



**Mental
Health
Europe**

Equal rights.
Better mental health.
For all.

December 2024

Kunstig intelligens i mental sundhedspleje



Denne undersøgelse fra Mental Health Europe er forfattet af Piers Gooding*, Hannah van Kolschooten** og Francesca Centola***

*Lektor, La Trobe Law School, La Trobe University.

** Underviser og forsker, Amsterdam Law School, University of Amsterdam, og AI-konsulent, Health Action International

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

Hvis du har spørgsmål, bedes du kontakte: f.centola@mentalhealtheurope.org

Dette dokument er oversat fra engelsk ved hjælp af maskinoversættelsessoftware; derfor garanteres der ikke for nøjagtigheden af oversættelsen.

Indhold

Ordliste.....	3
Sammenfatning.....	3
Om denne undersøgelse	4
Baggrund	4
Oversigt over AI-applikationer inden for mental sundhed	5
Muligheder.....	8
Risici	9
Sikkerhedsrisici.....	9
Privatlivets fred	10
Manglende informeret samtykke.....	10
Nye eller forstærkede uligheder	10
Depersonalisering af plejen	11
Overvågning	12
Forstærker det individualistiske syn på mental sundhed.....	12
Omdirigering af begrænsede ressourcer	12
Konklusioner og anbefalinger	13
Hvilken rolle kan levede erfaringer spille i udviklingen af AI-teknologier?	13
Anbefalinger til politiske beslutningstagere.....	14
Anbefalinger til civilsamfundet	15
Bibliografi	15

Ordliste

"AI-system" er defineret i artikel 3, stk. 1, i Den Europæiske Unions (EU) lov om kunstig intelligens (AI) som:

"et maskinbaseret system, der er designet til at fungere med forskellige niveauer af autonomi, og som kan udvise tilpasningsevne efter implementering, og som for eksplicite eller implicite mål udleder af det input, det modtager, hvordan man genererer output som f.eks. forudsigelser, indhold, anbefalinger eller beslutninger, der kan påvirke fysiske eller virtuelle miljøer."

Denne definition vil blive brugt i denne rapport.

Samskabelse: En samarbejdstilgang, der involverer alle aktører inden for mental sundhed, som arbejder sammen på lige fod for at udvikle og implementere politikker, tjenester, programmer, forskning og kommunikation, der fremmer positiv mental sundhed i henhold til en psykosocial model og en menneskerettighedsbaseret tilgang. ([Mental Health Europe](#))

Digital fænotyping henviser til brugen af data fra digitale enheder (såsom smartphones og wearables) til at identificere adfærdsmønstre, der er forbundet med bestemte mentale sundhedstilstande.

Generativ AI en delmængde af kunstig intelligens, der fokuserer på at skabe nyt indhold, lige fra tekst og billeder til lyd, video, 3D-modeller og syntetiske data. Dette felt bruger primært [maskinlæringsalgoritmer](#), især deep learning-modeller, til at skelne mellem mønstre i træningsdata og generere nye output.

Mennesker med levnet erfaring: mennesker, der oplever eller har oplevet psykiske problemer. ([Mental Health Europe](#))

Sammenfatning

Denne undersøgelse udforsker mulighederne, risiciene og de etiske overvejelser omkring brugen af kunstig intelligens (AI) i mental sundhedspleje og giver anbefalinger til ansvarlig implementering og regulering af dem.

Sundhedsvæsenet er en af de mest populære sektorer for implementering af AI i EU.¹ Inden for mental sundhed anvendes AI-systemer på forskellige måder, fra administrative opgaver til kommunikationsplatforme, professionel beslutningsstøtte, digitale terapier som chatbots og personlige sensorteknologier og patientovervågning. Ud over sundhedssektoren omfatter AI-applikationer, der bruger data om mental sundhed, også strafferet, forbrugerprodukter, uddannelse og beskæftigelse.

AI-systemer giver betydelige potentielle fordele, herunder forbedret adgang til mental sundhedsstøtte, især for underforsynede befolkningsgrupper, og reducerer den administrative byrde i sundhedssystemerne. På det kliniske plan er der fortalere for, at AI kan personalisere behandlinger, forbedre den diagnostiske nøjagtighed og understøtte rettidige indgreb. Derudover kan AI fremskynde forskning i mental sundhed og lette udviklingen af nye terapeutiske tilgange, f.eks. virtual reality-

behandlinger. Socialt orienterede muligheder omfatter forbedring af peer-support og uddannelse af behandlere, fremme af netværkssamarbejde og fremme af større samfundsforbindelser.

Men AI-systemer inden for mental sundhed udgør også alvorlige risici, som kan identificeres på tre niveauer: individuelt, kollektivt og samfundsmæssigt². På det individuelle niveau omfatter bekymringerne sikkerhedsrisici, krænkelser af privatlivets fred og utilstrækkeligt informeret samtykke. Bredere udfordringer omfatter styrkelse af uligheder eller skabelse af nye, overovervågning, forstærkning af individualistiske syn på mental sundhed, depersonalisering af pleje og omdirigering af begrænsede ressourcer.

Rapporten fremhæver nøgleprincipper og mulige handlingsforløb for politiske beslutningstagere for at imødegå disse risici ved at afbalancere innovation med etisk ansvar og menneskecentreret pleje i AI-drevne mentale sundhedssystemer. Der argumenteres for, at AI-værktøjer skal udvikles med tanke på etik, inklusion, nøjagtighed, sikkerhed og slutbrugernes reelle behov. Mulige løsninger omfatter primært robust regulering og tilsyn, gennemsigtighed og forklarlighed samt menneskerettighedscentrerede og samskabende tilgange.

Den aktive deltagelse og ledelse af de mest berørte samfund - mennesker med levede erfaringer - bør betragtes som et grundlæggende etisk og politisk princip, der styrer alle diskussioner og beslutninger om digitalisering og mental sundhed. At sikre, at de har en ligeværdig stemme i udformningen af politikker og innovation, er afgørende for at opnå retfærdige og afbalancerede resultater og forhindre en uddybning af ulighederne på sundhedsområdet i samfundet.

Om denne undersøgelse

Denne undersøgelse henvender sig til politiske beslutningstagere, tjenesteudbydere og alle andre interessenter, der er interesserede i anvendelsen og effekten af kunstig intelligens inden for mental sundhed.

Efter en indledende baggrund, der beskriver Mental Health Europes vision, giver undersøgelsen et overblik over anvendelsen af kunstig intelligens inden for mental sundhed. Derefter vurderes de muligheder og risici, der er forbundet med disse teknologier, og der gives anbefalinger til politiske beslutningstagere om, hvordan de kan håndtere risiciene. Undersøgelsen fremhæver den rolle, som levede erfaringer spiller i udviklingen af kunstig intelligens, og giver specifikke anbefalinger til civilsamfundet.

Baggrund

I 2022 udgav Mental Health Europe en rapport om [mental sundhed i den digitale tidsalder: Anvendelse af en menneskerettighedsbaseret, psykosocial tilgang som kompas](#). I rapporten fremlagde Mental Health Europe sin vision for mental sundhed i en digital verden: Digitalisering bør ikke betragtes som et mål i sig selv, men som et middel til at nå et større mål. Et middel til et mentalt sundere samfund, øget autonomi og handlekraft i forhold til vores egen mentale sundhed og en bedre tilpasset mental sundhedspleje.

Vi understregede vigtigheden af, at nye teknologier ikke kun anvendes, fordi de er rentable eller gennemførlige, men fordi de opfylder et reelt behov. Det er særligt vigtigt i betragtning af de kommercielle interesser, der er på spil. Vi slog til lyd for en samskabelsestilgang og en menneskerettighedsramme, der skal fungere som kompas for enhver udvikling inden for digital mental

sundhed. Kun med denne samarbejdstilgang - i hvert trin af processen, fra design til evaluering - kan digitale teknologier tilpasses reelle behov og arbejde hen imod at realisere en vision om et samfund, hvor alle kan nyde deres menneskerettigheder fuldt ud og trives.

Siden 2022 er der sket mange udviklinger på EU-plan, herunder vedtagelsen af [AI-loven](#), den første omfattende regulering af AI fra en større myndighed nogen steder. Denne forordning trådte i kraft i august 2024 og har til formål at sikre, at AI-systemer er sikre og respekterer grundlæggende rettigheder og værdier, at fremme tilliden til AI-teknologier, at støtte innovation og at forbedre EU's konkurrenceevne inden for AI. AI-loven gælder horisontalt for alle sektorer, hvor der anvendes AI, herunder sundhedssektoren

Ambitionen med denne undersøgelse er at fokusere på virkningen af AI-applikationer inden for mental sundhedspleje, fremhæve de risici, der er specifikke for denne sektor, og foreslå foranstaltninger til at håndtere dem. På den måde lægger den grunden til en kritisk vurdering af, om AI-loven er egnet til formålet i forbindelse med mental sundhedspleje, og hvordan potentielle huller kan håndteres af de politiske beslutningstagere. Undersøgelsen understreger vigtigheden af samarbejde mellem politiske beslutningstagere, mennesker med levede erfaringer, civilsamfundsorganisationer og andre relevante interessenter for at sikre, at disse huller håndteres effektivt i løbet af forordningens implementeringsfase.

Oversigt over AI-applikationer inden for mental sundhedspleje

I psykiatrien bruges AI-systemer på forskellige måder. Tabellerne nedenfor giver eksempler på AI-anvendelser og skelner mellem, om de er rettet mod brugere eller fagfolk i psykiatrien.

AI-applikationer til

		Eksempler
Digitale terapier	Digital fænotypebestemmelse: AI analyserer data fra digitale enheder (f.eks. smartphones, wearables) for at identificere adfærdsmønstre, der er forbundet med mentale sundhedstilstande (f.eks. depression, angst).	MindLAMP2 (hvor LAMP står for Learn, Assess, Manage, Prevent) er en open source og frit tilgængelig app, som genererer flere, tilpassede datastrømme fra en bruger (f.eks. enhedens bevægelse, lokaliseringsaktivitet) og tilbyder aktiviteter, der tilskynder til mindfulness og selvrefleksion. MindLAMP2 er blevet brugt i en række forskellige kliniske og forskningsmæssige anvendelser inden for mental sundhed.
	Chatbots: Interaktive AI-værktøjer, der fungerer som virtuelle rådgivere, efterligner terapeutiske møder og giver råd. De kan også guide brugerne	Woebot og Wysa er AI-chatbots, der er designet til at efterligne terapeutiske interaktioner og tilbyde rådgivning om mental sundhed og mestringsstrategier.

	til at få adgang til sociale tjenester eller støttesystemer.	
Personlig medicin	AI udnytter genetiske, miljømæssige og adfærdsmæssige data til at skabe behandlingsplaner for individuelle brugere.	IBM Watson Health behandler patientdata for at anbefale personlige behandlingsmuligheder for depression baseret på en kombination af genetiske og adfærdsmæssige faktorer.
Patientovervågning og -kontrol	<ul style="list-style-type: none"> • AI kan spore individuelle sundheds- eller krisedata (f.eks. identificere dem, der er i risiko for selvmord eller psykose). • Overvåger overholdelse af medicinering ved at markere, når patienter holder op med at tage ordinerede behandlinger. 	AI-baserede selvmordsadvarsler hos Facebook/Meta, dvs. mønstergenkendelsessoftware til at opdage brugere, der udtrykker selvmordstanker.
Sundhedsinformatik for servicebrugere	AI kan hjælpe brugerne med at navigere i sundhedssystemerne, bl.a: <ul style="list-style-type: none"> • Håndtering af personlige sundhedsjournaler. • Hjælpe med lovpligtig rapportering, udfyldelse af formularer og ansøgning om sociale ydelser ved hjælp af værktøjer som ChatGPT. 	MyChart hjælper brugerne med at administrere deres personlige sundhedsjournaler og kommunikere med udbydere.
Uddannelse og informationsdeling	AI kan forbedre kendskabet til mental sundhed ved at formidle præcis information og bekæmpe stigmatisering via platforme som apps eller chatbots.	AI-platforme som Ada Health tilbyder undervisning i mental sundhed og stigma-reducerende indhold, der er skræddersyet til mennesker.

Anvendelser for fagfolk inden for mental sundhed

		Eksempler
Digitale terapier	Nogle AI-integrerede værktøjer (f.eks. digital fænotypebestemmelse, chatbots) anbefales eller ordineres af praktiserende læger som en del af behandlingsplanerne.	MindLAMP2, Woebot og Wysa.

Professionel beslutningsstøtte	AI præsenterer data for behandlere inden for mental sundhed for at understøtte beslutningstagning eller udløse handlinger uden at kræve input.	Tempus behandler genomiske og kliniske data for at hjælpe psykiatere med at skabe personlige behandlingsplaner.
Overvågning af patienter og befolkning	<ul style="list-style-type: none"> • AI-systemer analyserer data for at identificere risikopersoner eller befolkningsgrupper (f.eks. selvmordsforebyggelse). • Spør sundhedstendenser i samfundet ved hjælp af administrative data eller indsigter fra bærbare enheder. 	Crisis Text Line bruger AI til at prioritere højrisikosager baseret på sprogmønstre i sms'er. AI-systemer som Babylon Health markerer tendenser i befolkningens sundhed, f.eks. stigende forekomst af depression i bestemte geografiske områder.
Støtte i juridiske og administrative roller	AI hjælper fagfolk med at overholde lovgivningen, rapportere og analysere sundhedsresultater for hele systemet.	DocuSign og lignende AI-værktøjer hjælper fagfolk med lovpligtig dokumentation og compliance-opgaver.
Informationsdeling og administration	<ul style="list-style-type: none"> • Håndtering af elektroniske patientjournaler. • Støtte til klinisk kodning og planlægning af aftaler. 	Epic Systems leverer AI-drevet klinisk beslutningsstøtte integreret i elektroniske patientjournalssystemer (EHR) for at strømline journalhåndtering og aftaleplanlægning; Zocdoc bruger AI til at matche folk med passende praktiserende læger baseret på specialisering, forsikring og tilgængelighed.
Kommunikationsværktøjer	<ul style="list-style-type: none"> • Facilitering af servicerelateret kommunikation, som f.eks. at matche brugere med behandlere baseret på ekspertise, placering og andre kriterier. • Giver forklaringer i almindeligt sprog på komplekse medicinske oplysninger til sundhedspersonale. 	MOST (Moderated Online Social Therapy) er en digital mental sundhedsplatform for unge (12-25), der kombinerer interaktive terapeutiske værktøjer, et modereret peer-netværk og støtte i realtid fra klinikere og peer-medarbejdere; Galaxy. AI's Medical Language Simplifier konverterer kompleks medicinsk terminologi til almindeligt, forståeligt sprog.

Muligheder

Beviserne for fordelene ved AI-systemer i forbindelse med mental sundhed er begrænsede (diskuteres senere), da området i høj grad befinder sig på et eksperimentelt stadie. De fleste kommentarer og undersøgelser handler derfor om, hvilke *muligheder* AI-systemer kan tilbyde. Hvad der betragtes som en mulighed, er forskelligt fra person til person og fra gruppe til gruppe - hvad én person eller gruppe ser som en mulighed, kan af andre ses som et misforstået eller uønsket mål.

Det ser ikke ud til at være muligt at skelne skarpt mellem muligheder, der primært er til gavn for brugere og fagfolk inden for mental sundhed, da de ser ud til at være sammenflettede. Det samme kan siges om muligheder for samfundet mere generelt. De primære muligheder for brugere og fagfolk i psykiatrien, der er nævnt i forskningen om AI-systemer i psykiatrien, omfatter:

- **Forbedret tilgængelighed af mental sundhedsstøtte**, hvor chatbots og onlineplatforme kan hjælpe med at overvinde geografiske barrierer og yde støtte til personer i fjernliggende områder, folk, der søger støtte uden for typisk arbejdstid, eller dem, der har svært ved at få adgang til traditionelle ansigt-til-ansigt-tjenester. Øget tilgængelighed kan især hjælpe underforsynede befolkningsgrupper.³ Et eksempel er den flersprogede chatbot ChatPal, som blev udviklet specifikt til at fremme mental trivsel blandt personer, der bor i tyndt befolkede områder, hvor traditionelle tjenester er begrænsede på grund af den geografiske afstand.⁴
- **Potentiale til at reducere administrative omkostninger og afhjælpe mangel på arbejdskraft:** AI kan hjælpe psykiatrien, f.eks. med administrative opgaver som planlægning af aftaler på , styring af brugerflowet og generering af rapporter, hvilket frigør tid til at fokusere på at yde direkte støtte.⁵

Fra et mere klinisk perspektiv omfatter de muligheder, der diskuteres i forskningen:

- **Personliggørelse af behandling:** AI-systemer kan analysere store mængder data for at identificere mønstre og forudsige, hvordan en person vil reagere på behandlingen, og dermed hjælpe med at skabe personlige løsninger, der er skræddersyet til individuelle behov.⁶
- **Rettidig støtte:** AI-systemer kan hjælpe med at forudse, hvornår en person kommer til at opleve kriser, hvilket potentielt kan lette rettidig støtte og forhindre eskalering. Dette mål ser ud til at kræve store mængder personlig information genereret af wearables, elektroniske optegnelser, aktivitet på sociale medier og så videre.⁷
- **Forbedret diagnostisk nøjagtighed:** AI-systemer er blevet brugt til at forbedre den diagnostiske nøjagtighed inden for nogle medicinske områder, f.eks. mammografiscreening.⁸ Nogle fortalere mener, at dette kan være muligt inden for mental sundhed ved at analysere komplekse datasæt og identificere subtile mønstre, der kan hjælpe med at stille diagnoser, som mennesker måske overser.⁹
- **Udvikling af nye behandlingsmetoder:** Kunstig intelligens kan fremme udviklingen af nye terapeutiske værktøjer og interventioner, f.eks. virtual reality-behandlinger. Disse teknikker kan tilbyde alternative støttemuligheder som supplement til, forbedring af eller erstatning for traditionel støtte.¹⁰
- **Støtte til forskning og vidensgenerering:** AI kan fremskynde forskning i mental sundhed ved at analysere store datasæt og identificere mønstre, f.eks. risikofaktorer, behandlingsmål, mønstre for brug af tjenester for bestemte socioøkonomiske grupper.¹¹

Nogle har valgt en mere social tilgang til krisestøtte i deres søgen efter muligheder for AI-systemer, og de mener, at et snævert fokus på diagnose og opsporing kan være misforstået. Denne sociale tilgang kan karakterisere en vigtig mulighed som **netværkssamarbejde**, hvor AI-systemer bruges til at forbedre sociale forbindelser, peer-støtte eller til uddannelse og supervision af psykiatriske behandlere.¹² Eksempler omfatter udvikling af bedre værktøjer til at hjælpe tjenesteudbydere med at forbedre personalets færdigheder og empatiske forståelse; netværksbaserede, interaktive medier til at engagere støttearbejdere i deres supervision og hjælpe med at forbedre sig kollektivt; og teknologier, der gør det lettere for tjenestebrugere at komme i kontakt med jævnaldrende og lokale eller online fællesskaber. Denne sociale orientering trækker ofte på arbejde udført af mennesker med erfaring med psykiske problemer og psykosociale handicap.

Disse muligheder præsenteres med en vis forsigtig optimisme af de fleste kommentatorer, og der kan være en vis uenighed om, præcis hvor mulighederne ligger, hvilket er forventeligt i en periode med eksperimenter

Risici

Vi vil nu præsentere de risici, der er nævnt i litteraturen. De medfører skade på individuelt såvel som kollektivt og samfundsmæssigt niveau¹³. De bringer sundhedsvæsenets kerneværdier i fare, såsom autonomi, værdighed og tillid.¹⁴

Sikkerhedsrisici

Sundhedsrelaterede risici kan opstå på grund af fejl eller misinformation, der genereres af eller forværres af AI-systemer, samt på grund af AI-systemers manglende evne til at forstå forskellige kontekstuelle faktorer. En væsentlig bekymring er **bias i træningsdata**. Hvis disse data er ufuldstændige, ikke-repræsentative eller afspejler eksisterende samfundsmæssige fordomme, kan AI'en videreføre disse fordomme og føre til unøjagtige eller diskriminerende diagnoser.¹⁵ Hvis data, der bruges til at træne AI'en, f.eks. underrepræsenterer visse etniske grupper eller socioøkonomiske baggrunde, kan AI'en fejlfortolke symptomer eller adfærd, der er almindelig inden for disse grupper.

Et andet problem er **potentialet for falske positive og negative**. AI kan producere falske positive, der foreslår en bestemt reaktion, som f.eks. at markere en selvmordsrisiko, når den ikke er til stede, eller falske negative, som f.eks. ikke at identificere en væsentlig risiko eller krise.¹⁶ Disse fejl har betydelige konsekvenser og kan potentielt føre til unødvendige eller utilstrækkelige reaktioner.

Udbredelsen af offentligt tilgængelige AI-værktøjer, som f.eks. chatbots, medfører yderligere risici. Disse værktøjer kan give **skadelig rådgivning eller misinformation**, især når de bruges til mental sundhedsstøtte.¹⁷ De kan fejlfortolke brugerinput, tilbyde unøjagtige oplysninger eller give upassende svar, hvilket potentielt kan skabe en falsk følelse af sikkerhed for personer, der søger hjælp.

Endelig **udgør den indbyggede kompleksitet i mental sundhed en betydelig udfordring for AI-systemer**. Psykiske problemer manifesterer sig ofte unikt hos enkeltpersoner, formet af subjektive oplevelser, kulturelle sammenhænge, miljøfaktorer og personlige historier.¹⁸ Disse nuancer er vanskelige, hvis ikke umulige, at reducere til beregningsmodeller, hvilket gør det udfordrende for AI at fortolke individuelle behov nøjagtigt eller yde pålidelig støtte. Denne begrænsning er især problematisk ved kritiske hændelser, hvor nøjagtig forståelse og passende reaktioner er afgørende

Privatlivets fred

De risici for privatlivets fred, der er forbundet med at bruge AI i psykiatrien, stammer fra **den følsomme karakter af psykiatrirelaterede data** og potentialet for skade, hvis disse oplysninger håndteres forkert eller misbruges

Den stigende afhængighed af digitale platforme og sammenkoblede systemer i sundheds- og socialektoren har øget risikoen for **databrud og uautoriseret adgang**,¹⁹ som f.eks. den storstilede hacking af psykoterapeutiske journaler, der påvirkede op til 30.000 mennesker i Finland²⁰. Dette er især bekymrende for AI-systemer, der indsamler og analyserer store mængder følsomme persondata, herunder lægejournaler, terapinotater og endda aktivitet på sociale medier.

Et andet vigtigt spørgsmål er **deling af data med tredjeparter**. Der er stigende bekymring for, at AI-systemer videregiver personlige oplysninger til enheder som forsikringselskaber, arbejdsgivere eller marketingfirmaer uden udtrykkeligt samtykke. En sådan praksis kan resultere i diskrimination, underminere tilliden til mentale sundhedstjenester og i tilfælde, hvor der ikke er indhentet korrekt informeret samtykke, udgøre vildledende eller bedragerisk adfærd.

Derudover udgør **manglen på gennemsigtighed og kontrol** over, hvordan AI-systemer håndterer personlige data, en alvorlig udfordring. Folk er ofte uvidende om eller ude af stand til at kontrollere, hvordan deres oplysninger indsamles, bruges eller deles. Denne mangel på indflydelse kan øge bekymringen for privatlivets fred, afskrække folk fra at søge støtte og afholde dem fra at videregive følsomme oplysninger, der er nødvendige for effektiv støtte.

Manglende informeret samtykke

Informeret samtykke er et grundlæggende princip for beskyttelse af menneskerettigheder i sundhedssektoren.²¹ AI-systemer udgør en risiko for **utilstrækkeligt eller kompromitteret informeret samtykke**, når de bruges i psykiatrien (og sundhedsvæsenet generelt).

Det kan være svært for folk, herunder brugere og behandlere, **fuldt ud at forstå, hvordan AI-systemer fungerer**, hvilken type data de indsamler, og hvad der kan ske med disse data. Denne manglende forståelse gør det vanskeligt at give et reelt informeret samtykke, da folk måske ikke er klar over, hvad de siger ja til. Nogle etikere har foreslået, at generativ AI i teorien kan lette informeret samtykke, hvis den kan give oplysninger, der i det mindste er mere nøjagtige, tilgængelige og troværdige end dem, der tilbydes af behandlere inden for mental sundhed.²² Samtykke kan blive kompromitteret, når data indsamlet af AI-systemer også kan bruges til **andre formål end dem, der oprindeligt blev givet samtykke** til, herunder sekundær brug af offentlige afdelinger eller kommerciel datamæglervirksomhed²³

Mangel på gennemsigtig rapportering om AI-modeller underminerer deres reproducerbarhed og forhindrer identifikation af potentielle skævheder eller fejl.²⁴

Nye eller forstærkede uligheder

Bekymringer om potentialet for **bias og diskrimination** ved brug af AI i mental sundhed omfatter brugen af **biased datasæt**. AI-algoritmer lærer af de data, de er trænet på. Hvis disse datasæt indeholder fordomme eller diskriminerende indhold, vil AI-systemerne sandsynligvis arve og reproducere dem i deres forudsigelser og anbefalinger. Selv om dette kan være tilfældet med

eksisterende mentale sundhedstjenester, hvor der findes kendte skævheder og uligheder, såsom at visse grupper behandles mindre gunstigt end andre, kan AI-systemer forværre eksisterende uligheder og videreføre dem i større skala. De tidligere nævnte problemer med gennemsigtighed kan også gøre det vanskeligt at identificere og håndtere bias i AI-systemer.²⁵

AI-systemer kan også skabe nye former for diskriminerende indhold, f.eks. AI-genererede billeder, der forstærker skadelige stereotyper. En undersøgelse viste f.eks., at når man brugte ordet skizofreni til at generere billeder, frembragte det billeder med "groteske, unaturlige ansigtstræk ... blod og udtryk for rædsel".²⁶ På samme måde viste en forskningsgruppe hos Google, at sociale holdninger til mennesker med psykiske lidelser, der beskrev dem som dårlige og endda voldelige, blev indkodet i AI-systemer, der var designet til at opdage hadefulde tale i skreven tekst.²⁷

Datasæt, der bruges til at træne AI-modeller, mangler ofte mangfoldighed og repræsenterer ikke visse befolkningsgrupper tilstrækkeligt. Det kan føre til AI-systemer, der er mindre præcise eller effektive for grupper, der lever med psykiske problemer, og som fejlfortolker individuelle eller samfundsmæssige erfaringer, hvilket potentielt kan føre til fejldiagnoser og uhensigtsmæssig eller utilstrækkelig pleje. Denne mulighed forværres også af den generelle mangel på inddragelse af mennesker med levede erfaringer i skabelsen, designet, udviklingen og styringen af teknologier, der angiveligt er designet til at gavne dem.

I nogle tilfælde, selv når datasættene er repræsentative og af høj kvalitet, vil dette være utilstrækkeligt til at løse **diskrimination, der opretholdes af menneskelige systemer og institutioner, der bruger AI-systemer**, såsom forsikringsbaseret diskrimination, hvor forsikringspræmier hæves, hvis en person forudsiges at have et mentalt helbredsproblem.

Ud over bias kan AI-værktøjer styrke uligheder, hvis de **ikke er tilgængelige** for visse grupper (f.eks. mennesker med handicap eller lav digital kompetence).

Depersonalisering af plejen

AI-systemer som chatbots mangler empati, et element, der generelt anses for at være afgørende for at opbygge tillid og terapeutiske relationer, selv om de måske kan efterligne det. Robotter, chatbots og animationer kan simulere følelser som trivsel, empati og nysgerrighed, men det er blot programmerede reaktioner, og de kan forlede folk til at tillægge maskiner menneskelignende empati.

AI-systemer har betydelige begrænsninger, når de skal reagere på menneskelige følelser. AI-systemer er værktøjer. Selv om de er dygtige til at behandle data, der kan hjælpe enkeltpersoner eller tjenesteudbydere med at belyse aspekter af en persons oplevelse, kan de ikke forstå den nuancerede kompleksitet i menneskelige følelser. Det kan føre til ufølsomme eller uhensigtsmæssige reaktioner, især når det drejer sig om udtryk for nød eller sårbarhed,²⁸ og kan fremmedgøre brugerne.

Der er også en risiko for, at brugen af kunstig intelligens i psykiatrien kan føre til en **depersonalisering af plejen**. Ved at reducere den menneskelige interaktion og forbindelse kan AI prioritere bekvemmelighed og effektivitet på bekostning af den personlige, empatiske støtte, som mange mennesker i krise har brug for.

Desuden kan overdreven afhængighed af AI **potentielt underminere mentalhygiejnikeres empati**. Hvis AI-systemer bruges til at overvåge brugere eller erstatte direkte interaktion, kan psykiatriske behandlere få færre muligheder for at øve og opdyrke kommunikationsevner og blive mindre opmærksomme på, hvad der er brug for i plejemøder.²⁹

Overvågning

Nogle kommentatorer har udtrykt bekymring over det potentielle misbrug af AI til overvågning og kontrol, især når der er tale om data om mental sundhed. Et område, der vækker bekymring, er brugen af AI-systemer **til at forudsige selvmord og selvskade**. Mens sådanne teknologier ofte promoveres som værktøjer til forebyggelse og indgriben, er de forbundet med betydelige risici, herunder krænkelse af privatlivets fred, unøjagtigheder i forudsigelser og muligheden for uberettigede indgreb. Disse indgreb kan omfatte ufrivillig behandling eller unødvendig inddragelse af retshåndhævende myndigheder.³⁰

Biometriske overvågningsteknologier udgør også en overvågningsrisiko. Ved at indsamle og analysere fysiologiske data som ansigtsudtryk, stemmemønstre, gangart og øjenbevægelser kan disse systemer forsøge at udlede mentale tilstande og adfærd. Sådanne anvendelser slører grænsen mellem overvågning og invasiv overvågning

Et andet spørgsmål er **deling af følsomme data om mental sundhed med retshåndhævende myndigheder og regeringsorganer**, såsom ikke-kriminelle data om selvmordsforsøg, der bruges af retshåndhævende myndigheder eller grænsemyndigheder³¹

Denne praksis risikerer at skabe "et marked for overvågning i den mentale sundhedskontekst, der fastholder og endda udvider de værste magtubalancer, uligheder og skader i den nuværende mentale sundhedspraksis"³².

Forstærker det individualistiske syn på mental sundhed

De fleste AI-drevne teknologier inden for mental sundhed ser ud til at være rettet mod opsporing og diagnosticering, hvilket sætter fokus på den person, der er identificeret som havende brug for ekspertintervention³³. Det kan føre til en overbetoning af biologiske determinanter og et uretfærdigt fokus på det individuelle ansvar for mental sundhed³⁴. Denne indramning forstærker det individualistiske syn på mental sundhed og usynliggør de bredere socioøkonomiske, relationelle og miljømæssige faktorer, som - i samspil med personlige sundhedsproblemer - former folks mentale sundhed.

Omdirigering af begrænsede ressourcer

AI fremmes ofte af et meget aktivt marked af firmaer, der forsøger at sælge værktøjer til regeringer og serviceudbydere. Disse virksomheder kan være meget strategiske, retorisk sofistikerede og velorganiserede i forhold til at påvirke den politiske beslutningsproces, herunder indkøb.³⁵

AI-markedsføring har inkluderet besynderlige påstande om at bruge AI til at løse de langvarige og komplekse problemer med mental sundhed, såsom påstande om, at "AI-hjernechips ... kunne "løse" autisme og skizofreni".³⁶ En sådan AI-hype, især når evidensen overdrives, kan ændre, hvordan finansieringen styres, og trække ressourcer væk fra der, hvor der er mest brug for dem.

Påstande om, at kunstig intelligens og andre digitale teknologier nødvendigvis vil øge effektiviteten inden for mental sundhedspleje, understøttes ofte ikke af beviser, selv når teknologierne præsenteres

som en lavere serviceomkostning i forhold til støtte ansigt til ansigt.³⁷ Interesseorganisationen *Privacy International* har ligeledes hævdet, at der ikke er meget dokumentation for, at AI nødvendigvis vil føre til mere effektive sundhedssystemer, på trods af en udbredt antagelse - støttet af teknologileverandører - om, at dette vil være tilfældet.³⁸

Selv hvis det viser sig, at AI-systemer øger effektiviteten, bør effektivitet ikke være det vigtigste eller eneste mål, da andre mål kan være lige så vigtige eller endda vigtigere, f.eks. at skabe omsorgsfulde systemer.

Konklusioner og anbefalinger

De kilder, der er undersøgt i forbindelse med denne rapport, anerkender behovet for solid forskning for at understøtte påstande om AI-systemer og for at håndtere potentielle risici og udfordringer. De understreger behovet for etiske overvejelser, omhyggelig implementering og former for styring og tilsyn, der involverer mennesker med erfaring med psykiske problemer, for at maksimere fordelene og samtidig minimere skaderne.

Når man regulerer brugen af kunstig intelligens i psykiatrien, er det afgørende at reflektere over menneskelige relationer. "Et almindeligt mål med kunstig intelligens er at nedbryde opgaver til individuelle komponenter, som kan udføres gentagne gange. Men pleje er ikke kun opgaver, det er også følelser; det er en grundlæggende del af menneskelige relationer, og det er en meget kompleks social interaktion".³⁹

Det er blevet hævdet, at EU's nuværende lovgivningsmæssige rammer næppe tager højde for AI's unikke indvirkning på menneskelige interaktioner og følelser, som er en integreret del af mental sundhedspleje⁴⁰. Dette tilsyn ville styrke den begrænsede ansvarlighed og det begrænsede ansvar, som AI-udviklende virksomheder har inden for mental sundhed⁴¹. Det er også afgørende at tage fat på virksomhedernes interessekonflikt og sætte mental sundhedspleje af god kvalitet i centrum i stedet for profitovervejelser.

Hvilken rolle spiller levede erfaringer i af AI-teknologier?

At inddrage mennesker med egen erfaring i udviklingen af AI-systemer til mental sundhed er afgørende for at skabe teknologier, der er relevante, effektive og etiske. Dette perspektiv kan hjælpe med at identificere prioriteter, der giver genlyd hos brugerne, såsom privatlivets fred, retfærdighed og menneskelig forbindelse.⁴² At inddrage disse stemmer hjælper også med at tackle fordomme og diskrimination, magtubalancer og tilgængelighedsbarrierer, hvilket forbedrer sandsynligheden for, at AI-systemer er effektive, inkluderende og troværdige.

Nogle forskere er fortalere for at inddrage personer med levede erfaringer i hele forsknings- og udviklingsprocessen, fra definition af problemer til formidling af viden og løbende styring af systemerne.⁴³ Anbefalinger til integration af perspektiver fra levede erfaringer omfatter co-design af AI-systemer, reflekterende praksis, gennemsigtig beslutningstagning, prioritering af deres lederskab, sikring af mangfoldighed og rimelig compensation for bidrag. Den aktive inddragelse af dem, der er mest påvirket af algoritmiske og datadrevne teknologier, bør ikke blot ses som et nødvendigt skridt i "interessentinddragelsen", men snarere som en etisk nødvendighed. Når udviklere, offentlige

indkøbsmyndigheder, psykiatriske tjenester og andre inddrager ekspertise i levede erfaringer på disse måder, er der større sandsynlighed for, at AI-systemer afspejler behov i den virkelige verden og fremmer ikke bare accept, men ægte tillid til og nytteværdi af AI-systemer.

Anbefalinger til politiske beslutningstagere

For at imødegå **sikkerhedsrisikoen**

- ✓ Sørg for, at **tjenesterne er af høj kvalitet ved** at integrere AI med en høj standard for videnskabelig integritet, herunder robust validering før udbredt brug.
- ✓ Etablering af et uafhængigt tilsynsorgan til overvågning og behandling af klager over misbrug af kunstig intelligens i psykiatrien (**ansvarlighed**).
- ✓ **Indfør obligatorisk rapportering af bivirkninger** og kræv, at AI-udviklere og sundhedsudbydere rapporterer og sporer fejl, bias og utilsigtede skader forårsaget af AI-systemer.

For at imødegå **risikoen for privatlivets fred**

- ✓ Håndhæv robuste databeskyttelsesforanstaltninger, herunder klare samtykkeprotokoller og lovgivningsmæssige rammer, der prioriterer individuelle og kollektive rettigheder til privatlivets fred.
- ✓ **Forbyd uautoriseret datadeling** af mentale sundhedsdata, og forhindr AI-systemer i at overføre mentale sundhedsdata til tredjeparter (f.eks. forsikringsselskaber, arbejdsgivere, annoncører) uden udtrykkeligt, informeret samtykke.

At afhjælpe **manglen på passende informeret samtykke**

- ✓ Sikre **gennemsigtighed og forklaring** af AI, så folk kan få adgang til klare oplysninger om, hvordan disse systemer fungerer, hvilke data de bruger, hvordan de træffer beslutninger, og hvem disse data kan deles med

At imødegå risikoen for **at skabe nye uligheder eller forstærke dem**

- ✓ Gør det obligatorisk **at inkludere forskellige befolkningsgrupper i AI-træningsdatasæt** for at reducere bias og forbedre nøjagtigheden for alle samfund.
- ✓ Sørg for **aktiv inddragelse af mennesker fra samfund i sårbare og/eller marginaliserede situationer mennesker med levede erfaringer** i design, udvikling og afprøvning af AI-teknologier for at reducere fordomme.
- ✓ Tilskud til adgang til validerede AI-værktøjer til mental sundhed for underprivilegerede samfund for at sikre **digital inklusion**.
- ✓ Udvikle og håndhæve **retningslinjer for tilgængelighed** for AI-værktøjer for at sikre brugervenlighed for personer med handicap, sprogbarrierer eller lave digitale færdigheder.

For at imødegå **afpersonalisering af plejerisiko**

- ✓ Kræv **menneskeligt tilsyn** med AI-drevet mental sundhedspleje for at sikre, at empati forbliver centralt i behandlingen.

For at imødegå **overvågningsrisikoen**

- ✓ Kræv, at AI-udviklere informerer enkeltpersoner om deres interaktion med AI og sikrer en **opt-in-tilgang** eller - hvis en sådan tilgang ikke er mulig - i det mindste sikrer muligheden for at fravælge AI-systemer.
- ✓ **Reguler retshåndhævelsens brug af data om mental sundhed** og begræns delingen af mental sundhedsrelateret AI-indsigt med politi, immigrationsmyndigheder og andre offentlige instanser.

For at imødegå **risikoen for at forstærke individualistiske opfattelser af mental sundhed**

- ✓ **Modstå fristelsen til "techno-solutionism"**, troen på, at alle problemer kan løses med en nem teknologisk løsning. Der bør investeres ressourcer og kræfter i at tage fat på de bredere determinanter, der former mental sundhed, og flytte fokus fra individet til den bredere kontekst omkring ham.

For at imødegå **risikoen for at omdirigere**

- ✓ Kræv, at AI-udviklere prioriterer **plejekvalitet frem for profit og** beviser, at AI opfylder et reelt behov og ikke skaber skade. Offentlig finansiering kan bruges som løftestang i denne retning. Samskabelse kan fungere som et værktøj til at sikre, at AI-udviklinger opfylder reelle behov.

Anbefalinger til civilsamfundet

- ✓ **Overvåg AI i psykiatrien:** Spor udbredelsen af AI-værktøjer i psykiatrien, dokumenter tilfælde af bias, krænkelse af privatlivets fred eller skade, og advokér for politiske indgreb.
- ✓ **Centrer levet erfaring i AI-udvikling:** Facilitere strukturerede diskussioner og konsultationer med personer, der har erfaring med psykiske problemer, for at afgøre, hvilke AI-applikationer der er gavnlige, hvilke der er skadelige, og hvordan AI kan designes til at understøtte snarere end underminere mental sundhed.
- ✓ **Øg bevidstheden om AI-risici og -rettigheder:** Iværksæt offentlige kampagner i samarbejde med organisationer for mental sundhed og digitale rettigheder for at fremhæve risiciene ved AI i mental sundhedspleje og understrege behovet for menneskerettighedsorienteret regulering.
- ✓ **Engager dig i politiske og lovgivningsmæssige processer:** Deltag aktivt i høringer, ekspertpaneler og politiske tiltag for at sikre, at AI-regler prioriterer mentale sundhedsrettigheder, tilgængelighed og etiske standarder og forhindrer, at AI forstærker stigmatisering eller begrænser adgangen til behandling.

Bibliografi

¹ Hannah van Kolfshoeten og 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 Kolfshoeten og Janneke van Oirschot, *When People Become Data Points: The Potential Impact of AI in Mental Healthcare* (Health Action International, december 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, 'A Multilingual Digital Mental Health and Well-Being Chatbot (ChatPal): Pre-Post Multicenter Intervention Study' (2023) 25(1) *Journal of Medical Internet Research* e43051.

⁵ Verdenssundhedsorganisationen, 'Kunstig intelligens i mental sundhedsforskning: New WHO Study on Applications and Challenges' (6. februar 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 og Avinash De Sousa, 'Artificial Intelligence in Positive Mental Health: A Narrative Review' (2024) 6 *Frontiers in Digital Health* 1280235.

⁸ Se f.eks. 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 og Adam Rodman, 'Generativ kunstig intelligens i psykiatrien: 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 og De Sousa (nr. 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, "Modige nye apps: The Arrival of Surveillance Psychiatry", *Mad In America* (9. august 2019) <<https://www.madinamerica.com/2019/08/brave-new-apps-the-arrival-of-surveillance-psychiatry/>>.

¹³ van Kolfshoeten & van Oirschot (n 2).

¹⁴ Hannah van Kolfshoeten, "EU-regulering af kunstig intelligens: Challenges for patients' rights", (2022), 59, *Common Market Law Review*, Issue 1, s. 81-112.

¹⁵ Rhonda L Wilson et al, 'Kunstig intelligens: An Eye Cast towards the Mental Health Nursing Horizon' 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 og Rodman (n 9).

¹⁸ Tornero-Costa et al. (n 11).

¹⁹ Thakkar, Gupta og De Sousa (nr. 7).

²⁰ Erstatning til Vastaamo-ofre er usikker. (2021, 20. juni). *Yle Uutiset*. https://yle.fi/uutiset/osasto/news/compensation_uncertain_for_vastaamo_victims/11991155

²¹ Se fx Den Europæiske Menneskerettighedskonvention, artikel 3 og 8; Konventionen om rettigheder for personer med handicap, artikel 3, 12 og 25.

²² Blease og Rodman (n 9).

²³ Se fx Federal Trade Commission (USA), 'FTC to Ban BetterHelp from Revealing Consumers' Data, Including Sensitive Mental Health Information, to Facebook and Others for Targeted Advertising' (2. marts 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 Verdenssundhedsorganisationen (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), July-December 2023, lsad031, <https://doi.org/10.1093/jlb/lsad031>.

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 Chatbots Fail to Spot Sexual Abuse", *BBC News* (online, 11. december 2018) <<https://www.bbc.com/news/technology-46507900>>.

²⁹ Se fx 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 og Rodman (nr. 9).

³⁰ Marks (nr. 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. september 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* (rapport, University of Melbourne, 1. september 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' i Lieketseng Ned et al (eds), *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, 'Digitale interventioner til screening og behandling af almindelige psykiske lidelser eller symptomer på almindelig psykisk sygdom hos voksne: 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. juli 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, 'Regulering af AI inden for mental sundhed: Ethics of Care Perspective' (2024) 11(1) *JMIR Mental Health* e58493.

⁴¹ Ibid.

⁴² Ian Muchamore, Panos Karanikolas og 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) (online først) *Australian Journal of Social Issues* <<https://onlinelibrary.wiley.com/doi/abs/10.1002/ajs4.355>>.



Brugen af disse materialer skal omfatte en kildeangivelse. Ingen af disse materialer må ændres eller tilpasses uden udtrykkelig skriftlig tilladelse fra forfatterne.

Mental Health Europe (december 2024)

De synspunkter og meninger, der kommer til udtryk, er udelukkende forfatterens/forfatternes og afspejler ikke nødvendigvis Den Europæiske Unions eller Europa-Kommissionens CERV-programms synspunkter. Hverken EU eller den bevilgende myndighed kan holdes ansvarlig for dem.



Co-funded by
the European Union