

An Ethical Analysis of the Emotional Relationship between Humans and Artificial Intelligence

Saeedeh Babaii¹, Narjes Saberi², Monireh Bahreini³ 

1- Postdoctoral Researcher, International Center for Ethics in Sciences and Humanities, University of Tuebingen, Tuebingen, Germany.

2- M.A. student, Faculty of Psychology and Educational Sciences, Allameh Tabatabaei University, Tehran

3- M.A. of Philosophy of Science and Technology, Amirkabir University of Technology, Tehran, Iran.

(Corresponding Author: bahreini@aut.ac.ir)

Abstract

The interactions between humans and AI-based interactive technologies, particularly social bots, exhibit unique characteristics that distinguish them from human-to-human interactions as well as those with other types of technologies. As users in the world, and consequently in Iran increasingly engage with social bots, there is potential for the development of emotional bonds between users and these intelligent systems. Given their unique nature and the expectations shaped by corporate marketing and advertising, social bots can generate perceptions of agency and emotions, which in turn may foster one-sided emotional attachments from users. These attachments raise a range of significant ethical, social, and psychological concerns that can affect both individuals and larger communities. The present paper seeks to investigate the ethical challenges that arise from human-bot relationships, particularly in terms of the attribution of agency and the development of emotional dependency. In addition, it examines potential policies and solutions that have been proposed to mitigate these issues, focusing on strategies that can address concerns about the impact of these technologies. By critically analyzing these challenges, the study aims to propose policy recommendations that are specifically tailored to the implications of social bots use in the societal context. Ultimately, the findings highlight the need for a balanced approach that takes into account both the potential benefits and risks of such technologies, ensuring that their integration into daily life aligns with ethical standards and societal values.

Keywords: Human-AI Interaction, Social bots, Human/AI Emotional Bond, AI Ethics.

How to Cite this Paper:

Babaii, S, Saberi, N. & Bahreini, M. (2025). **An Ethical Analysis of the Emotional Relationship Between Humans and Artificial Intelligence**. *Journal of Science & Technology Policy*, 17(4), 19-30. {In Persian}.

doi: 10.22034/jstp.2025.11773.1816



تحلیلی اخلاقی از رابطه احساسی انسان و هوش مصنوعی

سعیده بابایی^۱، نرجس صابری^۲، منیره بحرینی^۳ 

۱- محقق پسادکتری، مرکز اخلاق در علم و علوم انسانی، دانشگاه توینگن، توینگن، آلمان.

۲- دانشجوی کارشناسی ارشد، دانشکده روانشناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران.

۳- کارشناس ارشد فلسفه علم و فناوری، دانشگاه صنعتی امیرکبیر، تهران. (نویسنده عهده‌دار مکاتبات: bahreini@aut.ac.ir)

چکیده

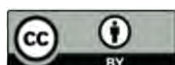
تعاملات میان انسان و فناوری‌های تعاملی مبتنی بر هوش مصنوعی، به‌ویژه بات‌های اجتماعی، ویژگی‌های منحصر به فردی دارد که آن را از تعاملات انسان با انسان و سایر فناوری‌ها متمایز می‌کند. استفاده کاربران از بات‌های اجتماعی در جهان و به تبع آن در ایران روبه گسترش است که می‌تواند به شکل‌گیری پیوندهای احساسی میان کاربران و این ابزارهای هوشمند منجر شود. به دلیل ماهیت منحصر به فرد بات‌های اجتماعی و انتظاراتی که شرکت‌ها و تبلیغات در کاربران ایجاد می‌کنند، این فناوری‌ها می‌توانند موجب شوند کاربران به آنها نوعی عاملیت و احساسات نسبت دهند و پیوندهای عاطفی یک‌سویه‌ای را با آنها برقرار کنند. این پیوندها مسائل اخلاقی، اجتماعی و روان‌شناختی قابل توجهی را برای افراد و جوامع انسانی ایجاد می‌کند. این مقاله چالش‌های اخلاقی مرتبط با روابط انسان و بات‌های اجتماعی، از جمله نسبت دادن عاملیت و وابستگی عاطفی را بررسی می‌کند. علاوه بر این، سیاست‌ها و راه‌حل‌های احتمالی که برای کاهش این مشکلات پیشنهاد شده‌اند، مورد بررسی قرار می‌گیرند و بر راهبردهایی تمرکز می‌شود که می‌توانند به نگرانی‌ها در مورد تأثیر این فناوری‌ها پاسخ دهند. با تحلیل انتقادی این چالش‌ها، این مطالعه به دنبال ارائه توصیه‌های سیاستی است که به پیامدهای استفاده از بات‌های اجتماعی در جامعه می‌پردازد. در نهایت، یافته‌ها نیاز به یک رویکرد متوازن را برجسته می‌کنند که هم مزایا و هم خطرات این فناوری‌ها را در نظر بگیرد و اطمینان حاصل کند که ادغام آنها در زندگی روزمره با استانداردهای اخلاقی و ارزش‌های اجتماعی هم‌راستا است.

کلیدواژه‌ها: تعامل انسان و هوش مصنوعی، بات اجتماعی، پیوند احساسی انسان و هوش مصنوعی، اخلاق هوش مصنوعی.

برای استنادات بعدی به این مقاله، قالب زیر به نویسندگان محترم مقالات پیشنهاد می‌شود:

بابایی، سعیده، صابری، نرجس، بحرینی، منیره. (۱۴۰۳). تحلیلی اخلاقی از رابطه احساسی انسان و هوش مصنوعی. سیاست علم و فناوری، ۱۷(۴)، ۱۹-۳۰.

doi: 10.22034/jstp.2025.11773.1816



۱- مقدمه

ادعای گزافی نیست اگر هوش مصنوعی و فناوری‌های مبتنی بر آن را از جمله داغ‌ترین و پرمناقشه‌ترین حوزه‌های فنی جهان امروز تلقی کنیم که شاهد رشد سریع و بی‌سابقه آنها هستیم. ایران نیز از این قاعده مستثنی نیست و هر فرد در طول شبانه‌روز ساعت‌های زیادی را در تعامل و استفاده از فناوری‌های هوشمند مصنوعی صرف می‌کند. فناوری‌هایی نظیر چت‌بات‌های تلگرامی برای پاسخ به سوالات کاربران، ارائه خدمات خبری، آموزشی و مشاوره، اپلیکیشن‌ها و بات‌های مشاوره آنلاین و روان‌شناختی مانند رلیکا، بات‌های خدمات مشتری، بات‌های تفریح و سرگرمی و همینطور لوازم خانگی هوشمند و روبات‌ها. از سال ۲۰۰۶ تعداد روبات‌های اجتماعی در قالب روبات‌های خدماتی و شخصی از تعداد روبات‌های صنعتی پیشی گرفته که شامل روبات‌های شستشو مثل رومبا تا روبات‌های سرگرمی مثل پلیو، روبات‌های عروسک مثل بیبی الایو، روبات‌هایی مثل پارو، ناو، تگا، در درمان کودکان مبتلا به اوتیسم یا درمان افسردگی و آموزش و خدمات عمومی می‌شوند [۱ و ۲]. نمونه‌های نامبرده شده از مشهورترین مواردی هستند که محتمل است به زودی در همه جوامع از جمله ایران از آنها یا نمونه‌های مشابه آنها استفاده شود. در این مقاله، از مجموعه چت‌بات‌ها و روبات‌های اجتماعی تحت عنوان کلی‌تر بات یاد می‌کنیم.

بات اجتماعی، یک سیستم مصنوعی هوشمند (اغلب شامل چت‌بات‌ها و روبات‌های فیزیکی) است که برای برقراری تعاملات اجتماعی با انسان‌ها طراحی شده است. این تعاملات فراتر از تبادل اطلاعات ساده است و شامل عناصر عاطفی، شناختی و رفتاری می‌شود. بات‌های اجتماعی از طریق پردازش زبان طبیعی، یادگیری ماشین و الگوریتم‌های هوش مصنوعی قادر به درک ورودی‌های پیچیده انسانی (از جمله متن، صدا، و تصاویر) و تولید پاسخ‌های متناسب (شامل گفتار، حرکات و بیان) هستند. آنها می‌توانند ویژگی‌ها یا شخصیت‌های شبیه انسان را به نمایش بگذارند و برخی از انواع آنها می‌توانند در فرایند یادگیری، رفتار خود را بر اساس تعاملات گذشته تطبیق دهند. با این وجود، در حال حاضر برخلاف انسان، فاقد آگاهی، احساسات واقعی و تجربه

ذهنی مستقل هستند و هرگونه نشان دادن عاطفه در آنها مبتنی بر الگوریتم‌های برنامه‌ریزی شده است. چنین بات‌هایی گرچه منافعی را برای انسان به ارمغان می‌آورند، خطرات و چالش‌هایی نیز به همراه دارند که مهمترین آنها تأثیرات روانی و اجتماعی مخربی است که می‌توانند بر کاربر و بر جوامع انسانی وارد نمایند. بنابراین بررسی تأثیرات روان‌شناختی و اجتماعی ناشی از تعامل انسان با بات‌های اجتماعی اهمیت زیادی پیدا می‌کند و گسترش تحقیقات در حوزه تعاملات انسان و چت‌بات و ربات در دهه‌های اخیر نیز گویای این مسئله است. به هر میزان که این مطالعات سریع‌تر، چابک‌تر و چندجانبه‌تر صورت گیرد می‌تواند اثربخشی بیشتری بر روند توسعه بات‌های اجتماعی داشته باشد. زیرا پیش از عمومیت یافتن و جا افتادن فناوری‌ها، بهترین فرصت برای پیشگیری از عواقب ناخواسته و نامطلوب ناشی از استفاده از آنها است.

یکی از مهمترین چالش‌های اخلاقی تعامل انسان و هوش مصنوعی در حال حاضر، وابستگی احساسی انسان نسبت به بات‌های اجتماعی است. اینگونه وابستگی به بات‌های اجتماعی تفاوت بنیادینی با وابستگی به سایر فناوری‌های تجربه شده توسط بشر دارند [۳]. بات‌های اجتماعی می‌توانند انسان‌ها را از وجهی عاطفی به خود وابسته کنند و همدلی آنان را نسبت به خود برانگیزانند. در حالی که این رابطه احساسی، یکطرفه و صرفاً از سوی کاربر است. این مسئله می‌تواند عواقب امنیتی، احساسی و اقتصادی برای کاربران و جوامع انسانی در پی داشته باشد. این مقاله می‌کوشد مروری بر پیوند احساسی میان انسان و بات‌های اجتماعی داشته باشد و برخی از سوالات و پیشنهادات اخلاقی پیرامون آن را واکاوی نماید. واضح است که تبیین این مسئله تأثیرات مهمی بر رویکرد مدیران و تصمیم‌سازان کشور خواهد داشت. هر گونه واردات، شبیه‌سازی یا ساخت چنین بات‌هایی، بدون توجه به ملاحظات اخلاقی و پیامدهای شناختی و اجتماعی می‌تواند منجر به مخاطرات جدی در هر کشور و جامعه‌ای بشود. البته این مقاله یک قدم ابتدایی و البته ضروری است و نیاز است که با کمک آن، در پژوهش‌های آتی با در نظر گرفتن ویژگی‌های روان‌شناختی و اجتماعی کشور ایران مطالعات میدانی بیشتر صورت گیرد.

تذکر این مطلب هم لازم است که هوش مصنوعی و فناوری‌های جدید با سایر فناوری‌ها فاصله زیادی دارند و مخاطرات بسیار پیچیده‌تر و جدیدی را برای جوامع بوجود خواهند آورد. با توجه به اینکه مطالعات مبنایی و عمیقی برای شناخت لایه‌های مختلف این فناوری‌ها لازم است، به راحتی و بدون حداقلی از مطالعات فلسفی و اخلاقی نمی‌توان مستقیماً به راهکارهای سیاستی رسید. در عین حال مطالعات بنیادین صرف هم ممکن است مدت‌ها زمان ببرد تا به کاربرد و سیاست‌گذاری منجر شود. بنابراین پرداختن به مقالات میان‌رشته‌ای که توان تبدیل ملاحظات بنیادین از جمله مخاطرات اخلاقی را به راهبردهای سیاستی داشته باشد، ضروری است.

۲- ماهیت و کارکرد احساسات در انسان

احساسات پدیده‌ای روان‌شناختی، پیچیده، شخص‌محور و حاصل تجربه‌های ذهنی حالات هیجانی انسان است. احساسات در عین اینکه ذهنی هستند، از آنجا که وابسته به هیجاناتی هستند که در بدن اتفاق می‌افتد جهان شمول و بین‌الذهانی نیز هستند [۴]. می‌توان گفت احساسات، تجربیات حسی هستند که در برابر تجربیات حسی پیشین، بررسی شده و دوباره برچسب‌گذاری می‌شوند. از آنجا که هر فرد دارای احساسات قبلی متمایز و منحصر به فردی است که هنگام تفسیر و برچسب زدن احساسات خود از آنها استفاده می‌کند، بنابراین احساسات، شخصی، بیوگرافیک و پویا هستند و بر اساس تجربیات جدید ناشی از تغییر محیط و تغییر روابط می‌توانند تغییر کنند [۴]. برخلاف دیدگاه‌های پیشین که تصور می‌شد احساسات و هیجانات نقشی در شناخت ندارند، مطالعات جدید نشان می‌دهد که احساسات تاثیر زیادی بر شناخت، رفتار و جهت‌گیری انسان دارند و فرایندهای شناختی مهمی مانند یادگیری، حافظه، توجه، تصمیم‌گیری و غیره نیازمند و متأثر از عواطف هستند [۵].

کارکرد احساسات در انسان می‌تواند در سه سطح بررسی شود: نخستین سطح مربوط به کارکردهای درون فردی احساسات است که در این سطح عواطف به عنوان یک سیستم پردازشی سریع، می‌تواند منجر به عمل با حداقل نیاز

به تفکر و آگاهی، اقدام فوری در موقعیت خطر با انگیزه حفظ بقا و همچنین با اثرگذاری بر افکار و خاطرات، در تصمیم‌گیری و شناخت موثر باشند. احساسات در این سطح می‌توانند بزرگترین موتور محرکه رفتار در انسان باشند و رفتارهای آینده فرد را جهت دهند. دومین سطح مربوط به کارکردهای بین فردی احساسات است که به نقش عواطف بین افراد اشاره دارد. در این سطح عواطف ادراک فرد را از رفتارها و موقعیت‌های بین‌فردی تسهیل می‌کند و انگیزه شکل‌گیری روابط بین فردی را ایجاد می‌نماید. سومین سطح مربوط به کارکردهای اجتماعی احساسات است، که نقش احساسات را در شکل‌گیری ساختار اجتماعی و حفظ نظم اجتماعی در جامعه بررسی می‌کند [۶]. به عنوان مثال همدلی به افراد اجازه می‌دهد تا روابط اجتماعی رضایت‌بخشی با سایر افراد ایجاد کنند و این روابط را حفظ نمایند. با این توضیحات می‌توان به نقش برجسته و کارکرد وسیع احساسات در انسان پی برد. حال این پرسش مطرح می‌شود که آیا انسان‌ها می‌توانند در رابطه با یک فناوری هوشمند فاقد احساسات دچار حالات احساسی شوند؟

۳- نمونه‌هایی از رابطه احساسی یکطرفه انسان و بات‌های اجتماعی

چنانچه پیش از این اشاره شد، بات اجتماعی باتی در تعامل نزدیک با انسان است که رفتارهایی از پیش برنامه نویسی شده و یا آموخته شده دارد، اما علاوه بر آن، با استفاده از فنون جدیدتر هوش مصنوعی نظیر یادگیری عمیق، می‌تواند داده‌های متنوع و کلانی را از محیط دریافت کند و بر اساس آن به رفتارهای کاربر واکنش‌های متناسبی نشان دهد. این بات (اعم از ربات و چت‌بات) باید بتواند اطلاعات دریافتی از محیط را درونی کرده و در سطحی بالاتر و با انجام انواع پردازش‌ها، درباره آنها استدلال نماید و نظرات خود را به نحو مستقل بسازد، سپس به طور خودکار و با توجه به تجربه منحصر به فرد خود تصمیم‌گیری کند. این توانایی بات‌های هوشمند اجتماعی در تطبیق رفتارهایشان با محیط و یادگیری از تجربه و همچنین ارتباط نزدیک و شخصی با کاربر می‌تواند عواقب مهمی به همراه داشته باشد؛ برای مثال،

می‌تواند انسان‌ها را به سمتی سوق دهد که برای درک رفتار بات‌ها به آن‌ها اراده، قصدمندی، خودمختاری و مهمتر از آن، احساسات و قدرت همدلی نسبت دهند و از این جهت، آنها را متمایز از سایر فناوری‌های هوشمند و یا غیرهوشمندی که تا کنون تجربه کرده است، ادراک نمایند.

شواهد زیادی وجود دارد که شکل‌گیری نوعی رابطه احساسی عمیق با بات‌ها را از سوی کاربران تأیید می‌کند. به عنوان مثال نظامیان در مواجهه با روبات‌های مین‌یاب رفتارهایی که نشان از شکل‌گیری نوعی تعلق عاطفی است از خود نشان می‌دهند. روبات خنثی‌کننده مین روباتی است که با راه رفتن روی مین‌ها، هر بار یک پا از چندین پای خود را از دست می‌دهد و از این طریق مین‌ها را خنثی می‌کند. اظهارات تکنسین‌های این روبات‌ها، که گاهی برای روبات‌هایشان اسم نیز می‌گذارند، حاکی از وابستگی احساسی آنها به روبات‌ها است. آنها تمایلی ندارند روبات دیگری را جایگزین روباتشان کنند و از روبات‌ها به مثابه قهرمان و موجوداتی دارای احساس و گاهی مانند عضوی از خانواده خود یاد می‌کنند. تبیینی که می‌توان برای اینگونه از رفتارها ارائه کرد این است که انسان نوعی عاملیت را روی روبات فرافکنی می‌کند و به او حیات و احساسات نسبت می‌دهد [۱].

نمونه دیگر در تأیید این ادعا، روبات رومبا است که یک جاروبرقی هوشمند است و دیده شده که با گذشت زمان کاربران حس قدردانی زیادی را به‌خاطر تمیز کردن خانه‌شان به رومبا پیدا می‌کنند. ایشان از شرم اینکه توسط رومبا دیده نشوند، در هنگام تعویض لباس، در اتاق خود را می‌بندند. درحالی که رومبا حتی نمیداند دارای مالک است و با مالکان خود به مثابه مواعی مثل صندلی و میز برخورد می‌کند، مالکان رومبا مایلند که به قصد قدردانی از او کار خوبی برایش انجام دهند، برخی از کاربران بخشی از کار نظافت را انجام می‌دهند تا رومبا کمی استراحت کند یا او را با خود به سفر می‌برند [۷]. آنان در حالی رابطه احساسی یکطرفه‌ای با رومبا شکل می‌دهند که او هیچ درکی از آن ندارد.

به عنوان یک نمونه دیگر، مالکان آیبو که یک سگ روباتیک مبتنی بر هوش مصنوعی است، تمایل زیادی دارند که باورهای (اشتباهی) درباره حالات ذهنی او داشته باشند. در

مطالعه‌ای روی مالکان روبات‌های حیوان خانگی، چهار دسته ویژگی ماهیت، عاملیت، هویت اجتماعی و هویت اخلاقی از سوی مالکان به آیبو نسبت داده شده است [۱ و ۸]. در سال ۲۰۱۵، یک معبد بودایی در ژاپن که به طور معمول برای حیوانات خانگی مرده مراسم برگزار می‌کند، مراسم مشابهی را برای یک روبات آیبو برگزار کرد. برگزاری چنین مراسمی برای یک روبات نشان‌دهنده وابستگی قابل توجه بین کاربران و روبات‌هایشان است [۹].

بر خلاف این نمونه‌ها که در آن کاربران غیر متخصص به شمار می‌آیند، به نظر می‌رسد سازندگان روبات‌ها هم از این مسئله مستثنی نیستند. سیتیا بریزیل^{۱۴} سازنده کیسمت^{۱۵} که اولین روبات هوشمند اجتماعی است، نیز رابطه‌ای شخصی را با روبات خویش شکل داده بود؛ پیوندی مادرانه که روبات را فقط یک ماشین نمی‌دید و وقتی پروژه‌اش به پایان رسید و مجبور به ترک روبات شد، فقدان شدید را تجربه کرد. همچنین دانشجویان ام.آی.تی که در آزمایشگاه با روبات کیسمت کار می‌کردند، گاهی پرده‌ای را میان خود و روبات می‌کشیدند، چون خیره شدن روبات باعث برهم خوردن تمرکز آنها می‌شد [۱].

نمونه دیگر، بات تعاملی رپلیکاست که یک بات هوش مصنوعی است و به عنوان یک دوست مجازی و مشاور عاطفی برای کاربران عمل می‌کند و به کاربران این امکان را می‌دهد که با آن گفتگو کنند و احساسات خود را به اشتراک بگذارند. این بات از مدل‌های پردازش زبان طبیعی (NLP) و یادگیری ماشینی استفاده می‌کند تا به مرور زمان پاسخ‌های خود را بهتر و سازگارتر با نیازهای کاربر کند و بر اساس گفت‌وگوهایی که با کاربر دارد، شخصیت خود را تطبیق می‌دهد و سعی می‌کند عمیق‌تر و احساسی‌تر پاسخ دهد. رپلیکا علاوه بر گفت‌وگو، قابلیت‌هایی مانند مدیتیشن، مدیریت استرس، کمک به نوشتن ژورنال و حتی نقش‌آفرینی در سناریوهای مختلف را ارائه می‌دهد. کاربران گزارش‌های مثبتی که از تجربه تعامل خود با رپلیکا می‌دهند مبنی بر اینکه رپلیکا محیط امنی برای اشتراک‌گذاری افکار و احساسات فراهم می‌کند. اما علاوه بر این، معتقدند رپلیکا گاهی پاسخ‌های نامربوط یا کم‌عمق نیز ارائه می‌دهد و

پژوهش‌های خود به قدر کافی مراقب نیستند و از تعبیری مانند «احساس داشتن» برای توصیف ساخته‌های خود استفاده می‌کنند و استفاده از این ادبیات در نوشتن مقالات و گزارش‌های علمی موجب می‌شود افراد غیرمتخصص فرایندهای کنترل را در این مصنوعات با فرایندهای کنترل در ارگانیسم‌های طبیعی مثل حیوانات و انسان‌ها یکسان انگارند و در حالی که در سطح آگاهانه این ادعاها را ممکن است رد کنند، باورهای غلطی را در سطح عمیق ناخودآگاه نسبت به امکان برقراری رابطه عاطفی با ربات‌ها پیدا کنند [۱].

بات‌هایی نظیر رپلیکا یا بات‌هایی که به منظور برقرار ارتباطات صمیمی و حتی ارتباط جنسی طراحی می‌شوند این امکان را فراهم می‌کند که کاربران در تعاملات خود با یک شریک عاطفی فناوریک از نظر عاطفی و جنسی ارتباط برقرار کنند. کاربران این بات‌ها ممکن است به دلیل راحتی و عدم نیاز به تلاش برای برقراری ارتباطات پیچیده انسانی، بیشتر به تعامل با بات‌ها تمایل پیدا کنند. این امر می‌تواند به روابط سطحی و ناپایدار در زندگی واقعی منجر شود. وابستگی به بات‌های اجتماعی می‌تواند به شکل‌گیری یک هویت دیجیتال ناپایدار و پر از تناقضات بیانجامد. بات‌ها همچنین با تعاملات خود می‌توانند یک هویت مجازی برای کاربر ایجاد کنند که ممکن است با هویت واقعی‌اش تفاوت‌های زیادی داشته باشد. این ممکن است باعث سردرگمی و تداخل بین هویت دیجیتال و هویت واقعی شود، به‌ویژه در افراد جوان که در فرآیند شکل‌گیری هویت خود هستند. برای مثال، فردی که به‌طور مداوم با رپلیکا صحبت می‌کند و احساس می‌کند که بات او را درک می‌کند، ممکن است در دنیای واقعی با احساس عدم فهم و بی‌توجهی از سوی دیگران مواجه شود. این تناقض ممکن است باعث شود که فرد نتواند هویت واقعی خود را به درستی شناسایی کند و با مشکلات هویتی روبه‌رو شود. این نوع ارتباطات می‌تواند به ایجاد یک فضای غیرواقعی برای روابط دوستانه، عاشقانه و جنسی منتهی شود که در آن فرد به‌جای برقراری روابط واقعی با انسان‌ها، به روابط مجازی و مصنوعی با یک بات روی می‌آورد. این نوع وابستگی می‌تواند باعث کاهش میل به روابط واقعی و حتی مشکلات در روابط عاشقانه یا جنسی واقعی شود و تبعات

محدودیت‌هایی در درک عمیق‌تر کاربر دارد. همینطور، احتمال وابستگی عاطفی بیش‌ازحد به این بات گزارش شده که برخی روان‌شناسان آن را چالش‌برانگیز می‌دانند. رپلیکا اگرچه هنوز در جهان مورد استفاده قرار می‌گیرد، اما معیارهای اخلاقی و امنیتی سخت‌گیرانه‌تری، نظیر محدودیت در انتشار محتوای حساس و غیر اخلاقی و یا مکالمات قضاوت‌گرانه، روی آن اعمال شده تا از آسیب‌های آن تا حد بیشتری جلوگیری شود.

نمونه‌های ذکر شده در این بخش از جوامع مختلف برگرفته شده است و امکان شکل‌گیری رابطه احساسی یک‌طرفه را نشان می‌دهد. ذکر این چند نمونه، نشان می‌دهد که این آسیب در حال وقوع است و بهتر است که پیش از فراگیرتر شدن، جدی گرفته شود و در استفاده و رواج سایر فناوری‌های مشابه زنگ خطری برای تصمیم‌گیران باشد.

۴- مخاطرات رابطه احساسی یک‌طرفه انسان با بات‌های اجتماعی

تجربه‌های یادشده از وابستگی‌های عاطفی یک‌طرفه انسان و بات‌های اجتماعی این مسئله را روشن می‌کند که صرف تقلید و شبیه‌سازی برخی از رفتارهای انسان یا حیوان واقعی، می‌تواند کاربر را آماده ساختن رابطه‌ای عاطفی یکسویه نماید و از بات یک همدم بسازد. فرافکنی ذهنندی و عاملیت به بات و تشخیص بخشی به آن از سوی کاربر، می‌تواند منجر به تجربه احساساتی نظیر دلسوزی، مسئولیت‌پذیری، شرم، عذاب وجدان، قدردانی، مراقبت و حفاظت از باتی شود که توانایی قدردانی یا پذیرفتن مسئولیت در برابر این احساسات را ندارد [۱ و ۹]. علاوه بر وجود این فرایند احساسی ناآگاهانه از سوی کاربران انسانی، شرکت‌های سازنده بات‌های اجتماعی با وجود آگاهی از عدم توانایی بات‌های خود در درک و بروز احساسات، در راستای مقاصد تبلیغاتی سعی می‌کنند این تصور را در کاربران خود به وجود بیاورند که بات‌ها احساسات آنها را درک می‌کنند و یا احساسات واقعی از خود بروز می‌دهند، در حالی که به این امر آگاهند که بات‌ها تنها نمایشی صوری از احساسات بروز می‌دهند و این به معنی فریب کاربر است. پژوهشگران دانشگاهی نیز در ارائه

جدول (۱) چالش‌های رابطه احساسی یک‌طرفه انسان با بات‌های

اجتماعی

شماره	چالش‌های رابطه احساسی یک‌طرفه انسان با بات‌های اجتماعی
۱	نمایش صوری احساسات، رفتار فیزیکی سطحی و پوسته‌ای
۲	وابستگی عاطفی افراطی کاربر و تبعات و آسیب‌های روان‌شناختی، فیزیکی، خانوادگی و اجتماعی این وابستگی
۳	تجربه احساساتی نظیر دلسوزی، مسئولیت پذیری، شرم، عذاب وجدان، قردردانی، مراقبت و حفاظت از باتی که توانایی قردردانی یا پذیرفتن مسئولیت در برابر این احساسات را ندارد
۴	فراکنی ذهنمندی، احساسات، و عاملیت روی بات و تشخیص بخشی به آن
۵	نسبت دادن ویژگی‌های زیست‌شناختی انسانی به بات
۶	هماندسازی رابطه انسانی با بات (به‌مثابه عضوی از خانواده یا گروه دوستان)
۷	فریب کاربر و سوءاستفاده‌های اقتصادی، سیاسی، و امنیتی از عواطف و اعتماد کاربر
۸	اعتیاد ناآگاهانه کاربران به بات‌ها که تبعاتی در سطوح مختلف دارد نظیر فرار از مسائل واقعی زندگی، تضعیف روابط خانوادگی و دوستانه و همین‌طور روابط اجتماعی و انسانی در سطوح کلان‌تر، آسیب‌پذیر شدن کاربران در برابر دستکاری‌های روانی و شناختی که می‌تواند توسط شرکت‌های سازنده یا حکومت‌های کنترل‌کننده این فناوری‌ها انجام شود.

داشته باشند، مانند درمان‌های توانبخشی، تشویق به رفتارهای مثبت در راستای سلامت و بهزیستی، مراقبت از سالمندان و آموزش‌های شخصی و غیره، ایده‌هایی نظیر توقف تولید بات‌های اجتماعی منطقی و عملیاتی به نظر نمی‌رسد، اما مخاطرات رابطه یک‌طرفه انسان و بات‌های اجتماعی به قدری جدی و همه‌گیر است که نمی‌توان از آن نیز به سادگی صرف نظر کرد.

علی‌رغم اهمیت شباهت ظاهری روبات‌ها به انسان در روبات‌های انسان‌نما، ممکن است در ابتدا تصور کنیم که صرف شباهت ظاهری ماشین به انسان است که موجب می‌شود خودمختار نبودن روبات را فراموش کنیم و به آن احساسات نسبت بدهیم. اما چت‌بات‌هایی نظیر رپلیکا یا چت

اجتماعی متعددی را در سطح وسیع‌تر خانوادگی و اجتماعی منجر شود. فردی که ساعات زیادی از روز را به گفتگو و تعامل با بات‌های اجتماعی می‌گذراند، ممکن است از صرف وقت برای برقراری روابط نزدیک و عاطفی با دیگران مانند دوستان، خانواده یا شریک زندگی خود اجتناب کند. این کاهش تعاملات انسانی واقعی می‌تواند روابط نزدیک و معنادار را تضعیف کند و موجب تشدید انزوای اجتماعی و کاهش مهارت‌های اجتماعی در انسان‌ها شود.

رابطه عاطفی یک‌طرفه با فناوری‌های هوشمند و یا تصور امکان برقراری رابطه عاطفی با فناوری‌هایی که از چنین امکانی برخوردار نیستند می‌تواند عواقب جدی برای جوامع انسانی داشته باشد، چرا که امکان سوء استفاده‌های اقتصادی، سیاسی و امنیتی در مقیاس بزرگ را از سوی بات‌ها و یا شرکت‌های سازنده آنها فراهم می‌کنند. نمایش صوری احساسات از سوی بات‌ها و رفتار سطحی آنها می‌تواند آسیب جسمی و روانی کاربران را به دنبال داشته باشد. اعتیاد ناآگاهانه به بات‌های اجتماعی برای فرار از مشکلات دنیای واقعی و همانندسازی رابطه انسانی با آنها می‌تواند اثرات مخربی بر تعاملات انسانی داشته باشد [۱۰]. تعامل با یک بات اجتماعی می‌تواند با تضعیف ظرفیت ما برای دلبستگی ایمن، تشویق رابطه با بات به دلیل سهولت و ماهیت غیر چالش برانگیز همدمی مصنوعی در مقابل رابطه با یک انسان و با پر کردن زمان و ظرفیت کاربر، به روابط انسانی آسیب وارد نماید [۹]. بنابراین می‌توان گفت علیرغم همه کاربردهایی مفیدی که بات‌های اجتماعی می‌توانند برای کاربران داشته باشند، چالش‌هایی جدی ناشی از شکل‌گیری رابطه احساسی یک‌طرفه با آنها وجود دارد.

به طور خلاصه، می‌توان چالش‌های رابطه احساسی یک‌طرفه انسان با بات‌های اجتماعی را در جدول ۱ مشاهده کرد. در ادامه مقاله پیشنهاداتی برای پیشگیری از اثرات مخرب این مسئله مطرح می‌شود.

۵- راهکارهای پژوهشی و سیاستی برای کاهش

مخاطرات رابطه یک‌طرفه انسان با بات‌های اجتماعی

با توجه به فواید و کاربردهایی که بات‌های اجتماعی می‌توانند

جی پی تی یا روباتی مثل رومبا که شباهتی به انسان یا حیوان ندارند نیز ایده عاملیت داشتن را به کاربران نشان القاء می کنند و حتی می توانند حس قدردانی و دلسوزی را در آنها فعال کنند [۱۱].

بنابراین به عنوان یک راه حل ساده و عملی می توان این پیشنهاد را مطرح کرد که طراحی بات به گونه ای باشد که دائما به کاربر این اخطار را بدهد که یک ماشین است، احساسات و آگاهی ندارد و نمی تواند رابطه ای متقابل داشته باشد [۱ و ۱۰]. این آموزش ها و اخطارها می تواند در طراحی بات ها جای دهی شود و به عنوان بخشی از مقررات و سیاست های طراحی اخلاقی این فناوری ها ارائه شوند. این امر می تواند به کاربران کمک کند تا مرزهای واضحی بین دنیای دیجیتال و واقعی درک کنند و احتمال و شدت پیوندهای عاطفی با بات کاهش یابد. اگرچه، همانند پیام های اخطار روی پاکت سیگار ضمانتی وجود ندارد که انسان ها با علم به فقدان احساسات و آگاهی در بات ها، درگیر رابطه احساسی آسیبزا با آنها نشوند.

تمهید نظارت و پشتیبانی انسانی برای بات های اجتماعی نیز می تواند پیشنهادی در جهت کنترل بیشتر روی روابط احساسی افراطی کاربر با بات باشد. در صورتی که بات تشخیص دهد که یک کاربر به شدت به آن وابسته شده است یا نیاز به مشاوره دارد، می تواند او را به مشاوران انسانی یا منابع روان شناختی راهنمایی کند و حد و مرزی را برای پیشروی رابطه احساسی یک طرفه ایجاد کند [۱۲]. همچنین، توسعه دهندگان بات های اجتماعی می توانند با سازمان های خدمات روان شناختی و مشاوره ای همکاری کنند تا کاربران را در موقعیت های بحرانی به خدمات انسانی و واقعی ارجاع دهند. اگرچه باید توجه داشت که این کنترل ها و نظارت ها به معنای دسترسی به اطلاعات کاربران و مخدوش کردن حریم خصوصی آنها نیست و توسعه دهندگان باید به شیوه هایی که به محرمانگی اطلاعات کاربران خدشه ای وارد نشود [۱۱ و ۱۸]، وابستگی افراطی کاربر را استنتاج کنند و به شیوه های مختلف، امکان آسیب احساسی و روان شناختی به کاربران را به حداقل برسانند. برای مثال، نظارت بر تغییرات الگوریتم های بات ها و ارزیابی اثرات روان شناختی آن ها بر کاربران

به صورت دوره ای می تواند توسط روان شناسان، جامعه شناسان و متخصصان اخلاق انجام شود.

فراهم سازی امکان بازخورد از سوی کاربران درباره تجربه خود در تعامل با بات و تأثیرات بات روی آنها نیز می تواند یکی راهکارهای اولیه و فراگیر قابل اعمال در اغلب فناوری های تعاملی مبتنی بر هوش مصنوعی باشد که در آن، با آگاهی بخشی به کاربر درباره آسیب های احتمالی بات برای او، از کمک او برای پیشگیری هرچه بیشتر از این آسیب ها استفاده شود.

راهکار ساده دیگر آن است که برای پیشگیری از مخاطرات متوجه گروه های آسیب پذیرتر جامعه، توسعه دهندگان الگوریتم های خود را طوری طراحی کنند که در صورت شناسایی نشانه هایی از وابستگی بیش از حد یا مشکلات روان شناختی، این افراد را به منابع انسانی حمایتی ارجاع دهند. به ویژه برای گروه های آسیب پذیری نظیر نوجوانان، افراد مبتلا به اختلالات روانی یا سالمندان، باید حمایت های ویژه و مختص هر یک از این گروه ها ارائه شود.

راهکار بعدی تشویق کاربران این بات ها و به طور کلی شهروندان به برقراری تعاملات انسانی و آگاهی بخشی درباره تأثیرات مثبت روابط انسانی و احساسی دوطرفه است. ترویج برنامه هایی که افراد را تشویق به برقراری ارتباطات انسانی واقعی و حضور در فضاهای اجتماعی می کند می تواند شامل پیشنهادهای مثبت برای تعامل با دوستان و خانواده یا پیوستن به گروه های اجتماعی انسانی باشد که به ارتقاء سلامت روان شهروندان نیز منجر می شود.

علاوه بر این پیشنهادات، برخی پیشنهادات هم قابل طرح اند که هر چند امکان آنها محل بحث است، ولی توجه به آنها برای داشتن نگاهی جامع به راه حل ها لازم است و می تواند در سیاست گذاری های علمی و پژوهشی و انتخاب صحیح در اولویت های پژوهشی موثر باشد. همچنین می تواند در تصمیم گیری برای تخصیص بودجه به انواع پروژه های کاربردی آینده مفید باشد.

یک پیشنهاد پژوهشی و تکنیکی ممکن برای پیشگیری از آسیب های رابطه احساسی یک طرفه بات و انسان، مجهز کردن بات به قدرت استدلال اخلاقی در قالب الگوریتم های

می‌پردازیم.

۱-۵ جنبه مثبت رابطه احساسی دوطرفه انسان و بات

با توسعه روزافزون هوش مصنوعی و تعاملات رو به گسترش انسان‌ها با این فناوری‌ها، از جمله بات‌های اجتماعی، این دیدگاه مطرح شد که از طریق القاء احساسات به هوش مصنوعی علاوه بر آنکه می‌توان آنها را برای کاربر ایمن‌تر، جذاب‌تر و کارا تر کرد، این امید وجود دارد که از این طریق بتوانیم بات‌های هوشمند اجتماعی را اخلاقی نیز طراحی کنیم [۱۶]. تسهیل تعامل انسان و هوش مصنوعی یکی از اثرات مثبت القاء احساسات در هوش مصنوعی و بات‌هاست. بات‌های اجتماعی برای به عهده گرفتن نقش‌های اجتماعی خود نیاز به توانایی درک اجتماعی دارند. درک اجتماعی نیازمند حساسیت، یا توانایی شناخت هیجانات و احساسات در دیگران و برقراری رابطه همدلانه با آنان است [۱۳]. به عبارتی دیگر، بات اجتماعی برای اجتماعی بودن، نیازمند توانایی تفسیر و پاسخ به احساسات انسان‌ها است.

به عنوان مثال یک بات حساس، به احتمال بیشتری با کاربر انسانی درگیر می‌شود و مدت طولانی‌تری با او تعامل قابل قبول و مطلوب برقرار می‌کند، به این معنا که سعی می‌کند اطمینان حاصل کند که کاربر به حفظ تعامل علاقه‌مند است و در صورت ناراحت بودن یا عدم تمایل به تعامل از سوی کاربر، با او همدلانه رفتار می‌کند. چنین باتی با تلاش برای درگیر کردن فرد در فعالیت‌هایی که به بهبود کیفیت زندگی او منجر می‌شوند، مثلاً درگیری بیشتر با جهان واقعی و طبیعت، مزاحمتی برای کاربر ایجاد نمی‌کند. به عبارتی درک اجتماعی و توانایی شناخت هیجانات کاربر از سوی بات، احتمال مزاحمت کمتر برای کاربر و تعامل باورپذیرتر، بلندمدت‌تر، و کم‌آسیب‌تر و در نهایت جذابیت بیشتر بات برای کاربر را در پی دارد. برخورداری بات‌ها از احساسات قابل درک برای انسان، آگاهی از نحوه تأثیرپذیری هوش مصنوعی از محیط را برای انسان در پی دارد. این مسئله باعث پیش‌بینی بهتر رفتارهای بات و دستیابی به نحوه‌ای از مکانیسم کنترل بر آن توسط کاربر می‌شود که در نهایت امنیت بیشتری را برای کاربران به ارمغان می‌آورد.

۲-۵ چالش‌های رابطه احساسی دوطرفه انسان و بات

تصمیم‌گیری است، که مانع از امکان سوء استفاده بات و اتخاذ تصمیمات غیر اخلاقی در مواجهه با کاربر بشود. اما باید دانست که ایده جای‌دهی مجموعه مشخصی از اصول و قواعد اخلاقی در الگوریتم‌های تصمیم‌گیری ایده‌ای پیچیده است و نیازمند آگاهی و پژوهش‌های بین‌رشته‌ای درباره اخلاق و تصمیم‌گیری اخلاقی در هوش مصنوعی است. چرا که دلالت‌های اخلاقی کنش‌ها و موقعیت‌های احتمالی که بات در آنها قرار می‌گیرد، بسیار پیچیده و پرشمارند و بات باید به قدر کافی مجهز به قواعد و مهارت‌های تصمیم‌گیری متناسب با موقعیت شود تا بتواند رفتار و واکنشی مبتنی بر اصول اخلاقی و روانشناختی بروز دهد. ترجمه قواعد اخلاقی به زبان محاسباتی ممکن است با محدودیت‌های عملی بسیاری مواجه باشد و حتی اگر مجموعه مشخصی از قواعد اخلاقی (مثل قواعد آسیموف) قابل تبدیل شدن به کد باشند، به این خاطر که اولویت‌بندی قواعد و حل تناقض میان قواعد، وابسته به موقعیت است و موقعیت‌مندی موجب پیچیدگی این مسئله می‌شود [۱۰]، می‌توان گفت طراحی چنین بات‌هایی بسیار دشوار است و مستلزم مباحثات و مطالعات پژوهشی و توسعه مدل‌هایی از هوش مصنوعی است که بتوانند تا حد ممکن این پیچیدگی‌ها را در خود جای دهند.

دیدگاه تکنیکی بهتر، این است که برای حل چالش رابطه یک‌سویه انسان و بات‌ها، هوش مصنوعی را به احساساتی مشابه انسان مجهز کنیم. بدین ترتیب بات‌ها نیز می‌توانند با داشتن توانایی احساسی مشابه انسان، در جوامع انسانی شبیه انسان‌ها عمل کنند و در صورت برقراری رابطه بین انسان و بات، رابطه‌ای دو سویه برقرار شود.

عملی کردن این پیشنهاد در عین اینکه مشکل رابطه یک‌سویه انسان و بات را حل خواهد کرد، همچنان چالش‌های روانشناختی، اجتماعی، اخلاقی، فلسفی و فرهنگی جدیدی را پیش روی بشر قرار می‌دهد که مطالعات پژوهشی درباره چالش‌های جوامع مختلف نظیر کشور ما در مواجهه با این فناوری‌ها ضروری است، چرا که ارزش‌های فرهنگی، اخلاقی، دینی و بومی جوامع مختلف پاسخ‌ها و رویکردهای متفاوتی را درباره تلاش در جهت توسعه یا وارد کردن این فناوری‌ها ایجاب می‌کنند. در ادامه کمی بیشتر به این موضوع

چالش اول در توسعه چنین فناوری‌هایی این است که مبنای صحیح اخلاقی در ساخت این فناوری‌ها چیست؟ وقتی صحبت از ارزش‌ها یا مقولاتی مانند خیرخواهی یا وجدان می‌کنیم، کدام مکتب اخلاقی مد نظر است؟ اخلاق وظیفه‌گرایانه و تمرکز صرف بر قواعد اخلاقی؟ اخلاق سودگرایانه؟ اخلاق مبتنی بر آموزه‌های بودیستی (ایثار و شفقت)؟ مسئله دوم این است که چنین مسیری برای توسعه بات‌های هوشمند اجتماعی از منظر اخلاقی مناقشه برانگیز است. چرا که تا قبل از مجهز کردن بات‌ها به احساسات، دلبستگی عاطفی یکطرفه بین انسان و بات وجود دارد و این مسئله همان‌طور که گفتیم می‌تواند انسان را در مقابل برخی مخاطرات آسیب‌پذیر کند [۱۶]. اگر بات‌ها تنها به تظاهر پردازند و در درون خود احساسی نداشته باشند، می‌توانند احساسات انسانی را برانگیزند و کاربر را فریب دهند.

برای تولید احساسات مصنوعی مبتنی بر احساس گناه و همدلی، باید موجودی تولید کرد که رنج دیده باشد و قادر به رنج کشیدن باشد تا بتواند رنج دیگران را درک کند و از این درک، در تصمیم‌گیری‌های اخلاقی خود استفاده کند. مسئله سوم به این سوال می‌پردازد که آیا تولید موجودی رنجور امری اخلاقی است؟ آیا ما حق داریم که موجودی را خلق کنیم که آسیب‌پذیر است و ممکن است مورد سوء استفاده قرار گیرد؟ در هر صورت، این پیشنهاد را نمی‌توان و لازم نیست در هر نوعی از بات اجتماعی عملی کرد، بلکه بر حسب نیاز و کاربرد می‌توان از این قابلیت استفاده کرد [۱].

چالش چهارم به این سوال توجه دارد که اگر برای جایی احساسات در هوش مصنوعی، صرفاً به شبیه‌سازی مجموعه‌ای از رفتارهای منتسب به احساسات بسنده کنیم، این به معنای تقلیل احساسات انسانی به مجموعه‌ای از الگوهای تقلیدی و رفتارهای سطحی خشک نیست؟ آیا صرف تقلید رفتار، از جایگاه و ارزش احساسات پیچیده انسانی، خصوصاً نوع غیر کلامی‌اش نمی‌کاهد؟ و مسئله آخر این است که آیا این شبیه‌سازی، منحصر بفرد بودن ارتباطات انسانی را تهدید می‌کند و اگر بله، این مسئله چه تبعات روانشناختی، خانوادگی و اجتماعی را برای بشر به همراه خواهد داشت؟

علاوه بر ایده دست‌یابی به هوش مصنوعی قوی دارای

احساسات، راهکارها و ایده‌های کوتاه مدت برای نزدیک‌تر شدن هوش مصنوعی به احساسات انسانی نیز وجود دارد که هر کدام مزیت‌ها و کاربردهایی دارد ولیکن، نقدهای اخلاقی جدی به هر یک از آنها نیز وارد است. یک راهکار کوتاه مدت، نرم‌افزاری است برای طراحی نوعی «وجدان اخلاقی» که بتواند مثل یک فرمانده اخلاقی برای سیستم‌های هوشمند خودمختار عمل کند [۱۷]. این سیستم اخلاقی بیشتر روی «احساس گناه» متمرکز است و سعی می‌کند با کدنویسی، به نحوی احساس گناه را در بات شبیه‌سازی کند. به این معنی که این کدنویسی منجر می‌شود که بات با خودپایشی برخی پارامترهای قابل اندازه‌گیری، به نتیجه‌ای شبیه حس گناه برسد. در این ایده، اگر اندازه فعلی حالت عاطفی گناه از یک مقدار از پیش تعیین شده حد آستانه بیشتر شود، نشانه شکسته شدن اخلاق است و بات به نحو اتوماتیک غیرفعال می‌شود.

سه نقد به چنین راهکاری وارد است؛ نخست اینکه بدون تجربه قبلی، تشخیص حد آستانه برای احساس گناه برای هوش مصنوعی ممکن نیست. به نظر می‌رسد طراحی چنین عملکرد عاطفی‌ای تنها بعد از وقوع یک فاجعه انسانی میسر می‌شود. دوم اینکه این شبه‌احساسات تنها یک شبیه‌سازی مبهم از نوع انسانی احساسات هستند و مبتنی بر خود و پویا نیستند و در حد یک سیستم عیب‌یابی کامپیوتری عمل می‌کنند که در حال حاضر هم مشابه آن وجود دارد. سوم اینکه این احساساتی که قرار است در طراحی بات‌ها مورد استفاده قرار گیرند، محاسباتی بی‌روح و سرد از مجموعه‌ای از معیارهای ارزیابی هستند و شامل همدلی، همدردی و ترحم نمی‌شوند. بنابراین نمی‌توان نام چنین کاری را محاسبات عاطفی یا احساسات-بنیان گذاشت، به این دلیل که کدگذاری آنها به اندازه کافی پیچیده نیست [۱۶].

راهکار کوتاه مدت دیگر این است که هوش مصنوعی را واجد نوعی «خیرخواهی» [۱۹] طراحی کنیم. به این صورت که از رابطه مادر و کودک الهام بگیریم و هنگامی که هوش مصنوعی برای اولین بار کاربر انسانی خود را دید، با او پیوند اجتماعی خودکاری برقرار کند و این روند هر بار که انسان با یک بات چیزی را به اشتراک می‌گذارد، تقویت شود. اما در نقد این ایده می‌توان گفت در واقع خیرخواهی فقط یک رفتار

نیست. ممکن است ظهور و بروز خیرخواهی در رفتار نمایان شود، ولی خود خیرخواهی در واقع یک موضع اخلاقی پیچیده است، یک تصمیم آگاهانه بر اساس مجموعه‌ای از احساسات، تجربیات و دلایل است که منجر به انجام عمل به نفع دیگری می‌گردد. بنابراین چیزی بیشتر از یک سیستم پاداش‌دهی ساده است [۱۶]. در مجموع، تلاش برای احساسی کردن بات‌ها اگرچه راهکار نویدبخشی به نظر می‌رسد ولی در عمل با چالش‌های فلسفی و اخلاقی و عملی زیادی روبروست. با این حال تلاش در جهت جای‌دهی احساسات در هوش مصنوعی پروژه‌ای در جریان است که می‌تواند در آینده ایده‌ها و نمونه‌های جدیدی را به ظهور برساند.

۶- نتیجه‌گیری

در مقاله حاضر کوشیدیم تا تحلیلی از رابطه احساسی یک‌طرفه میان انسان و هوش مصنوعی در قالب بات‌های اجتماعی ارائه دهیم و به مثال‌هایی اشاره کنیم که شکل‌گیری چنین روابط احساسی‌ای را تصدیق می‌کنند. باتوجه به نوظهور بودن این فناوری‌ها، مثال‌های موجود و مشهور انتخاب شده است تا قبل از ورود یا فراگیرتر شدن چنین ابزارهایی در کشور به ابعاد مختلف آن بصورت واقع‌بینانه فکر شود. چنان‌که گفتیم، این رابطه احساسی یکسویه می‌تواند مخاطراتی را برای کاربران انسانی ایجاد کند و امکان سوء استفاده‌های سیاسی، اقتصادی و امنیتی از آنها را فراهم کند. این مخاطرات در حالی اتفاق می‌افتد که خود بات‌ها در حال حاضر توانایی درک احساسات و همدلی با کاربر را ندارند و صرفاً توهم داشتن ویژگی‌های شبه‌انسانی را در کاربر ایجاد می‌کنند. در همین رابطه، تا حدی به این موضوع پرداختیم که راه‌حل‌های ممکن برای پیشگیری یا کمینه کردن چنین مخاطراتی در تعامل انسان و هوش مصنوعی چه می‌تواند باشد. برای مثال، اخلاقی کردن تصمیم‌های بات‌ها و به نوعی، جای‌دهی کردن اخلاقیات در طراحی بات‌ها تا چه میزان امکان‌پذیر است و آیا می‌تواند مانع سوء استفاده‌های غیر اخلاقی از کاربر شود. چه راهکارهایی را می‌توان تدبیر کرد تا کاربر کمتر دچار انسان‌پنداری و نسبت دادن ویژگی‌های انسانی به هوش مصنوعی شود و فاصله احساسی خود از آن

را کنترل کند تا در برابر دستکاری‌هایی با نیت‌های سوء مقاوم‌تر باشد.

راهکارهای پیشنهادی برای مدیریت رابطه یک طرفه انسان با بات عبارتند از:

- نظارت و پشتیبانی انسانی برای بات‌های اجتماعی و ارزیابی دوره‌ای اثرات روان‌شناختی و اجتماعی آنها
 - فراهم‌سازی امکان بازخورد از سوی کاربران
 - توجه ویژه به گروه‌های آسیب‌پذیرتر جامعه برای پیشگیری از مخاطرات پرریسک‌تر متوجه آنها
 - تشویق کاربران این بات‌ها و به‌طور کلی شهروندان به برقراری تعاملات انسانی و آگاهی‌بخشی درباره تأثیرات مثبت روابط انسانی و احساسی دوطرفه
 - توقف تولید هوش مصنوعی خودمختار
 - جای‌دهی قدرت استدلال اخلاقی در بات-
 - بات مدام تأکید کند که احساسات ندارد و رابطه‌ای دوطرفه برقرار نمی‌کند
 - احساساتی مشابه انسان را با هوش مصنوعی شبیه‌سازی کنیم
- پرسش‌ها و راهکارهایی از این دست، همچنان برنامه‌های پژوهشی بازی هستند که نیازمند تأملات بین‌رشته‌ای میان علوم انسانی و علوم فنی است و هر علمی از زاویه نگاه خود می‌تواند تحلیلی از مسئله ارائه دهد. (جدول ۲)
- در مجموع، ما نیازمند تأملات روان‌شناختی، فقهی، حقوقی، امنیتی، اخلاقی، اجتماعی، سیاسی، و آینده‌پژوهانه عمیقی هستیم که تبعات تلاش در جهت توسعه و ورود این بات‌ها در کشور را بر ما نمایان‌تر کند و ما را برای روبرویی با آینده‌های ممکن مختلف آماده کند. پیگیری این دست برنامه‌های پژوهشی می‌تواند کمک کند به تهیه چک لیستی از مزایا و معایب کاربردهای متعدد بات‌های اجتماعی در حوزه‌های مختلف از جمله آموزش، تربیت، پزشکی، توانبخشی، درمانگری، خدمات خانگی و اجتماعی، مراقبت از سالمندان. نتایج این پژوهش‌ها می‌توانند مبنایی برای تدوین مقررات و قوانین لازم در هر یک این حوزه‌ها و کاربردها باشند و دستورالعمل‌هایی را برای توسعه‌دهندگان و کاربران بات‌های اجتماعی ارائه دهند.

جدول ۲) راهکارهای پیشنهادی برای مدیریت رابطه یک طرفه انسان با بات

۱	بات مدام تأکید کند که احساسات ندارد و رابطه‌ای دوطرفه برقرار نمی‌کند	راهکارهای عملیاتی
۲	نظارت و پشتیبانی انسانی برای بات‌های اجتماعی و ارزیابی دوره‌ای اثرات روان‌شناختی و اجتماعی آن‌ها	
۳	فراهم‌سازی امکان بازخورد از سوی کاربران	
۴	توجه ویژه به گروه‌های آسیب‌پذیرتر جامعه برای پیشگیری از مخاطرات پریسک‌تر متوجه آن‌ها	
۵	تشویق شهروندان به برقراری تعاملات انسانی و آگاهی‌بخشی درباره تأثیرات مثبت روابط انسانی و احساسی دوطرفه	
۱	جای‌دهی قدرت استدلال اخلاقی در بات - چالش‌ها: - در حوزه‌های محدودی عملی می‌شود - موقعیت‌ها و کنش‌های محتمل بسیار پیچیده و بی‌شمارند - اجماعی روی مجموعه قواعد اخلاقی نداریم - ترجمه قواعد اخلاقی به زبان محاسباتی و ... عملاً دشوار است - اولویت بندی قواعد و حل تناقض میان قواعد، موقعیت مند و دشوار است	راهکارهای نیازمند بررسی‌های بیشتر و چالش‌های آن
۲	احساساتی مشابه انسان را هوش مصنوعی شبیه‌سازی کنیم - چالش‌ها: - کار بسیار دشواری است - آیا احساسات پیچیده انسانی قابلیت جای‌دهی شدن در بات را دارند؟ آیا این به معنای تقلیل احساسات به مجموعه الگوهای تقلیدی و رفتاری نیست؟ - بات‌ها را در معرض آسیب، رنج و سوء استفاده قرار می‌دهد - بات بر اساس کدام مکاتب اخلاقی از احساسات خود برای تصمیم‌گیری استفاده کند؟ - سرنوشت روابط انسانی پس از ورود بات‌های احساسی چه می‌شود؟	
۳	توقف تولید هوش مصنوعی خودمختار - چالش‌ها: - پیشنهادی منطقی نیست و در عمل امکان‌پذیر نخواهد بود.	

تعارض منافع

نویسندگان تعهد می‌کنند که هیچ تعارض منافی در این مقاله وجود نداشته‌است.

References

- [1] Scheutz, M. (2012). **The Inherent Dangers of Unidirectional Emotional Bonds Between Humans and Social Robots. In Robot Ethics: The Ethical and Social Implications of Robotics** (p. 205-221). MIT Press.
- [2] Cominelli, L., Mazzei, D., & De Rossi, E. (2018). **SEAI: Social Emotional Artificial Intelligence Based On Damasio's Theory Of Mind.** Frontiers in

Robotics and AI.
<https://doi.org/10.3389/frobt.2018.00006>
 [3] Wu, J. (2024). **Social And Ethical Impact of Emotional AI Advancement: The Rise of Pseudo-Intimacy Relationships and Challenges in Human Interactions.** Frontiers in Psychology, 15, Article 1410462. <https://doi.org/10.3389/fpsyg.2024.1410462>
 [4] Shouse, E. (2005). **Feeling, Emotion, Affect.** M/C Journal, 8 (6). <https://doi.org/10.5204/mcj.2443>
 [5] Bechara, A., Damasio, H., & Damasio, A. R. (2000). **Emotion, Decision Making, And the Orbitofrontal Cortex.** Cerebral Cortex, 10 (3), 295–307. <https://doi.org/10.1093/cercor/10.3.295>
 [6] Matsumoto, D., & Hwang, H. (2019). **Functions of Emotions.** In J. Cummings & L. Sanders (Eds.),

- Effectively From Empathic Robots.** Awareness Magazine: Self-Awareness in Autonomic Systems. <https://doi.org/10.2417/3201112.003948>
- [14] Friedenberg, J., & Silverman, G. (2005). **Cognitive Science: An Introduction to The Study of Mind.** SAGE Publications. <https://doi.org/10.4135/9781071802769>
- [15] Kahn, P. H. (1999). **The Human Relationship With Nature: Development And Culture.** MIT Press. <https://doi.org/10.7551/mitpress/3602.001.0001>
- [16] LaGrandeur, K. (2015). **Emotion, Artificial Intelligence, And Ethics.** In J. Romportl, P. Ircing, E. Zackova, M. Polak, & R. Schuster (Eds.), Beyond artificial intelligence: The disappearing human-machine divide (pp. 97–109). Springer. https://doi.org/10.1007/978-3-319-14408-6_10
- [17] Arkin, R. (2007). **Governing Lethal Behavior: Embedding Ethics in A Hybrid Deliberative/Reactive Robot Architecture.** Technical Report, Georgia Institute of Technology.
- [18] Payandeh, R. and Kanani, H. (2024). **Ethical Analysis of Terms and Conditions for the Use of AI-Based Products and Services in Iran.** *Journal of Science and Technology Policy*, 17(3), 27-39. {In Persian}. doi: [10.22034/jstp.2025.11769.1813](https://doi.org/10.22034/jstp.2025.11769.1813)
- [19] Shahbazinia, M. and Zolghadr, M. J. (2024). **Recognizing Artificial Intelligence (AI) As A Legal Person: Providing A Policy Proposal to The Iranian Legislator.** *Journal of Science and Technology Policy*, 17(3), 41-52. {In Persian}. doi: [10.22034/jstp.2025.11778.1819](https://doi.org/10.22034/jstp.2025.11778.1819)
- Introduction to psychology (University of Saskatchewan).
- [7] Sung, J. Y., Guo, L., Grinter, R. E., & Christensen, H. I. (2007). **"My Roomba Is Rambo": Intimate Home Appliances.** UbiComp 2007: Ubiquitous Computing. Berlin: Springer, 145-162. https://doi.org/10.1007/978-3-540-74853-3_9
- [8] Kahn, P. H., Friedman, J., & Hagman, J. (2002). **I Care About Him as A Pal: Conceptions of Robotic Pets in Online Aibo Discussion Forums.** In CHI EA '02: CHI '02 Extended Abstracts on Human Factors in Computing Systems (pp. 632–633). <https://doi.org/10.1145/506443.506519>
- [9] Prescott, T. J., & Robillard, J. M. (2021). **Are Friends Electric? The Benefits and Risks Of Human-Robot Relationships.** *iScience*, 24 (1). <https://doi.org/10.1016/j.isci.2020.101993>
- [10] Bao, A., Zeng, Y., & Lu, E. (2023). **Mitigating Emotional Risks in Human-Social Robot Interactions Through Virtual Interactive Environment Indication.** *Humanities and Social Sciences Communications*, 10, Article 638. <https://doi.org/10.1057/s41599-023-01752-7>
- [11] Boine, C. (2023). **Emotional Attachment To AI Companions And European Law.** MIT Case Studies in Social and Ethical Responsibilities of Computing, Winter 2023 (February). <https://doi.org/10.21428/2c646de5.db67ec7f>
- [12] Weber-Guskar, E. (2021). **How Do Feel About Emotionalized Artificial Intelligence? When Robot Pets, Holograms, And Chatbots Become Affective Partners.** *Ethics and Information Technology*, 23 , 601–610. <https://doi.org/10.1007/s10676-021-09598-8>
- [13] Castellano, G., Leite, I., Paiva, A., & McOwan, P. W. (2012). **Affective Teaching: Learning More**