



Multimodaalisen tunteiden tunnistamisen eettiset näkökulmat

Jenna Kauranen

Haaga-Helia ammattikorkeakoulu

Tradenomi

AMK-opinnäytetyö

2023

Tiivistelmä

Tekijä Jenna Kauranen
Tutkinto Tradenomi
Opinnäytetyön nimi Multimodaalisen tunteiden tunnistamisen eettiset näkökulmat
Sivu- ja liitesivumäärä 46 + 2
<p>Tunteet ovat merkittävä osa elämää, ja niillä on suuri vaikutus ihmisten välisessä vuorovaikutuksessa. Multimodaalinen tunteiden tunnistaminen tavoittelee ihmisen tunteiden tunnistamista saman aikaisesti usean signaalin välityksellä, kuten ilmeistä ja puheesta. Usein tunteiden tunnistamiseen hyödynnetään tekoälysovelluksia. Teknologia mahdollistaa entistä paremman tunneälyn koneille, joka parantaa koneen ja ihmisen välistä vuorovaikutusta.</p> <p>Opinnäytetyössä tutkittiin, kuinka multimodaalisen tunteiden tunnistamisen eettiset näkökulmat koetaan opetus- ja koulutuslalla. Tutkimuksessa selvitettiin mitä signaaleita tai teknologioita tunteiden tunnistamiseen on olemassa, mitä tunnetietoja niillä saadaan kerättyä, minkälaisia mahdollisuuksia ja toisaalta myös haasteita nähdään sekä minkälaisia asioita tulee ottaa huomioon eettisistä näkökulmista.</p> <p>Tutkimus toteutettiin laadullisena tutkimuksena kevään 2023 aikana. Teoriaosuudessa tutkittiin aihetta monipuolisesti ensin yleisesti affektiivisen tietojenkäsittelyn näkökulmasta, jonka jälkeen aiheeseen syvennyttiin multimodaalisen tunteiden tunnistamisen ja eettisen viitekehyksen näkökulmista. Keskeinen lähdeaineisto muodostui tuoreista multimodaalista tunteiden tunnistamista käsittelevistä tieteellisistä artikkeleista.</p> <p>Empiirisessä osuudessa haastateltiin kahdeksaa henkilöä, joilla tiedettiin etukäteen olevan tietoa jostain tutkimuksen aihealueesta. Haastateltavat olivat aiheen asiantuntijoita, opiskelijoita ja huoltajia. Haastateltavien valinnoilla pyrittiin kattamaan kaikki opetusasteet varhaiskasvatuksesta jatkuvaan oppimiseen saakka. Kaikki haastattelut nauhoitettiin ja litteroitiin, joita sitten analysoitiin.</p> <p>Tutkimuksen tuloksena syntyi käsitys siitä, kuinka multimodaalinen tunteiden tunnistaminen koetaan eettisistä näkökulmista opetus- ja koulutuslalla. Keskeisiä eettisiä teemoja ovat yksityisyys, oikeudenmukaisuus ja syrjimättömyys. Kaikki nämä yhdistyvät dataan, johon liittyvät myös merkittävimmät haasteet. Tärkeäksi koettiin tiedot, kuka dataa käsittelee, mihin tarkoitukseen ja käyttävätkö kolmannet osapuolet tietoja. Kyse on erittäin sensitiivisestä datasta, joka saattaa paljastaa henkilökohtaisia tietoja, joita ei haluta paljastettavan itsestä.</p> <p>Multimodaalisesta tunteiden tunnistamisesta tulisi olla itselle jotakin hyötyä, jotta siihen oltaisiin ylipäättään suostuvaisia. Teknologialla nähdään mahdollisuuksia opetus- ja koulutuslalla, kuten yksilöllisemmän oppimiskokemuksen mahdollistamisen ja opetuksen tehokkuuden parantamisen. Lisäksi olisi mahdollisuus tutkia, mitä oppimisprosessin aikana tapahtuu.</p>
Asiasanat Tekoäly, tietojenkäsittely, tunteet, etiikka

Abstrack

Author Jenna Kauranen
Degree programme Tradenomi
Thesis title Multimodaalisen tunteiden tunnistamisen eettiset näkökulmat
Number of pages and appendix pages 46 + 2
<p>Emotions are a significant part of life, and they have a huge impact on human interaction. Multimodal emotion recognition aims to recognize human emotions at the same time through multiple signals, such as facial expressions and speech. Artificial intelligence applications are often used to identify emotions. The technology enables even better emotional intelligence for machines, which can improve the computer-human interaction.</p> <p>The thesis studied how the ethical aspects of multimodal emotion recognition are experienced in education sector. The study answers the questions what signals or technologies there is for recognizing emotions, what kind of emotional data can be collected, what kind of opportunities and challenges are seen and what kind of issues should be notice from an ethical perspective.</p> <p>The study is a qualitative research and it was done during the spring 2023. In the theoretical part, the topic was explored first in a generally from the perspective of affective computing. Then the study goes deeper into a multimodal emotion recognition and ethical framework. The most important sources of data to understand the topic includes recent scientific articles dealing with multimodal emotion recognition.</p> <p>The empirical part consisted of interviews where eight people were interviewed. All the interviewees were known in advance to have knowledge about some topic area of the research. The interviewees were experts, students and mothers. The interviewees were chosen to cover all levels of education, from early childhood education to continuous learning. All interviews were recorded and transcribed. Then they were manually analyzed.</p> <p>As a result is an understanding of how multimodal emotion recognitios is experienced from ethical perspectives in education sector. The main ethical themes are privacy, justice, equality and non-discrimination. All of these are connected to the data, which is also associated with the most significant challenges. Information about who processes the data, for what purpose and whether third parties use the data was considered important. It's about very sensitive data that may reveal peronal information that you don't want to be revealed about yourself.</p> <p>Multimodal emotion recognition should be of some benefit to oneself that people would agree to it. Technology has opportunities in education sector, such as enabling a more individualized learning experience and improving the effectiveness of teaching. In addition, there would be an opportunity to study what happens during in the learning process.</p>
Keywords Artificial Intelligence, computing, feelings, ethics

Sisällys

1	Johdanto	1
1.1	Tausta ja tavoitteet	1
1.2	Sisältö, lopputulos ja rajaus	2
1.3	Käsitteet	3
2	Affektiivinen tietojenkäsittely	6
2.1	Affektiivinen tietojenkäsittely on monitieteinen tutkimusalue	6
2.2	Affektiivisen tietojenkäsittelyn tekniikat	7
2.3	Affektiivisen tietojenkäsittelyn haasteita	8
2.4	Koulutusteknologia ja tunnetekoäly opetuslalla Kiinassa	9
2.5	Psykologiset mallit tunteiden mittaamiseen	10
3	Multimodaalinen tunteiden tunnistaminen	12
3.1	Multimodaalisen tunteiden tunnistamisen merkitys	12
3.2	Tunteiden tunnistaminen ihmisen fyysisistä ja psykologisia muutoksista	13
3.2.1	Tunteiden tunnistaminen kasvojen ilmeistä	13
3.2.2	Tunteiden tunnistaminen puheesta	16
3.2.3	Tunteiden tunnistaminen aivosähkökäyrästä	17
3.3	Tunteiden tunnistaminen digitaalisen median lähteistä	18
3.4	Tunneteknologia keräämisen vaiheet ja luokat	18
3.5	Multimodaalisen tunteiden tunnistamisen haasteita	19
3.6	Multimodaalisen tunteiden tunnistamisen hyödyt	20
3.7	Multimodaalisen tunteiden tunnistamisen käyttökohteita opetus- ja koulutuslalla	21
4	Tekoällyn etiikka koulutuksen näkökulmasta	23
4.1	Tekoällyjärjestelmien eettiset vaikutukset	23
4.2	Tekoällyä koskeva Euroopan Unionin politiikka ja muu lainsäädäntö	24
4.3	Tekoällyn käyttö koulutuslalla	25
4.4	Opetus- ja koulutuslalan eettinen viitekehys	27
5	Tutkimus multimodaalisesta tunteiden tunnistamisesta ja sen eettisistä näkökulmista	30
5.1	Tutkimuksen esittely	30
5.2	Tutkimuskysymykset, tutkimusmenetelmät, aineisto	30
5.3	Tutkimusaineiston analysointi	31
5.4	Tutkimuksen tulokset	32
5.4.1	Olemassa olevat signaalit ja tekniikat	33
5.4.2	Multimodaalisen tunteiden tunnistamisen mahdollisuudet	34
5.4.3	Multimodaalisen tunteiden tunnistamisen haasteet	36
5.4.4	Multimodaalisen tunteiden tunnistamisen eettiset näkökulmat	38

6	Pohdinta ja johtopäätökset	42
6.1	Johtopäätökset.....	42
6.2	Keskeiset tulokset	43
6.3	Tutkimuksen luotettavuus.....	45
6.4	Suositukset ja kehittämismahdollisuudet	46
	Lähteet.....	47
	Liitteet	53
	Liite 1. Saatekirje.....	53
	Liite 2. Haastattelukysymykset	54

1 Johdanto

Ihmiset ovat tunteellisia olentoja (Zhao, Jia, Yang, Ding & Keutzer 2021). Tunteet ovat tärkeitä ja ne vaikuttavat ihmisten elämään monella tavalla. Tunteilla on merkittävä rooli ihmisten välisessä kommunikaatiossa ja vuorovaikutuksessa, ja ne mahdollistavat itsensä ilmaisun myös sanallisen alueen ulkopuolella (Sebe, Cohen, Huang 2005).

Tässä työssä tutkittiin multimodaalisen tunteiden tunnistamisen eettisiä näkökulmia opetus- ja koulutusalailla. Aihetta lähestyttiin affektiivisen tietojenkäsittelytieteen kautta, jonka yksi tutkimuskohde on multimodaalinen tunteiden tunnistaminen. Multimodaalinen tunteiden tunnistaminen pyrkii tunnistamaan ihmisen tunteita useiden eri modalityettien välityksellä, kuten kuvan, äänen tai tekstin perusteella (Zhao, Yang & Zhang 2022). Yleisesti ottaen tavoitteena on ihmisen emotionaalisen tilan havaitseminen luonnollisessa tilanteessa (Sebe ym. 2005).

Viime aikaiset tekniset edistysaskeleet ovat mahdollistaneet ihmiskäyttäjille sellaisia tapoja kommunikoida koneiden kanssa, joita ei aiemmin voitu edes kuvitella. Tietokoneiden kyvystä ymmärtää ihmisen tunteita on tullut toivottavaa useissa sovelluksissa. (Sebe ym. 2005) Kuten voidaan todeta tietokoneiden kyky ymmärtää ihmisten tunteita on ollut ajankohtaista jo vuonna 2005. Aihe on ajankohtainen edelleen, sillä teknologia kehittyy valtavalla vauhdilla mahdollistaen aina vain uusia asioita.

Tunteisiin liittyvää tekoälyteknologiaa voidaan hyödyntää monilla aloilla useissa sovelluksissa, joten teknologiajätit ja startupit ovat alkaneet investoida tietokonenäköön ja äänianalyysiin ihmisten tunteiden tunnistamiseksi. Sen seurauksena affektiivisen tietojenkäsittelyn markkinoiden on arvioitu kasvavan yli 35% vuosittain vuosina 2023-2027. (Dilmegani 2023.) Näiden järjestelmien yleistymisen ja arkipäiväistyminen nostavat esiin eettisiä kysymyksiä, joita on syytä pohtia.

Järjestelmät automaattiseen tunteiden tunnistamiseen ja mieliala-analyysiin voivat olla suuren edistyksen edesauttajia, mutta ne voivat olla myös suuren harmin mahdollistajia. Siksi on välttämätöntä, että affektiivisen tietojenkäsittelytieteen yhteisö osallistuu aktiivisesti luomustensa eettisiin seuraamuksiin. (Mohammad 2022.)

1.1 Tausta ja tavoitteet

Opinnäytetyö toteutetaan osana AI Forum -hanketta ja siihen liittyvää ”multimodal emotion recognition” -osahanketta. Opinnäytetyön hyöty muodostuu aiheen tutkimisesta eettisestä näkökulmasta. Tavoitteena on tutkia tekoälyn tarjoamia mahdollisuuksia ja haasteita multimodaaliseen tunteiden tunnistamiseen, sekä miten nämä asiat koetaan eettisesti.

1.2 Sisältö, lopputulos ja rajaus

Multimodaalinen tunteiden tunnistaminen on hyvin laaja aihepiiri, johon liittyen on tehty paljon tutkimuksia. Työni kannalta merkittävimmät tutkimukset liittyvät multimodaalisen tunteiden tunnistamisen keskeisiin näkökulmiin, joihin kuuluvat laajasti käytetyt tunneteoriat, tunteiden tunnistamiseen käytettävät teknologiat ja niiden käyttökohteet sovelluksissa. Tarkastelen teemoja peilaten niitä eettisiin viitekehyksiin.

Opinnäytetyössä selvitetään multimodaaliseen tunteiden tunnistamiseen liittyviä asioita seuraavien tutkimuskysymyksien kautta:

- Mitä signaaleita tai teknologioita tunteiden tunnistamiseen on olemassa,
- Mitä tunnetietoja niillä saadaan kerättyä,
- Minkälaisia mahdollisuuksia ja toisaalta myös haasteita tekoäly tarjoaa multimodaaliseen tunteiden tunnistamiseen opetus- ja koulutusalailla,
- Minkälaisia asioita tulee huomioida eettisistä näkökulmista?

Opinnäytetyössä huomioidaan tekoälyyn liittyvä lainsäädäntö, kuten Euroopan Unionin tekoälystrategia ja Euroopan Unionin yleinen tietosuojasetus (GDPR).

Opinnäytetyöprojektissa tutkimus rajataan koskettamaan koulutus- ja opetusalaa, jolloin muut alat jäävät tutkimuksen ulkopuolelle. Mukaan otetaan kaikki koulutustasot päiväkodista jatkuvaan oppimiseen. Saadakseni näkökulmia mahdollisuuksiin tutkin aihetta maailman laajuisesti, jolloin syntyy käsitys mitä kaikkea aiheen parissa on jo tehty.

Empiirisen osuuden tutkimus toteutettiin laadullisena tutkimuksena. Tutkimusaineistoa kerättiin haastattelemalla henkilöitä siitä, kuinka he kokevat multimodaalisen tunteiden tunnistamisen ja siihen liittyvät eettiset näkökulmat. Tutkimustuloksena syntyi käsitys siitä, kuinka opetus- ja koulutusalailla koetaan multimodaaliseen tunteiden tunnistamiseen liittyvät asiat eettisestä näkökulmasta. Opinnäytetyön tutkimustuloksia voidaan hyödyntää AI Forum hankkeen puitteissa.

Peittomatriisi (taulukko 1) havainnollistaa tutkimuksen teoria- ja empiriaosuuden linkittymistä tutkimuskysymyksiin.

Taulukko 1. Peittomatriisi

Tutkimuskysymykset	Teoreettinen viitekehys	Haastattelukysymys	Tulokset
Mitä signaaleita tai teknologioita tunteiden tunnistamiseen on olemassa?	2.2, 3.2, 3.3	4, 5	5.4.1
Mitä tunnetietoja niillä saadaan kerättyä?	3.4	4, 5	5.4.1
Minkälaisia mahdollisuuksia tekoäly tarjoaa multimodaaliseen tunteiden tunnistamiseen opetus- ja koulutusosalalla	2.4, 3.6, 3.7	8, 10, 11	5.4.2
Minkälaisia haasteita multimodaaliseen tunteiden tunnistamiseen liittyy?	2.3, 3.5	9, 12, 13	5.4.3
Minkälaisia asioita tulee huomioida eettisistä näkökulmista?	Luku 4	6, 7, 8, 9	5.4.4

1.3 Käsitteet

Automaattinen tunteiden tunnistus ("Automatic Emotion Recognition", AER) viittaa järjestelmään, jolla tunnistetaan ihmisen tunteita. Se hyödyntää teknologioita, kuten signaalinkäsittelyä, koneoppimista ja tietokonenäköä. (IGI Global s.a.)

Affektiivinen tietojenkäsittely ("Affective Computing") on sellaisten järjestelmien tutkimusta ja kehittämistä, jotka voivat tunnistaa, tulkita, käsitellä ja simuloida ihmisen tunteita. Se on monitieteinen ala, joka käsittelee tietojenkäsittelytieteitä, psykologiaa ja kognitiivisia tieteitä. (Banafa 2016.)

Aivosähkökäyrä ("Electroencephalogram", EEG) on testi, joka mittaa aivojen sähköistä aktiivisuutta. Siinä käytetään pieniä metallilevyjä (elektrodeja), jotka kiinnitetään päänahkaan. Aivosolut kommunikoivat sähköisten impulssien avulla ja ne ovat aktiivisia koko ajan. Ne ovat aktiivisia jopa unen aikana. (Mayo Foundation for Medical Education and Research s.a.)

Ilmeidenkoodausjärjestelmä ("Facial Action Coding System", FACS) on Ekmanin ja Friesenin kehittämä standardi. Se on kattava anatomiaan perustuva järjestelmä, joka kuvaa kaikkia visuaalisesti havaittavia ilmeitä. Se pilkkoo kasvojen ilmeet yksittäisiksi lihasliikkeen osiksi, joita kutsutaan toimintayksiköiksi. (Paul Ekman Group LLC s.a.a.)

Ilmeidentunnistus ("Facial Emotion Recognition", FER) on teknologia, jota käytetään analysoimaan tunteita eri lähteistä, kuten kuvista ja videoista. Se kuuluu affektiivisen tietojenkäsittelyn teknologiaperheeseen ja se perustuu usein tekoälyteknologiaan. (Vemou & Horvath 2021.)

Kasvojen solmupisteet ("Facial landmarks", FLs) ovat kasvojen alueella olevia visuaalisesti näkyviä kohtia, kuten nenänpää, kulmakarvat ja suu. Jokaista kahden pisteen parittaista sijaintia käytetään ilmeidentunnistuksen piirrevektorina. (Ko 2018.)

Koneoppinen ("Machine Learning") tarkoittaa toimintaa, jossa joukko algoritmeja oppivat tekemään päätöksiä tai jäsentämään dataa. Dataan perustuvia koneoppimisen muotoja ovat ohjattu ja ohjaamaton oppiminen. Koneoppimisen muotoa, jossa algoritmit oppivat yrityksen ja erehdyksen kautta, kutsutaan vahvistusoppimiseksi. (Rusanen, Nurminen, Räisänen, Tarkoma & Halmetoja s.a.)

Mikroilmeet ("Micro Expressions", MEs) kuvastavat spontaaneita ja hienovaraisia kasvojen liikkeitä, jotka ilmestyvät tahattomasti. Niillä on taipumus paljastaa lyhyessä ajassa ihmisen aidot ja piilevät tunteet. (Ko 2018.)

Multimodaalinen tunteiden tunnistaminen ("Multimodal Emotion Recognition") pyrkii tunnistamaan ihmisen tunteita eri modaaliteettien välityksellä, jotka liittyvät tunteiden ilmaisuun. Näitä ovat esimerkiksi ääni, kuva ja teksti. (Xiaoming, Yijiao & Shiqing 2021).

Perustunteet ("Basic Emotions", BE) ovat ihmisen seitsemän perustunnetta, jotka ovat ilo, yllättyneisyys, viha, suru, pelko, inho ja neutraali (Ko 2018).

Syväoppiminen ("Deep Learning") on koneoppimistekniikka, joka opettaa tietokoneita oppimaan esimerkin avulla. Syväoppimisessa tietokonemalli oppii suorittamaan luokittelutehtäviä suoraan kuvista, tekstistä tai äänestä. Nämä mallit voivat saavuttaa huippuluokan tarkkuuden ja joskus ylittämään ihmistasoisen suorituskyvyn. Malleja opetetaan käyttämällä suurta määrää luokiteltua dataa ja neuroverkkojen arkkitehtuuria, jotka sisältävät monia kerroksia. (The MathWorks, Inc. s.a.)

Tekoäly ("Artificial Intelligence", AI) on yleistermi, joka kuvaa erilaisia tekniikoita, joita käytetään tietokoneiden älykkään toiminnan luomiseksi. Tekoälylle ei ole varsinaisesti sovittua määritelmää, mutta keskeisiä tekoälyn osia ovat kyky suorittaa tehtäviä ilman valvontaa ja kyky oppia suorituskyvyn parantamiseksi. (Rusanen ym. s.a.)

Tietokonenäkö ("Computer Vision") on tekoälyn ala, jonka avulla tietokoneet ja järjestelmät voivat saada merkityksellistä tietoa digitaalisista kuvista, videoista ja muista visuaalisista syötteistä. Pohjautuen tähän tietoon ne voivat ryhtyä toimiin tai antaa suosituksia. Jos tekoälyn ajatellaan mahdollistavan tietokoneiden ajattelemisen, niin tietokonenäkö mahdollistaa näkemisen, havainnoin ja ymmärtämisen. (IBM s.a.)

Tunnetekoäly ("Emotion AI", "Affective AI", "Affective Computing") viittaa tekoälyyn, joka havaitsee ja tulkitsee ihmisen emotionaalisia signaaleita esimerkiksi tekstistä, äänestä ja videoista, tai näiden yhdistelmistä. Tekstin analysointiin voidaan hyödyntää luonnollisen kielen käsittelyä ja tunteiden analysointia. Äänen analysointiin voidaan käyttää äänitunnetekoälyä. Videoissa voidaan puolestaan hyödyntää esimerkiksi ilmeanalyysiä, kävelyanalyysiä ja fysiologisia signaaleita. (Gossett 2023.)

Tunteiden tunnistaminen puheesta ("Speech Emotion Recognition", SER) on toiminto, jolla pyritään tunnistamaan ihmisen tunteita ja tunnetiloja puheesta. Menetelmässä hyödynnetään tosiasiaa, että äänensävy ja -korkeus usein heijastavat tunteita. (Burnwal 2020.)

2 Affektiivinen tietojenkäsittely

Tässä luvussa selvitetään mitä on affektiivinen tietojenkäsittely, miksi se on merkittävää ja minkälaisia tekniikoita siinä hyödynnetään. Sen jälkeen tarkastellaan minkälaisia käyttökohteita opetuslalla on maailmanlaajuisesti. Tutustumme myös miten affektiivista tietojenkäsittelyä hyödynnetään Kiinassa ja minkälaisia haasteita siellä on kohdattu.

Luvussa perehdytään eri kappaleissa siihen, minkälaista tutkimusta affektiivisesta tietojenkäsittelystä on tehty, jotta saadaan hyvä yleiskuva aiheesta. Luvun lopuksi tutustutaan muutamiin yleisimpiin psykologisiin malleihin tunteiden luokitteluksi, joiden ymmärtäminen on tärkeää multimodaalisen tunteiden tunnistamisen menetelmien ymmärtämiseksi.

2.1 Affektiivinen tietojenkäsittely on monitieteinen tutkimusalue

Tunteet ovat suuressa osassa elämässämme. Näin ollen affektiivinen tietojenkäsittely, joka käsittelee tunteita ja tietojenkäsittelyä usein tekoälyn avulla, on valtavan tärkeä ja vilkas työlinja. (Mohammad 2022.) Se on ala, joka yhdistää tietojenkäsittelytieteitä, psykologiaa ja kognitiivista tiedettä (Dilmegani 2022, Banafa 2016).

Affektiivinen tietojenkäsittely on laaja monitieteinen tutkimusalue, joka tutkii perustavaa laatua olevia kysymyksiä (esimerkiksi mitä ovat tunteet) ja kaupallisia sovelluksia (esimerkiksi voivatko tietokoneet tunnistaa kuluttajan tunteet) (Mohammad 2022). Se on nouseva teknologia, jonka avulla järjestelmien on mahdollista tunnistaa, käsitellä ja simuloida ihmisten tunteita (Dilmegani 2022.) Koneen tulisi tulkita ihmisen tunteita ja mukauttaa toimintaansa sen mukaisesti (Banafa 2016).

On suuri kysymys, voivatko tietokoneet tunnistaa tunteita (Dilmegani 2022). Vaikka ihmisillä saattaa olla tällä hetkellä ylivoima tunteiden tulkitsemisessa, koneet valtaavat alaa käyttämällä omia vahvuuksiaan, kuten kykyä analysoida suuria määriä dataa (Somers 2019). Viime aikaiset tutkimukset osoittavat, että tekoäly voi tunnistaa tunteita hyödyntämällä kasvojen ilmeitä ja puheen tunnistusta (Dilmegani 2022). Koneet voivat analysoida kuvia ja poimia ihmisten kasvoilta mikroilmeitä, jotka saattavat tapahtua liian nopeasti, jotta ihminen pystyisi niitä havaitsemaan (Somers 2019).

Ihmiset ilmaisevat tunteitaan yllättävän samankaltaisesti eri kulttuureissa, jolloin tietokoneet voivat tunnistaa näitä vihjeitä luokitellakseen tunteita tarkasti. Tietokoneet jopa ylittivät ihmiset tutkimuksissa, jotka tehtiin vuosina 2003-2006. Tietokoneet saavuttivat silloin 70-80 prosentin tarkkuuden, kun ihmisen tarkkuus oli noin 60 prosenttia. (Dilmegani 2022.)

Kansainvälinen ihmisoikeusjärjestö Article 19 kirjoittaa empiirisestä tutkimuksesta, joka pyrki selvittämään onko oletamus pätevä, että ilmeet ovat seurausta tunteista. Lopputuloksena todettiin, että raportoidut meta-analyysit osoittivat, että kaikki kuusi tunnetta (onnellisuus/huvittuneisuus, yllättynyt, inho, suru, viha ja pelko) liittyivät keskimäärin vain heikosti kasvojen ilmeisiin, jotka ovat asetettu niiden yleismaailmallisiksi ilmeiksi. (ARTICLE 19 2021.) Toisaalta jotkin tutkimukset osoittavat, että kasvojen ilmeet visuaalisissa kanavissa ovat tärkein tunnetietoja sisältävä vihje, joka korreloi hyvin kehon ja äänen kanssa (Zeng, Pantic, Roisman, & Huang 2009).

Edellä mainittuun Article 19 -ihmisoikeusjärjestön viittaukseen yleismaailmallisiin ilmeisiin liittyen on hyvä huomioida, että ei ole olemassa yhtä totuutta siitä, mitkä ovat yleismaailmalliset ilmeet. Teorioita yleismaailmallisista ilmeistä ja perustunteista on useita (Zhao ym. 2021, Paul Ekman Group LLC s.a.b). Perehdymme tunteisiin tarkemmin luvussa 2.5.

2.2 Affektiivisen tietojenkäsittelyn tekniikat

Affektiivista tietojenkäsittelyä kutsutaan myös tunnetekoälyksi tai affektiiviseksi tekoälyksi. Se on teknologia, jonka avulla tunnistetaan, käsitellään ja simuloidaan ihmisen tunteita. (Dilmegani 2022, Gosset 2023, Somers 2019.) Affektiivisen tietojenkäsittelyn järjestelmät tunnistavat automaattisesti tunteita. (Dilmegani 2022).

Affektiivisen tietojenkäsittelyn nousun mahdollistavat korkearesoluutioisten kameroiden, nopean internetin ja koneoppimismahdollisuuksien, erityisesti syväoppimisen, yleistyminen. Kaikki nämä ovat kehittyneet huomattavasti 2010-luvun jälkeen. (Dilmegani 2022.)

Tehokkaimmat tietokonejärjestelmät käyttävät merkittävää harjoitusdataa koneoppimismallien kouluttamiseen. Ne tunnistavat tunteita puheesta tai videosta. Syväoppimisjärjestelmien suorituskyky kasvaa sitä mukaa mitä enemmän tietoa on. Siksi tällä alalla toimivat yritykset yrittävät laajentaa merkittävien tietojoukkojensa laajutta parantaakseen mallejaan. (Dilmegani 2022.)

Affektiiviset laskentaratkaisut, jotka hyödyntävät kuvia, käyttävät ilmeiden normalisoimiseksi seuraavia tekniikoita: kasvojen irroittaminen taustasta, kasvojen geometrian (esimerkiksi silmien, nenän ja suun sijaintien) arviointi, pään pyörittämisen sekä muiden pään liikkeiden vaikutuksen poistaminen. (Dilmegani 2022.)

2.3 Affektiivisen tietojenkäsittelyn haasteita

Automaattisesti tunteita tunnistavat ("Automatic Emotion Recognition", AER) järjestelmät voivat olla uskomattoman tehokkaita ja valtavan edistyneen edistäjiä. Ne voivat myös olla suuren vahingon aiheuttajia. Tunteiden tunnistaminen vaikuttaa ihmisiin ja sisältää siten eettisiä näkökohtia. Siksi on välttämätöntä, että yhteisö osallistuu luomustensa eettisiin seurauksiin. (Mohammad 2022.)

Amazonin koneoppimisen asiantuntijat paljastivat suuren ongelman: heidän uusi rekrytointijärjestelmänsä ei pitänyt naisista. Viisi hankkeeseen perehtynyttä henkilöä kertoi Reutersille, että tiimi oli rakentanut tietokoneohjelmia jo vuodesta 2014 lähtien tarkistaakseen työnhakijoiden ansioluetteloita tavoitteenaan huippulahjakkuuksien etsinnän koneistaminen. Yrityksen kokeellinen rekrytointityökalu käytti tekoälyä antamaan työnhakijoille arviot asteikolla yhdestä viiteen, samalla tavalla kuin ostajat arvioivat tuotteita. Vuoteen 2015 mennessä yritys huomasi, ettei uusi järjestelmä arvioinut ehdokkaita ohjelmistokehittäjän tai muihin teknisiin työtehtäviin sukupuolineutraalilla tavalla. Tämä johtui mallin kouluttamisesta tarkkailemalla kaavoja yritykselle toimitetuista ansioluetteloista yli kymmenen vuoden ajalta. Niistä suurin osa oli tullut miehiltä, joka kuvaa miesten dominoivaa asemaa teknisellä alalla. Itse asiassa järjestelmä opetti, että miehet ovat parempia. Se rankaisi ansioluetteloita, jotka sisälsivät sanan "naisten", kuten "naisten shakkikerhon kapteeni". Sukupuoleen liittyvät haasteet, eivät olleet ainoita ongelmia. Kun tekniikka palautti tuloksia lähes sattuman varaisesti, Amazon sulki projektin. (Dastin 2018.)

Vuonna 2016 Microsoftin Tay-niminen chatbot paljasti verkkokeskusteluiden vaarat. Se oppi kielen ihmisiltä Twitterissä, mutta se oppi myös arvot. Tay oli suunniteltu käymään keskusteluja ihmisten kanssa tveettien ja suorien viestien välityksellä samalla, kun se jäljitteli teini-ikäisen tytön tyyliä ja slangia. Tay oli kokeilu koneoppimisesta, luonnollisen kielen käsittelystä ja sosiaalisista verkostoista. Se suunniteltiin oppimaan lisää kielestä mahdollistaen keskustelut mistä tahansa aiheesta. Kun Tayn julkaisemisesta oli kulunut 16 tuntia, se oli twiitannut yli 95 000 kertaa, joista huolestuttava osuus oli loukkavia. Twitterin käyttäjät alkoivat osoittaa raivoaan, eikä Microsoftilla ollut muuta vaihtoehtoa kuin jäädyttää tili. (Schwartz 2019.)

Stanfordin yliopiston professori Michael Kosinki julkaisi vuonna 2017 tutkimuksen, joka viittaa siihen, että tekoäly voisi havaita valokuvien perusteella ihmisen seksuaalisen suuntautuneisuuden. Sen lisäksi käyttämällä kuvia tekoäly pystyisi tunnistamaan poliittiset näkemykset, arvioimaan onko heillä korkea älykkyysosamäärä, ovatko he alttiita rikolliselle käyttäytymiselle ja monia muita yksityisiä sekä henkilökohtaisia tietoja. Näillä tiedoilla voi olla suuria sosiaalisia seurauksia. (Levin 2017.)

Tekoäly soluttautuu nopeasti kaikille yhteiskunnan alueille. Päätökset, joita perinteisesti ihmiset ovat tehneet, suoritetaan nopeasti algoritmien avulla. Jopa tekoälypohjaiset teknologiat, joita ei ole suunniteltu suorittamaan vaativia tehtäviä, kuten määrittämään kauanko joku on vankilassa, voidaan käyttää välillisesti sellaisiin tehtäviin. Esimerkiksi kasvojentunnistusohjelmistoa itsessään ei pitäisi kouluttaa päättämään yksilön kohtaloa rikosoikeudellisessa järjestelmässä, mutta on hyvin todennäköistä, että sitä käytetään epäiltyjen tunnistamiseen. Virheellä kasvojen tunnistusalgoritmin lähtödatassa, jota käytetään muiden tehtävien syötteenä, voi olla vakavia seurauksia. Esimerkiksi joku voi joutua perusteettomasti syytteeseen rikoksesta. (Buolamwini & Gebru 2018.)

Tutkimukset osoittavat, että koneoppimisen algoritmit voivat syrjiä perustuen luokkiin, kuten rotu tai sukupuoli. Gender Shades -projektissa käytettiin ihotautilääkäreiden hyväksymää Fitzpatrick Skin Type -luokitusjärjestelmää sukupuolen ja ihotyypin merkitsemisessä kahteen kasvoanalyysin vertailuarvoon: IJB-A ja Adience. He havaitsivat, että kyseiset tietojoukot koostuvat ylivoimaisesti vaaleaihoisista subjekteista. He esittelivät uuden tietojoukon kasvoanalyysiin, joka on tasapainossa sukupuolen ja ihotyypin mukaan. He arvioivat kolme kaupallista sukupuolen luokittelujärjestelmää käyttäen omaa tietojoukkoaan ja osoittavat, että tummaihoiset naiset ovat eniten väärin luokiteltu ryhmä (virheprosentti jopa 34,7%). Vaaleaihoisilla miehillä korkein virheprosentti on 0,8%. He toteavat, että luokittelujärjestelmät vaativat kiireellistä huomiota, jos kaupalliset yritykset aikovat rakentaa aidosti reilun, avoimen ja vastuullisen kasvojenanalyysialgoritmin. (Buolamwini & Gebru 2018.)

2.4 Koulutusteknologia ja tunnetekoäly opetuslalla Kiinassa

Multimodaalista tunteiden tunnistamista tutkittaessa ilmeni, että Kiina mainittiin useissa lähteissä. Kiinassa multimodaaliseen tunteiden tunnistamiseen käytettävät teknologiat ovat jo käytössä opetus- ja koulutuslalla, joten on kiinnostavaa tutkia, kuinka siellä teknologioita käytetään ja kuinka niihin suhtaudutaan.

Koulutusteknologia ("Educational technology", edtech) on sekoitus lupauksia opiskelijoiden turvallisuudesta, oppimisen edistymisestä ja tehokkaista opetusmenetelmistä. On selvää, että näille työkaluille on olemassa kasvavat markkinat ympäri maailmaa, kuten myös Kiinassa. Käytetyt kasvojentunnistusteknologiat koulutusteknologioiden lisäksi tai upotettuina niihin, ovat herättäneet keskustelua opiskelijoiden yksityisyydestä ja näiden teknologioiden syrjimisestä mahdollisuudesta erilaisissa kansallisissa konteksteissa. Testausohjelmistojen myynti on kasvanut huomattavasti koronapandemian sulkiessa kouluja, jolloin opettajat ovat huolestuneet, kuinka nämä alustat valvovat oppilaita havaitakseen merkkejä huijaamisesta. Järjestelmien suunnittelusta johtuen tummemman ihon omaavat opiskelijat ovat kohdanneet teknisiä haasteita järjestelmien yrittäessä tunnistaa heidän kasvonsa. (ARTICLE 19 2021.)

Ennen pandemiaa kasvojen- ja ilmeidentunnistaminen kouluissa painottui kolmeen yleiseen käyttötarkoitukseen. Nämä olivat läsnäolotarkastusten suorittaminen kasvojen perusteella, oppilaiden huomion havaitseminen tai heidän kiinnostuksensa luentoon sekä turvallisuusuhkista ilmoittaminen. Koulutusteknologian painotukseen Kiinan paikallisella ja alueellisella tasolla koulutusalaalla on kaksi pääsyttä. Ensinnäkin lupaukset edistymisen seurannasta, kokeisiin valmistautumisesta ja oppimisen korkeammasta laadusta resonoivat vahvasti vanhempien kanssa, jotka ovat halukkaita panostamaan suuria määriä rahaa kilpailukykyiseen koulutusjärjestelmään. Toiseksi uuden koulutusteknologian kehittämiseen ja käyttöönnottoon kannustavat valtio- ja kansallisen tason tekoälytavoitteet ja koulutuspainotteiden kannustimet. (ARTICLE 19 2021.)

Oletettu syy-yhteys opetetun sisällön ja opiskelijan ulkoisen kiinnostuksen ilmaisun välillä on perusta argumentille yksilöllisestä oppimisesta, jota monet yritykset toistavat. Toisessa tutkimuksessa käytetään syväoppimisalgoritmeja opiskelijoiden reaaliaikaisten ilmeisen tunnistamiseen suurilla avoimilla verkkokursseilla (MOOC). Kirjoittajat uskovat tunteiden tunnistamisen hyödyttävän verkkokurssia erityisesti siksi, että opettajat eivät ole samassa sijainnissa opiskelijoidensa kanssa sekä opiskelijoiden ja opettajien välistä vuorovaikutusta on parannettava. Vaikka ainakin yksi tutkimus tunnustaa, että opiskelijoiden ilmeet rinnastettuna oppimisen sitoutumisen tasoon, on erittäin rajoittunut lähestymistapa. Merkittävin vastaus tähän puutteeseen on ollut uusien versioiden luominen, jotka oppivat menneistä tiedoista unimodaalisessa tunteiden tunnistamisessa multimodaalisilla vaihtoehdoilla. (ARTICLE 19 2021.)

2.5 Psykologiset mallit tunteiden mittaamiseen

Psykologiassa on kaksi edustavaa mallityyppiä tunteiden mittaamiseksi: kategorialliset tunnetilat ("categorical emotion states", CES) ja ulottuvuuden tunnetila ("dimensional emotion space", DES). Ulottuvuuden tunnetilan mallit voivat teoriassa mitata kaikkia tunteita eri koordinaattipisteinä jatkuvassa karteesisessä avaruudessa, mutta absoluuttiset jatkuvat arvot ovat yli käyttäjien ymmärryksen. Nämä kaksi määritelmää tunteista liittyvät toisiinsa ja voivat muuttua kategorialliset tunnetilasta ulottuvuuden tunnetilaan. (Zhao ym. 2021.)

Henkilöt, kuten Paul Ekman, omaksuivat ajatuksen yleismaailmallisuudesta 1960-luvulla, joka kehittyi myöhemmin ajatukseksi tunneteorian perusajatuksesta ("Basic Emotions Theory", BET). Siinä ajatuksena on, että jotkin kategoriset tunteet ovat perustavanlaatuisempia, yleismaailmallisempia ja tärkeämpiä kuin jotkin muut tunteet. Tietysti on erilaisia psykologeja, jotka määrittelevät eri tunteet perustunteiksi, josta syystä vastaavia teorioita on useita. (Mohammad heinäkuu 2021, 6-10 min.)

Kategorialliset tunnetilan mallit määrittelevät tunteet muutamaaan peruskategoriaan, kuten Ekmanin perustunteet, Mikelsin kahdeksan tunnetta (positiiviset: huvi, kunnioitus, tyytyväisyys, jännitys, negatiiviset: viha, inho, pelko ja suru), Plutchikin tunnepyörä (kahdeksan perus tunneluokkaa kolmella intensiteetillä) ja Parrotin puu hierarkisella ryhmittelyllä (ensisijainen, toissijainen ja kolmannen asteen luokat. Pyskologisten teorioiden kehitys motivoi kategorisia tunnetiloja olemaan yhä monipuolisempi ja hienojakoisempi. (Zhao ym. 2021.)

Vaikka asiantuntijoiden näkemykset tunteiden määrästä vaihtelevat, suurin osa tunnetutkijoista ovat samaa mieltä siitä, että ydintunteita on vähintään viisi. Ekmanin tutkimus osoittaa, että tähän mennessä vahvimmat todisteet ovat seitsemästä yleismaailmallisesta tunteiden ilmaisusta. Nämä tunteet ovat viha, halveksunta, inho, nautinto, pelko, suru ja yllättyneisyys. (Paul Ekman Group LLC s.a.b.)

Ekmanin ja Friesenin kehittämä ilmeiden koodausjärjestelmä ("Facial Action Coding System", FACS) on standardi emotionaalisten ilmeiden mittaamiseen. Tämä anatomiaan perustuva järjestelmä mahdollistaa tunteiden arvioimisen perustuen 46 tarkkailtavaan toimintayksikköön. FACS-koodaus mahdollistaa päätelmien tekemisen perustunteille, koska tutkimukset ovat osoittaneet, että tiettyjen toimintayksiköiden yhdistelmä liittyy tiettyihin tunteisiin. Tähän mennessä perustunneteoriat (eli FACS) on muokannut huomattavasti kaikkia ilmeiden mittaamenetelmiä. (Stöckli, Schulte-Mecklenbeck, Borer & Samson 2017.) Toisaalta tunteet ovat monimutkaisia ja hienovaraisia, joita ei voida heijastaa hyvin rajoitetuilla erillisillä luokilla (Zhao ym. 2021).

Lisäksi jotkin tunneteorian tärkeistä periaatteista ovat johtaneet huonoon valoon, kuten kasvojen ilmeiden ja tunteiden yhdistäminen, joka on lähes epäilyttävää tai parhaimmillaan kiistanalaista. Siitä huolimatta ajatus työskentelystä kategoristen tunteiden parissa on järkevää, koska käytämme niitä puhuessamme tunteista. (Mohammad heinäkuu 2021, 6-10 min.)

3 Multimodaalinen tunteiden tunnistaminen

Tässä luvussa selvitetään, minkälainen merkitys multimodaalisella tunteiden tunnistamisella on ihmisen ja koneen välisessä vuorovaikutuksessa. Sen jälkeen selvitetään minkälaisista signaaleista tunteita voidaan tunnistaa. Katsauksessa perehdytään tarkemmin signaaleihin, jotka liittyvät ihmisen fyysisiin ja psyykkisiin muutoksiin, kuten tunteiden tunnistamiseen ilmeistä, puheesta ja aivosähkökäyrästä. Sen jälkeen esitellään hyvin lyhyesti muutamia signaaleita, jotka ovat lähtöisin digitaalisen median lähteistä, kuten ääni, teksti ja kuva.

Ymmärtääkseen kokonaisuuden, mitä kaikkea multimodaaliseen tunteiden tunnistamiseen liittyy käytettäessä tunneataa hyödyntäviä sovelluksia, tulisi tietää peruseriaatteet datan keräämisen vaiheista ja datan luokittelusta. Tässä luvussa näitä asioita käydään lyhyesti läpi. Sen jälkeen selvitetään, millaisia haasteita ja hyötyjä multimodaaliseen tunteiden tunnistamiseen liittyy. Luku päättyy katsaukseen, millaisia käyttökohteita multimodaalisen tunteiden tunnistamisen sovelluksilla voisi olla opetus- ja koulutusallalla.

3.1 Multimodaalisen tunteiden tunnistamisen merkitys

Multimodaalinen tunteiden tunnistaminen pyrkii tunnistamaan ihmisen tunteita eri modaaliteettien välityksellä, jotka liittyvät tunteiden ilmaisuun. Näitä ovat esimerkiksi ääni, kuva ja teksti. Aihe on merkittävä monilla aloilla, joihin kuuluvat esimerkiksi ihmisen ja koneen välinen vuorovaikutus, tekoäly ja affektiivinen tietojenkäsittely. (Xiaoming ym. 2021).

Ihminen on tunteellinen olento. Useita modaaliteetteja on usein mukana, kun tunteita ilmaistaan. Tehdäänpä niin suoraan (esimerkiksi ilme ja puhe) tai epäsuorasti (kuten teksti ja kuva). Entistä tärkeämmäksi on kasvamassa tunneällyn mahdollistaminen koneille, kuten tunteiden tunnistaminen, tulkinta, käsittely ja simulointi. (Zhao ym. 2021.)

Jotkin tunnetilat motivoivat toimintaan, kun toiset rikastavat ihmisten kommunikoinnin merkitystä (Zeng, Pantic, Roisman & Huang 2009). Opiskelijoiden tunnekokemukset voivat vaikuttaa heidän kykyynsä oppia, heidän sitoutumiseensa opiskeluun ja jopa heidän uravalintoihinsa (Trezise 2017). Tunteiden tärkeys tekoälyssä on tunnistettu vuosikymmeniä sitten (Zhao ym. 2021). Minskyn mukaan (Minsky 1985) kysymys ei ole siitä voiko älykkäillä koneilla olla tunteita, vaan siitä voisivatko koneet olla älykkäitä ilman tunteita.

Tunneäly on laajasti määritelty älykkyydeksi, joka on kykyä tunnistaa omia ja muiden tunteita, muodostaa ja mukauttaa tunteita, sekä soveltaa tunnetietoa tavoitteen saavuttamiseksi ja ongelman ratkaisemiseksi. Tunteet ovat avainasemassa kommunikoinnissa, joka on edelleen laajasti laiminlyöty ihmisen ja koneen välisessä vuorovaikutuksessa. (Schuller & Schuller 2018.)

Emotionaalisesti älykkäät koneet voivat tarjota entistä harmonisemman ja henkilökohtaisemman palvelun ihmisille (Zhao ym. 2021). Varsinkin ihmiskeskeisillä käyttöliittymillä tulee olla kyky havainnoida vivahteet ja muutokset käyttäjän käytöksessä, erityisesti tunteellisessa käytöksessä, ja aloittaa vuorovaikutus perustuen tähän tietoon sen sijaan, että se vain vastaisi käyttäjän kommentoihin (Zeng ym. 2009).

3.2 Tunteiden tunnistaminen ihmisen fyysisistä ja psykologisia muutoksista

Ensimmäinen askel älykkäiden koneiden ihmismäisten tunteiden ilmaisemiseen on tunnistaa ja ymmärtää ihmisten tunteita. Fyysisiä ja psykologisia muutoksia ihmisissä voidaan suoraan tarkkailla, kerätä ja tallentaa. Nämä signaalit voivat olla joko helposti tukahdutettavissa tai piilotettavissa, tai ne voivat olla vaikeasti tai epäkäytännöllisesti kaapattavissa. Fyysisiä ja psykologisia muutoksia ovat esimerkiksi silmien liike, puhe, toiminta ja fysiologiset signaalit, kuten aivosähkökäyrä. (Zhao ym. 2021.) Ilmeet ja puhe ovat yleisimmin käytetyt signaalit tunteiden tulkitsemiseen (Pan, Fang, Zhang, Chen, Zhang & Wang 2023).

3.2.1 Tunteiden tunnistaminen kasvojen ilmeistä

Multimodaalisesta tunteiden tunnistamisesta ja siihen yhdistetyistä teknologioista puhuttaessa, saatetaan usein puhua kasvojentunnistusteknologiasta, vaikka tarkoitettaisiin teknisesti ilmeidentunnistusjärjestelmää. Nämä liittyvät läheisesti toisiinsa, sillä ilmeidentunnistusjärjestelmät hyödyntävät kasvojentunnistusteknologiaa. Näistä syistä pyrin käsittelemään tässä luvussa molempia tekniikoita.

Kasvojentunnistus ("Face Recognition", "Facial recognition system") tarkoittaa määritelmän mukaan teknologiaa, joka pystyy tunnistamaan tai varmentamaan henkilön kasvot kuvasta, videosta tai mistä tahansa audiovisuaalisesta elementistä (Electronic Identification 2022, Kaspersky s.a.). Toisin sanoen kasvojentunnistus on keino tunnistaa tai vahvistaa henkilöllisyyttä kasvojen avulla (Kaspersky s.a.).

Yleisesti ottaen tätä tunnistusta käytetään, jotta päästään sovellukseen, järjestelmään tai palveluun. Se on biometrinen tunnistusmenetelmä, joka hyödyntää kasvojen ja pään mittauksia, varmentaakseen henkilön henkilöllisyyden kasvojen biometrinen kuvioiden ja tietojen avulla. (Electronic Identification 2022.) Keskeisiä tekijöitä tunnistettavien kasvonpiirteiden löytämiseksi ovat silmien välinen etäisyys, silmäkuoppien syvyys, etäisyys otsasta leukaan sekä poskipäiden, huulien, korvien ja leuan muodot (Kaspersky s.a.).

Kasvojentunnistuksen hyötyjä ovat esimerkiksi tietojen nopeampi käsittely, jolloin henkilöllisyys voidaan vahvistaa nopeasti ja tehokkaasti. Haittoja ovat muun muassa kasvojentunnistustiedoissa mahdollisesti olevat virheet, joiden seurauksena ihmisiä saatetaan syyttää teoista, joita he eivät ole tehneet. Kiistanalaisimpia ovat kuitenkin etiikkaa ja yksityisyyttä koskevat kysymykset. (Kaspersky s.a.).

Ruotsin tietosuojaviranomainen määräsi vuonna 2019 kunnalle 200 000 kruunun (noin 20 000 euron) sakon, koska eräs pohjoisruotsalainen koulu oli käyttänyt kasvojentunnistusteknologiaa seuratakseen oppilaiden läsnäoloja. Koulu suoritti järjestelmän testauksen yhdessä koululuokassa rajoitetun ajan. Ruotsin tietosuojaviranomainen totesi testauksen rikkoneen useita yleisen tietosuoja-asetuksen artikloja. Koulun katsottiin käsitelleen arkaluontoisia biometrisiä tietoja vastoin lakia, eikä se ollut tehnyt riittävää vaikutustenarviointia. Koulu on perustanut tietojen käsittelyn suostumukseen, mutta Ruotsin tietosuojaviranomaisen mukaan suostumus ei ollut pätevä oikeusperusta, sillä rekisteröidyn ja rekisterinpitäjän välillä oli selkeä epätasapaino. (Europeiska datatillsynsmannen 2019.) Kiinassa sen sijaan joissakin oppilaitoksissa on käytössä kasvojentunnistusjärjestelmiä, jotka tunnistavat ja kirjaavat automaattisesti opiskelijoiden läsnäolotiedot (Zhang 2019).

Kasvokuvia ja videoita voidaan ottaa missä vain, josta on kiittäminen kameroiden pientä kokoa. Valvontakamerat julkisilla paikoilla tai kauppoissa eivät ole ainoita kasvokuvia nappaavia kameroita, sillä henkilöiden omat mobiililaitteet voivat tallentaa ilmeitä käytön aikana. Näissä tapauksissa läpinäkyvyyteen liittyvät kysymykset nousevat huolenaiheeksi henkilökohtaisen datan keräämiseen ja käsittelyyn liittyen. Tilanteissa, joissa ilmeitä on tallennettu etänä käsiteltäviksi, ei ole aina selvää mikä järjestelmä tai sovellus dataa käsittelee, mihin tarkoitukseen ja ketkä siitä vastaavat. Tämän tuloksena henkilöillä ei ole mahdollisuutta antaa suostumustaan tai valvoa henkilötietojensa käsittelyä, johon liittyy esimerkiksi tietojen jakaminen kolmansien osapuolten kanssa. (Vemou & Horvath 2021.)

Ilmeidentunnistus ("Facial Emotion Recognition" tai "Facial Expression Recognition", FER) on teknologia, joka analysoi ilmeitä erilaisista lähteistä, kuten kuvista ja videoista, paljastaakseen informaatiota henkilön tunnetilasta (Vemou & Horvath 2021). Ilmeidentunnistus on tärkeä aihe tietokonenäön ja tekoälyn alalla, koska sillä on merkittävä akateeminen ja kaupallinen potentiaali (Ko 2018).

Kasvojen ilme on yhden tai useamman kasvojen alueen tai yksikön eristetty liike tai tällaisten liikkeiden yhdistelmä. Yleisesti ollaan samaa mieltä siitä, että ilmeet voivat sisältää informatiivisia tunnetilaan liittyviä vihjeitä. Ne ovat tunnustettu yhdeksi luonnollisemmaksi ja voimakkaammaksi signaaliksi tunnetilojen ja aikeiden välittämisessä. Ilme on myös sanattoman viestinnän muoto, jolla välitetään sosiaalista tietoa ihmisten välillä. (Zhao ym. 2021.)

Ilmeet ovat merkittäviä tekijöitä ihmisten välisessä vuorovaikutuksessa, jotka auttavat ymmärtämään toistemme aikeita. Yleisesti ottaen ihmiset osoittavat tunteitaan toisille ihmisille käyttäen ilmeitä ja äänen painoa. (Ko 2018). Muihin lajeihin verrattuna ihmiset ovat kehittäneet erittäin hienostuneen kommunikaatiojärjestelmän sosiaaliseen kanssakäymiseen (Kaulard, Cunningham, Bühlhoff & Wallraven 2012). Useiden ei-sanallisten tapojen joukosta ilmeet ovat yksi tärkeimmistä kanavista välittää tunteita ihmisten välisessä vuorovaikutuksessa (Ko 2018, Kaulard ym. 2012). Ei ole yllättävää, että ilmeiden tunnistamiseen liittyvä tutkimus, on saanut paljon huomiota viime vuosikymmeninä, sillä niitä käytetään havainto- ja kognitiivisten tieteiden lisäksi myös affektiivisessä tietojenkäsittelyssä ja tietokoneaniomaatioissa (Kaulard ym. 2012).

Ilmeidentunnistusanalyysi sisältää kolme vaihetta, jotka ovat kasvojetunnistus, ilmeidentunnistus ja tunnetilan luokittelu. Ilmeidentunnistus perustuu analyysiin kasvojen solmupisteistä, kuten nenänpää ja kulmakarvat. (Vemou & Horvath 2021.) Ilmeet voidaan luokitella perustunteisiin tai yhdistettyihin tunteisiin (Du, Tao & Martinez 2014).

Ilmeisiin liittyvät tutkimukset ovat keskittyneet tutkimaan perustunteita, kuten viha, inho, pelko, ilo, surullisuus ja yllättynyt. Ilmeitä on olemassa kuitenkin enemmän. Tunteita voidaan luokitella myös yhdistämällä kaksi perustunnetta, kuten iloisesti yllättynyt tai vihaisesti yllättynyt, jolloin voidaan puhua yhdistetyistä tunteista. (Du ym. 2014.) Joissakin tapauksissa ilmeet voivat linkittyä psyykkiseen tai henkiseen tilaan, kuten väsymys tai kyllästyneisyys (Vemou & Horvath 2021).

Ilmeidentunnistukselle on useita potentiaalisia käyttökohteita sovelluksissa. Opetusalalla voidaan seurata opiskelijoiden tarkkaavaisuutta, havaita tunnereaktiot opetusohjelmaan ja muokata opetuspolkua, suunnitella tunteet huomioiva tutorointijärjestelmä sekä havaita verkko-opetukseen sitoutuminen. (Vemou & Horvath 2021.)

Tunneilmaisuiden muuttaminen datalähteeksi tunteiden tunnistamiseksi koskettaa selvästi ihmisten yksityisempiä tietoja. Tunteiden tunnistaminen on häiritsevä teknologia, joka nostaa esiin tärkeitä kysymyksiä, jotka liittyvät tarpeellisuuteen ja suhteellisuuteen. On arvioitava huolellisesti, tarvitaanko ilmeidentunnistamista tavoitteen saavuttamiseksi vai olisiko muita vaihtoehtoja olemassa. Suhteellisuuteen vaikuttavat monet asiat, kuten kerätyn datan tyyppi, päätelmien tyyppi, datan säilytysaika ja datan mahdollinen jatkokäsittely. (Vemou & Horvath 2021.)

Ilmeidentunnistamiseen perustuvat tunneanalyysit eivät välttämättä ole kovinkaan tarkkoja, sillä ilmeet voivat vaihdella hieman yksilöiden välillä, ne voivat sekoittua muihin samaan aikaan koettuihin tunnetiloihin tai ne eivät välttämättä ilmaise tunnetta ollenkaan. Toisaalta kaikkia tunteita ei voida päätellä ilmeiden perusteella, jolloin päätelmät voivat johtaa väärin lopputuloksiin. (Vemou & Horvath 2021.)

Vaikka tunteiden tulkitseminen olisi tarkka, lopputuloksen käyttäminen voi johtaa väärin päätelmiin henkilöstä. Ilmeiden tunnistaminen ei selitä, mikä tunteen aiheutti. Syy voi olla jokin äskettäin tai kauempana menneisyydessä tapahtunut tapahtuma. Jokatapauksessa lopputuloksia kohdellaan usein faktoina ja ne syötetään prosesseihin, jolloin ne vaikuttavat henkilön elämään. Sen sijaan, että tuloksia käytettäisiin arvioinnin laukaisemiseksi, jotta voitaisiin selvittää lisää henkilöiden tilanteesta tietyssä kontekstissa. (Vemou & Horvath 2021.)

Ilmeidentunnistamisen tarkkuus voi olla tärkeässä roolissa, jotta ihon väri tai etninen tausta eivät johda syrjintään. Sosiaaliset normit ja kulttuuriset erot vaikuttavat tunneilmaisun tasoon joidenkin tunteiden kohdalla. Esimerkiksi tutkimukset, joissa testattiin ilmeidentunnistamisen algoritmeja, paljastivat niiden osoittavan enemmän negatiivisia tunteita afrikkalaisten ihmisten kasvoille kuin muiden. Oikean datajoukon valitseminen on ratkaisevan tärkeää syrjinnän estämiseksi. Jos harjoitusdata on liian yksipuolinen, teknologia saattaa syrjiä vähemmän edustettuina olevia väestöryhmiä. (Vemou & Horvath 2021.)

3.2.2 Tunteiden tunnistaminen puheesta

Puhe on modaaliteetti, jonka välityksellä puhujat voivat ilmaista aikeitaan kuten kysyä tai julistaa käyttämällä erilaisia intonaatioita, äänenvoimakkuutta ja tempoa. Tunteet voidaan paljastaa erityisesti, kun ihmiset puhuvat toisilleen tai mutisevat itsekseen. (Zhao ym. 2021.) Puhe on merkittävä kommunikoinnin modaaliteetti ihmisten välisessä vuorovaikutuksessa (Zeng ym. 2009).

Äänen- ja puheentunnistus ("Voice Recognition" ja "Speaker Recognition") ovat koneen tai järjestelmän kyky vastaanottaa ja tulkita sanelua, tai ymmärtää ja suorittaa puhekomentoja. Äänentunnistusjärjestelmien avulla ihmiset voivat puhumalla olla vuorovaikutuksessa teknologian kanssa, joka mahdollistaa "handsfree"-pyynnöt, muistutukset ja muut yksinkertaiset tehtävät. (Gillis, A. s.a.) Tunteiden tunnistaminen on puheen tunnistuksen osa, joka kasvattaa suosiotaan ja jonka tarve kasvaa valtavasti (Burnwal 2020).

Yksi tyypillinen automaattisen puheentunnistuksen ("Automatic Speech Analysis", ASA) tehtävä on tunnistaa puheesta erilaisia tiloja ja ominaisuuksia. Ei-kielellisiä lyhytaikaisia tiloja ovat esimerkiksi tunteet. Keskipitkän aikavälin tiloja ovat esimerkiksi asenne. Pitkäaikaisia ominaisuuksia ovat persoonallisuus, ikä ja sukupuoli. Puhedatan variaatioihin vaikuttavat vahvasti monet tekijät, kuten puhujan ikä, sukupuoli ja kulttuurinen tausta. (Zhang, Cummins & Schuller 2017).

Tunteiden tunnistaminen puheesta ("Speech Emotion Recognition", SER) on tehtävä, jonka tarkoituksena on tunnistaa tunteita puhesignaaleista. Se on hyvin tärkeää ihmisen ja tietokoneen välisen vuorovaikutuksen edistämiseksi. Yksilön tunteiden ymmärtäminen kommunikoinnin aikana on tärkeää keskustelun ymmärtämiseksi ja siihen kunnolla vastaamiseksi. Huolimatta kaikkein suurimmista koneoppimisen ongelmista, tunteiden tunnistaminen puheesta on alkanut saamaan etua syväoppimisen tarjoamista työkaluista. (Abbaschian, Sierra-Sosa & Elmaghraby 2021.)

Tunteiden tunnistaminen puheesta on hyödyllinen verkossa tapahtuvissa vuorovaikutteisissa opetusohjelmissa ja kurseissa. Opiskelijan tunnetilan ymmärtäminen auttaa konetta päättämään, kuinka se esittää loppuosan kurssin sisällöstä. (Abbaschian ym. 2021.)

3.2.3 Tunteiden tunnistaminen aivosähkökäyrästä

Psykologista signaalia edustava aivosähkökäyrä ("Electroencephalography", EEG) on tärkeä menetelmä aivojen sähköisen ja emotionaalisen toiminnan tallentamiseen (Zhao ym. 2021). Aivosähkökäyrätutkimus mittaa hermotoiminnan aiheuttamia pieniä muutoksia aivojen sähkökentässä (Subramanian, Wache, Abadi, Vieriu, Winkler & Sebe 2018).

Keskushermoston aivosähkökäyräsignaalit voivat heijastaa tarkemmin ja objektiivisemmin muutoksia ihmisen tunteissa verrattuna muihin signaaleihin (Pan ym. 2023). Toisaalta muihin signaaleihin verrattuna aivosähkökäyräsignaalien kokoelma on tyypillisesti vaikeampaa ja luonnottomampaa riippumatta siitä ovatko elektrodit sijoitettu ei-invaasisesti päänäköä pitkin vai invaasisesti elektrokortikografiaa hyödyntäen (Zhao ym. 2021).

Aivosähkökäyriin perustuvalla tunteiden tunnistamisella ovat laajat sovellusmahdollisuudet. Koulutuksessa käytettävissä virtuaalitodellisuuden ("Virtual Reality", VR) sovelluksissa on todettu tunnetiloilla olevan vaikutusta muistiin. Positiivisella mielialalla on hyödyllinen vaikutus tilaoppimiseen. Siitä syystä tunne tulisi tunnistaa virtuaalitodellisuudessa oppimisen aikana. (Li, Zhang, Tiwari, Song, Hu, Yang, Zhao, Kumar & Marttinen 2022).

Vaikka alalla on monia saavutuksia, niin useita ongelmia ja haasteita on vielä tutkittavaksi sekä ratkaistavaksi. Vielä on tilaa esimerkiksi tutkimukselle tehokkaista aivosähkökäyrien lähestymistavoista. Toiseksi tunteiden tunnistamisen laskennallisten menetelmien kehittäminen vaatii syvempää ymmärtämistä tunneprosessista, vaikka lukuisia huippuluokan tekoälymalleja onkin tutkittu. Kolmantena aivosähkökäyräsignaali korkealla näytteenottotaajuudella on erittäin pitkä, jonka takia mallin laskentataakka kasvaa reilusti. (Li ym. 2022.)

3.3 Tunteiden tunnistaminen digitaalisen median lähteistä

Mobiililaitteiden ja sosiaalisten verkostojen suosio mahdollistaa kokemusten jakamisen ja mielipiteiden ilmaisun verkossa käyttäen tekstiä, kuvia, ääntä ja videoita. Näitä tietotyyppejä käytetään tiedon tallentamiseen ja siirtämiseen eri digitaalisten laitteiden välillä. Vaikka yhden tietyn modaalin tehokuutta luotettavana kanavana tunteiden ilmaisussa ei voida taata, yhdessä useat modaalit parantaisivat merkittävästä luotettavuutta ja kestävyyttä. (Zhao ym. 2021.)

Teksti on muoto, jolla tallennetaan luonnollista kieltä, joka voi epäsuorasti sisältää informatiivisia tunteita. Tekstissä on eri tasoisia kielellisiä komponentteja mukaan lukien sana, lause, kappale ja artikkeli. Nämä ovat hyvin tutkittuja: monet valmiit algoritmit ovat kehitetty jakamaan teksti pieniksi paloiksi. Sitten jokaisen kielipalasan tunnepitoinen ominaisuus tunnistetaan julkisesti saatavilla olevan sanakirjan avulla ja tekstin herättämä tunne voidaan päätellä. (Zhao ym. 2021.)

Kuva on värillisten pisteiden jakautumista avaruudessa. On hyvin tunnettua, että ”kuva on arvokkaampi kuin tuhat sanaa”. Psykologiassa on osoitettu, että kuvat voivat herättää tunteita. Verkossa jaettujen kuvien määrä on kasvanut räjähdysmäisesti. Niillä on kyky kuvailla voimakkaasti tilanteita. Nämä tekijät mahdollistavat, että laajan tutkimustyön kanssa kuvista voi tulla ratkaisevia epäsuoria ärsykeitä. (Zhao ym. 2021.)

Video luonnollisesti sisältää useita modaaliteetteja samaan aikaan, kuten visuaalisuuden ja ääni- sekä tekstitiedon. Se tarkoittaa, että ajalliset, tilalliset ja monikanavaiset esitykset, voidaan oppia ja käyttää tunteiden tunnistamiseen videoilta. (Zhao ym. 2021.)

3.4 Tunnetiedon keräämisen vaiheet ja luokat

Multimodaalisen tunteiden tunnistamisen tietojoukon rakentamiseen sisältyy yleensä kaksi vaihetta: datan kerääminen ja tunteiden merkitseminen. Kerätty data voidaan jakaa karkeasti kahteen luokkaan, jotka ovat olemassa olevan datan valinta ja uuden nauhoittaminen tiettyssä ympäristössä. Osa datasta valitaan elokuvista, videoista ja TV-ohjelmista sosiaalisessa mediassa, kuten Youtubesta ja Weibosta. Toisaalta osa datasta on vasta nauhoitettu erilaisilla antureilla tietynlaiseksi suunnitelluissa ympäristöissä. (Zhao ym. 2021.)

Nykyiselle tunneanalyysin alalle on mahdollisuus ja haaste, kuinka analysoida tunteita multimodaalisesta datasta. Seuraavaksi esitellään yleisesti käytössä olevat tietojoukot multimodaaliseen tunneanalyysiin. (Datatang 2022.) Datatang esittelee yhteensä kymmenen tietojoukkoa, mutta niistä opinnäytetyöhön valikoitui esimerkin omaisesti kaksi ensimmäistä, joista selviää minkälaista tunnetietoa niillä saadaan kerättyä. Nämä eivät suoranaisesti liity opetus- ja koulutusalaan, mutta antavat viitteitä siitä millaisia tietojoukkoja voitaisiin tulevaisuudessa kehittää.

”Yelp dataset” tulee osoitteesta <http://Yelp.com>, joka on arvosteluihin keskittyvä internet-sivusto. Se kerää Yelp arvosteluja ravintoloista ja ruoista viidessä kaupungissa: Boston, Chicago, Los Angeles, New York ja San Francisco. Siellä on yhteensä 44 305 kommenttia ja 244 569 kuvaa. Tietojoukon tunnustamiseksi on jokaiselle kommentille mielialatrendin määrittely viisi portaisella asteikolla (1-5). (Datatang 2022.)

”Tumblr Dataset” on Tumblr-mikroblogipalvelulta kerätty multimodaalinen tunnetietojoukko. Käyttäjien julkaisema multimediasisältö sisältää yleensä kuvia, tekstiä ja tunnisteita (”tags”). Tietojoukon tarkoituksena on etsiä twiittejä vastaavilla tunnetunnisteilla valitun viidentoista tunteen mukaan ja valita vain ne osat, joissa on sekä tekstiä että kuvia. Sitten suoritetaan tietojoukon käsittely sellaisten tekstien poistamiseksi, jotka alun perin sisältävät vastaavia tunnesanoja ja jotka eivät ensisijaisesti ole englanninkielisiä. Koko tietojoukossa on yhteensä 256 897 multimodaalista twiittiä, joiden tunteet ovat merkitty viideksitoista tunteeksi. (Datatang 2022.)

3.5 Multimodaalisen tunteiden tunnistamisen haasteita

Multimodaalisella tunteiden tunnistamisella on useita etuja verrattuna unimodaaliseen tunteiden tunnistamiseen, mutta sillä on myös enemmän haasteita (Zhao ym. 2021). Opinnäytetyössän pureudutaan viiteen oleellisimpaan multimodaaliseen tunteiden tunnistamiseen liittyvään haasteeseen.

Affektiivinen kuilu on yksi multimodaalisen tunteiden tunnistamisen suurimmista haasteista, joka mittaa epä johdonmukaisuuksia poimittujen ominaisuuksien ja havaittujen korkean tason tunteiden välillä. Affektiivinen kuilu on vielä haastavampi kuin semanttinen kuilu objektiivisen multimedian analyysissä. Vaikka semanttinen kuilu olisi ylitetty, affektiivinen kuilu saattaa olla edelleen olemassa. Esimerkiksi kukkiva ruusu ja kuihtunut ruusu sisältävät molemmat ruusun, mutta saattavat herättää erilaisia tunteita. Samalle lauseelle eri ääni-intonaatiot voivat vastata täysin eri tunteita. Erottelevien korkean tason ominaisuuksien purkaminen, ja varsinkin tunteisiin liittyvien, voi auttaa kuroma umpeen affektiivisen kuilun. Suurin haaste on siinä, kuinka arvioida liittyvätkö puretut ominaisuudet tunteisiin. (Zhao ym. 2021.)

Subjektivisuushaaste: eri katsojilla saattaa olla aivan erilaiset tunnereaktiot samaan kuvaan, jonka taustalla ovat monet henkilökohtaiset ja tilannesidonnaiset tekijät, kuten kulttuuristausta, persoonallisuus ja sosiaalinen konteksti (Zhao, Yao, Yang, Jia, Ding, Chua, Schuller & Keutzer 2022). Toisaalta vaikka tunne olisi sama, henkilöiden fyysiset ja psyykkiset muutokset voivat olla melko erilaisia (Zhao ym. 2021). Yksi ratkaisu on ennustaa jokaiselle katsojalle henkilökohtaiset tunnehavainnot, sillä jokaiselle kuvalle voidaan määrittää useita tunnisteita (Zhao ym. 2022).

Datan epätäydellisyys: Datan keräämisessä on monia väistämättömiä tekijöitä, kuten anturilaitteen vika. Tietyn modaliteetin tieto saattaa olla vioittunut, joka johtaa puuttuvaan tai epätäydelliseen dataan. Tietojen epätäydellisyys on yleinen ilmiö todellisissa multimodaalisen tunteiden tunnistamisen tehtävissä. Tehokkaiden fuusiomenetelmien suunnittelu, jotka voivat käsitellä datan epätäydellisyyttä, on laajalti käytetty strategia. (Zhao ym. 2021.)

Tunnisteiden kohina tai puuttuminen: Olemassa olevat multimodaalisen tunteiden tunnistamisen menetit, varsinkin syväoppimiseen pohjautuvat, vaativat paljon luokiteltua harjoitusdataa. Tunteiden luokittelu on kallista ja aikaa vievää, mutta myös epäjohdonmukaista. Se johtaa suureen määrään dataa, jossa luokiteltua tunne-dataa on vain vähän tai ei ollenkaan. (Zhao ym. 2021.) Kysymykset tämän kaltaisen datan käsittelyyn ovat tutkimuksen arvoisia (Zhao ym. 2022).

Modaliteettien välinen epäjohdonmukaisuus: Eri modaliteetit samassa näytteessä voivat olla keskenään ristiriidassa ja ilmaista siten eri tunteita. Esimerkiksi kasvojen ilmeet ja puhe voidaan helposti peittää tunnistamisen estämiseksi, mutta EEG-signaalit ovat keskushermoston ohjaamia voiden ilmaista tiedostomattomia kehon muutoksia. Kun ihmiset postaavat tweettejä sosiaaliseen mediaan, on hyvin yleistä, että kuvat eivät semanttisesti vastaa tekstiä. Silloin tehokkaan multimodaalisen tunteiden tunnistamisen metodin odotetaan arvioivan automaattisesti mitkä modaliteetit ovat luotettavampia, esimerkiksi määrittelemällä painoarvon jokaiselle modaliteetille. (Zhao ym. 2021.)

3.6 Multimodaalisen tunteiden tunnistamisen hyödyt

Viimeisen vuosikymmenen aikana suurin osa tutkimuksista on keskittynyt unimodaaliseen tunteiden tunnistamiseen ("Unimodal Emotion Recognition", UER), joka käyttää vain yhtä modaliteettia (Pan ym. 2023). Viime vuosina tutkimus on kuitenkin siirtynyt enemmän multimodaaliseen tunteiden tunnistamiseen johtuen joistakin ongelmista, joita unimodaaliseen tunteiden tunnistamiseen liittyy: unimodaalinen data saattaa jostain syystä puuttua, ulkoisesta käytöksestä saatu informaatio ei välttämättä kuvasta todellista tunnetta ja tunnistuskykyä voidaan parantaa yhdistämällä useampien modaliteettien välityksellä kerättyä dataa (Li ym. 2022).

Unimodaaliseen tunteiden tunnistamiseen verrattuna multimodaalisella tunteiden tunnistamisella on useita etuja. Ensinnäkin datan täydentävyys, sillä eri menetelmien vihjeet voivat lisätä tai täydentää toisiaan. Ystävältä saatu viesti, jossa lukee ”mikä hieno sää!”, voidaan tulkita positiiviseksi tunteeksi. Jos viestiin liittyy kuitenkin kuva myrskystä, voidaan tekstin päätellä olevan sarkasmia ja että tarkoitus on ilmaista negatiivista tunnetta. (Zhao ym. 2021.) Toisaalta on hyvä huomioida, että henkilön ulkoisesta käytöksestä päätelty tieto, ei välttämättä vastaa hänen todellista tunnetta, sillä tunteet voidaan piilottaa sosiaalisissa tilanteissa (Li ym. 2022).

Toinen etu on mallin kestävyys. Datakokoelmassa monien tavallisesti esiintyvien tekijöiden johdosta, kuten anturilaitteen vika, jotkin datamuodot eivät välttämättä ole saatavilla. Esimerkiksi puhetta, ilmeitä ja eleitä sisältävässä CALLAS-tietojoukossa (”CALLAS dataset”) eletiedot puuttuvat joiltakin hetkellisesti liikkumattomilta käyttäjiltä. (Zhao ym. 2021.) Monitoroituun signaaliin voivat vaikuttaa erilaiset ulkoiset tekijät, kuten melu (Li ym. 2022). Häiriötekeijöistä huolimatta opittu multimodaalisen tunteiden tunnistamisen malli voi silti toimia muiden saatavilla olevien modaliteettien avulla (Zhao ym. 2021).

Kolmantena etuna on suorituskyvyn ylivoima, sillä eri menetelmien avulla kerättyjen täydentävien tietojen huomioon ottaminen voi johtaa parempaan tunnistustehokkuuteen. Meta-analyysi osoittaa, että verrattuna parhaaseen yhteen modaliteettiin, multimodaalinen tunteiden tunnistaminen saavuttaa keskimäärin 9,83% suorituskyvyn parannuksen. (Zhao ym. 2021.) Tutkijoiden tulisi ottaa mukaan tunteiden tunnistamiseen mahdollisimman monta modaliteettia, mukaan lukien aivosähkökäyrä, ilmeet, eleet ja niin edelleen (Li ym. 2022).

3.7 Multimodaalisen tunteiden tunnistamisen käyttökohteita opetus- ja koulutuslalla

Tunteiden tunnistaminen useista signaaleista on erittäin tärkeää monissa reaali maailman sovelluksissa. Yleisesti ottaen tunteet ovat tärkein osa olemassa olomme laatua ja tarkoitusta, joka tekee elämästä elämisen arvoista. Digitaalisen datan emotionaalinen vaikutus perustuu siihen, että se voi parantaa olemassa olevien tekniikoiden käyttäjäkokemusta ja siten vahvistaa tiedonsiirtoa ihmisten ja tietokoneiden välillä. (Zhao ym. 2021.)

Affektiivisen tietojenkäsittelyn alalta syntyneen tunteiden tunnistamisen ennustetaan olevan jopa 65 miljardin dollarin ala vuoteen 2024 mennessä (ARTICLE 19 2021). Maailmalla affektiivista tietojenkäsittelyä hyödynnetään jo monilla aloilla, kuten markkinoinnissa, puhelinkeskuksissa ja autoteollisuudessa (Somers 2019). Koulutuksessa affektiivisen laskentateknologian sovelluksia voidaan hyödyntää esimerkiksi tehokkuuden mittaamisessa ja **autististen lasten tukemisessa** (Dilmegani 2023).

Tehokkuuden mittaamisessa sensoreita, kuten videokameroita ja mikrofoneja, voidaan hyödyntää oppilaiden tunnetilojen mittaamiseen opetustunnin aikana. Tekoäly voi arvioida, kuinka tyytyväisiä tai turhautuneita, oppilaat ovat tunnin aikana tehtävien ollessa joko liian haastavia tai liian yksinkertaisia. Tuloksena opettajat voivat räätälöidä luokan kuormitusta vastaavalla tavalla. Samaa lähestymistapaa voidaan hyödyntää myös verkko-opetusjärjestelmien prototyyppien testauksessa. (Dilmegani 2023.)

Entropik Tech on maailman johtava tunnetekoälyalusta. He tarjoavat kahta koulutusteknologiaa hyödyntävää ratkaisua. Ensimmäinen on **opiskelijoiden tunneanalyysi**, joka tulkitsee opiskelijoiden huomion huiput ja laskut, mittaa heidän kognitiivisen sitoutumisen sekä tallentaa toiminnan opiskelumoduulissa kokonaisuudessaan. Lisäksi se mittaa opettajan opetustaidot. Toinen ratkaisu on **tuutoreiden suorituskyvyn mittaus**, joka mittaa verkkokurssin luokahuoneen sitoutumista, vertaa tuutoreiden sitoutumista ja kognitiivista suorituskykyä sekä tunnistaa kitkaa aiheuttavat kohdat tai vähäiselle huomiolle jääneet alueet. Entropikin tunnetekoälyratkaisu perustuu teknologioihin, kuten silmänseuranta ja ilmeidentunnistus. (Entropik Tech s.a.)

Oppimiskokemuksessa kognitiiviset ja emotionaaliset ulottuvuudet liittyvät usein toisiinsa. Tunteet ja sitoutuminen voivat johtaa parempiin lopputuloksiin kognitiivisella tasolla. Aikaisemmat tutkimukset ovat osoittaneet, että **virtuaalitodellisuus** tarjoaa läsnäolon ja syventymisen, joka voi johtaa emotionaalisesti kiinnostaviin oppimistilanteisiin. Virtuaalitodellisuudesta onkin tullut entistä suosittuampi opetusväline. Se tarjoaa tilanteiden kokemiseen alustan, jolla ei ole riskiä itselle tai muille. (Vesisenaho, Juntunen, Häkkinen & Pöysä-Tarhonen 2019.)

4 Tekoälyn etiikka koulutuksen näkökulmasta

Tässä luvussa selvitetään, mitä tekoälyn etiikalla takoitetaan ja millaisia eettisiä vaikutuksia tekoälyjärjestelmillä voi olla. Luvussa huomioidaan yleinen tietosuoja-asetus ja selvitetään miten se tulee huomioida opetus- ja koulutuslalla henkilötietoja käsiteltäessä. Lisäksi käsitellään Euroopan Unionin ehdotusta tekoälysäädöksestä ja tarkastellaan millaisia asioita sen myötä tulisi huomioida tekoälyjärjestelmien käyttöön liittyen.

Luvussa pohditaan tekoälyjärjestelmien käyttöä opetus- ja koulutuslalla eettisistä näkökulmista. Eettisessä viitekehyksessä pyritään huomioimaan näkökulmat monipuolisesti niin oppijoiden, opettajien kuin koulujenkin näkökulmista. Kappaleessa pohditaan myös, millaisia asioita tulisi huomioida, jotta tekoälyjärjestelmiä voidaan käyttää eettisesti kestäväällä tavalla.

4.1 Tekoälyjärjestelmien eettiset vaikutukset

Tekoälyteknologiat muovaavat yhteiskuntaa vaikuttaen paljon ihmisten arkeen. Monet oikeudelliset ja yhteiskunnalliset kysymykset ovat tuoneet esille, että tekoälyteknologioilla voi olla ei-toivottuja vaikutuksia. Algoritmit voivat esimerkiksi pahentaa olemassa olevia vinoumia, ja ne voivat syrjiä. Niillä voi olla kohtalokkaita seurauksia, sillä ne voivat jopa uhata ihmisten turvallisuutta ja manipuloida. (Rusanen ym. s.a.)

Tekoälyn etiikka tarkoittaa arvoja, jotka ohjaavat ihmisiä kehitettäessä älykkäitä teknologioita. Tekoäly tekee valintoja sen perusteella, miten se on ohjelmoitu ja millä datalla se on opetettu. Tekoälyn herättämiä kysymyksiä ovat esimerkiksi miten tietoja käytetään, kuka on vastuussa koneen tekemistä päätöksistä ja käytetäänkö vaikkapa kasvojentunnistusta ihmisoikeuksia kunnioittavalla tavalla. Laajemmin kyse on siitä, kuinka haluamme oppivia teknologioita käyttää. (Helsingin yliopisto s.a.) Anna-Mari Rusasen mukaan (Pekkarinen 2021) jokainen tarvitsee tekoälyn lukutaitoa, jonka osa on eettisten kysymysten ymmärtäminen.

Tekoälyn käyttö koulutuksessa ei ole enää kaukaista tulevaisuutta vaan arkipäivää. Se muuttaa jo nyt kouluissa ja yliopistoissa tapaa, jolla opettajat työskentelevät ja oppilaat sekä opiskelijat oppivat. Tekoäly tekee koulutusympäristöistä yhä mukautuvampia, sillä se auttaa opettajia vastaamaan oppilaan erityistarpeisiin. Siitä onkin tulossa perustyökalu yksilöityyn ohjaukseen ja arviointiin. On myös huomattu, että tekoälyn avulla on saatavissa arvokasta tietoa opiskelijoiden kehityksestä. Tekoälyn vaikutus koulutusjärjestelmiin väistämätön ja sen merkitys tulee korostumaan tulevaisuudessa. (Euroopan komissio 2022.)

Tekoälyjärjestelmien kehittyessä ja datan käytön lisääntyessä on tärkeää ymmärtää niiden vaikutuksista ympäröivään maailmaan, erityisesti koulutukseen. Opettajilla ja koulujen johtajilla on oltava vähintään perustiedot tekoälyn ja datan käytöstä, jotta kyseiseen teknologiaan voidaan suhtautua myönteisesti, mutta harkiten ja eettisesti kestäväällä tavalla. (Euroopan komissio 2022.)

Tekoälyn etiikassa tarkastellaan keinoja, joilla voidaan minimoida eettiset riskit, joita tekoälyn käytöstä voi aiheutua yhteiskunnassa. Riskejä voivat aiheuttaa esimerkiksi teknologian suunnittelu, epäasianmukainen soveltaminen tai tahallinen väärinkäyttö. Riskien minimoiminen koskettaa niin kehittäjiä, valmistajia, viranomaisia kuin toimijoitakin. (Rusanen ym. s.a.)

4.2 Tekoälyä koskeva Euroopan Unionin politiikka ja muu lainsäädäntö

Yleinen tietosuoja-asetus

Yleinen tietosuoja-asetus asettaa yrityksille ja organisaatioille henkilötietojen keräämiseen, säilytykseen ja hallintaan tarkat vaatimukset. Näitä vaatimuksia sovelletaan eurooppalaisiin organisaatioihin, jotka käsittelevät ihmisten henkilötietoja Euroopan Unionissa. Vaatimuksia sovelletaan myös Euroopan Unionin ulkopuolisiin organisaatioihin, joiden suorittama tietojen käsittely kohdistuu Euroopan Unionin alueella asuviin ihmisiin. (Euroopan komissio s.a.)

Koulujen on varmistettava, että tiedot säilytetään luottamuksellisesti ja turvallisesti. Niillä on oltava myös asianmukaiset toimintaperiaatteet ja menettelyt, joilla suojataan henkilötiedot ja varmistetaan niiden käyttäminen eettisellä tavalla yleisen tietosuoja-asetuksen mukaisesti. (Euroopan komissio 2022.)

Euroopan komission ehdotus tekoälysäädöksestä

Euroopan komissio on tehnyt vuonna 2021 ehdotuksen tekoälysäädöksestä, joka on tekoälyä koskeva kattava oikeudellinen kehys. Ohjeet perustuvat tekoälyä ja dataa käsittelevän Euroopan Unionin sääntelyn sekä politiikan kehitykseen, joihin kuuluvat yleistä tietosuoja-asetusta ja datasäädöstä koskevat ehdotukset. Näissä otetaan huomioon myös koulutuksen erityispiirteet. Ohjeissa annetaan tietoa ja käytännön ohjausta opettajille, jotka ovat työssään yhä useammin tekemisissä tekoälyn kanssa. (Euroopan komissio 2022.)

Ehdotus uusista säännöistä ja toimenpiteistä pyrkii tekemään Euroopasta johtavan alueen luotettavan tekoälyn kehittämisessä. Tekoälyasetuksen tarkoituksena on varmistaa ihmisten ja yritysten turvallisuus sekä perusoikeuksien toteutuminen tekoälyä hyödynnettäessä. Asetuksessa tekoälyjärjestelmät luokitellaan riskikategorioihin niistä aiheutuvan riskin perusteella: ei-hyväksyttävä riski, suuri riski, vähäinen riski vai minimaalinen riski. Sovellettavat säännöt määräytyvät tekoälyjärjestelmälle määritellyn riskikategorian mukaisesti. (KPMG Oy Ab s.a.)

Sellaiset tekoälyjärjestelmät kielletään kokonaan, joista aiheutuvaa riskiä ei pidetä hyväksyttävänä ja jotka muodostavat selvän uhan ihmisten turvallisuudelle, toimentulolle tai oikeuksille. Tähän kategoriaan sisältyvät esimerkiksi tekoälyjärjestelmät ja -sovellukset, joilla voidaan manipuloida ihmisen käyttäytymistä vapaan tahdon vastaisesti. Kategoriaan kuuluvat myös järjestelmät, jotka pisteyttävät sosiaalisten tekijöiden perusteella ihmisiä. (KPMG Oy Ab s.a.)

Tekoälyssäädöksessä vahvistetaan käyttötarkoituksen perusteella suuririskisinä pidettäville tekoälyjärjestelmille oikeudellisesti sitovat vaatimukset. Suuririskisiä tekoälyjärjestelmiä ovat esimerkiksi jotkin yleissivistävässä ja ammatillisessa koulutuksessa käytettävät tekoälyjärjestelmät. Tulevaisuudessa tekoälyssäädöstä sovellettaessa, tekoälyjärjestelmiä käyttävät oppilaitokset ja koulut tietävät palveluntarjoajan saaman sertifiointin perusteella, että suuririskiset tekoälyjärjestelmät täyttävät tietyt vaatimukset ja ovat luotettavia. (Euroopan komissio 2022.)

4.3 Tekoälyn käyttö koulutusalailla

Tietojenkäsittelytieteeseen perehtymättömät saattavat vierastaa tekoälyyn ja tietojärjestelmiin liittyvää erikoissanastoa. Joskus myös asiaan perehtyneilläkin saattaa olla haastavaa ymmärtää täysin tekoälyn toimintaa, koska kyseessä on laaja ja monimutkainen ala. Tekoälyjärjestelmien sisäistä toimintalogiikkaa voi olla vaikea hahmottaa, joten sitä kutsutaan toisinaan niin sanotuksi ”mustan laatikon” ongelmaksi. (Euroopan komissio 2022.)

Järjestelmästä tekevät ”mustan laatikon” järjestelmän monimutkaisuus, selitettävien ratkaisujen kehittämisen vaikeus ja riskit. Nykyisten tekoälyjärjestelmien neuroverkkojen toiminta koodataan tuhansiin, ehkä jopa miljooniin, numeerisiin kertoimiin. Neuroverkon toiminta perustuu näiden arvojen monimutkaiseen vuorovaikutukseen, jonka takia sen toimintaa on käytännössä mahdoton täysin ymmärtää. Käytetyt tekoälymallit saattavat tukea jonkinlaista selitettävyyttä, mutta saattaa olla vaikea kehittää käyttäjäkokemusta, joka perustuisi käyttäjille helposti ymmärrettäviin selityksiin. Monia tekoälyalgoritmeja voidaan huijata suunnittelemalla syöte, joka aiheuttaa järjestelmään toimintahäiriön, jonka takia järjestelmät saatetaan suunnitella tarkoituksella mustiksi laatikoiksi. (Rusanen ym. s.a.)

Tekoäly on kokoelma menetelmiä ja tekniikoita, joista tekoälyjärjestelmä koostuu. Opettajilla ei ole tarvetta ymmärtää täydellisesti tekoälyjärjestelmän toimintaa. Heidän olisi kuitenkin hyvä olla tietoisia tekoälyjärjestelmien perusmekanismeista ja rajoituksista, sekä niiden käyttömahdollisuuksista opetuksen ja oppimisen tukemiseen turvallisella sekä eettisellä tavalla. (Euroopan komissio 2022.)

Tekoäly muuttaa tapoja oppia, työskennellä ja elää. Se vaikuttaa vääjäämättä myös koulutukseen. Kaikilla olisi oltava mahdollisuus osallistua tekoälyn kehittämiseen ja hyötyä siitä. Painottamalla eettisiä periaatteita keskustelussa tekoälyn roolista koulutuksessa voidaan edistää myös tekoälyjärjestelmien ja ratkaisujen eettistä, luotettavaa, oikeudenmukaista ja osallistavaa kehittämistä sekä käyttöä. (Euroopan komissio 2022.)

Tekoäly voi aiheuttaa syrjintää tai uudenlaista eriarvoisuutta. Se voi myös pahentaa niiden nykyisiä muotoja. Asianmukaisesti suunniteltuna ja käytettynä tekoälyn avulla voidaan parantaa esteettömyyttä ja osallisuutta arjessa, työelämässä ja koulutuksessa. (Euroopan komissio 2022.)

Tekoälyjärjestelmien kehittyessä ne täydentävät ja korvaavat yhä useammin ihmisten suorittamia tehtäviä. Tämä voi aiheuttaa eettisyyteen ja luottamukseen liittyviä kysymyksiä, kuten onko tekoälyn avulla mahdollista tehdä oikeudenmukaisia päätöksiä, entä voidaanko näitä päätöksiä varten kerätyt tiedot suojata. Ehdotettu Euroopan Unionin tekoälysäädös auttaa varmistamaan, että palveluntarjoajat noudattavat pakollisia vaatimuksia kehittäessään suuririskiseksi luokiteltuja tekoälyjärjestelmiä. Tällä voidaan pienentää riskejä ja varmistaa järjestelmien luotettavuus. Koulutusviranomaiset ja koulut voisivat silloin varmistaa, että tekoälyjärjestelmät ovat sääntelykehyksen mukaisia ja keskittyä siihen, miten tekoälyä voidaan käyttää eettisellä sekä tietosuojasäännösten mukaisella tavalla. (Euroopan komissio 2022.)

Tekoälyllä ei voida korvata opettajia, mutta sen avulla voidaan tukea heidän työtään. Tekoäly voi auttaa suunnittelemaan oppimiskokemuksia, joissa oppijat voivat muun muassa käyttää luovuuttaan, ratkaista ongelmia ja tehdä tehokasta yhteistyötä. Lisäksi tekoälyn avulla on mahdollista automatisoida hallinnollisia rutiinitehtäviä, jolloin oppimisympäristöä varten jää enemmän aikaa. (Euroopan komissio 2022.)

Tekoälyjärjestelmien käyttö opetustilanteissa kasvaa kaikkialla Euroopassa. Tekoälyä voidaan käyttää erilaisin tavoin opetus-, oppimis- ja arviointikäytännöissä. Tekoäly tarjoaa mahdollisuuden parantaa opetus- ja oppimiskäytäntöjä sekä auttaa kouluja parantamaan niin organisaatiotaan kuin toimintatapaansaakin. Tekoälyn vaikutuksesta koulutuksessa on tehty vain vähän näyttöön perustuvaa tutkimusta, joten tekoälyyn on tärkeää suhtautua kriittisesti ja sen käyttöä tulee valvoa. (Euroopan komissio 2022.)

Tekoälyjärjestelmiä on mahdollista käyttää eri tavoin opetuksen tukemiseen ja oppimisen helpottamiseen. Puhuttaessa arvioinnissa, koulujen hallinnollisissa, opetuksessa tai oppimisessa käytettävistä tekoälyjärjestelmistä, erotetaan usein toisistaan opiskelijoille suunnatut, opettajille suunnatut ja järjestelmälähtöiset tekoälyjärjestelmät. (Euroopan komissio 2022.)

Opiskelijoille suunnatut tekoälyjärjestelmät voidaan jakaa vielä tekoälyn käyttöön opetuksessa ja tekoälyn käyttöön oppimisen tukemisessa. Tekoälyä voidaan käyttää opetuksessa, esimerkiksi seuraavien järjestelmien avulla: älykkäät tutorointijärjestelmät, vuoropuheluun perustuvat tutorointijärjestelmät ja kieltenopiskeluun tarkoitetut sovellukset. Oppimisen tukevia järjestelmiä ovat kokeilevat oppimisympäristöt, kirjoitustehtävien formatiivinen arviointi ja tekoäly yhteistoiminnallisen oppimisen tukena. (Euroopan komissio 2022.)

Opettajille suunnattuja tekoälyjärjestelmiä ovat esimerkiksi kirjoitustehtävien summatiivinen arviointi ja esseiden pisteytys, opiskelijafoorumien seuranta, tekoälyyn perustuvat opetusavustajat sekä pedagogisia resursseja koskevat suositukset. (Euroopan komissio 2022.)

Järjestelmälähtöisissä tekoälyjärjestelmissä tekoälyä käytetään diagnostiikan tai koko järjestelmän kattavan suunnittelun tukemiseen. Näitä ovat esimerkiksi koulutusta koskevien tietojen louhinta resurssien kohdentamiseksi, oppimisvaikeuksien diagnosointi ja ohjauspalvelut. (Euroopan komissio 2022.)

4.4 Opetus- ja koulutusalan eettinen viitekehys

Käsitteellä ”arvo” tarkoitetaan jonkin asian tai teon tärkeysastetta, kun normit ovat puolestaan arvopohjaisia periaatteita, käskyjä ja määräyksiä. Normeilla määritetään, mitä pitäisi tehdä tai mitä odotetaan. Normeja on useanlaisia, kuten sosiaalisia tai moraalisia. Sosiaaliset normit kertovat esimerkiksi, mikä jäsenten mielestä on asianmukaista toimintaa heidän keskuudessaan. Moraaliset normit ovat sosiaalisia normeja pidemmälle meneviä. Ne ovat ohjailevia tai kieltäviä sääntöjä. (Rusanen ym. s.a.)

Yhteisön hyväksymien arvojen ja normien pohjalta pyritään rakentamaan konsensus eettisille kehyksille. Tekoälyn eettisen kehyksen kehittämiseen voivat osallistua eri organisaatiot. (Rusanen ym. s.a.) Vuoden 2019 tehdyn tutkimuksen mukaan (Jobin, Ienca & Vayena 2019) seuraavien viiden eettisen periaatteen ympärillä on tapahtunut maailman laajuinen lähentyminen: läpinäkyvyys, oikeudenmukaisuus ja reiluus, pahan välttäminen, vastuullisuus ja yksityisyys.

Euroopan komissio on määritellyt (Euroopan komissio 2022) neljä keskeistä näkökohtaa, jotka muodostavat perustan tekoälyn ja datan eettiselle käytölle opetuksessa, oppimisessa ja arvioinnissa. Nämä näkökohdat ovat ihmisen toimijuus, ihmisläheisyys, oikeudenmukaisuus ja perusteltujen valintojen tekeminen. (Euroopan komissio 2022.)

Ihmisen toimijuudesta puhuttaessa tarkoitetaan yksilön kykyä tulla täysivaltaiseksi yhteiskunnan jäseneksi. Aktiivisella toimijalla on mahdollisuus itse päättää elämänvalinnoistaan ja ottaa vastuuta omasta toiminnastaan. Toimijuus on laajasti käytettyjen käsitteiden taustalla, kuten itsenäisyys, itsemääräämisoikeus ja vastuu. (Euroopan komissio 2022.)

Ihmisläheisyydellä tarkoitetaan, että huomioidaan ihmiset, heidän identiteettinsä, koskemattomuutensa ja ihmisarvonsa. Aito inhimillinen yhteys edellyttää, että huomioidaan ihmisten hyvinvointi, kunnioitus, merkityksellinen vuorovaikutus, sosiaalinen yhteenkuuluvuus ja turvallisuus. Inhimillinen yhteys tarkoittaa, että ihmisiä lähestytään heidän itseisarvonsa pohjalta eikä dataobjekteina tai keinoina päästä johonkin tavoitteeseen. Tämä on olennainen osa tekoälyn ihmiskeskeistä lähestymistapaa. (Euroopan komissio 2022.)

Oikeudenmukaisuus tarkoittaa, että yhteiskuntajärjestelmässä kaikkia kohdellaan oikeudenmukaisesti. Selkeät menettelyt ovat tarpeellisia, jotta kaikilla ovat samanlaiset mahdollisuudet. On tärkeää huomioida esimerkiksi osallisuus, syrjimättömyys ja tasapuolisuus samoin kuin oikeuksien ja velvollisuuksien oikeudenmukainen jakautuminen. (Euroopan komissio 2022.)

Perustelujen valintojen tekeminen tarkoittaa sitä, että kouluympäristössä eri sidosryhmien tekemien tarpeellisten yhteisten päätösten tukena käytetään tietämystä, faktoja ja dataa tarpeellisten yhteisten päätösten tukena. Tämän toteutuminen edellyttää läpinäkyvyyttä. Se juontaa juurensa osallistaviin ja yhteistyöhön perustuviin päätöksentekomalleihin sekä selitettävyyteen. (Euroopan komissio 2022.)

Tekoälyn etiikan viisi periaatetta ovat hyvää tekemisen periaate/vahinkojen välttämisen periaate, vastuullisuusperiaate, läpinäkyvyyden periaate, oikeudenmukaisuuden periaate ja perusihmisoikeuksien kunnioittamisen periaate (Rusanen ym. s.a.). Myös Euroopan komissio (Euroopan komissio 2022) on määrittänyt eettisiä vaatimuksia, jotka ovat samansuuntaisia edellä mainittujen kanssa.

Euroopan komissio (2022) on listannut AI HLEG -ryhmän eettisiin ohjeisiin perustuvat vaatimukset seuraavasti:

- ihmisen toimijuus ja ihmisen suorittama valvonta,
- läpinäkyvyys,
- monimuotoisuus, syrjimättömyys ja oikeudenmukaisuus,
- yhteiskunnallinen ja ekologinen hyvinvointi,
- yksityisyyden suoja ja datan hallinta,
- tekninen luotettavuus ja turvallisuus,
- vastuuvollisuus.

Keskeiset eettiset vaatimukset auttavat varmistamaan, että koulutuksessa käytettävät tekoälyjärjestelmät ovat luotettavia ja niissä huomioidaan niihin liittyvät huolenaiheet. Kaikkien koulutuksessa käytettävien tekoälyjärjestelmien olisi hyvä olla vaatimusten mukaisia, sillä niissä puututaan tärkeisiin huolenaiheisiin. Huolenaiheita ovat esimerkiksi koulutustuloksiin vaikuttavat vääristymät ja virheiden riskit. (Euroopan komissio 2022.)

Koulun on tärkeää laatia ja ottaa käyttöön yhteistyöhön perustuvan sekä refleктоivan koulun sisäisen arviointiprosessin, kun tekoälyn ja datan käyttöä harkitaan. Tämän edellytyksenä on, että opettajat ottavat selvää, miten he voivat käyttää tekoälyjärjestelmiä opetuksen ja oppimisen tukemiseen. Tekoälyn ja datan käytön seurauksia sekä vaikutuksia koulutuksessa voi olla hankala ennustaa. Tästä syystä teknologian kehittämisessä, käyttöönotossa ja arvioinnissa on edettävä vaiheittain. Välineet tulee ottaa käyttöön asteittain, jolloin niiden mahdollisia yhteiskunnallisia vaikutuksia voidaan seurata jatkuvasti. Silloin voidaan mennä taaksepäin, mikäli tahattomia seurauksia ilmenee. Tekoälyn eettinen käyttö koulutuksessa vaatii toimijuutta niin opiskelijoilta, opettajilta kuin koulun hallinnoltakin. Lisäksi se vaatii myös institutionaalisen tason toimijuutta. (Euroopan komissio 2022.)

5 Tutkimus multimodaalisesta tunteiden tunnistamisesta ja sen eettisistä näkökulmista

AI Forum tekee yhteistyötä tekoälyn saralla. Se on viiden korkeakoulun monitieellinen yhteistyöverkosto, johon kuuluvat Haaga-Helian ammattikorkeakoulu, Jyväskylän yliopisto, Oulun yliopisto, Satakunnan ammattikorkeakoulu ja Turun yliopisto. AI Forumin tavoitteena on asiantuntemuksen ja tutkimusyhteistyön vahvistaminen tekoälyn alalla. Verkosto keskittyy yhteistyössä yrityskumppaneiden kanssa erityisesti tekoälyn sovelluksiin digipalveluissa ja tekoälyn tuomaan muutokseen työelämässä. (AI Forum s.a.) Opinnäytetyö tehtiin toimeksiantona AI Forumille.

5.1 Tutkimuksen esittely

Tutkimus toteutettiin laadullisena (kvalitatiivinen) tutkimuksena kevään 2023 aikana. Siinä pyrittiin ymmärtämään, kuinka multimodaalinen tunteiden tunnistaminen ja siihen liittyvät eettiset näkökohdat koetaan opetus- ja koulutusalan näkökulmasta. Lisäksi tutkittiin millaisia mahdollisuuksia ja toisaalta myös haasteita nähdään.

5.2 Tutkimuskysymykset, tutkimusmenetelmät, aineisto

Methodinen etu haastattelussa on siinä, että haastateltaviksi voidaan valita sellaisia henkilöitä, joilla tiedetään olevan kokemusta joko tutkittavasta ilmiöstä tai muuten tietoa aiheesta. Tällöin puhutaan tarkoituksenmukaisesta, eli harkinnanvaraisesta näytteestä. Tiedonkeruutilanteessa haastattelu antaa tutkijalle mahdollisuuden suunnata tiedonhankintaa siihen suuntaan, joka on tutkimuskysymysten kannalta olennaista. (Puusa & Juuti 2020, 267-268.)

Tutkimuksessa haastateltiin asiantuntijoita, opiskelijoita ja huoltajia. Haastatteluihin valikoitui kahdeksan henkilöä eri aloilta, joilla tiedettiin etukäteen olevan joko kokemusta tutkittavasta ilmiöstä tai tietoa siihen liittyvästä jostakin aihealueesta. Haastateltavista neljä ovat multimodaalisen tunteiden tunnistamisen asiantuntijoita, joilla on aiheesta erittäin paljon tietoa. Lisäksi haastateltiin opiskelijaa, jolle aihealue on osittain tuttu ja jolla on näkemyksiä multimodaaliseen tunteiden tunnistamiseen opiskelijan näkökulmasta. Oikeustieteiden alalla työskentelevälle henkilölle multimodaalinen tunteiden tunnistaminen ei ollut ennestään tuttu aihe, mutta hänellä tiedettiin olevan näkemyksiä aiheeseen lainsäädännön näkökulmasta. Haastatelluilla vanhemmilla ei ollut myöskään multimodaalisesta tunteiden tunnistamista etukäteen tietoa, mutta tutkimuksessa haluttiin selvittää, kuinka he kokevat asian vanhempina ja millaiset asiat heitä pohdituttavat.

Haastatteluissa edettiin ennalta suunnitellun haastattelurungon mukaisesti, joka myös lähetettiin tutkimukseen osallistuville etukäteen tutustuttavaksi. Näin tutkimukseen osallistuvilla oli halutessaan mahdollisuus valmistautua haastatteluun etukäteen. Tavoitteena oli muodostaa mahdollisimman monipuolinen kuva tutkittavasta ilmiöstä. Haastattelu muokkautui jonkin verran tutkimukseen osallistuvan kokemuksen mukaisesti, jolloin haastatteluissa saatettiin painottaa johonkin aihealueeseen toista enemmän. Haastattelut mahdollistivat myös tarkentavien kysymysten esittämisen ohjailematta tai rajaamatta keskustelua kuitenkaan kovinkaan paljoa. Multimodaalinen tunteiden tunnistaminen on käsitteenä monelle vieras, jonka takia esittelin lyhyesti aihetta haastattelun aluksi varmistaakseni, että puhumme haastattelussa samasta asiasta. Saatekirje ja haastattelukysymykset ovat liitteinä (liite 1 ja liite 2).

Kaikki haastattelut toteutettiin yksilöhaastatteluina puolistrukturoidun haastattelumenetelmän mukaisesti, joka mahdollisti tutkimusten kannalta oleellisten aihealueiden valitsimisen etukäteen antamatta kuitenkaan valmiita vastausvaihtoehtoja. Valmistauduin haastatteluihin opiskelemalla itse aihetta ja koostamalla opinnäytetyön teoriaosuutta. Valitsin haastattelumentelmän sillä perusteella, että opinnäytetyö on varsin rajattu niin ajallisesti kuin työmäärältään. Kokemattomuuteni niin tutkijana kuin multimodaalisen tunteiden tunnistamisen parissa rajasi teemahaastattelun pois vaihtoehtoista tutkimusmenetelmäksi. Haastatteluista osa suoritettiin kasvokkain ja osa Teamsin välityksellä. Kaikki haastattelut nauhoitettiin ja litteroitiin tekstiksi.

5.3 Tutkimusaineiston analysointi

Aineiston hankinnan jälkeen laadullisen aineiston analysointi jatkuu usein kahdella tavalla. Ensinnäkin tutkija pyrkii lukemaan aineistoa yhä uudelleen tavoitellessaan vihjeitä siitä, millaisiin teemoihin, luokkiin tai kategorioihin aineistoa voisi ryhmitellä. Toisaalta hän hankkii lisää tietoa eri lähteistä perehtymällä aihetta käsitteleviin teorioihin ja tutkimuksiin. Kirjallisuudesta hankitut tiedot vaikuttavat osaltaan siihen, millaisia kategorioita tutkija pyrkii löytämään tutkimuksesta. Aineiston analyysin tavoitteena on tutkimuksen kohteena olevan ilmiön kuvailu, tulkitseminen ja ymmärtäminen. (Puusa & Juuti 2020, 370-371.)

Tutkimusaineistoa analysoitiin lukemalla haastatteluiden litterointeja useampaan kertaan ja koodamalla teemoja eri väreillä. Samalla kirjoitettiin Word-tiedostoon muistiinpanoja ranskalaisilla viivoilla. Sen jälkeen pyrittiin löytämään aineistosta kategorioita ja teemoja, joihin syvennyttiin tarkemmin. Samalla etsittiin lähdeaineistosta lisätietoa teemoihin liittyen, joiden pohjalta täydennettiin myös opinnäytetyön teoriaosuutta.

5.4 Tutkimuksen tulokset

Osana tutkimusaineiston analysointia tein aiheista ja teemoista useita miellekarttoja, joiden perusteella jäsentelin aineistoa ja pyrin löytämään keskeisiä teemoja. Oheisessa miellekartassa (kuva 1) on jäsenneilty tutkimusaineiston teemoja tutkimuskysymysten mukaisesti.



Kuva 1. Tutkimustulosten analysointi miellekarttaa hyödyntäen

Haastatteluissa multimodaalisen tunteiden tunnistamisen tekniikoiksi mainittiin kasvojen- ja puheentunnistus sekä aivosähkökäyrä. Multimodaaliselle tunteiden tunnistamiselle nähdään mahdollisuuksia oppimistulosten ja opetuksen parantamisessa, autismin parissa työskentellessä ja virtuaalitodellisuuden parissa. Haasteita puolestaan nähdään yksilöllisten ja kulttuuristen erojen huomioimisessa, teknologian väärinkäytössä, teknisissä haasteissa sekä dataan liittyvissä kysymyksissä. Keskeiset eettiset näkökulmat ovat yksityisyys, syrjimättömyys ja oikeudenmukaisuus. (Kuva 1.)

Haastateltavat kokevat tunteet hyvin yksityiseksi asiaksi, joita ei välttämättä aina haluta jakaa muiden kanssa. Joissain tilanteissa voitaisiin olla suostuvaisia tunteiden tunnistamiseen, mutta silloinkin asiaan vaikuttaa merkittävässä määrin dataan liittyvät kysymykset. Kuka dataa kerää ja mihin tarkoitukseen, mitä hyötyä datan keräämisestä olisi ja kenelle? Järjestelmien tulee olla niin hyvin koodattuja, etteivät ne aiheuta syrjintää.

5.4.1 Olemassa olevat signaalit ja tekniikat

Eri multimodaalisten tunteiden tunnistamisteknologioihin liittyy haastateltavien pohdintojen mukaisesti erilaisia eettisiä kysymyksiä. Haastatteluissa esiin nousseet keskeiset tekniikat multimodaalisten tunteiden tunnistamiseksi olivat kasvojen tunnistus, puheen tunnistus ja aivosähkökäyrä. Haastateltavat pohtivat aihetta hyvin pitkälti nimenomaan ihmisen fyysisten ja psyykkisten muutosten näkökulmasta.

Kasvojentunnistusjärjestelmät, joiden avulla voidaan tallentaa esimerkiksi ilmeitä, lienee teknologioista tutuin. Siihen liitetään haastatteluiden perusteella usein yksityisyyteen, syrjimättömyyteen ja oikeudenmukaisuuteen liittyviä eettisiä kysymyksiä.

Kasvojentunnistusjärjestelmien tiedetään olleen epäreiluja tiettyjä kasvojenpiirteitä tai ihonvärejä kohtaan (ks. 3.2.1). Yksi haastateltavista pohti myös esimerkiksi silmälasien ja parran vaikutusta kasvojen tunnistukseen.

Kasvojen koetaan olevan todella henkilökohtainen ja iso osa elämää, jolloin kasvoihin liittyvää dataa ei välttämättä haluta jakaa järjestelmien analysoitavaksi. Toisaalta haastateltavat kokivat, että tunnistettaessa kasvoja ja ilmeitä ihminen pystyy itse vaikuttamaan siihen millaisen kuvan itsestään antaa. Sen sijaan fysiologisiin muutoksiin, kuten silmien pupillien laajentumiseen, koettiin pystyttävän vaikuttavan huomattavasti vähemmän (vrt 3.2). Osa haastateltavista koki tämä negatiivisena asiana, sillä järjestelmät saattaisivat paljastaa itsestä tietoja, joita ei haluta muiden tietävän.

Useampi haastateltavista pohti kasvojentunnistusjärjestelmiin liittyen sitä, miten ihmisten erilaisuus vaikuttaa tulokseen. Ihmisellä saattaa olla taustalla jotakin, jonka takia hän saattaa ilmaista tunteitaan eri tavalla kuin muut. Toisaalta tunteiden osoittamiseen vaikuttavat myös esimerkiksi persoonallisuus ja kulttuurinen tausta (ks. 3.5). Multimodaalista tunteiden tunnistamista tekevän tutkijan ja näitä tietoja muutoin hyödyntävien henkilöiden tulisi ymmärtää hyvin laajasti aihetta virheellisten tulkintojen välttämiseksi.

Puheentunnistusjärjestelmiin liittyen haastateltavat pohtivat, kuinka ne voidaan liittää esimerkiksi sukupuoleen liittyviin stereotyyppioihin (vrt. 3.2.2). Myös puheentunnistusjärjestelmien kohdalla syrjimättömyyteen ja oikeudenmukaisuuteen liittyvät eettiset kysymykset koettiin keskeisiksi.

Aivosähkökäyrän osalta osa haastateltavista pohti, minkälaista dataa sen avulla henkilöstä olisi saatavilla. Kun dataa yhdistetään vielä muista signaaleista kerättyyn dataan, niin saatetaan saada selville ihmisestä jotakin sellaista mitä ei vielä osata välttämättä ajatella. Useat haastateltavat nostivat esiin huolen tämän tyyppisen datan seurauksista väärissä käsissä. On siis tärkeää huomioida kuka tietoa käyttää ja mihin tarkoitukseen.

Aivosähkökäyrätutkimuksessa tutkittavaan henkilöön kiinnitetään mittauslaitteita, joihin liittyy omat eettiset kysymyksensä (vrt 3.2.3). Vaikka mittauslaitteet ovat turvallisia käyttää, liittyy niihin silti riski henkilön satuttamisesta. Mittauslaitteiden kiinnittäminen voi myös estää henkilöä liikkumasta vapaasti, joka tulee myös huomioida tämän tyyppistä tutkimusta tehtäessä.

Haastatteluissa oma kokemus koettiin myös tärkeänä yhtenä modaliteettina. Toiset kokevat oman kokemuksen olevan hyödyllinen lisä, jolloin sovellukset voisivat palvella entistä paremmin omia tarpeita. Toiset suhtautuivat asiaan epäilevämmiin ja pohtivat, kuinka paikkansa pitävä ihmisen oma kuvaus hänen sen hetkisestä tunteestaan olisi. Jotta omasta kokemuksesta olisi hyötyä, pitäisi tutkittavien puhua totta.

5.4.2 Multimodaalisen tunteiden tunnistamisen mahdollisuudet

Haastateltavat kokevat, että multimodaalisesta tunteiden tunnistamisesta voisi olla hyötyä joissakin sovelluksissa. He pohtivat erilaisia käyttötarkoituksia sovelluksille, joista he kokisivat voivan olla heille itselleen hyötyä. Sellaisia voisivat olla esimerkiksi jonkinlaisia suosituksia tai ehdotuksia antavat sovellukset. Myös oppimistilanteet voisivat olla sellaisia, joissa multimodaalisen tunteiden tunnistamisen metodeista voisi olla hyötyä (ks. 3.7), samoin kuin virtuaalitodellisuus (vrt. 3.2.3).

Multimodaaliseen tunteiden tunnistamiseen opetus- ja koulutuslalla liittyy joukko erilaisia kysymyksiä, joihin kaikkiin ei ole vielä vastauksia. Multimodaalinen tunteiden tunnistaminen voisi kuitenkin auttaa niin opettajia kuin oppijoitakin. Useat haastateltavat pohtivat, että multimodaalinen tunteiden tunnistaminen voisi auttaa kehittämään yksilöllisempää opetusta ja auttaa oppijoita oppimaan paremmin (ks. 3.7). Tähän tosin suhtaudutaan hieman varauksella empiirisen todistusaineiston puuttuessa tai ollessa vielä varsin vähäistä, joka todistaisi asiaa kumpakaan suuntaan.

Oppijoiden näkökulmasta koetaan, että multimodaalinen tunteiden tunnistaminen mahdollistaisi voinnin ja mielialan seuraamisen. Silloin voitaisiin havaita muutokset mielialassa ja auttaa oppijoita oikealla hetkellä. Multimodaalisesta tunteiden tunnistamisesta voisi olla apua myös silloin, jos oppijalla on oppimisvaikeuksia. Mahdollisuuksia nähdään eri koulutustasoilla, vaikka täysi-ikäisten kohdalla yleiseen tietosuoja-asetukseen ja eettisiin kysymyksiin ovat helpompia. Haastateltavat pohtivat tekoälyjärjestelmien eettisyyttä tasapuolisuuden näkökulmasta, sillä järjestelmien tulisi antaa yhdenvertaiset mahdollisuudet kaikille (ks. 4.4). Yksi haastateltava (Mies, 42-vuotias, virkamies) tiivistä ajatuksiaan opiskelijoiden auttamisesta tekoälyjärjestelmien avulla seuraavasti: ”Jos aiotaan lukea opiskelijoiden reaktioita tiedon käyttämiseksi opiskelijoiden auttamiseksi, niin voidaanko ja aiotaanko auttaa kaikkia”.

Opettajien näkökulmasta mahdollisuuksiksi nähdään resurssien kohdentamisen paremmin ja opettajien kouluttamisen. Opettajia olisi mahdollista kouluttaa, kuinka tekoälyä käytetään opetuksessa, jolloin tehokkuus luokassa paranisi (ks. 4.3). Opettajalla olisi myös mahdollisuus seurata, kuinka oppilaat reagoivat opetukseen, jolloin he voivat mukauttaa opetustaan ja kehittää opetustyyliään. Kolmas näkökulma aiheeseen on, että multimodaalisen tunteiden tunnistamisen avulla olisi mahdollista tutkia, miten ihminen oppii oppimisprosessin aikana.

Haastateltavat esittivät useita mahdollisuuksia, kuinka multimodaalinen tunteiden tunnistaminen voisi auttaa lapsia. Puhuttaessa lapsista, eli tässä tapauksessa alle kahdeksantoista vuotiaista, tulee heidän kohdallaan huomioida erilaisia eettisiä kysymyksiä kuin aikuisten kohdalla. Järjestelmät eivät saa vahingoittaa lapsia ja niiden tulee olla reiluja. Sama tietysti koskee myös aikuisia, mutta lasten kohdalla näissä asioissa halutaan olla erityisen tarkkoja. Dataa käytetään sellaisella tavalla, jota lapset eivät välttämättä ymmärrä. Haastateltavilla oli eriäviä näkemyksiä siitä, pitäisikö lapsen itsensä antaa suostumus tunteiden tulkittamiseen, minkä ikäinen lapsi pystyy antamaan itse suostumuksensa, minkä ikäiseltä asiaa voidaan kysyä ja miten lapsi sen ymmärtää.

Yksi haastateltavista (Nainen, 37-vuotias, vanhempi) pohti eri ikäisten lasten kykyä päättää tunteidensa tunnistamiseen seuraavasti: ”Lapsen pitäisi saada itse päättää, mutta päiväkotikäinen ei ehkä ole kykeneväinen. Kouluikäinen pystyy jo vähän miettimään, että haluaako hän jotain vai ei. Toki asia pitäisi pystyä perustelemaan lapselle ikätaso huomioiden, että miksi se on hyväksi ja mitä siitä seuraa.”

Useat haastateltavat mainitsivat multimodaalisen tunteiden tunnistamisen mahdollisuudet autismin parissa (ks. 3.7). Sen lisäksi mainittiin myös muut sosiaaliset ongelmat, jolloin tekoäly voisi auttaa välittämään viestin, kuinka lapsi tuntee, mikäli hänellä itsellään on haasteita näyttää tai sanoittaa tunteitaan. Silloin tunteet voisi olla mahdollista saada näkyviksi ennen kuin ne ovat ääripäässä. Silloin tilanteeseen voitaisiin puuttua ennen kuin se eskaloituu. Toisaalta osa haastateltavista pohti, kuinka se vaikuttaa ihmisen kykyyn tunnistaa tunteita. Mitä tapahtuu vuorovaikutustaidoille?

Osa haastattelemistani vanhemmista suhtautuu multimodaaliseen tunteiden tunnistamiseen luottavammin kuin toiset. Yksi vanhemmista (Nainen, 36-vuotias, vanhempi) suhtautuu multimodaaliseen tunteiden tunnistamiseen epäilevästi tiivistäen ajatuksiaan seuraavasti: ”Lähtökohtaisesti arveluttaa se, että joku laite tulkitseisi lapseni tunteita sen sijaan, että siellä olisi länsäolevia aikuisia”. Toinen vanhemmista (Nainen, 37-vuotias, vanhempi) suhtautui asiaan sen sijaan positiivisemmin: ”Jos kehitetään jotain, niin mielestäni se on ihan hyvä. Jos sen kautta voidaan tuoda lapsen elämään jotain positiivista, niin sehän on hienoa.”

Osa haastateltavista olisivat valmiita suostumaan lapsen tunteiden tulkitsemiseen sellaisessa tilanteessa, jossa siitä olisi lapselle hyötyä, eikä se vahingoita lasta. Tunteiden tulkitsemisesta halutaan kuitenkin tieto etukäteen, jotta asiaan voidaan perehtyä. Toisaalta pohdittiin myös, että miten tunnedatan kerääminen esimerkiksi lapsilta vaikuttaa, jos he tietävät etukäteen, että heidän tunteitaan mitataan. Opetus- ja koulutusalailla on syytä huomioida hyväksyntää kysyttäessä henkilön auktoriteetti suhteessa tutkittavaan henkilöön, jotta henkilölle ei synny oloa, että häntä siihen painostettaisiin (vrt. 3.2.1). Eräs haastateltavista (Nainen, 31-vuotias, tutkija) pohti opettajan ja oppilaan välisen auktoriteetin vaikuttamista oppilaan päätökseen tutkimukseen suostumiseksi seuraavasti: ”Jos olemme opettajia ja jos me kysymme oppilaita tulemaan mittaukseen, se voi aiheuttaa heille paineita. He saattavat ajatella, että heidät pakotetaan tämänkaltaiseen ympäristöön.”

Ajatus oman lapsen tunteiden tulkitsemisesta päiväkodissa tai koulussa tekoälyn avulla herätti vanhemmissa monenlaisia ajatuksia. Jokin tieteellinen tutkimus lapsen parasta ajatellen voisi olla sellainen, johon voitaisiin suostua. Ei kuitenkaan voi olla niin, että pelkästään tulkitaan tunne, vaan siitä pitää seurata jotain kuten, että asiaa lähdetään työstämään lapsen kanssa. Järjestelmää ei pidä myöskään toteuttaa siksi, että sitten voitaisiin laittaa enemmän lapsia samaan tilaan tekoälyjärjestelmän monitoroidessa oppijoita. Sitten tulisi myös huomioida, että vähemmän tunteita näyttävät lapset eivät jää vähemmälle huomiolle.

Yksi keskeinen kysymys on, että kuka tietoa käyttää. Onko tunnedata opettajan käytössä, onko hänen lisäksi muilla pääsy dataan? Haastatellut pitivät tärkeänä, että dataa käsittelevällä tulee olla ammattitaitoa tunnedatan käsittelyyn. He kokevat, että tilanteet ovat niin moninaisia, että pitäisi tietää mitä kaikkea on tunteiden taustalla. Onko perheessä jotain muuta meneillään, joka vaikuttaa? Silloin pitäisi pystyä katsomaan kokonaisuutta laajemmin. Ihmisen läsnäolo ja vuorovaikutus koetaan niin tärkeiksi, ettei niitä voida tekoälyllä korvata.

5.4.3 Multimodaalisen tunteiden tunnistamisen haasteet

Tekoäly on ollut kauan, mutta multimodaaliseen tunteiden tunnistamiseen liittyvä teknologia on verrattain uusi asia. Haastateltavista kovinkaan moni ei välttämättä olisi kovinkaan suostuvainen siihen, että hänen tunteitaan tulkittaisiin multimodaalisen tunteiden tunnistamisen keinoin tällä hetkellä. Jokin tutkimus voisi olla sellainen, johon useampi saattaisi suostua. Tulevaisuudessa tilanne saattaa muuttua, kun saadaan tutkimustuloksia ja todisteita laajemmassa mittakaavassa.

Osalle haastateltavista multimodaalinen tunteiden tunnistaminen ei ollut ennestään tuttu aihe. Osa koki aiheen vieraana ja epäilyttävänä. Jos aiheesta olisi enemmän tieteellistä tutkimusta, niin se saattaisi vaikuttaa asiaan.

Varsinkin isojen ihmisjoukkojen tunteiden tunnistaminen koettiin pelottavana asiana, joka sai pohtimaan millaisiin käyttötarkoituksiin sellaista järjestelmää voitaisiin käyttää. Eräs haastateltava (Nainen, 59-vuotias, oikeustieteet) pohti järjestelmien mahdollisuutta ihmisjoukkojen manipulointiin (ks. 4.1) seuraavasti: ”Jos isojen ihmisjoukkojen tunteita voitaisiin tulkita tekoälyllä, niin se olisi pelottavaa. Avaisiko se mahdollisuuden ihmisjoukkojen manipulaatioon? Kun voitaisiin sanoa, että tämä joukko on nyt siinä mielentilassa, että kannattaa tehdä näin tai olla tekemättä.”

Osa haastateltavista kokivat haasteiden hallinnan hieman hankalaksi ja monimutkaiseksi asiaksi. Mitään yksiselitteistä kaiken kattavaa ratkaisua ei ole olemassa. Toisaalta koettiin, että haasteita voidaan hallita lainsäädännöllä, säännöillä ja ohjeilla (vrt. 4.2). Toisaalta sitten taas koettiin, että olemme tekoälyn kehityksessä vaiheessa, jossa haasteet eivät ole helposti hallittavissa. Yksi haastateltava (Mies, 42-vuotias, virkamies) pohti, että olemme tekoälyn hyödyntämisessä leviämisvaiheessa, jossa haasteita ei ole kovin helppoa hallita: ”Haasteita ei voida hallita kovin helposti. Olemme alussa sen ymmärtämisessä ja yleensä, kun on iso aloitus teknologiassa. Se menee ensin hieman villiksi ja sitten sitä hallitaan. Luulen, että olemme nyt leviämisvaiheessa. Olemme villissä normaalitilanteessa, jossa tekoälyä ymmärtämättömät ihmiset ottavat tekoälyn käyttöön, eivätkä tiedä mitä tekevät.”

Haasteiden hallintaan on kuitenkin olemassa monia keinoja. Ylätasolla Euroopan Unioni hallitsee haasteita ainakin datan osalta, sillä se on asettanut muun muassa tietosuojasetuksen. Heillä on tällä hetkellä työn alla tekoälyasetus, jolla haasteita pyritään myös hallitsemaan (ks. 4.2) Haastateltavat kokevat Euroopan Unionin määrittelemät säännöt tarpeelliksi. Niissä määritellään muun muassa, kuinka järjestelmiä voidaan käyttää.

Haasteita voidaan hallita osaltaan Suomessa myös lainsäädännöllä. Haastateltavat kokevat, että tarvittaisiin edelleen lakeja, jotka estävät järjestelmien pahat käyttötarkoitukset. Toisaalta osaa haastateltavista pohditutti, kuka lakeja sitten valvoo?

Yksi haastateltavista (Nainen, 59-vuotias, oikeustieteet) tiivisti ajatuksiaan lainsäädännöstä seuraavasti: ”Lainsäädännössä voidaan säädellä lähes mitä vain, mutta on eri asia mitä voidaan valvoa, tai panemaan täytäntöön, jotta varmistutaan säädösten noudattamisesta.” Lainsäädännön koettiin tulevan pikkaisen jälkijunassa, jonka koetaan johtuvan siitä, että lakeja ei useinkaan säädetä ennakoimaan tulevia tilanteita, vaan niillä reagoidaan jo olemassa oleviin haasteisiin.

Haastatellut näkivät, että haasteita voitaisiin hallita myös vahvalla yhteydellä tutkijoiden ja yritysten välillä. Olisi tärkeää tietää, mitkä ovat yritysten tulevaisuuden tarpeet, jotta kehitettävät mallit palvelevat käyttötarkoituksia. Tutkimustyön halutaan kohtaavan todellisen maailman kanssa, jotta se voi auttaa ihmisiä.

Tutkimuksen tulee olla vastuullista, eli tutkimuksessa tulee noudattaa sääntöjä ja suojata dataa. Toisaalta haluttaisiin enemmän standardeja ja kriteerejä, jotta tiedetään mitä voidaan tehdä ja mitä ei.

5.4.4 Multimodaalisen tunteiden tunnistamisen eettiset näkökulmat

Haastatteluissa esiin nostettuja eettisiä näkökulmia ovat yksityisyys, läpinäkyvyys, vastuullisuus, oikeudenmukaisuus, luottamuksellisuus, yhdenvertaisuus ja syrjimättömyys. Kuten aikaisemmin todettiin, joidenkin signaalien tai tekniikoiden kohdalla voi tulla aiheelliseksi pohtia myös ihmisen fyysiseen vahingoittamiseen ja liikkumisvapauden rajoittamiseen liittyvät eettiset kysymykset, jos esimerkiksi tutkittavalle henkilölle asetetaan joitain mittausvälineitä.

Ihmisen toimijuus

Tutkimuksessa yhdeksi keskeisimmäksi teemaksi nousi yksityisyys, jonka haastateltavat kokivat erittäin tärkeäksi asiaksi. Haastatteluissa haastateltavat pohtivat milloin tunteita halutaan näyttää muille ja milloin niitä puolestaan ei haluta näyttää. Haastateltavat kokevat, että jokaisella on oikeus omiin tunteisiinsa ja että jokaisen tunnetdata on hänen omaansa. Kaikki eivät halua jakaa tunteitaan muiden kanssa, ei ainakaan kaikissa tilanteissa. Varsinkin koettaessa negatiivisia tunteita saatetaan mielummin haluta olla kotona yksin, kuin näyttää tunteita muille. Yksi haastateltavista (Mies, 42-vuotias, virkamies) pohti, että varsinkin koettaessa negatiivisia tunteita, niitä ei haluta näyttää muille: ”Jos halutaan olla huolestuneita tai surullisia, niin se on jokaisen oikeus. Ollessaan poissa tolaltaan halutaanko lähteä ulos ja näyttää muille olevansa poissa tolaltaan, jos jotain kauheaa tapahtuu? Mahdollisesti ei, vaan lukittaudutaan yksin huoneeseen jäätelön kanssa ja katsellaan Netflixiä pari tuntia.”

Monet haluavat myös pystyä kontrolloimaan millaisen kuvan itsestään antavat muille ja siinä tunteiden hallinta on keskeistä. Haastateltavat kokevat, että tilanteet tulisivat olla sellaisia, joissa tiedetään, että tunteiden tulkintaa tehdään, eikä sitä saisi tehdä ilman lupaa.

Useampi haastateltavista koki, että tunnetdatan kerääminen ilman nimenomaista lupaa, on eettisesti arveluttavaa. Haastateltavat luottavat siihen, että sopimuksista pidetään kiinni, ainakin meidän yhteiskunnassamme. Toisaalta osa haastateltavista pohti erilaisia käytäntöjä, joita eri valtioissa voi olla. Se on huomionarvoista, sillä samaa internettiä käytetään kaikkialla, eikä se ole sidoksissa valtion rajoihin.

Datan käsittelyyn liittyy monenlaisia haasteita: missä data sijaitsee, onko se tallennettu pilvipalveluun vai säilötäänkö se jonnekin paikallisesti? Oma dataa halutaan kontrolloida ja sitä pitäisi pystyä tarvittaessa korjaamaan (vrt. 3.2.1).

Dataa tarvitaan paljon ja siihen liittyvätkin ehkä suurimmat haasteet. Datan puutteeseen vaikuttavat osaltaan laitteet, joilla dataa kerätään. Toisaalta useista modalityteeteistä kerätty data tulee yhdistää, joka on myös suuri haaste. Yksi haastateltavista (Mies, 31-vuotias, tutkija) pohti multimodaaliseen tunteiden tunnistamiseen vaadittavaa laitteistoa seuraavasti: "Multimodaalista tunnedataa on vaikea kerätä ihmisistä, sillä joskus vaaditaan korkean tarkkuuden välineitä signaalien keräämiseksi." (vrt. 2.2)

Haastatteluissa heräsi ajatuksia siitä, että esimerkiksi kaupallisten sovellusten ei haluttaisi tietävän, kuinka oikeasti tunnetaan (vrt. 3.2.1). Tunnetietoa ei haluta ikään kuin menettää yritykselle. Haastateltavat haluavat tietää, käsittelevätkö jotkin kolmannet osapuolet heidän dataansa. Haastatteluissa nousi esiin myös huoli siitä, että järjestelmiä ja malleja käyttäisivät sellaiset tahot, joilla ei ole tarpeeksi ymmärrystä niiden toiminnasta. Eräs haastateltavista (Mies, 31-vuotias, tutkija) tiivisti ajatuksiaan seuraavasti: "Monet menetelmistä ovat nykyään avointa lähdekoodia, jota kaikki voivat käyttää teknistä kehitystä varten. Silloin joku voi päästä käsiksi näihin malleihin ja käyttää niitä ilman asiantuntemusta tai asiantuntevaa taustaa."

Datan turvallisuuteen liittyen moni haastateltavista pohti tietoturvaan ja -vuotoihin liittyviä kysymyksiä. Mihin data menee, kuinka se suojataan, miten estetään väärinkäyttö? (vrt. 4.2) Data käsiteltynä väärissä käsissä voi olla ongelma, emmekä tiedä mitä silloin tapahtuu. Ei haluta, että näin sensitiivistä dataa joutuu väriin käsiin. Kaikkea henkilöstä kerättyä tietoa voidaan käyttää häntä vastaan. Alun perin datan käsittelijällä on voinut olla hyvä tarkoitus, mutta tietämättömyys on voinut johtaa johonkin muuhun, jos dataa pääsee vuotamaan tai järjestelmiä käytetään väärin. Varsinkin kasvojen tunnistamiseen liittyen nähdään uhkakuvia, joita dataan liittyy.

Myös sillä on merkitystä, mihin dataa tullaan käyttämään tulevaisuudessa. On olemassa hyviä ja huonoja käyttötarkoituksia, eikä tunnedataa haluta antaa mihin vain tarkoitukseen (vrt 2.3). Toisaalta useampi haastateltavista, joille multimodaalinen tunteiden tunnistaminen ei ollut ennestään tuttu aihe, pohti miksi tunnedatan kerääminen on ylipäätään tärkeää, miksi sitä pitäisi kerätä tai käsitellä, pitääkö kaikkea aina mitata?

Ihmisläheisyys

Haastateltavat kokivat, että multimodaalisesta tunteiden tunnistamisesta pitäisi olla jotain hyötyä itselle, jotta siihen suostuttaisiin. Sen pitäisi jotenkin auttaa itseä tai tuoda elämään jotain hyvää (vrt. 4.4). Useat haastateltavat sanoivat, että tekoälyjärjestelmät voivat tuoda paljon hyvää, mutta ne voivat aiheuttaa myös paljon harmia. Niillä on vaikutuksia niin yksilöiden elämään kuin yhteisöjekin toimintaan.

Eräs haastateltava (Mies, 42-vuotias, virkamies) pohti tekoälystä luotua mielikuvaa seuraavasti: ”Mielikuva tekoälystä tai uudesta teknologiasta on markkinoitu aina jonain voimaannuttavana ja hyödyllisenä. Jos se on voimaannuttava ja hyödyllinen, niin se voi olla myös päinvastoin.”

Osa haastateltavista pohti, kuinka tarkkaa data on ja minkälaiset taustatekijät siihen mahdollisesti vaikuttavat. Miten dataan vaikuttaa esimerkiksi väsymys tai aamuinen riitely? Yksilöillä voi olla myös taustalla jotain, jolloin he eivät näytä tunteita samalla tavalla kuin suurin osa ihmisistä (vrt. 3.5). Useampi haastateltava nosti esiin huolen siitä, että tekoälyjärjestelmillä ei ole mahdollisuuksia ottaa huomioon kaikkia taustatekijöitä, joita ihminen puolestaan pystyy. Eräs haastateltavista (Mies, 32-vuotias, opiskelija) pohti tunteiden ilmaisemista eri kulttuureissa seuraavasti: ”Meillä ovat samat tunteet, mutta eri kulttuureissa ne ilmaistaan hieman eri tavoin. Mikroeleet saattavat olla hieman tarkempia.”

Yksilölliset ja kulttuuriset erot vaikuttavat myös siihen, kuinka tietyssä tilanteessa reagoidaan (vrt. 3.5). Ihmiset saattavat reagoida samassa tilanteessa hyvinkin paljon toisistaan poiketen. Toiset saattavat myös muokata käytöstään enemmän tilanteen mukaan, kun toiset käyttäytyvät aina samalla tavalla. Tunteet ovat iso osa elämää. Vaikka muutoin saatetaan olla avoimia tiedon jakamiseen, niin tunteiden osalta saattaa tulla raja vastaan. Silloin mietitään tarkemmin suostutaanko tunteiden tulkitsemiseen. Tunne-datasta voi olla tunnistettavissa yksittäisen henkilön identiteetti, mutta myös muita yksityisiä tietoja.

Oikeudenmukaisuus

Multimodaalisten tunteiden tunnistamismenetelmien ja tekoälyyn käyttöön liittyy vahvasti ajatus, että sen tulisi olla reilu kaikille maailmassa ja että sen tulisi tuottaa reiluja lopputuloksia (vrt. 4.4). Pohdinnoissa korostuivat yhdenvertaisuus ja syrjimättömyys. Järjestelmien tulisi olla niin hyvin rakennettuja ja läpinäkyviä, että tiedetään miten niiden algoritmit toimivat. Hyvänä asiana pidettiin auditointeja, joilla voitaisiin estää ei-haluttujen elementtien pääseminen järjestelmiin.

Haastateltavat pohtivat, että joskus järjestelmiin voi päästä elementtejä, joiden ei tiedetä aiheuttavan esimerkiksi syrjintää, mutta ajan kanssa se saattaa vahvistua ja tulla silloin ilmi. Jonkinlaisella seurannalla voitaisiin varmistaa, että järjestelmiin ei pääse elementtejä, jotka johtaisivat eettisiin haasteisiin.

Haastateltavia mietitytti kysymykset, jotka liittyvät tekoälyjärjestelmän tekemisiin päätöksiin ja siihen, kuinka ne ovat ohjelmoitu. Haastatteluissa pohdittiin klassista asetelmaa, jossa tekoälyn tekemä päätös on toisen henkilön edun mukaista, mutta pahaksi toiselle. Kuka loppupeleissä tekee päätöksen, kuka on vastuussa? Tekoälyjärjestelmien opettamiseen liittyvistä teemoista keskeisimmät liittyivät vinoumiin ja vääristymiin (vrt. 4.1).

Haastateltavat kokivat tärkeiksi, että vinoumat tunnistetaan ja niihin reagoidaan. Järjestelmien kehitykseen suhtauduttiin erittäin vakavasti ja vinoumien kaltaiset haasteet halutaan minimoida.

Yksi haastateltavista (Mies, 42-vuotias, virkamies) pohti tekoälyjärjestelmien eettisyyttä seuraavasti: ”Vinoumat ja reiluus olisivat minulle kaikista kriittisimmät, koska tekoälyjärjestelmät ovat vain niin vinoumattomia kuin mitä ne ovat opetettu, eivätkä ne pysty tunnistamaan tietojoukkonsa ulkopuolelta.” Vinoumien lisäksi haastateltavat pohtivat virheen mahdollisuutta ja sitä, kuinka voidaan varmistaa datan paikkansa pitävyys. Tunteet voidaan ajatella hyvinkin subjektiivisena kokemuksena, joten kuka on oikea henkilö määrittämään mikä määritelmä on oikein millekin tunteelle?

Tekoälyjärjestelmien kehityksen pitää olla aina ihmiskeskeistä. Järjestelmät eivät voi korvata ihmistä, mutta ne voivat tukea heitä (ks. 4.3). Tekoälyjärjestelmiä käytettäessä tulee kuitenkin aina olla ihminen mukana, joka ohjaa järjestelmiä ja tarkistaa niiden tekemät päätökset. Eräs haastateltava (Mies, 42-vuotias, virkamies) pohti ihmisen ja koneen välistä suhdetta seuraavasti: ”Pitää olla aina ihminen silmukassa, koska muutoin päädytään tilanteeseen, jossa opetettu tekoäly oli opetettu tekoälyn toimesta. Se vaikuttaa huonolta silmukalta olla.”

6 Pohdinta ja johtopäätökset

Opinnäytetyön tarkoituksena oli tutkia multimodaalisen tunteiden tunnistamisen eettisiä näkökulmia opetus- ja koulutusalan näkökulmasta. Tavoitteena oli selvittää mitä signaaleita/tekniikoita on, mitä tunnedataa niillä saadaan kerättyä, millaisia mahdollisuuksia ja toisaalta myös haasteita nähdään, sekä millaisia asioita tulee huomioida eettisistä näkökulmista. Aloitin tutkimuksen tutustumalla teoriaan tekoälystä ja tekoälyn etiikasta. Sen jälkeen syvensin teoriaan pohjautuvaa tietämystäni tutkimalla affektiivista tietojenkäsittelyä ja multimodaalista tunteiden tunnistamista. Pyrin ymmärtämään haastatteluilla, kuinka multimodaalinen tunteiden tunnistaminen ja siihen liittyvät eettiset kysymykset koetaan. Analysoidessani tutkimusaineistoa perehdyin teoriaan vielä syvemmin tutkien jotain aiheita entistä syvällisemmin.

6.1 Johtopäätökset

Mitä signaaleita tai teknologioita tunteiden tunnistamiseen on olemassa? Multimodaalisessa tunteiden tunnistamisessa voidaan käyttää monia signaaleita ja teknologioita, joista merkittäviä ovat ilmeiden- ja puheentunnistaminen sekä aivosähkökäyrä. Varsinkin ilmeidentunnistusta ja siinä hyödynnettävää kasvojentunnistusta on tutkittu paljon. Tutkimusten perusteella ilmeistä on mahdollista tunnistaa tunteita (Dilmegani 2022). Tosin on hyvä huomioida, että ilmeet eivät aina kuvasta aitoa tunnetta, joten tunteiden tunnistaminen ilmeiden perusteella voidaan kyseenalaistaa (ARTICLE 19 2021). Haastatteluissa yhdeksi merkittäväksi modaliteetiksi mainittiin henkilön oma kokemus, jonka avulla kerättyä dataa voitaisiin täydentää.

Mitä tunnetietoja saadaan kerättyä? Tunneteorian perusajatus on, että jotkin kategoriset tunteet olisivat perustavanlaatuisempia ja yleismaailmallisempia kuin toiset ja että työskentely niiden parissa on järkevää, sillä käytämme niitä puhuessamme tunteista (Mohammad heinäkuu 2021, 6-10 min). Tutkimuksessa perustunteet toistuivat useissa teorialähteissä ja haastatteluissa. On olemassa myös muita malleja tunteiden luokitteluun ja niiden käyttäminen voi olla järkevää, joidenkin tekniikoiden kohdalla.

Minkälaisia mahdollisuuksia tekoäly tarjoaa multimodaaliseen tunteiden tunnistamiseen opetus- ja koulutusallalla? Mahdollisuuksia nähdään esimerkiksi yksilöllisemmän opetuksen mahdollistamisessa. Kansainvälisen ihmisoikeusjärjestön Article 19 mukaan (ARTICLE 19 2021) monet koulutusteknologiaa markkinoivat yritykset toistavat oletetun syy-yhteyden opetetun sisällön ja opiskelijan ulkoisen kiinnostuksen välillä mahdollistavan yksilöllisemmän oppimisen saavuttamisen.

Haastatteluiden perusteella asiaan suhtaudutaan ainakin vielä toistaiseksi kriittisesti, sillä multimodaalisen tunteiden tunnistamisen hyödyistä ei ole vielä juuri tutkimusnäyttöä.

Yksilöllisemmän opetuksen lisäksi haastatteluissa mahdollisuuksiksi pohdittiin virtuaalitodellisuutta ja oppimisprosessin tutkimista.

Minkälaisia haasteita multimodaaliseen tunteiden tunnistamiseen liittyy? Keskeisiä haasteita datan käsittelyyn liittyen ovat affektiivinen kuilu, subjektiivisuushaaste, datan epätäydellisyys, tunnisteiden kohina tai puuttuminen sekä modaliteettien välinen epäjohdonmukaisuus (Zhao ym. 2021). Tekoälyjärjestelmät saattavat esimerkiksi olla syrjiviä (Dastin 2018) tai rasistisia (Schwartz 2019). Sekä teorialähteissä että haastatteluissa esiin nousivat haasteet, jotka pohjautuvat käytettyyn harjoitusdataan ja algoritmeihin. Esimerkiksi ilmeidentunnistuksessa tarkkuus on merkittävässä roolissa, jotta ihon väri tai etninen tausta eivät johda syrjintään (Vemou & Horvath 2021). Myös haastateltavat pohtivat näitä samoja kysymyksiä. Lisäksi haastatteluissa keskeisiksi haasteiksi koettiin yksityisyyteen liittyvät kysymykset, datan päätyminen vääriin käsiin ja järjestelmien väärinkäyttö. Haasteisiin pyritään vastaamaan lainsäädännöllä, kuten Euroopan Unionin tekoälysäädöksellä, joka koetaan hyväksi asiaksi.

Minkälaisia asioita tulee huomioida eettisistä näkökulmista? Euroopan komission määrittelemät (Euroopan komissio 2022) keskeiset näkökohdat ovat ihmisen toimijuus, ihmisläheisyys, oikeudenmukaisuus ja perusteltujen valintojen tekeminen. Haastatteluissa multimodaaliseen tunteiden tunnistamiseen liittyvät keskeiset eettiset kysymykset liittyvät oikeudenmukaisuuteen, syrjimättömyyteen ja yksityisyyteen.

Näiden näkökohtien ja haastatteluiden välillä yhtymäkohtia on nähtävissä ennen kaikkea oikeudenmukaisuuden kohdalla. Molemmissa painotetaan, että kaikilla tulee olla yhdenmukaiset mahdollisuudet ja että järjestelmien tulee olla syrjimättömiä. Muita yhtä selkeitä yhtymäkohtia ei ole havaittavissa.

6.2 Keskeiset tulokset

Multimodaalisessa tunteiden tunnistamisessa oman tunnetilan luokittelua voitaisiin hyödyntää yhtenä modaliteettina. Jos sovellusta käytetään omasta tahdosta sellaiseen tarkoitukseen, josta on itselle hyötyä, niin oman tunnetilan luokittelu voisi olla melko paikkansa pitävä menetelmä. En oikeastaan löytänyt tieteellistä tutkimusta oman tunnetilan hyödyntämisestä multimodaalisessa tunteiden tunnistamisessa, joten aihetta voisi tutkia enemmän. Omia tunteita voisi olla mahdollista luokitella esimerkiksi perustunteiden mukaisesti, jotka ovat selkeästi määriteltävissä.

Pohdittaessa, tulisiko multimodaalisia tunteiden tunnistamisen menetelmiä hyödyntäviä järjestelmiä ottaa ylipäätään käyttöön opetus- ja koulutusalailla, tulee arvioida järjestelmien hyötyjen ja haittojen suhdetta. Koen mielenkiintoisina tutkimukset, joista selviäisi millaisia tuloksia maailmalla saadaan näiden järjestelmien käytön vaikutuksista oppimistuloksiin ja opetuksen tehokkuuteen. Kiinnostavia olisivat myös tutkimustulokset siitä, millaisia muita vaikutuksia näillä järjestelmillä on esimerkiksi oppilaiden vuorovaikutustaitoihin ja sosiaalisiin tunteitaitoihin. Meillä Suomessa opetus- ja koulutus on hyvin pitkälti maksutonta, joka vaikuttaa osaltaan suhtautumiseemme opetukseen. Voi olla, että olisimme lähtökohtaisesti valmiimpia panostamaan enemmän oppimistuloksia parantavien järjestelmien käyttöönottoon, jos maksaisimme opetuksesta suuria summia samalla tavalla kuin joissain muissa yhteiskunnissa. Jos vaakakuppiin asetetaan oppimistulokset ja yksityisyys, niin kumpi on meille tärkeämpää.

Opetus- ja koulutusalailla järjestelmien käyttöönottoon ja riskien arviointiin on syytä kiinnittää erityisen paljon huomiota. Toisaalta vielä ei välttämättä ole tietoa siitä, mihin kaikkeen tekoälyjärjestelmät pystyvät esimerkiksi viiden vuoden päästä. Viittasin luvussa 2.3 yhdysvaltalaisen professori Kosinkin tutkimukseen (Levin 2017), jonka mukaan kourallisesta kuvista voisi päätellä henkilöstä hyvinkin henkilökohtaisia tietoja liittyen esimerkiksi poliittisiin näkemyksiin, älykkyyssosamäärään ja alttiuteen rikolliseen käyttäytymiseen. On tämä totta tai ei, niin se on jo ajatuksena pelottava. Kuinka varmistetaan, että järjestelmiä käytetään vain käyttötarkoituksen mukaisesti ja henkilön suostumuksella?

Puhuttaessa tekoälyjärjestelmistä data on hyvin keskeisessä asemassa. Voisi tiivistää, että data vaikuttaa kaikkeen. Jotta tekoälyjärjestelmiin voidaan suhtautua luottavaisesti, niiden tulee olla hyvin ohjelmoituja, niiden käyttötarkoitus tulee olla hyvin perusteltu ja niissä tulee huomioida tietoturvaan liittyvät kysymykset. Dataa käsiteltäessä on tärkeää huomioida, mihin data on tallennettu, ketkä sitä käsittelevät ja miten estetään datan väärinkäyttö. Näihin asioihin otetaan hyvin laajasti kantaa yleisessä tietosuojasetuksessa. Vaikka lainsäädäntö koetaan hyväksi asiaksi, saattaa se rajoittaa myös näiden järjestelmien käyttöä, jolloin niitä ei välttämättä oteta käyttöön oppilaitoksessa siinä laajuudessa kuin se olisi mahdollista.

Tekoälyjärjestelmien ohjelmointiin käytettävällä harjoitusdatalla on suuri merkitys, jotta voidaan varmistaa, että järjestelmät ovat oikeudenmukaisia ja syrjimättömiä. Tunnettu data on hyvin sensitiivistä ja se on usein liitännäinen tiettyyn tilanteeseen. Jotta dataa voidaan tulkita oikein, tulee tuntea monia taustatekijöitä, jotka vaikuttavat dataan.

Tutkimukseeni lukemassani teoriassa en huomannut, että niissä olisi otettu kantaa siihen, millainen koulutus henkilöllä tulisi olla, jotta hänellä olisi tarvittava ammattitaito tehdä päätelmiä multimodaalisen tunteiden tunnistamisen menetelmien avulla kerätyn datan perusteella.

Haastatteluissa tämä näkökulma kuitenkin pohditutti ja asiaan kaivattiin jonkinlaista yhtenäistä linjaa.

Kyse on niin henkilökohtaisesta datasta, ettei sitä mielellään jaeta mihinkään yleishyödylliseen tai vasta seuraavia opiskelijoita hyödyttäviin tarkoituksiin. Tässä kohtaa täytyy tosin todeta, että tutkimuksessani pääpaino oli nimenomaan ihmisen fyysisten ja psyykkisten muutosten avulla tunnistettavista tunteista. Tutkimuksessani ei selvinnyt koettaisiinko esimerkiksi tunteiden tunnistaminen tekstistä, kuten palautetuista tehtävistä, vähemmän henkilökohtaiseksi ja oltaisiinko siihen suostuvaisempia.

6.3 Tutkimuksen luotettavuus

Lähdin tutkimaan aihetta noviisina, jolla ei ollut oikeastaan aikasempaa kokemusta tekoälystä tai multimodaalisesta tunteiden tunnistamisesta. Opinnäytetyö on ollut vaiherikas projekti, joka on vaatinut useiden laajojen kokonaisuuksien opiskelemista varsin lyhyessä ajassa. Jokainen projektin vaihe on vienyt suunniteltua enemmän aikaa, joten aika onkin tuntuvan loppuvan jatkuvasti kesken. Olen jäsentänyt työtä matkan varrella moneen kertaan, jotta se olisi johdonmukainen ja selkeä. Teoriaosuuden koostamista ja täydentämistä olisi voinut jatkaa pidemmällekin.

Olen koostanut teoriaosuuden hyödyntämällä monipuolisesti erilaisia lähteitä. Affektiivista tietojenkäsittelyä käsittelevässä luvussa olen hyödyntänyt julkisia internet-sivuja koostaakseni asiasta mahdollisimman selkeän yleiskuvauksen, jonka jälkeen olen syventynyt syvemmin multimodaaliseen tunteiden tunnistamiseen. Varsinkin multimodaalista tunteiden tunnistamista käsittelevässä luvussa olen perehtynyt useisiin tuoreisiin tieteellisiin artikkeleihin, joten tutkimukseni tuotosta voidaan pitää luotettavana ja uskottavana.

Empiiriseen osuuteen valitsin haastateltaviksi henkilöitä, joilla tiesin etukäteen olevan annettavaa aiheeseen liittyen. Haastateltaviksi ei valikoitunut yhtään opettajaa, mutta haastateltavien joukossa oli muita henkilöitä, joilla oli muunlaisia kytköksiä opetus- ja koulutusalaan. Koen, että sain haastatteluiden perusteella kattavan käsityksen aiheesta. Haastateltavia oli kahdeksan, joten haastatteluiden perusteella muodostettu käsitys ilmiöstä on varsin suppea, eikä sen perusteella ole tehtävissä yleistyksiä kovinkaan laajassa mittakaavassa.

Onnistuin mielestäni tutkimuksessani ja sain vastauksia tutkimuskysymyksiin. Teoria ja empiria tukevat hyvin toisiaan, jolloin lopputulos on selkeä kokonaisuus. Tutkimuksen rajauksen mukaisesti mukana ovat kaikki koulutusasteet varhaiskasvatuksesta jatkuvaan oppimiseen, eikä mitään rajattu ulkopuolelle.

Tutkimuksessa perehdyttiin millä tavoin multimodaalisen tunteiden tunnistamisen menetelmiä hyödynnetään maailmalla, joka keskittyy Kiinaan, sillä siellä olla asian parissa pitkällä. Lopputuloksena on käsitys, kuinka multimodaalisen tunteiden tunnistamisen eettiset näkökulmat koetaan opetus- ja koulutusosalalla.

6.4 Suositukset ja kehittämismahdollisuudet

Toivon tutkimuksen herättävän kiinnostusta multimodaaliseen tunteiden tunnistamiseen, jotta aiheesta voidaan käydä enemmän julkista keskustelua. Keskustelun myötä on mahdollisuus lisätä tietoisuutta aiheen parissa, joka on mielestäni tärkeää uusia teknologioita kehitettäessä ja otettaessa käyttöön. Mikäli multimodaalisen tunteiden tunnistamisen menetelmiä otetaan käyttöön opetus- ja koulutusosalalla, vaatii se jonkinlaista tietoutta aiheesta kaikille, jotka jollakin tavalla olisivat sidoksissa näihin järjestelmiin. Silloin tulisi nimittäin ajankohtaiseksi päättää suostuisiko multimodaaliseen tunteiden tunnistamiseen vai ei. Käytännössä se koskisi siis hyvin laajasti oppijoita, huoltajia, opettajia, tutkijoita ja järjestelmiä kehittäviä tahoja.

Ehdotan jatkotutkimusta multimodaalisen tunteiden tunnistamisen mahdollisuuksista ja haasteista korkeakouluopiskelijoille. Uskon, että multimodaalista tunteiden tunnistamista hyödyntävät järjestelmät tullaan ottamaan käyttöön opetus- ja koulutusosalalla jossain vaiheessa, joten lisätutkimusta aiheen parista tarvittaisiin. Rajaisin jatkotutkimuksen nimenomaan aikuisille, jotka voivat itse tehdä päätöksiä käyttävätkö he näitä järjestelmiä vai eivät. Erittäin mielenkiintoiseksi suuntaukseksi näen virtuaalitodellisuuden hyödyntämisen opetuksessa, joten voisi tutkia, miten siinä voisi hyödyntää multimodaalista tunteiden tunnistamista.

Lähteet

Abbaschian, B., Sierra-Sosa, D. & Elmaghraby, A. 2021. Deep Learning Techniques for Speech Emotion Recognition, from Databases to Models. Luettavissa: <https://www.mdpi.com/1424-8220/21/4/1249>. Luettu: 14.5.2023.

AI Forum s.a. Tutkimusyhteistyötä tekoälyn saralla. Luettavissa: <https://www.aiforum.fi/>. Luettu: 8.3.2023.

ARTICLE 19 2021. Emotional entanglement: China's emotion recognition market and its implications for human rights. Luettavissa: <https://www.article19.org/wp-content/uploads/2021/01/ER-Tech-China-Report.pdf>. Luettu: 9.3.2023.

Banafa, A. 2016. What is Affective Computing? Luettavissa: <https://www.bbvaopenmind.com/en/technology/digital-world/what-is-affective-computing/>. Luettu: 18.5.2023.

Buolamwini, J. & Gebru, T. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Luettavissa: <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>. Luettu: 27.4.2023.

Burnwal, S. 2020. Speech Emotion Recognition. Luettavissa: <https://www.kaggle.com/code/shivamburnwal/speech-emotion-recognition>. Luettu: 18.5.2023.

Dastin, J. 2018. Amazon scraps secret AI recruiting tool that showed bias against women. Luettavissa: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>. Luettu: 26.4.2023.

Datang. 2022. 10 Open Source Multimodal Emotion Recognition Datasets. Luettavissa: [10 Open Source Multimodal Emotion Recognition Datasets | by Datang | Medium](https://www.datang.com/10-Open-Source-Multimodal-Emotion-Recognition-Datasets-by-Datang-Medium). Luettu: 24.4.2023.

Dilmegani, C. 2022. Affective Computing: In-Depth Guide to Emotion AI in 2023. Luettavissa: <https://research.aimultiple.com/affective-computing/>. Luettu: 9.3.2023.

Dilmegani, C. 2023. Top 24 Affective Computing (Emotion AI) Use Cases in 2023. Luettavissa: <https://research.aimultiple.com/affective-computing-applications/>. Luettu: 9.3.2023.

Du, S. Tao, Y. & Martinez, A. 2014. Compound facial expressions of emotion. Luettavissa: <http://dx.doi.org/10.1073/pnas.1322355111>. Luettu: 12.5.2023.

Paul Ekman Group LLC s.a.a Facial Action Coding System. Luettavissa:

<https://www.paulekman.com/facial-action-coding-system/>. Luettu: 18.5.2023.

Paul Ekman Group LLC s.a.b Universal Emotions. Luettavissa:

<https://www.paulekman.com/universal-emotions/>. Luettu: 20.2.2023.

Electronic Identification. 2022. Facial Recognition: how it works and its safety. Luettavissa:

<https://www.electronicid.eu/es/blog/post/face-recognition/en>. Luettu: 9.5.2023.

Entropik Tech s.a. Want to measure tutors' performance and the impact of online classroom experience? Luettavissa: <https://entropiktech.com/edtech-content-testing-solutions/>. Luettu: 22.5.2023.

Euroopan komissio 2022. Tekoälyn ja datan käyttö opetuksessa ja oppimisessa – eettiset ohjeet opettajille. Euroopan unionin julkaisutoimisto. Luxemburg. Luettavissa:

<https://data.europa.eu/doi/10.2766/153756>. Luettu: 23.2.2023.

Euroopan komissio s.a. Yleinen tietosuoja-asetus. Luettavissa:

https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/index_fi.htm. Luettu: 27.4.2023.

Europeiska datatillsynsmannen 2019. Facial recognition in school renders Sweden's first GDPR fine. Luettavissa: https://edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_sv. Luettu: 10.5.2023.

Gillis, A. s.a. voice recognition (speaker recognition). Luettavissa:

<https://www.techtarget.com/searchcustomerexperience/definition/voice-recognition-speaker-recognition>. Luettu: 13.5.2023.

Gossett, S. 2023. Emotion AI: 3 Experts on the Possibilities and Risks. Luettavissa:

<https://builtin.com/artificial-intelligence/emotion-ai>. Luettu: 18.5.2023.

Helsingin yliopisto s.a. Eettinen tekoäly ja algoritmit. Luettavissa:

<https://www.helsinki.fi/fi/ajankohtaista/ihmisten-teknologia/eettinen-tekoaly-ja-algoritmit>. Luettu: 27.4.2023.

IBM s.a. What is computer vision? Luettavissa: <https://www.ibm.com/topics/computer-vision>.

Luettu: 18.5.2023.

- IGI Global s.a. What is Automatic Emotion Recognition. Luettavissa: <https://www.igi-global.com/dictionary/automatic-emotion-recognition-based-on-non-contact-gaits-information/59395>. Luettu: 18.5.2023.
- Jobin, A., Ienca, M. & Vayena, E. 2019. The global landscape of AI ethics guidelines. Luettavissa: <https://www.nature.com/articles/s42256-019-0088-2>. Luettu: 15.5.2023.
- Kaspersky s.a. Kasvojentunnistus – määritelmä ja selitys. Luettavissa: <https://www.kaspersky.fi/resource-center/definitions/what-is-facial-recognition>. Luettu: 10.5.2023.
- Kaulard, K., Cunningham, DW., Bühlhoff, HH. & Wallraven, C. 2012. The MPI facial expression database—A validated database of emotional and conversational facial expressions. PLoS ONE 7(3). Luettavissa: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0032321#abstract0>. Luettu: 15.5.2023.
- Ko, B. 2018. A Brief Review of Facial Emotion Recognition Based on Visual Information. Luettavissa: <http://dx.doi.org/10.3390/s18020401>. Luettu: 12.5.2023.
- KPMG Oy Ab s.a. Euroopan komission ehdotus uudeksi... Luettavissa: <https://kpmg.com/fi/fi/home/Pinnalla/2021/04/euroopan-komission-ehdotus-uudeksi-tekoalasetukseksi-julkaistiin.html>. Luettu: 27.4.2023.
- Levin, S. 2017. Face-reading AI will be able to detect your politics and IQ, professor says. Luettavissa: [Face-reading AI will be able to detect your politics and IQ, professor says | Facial recognition | The Guardian](https://www.theguardian.com/technology/2017/mar/30/face-reading-ai-will-be-able-to-detect-your-politics-and-iq-professor-says). Luettu: 30.3.2023.
- Li, X., Zhang, Y., Tiwari, P., Song, D., Hu, B., Yang, M., Zhao, Z., Kumar, N. & Marttinen, P. 2022. EEG based Emotion Recognition: A Tutorial and Review. Luettavissa: [https://www.researchgate.net/publication/359491630 EEG based Emotion Recognition A Tutorial and Review](https://www.researchgate.net/publication/359491630_EEG_based_Emotion_Recognition_A_Tutorial_and_Review). Luettu: 13.5.2023.
- Mathur, V. 2022. Multi-Modal Learning: Types, Benefits & Examples. Luettavissa: <https://www.analyticssteps.com/blogs/multi-modal-learning-types-benefits-examples>. Luettu: 20.5.2023.
- The MathWorks, Inc. s.a. What Is Deep Learning? Luettavissa: <https://www.mathworks.com/discovery/deep-learning.html>. Luettu: 18.5.2023.

Mayo Foundation for Medical Education and Research s.a. EEG (electroencephalogram).

Luettavissa: <https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875>. Luettu: 18.5.2023.

Minsky, M. 1985. The Society of Mind. Simon & Schuster, Inc. New York. Luettavissa:

<http://www.acad.bg/ebook/ml/Society%20of%20Mind.pdf>. Luettu: 13.5.2023.

Mohammad, S. 2022. Ethics Sheet for Automatic Emotion Recognition and Sentiment Analysis.

Luettavissa: https://doi.org/10.1162/coli_a_00433. Luettu: 20.2.2023.

Mohammad, S. Heinäkuu 2021. Ethics Sheets for AI Tasks and a Case Study for Automatic Emotion Recognition. UBC Language Sciences. Video katsottavissa:

<https://www.youtube.com/watch?v=QXeUIJqEBbM>. Katsottu 20.2.2023.

Pan, J., Fang, W. Zhang, Z., Chen, B., Zhang, Z. Wang, S. 2023. Multimodal Emotion Recognition based on Facial Expressions, Speech, and EEG. IEEE Open Journal of Engineering in Medicine and Biology. IEEE. Luettavissa:

<https://ieeexplore.ieee.org/abstract/document/10026861>. Luettu: 13.5.2023.

Pekkarinen, A. 2021. Mitä tekoälyn etiikka tarkoittaa? Kolme syytä opetella perusasiat. Luettavissa:

<https://www.helsinki.fi/fi/uutiset/tekoaly/mita-tekoalyn-etiikka-tarkoittaa-kolme-syyta-opetella-perusasiat>. Luettu: 27.4.2023.

Puusa, A. & Juuti P. 2020. Laadullisen tutkimuksen näkökulmat ja menetelmät. Gaudeamus.

Helsinki. E-kirja. Luettu: 29.3.2023.

Rusanen, A-M., Nurminen, J., Räisänen, S., Tarkoma, S. & Halmetoja, S. s.a. Tekoälyn etiikka.

Helsingin yliopisto. Luettavissa: <https://ethics-of-ai.mooc.fi/fi/>. Luettu: 15.5.2023.

Schuller D. & Schuller B. 2018. The Age of Artificial Emotional Intelligence. THE IEEE COMPUTER SOCIETY. Luettavissa:

<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8481266>. Luettu: 13.5.2023.

Schwartz, O. 2019. In 2016, Microsoft's Racist Chatbot Revealed the Dangers of Online Conversation The bot learned language from people on Twitter—but it also learned values.

Luettavissa: <https://spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation>. Luettu: 26.4.2023.

- Sebe, N., Cohen, I. & Huang, T. 2005. Handbook of Pattern Recognition and Computer Vision, pp. 387-409 (2005). 3. painos. Luettavissa: https://doi.org/10.1142/9789812775320_0021. Luettu: 24.4.2023.
- Somers, M. 2019. Emotion AI, explained. Luettavissa: <https://mitsloan.mit.edu/ideas-made-to-matter/emotion-ai-explained>. Luettu: 9.5.2023.
- Stöckli, S., Schulte-Mecklenbeck, M., Borer, S. & Samson, A. 2017. Facial expression analysis with AFFDEX and FACET: A validation study. Luettavissa: <https://doi.org/10.3758/s13428-017-0996-1>. Luettu: 18.5.2023.
- Subramanian, R., Wache, J., Abadi, M., Vieriu, R., Winkler, S. & Sebe, N. 2018. ASCERTAIN: Emotion and Personality Recognition Using Commercial Sensors. Luettavissa: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7736040>. Luettu: 13.5.2023.
- Trezise, K. 2017. Emotions in classrooms: The need to understand how emotions affect learning and education. Luettavissa: <https://neurosciencecommunity.nature.com/posts/18507-emotions-in-classrooms-the-need-to-understand-how-emotions-affect-learning-and-education>. Luettu: 13.5.2023.
- Vemou, K. & Horvath, A. 2021. Facial Emotion Recognition. Luettavissa: https://edps.europa.eu/system/files/2021-05/21-05-26_techdispatch-facial-emotion-recognition_ref_en.pdf. Luettu: 12.5.2023.
- Vesisenaho, M., Juntunen, M., Häkkinen, P. & Pöysä-Tarhonen, J. 2019. Virtual Reality in Education: Focus on the Role of Emotions and Physiological Reactivity. Luettavissa: https://www.researchgate.net/publication/330928604_Virtual_Reality_in_Education_Focus_on_the_Role_of_Emotions_and_Physiological_Reactivity. Luettu: 14.5.2023.
- Xiaoming, Z, Yijiao, Y. & Shiqing, Z. 2021. Survey of Deep Learning Based Multimodal Emotion Recognition. Luettavissa: <http://fcst.ceaj.org/EN/10.3778/j.issn.1673-9418.2112081>. Luettu: 14.5.2023.
- Zeng, Z., Pantic, M., Roisman, G. & Huang, T. 2009. A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. Luettavissa: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4468714>. Luettu: 13.5.2023.
- Zhang, S. 2019. Schools using facial recognition system sparks privacy concerns in China. Luettavissa: <https://en.jmdedu.com/Article/52>. Luettu: 10.5.2023.

Zhang, Z., Cummins, N. & Schuller, B. 2017. Advanced Data Exploitation in Speech Analysis. Luettavissa: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7974862>. Luettu: 16.5.2023.

Zhao, S., Jia, G., Yang, J., Ding, G. & Keutzer, K. 2021. Emotion Recognition from Multiple Modalities: Fundamentals and Methodologies. IEEE SIGNAL PROCESSING MAGAZINE. Luettavissa: <https://arxiv.org/pdf/2108.10152.pdf>. Luettu: 20.2.2023.

Zhao, X., Yang, Y., Zhang, S. 2022. Survey of Deep Learning Based Multimodal Emotion Recognition. Journal of Frontiers of Computer Science and Technology, 2022, 16(7): 1479-1503. Luettavissa: <http://fcst.ceaj.org/EN/10.3778/j.issn.1673-9418.2112081>. Luettu: 24.4.2023.

Zhao, S., Yao, X., Yang, J., Jia, G., Ding, G., Chua, T., Schuller, B. & Keutzer, K. 2022. Affective Image Content Analysis: Two Decades Review and New Perspectives. Luettavissa: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9472932>. Luettu: 14.5.2023.

Liitteet

Liite 1. Saatekirje

Saatekirje

Hei,

teen tutkimusta multimodaalisesta tunteiden tunnistamisesta osana opinnäytetyötäni Haaga-Helian ammattikorkeakoulussa. Teen opinnäytetyötä toimeksiantona viiden suomalaisen korkeakoulun monitieteelliselle yhteistyöverkostolle, AI Forumille (www.aiforum.fi). Opinnäytetyön on määrä olla kokonaisuudessaan valmis toukokuun loppuun mennessä, jolloin se tullaan tallentamaan Theseus - verkkokirjastoon.

Haastatteluilla pyrin ymmärtämään, kuinka multimodaalinen tunteiden tunnistaminen ja siihen liittyvät eettiset näkökulmat koetaan opetus- ja koulutusalan näkökulmasta. Lisäksi olen kiinnostunut millaisia mahdollisuuksia ja toisaalta myös haasteita nähdään. Haastattelu kestää noin 30-45 minuuttia ja se olisi tarkoitus nauhoittaa. Säilytän tallenteita Haaga-Helian Office 365 -palveluissa.

Tutkimuksen aikana kerättyjä henkilötietoja käsitellään luottamuksellisesti, eikä niitä käytetä muuhun tarkoitukseen kuin opinnäytetyön tekemiseen. Haastattelujen tallenteet poistetaan opinnäytetyön julkaisun jälkeen.

Vastaan mielelläni kaikkiin tutkimustani ja opinnäytetyötäni koskeviin kysymyksiin.

Ystävällisin terveisin,
Jenna Kauranen
Tradenomi-opiskelija

Liite 2. Haastattelukysymykset

Haastattelurunko

1. Sopiiko, että nauhoitan haastattelun ja käytän sen sisältöä tutkimuksessani?
2. Kuka olet (sukupuoli, ikä, ammatti/nimike)? Viitataan opinnäytetyössä haastatteluihin ilman nimitietoja (esim. nainen, 00-vuotta, opiskelija).
3. Avaan tässä mitä multimodaalisella tunteiden tunnistamisella tarkoitetaan.
4. Onko aihe ”multimodaalinen tunteiden tunnistaminen” sinulle ennestään tuttu?
5. Onko sinulla kokemuksia tunteiden tunnistamisesta eri menetelmien avulla?
6. Millaisia eettisiä kysymyksiä olet pohtinut?
7. Onko eri tunteiden tunnistamisen menetelmillä eroja eettisistä näkökulmista tarkasteltaessa?
8. Missä tilanteissa voisit ajatella suostuvasi tunteiden tulkitsemiseen?
9. Milloin tai missä tilanteessa et haluaisi tunteitasi tulkittavan?
10. Miten multimodaalinen tunteiden tunnistaminen voisi auttaa kehittämään opetus- ja koulutusala?
11. Millaisia muita mahdollisuuksia näet multimodaalisella tunteiden tunnistamisella tulevaisuudessa?
12. Millaisia haasteita multimodaaliseen tunteiden tunnistamiseen liittyy?
13. Kuinka haasteita voidaan hallita?
14. Tuleeko sinulle muuta mieleen aiheesta, jonka haluaisit kertoa?