



New Media and Cognitive Agency: Exploring the Impact of Artificial Intelligence on Political and Social Agency

Razieh Mehrabi Koushki 

Assistant Professor, Faculty member of the Department of International Studies in
Science and Technology, Technology Studies Institute, Tehran, Iran.

Email: mehrabi@tsi.ir

Article Info

ABSTRACT

Article type:

Research Article

Article history:

Received:

12 September 2024

Revised version

received:

23 October 2024

Accepted:

30 October 2024

Available online:

5 April 2025

Keywords:

Artificial Intelligence,

Democracy,

Cognitive Bubbles,

Misinformation,

Social Robots,

Social Media

Objective

Democracy relies on the premise that citizens participate in civil and political affairs based on informed and deliberate will, and democratic systems are tasked with safeguarding this will. However, the rise of social networks, powered by artificial intelligence technologies, has disrupted democratic order by facilitating the production and dissemination of false, fake, and biased information. The integration of AI into social media has exacerbated a critical issue: the political and civil fate of societies is increasingly shaped by the will of AI algorithms rather than the informed judgment of citizens. In a democratic system, decisions should reflect the public's calculated and reasoned judgment, not their impulsive or manipulated reactions. This study addresses a fundamental question: "How does artificial intelligence, operating within social networks, influence political-social agency and manipulate public knowledge? By drawing on existing theoretical and experimental studies and analyzing key themes, we aim to explore this understudied and emerging phenomenon. Our findings reveal that AI algorithms manipulate citizens' control over their political and civic beliefs and behaviors through several strategies, including gradual cognitive damage, cognitive hacking, erosion of individual knowledge, amplification of emotional reasoning, and creation of cognitive distance. To mitigate these effects, it is essential to empower citizens' cognitive agency by fostering self-awareness and critical thinking skills. By doing so, control over the cognitive environment can be partially reclaimed from AI-driven social media platforms, ensuring a more informed and autonomous public.

Cite this article: Mehrabi Koushki, Rezieh (2025). "New Media and Cognitive Agency:

Exploring the Impact of Artificial Intelligence on Political and Social Agency", *Fasl'nāmeḥ-ye siyāsat (Politics Quarterly)* 55, (1): 245-273,

DOI: <http://doi.org/10.22059/JPQ.2025.382279.1008210>



The Author(s).

Publisher: University of Tehran Press.

<http://doi.org/10.22059/JPQ.2025.382279.1008210>

Introduction

Knowledge serves as the foundation of culture, history, and a wide range of human actions, including the formation of moral judgments, choices, and decisions in both individual and social life. In a democratic system, the fundamental assumption is that citizens participate in civil and political affairs based on this knowledge, and democratic institutions are tasked with facilitating the natural realization of collective consciousness and will in societal and political spheres. However, rapid technological advancements, particularly in artificial intelligence (AI), have introduced an intelligent yet non-human agent that profoundly impacts various aspects of human life, including how individuals understand and interpret the world around them. AI is unique in its ability to mimic and perform functions of human intelligence, such as reasoning, problem-solving, discovering meaning, generalizing, learning from past experiences, and making targeted decisions by identifying hidden patterns, rules, and relationships within data. It can also anticipate future trends, making it a powerful tool for influencing human behavior. With its integration into modern social media platforms, AI has disrupted democratic systems by enabling the production and dissemination of false, fake, and biased information. The presence of AI-driven technologies in social media has intensified a critical issue: the political and civic fate of societies is increasingly determined by the will and algorithms of artificial intelligence rather than the deliberate thought and will of individuals. This undermines the democratic ideal, which relies on the public's calculated and reasoned judgment, not their impulsive or manipulated reactions.

Today, the adverse effects of artificial intelligence (AI) on public goods—such as justice, social equality, human rights, freedom of expression, and, more broadly, democracy and civil engagement—have emerged as pressing challenges for experts and sociologists to understand and address. These issues highlight the growing influence of AI in shaping societal structures and individual behaviors.

This research focuses on a central question: How does a technological factor like AI, operating within the context of social networks, influence people's political and social awareness and, ultimately, determine their behavior? In other words, how do intelligent algorithms shape socio-political beliefs, justify certain choices and decisions, and invalidate others? By addressing this question, the study aims to uncover how AI technology, embedded in social media platforms, undermines human "agency" in receiving, evaluating, and interpreting information. Over time, AI itself becomes the dominant agency, steering public opinion and decision-making processes.

Methodology

To address the above question, this research adopts a qualitative approach, collecting and analyzing data from a range of sources. Using the library and documentary method, the study compiles the latest theoretical and experimental findings from the past eight years. Through content analysis, it identifies and evaluates the most significant cognitive effects of artificial intelligence on political-social agency, supported by available evidence. Finally, the research offers practical suggestions and solutions aimed at restoring human agency in political and social life, thereby contributing to the reinvigoration of democratic order.

Results and Discussion

Artificial intelligence serves as a powerful tool for both malicious government agencies and foreign actors seeking to infiltrate and disrupt. It targets political-social agency on social media through various strategies, undermining individuals' independence in decision-making and taking control of their rational, internal, and measured processes of activism. Below, we discuss the most significant strategies identified in this study:

1. Gradual Cognitive Damage:

Smart technologies in the digital space function like a "digital Pavlov," conditioning human behavior and agency through free, diverse, and highly engaging services. This process rapidly captures users' attention, shapes their worldview, and earns their trust. Prolonged exposure to these platforms can make individuals more compliant, susceptible to external persuasion, and prone to uncritical acceptance of information.

2. Cognitive Hacking Through Micro-Profiling:

AI algorithms collect personal data by identifying emotional vulnerabilities, fears, political preferences, and social interests. Using this information, they tailor and deliver political messages aligned with individuals' personality types. This enables AI to target undecided or swing voters, producing and disseminating specific information to sway their decisions in favor of a particular political agenda.

3. Epistemic Bubbles:

Artificial intelligence categorizes users into homogeneous groups through homogenizing algorithms, delivering information that aligns with their existing intellectual and value frameworks. While this may seem like personalized content, it effectively censors or restricts access to diverse perspectives. Users become trapped in "filter bubbles" or information caves, where they only encounter echoes of their own beliefs and approved thoughts, limiting exposure to alternative viewpoints.

4. Erosion of Individual Knowledge and Trust:

AI can undermine trust in an individual's cognitive capacities, gradually eroding their reasoning and critical thinking skills. Over-reliance on AI-driven information can lead individuals to forget that agency lies in their own knowledge and awareness, with technology serving merely as a tool for efficiency rather than a replacement for independent thought.

5. Amplification of Emotional Knowledge:

Exposure to a high volume of false, hateful, or fake news and information can lead to the overflow of emotional behaviors into offline actions, fostering external conflicts. In this environment, knowledge becomes intertwined with exaggerated emotions, undermining critical rationality and progressive politics, which are essential for a well-informed and balanced society.

6. Creation of False Epistemic Consensus and Minority Dominance:

Certain AI mechanisms, such as astroturfing, create platforms where fringe views are amplified to appear mainstream, dictating values and choices to the majority. This distorts the worldview that serves the interests of most people, pushing it into the background or dismantling it entirely. The result is a disoriented society led by unstable and superficially competent individuals. This cognitive disorder is further exacerbated by social and political bots, as well as false campaigns, which manipulate public perception and decision-making.

7. Creating Epistemic Distance from Socio-Political Reality:

Epistemic distance arises when information technologies create a gap between reality and cognitive agency. In this process, data passes through multiple intermediaries before reaching individuals, creating a space uniquely suited for constructing and maintaining a fabricated world. AI exacerbates this issue by continuously producing and republishing false content, such as deepfake videos, targeting rival political groups. The situation becomes particularly dire when AI-generated content is so close to the truth that distinguishing it from reality becomes nearly impossible. This blurring of lines between fact and fiction further erodes trust in information and undermines the ability of individuals to engage meaningfully with socio-political realities.

Conclusion

New media technologies, with their complex technical mechanisms, possess the ability to control and direct people's will and choices, shaping their understanding and knowledge of social and political realities—often without their awareness. This study observed that artificial intelligence manipulates human agency through political bots, echo chambers, deepfakes, and targeted misinformation, effectively controlling individuals' thinking, reasoning, and decision-making in political matters using various strategies. To address these challenges, the most critical and effective solution is to strengthen human agency by fostering analytical and critical thinking skills. Equipping individuals with the ability to discern and evaluate information independently can mitigate the influence of AI-driven manipulation.

رسانه‌های نوین و عاملیت معرفتی: بررسی تأثیر هوش مصنوعی بر عاملیت سیاسی - اجتماعی

راضیه مهرابی کوشکی 

استادیار، مطالعات و همکاری‌های بین‌المللی علم و فناوری، پژوهشکده مطالعات فناوری، تهران، ایران.

رایانامه: mehrabi@tsi.ir

اطلاعات مقاله چکیده

نوع مقاله:

پژوهشی

تاریخ دریافت:

۱۴۰۳/۰۶/۲۲

تاریخ بازنگری:

۱۴۰۳/۰۸/۲

تاریخ پذیرش:

۱۴۰۳/۰۸/۹

تاریخ انتشار:

۱۴۰۴/۰۱/۱۶

کلیدواژه‌ها:

هوش مصنوعی،
مردم‌سالاری،
حباب‌های معرفتی،
الگوریتم‌های تولید محتوا،
اطلاعات نادرست،
ربات‌های اجتماعی،
رسانه‌های اجتماعی

مردم‌سالاری بر این فرض بنیادین بنا شده که شهروندان بر اساس نوعی آگاهی که مبتنی بر اراده و خواست ایشان است، در امور مدنی و سیاسی مشارکت می‌کنند و نظام‌های مردم‌سالار وظیفه‌دارند به تحقق اراده جمعی در جامعه و سیاست کمک کنند. امروزه رسانه‌های نوین به‌ویژه شبکه‌های اجتماعی به کمک فناوری هوش مصنوعی به تولید و توزیع اطلاعات نادرست، جعلی و مغرضانه می‌پردازند و نظم دموکراتیک را دچار بحران کرده‌اند. به بیان دیگر مداخله و حضور فناوری‌های هوشمند در رسانه‌های اجتماعی به این مسئله بنیادین دامن زده که این اراده و انتخاب الگوریتم‌های هوش مصنوعی است که سرنوشت سیاسی و مدنی جوامع را تعیین می‌کند نه آگاهی و اراده شهروندان. درحالی‌که مشروعیت نظم دموکراتیک مشروط بر این است که بازتاب دآوری حساب‌شده و سنجیده عموم مردم و نه واکنش سرسری و غیرارادی آنان باشد. پرسش اساسی این مقاله این است که هوش مصنوعی در بستر رسانه‌های نوین به‌ویژه شبکه‌های اجتماعی چگونه بر عاملیت سیاسی - اجتماعی اثر می‌گذارد و اراده و معرفت سیاسی - اجتماعی او را دست‌کاری می‌کند؟ در این زمینه با مراجعه با اطلاعات پراکنده موجود در ادبیات نظری و تجربی و احصای مهم‌ترین مضامین مرتبط، تلاش می‌شود این اثرگذاری کمتر بررسی‌شده به شکل یک الگو بازنمایی شود. بررسی کیفی داده‌ها نشان داد الگوریتم‌های هوش مصنوعی در رسانه‌های اجتماعی از طریق راهبردهای متنوعی، کنترل شهروندان بر معرفت، باورها و رفتارهای سیاسی و مدنی‌شان را دست‌کاری می‌کند که مهم‌ترین آن عبارت است از: آسیب معرفتی تدریجی، هک شناختی، فرسایش معرفت فردی، تضعیف معرفت علّی، افزایش معرفت احساسی و ایجاد فاصله معرفتی. به‌نظر می‌رسد لازمه کاستن از این اثرگذاری، فعال کردن عاملیت معرفتی شهروندان از طریق خودآگاهی آنان و آموزش تفکر انتقادی است، به‌نحوی که یک نظم دموکراتیک درونی و پایدار در افراد ایجاد شده و کنترل محیط معرفتی تا حدی از رسانه‌های اجتماعی هوشمند خارج شود.

استناد: مهرابی کوشکی، راضیه (۱۴۰۴). رسانه‌های نوین و عاملیت معرفتی: بررسی تأثیر هوش مصنوعی بر عاملیت سیاسی

- اجتماعی، فصلنامه سیاست، ۵۵ (۱)، ۲۷۳-۲۴۵.

<http://doi.org/10.22059/JPQ.2025.382279.1008210>



نویسندگان.

ناشر: انتشارات دانشگاه تهران.

۱. مقدمه

دانش و آگاهی نه تنها زیربنای فرهنگ، تاریخ و دامنه وسیعی از اعمال انسانی است بلکه مبنای شکل‌گیری قضاوت‌های اخلاقی، انتخاب‌ها و تصمیمات بشر در زندگی فردی و اجتماعی محسوب می‌شود. این دانش و آگاهی و اعمال وابسته آن، هسته اصلی دانشی است که ما را انسان تعریف می‌کند. شناخت‌شناسی یا «معرفت‌شناسی» نه تنها به صورت بندی دانش و ابزار کسب، انتشار و حفظ آن کمک می‌کند بلکه زیربنای فهم و درک انسان را به عنوان «عاملیت»^۱ اخلاقی، فرهنگی، سیاسی و اجتماعی آماده کرده و نقش رقابتی «عاملیت انسانی» را نیز تبیین می‌کند (Reider, 2016). از سوی دیگر امروزه پیشرفت‌های پارادایمی و نوظهور در حوزه فناوری‌های ارتباطی و اطلاعاتی، سبب شده فناوری به عنوان یک «عاملیت غیرانسانی» اما هوشمند، بر حوزه‌های مختلف زندگی بشر از جمله فهم و شناخت او از جهان پیرامونی تأثیر بنیادین و برافکن داشته باشد. این مهم در خصوص هوش مصنوعی (AI) ویژه است، زیرا به کمک فناوری‌هایی چون یادگیری ماشین و یادگیری عمیق (Goodfellow et al., 2016; Bishop, 2006) قادر است عملکردهایی از هوش انسانی مانند استدلال، حل مسئله، کشف معنا و تعمیم را تقلید و اجرا کند. به نحوی که تا حدی مستقل از انسان و در شرایط متغیر، می‌تواند از تجارب گذشته و دانش در دسترس بیاموزد و با شناسایی الگوها، قوانین و روابط پنهان در داده‌ها تصمیمات هدفمند بگیرد و روندهای آتی را پیش‌بینی کند (Copeland, 2023; Nilsson, 1998).

AI در بستر رسانه‌های نوین اجتماعی در عملکردهایی همچون بینایی و گفتار به انسان نزدیک شده و می‌تواند با پردازش زبان طبیعی انسان، به یادگیری و حل مسئله بپردازد. همچنین می‌تواند با تحلیل کلان داده‌های موجود که ناشی از کنشگری افراد در شبکه‌های اجتماعی و ردپای دیجیتال آنهاست، ذائقه فردی و اجتماعی ایشان را شناسایی و متناسب با پروفایل روان‌شناختی و رفتاری او توصیه‌های شخصی‌سازی شده و منحصر به فرد ارائه کند (Buhmann & Fieseler, 2023). اما این ویژگی‌های هوشمند لزوماً تأثیرات مثبت به همراه ندارد. برای نمونه به دلیل وجود سوگیری در داده‌های آموزشی که الگوریتم‌های هوشمند دریافت می‌کند، می‌تواند توصیه‌ها، تصمیم‌گیری‌ها و اقدامات آن توأم با پیامدهای ناخواسته، منفی و مخرب باشد (Crawford & Calo, 2016; Martin, 2019). در حوزه زندگی اجتماعی، حقوق بشر، آزادی بیان و نامطلوب گسترده‌ای بر کالاهای عمومی چون عدالت، برابری اجتماعی، حقوق بشر، آزادی بیان و در اینجا بر دموکراسی و انتخابات داشته باشد و به طور ویژه فرایندهای شناختی مربوط به انتخابات عادلانه، آزاد و دموکراتیک را تضعیف و دست‌کاری کند (Eubanks, 2018: 142).

1. agent

چگونگی اثرگذاری هوش مصنوعی بر عاملیت انسانی و اراده و ارزیابی سنجیده و درونی او در موقعیت‌های اجتماعی و سیاسی چندان شفاف نیست و پژوهش‌های موجود، هریک بر بخشی از آن تمرکز داشته‌اند. یافتن راهبردهای اثرگذاری AI به‌ویژه در بستر رسانه‌های نوین اجتماعی و چگونگی مداخله آن در نظم دموکراتیک جوامع و نحوه ایجاد اختلال در روند مطلوب شکل‌گیری آگاهی و تعاملات مردم‌سالارانه، موضوع این پژوهش است. به‌عبارت دیگر سؤال اصلی این است که چگونه یک عامل فناورانه چون AI در بستر شبکه‌های اجتماعی، بر آگاهی سیاسی و اجتماعی افراد اثر می‌گذارد و رفتار آنها را تعیین می‌کند؟ در چارچوب معرفت‌شناسی سؤال این است که چگونه الگوریتم‌های هوشمند، باورهای سیاسی اجتماعی را می‌سازند و یک انتخاب و تصمیم اجتماعی را موجه و دیگری را بی‌اعتبار می‌کنند؟ پرداختن به این سؤال نشان می‌دهد که فناوری AI در بستر رسانه‌های اجتماعی خود به عاملیت اصلی تبدیل می‌شود و «عاملیت» انسانی را در دریافت و ارزیابی اطلاعات و آگاهی تضعیف می‌کند. برای پاسخ به این پرسش، با مراجعه به داده‌های ثانویه انباشته و پراکنده در مطالعات هشت سال اخیر و مرور نظام‌مند آن، تلاش می‌شود الگوی این اثرگذاری احصا و تبیین و در یک نظم بصری ارائه شود.

۲. روش پژوهش

روش این پژوهش کیفی است و از شیوه مطالعه مروری و تحلیل محتوا برای جمع‌آوری و تحلیل داده‌ها استفاده شده است. با توجه به ماهیت پیچیده تأثیر هوش مصنوعی بر معرفت انسانی تلاش شد تحلیل‌های نظری و تجربی موجود، به‌صورت نظام‌مند مرور شود تا از بین داده‌های اغلب پراکنده یک الگوی تبیینی در بیان راهبردهای اثرگذاری AI بر عاملیت سیاسی- اجتماعی احصا شود. در این روش، پژوهشگر به‌جای گردآوری داده‌های جدید، از داده‌های موجود استفاده می‌کند (تحلیل ثانویه) تا ضمن پاسخ به پرسش تحقیق، یافته‌های پژوهش‌های قبلی را در یک نظم جدید بازنمایی کند. در این زمینه و برای جمع‌آوری داده‌های نظری و تجربی انجام‌گرفته، یک جست‌وجوی نظام‌مند در پایگاه داده معتبر گوگل اسکولار انجام گرفت و مقالات مرتبط در بازه زمانی ۲۰۱۶ تا ۲۰۲۳ بررسی شد. سپس با استفاده از روش تحلیل محتوای کیفی، داده‌های استخراج‌شده کدگذاری و تحلیل شد تا مضامین محوری و الگوی اصلی به‌دست آید. در روند تحلیل محتوا و کدگذاری مضامین اصلی، تلاش شد کدهای انتخاب شود که فضای معنایی راهبرد موردنظر را از منظر معرفت‌شناسی پوشش دهد. در نهایت پس از تبیین و توصیف راهبردهای اثرگذاری هوش مصنوعی بر عاملیت انسانی، نتیجه به شکل یک الگوی بصری بازنمایی شد. نوآوری انجام‌گرفته در این پژوهش نظم بخشیدن به اطلاعات پراکنده در تبیین

تأثیرپذیری عاملیت انسانی در هوش مصنوعی در بستر مشارکت سیاسی و اجتماعی و احصای یک مدل تبیینی برای انجام مطالعات بیشتر در کشور است.

شایان ذکر است در انتخاب منابع مورد بررسی، بر منابعی تمرکز شد که به صورت تجربی و نظری، بر تأثیر هوش مصنوعی بر دموکراسی با توجه به مقوله عاملیت معرفتی و سیاسی-اجتماعی توجه داشتند. در این زمینه با ترکیب کلید واژگان دموکراسی، هوش مصنوعی، عاملیت معرفتی، عاملیت سیاسی و اجتماعی این منابع احصا و موارد غیرمرتبط حذف شد. در ضمن برای ایجاد عمق بیشتر در تحقیق، به منابع مورد استفاده در منابع مذکور نیز مراجعه شد. همچنین فرایند تحقیق و جمع آوری منابع تا زمان اشباع داده‌ها و رسیدن به مدل تبیینی ادامه داشت. در مجموع همان‌طور که بیان شد، منابع تجربی متعددی وجود دارند که تأثیرات فنی هوش مصنوعی را رصد تجربی و علمی کرده‌اند، لیکن موارد محدودی از منظر عاملیت معرفتی وارد این موضوع شده‌اند.

۳. پیشنهاد

نگرانی در زمینه پیامدها و تهدیدات فناوری هوش مصنوعی برای جوامع مردم‌سالار و جامعه مدنی قریب هشت سال است که در سطح مطالعات سیاسی و اجتماعی مطرح شده و از انتخابات ۲۰۱۹ در آمریکا همه‌پرسی بر گزیت ۲۰۱۶ انگلستان تاکنون رو به فزونی نهاده است. در این زمینه اغلب پژوهش‌های خارجی تأثیرات هوش مصنوعی بر مردم‌سالاری را با توجه به راهبردهای فنی و تأثیرات الگوریتم‌های مختلف بررسی کرده‌اند. برای نمونه هائو (۲۰۱۹) توضیح می‌دهد چگونه هوش مصنوعی از طریق فناوری «فیلترهای حباب» دموکراسی را تهدید می‌کند (Hao, 2019). مارتین-روزومیلوویچ و کوزل (۲۰۱۹) نیز در تحقیقی نشان می‌دهند چگونه فناوری ربات‌های اجتماعی نامزدهای انتخاباتی زن را حذف قرار می‌دهند و با تولید اطلاعات مغرضانه و جعلی سبب کاهش مشارکت سیاسی و اجتماعی آنان می‌شوند (Martin-Rozumiłowicz & Kužel, 2019). همچنین فناوری جعل عمیق با محتوای جنسی در رسانه‌های اجتماعی می‌تواند برای ایجاد شرمساری و جلوگیری از مشارکت سیاسی و مدنی زنان استفاده شود (McGonagle et al., 2019). باربارا (۲۰۲۰) نیز نشان می‌دهد چگونه طرفداران افراطی تحت فناوری اتاقک‌های پژواک که نوعی فناوری هوشمند خوشه‌بندی مخاطبان در شبکه‌های اجتماعی است، یکدیگر را تقویت می‌کنند و قطبی‌سازی در فضای انتخاباتی را افزایش می‌دهند و در کنار آن اطلاعات و احساسات علیه و بر ضد مقامات رسمی افزایش می‌یابد (Barberá, 2020). اینها نمونه‌هایی از کوه یخی است که در کمتر از ده سال در حال انباشت تبیین علمی تأثیرات هوش مصنوعی بر دموکراسی است، اما غالب این پژوهش‌ها به تبیین‌های

فنی روی آورده و کمتر از منظر عاملیت معرفتی بر این موضوع متمرکز شده‌اند. در مجموع می‌توان منابع زیر را به‌عنوان مهم‌ترین و جدیدترین مراجع تبیین تأثیر هوش مصنوعی بر عاملیت سیاسی - اجتماعی دانست.

جدول ۱: مهم‌ترین منابع تبیین کننده تأثیر هوش مصنوعی بر عاملیت سیاسی - اجتماعی

نام نویسنده	مسئله پژوهش	محورهای اصلی در تبیین موضوع
کوکلیبرگ (۲۰۲۳)	هوش مصنوعی چگونه عاملیت سیاسی و دموکراسی‌ها را به خطر می‌اندازد.	بررسی تأثیر اخبار جعلی، تجزیه و تحلیل داده‌های بزرگ، حباب‌های معرفتی، نتایج ساده دانش آماری، تضعیف اعتماد به ظرفیت‌های معرفتی فرد
رزنبرگ (۲۰۲۳)	تاکتیک‌های دست‌کاری و کنترل معرفت انسانی توسط هوش مصنوعی	بررسی تأثیر هدفمند و شخصی‌سازی‌شده، تکنیک‌های متقاعدسازی و ایجاد اعتماد
کرزانوسکی و پولاک (۲۰۲۲)	چگونگی اثرگذاری اینترنت به‌عنوان یک عاملیت غیرانسانی بر جامعه	بررسی تأثیر بر جهان‌بینی، هستی‌شناختی، فاصله معرفتی با حقایق، خودمختاری معرفتی
گان و لینچ (۲۰۲۱)	بررسی ارتباط بین عاملیت معرفتی و فناوری اینترنت	بررسی اثرگذاری از طریق تضعیف مسئولیت در عاملیت معرفتی، محیط معرفتی آلوده به اخبار نادرست، حباب‌های فیلترکننده اطلاعات

پژوهش‌ها در ایران نیز بیشتر بر جنبه‌های حقوقی، فقهی و اجتهادی تأثیرات غیرفنی هوش مصنوعی توجه داشته است (Esmaili & Valipour, 2021; Takhshid, 2021). از این رو پژوهش حاضر یکی از محدود پژوهش‌هایی است که در بستر مطالعات بین‌رشته‌ای و از منظر معرفت‌شناسی سیاسی به موضوع تأثیرات هوش مصنوعی پرداخته و تلاش کرده است با تحلیل ثانویه اطلاعات پراکنده موجود، به یک الگوی تبیینی نزدیک شود.

۴. چارچوب مفهومی و نظری

برای نزدیک شدن به نحوه تأثیرگذاری هوش مصنوعی بر عاملیت سیاسی و اجتماعی ابتدا لازم است مفاهیم مربوطه و نظریات موجود مرور شود. AI یکی از آنهاست که در بیان مسئله تعریف شد، اما این فناوری نوظهور و پیچیده در شبکه‌های اجتماعی به شکل فناوری‌ها و تکنیک‌های متنوعی به کار می‌رود. الگوریتم‌های توصیه‌گر، شخصی‌ساز محتوا، ترجمه چندزبانه، قرار دادن آگهی به صورت خودکار، چت‌بات‌ها، تشخیص روند، نظر کاو، تعدیل محتوا، تعامل کاربر با محتوا، درآمدزایی محتوا و درک روندها در رسانه‌های اجتماعی از مهم‌ترین آنهاست. این فناوری‌ها هرچند به لحاظ سطح هوشمندی قابلیت‌ها و وظایف محدود و مشخصی دارند. اما قادرند الگوهای آماری در داده‌ها را با کمترین مداخله انسانی شناسایی و در جهت اهداف تعریف‌شده تصمیم‌گیری کنند. این اهداف در زمینه‌های رسانه‌های اجتماعی، عبارت است از جست‌وجوی محتوا، ارائه توصیه‌ها، تشخیص تصاویر یا گفتار، پروفایل کاربران و هدف قرار دادن آنها با تبلیغات شخصی، تجزیه و تحلیل احساسات در متن یا ایجاد محتوای جدید (Tsalakanidou, 2022).

۴.۱. عاملیت معرفتی

«عاملیت معرفتی»^۱ اصطلاحی آشنا در معرفت‌شناسی است که درباره چگونگی کنترل فرد بر باورهای خود و چگونگی شکل‌گیری و تجدیدنظر در آنها بحث می‌کند (Schlosser, 2019). همچنین بر این موضوع متمرکز است که چگونه نظام معرفتی فرد از فضایی چون عدالت، مسئولیت‌پذیری، اعتدال و صداقت انباشته یا تهی می‌شود (Elgin, 2013). امر مطلوب در حوزه عاملیت معرفتی آن است که انسان به عنوان «عامل» بر باورهای خود کنترل و اراده داشته باشد و بتواند تغییرات مطلوب را در زندگی ارتباطی و معرفتی خود ایجاد کند (Gunn, 2018)، از این‌رو بحث اصلی نشان دادن توانایی یک عامل انسانی در انتخاب، تأمل و شکل‌دهی آزادانه به باورها است (Elgin, 2013: 144; Krzanowski & Polak, 2022).

این امر کلی نیز پذیرفته شده که کسب معرفت و آگاهی امری صرفاً طبیعی و مبتنی بر عملکرد نورون‌های عصبی مغز نیست (Collin, 2013) و عاملیت‌های دیگر، چگونگی کسب آگاهی انسان را تحت تأثیر قرار می‌دهند. این عاملیت‌ها اغلب از بستر زندگی اجتماعی بر روی «تصمیم‌ها»، «ارزیابی‌ها» یا «انتخاب‌های» مختلف فرد اثر می‌گذارند. از این‌رو معرفت و کسب آگاهی مقوله‌ای اجتماعی است که نه تنها بر شکل‌گیری و تغییر باورها، تصمیمات و هنجارهای

1. Epistemic Agency

فردی تأثیر می‌گذارد که حتی می‌تواند به شکل «دانش جمعی»^۱ بازتعریف شود (Goldman & Whitcomb, 2011: 37; Hannon & De Ridder, 2021). یکی از این عوامل اجتماعی، رسانه‌ها و شبکه‌های اجتماعی است که به کمک فناوری‌های جدید عاملیت و اراده انسان بر باورها، ارزیابی‌ها و شناخت او از حقایق سیاسی-اجتماعی را تحت تأثیر قرار می‌دهد. البته این به معنای آن نیست که فرد مسئول جهان‌بینی و شکل دادن به باورهای خود نیست، بلکه فرض این است که عاملیت معرفتی می‌داند که باورهای او تا حد زیادی توسط مدارس، والدین، جامعه، فرهنگ، رسانه‌ها و سازمان‌های مذهبی شکل می‌گیرد، اما به‌عنوان موجود منطقی و عقلانی توانایی‌های مهمی برای ارزیابی این تأثیرات دارد و به‌طور تأملی با آنها درگیر می‌شود (Krzanowski & Polak, 2022: 41).

۴.۲. نظم دموکراتیک

نظم دموکراتیک معطوف به وضعیتی سیاسی-اجتماعی است که شهروندان در امور سیاسی و اجتماعی مشارکت دارند. در یک برداشت محدود، این نظم محدود به مشارکت انتخاباتی و حضور چهار تا پنج سال یک‌بار شهروندان در انتخابات می‌شود که نمایندگان خود در دولت یا مجلس را انتخاب می‌کنند (دموکراسی نمایندگی^۲). درحالی‌که در مفهوم وسیع‌تر این مشارکت مستقیم‌تر، پررودادتر و متنوع‌تر است و از شهروندان انتظار می‌رود به مناسب‌های مختلف دیدگاه و رأی خود را ابراز کنند (دموکراسی مشورتی^۳). بحث ما در اینجا بر معنای دموکراسی قوی و ضعیف متمرکز نیست و تنها بر این مسئله تأکید می‌شود که آن نظامی که به حق انتخاب و مشارکت شهروندان در سرنوشت سیاسی و اجتماعی احترام می‌گذارد و شرایط بروز آن را مهیا می‌سازد، دارای نظم دموکراتیک است (Barber, 1984). از سوی دیگر امر پایه‌ای در همه انواع دموکراسی، اعم از نمایندگی یا مشورتی، آن است که بروز انتخاب و مشارکت شهروندان بر اساس داوری منصفانه و دسترسی به آگاهی راستین است (Godin, 2023: 11). مسئله در اینجا آن است که آیا در شرایط کنونی که افراد در معرض انواع فناوری‌های ارتباطی و اطلاعاتی هستند، این عاملیت معرفتی مصون از خطا و انحراف است؟ آیا در پرتو مداخله هوش مصنوعی در تولید و توزیع اطلاعات عاملیت معرفتی یعنی شهروندان می‌تواند از تأثیرات شناختی آن در امان بماند و اراده و آگاهی راستین خود را در تعاملات سیاسی و اجتماعی ابراز کند؟ آنچه هابرماس و همکاران (۱۹۹۰) تحت عنوان عقلانیت ارتباطی بیان می‌کند، آیا در بستر هوش مصنوعی امکان تحقق دارد؟ (Habermas *et al.*, 1990). پاسخ اولیه آن است که تحقق چنین ایده‌آلی به‌صورت مطلق

1. Collective knowledge
2. Representative democracy
3. Deliberative democracy

غیرممکن است لیکن باید عاملیت‌های اثرگذار و بعضاً جایگزین اراده و آگاهی شهروندان را رصد و شناسایی کرد و سازوکار شناختی آنان را احصا کرد تا با تدابیر مقتضی از نظم دموکراتیک در برابر اختلالات شناختی محافظت کرد و فضای شناختی آزاد و برابر را برای افکار عمومی جهت مشارکت مدنی و سیاسی مهیا ساخت. البته در نظم‌های اقتدارگرا، تمامیت‌خواه و غیردموکراتیک، به حفاظت از عاملیت معرفتی شهروندان نیازی نیست، زیرا قرار نیست آنها کنش مدنی دیگری جز حمایت از رژیم داشته باشند، از این رو نحوه شکل‌گیری باورهای سیاسی خود بنیاد افراد، مهارت تأمل در خصوص آنها و بحث در مورد آن نمی‌تواند موضوع و دغدغه نظام‌های اقتدارگرا و طرفداران آنها باشد.

۳.۴. نظریات و چشم‌اندازها در خصوص هوش مصنوعی و عاملیت سیاسی – اجتماعی

در اینجا به اجمال چشم‌اندازهایی که تبیین‌کننده چگونگی اثرگذاری AI بر فرد و به تعبیری دست‌کاری عاملیت سیاسی و اجتماعی است بررسی می‌شود، چه آنکه به‌طور مستقیم وارد موضوع شده‌اند و چه آنان که غیرمستقیم فناوری به‌ویژه رسانه‌های ارتباط جمعی نوین را به‌عنوان عاملیت معرفتی در رقابت و جانشینی عاملیت انسانی معرفی می‌کنند؛ عاملی که اراده و انتخاب افراد در امور سیاسی و اجتماعی را مختل می‌کند و به‌تبع، آن نظم دموکراتیک را به نظم دستوری و اقتدارگرا سوق می‌دهد.

دانش و آگاهی (اعم از صادق و کاذب) باورهای انسانی را عقلانی و توجیه می‌کنند. منابع این آگاهی و دانش، به‌طور سنتی فرهنگ و منابع روزمره جامعه‌پذیری یعنی خانواده، مدرسه و دانشگاه بوده است، اما به‌تدریج با انقلاب اطلاعات و ارتباطات و امروزه با آمدن هوش مصنوعی در بستر تعاملات انسانی، بشر عاملیت معرفتی خود را از دست داده و عوامل پیچیده فناورانه، باورها و انتخاب‌های او را به شکل ناملموس جهت می‌دهد (Krzanowski & Polak, 2022: 42). از این رو برخی صاحب‌نظران معتقدند که بشر امروز در حال تجربه یک بحران معرفتی با نتایج نامشخص است (Waltzman, 2017; Zuboff, 2023: 20) و این بحران در حال گسترش به حوزه‌های عمومی و تضعیف عاملیت سیاسی است (Dahlgren, 2018: 20). هنگامی که شهروندان عاملیت معرفتی خود را از دست می‌دهند، منابع اقتدارگرا و ذی‌نفعان اصلی آن، به‌راحتی می‌توانند تبلیغات، اطلاعات نادرست و اطلاعات نادرست را که از اهداف، منافع یا سیاست‌های استبدادی آنان حمایت می‌کند، به‌کار گیرند و به این ترتیب نظم دموکراتیک مختل می‌شود (Bozdog & Van Den Hoven, 2015: 250).

در ابتدای عصر انقلاب ارتباطات و رسانه‌های نوین، دیدگاه غالب نظریات، معطوف بر این باور بود که اینترنت و فناوری اطلاعات به‌عنوان یک «اکولایزر بزرگ»^۱ برای جوامع در حاشیه و تحت ستم عمل می‌کند و می‌تواند دانش راستین و دموکراسی را در میان آنان تقویت نماید، بدون آنکه محصور در اقتدارگرایی دولتی و سانسور شود (Ross, 2017; Xenos et al., 2014: 21; Persily & Tucker, 2020: 18) اما با پیوستن فناوری‌های نوظهور هوشمند به اینترنت این چشم‌انداز خوش‌بینانه کمرنگ شد، زیرا رصد و تحلیل داده‌های انباشت‌شده در شبکه‌های اجتماعی توسط الگوریتم‌های هوشمند و ربات‌های اجتماعی که دارای سوگیری و اهداف اقتدارگرایانه هستند، دموکراسی و زیست مدنی را دچار اختلال می‌سازد (Dvorak, 2016).

نظریات، اختلال در نظم دموکراتیک متأثر از هوش مصنوعی را به طرق مختلف رصد کرده‌اند. برای نمونه سوگیری مندرج در مجموعه دانش و اطلاعات انباشت‌شده در فضای مجازی به‌عنوان داده‌های ورودی الگوریتم‌های هوش مصنوعی، می‌تواند الگوهای تبعیض‌آمیز و اقتدارگرایی اجتماعی، قانونی و سیاسی گذشته نسبت به زنان، اقلیت‌های قومی و نژادی و مذهبی را در بطن الگوریتم‌های هوشمند جای دهد و تبعیض‌های تاریخی و متصلب را از گذشته به حال و آینده منتقل سازد (Eubanks, 2018: 156; Mehrabi et al., 2022: 2). در این زمینه هوش مصنوعی ممکن است منعکس‌کننده سوگیری علیه کنشگری اجتماعی زنان به‌دلیل سوگیری موجود در تیم‌های توسعه‌دهنده این فناوری باشد، افرادی که اغلب مرد هستند و به نوبه خود این تعصب را از نسل‌های قبلی به ارث برده‌اند (Sun et al., 2019; Caliskan et al., 2017).

برخی دیگر اختلال در نظم دموکراتیک را با مفهوم «الگوکراسی»^۲ تبیین می‌کنند، یک سیستم حکمرانی سازمان‌یافته که بر اساس الگوریتم‌های برنامه‌ریزی‌شده کامپیوتری تمام وجوه زندگی شخصی و عمومی افراد را تحت کنترل، رصد و بررسی انبوه دارد و عوامل اقتدارگرایی حاکمیتی می‌توانند از این شیوه برای انقیاد و کنترل شهروندان استفاده کنند (Danahar, 2016: 247).

برخی دیگر ضعیف شدن عاملیت سیاسی-اجتماعی از طریق هوش مصنوعی را آن‌قدر وخیم ارزیابی می‌کنند که نظم دموکراتیک به «پایان» خود رسیده است، زیرا الگوریتم‌های خودسازمان‌دهنده هوش مصنوعی نقش انسان در تصمیمات سیاسی را کاهش می‌دهد. در این زمینه ابتدا در آزادی فکر و بیان عقیده سیاسی اختلال ایجاد می‌کند و سپس الگوهای رفتاری افراد را تجزیه و تحلیل و پیش‌بینی کرده و در نتیجه پیام‌ها و اطلاعات را به‌گونه‌ای سفارشی جهت تغییر نظر افراد نسبت به یک حزب، جریان سیاسی و نامزد انتخاباتی منتشر می‌کند. این اثر، به

1. Grand Equaliser
2. Algocracy

ترویج نظرهای خاص و حذف نظرهای مخالف منجر می‌شود (Brkan, 2019: 66; Zuiderveen & Borgesius *et al.*, 2018).

در برخی نظریات نیز هوش مصنوعی نه تنها خود به طور مستقیم عاملیت سیاسی و اجتماعی را تضعیف می‌کند که به عنوان ابزاری در دست عاملیت‌های بدخیم و خارجی نظم شناختی یک جامعه را نسبت به رهبران سیاسی و روندهای دموکراتیک موجود برهم می‌زند. در این زمینه هوش مصنوعی به عنوان ابزار خطرناک شناختی در ایجاد شایعه و تولید اخبار جعلی توسط قدرت‌های خارجی به کار می‌رود تا ثبات سیاسی، اعتماد عمومی به رسانه‌ها و نهادهای ملی، اعتبار رهبران سیاسی را کاهش دهد و به تعمیق شکاف‌های اجتماعی و مدنی، افزایش تنش‌های سیاسی و ایدئولوژیک کمک کند از طریق حملات سایبری، تحلیل اطلاعات کلان شهروندان از طریق هوش مصنوعی و نفوذ به اطلاعات شبکه‌های دولتی و شخصی ضمن نقض حریم خصوصی افراد و به خطر انداختن امنیت ملی، انتخابات را بدون هیچ ردپایی تحریف کنند (Susar & Aquaro, 2019: 419).

۵. راهبردهای اثر گذاری AI بر عاملیت سیاسی و اجتماعی

هوش مصنوعی چه ابزار عاملیت‌های بدخیم دولتی باشد، چه ابزار نفوذ و تخریب بازیگران متخاصم خارجی، می‌تواند به صورت‌ها و راهبردهای مختلف عاملیت سیاسی- اجتماعی را در رسانه‌های اجتماعی هدف قرار دهد، استقلال او در تصمیم‌گیری را تضعیف کند و روند کنشگری عقلانی، درونی و سنجیده او را در دست بگیرد. در ادامه به مهم‌ترین راهبردهای احصاشده از این حوزه پرداخته می‌شود.

۵.۱. آسیب معرفتی تدریجی

فناوری‌های هوشمند در رسانه‌های اجتماعی همچون اسلاف پیشین خود (رادیو و تلویزیون)، به تدریج بر حوزه شناخت افراد وارد می‌شود و آن را منقاد و مطیع می‌سازد. قرار گرفتن طولانی در بستر این رسانه‌ها، می‌تواند افراد را مطیع‌تر و بیشتر مستعد اقناع بیرونی و پذیرش نسنجیده ارزش‌ها و باورهای دیگران بسازد. پژوهش‌ها نشان می‌دهد این نفوذ تدریجی در ابتدا با جذب اعتماد از طریق ارائه رایگان اطلاعات یا خدمات مختلف پلتفرم‌ها صورت می‌گیرد. به تدریج فرد باور می‌کند که فضای سکوه‌ای اجتماعی و اینترنت مملو از اطلاعات قابل اعتماد است، به نحوی که حتی خصوصی‌ترین جنبه‌های زندگی روزمره خود را در آن جست‌وجو و دنبال می‌کنند (Wylie, 2019).

فناوری‌های هوشمند در فضای اجتماعی همچون یک «پاولوف دیجیتال»^۱ عمل و عاملیت انسانی را از طریق خدمات رایگان و خدمات متنوع و جذاب شرطی می‌کند (Zuboff, 2023: 15). این نفوذ معرفتی را ابزار «سرمایه‌داری نظارتی»^۲ معرفی می‌کند که کنترل انسان و انتخاب‌های او، غایت مطلوب آن است. روند مذکور به سرعت جهان معرفتی کاربرانی را در تسخیر خود می‌گیرد. بررسی تجربی این موضوع نشان می‌دهد افراد به تدریج به سیستم‌های هوشمند فعلی از جمله الگوریتم‌های مکالمه‌ای مانند ChatGPT و LaMDA اعتماد کرده و در تعامل با آن احساس امنیت می‌کنند. فناوری که در حال حاضر به سمت شخصیت‌های دیجیتالی واقعی در حال توسعه هستند و ربات‌های سخنگوی در آینده نه‌چندان دور، با توان اقناع بالا، بر معرفت انسانی تسلط خواهند یافت (Rosenberg, 2023: 2; Savitz, 2023).

۲.۵. هک شناختی از طریق پروفایل‌های خرد

در «هک شناختی»^۳ الگوریتم‌ها به شناسایی آسیب‌پذیری‌های عاطفی و ترس‌های افراد و نیز سلاقی و علایق سیاسی و اجتماعی آنها می‌پردازند تا درک آنها از حقایق، افراد یا موقعیت‌های سیاسی را تغییر دهد. این عملکرد به‌طور گسترده از پروفایل‌های شخصی که در بستر پلتفرم‌های رسانه‌های اجتماعی مانند فیس‌بوک، متا، اینستاگرام و غیره ایجاد شده‌اند، استفاده و اطلاعات را جمع‌آوری می‌کند (Dabrowa, 2022: 3; Schweitzer et al., 2023: 2). برای مثال در سال ۲۰۱۶ داده‌های کاربران فیس‌بوک توسط کمپین ریاست جمهوری ترامپ برای هدف قرار دادن رأی‌دهندگان استفاده شد و مطابق با تیپ شخصیتی افراد پیام مناسب به آنان ارسال می‌شد. AI به همین طریق می‌تواند رأی‌دهندگان مردد و نوسانی را شناسایی کرده و با تولید اطلاعات مختلف آنان را به رأی دادن به نفع یک جریان سیاسی خاص ترغیب کند (Henderson, 2023: 7).

۳.۵. ایجاد حباب‌های معرفتی

هوش مصنوعی از طریق الگوریتم‌های همگن‌ساز، کاربران را در طبقات همگن دسته‌بندی کرده و اطلاعات سازگار با آن فضای فکری و ارزشی را منتشر می‌کند، از این‌رو خواسته یا ناخواسته به‌جای افزایش اطلاع‌رسانی و دسترسی گسترده با اطلاعات متنوع، در عمل، دسترسی به اطلاعات را سانسور یا محدود می‌سازد. در این حباب‌های فیلترکننده افراد در غارهای اطلاعاتی

1. Digital Pavlov
2. Surveillance Capitalism
3. Cognitive Hacking

محبوس می‌شوند و تنها پژواک اندیشه‌ها و اطلاعات مورد تأیید آنان شنیده می‌شود. اتاق‌های پژواک مذکور این پتانسیل را دارند که سوگیری را در شناخت افزایش دهند، یعنی تنها آن دسته از نظرهایی را که با دیدگاه ما موافق‌اند، به‌عنوان واقعیت ببینیم و بپذیریم. این ویژگی به فرایندهای تعقل سنجیده و دموکراتیک که مستلزم تعامل و مواجهه با حداکثر دیدگاه‌های سیاسی و اجتماعی در تقابل است (Nguyen, 2020: 142). البته «حباب‌های معرفتی»^۱ امری جدید نیست و فقط فناوری آن را تقویت می‌کند. افراد در جوامع کوچک، بسته یا سنتی با افراد همفکر زندگی خود می‌کنند که عاملیت معرفتی آنها را کاهش می‌دهد. در فضای مدرن نیز تا حدودی این اتفاق از طریق رسانه‌های جمعی قدیمی مانند تلویزیون رخ می‌داد. برای مثال تماشای فاکس نیوز در ایالات متحده افراد را در یک حباب معرفتی خاص قرار می‌دهد. اما اکنون هوش مصنوعی در ترکیب با رسانه‌های اجتماعی دیجیتال تأثیری سریع و بسیار گسترده‌تر به این پدیده شکل می‌دهد و سیلوهای جدید و مؤثرتری ایجاد می‌کند (Coeckelbergh, 2022: 1343). این حباب‌ها و کاهش طنین صداهای مختلف در آن ذاتاً برای دموکراسی نامطلوب است و نقش مشارکت‌کنندگان را به حامیان و پیروان افراطی تقلیل می‌دهد (Bozdag & Van Den Hoven, 2015: 249).

۴.۵. فرسایش معرفت^۲ فردی و بی‌اعتمادی به آن

هوش مصنوعی می‌تواند اعتماد به ظرفیت‌های معرفتی شخص را از بین ببرد و به‌تدریج مهارت‌های او در زمینه تعقل و استدلال را تضعیف کند، وقتی هوش مصنوعی به‌طور روزافزون و «باورپذیرتری» استفاده شود، فرد شروع به زیر سؤال بردن ظرفیت‌های خود به‌عنوان یک عامل معرفتی برای تشخیص حقیقت از دروغ می‌کند. این ویژگی که به «شرم معرفتی»^۳ موسوم است (Coeckelbergh, 2022: 1343). از ایده «شرم خلاق»^۴ گونتر آندرس گرفته شده که توصیف احساسی است مبنی بر آنکه «ماشین‌ها در مقایسه با بدن فانی و ناقصمان، بسیار کامل‌تر از ما هستند و به‌طور تحقیرآمیزی کیفیت بالاتری نسبت به انسان دارند» (Anders, 2016: 30). به همین شکل افراد در مواجهه با هوش مصنوعی به‌عنوان یک فناوری توانمند، ظرفیت معرفتی خود را دست‌کم می‌گیرند و نسبت به آن بی‌اعتماد و شرمگین می‌شوند. البته این حالت با آگاهی و تقویت استدلال تحلیلی و انتقادی قابل رفع است و فرد تلاش خواهد کرد کنترل معرفت خود را باز پس گیرد. اما واقعیت این است که امروزه AI باورهای سیاسی فرد را از طریق ردپای

1. Epistemic Bubbles

2. Corrosion of Epistemic Agency

3. Epistemic Shame

4. Günter Anders

دیجیتال او به خوبی و حتی گاهی بهتر از خود فرد شناسایی می‌کند و این امکان کنترل توسط کسانی را فراهم می‌کند که توسعه‌دهنده یا مالک این فناوری هستند. از سوی دیگر، وقتی دولت‌ها و افراد در معرض توانمندی‌های بی‌نظیر هوش مصنوعی در جمع‌آوری و تحلیل اطلاعات قرار می‌گیرند، اداره و مدیریت بسیاری از امور را به آن می‌سپارند. گاهی اعتماد بیش از حد سبب می‌شود که انسان فراموش کند که عاملیت دانش و آگاهی خود اوست و فناوری ابزار و دستکاری برای کارآمدی بیشتر است.

۵.۵. تضعیف معرفت علی

یکی دیگر از شیوه‌هایی که عاملیت معرفتی شهروندان را تحت تأثیر قرار می‌دهد، ارائه اطلاعات و آگاهی به صورت علم داده و آمار است که به صورت یک پیش‌فرض مورد موثق افراد قرار گرفته و توصیه‌ها و تحلیل‌های مبتنی بر آن استفاده می‌شود. بدون آنکه معرفت افراد به سوی کشف روابط علی موضوع هدایت شود، افراد ممکن است توسط داده‌های آماری و نتایج یک جست‌وجو از طریق الگوریتم‌های هوش مصنوعی در معرض محبوبیت یا عدم محبوبیت یک سیاست یا فرد قرار بگیرند، بدون آنکه شواهد علی آن وضعیت و چرایی آن تبیین شود، البته افراد نیز برای کشف آن تشویق نمی‌شوند (Coeckelbergh, 2022: 1347). نگرانی در اینجا این نیست که آمار دروغ می‌گویند، بلکه ابتدای دانش و آمار ارائه شده بر رابطه همبستگی است که یک مسیر معرفتی «آسان»^۱ را از طریق آزمون همبستگی‌ها ارائه می‌کند، به نحوی که به تلاش زیادی از سوی شهروندان به عنوان عوامل معرفتی نیاز ندارد. همبستگی می‌تواند درست باشد، اما پیچیدگی واقعیت و عوامل سازنده آن را پنهان می‌کند. برای نمونه آمار نشان می‌دهد در دو جامعه تمایل به مشارکت سیاسی افزایش یافته است، اما دقت بیشتر در عوامل سازنده نشان می‌دهد جامعه نخست به دلیل پروپاگاندا و سیاست‌های پوپولیستی حکمرانان به مشارکت بیشتر در انتخابات ترغیب شده‌اند (عوامل بیرونی)، در حالی که جامعه دوم به دلیل فرهنگ مشارکتی و فعال تمایل بالایی به مشارکت در انتخابات دارند (عوامل درونی). البته فقط افراد آگاه و متخصصان می‌دانند که چه اتفاقی در حال وقوع است و عموم شهروندان در لایه‌های سطحی اطلاعات آماری متوقف می‌شوند و بر اساس اعتماد به آن دست به انتخاب می‌زنند.

۵.۶. افزایش معرفت احساسی و هیجانی

در نتیجه قرار گرفتن افراد با حجم بالای اخبار و اطلاعات نادرست، نفرت افکن و جعلی، این احتمال وجود دارد که کنش‌های آفلاین شهروندان، به‌سوی رفتارهای هیجانی سرریز، و منازعات بیرونی (آفلاین) را تقویت کند. گفتار نفرت‌افکن بلندتر از سایر پیام‌ها هیجان و احساس توده‌ها را برمی‌انگیزد به عبارت دیگر محتوای نفرت‌افکن، با توجه به نقطه‌زنی احساسی آنها، بلندتر از حد معمول شنیده می‌شود (Gagliardon *et al.*, 2016). انتشار چنین فضایی به زندگی روزمره افراد، چندین پیامد دارد: نخست ایجاد هراس و «ترس‌های اشتراکی»^۱ به‌ویژه در میان اقلیت‌ها و کاربران در حاشیه و ضعیف است. بر اساس پژوهش‌ها زنان این احساس را بیشتر تجربه می‌کنند (Costello *et al.*, 2019: 428). دو تحقیق آوان و زبی^۲ در سال‌های ۲۰۱۵ و ۲۰۲۰ نشان می‌دهد مسلمانان ساکن در کشورهای غربی از آن بیم دارند که تهدیدات فضای مجازی در فضای حقیقی زندگی‌شان رخ دهد و عملی شود (Awan & Zempi, 2015 & 2020). دومین پیامد انزوا و کناره‌گیری از فعالیت‌های مدنی است که خود را به‌صورت انزوا و ترس از بیان عقاید در فضای عمومی و مشارکت مدنی در دو فضای آفلاین و آنلاین نشان می‌دهد. در واقع انزوا و عدم ورود به مباحث سیاسی از نتایجی است که به مرور زمان در افراد بیشتر می‌شود و این ناشی از غلبه احساسات غلو شده و نفرت‌افکن در فضای شبکه‌های مجازی است که به کمک الگوریتم‌های هوشمند تشدید می‌شود (Barnidge *et al.*, 2019). سوم بسیج سیاسی خشونت‌بار است. تولید و توزیع هوشمند اطلاعات نادرست، جعلی و مغرضانه نه‌تنها انزوا که خشونت محیطی و تعصب‌محور را افزایش می‌دهد (Matsuda, 2018)، بسیج سیاسی و خشونت‌بار اوباش علیه مسلمانان سریلانکا و مردم روهینگا در میانمار در بستر فیس‌بوک نمونه‌ای از آن است (Vindu *et al.*, 2018). اساساً فضای گمنامی این اجازه را به افراد می‌دهد تا در خشونت گفتاری حد و مرضی را نگه ندارند و در برخی موارد با کنار رفتن هنجارها و نرم‌های اجتماع آفلاین تنش‌های بین گروهی افزایش می‌یابد؛ یعنی افراد جسورانه و بدون ترس از پیامد نفرت‌شان خودشان را اظهار می‌کنند. بر اساس تحقیق مولر و شوارز^۳ (۲۰۱۹) با استفاده بیشتر از توئیتر در آغاز پویش انتخاباتی ترامپ جرائم نفرت‌افکن علیه مسلمانان افزایش یافت (Muller & Schwarz, 2019).

1. Communal Fear
2. Awan & Zempi
3. Muller & Schwarz

۵.۷. ایجاد اجماع معرفتی دروغین و احتمال سلطه اقلیت بر اکثریت

در علوم سیاسی، «آستروتورفینگ»^۱ به ایجاد یک کمپین اطلاعاتی کنترل‌شده برای ایجاد تصورات نادرست اطلاق می‌شود که امروزه اغلب توسط ربات‌های هوشمند ساخته می‌شوند. این کمپین‌ها، به‌ظاهر از حمایت گسترده افراد انسانی برخوردارند (Longley, 2020; Chan, 2024: 507). گاهی به‌نحوی است که به‌شکل یک دانش جمعی به شهروندان القا می‌شود (Goldman & Whitcomb, 2011). درحالی‌که در نظم دموکراتیک اصل بر این است که آحاد جامعه از جمله نیروها و هویت‌های اجتماعی خاموش و اقلیت، بدخواهان، محرومان و خرده‌گروه‌های ضداجتماعی و ناسازگار نیز دارای حق و حقوقی است و کنش‌های سیاسی و اجتماعی آنان چنانچه با نفع اکثریت اعضای جامعه باشد، پذیرفته است. در خصوص تأثیرگذاری هوش مصنوعی بر این عاملیت‌ها دو تبیین وجود دارد؛ رویکرد نخست این است که اگر هوش مصنوعی ابزار اکثریت و دولت‌های اقتدارگرا باشد، این عاملیت‌ها، بیش از پیش به حاشیه رانده شده و حذف می‌شوند، آنچنان‌که در نظم تکنوکراسی انتظار می‌رود. اما رویکرد دوم پرده از اتفاقی برمی‌دارد که اقلیت را بر اکثریت حاکم می‌سازد و آن این است که گاهی الگوریتم‌های هوش مصنوعی به‌نحوی متناقض ابزار اکثریت‌نمایی اقلیت بر اکثریت می‌شود و از همین نقطه نظم دموکراتیک فرو می‌پاشد. در واقع برخی سازوکارهای فنی هوش مصنوعی بستری را فراهم می‌کند که حاشیه به جریان اصلی تبدیل شود و شروع به دیکته کردن ارزش‌ها و انتخاب‌های خود به دیگران کند؛ یعنی می‌تواند دیدگاه‌های حاشیه‌ای را چنان تقویت کند که آنها به جریان اصلی تبدیل شوند. در این شرایط جهان‌بینی‌ای که به نفع بیشتر مردم بود، تحریف می‌شود و به پس‌زمینه می‌رود یا از هم پاشیده می‌شود و یک جامعه سرگردان شکل می‌گیرد که رهبریت آن به دست افراد بی‌ثبات، فاقد صلاحیت سطحی و متزلزل می‌افتد. این اختلال معرفتی را ربات‌های اجتماعی و سیاسی و کمپین‌های دروغی به‌خوبی مدیریت می‌کند (Anderson & Rainie, 2020; Dumbrava, 2021).

۵.۸. ایجاد فاصله معرفتی با واقعیت سیاسی-اجتماعی

«ایجاد فاصله معرفتی»^۲ راهبرد دیگر اثرگذاری هوش مصنوعی بر عاملیت سیاسی-اجتماعی است. فاصله معرفتی زمانی مطرح است که فناوری‌های اطلاعاتی میان واقعیت و عاملیت معرفتی فاصله ایجاد می‌کنند و داده‌ها پیش از رسیدن به عاملیت انسانی از چندین واسطه عبور می‌کند (Nilsson, 2021: 102)، یعنی داده‌ها پیش از رسیدن به عاملیت سیاسی و اجتماعی توسط

1. Astroturfing

2. Epistemic Distance

رسانه‌های اجتماعی چندین باز تحلیل، هضم و آماده شده‌اند و به این شکل الگوریتم‌های هوش مصنوعی فاصله معرفتی انسان از واقعیت اجتماعی را افزایش می‌دهد. به بیان دیگر تماس ما با واقعیت به واسطه برخی رابط‌های غیرشفاف چون الگوریتم‌های هوشمند انجام می‌شود، از این رو ما مالکیت جهان‌بینی خود را از دست می‌دهیم و کمتر خودمان استدلال می‌کنیم (Krzanowski & Polak, 2022: 45).

برای دموکراسی حقیقت یک شرط لازم است و هرچه فاصله فرد از حقیقت سیاسی دور شود، نظم دموکراتیک بیشتر تهدید می‌شود (Coeckelbergh, 2022: 1347). هانا آرنت نیز در تبیین خاستگاه توتالیتاریسم می‌گوید تحریف حقیقت از ویژگی‌های رژیم‌های توتالیتار است، چنین فضایی ظرفیت بی‌نظیری برای ایجاد و حفاظت از دنیای ساختگی از طریق دروغ‌گویی مداوم دارد (هدف ساختارهای تمامیت‌خواه کنترل محیط معرفتی است، به گونه‌ای که از استقرار و حفظ قدرت تمامیت‌خواه حمایت کند، استفاده از هوش مصنوعی امکان تحریف حقیقت را تسریع می‌کند (Arendt, 1973). تولید و باز نشر هدفمند ویدیوهای جعلی (دیپ‌فیک) علیه جریان‌های سیاسی رقیب از آن جمله است و وضعیت وقتی وخیم‌تر می‌شود که این قابلیت هوش مصنوعی آنقدر به حقیقت نزدیک است که تشخیص جعلی بودن آن تقریباً غیرممکن است (Langguth et al., 2021: 5). بررسی تجربی این موضوع نیز امری مشکل است، زیرا افراد از این حالت‌های معرفتی آگاه نیستند که بتوانند درباره آن اظهار نظر کنند. عمده این مباحث ناشی از استنتاج علمی و ارزیابی ثانویه محققان است.

۵.۹. اعتبار بخشیدن به معرفت غیر مستحق

هوش مصنوعی از طریق فعالیت گسترده ربات‌های اجتماعی و ایجاد اجماع سوری می‌تواند به نامزدهای انتخاباتی اعتبار غیر مستحق اعطا کند و شکاف فناورانه و عدم دسترسی افراد شایسته با تاکتیک‌های هوش مصنوعی در زمان انتخابات یا اجرای یک سیاست سبب می‌شود در منظر عموم کمتر دیده شده و صدایی از ایشان در فضای عمومی بازتاب پیدا نکند. در این شرایط اغلب افراد فاقد صلاحیت و به پشتوانه رسانه و ثروت در حجم و سرعت بالا، شایستگی خود را به شهروندان القا می‌کنند. همان‌طور که بیان شد الگوریتم‌های تولیدکننده حباب معرفتی که دارای ترجیحات جناحی و سیاسی‌اند، افراد را بیش از پیش از شنیدن صدای طرف دیگر که ممکن است توانمندی و صلاحیت بیشتری داشته باشد، دور می‌سازد. به بیان دیگر «سوگیری ضمنی»^۱

1. implicit biases

الگوریتم‌های هوش مصنوعی موجب حذف تبعیض‌آمیز یک جریان سیاسی و اجتماعی و دیده شدن یک جریان ضعیف و سطحی شود (Marinucci et al., 2023; Johnson, 2021).



شکل ۱. الگوی اثرگذاری عاملیت هوش مصنوعی بر عاملیت سیاسی-اجتماعی

۶. نتیجه و پیشنهادها

مسئله پژوهش حاضر بر این گزاره بنا شد که فناوری‌های رسانه‌ای جدید چون هوش مصنوعی با سازوکارهای پیچیده فنی خود می‌توانند اراده و انتخاب افراد را کنترل و جهت‌دهی کنند و بدون آنکه فرد متوجه باشد، درک و دانش او از حقایق اجتماعی و سیاسی را برسانند. ملاحظه شد فناوری هوش مصنوعی به‌صورت ربات‌های سیاسی، اتاقک‌های پژواک، جعل عمیق و هدف‌گیری خرد، معرفت را در عاملیت انسانی دست‌کاری می‌کند و می‌تواند تحت راهبردهای مختلف کنترل فکر، استدلال و تعقل او را در موضوعات سیاسی و مدنی کنترل کند. بر اساس تحلیل ثانویه و مرور نظام‌مند مطالعات معرفت‌شناختی حوزه هوش مصنوعی که به تأثیرات معرفتی هوش مصنوعی بر اراده و اختیار انسان در موقعیت‌های سیاسی و اجتماعی پرداخته‌اند، الگوی تبیینی زیر احصا شد و نشان داد هوش مصنوعی به‌عنوان یک عاملیت هوشمند غیرانسانی تحت راهبردهای زیر می‌تواند معرفت انسانی را در بستر شبکه‌های اجتماعی دست‌کاری کرده و پیامدهای مختلفی بر رفتار اجتماعی او تحمیل کند.

جدول ۲. الگوی تبیینی اثرگذاری هوش مصنوعی بر عاملیت سیاسی - اجتماعی

پیامدهای معرفتی و رفتاری	راهبردهای اثر
- جذب اعتماد کاربران از طریق ارائه اطلاعات رایگان - ایجاد احساس امنیت به اطلاعات موجود در بستر اینترنت - نفوذ و کنترل تدریجی بر انسان و انتخاب‌های سیاسی و اجتماعی او	آسیب معرفتی تدریجی
- شناسایی و احصای سلاقی و علایق سیاسی هر فرد - ارسال پیام‌های سیاسی متناسب با تیپ شخصیتی افراد - ترغیب رأی‌دهندگان مردد و نوسانی به سمت یک معرفت سیاسی مشخص	هک شناختی از طریق پروفایل‌های خرد
- قرار دادن افراد در طبقات همگن معرفتی و ارزشی - سانسور یا محدود کردن اطلاعات در غارهای معرفتی - نشاندن صداها و ناهمسو و عدم تعامل با آنها - اختلال در مواجهه حداکثری با دیدگاه‌های سیاسی و اجتماعی - تبدیل شدن افراد به حامیان و پیروان افراطی در اتاق‌های معرفتی	ایجاد حباب‌های معرفتی
- تضعیف تدریجی مهارت‌های فرد در زمینه تعقل و استدلال - زیر سؤال بردن ظرفیت‌های معرفتی خود در تشخیص حقیقت از دروغ - ایجاد «شرم معرفتی» و بی‌اعتمادی نسبت به برتری هوش مصنوعی بر هوش انسانی - فراموش شدن نقش هوش مصنوعی به‌عنوان یک ابزار و دستیار و تبدیل آن به تصمیم‌ساز اصلی در زندگی افراد	فرسایش معرفت فردی و بی‌اعتمادی به آن
- عدم هدایت افراد به‌سوی کشف روابط علی، پیچیده و پنهان پدیده‌های سیاسی و اجتماعی - ترسیم یک مسیر معرفتی آسان مبتنی بر آزمون‌های آماری ساده به‌ویژه آزمون همبستگی	تضعیف معرفت علی
- سرریز احساسات غلو شده به‌سوی رفتارهای هیجانی و منازعات بیرونی از طریق اطلاعات مغرضانه، جعلی و نادرست - ایجاد هراس و «ترس‌های اشتراکی» در میان اقلیت‌های مذهبی و زنان - تشدید انزوا و کناره‌گیری از فعالیت‌های مدنی - ترس از بیان عقاید در فضای عمومی - سوم بسیج سیاسی خشونت‌بار و افزایش خشونت محیطی و تعصب‌محور	افزایش معرفت احساسی و هیجانی

• ایجاد تصورات نادرست که یک فرد یا سیاست از حمایت گسترده مردمی برخوردار است. • قدرت اجماع‌سازی سوری به شکل القای یک دانش جمعی • اکثریت‌نمایی اقلیت بر اکثریت • تبدیل‌شدن حاشیه به جریان اصلی و دیکته کردن ارزش‌ها و انتخاب‌های آنان بر اکثریت • ایجاد اکثریت نمای فاقد صلاحیت، سطحی و متزلزل	ایجاد اجماع معرفتی دروغی و احتمال سلطه اقلیت بر اکثریت
• هضم اطلاعات توسط واسطه‌های معرفتی غیرانسانی پیش از رسیدن به انسان • از دست دادن مالکیت جهان‌بینی • عادت به استدلال ماشین در رسیدن به شناخت واقعیت	ایجاد فاصله معرفتی با واقعیت سیاسی-اجتماعی
• افزایش تبعیض و بی‌عدالتی نسبت به افراد صلاحیت‌دار فاقد امکانات فناورانه پرسرعت و هوشمند • دیده نشدن صدای افراد دارای صلاحیت • سوگیری ضمنی هوش مصنوعی نسبت به یک جریان خاص ضعیف و سطحی	اعتبار بخشیدن به معرفت غیرمستحق

الگوی تبیینی مذکور البته تمام روایت اثرگذاری هوش مصنوعی بر افکار عمومی نیست و این فناوری نوظهور قابلیت‌هایی نیز دارد که می‌توان از آن برای کمک به دموکراسی به‌ویژه در زمان انتخابات و دعوت افراد به مشارکت مدنی بهره برد و تأثیرات منفی پیشین را تا حدی اصلاح نماید. برای نمونه هوش مصنوعی می‌تواند به تعادل بهتر محتوای رسانه و «شناسایی اطلاعات مغرضانه و جعلی» و ارائه پوشش جایگزین کمک کرده و از «قاب‌بندی» محتوا به سمت یک جریان خاص جلوگیری کند. البته این پژوهش بر چگونگی اثرگذاری هوش مصنوعی بر معرفت انسانی و پیامد منفی آن بر نظم دموکراتیک متمرکز بود و بررسی تأثیرات اصلاحی و جبرانی آن را به فرصت دیگری می‌سپارد.

در خصوص راه‌حل‌ها شاید مهم‌ترین و مؤثرترین شیوه نه‌گزینه سانسور باشد و نه محدودسازی توسعه و استفاده از هوش مصنوعی، بلکه گزینه مؤثر، تقویت مهارت عاملیت انسانی در تفکر تحلیلی و انتقادی است. آموزش انواع شناخت و محدودیت‌های آن، به‌نحوی که شهروندان تشویق شوند تا ظرفیت‌های استدلالی خود را از طریق فراگیری و تقویت تفکر انتقادی پرورش دهند. به بیان دیگر مقابله با انقیاد و تعصب فکری که از زیست در غارهای معرفتی و اتاق‌های پژواک مجازی حاصل شده است، نیازمند آن است تا فرد بیاموزد که تکبر معرفتی یک

فضیلت نیست و باید با سایر دیدگاه‌های سیاسی و اجتماعی که منافع اکثریت را محترم می‌دانند، در تعامل باشد (Gunn et al., 2021). توسعه هوش مصنوعی علی، مسئولیت‌پذیر و قابل اعتماد راه‌حل‌های دیگری است که معضلات مطرح‌شده را تا حدی کم می‌کند که مخاطب آن بیشتر شرکت‌های فناوری و خصوصی‌اند.

در هر صورت دور ساختن انسان از فناوری نمی‌تواند گزینه بادوام و مؤثری باشد، لازم است پیوستگی و همبستگی با فناوری‌های جدید اطلاعاتی را پذیرفت و چالش‌های آن را از طریق رصد تأثیرات آن و تقویت تفکر انتقادی، خلق و فعال مدیریت کرد. در واقع یکی از دستاوردهای این پژوهش آن است که دیگر عاملیت معرفتی تنها به انسان و تلاش فردگرایانه آن محصور و محدود نیست و به همان نسبت مسئولیت تمام تصمیمات سیاسی و اجتماعی خود را بر عهده ندارد، بلکه در یک سیستمی از عاملیت‌های معرفتی قرار دارد و فناوری‌های هوشمند به‌عنوان یکی از عوامل معرفتی، مشارکت‌کننده در تولید دانش و معرفت است که وضعیت کنونی ما را به‌عنوان دانا شکل می‌دهد. به بیان دیگر با توسعه هرچه بیشتر فناوری‌های نوظهور هوشمند چون هوش مصنوعی، «عامل معرفتی به طیف گسترده‌ای از موجودیت‌های انسانی و غیرانسانی مانند شرکت‌ها، رایانه‌های پیچیده، هوش مصنوعی و رسانه‌های اجتماعی اطلاق می‌شود که همه این عاملیت‌ها مسئول ارزیابی، کسب، انتشار و حفظ دانش و معرفت خواهند بود.

تعارض منافع

نویسنده اعلام می‌کند که تعارض منافع وجود ندارد و تمام مسائل اخلاق در پژوهش را شامل پرهیز از سرقت ادبی، انتشار و یا ارسال بیش از یک بار مقاله، تکرار پژوهش دیگران، داده‌سازی یا جعل داده‌ها، منبع‌سازی و جعل منابع، رضایت ناآگاهانه سوژه یا پژوهش شوند، سوء رفتار و غیره، به‌طور کامل رعایت کرده است.

سپاسگزاری

نویسنده از پژوهشکده مطالعات فناوری ریاست جمهوری و مجله سیاست به‌دلیل اهتمام در حمایت از انتشار مقاله کمال تشکر و قدردانی را دارد.

References

- Anders G. (2016). on promethean shame. In: Müller CJ, editor. Prometheanism. London: Rowman & Littlefield; pp.29-95.

- Anderson, J; Rainie, L. (2020). Many tech experts say digital disruption will hurt democracy. In: <https://www.pewresearch.org/internet/2020/02/21/many-tech-experts-say-digitaldisruption-will-hurt-democracy/>
- Arendt, H. (1973). The origins of totalitarianism (Vol. 244). Houghton Mifflin Harcourt.
- Awan, I; Zempi, I. (2015). We fear for our lives: Offline and online experiences of anti-Muslim hostility. Tell MAMA, October. www.tellmamauk.org/wp-content/uploads/resources/We%20Fear%20For%20Our%20Lives.pdf
- Barber, B.R (1984). Strong democracy: participatory politics for a new age. University of California Press, Oakland
- Barberá, P. (2020). Social media, echo chambers, and political polarization. Social media and democracy: The state of the field, prospects for reform, 34.
- Barnidge, M; Kim, B; Sherrill, L.A; Luknar, Ž; Zhang, J. (2019). Perceived exposure to and avoidance of hate speech in various communication settings. Telematics and Informatics, 44, 101263.
- Bishop, C.M. (2006). Pattern Recognition and Machine Learning. Springer New York, NY. Retrieved from <https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>
- Bozdog, E; Van Den Hoven, J. (2015). Breaking the filter bubble: democracy and design. Ethics and information technology, 17, 249-265.
- Brkan, M. (2019). Artificial intelligence and democracy: The impact of disinformation, social bots and political targeting. Delphi, 2, 66.
- Buhmann, A; Fieseler, C. (2023). Deep learning meets deep democracy: Deliberative governance and responsible innovation in artificial intelligence. Business Ethics Quarterly, 33(1), 146-179
- Caliskan, A; Bryson, J.J; Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. Science, 356(6334), 183-186.
- Chan, J. (2024). Online astroturfing: A problem beyond disinformation. Philosophy & Social Criticism, 50(3), 507-528.
- Coeckelbergh, M. (2023). Democracy, epistemic agency, and AI: political epistemology in times of artificial intelligence. AI and Ethics, 3(4), 1341-1350
- Collin, F. (2013). Two kinds of social epistemology. Social Epistemology Review and Reply Collective, 2(8), 79-104.
- Copeland, B.J. (2023). Artificial intelligence". Encyclopedia Britannica, 30 Jun. 2023, <https://www.britannica.com/technology/artificial-intelligence>. Accessed 1 July 2023.

- Costello, M; Hawdon, J; Bernatzky, C; Mendes, K. (2019). Social group identity and perceptions of online hate. *Sociological Inquiry*, 89(3), 427–452
- Crawford, K; Calo, R. (2016). There is a blind spot in AI research. *Nature*, 538(7625), 311-313.
- Dabrowa, P. (2022). Cognitive hacking as the new disinformation frontier. *Financial Times*. <https://www.Ft.com/content/52535b2b-cb23-4ab6-ac66-2859cf9d1ae9>.
- Dahlgren, P. (2018). Media, Knowledge and Trust: The Deepening Epistemic Crisis of Democracy. *Javnost - The Public*. 25. 1-9. 10.1080/13183222.2018.1418819.
- Danahar J. (2016). The threat of algocracy: reality, resistance and accommodation. *Philos. Technol.* 29:245–268.
- Dumbrava, C. (2021) 'Key social media risks to democracy. EPRS (European Parliamentary Research Service). https://www.europarl.europa.eu/RegData/etudes/IDAN/2021/698845/EPRS_IDA_698845_EN.Pdf
- Dvorak, John C.(2016) 'The false promises of the Internet.' <https://www.pcmag.com/opinions/the-false-promises-of-the-internet>
- Elgin, C.Z. (2013). Epistemic agency. *Theory and research in education*, 11(2), 135-152.
- Esmaili, M; Valipour, A (2021). Feasibility study of civil liability of general artificial intelligence due to causing damage in civil law, *Legal Thought*, Second Volume, Fall 1400, No. 6, pp. 1-16. **[In Persian]**
- Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. New York: St. Martin's Press
- Gagliardone, I; Pohjonen, M; Beyene, Z; and et al. (2016). Mechachal: Online debates and elections in Ethiopia: From hate speech to engagement in social media. Working paper. <https://eprints.soas.ac.uk/30572/>
- Godin, R.A. (2023). *Deliberate Democracy*, translated by Lila Sazgar, Tehran: Qoqnos Publications. **[In Persian]**
- Goldman, A; Whitcomb, D. (2011): *Social epistemology: essential readings*. Oxford University Press, New York.
- Goodfellow, I; Bengio, Y; Courville, A. (2016). *Deep learning*. MIT press.
- Gunn, H. (2018). *Social-Epistemic Agency*. Doctoral Dissertations. University of Connecticut, 1875. <https://opencommons.Uconn.Edu/dissertations/1875>
- Gunn, H; Lynch, M.P; Lackey, J. (2021). The internet and epistemic agency. *Applied epistemology*, 389-409.
- Habermas, J; Lenhart, T.C; Nicholson, S.W. (1990). (eds.) *Moral consciousness and communicative action*. MIT Press, Cambridge

- Hannon, M; De Ridder, J. (Eds.). (2021). The Routledge handbook of political epistemology. Routledge.
- Hao, K. (2019). 'Why AI is a threat to democracy—and what we can do to stop it'. MIT Technology Review. <https://www.Technologyreview.com/2019/02/26/66043/why-ai-is-a-threat-to-democracy-and-what-we-can-do-to-stop-it/> Accessed 26 Feb 2019
- Henderson, D.P; Cole, D.J; Baulis, N. (2023). How AI is changing democracy, Nudging, microtargeting, and epistemic bubbles, published Maxim Institute, <https://www.maxim.org.nz/content/uploads/2023/10/WEB-How-AI-is-changing-democracy-Oct-23.pdf>
- Johnson, G.M. (2021). Algorithmic bias: on the implicit biases of social technology. *Synthese*, 198(10), 9941-9961.
- Krzanowski, R; Langguth, J; Pogorelov, K; Brenner, S; Filkuková, P; Schroeder, D.T. (2021). Don't Trust Your Eyes: Image Manipulation in the Age of DeepFakes. *Frontiers in Communication*, 6, 632317.
- Longley, R. (2020). What Is Astroturfing in Politics? Definition and Examples. In: <https://www.thoughtco.com/what-is-astroturfing-definition-and-examples-5082082>
- Marinucci, L; Mazzuca, C; Gangemi, A. (2023). Exposing implicit biases and stereotypes in human and artificial intelligence: state of the art and challenges with a focus on gender. *AI & SOCIETY*, 38(2), 747-761.
- Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, 160(4): 835–50
- Martin-Rozumiłowicz, B; Kužel, R. (2019). Social media, disinformation and electoral integrity. International Foundation for Electoral Systems.
- Matsuda, M.J. (2018). Words That Wound: Critical Race Theory, Assaultive Speech, and the First Amendment. London: Routledge
- McGonagle, T; Bednarski, M; Francese Coutinho, M; Zimin, A. (2019). Elections and media in digital times. UNESCO Publishing.
- Mehrabi N; Morstatter F; Saxena N; Lerman K; Galstyan A. (2022). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35.
- Muller, K; Schwarz, C. (2019). Making America hate again? Twitter and hate crime under Trump. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3149103
- Nguyen, C.T. (2020). Echo chambers and epistemic bubbles. *Episteme*, 17(2), 141-161.
- Nilsson, N.J. (1998). Artificial Intelligence: A New Synthesis. Morgan Kaufmann
- Nilsson, J. (2021). Shaping epistemic distance: producing and withholding knowledge in market research, *Journal of Cultural Economy*, 14:1, 101-116

- Persily, N; Tucker, J.A. (Eds.). (2020). Social media and democracy: The state of the field, prospects for reform. Cambridge University Press
- Reider, P.J. (2016). Social Epistemology and Epistemic Agency Decentralizing Epistemic Agency (ed.), Rowman and Littlefield
- Rosenberg, L. (2023). The Manipulation Problem: Conversational AI as a Threat to Epistemic Agency. ArXiv preprint arXiv: 2306.11748.
- Ross, S. (2017). Whatever happened to the Internet's promise? In: <https://techonomy.com/whatever-happened-to-the-internets-promise/>
- Savitz, E.J. (2023). Microsoft brings chatgpt to Bing. There's finally some new competition in search. Microsoft Updates Bing To Add ChatGPT AI Technology | Barron's. Barrons. Available at: <https://www.barrons.com/articles/chatgpt-microsoft-bing-51675799472>
- Schlosser, M. (2019). Agency. The Stanford encyclopedia of philosophy. Edward N. Zalta (Ed.). [https:// www. Plato. Stanf ord.edu/ archi ves/ win20 19/ entri es/ agency/](https://www.plato.stanford.edu/archives/win2019/entries/agency/). Accessed 8 Nov 2022.
- Schweitzer, S; Dobson, K.S; Waytz, A. (2023). Political Bot Bias in the Perception of Online Discourse. Social Psychological and Personality Science, 19485506231156020.
- Sun, T; Gaut, A; Tang, S; Huang, Y; El Shereif, M; Zhao, J; Mirza, D; Belding, E; Chang, K.W; Wang, W.Y. (2019). "Mitigating Gender Bias in Natural Language Processing: Literature Review." In: A. Korhonen, D. Traum, and L. Marquez (eds.), Proceedings of the 57th Annual Meeting of the Association of Computational Linguistics, pp. 1630–40. Available at: <https://www.aclweb.org/anthology/P19-1159.pdf>
- Susar, D; Aquaro, V. (2019). Artificial intelligence: Opportunities and challenges for the public sector. In Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance (pp. 418-426).
- Takhshid, Z (2021). An Introduction to the Challenges of Artificial Intelligence in the Field of Civil Liability. Private Law (University of Tehran) spring and summer - Issue 38, 227 to 250. **[In Persian]**
- Tsalakanidou, F. (2022). AI technologies and applications in media: State of Play, Foresight, and Research Directions, AI4Media - A European Excellence Centre for Media, Society and Democracy
- Vindu G; Kumar, H; Frenkel, S. (2018). In Sri Lanka, Facebook contends with shutdown after mob violence. New York Times, March 8. www.nytimes.com/2018/03/08/technology/sri-lanka-facebook-shutdown.html
- Waltzman, R. (2017). The weaponization of information: The need for cognitive security. In: https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT473/RAND_CT473.pdf
- Wylie, C. (2019). Mindf*ck: Inside Cambridge Analytica's plot to break the world. Profile Books.

- Xenos, M.A; Vromen, A; Loader, B.D. (2014). The great equalizer? Patterns of social media use and youth political engagement in three advanced democracies. In *The networked young citizen* (pp. 17-38). Routledge.
- Zempi, I; Awan, I. (2020). Offline and Online Experiences of Anti-Muslim crime. For the Special Rapporteur on freedom of religion or belief [Preparation for the report to the 46th Session of Human Rights Council].
- Zuboff, S. (2023). The age of surveillance capitalism. In *Social theory re-wired* (pp. 203-213). Routledge.
- Zuiderveen Borgesius, F; Möller, J; Kruikemeier, S; Ó Fathaigh, R; Irion, K; Dobber, T; De Vreese, C.H. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review*, 14(1), 82-96. 2, 66.

COPYRIGHTS

©2023 by the University of Tehran. Published by the University of Tehran Press. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0>

