

ضرورت‌ها و چالش‌های حکمرانی هوش مصنوعی در سطح جهان

حمید محسنی^۱، کاظم فولادی قلعه^۲

^۱دکترای حکمت متعالیه از مؤسسه آموزشی و پژوهشی امام خمینی (ره)؛ دبیر کارگروه فلسفه سایبر آزمایشگاه پژوهشی فضای سایبر دانشگاه تهران، ایران
h.mohseni1297@mailfa.com

^۲استادیار گروه مهندسی کامپیوتر، دانشکده مهندسی دانشکدگان فارابی دانشگاه تهران؛ سرپرست آزمایشگاه پژوهشی فضای سایبر دانشگاه تهران، ایران
kfouladi@ut.ac.ir

چکیده

در عصر حاضر هوش مصنوعی در علوم، صنایع و جوانب مختلف زندگی انسان تأثیر گذاشته و در اداره دنیا و جوامع نقشه بسزایی خواهد داشت. به دلیل وجود برخی چالش‌های اساسی از ناحیه هوش مصنوعی در سطح دنیا، سازمان‌های بین‌المللی و مراکز علمی به مسئله حکمرانی هوش مصنوعی اهتمام داشته و در این حوزه بررسی می‌کنند. به‌طورکلی حکمرانی هوش مصنوعی به دنبال این است که میزان تولید و استفاده مفید از هوش مصنوعی را به نفع بشریت بالا ببرد. مسئله اصلی پژوهش حاضر این است که چالش‌ها و اهمیت حکمرانی هوش مصنوعی در سطح دنیا را با روش توصیفی-تحلیلی تبیین کند. در این مقاله بعد از تبیین مختصر حکمرانی هوش مصنوعی، چرایی حکمرانی بر انواع سیستم‌های هوش مصنوعی پیشرفته (قوی، عمومی و ابر هوش)، یادگیرنده، مولد، خودمختار و خود توسعه‌یابنده مورد تأکید قرار گرفته است و ضرورت حکمرانی هوش مصنوعی در سطح دنیا از نگاه اندیشمندان، مراکز علمی و برخی سازمان‌های بین‌المللی تبیین شده است. جمع‌بندی موارد ذکر شده خاطر نشان می‌کند که بنابر مبانی و ارزش‌های خاص در نظام اسلامی، چالش‌ها و اهمیت حکمرانی هوش مصنوعی در ایران اسلامی نیز مورد اهتمام قرار گیرد.

کلمات کلیدی: چالش‌های هوش مصنوعی، هوش مصنوعی خودمختار، نظام اسلامی ایران.

۱ مقدمه

با ورود فناوری‌های هوش مصنوعی در صنایع مختلف نگرانی‌هایی در مورد اخلاق هوش مصنوعی، شفافیت، و انطباق با سایر مقررات -مانند مقررات حفاظت از داده‌های عمومی- ایجاد شده است. بدون حکمرانی مناسب، سیستم‌های هوش مصنوعی می‌توانند خطراتی مانند تصمیم‌گیری مغرضانه، نقض حریم خصوصی و سوء استفاده از داده‌ها را به همراه داشته باشد. استفاده از هوش مصنوعی تقریباً در تمام بخش‌های صنعت از جمله مراقبت‌های بهداشتی، حمل‌ونقل، خدمات مالی، آموزش، ایمنی عمومی به سرعت در حال افزایش

بوده، همچنین سایر حوزه‌ها و علوم نیز از تأثیرات آن مصون نمانده است. در نتیجه اهمیت حکمرانی هوش مصنوعی نقش پررنگی پیدا کرده و بیش از گذشته موردتوجه قرار گرفته است. به همین دلیل در سطح دنیا با پیشرفت روزافزون هوش مصنوعی، اهمیت حکمرانی، سیاست‌گذاری و قانون‌گذاری در این حوزه بیش از پیش احساس می‌شود. سازمان‌های بین‌المللی، کشورها، دانشگاه‌ها، مؤسسات و مراکز علمی در سطح دنیا به دلیل وجود چالش‌ها و نگرانی‌های اساسی از ناحیه هوش مصنوعی در حوزه حکمرانی و قانون‌گذاری برای این موضوع نوین تحقیق و پژوهش می‌کنند. هدف از این تحقیق توجه دادن به چالش‌ها و تهدیدهای هوش مصنوعی و تأکید بر اهتمام مسئولان و دست‌اندرکاران بر تحقق حکمرانی آن در نظام اسلامی ایران است. از این رو پژوهش حاضر درصدد آن است که با روش توصیفی و تحلیلی چالش‌ها، نگرانی‌ها و اهمیت حکمرانی بر انواع سیستم‌های هوش مصنوعی پیشرفته، یادگیرنده، مولد، خودمختار و خود توسعه‌یابنده در سطح دنیا را مورد بررسی قرار دهد.

این مقاله در پنج قسمت سازمان‌دهی شده است. در قسمت دوم به پیشینه‌ی پژوهش در این زمینه می‌پردازیم. قسمت سوم به نگرانی‌ها از ناحیه‌ی هوش مصنوعی اختصاص دارد. در قسمت چهارم، اهمیت حکمرانی هوش مصنوعی در سطح دنیا مورد بررسی قرار می‌گیرد. در نهایت، قسمت پنجم به جمع‌بندی مطالب ذکر شده خواهد پرداخت.

۲ پیشینه‌ی پژوهش

پیش از ورود به بحث اصلی، پیشینه‌ی پژوهش مرتبط با مسئله حاضر مورد واکاوی قرار می‌دهیم. در این راستا، پژوهشی مستقل، جامع و کامل که نگرانی‌های متصور از ناحیه هوش مصنوعی و اهمیت حکمرانی آن را در ایران واکاوی کرده باشد، مشاهده نشد. ولی در برخی تحقیقات به بحث فناوری به‌طور عام و فضای مجازی و کمی مختصر درباره هوش مصنوعی پرداخته شده است. اما در سطح دنیا برخی قوانین، توصیه‌نامه‌ها، گزارش‌ها و پژوهش‌هایی در حوزه تبیین چالش‌ها و اهمیت حکمرانی هوش مصنوعی وجود دارد که به برخی از آنها در پژوهش حاضر اشاره خواهد شد. البته برخی کارها در زمینه ترجمه و نقد این آثار در ایران انجام شده است که می‌توانند به‌عنوان دریچه‌هایی از پیشینه تحقیق حاضر محسوب شوند.

اخیراً یونسکو سندی با عنوان «توصیه‌نامه اخلاق هوش مصنوعی» تصویب کرده است. در این سند بین‌المللی بیشتر، نگرانی‌های اخلاقی از ناحیه هوش مصنوعی مورد تأکید قرار گرفته و سعی شده است با تعیین بایدها و نبایدهای اخلاقی، راهکاری برای مدیریت این نوع تهدیدات در سطح دنیا ارائه شود (UN-ESCO, 2021). یونسکو به‌عنوان بازوی فرهنگی سازمان ملل سعی می‌کند با رهبری اخلاقی انسان‌محور، حکمرانی هوش مصنوعی در سطح دنیا را در دست بگیرد. اما این سند از منظر مبانی اسلامی دارای اشکالات اساسی هست - برای آشنایی با این اشکالات مراجعه شود به (محسنی و همکاران، ۱۴۰۲)؛ همچنین (محسنی و فولادی قلعه، ۱۴۰۱).

در گزارش (محسنی و همکاران، ۱۴۰۱) برخی از چالش‌های هستی‌شناختی، معرفت‌شناختی، انسان‌شناختی، ارزش‌شناختی، دین‌شناختی، اقتصادی، حقوقی و چالش‌های سیاسی و حکمرانی هوش

مصنوعی از منظر مبانی فلسفه اسلامی و آموزه‌های دینی تبیین شده است. هرچند به‌طور کلی در انتهای این گزارش علمی - که برای مرکز تحقیقات اسلامی مجلس شورای اسلامی انجام شده است -، برخی توصیه‌های سیاستی پیشنهاد شده، اما به‌طور مفصل و جدی به چالش‌ها و اهمیت حکمرانی هوش مصنوعی در سطح دنیا پرداخته نشده است.

گزارش دیگری با عنوان «درآمدی بر حکمرانی هوش مصنوعی» توسط مرکز پژوهش‌های مجلس شورای اسلامی تهیه شده است (مرکز پژوهش‌های مجلس، ۱۳۹۸) که هدف از آن آشنایی نمایندگان مجلس با مباحث حکمرانی هوش مصنوعی بوده است. این اثر در واقع گزارشی مختصر از گزارش مهم جمعی از پژوهشگران زیر نظر آلن دافو در مرکز حکمرانی هوش مصنوعی و مؤسسه آینده بشریت دانشگاه آکسفورد است که با عنوان «حکمرانی هوش مصنوعی: یک دستور جلسه تحقیقاتی» تنظیم شده است (Dafoe, 2018). این سند بر خطرات مهم «هوش مصنوعی پیشرفته»^۱ تمرکز دارد. در این سند خطرات این نوع هوش مصنوعی بیشتر از فرصت‌هایش مورد توجه قرار گرفته است و سعی کرده است با تبیین حکمرانی و ساختارهایی گذار به مرحله هوش مصنوعی پیشرفته را ترسیم کند.

افزون بر این، پژوهش‌های دیگری در سطح دنیا انجام شده است که در ادامه برخی از آنها تبیین خواهد شد. اما در ایران تاکنون پژوهش مستقل و جدی در این حوزه انجام نشده است. از این رو در این مقاله برخی چالش‌های مهم از ناحیه هوش مصنوعی تبیین شده و در ادامه با پرداختن به برخی پژوهش‌های انجام شده در مؤسسات و مراکز علمی در سطح دنیا، اهمیت این مسئله نیز روشن خواهد شد.

۳ نگرانی‌ها از ناحیه هوش مصنوعی

چرا می‌بایست از ناحیه هوش مصنوعی احساس نگرانی و خطر کرد؟ به‌طور کلی هوش مصنوعی می‌تواند از جهات بسیاری دگرگون‌کننده باشد. برخی از نمونه‌های متحول‌کننده به بخش‌های خاصی همچون امنیت سایبری، سلاح‌های خودمختار، اقتصاد، استراتژی، مهندسی و سایر زمینه‌های علم و فناوری مربوط می‌شود. با افزایش توانایی‌ها، هوش مصنوعی می‌تواند صنایع و نظم جهانی را متحول کند. به هوش مصنوعی که به توانایی متحول‌کننده دیگر نرسیده باشد، اغلب «هوش مصنوعی محدود»^۲ گفته می‌شود. اما برخی دیگر با توانایی‌های دگرگون‌کننده در هوش مصنوعی، فرا انسانی هستند (Dafoe, 2018: 19).

افزون بر این در قرن حاضر هوش مصنوعی ممکن است فراتر از انسان رفته و در انجام بسیاری از وظایف مهم بهتر از انسان‌های [معمولی] عمل کند. به هوش مصنوعی که به قابلیت‌های شناختی و رفتاری انسان برسد یا از حد توانایی‌های انسان بگذرد - از منظر برخی محققان و فلاسفه - «هوش مصنوعی قوی»^۳، و یا با تفسیری دیگر «ابر هوش مصنوعی»^۴ گفته می‌شود. «هوش مصنوعی عمومی»^۵ به نوعی از هوش

^۱ Advanced AI

^۲ Artificial Narrow Intelligence (ANI)

^۳ Strong Artificial Intelligence (SAI)

^۴ Artificial Superintelligence (ASI)

^۵ Artificial General Intelligence (AGI)

مصنوعی گفته می‌شود که برای حل هر نوع مسئله‌ای توانایی دارد و قابلیت‌های آن مانند ANI محدود به یک نوع مسئله‌ی خاص نیست. همچنین هوش مصنوعی انواع مختلفی همچون هوش مصنوعی مولد^۶، هوش مصنوعی یادگیرنده (یادگیری ماشین و ماشین‌های یادگیرنده)، سیستم‌های خودمختار^۷، خودتصمیم‌گیر و خودتوسعه‌یابنده و دیگر موارد را دربر می‌گیرد. به‌طور کلی پیشرفت هوش مصنوعی به سمتی در حرکت است که دارای ظرفیت‌هایی بیش از این خواهد بود؛ به این نوع هوش مصنوعی اصطلاحاً «هوش مصنوعی پیشرفته» گفته می‌شود.

بنابراین این فناوری برتر، گسترده و اثرگذار در سایر علوم و صنایع، فرصت‌ها و خطرات بزرگی برای بشریت به همراه خواهد داشت. برخی کارشناسان این احتمال را پیش‌بینی می‌کنند که در یک تا دو دهه آینده قابلیت‌هایی از هوش مصنوعی ظهور می‌کند که می‌تواند میزان تحول در رفاه، ثروت و قدرت را به انقلاب هسته‌ای یا حتی با انقلاب صنعتی و بالاتر از آن مقایسه کرد (Dafoe, 2018: 5). در همین راستا برخی پژوهشگران درباره زمان پیشی گرفتن هوش مصنوعی از انسان‌ها پژوهش‌هایی را انجام داده‌اند - به‌طور نمونه رجوع شود به: (Grace, Katja, et al., 2017). هرچند از منظر برخی پژوهشگران، هوش مصنوعی تنها جنبه فنی و مهندسی دارد و به همین دلیل با این نوع فناوری تنها مواجهه ابزاری دارند. در این رویکرد بیشتر فواید و منافع ابزاری هوش مصنوعی در نظر گرفته می‌شود (ویتابی، ۱۴۰۱). اما از منظر دیگر متفکران به‌زودی هوش مصنوعی همه زندگی و تلاش‌های بشری را تحت تأثیر خود قرار خواهد داد. تحولات دوران‌ساز ناشی از هوش مصنوعی در جامعه، اقتصاد، سیاست، فرهنگ، علوم انسانی و حکمرانی اثرات بسیاری خواهد گذاشت (کیسینجر و همکاران، ۱۴۰۱). به‌طور کلی همین هوش مصنوعی محدود نیز در زندگی بسیاری از انسان‌ها تأثیرگذار است (Russell, Norvig, 2020). بنابراین با پیشرفت روزافزون هوش مصنوعی به‌زودی همه ابعاد زندگی در جوامع بشری تحت الشعاع این فناوری نوین قرار خواهد گرفت.

از این رو، دلیل تمرکز برخی پژوهشگران غربی بر حکمرانی هوش مصنوعی، وجود نگرانی‌هایی در حوزه‌های مرتبط با عدالت، کیفیت داده‌ها و خودمختاری^۸ (استقلال) می‌باشد^۹. افزون بر این، مسئله دیگری که موضوع حکمرانی هوش مصنوعی را ضروری می‌کند، پیشرفت سریع در «یادگیری ماشین»^{۱۰} است. امروزه یادگیری ماشین قلب فناوری هوش مصنوعی به شمار می‌آید و منشأ نقاط قوت و ضعف این فناوری می‌باشد. به‌طور نمونه، یکی از مشکل‌های شناخته‌شده که در سیستم‌های یادگیری وجود دارد این است که می‌تواند موجب شود رفتاری ناخواسته و یا مضر در دنیای واقعی از ماشین سر بزند (Amodei et al., 2016: 1). به‌همین دلیل برخی راهبردان‌دانشان غربی همچون کیسینجر از این ناحیه احساس نگرانی کرده‌اند. از منظر آنها به‌دلیل یادگیری ماشینی و شکل‌گیری منطق و شناختی جدید در ماشین، نگرانی‌ها و خطراتی نظیر گمراهی انسان، ارائه اطلاعات نادرست و رفتارهای ناعادلانه وجود دارد (کیسینجر و همکاران، ۱۴۰۱: ۲۲-۲۳).

⁶Generative AI

⁷Autonomous Systems

⁸Autonomy

⁹<https://www.techtarget.com/searchenterpriseai/definition/AI-governance>

¹⁰Machine Learning

سیستم‌های هوش مصنوعی و الگوریتم‌هایی که از عملیات آنها پشتیبانی می‌کنند می‌توانند در تصمیم‌گیری‌های مهم تأثیر بگذارند. به‌طور نمونه در تصمیمات بالینی در تشخیص پزشکی، در تصمیم‌های پلیسی، در شناخت مجرمان، و همچنین با شخصی‌سازی و فیلتر کردن محتوا می‌توانند در رفتار کاربران تأثیر بگذارند. همچنین، عدم قطعیت در تعاملات انسان-ماشین، می‌تواند سیستم‌های هوش مصنوعی را به رفتارهای غیرمنتظره سوق دهد. این فرایند می‌تواند خطراتی را برای کاربرانش ایجاد کند. افزون بر این، در بسیاری از سیستم‌های هوش مصنوعی سوگیری‌ها در داده‌ها و الگوریتم‌ها، نتایج تبعیض‌آمیز و غیراخلاقی را برای افراد به همراه دارد. مواردی همچون امتیازدهی اعتبار و مجازات کیفری از این قبیل هستند. همچنین ماهیت خودمختاری سیستم‌های هوش مصنوعی منجر به مسائل و چالش‌های بالقوه ناشی از مستقل بودن کنترل ماشین از انسان در تصمیم‌گیری را در پی دارد. از آنجا که تصمیمات یادگیری ماشین (ML) به‌شدت مبتنی بر داده‌ها و غیرقابل‌پیش‌بینی هستند، سیستم می‌تواند رفتارهای بسیار متفاوتی را در پاسخ به ورودی‌های تقریباً یکسان نشان دهد که این تعیین رفتارهای «صحیح» و تأیید ایمنی آنها پیش از عمل را دشوار می‌کند. بنابراین از یک سو استقلال تصمیم‌گیری هوش مصنوعی، به‌طور قابل‌توجهی کنترل انسان بر تصمیمات خود را کاهش می‌دهد؛ و از سوی دیگر از آنجا که هوش مصنوعی که عمدتاً به فرآیندهای ML برای یادگیری و تطبیق قوانین خود متکی است، انسان‌ها دیگر کنترلی برای هدایت این نوع سیستم‌ها در دست ندارند (Taeihagh, 2021: 1-5).

همچنین «هوش مصنوعی متحول‌کننده»^{۱۱} که به‌عنوان هوش مصنوعی پیشرفته شناخته می‌شود، می‌تواند تغییرات اساسی را در ثروت، سلامت و قدرت ایجاد کرده و این زمینه‌ها را متحول کند. این امر می‌تواند با تحول در کار از طریق تغییر در چشم‌انداز اشتغال و امنیت شغلی باشد. هوش مصنوعی متحول‌کننده می‌تواند نابرابری اقتصادی در داخل و در سطح جهانی را افزایش دهد. افزون بر این می‌تواند ابزارهای سرکوب و کنترل دولتی و قدرت بخشیدن به دولت‌های مستبد را فراهم آورد. هوش مصنوعی پیشرفته می‌تواند بر شکل‌های جدیدی از پاسخگویی و تصمیم‌گیری دموکراتیک تأثیر بگذارد. افزون بر این هوش مصنوعی می‌تواند امنیت بین‌المللی را به‌وسیله شاخصه‌های راهبردی و کلیدی متحول کند. به‌طور نمونه این فناوری می‌تواند امنیت تلافی‌جویانه هسته‌ای، توازن تهاجمی-دفاعی، حفظ حریم خصوصی و پیش‌بینی توازن قوا در آینده را متحول کند. همچنین هوش مصنوعی می‌تواند عملیات و راهبردهای جدیدی را فعال کند. به‌طور نمونه هوش مصنوعی می‌تواند اقتناع عمومی، عملیات سایبری، فرماندهی و کنترل، نبرد هوایی و زیردریایی، مهندسی و علم را متحول کند. وجود این مزایا و حتی صرف تصور آنها می‌تواند به‌قدر کافی منجر به تغییر اساسی قدرت در دولت‌ها شده و فرایند فناوری و نظم جهانی را متحول کند (Dafoe, 2018: 9-10).

همچنین از منظر بلندمدت و در رویکرد انتزاعی‌تر، ظهور ماشین «ابر هوش» (نوعی هوش مصنوعی که در همه کارهای مهم بسیار بهتر از انسان شناخته شده و عمل می‌کند) دارای اهمیت است. هرچند ابر هوش مصنوعی می‌تواند فرصت‌های فوق‌العاده‌ای برای مواردی همچون کاهش ریشه‌ای بیماری‌ها، فقر، درگیری بین فردی و مخاطرات فاجعه‌باری همانند تغییرات آب و هوایی ارائه دهد. با این حال ابر هوش و به‌طور کلی

¹¹ Transformative AI (TAI)

هوش مصنوعی پیشرفته نیز ممکن است فاجعه‌هایی ایجاد کند. از جمله نابرابری شدید، استبداد جهانی و بی‌ثباتی که منجر به جرقه‌های درگیری هسته‌ای در سطح جهان شود (Dafoe, 2018: 9-10). در نتیجه هوش مصنوعی در عدالت، امنیت (اعم از امنیت جانی، مالی و نظامی)، سیاست، حکمرانی و صنایع تأثیر می‌گذارد.

خلاصه اینکه در بخش حاضر تلاش شده است برخی نگرانی‌های مهم از انواع سیستم‌های هوش مصنوعی پیشرفته، یادگیرنده، مولد، خودمختار و خود توسعه‌یابنده مورد تأکید گیرد. در این نوع فناوری‌ها به‌نوعی خود ماشین به‌طور مستقل تشخیص داده و عمل می‌کند. از این رو ممکن است به‌دلیل برخی عوامل همچون سوگیری‌های الگوریتمی، تحلیل ناقص داده‌ها، عدم قطعیت در تصمیم‌گیری صحیح، قابل پیش‌بینی نبودن، عدم شفافیت یا ابهام ذاتی، و یادگیری، تصمیم‌ها و رفتارهای آسیب‌زا و زیانباری از سیستم سر بزند. همچنین این داده‌کاوی‌ها، استنتاج‌ها و تصمیم‌های ماشینی می‌توانند بر شناخت، تصمیم و رفتار انسان نیز تأثیرگذار باشد؛ زیرا انسان‌ها در بسیاری از علوم، صنایع و حرفه‌ها با اعتماد بر اطلاعات و استنتاج به‌دست‌آمده از هوش مصنوعی عمل می‌کنند. از این‌رو تصمیم‌های غلط ماشین بر جهت‌گیری و رفتار ناصحیح انسان نیز اثرگذار است. به همین دلیل حکمرانی و برنامه‌ریزی برای کاهش آسیب‌ها و افزایش منافع و بهره‌وری در این حوزه‌ها امری ضروری و اجتناب‌ناپذیر می‌نماید.

۴ اهمیت حکمرانی هوش مصنوعی در سطح دنیا

از یک‌سو تحولات هوش مصنوعی ممکن است از ناحیه برخی نادیده گرفته شود؛ ولی از سوی دیگر در رویکرد واقع‌بینانه اثرات آن جهان‌شمول بوده و اثرات بیرونی و بین‌نسلی داشته و یکی از مهم‌ترین مسائل جهانی در قرن بیست و یکم به‌حساب می‌آید. از این رو برای مواجهه صحیح با این فناوری دگرگون‌کننده، نیازمند پژوهش درباره «حکمرانی هوش مصنوعی»^{۱۲} هستیم. حکمرانی هوش مصنوعی در این زمینه مطالعه می‌کند که چگونه انسان‌ها می‌توانند به بهترین نحو به سمت «هوش مصنوعی پیشرفته» با تمرکز بر ابعاد سیاسی، اقتصادی، نظامی، حکمرانی و اخلاقی حرکت کنند. حکمرانی هوش مصنوعی بر مؤسسات و بسترهایی که هوش مصنوعی در آن ساخته یا استفاده می‌شود تمرکز دارد. به‌طور خاص، حکمرانی هوش مصنوعی به دنبال این است که میزان تولید و استفاده مفید افراد از هوش مصنوعی پیشرفته را به‌نفع بشریت رقم بزند (Dafoe, 2018: 5-6).

به‌طور کلی برخی از مسائل و چالش‌هایی که حکمرانی هوش مصنوعی به‌دنبال رفع آن است به حوزه‌های ذیل مربوط می‌شود:

- ارزیابی ایمنی هوش مصنوعی؛
- تعیین بخش‌هایی که برای هوش مصنوعی خودمختار (مستقل) مناسب هستند؛
- ایجاد ساختارهای قانونی پیرامون استفاده از فناوری هوش مصنوعی؛

¹² AI Governance

- تعریف قوانینی پیرامون کنترل و دسترسی به داده‌های شخصی؛
- پرداختن به سؤالات اخلاقی مرتبط با هوش مصنوعی.

برخی رهبران فناوری و کارشناسان هوش مصنوعی همانند ایلان ماسک، استیو ورنیاک و دیگران به دلیل برخی نگرانی‌ها - در سال ۲۰۲۳- با امضای نامه‌ای سرگشاده خواستار توقف موقت تحقیقات هوش مصنوعی شدند تا زمینه تدوین مقررات و قوانین فراهم شود. در همین سال سام آلتمن مدیرعامل OpenAI در کنگره آمریکا حاضر شد و خواستار نظارت بر هوش مصنوعی شد.^{۱۳}

افزون بر این در سال ۲۰۱۸ بیست‌هزار نفر پژوهشگر نامه‌ای را امضا کردند و به سیاست‌گذاران در مورد خطرات سلاح‌های خودمختار - که ممکن است از الگوریتم‌های هوش مصنوعی برای حمله به اهداف خاص استفاده کنند- هشدار دادند (Wang, Siau, 2018: 3). از این رو در کشورهای مختلف دنیا و از منظر سازمان‌های بین‌المللی، حکمرانی و قانون‌گذاری هوش مصنوعی از مسئله پیشرفت صرف و بی‌قیدوبند آن مهم‌تر است.

برخی پژوهشگران اشاره می‌کنند که هرچند انواعی از حکمرانی در سطح ملی و منطقه‌ای کم‌وبیش مشاهده می‌شود ولی آنها سعی می‌کنند که ضرورت وجود یک حکمرانی جهانی برای هوش مصنوعی را برجسته کنند (Tallberg et al., 2023). به دلیل اینکه در دهه اخیر هوش مصنوعی به موضوعی فراگیر تبدیل شده است، سازمان‌های بین‌دولتی، ملی و منطقه برای پیشرفت سیاست‌ها و راهبردها در حوزه حکمرانی هوش مصنوعی در حال تلاش هستند. نقش‌آفرینان این حوزه دریافته‌اند که برای رسیدگی به نگرانی‌های اخلاقی و اجتماعی پیرامون هوش مصنوعی می‌بایست شیوه‌ها و راه‌حل‌هایی را پیدا کنند. پس، سیاست‌گذاری و حکمرانی برای به حداکثر رساندن مزایای هوش مصنوعی امری ضروری است. برای همین مدیریت فعال و آگاهانه فناوری‌های هوش مصنوعی به اولویت بسیاری از دولت‌ها در سراسر دنیا تبدیل شده است. تجزیه و تحلیل شاخص‌های هوش مصنوعی از سوابق قانونی در ۲۵ کشور دنیا نشان می‌دهد که تعداد لوایح حاوی هوش مصنوعی که به قانون تبدیل شده است از تنها یک مورد در سال ۲۰۱۶ به ۱۸ مورد در سال ۲۰۲۱ افزایش یافته است. اسپانیا، بریتانیا و ایالات متحده هر کدام با تصویب سه لایحه در سال ۲۰۲۱ بیشترین آمار تصویب قوانین هوش مصنوعی را به خود اختصاص داده بودند (Clark, Perrault, 2022: 174-175).

سازمان‌های دولتی ایالات متحده که در زمینه حکمرانی هوش مصنوعی کار می‌کنند، از جمله دفتر ابتکار هوش مصنوعی ملی دفتر سیاست علم و فناوری کاخ سفید است که در سال ۲۰۲۱ راه‌اندازی شده است. همچنین کمیته مشاوره ملی هوش مصنوعی در سال ۲۰۲۲ به‌عنوان بخشی از دفتر ابتکار ملی هوش مصنوعی، برای مشاوره به رئیس‌جمهور درباره مسائل هوش مصنوعی ایجاد شده است.^{۱۴} دفتر ابتکار هوش مصنوعی ملی دفتر سیاست علم و تکنولوژی کاخ سفید یک چارچوب حکمرانی هوش مصنوعی در زمینه‌های مختلف ایجاد کرده است. مأموریت ابتکار ملی هوش مصنوعی، تضمین رهبری مستمر ایالات متحده در پژوهش

¹³<https://www.techtarget.com/searchenterpriseai/definition/AI-governance>

¹⁴<https://www.techtarget.com/searchenterpriseai/definition/AI-governance>

و توسعه است. همچنین هدف دولت آمریکا رهبری جهان در توسعه و استفاده از هوش مصنوعی قابل اعتماد در بخش عمومی و خصوصی و در تمام بخش‌های اقتصادی و جامعه است.^{۱۵}

افزون براین، در حال حاضر ابتکارات زیادی در زمینه اخلاقی، اجتماعی و جنبه‌های حقوقی هوش مصنوعی در سراسر مراکز دانشگاهی و تحقیقاتی جهان در حال انجام است. از جمله مراکز اختصاصی همچون مرکز Leverhulme مرکز آینده اطلاعات در کمبریج است. مأموریت مرکز اطلاعات آینده لوور هولم ایجاد یک جامعه بین‌رشته‌ای جدید از محققان است که ارتباط‌های قوی بین فناوران و سیاست‌گذاران داشته باشند. همچنین هدف اصلی آن‌ها این است که بهترین فرصت‌های هوش مصنوعی در طول قراردادهای آینده را فراهم آورند.^{۱۶} مطالعه صدساله هوش مصنوعی (AI100) در دانشگاه استنفورد پژوهش بلندمدت دیگری است که در پاییز ۲۰۱۴ راه‌اندازی شد. این یک پژوهش طولانی‌مدت در زمینه هوش مصنوعی و تأثیر آن بر مردم، و جوامع انسانی است. همچنین این مطالعات چالش‌های و فرصت‌های فنی-اجتماعی که این پیشرفت در عرصه‌هایی مانند اخلاق، اقتصاد با طراحی سیستم سازگار با شناخت انسان ایجاد می‌کند را بررسی می‌کنند (Stone, et al., 2016). هدف آنها مطالعه و پیش‌بینی چگونگی تأثیرات هوش مصنوعی در تمام جنبه‌های نحوه کار، زندگی و بازی افراد است.^{۱۷} این پروژه‌ها نشان می‌دهند پژوهشگران به مشکلات، تهدیدات و خطرات بالقوه، احتمالی، بلکه چالش‌های بالفعل و موجود هوش مصنوعی واقف بوده و در این زمینه تحقیق و پژوهش می‌کنند.

همچنین پژوهش‌هایی تحت عنوان «هوش مصنوعی باز» (OpenAI) سعی دارند کدهای باز را تعریف کنند. آنها براین باور هستند که این راه بهترین شیوه برای مبارزه با خطرات مخرب هوش مصنوعی است.^{۱۸} محققان هوش ماشینی مؤسسه MIRI هدف خود را «همسوسازی هوش مصنوعی پیشرفته با منافع انسانی» عنوان می‌کند.^{۱۹} شرکت‌های بزرگ و کوچک‌تر نیز ابتکاراتی دارند، همانند مشارکت در هوش مصنوعی به نفع مردم و جامعه، با همکاری آمازون، دیپ مایند^{۲۰}، فیس‌بوک، گوگل، آی بی ام و مایکروسافت^{۲۱}. توسط افراد حرفه‌ای نیز پژوهش‌هایی در حال انجام است، همانند تحقیقاتی که سوگیری در هوش مصنوعی را بررسی می‌کند (Boddington, 2017: 3). همچنین سازمان‌هایی همچون مغز گوگل^{۲۲}، مرکز هوش مصنوعی سازگار با انسان^{۲۳}، مؤسسه تحقیقاتی هوش ماشینی^{۲۴} در دانشگاه پرکلی کالیفرنیا^{۲۵} در حوزه ایمنی هوش مصنوعی کار می‌کنند (Dafoe, 2018: 33).

استنفورد مؤسسه هوش مصنوعی انسان‌محور را در سال ۲۰۱۹ تأسیس کرده است، که هدف اصلی آن

¹⁵<https://www.ai.gov>

¹⁶<http://lcfi.ac.uk>

¹⁷<https://ai100.stanford.edu>

¹⁸<http://open.ai>

¹⁹<https://intelligence.org>

²⁰DeepMind

²¹<https://www.partnershiponai.org>

²²Google Brain

²³The Center for Human Compatible AI

²⁴The Machine Intelligence Research Institute

²⁵UC Berkeley

پیشبرد پژوهش، آموزش و سیاست‌گذاری هوش مصنوعی برای بهبود شرایط انسانی است و اخیراً یک گزارش پژوهشی مفصل درباره شاخص هوش مصنوعی در سال ۲۰۲۳ منتشر کرده است.^{۲۶} مؤسسه آینده زندگی^{۲۷} نیز در سال ۲۰۱۵ تأسیس شد و مأموریت آن هدایت فناوری تحول‌آفرین به سمت سودمندی زندگی و دوری از خطرات شدید در مقیاس بزرگ است.^{۲۸} رصدخانه هوش مصنوعی OECD.AI^{۲۹} خود را یک مرکز جهانی برای سیاست‌گذاری هوش مصنوعی معرفی کرده که با تبیین اصولی سعی دارد تا به کشورها کمک کند خطرات هوش مصنوعی را مدیریت کرده و از منافع آن بهره‌مند شوند.^{۳۰} این رصدخانه بیش از ۷۰۰ ابتکار سیاست ملی هوش مصنوعی را از ۶۰ کشور و منطقه ثبت کرده است (Tallberg, et al., 2023: 7-8). همان‌طور که تبیین شد، در برخی مراکز علمی دنیا به‌طور گسترده در زمینه هدایت فناوری هوش مصنوعی تحقیق و پژوهش می‌شود. از یک‌سو امروزه مسئله فناوری‌های نوین هوش مصنوعی، همه‌ی جنبه‌های زندگی جوامع بشری را تحت تأثیر خود قرار داده است. از سوی دیگر تهدیدات هوش مصنوعی امری واقعی بوده، به همین دلیل حکمرانی هوش مصنوعی امری اجتناب‌ناپذیر و ضروری است (Wang, Siau, 2018: 3).

البته در سطح دنیا با مبانی، مسائل، رویکردها و دغدغه‌های خاص خودشان به دنبال چالش‌ها و حکمرانی هوش مصنوعی هستند؛ اما از یک‌سو نظام جمهوری اسلامی ایران، مبانی، مسائل و رویکردهای خاص خود را دارد و از سوی دیگر به دلیل اهمیت هوش مصنوعی و تأثیر آن در زمینه‌های مختلف همچون فرهنگ، دین، اخلاق، اقتصاد، امنیت، حکمرانی و سیاست، هدف‌گذاری، برنامه‌ریزی و حکمرانی ویژه خود را می‌طلبد. از این‌رو ضروری است این مسئله در نظام جمهوری اسلامی ایران که مبانی، ارزش‌ها، سیاست‌ها و حکمرانی مستقل و خاص خود را دارد، نیز مورد توجه و پژوهش قرار گیرد (برای آشنایی بیشتر با برخی مبانی به (محسنی و فولادی قلعه، ۱۴۰۲) مراجعه شود).

مقام معظم رهبری اخیراً بر پیشرفت در حوزه هوش مصنوعی تأکید کرده‌اند: «پیشنهاد می‌کنم یکی از مسائلی که مورد تکیه و توجه و تعمیق واقع می‌شود، مسئله هوش مصنوعی باشد که در اداره آینده دنیا نقش خواهد داشت»^{۳۱} تنها با دقت و تأمل در همین عبارت ایشان مشخص می‌شود که مسئله «اداره» و «حکمرانی» آینده دنیا از طریق هوش مصنوعی بسیار مهم است. وقتی از منظر معظم‌له حکمرانی هوشمند مسئله فراگیر و مهم آینده دنیا هست، به تبع آن مشخص شدن تهدیدات و چالش‌های پیش رو، باید و نبایدها، هدایت و جهت‌گیری این فناوری و در یک بیان «حکمرانی هوش مصنوعی» نیز به «طریق اولی» مهم بوده و ضرورت دارد. افزون بر این معظم‌له به‌تازگی مسئله مهم توجه به لایه‌های عمیق هوش مصنوعی را مطرح کرده و در ضمن آن مسئله مهم دیگری را بیان کردند که به حوزه حکمرانی هوش مصنوعی مربوط می‌شود. ایشان در یک رویکرد آینده‌پژوهانه زمینه‌سازی نظام سلطه برای حکمرانی هوش مصنوعی در آینده را با تأسیس آژانس

²⁶<https://hai.stanford.edu>

²⁷Future of Life Institute

²⁸<https://futureoflife.org/resource/ai-policy>

²⁹OECD AI Policy Observatory

³⁰<https://oecd.ai/en/wonk>

³¹بیانات، دیدار جمعی از نخبگان و استعدادهای برتر علمی کشور، ۱۴۰۰/۸/۲۶

هوش مصنوعی برجسته نمودند.^{۳۲} افزون بر این، این مسئله را در ذیل بحث حکمرانی قانونمند فضای مجازی مطرح کرده‌اند.

از این رو بنا بر اهمیت حکمرانی و ریل‌گذاری هوش مصنوعی در سطح دنیا و تهدیدات جدی از ناحیه هوش مصنوعی، و ضرورت داشتن حکمرانی مستقل می‌بایست در نظام ایران اسلامی نیز در این زمینه تحقیقات و پژوهش‌های مستقل و چشم‌گیری انجام شود.

۵ جمع‌بندی

در پژوهش حاضر تلاش شد اهمیت حکمرانی بر انواع سیستم‌های هوش مصنوعی پیشرفته (قوی، عمومی و ابر هوش)، یادگیرنده، مولد، خودمختار و خود توسعه‌یابنده در سطح دنیا تبیین شود. در این نوع فناوری‌ها به‌نوعی خود ماشین به‌طور مستقل تشخیص داده و عمل می‌کند. از این رو ممکن است به دلیل برخی عوامل همچون سوگیری‌های الگوریتمی، تحلیل ناقص داده‌ها، عدم قطعیت در تصمیم‌گیری صحیح، قابل پیش‌بینی نبودن، کدورت یا ابهام ذاتی، و یادگیری، تصمیم‌ها و رفتارهای آسیب‌زا و زیانباری از سیستم سر بزند. همچنین داده‌کاوی‌ها، استنتاج‌ها و تصمیم‌های ماشینی می‌توانند بر شناخت، تصمیم و رفتار انسان نیز تأثیرگذار باشد؛ زیرا انسان‌ها در بسیاری از علوم، صنایع و حرفه‌ها با اعتماد بر اطلاعات و استنتاج به‌دست‌آمده از هوش مصنوعی عمل می‌کنند. از این رو تصمیم‌های غلط ماشین بر جهت‌گیری و رفتار ناصحیح انسان نیز اثرگذار است. به بیان دیگر، به دلیل اینکه هوش مصنوعی در جنبه‌های مختلفی همچون در عدالت، امنیت، سیاست، حکمرانی، صنایع و سایر علوم تأثیر می‌گذارد؛ و یا حتی ممکن است مورد هک و نفوذ قرار گیرد؛ حکمرانی و برنامه‌ریزی برای کاهش نگرانی‌ها و آسیب‌ها و افزایش منافع، ایمنی و بهره‌وری در سطح دنیا اهمیت دارد. بدین‌سان اگر می‌خواهیم در جامعه اسلامی پیشرفت تکنولوژی هوش مصنوعی در راستای نظام اسلامی باشد، می‌بایست چالش‌های آن تبیین شده و این فناوری مبتنی بر مبانی نظام اسلامی حکمرانی و ریل‌گذاری شود. این امر زمانی محقق خواهد شد که برخی چالش‌ها و مسائل هوش مصنوعی و مواجهه و اهتمام غربی‌ها و سازمان‌ها و مراکز علمی مهم دنیا با این مسئله مشخص شود. در این صورت -دست‌کم- با توجه به اهمیت دادن کشورهای دنیا به مسئله حکمرانی در حوزه هوش مصنوعی، ضرورت حکمرانی آن در نظام ایران اسلامی نیز نمایان خواهد گشت.

سپاسگزاری

از مؤسسه شناخت (قم) بابت پاره‌ای از حمایت‌ها در فرایند گردآوری مطالب پژوهش حاضر تشکر و قدردانی می‌کنیم.

^{۳۲}بیانات در نخستین دیدار هیئت دولت چهاردهم، ۱۴۰۳/۶/۶

مراجع

- [۱] کیسینجر، هنری، و همکاران (۱۴۰۱). عصر هوش مصنوعی و آینده بشریت، ترجمه: رحمن قهرمان پور، تهران، روزنه.
- [۲] محسنی، حمید، و فولادی قلعه، کاظم (آبان ۱۴۰۱). «سازگاری یا ناسازگاری توصیه‌نامه اخلاق هوش مصنوعی یونسکو با نظریه اخلاقی اسلامی»، در مجموعه مقالات نخستین کنفرانس فضای سایبر، دانشگاه تهران، صص ۱۳۳-۱۴۰.
- [۳] محسنی، حمید، و فولادی قلعه، کاظم (آبان ۱۴۰۲). «پیشرفت مطلق یا هدفمند فناوری هوش مصنوعی در اندیشه مقام معظم رهبری»، در مجموعه مقالات دومین کنفرانس فضای سایبر، دانشگاه تهران، صص ۶۶۵-۶۷۵.
- [۴] محسنی، حمید، و همکاران (۱۴۰۲). «بررسی توصیه‌نامه اخلاقی هوش مصنوعی یونسکو و تبیین نقاط و ضعف و قوت آن»، در ک: نقد و تکمیل پیش‌نویس نهایی توصیه‌نامه درباره اخلاق هوش مصنوعی، دفتر فقه معاصر حوزه علمیه قم، صص ۱۸۳-۲۱۷.
- [۵] محسنی، حمید، و همکاران (زمستان ۱۴۰۱). «چالش‌های بنیادین در مسیر سیاست‌گذاری و قانون‌گذاری هوش مصنوعی»، گزارش تحقیقاتی، قم، مرکز تحقیقات اسلامی مجلس شورای اسلامی.
- [۶] مرکز پژوهش‌های مجلس (۱۳۹۸). درآمدی بر حکمرانی هوش مصنوعی؛ خلاصه راهبردی از: Allan Dafoe, AI Governance: A research agenda, Oxford university, 2018, مرکز پژوهش‌های مجلس شورای اسلامی.
- [۷] ویتبای، بلی (۱۴۰۱). هوش مصنوعی به زبان ساده، ترجمه حسین مجدفر و آوا بهرامی، چ ۳، تهران، سبزان.
- [8] Amodei, Dario, et al. (2016). "Concrete Problems in AI Safety". ArXiv:1606.06565 [Cs], June 21.
- [9] Boddington, Paula (2017). Towards a Code of Ethics for Artificial Intelligence, Springer International Publishing.
- [10] Clark, Jack and Perrault, Ray (2022). Artificial Intelligence Index Report 2022, HAI, Stanford University, <https://aiindex.stanford.edu/ai-index-report-2022>
- [11] Dafoe, Allan (2018). AI Governance: A research agenda, Oxford university.
- [12] Grace, Katja, et al. (2017). "When Will AI Exceed Human Performance? Evidence from AI Experts". ArXiv:1705.08807 [Cs], May 24.
- [13] Jonas Tallberg et al. (2023). The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research, International Studies Review, Volume 25, Issue 3, September 2023, <https://doi.org/10.1093/isr/viad040>
- [14] Russell, S., Norvig, P. (2020). Artificial Intelligence: A Modern Approach, 4th Ed., Prentice Hall.
- [15] Taeiagh, Araz (2021). Governance of artificial intelligence, Policy and Society, Volume 40, 2021 - Issue 2: Governance of AI and Robotics, pp. 137-157, <https://doi.org/10.1080/14494035.2021.1928377>
- [16] Stone, Peter, et al. (2016). Artificial Intelligence and Life in 2030: The One Hundred Year Study on Artificial Intelligence, Report of the 2015 study panel.

- [17] UNESCO (2021). Recommendation on the ethics of artificial intelligence, URL: <https://en.unesco.org/artificial-intelligence/ethics>, retrieved: 2022 Feb. 6.
- [18] Wang, Weiyu, Siau, Keng (2018). "Artificial Intelligence: A Study on Governance, Policies, and Regulations", MWAIS 2018 Proceedings, <http://aisel.aisnet.org/mwais2018/40>
- [19] <https://www.techtarget.com/searchenterpriseai/definition/AI-governance>
- [20] <https://www.oxfordinsights.com/government-ai-readiness-index-2020>
- [21] www.khamenei.ir