Iida Oikarainen ja Saana Sammalo

Pilaako tekoäly luotettavuuden?

Suomalaisten mediankäyttäjien rakentamat diskurssit tekoälyn luotettavuudesta

Viestinnän- ja markkinoinnin yksikkö
Viestintätieteet
Viestintätieteiden kandidaattiohjelma

Vaasa 2025

VAASAN YLIOPISTO
Viestinnän- ja markkinoinnin yksikkö
Tekijä: Iida Oikarainen ja Saana Sammalo
Tutkielman nimi: Pilaako tekoäly luotettavuuden?: Suomalaisten mediankäyttäjien rakentamat diskurssit tekoälyn luotettavuudesta
Tutkinto: Humanististen tieteiden kandidaatti
Oppiaine: Viestintätieteet
Työn ohjaaja: Maria Eronen-Valli
Valmistumisvuosi: 2025 **Sivumäärä:** 39

TIIVISTELMÄ:
Generatiiviseen tekoälyyn liittyvä luottamus on ajankohtainen keskustelunaihe nykyajan digitaalisessa mediaympäristössä. Luottamusta on perinteisesti tarkasteltu ihmisten välisenä ilmiönä, mutta viime vuosina tutkimus on laajentunut koskemaan myös generatiivisia tekoälyjärjestelmiä. Tämä tutkielma ei pyri arvioimaan tekoälyn objektiivista luotettavuutta, vaan tarkastelee, miten luotettavuuden käsitettä merkityksellistetään ja tuotetaan kielellisesti.

Tutkimuksemme tavoitteena on selvittää, millaisia diskursseja suomalaiset mediankäyttäjät rakentavat generatiivisen tekoälyn luotettavuudesta YouTuben verkkokommenteissa. Teoreettinen viitekehyksemme käsittelee luottamuksen psykologista ulottuvuutta, sen eri muotoja ja konteksteja sekä tekoälylukutaidon merkitystä. Lähestymme aihetta sosiaalisen konstruktionismin näkökulmasta, jonka mukaan todellisuus rakentuu kielellisessä vuorovaikutuksessa. Lisäksi tarkastelemme YouTubea yhteisöllisenä ja vuorovaikutteisena sosiaalisen median alustana.

Tutkimusaineistomme koostuu generatiivista tekoälyä käsittelevistä YouTube-kommenteista. Aineisto analysoitiin laadullisen diskurssianalyysin menetelmin. Analyysissa tunnistettiin viisi keskeistä diskurssia: affektiivinen, teknologinen, reflektiivinen (itsereflektiota sisältävä ilmaisutapa), inhimillinen sekä eettinen ja yhteiskunnallinen diskurssi. Diskurssit limittyvät toisiinsa, ja yksittäiset kommentit voivat sisältää useita merkitysrakenteita samanaikaisesti. Tulokset osoittavat, että diskurssit generatiivisesta tekoälystä korostavat enemmän epäluottamusta kuin luottamusta.

Tutkimuksemme korostaa tekoälylukutaidon roolia ja merkitystä nykyhetken sekä tulevaisuuden mediataitona. Generatiivisen tekoälyn yleistyminen ja jatkuva kehittyminen hämärtää ihmisen ja tekoälyn tuottaman sisällön rajapintaa. Tekoälylukutaito vahvistaa käyttäjien kriittistä arviointikykyä ja tarjoaa pohjan kestävämmälle luottamukselle tekoälyjärjestelmiä kohtaan.

AVAINSANAT: Generatiivinen tekoäly, luottamus, YouTube, verkkokeskustelu, diskurssi, diskurssianalyysi, tekoälylukutaito

Sisällys

1 Johdanto 5

2 Tutkimuksen rajaukset 6

2.1 Tutkimuksen tavoite 6

2.2 Tutkimusaineisto 7

2.3 Tutkimusmenetelmä 8

3 Luottamus ja generatiivinen tekoäly 11

3.1 Luottamus käsitteenä ja psykologisena ilmiönä 11

3.2 Luottamuksen muodot ja kontekstit 12

3.3 Tekoälylukutaidon merkitys 14

4 YouTube sosiaalisen vuorovaikutuksen kontekstina 16

4.1 Sosiaalinen konstruktionismi 16

4.2 YouTube sosiaalisen median keskustelualustana 18

4.3 Yhteisöllinen vuorovaikutus YouTubessa 19

5 Generatiivisesta tekoälystä rakentuvat diskurssit 21

5.1 Affektiivinen diskurssi 22

5.1.1 Torjuva diskurssi 23

5.1.2 Pelon ja huolen diskurssi 24

5.2 Teknologinen diskurssi 25

5.2.1 Käytännöllinen diskurssi 25

5.2.2 Teknologinen metadiskurssi 26

5.3 Reflektiivinen diskurssi 27

5.3.1 Kaksijakoinen diskurssi 27

5.3.2 Yhteisöllisen reagoinnin diskurssi 28

5.4 Inhimillinen diskurssi 29

5.4.1 Ihmisen ainutlaatuisuuden diskurssi 29

5.4.2 Ihmisen ja tekoälyn välinen diskurssi 30

5.5 Eettinen ja yhteiskunnallinen diskurssi 31

5.5.1 Vastuun ja moraalin diskurssi 31

5.5.2 Ekologisuuden diskurssi 33

6 Johtopäätökset 34

Lähteet 36

1 Johdanto

Tekoälyteknologiat ovat jo vuosikymmenten ajan olleet osa arkeamme, ja niiden käyttäminen sekä kehitys ovat viime vuosina kasvaneet merkittävästi (Lalot & Bertram, 2024, s. 1). Erityisesti generatiivinen tekoäly on synnyttänyt vilkasta keskustelua (Sihvonen, 2025, s. 72) ja herättänyt kysymyksen: voiko tekoälyn tuottamaan sisältöön todella luottaa, vai onko kyseessä vain illuusio ihmismäisestä luovuudesta? Generatiivinen tekoäly (*GenAI*) on edistynyt tekoälyjärjestelmä, jolla on kyky analysoida olemassa olevaa dataa sekä luoda uutta sisältöä – kuten kuvia, tekstiä, ääntä tai videoita – hyödyntämällä generatiivisia malleja, koneoppimisalgoritmeja ja neuroverkkoja (Sengar ja muut, 2024, s. 4; AlAli & Wardat, 2024, s. 786). Näiden mallien avulla generatiivinen tekoäly pystyy luomaan kontekstuaalisesti merkityksellistä sisältöä, joka jäljittelee ihmisen luovuutta (Sengar ja muut, 2024, s. 4; AlAli & Wardat, 2024, s. 786).

Aiemmat tutkimukset ovat korostaneet tarvetta tarkastella tekoälyyn liittyviä merkityksiä empiirisesti ja huomioida sekä mikro- että makrotason yhteisvaikutuksia viestinnän käytännöissä (Laapotti ja muut, 2024, s. 151). Lisäksi käyttäjien näkökulmien syvällinen ymmärtäminen on keskeisessä roolissa teknologioiden, kuten generatiivisen tekoälyn, kohdalla (Young ja muut, 2021). Tässä tutkimuksessa tuomme lisäymmärrystä käyttäjien suhtautumisesta generatiiviseen tekoälyyn sekä siihen liittyvästä luottamuksesta. Näiden tulkintojen tutkiminen on tärkeää, sillä sosiaalisen median keskusteluissa rakentuvat käsitykset vaikuttavat suoraan siihen, miten generatiivisella tekoälyllä tuotettua sisältöä ymmärretään, arvotetaan ja joko hyväksytään tai torjutaan osana digitaalista mediaympäristöä. Tämän tutkimusaukon täyttämiseksi tutkimuksemme tarttuu empiirisellä otteella ajankohtaiseen aiheeseen, ja syventää viestinnän sekä mediatutkimuksen ymmärrystä siitä, miten mediankäyttäjät arvioivat ja tulkitsevat generatiivisen tekoälyn luotettavuutta sekä siihen liittyviä merkityksiä. Samalla tutkimus avaa näkökulman siihen, miten yksittäisten mediankäyttäjien käsitykset heijastavat laajemmin luottamusta ja sisällön luotettavuutta digitaalisessa mediaympäristössä.

2 Tutkimuksen rajaukset

Seuraavassa kappaleessa käsittelemme tutkimukselle keskeisimmät valinnat, tavoitteet, aineiston sekä tutkimusmenetelmän. Lisäksi varmistamme, että teemme analyysiä koskevat valinnat objektiivisesti, perusteellisesti sekä tutkimuseettisten periaatteiden mukaisesti säilyttääksemme tutkimuksen puolueettomuuden.

2.1 Tutkimuksen tavoite

Tutkimuksen tavoitteena on selvittää, millaisia käsityksiä suomalaiset mediankäyttäjät rakentavat generatiivisen tekoälyn luotettavuudesta YouTuben verkkokommenteissa. Tutkimus keskittyy analysoimaan YouTube-kommenteissa esiintyviä käsityksiä, ja erityisenä kiinnostuksen kohteena on se, millä tavoin generatiiviseen tekoälyyn sisällöntuotantovälineenä liitetään merkityksiä sen luotettavuudesta.

Tutkimus tarkastelee, miten mediankäyttäjien kommentit jäsentävät keskustelua generatiivisen tekoälyn luotettavuudesta sekä siitä, miten nämä näkemykset muovaavat käsitystä tekoälyn asemasta digitaalisessa mediassa ja arjessa. Tutkimus tarjoaa ymmärrystä myös siitä, kuinka sosiaalinen todellisuus rakentuu kommenteissa esitettyjen merkitysten, tulkintojen sekä vuorovaikutuksen kautta. Tavoitteena ei ole siten arvioida tekoälyn todellista luotettavuutta eikä pyrkiä arvioimaan sen objektiivista tai annettua olemusta. Huomioimme teoreettisten viitekehysten väliset erot tarkastelemalla verkkokeskusteluja ilmiönä, jossa yksilölliset ja subjektiiviset tulkinnat ovat osa yhteisesti rakentuvaa merkityksenantoa.

Vaikka tutkimuksemme keskittyy YouTubeen, se tarjoaa pohjan ymmärrykselle siitä, kuinka luotettavana mediankäyttäjät pitävät generatiivista tekoälyä. Samalla tutkimus avaa mahdollisuuden laajentaa analyysiä myös muihin sosiaalisen median alustoihin, mikä voi syventää ymmärrystä mediankäyttäjien näkemyksistä. Lisäksi tutkimuksemme tuottaa arvokasta tietoa ja ymmärrystä viestinnän alalle, erityisesti teknologisen

viestinnän ja verkkokeskustelujen tutkimuskenttään luoden perustan jatkotutkimuksille mediankäyttäjien käsityksistä generatiivisesta tekoälystä ja sen luotettavuudesta.

2.2 Tutkimusaineisto

Tutkimuksen aineistona käytetään ZoneVD:n videon *"Tekoäly pilaa YouTuben"* (26.8.2025) kommentteja. Videon kommentit kerättiin 7.10.2025 YouTuben Data Tools -työkalun avulla, joka muodosti niistä automaattisesti Excel-tiedoston. Tähän mennessä videolla oli yli 31 000 näyttökertaa ja 491 kommenttia. Järjestimme kommentit Excelissä julkaisuajan perusteella vanhimmasta uusimpaan, koska se mahdollisti selkeästi rajatun aineiston. Lopullisena aineistona käytämme videoon tulleita vanhimpia kommentteja sekä niihin tulleita vastauksia. Aineiston kooksi rajasimme 50 kommenttia, joista kaikki liittyvät generatiiviseen tekoälyyn. Aineiston ulkopuolelle jätettiin kommentit, jotka eivät liittyneet suoraan tutkimuskysymykseemme. Tämä rajattu sekä valikoitu otanta mahdollistaa sen, että aineisto säilyy tutkimuksen kannalta monipuolisena ja relevanttina, mutta tarjoaa syvällisen tarkastelun eri käsityksien rakentumisesta.

Tutkimusaineiston videon valintaan vaikuttivat sen ajankohtaisuus, suomalaisten mediankäyttäjien laaja yleisöosallistuminen sekä aiheen suora yhteys tutkimuskysymykseemme. Lisäksi video on vapaasti saatavilla, sillä se ei ole maksumuurin takana. Tämä tekee tutkimusaineistosta läpinäkyvän ja tutkimuksesta helposti toistettavan. YouTube-videon kommenttiosio tarjoaa julkisen alustan, jossa yleisö voi reagoida videossa esitettyihin näkemyksiin sekä jakaa omia ajatuksiaan ja kokemuksiaan generatiivisen tekoälyn käytöstä sekä sen vaikutuksista.

ZoneVD:n videolla (26.8.2025) tarkastellaan kriittisesti tekoälyn kasvavaa roolia YouTubessa ja sen vaikutuksia sisällöntuotantoon, tekijöiden luovuuteen ja vuorovaikutukseen katsojien kanssa. Hän mainitsee, että generatiivinen tekoäly voi nykyään automatisoida esimerkiksi videon tekijöiden kommenttivastaukset sekä videoiden pikkukuvien tuottamisen, mikä nopeuttaa tuotantoprosesseja. Samalla ne

kuitenkin kaventavat videon tekijän henkilökohtaista panosta ja vähentävät tekemisen iloa. Videossa käsitellään myös sitä, kuinka ihmiset yhä enemmän ulkoistavat ajatteluaan generatiiviselle tekoälylle, kuten ChatGPT:lle, mikä voi johtaa oman ajattelun "laiskistumiseen". Näiden lisäksi hän nostaa esiin, että generatiivisen tekoälyn tuottaman sisällön erottaminen ihmisten tekemästä sisällöstä voi tulevaisuudessa muuttua yhä vaikeammaksi. Videon lopussa ZoneVD rohkaisee katsojia jakamaan kommenttikentässä omia kokemuksiaan ja huoliaan generatiivisen tekoälyn vaikutuksista. Video toimii keskustelun avauksena sille, millaisia käsityksiä suomalaiset mediankäyttäjät rakentavat generatiivisen tekoälyn luotettavuudesta.

2.3 Tutkimusmenetelmä

Tutkielmamme on kvalitatiivinen empiirinen tutkimus, ja analyysimenetelmänä käytämme diskurssianalyysiä. Diskurssianalyysilla tarkoitetaan kielenkäytön tutkimista sosiaalisena toimintana, jossa kiinnostuksen kohteena on, miten merkityksiä rakennetaan, millaisia asioita nostetaan esiin ja mitä jätetään sanomatta (Pietikäinen & Mäntynen, 2019, s. 13–22). Menetelmä soveltuu tutkimukseemme, koska generatiiviseen tekoälyyn liittyvä luottamus on viestinnällinen ja kiistanalainen aihe, jota mediankäyttäjät käsittelevät erilaisin ilmaisutavoin YouTuben kommenttiosioissa. Diskurssianalyysin avulla voimme tarkastella, millaisin kielellisin keinoin käyttäjät tuottavat generatiiviseen tekoälyyn liittyviä merkityksiä, ja miten nämä merkitykset rakentavat todellisuutta generatiivisen tekoälyn roolista mediassa.

Lähdemme siitä, että kielenkäyttö, sosiaalinen toiminta ja konteksti muodostavat jatkumon, eivät erillisiä tasoja (Pietikäinen & Mäntynen, 2019, s. 15). Tämä mahdollistaa sen, että tarkastelemme kommentteja sekä yksittäisinä puheenvuoroina että osana laajempaa yhteiskunnallista keskustelua generatiivisen tekoälyn luotettavuudesta. Diskurssianalyyttinen lähestymistapa on luonteeltaan konstruktionistinen, eikä kielenkäyttö ainoastaan kuvaa todellisuutta, vaan osallistuu sen rakentamiseen (Jokinen ja muut, 2016, s. 17). Näin ollen analyysi ei pyri etsimään yhtä objektiivista totuutta, vaan

eri tavoin rakentuneita käsityksiä ja merkitysjärjestelmiä. Diskurssintutkimuksen prosessi on luonteeltaan monivaiheinen ja mutkitteleva, sillä analyysi ei etene suoraviivaisesti rajattuina jaksoina vaan sisältää vaiheiden vuorottelua, takaisin palaamista ja jatkuvaa tulkintaa (Pietikäinen & Mäntynen, 2019, s. 207–213). Tutkimuksen eri vaiheet myös limittyvät toisiinsa prosessin aikana (Pietikäinen & Mäntynen, 2019, s. 214; Günther & Hasanen, 2021), mikä tekee laadullisesta diskurssianalyysista luonteeltaan joustavan ja tulkinnallisen tutkimustavan.

Analyysissämme sovellamme Pietikäisen ja Mäntysen (2019, s. 253–311) esittämiä diskurssianalyysin vaiheita. Aluksi muodostamme teoreettisen viitekehyksen, ja perehdymme aineistoon kokonaisuutena, minkä pohjalta hahmotamme alustavat teemat ja sävyt YouTube-kommenteista. Seuraavaksi tarkastelemme aineistoa systemaattisesti laadullisen analyysin menetelmin merkiten toistuvia ilmaisuja ja erilaisia diskursseja. Analyysin myöhemmissä vaiheissa syvennämme havaintoja vertailemalla eri diskurssien välisiä suhteita ja vastakkainasetteluja. Kiinnostavaa on se, esitetäänkö generatiivinen tekoäly uhkana vai mahdollisuutena. Diskurssien luokittelu, nimeäminen ja kuvaaminen perustuvat havaintojen toistuvuuteen sekä merkitysten rakentumiseen niiden ensisijaisen painotuksen perusteella. Jaottelun toteutamme itse aineiston pohjalta, eikä se perustu valmiisiin malleihin tai jaotteluihin. Lopuksi kokoamme löydetyt diskurssit analyysiosioon pääkategorioihin esimerkkilainauksin, jotka havainnollistavat, miten suomalaiset mediankäyttäjät rakentavat käsityksiä generatiivisen tekoälyn luotettavuudesta. Diskurssianalyysin avulla voimme tunnistaa analyysistämme keskeisimmät luottamusdiskurssit, ja sen miten luottamus ilmenee YouTube-kommenteissa.

Tämän pohjalta analyysimme on sekä havaintojen tekemistä että niiden merkitysten tulkintaa osana laajempaa yhteiskunnalista ja verkkokeskustelun kontekstia. Ymmärrämme diskurssin käsitteen kielenkäytön sosiaalisena toimintana, joka rakentaa ja ylläpitää tietynlaisia merkitysjärjestelmiä (Pietikäinen & Mäntynen, 2019, s. 36; Jokinen ja muut, 2016, s. 34). Pietikäisen ja Mäntysen mukaan (2019, s. 36) diskurssit

taas voidaan nähdä kiteytyneinä merkityksellistämisen käytäntöinä, jotka määrittävät, miten ja millä oikeutuksella ilmiöistä voidaan puhua. Kun taas Jokisen ja muiden (2016, s. 34) mukaan diskurssit ovat jokseenkin yhtenäisiä merkityssuhteiden rakenteita, jotka syntyvät sosiaalisissa käytännöissä ja samalla muovaavat sosiaalista todellisuutta.

3 Luottamus ja generatiivinen tekoäly

Luottamusta on tutkittu laajasti eri tieteenaloilla, kuten psykologiassa ja filosofiassa, mutta tutkimukset ovat keskittyneet usein työpaikkadynamiikkaan, sosiaaliseen luottamukseen tai läheisiin ihmissuhteisiin (Lalot & Bertram, 2024, s. 2; Glikson & Woolley, 2021). Luottamus ei ole kuitenkaan pelkästään ihmisten välinen ilmiö, vaan viimeisten vuosien aikana luottamus on noussut keskeiseksi tutkimusaiheeksi myös generatiivisen tekoälyn kohdalla (Huynh & Aichner, 2025; Wolf ja muut, 2024; Lalot & Bertram, 2025). Braunerin ja muiden mukaan (2023, s. 4) luottamuksella on merkitystä generatiivisen tekoälyn hyväksymisessä ja käytössä.

Tässä luvussa tarkastelemme luottamusta ensin perinteisenä psykologisena ilmiönä, jonka sovellamme generatiivisen tekoälyn kontekstiin. Tarkastelemme myös luottamuksen muotoja ja sitä, miten luottamus voi rakentua tai heikentyä dynaamisessa vuorovaikutuksessa. Lopuksi käsittelemme tekoälylukutaidon merkitystä, sillä mediankäyttäjien kyky arvioida ja ymmärtää tekoälyn toimintaa heijastuu siihen, miten luottamus ilmentyy julkisessa verkkokeskustelussa.

3.1 Luottamus käsitteenä ja psykologisena ilmiönä

Tieteellisissä tutkimuksissa ei ole toistaiseksi yhtä yleisesti hyväksyttyä tieteellistä määritelmää luottamukselle (Rousseau ja muut, 1998, s. 394). Tässä tutkimuksessa omaksumme Rousseau'n ja muiden, (1998, s. 395) yleisesti hyväksytyn ja laajasti siteeratun määritelmän: "Luottamus on psykologinen tila, joka koostuu aikomuksesta hyväksyä haavoittuvuus, joka perustuu myönteisiin odotuksiin toisen aikomuksista tai käyttäytymisestä." Sovellamme kyseistä määritelmää tutkimukseemme, koska "toinen" voi viitata myös ei-inhimillisiin toimijoihin – tässä tapauksessa generatiiviseen tekoälyyn (Lee & See, 2004, s. 51).

Luottamus psykologisena tilana kuvaa yksilön halukkuutta altistua haavoittuvuudelle keskinäisen riippuvuuden ja riskin tilanteissa, perustuen myönteisiin odotuksiin (Rousseau ja muut, 1998, s. 394–395). Luottamus on siten keskeinen tekijä ihmisten halukkuudessa toimia myös epävarmoissa tilanteissa (Hoff & Bashir, 2015), joka ilmenee myös käyttäjän ja generatiivisen tekoälyn kohdalla. Generatiivisen tekoälyn vastaukset perustuvat probablistiseen mallinnukseen eli ne tuottavat vastauksia todennäköisyyksien perusteella (Wolf ja muut, 2024), jolloin vastaukset voivat vaihdella samankin kysymyksen kohdalla, eikä niiden paikkansapitävyyttä voi täysin arvioida. Näin käyttäjä altistaa itsensä generatiivisen tekoälyn haavoittuvuudelle, mutta samanaikaisesti hän toimii myönteisten odotusten varassa luottaen siihen, että generatiivinen tekoäly toimii rehellisesti ja eettisesti oikein.

Luottamus ei ole staattinen tila, vaan dynaaminen prosessi, joka rakentuu, vakiintuu ja voi myös heiketä tai purkautua ajan myötä tai erilaisten kokemuksien kautta (Rousseau ja muut, 1998, s. 396–397). Aiemmat tutkimukset ovat osoittaneet, että luottamus generatiivisen tekoälyn ja käyttäjän välillä on myös dynaaminen prosessi (Wolf ja muut, 2024; Lee & See, 2004, Huynh & Aichner, 2025; Hoff & Bashir, 2015), jossa myönteiset käyttökokemukset vahvistavat luottamusta, kun taas virheet, epäluotettavat tulokset tai eettiset rikkomukset voivat murentaa luottamuksen nopeasti.

3.2 Luottamuksen muodot ja kontekstit

Luottamus ilmiönä ei ole yksiselitteinen, ja ymmärtääksemme, miten ihmiset luottavat generatiiviseen tekoälyyn, on tarkasteltava luottamuksen eri muotoja. Rousseau'n ja muiden (1998, s. 398–401) mukaan luottamus voidaan jakaa kolmeen päämuotoon, jotka ovat laskelmoiva luottamus (*calculus-based trust*), suhteellinen luottamus (*relational trust*) sekä institutionaalinen luottamus (*institution-based trust*). Aiemmat tutkimukset ovat kuitenkin osoittaneet, että luottamus tekoälyyn on rakenteeltaan erilainen kuin luottamus ihmisten välillä (Bochniarz ja muut, 2022, s. 3).

Laskelmoiva luottamus perustuu rationaaliseen arvioon toisen osapuolen hyödyistä ja riskeistä (Rousseau ja muut, 1998, s. 399). Generatiivisen tekoälyn kontekstissa tämä tarkoittaa käyttäjän tietoista arviointia generatiivisen tekoälyn tuottaman sisällön tarkkuudesta, hyödyllisyydestä ja uskottavuudesta. Laskelmoiva luottamus on luonteeltaan ehdollista, sillä käyttäjät yleensä huomioivat myös mahdolliset riskit (Rousseau ja muut, 1998, s. 399), kuten virheellisen sisällön. Laskelmoiva luottamus perustuu jatkuvaan arviointiin ja vahvistamiseen – eli periaatteeseen "*trust but verify*" (Rousseau ja muut, 1998, s. 399).

Suhteellinen luottamus rakentuu toistuvien vuorovaikutusten ja kokemusten kautta, kun osapuolten välinen luotettavuus ja johdonmukaisuus vahvistuvat ajan myötä (Rousseau ja muut, 1998, s. 399). Ihmisillä on luontainen taipumus nähdä toiset ihmiset aidosti ja vilpittömästi toimivina (Sihvonen, 2025, s. 61). Generatiivisen tekoälyn kohdalla tämä voi tarkoittaa sitä, että mitä useammin tekoäly tuottaa hyödyllistä, johdonmukaista ja käyttäjän tarpeisiin mukautuvaa tietoa, sitä vahvemmin käyttäjä alkaa luottaa sen kykyihin. Tällöin generatiivinen tekoäly voidaan nähdä jopa "vuorovaikutuskumppanina", eikä pelkästään työkaluna. Glikson ja Woolley (2020, s. 630) esittävät kuitenkin, että vaikka luottamus ihmisiin yleensä kasvaa toistuvan vuorovaikutuksen kautta, luottamus teknologiaan usein vähenee virheiden tai toimintahäiriöiden kohtaamisen perusteella.

Instituutioihin perustuva luottamus puolestaan rakentuu yhteiskunnallisten ja organisatoristen rakenteiden varaan, jotka tekevät luottamisesta mahdollisen ja turvallisen (Rousseau ja muut, 1998, s. 400). Instituutiot nähdään ennen kaikkea luottamuksen mahdollistajina, ei pelkästään kontrollijärjestelminä (Rousseau ja muut, 1998, s. 400). Generatiivisen tekoälyn yhteydessä tämä tarkoittaa luottamusta erityisesti teknologiayritysten, mediaympäristöjen ja instituutioiden vastuullisuuteen, läpinäkyvyyteen ja kykyyn hallita riskejä. Näkökulma on tutkimuksemme kannalta keskeinen, sillä se selittää, miten generatiivisen tekoälyn luotettavuutta voidaan arvioida yksilöllisten käsityksien lisäksi laajempien yhteiskunnallisten rakenteiden kautta.

Rousseau ja muut (1998, s. 398–401) korostavat kuitenkin artikkelissaan, että luottamus ei ole yhtenäinen – se vaihtelee tehtävän, kontekstin ja aiemman kokemuksen mukaan. Käyttäjä saattaa luottaa generatiiviseen tekoälyyn tiettyjen asioiden kohdalla, kuten arkisissa toiminnoissa, mutta ei esimerkiksi kriittistä ajattelua tai eettistä harkintaa vaativissa tehtävissä. Lisäksi useat luottamuksen muodot saattavat olla läsnä samanaikaisesti (Rousseau ja muut, 1998). Vaikka käyttäjä tarkistaa vastaukset rationaalisesti, hän voi silti kokea jatkuvaa, myönteistä vuorovaikutusta generatiiviseen tekoälyyn. Tämä monitasoinen ja dynaaminen näkemys luottamuksesta tarjoaa vahvan teoreettisen pohjan diskurssien tutkimiselle, sillä se mahdollistaa luottamuksen eri muotojen tarkastelun ja niiden kontekstuaalisen ulottuvuuksien analysoinnin.

3.3 Tekoälylukutaidon merkitys

Tutkimuksemme keskittyy tarkastelemaan luottamusta julkisen verkkokeskustelun lisäksi myös yhteiskunnallisessa kontekstissa. Euroopan komission korkean tason asiantuntijaryhmä HLEG (*High-Level Expert Group on Artificial Intelligence*) mukaan luotettavan tekoälyn tulee olla laillinen, eettinen sekä kestävä, sillä käyttäjien luottamus on edellytys sen hyväksymiselle ja vastuulliselle hyödyntämiselle (European Commission, 2019). Tämä koskee erityisesti generatiivista tekoälyä, joka ei ainoastaan automatisoi prosesseja, vaan tuottaa myös uutta sisältöä, ja siten vaikuttaa suoraan ihmisten informaation käsittelyyn ja tiedon tulkintaan. Lisäksi laaja tekoälyn käyttäminen johtaa sekä kulttuurin että yhteiskunnan tasoilla kasvavaan tekoälyriippuvuuteen, mikä tekee samalla informaatioympäristöstä entistä epäluotettavampaa (Sihvonen, 2025, s. 77). Tämä korostaa tekoälylukutaidon merkitystä nyt ja tulevaisuudessa.

Tekoälylukutaito (*AI literacy*) voidaan ymmärtää yksilön perusymmärryksenä siitä, miten tekoäly kerää, yhdistelee ja tulkitsee dataa sekä, miten se tuottaa ja välittää sisältöjä (Hermann, 2021, s. 1270). Ng:n ja muiden (2021, s. 504–507) mukaan tulevaisuudessa ihmisten on tunnettava ja ymmärrettävä tekoälyn periaatteet sekä kyettävä käyttämään tällaisia teknologioita tehokkaasti. Lisäksi ihmisten tulisi pystyä arvioimaan tekoälyn

tuottamia tuloksia ja oltava tietoisia mahdollisista eettisistä kysymyksistä (Ng ja muut, 2021, s. 504–507). Hermann (2021, s. 1270–1271) korostaa, että tällainen ymmärrys on edellytys eettisesti kestävän tekoälyn hyödyntämiselle, sillä ilman selkeyttä ja läpinäkyvyyttä yksilöiden on vaikea arvioida tekoälyn toiminnan oikeudenmukaisuutta, hyödyllisyyttä tai mahdollisia haittavaikutuksia. Kunz ja Wirtz (2023, s. 35) myös huomauttavat, että tekoälylukutaito on erittäin tärkeää, kun tekoäly integroituu yhä enemmän ihmisten käyttämiin teknologioihin, kuten sosiaaliseen mediaan.

Aiempien tutkimuksien lisäksi, tekoälylukutaidon merkitys on korostunut myös yhteiskunnallisella tasolla. Myös Euroopan unionin näkökulmasta tekoälylukutaito on keskeinen osa vastuullista tekoälyn käyttöönottoa ja kehittämistä. EU:n tekoälyasetuksessa (Euroopan parlamentti ja neuvosto, asetus (EU) 2024/1689, art. 3) tekoälylukutaito määritellään osaamiseksi, tietämykseksi ja ymmärrykseksi, joiden avulla tarjoajat, käyttöönottajat ja henkilöt, joihin tekoälyn vaikutukset kohdistuvat, voivat tietoon perustuen käyttää tekoälyjärjestelmiä sekä ymmärtää niiden mahdollisuuksia, riskejä ja mahdollisia haittoja. Tällä korostetaan, että tekoälylukutaito ei ole ainoastaan teknistä tietoa, vaan siihen sisältyy myös oikeudellinen ja eettinen ymmärrys, joka tukee käyttäjien kykyä toimia vastuullisesti tekoälyä hyödyntävissä ympäristöissä.

Tekoälylukutaidon kehittäminen ja sen merkityksen ymmärtäminen voi toimia keinona vahvistaa käyttäjien kriittistä suhtautumista tekoälyyn, sekä lisätä heidän kykyään arvioida sen tuottamaa sisältöä. Hermannin (2021, s. 1271) mukaan tekoälylukutaito auttaa yksilöitä tunnistamaan, miten tekoäly voi vaikuttaa heidän valintoihinsa ja päätöksiinsä, ja siten tukee heidän autonomiaansa, toimijuuttaan ja eettistä arviointikykyään. Tekoälylukutaito voi siten myös vaikuttaa luottamukseen tekoälyä kohtaan ja tukea sen vastuullista hyödyntämistä.

4 YouTube sosiaalisen vuorovaikutuksen kontekstina

Tässä luvussa käsittelemme sosiaalisen todellisuuden rakentumista sosiaalisen konstruktionismin kautta, ja käsittelemme sosiaaliselle medialle, erityisesti YouTube-alustalle ominaista keskustelukulttuuria ja vuorovaikutusta. Lisäksi tarkastelemme sitä, miten luottamus ilmenee kielellisesti verkkokeskusteluissa, sekä millaisia merkityksiä ja tulkintoja niiden ympärille muodostuu.

4.1 Sosiaalinen konstruktionismi

Bergerin ja Luckmannin vuonna 1967 esittämä klassinen näkemys sosiaalisesta konstruktionismista on yksi käytetyimmistä, ja sen tavoitteena on ymmärtää, miten näennäisesti objektiiviset ilmiöt, kuten organisaatiot, teknologiat ja yhteiskunnalliset rakenteet, rakentuvat ihmisten välisissä merkityksenannon prosesseissa (Eriksson & Kovalainen, 2016). Heidän mukaansa todellisuus on sosiaalisesti tuotettu ja jatkuvasti uusiutuva ilmiö, joka saa muotonsa yhteisöllisessä vuorovaikutuksessa.

Sosiaalinen konstruktionismi ei kuitenkaan muodosta yhtä yhtenäistä teoreettista koulukuntaa, vaan se muodostuu useista toisiinsa limittyvistä ja osin ristiriitaisista suuntauksista (Burr, 1995, s. 2; Niska ja muut, 2024, s. 9). Burr (1995, s. 2–3) esittää kuitenkin, että sosiaalisella konstruktionismilla on tyypillisiä perusoletuksia, joita ovat kriittinen suhtautuminen itsestään selvinä pidettyihin tietorakenteisiin, tiedon historiallinen ja kulttuurinen sidonnaisuus, tiedon sosiaalinen rakentuminen vuorovaikutuksessa sekä tiedon ja toiminnan välinen yhteys. Konstruktionistisesta näkökulmasta tieto ei heijasta objektiivista todellisuutta, vaan rakentuu ihmisten välisessä vuorovaikutuksessa, erityisesti kielen ja diskurssien kautta (Burr, 1995; Niska ja muut, 2024; Stevanovic ja muut, 2024, s. 92–93).

Jokisen ja muiden (2016, s. 251–253) mukaan diskurssianalyysin teoreettinen perusta on sosiaalisessa konstruktionismissa. Heidän mukaansa todellisuutta ei voida kohdata

sellaisenaan, vaan se on aina kulttuurisesti ja kielellisesti merkityksellistetty. Kieli ei ainoastaan kuvaa todellisuutta, vaan se on todellisuuden kanssa jatkuvassa vuorovaikutuksessa rakentaen toisiaan (Jokinen ja muut, 2016, s. 252; Niska ja muut, 2024, s. 9). Näin ollen merkitykset syntyvät, vakiintuvat ja muuttuvat sosiaalisessa vuorovaikutuksessa, jolloin myös tutkijat osallistuvat omalla kielenkäytöllään todellisuuden tuottamiseen (Jokinen ja muut, 2016, s. 251–253). Sosiaalinen konstruktionismi tarjoaa siten arvokkaita välineitä tutkimukselle, joka kohdistuu sosiaalisiin tiedontuottamisen, merkityksenannon ja vuorovaikutuksen prosesseihin (Niska ja muut, 2024, s. 9).

Tässä tutkimuksessa lähestymme sosiaalista konstruktionismia näkökulmasta, joka korostaa kielen ja vuorovaikutuksen merkitystä todellisuuden rakentumisessa (Niska ja muut, 2024; Erikson & Kovalainen, 2016), eikä pyri arvioimaan todellisuuden objektiivista tai annettua olemusta. Näin ollen tutkimuksen tavoitteena ei ole arvioida tekoälyn todellista luotettavuutta, vaan tarkastella, miten luotettavuuden käsitettä merkityksellistetään ja tuotetaan kielellisesti digitaalisessa mediaympäristössä. Samalla tutkimuksemme huomioi laajemman kontekstin, sillä yksittäisissä YouTube-kommenteissa rakentuvat merkitykset kytkeytyvät usein myös laajempiin yhteiskunnallisiin ja kulttuurisiin diskursseihin teknologiasta ja tekoälystä (Stevanovic ja muut, 2024; Niska ja muut, 2024).

Sosiaalinen konstruktionismi toimii näin teoreettisena viitekehyksenä, joka auttaa ymmärtämään, miten YouTube-kommenteissa rakennetaan ja kiistetään erilaisia käsityksiä tekoälyn luotettavuudesta. Kommentit eivät ole vain yksittäisiä mielipiteitä, vaan ne muodostavat osan yhteisöllistä keskustelua, jossa tekoälyn roolista mediassa neuvotellaan kielen avulla. Tämän vuoksi analyysimme keskittyy siihen, millaisin sanavalinnoin, vertauksin ja ilmaisutavoin käyttäjät rakentavat tekoälyyn liittyvää luottamusta sekä miten nämä diskurssit heijastavat laajempia yhteiskunnallisia ja kulttuurillisia tulkintoja teknologiasta.

4.2 YouTube sosiaalisen median keskustelualustana

YouTube toimii maailmanlaajuisesti yhtenä merkittävimpänä käyttäjien tuottaman videosisällön jakelualustana ja on samalla keskeinen väline multimediasisällön leviämiselle verkossa (Striewski ja muut, 2024, s. 1; Wattenhofer ja muut, 2015). Vuonna 2025 YouTube on maailman neljänneksi käytetyin sosiaalisen median alusta Facebookin, WhatsAppin ja Instagramin jälkeen, ja sillä on 2,58 miljardia käyttäjää kuukaudessa (Datareportal, 2025). Vaahensalon (2021, s. 125) mukaan sosiaalisen median alustoilla vakiintuvat omat käyttö- ja keskustelukulttuurinsa, jotka muodostuvat erilaisten toiminta- ja puhetapojen sekä sanattomien käytössääntöjen kautta. Tämä tarkoittaa, että myös YouTuben kommenttiosioissa esiintyvä keskustelukulttuuri on sidoksissa alustan erityisiin toimintoihin ja käytäntöihin.

YouTube voidaan määritellä internetsivustoksi, joka mahdollistaa videoiden lataamisen, niiden katselun ja käyttäjien välisen vuorovaikutuksen (Hautakangas, 2008, s. 10). Lisäksi se voidaan nähdä kulttuurisena ja sosiaalisena tilana, joka heijastaa käyttäjälähtöisen (Web 2.0) aikakauden osallistuvan median ilmiötä (Hautakangas, 2008, s. 10; Lange, 2019, s. 3–4). YouTuben käyttäjät ovat tavallisesti henkilöitä, jotka tuottavat ja/tai kommentoivat sisältöä alustalla ja siten aloittavat tai osallistuvat keskusteluun (Striewski ja muut, 2024, s. 2; Lange, 2019). Käyttäjien osallistuminen sisältää erilaisia aktiivisuuden tasoja, jotka vaihtelevat passiivisesta katsojasta aktiiviseen osallistumiseen, kuten kommentointiin ja oman näkemyksen jakamiseen (Lange, 2019, s. 36–37). Kommentointi edellyttää rekisteröitymistä, mutta niiden lukeminen on kaikille käyttäjille mahdollista (Lehti ja muut, 2016). Tässä tutkimuksessa keskitymme kommentteihin, ja otamme huomioon aktiivisten osallistujien kirjoittamat kommentit.

YouTuben käyttöön vaikuttavat alustan normit ja sen tarjoama anonymiteetti. Spearsin ja Postmesin (2015, s. 61–64) mukaan nimettömyys verkkoympäristöissä lisää vuorovaikutuksen normien noudattamista. Lehti ja muut (2016) huomauttavat, että YouTuben kommenttiketjujen implisiittiset normit näyttävät usein sallivan ja jopa vahvistavan aggressiivista kielenkäyttöä. Lisäksi anonymiteetti ja

ryhmäkäyttäytymismallit voivat lisätä polarisaatiota ja ryhmien välistä vastakkainasettelua (Lehti ja muut, 2016), mikä näkyy erityisesti silloin, kun keskustelut kytkeytyvät kiistanalaisiin aiheisiin. Tällöin ryhmänormien vahva noudattaminen voi kärjistää puhetapoja (Lehti ja muut, 2016), vaikuttaen siihen, millainen kuva jostakin ilmiöstä alustalla rakentuu.

Vaahensalo (2021, s. 118) korostaa, ettei sosiaalisella medialla ole yhtä universaalia historiaa, vaan sen taustat kietoutuvat aina tiettyihin maantieteellisiin sijainteihin, kieliin, kulttuureihin ja identiteetteihin. Tämä ilmiö näkyy myös YouTubessa, jossa kommentit heijastavat sekä käyttäjien yhteiskunnallisia ja kielellisiä konteksteja että laajempaa kulttuurista taustaa. Kommenttiosio toimii näin eräänlaisena diskursiivisena tilana, jossa yksilölliset puhetavat ja yhteisön normit ilmenevät rinnakkain. Samalla se tarjoaa mahdollisuuden vahvistaa tai haastaa vallitsevia käsityksiä, esimerkiksi generatiivisen tekoälyn luotettavuudesta.

4.3 Yhteisöllinen vuorovaikutus YouTubessa

Sosiaalisen median alustat tarjoavat käyttäjilleen monimuotoisia tiloja vuorovaikutukseen, mielipiteiden ilmaisuun ja merkitysten rakentamiseen, joissa yksilöt toimivat vuorovaikutteisesti (Vaahensalo, 2021, s. 112). Teknisten ominaisuuksien lisäksi sosiaaliset rakenteet mahdollistavat laajan osallistumisen, sisältöjen moninaisuuden sekä sosiaalisen vuorovaikutuksen (Hautakangas, 2008, s. 11; Lange, 2019, s. 3–4). YouTubessa tämä näkyy erityisesti kommenttien kautta, sillä kommenttikenttä toimii julkisena alustana, jossa käyttäjät reagoivat videoihin, käyvät keskustelua ja osallistuvat kollektiiviseen merkitysten muodostamiseen. Vuorovaikutus YouTubessa ei rajoitu pelkästään sisällön kuluttamiseen, vaan se on sosiaalinen ja kulttuurinen prosessi, joka yhdistää audiovisuaalisen viestinnän, yhteisöllisen jakamisen ja keskustelun (Lange, 2019; Wattenhofer ja muut, 2015). Dynel (2014) korostaa YouTube-viestinnän keskeisenä piirteenä sitä, että alustalla on periaatteessa rajaton määrä mahdollisia osallistujia, jotka

eivät yleensä tunne toisiaan. Tämä laajentaa vuorovaikutuksen mittakaavan yksittäisen keskustelun sijaan laajaksi yleisöksi.

Lehti ja muut (2016) kuvaavat YouTuben kommenttiketjuja “massiivisena polylogina”, jossa kommentit muodostavat moniäänisen, jatkuvasti muotoutuvan ja useita rinnakkaisia keskusteluja sisältävän rakenteen. Lisäksi kommenttiketjujen osallistujarakenne on erityisen dynaaminen: käyttäjät voivat milloin tahansa liittyä keskusteluun, poistua siitä, aloittaa uuden kommentin tai vastata olemassa olevaan ketjuun (Lehti ja muut, 2016). Kommenttiketju säilyy niin kauan kuin video on alustalla, mikä mahdollistaa keskustelun jatkumisen jopa kuukausien tai vuosien ajan (Lehti ja muut, 2016). Kommenttiosio voidaan sulkea myös videontekijän toimesta. Striewski ja muut (2024, s. 8) toteavat, että YouTuben tarjoamat affordanssit muokkaavat vuorovaikutusta. Tämän vuoksi alusta näyttäytyy dynaamisena tilana, jossa yhteisöllisyys rakentuu käyttäjien toiminnan kautta.

Tutkimuksemme kannalta olennaista on, että YouTuben kommenteissa ja käyttäjien vuorovaikutuksessa syntyvät käsitykset tekoälyn riskeistä ja hyödyistä ovat vahvasti kontekstisidonnaisia. Brauner ja muut (2023, s. 3) osoittavat, että ihmiset arvioivat tekoälyä enemmän tilanteen ja käyttöyhteyden kuin yksilöllisten taustatekijöiden perusteella, mikä tekee alustakohtaisista keskusteluista erityisen merkityksellisiä. YouTube-kommentit tarjoavat näin kontekstin tarkastella, miten käyttäjät osallistuvat keskusteluun, jakavat näkemyksiään, arvottavat generatiivisen tekoälyn luotettavuutta, sekä sitä, millaisia diskursseja sen ympärille muodostuu.

5 Generatiivisesta tekoälystä rakentuvat diskurssit

Tässä luvussa esittelemme tutkimusaineistomme analyysin ja keskeisimmät havainnot. Kuten aiemmin mainitsimme, analyysin tavoitteena on tunnistaa ja kuvata, millaisia diskursseja suomalaiset mediankäyttäjät rakentavat generatiivisen tekoälyn luotettavuudesta. Analyysi on toteutettu diskurssianalyyttisella lähestymistavalla hyödyntäen Pietikäisen ja Mäntysen (2019, s. 253–311) esittämiä analyysin vaiheita. Aineistoon perehdyttiin kokonaisuutena, jonka jälkeen kommentit koodattiin laadullisen analyysin menetelmin ja ryhmiteltiin viiteen pääkategoriaan (Junila, 2021). Olemme toteuttaneet diskurssien jaottelun itse aineiston pohjalta, eikä se perustu valmiisiin malleihin tai jaotteluihin. Diskurssien muodostaminen on edellyttänyt tulkinnallisia valintoja, sillä laadullisessa analyysissa tutkijan rooli ja tekemät rajaukset vaikuttavat analyysin lopputulokseen.

Tutkimuseettisesti on huomioitu käyttäjien anonymiteetti ja vastuullinen aineiston käsittely (Kuula-Luumi, 2021), joten käyttäjien nimiä ei tuoda esille analyysissä. Seuraavissa alaluvuissa esittelemme aineistosta tunnistetut diskurssit viiden pääkategorian kautta, joita ovat affektiivinen, teknologinen, reflektiivinen, inhimillinen sekä eettinen ja yhteiskunnallinen diskurssi (ks. Kuvio 1.). Jokainen kategoria kuvastaa erilaista tapaa puhua generatiivisesta tekoälystä, sen vaikutuksista ja riskeistä, sekä sitä, miten luottamus tai sen puute ilmentyy. Analyysin tueksi tuomme esiin kunkin diskurssin yhteydessä numeroidut esimerkkikommentit.

Kuvio 1. Generatiivisesta tekoälystä rakentuvat diskurssit.

Analyysimme kannalta keskeinen havainto on se, että vaikka kommenteissa käytetään usein yleistermiä "tekoäly" tai "AI", ne viittaavat generatiiviseen tekoälyyn. Kontekstin luhistumisen kannalta on oleellista tuoda myös esille, että kaikissa kommenteissa ei suoraan mainita tekoälyä, sillä ne ovat osana laajempaa kommenttiketjua. Kyseiset kommentit kuitenkin käsittelevät tätä aihetta. Osa kommenteista on lyhennetty niiden pituuden vuoksi siten, että ne edustavat tutkimuksemme kannalta olennaista diskurssia. Lisäksi osa kommenteista voi ilmentää useita eri diskursseja, ja yksittäisen kommentin eri osat voivat liittyä eri kategorioihin. Analyysissämme kommenttien luokittelu on kuitenkin toteutettu sen perusteella, mihin diskurssiin niiden keskeinen merkitys tai painotus ensisijaisesti sijoittuu.

5.1 Affektiivinen diskurssi

Affektiivinen diskurssi kattaa aineistossamme kommentit, joissa generatiiviseen tekoälyyn suhtaudutaan tunteiden, arvojen ja asenteiden kautta. Se kuvaa, miten nämä affektiiviset näkökulmat liikkuvat, muuntuvat ja vahvistuvat diskurssin sekä kielen kautta (Milani & Richardson, 2021). Kommentit ilmentävät sekä henkilökohtaisia mielipiteitä että laajemmin yhteisöllisesti jaettuja merkityksiä generatiivisesta tekoälystä.

Tärkeä havainto on, että affektiivisuus esiintyy lähes kaikissa aineistomme kommenteissa jossain muodossa. Päätimme nostaa sen kuitenkin erilliseksi pääkategoriaksi, sillä kommenteissa tunnepitoiset ilmaisut olivat hallitsevia ja vaikuttivat sekä keskustelun sävyyn että siihen, millaisia diskursseja generatiivisesta tekoälystä rakennetaan. Affektiivisuus toimii siten keskeisenä analyysin kohteena, sillä se kuvaa käyttäjien tunteellista ja asenteellista suhtautumista generatiiviseen tekoälyyn.

5.1.1 Torjuva diskurssi

Torjuvassa diskurssissa generatiivinen tekoäly nähdään kielteisenä ilmiönä. Aineistomme kommentit rakentuvat voimakkaiden tunteiden ja kielteisten asenteiden kautta. Lisäksi useat kommentit toistavat kokemuksia generatiivisen tekoälyn turhuudesta herättäen negatiivisia tunteita (ks. kommentit 1–3).

(1) En tykkää tekoälystä yhtään (26.8.2025)

(2) Tekoäly on perseestä (26.8.2025)

(3) Tekoäly ei tarvita yhtään mihinkään. Tekoäly on perseestä (27.8.2025)

Torjuvassa diskurssissa generatiivinen tekoäly näyttäytyy pääosin hyödyttömänä ja kielteisenä. Kommenttien ilmaisut, kuten *"en tykkää yhtään"* ja *"perseestä"* korostavat kommentoijien tunteellista sävyä ja torjuvaa asennetta generatiivista tekoälyä kohtaan (Laapotti ja muut, 2024, s. 148). Vaikka monet tähän kategoriaan liittyvät kommentit ovat lyhyitä ja suoraviivaisia, ne ilmentävät selkeästi torjuvaa ja kielteistä asennoitumista generatiivista tekoälyä kohtaan.

Kommentit rakentavat yhdessä diskurssin, jossa generatiivinen tekoäly näyttäytyy epäluotettavana ja ei-toivottuna. Torjuva diskurssi ei kuitenkaan rajoitu pelkästään käyttäjien henkilökohtaisiin mielipiteisiin, vaan rakentaa julkisen keskustelun alustalla laajempaa kielteistä käsitystä generatiivisesta tekoälystä. Kommenttien tunnepitoisuus luo kollektiivisen tulkintakehyksen, jossa generatiivinen tekoäly nähdään turhana ja lähtökohtaisesti kielteisenä. Tämä korostaa, että käyttäjien tunteet tai yksittäiset mielipiteet voivat sosiaalisen median alustoilla vahvistua ja muodostaa siten yhteisöllisiä merkityksiä, jotka vaikuttavat mahdollisesti keskustelun sävyyn ja yleiseen käsitykseen generatiivisesta tekoälystä.

5.1.2 Pelon ja huolen diskurssi

Pelon ja huolen diskurssi korostaa aineistossamme generatiivisen tekoälyn aiheuttamaa epävarmuutta, tulevaisuuteen liittyviä uhkakuvia sekä välittömiä huolenaiheita. Tässä diskurssissa painottuvat erityisesti pelon ja huolen affektiiviset ulottuvuudet sekä niihin liittyvät arvotukset.

(4) Aika pelottavaa on kyllä! [...] (26.8.2025)

(5) just tota pelkään, että kaikki tyhmistyy niin paljon, että esim. jos joku kysyy vaikka mitä kuuluu niiin kysytääb tekoälyltä mitä minulle kuuluu. (26.8.2025)

Tunnistimme aineistomme kommenteista toistuvia huolenaiheita, jotka luovat kollektiivista ymmärrystä generatiivisesta tekoälystä ja sen mahdollisista uhista. Pelon ja huolen diskurssi korostaa generatiivisen tekoälyn herättämää epävarmuutta ja uhkaa (ks. kommentti 4) – tilanteita, joissa luottamus on luonteeltaan heikkoa tai haavoittuvuutta ei haluta hyväksyä. Aineistomme kommenttien affektiiviset sävyt viittaavat siihen, että käyttäjät arvioivat rationaalisesti generatiivisen tekoälyn riskit suuremmiksi kuin sen tarjoamat hyödyt, mikä johtaa luottamuksen puutteeseen.

Aineistomme kommentit ilmentävät myös konkreettisesti huolta ja pelkoa siitä, miten generatiivinen tekoäly vaikuttaa erityisesti sosiaaliseen vuorovaikutukseen ja ajattelun itsenäisyyteen (ks. kommentti 5). Lisäksi käyttäjät ilmaisevat pelkoa generatiivisen tekoälyn hallitsemattomasta kehityksestä, ihmisten korvaamisesta sekä ihmisten ja tekoälyteknologioiden välisestä rajan hämärtymisestä. Pelon ja huolen tunteet ovat relevantteja, sillä Glikson ja Woolley (2020, s. 642–643) huomauttavat, että tekoälyn kehittäjät voivat vaikuttaa käyttäjien tunteisiin manipuloimalla tekoälyn esitystapoja ja käyttäytymistä. Toisaalta näiden kommenttien taustalla voi olla puutteellinen tekoälylukutaito, joka voi lisätä pelkoa ja huolta.

5.2 Teknologinen diskurssi

Teknologinen diskurssi käsittelee tapoja, joilla käyttäjät jäsentävät generatiivista tekoälyä teknologisena ilmiönä ja välineenä. He pohtivat tekoälyn konkreettisia hyötyjä, rajoituksia ja luotettavuutta, mutta myös sen asemaa osana laajempaa teknologista murrosta. Diskurssissa rakentuu käsitys tekoälystä paitsi työkaluna, myös yhteiskunnallisesti merkittävänä teknologiana, jonka kehitykseen suhtaudutaan varauksella, mutta myös innostuneesti.

5.2.1 Käytännöllinen diskurssi

Käytännöllisessä diskurssissa generatiivinen tekoäly näyttäytyy ensisijaisesti hyödyllisenä välineenä, joka helpottaa arkea ja sujuvoittaa tekemistä. Käyttäjät lähestyvät generatiivista tekoälyä toiminnallisesta näkökulmasta: sitä käytetään esimerkiksi kuvien, videoiden ja informaation tuottamiseen. Gliksonin ja Woolleyn (2020, s. 630) mukaan tekoäly nähdään tarjoavan uusia mahdollisuuksia myös tulevaisuudessa eri aloille.

> (6) Tekoälyllä on hyvä tehdä kuvia ja videoita ja saada informaatiota kyllä sitä tulee käytettyä (26.8.2025)

Generatiivista tekoälyä kuvataan kommenteissa luontevana osana käyttäjien digitaalista työskentelyä (ks. kommentti 6). Siihen suhtaudutaan muihin diskursseihin verrattuna myönteisemmin, ja sen potentiaalia tarkastellaan teknisenä edistysaskeleena. Tällaisella ilmaisutavalla generatiivinen tekoäly nähdään tehokkuuden ja luovuuden tukijana eikä uhkana. Käytännöllinen diskurssi ilmentää siten myös pragmaattista suhtautumista generatiiviseen tekoälyyn. Teknologiaa ei arvioida moraalisin tai emotionaalisin perustein, vaan sen toimivuutta, käyttökelpoisuutta ja hyödyllisyyttä korostetaan.

Generatiivinen tekoäly hyödyntää datasta oppimista, mikä voi johtaa käyttäytymisen vaihteluun, vaikka tavoitefunktio pysyisi samana (Glikson & Woolley, 2020, s. 631). Tämä korostaa käyttäjien tekoälylukutaitoa ja sitä, voiko generatiivisen tekoälyn tuottamaan

sisältöön aina luottaa. Käytännöllinen diskurssi asettuu myös vastakohdaksi pelon ja huolen diskurssille, jossa generatiivinen tekoäly koettiin hallitsemattomana ja uhkaavana.

5.2.2 Teknologinen metadiskurssi

Teknologisessa metadiskurssissa käyttäjät tarkastelevat generatiivista tekoälyä teknologiana itsessään: sen rajoituksia, toimintaperiaatteita, kehityssuuntaa sekä yhteiskunnallisia vaikutuksia. Tähän diskurssiin kuuluu sekä kriittistä että pohdiskelevaa puhetta generatiivisen tekoälyn luotettavuudesta ja vaikutuksista tiedon tuotantoon ja vuorovaikutukseen (ks. kommentti 7).

(7) Kommentit täynnä tekoäly kommentteja (26.8.2025)

(8) [...] Niin monesti nähnyt asioita joissa AI alkaa ns. "arvailemaan" tai liittämään asiaan liittyviä asioita omalla "logiikallaan". Välillä tulokset on absurdeja, mutta välillä sellaisia joita monet luulevat todeksi koska se kuulostaa siltä. Ja juuri tuo että AI:lla ei ole kykyä ymmärtää lähteen luotettavuutta. Ja varsinkin jos AI kopio epäluotettavaa lähdettä on todennäköisempää että "muut" AI tulee jatkossa käyttämään tätä virheellistä tietoa. [...] (27.8.2025)

Teknologisessa metadiskurssissa kiinnitetään huomiota generatiivisen tekoälyn luotettavuuden ja aitouden ongelmiin. Käyttäjät pohtivat, miten tekoälyn tuottama sisältö voi olla ”harhaanjohtavaa” ja perustua ”arvailuun”, tai miten tekoäly ei ymmärrä lähteen luotettavuutta (ks. kommentti 8). Tällaiset huomiot ilmentävät teknologista tietoisuutta: käyttäjät ymmärtävät tekoälyn toimivan algoritmisten oletusten varassa, ja pitävät tärkeänä erottaa ihmisen tuottaman sisällön tekoälyn tuottamasta sisällöstä. Kommenteissa ilmennetään myös turhautumista siitä, miten tekoäly muuttaa keskustelukulttuuria ja luo epäaitouden tunnetta sosiaalisessa mediassa.

5.3 Reflektiivinen diskurssi

Reflektiivinen diskurssi käsittelee käyttäjien pohdintaa generatiivisesta tekoälystä, sen merkityksestä ja vaikutuksista. Kommentit ilmentävät tietoista tarkastelua, itsearviointia ja keskustelua, joissa saman kommentin sisällä voi esiintyä sekä myönteisiä että kielteisiä näkökulmia.

5.3.1 Kaksijakoinen diskurssi

Kaksijakoinen diskurssi kuvaa kommentteja, joissa käyttäjät ilmentävät samanaikaisesti keskenään ristiriitaisia näkökulmia generatiivisesta tekoälystä. Diskurssi tuo esiin käyttäjien tietoisen tai tiedostamattoman ristiriitaisuuden (ambivalenssin), jossa käyttäjät pohtivat generatiivisen tekoälyn hyötyjä ja ongelmia samanaikaisesti.

> (9) Oon kuullu monia uutisjuttuja siitä, mitä kaikkee ikävää tekoäly on saannu ai kaa [...] Mutta samaa aikaa, tekoäly osaa olla myös hyödyllinen työkalu joillekkin ja mm. tekoälyllä luotua musiikkia olen kuunnellu viime kuukausina aika usein ja yleensä ne kappaleet on aika hyviikin musiikin suhteen, lisäksi tekoälyn luoma laulaminen on kuulostanu hyvältä [...] tekoäly on antanu niin nautintoo, kun taas samaa aikaa sitä suositaan, vihataan, ja paheksutaan. Joten en osaa itse päättää, asetuntko tekoälyy vastaa, vai sen puolelle, vai pysynkö puolueettomana[...] (26.8.2025)

(10) Jossain asioissa hyvä ja jossain asioissa taas huono! (26.8.2025)

Aineistomme kommenteissa kaksijakoisuus ilmentyy eksplisiittisesti. Kommentit tuovat selkeästi esiin ambivalenssin eli osa käyttäjistä korostaa generatiivisen tekoälyn mahdollisuuksia huomioiden samalla epäluotettavat tai haitalliset vaikutukset (ks. kommentti 9 ja 10). Käyttäjät muodostavat käsityksen generatiivisen tekoälyn luotettavuudesta arvioimalla sen toimintaa ja vaikutuksia. Luottamus ei kuitenkaan ole yksiselitteistä, vaan dynaamista ja kontekstisidonnaista, koska arviot riippuvat myös henkilökohtaisista kokemuksista ja muiden havainnoista.

Glikson ja Woolley (2021, s. 627–628) huomauttavat kuitenkin, että käyttäjät eivät aina ole tietoisia tekoälyn todellisesta teknisestä kehittyneisyydestä, joka voi vaikeuttaa generatiivisen tekoälyn luotettavuuden arviointia. YouTube-kommenteissa ilmenevä luottamus tai epäluottamus ei siten kuvaa generatiivisen tekoälyn todellista toimintakykyä, vaan käyttäjien subjektiivisia tulkintoja tai odotuksia.

5.3.2 Yhteisöllisen reagoinnin diskurssi

Reflektiivisen diskurssin kontekstissa yhteisöllinen reagointi ilmenee erityisesti siinä, miten käyttäjät arvioivat ja tulkitsevat keskustelua generatiivisesta tekoälystä ja sen roolista YouTube-kommenteissa. Käyttäjät eivät tarkastele generatiivista tekoälyä teknologisena ilmiönä, vaan tuovat uudenlaisen lähestymistavan arvioiden YouTuben verkkokommenteissa käytyä keskustelun dynamiikkaa.

(11) Oota pitää kysyä tekoälyltä mitä mieltä oon tästä (26.8.2025)

(12) Koomista miten ekat kommentit oli nimenomaa botteja :D (26.8.2025)

Käyttäjien näkemykset kiinnittävät huomiota botteihin ja tekoälyllä luotuihin kommentteihin. Heidän reaktionsa vaihtelevat turhautumisesta osittain humoristiseen sävyyn, jossa botit ja generatiivisen tekoälyn tuottama sisältö tunnistetaan (ks. kommentti 11). Yhteisöllisen reagoinnin diskurssi tuo esille sen, että käyttäjät käsittelevät generatiivista tekoälyä kollektiivisesti rakentuvan epävarmuuden kautta. Generatiivista tekoälyä ei nähdä tässä kontekstissa ainoastaan työkaluna, vaan osallistuvana toimijana. Sen tunnistettavuus (ks. kommentti 12) muovaavat käyttäjien käsityksiä ja odotuksia generatiivisen tekoälyn luotettavuudesta. Tämä heijastuu myös siihen, kuinka YouTuben verkkokommenteissa keskustelun aitous ja inhimillisyys muuttuvat, kun generatiivinen tekoäly osallistuu julkiseen keskusteluun. Tekoälyn läsnäolo muuttaa keskustelun luonnetta, koska huomio saattaa kohdistua enemmän itse kommenttien luotettavuuteen.

Tämä diskurssi tuo esille myös sen, että käyttäjät eivät voi tietää, onko sisältö tai kommentit todellisuudessa luotu tekoälyllä vai ei. Tekoälyyn viittaaminen (ks. kommentti 12) toimii eräänlaisena keinona jäsentää keskustelun aitoutta ja arvioida kommenttien alkuperää tilanteissa, joissa ihmisen ja tekoälyn tuottamaa sisältöä on yhä vaikeampi erottaa toisistaan. Tämä murentaa samanaikaisesti käyttäjien luottamusta tekoälyä kohtaan, ja he saattavat alkaa myös kyseenalaistamaan minkä tahansa sisällön olevan tekoälyn tuottamaa. Sama ilmiö tulee esille myös pelon ja huolen, käytännöllisyyden ja yhteisöllisen reagoinnin diskursseissa sekä teknologian metadiskurssissa. Tekoälylukutaidon merkityksen korostaminen voisi toimia keinona ymmärtää ja tunnistaa tekoälyllä luotua sisältöä.

5.4 Inhimillinen diskurssi

Inhimillinen diskurssi käsittelee sitä, miten käyttäjät jäsentävät generatiivisen tekoälyn ja ihmisyyden välistä suhdetta. Diskurssissa korostuvat pohdinnat ihmisen luovuudesta, ajattelusta ja toimijuudesta sekä generatiivisen tekoälyn roolista näiden rinnalla tai vastakohtana. Diskurssi ilmentää pohdintaa myös siitä, kuinka generatiivisen tekoälyn yleistyminen voisi kaventaa ihmisen ja tekoälyn välistä rajapintaa.

5.4.1 Ihmisen ainutlaatuisuuden diskurssi

Tässä alakategoriassa generatiivista tekoälyä tarkastellaan suhteessa ihmisen erityispiirteisiin, kuten luovuuteen, tunneälyyn ja omaperäiseen ajatteluun. Käyttäjät korostavat kommenteissa, että generatiivinen tekoäly ei kykene tuottamaan aidosti merkityksellistä tai inhimillistä sisältöä, sillä sen tuotokset perustuvat algoritmeihin, eivätkä kokemuksiin, tunteisiin tai tietoiseen tulkintaan.

(13) Vähän sama kun palkkaa artistin tekemään jonkun työn ei siitä koskaan tule ihan sellanen kuin itse visioi (26.8.2025)

(14) [...] Mutta juurikin ihmisen puolesta ajatteluun AI ei todellakaan ole hyvä :) (27.8.2025)

Generatiivisen tekoälyn nähdään uhkaavan ihmisen älyllistä kehittymistä ja ajattelua, jos sen käyttö vie tilaa ihmisten omalta ajattelulta, oppimiselta ja kriittiseltä pohdinnalta (ks. kommentti 14). Aineiston kommentit korostavat, että generatiivisella tekoälyllä luotu sisältö voi olla teknisesti onnistunutta, mutta niistä puuttuu inhimillisyydelle tyypillinen näkemyksellisyys ja merkityksellisyys. Tämä vahvistaa ajatusta siitä, että luovuus on ihmiselle ominaista ja korvaamatonta (ks. kommentti 13).

Kommentit rakentavat yhdessä diskurssin, jossa generatiivisen tekoälyn vaikutus ihmisen omaan ajatteluun on kaksijakoinen. Se voidaan nähdä helpottavana apuvälineenä ja työkaluna, mutta samalla se uhkaa ihmisen luovuutta, sosiaalista toimintaa ja kriittistä pohdintaa. Tämä diskurssi ja keskustelu korostaa ihmisen ainutlaatuisuutta ja inhimillisyyttä, jossa generatiivisen tekoälyn tuottama sisältö nähdään usein merkityksettömänä, mekaanisena tai jopa harhaanjohtavana.

5.4.2 Ihmisen ja tekoälyn välinen diskurssi

Tässä alakategoriassa keskustelu keskittyy generatiivisen tekoälyn rooliin ihmisen rinnalla. Käyttäjät pohtivat, toimiiko generatiivinen tekoäly tehokkaana apuvälineenä vai uhkaako se syrjäyttää ihmisen itsenäisen toiminnan. Generatiivista tekoälyä tarkastellaan joko inhimillisen työn tukena tai sen korvaajana, mikä herättää kysymyksiä ihmisarvosta, toimijuudesta ja näiden merkityksestä myös tulevaisuudessa.

(15) [...] en muuten pääsis koulua läpi (26.8.2025)

(16) [...] se on myös perseestä. Ihmisistä tulee tyhmiä eikä ite opi mitään (26.8.2025)

Kuten ihmisen ainutlaatuisuutta korostavassa diskurssissa, aineistomme kommentit osoittavat myös tästä näkökulmasta sen, että generatiivinen tekoäly voi heikentää

ihmisen omaa oppimista ja itsenäistä ajattelua (ks. kommentit 15 ja 16). Käyttäjät pohtivat, että liiallinen luottamus tekoälyyn voi jopa syrjäyttää ihmisen aktiivisen osallistumisen ja ajattelun, joka uhkaa siten oppimista ja älyllistä kehittymistä. Generatiivista tekoälyä ei nähdä hyödyllisenä työkaluna, vaan ihmisten oman ajattelun korvaajana, kuten aiemmassa diskurssissa. Sen liiallisen käytön uskotaan murentavan ihmisen ajattelun, oppimisen ja osaamisen, joka vaikuttaa myös eettiseen ja yhteiskunnalliseen ulottuvuuteen.

Kuten yhteisöllisen reagoinnin diskurssi, myös tämä näkökulma korostaa tekoälylukutaidon merkitystä. Käyttäjien huoli ei kohdistu ainoastaan generatiivisen tekoälyn käyttöön, vaan siihen, ymmärtävätkö ihmiset sitä, että sen liiallinen käyttö voi johtaa tekoälyriippuvuuteen. Ilman tätä ymmärrystä tekoäly voi näyttäytyä uhkana, joka heikentää käyttäjien kykyä arvioida sen toimintaa kriittisesti. Lisäksi tämä diskurssi korostaa, että ihmisen asema tulisi edelleen olla aktiivinen ja reflektiivinen tekoälyn rinnalla, eikä tekoälyn toiminnalle ja päätöksille tule antaa liiallista valtaa erityisesti eettisten ja tiedollisesti kestävien ratkaisujen suhteen.

5.5 Eettinen ja yhteiskunnallinen diskurssi

Tämä diskurssi käsittelee generatiivisen tekoälyn liittyviä arvoja, vastuuta ja yhteiskunnallisia vaikutuksia. Käyttäjät pohtivat generatiivisen tekoälyn eettistä hyväksyttävyyttä ja sen seurauksia yksilöille, yhteisölle sekä ympäristölle. Diskurssissa korostuu huoli siitä, kuka kontrolloi generatiivisen tekoälyn käyttöä tai miten väärinkäytöstä tulisi kantaa vastuuta ja millä tavoin se muovaa luottamusta.

5.5.1 Vastuun ja moraalin diskurssi

Tässä alakategoriassa tarkastellaan generatiivisen tekoälyn käytön hallintaa, vastuunkantoa ja moraalisia rajoja. Käyttäjät pohtivat, kenelle generatiivisen tekoälyn

väärinkäytön seuraukset lopulta kuuluvat sekä millaisia eettisiä kysymyksiä tekoälyn käyttö arjen tilanteissa herättää. Pohdinta ulottuu myös yksittäisten käyttäjien moraaliseen vastuuseen, teknologian taustalla olevien kehittäjien rooliin sekä uskonnollisiin ja filosofisiin merkityksiin.

> (17) Ja se erät teko äly voidaan käyttää rikos asiaan ja tietokalastetuun ja tekoälystä hyvät ja huonot puolet jopa YouTubessa pyörii botti kommentti ja käyttäjiä jolla linkki kalastus sivustolle (26.8.2025)
>
> (18) Siis just toi että kouluissa oppilaat käyttää tekoälyä salaa, vaikka opettaja vahvasti kieltää sen. Nehän periaatteessa valehtelevat itselleen […]. (5.9.2025)

Kommenteissa ilmentyvät näkemykset heijastavat laajempaa huolta generatiivisen tekoälyn kontrollin puutteesta sekä siitä, kuinka tekoälyllä mahdollistetut väärinkäytöt kohdistuvat esimerkiksi haavoittuvassa asemassa oleviin käyttäjiin (ks. kommentti 17). Lisäksi kommenteista välittyy ajatus siitä, kuinka generatiivisen tekoälyn käyttö uhkaa oppimista ja henkilökohtaista vastuuta (ks. kommentti 18). Tällöin generatiivinen tekoäly nähdään toiminnan mahdollistajana, mutta vaikutukset koskevat käyttäjän eettistä vastuuta. Kommentit osoittavat, että generatiivisen tekoälyn käyttö olisi siten myös moraalisesti kyseenalaista. Hermannin (2021, s. 1270–1271) mukaan tekoälylukutaito on edellytys eettisesti kestävälle tekoälyn käyttämiselle.

Käyttäjät rakentavat keskustelun kautta diskurssia, jossa generatiivinen tekoäly synnyttää uudenlaista vinoumaa: se pystyy mahdollistamaan toiminnan, jonka seuraukset eivät ole helposti jäljitettävissä, eikä vastuunkanto ole selkeästi määriteltävissä. Wolfin ja muiden (2024, s. 8) mukaan tekoälyllä itsellään ei ole intentioita tai odotuksia, mutta sen taustalla toimivilla organisaatioilla sekä kehittäjillä on. Tämä haastaa tulkintoja siitä, kenelle vastuu todellisuudessa kuuluu.

5.5.2 Ekologisuuden diskurssi

Tässä alakategoriassa generatiivista tekoälyä tarkastellaan ekologisesta näkökulmasta. Käyttäjät ilmaisevat huolta tekoälyn laajamittaisen käytön ympäristövaikutuksista, kuten energiankulutuksesta ja resurssien tuhlauksesta. Havainnot nostavat esiin kysymyksen siitä, onko generatiivisen teköälyn käyttö kestävää suhteessa sen tuottamaan hyötyyn.

> (19) [...] miten jengi ei tajua miten haitallista se on ympäristölle käyttää tekoälyä. 🙄 (26.8.2025)

> (20) [...] kuluttaa vettä (27.8.2025)

Kommentit tuovat esille yhteisen huolen siitä, kuinka monet käyttäjät eivät ymmärrä tekoälyn käytön haitallisuutta ja seurauksia ympäristölle (ks. kommentti 19). Kommenttien sanavalinnat, kuten "*miten jengi ei tajua*" osoittaa myös arvioivan ja kritisoivan asenteen muita käyttäjiä kohtaan, joka vaikuttaa keskusteluun generatiivisen tekoälyn ekologisista vaikutuksista ja mahdollisista rajapinnoista. Generatiivisen tekoälyn käyttö voidaan siten nähdä myös eettisesti merkityksellisenä toimintana, jonka hyötyjä ja haittoja tulisi tarkastella suhteessa sen ympäristövaikutuksiin. Ekologisuuden diskurssi muodostaa myös vastakkainasettelun kommenteille, jotka korostavat tekoälyn hyödyllisyyttä ja arkea helpottavia toimintoja. Käyttäjien keskustelu korostaa myös laajemmin teknologian käyttöönottoon liittyvää kestävyysnäkökulmaa. Generatiivisen tekoälyn käytön seuraukset ilmenevät luonnon kustannuksina (ks. kommentti 20), joita yhteiskunta joutuu kantamaan.

6 Johtopäätökset

Tutkimuksessa analysoimme YouTuben verkkokommenteissa rakentuvia diskursseja ja muodostimme käsityksen siitä, millaisia näkemyksiä suomalaiset mediankäyttäjät liittävät generatiivisen tekoälyn luotettavuuteen ja sillä tuotettuun sisältöön. Tarkastelu rajautui generatiiviseen tekoälyyn yhteiskunnallisena ja julkisena keskustelunaiheena, joka näyttäytyi kommenttien perusteella sekä uhkia että mahdollisuuksia sisältävänä ilmiönä. Tutkimuksemme osoittaa, että generatiivinen tekoäly herättää monenlaisia mielipiteitä, tunteita ja näkökulmia, joiden ympärille muodostuu erilaisia, mutta osin päällekkäisiä diskursseja.

Tutkimuksemme kannalta keskeisin havainto on se, että analyysimme diskurssit generatiivisesta tekoälystä korostavat enemmän epäluottamusta kuin luottamusta. Aineistomme kommentit ilmensivät harvoin vain yhtä Rousseuau'n ja muiden (1998) luottamuksen muotoa, vaan ne kytkivät toisiinsa useita eri ulottuvuuksia. Lisäksi tutkimuksemme vahvisti aiempien tutkimuksien (Wolf ja muut, 2024; Lee & See, 2004, Huynh & Aichner, 2025; Hoff & Bashir, 2015) käsitystä siitä, että luottamus on dynaaminen prosessi käyttäjän ja generatiivisen tekoälyn välisessä suhteessa.

Näiden pohjalta tutkimuksemme tuo esiin uuden näkökulman generatiiviseen tekoälyyn kohdistuvaan luottamukseen: käyttäjien kommentit ilmaisevat Rousseau'n ja muiden (1998) luottamuksen muotojen yhdistelmää, jolloin luottamus generatiivista tekoälyä kohtaan näyttäytyy monikerroksisena ilmiönä, jossa yksilön tulkinnat, kokemukset ja konteksti ovat keskeisiä. Aiemmat kommentit ja muiden näkemykset voivat vaikuttaa omaan suhtautumiseen ja luottamukseen. Näin diskurssit muodostuvat useiden tekijöiden ja kontekstin vuorovaikutuksessa korostaen luottamuksen monitasoista, julkista ja vuorovaikutteista luonnetta. YouTube ei ole vain paikka mielipiteiden jakamiselle, vaan konteksti, jossa generatiivisen tekoälyn luotettavuutta koskevat diskurssit muotoutuvat, vahvistuvat ja samalla niitä haastetaan.

Analyysimme osoittaa, että käyttäjien käsitykset generatiivisesta tekoälystä synnyttävät erilaisia diskursseja puhetapojen ja kokemusten kautta luoden sosiaalista todellisuutta. Sosiaalisen konstruktionismin näkökulmasta generatiivisen tekoälyn luotettavuus ei näyttäydy objektiivisena tai teknisesti määrittyvänä ominaisuutena, vaan ilmiönä, joka muotoutuu jatkuvasti diskurssien kautta. Tällöin tekoäly voi näyttäytyä yhtä aikaa uhkana, hyödyllisenä välineenä, moraalisena ongelmana tai ihmisyyttä haastavana ilmiönä. Diskurssit eivät heijasta tekoälyn todellista luotettavuutta, vaan ennemmin tapoja, joilla käyttäjät neuvottelevat sen roolista. Kommenteissa rakentuvat käsitykset ovat myös osa laajempaa yhteiskunnallista keskustelua teknologisesta murroksesta, sekä eettisistä ja inhimillisistä rajoista.

Analyysimme kommentteihin on voinut vaikuttaa videon lähtökohtaisesti negatiivissävytteinen sisältö sekä otsikko, joka on voinut ohjata keskustelun sävyä sekä vaikuttaa analyysimme diskurssien rakentumiseen. Tutkimuksemme muut rajoitukset liittyvät yhteen sosiaalisen median alustaan sekä videoon, aineiston kokoon ja sen rajaukseen. Laajemmalla otannalla, eri aineistojen ja alustojen vertailulla olisi todennäköisesti saatu kattavampi käsitys käyttäjien generatiivisen tekoälyn luotettavuuteen liittyvistä diskursseista, mikä olisi vahvistanut analyysin luotettavuutta ja tulosten yleistettävyyttä.

Tutkimuksemme osoittaa, että tekoälylukutaito on yksi tulevaisuuden keskeisimmistä taidoista medialukutaidon rinnalla. Generatiivisen tekoälyn yleistyessä raja ihmislähtöisen ja tekoälyn tuottaman sisällön välillä hämärtyy, mikä näkyy käyttäjien kommenteissa epävarmuutena ja luottamuksen puutteena. Tämä korostaa tekoälylukutaitoa, joka mahdollistaa tekoälyllä luotujen sisältöjen kriittisen arvioinnin (Hermann, 2021, s. 1271). Tekoälylukutaito tukee myös käyttäjien toimijuutta ja tarjoaa perustan kestävämmälle luottamukselle tekoälyjärjestelmiä kohtaan. Yhteiskunnallisesta näkökulmasta yksilön kyky tunnistaa generatiivisen tekoälyn tuottama sisältö on olennainen osa tulevaisuuden mediataitoja, sillä luottamus tekoälyyn edellyttää riittävää tekoälylukutaitoa ja kriittistä ajattelua.

Lähteet

AlAli, R. & Wardat, Y. (2024). Opportunities and Challenges of Integrating Generative Artificial Intelligence in Education. *International Journal of Online Research, 10(1)*, 1–15. https://doi.org/10.61707/8y29gv34

Brauner, P., Alexander, H., Ralf, P. & Martina, Z. (2023). *What does the public think about artificial intelligence?—A criticality map to understand bias in the public perception of AI.* Frontiers Computer Science, 5, 1113903. https://doi.org/10.3389/fcomp.2023.1113903

Bochniarz, K. T., Czerwiński, S. K., Sawicki, A. & Atroszko, P. (2022). Attitudes to AI among high school students: Understanding distrust towards humans will not help us understand distrust towards AI. *Personality and Individual Differences, 185*, 1–7. https://doi.org/10.1016/j.paid.2021.111299

Burr, V. (1995). *An Introduction to Social Constructionism* (1. painos). London: Routledge. https://doi.org/10.4324/9780203133026

Datareportal. (2025). *Global Social Media Statistics*. Noudettu 2.12.2025 osoitteesta https://datareportal.com/social-media-users

Dynel, M. (2014). Participation framework underlying YouTube interaction. *Journal of Pragmatics, 73*, 37–52. https://doi.org/10.1016/j.pragma.2014.04.001

Eriksson, P, & Kovalainen, A. (2016). *Qualitative methods in business research* (2. painos). Sage Publications.

Euroopan parlamentti ja neuvosto. (2024). Asetus (EU) 2024/1689 tekoälyä koskevista yhdenmukaistetuista säännöistä ja asetusten (EY) N:o 300/2008, (EU) N:o 167/2013, (EU) N:o 168/2013, (EU) 2018/858, (EU) 2018/1139 ja (EU) 2019/2144 sekä direktiivien 2014/90/EU, (EU) 2016/797 ja (EU) 2020/1828 muuttamisesta (tekoälysäädös). EUVL L 325, 12.7.2024, 1–56. Noudettu 18.10.2025 https://eur-lex.europa.eu/legal-content/FI/TXT/?uri=CELEX%3A32024R1689

European Comission. (2019, 8. huhtikuuta). *Ethics guidelines for trustworthy AI.* Noudettu 18.10.2025 osoitteesta https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Glikson, E. & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *The Academy of Management annals, 14(2)*, 627–660. https://doi.org/10.5465/annals.2018.0057

Günther, K. & Hasanen, K. (2021). *Johdanto: Tutkimuksen kulku.* Teoksessa Laadullisen tutkimuksen käsikirja. (toim.) Vuori, J. Tietoarkisto. Noudettu 25.11.2025 osoitteesta https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/laadullisen-tutkimuksen-prosessi/tutkimuksen-kulku/

Hautakangas, M. (2008). YOUTUBE - Uusi media, uusi julkisuus? *Lähikuva, 21(2)*, 8. https://doi.org/10.23994/lk.121006

Hermann, E. (2021). Artificial intelligence and mass personalization of communication content - An ethical and literacy perspective. *New Media & Society, 24(5),* 1258–1277. https://doi.org/10.1177/14614448211022702

Hoff, K. A. & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human factors, 57(3),* 407-434. https://doi.org/10.1177/0018720814547570

Jokinen, A., Juhila, K. & Suoninen, E. (2016). *Diskurssianalyysi: Teoriat, peruskäsitteet ja käyttö*. Vastapaino.

Junila, K. (2021). *Koodaaminen.* Teoksessa Laadullisen tutkimuksen käsikirja. (toim.) Vuori, J. Tietoarkisto. Noudettu 10.11.2025 osoitteesta https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/analyysitavan-valinta-ja-yleiset-analyysitavat/koodaaminen/

Kunz, W. H. & Wirtz, J. (2023). Corporate digital responsibility (CDR) in the age of AI: implications for interactive marketing. *Journal of Research in Interactive Marketing, 18(1)*, 31–37. https://doi.org/10.1108/jrim-06-2023-0176

Kuula-Luumi, A. (2021). *Laadullisen aineiston anonymisointi.* Teoksessa Laadullisen tutkimuksen käsikirja. (toim.) Vuori, J. Tietoarkisto. Noudettu 10.11.2025 osoitteesta https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/tutkimusetiikka/laadullisen-aineiston-anonymisointi/

Laapotti, T., Eronen-Valli, M. & Raappana, M. (2024). Viestinnällinen tutkimusmatka toimijuuteen ja tekoälyyn. *Media & viestintä 47(4),* 142–153. https://doi.org/10.23983/mv.152367

Lalot, F. & Bertram, A. (2025). When the Bot Walks the Talk: Investigating the Foundations of Trust in an Artificial Intelligence (AI) Chatbot. *Journal of experimental psychology. General*, *154(2),* 533–551. https://doi.org/10.1037/xge0001696

Lange, P. G. (2019) *Thanks for Watching: An Anthropological Study of Video Sharing on YouTube.* University Press of Colorado.

Lehti, L., Isosävi, J., Laippala, V., & Luotolahti, M. (2016). Linguistic analysis of online conflicts: A case study of flaming in the Smokahontas comment thread on YouTube. *WiderScreen, 19*(1–2). http://widerscreen.fi/numerot/2016-1-2/linguistic-anaead-on-youtube/

Ng, D. T. K., Leung, J. K. L., Chu, K. W. S. & Qiao, M. S. (2021). AI Literacy: definition, teaching, evaluation and ethical issues. *Proceedings of the Association for Information Science and Technology, 58(1),* 504–509. https://doi.org/10.1002/pra2.487

Niska, M., Venäläinen, S., Olakivi, A. & Cañada, J. A. (2024). *Sosiaalinen konstruktionismi: Miten tarkastella tulkintojen ja todellisuuden rakentumista*. Vastapaino, Tampere. https://doi.org/10.58181/VP9789523970458

Pietikäinen, S. & Mäntynen, A. (2019). *Uusi kurssi kohti diskurssia*. Vastapaino. Noudettu 1.10.2025 osoitteesta https://www.elibslibrary.com/reader/9789517688253

Rousseau, D. M., Sitkin, S. B., Burt, R. S. & Camerer, C. (1998). Not So Different After All: A Cross-Discipline View Of Trust. *The Academy of Management review, 23(3),* 393–404. https://doi.org/10.5465/amr.1998.926617

Sengar, S., Hasan, A. B., Kumar, S. & Carroll, F. (2024). *Generative Artificial Intelligence: A Systematic Review and Applications.* https://doi.org/10.48550/arxiv.2405.11029

Sihvonen, T. (2025). Arkista informaatiohäirintää vai luovaa ilmaisua? Tehostetut videosisällöt ja syvävääreennökset TikTokissa. *Media & viestintä, 48(1),* 59–86. https://doi.org/10.23983/mv.149489

Spears, R. & Postmes, T. (2015). Group identity, social influence, and collective action online: Extensions and applications of the SIDE model. Teoksessa: *The Handbook of the Psychology of Communication Technology*, toim. Sundar, S. S. 57–85. Wiley-Blackwell.

Stevanovic, M., Tiitinen, S. & Voutilainen, L. (2024). Keskusteluanalyysi ja sosiaalinen konstruktionismi. Teoksessa Niska, M., Venäläinen, S., Olakivi, A., & Cañada, J. A. (2024). *Sosiaalinen konstruktionismi: Miten tarkastella tulkintojen ja todellisuuden rakentumista*. Vastapaino, Tampere. https://doi.org/10.58181/VP9789523970458

Striewski, S., Zagovora, O., & Peters, I. (2024). Scientific discourse on YouTube: Motivations for citing research in comments. *arXiv*. https://doi.org/10.48550/arXiv.2405.12798

Vaahensalo, E. (2021). Kontekstualisointimalli sosiaalisen median lähdekritiikin avaimena. *Informaatiotutkimus, 40(3),* 110–141. https://doi.org/10.23978/inf.107897

Wattenhofer, M., Wattenhofer, R. & Zhu, Z. (2012). The YouTube Social network. *Proceedings of the International AAAI Conference on Weblogs and Social Media, 6(1).* https://doi.org/10.1609/icwsm.v6i1.14243

Wolf, M. J., Grodzinsky, F. & Miller, K. W. (2024). Generative AI and Its Implications for Definitions of Trust. *Information (Basel*), *15(9), 542,* 1-10. https://doi.org/10.3390/info15090542

Young, A. T., Amara, D., Bhattacharya, A. & Wei, M. L. (2021). Patient and general public attitudes towards clinical artificial intelligence: a mixed methods systematic review. *The Lancet Digital Health, 3*(9), e599–e611. https://doi.org/10.1016/S2589-7500(21)00132-1

ZoneVD. (2025, 26. elokuuta). *Tekoäly pilaa Youtuben* [video]. YouTube. Noudettu 7.10.2025 osoitteesta https://www.youtube.com/watch?v=fDQlxtIizdY