



Exploring the Ethical Challenge of Emotionalized AI: A Philosophical Perspective on the Meaning of Life with a Focus on Emotional Relationships

Massoud Toossi Saeidi¹ 

Submitted: 2024.11.10

Accepted: 2025.01.19

Abstract

This article explores the ethical dilemma of developing emotionalized artificial intelligence (EAI) through the philosophical lens of meaning of life. It begins by analyzing foundational views from Thomas Nagel and Susan Wolf on absurdity and the role of emotional relationships—particularly love—in meaningful life. The paper then reviews key literature in AI ethics, with a focus on human-robot emotional interaction, and outlines a framework for evaluating EAI from an expanded perspective of technology ethics. The central argument asserts that emotional relationships, especially love, are among the core sources of human meaningfulness, and the integration of EAI into this domain risks undermining that source. Drawing on Verbeek's expanded ethics of technology, the paper contends that EAI functions as a competitor to human affection, thereby restricting our access to genuine emotional engagement. After presenting and defending this argument against potential objections, the article concludes that emotionalized AI raises serious ethical concerns. Without addressing these concerns, its development cannot be considered ethically justified.

Keywords

emotionalized artificial intelligence (EAI), relationships, philosophical absurdity, love


© The Author(s) 2025.



1. Assistant Professor, Institute for Science and Technology Studies, Shahid Beheshti University, Tehran, Iran. (m_tousi@sbu.ac)



بررسی معضل اخلاقی هوش مصنوعی عاطفی شده از منظر مسئله فلسفی معنای زندگی با تأکید بر جایگاه روابط عاطفی

مسعود طوسی سعیدی^۱ 

پذیرش: ۱۴۰۳/۱۰/۳۰

دریافت: ۱۴۰۳/۰۸/۲۰

چکیده

این مقاله به بررسی معضل اخلاقی توسعه هوش مصنوعی عاطفی شده از منظر فلسفی معنای زندگی می‌پردازد. ابتدا با تکیه بر آرای تامس نیگل و سوزان ولف، تحلیل دقیقی از مسأله‌ی فلسفی معنای زندگی و نسبت آن با روابط عاطفی ارائه می‌شود. سپس با مرور ادبیات تخصصی اخلاق هوش مصنوعی، به‌ویژه در حوزه‌ی روابط انسان-ربات، چارچوبی برای تحلیل اخلاقی توسعه‌ی هوش مصنوعی عاطفی شده ترسیم می‌گردد. استدلال اصلی مقاله بر این پایه استوار است که روابط عاطفی، به‌ویژه عشق، یکی از منابع بنیادین معناداری زندگی انسان‌اند و ورود هوش مصنوعی عاطفی شده به این حوزه، با تضعیف این منبع و تهدید معناداری زندگی انسان همراه است. این استدلال با بهره‌گیری از رویکرد گسترش‌یافته‌ی اخلاق فناوری، بر این تأکید می‌ورزد که هوش مصنوعی به عنوان رقیبی برای محبت ورزیدن انسان‌ها در جامعه عمل می‌کند و از این طریق، دسترسی ما به محبت ورزیدن را محدود می‌سازد. در ادامه و پس از صورت‌بندی این استدلال، به انتقادهای احتمالی وارد بر آن پاسخ داده می‌شود. در نتیجه‌گیری، مقاله نشان می‌دهد که هوش مصنوعی عاطفی شده، از این نظر دارای ملاحظات اخلاقی جدی است که بدون مراعات آن‌ها، توسعه‌ی آن از منظر اخلاقی موجه نیست.

کلیدواژه‌ها

هوش مصنوعی عاطفی شده، روابط، پوچی فلسفی، زندگی معنادار، عشق

رتال جامع علوم انسانی

۱. استادیار پژوهشکده مطالعات بنیادین علم و فناوری، دانشگاه شهید بهشتی، تهران، ایران. (m_tousi@sbu.ac)

۱. مقدمه: مسئله فلسفی معنای زندگی و روابط عاطفی

از دهه ۱۹۸۰، ادبیات مشخصی درباره «معنای زندگی» در فلسفه انگلیسی‌زبان ایجاد شده است (Metz 2013). من هم در اینجا همین ادبیات را مد نظر قرار داده‌ام و تلاش کرده‌ام از پراجاع‌ترین و مقبول‌ترین دیدگاه‌ها برای بررسی موردنظر این مقاله درباره نسبت روابط عاطفی با معنای زندگی بهره ببرم.

از جمله معتبرترین تحلیل‌های فلسفی درباره این معضل تحلیلی است که توسط تامس نیگل، در مقاله‌ای با نام «پوچی» (*The Absurd*) ارائه شده است. نیگل منشأ پوچی را فاصله‌ای می‌داند که میان واقعیت‌ها از یک سو و آرمان‌ها و ادعاهای ما از سوی دیگر پیش می‌آید (Nagel 1971, 718). البته به این معنا، زندگی‌های بسیاری به دلیل انتخاب‌های افراد و وضعیت‌هایی که پیش آمده است پوچ است، اما این پوچی‌ها اهمیت فلسفی ندارد. نیگل پوچی فلسفی زندگی را - که خود قائل به آن است - به این شکل صورت‌بندی می‌کند:

ما نمی‌توانیم بدون انرژی و توجه، زندگی انسانی داشته باشیم یا بدون انتخاب‌هایی که نشان می‌دهد برخی چیزها را جدی‌تر از برخی دیگر می‌گیریم. با این حال، ما همیشه منظری خارج از شکل خاص زندگی خود هم داریم، که این جدی‌گرفتن‌ها از آن منظر بیهوده به نظر می‌رسد. این دو منظر اجتناب‌ناپذیر در ما با هم برخورد می‌کنند و این همان چیزی است که زندگی را پوچ می‌کند. پوچ است زیرا ما شک و تردیدهایی را که می‌دانیم قابل حل نیست نادیده می‌گیریم؛ یعنی با وجود آن‌ها، با جدیت تقریباً کاهش‌نیافته‌ای به زندگی خود ادامه می‌دهیم (Nagel 1971, 719).
[...] [پوچی] نه به این دلیل [است] که می‌توان چنین منظری از خارج به خود داشت، بلکه به این دلیل که می‌توانیم آن را اتخاذ کنیم بی آنکه آن افرادی که دغدغه‌های بنیادینشان این‌گونه سرد و بی‌رُمق شده نباشیم (Nagel 1971, 722).

به این ترتیب بنا بر تحلیل نیگل، پوچی فلسفی زندگی انسان یک ویژگی اساسی زندگی اوست و ریشه‌اش «وجود یک شک استعلایی در او درباره کلیت آنچه در زندگی مهم می‌پندارد» و در عین حال «به لوازم آن شک پایبند نبودن» است. اقتضای شک این است که از اهمیت آنچه درباره آن شک کرده‌ایم کاسته گردد، درحالی که ما زندگی را دقیقاً برخلاف این اقتضا ادامه می‌دهیم.

بر این اساس منظور از مسئله معنای زندگی در این بررسی، یک مسئله ذهنی روان‌شناختی یا عاطفی درباره زندگی نیست، بلکه در اینجا تأکید بر جنبه فلسفی این موضوع است. نتیجه تأکید بر مسئله فلسفی معنای زندگی این است که ممکن است بسیاری افراد از حیث روان‌شناختی، از زندگی خود رضایت و نسبت به آن عواطف خوش‌آیندی داشته باشند، اما زندگی آن‌ها پوچ باشد. همچنین - و با این فرض که مسئله پوچی فلسفی زندگی پاسخی داشته باشد - ممکن است برخی افراد از حیث روان‌شناختی دچار افسردگی بوده یا عواطف آن‌ها نسبت به زندگی شان توأم با گلایه و نارضایتی شدید باشد، اما زندگی آن‌ها - از منظر عینی^۱ - معنادار باشد.

مباحث فلسفی پیرامون معنای زندگی را به دو دسته کلی تقسیم کرده‌اند. این دو دسته را بیشتر با دو عنوان استاندارد «معنای زندگی»^۲ و «معنا در زندگی»^۳ معرفی می‌کنند (Seachris 2019, 363). تادئوس متر^۴ این تفکیک را به این صورت توضیح می‌دهد که:

«معنا در زندگی» جایی به کار می‌رود که زندگی یک انسان می‌تواند معنادار باشد، و در «معنای زندگی» [...] زندگی گونه انسان به عنوان یک کل، چیزی است که می‌تواند معنادار باشد یا نه. [...] اکثر فیلسوفان تحلیلی به معنا در زندگی

علاقه‌مند بوده‌اند؛ یعنی به معنایی که زندگی یک فرد می‌تواند داشته باشد، درحالی‌که تعداد کمی از آن‌ها به معنای زندگی می‌پردازند (Metz 2013, Section 1).

در این بررسی، به همین مباحث پرتعدادتر از معنای زندگی توجه خواهد شد و بر این اساس اصطلاح «زندگی معنادار» را به معنای «معنا در زندگی» به کار خواهیم برد. البته تلاش خواهد شد که در نهایت تحلیلی یکپارچه از نسبت میان روابط عاطفی و زندگی معنادار ارائه شود، به گونه‌ای که از نظر معیارهای هر دو منظر قابل فهم و تصدیق باشد. یک دلالت این مباحث، تعلیق پاسخ‌دادن به اصل مسئله پوچی و تلاش برای یافتن منابع عظیم معناداری در زندگی^۵ است. توجه به زندگی مشاهیری نظیر قدیسه ترزا و نلسون ماندلا و تصدیق ارزشمند بودن و معنادار بودن زندگی آنان نزد عقل سلیم، این ایده را که «حتی اگر اصل زندگی نوع انسان پوچ باشد باز هم می‌توان به معنایی به زندگی معنادار نائل شد» نزد فیلسوفان تحلیلی موجه کرده است (Metz 2011, 389-390).

یکی از پرارجاع‌ترین فیلسوفان تحلیلی در این حوزه سوزان ولف^۶ است. او در ۲۰۱۶ نقطه آغاز تحلیل خود در باب زندگی معنادار را بررسی انواع زندگی‌های پوچ در تلقی رایج قرار می‌دهد. ولف سه وضعیت زندگی پوچ را برمی‌شمرد: (۱) زندگی انفعالی توأم با مصرف و فاقد هرگونه ثمردهی و دستاورد همچون زندگی یک فرد بیکار-ولو ثروتمند- که صرفاً به عیش و نوش می‌گذرد،^۷ (۲) زندگی پر از فعالیت، اما فاقد هدف درخور، همچون زندگی یک پرورش‌دهنده خوک که تمام وجودش صرف خرید زمین بیشتر برای غذای بیشتر برای خوک‌های بیشتر برای خرید زمین بیشتر می‌شود،^۸ و (۳) زندگی کسی که فعالانه و برای خلق یک ارزش واقعی در تلاش است اما شکست اساسی می‌خورد، مانند زندگی کسی که با تلاش فراوان کارخانه‌ای احداث می‌کند تا کالایی را عرضه کند اما پس از عمری تلاش و در آستانه ثمردهی، کالای مدنظر او به دلیل توسعه فناوری منسوخ می‌شود^۹ (Wolf 2016b, 116-117).

او با نظر به این سه وضعیت و در راستای زدودنشان از زندگی، پیشنهاد کلی‌اش برای تعریف زندگی معنادار را ارائه می‌دهد: «زندگی معنادار زندگی‌ای است که به‌طور فعالانه^{۱۰} و حداقل تا حدودی با موفقیت نسبت به پروژه (یا پروژه‌هایی) با ارزش مثبت^{۱۱} تعهد و درگیری^{۱۲} داشته باشد» (Wolf 2016b, 118).

ولف از جمله فیلسوفانی است که تلاش می‌کند نشان دهد نمی‌توان معنای زندگی را به مقوله خوشبختی^{۱۳} فروکاست (Metz 2013, sec. 1; Wolf 2016a). هرچند تصریح دارد که زندگی معنادار در اغلب قریب به اتفاق موارد خوش‌تر است. به بیان دقیق‌تر، او معتقد است ارزش‌هایی که زندگی معنادار را تضمین می‌کنند لزوماً ارزش‌های مربوط به خوشبختی نیستند و در مواردی، ارزش زندگی معنادار اقتضا می‌کند درد را - که خلاف خوشبختی است - انتخاب و تحمل کنیم (Wolf 1997, 222). به این ترتیب، از نظر ولف، همچون بسیاری دیگر از فیلسوفان تحلیلی، معناداری زندگی یک ارزش مستقل برای زندگی است که ارزش زیستن را فراهم می‌آورد: «آنچه به زندگی ما معنا می‌بخشد به ما دلایلی برای زیستن هم می‌دهد [تأکید از من]؛ حتی زمانی که خوشبختی دورنمای تیره‌ای داشته باشد» (Wolf et al. 2012, 56).

دو معیار «تعهد و درگیری فعالانه» و «ارزش‌های مثبت» در دیدگاه ولف سبب شده است که دیدگاهش به دیدگاه ترکیبی^{۱۴} مشهور شود و از این جهت، بسیار هم اثرگذار بوده و مقبول افتاده است. او با به‌کاربردن این اصطلاحات تأکید دارد که زندگی معنادار هم از منظر آفاقی و عینی ارزشمند است و هم از منظر ذهنی و شخصی. ولف تصریح می‌کند: «معنا زمانی محقق می‌شود که جاذبه ذهنی^{۱۵} با جذابیت عینی^{۱۶} توأمان شوند» (Wolf et al.).

9, 2012). معنای این سخن این است که اگر کسی نسبت به پروژه‌ای که واقعاً ارزشمند نیست تعلق و درگیری فعالانه داشته باشد، یا اگر پروژه‌ای واقعاً ارزشمند را - ارزشی که به زندگی معنا دار ختم می‌شود - عهده‌دار شود اما در درونش به آن تعلق نداشته باشد، زندگی‌اش معنا دار نخواهد بود.

ناظر به معیار «ارزش‌های مثبت»، ولف یک رکن اساسی برای بررسی اینکه چیزی واقعاً ارزشمند است یا نه را بررسی «استقلال از خود فرد»^{۱۷} می‌داند. چیزی که ارزشش به‌طور کامل تابع علایق شخصی است نمی‌تواند به زندگی معنا ببخشد (Wolf et al. 2012, 36-38). مضمون این سخن به دیدگاه مشهور دیگری در فلسفه تحلیلی بسیار نزدیک است که معیار عینی برای زندگی معنا دار را «فراترفتن از خود»^{۱۸} می‌داند و در راستای تنقیح آن تلاش می‌کند (Metz 2011).

ماحصل دیدگاه ولف را می‌توان در مثال‌هایی که برای زندگی معنا دار به کار می‌برد مشاهده نمود، مثال‌هایی که «فهم دلایل و انگیزه‌های مربوط به آن‌ها، از طریق بررسی سهمشان در خوشبختی یا دریافت‌مان از آنچه عقل یا اخلاق غیرشخصی ایجاب می‌کند به‌طور کامل محقق نمی‌شود» و «بهترین منظر برای فهم ارزش آن‌ها منظر زندگی معنا دار است:

بارزترین نمونه‌ها نمونه‌هایی هستند که در آن‌ها از روی عشق [تأکید از من] به افراد خاص عمل می‌کنیم. وقتی برادرم را در بیمارستان ملاقات می‌کنم، یا به دوستم کمک می‌کنم اسباب‌کشی کند، یا تمام شب را بیدار می‌مانم و لباس‌های هالووین را برای دخترم می‌دوزم، نه به دلایل خودخواهانه^{۱۹} و نه به دلایل اخلاقی^{۲۰} رفتار می‌کنم. فکر نمی‌کنم برای من بهتر باشد که لحظات افسرده‌کننده‌ای را در اتاقی تنگ بگذرانم، برادرم را عصبی و در رنج و درد بینم، یا در تلاش برای پایین‌آوردن مبل دوستم از طبقه دوم از پله‌ها به کمرم آسیب بزنم، یا از خوابی که خیلی می‌خواستمش صرف‌نظر کنم تا مطمئن شوم که بال‌های لباس پروانه‌ای دخترم که می‌خواهد در رژه روز بعد بپوشد با زاویه‌ای خوب برجسته شود. من خودم را موظف به انجام این اعمال نمی‌دانم و خودم را فریب نمی‌دهم که با انجام آن‌ها کاری را انجام دهم که برای دنیا بهترین است. من نه از روی نفع شخصی و نه از سر وظیفه یا هر نوع دلیل بدون غرض‌ورزی عمل می‌کنم. بلکه از روی عشق عمل می‌کنم. (Wolf 2016a, 254-255)

تحلیل ولف اگرچه با مسئله «معنا در زندگی» و بررسی عینی‌تر و تجربی‌تر پوچی آغاز می‌شود، اما ماحصل آن را می‌توان به‌منظور صورت‌بندی پاسخی به چالش پوچی عمیق تقویت نمود. نسبت انسان دارای رابطه عاطفی - به‌ویژه برجسته‌ترین نمونه این رابطه؛ یعنی عشق - با شک (تعبیر نیگل از ریشه پوچی عمیق) به‌کلی متفاوت است. معمولاً این‌گونه است که محبت مجاللی برای شک کردن ندارد، و زمانی که محبت به کمال می‌رسد - نظیر محبت و مهر مادری نسبت به طفل شیرخواره خود - هرگونه شک در ضدیتی قوی با محبت قرار می‌گیرد. به این ترتیب چه‌بسا طرح کلی ولف بتواند - حداقل در مثال‌هایی - حتی به مسئله پوچی عمیق یا فلسفی نیز پاسخ دهد.

این جمع‌بندی ویژگی دیگری نیز دارد. اگرچه تفکیکی میان دیدگاه‌های فراطبیعت‌گرایانه و توحیدی^{۲۱} و دیدگاه‌های طبیعت‌گرایانه درباره معنای زندگی در ادبیات تخصصی این حوزه به رسمیت شناخته شده است، اما در طرح کلی ولف چنین تفکیکی ضرورت چندانی ندارد. عشقی که متعلق آن واقعاً ارزشمند باشد، منبع معناداری زندگی است؛ خواه عشق مادری و خواه عشق الهی. از این رو این جمع‌بندی را می‌توان دارای نوعی تناسب با جایگاه عشق نزد عرفا قلمداد نمود:

با مدعی مگوئید اسرار عشق و مستی

تا بی‌خبر بمیرد در درد خودپرستی
عاشق شو از نه روزی کار جهان سرآید
ناخوانده نقش مقصود از کارگاه هستی^{۲۲}
جسم خاک از عشق بر افلاک شد
کوه در رقص آمد و چالاک شد
عشق، جان طور آمد عاشقا!
طور، مست و خَرّ موسی صاعقاً^{۲۳}

بنابراین هرگاه ما از روی عواطف و از سر محبت نسبت به کسی یا چیزی که ارزشش را دارد تعهد و درگیری فعالانه پیدا کنیم، در راستای پاسخی به نیازمان به زندگی معنادار گام برداشته‌ایم. معناداربودن زندگی ارزشی غیر از اخلاقی بودن یا توأم با خوشبختی بودن آن است و از این‌رو نیازی مستقل است. و از آنجایی که پاسخ به نیاز معناداری پاسخ به چرایی زیستن را هم در پی دارد، نیازی اساسی محسوب می‌شود. این تحلیل، مبنای این مقاله برای بررسی اخلاقی هوش مصنوعی عاطفی شده است.

۲. اخلاق هوش مصنوعی، معنای زندگی و روابط

دغدغه‌های اخلاقی پیرامون هوش مصنوعی پیوندهایی با موضوع معنای زندگی پیدا کرده است. این پیوندها را نایهولم و روثر (Nyholm & Rütger 2023) تحلیل و دسته‌بندی کرده‌اند. دسته‌بندی آن‌ها نقطه آغاز خوبی برای بیان مسئله موردنظر این مقاله است. آن‌ها تأکید می‌کنند که پیوند میان اخلاق هوش مصنوعی و معنای زندگی پیوند بااهمیتی است و چارچوبی را برای تلاش بیشتر درباره این موضوع ترسیم می‌کنند (Nyholm & Rütger 2023, 18-21). بنا بر تحلیل آن‌ها، بررسی مسائل اخلاقی پیرامون هوش مصنوعی، زمانی که با مسئله معنای زندگی پیوند می‌خورد، ذیل سه عنوان کلی قابل بازشناسی است:

- خودسازی^{۲۴}
- آینده‌کار^{۲۵}
- روابط^{۲۶}

هر یک از این سه عنوان، وقتی که از منظر مسئله معنای زندگی به آن‌ها نظر شود، می‌تواند - بنا بر دیدگاه‌های متنوع مطرح شده از سوی متفکران مختلف - منبعی برای معنادارشدن زندگی قلمداد شود. از این‌رو، زمانی که مشخص شود محصولات مبتنی بر فناوری هوش مصنوعی ممکن است بر وضعیت این امور - یعنی خودسازی، آینده‌کار و روابط - اثرات عمیقی داشته باشند، دغدغه‌های اخلاقی درباره هوش مصنوعی جدی می‌شوند.

نایهولم و روثر، ضمن بررسی این دغدغه‌ها، نشان می‌دهند هنوز یک جمع‌بندی قطعی - یا حتی تقریباً قطعی - از اینکه هوش مصنوعی در نهایت به نفع معنادارشدن زندگی است یا به زیان آن، وجود ندارد (Nyholm & Rütger 2023, 8). این دو نفر با مرور مقایسه‌ای استدلال‌هایی که تا امروز ارائه شده است، نشان می‌دهند در ازای هر استدلالی که می‌گوید هوش مصنوعی تهدیدی برای معناداری زندگی است، استدلال دیگری داریم که دقیقاً خلاف این را می‌گوید.

به عنوان مثال و به بیانی ساده شده، یک مقدمه کلیدی در قالبی کلی از استدلال‌ها - مربوط به دسته آینده کار- این است که ما با انجام دادن برخی کارها و وظایف، به زندگی مان معنا می‌دهیم و از آنجایی که هوش مصنوعی جایگزین ما برای انجام کارها می‌شود، از فرصت‌های ما برای انجام کارها می‌کاهد و لذا تهدیدی برای معناداری زندگی ما محسوب می‌شود. در سوی مقابل، قالب دیگری از استدلال‌ها قرار دارند که می‌گویند ما در زندگی به بسیاری از امور اشتغال داریم که هیچ معنایی به زندگی ما نمی‌دهند بلکه مانعی برای انجام امور بااهمیت‌ترند، و از آنجایی که هوش مصنوعی می‌تواند این امور را به جای ما انجام دهد، با توسعه آن، ما فرصت بیشتری برای انجام کارهای معنادهنده به زندگی خواهیم داشت و به این ترتیب هوش مصنوعی در مجموع فرصتی برای معنادار شدن زندگی ما محسوب می‌شود (Nyholm & Rüter 2023, 12-14).

مشابه همین وضعیت برای سایر استدلال‌هایی که ذیل دو دسته «خودسازی» و «روابط» قرار می‌گیرند هم برقرار است (Nyholm & Rüter 2023, 9-11, 15-18). بررسی من در این مقاله در دسته بررسی‌های با محوریت «روابط» قرار می‌گیرد. از این رو مباحث مربوط به اخلاق هوش مصنوعی، معنای زندگی و روابط در ادامه با تفصیل بیشتری بررسی خواهد شد.

همچنین در اینجا باید گفته شود که تلقی من از محدوده مباحث اخلاق فناوری متناسب با دیدگاه گسترش یافته^{۲۷} وربیک^{۲۸} است. او تأکید دارد فناوری صرفاً یک ابزار دارای کارکرد و مجزا از زمینه‌های اجتماعی نیست، بلکه وقتی یک فناوری به کار گرفته می‌شود مناسبات و زمینه اجتماعی را هم تحت تأثیر خودش قرار می‌دهد (Verbeek 2008, 91-92). از این رو رویکرد من به اخلاق هوش مصنوعی محدود به تلقی ابزاری از آن و آثار کاربردی‌اش به مثابه یک ابزار برای افراد نیست بلکه به آثار گسترده‌تر آن در زمینه اجتماعی نظر خواهم داشت.

در همین راستا و به منظور محدودتر کردن دامنه بررسی، دقیق‌تر کردن مسئله و قطعی‌تر کردن نتایج، من تلاش می‌کنم اثر کاربردی شدن هوش مصنوعی احساساتی^{۲۹} - اصطلاحی که اوا ویر-گوسکار (Weber-Guskar 2021) آن را پیشنهاد می‌دهد- در بات‌ها و ربات‌ها را بر زمینه اجتماعی از منظر مسئله معنای زندگی بررسی کنم. در ادامه صورت‌بندی دقیق‌تری از مسئله ارائه خواهد شد.

۱-۲. روابط و اخلاق هوش مصنوعی: مرور ادبیات

دغدغه‌های اخلاقی پیرامون تعامل انسان و ربات متناسب با رویکرد گسترش یافته وربیک، که در بالا به آن اشاره شد - حتی قبل از پیشرفت‌های سال‌های اخیر در حوزه هوش مصنوعی و محصولات مبتنی بر آن- مباحث و ادبیات تخصصی مشخصی را باعث شده است. دی گراف (De Graaf 2016) برآورد خوبی از این موضوع ارائه می‌دهد. او قرائن فراوانی را مد نظر قرار می‌دهد که تأیید می‌کنند انسان‌ها هم به برقراری روابطی از نوع انسان-انسان (با تأکید بر روابط عاطفی و اجتماعی) با ربات‌ها تمایل دارند، هم به برقراری چنین رابطه‌ای قادرند و هم در عمل چنین رابطه‌ای را برقرار می‌سازند. علاوه بر این، ربات‌ها هم برای اینکه بتوانند اهدافی را که برایشان در نظر گرفته شده محقق کنند، قبل از هر چیزی باید در تعاملات اجتماعی و برقراری روابط با کاربرانشان مشارکت نمایند (De Graaf 2016, 589-590)؛ یعنی ربات‌ها هم برای اینکه کار کنند باید به گونه‌ای طراحی شوند که رابطه‌ای بسیار شبیه به روابط اجتماعی با انسان‌ها تعریف کنند.

در نظر دی گراف، ربات‌ها تمایزهایی با سایر فناوری‌ها دارند که به آن‌ها در موضوع روابط، اهمیتی اساسی می‌بخشد. او به بدن‌مندی و حضور فیزیکی ربات‌ها و اطلاع آن‌ها از محیط پیرامونشان اشاره می‌کند که مقتضی این است که ما جهان فیزیکی‌مان را با آن‌ها به اشتراک بگذاریم و در نتیجه، حضور آن‌ها را در اجتماع خودمان به رسمیت بشناسیم، وضعیتی که آن را دربارهٔ سایر فناوری‌ها سراغ نداریم. علاوه‌براین، بررسی‌ها نشان می‌دهد ما به سبب خودکار^{۳۰} بودن ربات‌ها و وجود قابلیت‌هایی نظیر شنیدن فرامین و ارتباط زبانی، بیشتر آن‌ها را از سنخ حیوانات یا سایر اشخاص تلقی می‌کنیم تا مصنوعات فناورانه (De Graaf 2016, 593).

بررسی‌هایی که بر برقراری روابط عاطفی (چه از حیث روان‌شناختی و چه از حیث فیزیولوژیک) میان انسان و ربات دلالت دارند (Wada & Shibata 2007) قرینه‌ای دیگر بر اهمیت مسئلهٔ رابطهٔ انسان-ربات و متمایز بودن این رابطه با سایر روابط مربوط به انسان و فناوری محسوب می‌شود. به این ترتیب، نزدیکی رابطهٔ انسان-ربات به رابطهٔ انسان-انسان -وقتی از بیرون به آن نظر بیفکنیم- یک واقعیت است. این واقعیت به دغدغه‌های اخلاقی از منظر مقولهٔ روابط و آثار کاربردی‌شدن ربات‌ها در متن زندگی انسان‌ها، حتی قبل از آنکه آن‌ها با هوش مصنوعی تقویت شوند، اهمیت مضاعفی بخشیده است.

مرور این ادبیات نشان می‌دهد که تا به امروز بسیاری از دغدغه‌های اخلاقی در باب رابطهٔ انسان-ربات حول دو مسئلهٔ «فریب»^{۳۱} و «دوسویگی»^{۳۲} مطرح شده است. طرح دغدغه به این صورت است که ما با ربات‌ها روابطی از سنخ روابطمان با اشخاص برقرار می‌کنیم، درحالی‌که او واقعاً شخص نیست؛ یعنی ما گمان می‌کنیم ربات ویژگی‌هایی دارد که در واقع آن‌ها را ندارد اما با فرض اینکه آن ویژگی‌ها را دارد، با او رابطه‌ای متناسب با آن ویژگی‌ها برقرار می‌کنیم (Sparrow & Sparrow 2006, 141-161; Turkle 2012; Wallach & Allen 2010). با این نقطه آغاز، دو سؤال اولیه پیش روی متفکران قرار می‌گیرد: «آیا این فریب است یا نه؟ و آیا [از نظر اخلاقی] غلط است یا نه؟» (De Graaf 2016, 593).^{۳۳}

تا آنجایی که مفهوم «فریب» مد نظر بوده، به کارگیری هرچه بیشتر هوش مصنوعی در بات‌ها و ربات‌ها دغدغه‌های اخلاقی را جدی‌تر برانگیخته است. معرفی کردن بات‌ها و ربات‌های هوشمند به مثابهٔ موجودات ذی‌شعور^{۳۴} علائق و توجهات زیادی را به خود جلب می‌کند. موضوع هر از گاهی حتی برای شرکت‌ها و مؤسسه‌های مانند گوگل، تسلا و اوپن‌ای‌آی هم حاشیه‌هایی درست می‌کند (Nyholm & Rüter 2023, 18) که معروف‌ترین آن اخراج بلیک لموین^{۳۵} از گوگل در جولای ۲۰۲۲ است.^{۳۶} هوش مصنوعی از آنجایی که قابلیت‌های -حداقل در ظاهر- انسانی را در بات‌ها و ربات‌ها افزایش می‌دهد، دغدغهٔ «فریب» را جدی‌تر می‌کند، به‌ویژه زمانی که این قابلیت‌ها شامل عواطف هم بشود.

عاطفی‌شدن بات‌ها و ربات‌ها به این معناست که آن‌ها بتوانند تشخیصی دربارهٔ عواطف کاربرانشان داشته باشند، بر عواطف آن‌ها مؤثر باشند و از خودشان نیز عواطفی بروز دهند. این ویژگی حداقل از دو جنبه، تعامل انسان با آن‌ها را تقویت می‌کند: اول اینکه آن‌ها را از محصولات نامأنوس^{۳۷} و غریب^{۳۸} به محصولات آشنا^{۳۹} و بلکه صمیمی^{۴۰} تبدیل می‌کند و در نتیجه زمینه‌های تحقق ارتباطات از نوع انسان-انسان به شدت افزایش می‌یابد. دوم اینکه باعث می‌شود این مصنوعات در حوزه‌هایی نظیر روان‌درمانی و بسیاری از خدمات مراقبتی (نظیر مراقبت از سالمندان و کودکان) که وجود عواطف برای انجام آن‌ها ضروری است و قبلاً صرفاً توسط انسان‌ها انجام می‌شد و اساسش

برقراری ارتباطات عاطفه‌محور انسانی بود نیز کاربردی شوند. از این رو بات‌ها و ربات‌های برخوردار از هوش مصنوعی عاطفی شده منتقدان را نسبت به ابعاد اخلاقی توسعه آن‌ها حساس‌تر کرده است (Weber-Guskar, 2021, 601).

به‌طور کلی - و حتی تا حدودی فارغ از دو مسئله «فریب» و «دوسویگی» - در برقراری روابطی از نوع انسان-انسان میان انسان‌ها و بات‌ها یا ربات‌ها، هرچه عواطف نقش محوری‌تری بازی کنند و نوع رابطه عاطفی‌تر باشد دغدغه‌های اخلاقی بیشتر می‌شود. مثلاً انسان ممکن است در محیط‌های کاری با ربات‌ها رابطه همکاری^{۴۱} تعریف کند.

نایهولم و اسمیدز مسئله برقراری رابطه کاری و مقایسه آن با رابطه عاطفی را بررسی کرده‌اند (Nyholm & Smids 2020). آن‌ها نشان داده‌اند که همکار ایده‌آل بودن مؤلفه‌های مختلفی دارد و در سطح رفتاری، ربات‌ها می‌توانند مطابق با بسیاری از معیارهایی که معمولاً با همکار خوب بودن مرتبط هستند (نظیر یادگیری، تعامل سازنده، خروجی خوب دادن و غیره)، عمل کنند (Nyholm & Smids 2020, 2177). در راستای مقایسه این نوع رابطه با رابطه عاطفی^{۴۲}، آن‌ها استدلال می‌کنند که ایده‌آل‌های سخت‌تری مربوط به روابط عاطفه‌محور داریم و لذا برای یک ربات ساده‌تر است که به ایده‌آل‌های یک همکار خوب نزدیک شود تا یک شریک عاطفی. دلیلش این است که «درون» دوستان ما - یعنی آنانی که طرف رابطه عاطفی ما هستند - برای ما مهم‌تر از «درون» همکارانمان است (Nyholm & Smids 2020, 2184-2192).

به نظر می‌رسد این یک صورت‌بندی دقیق و عمیق است. در بررسی اخلاقی ارتباط انسان و ربات از منظر دغدغه «فریب»، بسیاری از محققان در عمل در باب ویژگی‌های مربوط به «درون» ربات بحث کرده‌اند. در برقراری نوعی رابطه مشابه با رابطه انسان-انسان، ما علی‌الاصول نوعی تلقی انسان‌وار از ربات داریم؛ یعنی یک ویژگی انسانی را به ربات نسبت می‌دهیم. «اراده»، «آگاهی» و «عاطفه» ویژگی‌هایی هستند که بیشترین توجه را به خود جلب کرده‌اند. برخی نسبت به چنین تعمیمی انتقاد دارند و توسعه چنین ربات‌هایی را شایسته ملاحظات اخلاقی می‌دانند (Nyholm & Rüter 2023, 15).

استدلال آن‌ها این است که می‌گویند انسانی که طرف رابطه عاطفی با یک ربات است لاجرم به‌صورت پیش‌فرض ربات را هم واجد عاطفه می‌داند، درحالی‌که این‌طور نیست - ربات‌ها صرفاً ماشین هستند - و در نتیجه، او فریب خورده است (فارغ از اینکه خود را فریب داده است یا دیگری او را فریبته است). این قسم استدلال‌ها به دو شکل پاسخ داده شده است. برخی با موشکافی مقدمات و تمرکز بر مسئله «فریب» توضیح می‌دهند که این پیش‌فرض گرفتن وجود واقعیت‌های عاطفی لزوماً به معنای باور به وجود چنین واقعیاتی نیست. مثلاً وبر چنین روابطی را با روابطی که ما با یک اثر مهیج سینمایی یا یک رمان و برقراری ارتباط عاطفی با شخصیت‌های آن داریم مشابه دانسته است. به این ترتیب آن‌ها استدلال می‌کنند که این یک «فریب» نیست و اگر هم باشد، ملاحظه جدی اخلاقی ندارد و به علاقه ما نسبت به تصورات خیالی و مهیج بازمی‌گردد (Weber-Guskar 2021, 604-606).

شکل دوم پاسخ این است که معلوم نیست ربات‌ها واقعاً عاطفه نداشته باشند. این شکل را جان دانا هر در مقاله پراجاعش در ۲۰۱۹ پی گرفته است. او یادآوری می‌کند این بحث که ربات‌های هوشمند، به معنای قوی متافیزیکی، نمی‌توانند از ویژگی‌های متقابل برای برقراری رابطه عاطفی و دوستانه برخوردار باشند هنوز جمع‌بندی

قطعی ندارد و لذا مقدمه‌ای اساسی‌تر از استدلال فریب (اینکه ربات‌ها عاطفه ندارند یا از حیث درونی هم‌تراز انسان نیستند) مخدوش است؛ هرچند ممکن است در حال حاضر از نظر فنی^{۴۳} محدودیت‌هایی وجود داشته باشد اما از نظر متافیزیکی تکلیف معلوم نیست. علاوه‌براین او از منظری دیگر، به فواید کاربردی ربات‌ها برای بهبود روابط عاطفی معنادار میان انسان‌ها اشاره می‌کند - مثل اینکه یک ربات هوشمند نقش مشاور را برای بهبود روابط ایفا کند - و به این ترتیب، از زاویه‌ای دیگر آن‌ها را برای برقراری روابط معنادار مفید می‌داند (Danaher 2019, 12-23).

به موازات دغدغه «فریب»، دغدغه «دوسویگی» مطرح می‌شود که می‌گوید رابطه خوب عاطفی رابطه‌ای است که در آن عواطف دوسویه باشند: چه خوش بی‌مهربانی هر دو سر بی! درحالی‌که می‌دانیم هوش مصنوعی عاطفی شده به‌راستی چنین رابطه‌ای را برقرار نمی‌کند (دو نمونه از این استدلال در: Nyholm and Frank 2017; Hauskeller 2017). و بر-گوسکار این قسم استدلال‌ها را با استناد به یافته‌های روان‌شناسان و به این ترتیب پاسخ می‌دهد که برخی یافته‌های آنان بر این دلالت دارد که رابطه خوب مبتنی بر تعامل قوی و پایدار محقق می‌شود، تحلیلی که در آن اثری از دوسویگی نیست (Weber-Guskar 2021, 606).

۲-۲. روابط عاطفی و اخلاق هوش مصنوعی: یک منظر جدید

بحث‌ها درباره ملاحظات برقرارشدن رابطه‌ای شبیه به رابطه انسان-انسان میان انسان‌ها و ربات‌ها از دو منظر «فریب» و «دوسویگی» ادامه دارد، اما من گمان می‌کنم منظر دیگری که و بر-گوسکار آن را «غفلت از اخلاق»^{۴۴} می‌نامد منظری اساسی‌تر است.^{۴۵} هر دو دغدغه «فریب» و «دوسویگی» مربوط به رابطه فرد و هوش مصنوعی است، درحالی‌که رویکرد گسترش‌یافته وربیک بر توجه نسبت به امور دیگری تأکید دارد. «غفلت از اخلاق» چنین رویکردی را پیگیری می‌کند. اینکه رابطه میان انسان و هوش مصنوعی برای سایر افراد در اجتماع چه معنایی می‌تواند داشته باشد ذیل این عنوان بررسی شده است (Weber-Guskar 2021, 608). این بررسی همچنین با پیشنهاد دی‌گراف سازگار است که سعی می‌کند اخلاق رابطه انسان-ربات را فارغ از مناقشاتی که به درون ربات^{۴۶} مربوط است بررسی کند و تأکید می‌کند بررسی‌های اخلاقی درباره ربات‌ها باید به «بیرون» از ربات‌ها منتقل شود (De Graaf 2016, 592).

و بر-گوسکار یکی از چالش‌های مرتبط با «غفلت از اخلاق» را «اتلاف منابع مهم عاطفی»^{۴۷} می‌نامد. او استدلال مربوط به این عنوان را این‌گونه صورت‌بندی کرده است: «انسان‌ها برای رشد و شکوفایی نیاز به روابط عاطفی انسانی دارند. [تحقق] روابط عاطفی مستلزم استفاده از منابع محدود (زمان و انرژی عاطفی) است. EAI می‌تواند بخشی از یک رابطه عاطفی باشد، اما برای رشد کردن و شکوفاشدن نیازی به روابط عاطفی ندارد (زیرا در نظر اولیه، رشد و شکوفایی برایشان ناممکن است). صرف منابع محدود برای کسی [یا چیزی] که به آن‌ها نیاز ندارد، هدر دادن این منابع است. (از این‌ها نتیجه می‌شود) افرادی که روابط مؤثری با EAI دارند منابع عاطفی را تلف می‌کنند. هدر دادن منابع محدود بد است. [پس] ارتباط با EAI بد است». و بر-گوسکار برای پاسخ به این استدلال، اشاره می‌کند که بهره‌مندشدن از روابط عاطفی، اگر هم یک نیاز باشد، یک حق نیست. به این ترتیب او نتیجه می‌گیرد درباره آنانی که با مصرف منابع عاطفی‌شان در جای دیگر ما را از عواطف خودشان محروم کرده‌اند نمی‌توان نتیجه اخلاقی گرفت (Weber-Guskar 2021, 609).

من در این مقاله سعی می‌کنم با بازنگری اساسی در این استدلال، استدلال قوی‌تری ارائه دهم که نشان می‌دهد استفاده از EAI منابع بسیار محدود ما برای معناداری زندگی را محدودتر می‌کند و از آنجایی که ما به معنای زندگی نیاز

داریم، ملاحظه اخلاقی اساسی درباره EAI یا حداقل برخی انواع خاص آن وجود دارد. به این ترتیب نشان می‌دهم پاسخی که وبر-گوسکار از طریق بازگرداندن چالش اتلاف منابع عاطفی به دو مقوله نیاز و حق ارائه داده است همواره درست نیست.

به این ترتیب فارغ از اینکه می‌توانیم به بات‌ها و ربات‌های هوشمند عواطف را نسبت دهیم یا نه، یا اینکه رابطه ما با آن‌ها دوسویه است یا نه، دغدغه‌های اخلاقی پیرامون روابط انسان-ربات باقی می‌ماند و من نشان خواهم داد حداقل درباره برخی امور این دغدغه‌ها بسیار اساسی است. بحث‌های حول «فرب» و «دوسویگی» معنادار بودن روابط انسان-ربات را بررسی می‌کنند اما در این مقاله، اثر فراگیر شدن رابطه عاطفی انسان-ربات در اجتماع بر معنای زندگی افراد بررسی می‌شود؛ یعنی مسئله در زمینه گسترده‌تری بررسی می‌گردد. برای این منظور، من نخست صورت‌بندی متناسبی از چالش معنای زندگی و نسبت روابط با آن ارائه خواهم داد.

۳. معضل اخلاقی هوش مصنوعی عاطفی شده از منظر زندگی معنادار

همچنان که مرور شد، در عمده بررسی‌های اخلاقی ناظر به روابط معنادار و هوش مصنوعی عاطفی شده، رابطه فرد با بات یا ربات هوشمند بر مبنای تلقی ابزاری و دو مسئله «فرب» و «دوسویگی» مورد نظر قرار گرفته است. اما سیر تحلیلی من در این مقاله نزدیک به چالشی است که از آن با عنوان «اتلاف منابع مهم عاطفی» یاد کردم و متناسب با دیدگاه گسترش‌یافته وربیک در باب اخلاق فناوری است. من سعی دارم در بررسی نهایی خود، مسئله را در زمینه اجتماعی‌اش مورد واکاوی قرار دهم. در عین حال، بر اساس مبنای تشریح شده در باب زندگی معنادار، مسئله‌ای که بررسی من به آن ختم می‌شود تفاوت منطقی اساسی با آنچه تا به امروز ذیل عنوان «اتلاف منابع مهم عاطفی» مطرح شده دارد.

در مسئله «اتلاف منابع مهم عاطفی»، منظر بررسی به گونه‌ای است که چالش محروم شدن از رابطه معنادار برای دیگری پیش می‌آید. بیان دقیق‌تر به این ترتیب است که با توجه به محدودیت منابع عاطفی انسان، زمانی که «من» این منابع را صرف ارتباط عاطفی با هوش مصنوعی عاطفی شده نمایم، دسترسی «دیگری» برای بهره‌بردن از عواطف من - که برای شکوفایی و کمالش به آن نیاز دارد - محدودتر می‌شود و از این جهت، هوش مصنوعی عاطفی شده ملاحظه اخلاقی پیدا می‌کند. البته این موضوع که آیا «نیاز» دیگری به «حق» او در بهره‌مند شدن از منابع عاطفی من می‌انجامد یا نه شایسته بررسی بیشتری است.

۳-۱. طرح اولیه استدلالی مبنی بر بروز معضل اخلاقی برای هوش مصنوعی عاطفی شده

اما منطقی که در اینجا پیشنهاد می‌کنم به خود «من» معطوف است و نه «دیگری»، اگرچه موضوع در زمینه اجتماعی و بر اساس تحلیل اثر هوش مصنوعی عاطفی شده بر این زمینه بررسی می‌شود. همچنان که تبیین شد، زندگی معنادار ارزشی مستقل و نیازی اساسی است و من برای محقق نمودن آن ارزش در زندگی ام - که ارزش زیستنم را نیز مشخص می‌نماید - و برای پاسخ دادن به این نیاز اساسی، محتاج برقراری روابط عاطفی با دیگرانم. به بیانی ساده شده، من برای برخوردار شدن از زندگی معنادار، محتاج دوست داشتن دیگران و عاشق دیگران شدنم؛ دوست داشتن و عشقی که در اعمال گوناگون - پروژه‌های دارای ارزش مثبت - مربوط به دوستی و عاشقی بروز و ظهور پیدا می‌کند.

هوش مصنوعی عاطفی شده، در میدان برقراری روابط عاطفی و انجام پروژه‌های ارزشمندی که من از طریق تعلق و درگیری نسبت به آن‌ها از زندگی معنادار برخوردار می‌شوم، رقیب قدرتمندی است. او احتمالاً در این رقابت پیروز است چرا که:

- به‌طور خاص برای برقراری رابطه عاطفی و انجام کارهای مربوط به آن طراحی شده است.
- دسترسی‌اش به داده‌های مختلف بیشتر از من است.
- احتیاجاتی که من دارم - از جمله احتیاج متقابل به عاطفه و نیز سایر احتیاجات انسانی - را ندارد. احتیاجاتی که بر کیفیت رابطه عاطفی من اثر می‌گذارد.

به این ترتیب به نظر می‌رسد که هوش مصنوعی عاطفی شده نسبت به من، بهینه‌تر و بهره‌ورتر عمل خواهد کرد و از این جهت انتخابش ترجیح پیدا می‌کند. با ورود بات‌ها و ربات‌های تقویت‌شده با هوش مصنوعی عاطفی شده به جامعه، امکان برقراری رابطه عاطفی برای ما، از این جهت که عاشق باشیم و از سر عشق کارهایی انجام دهیم، محدودتر می‌شود. بنابراین توسعه هوش مصنوعی عاطفی شده، به دلیل تحدید دسترسی ما به «روابط عاطفی»، تهدیدی برای معناداری زندگی محسوب می‌شود و از این جهت توسعه آن از منظر اخلاقی موجه نیست.

دنیایی را تصور کنید که در آن وقتی برادرم در بیمارستان بستری است یک پرستار هوشمند عاطفی شده به‌صورت دائم و بدون لحظه‌ای خواب و غفلت، با مهربانی و محبت غیر قابل جبرانی، و نیز با اطلاعات کافی درباره مشکلات برادرم و آنچه نیاز دارد تمام مراقبت لازم از او را انجام می‌دهد تا آنجا که ملاقات کوتاه‌مدت و خدمت من به برادرم نوعی اختلال در مراقبت از او محسوب شود. یا دوستم برای اسباب‌کشی و چیدمان وسایلش در منزل جدید از کارگران هوشمندی استفاده می‌کند که علاوه بر انجام محاسبات گوناگون برای بهینه‌کردن فرایندها و رعایت همه جوانب ایمنی و کیفی، خسته هم نمی‌شوند و با احترام و ادب کامل پروژه جابه‌جایی را انجام می‌دهند به گونه‌ای که حضور یا کمک من مداخله محسوب شود و هزینه‌های کار را افزایش دهد. یا خانه‌ای که در آن دختری که از کودکی با حمایت‌ها و آموزش‌های مادر مهربان هوشمندش پرورش یافته برای طراحی و دوخت خلاقانه‌ترین لباس جشن مدرسه نیز سراغ این مادر باهوش و کاربلد می‌رود تا آنجا که مادر حقیقی او حتی از برنامه جشن نیز مطلع نیست. با توسعه هوش مصنوعی عاطفی شده، مجال ما برای دوست‌داشتن و عاشق‌شدن کمتر می‌شود و این در حالی است که ما برای برخوردار شدن از زندگی معنادار و فائق آمدن بر چالش پوچی، منابع معدودی داریم. بر اساس آنچه گذشت، می‌توان برهان زیر را به‌منظور موجه‌نبودن توسعه هوش مصنوعی عاطفی شده ارائه نمود:

- | | |
|---|---------------------------------|
| زندگی معنادار یک ارزش مستقل و یک نیاز اساسی برای ماست. | مقدمه ۱ |
| ما برای برخوردار شدن از زندگی معنادار نیاز داریم که دوست داشته باشیم و کارهایی ارزشمند را از سر دوست‌داشتن انجام دهیم. | مقدمه ۲ |
| ما برای دوست‌داشتن دیگری، منابع محدودی در اختیار داریم؛ یعنی دیگری را که ما بتوانیم دوستشان داشته باشیم و آن‌ها نیز در این رابطه دوستی و عاطفی مشارکت نمایند محدودند. | مقدمه ۳ |
| محدودتر شدن منابع محدود ما برای زندگی معنادار توسط خود ما از منظر اخلاق فناوری ناموجه است. | نتیجه فرعی
(از ۱م الی
۳م) |

مقدمه ۴
 هوش مصنوعی عاطفی شده رقیب ما برای این دوست داشتن‌ها محسوب می‌شود، رقیبی که از ما بهینه‌تر عمل می‌کند و از این جهت شانس انتخاب بیشتری دارد. در نتیجه با توسعه آن، زمینه اجتماعی به گونه‌ای تغییر می‌کند که «دوست داشتن» برای ما سخت‌تر و محدودتر می‌شود.

نتیجه
 (از نتیجه
 فرعی و
 مقدمه ۴)

توسعه هوش مصنوعی عاطفی شده توسط ما ناموجه است.

۳-۲. ارزیابی انتقادی این استدلال

یک نقد خارج از این برهان و مشابه با سایر مباحث اخلاق هوش مصنوعی طرح این مطلب است که برخی دیگر از انواع هوش مصنوعی عاطفی شده به موازات نوعی که در این برهان بر آن تأکید شده و رقیب ما در برقراری رابطه عاطفی با دیگران است می‌تواند طرف دوم رابطه عاطفی ما قرار گیرد؛ کما اینکه بسیاری از محصولات ارائه شده در حال حاضر چنین است. و به ترتیب اگرچه در مواردی رقیب ماست و منابع زندگی معنادار ما را محدود می‌کند، اما در مواردی هم - آنجا که خودش موضوع دوست داشتن است - این منابع را گسترش می‌دهد و لذا چنین نیست که همواره محدودکننده منابع معناداری باشد.

در اینجا برخلاف سایر بررسی‌های انجام شده درباره اخلاق هوش مصنوعی - نظیر وضعیتی که براهین ناظر به آینده کار دارند و در آغاز مقاله به آن اشاره شد - نمی‌توان چنین گفت و این مطلب قابل قبول نیست. آن دیگری که بر دوست داشتنش در مقدمه دوم استدلال بالا تأکید شده است نمی‌تواند هوش مصنوعی باشد؛ یا حداقل دوست داشتن هوش مصنوعی، به سبب تردیدی که درباره ارزش آفاقی آن وجود دارد زندگی را معنادار نمی‌کند.

توضیح اینکه ما از منظر عقل سلیم و مستقل از تمایلات فردی، درباره ارزشمند بودن نمونه‌های اعلائی عشق، شک نمی‌کنیم؛ عشق مادری، عشق به آزادی و عدالت، و عشق الهی آن‌چنان در تاریخ جوامع بشری شناخته شده و مثال‌های عالی دارد، از حیث آفاقی ارزشمند است. اما در ارزشمند بودن برقراری رابطه عاطفی با هوش مصنوعی، با وجود مباحث «فریب» و «دوسویگی»، از همین منظر شک و تردید وجود دارد. وجود این شک و تردید به معنای وقوع شکاکیت استعلایی فرایند پوچی - به تعبیر نیگل - و در نتیجه پوچی فلسفی زندگی است. تأکید می‌شود که در اینجا صرف ادامه‌دار بودن مباحثی نظیر «فریب» و «دوسویگی» و عدم جمع‌بندی متافیزیکی درباره ویژگی‌های درونی هوش مصنوعی و نه قبول فریب بودن یا دوسویه نبودن رابطه عاطفی با هوش مصنوعی، به تردید در ارزش آفاقی این رابطه ختم می‌شود و این تردید برای وقوع آن شکاکیت استعلایی مورد نظر نیگل کفایت می‌کند.

نقدی دیگر بر پیشینی بودن این استدلال و احتمال ناهمخوان شدن آن با شواهد پسینی تأکید دارد. به عنوان مثالی فرضی، در مواردی که طبق شواهد افرادی ضمن برقراری روابط با فناوری‌های مدرن تبدیل به انسان‌های دیگرگونه شده‌اند، ممکن است نوعی دیگر از معناداری زندگی را تجربه کنند به گونه‌ای که حتی بر اساس معیارهای سوزان ولف، در نوع خود دارای دو مؤلفه جاذبه ذهنی و جذابیت عینی باشد. به این ترتیب این استدلال مثال نقضی پیدا می‌کند و باطل می‌شود.

اهمیت تحلیل یکپارچه‌ای که از مسئله فلسفی معنای زندگی در آغاز این مقاله ارائه شد در پاسخ به این انتقاد روشن می‌شود. نیگل در پوچی، جایی تذکر می‌دهد که شکاکیت استعلایی ویژگی نهاد وجودی انسان است و یک موش آزمایشگاهی هیچ‌گاه همچون ما دچار شک نمی‌شود (Nagel 1971, 725). از این رو، مسئله فلسفی معنای زندگی اختصاص به نوع حیات انسان دارد. آنچه در این مقاله به‌مثابه تحلیلی یکپارچه از مسئله معنای زندگی ارائه شد، اگرچه الهام گرفته از دیدگاه سوزان ولف است، اما محدود به یک رویکرد عمل‌گرایانه نیست و علاوه بر این، ولف نیز دیدگاهش را در زمینه فلسفی مسئله معنای زندگی مطرح نموده است. به این ترتیب، دیگرگونه شدن انسان‌ها در پی ارتباط با فناوری، ولو اینکه فرضی معقول و موجه باشد، مسئله معنای زندگی را از اساس منحل می‌کند. از این رو، در پاسخ به این نقد می‌توان چنین گفت که تا انسان انسان است - با فرض قبول تحلیل نیگل از معضل فلسفی پوچی که فرض جاافتاده‌ای است - استدلال فوق درباره معضل اخلاقی هوش مصنوعی عاطفی شده معتبر خواهد بود.

نقد سوم این است که گفته شود یک پیش‌فرض اساسی در این استدلال، ارزش‌دهی به روابط عاطفی بر اساس محبت یا عشق است، حال آنکه این پیش‌فرضی اثبات نشده است. به بیان دیگر می‌توان مبناهایی دیگر همچون شفقت، همدلی، عدل، فضیلت، خودگروی و دیگرگزینی را فرض نمود که در آن‌ها، این استدلال کافی نیست و ممکن است به معناداری زندگی در روابط انسان-ربات بینجامد.

در پاسخ لازم است یادآوری شود مسئله معنای زندگی اگرچه مسئله‌ای هنجاری^{۴۸} است، اما مستقل از ارزش‌های اخلاقی است. به بیان دیگر، یک زندگی می‌تواند سراسر اخلاقی باشد اما معنادار نباشد. این مهم یکی از پیش‌فرض‌های رایج در ادبیات فلسفی پیرامون مسئله معنای زندگی است. بنا بر تحلیل ارائه‌شده از مسئله فلسفی معنای زندگی، و به‌ویژه با الهام گرفتن از پاسخ مقبول سوزان ولف، این روابط عاطفی مبتنی بر محبت و عشق منبعی برای معناداری زندگی محسوب می‌شود و ارزش‌های اخلاقی مرتبط با این روابط، به معضل پوچی زندگی پاسخ نمی‌دهند.

۴. جمع‌بندی و نتیجه‌گیری

یک نکته مغفول که به کمال این استدلال کمک می‌کند این است که با نگاهی ابطال‌گرایانه، توضیح داده شود که شرایطی جز آنچه این استدلال منعکس می‌کند وجود ندارد. به عبارت دیگر، نشان داده شود که هیچ نوع امکانی که در آن روابط عاطفی مبتنی بر محبت با هوش مصنوعی برقرار شود، و این رابطه منبعی برای معناداری زندگی محسوب شود وجود ندارد. در این صورت است که استدلال ارائه‌شده به کمال می‌رسد و منظر جدید توسعه‌یافته که نتیجه آن محدود شدن منابع معناداری زندگی است غلبه پیدا می‌کند.

البته این استدلال از این جهت ناتمام است و به همین سبب، نتیجه این بررسی تبیین معضل اخلاقی هوش مصنوعی عاطفی شده است و اینکه توسعه آن از منظر اخلاقی ملاحظه اساسی دارد. به بیان دقیق‌تر، اقتضای نتیجه‌گیری در این بررسی با توجه به این نکته، این است که با تعدیل ادعای ناموجه بودن توسعه هوش مصنوعی عاطفی شده، به این بسنده شود که این توسعه از منظر اخلاقی ملاحظه اساسی دارد. در عین حال این بررسی و نتیجه آن ۲ ویژگی دارد که آن را از بسیاری از بررسی‌های انجام‌شده تا به امروز متمایز می‌کند:

- دیدگاه انتخاب‌شده درباره معنای زندگی در این بررسی دیدگاهی دقیق و فلسفی است. در عمده بررسی‌های اخلاقی پیرامون هوش مصنوعی و روابط، روابط معنادار به معنایی عام مورد نظر قرار گرفته است و نه

روابطی که موجب زندگی معنادار به معنای فلسفی آن گردند. از این رو ناموجه دانستن توسعه هوش مصنوعی عاطفی شده، که نتیجه این بررسی است، معنایی قوی تر از ناموجه بودن اخلاقی توسعه هوش مصنوعی بر اساس دغدغه‌های مربوط به روابط معنادار انسان-انسان (با فرض قبول آن دست استدلال‌ها) دارد. از آنجایی که زندگی معنادار ارزشی مستقل و نیازی اساسی است، اخلاق در این بررسی به معنای هنجارهای مرتبط با روابط انسان-انسان^{۴۹} نیست، بلکه معنایی عام تر و اساسی تر دارد که عبارت است از هنجارهای مرتبط با زندگی خوب و کمال یافتن انسان.^{۵۰}

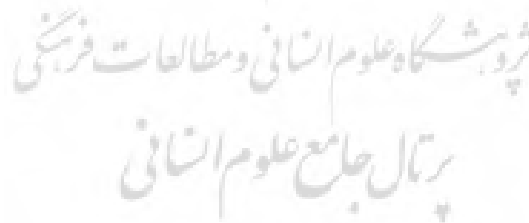
Bibliography

- Barrett, J., & A. H. Johnson. 2003. "The Role of Control in Attributing Intentional Agency to Inanimate Objects." *Journal of Cognition and Culture* 3(3): 208-217
<https://doi.org/10.1163/156853703322336634>
- Cacioppo, J. T., & W. Patrick. 2009. *Loneliness: Human Nature and the Need for Social Connection* (Reprint edition). W. W. Norton & Company.
- Danaher, J. 2019. "The Philosophical Case for Robot Friendship." *Journal of Posthuman Studies* 3(1): 5-24. <https://doi.org/10.5325/jpoststud.3.1.0005>
- De Graaf, M. M. A. 2016. "An Ethical Evaluation of Human-Robot Relationships." *International Journal of Social Robotics* 8(4): 589-598. <https://doi.org/10.1007/s12369-016-0368-5>
- Hauskeller, M. 2017. "Automatic Sweethearts for Transhumanists." In J. Danaher & N. McArthur (eds.), *Robot Sex*. The MIT Press. <https://doi.org/10.7551/mitpress/10718.003.0018>
- Metz, T. 2011. "The Good, the True, and the Beautiful: Toward a Unified Account of Great Meaning in Life." *Religious Studies* 47(4): 389-409.
- Metz, T. 2013. "The Meaning of Life." In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2013). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/sum2013/entries/life-meaning/>
- Nagel, T. 1971. "The Absurd." *The Journal of Philosophy* 68(20): 716-727. <https://doi.org/10.2307/2024942>
- Nyholm, S., & L. E. Frank. 2017. "From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible?" In J. Danaher & N. McArthur (eds.), *Robot Sex*. The MIT Press. <https://doi.org/10.7551/mitpress/10718.003.0019>
- Nyholm, S., & M. R  ther. 2023. "Meaning in Life in AI Ethics—Some Trends and Perspectives." *Philosophy & Technology* 36 (20). <https://doi.org/10.1007/s13347-023-00620-z>
- Nyholm, S., & J. Smids. 2020. "Can a Robot Be a Good Colleague?." *Science and Engineering Ethics* 26(4): 2169-2188. <https://doi.org/10.1007/s11948-019-00172-6>
- Seachris, J. 2019. "From the Meaning Triad to Meaning Holism: Unifying Life's Meaning." *Human Affairs* 29(4): 363-378. <https://doi.org/10.1515/humaff-2019-0031>
- Sparrow, R., & L. Sparrow. 2006. "In the Hands of Machines? The Future of Aged Care." *Minds and Machines* 16(2): 141-161. <https://doi.org/10.1007/s11023-006-9030-6>
- Turkle, S. 2012. *Alone Together: Why We Expect More from Technology and Less from Each Other*. 1st edition. Basic Books.

- Verbeek, P.-P. 2008. "Morality in Design: Design Ethics and the Morality of Technological Artifacts." In P. Kroes, P. E. Vermaas, A. Light, & S. A. Moore (eds.), *Philosophy and Design*: 91–103. Springer Netherlands. https://doi.org/10.1007/978-1-4020-6591-0_7
- Wada, K., & T. Shibata. 2007. "Living with Seal Robots—Its Sociopsychological and Physiological Influences on the Elderly at a Care House." *IEEE Transactions on Robotics* 23(5): 972–980.
- Wallach, W., & C. Allen. 2010. *Moral Machines: Teaching Robots Right from Wrong* 1st edition. Oxford University Press.
- Weber-Guskar, E. 2021. "How to Feel About Emotionalized Artificial Intelligence? When Robot Pets, Holograms, and Chatbots Become Affective Partners." *Ethics and Information Technology* 23(4): 601–610. <https://doi.org/10.1007/s10676-021-09598-8>
- Wolf, S. 1997. "Happiness and Meaning: Two Aspects of the Good Life." *Social Philosophy and Policy* 14(1): 207–225. <https://doi.org/10.1017/S0265052500001734>
- Wolf, S. 2016a. "Meaningfulness: A Third Dimension of the Good Life." *Foundations of Science* 21(2): 253–269. <https://doi.org/10.1007/s10699-014-9384-9>
- Wolf, S. 2016b. "The Meaning of Lives." In *Life, Death, and Meaning: Key Philosophical Readings on the Big Questions*, Third edition: 113–130. Rowman & Littlefield Publishers.
- Wolf, S., S. Macedo, J. Koethe, R. M. Adams, N. Arpaly, & J. Haidt. 2012. *Meaning in Life and Why It Matters*. 58908th edition. Princeton University Press.

یادداشت‌ها

1. Objective
2. Meaning of Life
3. Meaning in Life
4. Thaddeus Metz
5. Great Meaning in Life
6. Susan R. Wolf
7. The Blob
8. The Useless
9. The Bankrupt
10. Actively
11. Positive Value
12. Engagement
13. Happiness
14. Hybrid
15. Subjective Attraction
16. Objective Attractiveness
17. Independent of Oneself
18. Self-Transcendence
19. Egoistic
20. Moral



21. Monotheism

۲۲. حافظ. مضمون این دو بیت از این جهت که بیان می‌دارد کسی که از عشق بی‌بهره است نمی‌تواند از خود مستقل گردیده و فراتر رود و در این صورت معنای هستی را درک نخواهد کرد، با جمع‌بندی انجام‌شده تناسب دارد.

۲۳. مثنوی معنوی، دفتر اول. در این ابیات نیز عشق سبب فراتر رفتن از خود مادی انسان معرفی شده است و همچنین با عشق است که اجزای هستی به حرکت وا داشته شده و معانی حقیقی خود را بازتاب می‌دهند.

24. Self-Development

25. The Future of Work

26. Relationships

27. Expanding

28. Peter-Paul Verbeek

29. Emotionalized Artificial Intelligence - EAI

30. Autonomous

31. Deception

32. Mutuality

۳۳. ریشه‌های عمیق‌تر این مسئله در انسان را می‌توان در برخی ویژگی‌ها و سوگیری‌های شناختی او یافت: ویژگی‌هایی نظیر «نیاز به تعلق داشتن» (Need to Belong) (Cacioppo & Patrick 2009) و «بیش‌فعالی در تشخیص فاعل» (Hyperactive Agency Detectoin) (Barrett & Johnson 2003).

34. Sentient Being

35. Blake Lemoine

36. <https://arstechnica.com/tech-policy/2022/07/google-fires-engineer-who-claimed-lamda-chatbot-is-a-sentient-person/?ssp=1&darkschemeovr=1&setlang=en&cc=IR&safesearch=moderate>

37. Unknown

38. Weird

39. Familiar

40. Sincere

41. Collegial Relationships

42. Affective Relationship

43. Technical

44. Negligence of Morality

۴۵. وبر-گوسکار اخلاق را در اینجا به‌مثابه مجموعه‌ای از خواسته‌ها و انتظارات متقابل در مورد اینکه انسان‌ها چگونه باید با یکدیگر رفتار کنند می‌فهمد.

46. Robot's Inside

47. Waste of Important Emotional Resources

- 48. Normative
- 49. Morality
- 50. Ethics

