



Medical Ethics and Law
Research Center

Faṣlnāmah-i akhlāq-i pizishkī i.e., Quarterly Journal of Medical Ethics

2024; 18: e21



Shahid Beheshti
University of Medical Sciences

The Effect of Rationality, Intelligence and Emotions on the Moral-Legal Agency of Artificial Intelligence from the Perspective of Modern Cognitive Science

Mahmoud Abbasi¹, Ismail Keshavarz Safiei^{2*}

1. Medical Ethics and Law Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

2. Department of Law, Qazvin Branch, Islamic Azad University, Qazvin, Iran.

ABSTRACT

Background and Aim: The main objective of this article is to investigate the role of rationality, intelligence and emotions as mental-physical components in the moral-legal agency of artificial intelligence, based on the findings and achievements of contemporary cognitive science. Moral-legal agency of AI refers to the ability of AI systems to make independent decisions and to be held accountable for the consequences of their decisions.

Methods: This is a review study. The necessary resources and data were collected through library research and by accessing reputable domestic databases such as Civilica and Magiran and international databases such as Scopus, using books and articles from 2009-2024. A total of 72 articles and books (67 English and 5 Persian sources) were selected as the main sources for the final review. Of these 72 documents, 22 sources are directly related to the research findings, which are mentioned in the findings section.

Ethical Considerations: This research was conducted in full compliance with all ethical considerations regarding honesty and originality of texts.

Results: Currently, according to the prevailing view of cognitive scientists, due to the algorithmic and non-biological structure of AI systems and the lack of inherent rationality (lack of self, self-awareness, a perspective containing its own beliefs and preferences), the lack of high-level fluid intelligence and the lack of the biological and anatomical structure of the human nervous and brain system for emotional perception, it is not possible to speak of independent moral-legal agency for AI. Conversely, some researchers believe in minimal or participatory agency of AI. However, with the advancement of AI in the areas of human-like phenomenal and mental components (such as emotional perception, fluid intelligence, self and self-awareness), it may be possible in the future to attribute independent moral and legal agency to them.

Conclusion: Although AI currently lacks the mental capabilities necessary for independent agency, with the development and advancement of this technology in the future, if AI is equipped with human-like mental abilities (inherent rationality, fluid intelligence and emotional perception), achieving independent moral-legal agency will not be out of reach.

Keywords: Rationality; Intelligence; Emotions; Moral Agency; Artificial Intelligence

Corresponding Author: Ismail Keshavarz Safiei; **Email:** keshavarzsmail@gmail.com

Received: December 27, 2024; **Accepted:** May 04, 2025; **Published Online:** June 07, 2025

Please cite this article as:

Abbasi M, Keshavarz Safiei I. The Effect of Rationality, Intelligence and Emotions on the Moral-Legal Agency of Artificial Intelligence from the Perspective of Modern Cognitive Science. *Faṣlnāmah-i akhlāq-i pizishkī*, i.e., Quarterly Journal of Medical Ethics. 2024; 18: e21.



تأثیر عقلانیت، هوشمندی و عواطف در عاملیت اخلاقی - حقوقی هوش مصنوعی

از منظر علوم شناختی نوین

محمود عباسی^۱، اسماعیل کشاورز صفی‌ئی^{۲*} 

۱. مرکز تحقیقات اخلاق و حقوق پزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران.

۲. گروه حقوق، واحد قزوین، دانشگاه آزاد اسلامی، قزوین، ایران.

چکیده

زمینه و هدف: هدف اصلی این مقاله مطالعه بررسی نقش عقلانیت، هوشمندی و احساسات (عواطف) به عنوان مؤلفه‌های ذهنی - فیزیکی در عاملیت اخلاقی - حقوقی هوش مصنوعی از چشم‌انداز یافته‌ها و دستاوردهای علوم شناختی معاصر است. منظور از عاملیت اخلاقی - حقوقی هوش مصنوعی، توانمندی سیستم‌های دارای هوش مصنوعی جهت گزینش تصمیم‌گیری مستقل از انسان و مسئولیت‌پذیری آن‌ها در برابر پیامدهای تصمیمات خود می‌باشد.

روش: پژوهش حاضر از نوع مروری است. منابع و داده‌های مورد نیاز به شیوه کتابخانه‌ای و با مراجعه به تارنماها بانک‌های اطلاعاتی معتبر داخلی نظیر Magiran, Civilica و پایگاه داده‌های بین‌المللی نظیر ISC, Scopus از محل کتاب‌ها و مقالات سال‌های ۲۰۰۹-۲۴ گردآوری شده است. تعداد ۷۲ مقاله و کتاب (۶۷ منبع انگلیسی و ۵ منبع فارسی) به عنوان منابع کلی مقاله جهت بررسی نهایی انتخاب گردیدند، سپس از میان ۷۲ مستند، ۲۲ منبع به طور مستقیم با یافته‌های حاصل از پژوهش مرتبط هستند که در بخش یافته‌ها مورد اشاره قرار گرفته‌اند.

ملاحظات اخلاقی: این پژوهش با رعایت کامل تمامی ملاحظات اخلاقی در خصوص رعایت امانتداری، صداقت اصالت متون انجام گرفته است.

یافته‌ها: در وضعیت کنونی بنا بر نظر غالب دانشمندان علوم شناختی، به دلیل ساختار الگوریتمی و غیر بیولوژیکی سیستم‌های هوش مصنوعی و فقدان عقلانیت ماهوی و حیث التفاتی (فقدان خود، خودآگاهی، چشم‌انداز حاوی باورها و ترجیحات مخصوص خود)، فقدان برخورداری از هوش سیال سطح بالا، فقدان ساختار بیولوژیکی و آناتومیک سیستم عصبی - مغزی انسان جهت ادراک عاطفی، نمی‌توان از عاملیت اخلاقی - حقوقی مستقل هوش مصنوعی سخن به میان آورد. در نقطه مقابل برخی پژوهشگران بر عاملیت حداقلی یا مشارکتی هوش مصنوعی باور دارند. با این وجود با پیشرفت هوش مصنوعی در زمینه مؤلفه‌های پدیداری و ذهنی انسان مانند (نظیر ادراک عاطفی، برخورداری از هوش سیال، برخورداری از خود و خودآگاهی)، در آینده می‌توان برای آن‌ها نیز عاملیت مستقل اخلاقی و حقوقی قائل شد.

نتیجه‌گیری: هرچند در وضع کنونی هوش مصنوعی فاقد توانایی‌های ذهنی انسان جهت برخورداری از عاملیت مستقل است، اما با گسترش و پیشرفت این فناوری در آینده، در صورت مجهز شدن هوش مصنوعی به توانمندی‌های ذهنی انسان (عقلانیت ماهوی، هوش سیال و ادراک عاطفی)، دستیابی هوش مصنوعی به عاملیت اخلاقی - حقوقی مستقل دور از دسترس نخواهد بود.

واژگان کلیدی: عقلانیت؛ هوشمندی؛ عواطف؛ عاملیت اخلاقی؛ هوش مصنوعی

نویسنده مسئول: اسماعیل کشاورز صفی‌ئی؛ پست الکترونیک: keshavarzsmail@gmail.com

تاریخ دریافت: ۱۴۰۳/۱۰/۰۷؛ تاریخ پذیرش: ۱۴۰۴/۰۲/۱۴؛ تاریخ انتشار: ۱۴۰۴/۰۳/۱۷

خواهشمند است این مقاله به روش زیر مورد استناد قرار گیرد:

Abbasi M, Keshavarz Safiei I. The Effect of Rationality, Intelligence and Emotions on the Moral-Legal Agency of Artificial Intelligence from the Perspective of Modern Cognitive Science. Faşlnāmah-i akhlāq-i pizishkī, i.e., Quarterly Journal of Medical Ethics. 2024; 18: e21.

مقدمه

مقوله عقلانیت، عواطف و هوشمندی و تأثیر آن بر تصمیم‌گیری، قلمرو مشترکی میان عاملیت اخلاقی - حقوقی انسان و ماشین‌های هوشمند محسوب می‌شود. عقلانیت، عواطف و هوشمندی هر سه از قابلیت‌های ذهنی، بدنی انسان به شمار می‌روند، هرچند که تعریف یکسانی از این مفاهیم نزد اندیشمندان به چشم نمی‌خورد و مرزهای تمایز میان آن‌ها دقیق و شفاف نیست (۱، ۲). پژوهش‌ها و مطالعات اخیر حاکی از چالش‌ها و اختلاف نظرهای بنیادین میان پژوهشگران در خصوص مفاهیم سه‌گانه عقلانیت (عملی)، عاطفه (احساس) و هوشمندی و تأثیر آن‌ها در عاملیت اخلاقی - حقوقی انسان است (۳، ۴). به دنبال پیشرفت روزافزون فناوری‌های هوش مصنوعی، تلاش‌ها در راستای شبیه‌سازی هرچه بیشتر ماشین‌های هوشمند با مؤلفه‌های ذهنی انسان گسترش یافته است. مقوله هوشمندی در فناوری‌های هوش مصنوعی از سابقه‌ای طولانی برخوردار است و این در حالی است که عقلانیت، شرط ضروری عاملیت اخلاقی و اهلیت و مسئولیت‌پذیری حقوقی انسان‌ها در تصمیمات و اعمال او محسوب می‌شود (۵)، لیکن سخنی از عقلانیت مصنوعی به میان نیامده است. بر همین اساس از منظر برخی پژوهشگران علوم شناختی، هوش مصنوعی در وضع کنونی فاقد ذهن بازتابنده و قابلیت‌های عقلانیت ماهوی انسان است، چراکه ذهن بازتابنده متشکل از تمایلات فکری شامل امیال، ارزش‌ها و اهداف متنوع است و عقلانیت ماهوی نیز دربردارنده برخورداری و آگاهی از ارزش‌ها، امیال و اهداف متکثر و متنوع تصمیم‌گیرنده در هر موقعیت رفتاری جهت تصمیم‌گیری است. هر تصمیمی می‌تواند دارای محتوای اخلاقی باشد و این در حالیست که هوش مصنوعی فاقد ظرفیت و کیفیت اخلاقی است، چراکه فاقد نیروی تفکر و بازتابندگی راجع به تمایلات اخلاقی و فکری نظیر ارزش‌ها، باورها و غایات است (۵، ۶). از طرف دیگر دستاوردها و یافته‌های روان‌شناسان و عصب‌پژوهان معاصر حاکی از ارتباط تنگاتنگ میان هوش و عاطفه و تأثیر تعیین‌کننده هوش عاطفی در کمیت و کیفیت تصمیم‌گیری‌های انسان است (۷). همچنین یافته‌های برخی

متخصصان علوم مغز، عصب‌شناسان و روان‌شناسان شناختی معاصر نظیر Stanovich و همکاران حاکی از تفاوت میان عقلانیت و هوشمندی در عاملیت و تصمیم‌گیری‌های انسان است (۳)، بدین ترتیب که هوشمندی صرفاً دربردارنده کارکرد محاسباتی ذهن (شامل هوش سیال و متبلور و حافظه کاری) است، در حالی که عقلانیت علاوه بر کارکرد محاسباتی شامل ذهن بازتابنده نیز می‌شود که دربردارنده تمایلات فکری (باورها، امیال و اهداف) است (۴). با این وجود کاربست این یافته‌ها و ارتباط میان مؤلفه‌های عقلانیت، هوشمندی و عواطف در عاملیت ماشین‌های هوشمند، کمتر و یا بدون تحلیل همزمان مورد بررسی قرار گرفته است. مسأله‌ای که می‌توان در ارزیابی عاملیت اخلاقی و حقوقی هوش مصنوعی مؤثر باشد. در واقع آنچه از آن می‌توان به عنوان خلأ مطالعاتی یاد نمود، فقدان بررسی همزمان تأثیر این مؤلفه‌های متمایز و مرتبط عقلانیت، هوشمندی و عواطف در عاملیت و تصمیم‌گیری سیستم‌های دارای هوش مصنوعی است. با توجه به اینکه از دستاوردهای علوم شناختی به طور گسترده‌ای در فناوری‌های نوین هوش مصنوعی (نظیر هوش مصنوعی عصبی - نمادین) استفاده می‌شود، مقوله تأثیر انتقال این دستاوردها در عاملیت و مسئولیت‌پذیری مستقل هوش مصنوعی کمتر مورد توجه قرار گرفته است. هدف از این مقاله تعیین تمایز مؤلفه‌های ذهنی - فیزیکی عقلانیت، هوشمندی و عواطف در تصمیم‌گیری‌های اخلاقی - حقوقی هوش مصنوعی است. این پژوهش می‌تواند به منزله حلقه پیوندی میان دستاوردهای دو دانش علوم شناختی و هوش مصنوعی باشد و بر مقولات فلسفی، اخلاقی و حقوقی تأثیر بگذارد.

روش

پژوهش حاضر از نوع مروری است. منابع و داده‌های مورد نیاز به شیوه کتابخانه‌ای و با مراجعه به سایت‌ها و بانک‌های اطلاعاتی معتبر داخلی نظیر Magiran، Civilica، پایگاه داده‌های بین‌المللی نظیر Scopus، از محل کتاب‌ها و مقالات در بازه زمانی سال‌های ۲۴-۲۰۰۹ گردآوری شده است. پس از جستجوی اولیه و حذف مقالات و مستندات غیر

مرتبط، تعداد ۷۲ مقاله و کتاب (۶۷ منبع به زبان انگلیسی و ۵ منبع به زبان فارسی) توسط پژوهشگر بررسی دقیق شد که از بین آن‌ها تعداد ۲۲ منبع به طور مستقیم با یافته‌های حاصل از پژوهش مرتبط بودند و مورد بررسی و تحلیل قرار گرفتند.

ملاحظات اخلاقی

تحقیق حاضر از سوی کمیته اخلاق زیست‌پزشکی دانشگاه علوم پزشکی شهید بهشتی به تأیید رسیده است (کد کمیته اخلاق: IR.SBMU.RETEC.REC.1403.768).

یافته‌ها

عقلانیت، هوشمندی و عواطف هر سه از مؤلفه‌های ذهنی - بدنی انسان هستند و دانشمندان علوم شناختی معاصر، رهیافت‌های نوینی در خصوص تأثیر این مؤلفه‌ها بر عاملیت انسان در تصمیم‌گیری ارائه نموده‌اند. نتایج این یافته‌ها به فناوری‌های هوش مصنوعی نیز رسوخ نموده است. برآیند حاصل از نفوذ یافته‌های علوم شناختی به مقوله عاملیت هوش مصنوعی بیانگر آن است که در وضعیت کنونی بنا بر نظر غالب دانشمندان علوم شناختی، به دلیل ساختار الگوریتمی و غیر بیولوژیکی سیستم‌های هوش مصنوعی و فقدان عقلانیت ماهوی و حیث التفاتی (فقدان خود، خودآگاهی، چشم‌انداز حاوی باورها و ترجیحات مخصوص خود)، فقدان برخورداری از

هوش سیال سطح بالا، فقدان ساختار بیولوژیکی و آناتومیک سیستم عصبی - مغزی انسان جهت ادراک عاطفی، نمی‌توان از عاملیت اخلاقی - حقوقی مستقل هوش مصنوعی سخن به میان آورد (۳، ۷-۱۶).

در مقابل برخی پژوهشگران بر عاملیت حداقلی یا مشارکتی هوش مصنوعی باور دارند. مدافعان عاملیت حداقلی بر این باورند که با گذار از مؤلفه‌های ذهنی عاملیت عقلانی - اخلاقی به مؤلفه‌های عینی، محاسباتی و منطقی و کاهش ضریب خطا، هوش مصنوعی حتی می‌تواند عملکرد بهتری نسبت به انسان داشته باشد، لذا می‌توان برای آن عاملیت اخلاقی قائل شد (۷، ۸، ۱۴، ۱۶-۱۸). طرفداران عاملیت مشارکتی نیز معتقدند با مشارکت سیستم‌های هوش مصنوعی در تصمیم‌گیری‌های انسان، دیگر نمی‌توان از عاملیت سنتی و مدرنیستی انسان سخن گفت، بلکه با تفویض بسیاری از تصمیمات از انسان به هوش مصنوعی در شبکه پیچیده‌ای از روابط، می‌بایست به بازنگری و بازتعریف عاملیت و مسئولیت اخلاقی پرداخت (۱۹-۲۲).

با این وجود بنا بر نظر غالب دانشمندان علوم شناختی با پیشرفت هوش مصنوعی در زمینه مؤلفه‌های پدیداری و ذهنی انسان مانند (نظیر ادراک عاطفی، برخورداری از هوش سیال، برخورداری از خود و خودآگاهی)، در آینده می‌توان برای آن‌ها نیز عاملیت مستقل اخلاقی و حقوقی قائل شد (۱۷، ۲۳-۲۸).

جدول ۱: خلاصه مطالعات انجام شده در خصوص تأثیر عقلانیت، هوشمندی و عواطف در عاملیت هوش مصنوعی

نام نویسنده	سال چاپ	عنوان پژوهش	نتایج
Valtonen L (۸)	۲۰۲۳	عقلانیت در تصمیم‌گیری هوش مصنوعی	نوع عقلانیت به کاررفته در هوش مصنوعی بر پایه عقلانیت شکلی و الگوریتمی است که مبتنی بر پردازش داده‌ها از طریق محاسبه به منظور دستیابی به بهترین نتیجه است. همچنین نوع عاملیت به کاررفته در هوش مصنوعی عاملیت عقلانی است نه اخلاقی.
Stanovich K (۳)	۲۰۱۲	تمایز میان عقلانیت و هوشمندی	عقلانیت مفهومی گسترده‌تر از هوشمندی است و هوشمندی صرفاً دربردارنده ذهن الگوریتمی و محاسباتی (شامل هوش سیال و متبلور و حافظه کاری) است نه ذهن بازتابنده که دربردارنده تمایلات فکری (باورها، امیال و اهداف) است. به باور نویسندگان ماشین‌های هوشمند صرفاً از ذهن الگوریتمی برخوردارند نه ذهن بازتابنده.
Nishant و همکاران (۹)	۲۰۲۴	عقلانیت شکلی الگوریتم‌های هوش مصنوعی و مسأله سوگیری	عقلانیت به کاررفته در هوش مصنوعی، عقلانیت شکلی است نه ماهوی؛ بر همین اساس در مواجهه با حیات واقعی، الگوریتم‌های هوش مصنوعی نمی‌توانند مؤلفه‌های ماهوی پیش‌زمینه‌ای متمایز را به نحو درست ارزیابی کنند و دچار سوگیری می‌شوند.
Nyamori D (۷)	۲۰۲۴	برجسته‌سازی رابطه میان تصمیم‌گیری اخلاقی و هوش عاطفی	هوش عاطفی به طور قابل توجهی بر توانایی افراد برای ساختن تصمیمات درست اخلاقی تأثیر می‌گذارد، همانطور که توسط همبستگی مثبت بین هوش عاطفی و انواع مختلف مشهود است، افراد با هوش عاطفی بالاتر آگاهی عاطفی، مقررات و همدلی برتری را نشان دادند که آن‌ها را قادر می‌سازد معضلات اخلاقی را به طور مؤثرتر دنبال و حل کنند.
Vicci DH (۱۷)	۲۰۲۴	هوش عاطفی در هوش مصنوعی	بعضی از انواع هوش مصنوعی برای نشان دادن احساسات ساخته شده‌اند، اما در عصر حاضر چندان رایج نیستند. به نحوی که با اعطای سیستم‌هایی با درک احساسات به هوش مصنوعی، می‌توانیم تجربیات عاطفی را در آن‌ها ایجاد کنیم که به ایجاد همدلی و احساس ارتباط با کاربران در کمک کند. با ایجاد هوش مصنوعی بدن‌مند(تنانه) یا عاطفی، آن‌ها قادر به ارزیابی و قضاوت اخلاقی و تصمیم‌گیری به تشخیص خود هستند و از نظر اجتماعی و قانونی دارای عاملیت و مسئولیت خواهند بود.
Northoff G, Gouveia SS (۱۵)	۲۰۲۴	آیا هوش مصنوعی قادر به نمایش ذهن‌مندی بنیادین است؟ بحثی عصب‌شناختی - فلسفی	ذهن‌مندی انسان (برخورداری از تجربه ادراکی متعلق به خود و داشتن چشم‌انداز به محیط پیرامون) خاستگاهی عصبی - اکولوژیکی دارد که هوش مصنوعی در وضع کنونی خود به دلیل فقدان این ویژگی‌ها، فاقد ذهنیت انسان‌مانند در تصمیم‌گیری است.
Nath R, Manna R (۱۶)	۲۰۲۳	از پسانسان‌گرایی تا اخلاقیات هوش مصنوعی	رابطه بین انسان زیستی و پسانسان نقش فعالی در طراحی عوامل اخلاقی مصنوعی خواهد داشت. اینکه آیا پسانسان‌های آینده بر بشریت بیولوژیکی غلبه خواهند کرد یا هر دو آن‌ها به عنوان همتا برای تشکیل یک جامعه آرمان‌شهری دیجیتال و ایجاد ابعاد جدیدی از عقلانیت کار خواهند کرد، هنوز مورد انتظار است.

نام نویسنده	سال چاپ	عنوان پژوهش	نتایج
Arvan R, Maley CJ (۱۰)	۲۰۲۲	همه روان (روح) دارانگاری و هوش مصنوعی	هوش مصنوعی مبتنی بر دیجیتال از داشتن تجربیات آگاهانه کلان پدیداری منسجم ناتوان است، زیرا عملکرد عصبی و آگاهی پدیداری مغز در سطح خرد و کلان، هر دو ماهیت آنالوگ دارند که کیفیت آن متناسب به شدت تحریک و سرعت شلیک نورونی متفاوت خواهد بود، در حالی که عملکرد AI به صورت دیجیتال (ارقام یکنواخت) بوده و از سطح خرد بالاتر نمی‌رود، لذا فاقد آگاهی پدیداری در سطح کلان نیست.
Damasio A (۱۱)	۲۰۱۸	نظم عجیب اشیا، زندگی، احساس و ساختن فرهنگ	از نظرگاه Damasio ذهن ما دارای دو کارکرد نگاشتی (ثبتي) است. در نگاشت اول که دقیق و محاسباتی است، ما با ادراک، حرکت، خاطرات، استدلال، زبان‌های کلامی و زبان‌های ریاضی سروکار داریم. این سیگنال‌های سیناپسی به خوبی توسط هوش مصنوعی و رباتیک شبیه‌سازی شده است، اما نگاشت (ثبت) دوم، مربوط به عواطف و احساسات است که وضعیت زندگی در بدن زنده ما را توصیف می‌کند و به راحتی خود را به یک حساب محاسباتی تقلیل نمی‌دهد. هوش مصنوعی و رباتیک فعلی به این ثبت دوم نمی‌پردازند.
O'Regan JK (۲۳)	۲۰۱۲	چگونه رباتی که دارای آگاهی و عواطف است، بسازیم؟	Oregan در مقام دفاع از دستیابی هوش مصنوعی به مؤلفه‌های ذهنی انسان در آینده بر این باور است، اگرچه تجربه آگاهانه صرفاً بر پایه مفهوم خود صورت‌پذیر خواهد بود و دستیابی ربات‌های هوشمند به آگاهی پدیداری و خود، بسیار دشوار است، اما اجرای آن‌ها در روبات‌ها از نظر منطقی غیر ممکن نیست.
Kapoor A, Verma V (۲۵)	۲۰۲۴	فهم احساسات از طریق هوش مصنوعی	هوش مصنوعی عاطفی دارای توانایی تشخیص، تجزیه و تحلیل و تفسیر احساسات انسان از طریق ورودی‌های مختلف مانند متن، گفتار، حالات چهره و سیگنال‌های فیزیولوژیکی است، با پیشرفت‌های اخیر در هوش مصنوعی و یادگیری ماشینی، تجزیه و تحلیل احساسات از نظر دقت، کارایی بسیار پیشرفت کرده است.
Booch و همکاران (۲۴)	۲۰۲۱	تفکر سریع و آهسته در هوش مصنوعی	چنانچه بینش‌هایی در مورد علل برخی از قابلیت‌های انسانی که هنوز در هوش مصنوعی وجود ندارد، به دست آورده شود (به عنوان مثال، سازگاری، تعمیم‌پذیری، عقل سلیم و استدلال علی)، می‌توان با تعبیه این مؤلفه‌های علی، به قابلیت‌های مشابهی در یک سیستم هوش مصنوعی دست یافت.
Pandey P (۲۷)	۲۰۲۳	مهارت‌های هوش مصنوعی عاطفی در عصر هوش مصنوعی	با اعطای سیستم‌هایی با قابلیت احساسات به هوش مصنوعی، آن‌ها قادر به ارزیابی و اتخاذ قضاوت و تصمیمات اخلاقی به روش‌های خود هستند، به نحوی که از نظر اجتماعی و قانونی می‌توانند مسئول شناخته شوند.
Pan W, Yang Y (۲۸)	۲۰۲۲	آگاهی عاطفی مجازی و هوش مصنوعی اجتماعی	پیشرفت‌های روزافزون انتقال توانمندی‌های عاطفی انسان به هوش مصنوعی، این باور را برای برخی اندیشمندان حوزه علوم شناختی و هوش مصنوعی پدید آورده است که در آینده می‌توان ماشین‌هایی با قابلیت‌های کامل عاطفی انسانی پدید آورد.
Zeng و همکاران (۱۳)	۲۰۲۴	هوش مصنوعی الهام‌گرفته از مغز و خودبنیاد	هیچ احساسی از مفهوم خود در هوش مصنوعی وجود ندارد. هوش مصنوعی کنونی فقط به ظاهر هوشمندانه به پردازش اطلاعات مبادرت می‌ورزد و واقعاً خود را درک نمی‌کند یا به طور ذهنی از خود، آگاه نیست و جهان را با خود درک نمی‌کند، همانطور که هوش انسانی می‌کند.
Wróblewski Z, Fortuna P (۱۲)	۲۰۲۳	ذهن‌مندی اخلاقی و وضعیت اخلاقی هوش مصنوعی	در وضع کنونی به دلیل فقدان برهم‌کنش مؤلفه‌های سه‌گانه طبیعت انسانی (بدن، ذهن و نفس یا خود)، فقدان ذهنیت و مؤلفه‌های ذهنی و انتزاعی اشخاص آگاه نظیر خود (من)، آگاهی، حیث التفاتی و احساس وجود خود (خودآگاهی)، هوش

نام نویسنده	سال چاپ	عنوان پژوهش	نتایج
Marwala T (۱۴)	۲۰۱۷	انتخاب عقلانی و هوش مصنوعی	مصنوعی فعلی را نمی‌توان دارای عاملیت اخلاقی دانست.
Gouiaa R, Bazarna A (۱۸)	۲۰۲۳	بر هم‌کنش میان عقلانیت، سیاست و هوش مصنوعی در فرایندهای تصمیم‌گیری	آینده با وجود پیش‌بینی‌های دقیق و تصمیم‌گیری‌های سیاسی یا اقتصادی مطلوب و حل مسائل پیچیده توسط ماشین‌های هوشمند محاسبه‌گر، مکان معقول‌تری و مناسب‌تری خواهد بود.
Monteiro WR, Ayrosa EAT (۱۹)	۲۰۲۴	از آنالوگ تا الگوریتم: بررسی جامع تلاقی هوش مصنوعی با ادبیات مدیریت و نظریه بازیگر - شبکه	شکاف تحقیقاتی قابل توجهی در درک تعاملات انسان و ماشین و مدیریت کسب و کار در سازمان‌هایی که از هوش مصنوعی استفاده می‌کنند، وجود دارد. راه‌های بالقوه برای مشارکت‌های علمی آینده قابل شناسایی است و نیاز به رویکردی یکپارچه‌تر که بازیگران انسان و ماشین را در زمینه‌های سازمانی در نظر می‌گیرد، مورد تأکید است.
Venturini T (۲۰)	۲۰۲۳	برونو لاتور و هوش مصنوعی	ایده جداشدن ماشین‌ها از سازندگان‌شان مورد پذیرش نیست. هوش مصنوعی تنها از طریق حمایت دستیاران انسانی خود می‌تواند وجود داشته باشند. خطرات مرتبط با این فناوری‌ها از یک تکینگی غیر محتمل ناشی نمی‌شوند، بلکه ناشی از تعبیه آن‌ها در صنعت کسل‌کننده و استثمارگر اقتصاد دیجیتال است.
Reed M (۲۱)	۲۰۱۸	طبقه‌بندی هوش مصنوعی به عنوان عاملان اجتماعی	نظریه شبکه بازیگر Bruno Latour می‌تواند به ما در درک جایگاه این اشیا و هوش مصنوعی در فرهنگ کمک کند.
Waelbers K (۲۲)	۲۰۰۹	تفویض فناوری: مسئولیت در قبال موارد ناخواسته، علم و اخلاق مهندسی.	مسئولیت اخلاقی انسان در عاملیت مشارکتی او با فناوری‌ها بر پایه مفروضاتی چون توجه به نقش اجتماعی فرد به عنوان هدف مسئولیت، نقش اخلاقی افراد برای عمل، نقش واسطه‌گری مصنوعات فناورانه در تحقق اعمال و نیت انسان و پویایی چندین عمل متقابل استوار است، لذا امکان به کارگیری الگوهای مرسوم برای مسئولیت اخلاقی میسر نیست و نیازمند بازنگری در تفسیر مسئولیت اخلاقی هستیم.

بحث

بنا بر یافته‌های به دست‌آمده از این پژوهش و در پاسخ به پرسش اصلی تحقیق باید گفت مسأله عاملیت اخلاقی و حقوقی هوش مصنوعی در تصمیم‌گیری از منظر متخصصان علوم شناختی بر پایه مؤلفه‌هایی چون عقلانیت، هوشمندی و عواطف محل چالش و اختلاف نظر است تا جایی که می‌توان یافته‌های به دست‌آمده از سوی پژوهشگران را می‌توان در سه طیف اصلی طبقه‌بندی کرد.

در راستای تحلیل این نظریه‌پردازی‌ها، از آنجا که عقلانیت، هوشمندی و عواطف در هوش مصنوعی بر پایه شبیه‌سازی این عناصر از انسان مورد اقتباس قرار گرفته است، لازم است ابتدا

به تحلیل جایگاه هر یک از این مؤلفه‌ها در عاملیت و تصمیم‌گیری‌های انسان و هوش مصنوعی پرداخته شود، سپس طیف‌های سه‌گانه یافته‌های حاصل از پژوهش‌های دانشمندان علوم شناختی در زمینه کاربرست مؤلفه‌های عقلانیت، هوشمندی و عواطف در عاملیت و تصمیم‌گیری‌های هوش مصنوعی مورد تحلیل قرار خواهد گرفت.

۱. نقش عقلانیت در عاملیت انسان و هوش مصنوعی: از آنجا که هوشمندی و فرایند پردازش سیستم‌های هوش مصنوعی، برگرفته و شبیه‌سازی‌شده از فرایند پردازش ذهنی انسان است، ابتدا عقلانیت در عاملیت انسان و سپس سیستم‌های هوش مصنوعی مورد بررسی قرار می‌گیرد.

۱-۱. نقش عقلانیت در عاملیت انسان: اندیشمندان علوم شناختی، عقلانیت را استعداد یا آمادگی و توانایی تفکر تحلیلی برای عاملیت و تصمیم‌گیری تلقی می‌کنند، به نحوی که موجب حداکثرسازی منافع مورد انتظار یا پیروی از قواعد احتمالات باشد. عقلانیت از حیث نحوه و هدف تصمیم‌گیری را می‌توان به عقلانیت ماهوی و شکلی تقسیم نمود. عقلانیت ماهوی از چپستی و ماهیت ارزشی تصمیم سخن می‌گوید. در عقلانیت شکلی تمرکز بر اتخاذ فرایندی و مکانیسمی مناسب جهت گزینش رفتار صحیح است (۸). Herbert Simon از پیشگامان فکری هوش مصنوعی، به بهترین وجه ممکن تفاوت این دو عقلانیت را برشمرده است. از نظرگاه وی در عقلانیت ماهوی صحبت از این است که چه تصمیمی گرفته می‌شود، در عقلانیت شکلی سخن از این است که چگونه تصمیم گرفته می‌شود (۲۹). عقلانیت یکی از مهم‌ترین مؤلفه‌های عاملیت انسان محسوب می‌شود، اما یافته‌های اخیر نشان می‌دهد که عقلانیت تنها عامل تعیین‌کننده در فرایند شناخت و عاملیت در تصمیم‌گیری نیست. امروزه مجموعه بزرگی از تحقیقات شناختی در حوزه‌های مختلف عصب شناختی و هوش مصنوعی، تحت تأثیر یافته‌های Daniel Kahneman و Amos Tversky، ساختار ذهن انسان را بر پایه نظریه پردازش دوگانه ذهن مطرح می‌کنند. بر اساس این نظریه فرایند شناخت (تفکر) و عاملیت در تصمیم‌گیری ما صرفاً محصول عاملیت عقلانی ما نیست، بلکه توسط دو فرایند مجزا انجام می‌شود. اولی معمولاً به عنوان سریع، شهودی، هیجانی، خودکار و تداعی‌محور مشخص می‌شود، در حالی که دومی عمدی، تلاشگر، کنترل‌شده، عقلانی و قاعده‌مند است (۳۰-۳۲).

از دانشمندان علوم شناختی معاصر در چهارچوب نظریه پردازش دوگانه Kahneman و Tversky پیشگام طراحی مدل و ساختار ذهنی انسان در سال‌های اخیر محسوب می‌شود. به طور خلاصه بر اساس یافته‌های Stanovich انسان دارای سه مدل ذهنی هستند: اول: ذهن خودکار که همان سامانه یک Kahneman بوده و مسئول عملکردهای سریع،

خودکار، هیجانی و بدون پردازش عقلانی است. دو مدل ذهنی دیگر دارای عملکرد پردازش عقلانی و مبتنی بر کنترل توجه، حافظه کاری درگیرشونده هستند. این دو مدل ذهنی عبارتند از: ۱- ذهن الگوریتمی؛ ۲- ذهن بازتابنده یا انعطافی (۳۳، ۳۴)؛ دوم: ذهن الگوریتمی یا محاسباتی، مرتبط با توانایی شناختی و محاسباتی مغز انسان است که شامل هوش متبلور و سیال، حافظه کاری و کنترل توجه می‌باشد. به طور کلی ذهن الگوریتمی بر توانایی‌های شناختی انسان در یادگیری، پردازش اطلاعات و محاسبه محرک‌های محیطی، حل مسأله و تصمیم‌گیری در پاسخ به موقعیت‌ها و محرک‌های مختلف محیطی تمرکز دارد؛

سوم: ذهن بازتابنده یا انعطافی، با مقوله تمایلات فکری مرتبط است. تمایلات فکری شامل باورها، ساختار، چگونگی شکل‌گیری و دگرگونی باورها، اهداف و سلسله‌مراتب اهداف، تفکر ذهن باز، ملاحظه پیامدهای آینده تصمیم می‌شود (۳)؛ (۴).

۲-۱. نقش عقلانیت در عاملیت هوش مصنوعی: هوش مصنوعی نزد متخصصان این حوزه به توانایی ماشین برای اجرای عملکردهای قابل انتساب به هوش انسانی از طریق به کارگیری الگوریتم پیچیده جمع‌آوری و تحلیل داده‌ها اطلاق می‌شود (۱). امروزه از هوش مصنوعی به نحو فزاینده‌ای برای تصمیم‌گیری در مشاغل حرفه‌ای مختلف به عنوان جایگزین انسانی استفاده می‌شود. به عبارت دیگر کارکرد اصلی ماشین‌های هوشمند، تصمیم‌گیری در موقعیت‌های مختلف به جای انسان از طریق شبیه‌سازی با هوش انسانی است. از منظر Simon که از متفکران شناخته‌شده هوش مصنوعی محسوب می‌شود، تصمیم‌گیری از چهار مرحله تشکیل می‌شود:

۱- جمع‌آوری داده‌ها؛ ۲- شناسایی راه حل‌های جایگزین؛ ۳- انتخاب تصمیم صحیح از میان گزینه‌های موجود؛ ۴- اجرای تصمیم گزینش‌شده (۳۵). تاکنون تعاریف متعددی از هوش مصنوعی به عمل آمده است که در همه آن‌ها شاهد نوعی تقلید و شبیه‌سازی با عاملیت و عقلانیت انسان هستیم. نظیر: تفکر انسان‌گونه، رفتار انسان‌گونه، تفکر عقلانی (منطقی)،

انسان است، بررسی نوع و کیفیت هوشمندی شبیه‌سازی شده در ماشین‌های هوشمند و نقش آن در عاملیت و تصمیم‌گیری آن‌ها می‌تواند محل بررسی قرار گیرد.

۱-۲. نقش هوشمندی در عاملیت انسان: مفهوم هوش یکی از دشوارترین مفاهیم نزد متخصصان علوم شناختی از حیث تعریف آن است. با این وجود برآیند شاخص‌های هوشمندی نزد اندیشمندان حاکی از چهار ویژگی اساسی است: ۱- توانایی سازگاری با محیط و تطبیق با موقعیت‌های جدید؛ ۲- توانایی یادگیری و آموزش‌پذیری؛ ۳- توانایی تفکر انتزاعی؛ ۴- توانایی حل مسئله (۱، ۲). تلاش‌های مختلفی برای تعیین کمیت هوش انسان از طریق آزمون‌های روان‌سنجی صورت گرفته است. از جمله مهم‌ترین این تست‌ها، تست‌های ضریب هوشی (IQ) هستند. یکی از این آزمون‌های مشهور، نظریه ترکیبی روان‌سنجی Cattell-Horn-Carroll (CHC) است که بر پایه ترکیب نظریه هوش سیال (Fluid Intelligence) و هوش متبلور (Crystallized Intelligence)، Horn و Cattell و نظریه هوش سه لایه Carroll (The Three-Stratum Theory) استوار بود. هوش سیال توانایی حل مسائل استدلالی جدید و شناسایی راه حل در مواجهه با محرک‌ها و موقعیت‌های تجربه نشده است. از سوی دیگر، هوش متبلور شامل توانایی استدلال استنتاجی انتزاعات رابطه‌ای ثانویه با استفاده از انتزاعات رابطه‌ای اولیه آموخته‌شده قبلی است (۴۴-۴۶).

هچنین بر اساس نظریه هوش سه لایه Carroll، انسان‌ها از سه لایه (قشر) به عنوان نمایانگر توانایی شناختی باریک، گسترده و کلی برخوردارند. این عوامل تفاوت‌های پایدار و قابل مشاهده را در بین توانایی‌های افراد در انجام وظایف توصیف می‌کنند. این توانایی‌های چند لایه شامل مؤلفه‌هایی نظیر دانش کمی، خواندن و نوشتن، درک مطلب، دانش استدلال سیال، حافظه کوتاه‌مدت ذخیره‌سازی، بازیابی بلندمدت، پردازش دیداری، پردازش شنیداری، سرعت پردازش، سرعت واکنش و تصمیم‌گیری می‌باشد (۴۷). به عقیده Carroll این لایه‌های سه‌گانه، مصنوعات یک فرآیند الگوریتمی ریاضی

رفتار عقلانی (منطقی) (۳۶، ۳۷). تعریف چهارم رایج‌ترین تعریف از هوش مصنوعی نزد متخصصان است که بر مقوله عقلانیت عملی (عاملیت عقلانی) تمرکز دارد. مطالعه آثار اندیشمندان در زمینه عقلانیت حاکم بر هوش مصنوعی حاکی از آن است که ماشین‌های هوشمند کنونی به دلیل برخورداری از ذهن الگوریتمی و اتکا بر ساختار محاسباتی هدف - وسیله صرفاً از عقلانیت شکلی بهره می‌برند و فاقد عقلانیت ماهوی می‌باشند. به دلیل اینکه عملکرد حل مسئله‌محورانه ماشین‌های هوشمند در مواجهه با محرک‌های محیطی، بر پایه پردازش داده‌ها از طریق محاسبه اعمال مختلف و گزینش بهترین عمل از میان آن‌ها به منظور بهینه‌سازی (حداکثرسازی) هدف - وسیله‌محورانه منافع مورد هدف یا انتظار است (۶، ۳۸، ۳۹).

برخی پژوهش‌ها و آزمایش‌های صورت‌گرفته نظیر آزمایش اتاق چینی (Chinese Room) که توسط Searle، Damasio و همکاران، در خصوص عقلانیت حاکم بر هوش مصنوعی انجام گردیده است، نیز حاکی از آن است که ارزیابی‌های هوش مصنوعی در وضع کنونی صرفاً به صورت دیجیتالی و بر پایه عقلانیت شکلی، عینی و محدود متکی است و هوش مصنوعی در حال حاضر برخلاف انسان، فاقد عقلانیت کامل، فاقد ذهنیت (Subjectivity) و خودآگاهی لازم جهت تفکر عقلانی، فاقد قصدیت یا حیث التفاتی (Intentionality) و فاقد ساختار آنالوگی و نوروشیمیایی عملکرد مغز انسان جهت ادراک و فهم معنای تصمیم‌گیری‌های انتخابی است، لذا قادر به تفسیر و فهم تفاوت‌های جزئی و فرهنگی و برهم‌کنش مؤثر با زمینه‌های موضوع تصمیم‌گیری و انجام قضاوت‌های اخلاقی شبیه انسان نیست، در نتیجه توانایی عاملیت اخلاقی - حقوقی مستقل ندارد (۹-۱۱، ۱۵، ۱۶، ۴۰-۴۳).

۲. نقش هوشمندی در عاملیت انسان و هوش مصنوعی: هوشمندی یکی از عوامل مؤثر و تعیین‌کننده در پردازش شناختی و تصمیم‌گیری در انسان است. با توجه به اینکه هوش مصنوعی نیز به دنبال شبیه‌سازی هوشمندی خود از

نیستند، بلکه منعکس‌کننده عوامل فیزیولوژیکی هستند که تفاوت‌های انسان‌ها در توانایی را توضیح می‌دهند، نظیر نرخ شلیک عصبی که در میان انسان‌ها متفاوت است (۴۸).

۲-۲. نقش هوشمندی در عاملیت هوش مصنوعی: یکی از مسائل مطرح در زمینه هوش مصنوعی، نوع هوشمندی شبیه‌سازی‌شده از انسان در فناوری‌های هوش مصنوعی است. مسأله این است که آیا هر دو مدل هوشمندی انسان در فناوری‌های هوش مصنوعی به کار رفته است یا اینکه صرفاً یکی از آن‌ها قابلیت به کارگیری و شبیه‌سازی در هوش مصنوعی را دارد؟ بنا بر نظر برخی از پژوهشگران حوزه هوش مصنوعی، هوش سیال چیزی است که در مواجهه با موقعیت‌های جدید و سابقاً تجربه‌نشده برای حل یک مسأله یا کار مؤثر است. ماشین‌های هوشمند در حال حاضر امکان حل مسائل و مشکلاتی که هنوز دیده یا با آن مواجه نشده‌اند، ندارند (۴۹). به عبارت دیگر هوش مصنوعی عمدتاً مبتنی بر هوش متبلور و دانش مربوط به وظیفه‌ای است که قبلاً داده شده است نه هوش سیال. بر همین اساس این دسته از پژوهشگران معتقدند در صورتی که در آینده بتوان ماشین‌های هوشمندی با قابلیت هوش سیال آفرید که قادر به حل مسأله در خصوص امور سابقاً تجربه نشده باشند، می‌توان به اندازه انسان از هوشمندی آن‌ها سخن گفت. برخی دیگر از پژوهشگران سطح بالای هوشمندی انسان را ناشی از ساختار نورومورفیک و بیولوژیکی مغز انسان می‌دانند که چنین ساختاری در هوش مصنوعی وجود ندارد. بنابراین در خوش‌بینانه‌ترین حالت، ماشین‌های هوشمند صرفاً سطح بسیار پایینی از هوش سیال و متبلور انسان را بازنمایی می‌کنند. از منظر آنان یکی از ویژگی‌های کلیدی مغزهای بیولوژیکی که احتمال یادگیری سریع از نمونه‌ها یا آزمایشات محدود را امکان‌پذیر می‌کند، ویژگی‌های ساختاری است که در نتیجه تکامل در مغزهای بیولوژیکی وجود دارد که باید از طریق فرآیند یادگیری سفارشی شود. هدف یک کامپیوتر نورومورفیک نباید تقلید از مغز باشد، بلکه در عوض باید از زیست‌شناسی الهام بگیرد، اما در عین حال خود را به مدل‌ها

یا الگوریتم‌های خاص محدود نکند. وضعیت فعلی صنعت الگوریتم‌های هوشمندی مصنوعی بر پایه محاسبات Von Neumann پیاده‌سازی می‌شوند، لذا نمی‌تواند عامل سطح بالای (g-factor/mood) نظریه سه لایه C-H-C هوش را مدل‌سازی کنند. رویکرد هوش مصنوعی فعلی می‌تواند صرفاً برخی از فعالیت‌های شناختی سطح پایین‌تر هوش سیال و هوش متبلور را تحقق بخشند (۵۰).

۳. نقش عاطفه‌مندی در عاملیت انسان و هوش مصنوعی: عواطف نقش تعیین‌کننده‌ای در فرایند تصمیم‌گیری انسان دارند. مسأله توانایی ماشین‌های هوشمند در شبیه‌سازی ظرفیت‌های عاطفی انسان در تصمیم‌گیری یکی از چالش‌های پیش روی هوش مصنوعی محسوب می‌شود. در این قسمت ابتدا نقش عواطف در تصمیم‌گیری انسان مورد بحث قرار می‌گیرد، سپس به امکان‌سنجی کاربست عاطفه‌مندی در عاملیت سیستم‌های هوش مصنوعی پرداخته می‌شود.

۳-۱. نقش عواطف در عاملیت انسان: بنا بر سنتی غالب از منظر فیلسوفان یونان باستان تا عصر روشنگری، عواطف نقش مخربی در تصمیم‌گیری عقلانی داشته‌اند، به نحوی که با طرح‌اندازی ثنویت میان عقل / احساس، تصمیم‌گیری ایده‌آل، تصمیم‌گیری ناشی از استدلال عقلانی و عاری از هیجانات و انگیزش‌های عاطفی برشمرده شد. این ثنویت در لباسی دیگر در نظریه پردازش دوگانه Kahneman و متفکران متأثر از اندیشه‌های او نیز احیا گردیده است. در این نظریه سامانه: ۱- مرتبط با فرایندهای هیجانی، خودکار و شهودی در تصمیم‌گیری است و سامانه؛ ۲- با پردازش عقلانی، استدلالی و با درنگ و آگاهانه در ارتباط است. با این حال از اواخر قرن بیستم گرایش‌های نوینی حاکی از طرفداری از نقش ضروری و مثبت عواطف در فرایند تصمیم‌گیری عقلانی به وجود آمده است. برآیند نظریه‌پردازان دانشمندان علوم شناختی معاصر در زمینه نقش عواطف در تصمیم‌گیری حاکی از این است که عواطف به شیوه‌های مختلف در تصمیم‌گیری و تفکر عقلانی نقش مثبت دارند، نظیر بهبود دسترسی انسان به دلایل و استدلال‌های خوب، تسهیل در طراحی و برنامه‌ریزی تصمیم،

تکمیل استدلال نظری تصمیم و رفع نواقص آن و تسهیل در اقدام مناسب و سریع در جایی که استدلال محض و کند قادر به تصمیم‌گیری به موقع و انتخاب تصمیم نیست (۵۱). ترکیب یافته‌های جدید در زمینه عاطفه‌مندی و هوشمندی زمینه‌ساز الگوی جدیدی از هوشمندی موسوم به هوش عاطفی (Emotional Intelligence) گردیده است. از دیدگاه Goleman هوش عاطفی شامل ویژگی‌هایی مانند خودآگاهی (توانایی شناخت احساسات، نقاط قوت، ضعف، انگیزه‌ها، ارزش‌ها و اهداف خود)، خودتنظیمی (کنترل احساسات و انگیزه‌های مخرب فرد و سازگاری با شرایط در حال تغییر)، مهارت اجتماعی (مدیریت روابط برای کنارآمدن با دیگران)، همدلی (در نظرگرفتن احساسات دیگران به ویژه هنگام تصمیم‌گیری) و انگیزه (آگاهی از آنچه آن‌ها را برمی‌انگیزد) می‌شود (۵۲) از منظر Salovey و Mayer، هوش عاطفی شامل توانایی‌هایی نظیر درک احساسات، استفاده از احساسات، فهم احساسات، مدیریت احساسات می‌شود (۷، ۵۲، ۵۳)

۲-۳. نقش عواطف در عاملیت هوش مصنوعی: بنا بر گرایش غالب در میان عصب‌پژوهان، پایه و اساس عاملیت در انسان و کلیه ارگانیسم‌های آگاه، تجربه ادراکی آن‌هاست که صرفاً یک عملکرد مکانیکی محض نیست، بلکه حاصل برهم‌کنش بدن‌مند و عاطفی سه عنصر مغز، بدن و محرک‌های محیطی است. در همین راستا انسان‌ها به عنوان نوعی ارگانیسم آگاه دارای عنصری موسوم به خود یا خویش‌نشدن هستند که آن را احساس می‌کنند. برآیند یافته‌ها و دستاوردهای عصب‌پژوهانی نظیر Oregan و Damasio حاکی از آن است که پایه و اساس عاملیت انسان در برهم‌کنش با محیط پیرامونی، تجربه‌ای پدیداری و آگاهانه است که بر پایه مفهوم خود صورت می‌گیرد. مفهوم خود متضمن توانایی تمایز ارگانیسم آگاه بین خود و جهان پیرامونی، توانایی دانستن درباره خود (خودآگاهی) است (۴۱، ۵۴، ۵۵).

از منظر این عصب‌پژوهان، ماشین‌های هوشمند کنونی به دلیل ساختار غیر بیولوژیکی آن‌ها و فقدان عاطفه‌مندی که زیربنای احساس خویش‌نشدن در موجودات بیولوژیکی نیز محسوب

می‌شود، فاقد تجربه آگاهانه و پدیداری هستند. با این حال تحقق مؤلفه‌هایی نظیر احساس خویش‌نشدن و آگاهی پدیداری در ماشین‌های هوشمند در آینده دور از دسترس نیست (۱۱)، (۲۳).

باور عصب‌پژوهان پدیدارشناس نظیر Damasio به لزوم وجود عواطف و احساسات به عنوان زیربنای بدن‌مندانه شکل‌گیری مفهوم خود و تجربه آگاهانه در برهم‌کنش میان مغز، بدن و محیط و دسترس‌پذیری تحقق آن‌ها در ماشین‌های هوشمند آینده، در کنار رواج گسترده نظریه مدل پردازش دوگانه ذهنی Kahneman و Tversky، زمینه‌ساز اشتیاق روزافزون متخصصان فناوری‌های هوش مصنوعی به سمت کاربست و شبیه‌سازی مؤلفه‌های ذهنی و عاطفی ارگانیسم‌های آگاه در ماشین‌های هوشمند گردیده است. این اشتیاق دو رهاورد توسعه‌ای به همراه داشته است. توسعه اول پدیداری مدل‌های جدیدی از هوش مصنوعی موسوم به هوش مصنوعی عصبی - نمادین (Neuro-Symbolic AI) است. هوش مصنوعی عصبی - نمادین نوعی از هوش مصنوعی است که معماری‌های هوش مصنوعی عصبی و نمادین را برای رفع نقاط ضعف هر یک ادغام می‌کند و یک هوش مصنوعی قوی با قابلیت استدلال، یادگیری و مدل‌سازی شناختی ارائه می‌دهد (۵۶). تحت تأثیر نظریه پردازش دوگانه ذهنی Kahneman، شناخت در هوش مصنوعی شامل دو جزء توصیف می‌شود: ۱- سیستم ۱ سریع، بازتابی، شهودی و ناخودآگاه است؛ ۲- سیستم ۲ کندتر، گام به گام و واضح است. سیستم ۱ برای تشخیص الگو استفاده می‌شود. سیستم ۲ برنامه‌ریزی، استنباط و تفکر مشورتی را مدیریت می‌کند. در این دیدگاه، یادگیری عمیق به بهترین وجه نوع اول شناخت را مدیریت می‌کند، در حالی که استدلال نمادین به بهترین شکل نوع دوم را مدیریت می‌کند. هر دوی آن‌ها برای یک هوش مصنوعی قوی و قابل اعتماد که می‌تواند یاد بگیرد، استدلال کند و با انسان‌ها برای پذیرش مشاوره و پاسخ به سؤالات تعامل داشته باشد، مورد نیاز است (۲۴، ۳۰، ۳۲، ۵۷).

توسعه دوم ظهور هوش عاطفی انسان در ربات‌ها و ماشین‌های هوشمند است. هوش مصنوعی عاطفی که به عنوان تحلیل احساسات یا محاسبات عاطفی نیز شناخته می‌شود، به توانایی سیستم‌های هوش مصنوعی برای تشخیص، تجزیه و تحلیل و تفسیر احساسات انسان از طریق ورودی‌های مختلف (مانند متن، گفتار، حالات چهره و سیگنال‌های فیزیولوژیکی) اشاره دارد. با پیشرفت‌های اخیر در هوش مصنوعی و یادگیری ماشین (Machine Learning)، تحلیل احساسات پیشرفت قابل توجهی از نظر دقت، کارایی و مقیاس‌پذیری رخ داده است (۲۵). در این چهارچوب یک ربات مجهز به سیستم هوش مصنوعی می‌تواند معمولاً سه کار انجام می‌دهند: ۱- حس؛ ۲- محاسبه؛ ۳- عمل. بنابراین، یک ربات دارای هوش عاطفی، باید مجهز به حسگرهایی باشد که قادر به ثبت احساسات کاربران (حس)، ارزیابی احساسات گرفته‌شده برای تنظیم حالت داخلی آن (محاسبه) و در نهایت انجام وظایفی است که در آن اعمال توسط حالت «عاطفی» محاسبه‌شده تنظیم می‌شوند (عمل) (۵۸). در مرحله اول (تشخیص خودکار عواطف) فرآیند شناسایی احساسات انسانی از کانال‌های ارتباطی مختلف، مانند ژست‌های صورت یا بدن، سیگنال‌های فیزیولوژیکی، تفاوت‌های ظریف صدا، محتوای گفتار و... جمع‌آوری و برچسب‌گذاری این‌گونه سیگنال‌ها انجام می‌شود (۲۵، ۵۸-۶۱)؛ در مرحله دوم (محاسبه)، جنبه‌های کیفی ادراک عاطفی و احساسی انسان با روش‌های الگوریتمی به صورت کمی در ماشین‌های هوشمند شبیه‌سازی می‌شوند. بسیاری از پیشرفت‌های مهم در یادگیری ماشین و هوش مصنوعی مبتنی بر اصول بیولوژیکی مانند شبکه‌های عصبی یا الگوریتم‌های تکاملی هستند. بنابراین یک رویکرد محاسباتی به احساسات می‌تواند نمونه دیگری از الهامات زیستی در علم محاسبات محسوب می‌شود. بنابراین هوش مصنوعی عاطفی در رباتیک را می‌توان این‌گونه دید که چگونه یک سیستم، احساسات را با تمرکز بر نحوه ترجمه ورودی از طریق یک الگوریتم به یک خروجی پردازش می‌کند و اینکه آیا حاوی دانشی از رویدادهای گذشته است یا نه (۵۸-۶۱)؛ در مرحله سوم (توانایی ماشین‌های هوشمند نظیر ربات جهت ابراز

واکنش عاطفی یا عمل) هوش مصنوعی عاطفی در نحوه ابراز رفتار یا واکنش عاطفی بیرونی به محرک‌ها با انسان شبیه‌سازی می‌شود. ابراز واکنش‌های عاطفی مستلزم وجود ساختار و سیستم‌های داخلی نظیر حافظه و سیستم بیولوژیکی نظیر حالات چهره، عضلات، صدا، غدد درون‌ریز، به منظور ایجاد یک محیط داخلی بهینه برای مؤثرترین رفتار است (۵۸، ۶۲).

کاربست و شبیه‌سازی ساختار بیولوژیکی واکنش‌های عاطفی انسان در ماشین‌های هوشمند مستلزم یک هوش مصنوعی تکامل یافته است که به درجه بسیاری بالای انسان‌انگاری ربات‌ها و ماشین‌های هوشمند بستگی دارد (۲۶، ۵۸). بر همین اساس با وجود پیشرفت چشم‌گیر در سال‌های اخیر از نظر هوش مصنوعی، هنوز از امکان نسبت‌دادن قابلیت‌های همدلانه انسان به ربات‌ها فاصله داریم. با این حال افزودن توانمندی‌های هوش عاطفی به هوش مصنوعی، نه تنها می‌تواند شکاف بین انسان و ماشین‌ها را کاهش دهد، بلکه راه حل‌های پیشرفته‌ای برای مشکلات پیچیده انسان و پدیداری ربات‌های هوشمند اجتماعی ارائه می‌کنند که آن‌ها را قادر به درک و پاسخ به احساسات انسانی، ایجاد فرصت برای بهبود تعاملات، تصمیم‌گیری مستقل و حل مسأله می‌سازد (۲۶، ۵۸). برخی از محققان پیشنهاد می‌کنند که با اعطای سیستم‌هایی با قابلیت احساسات به هوش مصنوعی، آن‌ها قادر به ارزیابی و اتخاذ قضاوت و تصمیمات اخلاقی به روش‌های خود هستند، به نحوی که از نظر اجتماعی و قانونی می‌توانند مسئول شناخته شوند، اعطای چنین قابلیت‌هایی موجب پدیداری ماشین‌هایی باملاحظه و محتاط، مراقب دلسوز، درمانی، شفاف‌بخش و اخلاقی خواهد شد (۲۷).

بدین ترتیب ماحصل پژوهش‌های عصب‌پژوهان و متخصصان علوم شناختی از یک طرف و همچنین متخصصان فناوری‌های هوش مصنوعی از طرف دیگر بیانگر دو مطلب قابل توجه است: ۱- در وضع کنونی علیرغم پیشرفت‌های گسترده هوش مصنوعی عاطفی در شبیه‌سازی توانمندی‌های عاطفی انسان در ادراک، محاسبه و واکنش‌های عاطفی به محرک‌های محیطی، هوش مصنوعی کنونی دارای فاصله قابل توجهی از

قابلیت‌های عاطفی انسان در اداراک، محاسبه و تصمیم‌گیری عاطفی است تا جایی که برخی اندیشمندان علوم شناختی به دلیل اعتقاد به ساختار بیولوژیکی و نوروفورمیک واکنش‌های عاطفی، اساساً مخالف کاربست و انتقال چنین قابلیت‌هایی کیفی به الگوریتم‌های کمی و غیر بیولوژیکی ماشین‌های هوشمند هستند (۱۰، ۱۱، ۴۳، ۵۸)؛ ۲- پیشرفت‌های روزافزون انتقال توانمندی‌های عاطفی انسان به هوش مصنوعی، این باور را برای برخی اندیشمندان حوزه علوم شناختی و هوش مصنوعی پدید آورده است که در آینده می‌توان ماشین‌هایی با قابلیت‌های کامل عاطفی انسانی پدید آورد (۱۷، ۲۸).

۴. امکان‌سنجی عاملیت اخلاقی - حقوقی هوش مصنوعی:

با توجه به آنچه در بخش‌های پیشین اشاره شد، سه رویکرد از منظر علوم شناختی معاصر مطرح است: رویکرد اول منکر عاملیت اخلاقی هوش مصنوعی است؛ رویکرد دوم به عاملیت حداقلی هوش مصنوعی پایبند است؛ رویکرد سوم به مشارکت و ترکیب عاملیت تکنولوژی ماشین‌های هوشمند با انسان باور دارد. در ادامه به تحلیل سه رویکرد می‌پردازیم:

۴-۱. فقدان عاملیت اخلاقی هوش مصنوعی: همانطور که در بخش پیشین ذکر شد، عاملیت کامل انسان از دو قسمت عاملیت عقلانی و عاملیت اخلاقی تشکیل می‌شود. عاملیت اخلاقی با جنبه‌های عقلانیت ماهوی و عاملیت عقلانی با عقلانیت شکلی منطبق است. بنا بر نظر پژوهشگران و اندیشمندان عاملیت حاکم بر هوش مصنوعی، عاملیت کامل نیست، بدین‌ترتیب که هوش مصنوعی صرفاً از عقلانیت شکلی بهره‌مند بوده است. به عبارت دیگر آنچه در ماشین‌های هوشمند مشاهده می‌شود، شبیه‌سازی فرایند پردازش و محاسبه الگوریتمی داده‌ها و انتخاب بهترین گزینه به عنوان تصمیم است (۸، ۳۸، ۳۹).

بر همین اساس از منظر بسیاری از متخصصان حوزه فناوری‌های هوشمند، روان‌شناسان، عصب‌پژوهان و فیلسوفان ذهن، هوش مصنوعی به جهت عدم برخورداری از عقلانیت ماهوی و حیث التفاتی، از عاملیت اخلاقی برخوردار نخواهد

بود، چراکه عامل اخلاقی دارای حیث التفاتی (نظیر ادراک، آگاهی و فهم نسبت به خوب و بد، باور به چیزی نظیر درست و نادرست، میل، قصد یا نیت به چیزی نظیر خوردن، خریدن، کشتن و...) است و هوش مصنوعی در وضع کنونی فاقد حیث التفاتی نظیر آگاهی، خودآگاهی و قصد، میل و باور به چیزی است و کیفیات ذهنی و التفاتی که خاستگاهی بیولوژیکی دارند، قابل فروکاستن به جنبه‌های کمی و الگوریتمی در موجودات غیر بیولوژیکی (هوش مصنوعی) نیستند (۱۰، ۳۶، ۴۳، ۶۳-۶۶). از دیگر دلایل مخالفین عاملیت اخلاقی هوش مصنوعی از دیدگاه اندیشمندان علوم شناختی این است که انسان در تصمیم‌گیری‌های اخلاقی خود از عواطف و احساسات مدد می‌جوید. به عبارت دیگر بر مبنای یافته‌های نوین عصب‌شناسان، هیچ‌گونه مرزبندی میان عاطفه و عقلانیت در مغز انسان وجود ندارد و در فرایند ذهنی - مغزی تصمیم‌گیری در سیستم عصبی انسان، ترکیبی از برهم‌کنش عاطفه و عقلانیت (بدن، مغز) با محیط پیرامون دخیل هستند. از آنجا کیفیات عاطفی انسان برگرفته از ویژگی‌هایی بیولوژیکی و آناتومیک سیستم عصبی و مغزی انسان به عنوان یگ ارگانیسم (موجود زنده) است و چنین قابلیت‌هایی در ماشین‌های هوشمند و انعطاف‌ناپذیر وجود ندارد، در نتیجه هوش مصنوعی ظرفیت‌های کامل جهت عاملیت اخلاقی را ندارد، هرچند تصمیماتش می‌تواند دارای آثار اخلاقی - حقوقی باشد (۱۱، ۶۷). به طور خلاصه از منظر عصب‌شناسان، ماشین‌های هوشمند به دلایل زیر در وضع کنونی از عاملیت اخلاقی مستقلاً برخوردار نخواهند بود:

۱- تمرکز هوش مصنوعی بر مدل کامپیوتری و الگوریتمی کارکرد محاسباتی مغز انسان در ادراک و تصمیم‌گیری و عدم توجه به تمایلات فکری (باورها، امیال و ارزش‌ها) انسان در گزینش نهایی تصمیم (۶)؛

۲- فقدان لایه‌های عصب - بوم‌شناختی و ذهنی مغز انسان در ماشین‌های هوشمند (۴۰)؛

۳- فقدان برهم‌کنش مؤلفه‌های سه‌گانه طبیعت انسانی (بدن، ذهن و نفس یا خود)، فقدان ذهنیت و مؤلفه‌های ذهنی و انتزاعی اشخاص آگاه نظیر خود (من)، آگاهی، حیث التفاتی و

احساس وجود خود (خودآگاهی) (۱۲، ۱۳). در همین راستا در زمینه عاملیت و مسئولیت حقوقی ناشی از تصمیمات هوش مصنوعی، فقدان این ویژگی‌ها سبب گردید اتحادیه اروپا در ماده ۷ قطعنامه خود در خصوص رژیم مسئولیت مدنی ناشی از عملکرد سیستم‌های هوش مصنوعی مصوب ۲۰۲۰، عاملیت و مسئولیت‌پذیری مستقل سیستم‌های هوش مصنوعی ضعیف را منتفی نماید و صرفاً عوامل انسانی مرتبط با عملکرد هوش مصنوعی (تولیدکننده، اپراتور، کاربر و...) را مسئول جبران خسارت وارده از ناحیه سیستم بداند.

۴-۲. عاملیت حداقلی برای هوش مصنوعی: طرفداران
رویکرد دوم با تمرکز بر شبیه‌سازی رفتاری ماشین‌های هوشمند با انسان‌ها که دربردارنده جنبه‌های عینی و بیرونی است و با پرهیز از تعمق در کیفیات ذهنی و درونی تفکر انسانی، بر این باورند که اساساً عاملیت به معنای توانایی یک عامل برای عمل کردن و ایجاد آثار بیرونی (حداکثرسازی منافع مورد انتظار) است (۳۶). طرفداران این رویکرد معتقدند برخلاف نظر مخالفین، ماشین‌های هوشمند، حتی نسبت به انسان نیز در بسیاری از عملکردها موفق‌ترند، به دلیل اینکه برخلاف انسان تحت تأثیر عواطف و هیجانات تصمیم نمی‌گیرند و کمتر از انسان دچار اشتباه و سوگیری در محاسبه داده‌ها و گزینش بهترین تصمیم هستند (۸، ۱۶) تا جایی که برخی طرفداران نظریه انتخاب عقلانی، تصمیمات اقتصادی یا سیاسی اتخاذی از سوی هوش مصنوعی را با عقلانیت ابزاری فایده‌گرا (محاسبه الگوریتمی هزینه و فایده و گزینش بهینه‌سازانه فواید و منافع) سازگارتر می‌دانند، به نحوی که آینده با وجود پیش‌بینی‌های دقیق و تصمیم‌گیری‌های مطلوب و حل مسائل پیچیده توسط ماشین‌های هوشمند، مکان معقول‌تری و مناسب‌تری خواهد بود (۱۴، ۱۸، ۶۸). به عبارت دیگر به کارگیری هوش مصنوعی بر مبنای نظریه انتخاب عقلانی دو اثر عمده خواهد داشت:

۱- هوش مصنوعی از اطلاعات در قالب‌های مختلف استفاده می‌کند. متن، عکس، اینترنت و آن‌ها را ترکیب می‌کند تا انتظارات منطقی بهتری از آینده ایجاد کند و تصمیم‌گیری را

به انتخاب منطقی و عقلانی نزدیک‌تر کند (۷)؛ در انتخاب بهینه روش‌های بهینه‌سازی جهانی مانند ژنتیک، الگوریتم احتمال بیشتری برای شناسایی یک ابزار بهینه جهانی می‌دهد و در نتیجه تصمیم‌گیری به نظریه انتخاب عقلانی نزدیک‌تر است. بر همین اساس به باور برخی اندیشمندان، ماشین‌های هوشمند با برخورداری از عاملیت عقلانی (از نوع عقلانیت شکلی - ابزاری) دارای عاملیت حداقلی هستند (۸)، چراکه به دلیل پیروی کردن از منطق محض مبتنی بر محاسبه الگوریتمی حداکثر منافع مورد انتظار و تأثیرپذیری کمتر از هیجانات و سوگیری‌های شناختی (مؤلفه‌هایی که از منظر این اندیشمندان در تصمیم‌گیری انسان‌ها تأثیر منفی دارد) دارای عاملیت اخلاقی برتر از انسان هستند (۸). به عبارت دیگر می‌توان ارزش‌های اخلاقی انسانی نظیر دوستی، محبت، کمک به هم‌نوع، اجتناب از آسیب‌زدن به دیگری را در قالب داده‌های کدگذاری شده در ماشین‌های هوشمند نهادینه کرد، بدین ترتیب با تجهیز آن‌ها به کدهای اخلاقی، می‌توان برای آن‌ها عاملیت اخلاقی قائل شد، حتی اگر به‌مانند انسان از احساس و آگاهی ذهنی برخوردار نباشند (۱۶)، بدین ترتیب با گذار از مؤلفه‌های ذهنی عاملیت عقلانی - اخلاقی به مؤلفه‌های عینی، هوش مصنوعی، حتی می‌تواند عملکرد بهتری نسبت به انسان داشته باشد، لذا می‌توان برای آن عاملیت اخلاقی قائل شد، اما اینکه عاملیت اخلاقی به‌منزله عاملیت مسئول باشد، نظر قالب اندیشمندان حوزه اخلاق این است که مسئولیت‌پذیری از مؤلفه‌های ضروری عاملیت اخلاقی مستقل است. در چنین حالتی ما با این پرسش تردیدآمیز مواجه می‌شویم که چگونه می‌توان بدون به رسمیت‌شناخته‌شدن ماشین‌های هوشمند به عنوان یک شخصیت حقوقی دارای حقوق و تکالیف مستقل که ضرورت‌ها و مقدمات پیشینی مسئولیت محسوب می‌شوند، برای آن‌ها مسئولیت قائل شد؟ پرسشی که رویکرد دوم قادر به پاسخگویی آن نیست.

۴-۳. عاملیت مشارکتی هوش مصنوعی و انسان: بر
مبنای رویکرد سوم، ماشین‌های هوشمند خواه عاملان خودآگاه اخلاقی محسوب شوند، خواه صرفاً یک ابزار هوشمند

برای انسان و فاقد عاملیت اخلاقی، نوعی دیگری محسوب می‌شوند که قابلیت تصمیم‌گیری به جای انسان یا لغو تصمیمات انسانی را نیز دارند. خاستگاه رویکردی که به عاملیت مشارکتی انسان و هوش مصنوعی باورمند است، به نظریه کنشگر شبکه (Bruno Latour Actor-Network Theory) بازمی‌گردد. بر اساس این نظریه که از اندیشه‌های John Dewey در زمینه تعامل و برهم‌کنش انسان با محیط طبیعی - اجتماعی الهام می‌گیرد، در جهان اجتماعی و طبیعی شبکه‌هایی از روابط در حال تغییر وجود دارد، به گونه‌ای که هیچ چیز خارج از این روابط وجود ندارد. همه عوامل دخیل در یک موقعیت اجتماعی در یک سطح هستند. بنابراین اشیاء، ایده‌ها، فناوری‌ها و هر عامل مرتبط دیگر به همان اندازه انسان در ایجاد موقعیت‌ها و شبکه‌های فنی - اجتماعی دارای عاملیت خواهد بود (۶۹، ۷۰).

بنا بر این نظریه مصنوعات انسانی، آنچنان مستقل و خودمختار محسوب نمی‌شوند تا بتوانند اراده خود را بر انسان تحمیل کنند، اما در عین حال به‌مثابه یک ابزار یا وسیله محض نیز تلقی نمی‌شوند، بلکه به عنوان واسطه‌هایی هستند که بر اراده و عاملیت در تصمیم‌گیری انسان‌ها تأثیر گذارند تا جایی که می‌توان گفت انسان‌ها نه «به وسیله» مصنوعات، بلکه «به همراه» آن‌ها تصمیم می‌گیرند (۷۱، ۷۲). بر اساس پژوهش‌های صورت‌گرفته، کاربرست نظریه کنشگر شبکه در تکنولوژی‌های برخوردار از الگوریتم‌های هوش مصنوعی (نظیر Chat GPT)، منتج به این امر گردیده است که امروزه ماشین‌های هوشمند همراه با اپراتورها و تولیدکننده‌های انسانی، در شبکه‌های بسیاری پیچیده‌ای از روابط در محیط‌های اجتماعی در حال بازیگری و عاملیت هستند، به نحوی که این بازیگری و عاملیت منجر به نتایج مفید و مثبت در زمینه گسترش دانش و آموزش، پویایی در مدیریت سازمان‌ها، ایجاد نظام اجتماعی - فنی و هم نتایج منفی و نامطلوب نظیر سوگیری‌های الگوریتمی، افشای اسرار و حریم خصوصی، ارائه اطلاعات نادرست و... شده است (۱۹-۲۱). بر همین اساس پژوهشگران معاصر اخلاق تکنولوژی بر پایه مقدمات ذیل قائل به لزوم بازتعریف مسئولیت اخلاقی انسان‌ها

در عصر فناوری‌های نوین هستند: ۱- انسان‌ها به طور فزاینده‌ای از فناوری‌های تفویض‌شده استفاده می‌کنند؛ ۲- از آنجا که فناوری‌ها، میانجی انجام اهداف اولیه ما هستند، با دخالت فناوری، اهداف اولیه تغییر می‌کنند و نتایج اغلب متفاوت از آنچه در نظر گرفته شده، پدید می‌آید؛ ۳- از آنجایی که نتایج اغلب پیش‌بینی‌نشده و ناخواسته است، دیگر نمی‌توانیم به سادگی از مدل‌های سنتی (مدرنیستی) برای بحث مسئولیت اخلاقی استفاده کنیم، بنابراین باید مسئولیت اخلاقی را دوباره تفسیر کنیم (۲۲). الگوی بازتعریف‌شده مسئولیت اخلاقی انسان در عاملیت مشارکتی او با فناوری‌ها از منظر این پژوهشگران بر پایه مفروضاتی چون توجه به نقش اجتماعی فرد به عنوان هدف مسئولیت (به جای توجه محض بر اعمال)، نقش اخلاقی افراد (به جای آزادی مطلق‌تر برای عمل)، نقش واسطه‌گری مصنوعات فناورانه در تحقق اعمال و نیت انسان و توجه به پویایی چندین عمل متقابل استوار است (۲۲).

با این وجود رویکرد سوم نیز از نظر نویسنده به مانند رویکرد دوم در مسأله پاسخگویی به مسئولیت‌پذیری مستقل ماشین‌های هوشمند ناتوان است، چراکه فی‌المثل مشارکت ماشین‌های هوشمند در تحمل مجازات یا مسئولیت مدنی جبران خسارت به دلیل میانجی‌گری و مشارکت با انسان در انجام رفتار خطا، امکان‌پذیر نیست (۸). با این وجود به نظر نویسنده چنانچه خطایی انسانی در طراحی و تولید یا نظارت و کنترل بر عملکرد ماشین‌های هوشمند رخ نداده باشد، می‌بایست در کمیت و کیفیت مسئولیت عوامل انسانی دخیل در حادثه بازنگری نمود، به ویژه در مواقعی که به دلیل عملکرد مناسب ماشین‌های هوشمند، نوعی اعتماد و اتکای معقول نسبت به بازدهی مثبت در انجام وظایف توسط آن‌ها به وجود آمده باشد.

نتیجه‌گیری

گسترش و پیشرفت روزافزون فناوری‌های هوش مصنوعی، مسأله عاملیت اخلاقی - حقوقی مستقل سیستم‌های هوشمند را بیش از پیش در معرض اندیشه‌های اندیشمندان و

به هوش مصنوعی، نویدبخش امکان‌پذیری پدیداری مؤلفه‌های ذهنی و عاطفی انسان‌محور جهت برخورداری هوش مصنوعی از عاملیت مستقل اخلاقی و حقوقی در آینده خواهد بود. در این صورت زمینه برای پذیرش عاملیت، شخصیت و مسئولیت پذیری مستقل حقوقی هوش مصنوعی نیز فراهم خواهد شد.

مشارکت نویسندگان

محمود عباسی: طراحی ایده، مرور و بازبینی متن.
اسماعیل کشاورز صفی‌ئی: جمع‌آوری و تجزیه و تحلیل داده‌ها، نگارش اولیه، اصلاح متن و تدوین کلی مقاله.
نویسندگان نسخه نهایی را مطالعه و تأیید نموده و مسئولیت پاسخگویی در قبال پژوهش را پذیرفته‌اند.

تشکر و قدردانی

ابراز نشده است.

تضاد منافع

نویسندگان هیچ‌گونه تضاد منافع احتمالی را در رابطه با تحقیق، تألیف و انتشار این مقاله اعلام نکرده‌اند.

تأمین مالی

نویسندگان اظهار می‌نمایند که هیچ‌گونه حمایت مالی برای تحقیق، تألیف و انتشار این مقاله دریافت نکرده‌اند.

بیانیه هوش مصنوعی

نویسندگان اعلام می‌دارند که جهت تدوین، تحقیق و آماده‌سازی، تألیف و انتشار این مقاله، از هیچ یک از فناوری‌های هوش مصنوعی استفاده نکرده‌اند.

پژوهشگران این حوزه قرار داده است. یافته‌های حاصل از این پژوهش بیانگر آن است که پژوهشگران در زمینه عاملیت اخلاقی و حقوقی هوش مصنوعی قائل به سه نظر متفاوت هستند: از منظر دسته اول در وضع کنونی با وجود توانمندی‌های محاسباتی هوش مصنوعی پیشرفته که از نقطه نظر سرعت و دقت و کاهش ضریب خطا، کیفیت و عملکرد بهتری نسبت به انسان در تصمیم‌گیری از خود نشان داده‌اند، زیرساخت‌های ذهنی - فیزیکی لازم (عقلانیت، هوشمندی و عواطف) جهت عاملیت اخلاقی - حقوقی مستقل هوش مصنوعی به مانند انسان وجود ندارد، بر همین اساس هنوز تا مرحله گذار هوش مصنوعی از یک برنامه یا شی (محصول) هوشمند، معقول (منطقی) به شخص هوشمند، عاقل و عاطفی مستقل فاصله داریم؛ دسته دوم از اندیشمندان بر این باورند که با کنارگذاشتن عناصر ذهنی عاملیت (نظیر خود، آگاهی، خودآگاهی) و تمرکز بر مؤلفه‌های عینی (ادراک و سنجش محاسباتی محرک و موقعیت‌ها و گزینش بهترین پاسخ مناسب به آن) می‌توان از عاملیت حداقلی هوش مصنوعی سخن به میان آورد؛ دسته سوم ورود سیستم‌های هوش مصنوعی به عرصه فعالیت‌ها و عملکردهای انسانی را نوعی مشارکت کنشگرانه و تکنولوژیک بی‌سابقه ماشین همراه با انسان در عاملیت و گزینش تصمیم می‌دانند، بر همین اساس به لزوم بازنگری در جنبه‌های اخلاقی و حقوقی عاملیت انسان باور دارند، چراکه دیگر با عاملیت خالص و محض انسانی رو به رو نیستیم، بلکه نوعی عاملیت مشارکتی انسان - ماشین در سنجش و محاسبه داده‌های برآمده از محرک‌ها و گزینش پاسخ مناسب به آن‌ها حکمفرما است. در مجموع رویکرد دوم و سوم که قائل به عاملیت هوش مصنوعی هستند، از پاسخ به مسئولیت‌پذیری حقوقی هوش مصنوعی ناتوان هستند، چراکه لازمه مسئولیت‌پذیری حقوقی هوش مصنوعی، به رسمیت شناختن شخصیت حقوقی مستقل برای ماشین‌های هوشمند است. مسأله‌ای که پذیرش آن حداقل برای سیستم‌های هوش مصنوعی کنونی امکان‌پذیر نیست. با این وجود سرعت پیشرفت‌های چشم‌گیر فناوری هوش مصنوعی و همچنین پیشرفت‌های حاصله در زمینه کاربرست یافته‌های علوم شناختی

References

- Legg S, Hutter M. Universal intelligence: A definition of machine intelligence. *Minds and machines*. 2007;17:391-444.
- Mirdarikondi R. Analytical comparison of religious reason and cognitive intelligence. *Marifat*. 2011;20(4):35-52.
- Stanovich KE. On the distinction between rationality and intelligence: Implications for understanding individual differences in reasoning. *The Oxford handbook of thinking and reasoning*. 2012:343-65.
- Damasio A. *Descartes Error: Emotion, Reason and the Human Brain*. London: Penguin Vintage Publishing; 2006.
- Khazaei Z. Rationality of Ethics. *Journal of Religious Thought*. 2007;6(21):23-50.
- Stanovich KE, Toplak ME. The need for intellectual diversity in psychological science: Our own studies of actively open-minded thinking as a case study. *Cognition*. 2019;187:156-66.
- Nyamori D. Exploring the Relationship between Moral Decision-Making and Emotional Intelligence.
- Valtonen L. Rationality in artificial intelligence decision-making. 2023.
- Nishant R, Schneckenberg D, Ravishankar M. The formal rationality of artificial intelligence-based algorithms and the problem of bias. *Journal of Information Technology*. 2024;39(1):19-40.
- Arvan M, Maley CJ. Panpsychism and AI consciousness. *Synthese*. 2022;200(3):244.
- Damasio A. *The Strange Order of Things: Life, Feeling, and the Making of Cultures* New York: Pantheon Books, 336 s. *Filozofia*. 2018;73(6):481.
- Wróblewski Z, Fortuna P. Moral subjectivity and the moral status of artificial intelligence: A Philosophical and Psychological Perspective. *Ethos* (0860-8024). 2023;36(144).
- Zeng Y, Zhao F, Zhao Y, Zhao D, Lu E, Zhang Q, et al. Brain-inspired and self-based artificial intelligence. *arXiv preprint arXiv:240218784*. 2024.
- Marwala T. Rational choice and artificial intelligence. *arXiv preprint arXiv:170310098*. 2017.
- Northoff G, Gouveia SS. Does artificial intelligence exhibit basic fundamental subjectivity? A neurophilosophical argument. *Phenomenology and the Cognitive Sciences*. 2024:1-22.
- Nath R, Manna R. From posthumanism to ethics of artificial intelligence. *AI & Society*. 2023;38(1):185-96.
- Vicci DH. Emotional intelligence in artificial intelligence: A review and evaluation study. Available at SSRN 4818285. 2024.
- Gouiaa R, Bazarna A. The interaction between rationality, politics and artificial intelligence in the decision-making process. *Corporate Governance and Organizational Behavior Review*. 2023;7(4):263-71.
- Monteiro WR, Ayrosa EAT. From Analog To Algorithm: a comprehensive review of Artificial Intelligence's intersection with Management Literature and Actor-Network Theory. *Brazilian Journal of Information Science*. 2024(18):13.
- Venturini T. Bruno Latour and artificial intelligence. *Tecnoscienza–Italian Journal of Science & Technology Studies*. 2023;14(2):101-14.
- Reed M. The Classification of Artificial Intelligence as "Social Actors". 2018.
- Waelbers K. Technological delegation: Responsibility for the unintended. *Science and Engineering Ethics*. 2009;15:51-68.
- O'Regan JK. How to build a robot that is conscious and feels. *Minds and machines*. 2012;22:117-36.
- Booch G, Fabiano F, Horesh L, Kate K, Lenchner J, Linck N, et al. Thinking fast and slow in AI. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 2021. pp. 15042-15046.
- Kapoor A, Verma V. Emotion AI: understanding emotions through artificial intelligence. *International Journal of Engineering Science and Humanities*. 2024;14(Special Issue 1):223-32.
- Zuo G, Pan T, Zhang T, Yang Y. SOAR improved artificial neural network for multistep decision-making tasks. *Cognitive Computation*. 2021;13(3):612-25.
- Pandey P. Emotional Intelligence Skills in the Age of AI: A.
- Pan W, Yang Y. Virtual affective consciousness and raw social AI. *Computational Social Science: Routledge*; 2022. p. 3-9.
- Russell SJ. Rationality and intelligence. *Artificial intelligence*. 1997;94(1-2):57-77.
- Sun R. Dual-process theories, cognitive architectures, and hybrid neural-symbolic models. *Neurosymbolic Artificial Intelligence*. (Preprint):1-9.

31. Sheth A, Roy K, Gaur M. Neurosymbolic AI--Why, What, and How. arXiv preprint arXiv:230500813. 2023.
32. Kahneman D. Thinking, fast and slow: macmillan; 2011.
33. Stanovich KE. Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory. In two minds: Dual processes and beyond. 2009;55-88.
34. Burgoyne AP, Mashburn CA, Tsukahara JS, Hambrick DZ, Engle RW. Understanding the relationship between rationality and intelligence: A latent-variable approach. Thinking & Reasoning. 2023;29(1):1-42.
35. Russell S. Rationality and intelligence. Foundations of rational agency: Springer; 1999. p. 11-33.
36. Russell SJ, Norvig P. Artificial intelligence: a modern approach: pearson; 2016.
37. Nath R. A Cartesian critique of the artificial intelligence. 2010.
38. Lindebaum D, Vesa M, Den Hond F. Insights from "the machine stops" to better understand rational assumptions in algorithmic decision making and its implications for organizations. Academy of Management Review. 2020;45(1):247-63.
39. Raphael MW. Artificial intelligence and the situational rationality of diagnosis: human problem-solving and the artifacts of health and medicine. Sociology Compass. 2022;16(11):e13047.
40. Northoff G, Smith D. The subjectivity of self and its ontology: From the world-brain relation to the point of view in the world. Theory & Psychology. 2023;33(4):485-514.
41. Searle JR. Minds, brains, and programs. Behavioral and brain sciences. 1980;3(3):417-24.
42. Cavanna AE, Nani A. Consciousness: Theories in neuroscience and philosophy of mind: Springer; 2014.
43. Chiang CC, Shivacharan RS, Wei X, Gonzalez-Reyes LE, Durand DM. Slow periodic activity in the longitudinal hippocampal slice can self-propagate non-synaptically by a mechanism consistent with ephaptic coupling. The Journal of physiology. 2019;597(1):249-69.
44. Cattell RB. Theory of fluid and crystallized intelligence: A critical experiment. Journal of educational psychology. 1963;54(1):1.
45. Horn JL. Organization of abilities and the development of intelligence. Psychological review. 1968;75(3):242.
46. Horn JL. Intelligence—why it grows, why it declines. Human intelligence: Routledge; 2020. p. 53-74.
47. Flanagan DP, Ortiz SO, Alfonso VC. Essentials of cross-battery assessment: John Wiley & Sons; 2013.
48. Carroll JB. The Three-Stratum Theory of Cognitive Abilities. New York: Guilford Press; 1997.
49. Davidson I, Walker PB, editors. Towards fluid machine intelligence: Can we make a gifted AI? Proceedings of the AAAI Conference on Artificial Intelligence; 2019. pp. 9760-9764.
50. Ray KS. Quest for I (Intelligence) in AI (Artificial Intelligence): A non-elusive attempt. Artificial Intelligence-Latest Advances, New Paradigms and Novel Applications. 2021.
51. Jones K. Quick and smart? Modularity and the pro-emotion consensus. Canadian Journal of Philosophy Supplementary Volume. 2006;32:3-27.
52. Goleman D. Emotional intelligence, imagination, cognition and personality. Personality and Individual Differences. Vol. 1995;9:185-211.
53. Mayer JD, Salovey P, Caruso DR. Target Articles: "emotional intelligence: Theory, findings, and Implications". Psychological inquiry. 2004;15(3):197-215.
54. Damasio AR. The feeling of what happens: Body and emotion in the making of consciousness: Houghton Mifflin Harcourt; 1999.
55. O'Regan JK. Why red doesn't sound like a bell: Understanding the feel of consciousness: Oxford University Press; 2011.
56. Valiant LG, editor Knowledge infusion: In pursuit of robustness in artificial intelligence. IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (2008); 2008: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
57. Marcus G, Davis E. Rebooting AI: Building artificial intelligence we can trust: Vintage; 2019.
58. Marcos-Pablos S, Lobato F, García-Peñalvo F. Integrating emotion recognition tools for developing emotionally intelligent agents. 2022.
59. Russell JA. A circumplex model of affect. Journal of personality and social psychology. 1980;39(6):1161.

60. Zhou Q. Multi-layer affective computing model based on emotional psychology. *Electronic Commerce Research*. 2018;18(1):109-24.
61. Shu L, Xie J, Yang M, Li Z, Li Z, Liao D, et al. A review of emotion recognition using physiological signals. *Sensors*. 2018;18(7):2074.
62. Ivanović M, Budimac Z, Radovanović M, Kurbalija V, Dai W, Bădică C, et al. Emotional agents-state of the art and applications. *Computer science and information systems*. 2015;12(4):1121-48.
63. Beorlegui C. Emergentism. *Pensamiento Revista de Investigación e Información Filosófica*. 2009;65(246 S. Esp):881-914.
64. Dreyfus HL. *What computers still can't do: A critique of artificial reason*: MIT press; 1992.
65. Tononi G. An information integration theory of consciousness. *BMC neuroscience*. 2004;5:1-22.
66. Searle J. *Freedom and Neurobiology: Reflections on Free Will, Language and Political Power*. New York: Columbia University Press; 2007.
67. Okon-Singer H, Hendler T, Pessoa L, Shackman AJ. The neurobiology of emotion–cognition interactions: fundamental questions and strategies for future research. *Frontiers in human neuroscience*. 2015;9:58.
68. Naudé W. *Artificial Intelligence and the Economics of Decision-Making*. 2023.
69. Latour B. *Reassembling the social: An introduction to actor-network-theory*: Oxford university press; 2005.
70. Waelbers K, Dorstewitz P. Ethics in actor networks, or: What Latour could Learn from Darwin and Dewey. *Science and engineering ethics*. 2014;20(1):23-40.
71. Ingold T. When ANT meets SPIDER: Social theory for arthropods. *Material agency: towards a non-anthropocentric approach*. 2008:209-15.
72. Sharifzadeh R. Technology, agency and decision. *Strategy for Culture*. 2016;9(34):115-36.