2), 1–19.
<https://doi.org/10.23860/JMLE-2019-11-2-1>

- Terras, M. M., Ramsay, J. & Boyle, E. (2011). A Psychological Perspective on Media Literacy. Teoksessa S. D. Grahame (toim.), *Science Education in a Rapidly Changing World* (s. 75–85). Nova Science Publishers Inc.
- Tiernan, P., Costello, E., Donlon, E., Parysz, M. & Scriney, M. (2023). Information and Media Literacy in the Age of AI: Options for the Future. *Education Sciences*, 13(9), 1–11.
<https://doi.org/10.3390/educsci13090906>
- Traficom. (2022). *Tekoälyn mahdollistamat kyberhyökkäykset*.
https://www.kyberturvallisuuskeskus.fi/sites/default/files/media/publication/TRAFICOM_Teko%C3%A4lyn_mahdollistamat_kyberhy%C3%B6kk%C3%A4ykset%202022-12-12_web.pdf
- Tuomi, J. & Sarajärvi, A. (2018). *Laadullinen tutkimus ja sisällönanalyysi*. Tammi.
- Tutkimuseettinen neuvottelukunta, TENK. (2023). *Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsitleminen Suomessa. Tutkimuseettisen neuvottelukunnan HTK-ohje 2023*. https://tenk.fi/sites/default/files/2023-03/HTK-ohje_2023.pdf
- Vaaherkumpu, S. & Terävä, H. (27.5.2023). *Tekoälyn kehitys pelästyttää ja ala pyytää lisää sääntelyä – suomalaistutkijan mukaan merkittäviä haittoja on jo nähty*. Yle Uutiset | yle.fi. <https://yle.fi/a/74-20033804>
- Wach, K., Duong, C. D., Ejdys, J., Kazlauskaitė, R., Korzynski, P., Mazurek, G., Paliszkievicz, J. & Ziemba, E. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. *Entrepreneurial Business and Economics Review*, 11(2), 7–30.
<https://doi.org/10.15678/EBER.2023.110201>
- Wallin, A., Koro-Ljungberg, M. & Eskola, J. (2018). The method of empathy-based stories. *International Journal of Research & Method in Education*, 42(5), 525–535. <https://doi.org/10.1080/1743727X.2018.1533937>
- Wuyckens, G., Landry, N. & Fastrez, P. (2022). Untangling media literacy, information literacy, and digital literacy: A systematic meta-review of core concepts in media education. *Journal of Media Literacy Education*, 14(1), 168–182. <https://doi.org/10.23860/JMLE-2022-14-1-12>

LIITTEET

Liite 1: Kehyskertomukset

(A) Tuttavasi kertoo sinulle, että törmäsi taannoin netissä erääseen julkaisuun, ja epäili, että tekoäly on tehnyt sen. Hän kertoo, millainen julkaisu oli ja mitä hän oli ajatellut pohtiessaan, oliko se todella tekoälyn luoma. Käytä mielikuvitustasi ja kerro keskustelustanne.

(B) Tuttavasi kertoo sinulle, että törmäsi taannoin netissä erääseen julkaisuun, ja oli varma, että tekoäly on tehnyt sen. Hän kertoo, millainen julkaisu oli ja mitä hän oli ajatellut pohtiessaan, oliko se todella tekoälyn luoma. Käytä mielikuvitustasi ja kerro keskustelustanne.

Liite 2: Tietosuojailmoitus



Opinnäytetutkimuksen tietosuojailmoitus EU:n yleinen tietosuoja-asetus (2016/679), art. 12–14

Rekisterin nimi	Kertomuksia tekoälylukutaidosta [pro gradu -tutkielma]
Päiväys	10.10.2024
Rekisterinpitäjä(t)	Nea Auranen nea.auranen@tuni.fi
Muut henkilötietoja käsittelevät henkilöt	
Ohjaaja tai oppilaitoksen yhteyshenkilö	Reijo Kupiainen reijo.kupiainen@tuni.fi
Henkilötietojen käsittelytarkoitus ja käsittelyperuste	<p>Mahdollisia epäsuoria henkilötietojasi käsitellään aikuisten tekoälylukutaitoon liittyvässä opinnäytetutkimuksessa. Tutkimuksen tarkoituksena on selvittää, minkälaisia tekoälylukutaidon osa-alueita aikuisten tekoälyyn liittyvissä kertomuksissa ilmenee. Pyrkimyksenä on siis tarkastella tekoälylukutaitoa aikuisten kertomusten kautta.</p> <p>Tutkimukseen osallistuminen on vapaaehtoista. Henkilötietojen käsittelyperusteena on:</p> <p><input checked="" type="checkbox"/> suostumus. EU:n yleinen tietosuoja-asetus 6 artikla 1.a-kohta. Suostumuksen voi peruuttaa milloin tahansa ilmoittamalla tästä rekisterinpitäjälle. Suostumuksen peruuttaminen ei vaikuta ennen suostumuksen peruuttamista suoritettun käsittelyn lainmukaisuuteen. TAI</p> <p><input type="checkbox"/> yleisen edun mukainen tieteellinen tutkimus. EU:n yleinen tietosuoja-asetus 6 artikla 1 e-kohta. Ks. ohje Henkilötieto opinnäytetyössä</p>
Henkilötietojen säilytysaika	<p>Opinnäytteen valmistuttua aineisto ja henkilötiedot tuhoetaan. Arvioitu valmistusaika 12/2024.</p> <p>Siltä osin kuin ohjaajalla on pääsy aineistoon opinnäytetyön ohjaamista ja tarkastamista varten, ohjaajat ja tarkastajat käsittelevät henkilötietoja ainoastaan niin kauan kuin on tarpeellista työn hyväksymistä varten.</p>
Rekisterin tietosisältö ja tietolähteet	<p>Kuvaus rekisterissä käsiteltävistä henkilötietotyypeistä tietoryhmittäin:</p> <ul style="list-style-type: none">- Mahdolliset epäsuorat henkilötiedot vastaajan kirjoittamassa kertomuksessa.

	Tiedot kerätään tutkittavilta itseltään.
Henkilötietojen vastaanottajat	Henkilötietojasi ei luovuteta ulkopuolisille.
Rekisterin suojauksen periaatteet	Manuaalinen aineisto säilytetään lukitussa tilassa/kaapissa. Digitaalinen aineisto suojataan käyttäjätunnuksella ja salasanalla tai kaksivaiheisella käyttäjän tunnistuksella (MFA). Aineistosta poistetaan suorat tunnistetiedot.
Rekisteröidyn oikeudet	Tietosuojalainsäädännön mukaisesti sinulle kuuluu oikeus saada pääsy tietoihin, oikaista tietoja, oikeus tietojen poistamiseen (oikeus tulla unohdetuksi), rajoittaa tietojen käsittelyä ja vastustaa henkilötietojen käsittelyä. Jos haluat käyttää jotain oikeuttasi, ota yhteys rekisterinpitäjään.
Oikeus valittaa viranomaiselle	Sinulla on oikeus tehdä valitus henkilötietojen käsittelyä valvovalle viranomaiselle, jos epäilet henkilötietojasi käsiteltävän vastoin tietosuojalainsäädäntöä: tietosuoja.fi / sähköposti: tietosuoja@om.fi

Tampereen yliopisto
33014 Tampereen yliopisto
Puh. 0294 5211
Y-tunnus 2844561-8

Tampereen ammattikorkeakoulu
Kuntokatu 3, 33520 Tampere
Puh. 0294 5222
Y-tunnus 1015428-1

www.tuni.fi