

The Path of Artificial Intelligence Governance Transformation: From Policymaking to Accountability



Ali Mirarab¹ 

Received: 2025/01/13 • Revised: 2025/02/20 • Accepted: 2025/03/16 • Published online: 2025/05/04

Abstract

Artificial Intelligence (AI), as a revolutionary, transformative, and disruptive technology, is changing the world. AI governance and its integration into public policy are important topics in contemporary discussions. The potential use of AI in the public policy cycle provides opportunities to improve decision-making processes for governments. Critical analysis of AI technologies in public policymaking is essential to understand their impact on decision-making processes and social inequalities. Although government agencies, businesses, international organizations, and other stakeholders have made efforts in the field of AI, the integration of AI and governance is still in the early stages of development. This research, using a descriptive-analytical method, aims to study the trends and frameworks of AI governance and public policymaking, and examines how AI can improve policymaking processes and outcomes. AI can assist policymakers in identifying needs, developing programs, forecasting outcomes, and analyzing policy effectiveness.

1. Assistant Professor, Islamic Sciences and Culture Academy, Qom, Iran. alimirarab@isca.ac.ir.

* Mirarab, A. (2025). The Path of Artificial Intelligence Governance Transformation: From Policymaking to Accountability. *Journal of Islamic Governance Studies*, 1(1), pp. 456-491.
<https://doi.org/10.22081/jislamicgo.2025.70830.1000>

* **Publisher:** Islamic Propagation Office of the Seminary of Qom (Islamic Sciences and Culture Academy, Qom, Iran). ***Type of article:** Research Article

© The Authors



Studies show that policy documents around the world highlight the potential of AI to help achieve sustainable development goals, such as treating chronic diseases, reducing road fatalities, combating climate change, and predicting cybersecurity threats. AI not only offers opportunities but also brings risks and challenges. Some concerns include job displacement, algorithmic bias, data misuse, and security threats. AI policies must deal with these challenges and ensure that AI is developed and used responsibly and ethically.

Keywords

Artificial Intelligence, Public Policy, Governance, Ethics, Global Competition.

من صنع السياسات إلى تحمّل المسؤولية: مسار تحوّل حوكمة الذكاء الاصطناعي

على مير عرب^١ 



تاريخ الاستلام: ٢٠٢٥/٠١/١٣ • تاريخ التعديل: ٢٠٢٥/٠٢/٢٠ • تاريخ القبول: ٢٠٢٥/٠٣/١٦ • تاريخ النشر الإلكتروني: ٢٠٢٥/٠٥/٠٤

الملخص

الذكاء الاصطناعي (AI)، بوصفه تقنية ثورية وتحويلية ومزعزعة، يُحدث تحولاً جذرياً في العالم. إنّ حوكمة الذكاء الاصطناعي ودمجه في السياسات العامة تُعدّ من المواضيع البارزة في النقاشات المعاصرة. يوفّر الاستخدام المحتمل للذكاء الاصطناعي في دورة السياسة العامة فرصاً لتحسين عمليات اتخاذ القرارات الحكومية. إنّ التحليل النقدي لتقنيات الذكاء الاصطناعي في صنع السياسات العامة يُعدّ أمراً ضرورياً لفهم تأثيرها على عمليات اتخاذ القرار وعلى أوجه عدم المساواة الاجتماعية (التفاوتات الاجتماعية). على الرغم من أنّ الإدارات الحكومية والمؤسسات الاقتصادية والمنظمات الدولية والأطراف الأخرى قد بذلت جهوداً في مجال الذكاء الاصطناعي، إلّا أنّ دمج الذكاء الاصطناعي والحوكمة لا يزال في مراحله الأولى من التطور. تسعى هذه الدراسة بالاعتماد على المنهج الوصفي-التحليلي، وبهدف دراسة الاتجاهات والأطر الخاصة (المتعلقة) بحوكمة الذكاء الاصطناعي وصنع السياسات العامة، إلى تحليل كيفية

٤٥٨
 دراسات
 حكر في اسلامي

سال اول، شماره اول (پیاپی) بهار و تابستان ١٤٠٣

١. أستاذ مساعد في المعهد العالي للعلوم والثقافة الإسلامية، قم، إيران. alimirarab@isca.ac.ir

* مير عرب، على. (٢٠٢٥). من صنع السياسات إلى تحمّل المسؤولية: مسار تحوّل حوكمة الذكاء الاصطناعي. دراسات الحوكمة الإسلامية، ١(١)، صص ٤٥٦-٤٩١.

<https://doi.org/10.22081/jislamicgo.2025.70830.1000>

□ التصنيف: علمية محكمة، الناشر: مكتب الإعلام الإسلامي للحوزة العلمية في قم (المعهد العالي للعلوم والثقافة الإسلامية، قم، إيران) © الكتاب



تحسين العمليات وتنتاج صنع السياسات باستخدام الذكاء الاصطناعي. يمكن للذكاء الاصطناعي أن يساعد صانعي السياسات في تحديد الاحتياجات، تطوير البرامج، التنبؤ بالنتائج، وتحليل فعالية السياسات. تُظهر الدراسات أن الوثائق السياسية في جميع أنحاء العالم تؤكد على إمكانيات الذكاء الاصطناعي في المساعدة على تحقيق أهداف التنمية المستدامة، بما في ذلك علاج الأمراض المزمنة، والحد من الوفيات الناجمة عن حوادث الطرق، ومكافحة التغير المناخي، والتنبؤ بالتهديدات السيبرانية. يشهد التنافس والتعاون العالميان في مجال الذكاء الاصطناعي تزايداً مستمراً. فإنّ الذكاء الاصطناعي لا يوفر الفرص فحسب، بل ينطوي أيضاً على مخاطر وتحديات. وتشمل بعض المخاوف، الاستبدال الوظيفي، التحيز الخوارزمي، إساءة استخدام البيانات، والتهديدات الأمنية. يجب أن تعالج سياسات الذكاء الاصطناعي هذه التحديات، وأن تضمن تطويره واستخدامه بطريقة مسؤولة وأخلاقية.

الكلمات المفتاحية

الذكاء الاصطناعي، صناعة السياسات العامة، الحوكمة، الأخلاق، التنافس العالمي.

از سیاست گذاری تا مسئولیت پذیری: مسیر تحول

حکمرانی هوش مصنوعی

علی میرعرب^۱ 

تاریخ دریافت: ۱۴۰۲/۱۰/۲۴ • تاریخ اصلاح: ۱۴۰۳/۱۲/۰۲ • تاریخ پذیرش: ۱۴۰۳/۱۲/۲۶ • تاریخ انتشار آنلاین: ۱۴۰۴/۰۲/۱۴



چکیده

هوش مصنوعی (AI) به عنوان یک فناوری انقلابی، تحول آفرین و برهم زننده، در حال دگرگونی دنیا است. حکمرانی هوش مصنوعی و ادغام آن در سیاست عمومی، از موضوعات مهم در بحث های امروزی است. استفاده ی بالقوه از هوش مصنوعی در چرخه ی سیاست عمومی، فرصت هایی را برای بهبود فرآیندهای تصمیم گیری دولت ها فراهم می کند. تجزیه و تحلیل انتقادی فناوری های هوش مصنوعی در سیاست گذاری عمومی برای درک تأثیر آن ها بر فرآیندهای تصمیم گیری و نابرابری های اجتماعی ضروری است. اگرچه ادارات دولتی، نگاه های اقتصادی، سازمان های بین المللی و سایر طرفین، در حوزه ی هوش مصنوعی تلاش کرده اند، اما ادغام هوش مصنوعی و حکمرانی هنوز در مراحل اولیه ی توسعه قرار دارد. این پژوهش با استفاده از روش توصیفی-تحلیلی، و با هدف مطالعه بر روندها و چارچوب های حکمرانی هوش مصنوعی و سیاست گذاری عمومی، سعی دارد به بررسی چگونگی بهبود فرآیند و نتایج سیاست گذاری با استفاده از هوش مصنوعی بپردازد. هوش مصنوعی می تواند به سیاست گذاران در شناسایی نیازها، توسعه ی برنامه ها، پیش بینی نتایج و تحلیل اثربخشی

۴۶۰
 مطالعات
 حکمرانی اسلامی

سال اول، شماره اول (پیاپی ۱) بهار و تابستان ۱۴۰۳

۱. استادیار، پژوهشگاه علوم و فرهنگ اسلامی، قم، ایران. alimirarab@isca.ac.ir

* میرعرب، علی. (۱۴۰۴). از سیاست گذاری تا مسئولیت پذیری: مسیر تحول حکمرانی هوش مصنوعی. مطالعات حکمرانی اسلامی، ۱(۱)، صص ۴۵۶-۴۹۱. <https://doi.org/10.22081/jislamicgo.2025.70830.1000>

□ نوع مقاله: پژوهشی؛ ناشر: دفتر تبلیغات اسلامی حوزه علمیه قم (پژوهشگاه علوم و فرهنگ اسلامی، قم، ایران) © نویسندگان



سیاست‌ها کمک کند. مطالعات انجام‌شده نشان می‌دهد اسناد سیاستی در سراسر جهان بر پتانسیل هوش مصنوعی برای کمک به دستیابی به اهداف توسعه‌ی پایدار از جمله درمان بیماری‌های مزمن، کاهش تلفات جاده‌ای، مبارزه با تغییرات اقلیمی و پیش‌بینی تهدیدات سایبری کمک می‌کند. هوش مصنوعی نه تنها فرصت‌ها، بلکه خطرات و چالش‌هایی را نیز به همراه دارد. برخی از نگرانی‌ها شامل جابه‌جایی شغلی، سوگیری الگوریتمی، سوءاستفاده از داده‌ها و تهدیدات امنیتی است. سیاست‌های هوش مصنوعی باید به این چالش‌ها رسیدگی کند و اطمینان حاصل کند که هوش مصنوعی به طور مسئولانه و اخلاقی توسعه و استفاده می‌شود.

کلیدواژه‌ها

هوش مصنوعی، سیاست‌گذاری عمومی، حکمرانی، اخلاق، رقابت جهانی.

مقدمه

در سپتامبر ۲۰۱۷م، رئیس‌جمهور روسیه اعلام کرد که آینده متعلق به هوش مصنوعی است و «هر کسی که در هوش مصنوعی پیشرو باشد، بر جهان حکومت خواهد کرد (RT, 2017)». در اوایل سال ۲۰۱۸م، نخست‌وزیر وقت بریتانیا در سخنرانی خود در مجمع جهانی اقتصاد در داووس اعلام کرد که «ما در حال تثبیت جایگاه بریتانیا به عنوان رهبر جهانی در هوش مصنوعی هستیم (WEF, 2018)». در مارس ۲۰۱۸، رئیس‌جمهور فرانسه در حین ارائه‌ی یک راهبرد ملی هوش مصنوعی، برنامه‌ای را برای تبدیل کشورش به رهبر جهانی در تحقیق و نوآوری هوش مصنوعی اعلام کرد (Rabesandratana, 2018). در آوریل ۲۰۱۸م، بیانیه‌ی هوش مصنوعی برای اروپا، هدف اتحادیه‌ی اروپا را «تبدیل شدن به رهبر انقلاب هوش مصنوعی، به روش خود و بر اساس ارزش‌هایش» اعلام کرد (European Commission, 2018a). سند دسامبر ۲۰۱۸م در مورد چشم‌انداز اروپایی در هوش مصنوعی، موقعیت جهانی اتحادیه‌ی اروپا را به شرح زیر نشان می‌دهد:

«رقابت جهانی شدیدی در زمینه‌ی هوش مصنوعی بین ایالات متحده، چین و اروپا وجود دارد. ایالات متحده در حال حاضر پیشرو است، اما چین به سرعت در حال رسیدن است و هدف آن رهبری تا سال ۲۰۳۰م است. برای اتحادیه‌ی اروپا، این موضوع بیشتر به معنای برنده یا بازنده شدن در مسابقه نیست، بلکه به معنای یافتن راهی برای استفاده از فرصت‌های ارائه‌شده توسط هوش مصنوعی به روشی است که انسان‌محور، اخلاقی، امن و وفادار به ارزش‌های اصلی ما باشد» (European Commission, 2018b).

در سال‌های آینده، هوش مصنوعی به شیوه‌های تأثیرگذاری در حمل‌ونقل، بهداشت، انرژی، اخبار، رسانه‌های اجتماعی، هنر، آموزش، علم، تولید، اشتغال، نظارت، نیروی انتظامی و ارتش به کار گرفته خواهد شد. هوش مصنوعی به عنوان یک فناوری با کاربردهای عمومی (Garfinkel, 2022; Bresnahan, Trajtenberg, 1995)، تغییراتی ایجاد می‌کند که گسترده و عمیق است به طوری که پیش‌بینی آن‌ها دشوار است (Ding & Dafoe, 2021). مزایای این فناوری قابل توجه خواهد بود، اما خطرات بالقوه‌ای نیز به همراه خواهد داشت.

در دهه‌های آینده، تأثیرات هوش مصنوعی می‌تواند بسیار فراتر رود و به‌طور بالقوه رفاه، ثروت و قدرت را به میزان بیشتری نسبت به انقلاب هسته‌ای یا انقلاب صنعتی متحول کند. در همین راستا، محققان یادگیری ماشین (ML) امکان دستیابی گسترده به هوش مصنوعی در سطح هوش انسانی را طی یک یا دو دهه پیش‌بینی می‌کنند و معتقدند که این احتمال تا سال ۲۰۶۰ بیش از ۵۰ درصد است (Cotra, 2020) که پیامدهای آن عمیق خواهد بود، چه مثبت و چه منفی؛ بنابراین، توسعه و به‌کارگیری درست هوش مصنوعی اهمیت بسیار زیادی دارد و حوزه‌ی حکمرانی هوش مصنوعی^۱ به دنبال درک و ارائه‌ی اطلاعات در مورد این چالش است. حوزه‌ی حکمرانی هوش مصنوعی، به بررسی این موضوع می‌پردازد که چگونه بشریت به بهترین نحو می‌تواند به سمت سیستم‌های پیشرفته‌ی هوش مصنوعی حرکت کند.

برخلاف سایر فناوری‌های مخرب که به یک صنعت خاص محدود می‌شوند، هوش مصنوعی یک فناوری فراگیر است که می‌تواند از اصلاحات همه‌ی صنایع پشتیبانی کند؛ از این رو، هوش مصنوعی، دامنه‌ی وسیعی از پیامدهای اجتماعی را به همراه دارد که تغییرات عمیقی را در زمینه‌های مختلف از جمله سیاست، اقتصاد، جامعه و غیره ایجاد خواهد کرد و همچنین منجر به چالش‌هایی در حکمرانی می‌شود. کشورها و مناطق مختلف، از ایالات متحده و چین گرفته تا اتحادیه‌ی اروپا، فرانسه، بریتانیا و سایر کشورها، جاه‌طلبی‌های خود را برای رهبری در حوزه‌ی هوش مصنوعی اعلام کرده‌اند. اما چگونه می‌توان این توجه سیاسی و سیاست‌گذاری نسبت به هوش مصنوعی را توضیح داد؟ آیا تحولات جهانی به حکمرانی هوش مصنوعی منجر می‌شود؟ حکمرانی هوش مصنوعی چگونه خواهد بود؟ چه عواملی باعث سرعت گرفتن سیاست‌های مرتبط با هوش مصنوعی در سراسر جهان می‌شود؟ و چارچوب‌های اصلی سیاست‌های نوظهور هوش مصنوعی کدامند؟

1. AI Governance

۱. ضرورت توجه به مسائل حکمرانی و سیاست‌گذاری در حوزه‌ی هوش مصنوعی

پیشرفت‌های آینده، بسیار سریع بوده و دانشمندان انتظار دارند که شاهد تحقق قابلیت‌های فراانسانی در حوزه‌های استراتژیک در چند دهه‌ی آینده باشند. فرصت‌های ایجاد شده توسط پیشرفت‌های فناورانه بالاخص در حوزه‌ی هوش مصنوعی در حوزه‌های مختلف همانند پزشکی و سلامت، حمل‌ونقل، انرژی، آموزش، رشد اقتصادی و پایداری محیطی، بسیار عظیم است. در مقابل، مخاطرات ممکن نیز کم نبوده و پیشرفت‌های فناوری ما را با چالش‌های فراوانی در حوزه‌ی حکمرانی مواجه می‌کند؛ از جمله‌ی این مخاطرات می‌توان به جابه‌جایی نیروی کار، نابرابری، دگرگونی در قدرت ملی، بی‌ثباتی استراتژیک و نوعی از هوش مصنوعی که می‌تواند سلامت و دیگر ارزش‌ها را قربانی خود نماید، اشاره کرد. رهبران حکومت‌های مختلف و همچنین شرکت‌های بزرگ در این وضعیت جدید غالباً تقاضای مشاوره برای سیاست‌گذاری بهتر را می‌کنند، ولی هنوز هم توجه علمی به انقلابی که هوش مصنوعی ایجاد کرده است بسیار اندک است. به همین دلیل نیاز فوری به انجام پژوهش‌های متعدد درباره‌ی مسئله‌ی حکمرانی هوش مصنوعی وجود دارد. مسئله‌ای که خود را در بازاندیشی نهادها، سیاست‌ها و هنجارهای جهانی به‌منظور تضمین گسترش و کاربرد بهتر و سودمندانه‌تر هوش مصنوعی می‌یابد (شیخ‌الاسلامی، ۱۳۹۸).

حکمرانی شامل شکل‌دهی رفتار برای دستیابی به اهداف اجتماعی از طریق نهادها است. «نهادها» به تمام طیف ساختارهای اجتماعی که رفتار را شکل می‌دهند، از جمله هنجارها، آداب و رسوم، قوانین، سازمان‌ها، مقررات، نهادهای نظارتی و قانون‌گذاری اشاره دارد (North, 1991). با بررسی انواع پیامدهای بیرونی در گیر می‌توانیم مسئله‌ی حکمرانی را مفهوم‌سازی کنیم تا ببینیم چه ذی‌نفعان، منافع و سازوکارهایی باید در هر نهادی برای رسیدگی به آن‌ها گنجانده شود. چنین رویکرد کارکردگرایانه به نهادها در رشته‌ی اقتصاد و رویکردهای عقلانی در علوم سیاسی و سیاست‌گذاری، رایج است.

بنابراین، هنگامی که با مشکلی در حکمرانی روبه‌رو هستیم، می‌توانیم با پرسیدن

این که چه ویژگی‌هایی برای یک نهاد برای شکل‌دهی کافی رفتار به سمت اهداف اجتماعی مورد نظر لازم است، شروع کنیم. چه پیامدهای بیرونی، نیاز به درونی‌سازی دارند و بر کدام فضاها و سیاسی گسترش می‌یابند؟ آیا نهادهای موجود یا فرضی دارای موارد زیر هستند:

- حوزه‌ی جغرافیایی؟ (منظور قلمرو جغرافیایی و دامنه‌ی اثرگذاری است)
- حوزه‌ی موضوعی؟ (مقصود حوزه‌ی عملکرد و موضوعاتی است که نهاد بر آن‌ها تمرکز دارد)
- حوزه‌ی سیاسی؟ (به این معنا که آیا ذینفعان مرتبط را به طور کافی و مشروع نمایندگی می‌کنند؟)
- صلاحیت فنی؟ (توانایی نهاد در حل مسائل فنی مرتبط با موضوع)
- صلاحیت نهادی؟ (قابلیت نهاد در عمل به کارکردهایش)
- نفوذ؟ (مانند توانایی کافی در شکل‌دهی به مشوق‌های مادی)

هرچه فاصله‌ی یک حوزه‌ی نوظهور حکومتی از یک نهاد موجود، مشروع و باصلاحیت، بیشتر باشد، به احتمال زیاد، با چالش‌های بیشتری در مورد تطبیق یا ایجاد نهادهای مناسب روبه‌رو خواهیم شد. برخی از مسائل، با وجود این که ممکن است منافع اجتماعی روشنی به همراه داشته باشند، همچنان می‌توانند شکست بخورند، چرا که سازگاری نهادی مورد نیاز، بیش از حد زیاد است. به عنوان مثال در آمریکا، فواید دسترسی رایگان به تمام کتاب‌های بدون حق نشر در کتابخانه‌های محلی به صورت دیجیتال را در نظر بگیرید؛ این منفعت عمومی عمیق، نه به دلیل فقدان حزبی که مایل به اسکن و ارائه‌ی این سرویس باشد، بلکه به دلیل عدم توانایی کنگره برای به‌روزرسانی قانون کپی‌رایت جهت قانونی کردن آن ارائه نمی‌شود (Somers, 2017).

حکمرانی هوش مصنوعی در واقع مطالعه‌ی چگونگی مدیریت گذار به سیستم‌های هوش مصنوعی پیشرفته با تمرکز بر ابعاد گوناگون سیاسی، اقتصادی، نظامی حکومتی و اخلاقی است. حکمرانی هوش مصنوعی، اغلب در کنار سلامت و ایمنی هوش مصنوعی قرار داده می‌شود. البته هدف هر دوی آن‌ها کمک به بشریت برای استفاده‌ی

مفید هر چه بیشتر از هوش مصنوعی است. اما با این تفاوت که حکمرانی هوش مصنوعی بر نهادها و زمینه‌هایی که هوش مصنوعی ساخته شده و استفاده می‌شود، متمرکز است (Dafoe, 2018).

تحقیقات در مورد هوش مصنوعی در بخش دولتی هنوز در مراحل اولیه خود است و فقط اثرات مورد انتظار هوش مصنوعی را پوشش می‌دهد که ماهیت آن حدس و گمان است. دامنه‌ی وسیعی از تحقیقات تجربی در مورد کاربردهای هوش مصنوعی همراه با چالش‌های آن وجود دارد که توسط ذی‌نفعان بخش دولتی در نظر گرفته شده است (Medaglia, Sun, 2019). بسیاری از محققین و متخصصان فقط جنبه‌های فناورانه‌ی برنامه‌های هوش مصنوعی را ذکر کرده‌اند که بدون مدل قوی مدیریت دولتی که پیامدهای حاکمیت دولت اداری را به تصویر می‌کشد، مبهم به نظر می‌رسد. به موازات این مسأله، هنوز نیاز اساسی برای درک دامنه و تأثیر برنامه‌های کاربردی مبتنی بر هوش مصنوعی و چالش‌های مرتبط با آن به شیوه‌ای جامع وجود دارد.

مسئله‌ی حکمرانی هوش مصنوعی، ما را با پرسش‌های جدی مواجه می‌کند: برای به حداکثر رساندن ایمنی گذار به جهانی جدید که پیشرفت‌های رخ داده در هوش مصنوعی سهم فراوانی در شکل دادن به آن دارند چه باید کرد؟ چه توصیه‌هایی به آزمایشگاه‌ها، حکومت‌ها، سازمان‌های مردم‌نهاد و عامه‌ی مردم چه هم‌اکنون و چه در آینده می‌توانیم ارائه دهیم؟ چه مقررات بین‌المللی در حوزه‌ی پروتکلی، سازمانی، فناورانه و... برای ممانعت از درافتادن کشورها و شرکت‌های بزرگ در دام منافع کوتاه‌بینانه لازم داریم؟

۲. روش تحقیق

در این پژوهش از روش تحلیل محتوا برای بررسی حکمرانی هوش مصنوعی و سیاست‌گذاری عمومی استفاده شده است. جامعه‌ی آماری این پژوهش شامل کلیه‌ی منابع منتشر شده در موضوع تحقیق از سال ۱۳۹۸ به بعد برای منابع فارسی و از سال ۲۰۲۰م به بعد برای منابع انگلیسی است. برای جمع‌آوری داده‌ها از روش کتابخانه‌ای

استفاده شده است. نرم افزار اندنوت^۱ برای مدیریت منابع و نرم افزار NVivo برای تحلیل محتوا مورد استفاده قرار گرفته است.

برای تجزیه و تحلیل داده‌ها از روش تحلیل کیفی محتوا استفاده شده است. در این روش، داده‌ها، کدگذاری شده و سپس با استفاده از تکنیک‌های مختلف تحلیل محتوا، مضامین و مفاهیم کلیدی استخراج و تفسیر شده‌اند. برای افزایش روایی و پایایی پژوهش، از روش خودبازبینی محقق استفاده شده است.

این پژوهش به منابع منتشر شده در سال‌های ۱۳۹۸ تا ۱۴۰۲ و ۲۰۲۰ تا ۲۰۲۴م محدود می‌شود؛ همچنین ممکن است برخی از منابع مرتبط با موضوع تحقیق از قلم افتاده باشد. یکی دیگر از محدودیت‌های این پژوهش، با توجه به روش تحلیل محتوا که یک روش تفسیری است، آن است که ممکن است نتایج به تفسیر محقق بستگی داشته باشد.

در نهایت، با استفاده از روش‌های تحلیل و تفسیر داده‌ها، نتایج حاصل از مطالعات پیشین و یافته‌های جدید تحقیق مورد بررسی و ارزیابی قرار گرفت. در کل، فرآیند تحقیق به صورت غیرتجربی و با رویکرد کیفی انجام شد تا اهداف پژوهش به درستی دنبال شوند. نتایج این مطالعه می‌تواند به درک بهتر تحولات حکمرانی هوش مصنوعی و سیاست‌گذاری عمومی و همچنین به ارتقای تصمیم‌گیری در این حوزه کمک کند.

۳. چارچوب حکمرانی هوش مصنوعی

حکمرانی به قواعد و فرآیندهای تصمیم‌گیری جمعی در محیط‌هایی با تعدد بازیگران و سازمان‌ها اشاره دارد که در آن هیچ سیستم کنترل رسمی نمی‌تواند روابط بین آنها را به‌طور کامل تعیین کند. واسودها چاتری و گری استوکر تأکید دارند که حکمرانی شامل تصمیمات رسمی و غیررسمی، قراردادهای و آداب و رسوم است که توسط گروه‌های مختلف گرفته می‌شود و می‌تواند راهبردی یا مربوط به اجرای روزمره‌ی یک

1. Endnote

سیستم یا سازمان باشد. آنها معتقدند که در حکمرانی، تعاملات اجتماعی، بیشتر به مذاکره، ارتباطات و نفوذ هژمونیک متکی است تا نظارت مستقیم. تضاد و مخالفت نیز جزء ضروری این فرآیند است، زیرا حکمرانی نیازمند هماهنگی و تصمیم‌گیری در بستر تعدد دیدگاه‌ها و منافع مختلف است (Stoker, Chhotray, 2009).

در زمینه‌ی سیستم‌های اجتماعی-فنی، سوزانا بوراس و جاکوب ادلر (۲۰۱۴) حکمرانی را به عنوان مکانیزم‌هایی معرفی می‌کنند که از طریق آنها بازیگران اجتماعی و دولتی برای تنظیم مسائل مهم جامعه با هم تعامل و هماهنگی دارند. این مفهوم نشان می‌دهد که دولت باید فعالیت‌های خود را با طیف وسیعی از بازیگران، از جمله بخش خصوصی و جامعه‌ی مدنی، هماهنگ کند. این تغییر نقش دولت در حکمرانی فناوری‌ها در مفاهیمی مانند "دولت کارآفرین" نیز منعکس می‌شود که بر اهمیت سرمایه‌گذاری و خطرپذیری بخش دولتی برای شکل‌دهی به توسعه‌ی فناوری به سمت اهداف اجتماعی تأکید می‌کند. این اهداف شامل چالش‌های بزرگ یا مأموریت‌هایی در حوزه‌هایی مانند اهداف توسعه‌ی پایدار سازمان ملل متحد، محیط زیست و تغییرات آب و هوایی است (Mazucato, 2021).

حکمرانی علوم و فناوری‌های نوظهور با چالش‌های خاصی مواجه است، زیرا این حوزه‌ها با ابهامات زیادی درباره‌ی تحولات آتی، منافع اجتماعی و خطرات احتمالی همراه هستند. استفان کولمن و همکاران (۲۰۱۹) مفهوم «حکمرانی آزمایشی» را پیشنهاد کرده‌اند که در آن مداخلات دولتی و خصوصی به عنوان فرآیندی پویا و مقدماتی طراحی می‌شوند تا انعطاف‌پذیری و قابلیت بررسی داشته باشند. این نوع حکمرانی به‌ویژه برای علوم و فناوری‌های نوظهور که با ابهامات بیشتری همراه هستند، اهمیت دارد. بینش‌های موجود در ادبیات فناوری‌های نوظهور برای درک تحولات اخیر و جاری در حوزه‌ی هوش مصنوعی بسیار مفید است، زیرا نشان می‌دهد که هوش مصنوعی با همکاری و رقابت پویا و انتظارات مثبت و منفی همراه است. مطالعات مربوط به حکمرانی هوش مصنوعی از رویکردهای پژوهش و نوآوری مسئولانه بهره می‌برند که بر چهار بعد پیش‌بینی، بازتابندگی، شمول و پاسخ‌گویی استوار است (Digum, 2019). این

رویکردها به‌ویژه در تحلیل ابتکارات بین‌المللی و سیاست‌گذاری کاربردهای خاص هوش مصنوعی مانند وسایل نقلیه‌ی خودران و بررسی اسناد ملی و بین‌المللی هوش مصنوعی از سال ۲۰۱۶م به بعد اهمیت دارند (Galanos, 2019; Cath & etal., 2018).

چارچوب‌های موجود حکومتی و مقرراتی نیز به دلیل اطلاعات ناکافی مورد نیاز برای درک فناوری و عقب‌ماندگی قوانین نسبت به پیشرفت‌های هوش مصنوعی، برای مدیریت مشکلات اجتماعی ناشی از هوش مصنوعی مناسب نیستند. شرکت‌های بزرگ فناوری و توسعه‌دهندگان هوش مصنوعی مانند گوگل، فیس‌بوک، مایکروسافت و اپل، از نظر اطلاعاتی و منابعی، نسبت به دولت‌ها در تنظیم هوش مصنوعی، برتری قابل توجهی دارند که جایگاه سنتی دولت‌ها در توزیع و کنترل منابع در جامعه را به طور قابل توجهی تحت‌الشعاع قرار می‌دهد (Guihot & etal., 2017). عدم تقارن اطلاعاتی بین شرکت‌های فناوری و تنظیم‌گران، دشواری درک و اعمال قوانین جدید یا موجود بر کاربردهای هوش مصنوعی را برای تنظیم‌گران تشدید می‌کند (Taeihagh & etal., 2021). تنظیم‌گران در تلاش برای پرکردن این شکاف‌های اطلاعاتی با مشکل مواجه هستند و به دلیل پیشرفت‌های سریع در فناوری، عقب‌مانده‌اند؛ که این امر به نوبه‌ی خود منجر به تدوین قوانینی می‌شود که «خیلی کلی» یا «مبهم» هستند و برای تنظیم مؤثر فناوری، از دقت کافی برخوردار نیستند.

از اواخر ۲۰۱۶م، توجه سیاسی و سیاستی به هوش مصنوعی در سطح جهانی بی‌سابقه بوده است. هوش مصنوعی تفاوت‌های مهمی با فناوری‌های پیشین دارد، چرا که تأثیرات آن فراتر از سلامت و اقتصاد انسان بوده و شامل تأثیرات عمده‌ای بر نظام سیاسی، بازار کار، دولت و رفاه می‌شود. ایده‌ی این که هوش ماشین می‌تواند بر هوش انسانی غلبه کند، مورد توجه بسیاری قرار گرفته است؛ به‌ویژه در مواردی نظیر شکست برنامه‌ی کامپیوتری آلفاگو مقابل قهرمان جهان (Lee, 2018). سازمان‌های بین‌المللی مانند مجمع جهانی اقتصاد (WEF) و سازمان توسعه و همکاری اقتصادی (OECD) نیز نقش مهمی در توسعه‌ی سیاست‌های اخیر هوش مصنوعی ایفا کرده‌اند. در چارچوب WEF، تمرکز بر هوش مصنوعی بخشی از بحث‌های مربوط به انقلاب صنعتی چهارم بوده که توسط

کلاوس شواب، بنیان‌گذار و رئیس اجرایی WEF، ترویج شده است. شواب معتقد است که از سال ۲۰۰۰م، جهان شاهد انقلاب صنعتی چهارم بوده که با ادغام فناوری‌های جدید، در حوزه‌های فیزیکی، دیجیتال و زیستی مشخص می‌شود (Schwab, 2017).

برای رسیدگی به چالش‌های حکمرانی ناشی از عدم قطعیت و پیچیدگی توسعه‌ی هوش مصنوعی، اتخاذ رویکردهای نوآورانه‌ی حکمرانی مانند حکمرانی تطبیقی و حکمرانی ترکیبی یا "غیرمتمرکز"، مورد توجه زیادی قرار گرفته است (Dafoe, 2018). ویژگی بارز حکمرانی تطبیقی و ترکیبی، کاهش نقش دولت در کنترل توزیع منابع در جامعه است. این مسأله در حکمرانی ترکیبی به عنوان ترکیبی از بازیگران دولتی و غیر دولتی تعریف می‌شود، یا تمایز بین بخش دولتی و خصوصی را تا درجات مختلفی محو می‌کند؛ جایی که تنظیم مقررات به عنوان ترکیبی از استانداردهای صنعتی و «نظارت عمومی» وجود دارد (Guihot & et al., 2017).

حکمرانی ترکیبی می‌تواند به اشکال تنظیم مشترک، خودتنظیمی اجباری و فرا-تنظیم وجود داشته باشد (Hemphill, 2016) که همه بر نقش فزاینده‌ی بازیگران غیردولتی و نیاز به «ارزیابی مداوم تعادل قدرت» بین بازیگران بخش خصوصی و دولتی تأکید می‌کنند. به طور مشابه، حکمرانی تطبیقی بر نیاز به دور شدن از اقدامات "فرمان و کنترل" به سمت رویکردهای انعطاف‌پذیرتر تأکید می‌کند که با تعدیل و بهبود تکراری مقررات و سیاست‌ها در حین جمع‌آوری اطلاعات جدید مشخص می‌شود (Li & et al., 2021).

بسیاری از مطالعات اخیر در زمینه‌ی چارچوب‌های پیشنهادی برای حکمرانی هوش مصنوعی، تأکید بر ایجاد اجماع گسترده‌ی جامعه بر روی اصول اخلاقی هوش مصنوعی و اطمینان از پاسخ‌گویی است، اما نیاز به مطالعاتی وجود دارد که بررسی کنند چگونه این چارچوب‌ها را می‌توان به طور عملی اجرا کرد. چارچوب سه‌لایه‌ای گاسر و آلْمیدا^۱ (۲۰۱۷) شامل یک لایه فنی برای فرآیندهای سیستم هوش مصنوعی و ساختارهای داده، یک لایه برای طراحی اخلاقی هوش مصنوعی و لایه سوم شامل پیامدهای

1. Gasser & Almeida

اجتماعی هوش مصنوعی و نقش مقررات و قوانین است. رحوان^۱ (۲۰۱۸) پیشنهاد می‌کند رویکرد «انسان در حلقه^۲» را به رویکرد گسترده‌تر «جامعه در حلقه^۳» گسترش دهد، جایی که جامعه ابتدا مسئول یافتن اجماع در مورد ارزش‌هایی است که باید هوش مصنوعی و هزینه و فایده آن را بین ذی‌نفعان مختلف شکل دهد.

برخلاف مفهوم سنتی یک چرخه‌ی واحد که نیازمند تکرارهای متعدد برای تولید یک سیاست عمومی است، چرخه‌ی سیاست‌گذاری در عصر هوش مصنوعی یک جریان پویا و انعطاف‌پذیر را دنبال می‌کند. هوش مصنوعی می‌تواند با تجزیه و تحلیل داده‌ها، با تکنیک‌ها و شبیه‌سازی‌های هوش مصنوعی در هر مرحله از چرخه، بازخورد ایجاد کند و به این ترتیب، چرخه‌ی سنتی را به رویکردی افزایشی و مارپیچی تبدیل کند (Valle-Cruz & et al., 2020). در نتیجه، برای ارزیابی و کسب نتایج، نیازی به انتظار برای مرحله‌ی اجرا نخواهد بود. هوش مصنوعی می‌تواند سناریوها، شبیه‌سازی‌ها و راه‌حل‌های مختلفی را برای هر مرحله از چرخه‌ی سیاست‌گذاری عمومی تقریباً به صورت بلادرنگ ارائه دهد.

در چرخه‌ی سیاست‌گذاری پویا، تعیین دستور کار می‌تواند همزمان با ارزیابی سیاست انجام شود؛ تدوین و تصمیم‌گیری سیاستی تقریباً هم‌زمان با اجرای سیاست اتفاق بیفتد. کاربردهای هوش مصنوعی معتبر برای چرخه‌ی سیاست‌گذاری عمومی می‌توانند داده‌ها را برای کشف الگوهای مفید در پیش‌بینی رویدادها رصد کند. با این حال، خطراتی مانند تبعیض الگوریتمی، ابهام، کنترل اجتماعی، جایگزینی انسان، حریم خصوصی داده‌ها و شکاف دیجیتالی رو به افزایش، باید تحت حاکمیت هوش مصنوعی مدیریت شوند. هوش مصنوعی با پاندمی، ظهور اجباری داشته است. تلاش‌ها برای تولید واکسن، درک و ویروس، مدیریت انتشار و کنترل جمعیت، هم‌افزایی برای تمام پیشرفت‌های هوش مصنوعی ایجاد خواهد کرد.

1. Rahwan

2. human-in-the-loop

3. society-in-the-loop

۴. چالش‌ها و فرصت‌های حکمرانی هوش مصنوعی

با توسعه‌ی سریع فناوری هوش مصنوعی، مشکلات اجتماعی متعددی به طور مداوم پدیدار شده‌اند. از یک طرف، فناوری هوش مصنوعی نیز نوعی فناوری است که ممکن است با خطر استفاده‌ی غیرقانونی مواجه شود. علاوه بر این، اطلاعات زیادی در مورد حریم خصوصی و محرمانگی در فناوری هوش مصنوعی وجود دارد. در صورت استفاده بی‌رویه، نظم کاری عادی ادارات مربوطه را مختل کرده و مشکلات و تضادهای اجتماعی را تشدید می‌کند و جامعه را در محیطی ناپایدار قرار می‌دهد. از طرف دیگر، در عصر جدید، با توسعه‌ی سریع هوش مصنوعی، حاکمیت اجتماعی به مراتب دشوارتر از گذشته است. استفاده‌ی گسترده از انواع دستگاه‌ها و ماشین‌های هوشمند نهایی در جامعه‌ی کنونی، افکار عمومی جامعه را تغییر داده است. علاوه بر این، مردم در عصر حاضر بیشتر در معرض نوسانات ایدئولوژیک قرار دارند که منجر به آشفتگی افکار عمومی می‌شود که این خود مشکلی برای حاکمیت اجتماعی است (Valle-Cruz & et al., 2020).

با توجه به این موضوع، بخش‌های دولتی باید از نزدیک، توسعه‌ی هوش مصنوعی را دنبال کنند تا نوآوری و توسعه‌ی سیستم‌های حکمرانی اجتماعی را در عصر هوش مصنوعی سرعت بخشند و به سیستم امنیت اجتماعی اجازه دهند تا با پیشرفت فناوری هوش مصنوعی توسعه یابد. تنها از این طریق مردم می‌توانند گرمای ناشی از توسعه‌ی فناوری هوش مصنوعی را احساس کنند و بدین ترتیب این فناوری توسط اکثریت قریب به اتفاق مردم عادی پذیرفته شود. بخش‌های دولتی نه تنها باید به طور جدی به کاربرد فناوری هوش مصنوعی برای حل مشکلات ناشی از توسعه‌ی سیاسی و اقتصادی فعلی توجه کنند، بلکه باید به طور فعال برای اجتناب از آن‌ها تلاش کنند، تناقضات اجتماعی ناشی از هوش مصنوعی را پیش‌بینی کرده و سیاست‌های عمومی را از پیش برای کاهش یا اجتناب از این تناقضات تدوین کنند. تنها با هم‌سو شدن اصلاح و توسعه‌ی سیاست‌های عمومی با توسعه‌ی فناوری هوش مصنوعی می‌توان به این فناوری

خدمت کرد تا به جامعه و مردم خدمت کند و در عین حال از بروز مشکلات و تناقضات اجتماعی جلوگیری کند.

یکی از چالش‌های کلیدی در حکمرانی هوش مصنوعی، نیاز به برقراری تعادل بین مزایای بالقوه‌ی فناوری هوش مصنوعی و خطرات و نگرانی‌های اخلاقی مرتبط با آن است. سیستم‌های هوش مصنوعی این پتانسیل را دارند که صنایع را متحول کنند و کیفیت زندگی افراد را بهبود بخشند، اما همچنین پرسش‌های مهمی را در مورد حریم خصوصی، سوگیری و پاسخ‌گویی مطرح می‌کنند. سیاست‌های عمومی با تعیین خط‌مشی‌ها و مقررات حاکم بر توسعه و استفاده از سیستم‌های هوش مصنوعی، نقش اساسی در رسیدگی به این مسائل ایفا می‌کنند.

همچنین مدیریت داده، موضوع کلیدی دیگری در مباحث مربوط به حکمرانی هوش مصنوعی است. چرا که چالش‌های متعدد سازمانی و فناوریانه‌ای وجود دارند که کنترل مؤثر بر داده‌ها و تخصیص مسئولیت برای تصمیم‌گیری‌های مبتنی بر داده توسط سیستم‌های هوش مصنوعی را با مشکل مواجه می‌کنند. تکه‌تکه شدن داده‌ها و فقدان قابلیت همکاری بین سیستم‌ها، کنترل سازمان‌ها بر جریان داده‌ها، در کل، چرخه‌ی عمر آن را محدود می‌کند و نقش‌های مشترک بین طرف‌های مختلف در به اشتراک‌گذاری داده‌ها، زنجیره‌ی پاسخ‌گویی و علت بین تصمیمات/رویدادهای هدایت‌شده توسط هوش مصنوعی و طرف‌های دخیل در تسهیل آن تصمیم را مبهم می‌کند (Janssen & et al., 2020). شناسایی و حذف همه‌ی متغیرهایی را که با دسته‌های حساس داده‌های شخصی مرتبط هستند، تقریباً غیرممکن می‌سازد (Carabantes, 2020). ثانیاً، تصمیمات یادگیری ماشین، غیرقابل پیش‌بینی هستند، زیرا از داده استخراج شده و می‌توانند با تغییرات جزئی در ورودی‌ها به طور قابل توجهی تغییر کنند. عدم کنترل انسان بر رفتار سیستم‌های هوش مصنوعی، دشواری در تعیین مسئولیت و پاسخ‌گویی برای آسیب‌های ناشی از نقص نرم‌افزار را نشان می‌دهد، زیرا سازندگان و برنامه‌نویسان اغلب نمی‌توانند ورودی‌ها و قوانین طراحی را که می‌توانند منجر به نتایج ناامن یا تبعیض‌آمیز شوند، پیش‌بینی کنند. شفافیت و پاسخ‌گویی در فرآیندهای تصمیم‌گیری هوش مصنوعی، جنبه‌ی مهم

دیگری از حکمرانی هوش مصنوعی است. با پیچیده تر و خودمختارتر شدن سیستم های هوش مصنوعی، درک چگونگی رسیدن این سیستم ها به نتایجشان دشوارتر می شود. سیاست های عمومی می توانند با الزام شرکت ها و سازمان ها به شفافیت در مورد داده ها و الگوریتم هایی که استفاده می کنند و با ایجاد سازوکارهایی برای پاسخ گویی در صورت بروز خطا یا سوگیری در سیستم های هوش مصنوعی، به رفع این چالش کمک کنند.

علاوه بر شفافیت و پاسخ گویی، سیاست های عمومی همچنین می توانند نقش کلیدی در ترویج عدالت و انصاف در سیستم های هوش مصنوعی ایفا کنند. فناوری های هوش مصنوعی در صورت عدم طراحی و اجرای دقیق، پتانسیل تشدید نابرابری های موجود را دارند. سیاست های عمومی می توانند با اطمینان از آزمایش سیستم های هوش مصنوعی برای سوگیری و تبعیض، و با ترویج تنوع و شمول در توسعه و استقرار فناوری های هوش مصنوعی، به رفع این چالش کمک کنند.

علاوه بر این، حکمرانی هوش مصنوعی همچنین شامل پرسش های امنیت ملی و ژئوپلیتیک است. با اهمیت یافتن فزاینده ی فناوری های هوش مصنوعی در عملیات نظامی و اطلاعاتی، نیاز به همکاری و هماهنگی بین المللی برای اطمینان از استفاده ی مسئولانه و اخلاقی از این فناوری ها به طور فزاینده ی احساس می شود. سیاست های عمومی می توانند با تقویت گفت و گو و همکاری بین کشورها و با ایجاد هنجارها و توافق نامه هایی که بر استفاده از هوش مصنوعی در زمینه های نظامی و امنیتی حاکم است، به رفع این نگرانی ها کمک کنند.

در عین حال، سیاست های عمومی همچنین باید به پیامدهای اقتصادی فناوری هوش مصنوعی پردازند. با خودکارسازی وظایف توسط سیستم های هوش مصنوعی و جابه جایی کارگران، نیاز به سیاست هایی وجود دارد که از بازآموزی و ارتقای مهارت نیروی کار حمایت کند، و همچنین اقداماتی برای اطمینان از توزیع عادلانه ی مزایای فناوری هوش مصنوعی صورت گیرد. سیاست های عمومی می توانند با حمایت از برنامه های آموزشی و مهارتی، ترویج کارآفرینی و نوآوری و ایجاد شبکه های ایمنی اجتماعی برای حمایت از افرادی که به دلیل اتوماسیون مبتنی بر هوش مصنوعی جابه جا

می‌شوند، نقش کلیدی در شکل‌دادن به آینده‌ی کار در عصر هوش مصنوعی ایفا کنند. برای رسیدگی مؤثر به چالش‌های پیچیده‌ی حکمرانی هوش مصنوعی، سیاست‌های عمومی باید از طیف وسیعی از دیدگاه‌ها و تخصص‌ها مطلع شوند. این شامل مشارکت فناوران، اخلاق‌مداران، سیاست‌گذاران و سایر ذی‌نفعانی است که می‌توانند دیدگاه‌ها و بینش‌های ارزشمندی در مورد توسعه و استقرار فناوری‌های هوش مصنوعی ارائه دهند. فرآیندهای سیاست‌گذاری عمومی باید شفاف و فراگیر باشند و امکان مشارکت و بازخورد معنادار از طیف وسیعی از ذی‌نفعان را فراهم کنند.

علاوه بر مشارکت ذی‌نفعان، سیاست‌های عمومی همچنین باید آینده‌نگر و انعطاف‌پذیر باشند تا با سرعت بالای تحول فناوری هوش مصنوعی همگام شوند. با قدرت‌مندتر و پیچیده‌تر شدن سیستم‌های هوش مصنوعی، ضروری است که چارچوب‌های سیاست‌گذاری عمومی بتوانند در پاسخ به تحولات و چالش‌های جدید، سازگار و تکامل پیدا کنند. این امر نیازمند پایش و ارزیابی مستمر فناوری‌های هوش مصنوعی و همچنین آمادگی برای تجدید نظر و به‌روزرسانی سیاست‌ها در صورت لزوم است.

برای معضلات اجتماعی جدید ناشی از تأثیر توسعه‌ی هوش مصنوعی، لازم است نظام حکمرانی جدیدی ساخته شود و ابزارهای حکمرانی جدیدی توسعه یابد. چالش‌های حکمرانی ناشی از توسعه‌ی هوش مصنوعی نه تنها در عدم انطباق سیستم موجود، بلکه در شکاف‌های حکمرانی ناشی از مسائل جدید نیز منعکس می‌شود. با در نظر گرفتن نگرانی جدی در مورد بیکاری به عنوان مثال، این یک واقعیت انکارناپذیر است که کاربرد گسترده‌ی هوش مصنوعی و روبات‌ها، از نظر امکانات فناورانه، جایگزین کار دستی شده است. روزنامه‌نگاران، تحلیل‌گران بازار سهام یا حتی کارگران حقوقی ممکن است توسط ماشین‌ها جایگزین شوند. در جهانی با «اتوماسیون کامل»، درک مجدد رابطه بین کار و امنیت رفاه و بازسازی سیستم کار و امنیت رفاه، یکی از چالش‌های حکمرانی است که نیاز به راه حل فوری دارد (Esteve, Mikhaylov, Campion, 2018).

در نهایت، حکمرانی فناوری هوش مصنوعی مجموعه‌ای منحصر به فرد از چالش‌ها را

به وجود می آورد که نیازمند بررسی دقیق و همکاری بین دولت‌ها، صنعت و جامعه‌ی مدنی است. سیاست‌های عمومی، نقش اساسی در هدایت توسعه و استقرار فناوری‌های هوش مصنوعی به شیوه‌ای مسئولانه و اخلاقی ایفا می‌کنند و اطمینان حاصل می‌کنند که در عین به حداقل رساندن خطرات و آسیب‌های بالقوه، از مزایای هوش مصنوعی بهره‌مند شویم. با اتخاذ رویکردی متفکرانه و همه‌گیر به حکمرانی هوش مصنوعی، سیاست‌گذاران می‌توانند به شکل‌دهی آینده‌ای کمک کنند که در آن فناوری‌های هوش مصنوعی در خدمت منافع عمومی باشد و به ایجاد جامعه‌ای عادلانه‌تر و منصفانه‌تر کمک کند.

سیاست خوب می‌تواند حاکمیت هوش مصنوعی را به نفع بشریت به ارمغان آورد. هوش مصنوعی این پتانسیل را دارد که فرآیندهای سازمانی و نحوه‌ی انجام فعالیت‌ها را به طور کلی تغییر دهد (Fernandez-Cortez & et al., 2020). اجرای روش‌های هوش مصنوعی در بخش دولتی، پویایی و چابکی را در چرخه‌ی سیاست‌گذاری عمومی افزایش می‌دهد و فرآیندها و فعالیت‌ها را تسریع می‌کند. در این راستا، سرعت تصمیم‌گیری در چرخه‌ی سیاست‌گذاری به متغیری تبدیل شده که به طور قابل توجهی بر فرآیند آن تأثیر می‌گذارد، زیرا این فرآیند می‌تواند تحت فشار تعداد زیاد مشکلات پیش روی دولت‌ها قرار گیرد (Criado, 2021). علاوه بر این، توانایی‌های توسعه‌یافته برای ارزیابی نتایج و انتشار اطلاعات، فرآیند تدوین استراتژی‌های مبتنی بر سیاست‌های عمومی را کند می‌کند. با اجرای تکنیک‌های هوش مصنوعی در بخش دولتی، امکان تسریع فرآیندهای سیاست‌گذاری عمومی و حکمرانی وجود دارد (Valle-Cruz & et al., 2020).

پویایی و چابکی در چرخه‌ی سیاست‌گذاری عمومی، با هدف پرکردن شکاف پردازش اطلاعات و تصمیم‌گیری بین دولت، بخش‌های مختلف تولیدی و بازیگران اجتماعی است (Ulbricht, Katzenbach, 2019; Wood, Gritsenko, 2020). ماهیت مخرب هوش مصنوعی می‌تواند سازمان‌ها و فرآیند سیاسی را متحول کند. در این راستا، هوش مصنوعی می‌تواند یک تیغ دولبه باشد: اول به عنوان ابزاری برای بهبود کارآمدی سازمان و دوم به عنوان مکانیزمی برای تقویت رویه‌های دیوانسالارانه‌ی بد (Young &

etal., 2019). در مورد مورد اول، هوش مصنوعی امکان تجزیه و تحلیل کلان داده‌ها، یافتن الگوها و اطلاعاتی را که انسان‌ها نمی‌توانند تشخیص دهند، فراهم می‌کند. همچنین، هوش مصنوعی می‌تواند وظایف روتین را با دقت بیشتر و بدون وقفه انجام دهد و فعالیت‌های پرخطر یا سنگین را بدون به خطر انداختن جان انسان انجام دهد. اجرای هوش مصنوعی در بخش دولتی، آزادی انجام رویه‌های خلاقانه یا ابتکاری را برای تصمیم‌گیری در اختیار می‌گذارد. همچنین، هوش مصنوعی می‌تواند به انسان‌ها این امکان را بدهد که از مزایای آن برای تقویت قابلیت‌های خود و همچنین ایجاد چارچوبی برای مشارکت اجتماعی در چرخه‌ی سیاست‌گذاری از طریق روش‌هایی که پیچیدگی را ساده کرده و دسترسی به سازوکارهایی که تولید را مردم سالارانه می‌کند، بهره‌مند شوند (Wirtz & etal., 2020).

نتیجه‌گیری

در حالی که توسعه‌ی هوش مصنوعی به دهه‌ی ۱۹۵۰م برمی‌گردد، پیشرفت‌های چشم‌گیر در دسترسی به داده و قدرت پردازش، اخیراً اتفاق افتاده است. این موضوع به استفاده‌ی فزاینده از هوش مصنوعی در طیف گسترده‌ای از حوزه‌ها، از مبارزات انتخاباتی و بازار کار گرفته تا سلامت، آموزش و پرورش و امور نظامی، و موارد دیگر منجر شده است. فرصت‌های جدید فناوری و تهدیدهای مرتبط با آن باعث ایجاد بحث‌های داغ میان سیاست‌گذاران، سیاست‌مداران، کارشناسان و ذی‌نفعان درباره‌ی تأثیرات مثبت و منفی هوش مصنوعی بر سیاست، اقتصاد، بازار کار، عدالت، حریم خصوصی و سایر مسائل کلیدی جامعه شده است. سیاست عمومی و حکمرانی می‌توانند نقش مهمی در اطمینان از پیشرفت‌های مفید هوش مصنوعی و جلوگیری از تأثیرات مخرب آن داشته باشند. در حالی که اخیراً توسعه‌ی هوش مصنوعی عمدتاً توسط شرکت‌های خصوصی بزرگ جهانی و انگیزه‌های سود آن‌ها هدایت می‌شده، تحولات نوظهور سیاستی حاکی از آن است که دولت‌های ملی و سازمان‌های بین‌المللی، با همکاری طیف وسیعی از ذی‌نفعان، در حال آماده‌سازی چارچوب‌های حکمرانی برای

هوش مصنوعی هستند. هدف نهایی این است که به سمت اهدافی مانند پاسخ‌گویی، عدالت و شمولیت حرکت کنیم. سیاست‌ها و مباحث پرشتاب پیرامون هوش مصنوعی تاکنون عمدتاً از منظر اخلاق و فلسفه مورد بررسی قرار گرفته‌اند. اما توجه کمتری به جنبه‌ی سیاست‌گذاری و حکمرانی شده است، در حالی که این حوزه می‌تواند نقش مهمی در شکل‌دهی به توسعه و کاربرد این فناوری مطابق با منافع و ارزش‌های جامعه ایفا کند.

هوش مصنوعی کاربردهای مختلف و فراوانی دارد و در حوزه‌های گوناگون زندگی اجتماعی، که به نوعی میدان عمل سیاست‌گذاری حساب می‌شوند، می‌توانیم از هوش مصنوعی استفاده کنیم. این کاربرد هوش مصنوعی در سیاست‌گذاری و مدیریت دولت و تأثیراتش می‌تواند به صورت بنیادی بوده و پیش‌فرض‌ها و شاخص‌هایی که با استفاده از آن‌ها وضعیت و کیفیت یک حکومت ارزیابی می‌شود را تحت تأثیر قرار دهد. به طور مثال انواع مدل‌های بهینه‌سازی توزیع یارانه‌ها، فهم شهروندان از عدالت و کارکردهای نهادهای نظام اقتصادی را به شدت تحت تأثیر قرار خواهد داد و یا در مثال دیگر می‌توان به شاخص‌های مبارزه با فساد که یکی از شاخص‌های اندازه‌گیری حکمرانی خوب به حساب می‌آید، اشاره کرد که به صورت جدی تحت تأثیر استفاده از الگوریتم‌های جدید کشف فساد قرار می‌گیرند. به طور کلی می‌توان این بخش را با عنوان جهش در سیستم‌های خبره در حوزه‌ی سیاست‌گذاری نیز نام‌گذاری کرد.

سیاست‌گذار، همان‌گونه که باید وظایف خود را در حوزه‌های مختلف اقتصادی-اجتماعی، سیاسی و ... انجام دهد و در مورد موضوعات جاری، به ویژه اثرگذاری مربوط سایر فناوری‌ها، برنامه‌ریزی کرده و تصمیمات مشخص و نظام‌یافته‌ای اتخاذ کند؛ باید در مورد هوش مصنوعی نیز به عنوان یک مسأله‌ی جدی، سیاست‌گذاری‌های مناسبی داشته باشد. هوش مصنوعی در طبقه‌بندی انواع فناوری، در دسته‌ی فناوری‌های تحول‌آفرین و برانداز^۱ قرار می‌گیرد و با توجه به شمول اثرات آن به صنایع و نهادها و

1. disruptive

طبقات مختلف جامعه، باید سیاست‌گذاری در ارتباط با آن کاملاً جدی گرفته شود. در صورتی که سیاست‌گذاری و به تبع آن برنامه‌ریزی و اجرای مناسبی در مورد هوش مصنوعی صورت نگیرد، قطعاً جامعه و سیستم حاکمیتی با مشکل روبه‌رو خواهند شد. توسعه‌ی سریع سیاست‌گذاری هوش مصنوعی، در میان بسیاری از راهبردها و سایر اسناد سیاستی که در سال‌های اخیر شروع شده‌اند، در سیاست فناوری بی‌سابقه است. این را می‌توان با طیف گسترده‌ای از کاربردهای هوش مصنوعی توضیح داد که فراتر از مسائل معمول فناوری‌های نوظهور در مورد تأثیر بر ایمنی و رشد اقتصادی است و همچنین بر نظام سیاسی، بازار کار و رفاه تأثیر می‌گذارد. علاوه بر این، پرسش‌های مربوط به دستیابی ماشین‌ها به هوش انسانی یا فراتر رفتن از آن، بازتاب ویژه‌ای در تخیلات جمعی و فردی دارد. همچنین، مجامع بین‌المللی مانند مجمع جهانی اقتصاد و سازمان توسعه و همکاری اقتصادی توجه بیشتری را به سیاست‌های مربوط به هوش مصنوعی جلب کرده‌اند و یادگیری بین‌المللی را در این زمینه تسهیل نموده‌اند. با این حال، توسعه سیاست‌های هوش مصنوعی به طور ناهمگون در سراسر جهان توزیع شده و بر توسعه‌یافته‌ترین مناطق متمرکز است.

هوش مصنوعی بسیاری از معادلات و مفروضاتی که در حال حاضر، دنیای سیاست‌گذاری و اداره‌ی امور عمومی را شکل داده است را به شدت تحت تأثیر قرار خواهد داد. مردم سالاری، دیوان‌سالاری، تخصص و بسیاری دیگر از کلیدواژه‌هایی که شکل‌دهنده‌ی نظام‌های اداری هستند، و با اهدافی مانند گسترش خیر عمومی، اداره‌ی بهتر امور، شایسته‌سالاری و حفظ حاکمیت مردم مردم توسعه پیدا کرده‌اند، تحت تأثیر موضوع هوش مصنوعی باید به صورت جدی مورد بازتعریف قرار گیرند. علاوه بر این، دور از ذهن نیست که در نتیجه‌ی رشد هوش مصنوعی، هر کدام از این پایه‌های مدیریت در بخش عمومی به صورت کلی منسوخ شود. تکنولوژی هوش مصنوعی نه تنها مانند بقیه‌ی تکنولوژی‌ها در دو بعد فوق‌الذکر (کاربردهای هوش مصنوعی و سیاست‌گذاری برای هوش مصنوعی) بر سیاست‌گذاری اثرگذار است، بلکه موضوع جدی و بی‌سابقه‌ی حکمرانی غیربشر بر بشر است (Borgogno, Colangelo, 2019).

تجزیه و تحلیل آینده‌نگرانه و هوش مصنوعی، به دولت‌ها اجازه می‌دهد تا مشکلات احتمالی را قبل از رسیدن به بحران‌ها مورد بررسی قرار دهند. پیشرفت‌های اخیر در پردازش زبان طبیعی، یادگیری ماشین و تشخیص تصویر باعث شده است که دولت‌ها بتوانند به جای مقابله با مشکلات، آن‌ها را پیش‌بینی کنند، از کشف تقلب تا مقابله با بیماری و همه‌گیری مواد افیونی. تجزیه و تحلیل‌های پیش‌بینی‌کننده، اکنون در طیف گسترده‌ای از حوزه‌ها از جمله دفاع، امنیت، مراقبت‌های بهداشتی و خدمات انسانی و همچنین دیگر حوزه‌ها، مورد استفاده قرار می‌گیرد (محمدی و همکاران، ۱۳۹۹).

این پژوهش مطالعه‌ای بر روندها و چارچوب‌های نوظهور سیاست هوش مصنوعی ارائه می‌کند که از سال ۲۰۱۶م به سرعت در سراسر جهان در حال توسعه است. در حالی که آغاز توسعه‌ی هوش مصنوعی را می‌توان حداقل به دهه‌ی ۱۹۵۰م نسبت داد، این فناوری به تازگی به دلیل پیشرفت‌های فناورانه که امکان کاربردهای گسترده‌ای را فراهم کرده است، توجه قابل توجهی از سیاست‌گذاران را به خود جلب کرده است؛ بنابراین، هوش مصنوعی امروزه بسیاری از ویژگی‌های یک فناوری نوظهور را به همراه ابهامات مرتبط، پویایی همکاری و رقابت، و عملکرد اغراق‌آمیز انتظارات مثبت و منفی که بر سیاست‌گذاری در این حوزه تأثیر می‌گذارند، دارا است.

پیشنهادهای زیر برای ایجاد حکمرانی مؤثر هوش مصنوعی ارائه می‌شود:

• تدوین استانداردها و چارچوب‌های تنظیمی مناسب:

○ طراحی قوانین و مقررات شفاف و جامع برای تنظیم استفاده از هوش مصنوعی در حوزه‌های مختلف

○ تعیین معیارهای اخلاقی و حقوقی برای کاربرد و توسعه هوش مصنوعی

○ ایجاد چارچوب‌های تنظیمی برای حفاظت از حریم خصوصی و امنیت داده‌ها

• ایجاد سازوکارهای نظارتی و پاسخ‌گویی:

○ ایجاد نهادهای نظارتی مستقل برای پایش استفاده از هوش مصنوعی

○ تعیین مکانیزم‌های شفاف برای پاسخ‌گویی در قبال آسیب‌های ناشی از هوش

مصنوعی

- ایجاد سازوکارهای رسیدگی به شکایات و جبران خسارات احتمالی
 - توسعه‌ی آموزش و مهارت‌های مرتبط با هوش مصنوعی:
 - ارتقای سواد دیجیتال و آموزش مهارت‌های مورد نیاز در حوزه‌ی هوش مصنوعی برای شهروندان
 - توسعه‌ی برنامه‌های آموزشی و تخصصی برای نیروی کار و متخصصان
 - ایجاد همکاری‌های آموزشی بین دانشگاه‌ها، صنعت و دولت
 - تقویت همکاری‌های بین‌المللی:
 - ایجاد هماهنگی و همکاری بین‌المللی برای تدوین قوانین و استانداردهای فراملی
 - تبادل دانش و تجربیات در زمینه‌ی حکمرانی هوش مصنوعی
 - تقویت ظرفیت‌های نظارتی و اجرایی در سطح بین‌المللی
 - تأکید بر مسئولیت‌پذیری و اخلاق‌محوری:
 - توسعه‌ی چارچوب‌های اخلاقی برای طراحی و استفاده از هوش مصنوعی
 - ایجاد شفافیت و پاسخ‌گویی در توسعه و کاربرد هوش مصنوعی
 - ارزیابی مداوم تأثیرات اجتماعی، اقتصادی و زیست‌محیطی هوش مصنوعی
- اجرای این پیشنهادها می‌تواند به ایجاد حکمرانی مؤثر هوش مصنوعی و ارتقای مسئولیت‌پذیری در قبال این فناوری نوظهور کمک کند.

فهرست منابع

- شیخ‌الاسلامی، خالد. (۱۳۹۸). درآمدی بر حکمرانی هوش مصنوعی. مطالعات سیاسی (گروه بنیادین حکومتی). تهران: مرکز پژوهش‌های مجلس شورای اسلامی.
- محمدی، مهدی؛ کاکاوندی، یوسف؛ شریفیان، امیر؛ محمدی، بهنام. (۱۳۹۹). روندهای حکمرانی هوش مصنوعی. تهران: دانش بنیان فناور.
- Ala-Pietilä, P., & Smuha, N. A. (2021). *A framework for global cooperation on artificial intelligence and its governance*. Reflections on artificial intelligence for humanity, pp. 237-265.
- Borgogno, O., & Colangelo, G. (2019). *Data sharing and interoperability: Fostering innovation and competition through APIs*. Computer Law & Security Review, 35(5), p. 105314.
- Borrás, S., & Edler, J. (Eds.). (2014). *The governance of socio-technical systems: explaining change*. Edward Elgar Publishing.
- Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? *Journal of Econometrics* 65(1), pp. 83-108.
[https://doi.org/10.1016/0304-4076\(94\)01598-T](https://doi.org/10.1016/0304-4076(94)01598-T).
- Carabantes, M. (2020). Black-box artificial intelligence: an epistemological and critical analysis. *AI & Soc* (35), pp. 309-317. <https://doi.org/10.1007/s00146-019-00888-w>
- Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the 'good society': the US, EU, and UK approach. *Science and engineering ethics*, 24, pp. 505-528.
- Chhotray, V., & Stoker, G. (2009). *Governance theory and practice*. A cross-disciplinary approach. London: Palgrave Macmillan.

- Cihon, P., Maas, M. M., & Kemp, L. (2020). Fragmentation and the Future: Investigating Architectures for International AI Governance. *Global Policy*, 11(5), pp. 545–556.
- Cotra, A. (2020). *Draft report on AI timelines*. LessWrong. Retrieved from: <https://www.lesswrong.com/posts/KrJfoZzpSDpnr9va/draft-report-on-ai-timelines>.
- Dafoe, A. (2018). *AI governance: A research agenda*; future of humanity institute. Oxford, UK: University of Oxford.
- De Saille, S. (2015). Innovating innovation policy: the emergence of ‘Responsible Research and Innovation’. *Journal of Responsible Innovation*, 2(2), pp. 152-168.
- Dignum, V. (2019). *Responsible artificial intelligence: how to develop and use AI in a responsible way* (Vol. 1). Cham: Springer.
- Ding, J., & Dafoe, A. (2021, June). *Engines of power: Electricity, AI, and general-purpose military transformations*. arXiv.
- European Commission. (2018a). *Artificial intelligence for Europe*. Communication COM(2018) 237. Brussels: European Commission.
- European Commission. (2018b). *Artificial intelligence: A European perspective*. Luxembourg: Publications Office of the European Union.
- Fernandez-Cortez, V., Valle-Cruz, D., & Gil-Garcia, J. R. (2020). *Can artificial intelligence help optimize the public budgeting process?* Lessons about smartness and public value from the Mexican federal government. 2020 Seventh International Conference on EDemocracy & EGovernment (ICEDEG), pp. 312–315. <https://doi.org/10.1109/ICEDEG48599.2020.9096745>.
- Galanos, V. (2019). Exploring expanding expertise: artificial intelligence as an existential threat and the role of prestigious commentators, 2014–2018. *Technology Analysis & Strategic Management*, 31(4), pp. 421-432.
- Garfinkel, B. (2022). *The impact of artificial intelligence: A historical perspective*. In J. Bullock (Ed.), *The Oxford Handbook of AI Governance*. Oxford University Press.

- Gasser, U., & Almeida, V. A. (2017). A layered model for AI governance. *IEEE Internet Computing*, 21(6), pp. 58–62.
- Gritsenko, D., & Wood, M. (2020). Algorithmic governance: A modes of governance approach. *Regulation & Governance*, 6(1), pp. 45–62.
- Guihot, M., Matthew, A. F., & Suzor, N. P. (2017). Nudging robots: Innovative solutions to regulate artificial intelligence. *Vand. J. Ent. & Tech. L*, 20, 385.
- Hemphill, T. A. (2016). Regulating nanomaterials: A case for hybrid governance. *Bulletin of Science, Technology & Society*, 36(4), pp. 219–228.
- Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organising data for trustworthy artificial intelligence. *Government Information Quarterly*, 37(3), pp. 101493.
- Katzenbach, C., & Ulbricht, L. (2019). Algorithmic governance. *Internet Policy Review* 8(4), pp. 1–18.
- Kuhlmann, S., Stegmaier, P., & Konrad, K. (2019). The tentative governance of emerging science and technology—A conceptual introduction. *Research policy*, 48(5), pp. 1091–1097.
- Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications policy*, 44(6), 1 p. 01976.
- Lee, K. F. (2018). AI superpowers: China, Silicon Valley, and the new world order. Houghton Mifflin.
- Li, Y., Taeihagh, A., De Jong, M., & Klinke, A. (2021). Toward a commonly shared public policy perspective for analysing risk coping strategies. *Risk analysis*, 41(3), pp. 519–532. <https://doi.org/10.1111/risa.13505>
- Mazzucato, M. (2021). *Mission economy: A moonshot guide to changing capitalism*. London: Allen Lane.
- Mikhaylov, S. J., Esteve, M., & Campion, A. (2018). Artificial intelligence for the public sector: opportunities and challenges of cross-sector collaboration. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 376(2128).

- North, D. C. (1991). Institutions. *Journal of Economic Perspectives* 5(1), pp. 97–112.
<http://www.jstor.org/stable/1942704>.
- Rabesandratana, T. (2018). “Emmanuel Macron wants France to become a leader in AI and avoid ‘dystopia.’” *Science* (30). <https://doi.org/10.1126/science.aat7491>
- Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology*, 20(1), pp. 5–14.
- RT. (1 September 2017). “Whoever leads in AI will rule the world’: Putin to Russian children on Knowledge Day.” *Retrieved from*:
<https://www.rt.com/news/401731-ai-rule-world-putin/>.
- Schwab, K. (2017). The fourth industrial revolution. Crown Currency.
- Somers, J. (2017, April). *Torching the modern-day library of Alexandria*. The Atlantic. <https://www.theatlantic.com/technology/archive/2017/04/the-tragedy-of-google-books/523320/>
- Stilgoe, J., Owen, R., & Macnaghten, P. (2020). Developing a framework for responsible innovation. In *The Ethics of Nanotechnology, Geoengineering, and Clean Energy* (pp. 347-359). Routledge.
- Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, 36(2), pp. 368-383.
- Taeihagh, A., & Lim, H. S. M. (2019). Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transport reviews*, 39(1), pp. 103-128.
- Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. In *Regulation & Governance*. <https://doi.org/10.1111/rego.12392>
- Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4).

- Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly* 37(4), 101509.
- WEF. (25 January 2018). "Theresa May's Davos address in full." Retrieved from: <https://www.weforum.org/agenda/2018/01/theresa-may-davos-address/>.
- Wirtz, B. W., Weyerer, J. C., & Sturm, B. J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration* 43(9), pp. 818–829.
- Young, M. M., Bullock, J. B., & Lecy, J. D. (2019). Artificial discretion as a tool of governance: a framework for understanding the impact of artificial intelligence on public administration. *Perspectives on Public Management and Governance* 2(4), pp. 301–313.

References

- Sheikh al-Islami, K. (2019). *An introduction to AI governance*. Political Studies (Fundamental Governance Group). Tehran: Islamic Parliament Research Center. [In Persian]
- Mohammadi, M., Kakavandi, Y., Sharifian, A., & Mohammadi, B. (2020). *Trends in AI governance*. Tehran: Danesh Bonyan Fanavar. [In Persian]
- Ala-Pietilä, P., & Smuha, N. A. (2021). A framework for global cooperation on artificial intelligence and its governance. *Reflections on Artificial Intelligence for Humanity*, pp. 237-265.
- Borgogno, O., & Colangelo, G. (2019). Data sharing and interoperability: Fostering innovation and competition through APIs. *Computer Law & Security Review*, 35(5), 105314.
- Borrás, S., & Edler, J. (Eds.). (2014). *The governance of socio-technical systems: explaining change*. Edward Elgar Publishing.
- Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? *Journal of Econometrics*, 65(1), pp. 83-108. [https://doi.org/10.1016/0304-4076\(94\)01598-T](https://doi.org/10.1016/0304-4076(94)01598-T).
- Carabantes, M. (2020). Black-box artificial intelligence: an epistemological and critical analysis. *AI & Soc*, 35, pp. 309-317. <https://doi.org/10.1007/s00146-019-00888-w>
- Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the 'good society': The US, EU, and UK approach. *Science and Engineering Ethics*, 24, pp. 505-528.
- Chhotray, V., & Stoker, G. (2009). *Governance theory and practice: A cross-disciplinary approach*. London: Palgrave Macmillan.
- Cihon, P., Maas, M. M., & Kemp, L. (2020). Fragmentation and the future: Investigating architectures for international AI governance. *Global Policy*, 11(5), pp. 545-556.

- Cotra, A. (2020). Draft report on AI timelines. *LessWrong*. Retrieved from: <https://www.lesswrong.com/posts/KrJfoZzpSDpnr9va/draft-report-on-ai-timelines>.
- Dafoe, A. (2018). *AI governance: A research agenda*. Future of Humanity Institute. Oxford, UK: University of Oxford.
- De Saille, S. (2015). Innovating innovation policy: The emergence of 'Responsible Research and Innovation'. *Journal of Responsible Innovation*, 2(2), pp. 152-168.
- Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way* (Vol. 1). Cham: Springer.
- Ding, J., & Dafoe, A. (2021, June). Engines of power: Electricity, AI, and general-purpose military transformations. *arXiv*.
- European Commission. (2018a). *Artificial intelligence for Europe*. Communication COM(2018) 237. Brussels: European Commission.
- European Commission. (2018b). *Artificial intelligence: A European perspective*. Luxembourg: Publications Office of the European Union.
- Fernandez-Cortez, V., Valle-Cruz, D., & Gil-Garcia, J. R. (2020). Can artificial intelligence help optimize the public budgeting process? Lessons about smartness and public value from the Mexican federal government. *2020 Seventh International Conference on EDemocracy & EGovernment (ICEDEG)*, 312–315. <https://doi.org/10.1109/ICEDEG48599.2020.9096745>.
- Galanos, V. (2019). Exploring expanding expertise: Artificial intelligence as an existential threat and the role of prestigious commentators, 2014–2018. *Technology Analysis & Strategic Management*, 31(4), pp. 421-432.
- Garfinkel, B. (2022). The impact of artificial intelligence: A historical perspective. In J. Bullock (Ed.), *The Oxford Handbook of AI Governance*. Oxford University Press.
- Gasser, U., & Almeida, V. A. (2017). A layered model for AI governance. *IEEE Internet Computing*, 21(6), pp. 58–62.

- Gritsenko, D., & Wood, M. (2020). Algorithmic governance: A modes of governance approach. *Regulation & Governance*, 6(1), pp. 45–62.
- Guihot, M., Matthew, A. F., & Suzor, N. P. (2017). Nudging robots: Innovative solutions to regulate artificial intelligence. *Vand. J. Ent. & Tech. L*, 20, 385.
- Hemphill, T. A. (2016). Regulating nanomaterials: A case for hybrid governance. *Bulletin of Science, Technology & Society*, 36(4), pp. 219–228.
- Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organising data for trustworthy artificial intelligence. *Government Information Quarterly*, 37(3), 101493.
- Katzenbach, C., & Ulbricht, L. (2019). Algorithmic governance. *Internet Policy Review*, 8(4), 1–18.
- Kuhlmann, S., Stegmaier, P., & Konrad, K. (2019). The tentative governance of emerging science and technology—A conceptual introduction. *Research Policy*, 48(5), pp. 1091–1097.
- Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, 44(6), 101976.
- Lee, K. F. (2018). *AI superpowers: China, Silicon Valley, and the new world order*. Houghton Mifflin.
- Li, Y., Taeihagh, A., De Jong, M., & Klinke, A. (2021). Toward a commonly shared public policy perspective for analysing risk coping strategies. *Risk Analysis*, 41(3), pp. 519–532. <https://doi.org/10.1111/risa.13505>
- Mazzucato, M. (2021). *Mission economy: A moonshot guide to changing capitalism*. London: Allen Lane.
- Mikhaylov, S. J., Esteve, M., & Campion, A. (2018). Artificial intelligence for the public sector: Opportunities and challenges of cross-sector collaboration. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 376(2128).

- North, D. C. (1991). Institutions. *Journal of Economic Perspectives*, 5(1), pp. 97–112.
<http://www.jstor.org/stable/1942704>.
- Rabesandratana, T. (2018). Emmanuel Macron wants France to become a leader in AI and avoid ‘dystopia.’ *Science*, 30. <https://doi.org/10.1126/science.aat7491>
- Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology*, 20(1), pp. 5–14.
- RT. (1 September 2017). “Whoever leads in AI will rule the world”: Putin to Russian children on Knowledge Day. Retrieved from: <https://www.rt.com/news/401731-ai-rule-world-putin/>.
- Schwab, K. (2017). *The fourth industrial revolution*. Crown Currency.
- Somers, J. (2017, April). Torching the modern-day library of Alexandria. *The Atlantic*. <https://www.theatlantic.com/technology/archive/2017/04/the-tragedy-of-google-books/523320/>
- Stilgoe, J., Owen, R., & Macnaghten, P. (2020). Developing a framework for responsible innovation. In *The Ethics of Nanotechnology, Geoengineering, and Clean Energy* (pp. 347-359). Routledge.
- Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of Artificial Intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly*, 36(2), pp. 368-383.
- Taeihagh, A., & Lim, H. S. M. (2019). Governing autonomous vehicles: Emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transport reviews*, 39(1), pp. 103-128.
- Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. *Regulation & Governance*. <https://doi.org/10.1111/rego.12392>
- Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4).

Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly*, 37(4), 101509.

World Economic Forum. (2018, January 25). Theresa May's Davos address in full. Retrieved from <https://www.weforum.org/agenda/2018/01/theresa-may-davos-address/>

Wirtz, B. W., Weyerer, J. C., & Sturm, B. J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration*, 43(9), pp. 818–829.

Young, M. M., Bullock, J. B., & Lecy, J. D. (2019). Artificial discretion as a tool of governance: A framework for understanding the impact of artificial intelligence on public administration. *Perspectives on Public Management and Governance*, 2(4), pp. 301–313.