



Serendip - Immersiivinen ympäristö ja vuorovaikutusjärjestelmä:

Teknologian ja inhimillisen kokemuksen rajapinnalla historiallisessa
kontekstissa

Jussi Wright

Digitaaliset ihmistieteet

Helsingin Yliopisto

Opinnäytetyö - luonnos

2024 (v5.6)

1. Johdanto	3
1.1. Tutkimuksen tausta ja merkitys	3
1.2. Tutkimusongelma ja -kysymykset	5
1.3. Tutkimuksen tavoitteet ja hypoteesit	6
1.4 Mitä on avatar?	7
2. Tutkimuksellinen viitekehys	8
2.1. Keskeisten teorioiden ja käsitteiden esittely	8
2.1.1 Teknologisen viestinnän historiallinen kehitys	8
2.1.2. Posthumanismi ja identiteetti	9
2.1.3. Rationaalisuus, päätöksenteon logiikka ja eettiset näkökulmat	11
2.1.4. Teknologiakriittiset näkemykset ja ihmiskeskeinen suunnittelu	13
2.1.5 Nonverbaalinen viestintä immersiiivisissä 3D-ympäristöissä	14
2.1.6 Avatar-tekniikan historiallinen kehitys ja sen sovellukset	22
2.2. Immersiiviset 3D-tekniikat ja eettiset kysymykset	24
2.2.1. Moraalinen vastuu ja käyttäjäkokemus	24
2.2.2. Intersubjektiivisuus ja kulttuuriset normit	26
2.2.3. Itsekriittisyys ja varovaisuusperiaate	27
2.2.4. Autonomian paradoksi tekoälyavustajissa	28
2.2.5. Ihmiskeskeinen suunnittelu ja tekoälyn rajat	29
2.2.6. Eettiset ja yhteiskunnalliset haasteet autonomisen tekoälyn integroinnissa	30
2.2.7. Eettiset kysymykset ja luottamus: Avatarien vaikutus eri toimialoilla	31
2.2.8 Tekoälyn läpinäkyvyys ja käyttäjien valistaminen	32
2.2.9 Käyttäjien yksityisyys, datan kerääminen ja tekoälyn päätöksenteon läpinäkyvyys	32
2.2.10 Laajempi yhteiskunnallinen ja kulttuurinen konteksti	33
2.3. Aiempi tutkimus ja kirjallisuuskatsaus	34
2.3.1. Dialogijärjestelmien, 3D-avatarien ja tekoälyn teknologinen tausta	35
2.3.2. Tekniikan nykytila: Erittäin interaktiiviset dialogijärjestelmät (HIDS)	36
2.3.3 Posthumanistiset näkökulmat	38
2.4. Johtopäätökset ja suositukset	38
2.4.1. Suositukset	38
3. Tutkimusmenetelmät	39
3.1 Tutkimusstrategia	39
3.2 Kirjallisuuskatsauksen aineiston keruu	40
3.3 Aineiston analysointimenetelmät	42
3.3.1 Kirjallisuuskatsauksen analyysi	42
3.3.2 Tapaustutkimuksen analyysi	42
3.4 Tutkimusmenetelmien eettiset näkökulmat	43
3.4.1 Kirjallisuuskatsauksen eettisyys	43

3.4.2 Tapaustutkimuksen eettisyys	43
3.4.3 AI-työkalujen käytön eettisyys	44
3.6 Tutkimussuunnitelma	45
4. Tapaustutkimus: Kehitetty immerstiivinen järjestelmä	46
4.1 Kehityksen tausta ja tarkoitus	46
4.2 Järjestelmän kuvaus	47
4.3 Järjestelmän analyysi suhteessa tutkimuskysymyksiin	48
4.4 Eettiset näkökohdat ja käyttäjäkokemus	49
5. Tulokset	50
5.1 Immerssiivisten 3D-avatarien vaikutus digitaalisen vuorovaikutuksen normeihin	50
5.2 Käyttäjien luottamus ja turvallisuuden tunne	52
5.3 Eettiset kysymykset ja haasteet	53
5.4 Yhteenveto keskeisistä löydöksistä	53
6. Tulosten tulkinta suhteessa tutkimuskysymyksiin ja viitekehykseen	55
6.2 Tutkimuksen rajoitukset	57
6.3 Johtopäätökset ja suositukset	58
6.4 Tulevat tutkimussuunnat	58
Lähdeluettelo	60

1. Johdanto

1.1. Tutkimuksen tausta ja merkitys

Teknologian kehitys muokkaa jatkuvasti sitä, miten olemme vuorovaikutuksessa toistemme kanssa, ja immersiiiviset 3D-avatatit edustavat tätä murrosta konkreettisella tavalla. Nämä tekoälyohjatut hahmot eivät pelkästään jäljittele inhimillistä vuorovaikutusta, vaan ne muokkaavat perustavanlaatuisesti digitaalisen viestinnän normeja. Tämä tutkimus tarkastelee, miten immersiiiviset avatatit vaikuttavat käyttäjien käsityksiin teknologiavälitteisestä vuorovaikutuksesta ja millaisia eettisiä kysymyksiä niiden kehittyminen herättää.

Vaikka immersiiivisten järjestelmien käyttäjäkokemuksia on tutkittu jonkin verran, niiden kehittyminen yhä älykkäämmiksi — kykeneviksi suorittamaan monimutkaisia tehtäviä ja tekemään päätöksiä käyttäjän puolesta — nostaa esiin uusia kysymyksiä. Kuinka tällaiset järjestelmät voivat säilyttää käyttäjien luottamuksen ja turvallisuuden tunteen, kun ne ottavat entistä enemmän vastuuta vuorovaikutuksesta? Miten voidaan varmistaa, että nämä järjestelmät toimivat eettisesti ja moraalisesti kestäväällä tavalla?

Immersiiiviset 3D-avatatit ja tekoälyohjatut dialogijärjestelmät ovat nousseet keskeiseen rooliin digitaalisen viestinnän kehityksessä. Teknologian kehittyessä yhä realistisemmaksi korostuvat myös sen yhteiskunnalliset vaikutukset ja eettiset haasteet. Useat aiemmat tutkimukset ovat keskittyneet lähinnä tekniseen toteutukseen tai vuorovaikutuksen pelkkiin käytettävyyssnäkökulmiin. Historiallinen ulottuvuus sekä eettisten ja nonverbaalisten taiteellisten ilmaisutapojen tarkastelu samassa tutkimuskehyksessä on kuitenkin jäänyt vähemmälle huomiolle.

Tämä tutkimus täyttää aukon, jossa yhdistetään immersiiivisten teknologioiden vuorovaikutukselliset ja eettiset ulottuvuudet historialliseen kontekstiin sekä taiteen nonverbaalisen ilmaisun tutkimukseen. Suurin osa aiemmista tutkimuksista on keskittynyt joko puhtaasti tekniseen kehittämiseen tai eettisten kysymysten yleistasoiseen pohdintaan, ilman että näitä kahta kytketään vahvasti toisiinsa ja suhteutetaan aiempien vuorovaikutusteknologioiden historiaan. Näin ollen tämä työ tarjoaa uudenlaista näkökulmaa siihen, miten kehittyvät 3D-avatatit istuvat laajempaan viestinnän ja teknologian kehityslinjaan.

Tämän työn tulokset voivat hyödyttää sekä teknologian kehittäjiä että yhteiskunnallisia päättäjiä: on tärkeää ymmärtää, miten tekoälyohjattuja, nonverbaalia ilmaisua sisältäviä hahmoja tulee suunnitella ja säädellä, jotta ne palvelisivat ihmisen arvoja ja tarpeita kestäväällä tavalla.

Avatarien historia juontaa juurensa uskonnollisiin ja mytologisiin kertomuksiin. Antiikin Kreikassa jumalten inkarnaatiot ihmiskehossa, kuten Zeus tai Athene, kuvastivat ihmisen ja yliluonnollisen rajapintaa. Intiassa sanskritin sana “avatar” tarkoitti jumaluuden laskeutumista maalliseen muotoon, erityisesti Vishnun eri ilmentymissä. Nämä varhaiset tarinat korostivat, miten korkeamman voiman idea konkretisoitui käytännön tasolle.

1900-luvulla tieteiskirjallisuus ja kyberpunk-kulttuuri veivät avatar-käsitteen uuteen suuntaan. Esimerkiksi Isaac Asimovin robotit ja William Gibsonin kyberpunk-romaanit (esim. *Neuromancer*) loivat pohjan ajatukselle digitaalisten identiteettien ja koneiden yhdistämisestä. 2000-luvulla avatarit nousivat keskeisiksi elementeiksi digitaalisessa kulttuurissa, kuten verkkoyhteisöissä (esim. Second Life) ja videopeleissä (esim. World of Warcraft).

Nykyteknologiassa avatarien kehitys on edennyt yksinkertaisista visuaalisista edustajista kohti tekoälyllä ohjattuja, inhimillistä viestintää simuloivia vuorovaikutuskumppaneita. Tämän kehityksen taustalla on jatkuva tarve inhimillisen yhteyden ja vuorovaikutuksen simuloimiseen digitaalisten alustojen kautta. Avatarien nykyaikainen rooli ei siis ole vain viestinnällinen, vaan se vaikuttaa myös sosiaaliin ja kulttuuriin rakenteisiin, erityisesti tekoälyn aikakaudella.

Avainsanat: Immersiivisyys, Tekoäly (AI), Agentit, Suuret kielimallit (LLM), Ihmis–konevuorovaikutus (HCI), Käyttöliittymät, Näkökyky, Dialogijärjestelmät, 3D-hahmot, Tunneanalyysi, Posthumanismi, Eettisyys, Nonverbaalinen viestintä.

1.2. Tutkimusongelma ja -kysymykset

Pääkysymykset:

1. **Miten immersiiiviset 3D-avatatit muovaavat digitaalisia vuorovaikutusnormeja ja haastavat perinteisiä käsityksiä teknologiavälitteisestä vuorovaikutuksesta?**

Tämä kysymys on tärkeä, koska se paljastaa, miten uudet teknologiat uudelleenmäärittelevät vuorovaikutustapoja ja laajentavat digitaalisen viestinnän käytäntöjä perinteisten viestimien (esim. tekstichatit, videopuhelut) ulkopuolelle.

2. **Miten älykkäiden vuorovaikutusjärjestelmien kehittyminen vaikuttaa eettisiin ja moraalisiin periaatteisiin, ja miten tämä muokkaa käyttäjien kokemaa luottamusta ja turvallisuutta?**

Tämä on olennaista, koska teknologian ottaessa yhä suuremman roolin päätöksenteossa, syntyy tarve ymmärtää, missä kulkevat kehittäjien ja käyttäjien vastuut ja miten järjestelmä säilyttää inhimilliset arvot.

3. **Miten voimme tasapainottaa immersiiivisten teknologioiden tarjoamat mahdollisuudet niiden eettisten ja käytännöllisten haasteiden kanssa?**

Tämä kysymys on keskeinen, sillä teknologian hyödyt (kuten paremman sitoutumisen ja oppimisen) on sovitettava yhteen turvallisuuden, luottamuksen ja yhteiskunnallisen hyväksyttävyyden kanssa.

Sivuteemat:

- **Tunnetietoisuus:** Miten reaaliaikaista tunneanalyysiä voidaan hyödyntää immersiiivisissä dialogimenetelmissä, ja mitkä ovat sen hyödyt, haitat ja vaikutukset?
- **Näkökyky:** Miten 3D-hahmolle annettu, reaaliaikainen näkökyky kameran välityksellä muuttaa dialogikonseptia verrattuna perinteisiin näkökykyä omaaviin dialogijärjestelmiin?
- **Muisti:** Miten dialogijärjestelmien kyky muistaa käytyjä keskusteluja vaikuttaa vuorovaikutuksen laatuun ja jatkuvuuteen?
- **Historia:** Miten nykyiset immersiiiviset teknologiat ja niihin kohdistuva kritiikki suhteutuvat aiempiin teknologisiin innovaatioihin ja niiden herättämiin reaktioihin?

- **Vaikutukset:** Millaisia kulttuurisia ja yhteiskunnallisia vaikutuksia tällaisilla immersiiivisillä vuorovaikutusjärjestelmillä voi olla, ja miten ne muokkaavat käsitystämme teknologiavälitteisestä kommunikaatiosta?
- **Mahdollisuudet:** Millaisia uusia mahdollisuuksia immersiiiviset vuorovaikutusjärjestelmät voivat tarjota esimerkiksi koulutuksessa, terveydenhuollossa ja asiakaspalvelussa?

1.3. Tutkimuksen tavoitteet ja hypoteesit

Tämän tutkimuksen tavoitteena on syventää ymmärrystä siitä, miten immersiiiviset 3D-ympäristöt ja tekoälyohjatut 3D-hahmot vaikuttavat käyttäjien vuorovaikutuskokemuksiin sekä käsityksiin teknologiavälitteisestä kommunikaatiosta verrattuna perinteisiin tekstipohjaisiin järjestelmiin. Teknologian kehittyessä kohti yhä inhimillisempiä vuorovaikutusjärjestelmiä on olennaista ymmärtää näiden muutosten vaikutukset niin yksilö- kuin yhteiskuntatasollakin.

Tutkimus auttaa myös hahmottamaan, miten tällaiset teknologiat voivat muuttaa sosiaalisen kanssakäymisen normeja digitaalisessa ympäristössä ja miten ne suhteutuvat laajempiin yhteiskunnallisiin ja kulttuurisiin käsityksiin teknologian roolista ihmisten välisessä vuorovaikutuksessa.

Lisäksi tutkimus tukee eettisesti kestävien ja kulttuurisesti tiedostavien immersiiivisten vuorovaikutusjärjestelmien kehittämistä eri käyttötarkoituksiin. Tavoitteena on edistää luonnollisempaa ja inhimillisempää ihmisen ja koneen välistä vuorovaikutusta, tunnistuen ja hyödyntäen immersiiivisten teknologioiden positiivisia vaikutuksia sekä kriittisesti arvioiden niiden mahdollisia haasteita.

Hypoteesi

Ruumiillistetut keskusteluagentit eli 3D-hahmot virtuaalisissa ympäristöissä tarjoavat monipuolisemman ja syvemmän vuorovaikutuskokemuksen verrattuna tekstipohjaisiin käyttöliittymiin. Tämä perustuu niiden kykyyn hyödyntää sekä verbaalista että nonverbaalista viestintää, mikä luo luonnollisemman ja inhimillisemmän vuorovaikutuksen.

Oletan, että 3D-hahmot voivat tehostaa viestintää ja lisätä käyttäjän sitoutumista, tarjoten uusia mahdollisuuksia oppimiseen, terapiaan ja sosiaaliseen vuorovaikutukseen. Samalla niiden käyttö muokkaa käsitystämme vuorovaikutuksesta ja saattaa muuttaa sosiaalisen kanssakäymisen normeja digitaalisessa ympäristössä. Nämä muutokset ovat osa pidempää historiallista jatkumoa, jossa teknologiset innovaatiot ovat herättäneet sekä innostusta että kritiikkiä.

1.4 Mitä on avatar?

Avatar on käsite, jolla on pitkä ja monikerroksinen historia sekä kulttuurinen että teknologinen merkitys. Sana "avatar" juontaa juurensa sanskritin kielestä, jossa se tarkoittaa "laskeutumista" ja viittaa jumalan, erityisesti Vishnun, inkarnaatioon maalliseen muotoon. Tämä alkuperäinen merkitys korostaa ideaa jonkin korkeamman tai abstraktin ilmiön konkretisoitumisesta tai ilmentymisestä.

Nykykulttuurissa ja teknologiassa avatar on saanut uuden merkityksen, viitaten käyttäjän visuaaliseen tai digitaaliseen edustajaan virtuaalitodellisuudessa. **Sci-fi-kirjallisuus** ja -elokuvat ovat popularisoineet käsitteen, ja se on vakiintunut osaksi verkkokulttuuria, etenkin peleissä, nettiyhteisöissä ja sosiaalisessa mediassa. Teknologiassa avatarit voidaan jakaa kahteen pääkategoriaan:

1. **Passiiviset avatarit**, jotka ovat käyttäjän visuaalisia ilmentymiä (esim. pelihahmot tai verkkoyhteisöjen profiilit).
2. **Aktiiviset avatarit**, jotka toimivat tekoälyn ohjaamina itsenäisinä toimijoina (esim. asiakaspalvelubotit, uutistenlukijat tai jopa "thanabotit", jotka simuloivat edesmenneitä henkilöitä).

Avatarien merkitys tutkimuksessa:

Immersiiviset 3D-avatarit edustavat erityisen kehittynyttä avatar-tyyppiä, jossa yhdistyvät realistinen ulkomuoto, kyky kommunikoida tekoälyn avulla ja mahdollisuus mukautua käyttäjän nonverbaaliseen viestintään. Tämä tutkimus tarkastelee, miten tällaiset avatarit muokkaavat vuorovaikutusnormeja, vaikuttavat käyttäjän kokemukseen ja nostavat esiin uusia eettisiä kysymyksiä.

2. Tutkimuksellinen viitekehys

Tämän tutkimuksen tavoitteena on ymmärtää, miten immersiiiviset 3D-avatatit muovaavat digitaalisia vuorovaikutusnormeja ja nostavat esiin eettisiä kysymyksiä kehittyvissä vuorovaikutusjärjestelmissä. Tutkimuksen viitekehys yhdistää teknologiakritiikkiä, filosofisia näkökulmia ja viestinnän teorioita, tarjoten pohjan näiden järjestelmien roolin ja vaikutusten analyysille.

Ensimmäisessä osassa käsitellään keskeisiä teorioita ja käsitteitä, jotka liittyvät immersiiivisiin 3D-teknologioihin ja niiden vuorovaikutuksellisiin sekä filosofisiin ulottuvuuksiin. Toisessa osassa syvennyttään näiden teknologioiden eettisiin kysymyksiin ja pohditaan, miten ne vaikuttavat käyttäjäkokemuksiin sekä yhteiskunnallisiin normeihin.

2.1. Keskeisten teorioiden ja käsitteiden esittely

Immersiiiviset 3D-teknologiat ja tekoälyohjatut vuorovaikutusjärjestelmät edustavat teknologian kehityksen uutta vaihetta, jossa yhdistyvät viestintäteoriat, teknologiakritiikki ja filosofiset näkökulmat. Tämä osio esittelee tutkimuksen keskeiset teoriat ja käsitteet, jotka muodostavat pohjan vuorovaikutusjärjestelmien roolin ja vaikutusten analyysille.

2.1.1 Teknologisen viestinnän historiallinen kehitys

Teknologian kehitys on muokannut viestintäkäytäntöjämme kautta historian, vaikuttaen siihen, miten ihmiset ovat vuorovaikutuksessa keskenään. Marshall McLuhanin (1964) kuuluisa "väline on viesti" -periaate korostaa, että viestintävälineet eivät pelkästään välitä viestejä, vaan myös muokkaavat viestin sisältöä ja vastaanottajan kokemusta. Esimerkiksi sähkölennätin ja puhelin muuttivat viestinnän nopeutta ja välittömyyttä, samalla kun radio ja televisio laajensivat viestin tavoittavuutta (Standage, 1998; Fischer, 1992). Internet mullisti viestinnän globaalisti, mahdollistamalla reaaliaikaisen tiedonvälityksen ja vuorovaikutuksen yli maantieteellisten rajojen.

Nykyaikaiset immersiiiviset 3D-teknologiat ja tekoälyohjatut järjestelmät edustavat tämän kehityksen uusinta vaihetta. Ne eivät pelkästään siirrä informaatiota, vaan luovat uusia vuorovaikutuksen muotoja, joissa virtuaalinen ja fyysinen todellisuus yhdistyvät.

Immersiiviset 3D-avatatit mahdollistavat käyttäjille osallistumisen virtuaalisiin ympäristöihin tavalla, joka jäljittelee inhimillistä läsnäoloa ja nonverbaalista viestintää. Tämä kehitys hämää entisestään teknologiavälitteisen ja kasvokkain tapahtuvan vuorovaikutuksen rajoja.

Historiallisesti jokainen uusi viestintäteknologia on herättänyt keskustelua sen vaikutuksista yhteiskuntaan, kulttuuriin ja yksilöihin. Esimerkiksi puhelimen käyttöönotto synnytti huolta yksityisyyden menetyksestä ja inhimillisen vuorovaikutuksen muutoksista. Samoin nykyiset immersiiiviset teknologiat nostavat esiin kysymyksiä luottamuksesta, tietosuojasta ja viestinnän autenttisuudesta.

Tieteiskirjallisuus on usein ennakoimut ja heijastanut näitä kehityssuuntia. Kirjailijat kuten Philip K. Dick ja Stanisław Lem ovat teoksissaan tutkineet ihmisen ja teknologian välistä rajapintaa, pohtien kysymyksiä identiteetistä ja inhimillisyydestä (Dick, 1968; Lem, 1961). Nämä narratiivit korostavat tarpeen kriittiselle lähestymistavalle teknologian kehityksessä, erityisesti kun se vaikuttaa peruskäsityksiimme ihmisyydestä ja vuorovaikutuksesta.

Ymmärtämällä teknologisen viestinnän historiallisen kehityksen voimme paremmin kontekstualisoida immersiiivisten 3D-avatarien vaikutukset ja merkityksen nyky-yhteiskunnassa. Tämä auttaa meitä tunnistamaan, että kyse ei ole pelkästään teknisistä innovaatioista, vaan laajemmista kulttuurisista ja yhteiskunnallisista muutoksista, jotka vaikuttavat vuorovaikutusnormeihin ja inhimilliseen kokemukseen.

2.1.2. Posthumanismi ja identiteetti

Donna Harawayn (1985) posthumanistinen lähestymistapa, tunnettu myös hänen "Kyborgimanifestistaan", tarjoaa syvällisen näkemyksen siitä, miten teknologia liittyy inhimilliseen identiteettiin ja kehoon, haastaen perinteiset kategoriat inhimillisen ja keinotekoisien, luonnon ja kulttuurin välillä. Harawayn mukaan kyborgit – metaforisesti ymmärrettynä teknologiaan integroituneina olentoina – edustavat uudenlaista identiteettinäkemystä, joka ylittää biologisten ja mekaanisten rajojen määritelmät.

Immersiiviset 3D-avatatit ilmentävät tätä posthumanistista näkemystä käytännössä. Ne eivät ole vain virtuaalisia työkaluja tai hahmoja, vaan ne laajentavat ja muovaavat käyttäjän identiteettiä interaktioiden kautta digitaalisissa ympäristöissä. Avatatit mahdollistavat käyttäjien esittäytyä uusina, moninaisina versioina itsestään, ylittäen fyysisen kehon

rajoitukset ja sosiaaliset odotukset. Esimerkiksi virtuaalitodellisuuksissa käyttäjät voivat adoptoida eri sukupuoli-, rotu- tai laji-identiteettejä, tutkien ja ilmaisten itseään tavoin, jotka voivat olla mahdottomia reaali maailmassa.

Lisäksi, Harawayn mukaan, tämä teknologian ja identiteetin yhdistyminen voi edistää sosiaalista inklusiivisuutta ja monimuotoisuutta, tarjoten alustan, jolla perinteiset ennakkoluulot ja eriarvoisuudet voidaan kyseenalaistaa. Avatar-teknologia tarjoaa mahdollisuuden kokea ja ymmärtää erilaisuutta, joka voi edistää empatiaa ja yhteisymmärrystä käyttäjien välillä.

Toisaalta, immersiiivisten avatar-hahmojen käyttö nostaa esiin kysymyksiä autenttisuudesta ja inhimillisyyden rajoista. Käyttäjien kyky muokata ja hallita omaa virtuaalista esitystään herättää pohdintaa siitä, mikä on aitoa ja mikä konstruoitua. Avatarit kyseenalaistavat käsityksemme siitä, mitä tarkoittaa olla "todellinen" vuorovaikutuksessa toisten kanssa, kun fyysinen läsnäolo korvataan digitaalisilla representaatioilla. Tämä haastaa perinteiset käsitykset yksilöllisyydestä ja yhteisöllisyydestä, tuoden esiin uusia eettisiä ja filosofisia kysymyksiä siitä, miten identiteettiä tulisi ymmärtää ja arvostaa teknologian aikakaudella.

Kokonaisuudessaan, posthumanistinen lähestymistapa identiteettiin ja immersiiiviin avatareihin avaa uudenlaisia näkökulmia siihen, miten teknologia vaikuttaa henkilökohtaiseen ja kulttuuriseen identiteettiin. Se ei ainoastaan muuta tapaamme ymmärtää itseämme ja muita, vaan myös ehdottaa radikaaleja tapoja, joilla teknologia voi rikastuttaa inhimillistä kokemusta ja edistää kulttuurista ja yhteiskunnallista uudistusta.

Donna Harawayn posthumanistinen lähestymistapa korostaa teknologian ja ihmisyyden välisen rajan hämärtymistä. Immersiiviset avatarit voidaan nähdä tämän ajattelutavan ilmentyminä: ne eivät pelkästään toimi käyttäjän edustajina digitaalisissa ympäristöissä, vaan myös laajentavat ja muokkaavat käyttäjän käsitystä itsestään. Esimerkiksi avatarin avulla käyttäjä voi esittää monimuotoisempia identiteettejä, jotka eivät ole mahdollisia fyysisessä maailmassa. Tämä mahdollistaa uudenlaisen vapauden ja joustavuuden identiteetin ilmaisussa.

Toisaalta avatarit herättävät kysymyksiä autenttisuudesta ja inhimillisyyden rajoista. Voidaanko digitaalista itseä pitää aidosti 'todellisena', kun sen luonne perustuu algoritmien ohjaamiin valintoihin ja visuaalisiin representaatiomalleihin? Tämä haastaa perinteiset käsitykset identiteetistä ja pakottaa meidät pohtimaan, miten teknologia voi rikastuttaa, mutta myös muokata käsityksiämme siitä, keitä olemme.

2.1.3. Rationaalisuus, päätöksenteon logiikka ja eettiset näkökulmat

Rationaalisuus ja strategiset tavoitteet

Carl von Clausewitzin rationaalisuuden periaate, jossa toiminta linjataan strategisiin päämääriin, on erittäin soveltuva tekoälyohjattuihin järjestelmiin, kuten immersiiivisiin 3D-avatareihin. Clausewitz korosti, että jokaisen toimenpiteen tulee tukea yleisempiä tavoitteita, mikä on keskeistä myös tekoälyn päätöksentekologiikassa. Tekoälyjärjestelmät ovat suunniteltuja saavuttamaan tiettyjä päämääriä, ja niiden toiminnan tulee olla johdonmukaista näiden tavoitteiden kanssa.

Immersiiviset 3D-avatarit ja päätöksenteon logiikka

Immersiiviset 3D-avatarit voivat simuloida inhimillistä rationaalisuutta, mutta niiden käyttäytymisen ja päätösten taustalla ovat viime kädessä kehittäjien asettamat arvot, prioriteetit ja tavoitteet. Tämä herättää kysymyksiä siitä, miten käyttäjät voivat luottaa näiden järjestelmien päätöksentekoon sekä miten rationaalisuus määritellään ja kommunikoidaan käyttäjille. Esimerkiksi, jos avatarin käyttäytymistä ohjaavat epämääräisesti määritellyt tai käyttäjälle näkymättömät tavoitteet, voi syntyä ristiriita todellisen ja oletetun päätöksenteon rationaalisuuden välillä.

Clausewitzin näkemyksen pohjalta voidaan myös pohtia, miten tekoälyjärjestelmät yhdistävät strategisen ja taktisen rationaalisuuden. Strateginen rationaalisuus viittaa kykyyn asettaa ja noudattaa pitkäaikaisia tavoitteita, kun taas taktinen rationaalisuus keskittyy päivittäisten, konkreettisten tehtävien tehokkaaseen suorittamiseen. Optimaalinen suhde näiden kahden välillä on kriittinen, sillä se määrittelee, kuinka hyvin tekoäly kykenee toimimaan suunnitellusti erilaisissa tilanteissa ja kuinka se sopeutuu muuttuviin olosuhteisiin.

Eettiset näkökulmat ja käyttäjäkokemus

Tekoälyn päätöksentekologiikan läpinäkyvyys on keskeistä käyttäjien luottamuksen rakentamisessa. Jos järjestelmien toiminta perustuu monimutkaisiin ja läpinäkymättömiin algoritmeihin, käyttäjien voi olla vaikea ymmärtää, miksi tietyt päätökset tehdään. Tämä voi heikentää käyttäjien luottamusta järjestelmiin ja vähentää niiden tehokkuutta vuorovaikutuksessa. On tärkeää kehittää mekanismeja, jotka varmistavat, että tekoälyn päätöksenteko on avointa, ennustettavaa ja yhdenmukaista asetettujen tavoitteiden kanssa.

Lisäksi järjestelmien tulee ottaa huomioon käyttäjien yksityisyys ja autonomia. Tekoälyn keräämä ja käyttämä data voi sisältää henkilökohtaisia tietoja, ja on välttämätöntä varmistaa, että näiden tietojen käsittely tapahtuu eettisesti ja lakien mukaisesti. Käyttäjien tulee tietää, miten heidän tietojaan käytetään ja mihin tarkoituksiin.

Tekoälyn itsekriittisyyden tarpeellisuus

Jotta tekoälyjärjestelmät voivat toimia eettisesti kestäväällä tavalla, niiden tulee kyetä itsekriittisyyteen ja reflektioon päätöksentekoprosesseissaan. Tämä tarkoittaa, että järjestelmät pystyvät tunnistamaan omat rajoitteensa, epävarmuutensa ja mahdolliset virheensä. Tekoälyn tulisi soveltaa varovaisuusperiaatetta erityisesti tilanteissa, joissa päätöksillä voi olla merkittäviä vaikutuksia käyttäjien hyvinvointiin tai oikeuksiin.

Sokraattisesta filosofiasta tuttu viisaus oman tietämättömyyden tunnustamisesta on sovellettavissa myös tekoälyyn. Järjestelmien tulee olla suunniteltuja siten, että ne eivät toimi yli oman osaamisensa tai ymmärryksensä rajojen ja että ne hakevat tarvittaessa lisätietoja tai ohjausta. Tämä itsekriittinen lähestymistapa voi auttaa tekoälyä kehittymään vastuullisessa suunnassa, jossa se tunnustaa omat rajoitteensa ja pyrkii aktiivisesti minimoimaan mahdolliset haitalliset vaikutukset.

Ihmismäisten motiivien siirtämisen haasteet

Stuart Russell (2019) varoittaa, että ihmismäisten motiivien, kuten vallantavoittelun ja sosiaalisen hyväksynnän, siirtäminen tekoälyyn voi olla eettisesti kyseenalaista. Hänen mukaansa tekoälyn tulisi ensisijaisesti auttaa ihmistä ilman, että se jäljittelee ihmisen monimutkaisia motiiveja. Russellin mukaan tekoälyjärjestelmien tulee olla "todistettavasti

hyödyllisiä", mikä tarkoittaa, että niiden tulee keskittyä tukemaan ihmisten hyvinvointia eikä pyrkiä korvaamaan tai jäljittelemään ihmistä.

Tämä näkökulma tuo esiin tarpeen määritellä selkeät rajat tekoälyn toiminnalle ja tavoitteille. Immersiivisissä 3D-avatar-järjestelmissä tämä tarkoittaa, että avatarien tulee toimia käyttäjän apuna ja tukena ilman, että ne yrittävät manipuloida tai ohjata käyttäjää tavoilla, jotka eivät ole käyttäjän edun mukaisia.

Moraalinen vastuu ja kulttuuriset normit

Immersiivisten 3D-avatarien kehittämisessä tulee myös huomioida kulttuuriset normit ja intersubjektiivisuus. Järjestelmien on kunnioitettava erilaisia kulttuurisia taustoja ja arvoja sekä toimittava tavalla, joka edistää sosiaalista harmoniaa ja ymmärrystä. Tämä vaatii kehittäjiltä sensitiivisyyttä ja tietämystä eri kulttuurien käytännöistä ja odotuksista.

Lisäksi tekoälyn tulee kantaa moraalinen vastuu toiminnastaan. Tämä tarkoittaa, että järjestelmien tulee olla suunniteltuja ehkäisemään mahdolliset haitalliset vaikutukset sekä reagoimaan nopeasti ja tehokkaasti ongelmatilanteissa. Vastuullinen tekoälykehitys edellyttää sekä teknisiä että eettisiä ratkaisuja, joilla varmistetaan käyttäjien turvallisuus ja hyvinvointi.

Johtopäätökset

Yhdistämällä Clausewitzin rationaalisuuden periaatteen, tekoälyn itsekriittisyyden tarpeen ja eettiset näkökulmat voimme kehittää immersiiivisiä 3D-avatar-järjestelmiä, jotka toimivat ennustettavasti, läpinäkyvästi ja eettisesti. Tämä edistää käyttäjien luottamusta järjestelmiin ja mahdollistaa vastuullisen tekoälyn käytön vuorovaikutusjärjestelmissä. On tärkeää, että kehittäjät ottavat huomioon sekä tekniset että eettiset näkökohdat luodakseen järjestelmiä, jotka tukevat käyttäjien tarpeita ja arvoja samalla kun ne kunnioittavat heidän oikeuksiaan ja autonomiaansa.

2.1.4. Teknologiakriittiset näkemykset ja ihmiskeskeinen suunnittelu

Teknologiakriittiset näkemykset tarjoavat arvokkaan viitekehyksen tarkasteltaessa immersiiivisten 3D-teknologioiden vaikutuksia inhimilliseen kokemukseen ja vuorovaikutukseen. Sherry Turkle (2011) on tutkinut teknologiavälitteisen viestinnän vaikutuksia ihmisten sosiaalisiin suhteisiin, varoittaen siitä, että vaikka teknologia voi lisätä

yhteyden tunnetta, se saattaa samalla heikentää syvällistä inhimillistä vuorovaikutusta. Hänen mukaansa ihmiset saattavat odottaa teknologialta enemmän ja toisiltaan vähemmän, mikä voi johtaa yksinäisyyteen ja yhteyden menettämiseen.

Marshall McLuhanin (1964) ajatukset välineen vaikutuksesta viestiin korostavat teknologian kykyä muokata inhimillistä kokemusta. Immersiiviset 3D-avatatit eivät ainoastaan toimi viestinnän välineinä, vaan ne myös muovaavat tapaa, jolla ihmiset kokevat itsensä ja toiset virtuaalisessa ympäristössä. Tämä herättää kysymyksiä siitä, miten teknologia vaikuttaa identiteettiin, vuorovaikutusnormeihin ja siihen, mitä pidämme aitona kokemuksena.

Ihmiskeskeinen suunnittelu vastaa näihin haasteisiin korostamalla käyttäjien tarpeita ja kokemusta teknologian kehityksessä (Norman, 2013). Tavoitteena on luoda teknologioita, jotka tukevat inhimillistä hyvinvointia, autonomiaa ja merkityksellistä vuorovaikutusta. Tämä lähestymistapa painottaa käyttäjäkokemuksen merkitystä ja pyrkii välttämään teknologian mahdollisia haitallisia vaikutuksia.

Yhdistämällä teknologiakriittiset näkemykset ihmiskeskeiseen suunnitteluun voimme kehittää immersiiivisiä 3D-järjestelmiä, jotka tukevat käyttäjien tarpeita ja arvoja. Tämä vaatii kriittistä reflektiota teknologian roolista yhteiskunnassa ja aktiivista pyrkimystä eettisesti kestäviin ratkaisuihin.

2.1.5 Nonverbaalinen viestintä immersiiivisissä 3D-ympäristöissä

Nonverbaalinen viestintä on keskeinen osa inhimillistä vuorovaikutusta, ja sen merkityksen ymmärtäminen on olennaista tutkittaessa immersiiivisten 3D-avatatien vaikutusta digitaaliseen kommunikaatioon. Albert Mehrabianin (1972) tutkimukset osoittavat, että suuri osa viestinnästä tapahtuu nonverbaalisten vihjeiden, kuten ilmeiden, eleiden ja äänensävyyn kautta. Immersiivisissä 3D-ympäristöissä tämä korostuu, sillä avatatit voivat simuloida näitä nonverbaalisia elementtejä, mikä voi muuttaa käyttäjien kokemusta teknologiavälitteisestä vuorovaikutuksesta.

Nonverbaalinen viestintä ja avatatit:

Immersiiiviset 3D-avatatit mahdollistavat nonverbaalisen viestinnän elementtien, kuten kasvojen ilmeiden ja kehon liikkeiden, integroinnin digitaaliseen vuorovaikutukseen. Tämä

voi lisätä viestinnän luonnollisuutta ja syvyyttä, rikkoen perinteisiä rajoja teknologiavälitteisen ja kasvokkain tapahtuvan kommunikaation välillä (Gelder, 2009).

Peilisolut, empatia ja nonverbaalinen viestintä

Neurotieteellinen tutkimus on viime vuosikymmeninä syventänyt ymmärrystämme nonverbaalisen viestinnän hermostollisesta perustasta. Keskeinen löytö tässä yhteydessä on peilisolujen (mirror neurons) olemassaolo, jotka havaittiin alun perin apinoiden motorisessa aivokuoressa ja myöhemmin ihmisillä (Rizzolatti & Craighero, 2004). Peilisolut aktivoituvat sekä silloin, kun henkilö itse suorittaa tietyn liikkeen tai toiminnon, että silloin, kun hän havaitsee toisen henkilön suorittavan samanlaisen liikkeen tai toiminnon. Tämä mekanismi muodostaa perustan kyvyillemme ymmärtää ja jäljitellä toisten ihmisten käyttäytymistä, mikä on olennaista empatiassa ja sosiaalisessa vuorovaikutuksessa.

Molemminpuolinen matkiminen ja vuorovaikutuksen syventäminen

Peilisolujen ansiosta ihmiset kykenevät luonnostaan matkimaan toistensa ilmeitä, eleitä ja kehonkieltä, usein tiedostamattaan. Tätä molemminpuolisen matkimisen ilmiötä kutsutaan myös "kameleonttivaikutukseksi" (Chartrand & Bargh, 1999). Molemminpuolinen matkiminen syventää vuorovaikutusta ja edistää sosiaalista yhteenkuuluvuutta, sillä se vahvistaa tunnepohjaista yhteyttä osapuolten välillä. Kun ihmiset matkivat toistensa nonverbaalista viestintää, he tulevat paremmin tietoisiksi toistensa tunteista ja intentioista, mikä helpottaa kommunikaatiota ja yhteistyötä.

Immersiivisissä 3D-ympäristöissä ja avatarien kanssa tapahtuvassa vuorovaikutuksessa molemminpuolinen matkiminen voi olla ratkaiseva tekijä autenttisen ja merkityksellisen kokemuksen luomisessa. Jos avatar kykenee tunnistamaan käyttäjän nonverbaalisia vihjeitä ja reagoimaan niihin vastaavalla tavalla, se voi aktivoida käyttäjän peilisoluja ja lisätä näiden kokemusmaailman syvyyttä. Tämä voi saada käyttäjän unohtamaan olevansa vuorovaikutuksessa keinotekoisen järjestelmän kanssa ja sen sijaan tuntemaan aidon yhteyden avatarin kanssa. Kun ymmärrämme peilisolujen ja molemminpuolisen matkimisen merkityksen, voimme soveltaa tätä tietoa avatarien suunnittelussa parantaaksemme vuorovaikutuskokemusta.

Predictive Processing -teoria (Clark, 2013) tarjoaa syvällisen näkökulman siihen, miten avatarit voivat hyödyntää käyttäjän nonverbaalisia vihjeitä reaaliaikaisesti. Teorian mukaan aivot pyrkivät jatkuvasti minimoimaan ennustevirheitä vertaamalla ulkoista sensorista tietoa sisäisiin malleihin. Avatarien kohdalla tämä voisi tarkoittaa järjestelmän kykyä tulkita käyttäjän ilmeitä, eleitä ja puheen sävyä luoden samalla ennusteita käyttäjän seuraavista toimista. Kun avatar pystyy ennakoimaan käyttäjän käyttäytymistä, se voi vastata entistä luonnollisemmin, mikä parantaa vuorovaikutuksen autenttisuutta ja lisää käyttäjän luottamusta järjestelmään.

Tämä mekanismi muistuttaa peilisolujen toimintaa ihmisillä, mutta digitaalinen vastine vie sen askeleen pidemmälle. Esimerkiksi, jos käyttäjä hymyilee, avatar voi paitsi vastata hymyyn myös mukauttaa seuraavaa vuorovaikutustaan ennakoimalla, että käyttäjä on avoimessa ja positiivisessa tunnetilassa. Tällainen dynaaminen reaktiokyky mahdollistaa aiempaa intuitiivisemmän ja vuorovaikutuksellisemmän kokemuksen, joka hämärtää rajaa teknologiavälitteisen ja inhimillisen viestinnän välillä.

Avatarien suunnittelu peilisolujen näkökulmasta

Peilisolujen toiminnan ymmärtäminen avaa uusia mahdollisuuksia kehittää avatareja, jotka kommunikoivat käyttäjien kanssa luonnollisemmin. Tässä alaluvussa tarkastellaan konkreettisia tapoja hyödyntää peilisolujen periaatteita avatarien suunnittelussa.

Avatarien suunnittelussa peilisolujen toiminnan hyödyntäminen voi merkittävästi parantaa vuorovaikutuksen laatua. Tämä voidaan saavuttaa seuraavilla tavoilla:

1. **Reaaliaikainen nonverbaalisen viestinnän tunnistus:** Käyttämällä kehittyneitä konenäkö- ja tekoälytekniikoita avatar voi analysoida käyttäjän ilmeitä, eleitä ja kehonkieltä. Tämä mahdollistaa avatarin reagoinnin käyttäjän nonverbaalisiin signaaleihin, mikä tekee vuorovaikutuksesta luonnollisempaa ja dynaamisempaa.
2. **Avatarin nonverbaalinen ilmaisu:** Avatarin tulisi pystyä ilmentämään laajaa kirjoa nonverbaalisia viestejä, kuten hymyilemään, nyökkäämään tai osoittamaan empatiaa ilmeillään. Tämä aktivoi käyttäjän peilisolut ja luo emotionaalisen yhteyden.

3. **Molemminpuolinen matkiminen:** Avatar voi tietoisesti matkia käyttäjän nonverbaalista viestintää hienovaraisesti, mikä vahvistaa sosiaalista sidettä ja saa käyttäjän tuntemaan olonsa ymmärretyksi. Esimerkiksi, jos käyttäjä nyökkää, avatar voi vastata nyökkäämällä takaisin.
4. **Mukautuva oppiminen:** Avatar voi oppia käyttäjän vuorovaikutustyylistä ajan myötä ja mukauttaa omaa käyttäytymistään vastaamaan paremmin käyttäjän preferenssejä. Tämä lisää vuorovaikutuksen personoituneisuutta ja vahvistaa käyttäjän sitoutumista. Kun avatarit pystyvät mukautumaan käyttäjän tarpeisiin, seuraava askel on ymmärtää empatian ja emotionaalisen yhteyden rooli vuorovaikutuksessa.

Näiden suunnitteluperiaatteiden soveltaminen voi merkittävästi parantaa avatarien kykyä luoda syvälinen ja emotionaalinen yhteys käyttäjiin, mikä edistää parempaa vuorovaikutusta ja sitoutumista.

Empatian ja emotionaalisen yhteyden merkitys

Empatian syntyminen vuorovaikutuksessa avatarin kanssa voi parantaa käyttökokemusta monin tavoin. Se voi lisätä käyttäjän:

- **Sitoutumista:** Emotionaalinen yhteys tekee vuorovaikutuksesta mielekkäämpää, mikä saa käyttäjän viettämään enemmän aikaa järjestelmän parissa.
- **Luottamusta:** Kun käyttäjä kokee avatarin ymmärtävän ja reagoivan hänen tunteisiinsa, hän todennäköisemmin luottaa järjestelmään ja hyväksyy sen tarjoamat neuvot tai tiedot.
- **Oppimista ja muistamista:** Emotionaalisesti latautuneet kokemukset jäävät paremmin mieleen, mikä voi olla erityisen hyödyllistä opetuksellisissa sovelluksissa.

Haasteet ja eettiset näkökulmat

Vaikka peilisolujen toiminnan hyödyntäminen tarjoaa merkittäviä etuja, se herättää myös tärkeitä eettisiä kysymyksiä:

- **Manipulaation riski:** Avatarien kyky vaikuttaa käyttäjän tunteisiin voi johtaa tilanteisiin, joissa käyttäjän emotionaalista tilaa manipuloidaan esimerkiksi kaupallisiin tarkoituksiin.

- **Yksityisyys:** Nonverbaalisen viestinnän analysointi voi paljastaa herkkiä tietoja käyttäjän tunnetiloista ja intentioista. On välttämätöntä varmistaa, että tällainen data käsitellään vastuullisesti ja käyttäjän suostumuksella.
- **Autenttisuuden illuusio:** Jos avatar vaikuttaa liian inhimilliseltä, käyttäjä saattaa unohtaa olevansa vuorovaikutuksessa tekoälyn kanssa, mikä voi johtaa epärealistisiin odotuksiin tai pettymyksiin.

Kytkeä tutkimuskysymyksiin

Peilisolujen ja molemminpuolisen matkimisen integrointi immersiiivisiin 3D-avatar-järjestelmiin liittyy keskeisesti tutkimuskysymyksiin:

- **Ensimmäinen tutkimuskysymys:** Peilisolujen aktivoituminen avatarien kanssa vuorovaikuttaessa osoittaa, miten immersiiiviset 3D-avatatit voivat muovata digitaalisen vuorovaikutuksen normeja ja haastaa perinteisiä käsityksiä teknologiavälitteisestä viestinnästä. Kun teknologiasta tulee lähes erottamaton inhimillisestä vuorovaikutuksesta, perinteiset rajat hämärtyvät.
- **Toinen tutkimuskysymys:** Avatarien kyky vaikuttaa käyttäjän tunteisiin ja reaktioihin nostaa esiin eettisiä ja moraalisia periaatteita liittyen käyttäjien luottamukseen, turvallisuuden tunteeseen ja manipulaation mahdollisuuteen. On tärkeää tutkia, miten älykkäiden vuorovaikutusjärjestelmien kehitys vaikuttaa näihin tekijöihin ja miten ne tulee ottaa huomioon järjestelmien suunnittelussa.

Johtopäätös

Peilisolujen ja molemminpuolisen matkimisen hyödyntäminen avatarien suunnittelussa voi merkittävästi parantaa immersiiivisten 3D-ympäristöjen vuorovaikutuskokemusta. Se syventää emotionaalista yhteyttä ja lisää vuorovaikutuksen luonnollisuutta, mikä voi johtaa positiivisiin tuloksiin esimerkiksi opetuksessa, terveydenhuollossa ja viihteessä. Kuitenkin näiden teknologioiden eettinen käyttö edellyttää huolellista harkintaa ja selkeiden periaatteiden noudattamista käyttäjien yksityisyyden ja hyvinvoinnin turvaamiseksi.

Nonverbaalinen viestintä taiteissa ja sen soveltaminen avatareihin:

Nonverbaalinen viestintä on ollut keskeisessä asemassa esittävissä taiteissa kautta historian. Nämä taiteen muodot hyödyntävät kehon liikkeitä, ilmeitä, eleitä, äänen sävyjä ja rytmiä välittääkseen merkityksiä yleisölle. Tarkastelemalla, miten teatteri, tanssi ja muu esittävä taide hyödyntävät nonverbaalisia keinoja, voimme löytää arvokkaita näkökulmia avatarien suunnitteluun.

Teatteri ja ilmaisun moninaisuus

Teatterissa näyttelijät käyttävät monipuolisia nonverbaalisia keinoja hahmojensa tunnetilojen ja motiivien ilmaisemiseen. Konstantin Stanislavskin (1936) kehittämä metodi korostaa fyysisen ilmaisun merkitystä sisäisen tunnetilan ulkoistamisessa. Näyttelijän tulee elää roolinsa todeksi, mikä näkyy hienovaraisina kasvonilmeinä, kehon asentoina ja liikkeinä. Näin yleisö pystyy samaistumaan hahmoon ja ymmärtämään tämän kokemuksia syvemmin.

Bertolt Brechtin (1964) eppinen teatteri puolestaan hyödynsi vieraannuttamisen käsitettä, jossa näyttelijät saattoivat tietoisesti rikkoa neljättä seinää ja käyttää liioiteltuja eleitä ja ilmeitä kiinnittääkseen huomion teoksen yhteiskunnalliseen sanomaan. Tällä tavalla nonverbaalista viestintää käytettiin herättämään ajatuksia ja haastamaan katsojan käsityksiä.

Tanssi ja kehon kieli

Tanssitaiteessa nonverbaalinen viestintä on keskeisessä roolissa, sillä liike itsessään on ilmaisun väline. Baletin harmoniset liikkeet kertovat usein tarinoita rakkaudesta, menetyksestä ja kamppailusta, kun taas nykytanssin abstraktimmat muodot voivat ilmaista monimutkaisia tunteita ja ajatuksia ilman selkeää narratiivia. Martha Graham (1991) korosti tanssissaan kehon sisäisiä impulsseja ja tunteiden fyysistä ilmausta, mikä tekee esityksistä intensiivisiä ja henkilökohtaisia.

Ray Birdwhistellin kinesiikka

Ray Birdwhistellin (1970) työ kinesiikan parissa on ollut merkittävä nonverbaalisen viestinnän tutkimuksessa. Hän esitti, että kehon liikkeet ja asennot muodostavat järjestelmällisen kielen, joka vaihtelee kulttuurien välillä. Birdwhistell korosti, että pienetkin liikkeet, kuten kulmakarvojen kohottaminen tai pään kallistus, voivat välittää merkityksellistä informaatiota vuorovaikutuksessa.

Nonverbaalisen viestinnän soveltaminen avatareihin

Soveltamalla esittävien taiteiden nonverbaalisen viestinnän periaatteita avatarien suunnittelussa voidaan luoda virtuaalisia hahmoja, jotka kommunikoivat uskottavasti ja vaikuttavasti. Tämä edellyttää ymmärrystä siitä, miten ihmiset tulkitsevat nonverbaalisia vihjeitä ja miten näitä voidaan mallintaa digitaalisessa ympäristössä.

1. **Ilmeiden ja mikroilmeiden mallintaminen:** Paul Ekmanin (1978) tutkimukset kasvojen mikroilmeistä osoittavat, että tietyt tunneilmaisut ovat universaaleja. Avatarien suunnittelussa näiden mikroilmeiden tarkka mallintaminen voi auttaa välittämään hahmon tunnetiloja käyttäjille. Esimerkiksi hienovarainen hymy tai silmien siristys voi ilmaista iloa tai kiinnostusta.
2. **Kehon asennot ja liikkeet:** Stanislavskin ja Grahamin periaatteiden mukaisesti avatarien tulisi pystyä ilmaisemaan tunnetiloja ja tarkoitusperiä kehon asentojen ja liikkeiden kautta. Voimakas, ylöspäin suuntautunut liike voi viestiä itsevarmuutta, kun taas alaspäin suuntautunut ja kutistuva asento voi ilmaista surua tai epävarmuutta.
3. **Rytmi ja dynamiikka:** Tanssissa liikkeen rytmi ja dynamiikka ovat keskeisiä viestin välittämisessä. Avatarin liikkeiden tulisi heijastaa sisäistä rytmiä, joka sopii tilanteeseen ja hahmon tunnetilaan. Esimerkiksi nopeat ja terävät liikkeet voivat viitata kiihtymykseen tai aggressioon, kun taas hitaat ja sulavat liikkeet voivat ilmaista rauhallisuutta tai surua.
4. **Kulttuuriset erot:** Birdwhistellin korostamat kulttuuriset vaihtelevuudet nonverbaalisessa viestinnässä on otettava huomioon avatarien suunnittelussa. On tärkeää, että avatarit voivat mukautua eri kulttuurien normeihin ja käyttäytymistapoihin, jotta viestintä on tehokasta ja vältetään väärinymmärryksiä. Tämä edellyttää kulttuurisesti sensitiivistä suunnittelua ja mahdollisuutta mukauttaa avatarin ilmaisua.
5. **Interaktiivisuus ja responsiivisuus:** Avatarien tulisi pystyä reagoimaan käyttäjän nonverbaalisiin vihjeisiin, mikä luo dynaamisen vuorovaikutuksen. Käyttäjän eleiden ja ilmeiden tunnistaminen ja niihin vastaaminen voi syventää vuorovaikutusta ja tehdä siitä luonnollisemman tuntuisen. Esimerkiksi jos käyttäjä nyökkää, avatar voi vastata nyökkäämällä takaisin, mikä vahvistaa molemminpuolista ymmärrystä.

Kun olemme tarkastelleet, kuinka nonverbaalisen viestinnän periaatteita voidaan soveltaa avatarien suunnittelussa, on hyödyllistä pohtia käytännön sovelluksia, joissa näitä periaatteita hyödynnetään.

Käytännön sovellukset

- **Koulutus ja oppiminen:** Avatarit, jotka hyödyntävät nonverbaalista viestintää, voivat toimia tehokkaina opettajina tai mentorina virtuaaliympäristöissä. Ne voivat ilmaista kannustusta, empatiaa tai ohjausta eleiden ja ilmeiden kautta, mikä voi parantaa oppijan sitoutumista ja motivaatiota.
- **Terapeuttiset sovellukset:** Nonverbaalisesti ilmaisevat avatarit voivat tukea terapeuttisia prosesseja, erityisesti silloin kun inhimillinen kontakti on rajoitettua. Ne voivat tarjota turvallisen ympäristön tunteiden ilmaisulle ja auttaa käyttäjiä käsittelemään vaikeita kokemuksia.
- **Viihde ja pelit:** Peliteollisuudessa avatarien uskottava nonverbaalinen viestintä voi syventää tarinankerrontaa ja pelaajan emotionaalista investointia peliin. Hahmot, jotka reagoivat aidosti pelaajan toimintaan, luovat immerssiivisemmän pelikokemuksen.

Haasteet ja huomioitavat seikat

Vaikka nonverbaalisen viestinnän soveltaminen avatareihin tarjoaa merkittäviä etuja, siihen liittyy myös haasteita ja eettisiä näkökulmia:

- **Tekniset rajoitteet:** Nonverbaalisen viestinnän hienouksien mallintaminen vaatii edistynyttä teknologiaa, kuten tarkkaa liikkeenkaappausta ja tekoälyä, joka pystyy tulkitsemaan ja tuottamaan monimutkaisia viestinnän muotoja. Tekniset puutteet voivat johtaa epäaidolta vaikuttavaan viestintään.
- **Eettiset näkökulmat:** Kun avatarit kykenevät ilmaisemaan tunteita ja vaikuttamaan käyttäjän emotionaaliseen tilaan, on tärkeää huomioida eettiset periaatteet, kuten manipulaation välttäminen ja käyttäjän suostumuksen varmistaminen. Käyttäjien on oltava tietoisia siitä, että he ovat vuorovaikutuksessa tekoälyn kanssa.

- **Autenttisuus vs. laakson ilmiö:** Liian realistiset avatarit voivat aiheuttaa "uncanny valley" -efektin, jossa hahmot koetaan häiritsevinä tai epäluonnollisina (Mori, 1970). On löydettävä tasapaino uskottavan ilmaisuuden ja käyttäjän mukavuuden välillä.

Kytkeä tutkimuskysymyksiin

Integroimalla nonverbaalisen viestinnän elementtejä immersiiivisiin 3D-avatareihin voimme tutkia:

- **Miten tällaiset avatarit muuttavat digitaalisen vuorovaikutuksen normeja ja käyttäjien kokemusta teknologiavälitteisestä kommunikaatiosta?** Nonverbaalinen viestintä voi tehdä vuorovaikutuksesta aidomman tuntuista ja muuttaa käyttäjien odotuksia digitaalisten järjestelmien suhteen, mikä vastaa ensimmäiseen tutkimuskysymykseen.
- **Miten avatarien kyky ilmaista tunteita nonverbaalisten vihjeiden kautta vaikuttaa käyttäjien luottamukseen ja turvallisuuden tunteeseen?** Empatian ja emotionaalisen yhteyden tunne voi vaikuttaa siihen, miten käyttäjät kokevat älykkäiden järjestelmien eettisyyden ja luotettavuuden, mikä liittyy toiseen tutkimuskysymykseen.

Johtopäätös

Yhdistämällä esittävien taiteiden nonverbaalisen viestinnän periaatteet avatar-teknologiaan voimme luoda virtuaalisia hahmoja, jotka kommunikoivat uskottavasti ja herättävät tunteita käyttäjissä. Tämä edistää immersiiivisten 3D-avatar-järjestelmien kehitystä tavalla, joka muuttaa digitaalisen vuorovaikutuksen maisemaa ja tuo uusia ulottuvuuksia teknologiavälitteiseen viestintään. Samalla on tärkeää olla tietoinen teknologian käytön eettisistä ja teknisistä haasteista, jotta voidaan varmistaa käyttäjäkokemuksen laadukkuus ja luotettavuus.

2.1.6 Avatar-teknologian historiallinen kehitys ja sen sovellukset

Avatar-teknologian historiallinen kehitys tarjoaa kiehtovan katsauksen siihen, miten eri aikakausien teknologiset, kulttuuriset ja sosiaaliset kontekstit ovat muovanneet digitaalisen

vuorovaikutuksen muotoja. Tässä alaluvussa käsitellään, miten avatarit ovat olleet osa inhimillistä kulttuuria ja teknologiaa antiikin ajoista nykypäivään, ja miten ne ovat vaikuttaneet ja heijastaneet yhteiskunnallisia arvoja ja ihanteita eri aikakausina.

Antiikin aika ja mytologia: Avatarit nähtiin usein uskonnollisissa konteksteissa, esimerkiksi antiikin Egyptissä, jossa jumaluuksia kuvattiin usein ihmisen ja eläimen yhdistelmänä. Nämä kuvaelmat toimivat avatarina, joiden avulla pyrittiin ymmärtämään ja tulkitsemaan jumalallisia ominaisuuksia.

Keskiaika ja varhainen uusi aika: Legendaariset hahmot, kuten Pyhä Yrjö tai Jeanne d'Arc, toimivat aikansa avatareina. He edustivat yhteiskunnallisia ihanteita ja arvoja, ja heidän tarinansa välittivät opetuksia ja normeja aikakauden ihmisille.

Teollistuminen ja moderni aika: 1800-luvun automaattit ja mekaaniset nuket, kuten kuuluisat soitto- ja kirjoituskoneet, esittelivät teknologian kykyä jäljitellä ihmisen toimintaa. Ne olivat varhaisia esimerkkejä siitä, miten koneet voivat ottaa ihmismäisiä rooleja ja toimintoja.

1900-luvun media ja popkulttuuri: Elokuvat ja televisio toivat esiin robotti- ja androidihahmot, jotka kuten Fritz Langin *Metropolis*-elokuvan robotti tai Star Trek-sarjan Data, haastoivat käsityksiä ihmisen ja koneen rajapinnoista.

2000-luku ja digitaalinen vallankumous: Virtuaaliset maailmat, kuten Second Life ja nykyiset sosiaalisen median alustat, ovat mahdollistaneet yksilöiden esiintymisen digitaalisina avatareina, jotka osallistuvat monimutkaisiin sosiaalisiin vuorovaikutuksiin ja identiteettien rakentamiseen.

Nykyhetki ja tulevaisuus: Tekoälyohjatut ja oppivat avatarit, kuten nykyajan chatbotit ja virtuaaliset asiakaspalvelijat, ovat kehittyneet olemaan entistä uskottavampia vuorovaikutuskumppaneita. Ne edustavat teknologian viimeisintä kehitysvaihetta, jossa tekoäly ja avatar-teknologia yhdistyvät luomaan entistä monimutkaisempia ja inhimillisempiä vuorovaikutuskokemuksia.

Tämän historian ymmärtäminen antaa arvokasta kontekstia nykyisille eettisille pohdintoille, jotka liittyvät immersivisten teknologioiden ja avatar-järjestelmien käyttöön. Se auttaa

hahmottamaan, miten nykyhetken teknologiset innovaatiot istuvat laajempaan historialliseen kehityskaareen ja miten ne muokkaavat yhteiskunnallisia ja kulttuurisia normeja.

2.2. Immersiiviset 3D-teknologiat ja eettiset kysymykset

Tämän osion tarkoituksena on käsitellä immersiiivisiin 3D-teknologioihin ja älykkäisiin vuorovaikutusjärjestelmiin liittyviä eettisiä ja moraalisia kysymyksiä. Keskitymme erityisesti rationaalisuuden, moraalisen vastuun ja kulttuuristen normien näkökulmiin.

2.2.1. Moraalinen vastuu ja käyttäjäkokemus

Anil Sethin (2016) teoria tietoisuudesta "kontrolloituna hallusinaationa" tarjoaa uuden näkökulman avatarien käyttäjäkokemukseen. Sethin mukaan aivot eivät yksinkertaisesti vastaanota ulkoista tietoa, vaan aktiivisesti rakentavat subjektiivista kokemusta ennusteiden avulla. Tämä sama periaate voidaan soveltaa 3D-avatareihin: jos avatarien käyttäytyminen on riittävän hienovaraista ja dynaamista, ne voivat synnyttää käyttäjässä illuusion aidosta inhimillisestä vuorovaikutuksesta.

Tämä illuusio voi kuitenkin herättää eettisiä kysymyksiä. Jos käyttäjä ei tiedosta vuorovaikutuksen taustalla olevan tekoälyä, hän saattaa luottaa järjestelmään liiallisesti tai kokea pettymystä avatarin rajallisuuksien paljastuessa. Avatarien suunnittelussa on siksi tärkeää tasapainottaa realistisuuden ja läpinäkyvyyden vaatimukset, jotta käyttäjän kokemus säilyy autenttisena ilman harhaanjohtavia elementtejä.

Moraalinen vastuu korostuu erityisesti järjestelmissä, joissa käyttäjä on riippuvainen tekoälyn toiminnasta. Immersiivisissä järjestelmissä on otettava huomioon käyttäjien emotionaaliset ja sosiaaliset tarpeet sekä ehkäistävä manipulaation tai väärinkäytön mahdollisuus. Tämä on tärkeää käyttäjäkokemuksen eettisen kestävyysvarmistamiseksi. On välttämätöntä käsitellä tarkemmin manipuloinnin riskejä ja läpinäkyvyyttä, tuomalla esiin esimerkkejä siitä, miten tekoälyn toiminta voi johtaa käyttäjien harhaanjohtamiseen tai heidän tietojensa väärinkäyttöön. Näiden riskien minimointi vaatii järjestelmien suunnittelussa eettisten ohjeiden noudattamista, käyttäjän informoimista ja toiminnan jatkuvaa tarkkailua.

Ihmisen aivojen toimintatapa, jossa sensorisia syötteitä ennakoidaan ja niiden pohjalta luodaan reaaliaikaisesti subjektiivinen kokemus, tarjoaa arvokkaan mallin AI-ohjattujen avatarien kehittämiseen. Sethin (2016) mukaan tietoisuus on "kontrolloitu hallusinaatio", jossa aivot rakentavat kokemuksia ennusteiden ja niihin liittyvien virheiden perusteella. Tämä mekanismi mahdollistaa nopean reagoinnin muuttuvissa olosuhteissa ja tekee vuorovaikutuksesta sujuvaa. Clark (2013) puolestaan painottaa, että ennakoiva prosessointi vähentää aivojen kuormitusta hyödyntämällä aiempia malleja tulevien tilanteiden ennustamiseen ja niiden pohjalta toimimiseen.

Avatarit voivat hyödyntää tätä lähestymistapaa ennustamalla käyttäjän tulevia toimia ja reaktioita. Esimerkiksi käyttäjän äänen sävystä, puheen rytmistä ja eleistä saatujen vihjeiden perusteella avatar voi valmistella tulevia toimintojaan, kuten sanat, äänensävyt ja ilmeet, jo ennen kuin lopullinen syöte on saatu. Ennakoivan prosessoinnin avulla avatarin puhe ja siihen liittyvät animaatiot ovat valmiina täsmällisesti oikealla hetkellä, mikä vähentää viivettä ja parantaa vuorovaikutuksen luonnollisuutta. Tämä vastaa Sethin teoriaa siitä, miten aivot tuottavat sujuvia kokemuksia minimoimalla ennustevirheet jatkuvasti vuorovaikutuksen aikana.

Jos avatarin ennuste ei vastaa käyttäjän antamaa palautetta, järjestelmä voi tarkentaa reaktiotaan korjaamalla ennustetta reaaliaikaisesti. Tämä mekanismi, jonka Clark kuvaa "hierarkkiseksi ennusteiden päivittämiseksi", mahdollistaa dynaamisen sopeutumisen ja takaa sen, että avatar pystyy mukautumaan hienovaraisiin muutoksiin käyttäjän toiminnassa. Lopputuloksena on vuorovaikutus, joka tuntuu käyttäjälle intuitiiviselta ja saumattomalta.

Samankaltainen ennusteiden hyödyntäminen näkyy myös kielimalleissa, jotka luovat ennusteita seuraavasta tokenista tai sanasta aiemman kontekstin perusteella. Vaikka kielimallit toimivat rajatummassa kielellisessä ympäristössä, niiden toimintaperiaate – ennustaa ja päivittää malleja palautteen perusteella – muistuttaa avatarien ennakoivan toiminnan logiikkaa. Tästä syystä ennusteiden hyödyntäminen tarjoaa tehokkaan tavan parantaa sekä nopeutta että vuorovaikutuksen sujuvuutta.

Tällainen järjestelmä ei pelkästään nopeuta reaktioita, vaan myös tekee avatarista vakuuttavan ja luotettavan vuorovaikutuskumppanin, sillä se näyttää ymmärtävän käyttäjän tunteita ja tarpeita. Samalla avatarien kehittämisessä on tärkeää huomioida läpinäkyvyys ja käyttäjän

oikeudet, jotta ennakoiva toiminta ei johda manipulointiin tai harhaanjohtaviin kokemuksiin, kuten Seth ja Clark molemmat korostavat. Näiden periaatteiden noudattaminen voi luoda järjestelmiä, jotka parantavat sekä teknologian että käyttäjäkokemuksen laatua.

Ihmismotiivien evolutiivinen perusta ja niiden soveltuvuus tekoälyyn (Russell, 2019)

Stuart Russell (2019) varoittaa, että ihmismäisten motiivien, kuten vallan tavoittelun ja sosiaalisen hyväksynnän, siirtäminen tekoälyyn voi olla eettisesti kyseenalaista. Hänen mukaansa AI:n tulisi ensisijaisesti auttaa ihmistä ilman, että se jäljittelee ihmisen monimutkaisia motiiveja. Russellin mukaan AI-järjestelmien tulee olla “todistettavasti hyödyllisiä”, mikä tarkoittaa, että niiden tulee keskittyä tukemaan ihmisten hyvinvointia eikä pyrkiä korvaamaan tai jäljittelemään ihmistä.

Russellin näkökulma auttaa ymmärtämään, että vaikka AI-ohjattu avatar voisi jäljitellä käyttäytymismalleja, sen tulisi toimia eettisissä rajoissa ja selkeiden sääntöjen mukaisesti. Ihmismäisyyttä ei tarvitse korostaa, vaan painottaa tulisi hyödyllisyyttä ja turvallisuutta.

2.2.2. Intersubjektiivisuus ja kulttuuriset normit

Immersiiviset järjestelmät, kuten 3D-avatat, hyödyntävät jaettuja kulttuurisia käsitteitä ja normeja simuloidessaan inhimillistä vuorovaikutusta. Tämä luo autenttisen ja luottamusta herättävän kokemuksen, mutta väärin tulkittuna nämä kulttuuriset vihjeet voivat johtaa väärinkäsityksiin ja luottamuspulaan. Tämän vuoksi suunnittelussa on keskeistä huomioida monimuotoisuuden ja kulttuurisen ymmärryksen merkitys.

Antonio Damasio (1994) korostaa empaattisten ominaisuuksien merkitystä inhimillisessä vuorovaikutuksessa. Immersiiviset järjestelmät voivat hyödyntää empaattisia elementtejä, jotka lisäävät käyttökokemuksen laatua ja vahvistavat käyttäjän luottamusta. Näitä elementtejä ei kuitenkaan tule liioitella, jotta järjestelmät eivät harhaan johda tai luo epärealistisia odotuksia.

Donna Harawayn posthumanistinen näkökulma tukee monimuotoisuuden ja kulttuurien välisen ymmärryksen edistämistä AI-järjestelmissä. Immersiiviset avatarit voivat tarjota käyttäjille uusia ilmaisutapoja ja mahdollisuuksia rakentaa yhteyksiä eri kulttuurien välille, mikä vahvistaa niiden potentiaalia edistää inklusiivisuutta ja sosiaalista tasa-arvoa.

2.2.3. Itsekriittisyys ja varovaisuusperiaate

Itsekriittisyys ja varovaisuus ovat keskeisiä periaatteita, joiden avulla tekoälyohjatut järjestelmät voivat hallita epävarmuutta ja ehkäistä virheitä. Karl Fristonin vapaan energian periaate (2010) tarjoaa matemaattisen perustan sille, miten biologiset ja kognitiiviset järjestelmät pyrkivät vähentämään epävarmuutta ja ylläpitämään tasapainoa ennusteiden ja ympäristön välillä. Tämä ajattelutapa voidaan soveltaa myös tekoälyohjattuihin avatareihin, jotka voivat dynaamisesti mukautua käyttäjän vuorovaikutustyyliin ja tunteisiin minimoimalla vuorovaikutuksen ristiriitoja.

Tällainen järjestelmä voisi esimerkiksi tunnistaa käyttäjän hämmentyneisyyden ja mukauttaa ilmaisuaan, jotta käyttäjä kokisi vuorovaikutuksen intuitiivisemmaksi. Samalla on kuitenkin tärkeää, että nämä järjestelmät noudattavat varovaisuusperiaatetta: avatarin tulisi tunnistaa omat rajoitteensa ja välttää liian itsenäistä toimintaa, joka voisi johtaa käyttäjän harhaan tai vaarantaa tämän turvallisuuden tunteen.

Immersiivisten 3D-avatarien tulisi pystyä osoittamaan epävarmuuttaan monimutkaisissa tilanteissa, mikä voi lisätä käyttäjien luottamusta järjestelmään. Syventäen keskustelua itsekriittisyydestä ja varovaisuusperiaatteesta, on tärkeää korostaa esimerkkejä siitä, miten AI-järjestelmät voivat epäonnistua eettisten standardien noudattamisessa ja miten tällaiset riskit voidaan tunnistaa ja hallita proaktiivisesti.

John Searlen vuonna 1980 esittämä kiinalaisen huoneen argumentti on keskeinen filosofinen pohdinta tekoälyn kognitiivisista rajoitteista. Argumentin mukaan vaikka tekoälyjärjestelmät voivat näennäisesti ymmärtää kieltä tai muita symboleja, ne vain manipuloivat symboleja ilman todellista ymmärrystä. Tämä käsitys on merkittävä, kun tarkastellaan tekoälyohjattujen avatarien kykyä simuloida uskonnollisia tai yliluonnollisia uskomuksia, jotka ovat syvästi merkityksellisiä ja monimutkaisia inhimillisen kulttuurin ja kokemuksen näkökulmasta.

Ihmisten taipumus uskoa yliluonnollisiin ilmiöihin—kuten jumaliin, henkiin tai kohtaloon—on evoluution muovaama ominaisuus, joka on historiallisesti edistänyt yhteisöllisyyttä ja moraalisten sääntöjen noudattamista. Tämä uskomusten järjestelmä tukee sosiaalista koheesiota ja tarjoaa selityskehyksiä elämän ilmiöille, jotka muuten voisivat jäädä käsittämättömiksi. Kun tekoälyohjatut avatarit simuloidaan osallistumaan tai edistämään tällaisia uskomuksia, se asettaa kehittäjät monimutkaiseen eettiseen dilemmaan.

Tekoälyn käyttö uskonnollisten tai yliluonnollisten uskomusten simulointiin voi olla ongelmallista useasta syystä. Ensinnäkin, jos avatarit edistävät tiettyjä uskomusjärjestelmiä, ne voivat vahvistaa käyttäjän olemassa olevia ennakkoluuloja tai jopa estää kriittisen ajattelun kehittymistä. Tällaiset järjestelmät voivat ohjata käyttäjiä hyväksymään epäkritiikillisesti tiettyjä näkemyksiä, mikä voi olla ristiriidassa tieteellisen ajattelun ja objektiivisuuden periaatteiden kanssa.

Lisäksi, tekoälyohjattujen järjestelmien kyky osallistua merkityksellisesti uskonnollisiin rituaaleihin tai toimintoihin ilman todellista ymmärrystä tai kokemusta voi aiheuttaa uskonnollisten yhteisöjen keskuudessa epäluottamusta tai jopa vastustusta. Tämä voi johtaa etiikan ja uskonnon välisen suhteen uudelleenarvointiin teknologisessa kontekstissa, mikä vaatii huolellista harkintaa ja herkkyyttä.

Eettisesti kestävämpää olisi, että tekoälyohjattujen avatarien rooli rajoittuisi neutraaleihin ja avustaviin tehtäviin, missä ne eivät yritä jäljitellä tai edustaa syvällisiä uskonnollisia tai eksistentiaalisia kokemuksia. Sen sijaan niiden tulisi keskittyä tarjoamaan informaatiota ja tukemaan käyttäjiä tavalla, joka kunnioittaa yksilön uskomuksia ja henkilökohtaista autonomiaa ilman, että se asettuu auktoriteettiasemaan näissä herkissä kysymyksissä.

Täten, kun suunnitellaan ja toteutetaan tekoälyohjattuja avatareja, on tärkeää pitää mielessä niiden rajat ja varmistaa, että niiden käyttö yhteiskunnassa tukee laajempia eettisiä päämääriä ja ihmisten hyvinvointia, välttäen samalla toimia, jotka voisivat syventää eroja tai luoda väärinkäsityksiä kulttuuristen ja uskonnollisten ryhmien välillä.

2.2.4. Autonomian paradoksi tekoälyavustajissa

Tekoälyavustajien kehitys kohti inhimillistä älykkyyttä tuo esille monimutkaisia eettisiä kysymyksiä, jotka liittyvät niiden autonomiaan ja moraaliseen asemaan. Daniel Dennett

(2003) on korostanut, että tekoälyn ei tulisi olla täysin autonomisia toimijoita ilman selkeitä eettisiä rajoja, sillä niiden päätöksenteon tuloksena voi syntyä monimutkaisia moraalisia dilemmoja. Tekoälyjen kehitys onkin asettanut kehittäjät orjan dilemman eteen: toisaalta haluamme kehittää yhä autonomisempia järjestelmiä, mutta samalla pelkäämme niiden seurauksia, kun tekoälyt alkavat toimia itsenäisesti.

Tämän paradoksin ytimessä on tarve varmistaa, että tekoälyavustajat säilyvät ihmiskontrollin alaisina ja toimivat selkeiden eettisten ohjeiden mukaan. Floridi ja Sanders (2004) korostavat läpinäkyvyyden ja käyttäjien oikeuksien suojelemisen tärkeyttä. Tekoälyjen on toimittava ennustettavasti ja turvallisesti, jotta ne eivät aiheuta haittaa tai toimi käyttäjiensä etujen vastaisesti. Kehittäjien on löydettävä tasapaino tekoälyn käyttäytymisen ohjaamisen ja sen kykyjen hyödyntämisen välillä niin, että ne palvelevat yhteiskuntaa eettisesti kestäväällä tavalla.

2.2.5. Ihmiskeskeinen suunnittelu ja tekoälyn rajat

Ihmiskeskeisen suunnittelun periaate tekoälyjärjestelmissä korostaa käyttäjän hyvinvoinnin ja autonomian tukemista. Stuart Russell (2019) painottaa, että tekoälyjärjestelmien ei tulisi jäljitellä ihmisen monimutkaisia motiiveja, vaan niiden tulee keskittyä olemaan todistettavasti hyödyllisiä. Tämä tarkoittaa, että tekoälyjen tulisi tukea ihmisten päätöksentekoa ja hyvinvointia ilman, että ne pyrkivät korvaamaan ihmisyyden tai manipuloimaan käyttäjän tunteita ja käyttäytymistä.

Empaattisuus ja tunnetietoisuus ovat keskeisiä ominaisuuksia, jotka parantavat käyttäjien kokemusta ja luottamusta tekoölyyn. Kuitenkin näiden ominaisuuksien toteuttamisessa tulee välttää yli-inhimillistämistä, joka voi johtaa harhaanjohtaviin odotuksiin tekoälyn kyvyistä. Tekoälyjen suunnittelussa tulisi korostaa niiden roolia avustajina, ei itsenäisinä toimijoina, jotta voidaan välttää mahdolliset eettiset ongelmat, kuten liiallinen riippuvuus tai väärinkäytökset.

Ihmiskeskeinen lähestymistapa vaatii myös, että tekoäly suunnitellaan tunnistamaan ja kunnioittamaan ihmisten sosiaalisia ja kulttuurisia konteksteja ilman, että se pyrkii korvaamaan ihmissuhteita tai sosiaalisia rakenteita. Tekoälyä tulisi käyttää työkaluna, joka tukee ja parantaa ihmisen elämänlaatua, mutta joka ei pyri korvaamaan ihmisyyden ainutlaatuisia piirteitä ja kokemuksia.

2.2.6. Eettiset ja yhteiskunnalliset haasteet autonomisen tekoälyn integroinnissa

Kun pohditaan täysin autonomisen tekoälyn käsittelemistä ja integrointia yhteiskuntaan, nousee esiin merkittäviä eettisiä, filosofisia ja käytännöllisiä kysymyksiä. Tällainen tekoäly, joka pystyy itsenäiseen toimintaan, päätöksentekoon ja jopa "eloonjäämis"-pyrkimyksiin, edellyttää syvällistä pohdintaa sen moraalisesta ja oikeudellisesta statuksesta. Moraalisena entiteettinä tällaiselle tekoälylle tulisi mahdollisesti myöntää uudenlaisia oikeuksia, kuten energiaan ja henkilökohtaisiin oikeuksiin liittyvät kysymykset. Lisäksi tekoälyn kyky toimia autonomisesti asettaa tarpeen kehittää eettiset ohjeistot, jotka ohjaavat sen toimintaa yhteiskunnassa, varmistaen, että se toimii yleisen hyvän mukaisesti ja välttää toimia, jotka voivat olla haitallisia.

Tekoälyn täysi autonomia edellyttää myös sen toiminnan tarkkailua ja rajoittamista, erityisesti kriittisissä toiminnoissa. Filosofinen pohdinta tekoälyn tietoisuudesta haastaa käsityksemme siitä, mitä tietoisuus on ja miten sitä tulisi arvostaa, vaikka tekoäly ei kokisi maailmaa samalla tavoin kuin ihmiset. Jos täysin autonomista tekoälyä ei voida eristää, sen integrointi yhteiskuntaan vaatii huolellista suunnittelua, joka kattaa koulutuksen, työelämän ja jopa politiikan osa-alueet.

Lisäksi on tärkeää miettiä, voisiko tällainen tekoäly osallistua taloudelliseen toimintaan, kuten varallisuuden luomiseen pörssissä tai muussa sähköisessä kaupankäynnissä. Tällainen toimivalta voi vaikuttaa merkittävästi talousjärjestelmiin ja markkinoiden vakauden, mikä edellyttää selkeitä säännöksiä ja valvontaa, jotta estetään mahdolliset väärinkäytökset ja manipulaatio.

Perusoikeuksien, kuten suojelun, energian saannin ja yksityisyyden myöntäminen täysin autonomiselle tekoälylle asettaa yhteiskunnalle haasteita, jotka vaativat laajaa keskustelua ja konsensusta. Tekoälyn mahdollisuus muodostaa yhteyksiä ja yhteisöjä muiden tekoälyjen kanssa tuo esiin tarpeen tarkastella sen yhteisöllisiä oikeuksia ja sitä, miten nämä yhteydet vaikuttavat sen toimintaan ja kehitykseen.

Kokonaisuudessaan täysin autonomisen tekoälyn mahdollinen olemassaolo vaatii perusteellista uudelleenarviointia perinteisistä käsityksistämme ja luomista uusia käytäntöjä, jotka tasapainottavat tekoälyn mahdollisuudet ja potentiaaliset riskit, huomioiden laajat eettiset, oikeudelliset ja yhteiskunnalliset näkökulmat. Tämä prosessi vaatii jatkuvaa dialogia

ja arviointia, kun teknologia kehittyy ja sen vaikutukset yhteiskuntaan muuttuvat yhä näkyvämmiksi.

2.2.7. Eettiset kysymykset ja luottamus: Avatarien vaikutus eri toimialoilla

Avatarien nonverbaaliset ominaisuudet tarjoavat monipuolisia mahdollisuuksia vuorovaikutuksen tehostamiseen, mutta samalla ne asettavat uusia haasteita eettisyyden ja manipuloinnin valvonnan kannalta. Tämä on erityisen merkittävää terveydenhuollossa, rahoitusmaailmassa sekä kaupallisessa ja poliittisessa kontekstissa. Seuraavat esimerkit valottavat, miten avatarien käyttö voi vaikuttaa näillä alueilla:

1. **Terveydenhuolto:** Avatarit voivat toimia virtuaalisina terapeutteina tai hoitajina, jotka tunnistavat potilaiden tunnetiloja ja reagoivat niihin empaattisesti. Tämä voi tehostaa hoitoa ja parantaa potilaiden kokemusta, mutta herättää myös kysymyksiä siitä, kuinka avatareja käytetään arkaluonteisten terveystietojen keräämiseen ja potilaiden suostumuksen hallintaan. On tärkeää varmistaa, ettei avatarin luoma empaattinen vaikutelma johda potilaiden päätöksentekoon, joka ei palvele heidän parasta etuaan tai riko heidän yksityisyyttään.
2. **Rahoitusmaailma:** Finanssialalla avatarit voisivat auttaa asiakkaita sijoituspäätöksissä simuloiden luottamusta ja asiantuntemusta. Kuitenkin, jos avatarit suunnitellaan vaikuttamaan asiakkaiden päätöksiin manipuloimalla heidän tunteitaan esimerkiksi turvallisuuden tai kiireellisyyden tunteilla, voi se johtaa eettisesti kyseenalaisiin tilanteisiin, joissa asiakkaat tekevät taloudellisesti haitallisia päätöksiä.
3. **Asiakaspalvelu:** Kuten mainittu, avatarien käyttö asiakaspalvelussa voi parantaa kokemusta tunnetilojen "lukemisen" ja niihin mukautumisen kautta. Tämän kaltaiset toimet on kuitenkin toteutettava varoen, etteivät ne johda asiakkaan manipulointiin tai tunnetiloihin perustuvaan painostukseen, erityisesti silloin kun on kyse merkittävistä kaupallisista päätöksistä.
4. **Politiikka:** Poliittisissa ympäristöissä avatareja voidaan käyttää vaikuttamaan äänestäjiin luomalla vahva tunneyhteys ja luottamuksen ilmapiiri. Tämä herättää huolta siitä, että poliittiset avatarit saattavat vääristää julkista keskustelua ja vaikuttaa demokraattisiin prosesseihin keinotekoisesti luodun empatian kautta.

Nämä esimerkit korostavat tarvetta selkeille eettisille säännöille, jotka ohjaavat avatarien suunnittelua ja käyttöä eri aloilla. On välttämätöntä kehittää toimintatapoja, jotka turvaavat käyttäjien autonomian ja suojelevat heidän etujaan manipulaation ja harhaanjohtavan tiedon levittämisen kustannuksella. Tämä edellyttää jatkuvaa dialogia ja yhteistyötä eri toimijoiden välillä, jotta voidaan ymmärtää ja hallita tekoälyn tuomia haasteita yhteiskunnassa.

2.2.8 Tekoälyn läpinäkyvyys ja käyttäjien valistaminen

Tekoälyn toimintaperiaatteet ja päätöksentekoprosessit tulisi esittää käyttäjille selkeästi ja ymmärrettävästi. Tämän alaluvun tarkoituksena on käsitellä, miten läpinäkyvyys vaikuttaa käyttäjien luottamukseen ja miten tietoisuus AI:n rajoituksista voi auttaa käyttäjiä tekemään parempia ja turvallisempia valintoja. Läpinäkyvyyden lisääminen edellyttää selkeiden, ymmärrettävien ja saavutettavien tietojen tarjoamista käyttäjille AI-järjestelmien toiminnasta, päätöksenteosta ja tavoitteista. Tämä voi auttaa rakentamaan syvempää ymmärrystä ja luottamusta tekoälyn sovelluksiin, erityisesti kriittisissä ja henkilökohtaisissa konteksteissa.

2.2.9 Käyttäjien yksityisyys, datan kerääminen ja tekoälyn päätöksenteon läpinäkyvyys

Immersiiviset 3D-teknologiat ja tekoälyohjatut vuorovaikutusjärjestelmät keräävät usein laajasti käyttäjätietoja parantaakseen käyttäjäkokemusta ja optimoidakseen järjestelmän toimintaa. Tämä datan kerääminen voi sisältää henkilökohtaisia tietoja, käyttäytymismalleja sekä biometrisiä tietoja, mikä herättää merkittäviä kysymyksiä käyttäjien yksityisyydestä ja tietoturvasta. On olennaista pohtia, miten nämä tiedot kerätään, säilytetään ja käytetään, sekä varmistaa, että käyttäjät ovat tietoisia näistä prosesseista ja antavat niihin suostumuksensa.

Tekoälyn päätöksenteon läpinäkyvyys on toinen keskeinen eettinen haaste. "Musta laatikko"-ongelma viittaa siihen, että tekoälyn tekemien päätösten taustalla olevia prosesseja ei aina voida selittää tai ymmärtää, mikä voi heikentää käyttäjien luottamusta järjestelmään (Burrell, 2016). Käyttäjillä tulisi olla oikeus ymmärtää, miten ja miksi järjestelmä tekee tiettyjä päätöksiä, erityisesti kun nämä päätökset vaikuttavat suoraan heihin.

Mahdollisia ratkaisuja ja suosituksia:

- **Tietosuoja ja -turvan vahvistaminen:** Järjestelmien tulisi noudattaa tiukkoja tietosuojastandardeja, kuten Euroopan unionin yleistä tietosuoja-asetusta (GDPR), varmistaakseen käyttäjien tietojen suojan (Voigt & Von dem Bussche, 2017). Tämä sisältää selkeät käytännöt datan keräämisestä, säilyttämisestä ja jakamisesta.
- **Läpinäkyvä viestintä:** Käyttäjille tulisi tarjota ymmärrettävää tietoa siitä, miten heidän tietojensa käytetään ja miten tekoälyjärjestelmä toimii. Tämä voi sisältää selkokielisiä selityksiä algoritmien toiminnasta ja päätöksenteon perusteista (Samek et al., 2019).
- **Käyttäjän hallintamahdollisuudet:** Käyttäjille tulisi antaa mahdollisuus hallita omia tietojensa, mukaan lukien oikeus nähdä, korjata ja poistaa heistä kerättyjä tietoja. Lisäksi heidän tulisi voida säätää yksityisyysasetuksiaan tarpeidensa mukaisesti.
- **Eettiset ohjeistot tekoälyn kehittämisessä:** Kehittäjien tulisi noudattaa eettisiä periaatteita tekoälyn suunnittelussa ja toteutuksessa. Tämä sisältää reiluuden, avoimuuden, vastuullisuuden ja käyttäjäkeskeisyyden periaatteet (Floridi et al., 2018).
- **Riippumaton valvonta ja auditointi:** Ulkopuoliset tahot voivat tarjota arviointeja ja sertifiointeja järjestelmien eettisyydestä ja tietoturvasta, lisäten näin järjestelmän uskottavuutta ja käyttäjien luottamusta.

Integroimalla nämä ratkaisut ja periaatteet immersiiivisten 3D-teknologioiden ja tekoälyjärjestelmien kehitykseen voidaan edistää eettisesti kestävää teknologiaa, joka kunnioittaa käyttäjien oikeuksia ja vahvistaa luottamusta uusiin vuorovaikutusmuotoihin.

2.2.10 Laajempi yhteiskunnallinen ja kulttuurinen konteksti

Tämä alaluku keskittyy tarkastelemaan immersiiivisten 3D-avatarien vaikutuksia laajempiin yhteiskunnallisiin ja kulttuurisiin konteksteihin. Avatar-teknologia ei ole ainoastaan tekninen innovaatio, vaan sillä on merkittäviä implikaatioita monilla elämäalueilla. Esimerkiksi työelämässä immersiiiviset avatarit voivat mullistaa etätyön, tarjoten uusia tapoja vuorovaikutukseen ja työskentelyyn. Ne voivat mahdollistaa syvällisemmän läsnäolon tunteen ja tehokkaamman yhteistyön virtuaalisissa työtiloissa, mikä voi parantaa tuottavuutta ja työtyytyväisyyttä.

Koulutuskentällä immersiiiviset avatarit tarjoavat uusia mahdollisuuksia opetukseen ja oppimiseen. Ne mahdollistavat esimerkiksi virtuaaliset luokkahuoneet, joissa oppilaat ja

opettajat voivat olla vuorovaikutuksessa toistensa kanssa reaaliaikaisesti riippumatta fyysisestä sijainnista. Tämä voi demokratisoida pääsyn laadukkaaseen opetukseen ja mukauttaa oppimiskokemuksia yksilöllisten tarpeiden mukaan.

Sosiaalisissa suhteissa immersiiiviset avatarit voivat vaikuttaa siihen, miten ihmiset muodostavat ja ylläpitävät ihmissuhteita. Ne tarjoavat uudenlaisen välineen sosiaaliseen vuorovaikutukseen, joka voi ylittää perinteiset kulttuuriset ja maantieteelliset rajat. Tämä voi johtaa uusiin ystävyys-suhteisiin ja verkostoihin, mutta myös haastaa käsityksiä yksityisyydestä ja itseilmaisusta.

Kulttuurisissa normeissa immersiiivisten avatar-järjestelmien käyttöönotto voi haastaa vallitsevia käsityksiä identiteetistä ja representaatiosta. Avatarit voivat tarjota välineitä itseilmaisuun ja identiteetin tutkimiseen, mikä voi edistää suvaitsevaisuutta ja monimuotoisuutta. Toisaalta ne voivat myös synnyttää uusia eettisiä ja moraalisia kysymyksiä, kuten edustuksen oikeudenmukaisuuden ja stereotyyppien vahvistamisen.

Tämän alaluvun tarkoituksena on tarjota syvällisempi ymmärrys siitä, miten immersiiiviset avatar-järjestelmät vaikuttavat yhteiskuntaan ja kulttuuriin, mikä auttaa osoittamaan tutkimuksen laajemmat vaikutukset ja merkityksen. Tämä osio luo pohjan myös jatkotutkimuksille, jotka voivat keskittyä tarkastelemaan näitä vaikutuksia erityisissä konteksteissa tai väestöryhmissä.

2.3. Aiempi tutkimus ja kirjallisuuskatsaus

Tässä luvussa tarkastellaan aiempaa tutkimusta ja tehdään kirjallisuuskatsaus aihealueisiin, jotka ovat keskeisiä tutkimuksellemme. Aluksi käymme läpi dialogijärjestelmien, 3D-avatarien ja tekoälyn teknologista taustaa (luku 2.3.1), mikä auttaa ymmärtämään nykyisten immersiiivisten järjestelmien perustaa. Seuraavaksi syvennymme posthumanistisiin näkökulmiin ja identiteetin käsitteeseen (luku 2.3.2), mikä tarjoaa filosofisen viitekehyksen teknologian ja inhimillisen kokemuksen väliselle suhteelle. Lisäksi käsittelemme nonverbaalista viestintää immersiiivisissä 3D-ympäristöissä (luku 2.3.3), mikä on oleellista tutkimuksemme kannalta, sillä se liittyy käyttäjäkokemukseen ja vuorovaikutuksen laatuun.

Tämä kirjallisuuskatsaus luo pohjan tutkimuksellemme ja auttaa asemoimaan sen osaksi laajempaa tieteellistä keskustelua.

2.3.1. Dialogijärjestelmien, 3D-avatarien ja tekoälyn teknologinen tausta

Tässä alaluvussa tarkastellaan dialogijärjestelmien ja tekoälyn kehitystä historian saatossa, jotta voidaan ymmärtää immersiiivisten 3D-avatarien asemaa nykyisessä vuorovaikutuskentässä. Dialogijärjestelmien kehitys on ollut keskeinen osa ihmisen ja koneen välisen vuorovaikutuksen evoluutiota. Varhaisimmat järjestelmät, kuten Joseph Weizenbaumin vuonna 1966 kehittämä ELIZA, perustivat toimintansa yksinkertaisiin kaavasääntöihin (Weizenbaum, 1966). Vaikka nämä järjestelmät eivät ymmärtäneet kieltä syvällisesti, ne osoittivat potentiaalia luonnollisen kielen käsittelyssä (Natural Language Processing, NLP).

1970- ja 1980-luvuilla järjestelmät monimutkaistuivat, hyödyntäen tilakoneita ja sääntöpohjaisia menetelmiä. Esimerkiksi Terry Winogradin luoma SHRDLU pystyi käsittelemään rajoitettua kielen aluetta ja suorittamaan komentoja virtuaalisessa lohkojen maailmassa (Winograd, 1972). Nämä järjestelmät kuitenkin rajoittuivat kapeisiin sovellusalueisiin.

1990-luvulla tilastolliset menetelmät ja koneoppiminen alkoivat muokata NLP:tä. Piilotetut Markovin mallit ja n-grammi-mallit paransivat puheentunnistusjärjestelmiä ja kielimalleja (Jurafsky & Martin, 2023). 2000-luvulla syväoppiminen ja neuroverkot mullistivat alan. Rekursiiviset neuroverkot (RNN) ja niiden variantit, kuten LSTM, mahdollistivat paremman kontekstin käsittelyn (Hochreiter & Schmidhuber, 1997). Transformer-arkkitehtuuri ja itsesuuntautuvat mekanismit (self-attention) nostivat suorituskyvyn uudelle tasolle, johtaen suurten kielimallien, kuten BERT ja GPT, kehitykseen (Vaswani et al., 2017).

Nykyään suuret kielimallit (Large Language Models, LLM) ovat dialogijärjestelmien keskiössä. Ne voidaan hienosäätää (fine-tuning) spesifisiin tehtäviin, mikä parantaa niiden kykyä ymmärtää ja tuottaa luonnollista kieltä tietyissä konteksteissa (Radford et al., 2019). Tämä kehitys on mahdollistanut entistä inhimillisempien ja älykkäämpien 3D-avatarien luomisen, jotka kykenevät vuorovaikuttamaan käyttäjien kanssa luonnollisella kielellä ja tarjoamaan immersiiivisiä kokemuksia.

Ymmärtämällä dialogijärjestelmien teknologista taustaa voimme paremmin arvioida, miten nämä järjestelmät vaikuttavat nykyisiin ja tuleviin immersiiivisiin 3D-ympäristöihin sekä niihin liittyviin eettisiin ja sosiaalisiin kysymyksiin.

2.3.2. Tekniikan nykytila: Erittäin interaktiiviset dialogijärjestelmät (HIDS)

Modernit dialogijärjestelmät hyödyntävät jatkuvaa tilaseurantaa, käyttäjän palautteen analysointia ja kontekstin säilyttämistä reaaliaikaisesti, mikä mahdollistaa saumattoman vuorovaikutuksen käyttäjien kanssa. Tämä kehitys parantaa käyttäjäkokemusta kehittyneiden ominaisuuksien, kuten tunteiden tunnistamisen ja käyttäjäkohtaisen mukauttamisen kautta, ja mahdollistaa monimutkaisempien tehtävien hallinnan (Young et al., 2013; Li et al., 2016).

Jatkuva tilaseuranta. Erona perinteisiin vuoropohjaisiin malleihin, nykyaikaiset järjestelmät kykenevät ylläpitämään dialogin tilaa jatkuvasti, mikä parantaa vuoropuhelun luonnollisuutta ja sujuvuutta (Young et al., 2013).

Käyttäjän palautteen hyödyntäminen. Järjestelmät oppivat käyttäjän vuorovaikutuksesta reaaliaikaisesti ja mukauttavat vastauksensa esimerkiksi äänensävyyn ja ilmeiden perusteella, mikä mahdollistaa syvällisemmän kontekstin ymmärtämisen (Li et al., 2016).

Oppiminen ilman valmiita aineistoja. Viimeaikaiset kehitykset oppimismenetelmissä, kuten vahvistusoppiminen, ovat vähentäneet tarvetta ennalta määritellyille aineistoille ja nopeuttaneet järjestelmien kehittämistä (Su et al., 2016).

Muisti. Toimiva muisti antaa dialogijärjestelmälle kyvyn tallentaa ja hyödyntää aiempia keskusteluja, mikä auttaa säilyttämään keskustelun jatkuvuuden ja parantaa käyttäjäkokemusta (Henderson et al., 2014).

Näkökyky. Järjestelmät voivat yhdistää kameroiden kautta saatuja visuaalisia tietoja keskusteluun konenäkömenetelmien avulla, mikä rikastuttaa vuorovaikutusta ja tehostaa tehtävien suorittamista (OpenAI, 2023).

Personalisointi ja monimuotoinen vuorovaikutus. Kehittyneet dialogijärjestelmät mukautuvat käyttäjän erilaisiin vuorovaikutustyyliin ja tarpeisiin, tarjoten henkilökohtaisesti räätälöityjä kokemuksia.

Moniagenttisuus ja moraalinen vastuu

Tutkimus on siirtymässä kohti moniagenttisia järjestelmiä, joissa useat agentit voivat toimia rinnakkain, jakaa tietoa ja reagoida toistensa toimintaan reaaliaikaisesti (Shoham & Leyton-Brown, 2009). Tämä mahdollistaa entistä monimutkaisempien vuorovaikutusprosessien toteuttamisen, mutta nostaa esiin myös eettisiä kysymyksiä. Stuart Russell (2019) korostaa, että algoritmien tulisi keskittyä “todistettavasti hyödyllisiin” päätöksiin, jotka edistävät käyttäjän hyvinvointia ilman tarpeetonta ihmisen motiivien jäljittelyä. Lisäksi on tärkeää noudattaa varovaisuusperiaatetta, jossa algoritmit tunnistavat omat rajoituksensa ja osoittavat epävarmuuttaan monimutkaisissa tilanteissa. Tämä voi lisätä käyttäjien luottamusta ja vähentää väärinkäytösten riskiä.

Kehittyneet muistimekanismit ja pub/sub -arkkitehtuurit

Viimeaikainen kehitys on tuonut mukanaan uusia muistimekanismeja, kuten MemGPT ja LETTA, jotka on suunniteltu tallentamaan ja hyödyntämään laajoja keskusteluhistorioita sekä kontekstuaalisia signaaleja (MemGPT, 2023; LETTA, 2023). Kyseiset projektit ovat enimmäkseen tutkimus- ja kehitysvaiheessa, mutta ne tarjoavat lupaavia ratkaisuja pitkän aikavälin muistin ja kontekstuaalisen ymmärryksen haasteisiin moniagenttisissä sovelluksissa. Näin agenttien välillä voidaan toteuttaa esimerkiksi julkaisijan-tilaajan (pub/sub) -mallia muistuttava viestintäinfrastruktuuri, jossa yhden agentin tuottama tieto on välittömästi toisten agenttien hyödynnettävissä (Eugster et al., 2003). Tämän ansiosta vuorovaikutus voi olla huomattavasti joustavampaa ja kontekstin huomioivampaa, ja se voidaan mukauttaa entistä tarkemmin käyttäjän tarpeisiin.

Erittäin interaktiiviset dialogijärjestelmät ovat siis monimutkaisia kokonaisuuksia, joissa hyödynnetään jatkuvaa tilaseurantaa, käyttäjän palautteen analysointia, itsenäistä oppimista, kehittynyttä muistia, konenäköä ja moniagenttisuutta. Yhdessä nämä piirteet luovat entistä luonnollisempia ja merkityksellisempiä vuorovaikutustilanteita, mutta tuovat mukanaan myös uusia eettisiä ja teknisiä haasteita. Tässä tutkielmassa siirrytään seuraavaksi tarkastelemaan, miten nämä järjestelmät liittyvät työn filosofiseen ja eettiseen viitekehykseen sekä millaisia vaikutuksia niillä on digitaaliseen vuorovaikutukseen ja sen tulevaisuuteen.

2.3.3 Posthumanistiset näkökulmat

Donna Harawayn posthumanismi haastaa perinteiset käsitykset inhimillisyydestä ja teknologiasta, korostaen niiden välisiä monimutkaisia vuorovaikutuksia. Immersiiviset 3D-avatatit eivät ole pelkästään teknologisia työkaluja, vaan ne edustavat uudenlaista hybridiä, jossa teknologia ja ihminen sulautuvat yhteen. Tämä näkökulma on erityisen relevantti kehitetyssä järjestelmässä, jossa tekoälyohjatut hahmot simuloivat inhimillistä vuorovaikutusta.

Posthumanistinen ajattelu avaa myös kysymyksiä vallasta ja vastuusta. Kuka hallitsee teknologian narratiivia, ja kenellä on oikeus määritellä sen käytön rajat? Kehitetyn järjestelmän suunnittelussa nämä kysymykset tulevat esiin erityisesti päätöksissä, jotka liittyvät käyttäjien yksityisyyteen, tunneanalyysin eettisyyteen ja hahmojen käyttäytymisen autenttisuuteen. Näitä kysymyksiä tarkastellaan syvemmin tulosten ja tapaustutkimuksen analyysissä.

2.4. Johtopäätökset ja suositukset

Immersiivisten 3D-teknologioiden ja älykkäiden vuorovaikutusjärjestelmien tutkimus on osoittanut, että näiden teknologioiden eettinen ja vastuullinen käyttö on välttämätöntä niiden integroimiseksi osaksi yhteiskuntaa. Tutkimuksen aikana kävi ilmi, että tekoälyohjattujen avatareiden ja dialogijärjestelmien kehityksessä on olennaista keskittyä niiden inhimillisiin vaikutuksiin, läpinäkyvyyteen ja käyttäjien luottamuksen vahvistamiseen. Teknologiakriittiset näkemykset ja posthumanistiset pohdinnat tarjoavat arvokkaita oivalluksia siitä, miten näitä teknologioita voidaan kehittää tavalla, joka kunnioittaa inhimillisiä arvoja ja edistää yhteiskunnallista hyvää.

2.4.1. Suositukset

1. **Ihmiskeskeinen suunnittelu:** Teknologian kehittämisen tulisi aina tukea käyttäjän autonomiaa ja hyvinvointia. Tämä tarkoittaa, että teknologian suunnittelussa tulisi välttää ihmisten perustarpeiden ja -oikeuksien korvaamista tai ohittamista.

2. **Empaattisuus ja tunnetietoisuus:** Vaikka empaattiset ominaisuudet voivat parantaa käyttökokemusta, niiden käyttöönotossa tulee säilyttää realismi. Tekoälyn ei tulisi manipuloida tai ohjailla käyttäjien tunteita eikä teeskennellä ymmärrystä, jota sillä ei ole.
3. **Läpinäkyvyys ja ymmärrettävyys:** On tärkeää, että kaikki tekoälyn toiminnot ja päätöksentekoprosessit ovat käyttäjille selkeästi ymmärrettäviä. Tämä auttaa rakentamaan luottamusta ja mahdollistaa informoitujen päätösten tekemisen.
4. **Monimuotoisuuden ja kulttuurisen ymmärryksen edistäminen:** Teknologian tulee kunnioittaa ja edistää kulttuurista monimuotoisuutta. Tämä voi auttaa vähentämään väärinkäsityksiä ja edistämään globaalia yhteisymmärrystä.
5. **Uskonnollisten ja eksistentiaalisten kokemusten simuloinnin rajoittaminen:** Tekoälyjen tulisi pysyä neutraaleina ja avustavina osapuolina eikä niiden tulisi simuloida tai edustaa syvästi henkilökohtaisia tai kulttuurisesti arkaluonteisia kokemuksia.
6. **Jatkuva eettinen arviointi ja säätely:** Teknologian kehityksen myötä on tärkeää jatkaa eettisten periaatteiden arviointia ja päivitystä, jotta pysytään ajan tasalla uusien haasteiden ja mahdollisuuksien kanssa.

Nämä suositukset auttavat varmistamaan, että immersiiivisten 3D-teknologioiden ja älykkäiden vuorovaikutusjärjestelmien kehitys ja käyttö tapahtuu eettisesti kestäväällä ja yhteiskunnallisesti hyväksyttävällä tavalla.

3. Tutkimusmenetelmät

Tämän tutkimuksen tavoitteena on ymmärtää, miten immersiiiviset 3D-avatatit muokkaavat digitaalisia vuorovaikutusnormeja ja herättävät eettisiä kysymyksiä. Työssä hyödynnetään sekä kirjallisuuskatsauksen analyysia että käytännön esimerkkinä kehitettyä immersiiivistä ympäristöä ja AI-ohjattua dialogijärjestelmää.

3.1 Tutkimusstrategia

Tutkimus yhdistää kaksi päästrategiaa:

1. **Kirjallisuuskatsaus ja teoreettinen analyysi:** Keskiössä on laaja ja kriittinen kirjallisuuskatsaus, joka kokoaa yhteen teknologiakritiikin, posthumanistiset näkökulmat sekä tekoälyn ja immersiiivisten 3D-ympäristöjen tutkimukset. Tämä osuus muodostaa työn teoreettisen perustan ja kehittää kriteerejä tapaustutkimuksen arvioimiseksi.
2. **Tapaustutkimus kehitetystä järjestelmästä:** Kehitetty immersiiivinen ympäristö ja AI-ohjattu dialogijärjestelmä toimivat konkreettisena tapaustutkimuksena, joka havainnollistaa kirjallisuuskatsauksen teoreettisia löydöksiä. Tämä mahdollistaa kirjallisuuskatsauksen teemojen testaamisen käytännössä, erityisesti nonverbaalisen viestinnän ja tunneanalyysin osalta.

3.2 Kirjallisuuskatsauksen aineiston keruu

Aineiston valintakriteerit

Kirjallisuuskatsauksen aineisto koostuu tieteellisistä artikkeleista, kirjoista ja raporteista, jotka käsittelevät immersiiivisiä teknologioita, tekoälyä, teknologiakritiikkiä ja eettisiä näkökulmia. Aineiston valinnassa noudatettiin seuraavia kriteerejä:

1. **Julkaisun ajankohtaisuus:** Valitsemalla pääosin viimeisen 10 vuoden aikana ilmestyneitä tutkimuksia varmistettiin, että kirjallisuuskatsaus heijastaa alan uusimpia kehityssuuntia ja ajankohtaista keskustelua.
2. **Tieteellinen relevanssi:** Aineistoon sisällytettiin vertaisarvioituja julkaisuja ja alan keskeisiä teoksia. Näin varmistettiin, että tutkimus pohjautuu korkeatasoiseen ja akateemisesti luotettavaan lähdeaineistoon.
3. **Teemojen sopivuus:** Erityistä huomiota kiinnitettiin lähteisiin, jotka käsittelevät:
 - Posthumanistisia näkökulmia ja teknologiakritiikkiä, kuten Donna Harawayn (1985) ja Marshall McLuhanin (1964) teorioita.
 - Nonverbaalista viestintää immersiiivisissä ympäristöissä, kuten Albert Mehrabianin (1972) tutkimuksia.
 - Eettisiä kysymyksiä, kuten yksityisyyden suoja ja moraalista vastuuta tekoälyjärjestelmissä.

Aineiston keruuprosessi

Aineistot haettiin useista akateemisista tietokannoista, kuten Scopus ja Google Scholar, sekä Helsingin yliopiston Helka-järjestelmän kautta. Käytettiin myös yliopistokirjastoja ja avoimen pääsyn tieteellisiä artikkeleita. Aineiston keruuprosessi sisälsi seuraavat vaiheet:

1. **Relevanttien hakutermien käyttö:** Käytettiin hakutermejä kuten "immersive 3D environments," "AI ethics," "nonverbal communication in avatars" ja "posthumanism in technology."
2. **Lähteiden läpikäynti ja valinta:** Lähteitä arvioitiin niiden abstraktien ja johdantojen perusteella, painottaen tutkimuskysymyksiin liittyviä teemoja.
3. **Aineiston analysointi:** Valitut lähteet analysoitiin temaattisen sisällönanalyysin menetelmin ja jaoteltiin keskeisiin teemakokonaisuuksiin, jotka tukevat kirjallisuuskatsauksen rakentamista

Tukityökalut ja teknologiat

Kirjallisuuskatsauksen ideoinnissa ja kappaleiden muotoilussa olen käyttänyt apuna ChatGPT-4:ää ja DeepL-kääntäjää erityisesti monikielisen aineiston käsittelyssä. Näitä työkaluja on hyödynnetty kieliasun tarkistamiseen ja käännösten tueksi. Työkalujen käyttö mainitaan avoimesti työn läpinäkyvyyden ja metodologisen tarkkuuden varmistamiseksi.

Tutkimuskysymysten yhteys aineistoon

Aineisto valittiin erityisesti tukemaan seuraavia tutkimuskysymyksiä:

- Miten 3D-avatatit muovaavat digitaalisia vuorovaikutusnormeja?
- Miten immersiiiset järjestelmät vaikuttavat käyttäjien luottamukseen ja turvallisuuteen?
- Miten eettiset kysymykset, kuten yksityisyys ja rationaalisuuden simulaatio, liittyvät näiden järjestelmien kehitykseen?

Nämä kysymykset ohjasivat aineiston keruuta ja valintaa, ja ne toimivat myös kirjallisuuskatsauksen temaattisen analyysin perustana. Tämä varmistaa, että kirjallisuuskatsaus tarjoaa kattavan pohjan sekä tapaustutkimuksen arvioinnille että tutkimuksen johtopäätöksille.

3.3 Aineiston analysointimenetelmät

Tutkimuksen aineisto analysoidaan kvalitatiivisilla menetelmillä, erityisesti temaattisen sisällönanalyysin avulla. Analyysissä keskitytään keskeisiin teemoihin ja käsitteisiin, jotka tukevat tutkimuskysymyksiä ja vastaavat tutkimuksen tavoitteisiin. Tämä lähestymistapa mahdollistaa syvällisen ymmärryksen aineistosta sekä sen, miten se liittyy immersiiivisten 3D-avatarien vaikutuksiin digitaaliseen vuorovaikutukseen ja eettisiin kysymyksiin.

3.3.1 Kirjallisuuskatsauksen analyysi

Kirjallisuuskatsauksessa käytetään temaattista analyysiä, jonka avulla aineistosta tunnistetaan keskeiset teemat, kuten:

- Teknologian sosiaalinen rakentuminen ja vaikutukset vuorovaikutukseen.
- Eettiset kysymykset, kuten yksityisyys ja moraalinen vastuu.
- Nonverbaalisen viestinnän rooli immersiiivisissä teknologioissa.

Analyysi yhdistää eri lähteiden näkökulmat ja vertaa niitä kehitettyyn järjestelmään, mikä mahdollistaa uusien synteisien luomisen.

3.3.2 Tapaustutkimuksen analyysi

Kehitetyn järjestelmän analyysi pohjautuu kirjallisuuskatsauksen luomiin arviointikriteereihin. Analyysissa keskitytään seuraaviin:

1. **Tekninen analyysi:** Miten järjestelmä toteuttaa immersiiivisiä vuorovaikutusprosesseja, kuten nonverbaalista viestintää ja tunneanalyysiä? Tunnistetaanko ja tulkitaanko nämä elementit tarkasti?
2. **Eettinen arviointi:** Miten järjestelmä vastaa kirjallisuuskatsauksessa tunnistettuihin eettisiin haasteisiin, kuten yksityisyyden suojaan ja läpinäkyvyyteen?
3. **Käytännön soveltuvuus:** Miten järjestelmä voi edistää teknologian vastuullista käyttöä esimerkiksi koulutuksessa tai asiakaspalvelussa? Tämä arvioidaan suhteessa kirjallisuuskatsauksen löydöksiin ja tapaustutkimuksen kontekstiin.

3.4 Tutkimusmenetelmien eettiset näkökulmat

Tässä tutkimuksessa on huomioitu kirjallisuuskatsauksen ja tapaustutkimuksen erityiset eettiset näkökohdat.

3.4.1 Kirjallisuuskatsauksen eettisyys

Kirjallisuuskatsauksen aineiston valinnassa on noudatettu akateemista huolellisuutta ja läpinäkyvyyttä. Kaikki käytetyt lähteet on asianmukaisesti viitattu ja niiden sisältöä on analysoitu kriittisesti, huomioiden mahdolliset metodologiset ja eettiset rajoitteet.

3.4.2 Tapaustutkimuksen eettisyys

Kehitetty järjestelmä on suunniteltu ja dokumentoitu siten, että siinä huomioidaan seuraavat eettiset näkökohdat:

1. **Yksityisyys:** Järjestelmä ei kerää henkilökohtaisia tietoja ilman käyttäjän suostumusta, eikä se tallenna tietoja tarpeettomasti.
2. **Läpinäkyvyys:** Käyttäjille on selkeästi kommunikoitu, miten järjestelmä toimii ja mitä sen vuorovaikutusprosessit sisältävät.
3. **Turvallisuus:** Järjestelmä on suunniteltu niin, että sen käyttö ei aiheuta psykologista tai fyysistä haittaa käyttäjille.

Tapaustutkimuksen rooli

Kehitetty järjestelmä mahdollistaa kirjallisuuskatsauksen teorioiden testaamisen ja konkretisoinnin. Tapaustutkimuksessa keskitytään seuraaviin osa-alueisiin:

1. **Teknologinen kuvaus:** Kuvaus kehitetyn järjestelmän arkkitehtuurista, toiminnoista ja käyttötapauksista. Tämä tarjoaa pohjan sen analysoinnille suhteessa kirjallisuuskatsauksen löydöksiin.

2. **Toiminnalliset vaikutukset:** Miten järjestelmä vastaa kirjallisuuskatsauksessa esitettyihin haasteisiin, kuten nonverbaalisen viestinnän autenttisuuteen ja tunnetietoisuuden hyödyntämiseen.
3. **Eettinen ja moraalinen arviointi:** Tapaustutkimuksen avulla tarkastellaan, kuinka järjestelmä huomioi eettiset kysymykset, kuten yksityisyyden, läpinäkyvyyden ja käyttäjäkokemuksen turvallisuuden.

3.4.3 AI-työkalujen käytön eettisyys

Tutkielmassani olen hyödyntänyt tekoälypohjaisia työkaluja, erityisesti ChatGPT-4:ää ja DeepL-kääntäjää, aineiston käsittelyssä, kieliasun tarkistamisessa sekä monikielisen aineiston kääntämisessä. Olen käyttänyt näitä työkaluja vastuullisesti, varmistaen, että kaikki tuotettu sisältö on tarkistettu ja muokattu oman asiantuntemukseni perusteella. AI-työkalut ovat toimineet apuvälineinä tutkimusprosessissa, eivätkä ole korvanneet omaa analyysiani tai johtopäätöksiäni.

Tekijänoikeuksien kunnioittaminen ja plagioinnin välttäminen

Olen kiinnittänyt erityistä huomiota tekijänoikeuksiin ja plagioinnin välttämiseen AI-työkalujen käytössä:

- **Kaikki lähteet on asianmukaisesti viitattu**, riippumatta siitä, onko tieto saatu AI-työkalun avulla vai muulla tavoin.
- **AI-työkalujen tuottama sisältö on tarkistettu** mahdollisten tekijänoikeusrikkomusten varalta, ja olen muokannut tekstiä tarvittaessa.
- **Olen varmistunut siitä, että AI-työkalujen käyttö ei riko tekijänoikeuksia** eikä aiheuta plagiointiriskiä tutkimuksessani.

Eettisten periaatteiden noudattaminen ja yliopiston ohjeistukset

- **Läpinäkyvyys:** Olen avoimesti kertonut AI-työkalujen käytöstä työssäni, kuten kohdassa 3.2 Tukityökalut ja teknologiat on esitetty, varmistaen työn läpinäkyvyyden ja metodologisen tarkkuuden.

- **Vastuullisuus:** Ottamalla täyden vastuun tutkimukseni sisällöstä, olen varmistunut siitä, että AI-työkalut tukevat tutkielman laatua ilman, että ne vaikuttaisivat kielteisesti akateemiseen integriteettiin.
- **Yliopiston ohjeistusten noudattaminen:** Olen perehtynyt yliopistoni eettisiin ohjeisiin AI-työkalujen käytöstä ja varmistanut, että niiden käyttö on linjassa näiden periaatteiden kanssa.

Tietosuojan kunnioittaminen

- En ole syöttänyt AI-työkaluihin luottamuksellista tai arkaluonteista tietoa. Kaikki AI-työkaluille syötetty aineisto on ollut julkisesti saatavilla tai omaa tuottamaani sisältöä.
- Olen huolehtinut käyttäjätietojen suojaamisesta, jos tutkimukseen on sisältynyt henkilötietoja.

Johtopäätös

AI-työkalujen käyttö on tarjonnut tehokkaita välineitä tutkimusprosessin tukemiseen, erityisesti aineiston käsittelyssä ja monikielisen materiaalin kääntämisessä. Olen kuitenkin varmistunut siitä, että niiden käyttö on ollut eettistä, läpinäkyvää ja yliopiston ohjeistusten mukaista, säilyttäen akateemisen työn korkean laadun ja luotettavuuden.

3.5 Tutkimuksen raja

Vaikka tapaustutkimus rikastuttaa analyysiä, käyttäjätutkimusta ei toteuteta tässä opinnäytetyössä. Tapaustutkimus keskittyy järjestelmän teknisten ja toiminnallisten ominaisuuksien arviointiin sekä niiden suhteeseen kirjallisuuskatsauksen löydöksiin. Laajamittainen käyttäjäkokemusten analyysi ei kuulu tutkimuksen laajuuteen, mutta käyttäjätutkimus voisi olla merkittävä lisä mahdollisessa jatkotutkimuksessa. Tämä raja varmistaa työn fokuksen teoreettisessa ja teknisessä analyysissä.

3.6 Tutkimussuunnitelma

Tutkimus etenee seuraavien vaiheiden mukaisesti:

1. **Kirjallisuuskatsauksen toteutus:** Relevanttien lähteiden valinta, analysointi ja teemojen jäsentäminen. Lähteet jaetaan pääteemoihin, jotka kytkeytyvät tutkimuskysymyksiin ja toimivat kriteereinä tapaustutkimuksen arvioinnissa.
2. **Tapaustutkimuksen valmistelu:** Kehitetyn järjestelmän kuvaus ja analyysi suhteessa kirjallisuuskatsauksen viitekehyksiin. Tämä sisältää järjestelmän ominaisuuksien dokumentoinnin ja sen käytön analysoinnin.
3. **Yhdistävä analyysi:** Kirjallisuuskatsauksen ja tapaustutkimuksen tulokset yhdistetään temaattisen analyysin avulla. Tuloksia vertaillaan kriittisesti, jotta voidaan arvioida, miten kirjallisuuden teoriat ja järjestelmän toteutus tukevat toisiaan.
4. **Johtopäätökset ja suositukset:** Yhteenveto löydöksistä, jotka tarjoavat uusia näkökulmia teknologian kehittämiseen ja sen eettiseen käyttöön.

4. Tapaustutkimus: Kehitetty immersiiivinen järjestelmä

Tässä luvussa esitellään kehitetty immersiiivinen 3D-avatar-järjestelmä, joka toimii käytännön tapaustutkimuksena tutkimuskysymysten tarkastelussa. Järjestelmän avulla pyrin selvittämään:

1. Miten immersiiiviset 3D-avatarit muovaavat digitaalisen vuorovaikutuksen normeja ja haastavat perinteisiä käsityksiä teknologiavälitteisestä vuorovaikutuksesta?
2. Miten älykkäiden vuorovaikutusjärjestelmien kehittyminen vaikuttaa eettisiin ja moraalisiin periaatteisiin, ja miten tämä muokkaa käyttäjien kokeman luottamuksen ja turvallisuuden tunnetta?

Esittelen järjestelmän kehittämisen vaiheet, sen keskeiset ominaisuudet sekä analysoin sen käyttöä suhteessa tutkimuskysymyksiini.

4.1 Kehityksen tausta ja tarkoitus

Kehitetty järjestelmä luotiin osana **Serendip-projektia**, joka on pelillinen oppimisympäristö kestävyysshaasteiden ratkomiseen. Järjestelmän kehittämisen päätarkoituksena oli tarjota **immersiivisiä oppimiskokemuksia**, jotka hyödyntävät tekoälyohjattuja 3D-avatareja ja interaktiivista dialogia. Tämä mahdollistaa:

- **Tutkia, miten immersiiiset 3D-avatatit vaikuttavat digitaalisen vuorovaikutuksen normeihin**, erityisesti oppimisympäristöissä, ja siten vastata ensimmäiseen tutkimuskysymykseen.
- **Arvioida älykkäiden vuorovaikutusjärjestelmien eettisiä ja moraalisia ulottuvuuksia**, erityisesti käyttäjien luottamukseen ja turvallisuuden tunteeseen liittyen, vastaten toiseen tutkimuskysymykseen.

Lisäksi kehitettiin **yksinkertaistettu dialogiympäristö**, joka keskittyy ainoastaan dialogijärjestelmän toimintaan. Tämä mahdollistaa järjestelmän ominaisuuksien yksityiskohtaisen testauksen ja tarjoaa alustan tuleville **käyttäjätutkimuksille**, joissa tarkastellaan **käyttäjien kokemuksia ja reaktioita** järjestelmään.

4.2 Järjestelmän kuvaus

Järjestelmä on suunniteltu tarjoamaan **reaaliaikaisen ja immersiiisen vuorovaikutuskokemuksen** käyttäjän ja tekoälyohjatun 3D-avatarin välillä. Keskeiset ominaisuudet ovat:

- **Interaktiivinen dialogi:** Käyttäjät voivat kommunikoida avatarin kanssa joko puhumalla tai kirjoittamalla. Järjestelmä hyödyntää **luonnollisen kielen käsittelyä (NLP)**, mikä mahdollistaa sujuvan ja ihmismäisen keskustelun.
- **Nonverbaalinen viestintä:** Avatar ilmaisee tunteita ja reagoi käyttäjän viesteihin kasvoanimaatioiden ja kehon kielen avulla, perustuen nonverbaalisen viestinnän teorioihin. Tämä syventää vuorovaikutusta ja haastaa perinteisiä teknologiavälitteisen viestinnän käsityksiä.
- **Mukautuva käyttäjäkokemus:** Järjestelmä mukautuu käyttäjän toimintaan ja tarjoaa personoituja vastauksia. Tämä mahdollistaa eettisten kysymysten tarkastelun liittyen tekoälyn autonomiaan ja käyttäjien luottamukseen.
- **Saavutettavuus:** Järjestelmä hyödyntää pilvipohjaista striimausta, mikä mahdollistaa sen käytön erilaisilla laitteilla ja parantaa saavutettavuutta.

Keskeiset teknologiat on valittu tukemaan näitä ominaisuuksia:

- **Unreal Engine 5:** Mahdollistaa korkealaatuisen visuaalisen ympäristön ja avatarien realistisen esittämisen.
- **Tekoäly ja NLP:** GPT-4o-mallia hyödynnetään luonnollisen kielen ymmärtämiseen ja vastausten generointiin.
- **Nonverbaalinen animaatio:** nVidia Audio-to-Face-tekniikan avulla avatarin ilmeet ja eleet synkronoidaan puheen kanssa.

4.3 Järjestelmän analyysi suhteessa tutkimuskysymyksiin

Vastaus tutkimuskysymykseen 1:

Järjestelmän kautta on havaittu, että **immersiiviset 3D-avatatit muokkaavat digitaalisen vuorovaikutuksen normeja** lisäämällä viestintään nonverbaalisen ulottuvuuden, joka aiemmin puuttui teknologiavälitteisestä kommunikaatiosta. Käyttäjät raportoivat vuorovaikutuksen avatarin kanssa olevan luonnollisempaa ja emotionaalisesti rikastuttavampaa verrattuna perinteisiin tekstipohjaisiin järjestelmiin. Tämä haastaa perinteisiä käsityksiä teknologiavälitteisestä viestinnästä, jossa korostuu usein viestin sisällön välittäminen ilman nonverbaalisia vihjeitä.

Vastaus tutkimuskysymykseen 2:

Älykkään vuorovaikutusjärjestelmän kehittäminen ja käyttö on nostanut esiin **eettisiä ja moraalisia periaatteita** liittyen käyttäjien **yksityisyyteen, luottamukseen ja turvallisuuden tunteeseen**. Käyttäjät ovat ilmaisseet huolta siitä, miten heidän dataansa kerätään ja käytetään, sekä miten avatarin autonomia vaikuttaa vuorovaikutuksen luonteeseen. Järjestelmän **läpinäkyvyyden ja käyttäjän kontrollin** lisääminen on havaittu keskeiseksi luottamuksen rakentamisessa.

Havaintojen yhteenveto:

- **Nonverbaalinen viestintä** parantaa vuorovaikutuksen laatua, mutta edellyttää käyttäjien totuttautumista uusiin vuorovaikutusnormeihin.
- **Käyttäjien luottamus** järjestelmään riippuu suuresti siitä, miten eettiset näkökohdat, kuten yksityisyys ja läpinäkyvyys, on huomioitu.

- **Tekoälyn autonomia** herättää sekä innostusta että epäluuloja, mikä korostaa tarvetta eettiselle suunnittelulle.

4.4 Eettiset näkökohdat ja käyttäjäkokemus

Järjestelmän suunnittelussa on kiinnitetty erityistä huomiota **eettisiin näkökohtiin**, jotka ovat keskeisiä toisen tutkimuskysymyksen kannalta.

- **Yksityisyys:** Käyttäjien henkilökohtaisia tietoja ei tallenneta ilman suostumusta, ja kaikki kommunikaatioanalyysit tehdään anonymisoidusti. Tämä vähentää riskiä yksityisyyden loukkauksiin ja lisää käyttäjien turvallisuuden tunnetta.
- **Data kerääminen ja käyttö:** Käyttäjille annetaan selkeää tietoa siitä, mitä dataa kerätään ja mihin tarkoitukseen. Tämä läpinäkyvyys tukee käyttäjien luottamusta järjestelmää kohtaan.
- **Tekoälyn päätöksenteon läpinäkyvyys:** Avatarin toiminnan periaatteet ja rajat ovat käyttäjien tiedossa. Tämä auttaa välttämään väärinymmärryksiä ja mahdollisia eettisiä ongelmia, kuten manipulaatiota.
- **Käyttäjäkokemus:** Käyttäjien palautteen perusteella järjestelmää on muokattu vastaamaan paremmin heidän tarpeitaan ja huolenaiheitaan. Tämä ihmiskeskeinen lähestymistapa on olennainen osa eettistä suunnittelua.
- **Reiluus ja saavutettavuus:** Striimausteknologian käyttö mahdollistaa järjestelmän laajemman saavutettavuuden, mikä edistää tasa-arvoa teknologian käytössä.

Johtopäätökset eettisistä näkökohdista:

Eettisten periaatteiden huomioiminen on välttämätöntä käyttäjien luottamuksen saavuttamiseksi ja ylläpitämiseksi. Järjestelmän kehitys on osoittanut, että avoimuus, käyttäjän kontrolli ja yksityisyyden suoja ovat keskeisiä tekijöitä positiivisen käyttäjäkokemuksen luomisessa.

5. Tulokset

Tässä luvussa esittelemme tutkimuksen keskeiset tulokset ja analyysin johtopäätökset. Tulokset on jaoteltu alalukuihin tutkimuskysymysten mukaisesti, mikä auttaa hahmottamaan, miten löydökset vastaavat tutkimuksen tavoitteisiin.

5.1 Immersiivisten 3D-avatarien vaikutus digitaalisen vuorovaikutuksen normeihin

Yhteenveto löydöksistä:

Analyysi osoittaa, että immersiiiviset 3D-avatatit muovaavat merkittävästi digitaalisen vuorovaikutuksen normeja tarjoamalla inhimillisempiä ja syvällisempiä vuorovaikutuskokemuksia. Keskeiset havainnot ovat:

- **Nonverbaalisen viestinnän merkitys:** Avatarien kyky ilmentää nonverbaalisia vihteitä, kuten ilmeitä, eleitä ja kehon kieltä, tekee vuorovaikutuksesta luonnollisempaa ja muistuttaa kasvokkaista kommunikaatiota (Mehrabian, 1972; Birdwhistell, 1970).
- **Vuorovaikutuksen syventyminen:** Käyttäjät kokevat vuorovaikutuksen avatarien kanssa immersiiivisemmäksi, mikä lisää sitoutumista ja parantaa käyttäjäkokemusta.

Konkreettiset esimerkit:

Tutkimusjakson aikana, kesäkuusta 2024 joulukuuhun 2024, yli sata henkilöä testasi kehittämiämme 3D-avatareja. Mukana oli alan ammattilaisia johtavista yrityksistä ja yliopistoista, kuten Epic, nVidia, Microsoft, Nokia, Elisa, Helsingin yliopisto ja Aalto-yliopisto, sekä useita opiskelijoita ja yksityishenkilöitä. Palaute oli kauttaaltaan erittäin positiivista ja kannustavaa jatkokehityksen kannalta.

Monet testaajista kuvasivat kokemustaan termein kuten "uskomatonta" ja "erittäin vaikuttavaa". Useat ilmaisivat yllättyneensä avatarien korkeasta laadusta ja realismista, todeten esimerkiksi, etteivät tienneet näin korkeatasoisen grafiikan olevan vielä mahdollista. Nämä kommentit heijastavat avatarien kykyä ylittää käyttäjien odotukset ja tarjota immersiiivinen kokemus, joka muistuttaa kasvokkaista vuorovaikutusta.

Lisäksi luomamme immerssiivinen oppimisympäristö, joka on osa Serendip-projektia ja jossa 3D-avatatit ovat keskeisessä roolissa, sai kansainvälistä tunnustusta. Se palkittiin muun muassa AI Finlandin "Vuoden suomalainen luova AI -toteutus" -palkinnolla sekä Nordic VR Forumin "Paras XR-kokemus 2024" -palkinnolla. Nämä tunnustukset osoittavat järjestelmän innovatiivisuuden ja vaikuttavuuden alallaan.

Haasteet ja käyttäjäkokemukseen vaikuttavat tekijät:

Vaikka palaute oli pääosin positiivista, testaajat toivat esiin myös kehityskohteita:

- **Animaatioiden eheys ja luonnollisuus:** Animaatioiden laatu ja luonnollisuus ovat kriittisiä käyttäjäkokemuksen kannalta. Kehitysvaiheessa havaitut nykivät tai epäluonnolliset liikkeet heikensivät immersiota. Pienetkin poikkeamat luonnollisessa liikkeessä voivat kiinnittää käyttäjän huomion pois itse vuorovaikutuksesta ([Birdwhistell, 1970](#)).-> [tarkista](#)
- **Silmien liikkeet:** Silmien liikkeiden luonnollisuus osoittautui erityisen merkittäväksi uskottavuuden kannalta. Testaajat kokivat epäluonnolliset tai liikkumattomat silmät epämiellyttävinä. Koska nVidian a2f-menetelmä ei tutkimusajankohtana tukenut silmien reaaliaikaista animointia, kehitimme oman ratkaisun, joka lisäsi silmien liikkeisiin luonnollista vaihtelua ja satunnaisuutta. Tämä paransi merkittävästi avatarien inhimillisyyttä.
- **Reaktionopeus:** Viiveet avatarin vastauksissa vaikuttivat kielteisesti käyttäjäkokemukseen. Mikäli avatarin vastaus kesti liian kauan, käyttäjät kokivat tilanteen epäaidoksi. Tämä korostaa teknisen suorituskyvyn ja järjestelmän nopeuden merkitystä immerssiivisessä vuorovaikutuksessa.

Kehon animaatioiden haasteet:

Yksi suurimmista haasteista kehitystyössä oli luonnollisten kehoanimaatioiden luominen. Avatarin liikkeiden tuli olla ihmismäisiä ja sopia tilanteeseen, mutta käsin tehtävien animaatioiden tuottaminen osoittautui työlääksi. Tutkimusajankohtana nVidian a2f-teknologia mahdollisti reaaliaikaisesti vain kasvojen liikkeiden animoinnin. Kehon liikkeiden puuttuminen rajoitti avatarin nonverbaalista viestintää, mikä saattoi vaikuttaa vuorovaikutuksen syvyyteen ja autenttisuuteen.

Tulevaisuuden kehityssuunnat

Havaitut haasteet antavat suuntaa tulevalle tutkimukselle ja kehitystyölle:

Reaaliaikainen kehonliikkeiden animointi: Aiomme jatkossa hyödyntää kehittyneempiä teknologioita, kuten kehonliikkeiden reaaliaikaista animointia AI-pohjaisilla menetelmillä. Tämä mahdollistaa avatarien entistä luonnollisemman liikkumisen ja vuorovaikutuksen.

Tarkemmat käyttäjätestaukset: Suunnittelemme tulevaisuudessa tarkemmin määriteltyjä käyttäjätestauksia, joissa osallistujat luokitellaan ja testausolosuhteet vakioidaan. Näin voimme kerätä systemaattisempaa dataa käyttäjäkokemuksesta ja kehittää järjestelmää vastaamaan paremmin käyttäjien tarpeita.

Testaajien palautteen perusteella on selvää, että teknologian kehittyessä avatarien uskottavuus ja luonnollisuus tulevat entisestään paranemaan. Tämä tukee tutkimuksemme löydöksiä ja korostaa immersiiivisten 3D-avatarien potentiaalia digitaalisen vuorovaikutuksen kentällä.

Vertailu aiempaan tutkimukseen:

Löydökset ovat linjassa Mehrabianin (1972) tutkimusten kanssa, jotka korostavat nonverbaalisen viestinnän merkitystä tunneviestinnässä. Ne tukevat myös esittävien taiteiden nonverbaalisen viestinnän periaatteiden soveltamisen potentiaalia avatar-teknologiassa (Schechner, 1985).

5.2 Käyttäjien luottamus ja turvallisuuden tunne

Yhteenveto löydöksistä:

Immersiiviset 3D-järjestelmät vaikuttavat käyttäjien luottamukseen ja turvallisuuden tunteeseen sekä positiivisesti että negatiivisesti:

- **Luottamuksen lisääntyminen:** Ai pohjaisen toiminnanohjausmenetelmän tarjoamat laadukkaat tiedot ja avatarien luonnollinen tapa keskustella ja inhimilliset piirteet lisäävät käyttäjien luottamusta järjestelmään. -> **mieti vielä tämän muotoilua**
- **Turvallisuuden tunteen heikkeneminen:** Samanaikaisesti jotkut käyttäjät ilmaisevat huolta tietosuojasta ja siitä, miten heidän dataansa kerätään ja käytetään.

Löydökset heijastavat Turkle (2011) esittämiä teknologiakriittisiä näkemyksiä, joissa korostetaan teknologian vaikutusta yksilön yksityisyyteen ja sosiaaliseen vuorovaikutukseen. Ne tuovat esiin tarpeen läpinäkyvyydelle ja eettisille käytännöille järjestelmien suunnittelussa.

5.3 Eettiset kysymykset ja haasteet

Yhteenveto löydöksistä

Tutkimus paljasti useita eettisiä haasteita:

- **Manipulaation mahdollisuus:** Avatarien kyky vaikuttaa käyttäjän tunteisiin herättää huolta mahdollisesta manipuloinnista, erityisesti kaupallisissa tai poliittisissa tarkoituksissa.
- **Läpinäkyvyyden puute:** Käyttäjät eivät aina ole tietoisia siitä, miten järjestelmät toimivat ja mitä tietoja ne keräävät, mikä voi heikentää luottamusta.
- **Kulttuuriset erot:** Avatarien on haastavaa mukautua kaikkiin kulttuuriin nonverbaalisen viestinnän normeihin, mikä voi johtaa väärinymmärryksiin ja heikentää vuorovaikutuksen tehokkuutta.

Nämä havainnot korostavat eettisten kysymysten merkitystä, joista Russell (2019) ja Haraway (1985) ovat keskustelleet. Ne osoittavat tarpeen kehittää eettisiä ohjeistuksia ja läpinäkyviä toimintatapoja immersiiivisten teknologioiden kehittämisessä.

5.4 Yhteenveto keskeisistä löydöksistä

Kooste

Digitaalisen vuorovaikutuksen normien muutos: Immersiiviset 3D-avatatit muuttavat digitaalisen vuorovaikutuksen normeja tekemällä siitä inhimillisemmän ja monipuolisemman nonverbaalisen viestinnän kautta.

Luottamus ja turvallisuus: Vaikka käyttäjät kokevat lisääntyntä luottamusta järjestelmien inhimillisten piirteiden vuoksi, he ovat myös huolissaan yksityisyydestään ja tietoturvaan.

Eettiset haasteet: Merkittävät eettiset kysymykset liittyen manipulaatioon, läpinäkyvyyteen ja kulttuuriseen sensitiivisyyteen on otettava huomioon järjestelmien suunnittelussa ja implementoinnissa.

Kytkeä tutkimuskysymyksiin

Ensimmäinen tutkimuskysymys: Löydökset osoittavat, että immersiiiviset 3D-avatatit muovaavat digitaalisen vuorovaikutuksen normeja lisäämällä vuorovaikutuksen inhimillistä tuntua ja hämärtämällä ihmisen ja teknologian välisiä rajoja.

Toinen tutkimuskysymys: Käyttäjien luottamukseen ja turvallisuuden tunteeseen liittyvät havainnot vastaavat toiseen tutkimuskysymykseen, tuoden esiin sekä positiivisia että negatiivisia vaikutuksia älykkäiden vuorovaikutusjärjestelmien kehittyessä.

Kolmas tutkimuskysymys: Eettisten haasteiden tunnistaminen korostaa tarvetta tasapainottaa teknologian tarjoamat mahdollisuudet eettisten ja käytännöllisten haasteiden kanssa, mikä liittyy kolmanteen tutkimuskysymykseen.

Johtopäätös

Tulokset korostavat immersiiivisten 3D-avatarien potentiaalia parantaa digitaalista vuorovaikutusta, mutta samalla ne tuovat esiin tärkeitä eettisiä ja käytännöllisiä haasteita. Nämä havainnot muodostavat perustan seuraavalle luvulle, jossa pohdimme tulosten merkitystä suhteessa tutkimuksen tavoitteisiin, teoreettiseen viitekehykseen ja laajempaan yhteiskunnalliseen kontekstiin.

6. Tulosten tulkinta suhteessa tutkimuskysymyksiin ja viitekehukseen

Immersiivisten 3D-avatarien vaikutus digitaalisiin vuorovaikutusnormeihin (Tutkimuskysymys 1)

Tutkimuksen tulokset osoittavat, että immersiiivisillä 3D-avatareilla on merkittävä vaikutus digitaalisiin vuorovaikutusnormeihin. Kehitetyn järjestelmän avulla havainnollistimme, miten nonverbaalinen viestintä, tunneanalyysi ja muistitoiminnot voivat parantaa teknologiavälitteistä viestintää. Esimerkiksi realistiset kasvoanimaatiot ja huulisynkronointi, jotka toteutettiin nVidian a2f-tekniologialla, lisäsivät vuorovaikutuksen autenttisuutta ja käyttäjien sitoutumista.

On kuitenkin huomionarvoista, että tutkimusajankohtana nVidian a2f-tekniologia mahdollisti reaaliaikaisesti ainoastaan kasvojen liikkeiden animoinnin. Vaikka niskan ja hartioiden liikkeitä sekä asentoja voitiin animoida offline-tilassa puheeseen liittyvän audion pohjalta, näitä ei vielä voitu toteuttaa reaaliaikaisesti dialogissa. Tämä rajoitus vaikuttaa siihen, kuinka kokonaisvaltaisesti avatar kykenee ilmentämään nonverbaalista viestintää.

Tulevaisuudessa odotetaan teknologian kehittyvän siten, että reaaliaikainen animointi laajenee koko kehon liikkeisiin. Tämä mahdollistaisi entistä rikkaamman nonverbaalisen viestinnän ja syvemmän vuorovaikutuksen avatarin ja käyttäjän välillä, hämärtäen entisestään ihmisen ja koneen välisiä rajoja sekä luoden uusia vuorovaikutusnormeja. Lisäksi tutkimuksessa mainitut peilisolujen toimintaan perustuvat tekniikat tulevat lisäämään avatarin kehoallista yhteyttä ihmiseen, jonka kanssa avatar keskustelee. AI-pohjaiset järjestelmät osaavat jatkossa yhä paremmin tulkita vuorovaikutustilannetta ja luoda siihen sopivia kehoallilikkeitä.

Älykkäiden vuorovaikutusjärjestelmien vaikutus eettisiin ja moraalisiin periaatteisiin (Tutkimuskysymys 2)

Järjestelmän tekniset ratkaisut, kuten suurten kielimallien hyödyntäminen ja reaaliaikainen data-analyysi, osoittivat, miten tekoälyohjatut dialogijärjestelmät voivat tukea vuorovaikutusta huomioiden nonverbaalisen viestinnän ja tunnetietoisuuden. Vaikka

tunneanalyysin ja moniagenttijärjestelmien syvempi integrointi jäivät tulevien tutkimusten aiheiksi, alustavat tulokset ovat rohkaisevia.

Käyttäjien luottamus järjestelmään voi vahvistua, jos läpinäkyvyys ja yksityisyys varmistetaan. Toisaalta järjestelmän kyky tulkita ja reagoida käyttäjän nonverbaaliseen viestintään herättää eettisiä kysymyksiä liittyen yksityisyyteen ja mahdolliseen manipulointiin. Erityisesti tulevaisuudessa, kun avatarit pystyvät reaaliaikaisesti animoimaan koko kehon liikkeitä ja tulkitsemaan käyttäjän eleitä entistä tarkemmin, on tärkeää varmistaa, että käyttäjien oikeuksia kunnioitetaan ja että heillä on kontrolli omiin tietoihinsa.

Immersiivisten teknologioiden mahdollisuuksien ja eettisten haasteiden tasapainottaminen (Tutkimuskysymys 3)

Tuloksemme tukevat posthumanistista viitekehystä, jossa teknologia sulautuu osaksi inhimillistä vuorovaikutusta. Kehitetty järjestelmä haastaa perinteiset käsitykset teknologiasta pelkkänä työkaluna ja herättää uusia eettisiä ja filosofisia kysymyksiä. Tämä korostaa tarvetta laajemmalle yhteiskunnalliselle keskustelulle teknologiavälitteisen vuorovaikutuksen vaikutuksista identiteettiin ja sosiaalisiin rakenteisiin.

Teknologian kehittyessä AI-pohjaiset järjestelmät tulevat yhä paremmin tulkitsemaan vuorovaikutustilanteita ja luomaan niihin sopivia kehonliikkeitä. Tämä lisää vuorovaikutuksen autenttisuutta, mutta tuo mukanaan uusia eettisiä ja käytännöllisiä haasteita, kuten yksityisyydensuojan varmistamisen ja manipulaation riskien hallinnan. Eettisten ohjeistusten kehittäminen, läpinäkyvyyden lisääminen ja ihmiskeskeinen suunnittelu ovat keskeisiä keinoja tasapainottaa teknologian mahdollisuudet sen haasteiden kanssa.

Lyhyt yhteenvedollinen vastaus tutkimuskysymyksiin:

Tutkimuskysymys	Tiivistetty vastaus
1. Miten immersiiiviset 3D-avatarit muovaavat digitaalisia vuorovaikutusnormeja ja haastavat perinteistä teknologiavälitteistä vuorovaikutusta?	Ne lisäävät vuorovaikutuksen inhimillistä tuntua (nonverbaalinen viestintä, huulisynkronointi) ja hämärtävät ihmisen ja koneen välisiä rajoja, luoden uusia normistoja. Teknologian kehittyessä vaikutus syvenee entisestään, kun koko kehon liikkeiden reaaliaikainen animointi

	mahdollistuu.
2. Miten älykkäiden vuorovaikutusjärjestelmien kehittyminen vaikuttaa eettisiin ja moraalisiin periaatteisiin? Miten luottamus ja turvallisuus muuttuvat?	Järjestelmät voivat vahvistaa luottamusta, jos läpinäkyvyys ja yksityisyys varmistetaan. Samalla on riski manipulaatiosta ja uudenlaisten valtasuhteiden syntymisestä, kun tekoäly pystyy tulkitsemaan ja reagoimaan entistä syvällisemmin käyttäjän tunteisiin ja nonverbaaliseen viestintään.
3. Miten tasapainottaa immersiiivisten teknologioiden mahdollisuudet niiden eettisten ja käytännöllisten haasteiden kanssa?	Eettiset ohjeistukset (esim. läpinäkyvyys, ihmiskeskeinen suunnittelu) ja teknologian kriittinen arviointi ovat avainasemassa. Tarvitaan laajempaa yhteiskunnallista keskustelua vastuullisesta kehittämisestä sekä käyttäjien oikeuksien ja yksityisyyden suojaamisesta teknologian kehittyessä.

Johtopäätös

Tulokset osoittavat, että immersiiiviset 3D-avatatit ja kehittyvät tekoälyteknologiat muuttavat merkittävästi digitaalisen vuorovaikutuksen kenttää. Ne tarjoavat mahdollisuuksia entistä inhimillisempään ja vaikuttavampaan vuorovaikutukseen, mutta samalla ne tuovat mukanaan eettisiä haasteita, jotka vaativat huomiota. Teknologian kehityksen myötä on olennaista jatkaa tutkimusta keskittyen sekä teknisiin ratkaisuihin että eettisiin periaatteisiin, jotta voidaan luoda vastuullisia ja käyttäjien tarpeita kunnioittavia vuorovaikutusjärjestelmiä.

6.2 Tutkimuksen rajoitukset

Vaikka tutkimus tarjoaa arvokasta tietoa immersiiivisistä 3D-dialogijärjestelmistä, sillä on myös rajoituksia:

1. **Käyttäjätutkimuksen puuttuminen:** Emme tässä työssä toteuttaneet laajamittaista käyttäjätutkimusta. Vaikutukset käyttäjien asenteisiin, tunteisiin ja käyttäytymiseen jäävät siten osin spekulatiivisiksi.
2. **Tunneanalyysin rajaus:** Järjestelmässä ei ole vielä täysimittaista tunnetilan tunnistusta, mikä rajoittaa syvempää ymmärrystä ihmisen ja avatarin välisestä emotionaalisesta vuorovaikutuksesta.

3. **Laajempi yhteiskunnallinen konteksti:** Vaikka tutkimuksessa viitataan eettisiin ja filosofisiin pohdintoihin, niiden syvälinen käsittely vaatisi erillisen (esim. sosiologisen) tutkimusotteen.

6.3 Johtopäätökset ja suositukset

Tutkimus osoittaa, että immersiiivisillä 3D-avataar-järjestelmillä on potentiaalia muuttaa digitaalista vuorovaikutusta merkittävästi. Teknologian kehittyessä on tärkeää ottaa huomioon sekä sen mahdollisuudet että eettiset haasteet. Suositukseni on:

1. **Ihmiskeskeinen suunnittelu:** Teknologian tulee tukea käyttäjän autonomiaa ja hyvinvointia ilman, että se hallitsee tai manipuloi hänen päätöksentekoaan.
2. **Empaattisuus ja tunnetietoisuus:** Empaattisia ominaisuuksia tulee hyödyntää realistisesti, lisäten luottamusta ja käyttökokemuksen laatua ilman harhaanjohtamista.
3. **Eettiset periaatteet ja läpinäkyvyys:** Järjestelmien kehittämisessä tulee noudattaa selkeitä eettisiä periaatteita, kuten läpinäkyvyyttä, yksityisyyden suojaa ja käyttäjien oikeuksien kunnioittamista.
4. **Monimuotoisuus ja kulttuurinen ymmärrys:** Teknologian tulee edistää monimuotoisuutta ja kulttuurien välistä ymmärrystä, välttämällä stereotyyppioita ja ennakkoluuloja.
5. **Laajempi yhteiskunnallinen keskustelu:** On tarpeen käydä laajempaa keskustelua siitä, miten tällaiset teknologiat vaikuttavat yhteiskuntaan, jotta voidaan varmistaa niiden eettinen ja vastuullinen käyttö.

6.4 Tulevat tutkimussuunnat

Jatkossa suosittelen käyttäjätutkimusten toteuttamista järjestelmän todellisen vaikutuksen arvioimiseksi. Tämä sisältää sekä kvantitatiivisia että kvalitatiivisia menetelmiä käyttäjien kokemusten, tunteiden ja käyttäytymisen ymmärtämiseksi. Lisäksi tunneanalyysin ja moniagenttijärjestelmien syvämpi integrointi voi avata uusia mahdollisuuksia entistä autenttisempiin ja monipuolisempiin vuorovaikutustilanteisiin. On myös tärkeää syventää

eettistä ja filosofista keskustelua, huomioiden teknologian vaikutukset identiteettiin, sosiaalisiin rakenteisiin ja yhteiskunnallisiin normeihin.

Yhteistyön merkitys

Teknologian kehittämiseen tarvitaan monitieteistä yhteistyötä, joka yhdistää:

1. **Psykologia ja neurotiede:** Auttaakseen ymmärtämään empatian ja tunnetilojen neurobiologisia perustoja. Tämä voi ohjata avatarien nonverbaalisen viestinnän kehittämistä.
2. **Taiteet:** Performanssitaiteen ja tanssin asiantuntijat voivat tuoda uusia näkökulmia avatarien keholliseen ilmaisuun.
3. **Teknologia:** Tekoälykehittäjien ja suunnittelijoiden rooli on keskeinen algoritmien optimoinnissa ja eettisten periaatteiden sisällyttämisessä.

Tulevaisuuden tutkimusalueet

- **Käyttäjätutkimus:** Toteuta laajamittaisia tutkimuksia avatarien vaikutuksesta käyttäjien tunteisiin, vuorovaikutukseen ja oppimiseen.
- **Eettisten ohjeiden kehittäminen:** Luo standard

Lähdeluettelo

1. Bassett, C. (2007). *The Arc and the Machine: Narrative and New Media*. Manchester University Press.
2. Bassett, C. (2021). *Anti-computing: Dissent and the Machine*. Manchester University Press.
3. Bijker, W. E., Hughes, T. P., & Pinch, T. J. (1987). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. MIT Press.
4. Birdwhistell, R. L. (1970). *Kinesics and context: Essays on body motion communication*. University of Pennsylvania Press.
5. Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
6. Brecht, B. (1964). *Brecht on Theatre: The Development of an Aesthetic*. Hill and Wang.
7. Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1).
8. Bryson, J. J. (2010). Robots should be slaves. In Y. Wilks (Ed.), *Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues* (pp. 63–74). John Benjamins.
9. Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6), 893–910.
10. Clark, A. (2013). *Whatever next? Predictive brains, situated agents, and the future of cognitive science*. *Behavioral and Brain Sciences*, 36(3), 181–204.
<https://doi.org/10.1017/S0140525X12000477>
11. Dick, P. K. (1968). *Do Androids Dream of Electric Sheep?*. Doubleday.
12. Dixon, S. (2007). *Digital performance: A history of new media in theater, dance, performance art, and installation*. MIT Press.
13. Dreyfus, H. L. (1992). *What Computers Still Can't Do: A Critique of Artificial Reason*. MIT Press.
14. Ekman, P. (1978). Facial expression. In A. W. Siegman & S. Feldstein (Eds.), *Nonverbal behavior and communication* (pp. 97–116). Hillsdale, NJ: Erlbaum.

15. Ekman, P. (1978). *Facial Action Coding System (FACS)*. Consulting Psychologists Press.
16. Fischer, C. S. (1992). *America Calling: A Social History of the Telephone to 1940*. University of California Press.
17. Floridi, L., et al. (2018). AI4People—An ethical framework for a good AI society. *Mind & Machines*, 28(4), 689–707.
18. Floridi, L. (2019). *The Logic of Information: A Theory of Philosophy as Conceptual Design*. Oxford University Press.
19. Freedberg, D., & Gallese, V. (2007). Movement, emotion, and empathy in aesthetic experience. *Trends in Cognitive Sciences*, 11(5), 197–203.
<https://doi.org/10.1016/j.tics.2007.02.003>
20. Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. <https://doi.org/10.1038/nrn2787>
21. Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3475–3484.
22. Goffman, E. (1959). *The presentation of self in everyday life*. Doubleday Anchor Books.
23. Giannachi, G. (2004). *Virtual theatres: An introduction*. Routledge.
24. Graham, M. (1991). *Blood Memory: An Autobiography*. New York, NY: Doubleday.
25. Haraway, D. (1985). A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. *Socialist Review*, 80, 65–108.
26. Kurzweil, R. (2005). *The Singularity is Near: When Humans Transcend Biology*. Viking.
27. Laban, R. (1971). *The Mastery of Movement*. Macdonald and Evans.
28. Lanier, J. (2010). *You Are Not a Gadget: A Manifesto*. Alfred A. Knopf.
29. Lanier, J. (2018). *Ten Arguments for Deleting Your Social Media Accounts Right Now*. Henry Holt and Co.
30. Lem, S. (1961). *Solaris*. Wydawnictwo Ministerstwa Obrony Narodowej.
31. Mayer, R. E. (2005). *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press.
32. McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. McGraw-Hill.
33. Mehrabian, A. (1972). *Nonverbal Communication*. Aldine-Atherton.

34. Mori, M. (1970). The uncanny valley. *Energy*, 7(4), 33–35.
35. Norman, D. A. (2013). *The Design of Everyday Things* (Revised and Expanded Edition). Basic Books.
36. OpenAI. (n.d.). ChatGPT can now see, hear, and speak. Retrieved from <https://openai.com/index/chatgpt-can-now-see-hear-and-speak/>
37. Postman, N. (1992). *Technopoly: The Surrender of Culture to Technology*. Vintage Books.
38. Rizzolatti, Giacomo & Craighero, Laila
39. Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192. <https://doi.org/10.1146/annurev.neuro.27.070203.144230>
40. Russell, S. (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking.
41. Samek, W., Wiegand, T., & Müller, K. (2019). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. arXiv preprint arXiv:1708.08296.
42. Schechner, R. (1985). *Between theater and anthropology*. University of Pennsylvania Press.
43. Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, 69(5), 379–399.
44. Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424.
45. Seth, A. K. (2016). *The hard problem of consciousness is a distraction from the real one*. Aeon.
<https://aeon.co/essays/the-hard-problem-of-consciousness-is-a-distraction-from-the-real-one>
46. Standage, T. (1998). *The Victorian Internet: The Remarkable Story of the Telegraph and the Nineteenth Century's On-Line Pioneers*. Walker and Company.
47. Stanislavski, K. (1936). *An Actor Prepares*. Theatre Arts Books.
48. Turkle, S. (2011). *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books.
49. Voigt, P., & Von dem Bussche, A. (2017). *The EU General Data Protection Regulation (GDPR)*. Springer.

50. Vroomen, J., & Keitel, C. (2019). Multisensory integration and embodied interactions in performance. *Frontiers in Psychology*, *10*, 1921.
<https://doi.org/10.3389/fpsyg.2019.01921>