

فراخوان هفدهمین کنگره‌پیشگامان پیشرفت

فناورک‌ها و نووظهور

پیامدهای روان‌شناختی و اجتماعی هوش مصنوعی در مسیر پیشرفت ایران: دلالت‌هایی

برای نظام مشاوره و سلامت روان

احمد عباسی دره بیدی

دکتری مشاوره، نویسنده مسئول: A.abbasi.d2020@gmail.com

چکیده

گسترش کاربردهای هوش مصنوعی در حوزه سلامت روان، تحولی بنیادین در شیوه طراحی و ارائه مداخلات روان‌شناختی ایجاد کرده است؛ با این حال، پیامدهای روان‌شناختی، اخلاقی، اجتماعی و نهادی این فناوری همچنان با ابهام‌های مفهومی و تحلیلی همراه است. هدف پژوهش حاضر، تبیین نظام‌مند پیامدهای به‌کارگیری هوش مصنوعی در سلامت روان و ارائه مدلی مفهومی برای فهم تعامل انسان و فناوری در این حوزه است. پژوهش با رویکرد کیفی و ماهیت تحلیلی–تبیینی انجام شد و از رهیافت تحلیل مضمون بازتابی براون و کلارک (۲۰۰۶) بهره گرفت. داده‌ها شامل مقالات علمی، مرورهای نظام‌مند و اسناد پژوهشی معتبر منتشرشده در فاصله سال‌های ۲۰۱۸ تا ۲۰۲۴ در پایگاه‌های علمی بین‌المللی بود که به‌صورت هدفمند و نظری تا اشباع مفهومی انتخاب شدند. واحد تحلیل، مفاهیم و گزاره‌های استخراج‌شده از متون بود که طی فرایند شش‌گانه تحلیل مضمون، کدگذاری، پالایش و تفسیر شدند. یافته‌ها به شناسایی سه مضمون اصلی و نه مضمون فرعی انجامید که پیامدهای هوش مصنوعی در سطوح فناوریانه، روان‌شناختی، اخلاقی–حرفه‌ای، فرهنگی–اجتماعی و نهادی را تبیین می‌کنند. بر اساس این یافته‌ها، «تعادل پویا میان کارآمدی فناوریانه، عمق درمانی و اخلاق انسانی» به‌عنوان پدیده مرکزی استخراج شد. نتایج نشان می‌دهد که اثربخشی و مشروعیت مداخلات مبتنی بر هوش مصنوعی، وابسته به تنظیم هم‌زمان این سطوح و پرهیز از تقلیل سلامت روان به فرایندهای صرفاً الگوریتمی است. این پژوهش با ارائه چارچوبی یکپارچه، می‌تواند مبنایی برای توسعه پژوهش‌های آینده، آموزش حرفه‌ای و سیاست‌گذاری مسئولانه در حوزه سلامت روان دیجیتال فراهم آورد.

کلیدواژگان: **هوش مصنوعی، سلامت روان، تعامل انسان–فناوری، اتحاد درمانی ترکیبی، اخلاق هوش مصنوعی**

مقدمه

هوش مصنوعی در سال‌های اخیر فراثر از یک ابزار فناوریانه، به عاملی ساختاری در بازتعریف تجربه‌های روان‌شناختی و روابط اجتماعی انسان تبدیل شده است. نفوذ این فناوری در حوزه‌هایی چون آموزش، ارتباطات و به‌ویژه خدمات سلامت، الگوهای تعامل، تصمیم‌گیری و تنظیم هیجانی را دگرگون کرده و پیامدهایی مستقیم برای سلامت روان فردی و جمعی به همراه داشته است.

در حوزه سلامت روان، اگرچه هوش مصنوعی امکان دسترسی گسترده‌تر به خدمات مشاوره‌ای و درمانی را فراهم کرده است، اما هم‌زمان پرسش‌هایی جدی درباره کیفیت رابطه درمانی، حدود مسئولیت حرفه‌ای و اصالت مداخله روان‌شناختی مطرح شده است. در ایران، مسیر پیشرفت فناوریانه در بستری دنبال می‌شود که با افزایش شیوع آسیب‌های روانی، فرسودگی هیجانی و کاهش تاب‌آوری اجتماعی همراه است.

شواهد داخلی نشان می‌دهد نظام سلامت روان کشور، علی‌رغم برخورداری از نیروی انسانی متخصص، در مواجهه با فناوری‌های نوین از جمله هوش مصنوعی با چالش‌های نهادی، اخلاقی و سیاست‌گذاری روبه‌روست. برای نمونه، پژوهش صالحی‌نیا و همکاران (۱۴۰۳) فقدان چارچوب‌های اخلاقی شفاف، ضعف آمادگی سازمانی نبود سیاست‌گذاری مبتنی بر شواهد را از موانع اصلی ادغام مسئولانه هوش مصنوعی در خدمات سلامت روان ایران معرفی می‌کند.

در سطح بین‌المللی نیز پژوهش‌ها هشدار می‌دهند که به‌کارگیری هوش مصنوعی در مداخلات سلامت روان بدون توجه نظام‌مند به زمینه‌های فرهنگی و معنایی کاربران، می‌تواند به کاهش اثربخشی درمان و تضعیف رابطه درمانی بینجامد. مطالعات اخیر همچنین نشان می‌دهد فقدان چارچوب‌های حرفه‌ای و بومی‌شده، خطر بی‌اعتمادی مراجعان و مخدوش شدن مرزهای اخلاقی درمان را افزایش می‌دهد (چکرود و همکاران، ۲۰۲۴؛ ارمان و فرندال، ۲۰۲۴).

بر این اساس، نظام مشاوره و سلامت روان ایران با یک چالش دو وجهی مواجه است: از یک‌سو ضرورت بهره‌گیری آگاهانه از ظرفیت‌های هوش مصنوعی برای پاسخ به نیازهای فراینده جامعه، و از سوی دیگر خطر تقلیل فرایندهای عمیق انسانی مشاوره به تعاملات الگوریتمی.

مسئله محوری پژوهش حاضر آن است که پیامدهای روان‌شناختی و اجتماعی هوش مصنوعی در مسیر پیشرفت ایران چیست و این پیامدها چه دلالت‌هایی برای بازاندیشی در ساختار، کارکرد و آمادگی نظام مشاوره و سلامت روان کشور به همراه دارد؟ پرداختن به این مسئله، گامی ضروری برای طراحی سیاست‌های بومی و تحقق پیشرفت انسانی در عصر هوش مصنوعی محسوب می‌شود.

روش تحقیق

پژوهش حاضر از نظر ماهیت، کیفی و از حیث هدف، تحلیلی–تبیینی است و با بهره‌گیری از رهیافت تحلیل مضمون بازتابی به بررسی پیامدهای روان‌شناختی و اجتماعی به‌کارگیری هوش مصنوعی در حوزه مشاوره و سلامت روان می‌پردازد.

جامعه پژوهش را مقالات علمی، مرورهای نظام‌مند و اسناد پژوهشی معتبر منتشرشده در فاصله سال‌های ۲۰۱۸ تا ۲۰۲۴ در پایگاه‌های بین‌المللی از جمله اسکوپوس، وب آف ساینس، پایمد و سایینس دایرکت تشکیل می‌دهد.

نمونه‌گیری به‌صورت هدفمند و نظری تا اشباع مفهومی انجام شد و واحد تحلیل، مفاهیم و گزاره‌های استخراج‌شده از متون بود.

داده‌ها بر اساس مراحل شش‌گانه تحلیل مضمون (براون و کلارک، ۲۰۰۶) شامل آشنایی، کدگذاری، جستجو، بازبینی، تعریف و روایت‌سازی تحلیل شدند.

تفسیر داده‌ها به شکل بازتابی و با درآمیزی چارچوب‌های نظری روان‌شناسی، اخلاق فناوری و مطالعات انسان–فناوری صورت گرفت.

برای تضمین اعتبار، از معیارهای (لینکلن و گوبا، ۱۹۸۵)، شامل باورپذیری، انتقال‌پذیری، اتکاپذیری و تأییدپذیری استفاده شد و مستندسازی کامل مسیر تحلیل، بازبینی مکرر مضامین و تطبیق با پیشینه علمی انجام گرفت.

یافته ها

یافته‌های پژوهش حاصل تحلیل مضمون بازتابی متون علمی منتخب در حوزه به‌کارگیری هوش مصنوعی در مشاوره و سلامت روان است. فرایند تحلیل به شناسایی ۳ مضمون اصلی، ۹ مضمون فرعی و ۲۴ کد مفهومی انجامید که بیانگر الگوهای معنادار، تنش‌های نظری و پیامدهای روان‌شناختی، اخلاقی و نهادی استفاده از هوش مصنوعی در نظام‌های روان‌درمانی است. در مجموع، یافته‌ها نشان می‌دهند که هوش مصنوعی به‌طور هم‌زمان واجد ظرفیت‌های ارتقادهنده خدمات سلامت روان و حامل مخاطرات عمیق در سطح تجربه درمانی، عاملیت روان‌شناختی و عدالت درمانی است؛ وضعیتی که به‌ویژه در بسترهای فرهنگی غیرغربی، مانند ایران، حساسیت بیشتری می‌یابد (براون و کلارک، ۲۰۲۱)

مضمون اول: **بازپیکربندی تجربه درمانی در عصر هوش مصنوعی**

این مضمون به تغییرات بنیادین در ماهیت تجربه مشاوره و روان‌درمانی در پی ورود سامانه‌های هوشمند اشاره دارد. یافته‌ها نشان می‌دهد که مداخلات مبتنی بر هوش مصنوعی، به‌ویژه ابزارهای مبتنی بر درمان شناختی–رفتاری، در کاهش علائم خفیف و متوسط اضطراب و افسردگی کارآمد هستند؛ بااین‌حال، در ایجاد اتحاد درمانی عمیق، معنایردازی وجودی و تحول پایدار با محدودیت مواجهاند (لی و همکاران، ۲۰۲۲)

همچنین، استانداردسازی الگوریتمی مداخلات گرچه به افزایش دسترس‌پذیری و یکنواختی خدمات می‌انجامد، اما می‌تواند به تضعیف فردیت مراجع و نادیده‌گرفتن زمینه‌های فرهنگی، ارزشی و زیسته منجر شود؛ امری که در سلامت روان یک ریسک بالینی محسوب می‌شود (چکرود، ۲۰۲۴). در این میان، نقش درمانگر انسانی نیز در حال جابه‌جایی از «کنشگر مستقیم درمانی» به «ناظر، تنظیم‌گر و مفسر تصمیمات الگوریتمی» است؛ تغییری که پیامدهای حرفه‌ای، هویتی و اخلاقی قابل‌توجهی دارد (لامبرت و نورگراس، ۲۰۱۹).

مضمون دوم: **پیامدهای روان‌شناختی و شناختی تعامل انسان–هوش مصنوعی**

این مضمون به تأثیرات تعامل مستمر با سامانه‌های هوشمند بر ساختارهای روان‌شناختی کاربران می‌پردازد. یافته‌ها نشان می‌دهد که مواجهه درمانی با هوش مصنوعی می‌تواند به اضطراب فناوریانه، احساس کاهش کنترل و ابهام وجودی منجر شود، به‌ویژه زمانی که منطق تصمیم‌گیری الگوریتم‌ها برای کاربران شفاف نیست(سانن وو هرسیمیک، ۲۰۲۰).

از سوی دیگر، هوش مصنوعی نوعی عاملیت شناختی دوباره‌شده ایجاد می‌کند که در آن تصمیم‌سازی درمانی میان انسان و الگوریتم توزیع می‌شود؛ وضعیتی که می‌تواند استقلال شناختی مراجع را بازتعریف یا در برخی موارد تضعیف کند. همچنین، یافته‌ها بر پدیده توهم همدلی الگوریتمی دلالت دارند؛ به این معنا که همدلی ارائه‌شده توسط سامانه‌های هوشمند غالباً از نوع شناختی و شبیه‌سازی‌شده است و فاقد عمق عاطفی انسانی است، هرچند در کوتاه‌مدت می‌تواند برای کاربران رضایت‌بخش به نظر برسد (لئوس و بروکتر، ۲۰۲۴).

مضمون سوم: **چالش‌های اخلاقی، فرهنگی و نهادی**

مضمون سوم بر ناترازی میان شتاب توسعه فناوریانه و ضعف زیرساخت‌های اخلاقی، حقوقی و نهادی تأکید دارد. یافته‌ها نشان می‌دهد که سوگیری داده‌ها و الگوریتم‌ها می‌تواند به بازتولید نابرابری‌های روان‌درمانی در گروه‌های مختلف اجتماعی و فرهنگی منجر شود (چکرود و همکاران، ۲۰۲۴). افزون بر این، خلأ تنظیم‌گری حرفه‌ای و مسئولیت‌پذیری حقوقی، به‌ویژه در موارد خطای تشخیصی یا مداخله نامناسب، یکی از چالش‌های اساسی نظام سلامت روان مبتنی بر هوش مصنوعی است(فلوربیدی و همکاران، ۲۰۱۸)..

آینده‌پیشرفت ایران

(چالش‌ها، فرصت‌ها، راهکارها)

مهلت دریافت آثار و ایده‌های نوآورانه:

۱۴ بهمن ۱۴۰۴ و زمان برگزاری: بهار ۱۴۰۵

نتیجه گیری و پیشنهادات

پژوهش حاضر با هدف تبیین پیامدهای روان‌شناختی، اخلاقی و اجتماعی به‌کارگیری هوش مصنوعی در حوزه سلامت روان انجام شد و با بهره‌گیری از تحلیل مضمون بازتابی، کوشید تصویری مفهومی و چندلایه از این پدیده نوظهور ارائه دهد.

یافته‌ها نشان داد که هوش مصنوعی در سلامت روان نه صرفاً یک ابزار فناوریانه، بلکه پدیده‌ای تعاملی است که در هم‌تیدگی میان سطوح فناوریانه، روان‌شناختی، اخلاقی–حرفه‌ای، فرهنگی–اجتماعی و نهادی معنا می‌یابد. این هم‌زمانی سلوح، منجر به شکل‌گیری الگوی «تعادل پویا میان کارآمدی فناوریانه، عمق درمانی و اخلاق انسانی» به‌عنوان پدیده مرکزی پژوهش شد.

در سطح فناوریانه، نتایج حاکی از آن است که سامانه‌های هوشمند، به‌ویژه چت‌بات‌های درمان‌محور، از طریق افزایش دسترس‌پذیری، کاهش هزینه و امکان پایش مستمر، ظرفیت‌های مهمی برای نظام سلامت روان ایجاد کرده‌اند؛ یافته‌ای که با نتایج فیتزپاتریک و همکاران (۲۰۱۷) و لی و همکاران (۲۰۲۲) هم‌خوان است.

با این حال، پژوهش حاضر نشان می‌دهد که همین کارآمدی می‌تواند به استانداردسازی افراطی مداخلات و کم‌رنگ شدن تفاوت‌های فردی و زمینه‌ای مراجعان بینجامد؛ امری که در سلامت روان، پیامدهای بالینی و اخلاقی قابل‌توجهی دارد.

در سطح روان‌شناختی، یافته‌ها بر ماهیت دوگانه تجربه درمانی مبتنی بر هوش مصنوعی دلالت دارند. پاسخ‌های سریع و ساختاریافته الگوریتم‌ها می‌تواند در کوتاه‌مدت احساس حمایت و نظم شناختی ایجاد کند، اما در بلندمدت، فقدان تجربه زیسته، آگاهی اخلاقی و حضور انسانی، به شکل‌گیری آنچه «توهم همدلی الگوریتمی» نام‌گذاری شد، منجر می‌شود. این نتیجه با هشدارهای شام و همکاران (۲۰۱۸) و تحلیل لئوس و بروکتر (۲۰۲۴) هم‌راساست و نشان می‌دهد که همدلی شبیه‌سازی‌شده، جایگزین همدلی عاطفی و رابطه‌محور انسانی نمی‌شود.

همچنین، مفهوم «عاملیت شناختی دوباره‌شده» نشان می‌دهد که توزیع تصمیم‌سازی میان انسان و الگوریتم می‌تواند استقلال روان‌شناختی مراجع را بازتعریف یا تضعیف کند؛ نکته‌ای که در ادبیات تجربی کمتر به‌صورت نظام‌مند مورد توجه قرار گرفته است.

در سطح اخلاقی–حرفه‌ای، یافته‌ها بر افزایش ابهام در مسئولیت‌پذیری درمانی و تغییر نقش درمانگر تأکید دارند. همسو با دیدگاه فلوربیدی و همکاران (۲۰۱۸)، نتایج نشان می‌دهد که اتکای فراینده به خروجی‌های الگوریتمی می‌تواند به انتقال ضمنی مسئولیت تصمیم‌گیری از انسان به سامانه هوشمند بینجامد و مرزهای اخلاقی حرفه‌ای را مخدوش سازد. از این منظر، نقش درمانگر نه حذف، بلکه به‌عنوان تنظیم‌گر، ناظر و تفسیرکننده تصمیمات الگوریتمی بازتعریف می‌شود.

در سطح فرهنگی–اجتماعی، پژوهش حاضر بر ضرورت بومی‌سازی مداخلات مبتنی بر هوش مصنوعی تأکید می‌کند. یافته‌ها نشان می‌دهد که فقدان انطباق با ارزش‌ها، هنجارها و الگوهای معنایی بستر فرهنگی ایران می‌تواند به سوء‌تعبیر نشانه‌های روانی، کاهش پذیرش اجتماعی و تضعیف اعتماد نهادی منجر شود؛ نتیجه‌ای که با دیدگاه‌های اسمعیلی و قاسمی‌نیا (۱۴۰۴) درباره زمینه‌مندی فرهنگی سلامت روان هم‌سو است.

افزون بر این، در سطح نهادی، نبود چارچوب‌های شفاف حقوقی و تنظیم‌گر، یکی از عوامل اصلی شکنندگی نظام سلامت روان مبتنی بر هوش مصنوعی شناسایی شد؛ به‌گونه‌ای که اختلال در این سطح، سایر سطوح مدل را نیز تحت تأثیر قرار می‌دهد.

از حیث نوآوری، تمایز اصلی پژوهش حاضر با مطالعات پیشین، بیانی چندسطحی پیامدهای هوش مصنوعی در نظام مشاوره و سلامت است که روابط میان فناوری، روان‌شناسی، اخلاق، فرهنگ و نهادها را به‌صورت یکپارچه تبیین می‌کند؛ در حالی که بخش عمده‌ای از پیشینه موجود، این ابعاد را به‌صورت منفک یا تک‌سطحی بررسی کرده‌اند.

با وجود این، پژوهش حاضر با محدودیت‌هایی از جمله ماهیت کیفی و تفسیری تحلیل و عدم تعمیم‌پذیری آماری یافته‌ها مواجه است.

همچنین، شتاب تحولات فناوریانه در حوزه هوش مصنوعی می‌تواند برخی مصادیق مورد بررسی را در آینده دستخوش تغییر کند.

با این حال، غنای تحلیلی داده‌ها و انسجام مفهومی مدل ارائه‌شده، قابلیت انتقال نظری مناسبی برای پژوهش‌های آتی فراهم می‌سازد.

در جمع‌بندی نهایی، می‌توان گفت هوش مصنوعی در سلامت روان نه تهدیدی مطلق و نه راه‌حلی کامل است، بلکه پدیده‌ای زمینه‌مند و تعاملی است که اثربخشی، ایمنی و مقبولیت آن منوط به حفظ تعادل میان کارآمدی فناوریانه، عمق درمانی و اخلاق انسانی است. بدین ترتیب در ادامه پیشنهادهاتی مرتبط ارائه شده است.

پیشنهادها

– توسعه مدل «اتحاد درمانی ترکیبی» (Hybrid Therapeutic Alliance)

پیشنهاد می‌شود پژوهش‌های آینده، مفهوم «اتحاد درمانی ترکیبی» را به‌عنوان یک سازه نظری مستقل توسعه دهند؛ به این معنا که رابطه درمانی نه صرفاً میان درمانگر و مراجع، بلکه در تعامل سه‌گانه «انسان–الگوریتم–مراجع» تعریف شود. در این چارچوب، ابعاد اعتماد، همدلی، ادراک حمایت و وابستگی شناختی می‌تواند به‌صورت مجزا operationalize و آزمون شود.

– نظریه‌پردازی درباره «توهم همدلی الگوریتمی»

یافته‌های پژوهش ظرفیت شکل‌دهی به یک چارچوب نظری جدید درباره «توهم همدلی الگوریتمی» را دارد. پیشنهاد می‌شود این مفهوم در پیوند با نظریه‌های دلبستگی، تعامل انسان–ایانه و روان‌شناسی اجتماعی بررسی شود تا مشخص گردد کاربران چگونه پاسخ‌های ساختاریافته و زبانی الگوریتم را به‌عنوان همدلی واقعی ادراک می‌کنند و این ادراک چه پیامدهایی بر وابستگی عاطفی و تصمیم‌گیری درمانی دارد.

– طراحی چارچوب نظری چندسطحی اخلاق هوش مصنوعی در سلامت روان

نتایج پژوهش نشان می‌دهد که اخلاق در این حوزه صرفاً یک مسئله فردی یا حرفه‌ای نیست، بلکه پدیده‌ای نهادی و فرهنگی نیز محسوب می‌شود. بنابراین پیشنهاد می‌شود مدلی نظری تدوین شود که در آن ابعاد اخلاقی در سطوح خرد (درمانگر–مراجع)، میانی (سازمان‌ها و پلتفرم‌ها) و کلان (قانون‌گذاری و فرهنگ) به‌صورت یکپارچه تحلیل شوند.

– بومی‌سازی نظریه‌های سلامت روان دیجیتال در بافت ایرانی

بخش زیادی از نظریه‌های موجود بر مبنای زمینه‌های غربی توسعه یافته‌اند. پیشنهاد می‌شود پژوهشگران ایرانی، مدل‌های بومی مبتنی بر ارزش‌های فرهنگی، الگوهای ارتباطی، معنویت و ساختار خانواده در ایران طراحی کنند تا تحلیل تعامل انسان و هوش مصنوعی در سلامت روان، با واقعیت فرهنگی جامعه ایرانی انطباق بیشتری یابد.

تدوین دستورالعمل‌های اخلاقی و حرفه‌ای برای استفاده از هوش مصنوعی

ضروری است سازمان نظام روان‌شناسی، وزارت بهداشت و مراکز مشاوره، دستورالعمل‌های شفاف درباره حدود استفاده از سامانه‌های هوشمند، مسئولیت حقوقی درمانگر، حفظ محرمانگی داده‌ها و نحوه مداخله در شرایط بحرانی تدوین کنند.

– استفاده از مدل «انسان در حلقه» (Human-in-the-loop)

در کاربردهای بالینی، تصمیم‌گیری نهایی نباید به الگوریتم واگذار شود. پیشنهاد می‌شود تمامی سامانه‌های هوش مصنوعی سلامت روان به‌گونه‌ای طراحی شوند که درمانگر انسانی در فرایند ارزیابی، تفسیر و تأیید مداخلات حضور فعال داشته باشد.

– ایجاد نهادهای ارزیابی و پایش سامانه‌های هوش مصنوعی سلامت روان

ضروری است مراکز مستقلی برای اعتبارسنجی، ارزیابی ایمنی، سنجش سوگیری فرهنگی و پایش عملکرد سامانه‌های هوش مصنوعی ایجاد شود تا کیفیت خدمات، امنیت داده‌ها و پیامدهای روان‌شناختی این فناوری‌ها به‌صورت مستمر بررسی گردد.

منابع

صالحی‌نیا، رضاه نصیری سنگری، مرضیه؛ عباسیان، حسین و صالحیان، سجاد (۱۴۰۳). بررسی ملاحظات اخلاقی در استفاده از هوش مصنوعی در حوزه سلامت: مروری روایتی. اخلاق و تاریخ پزشکی ایران. ۱۷ (۵): ۱–۲.

http://ijme.tums.ac.ir/article-1-6894-fa.html

اسمعیلی، مضمونه؛ قاسمی نیایی، فاطمه (۱۴۰۴). مطالعه مروری چالش‌ها و فرصت‌های استفاده از هوش مصنوعی در مشاوره و روان‌درمانی. مجله روانشناسی و روانپزشکی شناخت. ۱۲ (۵): ۱۶۷–۱۸۲. http://shenakht.muk.ac.ir/article-1-2424-fa.html

http://irh.mazums.ac.ir/article-1-1213-fa.html

نصیری، سمیه (۱۴۰۴). سبب محمد مهدی (۱۴۰۳). واکوابی ابعاد خطمشی به‌کارگیری هوش مصنوعی در نظام سلامت ایران. مدیریت سازمان‌های دولتی ۱۲ (شماره۲) پیاپی ۴۴)، ۱۸۱–۲۰۰. https://doi.org/10.30473/ipom.2024.71179.4977

عباسی، محمود و تیموری، مهرداد. (۱۴۰۲). مروری بر چالش‌های اخلاقی و حقوقی کاربرد هوش مصنوعی در نظام سلامت. مجله اخلاق پزشکی، ۱۷(۴۸)، ۱–۱۱. https://doi.org/10.22037/mejv.17148.44053

رضایی، فاطمه. (۱۴۰۳). آینده روان‌درمانی: بررسی ادغام و پیامدهای هوش مصنوعی در شیوه‌های درمانی معاصر. فصلنامه توسعه پژوهش‌های روان‌شناسی و مدیریت ناده. https://doi.org/10.61838/kman.pdmd.4.1.6

Bendig, E., Erb, B., et al. (2019). Chatbots in clinical psychology and psychotherapy. *Verhaltenstherapie*, 32, 64–76. https://doi.org/10.1159/000501812

Boucher, E. M., Haraake, N. R., et al. (2021). Artificially intelligent chatbots in digital mental health interventions. *Expert Review of Medical Devices*, 18(sup1), 37–49. https://doi.org/10.1080/17434440.2021.2013200

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qps063oa

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in thematic analysis. *Qualitative Research in Psychology*, 18(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic analysis. In P. Liamputtong (Ed.), *Handbook of research methods in health social sciences*. Springer. https://doi.org/10.1007/978-981-10-5251-4_103

Bullock, J., Young, M., Wang, Y., & Webb, A. (2024). *Using artificial intelligence in public governance*. Oxford University Press. https://academic.oup.com/book/52377

Chekroud, A. M., Bondar, J., Delgadillo, J., & McMillan, D. (2024). The illusory generalizability of clinical prediction models. *Science*, 384(6696), 112–118. https://doi.org/10.1126/science.adg8538

Chekroud, A. M., Bondar, J., Delgadillo, J., et al. (2021). The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry*, 20(2), 154–170. https://doi.org/10.1002/wps.20882

Chekroud, A. M., et al. (2024). The illusory generalizability of clinical prediction models. *Science*, 384(6696). https://doi.org/10.1126/science.adg8538

Chekroud, A. M., Havrilenko, M., et al. (2024). Illusory generalizability of clinical prediction models. *Science*, 383, 164–167. https://doi.org/10.1126/science.adg8538

Cognition and the web: Extended, distributed and embodied. *Synthese*, 197, 469–491. https://doi.org/10.1007/s11229-018-01964-3

Erman, S., & Furendal, M. (2024). Artificial agency and moral responsibility. *Ethics and Information Technology*, 26(2), 1–14. https://doi.org/10.1007/s10676-024-09764-9

Erman, S., & Furendal, M. (2024). Responsibility gaps in AI-assisted decision-making. *Ethics and Information Technology*. https://doi.org/10.1007/s10676-024-09702-9

Florida, L., Cowsils, J., Beltrametti, M., et al. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Heersmink, R., & Sutton, J. (2020). Cognition and the web: Extended minds and collective intelligence. *Synthese*, 197(3), 1129–1146. https://doi.org/10.1007/s11229-017-1589-8

Jobsan, A., Jenca, M., & Vayena, E. (2020). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2

Journal of Culture, Behavior and Learning. (2024). AI, society and mental wellbeing. https://journalcbl.com/index.php/cbl/article/view/101

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.

Ly, K. H., Ly, A. M., & Andersson, G. (2022). Effectiveness of artificial intelligence–based psychological interventions. *Journal of Affective Disorders*, 305, 141–150. https://doi.org/10.1016/j.jad.2022.06.073

Norcross, J. C., & Lambert, M. J. (2019). Psychotherapy relationships that work. *Oxford University Press*. https://academic.oup.com/book/32162

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *Qualitative Research in Psychology*, 14(1), 1–13. https://doi.org/10.1080/14780887.2016.1217097

Yonatan-Leus, R., & Brukner, R. (2024). Artificial empathy. *Counselling & Psychotherapy Research*. https://doi.org/10.1002/capr.12701

Yonatan-Leus, R., & Brukner, R. (2024). Human versus artificial empathy in counseling interactions. *Counselling and Psychotherapy Research*. https://doi.org/10.1002/capr.12701