



PLEASE NOTE! THIS IS PARALLEL PUBLISHED VERSION /
SELF-ARCHIVED VERSION OF THE OF THE ORIGINAL ARTICLE

This is an electronic reprint of the original article.
This version *may* differ from the original in pagination and typographic detail.

Author(s): Chichaeva, Julija

Title: Chatbotit terveydenhuollossa – eettisiä näkökulmia

Year: 2024

Version: Final draft

Please cite the original version:

Chichaeva, J. (2024). Chatbotit terveydenhuollossa – eettisiä näkökulmia. Nordic journal of rehabilitation. Artikkeleli 11.3.2024. <http://urn.fi/urn:nbn:fi:jamk-issn-2954-1069-10>

Chatbotit terveydenhuollossa – eettisiä näkökulmia

Julija Chichaeva ⁽¹⁾

⁽¹⁾Kuntoutusinstituutti, Jyväskylän ammattikorkeakoulu

Avainsanat: chatbotit, tekoäly, kuntoutus, etiikka (YSO)

Johdanto

Keskustelevan tekoälyn sovellukset, kuten esimerkiksi chatbotit näyttäisivät vakiinnuttaneen paikkansa asiakaspalvelijoiden avustajina toimialasta riippumatta. Ne on tunnistettu potentiaalisina apukäsinä myös terveysalalla, esimerkiksi asiakkaan ohjauksen ja neuvonnan tukena (esim. Singh 2023). Kirjallisuudessa keskusteleva tekoäly on määritelty mm. keskustelevaksi agentiksi, keskustelurobotiksi tai chatbotiksi, joka mahdollistaa interaktiivisen dialogin käyttäjän kanssa. Keskustelevan tekoälyn sovelluksia voidaan käyttää erilaisilla teknologisilla laitteilla kuten älypuhelimella, tabletilla tai tietokoneella ja ne on luotu simuloimaan keskustelua ihmisen kanssa. Ne voivat viestiä monipuolisesti käyttäen tekstiä, ääntä, kuvia tai videoita. (Tudor Car ym. 2020.) Keskustelevan tekoälyn sovelluksia on käytetty mm. keräämään asiakkaan terveyteen liittyviä tietoja sekä antamaan terveyteen ja hoitoon liittyviä ohjeita ja neuvontaa (Luxton 2020; Singh 2023; Tudor Car ym. 2020).

Eettisten näkökulmien tarkastelu kuuluu oleellisena osana sekä terveydenhuollon työhön että ihmiskeskeiseen teknologiakehitykseen. Vaikka on viitteitä siitä, että keskustelevan tekoälyn sovellukset edustavat potentiaalista ratkaisua vastaamaan terveyspalveluiden kasvavaan tarpeeseen (esim. Singh 2023 ja Tudor Car ym. 2020), saattaa se uhata terveydenhuollon asiakkaiden mieltymyksiä, turvallisuutta ja yksityisyyttä (Rigby 2019). Eettiset perusperiaatteet antavat ammattilaiselle suuntaviivoja toimia oikeudenmukaisesti, luottamuksellisesti, läpinäkyvästi, ammatilliset standardit huomioon ottaen sekä tukien ihmisen itsemääräämisoikeutta ja hyvinvointia.

Tämän artikkelin tarkoituksena on tarkastella terveysalalla toimivien chatbottien kehitysprosessiin ja käyttöönottoon liittyviä eettisiä käytänteitä ja haasteita aiemman kirjallisuuden sekä AIRE Tulevaisuuden älykkäät kuntoutuspalvelut -hankkeessa (2021–2023) kerättyjen, chatbotin pilotointiin liittyvien kuntoutujien kokemusten (Chichaeva 2023) pohjalta. Tämän hankkeentarkoituksena oli lisätä ymmärrystä keskustelevan tekoälyn hyödyistä ja haasteista osana kuntoutuspalveluja, sekä pilotoida ja kehittää tälle toimialalle soveltuvia keskustelevan tekoälyn ratkaisuja. AIRE-hankkeessa keskustelevalla tekoälyllä tarkoitetaan chatbottia, joka ymmärtää käyttäjän luonnollista kieltä, tarjoaa valmiita vastausvaihtoehtoja painikkeina ja vastaa tekstiä, kuvia ja linkkejä hyödyntäen.

Eettisiä näkökulmia chatbottien käytössä

Neljä yleisimmin terveydenhuollossa sovellettavaa eettistä periaatetta ovat pahan välttäminen (nonmaleficence), hyvän tekeminen (beneficence), ihmisyyden ja ihmisen autonomian kunnioittaminen (autonomy) sekä oikeudenmukaisuus (fairness, justice, equality) (Summers 2009). Teknologian ja tekoälyn nopean kehityksen ja käyttöönoton myötä eettiset näkökulmat ovat

nousseet globaalin kiinnostuksen kohteeksi myös tällä toimialalla. Kirjallisuuden mukaan tekoälyn vastuulliseen kehittämiseen liittyvissä ohjeistuksissa näyttäytyy edellisten lisäksi seitsemän eettistä periaatetta: läpinäkyvyys (transparency), vastuullisuus (accountability, responsibility), yksityisyys (privacy), luottamus (trust), kestävyys (sustainability), ihmisarvo (dignity) ja solidaarisuus (solidarity) (Jobin ym. 2019; Ryan & Stahl 2021).

Pahan välttämisen periaatteen mukaan terveydenhuollon ammattilaisen tulee välttää tuottamasta vahinkoa ihmiselle, kun taas **hyvän tekeminen** on laajempi käsite ja kuvaa terveydenhuollon ammattilaisen velvollisuutta auttaa muita. Lisäksi hyvän tekemisen periaatteen mukainen kohtaaminen on myötätuntoista, ymmärtäväistä ja yksilöllistä. (Summers 2009.) Koska keskusteleva tekoäly toimii itsenäisesti, eikä välttämättä pysty aina tulkitsemaan käyttäjän kieltä oikein, vastaaminen ihmisen parhaan edun mukaisesti kullakin hetkellä voi olla epävarmaa Coghlan ym. 2023; Palanica ym. 2019). Terveydenhuollon ammattilaisten käsitysten mukaan chatbotit eivät esimerkiksi pysty huomioimaan kaikkien potilaiden yksilöllisiä tarpeita ja ominaisuuksia, toimimaan hätätilanteissa, tekemään luotettavia diagnooseja tai näyttämään inhimillisiä tunteita (Palanica ym. 2019). Miner ym. (2016) poikittaistutkimuksessa havaittiin, että keskusteleva tekoälysovellus ei välttämättä osaa reagoida tarkoituksenmukaisesti ihmisen kertoessa mielenterveyden haasteista, itsetuhoisista ajatuksista tai lähisuhdeväkivallasta. Myös ihmiselle, joilla on kognitiivisia toimintarajoitteita, tulisi harkita tarkoin keskustelevan tekoälyn sopivuutta (Luxton 2014). AIRE-hankkeessa kyseinen eettinen haaste pyrittiin ratkaisemaan seulomalla ne tilanteet, jossa chatbot oli liian vastuullisen päätöksenteon äärellä. Näissä tapauksissa keskusteleva chatbot ohjasi akuuttia ja/tai henkilökohtaista kontaktia tarvitsevan henkilön ihmisasiakaspalvelijan pariin.

Tulevaisuuteen liittyvinä uhkakuvina kuvataan kielteiset vaikutukset kliiniseen työhön pitkällä aikavälillä, johtuen esimerkiksi chatbottien tekemistä nopeista ja mahdollisesti virheellisistä diagnooseista, potilaiden itsehoidon vastuun lisääntymisestä, terveydenhuollon luotettavuuden rappeutumisesta sekä hoitoon pääsyn viivästyttämisestä (Parviainen & Rantala 2022).

Oikeudenmukaisuuden periaate velvoittaa toimimaan samalla tavalla jokaista yksilöä kohtaan tukien tasa-arvoa ja yhdenvertaisuutta (Summers 2009). Keskustelevan tekoälyn sovelluksien etuna pidetään niiden muovailtavuutta käyttäjän tarpeisiin, esim. ulkonäön, puhutavan ja sosioekonomisen aseman suhteen, sillä nämä ominaisuudet saattavat lisätä hoitoon sitoutumista (Luxton 2014). Tekoäly saattaa myös tuntua käyttäjistä ennakkoluulottomammalta ihmisasiakaspalvelijaan verrattuna, jolloin arkaluontoisista aiheista voi olla helpompi puhua (Laranjo ym. 2018; Luxton 2014). Toisaalta keskustelevan tekoälyn hyödyntämisen eettisinä haasteina ovat tekoälyn ulkonäköön ja käyttäytymiseen liittyvät suunnitteluvirheet sekä ohjelmiston virheet algoritmeissa, jotka saattavat johtaa systemaattisiin virheisiin tekoälyn koulutuksessa ja uuden tiedon tuottamisessa, joiden vuoksi tekoäly saattaa toimia epätasa-arvoisesti eri käyttäjiä kohtaan. Näiden taustalla saattaa olla puutteellinen tai virheellinen data (puuttuva tieto, virheellinen luokittelu, mittausvirheet, pieni otoskoko) tai ohjelmoijien/organisaatioiden arvot, joihin nojaten algoritmien dataa kerätään, valikoidaan tai käytetään. (Gianfrancesco ym. 2018.)

Vaikka keskustelevan tekoälyn sovellusten ajatellaan lisäävän tasa-arvoista pääsyä terveyspalveluiden pariin (esim. Luxton 2014), voivat digitaaliset palvelut myös lisätä eriarvoisuutta. Teknologioiden rajoitettu saatavuus, puutteellinen koulutus ja teknologiaosaaminen voivat rajoittaa keskustelevan tekoälyn sovelluksien käyttöä erilaisissa väestöryhmissä (Luxton 2020). Huomio tulisi suunnata myös haavoittuvaisessa asemassa oleviin ihmisiin, kuten mielenterveyskuntoutujiin. He saattavat joutua epätasa-arvoiseen asemaan, jääden esimerkiksi palveluiden ulkopuolelle

käyttäessään uutta ja heidän kohderyhmällensä edelleen kehitysvaiheessa olevaa teknologiaa. (Coghlan ym. 2023.)

AIRE-hankkeessa kerättyjen kuntoutujien kokemuksissa chatbotin käytön esteinä nähtiin omat asenteet teknologiaa kohtaan, aiemmat kielteiset kokemukset sekä rajalliset taidot käyttää ja oppia käyttämään nopeasti kehittyvää teknologiaa (Chichaeva 2023). Näin teknologian avulla yhä paremmin saatavilla olevat, ajasta ja paikasta riippumattomat terveydenhuollon palvelut saattavat jäädä tietyltä väestöryhmältä kokonaan saavuttamatta.

Luottamus voidaan tunnistaa keskeiseksi vaatimukseksi tekoälyn eettiselle käyttöönotolle ja käytölle (Jobin ym. 2019). Loppukäyttäjän näkökulmasta luotettavan chatbotin taustalla on luotettava organisaatio, joka toimii tarkoituksenmukaisesti ja asiantuntevasti sekä vastaa käyttäjälle oikein esitettyyn kysymykseen. Luottamus voi olla myös riippuvaista loppukäyttäjän taipumuksesta luottaa teknologiaan yleisesti. (Chichaeva 2023; Nordheim 2019.) Vaikka chatbottien inhimillisten piirteiden puuttuminen nähdään niiden heikkoutena (Palanica ym. 2019), käyttäjä saattaa kokea robotin luotettavammaksi keskustelukumppaniksi esimerkiksi intiimeissä terveysongelmissa (Parviainen & Rantala 2022). Keskustelevan tekoälyn sovelluksen on väitetty olevan ihmistä luotettavampi myös sen puuttuvien inhimillisten ominaisuuksien vuoksi, sillä chatbot ei kärsi väsymyksestä, loppuun palamisesta tai kognitiivisista haasteista (Luxton 2014). Luotettavan chatbotin kehitystyö saattaa kuitenkin lisätä työmäärää ja kustannuksia, sillä toimiva ja luotettava chatbot vaatii paljon monialaista asiantuntijuutta ja rahoitusta (Parviainen & Rantala 2022) sekä AIRE-hankkeessa syntyneen kokemuksen mukaan aina myös ihmisen työparikseen.

Autonomian periaate velvoittaa terveysalan ammattilaisen huomioimaan ihmisen oman tahdon ja osallistamaan tämän omaa terveyttä koskevaan päätöksentekoon (Summers 2009). Loppukäyttäjän toimijuus tulisi olla tekoälyn kehitystyön keskiössä, säilyttäen loppukäyttäjän itsenäisen ja riippumattoman päätöksenteon kaikissa tilanteissa (Tekoälyä käsittelevä korkean tason asiantuntijaryhmä 2018). Chatbottien etu on kyky toimia nopeasti ja ajasta riippumatta. Ne ovat päteviä antamaan tietoa käyttäjän sairaudesta oirekuvan perusteella ja yhdistämään käyttäjän oikeiden palvelujen pariin. (Parviainen & Rantala 2022.) Nämä ominaisuudet saattavat parantaa ihmisen mahdollisuutta toimia oman terveytensä ja omien tavoitteidensa hyväksi sekä vähentää riippuvuutta terveydenhuollon yhä niukkenevista resursseista. Terveydenhuollon palveluiden saatavuus ihmisasiakaspalvelijan vastaanotolla on kuitenkin edelleen merkityksellistä digitaalisten palvelujen rinnalla, jotta myös ne, jotka eivät esimerkiksi pysty toimintakyvyn rajoituksen vuoksi käyttämään chatbottia, voisivat saada välineitä sairautensa hoitoon ja toimijuutensa tukemiseen.

Yksilön **yksityisyys** ja henkilötietojen käsittely on ensiarvoisen tärkeää tekoälyn suunnittelussa ja käytössä (Tekoälyä käsittelevä korkean tason asiantuntijaryhmä 2018). Ylläpitäjän tulisi olla selvillä yksityisyyteen liittyvästä lainsäädännöstä ja viestiä loppukäyttäjälle läpinäkyvästi ja saavutettavasti, miten tietosuoja toteutuu sekä sen mahdollisista rajoituksista. Käyttäjien yksityisyys nähdään keskustelevan tekoälyn sovellusten haasteena, sillä chatbot saattaa kerätä suuria määriä yksityistä ja arkaluontoista dataa. (Luxton 2020.)

Läpinäkyvyyden peruseriaate läpileikkaa monia edellä mainittuja eettisiä vaatimuksia, kuten ihmisoikeus, yksityisyys, ihmisarvo ja autonomia. Loppukäyttäjälle tulee selvittää ymmärrettävällä kielellä kaikki sellaiset tiedot, joilla voidaan varmistua, että tekoäly ei manipuloi, petä tai pakota tahdon vastaiseen toimintaan (Tekoälyä käsittelevä korkean tason asiantuntijaryhmä 2018). Käyttäjän tulisi siis ymmärtää kenen kanssa on tekemisissä eli mikä on taustalla toimiva

organisaatio ja, että kyseessä on tekoäly, ei ihminen, mikä on tekoälyn tarkoitus ja mihin sen toiminta perustuu. Siitä huolimatta, että AIRE-hankkeen chatboteissa pyrittiin läpinäkyvyyteen, loppukäyttäjien kokemusten perusteella niiden tarkoitus, toimintalogiikka sekä rajoitukset saattoivat jäädä käyttäjälle epäselviksi (Chichaeva 2023). Terveydenhuollossa hyödynnettävien chatbottien kehittäjien haasteeksi jääkin viestiä erilaisille loppukäyttäjille saavutettavasti niin, että läpinäkyvyyden periaate toteutuu. Erityisesti huomioon tulisi ottaa ne ihmiset, joilla on toimintakyvyn rajoituksia tai kognitiivisia haasteita.

Suuntaviivoja tulevaisuuteen

Vaikka tekoälyn eettiselle käyttönotolle terveydenhuollossa on yritetty hahmotella peruseriaatteita (Coghen ym. 2023; Beil ym. 2019; Rigby 2019), vakiintuneita eettisiä ohjeita ei edelleenkään ole olemassa (Parviainen & Rantala 2022). Euroopan komission korkean tason asiantuntijaryhmän (2018) linjaamien eettisten ohjeiden mukaan luotettavan tekoälyn tulisi noudattaa kaikkia sovellettavia lakeja ja määräyksiä, toimia eettisten arvojen ja periaatteiden mukaisesti sekä toimia teknisesti ja sosiaalisesti luotettavasti. Suomessa terveydenhuollossa toimivia chatbotteja ei kuitenkaan välttämättä määritellä lääkinnällisinä laitteina, jolloin niiden lainmukaisuus ja turvallisuus eivät ole lääkealan turvallisuus- ja kehittämiskeskuksen Fimean valvonnan alaisia. Lääkinnällisen laitteen arviointi tehdään ohjelmistokehittäjän toimesta, useimmiten tapauskohtaisesti, ohjelmiston käyttötarkoituksen ohjailemana (Fimea 2023). Tarvitaan lisää tutkimusta, joka ei selvitä pelkästään keskusteleavan tekoälyn potentiaalisia hyötyjä vaan myös näiden sovellusten läpinäkyvyyttä, asianmukaista valvontaa ja turvallisuutta (McGreevey ym. 2020). Määrittelemistä vaatii myös se, millaiset terveystalvet voidaan automatisoida ja toteuttaa chatbotin avustuksella ja millaisissa tapauksissa terveydenhuollon ammattilainen ei ole korvattavissa tekoälyllä.

Yhteyshenkilö:

Julija Chichaeva, tutkija
Kuntoutusinstituutti, Jyväskylän ammattikorkeakoulu
PL 207, 40101 Jyväskylä
etunimi.sukunimi(a)jamk.fi

Lähteet

Beil, M., Proft, I., van Heerden, D., Sviri, S. and van Heerden P. V. (2019). Ethical considerations about artificial intelligence for prognostication in intensive care. *Intensive Care Medicine Experimental* 7 (1), 70. <https://doi.org/10.1186/s40635-019-0286-6>.

Chichaeva, J. 2023. Kuntoutujien kokemuksia Chatbotin käytöstä. Jyväskylän ammattikorkeakoulu. ISBN: 978-951-830-716-0.

Coghlan, S., Leins, K., Sheldrick, S., Cheong, M., Gooding, P. & D'Alfonso, S. (2023). To chat or bot to chat: Ethical issues with using chatbots in mental health. *DIGITAL HEALTH* 9. <https://doi.org/10.1177/20552076231183542>.

Fimea. (2023). Ohjelmisto lääkinnällisenä laitteena. Ohjelmistot. Fimea. Viitattu 9.8.2023. https://www.fimea.fi/laakinnalliset_laitteet/mita-ovat-laakinnalliset-laitteet-erikoislaiteryhmat/ohjelmistot.

Tekoälyä käsittelevä korkean tason asiantuntijaryhmä. (2018). Luotettavaa tekoälyä koskevat eettiset ohjeet. Euroopan komissio. Viitattu 1.8.2023. https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/JURI/DV/2019/11-06/Ethics-guidelines-AI_FI.pdf.

Gianfrancesco, M. A., Tamang, S., Yazdany, J. & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med.* 178 (11), 1544–7. <https://doi.org/10.1001/jamainternmed.2018.3763>.

Jobin, A., Ienca, M. & Vayena, E. (2019). “The global landscape of AI ethics guidelines”, *Nature Machine Intelligence* 1 (9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>.

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A. & Coiera, E. (2018). Conversational agents in healthcare: a systematic review. *J Am Med Inform Association* 25 (9), 1248–1258. <https://doi.org/10.1093/jamia/ocy072>.

Luxton, D. D. (2020). Ethical implications of conversational agents in global public health. *Bull World Health* 98 (4), 285–287. <https://doi.org/10.2471/BLT.19.237636>.

Luxton, D. D. (2014). Recommendations for the ethical use and design of artificial intelligent care providers. *Artificial Intelligence* 62 (1), 1–10. <https://doi.org/10.1016/j.artmed.2014.06.004>.

McGreevey, J. D., Hanson, C. & Koppel R. (2020). Clinical, Legal, and Ethical Aspects of Artificial Intelligence–Assisted Conversational Agents in Health Care. *JAMA* 324 (6), 552–553. <https://doi.org/10.1001/jama.2020.2724>.

Miner, A. S., Milstein, A., Schueller, S., Hegde, R., Mangurian, C. & Linos, E. (2016). Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA Intern Med.* 176 (5), 619–625. <https://doi.org/10.1001/jamainternmed.2016.0400>.

Nordheim, C.B., Følstad, A. & Bjørkli C. A. (2019). An initial model of trust in chatbots for customer service—Findings from a questionnaire study. *Interacting with Computers* 31 (3), 317–335. <https://doi.org/10.1093/iwc/iwz022>.

Parviainen, J. & Rantala, J. (2022). Chatbot breakthrough in the 2020s? An ethical reflection on the trend of automated consultations in health care. *Med. Health Care and Philos.* 25, 61–71. <https://doi.org/10.1007/s11019-021-10049-w>.

Palanica, A., Flaschner, P., Thommandram, A., Li, M. & Fossat, Y. (2019). Physicians’ perceptions of chatbots in health care: cross-sectional web-based survey. *Journal of Medical Internet Research* 21 (4), e12887. <https://doi.org/10.2196/12887>.

Rigby, M. J. (2019). Ethical dimensions of using artificial intelligence in health care. *AMA Journal of Ethics* 21 (2), e121-124. <https://doi.org/10.1001/amajethics.2019.121>.

Ryan, M. & Stahl, B. (2020). Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications. *Journal of Information, Communication and Ethics in Society* 19 (1), 61-86. ISSN: 1477-996X.

Singh, B., Olds, T., Brinsley, J., Dumuid, D., Virgara, R., Matricciani, L., Watson, A., Szeto, K., Eglitis, E., Miatke, A., Simpson, C. E. M., Vandelanotte, C. & Maher, C. (2023). Systematic review and meta-analysis of the effectiveness of chatbots on lifestyle behaviours. *npj Digit. Med.* 6 (118). <https://doi.org/10.1038/s41746-023-00856-1>.

Summers, J. (2009). Principles of healthcare ethics. In E. Morrison (ed.) *Health Care Ethics: Critical Issues for the 21st Century*. 2nd edition. Sudbury: Jones and Bartlett Publishers.

Tudor Car, L., Dhinakaran, A. D., Kyaw, M. B., Kowatsch, T., Joty, S., Theng, Y-L. & Atun, R. (2020). Conversational Agents in Health Care: Scoping Review and Conceptual Analysis. *Journal of Medical Internet Research* 22 (8), e17158. <https://doi.org/10.2196/17158>.

Julkaisun pysyvä osoite:

URN: <http://urn.fi/urn:nbn:fi:jamk-issn-2954-1069-10>