



**Mental
Health
Europe**

Equal rights.
Better mental health.
For all.

joulukuu 2024

Tekoäly mielenterveystyössä



Tämän Mental Health Europe -tutkimuksen ovat kirjoittaneet Piers Gooding*, Hannah van Kolschooten** ja Francesca Centola***

*Associate Professor, La Trobe Law School, La Trobe University.

** Luennoitsija-tutkija, Amsterdamin yliopiston oikeustieteellinen tiedekunta, Amsterdamin yliopisto, ja AI-konsultti, Health Action International.

*** Policy and Knowledge Officer, Mental Health Europe.

Jos sinulla on kysyttävää, ota yhteyttä osoitteeseen: f.centola@mentalhealtheurope.org.

Tämä asiakirja on käännetty englannista konekäännösohjelmistolla, joten käännöksen tarkkuutta ei voida taata.

Sisältö

Sanasto.....	3
Yhteenveto	3
Tietoa tästä tutkimuksesta.....	4
Tausta	4
Katsaus tekoälysovelluksiin mielenterveystyössä	5
Mahdollisuudet.....	8
Riskit.....	9
Turvallisuusriskit.....	9
Yksityisyys.....	10
Tietoon perustuvan suostumuksen puuttuminen.....	10
Uudet tai vahvistuneet eriarvoisuudet	11
Hoidon muuttuminen persoonattomaksi.....	11
Valvonta	12
Vahvistetaan yksilökeskeisiä näkemyksiä mielenterveydestä	12
Rajallisten resurssien ohjaaminen muualle	13
Päätelmät ja suositukset	13
Mikä rooli eletyllä kokemuksella on tekoälyteknologian kehittämisessä?.....	14
Suositukset poliittisille päättäjille.....	14
Suositukset kansalaisyhteiskunnalle	16
Kirjallisuusluettelo.....	16

Sanasto

"**Tekoälyjärjestelmä**" määritellään Euroopan unionin (EU) tekoälylain 3 artiklan 1 kohdassa seuraavasti:

"konepohjainen järjestelmä, joka on suunniteltu toimimaan vaihtelevalla autonomian tasolla ja joka voi olla mukautuva käyttöönoton jälkeen ja joka eksplisiittisiä tai implisiittisiä tavoitteita varten päättelee saamastaan syötteestä, miten tuottaa tuotoksia, kuten ennusteita, sisältöä, suosituksia tai päätöksiä, jotka voivat vaikuttaa fyysisiin tai virtuaalisiin ympäristöihin."

Tätä määritelmää käytetään tässä kertomuksessa.

Yhteiskehittäminen: Yhteistoiminnallinen lähestymistapa, jossa kaikki mielenterveysalan toimijat työskentelevät tasavertaisesti yhdessä kehittääkseen ja toteuttaakseen politiikkoja, palveluja, ohjelmia, tutkimusta ja viestintää, jotka edistävät myönteistä mielenterveyttä psykososiaalisen mallin ja ihmisoikeuksiin perustuvan lähestymistavan mukaisesti. ([Mental Health Europe](#))

Digitaalisella fenotyyppityksellä tarkoitetaan digitaalisten laitteiden (kuten älypuhelinien ja puettavien laitteiden) tietojen käyttöä tiettyihin mielenterveyden tiloihin liittyvien käyttäytymismallien tunnistamiseksi.

Generatiivisella tarkoitetaan tekoälyn alaryhmää, joka keskittyy uuden sisällön luomiseen tekstistä ja kuvista ääniin, videoihin, 3D-malleihin ja synteettiseen dataan. Tällä alalla käytetään ensisijaisesti [koneoppimisalgoritmeja](#), erityisesti syväoppimismalleja, havaitsemaan malleja harjoitusdatassa ja tuottamaan uusia tuotoksia.

Elämäkokemusta omaavat ihmiset: mielenterveysongelmia kokeneet tai kokeneet ihmiset. ([Mental Health Europe](#))

Yhteenveto

Tässä tutkimuksessa tarkastellaan mahdollisuuksia, riskejä ja eettisiä näkökohtia, jotka liittyvät tekoälyjärjestelmien käyttöön mielenterveyspalveluissa, ja annetaan suosituksia niiden vastuullisesta käyttöönotosta ja sääntelystä.

Terveydenhuolto on yksi suosituimmista aloista, joilla tekoälyä otetaan käyttöön EU:ssa.¹ Mielenterveysalalla tekoälyjärjestelmiä käytetään monin tavoin hallinnollisista tehtävistä viestintäalustoihin, ammatilliseen päätöksenteon tukemiseen, digitaalisiin terapioihin, kuten chatbotteihin ja henkilökohtaisiin aistinteknologioihin, sekä potilaan seurantaan. Terveydenhuollon lisäksi mielenterveyteen liittyviä tietoja hyödyntävät tekoälysovellukset ulottuvat myös rikosoikeuteen, kuluttajatuotteisiin, koulutukseen ja työllisyyteen.

Tekoälyjärjestelmät tarjoavat merkittäviä potentiaalisia hyötyjä, kuten mielenterveystuen saatavuuden parantaminen erityisesti heikommassa asemassa oleville väestöryhmille ja hallinnollisen taakan vähentäminen terveydenhuoltojärjestelmissä. Kliinisesti tekoälyä kannattavat tahot edistävät hoitojen yksilöllistämistä, diagnoosien tarkkuuden parantamista ja oikea-aikaisten toimenpiteiden tukemista.

Lisäksi tekoäly voi nopeuttaa mielenterveystutkimusta ja helpottaa uusien terapeuttisten lähestymistapojen, kuten virtuaalitodellisuushoitojen, kehittämistä. Sosiaalisesti suuntautuneita mahdollisuuksia ovat vertaistuen ja ammattilaisten koulutuksen parantaminen, verkostoituneen yhteistyön edistäminen ja laajempien yhteisöyhteyksien luominen.

Mielenterveysalan tekoälyjärjestelmät aiheuttavat kuitenkin myös vakavia riskejä, jotka voidaan yksilöidä kolmella tasolla: yksilöllisellä, kollektiivisella ja yhteiskunnallisella². Yksilötasolla huolenaiheita ovat muun muassa turvallisuusriskit, yksityisyyden loukkaukset ja puutteellinen tietoinen suostumus. Laajempiin haasteisiin kuuluvat epätasa-arvon vahvistuminen tai uusien epätasa-arvojen syntyminen, liiallinen valvonta, yksilökeskeisten mielenterveyskäsitysten vahvistuminen, hoidon muuttuminen persoonattomaksi ja rajallisten resurssien ohjautuminen muualle.

Raportissa tuodaan esiin keskeisiä periaatteita ja mahdollisia toimintatapoja, joiden avulla poliittiset päättäjät voivat puuttua näihin riskeihin tasapainottamalla innovointia, eettistä vastuuta ja ihmisläheistä hoitoa tekoälyyn perustuvissa mielenterveysjärjestelmissä. Raportissa todetaan, että tekoälytyökaluja on kehitettävä eettisyys, osallisuus, tarkkuus, turvallisuus ja loppukäyttäjien todelliset tarpeet huomioon ottaen. Mahdollisia ratkaisuja ovat pääasiassa vankka sääntely ja valvonta, avoimuus ja selitettävyyys sekä ihmisoikeuskeskeiset ja yhteiskehittämiseen perustuvat lähestymistavat.

Eniten kärsivien yhteisöjen - ihmisten, joilla on omakohtaisia kokemuksia - aktiivista osallistumista ja johtajuutta olisi pidettävä eettisenä ja poliittisena perusperiaatteena, joka ohjaa kaikkia digitalisaatiota ja mielenterveyttä koskevia keskusteluja ja päätöksiä. Sen varmistaminen, että heillä on yhtäläinen ääni poliitikkojen ja innovaatioiden muotoilussa, on olennaisen tärkeää oikeudenmukaisten ja tasapainoisten tulosten saavuttamiseksi ja terveyserojen syvenemisen estämiseksi yhteiskunnassa.

Tietoa tästä tutkimuksesta

Tutkimus on suunnattu poliittisille päättäjille, palveluntarjoajille ja muille sidosryhmille, jotka ovat kiinnostuneita tekoälyn sovelluksista ja vaikutuksista mielenterveyspalveluissa.

Mental Health Europen visiota esittelevän taustatiedon jälkeen tutkimuksessa luodaan yleiskatsaus tekoälyn sovelluksiin mielenterveytyössä. Sen jälkeen siinä arvioidaan näihin teknologioihin liittyviä mahdollisuuksia ja riskejä ja annetaan poliittisille päättäjille suosituksia siitä, miten riskeihin voidaan puuttua. Tutkimuksessa korostetaan eletyn kokemuksen merkitystä tekoälyn kehittämisessä ja annetaan erityisiä suosituksia kansalaisyhteiskunnalle.

Tausta

Vuonna 2022 Mental Health Europe julkaisi raportin [Mental health in the digital age: Ihmisoikeuksiin perustuvan psykososiaalisen lähestymistavan soveltaminen kompassina](#). Raportissa Mental Health Europe esitti näkemyksensä mielenterveydestä digitaalisessa maailmassa: digitalisaatiota ei pitäisi pitää itsetarkoituksena vaan keinona kohti suurempaa päämäärää. Se on keino kohti henkisesti terveempää yhteiskuntaa, suurempaa itsemääräämisoikeutta ja vaikutusvaltaa omaan mielenterveyteemme sekä paremmin sopivaa mielenterveyspalvelua.

Korostimme, että on tärkeää, että uutta teknologiaa ei sovelleta vain siksi, että se on kannattavaa tai toteutettavissa, vaan siksi, että se vastaa todelliseen tarpeeseen. Tämä on erityisen tärkeää, kun otetaan huomioon kaupalliset intressit. Kannatimme yhteiskehittämisen lähestymistapaa ja

ihmisoikeuskehystä digitaalisen mielenterveyden kehityksen suunnannäyttäjänä. Ainoastaan tällaisella yhteistyöhön perustuvalla lähestymistavalla - prosessin kaikissa vaiheissa suunnittelusta arviointiin - digitaaliteknologia voi vastata todellisiin tarpeisiin ja edistää visiota yhteiskunnasta, jossa kaikki voivat nauttia ihmisoikeuksistaan ja menestystä.

Vuodesta 2022 lähtien EU:n tasolla on tapahtunut paljon kehitystä, kuten [tekoälylaki](#), joka on ensimmäinen kattava tekoälyä koskeva sääntely, jonka on antanut merkittävä sääntelyviranomaisena. Elokuussa 2024 voimaan tulevalla asetuksella pyritään varmistamaan, että tekoälyjärjestelmät ovat turvallisia ja kunnioittavat perusoikeuksia ja -arvoja, edistämään luottamusta tekoälyteknologioihin, tukemaan innovointia ja parantamaan EU:n kilpailukykyä tekoälyn alalla. Tekoälylakia sovelletaan horisontaalisesti kaikkiin aloihin, joilla tekoälyä käytetään, myös terveydenhuoltoon

Tämän tutkimuksen tavoitteena on keskittyä tekoälysovellusten vaikutuksiin mielenterveyspalveluissa, korostaa alalle ominaisia riskejä ja ehdottaa toimenpiteitä niiden torjumiseksi. Näin se luo pohjan kriittiselle arvioinnille siitä, soveltuuko tekoälylaki tarkoitukseensa mielenterveystyön osalta ja miten poliittiset päättäjät voivat puuttua mahdollisiin puutteisiin. Tutkimuksessa korostetaan, että on tärkeää tehdä yhteistyötä poliittisten päätöksentekijöiden, elävän kokemuksen omaavien ihmisten, kansalaisyhteiskunnan järjestöjen ja muiden asiaankuuluvien sidosryhmien välillä, jotta voidaan varmistaa, että näihin puutteisiin puututaan tehokkaasti asetuksen täytäntöönpanovaiheessa.

Katsaus tekoälysovelluksiin mielenterveyspalveluissa terveydenhuolto

Mielenterveystyössä tekoälyjärjestelmiä käytetään monin eri tavoin. Alla olevissa taulukoissa on esimerkkejä tekoälysovelluksista, ja niissä on eritelty, onko niiden kohderyhmänä palvelun käyttäjät vai mielenterveysalan ammattilaiset.

Tekoälysovellukset palvelujen

Käyttötarkoitu		Esimerkkejä
Digitaaliset terapiat	Digitaalinen fenotyyppitys: Tekoäly analysoi digitaalisista laitteista (esim. älypuhelimista, puettavista laitteista) saatuja tietoja mielenterveyden tiloihin (esim. masennus, ahdistus) liittyvien käyttäytymismallien tunnistamiseksi.	MindLAMP2 (jossa LAMP on lyhenne sanoista Learn, Assess, Manage, Prevent) on avoimen lähdekoodin ja vapaasti saatavilla oleva sovellus, joka tuottaa useita räätälöitävissä olevia tietovirtoja käyttäjältä (esim. laitteen liikkeitä, sijaintitoiminnot) ja tarjoaa aktiviteetteja, jotka rohkaisevat tarkkaavaisuuteen ja itsetutkiskeluun. MindLAMP2:ta on käytetty useissa mielenterveyden kliinisissä ja tutkimuksellisissa käyttötapauksissa.
	Chatbotit: Vuorovaikutteiset tekoälytyökalut, jotka toimivat virtuaalisina neuvonantajina, jäljittelevät terapeuttisia kohtaamisia ja antavat neuvoja.	Woebot ja Wysa ovat tekoälychatbotteja, jotka on suunniteltu jäljittelemään terapeuttista vuorovaikutusta ja tarjoamaan mielenterveysneuvoja ja

	Ne voivat myös ohjata sosiaalipalvelujen tai tukijärjestelmien piiriin.	selviytymisstrategioita.
Henkilökohtainen lääketiede	Tekoäly hyödyntää geneettisiä, ympäristöön liittyviä ja käyttäytymiseen liittyviä tietoja räätälöityjen hoitosuunnitelmien luomiseksi yksittäisille palvelunkäyttäjille.	IBM Watson Health käsittelee potilastietoja suositellakseen masennuksen yksilöllisiä hoitovaihtoehtoja, jotka perustuvat geneettisten ja käyttäytymiseen liittyvien tekijöiden yhdistelmään.
Potilaan seuranta ja valvonta	<ul style="list-style-type: none"> • Tekoälyllä voidaan seurata yksittäisiä terveys- tai kriisitietoja (esim. tunnistaa itsemurha- tai psykoosiriskissä olevat henkilöt). • Valvoo lääkityksen noudattamista merkitsemällä, kun potilaat lopettavat määrättyjen hoitojen ottamisen. 	Tekoälypohjaiset itsemurhahälytykset Facebookissa/Metassa, eli hahmontunnistusohjelmisto, jolla havaitaan itsemurha-aikeita ilmaisevat käyttäjät.
Terveysinformaatiot palvelunkäyttäjille	Tekoäly voi tukea palvelujen käyttäjiä s terveydenhuoltojärjestelmissä navigoinnissa, mukaan lukien: <ul style="list-style-type: none"> • Henkilökohtaisten terveystietojen hallinta. • Avustaminen viranomaisraportoinnissa, lomakkeiden täyttämässä ja sosiaalipalvelujen tai etuuksien hakemisessa esimerkiksi ChatGPT:n kaltaisten työkalujen avulla. 	MyChart auttaa palvelunkäyttäjiä hallinnoimaan henkilökohtaisia terveystietojaan ja kommunikoidaan palveluntarjoajien kanssa.
Koulutus ja tietojen jakaminen	Tekoäly voi parantaa mielenterveysosaamista levittämällä täsmällistä tietoa ja torjumalla leimautumista esimerkiksi sovellusten tai chatbottien avulla.	Ada Healthin kaltaiset tekoälyalustat tarjoavat mielenterveysvalistusta ja leimautumista vähentävää, ihmisille räätälöityä sisältöä.

Käyttö mielenterveysalan ammattilaisille

Käyttötarkoitu		Esimerkkejä
Digitaaliset terapiat	Lääkärit suosittelvat tai määräävät joitakin tekoälyyn integroituja työkaluja (esim. digitaalinen fenotyyppitys,	MindLAMP2, Woebot ja Wysa.

	chatbotit) osana hoitosuunnitelmia.	
Ammatillinen päätöksenteon tuki	Tekoäly esittää mielenterveysalan ammattilaisille tietoja päätöksenteon tueksi tai käynnistää toimia ilman, että he joutuvat antamaan tietoja.	Tempus käsittelee genomitietoja ja kliinisiä tietoja auttaakseen mielenterveysalan ammattilaisia luomaan yksilöllisiä hoitosuunnitelmia.
Potilaan ja väestön seuranta	<ul style="list-style-type: none"> • Tekoälyjärjestelmät analysoivat tietoja riskihenkilöiden tai väestökohorttien tunnistamiseksi (esim. itsemurhien ehkäisy). • Seuraa yhteisön terveystrendejä hallinnollisten tietojen tai puettavien laitteiden avulla. 	Crisis Text Line käyttää tekoälyä priorisoidakseen korkean riskin tapaukset tekstiviestien kielellisten mallien perusteella. Babylon Healthin kaltaiset tekoälyjärjestelmät merkitsevät väestön terveystrendejä, kuten masennuksen yleistymistä tietyillä maantieteellisillä alueilla.
Tuki oikeudellisissa ja hallinnollisissa tehtävissä	Tekoäly auttaa ammattilaisia sääntelyn noudattamisessa, raportoinnissa ja koko järjestelmän laajuisten terveystulosten analysoinnissa.	DocuSign ja vastaavat tekoälytyökalut auttavat ammattilaisia lakisääteisessä dokumentoinnissa ja vaatimustenmukaisuustehtävissä.
Tietojen jakaminen ja hallinto	<ul style="list-style-type: none"> • Sähköisten terveystietojen hallinta. • Kliinisen koodauksen ja ajanvarauksen tukeminen. 	Epic Systems tarjoaa tekoälyohjattua kliinistä päätöksentekotukea, joka on integroitu sähköisiin terveystietojärjestelmiin (EHR-järjestelmiin) tietojen hallinnan ja ajanvarauksen sujuvoittamiseksi; Zocdoc käyttää tekoälyä sovittamaan yhteen ihmisiä sopivien lääkkeiden kanssa erikoistumisen, vakuutuksen ja saatavuuden perusteella.
Viestintävälineet	<ul style="list-style-type: none"> • Palveluihin liittyvän viestinnän helpottaminen, kuten käyttäjien ja ammattilaisten yhdistäminen asiantuntemuksen, sijainnin ja muiden kriteerien perusteella. • Selkokielisten selitysten tarjoaminen monimutkaisista lääketieteellisistä tiedoista terveydenhuollon ammattilaisille. 	MOST (Moderated Online Social Therapy) on nuorille (12-25-vuotiaille) tarkoitettu digitaalinen mielenterveysalusta, jossa yhdistyvät vuorovaikutteiset terapiatyökalut, moderoitu vertaisverkosto sekä klinikkojen ja vertaistyöntekijöiden reaaliaikainen tuki; Galaxy. AI's Medical Language Simplifier muuntaa monimutkaisesta lääketieteellisestä terminologiasta selkeäksi ja ymmärrettäväksi kieleksi.

Mahdollisuudet

Näyttöä tekoälyjärjestelmien hyödyistä mielenterveystyössä on vain vähän (tätä käsitellään myöhemmin), koska ala on suurelta osin kokeiluvaiheessa. Useimmissa kommentteissa ja tutkimuksissa pohditaan, millaisia *mahdollisuuksia* tekoälyjärjestelmät voivat tarjota. Se, mitä pidetään mahdollisuutena, vaihtelee yksilöiden ja ryhmien välillä - mitä yksi henkilö tai ryhmä pitää mahdollisuutena, voi olla toisille harhaanjohtava tai ei-toivottu tavoite.

Ei näytä olevan mahdollista erottaa selkeästi toisistaan mahdollisuuksia, jotka hyödyttävät ensisijaisesti palvelujen käyttäjiä ja mielenterveysalan ammattilaisia, sillä ne näyttävät kietoutuvan toisiinsa. Samaa voidaan sanoa yhteiskunnan mahdollisuuksista yleisemmin. Mielenterveysalan tekoälyjärjestelmiä koskevassa tutkimuksessa on havaittu, että ensisijaisia mahdollisuuksia palvelujen käyttäjille ja mielenterveysalan ammattilaisille ovat muun muassa seuraavat:

- **Mielenterveystuen saatavuuden parantaminen**, kun chatbotit ja verkkoalustat voivat auttaa poistamaan maantieteelliset esteet ja tarjoamaan tukea syrjäseuduilla asuville henkilöille, ihmisille, jotka etsivät tukea tavanomaisten työaikojen ulkopuolella, tai niille, joiden on vaikea päästä perinteisiin kasvokkain tarjottaviin palveluihin. Saavutettavuuden parantaminen voisi auttaa erityisesti heikommassa asemassa olevia väestöryhmiä.³ Yksi esimerkki on monikielinen chatbot ChatPal, joka on kehitetty erityisesti edistämään henkistä hyvinvointia harvaan asutuilla alueilla asuvien henkilöiden keskuudessa, joilla perinteiset palvelut ovat maantieteellisen etäisyyden vuoksi rajallisia.⁴
- **Mahdollisuus vähentää hallintokustannuksia ja puuttua työvoimapulaan**: Tekoäly voi auttaa mielenterveyspalveluita esimerkiksi hallinnollisissa tehtävissä, kuten tapaamisten aikatauluttamisessa, palvelunkäyttäjävirtojen hallinnoinnissa ja raporttien tuottamisessa, jolloin aikaa vapautuu suoran tuen tarjoamiseen.⁵

Kliinisestä näkökulmasta katsottuna tutkimuksessa on käsitelty muun muassa seuraavia mahdollisuuksia:

- **Hoidon yksilöllistäminen**: Tekoälyjärjestelmät voisivat analysoida suuria tietomääriä, jotta ne voisivat tunnistaa malleja ja ennustaa, miten henkilö reagoi hoitoon, ja auttaa luomaan yksilöllisiin tarpeisiin räätälöityjä vastauksia.⁶
- **Oikea-aikainen tuki**: Tekoälyjärjestelmät voisivat auttaa ennakoimaan, milloin henkilö joutuu kriisiin, mikä voisi helpottaa oikea-aikaisen tuen antamista ja ehkäistä kriisin kärjistymistä. Tämä tavoite näyttäisi edellyttävän suuria määriä henkilökohtaisia tietoja, joita syntyy muun muassa puettavista laitteista, sähköisistä tallenteista ja sosiaalisen median toiminnasta.⁷
- **Parempi diagnostinen tarkkuus**: Tekoälyjärjestelmiä on käytetty diagnoositarkkuuden parantamiseen joillakin lääketieteen aloilla, kuten mammografian seulontamittareissa.⁸ Joidenkin kannattajien mukaan tämä voi olla mahdollista myös mielenterveysalalla analysoimalla monimutkaisia tietokokonaisuuksia ja tunnistamalla hienovaraisia kuvioita, jotka ihmiset saattavat jäädä huomaamatta diagnoosin tekemisessä.⁹
- **Uusien hoitomenetelmien kehittäminen**: Tekoäly voi helpottaa uusien terapeuttisten välineiden ja toimenpiteiden, kuten virtuaalitodellisuushoitojen, kehittämistä. Nämä tekniikat

voisivat tarjota vaihtoehtoisia täydentämään, parantamaan tai korvaamaan perinteistä tukea.¹⁰

- **Tutkimuksen ja tiedon tuottamisen tukeminen:** Tekoäly voi nopeuttaa mielenterveystutkimusta analysoimalla suuria tietokokonaisuuksia ja tunnistamalla malleja, kuten riskitekijöitä, hoitokohteita ja tiettyjen sosioekonomisten ryhmien palvelujen käytön malleja.¹¹

Jotkut ovat omaksuneet sosiaalisemman lähestymistavan etsiessään mahdollisuuksia tekoälyjärjestelmille ja katsovat, että keskittyminen vain diagnoosiin ja havaitsemiseen voi olla väärin. Sosiaalinen lähestymistapa voisi luonnehtia yhdeksi keskeiseksi mahdollisuudeksi **verkostoitunutta yhteistyötä**, jossa tekoälyjärjestelmiä käytetään sosiaalisten yhteyksien parantamiseen, vertaistukeen tai mielenterveysalan ammattilaisten koulutukseen ja valvontaan.¹² Esimerkkeinä voidaan mainita sellaisten parempien välineiden kehittäminen, joiden avulla palveluntarjoajat voivat parantaa henkilöstön taitoja ja empaattista ymmärrystä; verkottunut, vuorovaikutteinen media, jonka avulla tukityöntekijät voivat osallistua valvontaan ja kehittyä yhdessä; ja teknologiat, jotka helpottavat palvelunkäyttäjien yhteydenpitoa vertaisiinsa ja paikallis- tai verkkoyhteisöihin. Tämä sosiaalinen suuntautuminen perustuu usein mielenterveysongelmista ja psykososiaalisesta vammaisuudesta omaehtoista kokemusta omaavien ihmisten työhön.

Useimmat kommentoijat suhtautuvat näihin mahdollisuuksiin varovaisen optimistisesti, ja siitä, missä mahdollisuudet tarkalleen ottaen piilevät, saattaa olla erimielisyyttä, mikä on odotettavissa kokeiluvaiheessa

Riskit

Seuraavaksi esitellään kirjallisuudessa havaitut riskit. Niihin liittyy haittoja sekä yksilö- että kollektiivisella ja yhteiskunnallisella tasolla¹³. Ne vaarantavat terveydenhuollon keskeiset arvot, kuten autonomian, ihmisarvon ja luottamuksen.¹⁴

Turvallisuusriskit

Terveyteen liittyviä riskejä voi aiheutua tekoälyjärjestelmien tuottamista tai niiden pahentamista virheistä tai disinformaatiosta sekä siitä, että tekoälyjärjestelmät eivät kykene ymmärtämään erilaisia asiayhteyden liittyviä tekijöitä. Yksi merkittävä huolenaihe on **harjoitteluaineiston vääristyminen**. Jos nämä tiedot ovat epätäydellisiä, epäedustavia tai heijastavat olemassa olevia yhteiskunnallisia ennakkoluuloja, tekoäly voi ylläpitää näitä ennakkoluuloja ja johtaa epätarkkoihin tai syrjiviin diagnooseihin.¹⁵ Jos esimerkiksi tekoälyn kouluttamiseen käytetyt tiedot eivät edusta riittävästi tiettyjä etnisiä ryhmiä tai sosioekonomista taustaa, tekoäly saattaa tulkita väärin näissä ryhmissä yleisiä oireita tai käyttäytymistä.

Toinen ongelma on **väärien positiivisten ja negatiivisten tulosten mahdollisuus**. Tekoäly voi tuottaa vääriä positiivisia tuloksia, jolloin se ehdottaa tiettyä reaktiota, kuten itsemurhariskin merkitsemistä, vaikka sitä ei ole, tai vääriä negatiivisia tuloksia, jolloin se ei tunnista merkittävää riskiä tai kriisiä.¹⁶ Näillä virheillä on merkittäviä seurauksia, sillä ne voivat johtaa tarpeettomiin tai riittämättömiin palvelureaktioihin.

Julkisesti saatavilla olevien tekoälytyökalujen, kuten chatbottien, yleistyminen tuo mukanaan lisäriskejä. Nämä työkalut voivat antaa **haitallisia neuvoja tai väärää tietoa**, erityisesti kun niitä

käytetään mielenterveystukeen.¹⁷ Ne saattavat tulkita käyttäjän antamia tietoja väärin, tarjota epätarkkoja tietoja tai antaa epäasianmukaisia vastauksia, mikä saattaa luoda vääränlaisen turvallisuuden tunteen apua hakeville henkilöille.

Mielenterveyden luontainen monimutkaisuus on merkittävä haaste tekoälyjärjestelmille. Mielenterveysongelmat ilmenevät usein yksilöissä ainutlaatuisesti, ja niitä muokkaavat subjektiiviset kokemukset, kulttuuriset kontekstit, ympäristötekijät ja henkilökohtainen historia.¹⁸ Näitä vivahteita on vaikea, ellei jopa mahdotonta, vähentää laskennallisiin malleihin, minkä vuoksi tekoälyn on haastavaa tulkita tarkasti yksilöllisiä tarpeita tai tarjota luotettavaa tukea. Tämä rajoitus on erityisen ongelmallinen kriittisissä tilanteissa, joissa tarkka ymmärtäminen ja asianmukaiset reaktiot ovat ratkaisevan tärkeitä

Yksityisyys

Tekoälyn käyttöön mielenterveystyössä liittyvät **tietosuojariskit** johtuvat **mielenterveyteen liittyvien tietojen arkaluonteisuudesta** ja mahdollisista vahingoista, jos näitä tietoja käsitellään väärin tai käytetään väärin

Yhä suurempi riippuvuus digitaalisista alustoista ja toisiinsa kytketyistä järjestelmistä terveydenhuollossa ja sosiaalihuollossa on lisännyt **tietoturvaloukkausten ja luvattoman pääsyn** riskiä,¹⁹ kuten psykoterapeuttien tietoihin tehty laajamittainen hakkerointi, joka koski jopa 30 000 ihmistä Suomessa²⁰. Tämä on erityisen huolestuttavaa tekoälyjärjestelmien osalta, jotka keräävät ja analysoivat valtavia määriä arkaluonteisia henkilötietoja, kuten sairauskertomuksia, terapiamuistiinpanoja ja jopa sosiaalisen median toimintaa.

Toinen merkittävä kysymys on **tietojen jakaminen kolmansien osapuolten kanssa**. Yhä useammat ovat huolissaan siitä, että tekoälyjärjestelmät voivat luovuttaa henkilötietoja vakuutusyhtiöiden, työnantajien tai markkinointiyriyten kaltaisille tahoille ilman nimenomaista suostumusta. Tällaiset käytännöt voivat johtaa syrjintään, heikentää luottamusta mielenterveyspalveluihin ja olla harhaanjohtavaa tai vilpillistä toimintaa, jos tietoon perustuvaa suostumusta ei ole saatu asianmukaisesti.

Lisäksi **avoimuuden ja valvonnan puute** sen suhteen, miten tekoälyjärjestelmät käsittelevät henkilötietoja, on vakava haaste. Yksityishenkilöt eivät useinkaan tiedä tai pysty valvomaan, miten heidän tietojensa kerätään, käytetään tai jaetaan. Tämä toimivaltaa vailla oleminen voi lisätä huolta yksityisyyden suojasta, lannistaa ihmisiä etsimästä tukea ja estää heitä paljastamasta arkaluonteisia tietoja, jotka ovat välttämättömiä tehokkaan tuen saamiseksi.

Tietoon perustuvan suostumuksen puuttuminen

Tietoon perustuva suostumus on perusperiaate ihmisoikeuksien suojelussa terveydenhuoltoalalla.²¹ Tekoälyjärjestelmät aiheuttavat riskin **siitä, että tietoon perustuva suostumus on riittämätön tai vaarantunut**, kun niitä käytetään mielenterveyspalveluissa (ja terveydenhuollossa yleensä).

Ihmisten, myös palvelujen käyttäjien ja mielenterveysalan ammattilaisten, voi olla vaikea **ymmärtää täysin, miten tekoälyjärjestelmät toimivat**, millaista tietoa ne keräävät ja mitä näille tiedoille voi tapahtua. Tämä ymmärtämättömyys vaikeuttaa todellisen tietoon perustuvan suostumuksen antamista, koska ihmiset eivät välttämättä ymmärrä, mihin he ovat suostumassa. Jotkut eetikot ovat ehdottaneet, että generatiivinen tekoäly voisi teoriassa helpottaa tietoon perustuvan suostumuksen antamista, jos se pystyy tarjoamaan tietoa, joka on ainakin tarkempaa, helpommin saatavilla ja

luotettavampaa kuin mielenterveysalan ammattilaisten tarjoama tieto.²² Suostumus voi vaarantua, jos tekoälyjärjestelmien keräämiä tietoja voidaan käyttää myös **muihin tarkoituksiin kuin niihin, joihin alun perin annettiin suostumus**, kuten valtion virastojen toissijaiseen käyttöön tai kaupalliseen tiedonvälitykseen.²³

Tekoälymalleja koskevan avoimen raportoinnin puute heikentää niiden toistettavuutta ja vaikeuttaa mahdollisten vääristymien tai virheiden tunnistamista.²⁴

Uudet tai vahvistuneet eriarvoisuudet

Mielenterveysalan tekoälyä käytettäessä voidaan esiintyä **ennakkoluuloja ja syrjintää**, ja siihen liittyy muun muassa **ennakkoluuloista koostuvien tietokokonaisuuksien** käyttö. Tekoälyalgoritmit oppivat niistä tiedoista, joihin ne on koulutettu. Jos nämä tietokokonaisuudet sisältävät ennakkoluuloja tai syrjivää sisältöä, tekoälyjärjestelmät todennäköisesti perivät ja toistavat ne ennusteissaan ja suosituksissaan. Tämä voi päteä nykyisiin mielenterveyspalveluihin, joissa on tunnettuja ennakkoluuloja ja epätasa-arvoisuutta, kuten tiettyjen ryhmien kohtelu muita epäsuotuisammin, mutta tekoälyjärjestelmät voivat pahentaa nykyisiä eroja ja ylläpitää niitä laajemmassa mittakaavassa. Aiemmin mainitut avoimuuteen liittyvät ongelmat voivat myös vaikeuttaa tekoälyjärjestelmien ennakkoluulojen tunnistamista ja niihin puuttumista.²⁵

Tekoälyjärjestelmät voivat myös luoda uudenlaista syrjivää sisältöä, kuten tekoälyn tuottamia kuvia, jotka vahvistavat haitallisia stereotyyppioita. Esimerkiksi eräässä tutkimuksessa havaittiin, että käyttämällä sanaa skitsofrenia kehotuksena kuvien luomiseen tuotettiin kuvia, joissa oli "irvokkaita, epäluonnollisia kasvonpiirteitä... verta ja kauhun ilmaisuja".²⁶ Vastaavasti Googlen tutkimusryhmä osoitti, että mielenterveysongelmista kärsiviä ihmisiä kohtaan vallitsevat sosiaaliset asenteet, joissa heitä kuvataan pahoina ja jopa väkivaltaisina, koodattiin tekoälyjärjestelmiin, jotka on suunniteltu havaitsemaan vihapuhetta kirjoitetussa tekstissä.²⁷

Tekoälymallien kouluttamiseen käytetyistä tietokannoista puuttuu usein monimuotoisuutta, eivätkä ne edusta riittävästi tiettyjä populaatioita. Tämä voi johtaa siihen, että tekoälyjärjestelmät eivät ole yhtä tarkkoja tai tehokkaita mielenterveysongelmista kärsivien ryhmien kohdalla ja että ne tulkitsevat väärin yksilön tai yhteisön kokemuksia, mikä voi johtaa väärin diagnooseihin ja epätarkoituksenmukaiseen tai riittämättömään hoitoon. Tätä mahdollisuutta pahentaa myös se, että kokemusta omaavia ihmisiä ei yleensä oteta mukaan sellaisten teknologioiden luomiseen, suunnitteluun, kehittämiseen ja hallintaan, joiden väitetään olevan suunniteltu heitä hyödyttämään.

Joissakin tapauksissa, vaikka tietokokonaisuudet olisivat edustavia ja korkealaatuisia, tämä ei riitä ratkaisemaan **ihmisten järjestelmien ja tekoälyjärjestelmiä käyttävien instituutioiden harjoittamaa syrjintää**, kuten vakuutusperusteista syrjintää, jossa vakuutusmaksuja korotetaan, jos henkilöllä ennustetaan olevan mielenterveysongelmia.

Ennakkoluulojen lisäksi tekoälytyökalut voivat vahvistaa eriarvoisuutta, jos ne **eivät ole** tiettyjen ryhmien (esim. vammaisten tai heikosti digitaalista lukutaitoa omaavien) **saatavilla**.

Hoitotoimenpiteiden muuttuminen persoonattomiksi

Tekoälyjärjestelmistä, kuten chat-roboteista, puuttuu empatia, jota pidetään yleisesti ratkaisevana luottamuksen ja terapeuttien suhteiden rakentamisessa, vaikka ne saattavatkin pystyä matkimaan sitä. Robotit, chatbotit ja animaatiot voivat simuloida tunteita, kuten surua, empatiaa ja uteliaisuutta,

mutta ne ovat vain ohjelmoituja reaktioita, ja ne voivat johtaa ihmisiä harhaan, kun he luulevat koneille ihmismäistä empatiaa.

Tekoälyjärjestelmillä on huomattavia rajoituksia ihmisten tunteisiin vastaamisessa. Tekoälyjärjestelmät ovat työkaluja. Vaikka ne ovat taitavia käsittelemään tietoja, jotka voivat auttaa yksilöitä tai palveluntarjoajia valaisemaan henkilön kokemuksen osa-alueita, ne eivät kykene ymmärtämään ihmisen tunteiden monivivahteista monimutkaisuutta. Tämä voi johtaa epäherkkiin tai epäasianmukaisiin reaktioihin, erityisesti kun puhutaan hädän tai haavoittuvuuden ilmauksista,²⁸ ja saattaa vieraannuttaa käyttäjiä.

On myös olemassa vaara, että tekoälyn käyttö mielenterveyspalveluissa voi johtaa **hoidon depersonalisaatioon**. Vähentämällä inhimillistä vuorovaikutusta ja yhteydenpitoa tekoäly saattaa asettaa etusijalle mukavuuden ja tehokkuuden monien kriisitilanteissa olevien henkilöiden tarvitseman henkilökohtaisen ja empaattisen tuen kustannuksella.

Lisäksi liiallinen riippuvuus tekoälystä **voi mahdollisesti heikentää mielenterveyskuntoutujien empatiakykyä**. Jos tekoälyjärjestelmiä käytetään palvelunkäyttäjien seurantaan tai korvaamaan suoraa vuorovaikutusta, mielenterveyskuntoutujilla voi olla vähemmän mahdollisuuksia harjoitella ja kehittää kommunikointitaitoja, ja heistä voi tulla vähemmän herkkiä sille, mitä hoitokohtamisissa tarvitaan.²⁹

Valvonta

Jotkut kommentoijat ovat esittäneet huolensa tekoälyn mahdollisesta väärinkäytöstä valvontaan ja ohjaukseen, erityisesti kun kyseessä ovat mielenterveystiedot. Yksi huolenaihe on tekoälyjärjestelmien käyttö **itsemurhien ja itsensä vahingoittamisen ennustamiseen**. Vaikka tällaisia tekniikoita mainostetaan usein ennaltaehkäisyyn ja puuttumisen välineinä, niihin liittyy merkittäviä riskejä, kuten yksityisyyden suojan loukkauksia, ennusteiden epätarkkuutta ja aiheettomien toimenpiteiden mahdollisuutta. Tällaisia toimenpiteitä voivat olla esimerkiksi tahdonvastainen hoito tai lainvalvontaviranomaisten tarpeeton osallistuminen.³⁰

Biometriset valvontatekniikat aiheuttavat myös valvontariskejä. Keräämällä ja analysoimalla fysiologisia tietoja, kuten kasvojen ilmeitä, äänenkäyttöä, kävelyä ja silmien liikkeitä, nämä järjestelmät voivat yrittää päätellä mielentiloja ja käyttäytymistä. Tällainen käyttö hämärtää seurannan ja invasiivisen valvonnan välistä rajaa

Toinen ongelma on **arkaluonteisten mielenterveystietojen jakaminen lainvalvontaviranomaisten ja valtion virastojen kanssa**, kuten itsemurhayrityksiä koskevien muiden kuin rikosoikeudellisten tietojen käyttäminen lainvalvonta- tai rajaviranomaisten toimesta³¹.

Nämä käytännöt ovat vaarassa luoda "mielenterveysalan valvonnan markkinat, jotka ylläpitävät ja jopa laajentavat nykyisten mielenterveyskäytäntöjen pahinta vallan epätasapainoa, eriarvoisuutta ja haittoja"³².

Vahvistetaan yksilökeskeisiä näkemyksiä mielenterveydestä

Suurin osa mielenterveysalan tekoälyyn perustuvasta teknologiasta näyttää olevan suunnattu havaitsemiseen ja diagnosointiin, mikä kiinnittää huomion yksilöön, jonka on todettu tarvitsevan asiantuntijatoimia³³. Tämä voi johtaa biologisten tekijöiden ylikorostamiseen ja

epäoikeudenmukaiseen keskittymiseen yksilön vastuuseen mielenterveydestä³⁴ . Tämä kehys vahvistaa yksilökeskeistä näkemystä mielenterveydestä ja tekee näkymättömäksi laajemat sosioekonomiset, ihmissuhde- ja ympäristötekijät, jotka vuorovaikutuksessa henkilökohtaisten terveysongelmien kanssa muokkaavat ihmisten mielenterveyttä.

Rajallisten resurssien ohjaaminen muualle

Tekoälyä edistävät usein erittäin aktiiviset markkinat, joilla yritykset yrittävät myydä työkaluja hallituksille ja palveluntarjoajille. Nämä yritykset voivat olla erittäin strategisia, retorisesti hienostuneita ja hyvin organisoituneita vaikuttaessaan poliittiseen päätöksentekoon, myös hankintoihin.³⁵ .

Tekoälyn markkinointiin on sisällynyt mielettömiä väitteitä tekoälyn käytöstä mielenterveyden pitkäaikaisten ja monimutkaisten ongelmien ratkaisemiseksi, kuten väitteet, joiden mukaan "tekoälyavosirut... voisivat "ratkaista" autismin ja skitsofrenian".³⁶ Tällainen tekoälyhypetys, erityisesti jos näyttöä liioitellaan, voi muuttaa rahoituksen kohdentamista ja viedä resursseja pois sieltä, missä niitä eniten tarvitaan.

Väitteet siitä, että tekoäly ja muut digitaaliset teknologiat lisäävät välttämättä mielenterveytyön tehokkuutta, eivät useinkaan saa tukea todisteista, vaikka teknologiat esitettäisiin alhaisempina palvelukustannuksina verrattuna henkilökohtaiseen tukeen.³⁷ Myös *Privacy International* -järjestö on väittänyt, että on vain vähän näyttöä siitä, että tekoäly johtaa välttämättä tehokkaampiin terveydenhuoltojärjestelmiin, vaikka teknologiatoimittajat ovatkin yleisesti olettaneet, että näin tapahtuu.³⁸

Vaikka tekoälyjärjestelmien havaittaisiin lisäävän tehokkuutta, tehokkuuden ei pitäisi olla tärkein tai ainoa tavoite, sillä muut tavoitteet voivat olla yhtä tärkeitä tai jopa tärkeämpiä, kuten huolehtivien järjestelmien luominen.

Päätelmät ja suositukset

Tätä raporttia varten tutkituissa lähteissä todetaan, että tekoälyjärjestelmiä koskevien väitteiden tueksi tarvitaan vankkaa tutkimusta ja että mahdollisia riskejä ja haasteita on käsiteltävä. Niissä korostetaan, että tarvitaan eettisiä näkökohtia, huolellista toteutusta sekä hallinto- ja valvontamuotoja, joihin osallistuu mielenterveysongelmista kärsiviä ihmisiä, jotta hyödyt voidaan maksimoida ja haitat minimoida.

Kun säännellään tekoälyn käyttöä mielenterveyspalveluissa, on ratkaisevan tärkeää pohtia ihmissuhteita. Yksi tekoälyn yleisistä tavoitteista on pilkkoa tehtävät yksittäisiin osiin, jotka voidaan suorittaa toistuvasti. Hoito ei kuitenkaan ole vain tehtäviä, vaan myös tunteita; se on olennainen osa ihmissuhteita ja erittäin monimutkaista sosiaalista vuorovaikutusta" .³⁹

On väitetty, että EU:n nykyisessä sääntelykehyksessä ei juurikaan oteta huomioon tekoälyn ainutlaatua vaikutusta ihmisten väliseen vuorovaikutukseen ja tunteisiin, jotka ovat olennainen osa mielenterveyshoitoa⁴⁰ . Tämä valvonta vahvistaisi tekoälyä kehittävien yritysten rajallista vastuuvollisuutta ja vastuuta mielenterveyden alalla⁴¹ . On myös ratkaisevan tärkeää puuttua

yri­tysten eturistiriitoihin ja asettaa hyvälaatuinen mielenterveyshoito voittoa tavoittelevien näkökohtien edelle.

Mikä rooli eletyllä kokemuksella on tekoäly­teknologian kehittämisessä?

Elämäkokemusta omaavien ihmisten ottaminen mukaan mielenterveysalan tekoälyjärjestelmien kehittämiseen on ratkaisevan tärkeää, jotta voidaan luoda teknologioita, jotka ovat merkityksellisiä, tehokkaita ja eettisiä. Tämä näkökulma voi auttaa tunnistamaan palvelujen käyttäjien kanssa samansuuntaisia prioriteetteja, kuten yksityisyys, oikeudenmukaisuus ja inhimillinen yhteys.⁴² Näiden äänien mukaan ottaminen auttaa myös puuttumaan ennakkoluuloihin ja syrjintään, vallan epätasapainoon ja saavutettavuuden esteisiin, mikä parantaa todennäköisyyttä, että tekoälyjärjestelmät ovat tehokkaita, osallistavia ja luotettavia.

Jotkut tutkijat kannattavat sitä, että elämykselliset henkilöt otetaan mukaan koko tutkimus- ja kehittämisprosessiin aina ongelmien määrittelystä tiedon levittämiseen ja järjestelmien jatkuvaan hallintointiin.⁴³ Suosituksia elävän kokemuksen näkökulmien integroimiseksi ovat muun muassa tekoälyjärjestelmien yhteissuunnittelu, reflektiiviset käytännöt, läpinäkyvä päätöksenteko, heidän johtamisensa priorisointi, monimuotoisuuden varmistaminen ja osallistumisen oikeudenmukainen korvaaminen. Algoritmisen ja datapohjaisen teknologian vaikutuspiiriin eniten kuuluvien aktiivista osallistumista ei pitäisi nähdä pelkästään "sidosryhmien sitouttamisen" edellyttämänä askeleena, vaan pikemminkin eettisenä välttämättömyytenä. Kun kehittäjät, julkisista hankinnoista vastaavat viranomaiset, mielenterveyspalvelut ja muut tahot sisällyttävät eletyn kokemuksen asiantuntemuksen tällä tavoin, tekoälyjärjestelmät heijastavat todennäköisemmin todellisen maailman tarpeita ja edistävät tekoälyjärjestelmien hyväksynnän lisäksi aitoa luottamusta ja hyödyllisyyttä.

Suositukset poliittisille päättäjille

Turvallisuusriskin poistamiseksi

- ✓ Varmistetaan **palvelujen korkea laatu** integroimalla tekoäly korkeatasoisen tieteellisen eheyden mukaisesti, mukaan lukien vankka validointi ennen laajamittaista käyttöä.
- ✓ Perustetaan riippumaton sääntelyelin, joka valvoo ja käsittelee mielenterveyspalveluissa tapahtuvaan tekoälyn väärinkäyttöön liittyviä valituksia (**vastuuvollisuus**).
- ✓ **Toteutetaan pakollinen raportointi haittavaikutuksista** ja vaaditaan tekoälyn kehittäjiä ja terveydenhuollon tarjoajia raportoimaan ja seuraamaan tekoälyjärjestelmien aiheuttamia virheitä, ennakkoluuloja ja tahattomia haittoja.

Yksityisyysriskin poistaminen

- ✓ Voimaannuttaa vankat tietosuojatoimenpiteet, mukaan lukien selkeät suostumuspöytäkirjat ja sääntelykehykset, joissa asetetaan etusijalle yksilön ja kollektiiviset yksityisyyden suoja koskevat oikeudet.

- ✓ **Kielletään** mielenterveystietojen **luvatun jakaminen** ja estetään tekoälyjärjestelmiä siirtämästä mielenterveystietoja kolmansille osapuolille (esim. vakuutusenantajille, työnantajille, mainostajille) ilman nimenomaista, tietoon perustuvaa suostumusta.

Asianmukaisen tietoon perustuvan suostumuksen puuttuminen

- ✓ Varmistetaan tekoälyn **läpinäkyvyys ja selitettävyyys**, jotta ihmiset saavat selkeää tietoa siitä, miten nämä järjestelmät toimivat, mitä tietoja ne käyttävät, miten ne tekevät päätöksiä ja kenen kanssa tietoja voidaan jakaa

puuttua riskiin, että **syntyy uutta eriarvoisuutta tai että se vahvistuu.**

- ✓ Velvoitetaan **sisällyttämään tekoälyn harjoittelua varten laadittuihin tietokokonaisuuksiin erilaisia väestöryhmiä**, jotta voidaan vähentää vääristymiä ja parantaa tarkkuutta kaikkien yhteisöjen osalta.
- ✓ Varmistetaan, että **haavoittuvassa ja/tai syrjäytyneessä asemassa olevien yhteisöjen jäsenet ja ihmiset, joilla on elettyjä kokemuksia, osallistuvat aktiivisesti** tekoälyteknologian suunnitteluun, kehittämiseen ja testaamiseen, jotta ennakkoluuloja voidaan vähentää.
- ✓ Tukea validoitujen mielenterveyden tekoälytyökalujen käyttöä vähäosaisille yhteisöille **digitaalisen osallisuuden** varmistamiseksi.
- ✓ Kehitetään ja pannaan täytäntöön tekoälytyökalujen **esteettömyysohjeet**, joilla varmistetaan käytettävyys vammaisille, kielitaidottomille tai heikosti digitaalista lukutaitoa omaaville henkilöille.

Hoitoriskin **henkilöitymisen** estäminen.

- ✓ Vaaditaan **ihmisen valvontaa** tekoälyn perustuvassa mielenterveystyössä, jotta varmistetaan, että empatia pysyy hoidossa keskeisenä.

Valvontariskin käsittelemiseksi

- ✓ Vaaditaan tekoälyn kehittäjiä ilmoittamaan yksilöille heidän vuorovaikutuksestaan tekoälyn kanssa ja varmistamaan **opt-in-lähestymistapa** tai - jos tällainen lähestymistapa ei ole toteutettavissa - ainakin varmistamaan mahdollisuus jättäytyä pois tekoälyjärjestelmien käytöstä.
- ✓ **Säännellään mielenterveystietojen käyttöä lainvalvontaviranomaisten keskuudessa** ja rajoitetaan mielenterveyteen liittyvien tekoälyn liittyvien tietojen jakamista poliisin, maahanmuuttoviranomaisten ja muiden valtion virastojen kanssa.

puuttua riskiin, että **mielenterveyttä koskevat yksilökeskeiset näkemykset vahvistuvat.**

- ✓ **Vastustakaa "teknoratkaisujen" houkutusta**, eli uskoa siihen, että kaikki ongelmat voidaan ratkaista helpolla teknologisella ratkaisulla. Resursseja ja ponnisteluja olisi sijoitettava mielenterveyttä muokkaavien laajempien tekijöiden käsittelyyn, jolloin painopiste siirtyy yksilöstä häntä ympäröivään laajempaan kontekstiin.

uudelleen kohdentamisen riskin torjumiseksi

- ✓ Edellytetään, että tekoälyn kehittäjät asettavat **hoidon laadun voittojen edelle** ja todistavat, että tekoälyllä vastataan todelliseen tarpeeseen eikä siitä aiheudu haittaa. Julkista rahoitusta

voidaan käyttää vipuvoimana tähän suuntaan. Yhteiskehittämisen avulla voidaan varmistaa, että tekoälyn kehitys vastaa todellisiin tarpeisiin.

Suosituksukset kansalaisyhteiskunnalle

- ✓ **Monitor AI in Mental Healthcare:** Seuraa tekoälytyökalujen käyttöönottoa mielenterveysympäristöissä, dokumentoi ennakkoluuloja, yksityisyyden suojan loukkauksia tai haittatapauksia ja puolusta poliittisia toimenpiteitä.
- ✓ **Keskuksen kokemukset tekoälyn kehittämisestä:** Tavoitteena on selvittää, mitkä tekoälysovellukset ovat hyödyllisiä, mitkä haitallisia ja miten tekoäly voidaan suunnitella siten, että se tukee mielenterveyspalveluja eikä heikennä niitä.
- ✓ **Lisätään tietoisuutta tekoälyn riskeistä ja oikeuksista:** Käynnistetään julkisia kampanjoita yhteistyössä mielenterveys- ja digitaalisten oikeuksien järjestöjen kanssa tekoälyn riskien korostamiseksi mielenterveystyössä ja korostetaan ihmisoikeuskeskeisen sääntelyn tarvetta.
- ✓ **Osallistutaan poliittisiin ja lainsäädännöllisiin prosesseihin:** Osallistutaan aktiivisesti kuulemisiin, asiantuntijapaneelisiin ja poliittiseen päätöksentekoon sen varmistamiseksi, että tekoälyä koskevissa säädöksissä asetetaan etusijalle mielenterveystyökalut, saatavuus ja eettiset normit ja estetään tekoälyä vahvistamasta leimautumista tai rajoittamasta hoidon saantia.

Kirjallisuusluettelo

¹ Hannah van Kolschooten ja Janneke van Oirschot, "The EU Artificial Intelligence Act (2024): Implications for Healthcare" (2024) 149 *Health Policy* 105152 ("The EU Artificial Intelligence Act (2024)").

² Hannah van Kolschooten ja Janneke van Oirschot, *When People Become Data Points: The Potential Impact of AI in Mental Healthcare* (Health Action International, joulukuu 2024) <<https://haiweb.org/wp-content/uploads/2024/12/AI-in-Mental-Healthcare.pdf>>.

³ Shane Cross et al, "Use of AI in Mental Health Care: Community and Mental Health Professionals Survey" (2024) 11(1) *JMIR Mental Health* e60589.

⁴ Courtney Potts et al, "Monikielinen digitaalinen mielenterveys- ja hyvinvointikeskustelurobotti (ChatPal): Pre-Post Multicenter Intervention Study" (2023) 25(1) *Journal of Medical Internet Research* e43051.

⁵ Maailman terveysjärjestö, "Artificial Intelligence in Mental Health Research: New WHO Study on Applications and Challenges" (6. helmikuuta 2023) <<https://www.who.int/europe/news-room/06-02-2023-artificial-intelligence-in-mental-health-research--new-who-study-on-applications-and-challenges>>.

⁶ Ibid.

⁷ Anoushka Thakkar, Ankita Gupta ja Avinash De Sousa, "Artificial Intelligence in Positive Mental Health: A Narrative Review" (2024) 6 *Frontiers in Digital Health* 1280235.

⁸ Ks. esim. Nora Eisemann et al, "Nationwide Real-World Implementation of AI for Cancer Detection in Population-Based Mammography Screening" [2025] *Nature Medicine* 1.

⁹ Charlotte Blease ja Adam Rodman, "Generative Artificial Intelligence in Mental Healthcare: An Ethical Evaluation" (2024) 5(12) *Current Treatment Options in Psychiatry* <<https://link.springer.com/article/10.1007/s40501-024-00340-x>>.

¹⁰ Thakkar, Gupta ja De Sousa (n:o 7).

¹¹ Roberto Tornero-Costa et al, "Methodological and Quality Flaws in the Use of Artificial Intelligence in Mental Health Research: Systematic Review" (2023) 10(1) *JMIR Mental Health* e42045.

¹² Jonah Bossewitch, "Brave New Apps: The Arrival of Surveillance Psychiatry", *Mad In America* (9. elokuuta 2019) <<https://www.madinamerica.com/2019/08/brave-new-apps-the-arrival-of-surveillance-psychiatry/>>.

¹³ van Kolfshooten & van Oirschot (n:o 2).

¹⁴ Hannah van Kolfshooten, "EU:n tekoälysäätely: (2022), 59, Common Market Law Review, numero 1, s. 81-112.

¹⁵ Rhonda L Wilson et al, "Artificial Intelligence: n/a(n/a) *International Journal of Mental Health Nursing* <<https://onlinelibrary.wiley.com/doi/abs/10.1111/inm.13121>>.

¹⁶ Mason Marks, "Artificial Intelligence-Based Suicide Prediction" (2019) 21(98) *Yale J.L. & Tech.* 24.

¹⁷ Blease ja Rodman (n:o 9).

¹⁸ Tornero-Costa et al. (n 11).

¹⁹ Thakkar, Gupta ja De Sousa (n:o 7).

²⁰ Vastaamon uhrien korvaukset epävarmoja. (2021, 20. kesäkuuta). *Yle Uutiset*. https://yle.fi/uutiset/osasto/news/compensation_uncertain_for_vastaamo_victims/11991155

²¹ Ks. esim. Euroopan ihmisoikeussopimus, 3 ja 8 artikla; vammaisten henkilöiden oikeuksia koskeva yleissopimus, 3, 12 ja 25 artikla.

²² Blease ja Rodman (n:o 9).

²³ Ks. esim. Federal Trade Commission (Yhdysvallat), "FTC to Ban BetterHelp from Revealing Consumers' Data, Including Sensitive Mental Health Information, to Facebook and Others for Targeted Advertising" (2.3.2023) <<https://www.ftc.gov/news-events/news/press->

releases/2023/03/ftc-ban-betterhelp-revealing-consumers-data-including-sensitive-mental-health-information-facebook>.

²⁴ Maaailman terveystajärjestö (n 5).

²⁵ Hannah van Kolfshoeten, "The AI cycle of health inequity and digital ageism: mitigating biases through the EU regulatory framework on medical devices" 2023) *Journal of Law and the Biosciences* 10(2), heinä-joulukuu 2023, lsad031, <https://doi.org/10.1093/jlb/lad031>.

²⁶ Morgan King, "Harmful Biases in Artificial Intelligence" (2022) 9(11) *The Lancet Psychiatry* e48.

²⁷ Ben Hutchinson et al, 'Social Biases in NLP Models as Barriers for Persons with Disabilities' [2020] *arXiv:2005.00813 [cs]* <<http://arxiv.org/abs/2005.00813>>.

²⁸ Geoff White, 'Child Advice Chatbot Fail to Spot Sexual Abuse', *BBC News* (online, 11.12.2018) <<https://www.bbc.com/news/technology-46507900>>.

²⁹ Ks. esim. Rhonda L Wilson et al, "Artificial Intelligence: An Eye Cast towards the Mental Health Nursing Horizon" (2023) 32(3) *International Journal of Mental Health Nursing* <<https://onlinelibrary.wiley.com/doi/abs/10.1111/inm.13121>>; Blease ja Rodman (n 9).

³⁰ Marks (n 16).

³¹ Office of the Privacy Commissioner of Canada, "Disclosure of Information about Complainant's Attempted Suicide to US Customs and Border Protection Not Authorized under the Privacy Act" (21. syyskuuta 2017) <https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-federal-institutions/2016-17/pa_20170419_rcmp/>.

³² Jonah Bossewitch et al, *Digital Futures in Mind: Reflecting on Technological Experiments in Mental Health and Crisis Support* (raportti, Melbournen yliopisto, 1. syyskuuta 2022) <<https://automatingmentalhealth.cc/>>.

³³ Ibid.

³⁴ Ibid.

³⁵ Piers Gooding, "Market Forces in Automated Mental Health Services: New Claims in Algorithmic Care and Disability Justice" teoksessa Lieketseng Ned et al (toim.), *The Routledge International Handbook of Disability and Global Health* (Routledge, 2024) 565.

³⁶ Isobel Asher Hamilton, 'Elon Musk Said His AI-Brain-Chips Company Could "solve" Autism and Schizophrenia', *Business Insider Australia* (14 November 2019) <<https://www.businessinsider.com.au/elon-musk-said-neuralink-could-solve-autism-and-schizophrenia-2019-11>>.

³⁷ Jacqueline Sin et al, "Digital Interventions for Screening and Treating Common Mental Disorders or Symptoms of Common Mental Illness in Adults: Systematic Review and Meta-Analysis" (2020) 22(9) *Journal of Medical Internet Research* e20581.

³⁸ Privacy International, "Our Analysis of the WHO Report on Ethics and Governance of Artificial Intelligence for Health", *Privacy International* (20.7.2021) <<http://privacyinternational.org/news-analysis/4594/our-analysis-who-report-ethics-and-governance-artificial-intelligence-health>>.

³⁹ Bossewitch et al (n 32).

⁴⁰ Tamar Tavory, "Tekoälyn sääntely mielenterveysalalla: Ethics of Care Perspective" (2024) 11(1) *JMIR Mental Health* e58493.

⁴¹ Ibid.

⁴² Ian Muchamore, Panos Karanikolas ja Piers Gooding, 'How Lived Experience Expertise Shapes Research and Development in Digital Mental Health' <<https://apo.org.au/node/328006>>.

⁴³ Kellie Gilbert et al, "Digital Futures in Mind: Why Lived Experience Collaboration Must Guide Digital Mental Health Technologies" (2024) (ensin verkossa) *Australian Journal of Social Issues* <<https://onlinelibrary.wiley.com/doi/abs/10.1002/ajs4.355>>.



Näiden aineistojen käytön yhteydessä on mainittava lähde. Mitään näistä materiaaleista ei saa muuttaa tai mukauttaa ilman tekijöiden nimenomaista kirjallista lupaa.

© Mental Health Europe (joulukuu 2024)

Esitetyt näkemykset ja mielipiteet ovat ainoastaan kirjoittajan (kirjoittajien) omia eivätkä välttämättä vastaa Euroopan unionin tai Euroopan komission CERV-ohjelman näkemyksiä ja mielipiteitä. Euroopan unioni tai tuen myöntävä viranomais ei ole vastuussa niistä.



Co-funded by
the European Union