ISSN: 2717-0934 Vol.7 , No.4, Serial No. 26, Winter 2026 http://journal.iiwfs.com/

The Journal of Fundamental and Applied Studies of the Islamic World

https://doi.org/10.22034/fasiw.2026.245618

Article Type: Research paper

Artificial Intelligence and National Identity in Iran: Perspectives and Policymaking Imperatives

***Seyed Mehdi Hosseini Taghiabad[1], Behrouz Ghezel[2]**

1. Assistant Professor, Department of Regional Studies, Faculty of Law and Political Science, University of Tehran, Tehran, Iran.
2. PhD Student of Regional Studies, Faculty of Law and Political Science, University of Tehran, Tehran, Iran.

Article Info.

Received: 29/09/2025
Accepted: 24/01/2026
Available Online: 01/03/2026

Keywords:
Artificial Intelligence, National Identity, Policymaking, Semantic Infrastructure, Digital Governance.

*** Corresponding Author:**
Dr. Seyed Mehdi Hosseini Taghiabad
Address: Assistant Professor, Department of Regional Studies, Faculty of Law and Political Science, University of Tehran, Tehran, Iran.
E-mail: mehdiabad@ut.ac.ir

Abstract

Artificial intelligence (AI), particularly in its new generative forms and large language models, is not merely a technical tool alongside other technologies but an active mechanism that shapes content selection, credibility ranking and even the linguistic formats through which social experience is expressed. AI therefore must be examined not only at the level of "application" but also through deeper layers of data, evaluative standards and access rules that structurally determine what is considered "sayable", "visible", and "reliable". From a public policy standpoint, AI can be understood both as a "technology of governance" and as an "emergent mode of governance," encompassing a set of classification and coordination mechanisms whose political and cultural implications extend far beyond many classical regulatory approaches. Accordingly, the central question of this article is as follows: How can Iran, in the age of artificial intelligence, align technological development with the active preservation and creative renewal of national identity (in its Iranian–Islamic–revolutionary configuration), without sliding into cultural closure or harmful selfcontained insularity? In response, the article, through an analytical and policyoriented approach, proposes a framework of "cultural justice in AI" accompanied by six complementary policy pillars: (1) defining identitysensitive domains and determining risk levels, (2) developing reliable Persian corpora and data assets, (3) designing cultural-identity evaluation benchmarks and auditing mechanisms, (4) participatory governance and data diplomacy, (5) enhancing AI literacy within identityrelated and educational institutions and (6) implementing policy pilots in selected domains with iterative learning.

How to Site:
Hosseini Taghiabad, S M and Ghezel, B. (2026). Artificial Intelligence and National Identity in Iran: Perspectives and Policymaking Imperatives. *The Fundamental and Applied Studies of the Islamic World*, 7(4), 206-235.

Authors retain the copyright and full publishing rights.
Published by Iranian Association of Geopolitics. This article is an open access article licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0)

فصلنامهٔ مطالعات بنیادین و کاربردی جهان اسلام،
سال هفتم، شمارهٔ بیست و ششم، زمستان ۱۴۰۴
https://doi.org/10.22034/fasiw.2026.245618

انجمن ژئوپلیتیک ایران
Iranian Association of Geopolitics (I.A.G)

نوع مقاله: علمی ـ پژوهشی

هوش مصنوعی و هویت ملی در ایران؛ چشم‌انداز و بایسته‌های سیاستگذاری

* سیدمهدی حسینی تقی‌آباد[1]، بهروز قزل[2]

۱. استادیار گروه مطالعات منطقه‌ای، دانشکده حقوق و علوم سیاسی، دانشگاه تهران، تهران، ایران.
۲. دانشجوی دکتری مطالعات منطقه‌ای، دانشکده حقوق و علوم سیاسی، دانشگاه تهران، تهران، ایران.

اطلاعات مقاله

تاریخ دریافت: ۰۷ مهر ۱۴۰۴
تاریخ پذیرش: ۰۴ بهمن ۱۴۰۴
تاریخ انتشار: ۱۰ اسفند ۱۴۰۴

کلیدواژه‌ها:
هوش مصنوعی، هویت ملی، سیاستگذاری، زیرساخت معنا، حکمرانی دیجیتال.

* نویسندهٔ مسئول:
دکتر سیدمهدی حسینی تقی‌آباد
نشانی: استادیار گروه مطالعات منطقه‌ای، دانشکده حقوق و علوم سیاسی، دانشگاه تهران، تهران، ایران.
پست الکترونیک:
mehdiabad@ut.ac.ir

چکیده

هوش مصنوعی، به‌ویژه در صورت‌بندی‌های جدیدِ مولد و مدل‌های زبانی بزرگ، نه‌تنها ابزاری فنی در کنار سایر فناوری‌ها، بلکه سازوکاری فعال است که در گزینش محتوا، رتبه‌بندیِ اعتبار و حتی قالب‌های زبانیِ بیانِ تجربه اجتماعی، به‌طور روزمره نقش‌آفرینی می‌کند. از این‌رو، این فناوری را نه صرفا در سطح «کاربرد»، بلکه می‌بایست با توجه به لایه‌های عمیق‌تری در سطح داده، استانداردهای ارزیابی و قواعد دسترسی در نظر گرفت که به‌صورت ساختاری، تعیین می‌کند که چه چیزی «قابل گفتن»، «قابل دیدن» و «قابل اتکا» است. در چشم‌انداز سیاستگذاریِ عمومی، می‌توان هوش مصنوعی را هم «فناوریِ حکمرانی» و هم «شکلی نوظهور از حکمرانی» دانست؛ شامل مجموعه‌ای از سازوکارهای طبقه‌بندی و هماهنگ‌سازی که پیامدهای سیاسی و فرهنگی آن، از بسیاری تنظیم‌گری‌های کلاسیک فراتر است. بر همین اساس، مسئله اصلیِ این مقاله نیز چنین صورت‌بندی شده است: ایران چگونه می‌تواند در عصر هوش مصنوعی، توسعه فناورانه را با صیانتِ فعال و بازآفرینیِ خلاقِ هویت ملی (در ترکیبِ ایرانی-اسلامی-انقلابی) هم‌راستا کند، بی‌آنکه به انسداد فرهنگی یا خودبسندگیِ زیان‌بار دچار شود؟ در پاسخ، مقاله کوشیده است تا با روشی تحلیلی و سیاست‌محور، قالب «عدالت فرهنگی در هوش مصنوعی» و شش محور سیاستیِ مکمل را پیشنهاد کند: (۱) تعریف حوزه‌های حساس هویتی و تعیین سطح ریسک، (۲) ایجاد پیکره‌داده‌ها و دارایی‌های داده‌ایِ فارسیِ قابل اتکا، (۳) طراحی معیارهای ارزیابی فرهنگی-هویتی و سازوکار ممیزی، (۴) حکمرانی مشارکتی و دیپلماسی داده، (۵) ارتقای دانش هوش مصنوعی در نهادهای هویتی و آموزشی و (۶) اجرای پایلوت‌های سیاستی در حوزه‌های منتخب و یادگیری تدریجی.

استناد به این مقاله:
حسینی تقی‌آباد، سیدمهدی و قزل، بهروز. (۱۴۰۴). هوش مصنوعی و هویت ملی در ایران؛ چشم‌انداز و بایسته‌های سیاستگذاری. *مطالعات بنیادین و کاربردی جهان اسلام*، ۷(۴)، ۲۰۶-۲۳۵.

Authors retain the copyright and full publishing rights.
Published by Iranian Association of Geopolitics. This article is an open access article licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0)

۱. مقدمه

تحـولِ اخیـرِ هـوش مصنوعـی[1] را بایـد بخشـی از دگرگونـی بزرگ‌تـری دانسـت کـه طـی آن، «زیرسـاخت‌های دیجیتـال» از سـطح انتقـالِ اطلاعـات بـه سـطح سـازمان‌دهیِ معنـا و داوری دربـاره حقیقـت و اعتبـار نزدیـک شـده اسـت (Sourati & Evans, 2023: 1682-1696). در تجربه روزمره، زمانـی، موتـور جسـت‌وجو و شـبکه اجتماعـی درگاهِ اصلـیِ مواجهـه بـا جهـانِ اطلاعـات به‌شـمار می‌رفـت؛ امـا اکنـون مدل‌هـای زبانـی و سـامانه‌های مولـد، خـود بـه «هم‌صحبـت»، «مفسـر»، «خلاصه‌سـاز» و «پیشـنهاددهنده» تبدیـل شـده اسـت و همیـن نقـشِ میانجی‌گرانـه باعـث می‌شـود کـه پرسـش از هویـت ملـی، دیگـر در سـطح نمادهـا و محتـوا متوقـف نمانـده، بلکـه بـه لایه‌هـای داده، طراحـی و تنظیم‌گـری پیونـد بخـورد (Lane et al., 2024). ایـن وضعیـت بـرای ایـران، به‌منزلـه جامعـه‌ای کـه همزمـان بـا فشـارهای ژئوپلیتیـک، مؤلفـه زبـان فارسـی به‌عنـوان سـرمایه وحـدت ملـی و حساسـیت‌های دینـی و انقلابـی در عرصـه عمومـی را داراسـت، اهمیتـی دوچنـدان دارد. چراکـه مسـئله، فقـط ایـن نیسـت کـه «هـوش مصنوعـی چـه محتوایـی تولیـد می‌کنـد»، بلکـه ایـن اسـت کـه چـه داده‌ای را معتبـر می‌دانـد، چـه تنوعـی از سـبک‌های زبانـی و روایت‌هـای تاریخـی را بازنمایـی می‌کنـد و در عمـل چـه چیـزی را به‌عنـوان نسـخه غالـبِ فهـمِ جهـان بـه کاربـر تحمیـل کـرده یـا پیشـنهاد می‌دهـد (Moller et al., 2025; Pan et al., 2024).

بـا این‌حـال، سیاسـتگذاری داخلـی در حـوزه هـوش مصنوعـی، ممکن اسـت بـه دوگانـه سـاده‌انگارانه «فرصت/تهدیـد» فروکاسـته شـده، یـا در سـطح برنامه‌هـای کلـیِ توسـعه اقتصـاد دیجیتـال باقـی بمانـد. در شـرایطی‌که در سـطح بین‌المللـی، بررسـی راهبردهـای ملـی نشـان می‌دهـد کـه دولت‌هـا معمـولا ترکیبـی از اخلاق‌محـوری، انعطـاف نهـادی و نوعـی «نامعینـی کارکـردی» را برمی‌گزیننـد تـا هـم مسـیر توسـعه فنـاوری همـوار شـده و هـم امـکان مداخلـه تنظیم‌گرانـه حفـظ شـود (Dubois & Walzak, 2025)؛ بـرای ایـران، اتـکا بـه کلی‌گویی‌هـای اخلاقـی بـدون سـازکارهای اجرایـی، می‌توانـد بـه شـکاف میـان «گفتمـان» و «واقعیـتِ فنـی» بینجامـد؛ از سـوی دیگـر، تنظیم‌گـری سـخت‌گیرانه نیـز بـدون مشـارکت اجتماعـی، زیرسـاخت داده و ارزیابـی بومـی، ممکـن اسـت هـم نـوآوری را فرسـوده کنـد و هـم بـه دورزدن مقـررات و شـکل‌گیری بازارهـای غیررسـمی منتهـی شـود. بنابرایـن، چارچوبـی کـه هـم بـه «هویـت» به‌مثابـه مسـئله سیاسـت عمومـی توجـه کـرده و هـم از نظـر فنـی و نهـادی، قابـل پیاده‌سـازی و قابـل ارزیابـی باشـد، یکـی از ضرورت‌هـای جـدی اسـت؛ و ایـن، همـان پرسـش اصلـی

1. Artificial intelligence (AI)

پژوهش حاضر است که در عصر هوش مصنوعی، ایران چگونه می‌تواند توسعه فناورانه را با صیانتِ فعال و بازآفرینیِ خلاقِ هویت ملی (در ترکیبِ ایرانی-اسلامی-انقلابی) هم‌راستا کند، بی‌آنکه به انسداد فرهنگی یا خودبسندگیِ زیان‌بار دچار شود؟

بر این اساس، پژوهش، با روشی تحلیلی-مفهومی و سیاست‌محور می‌کوشد تا یک «چارچوب تصمیم‌پذیر» برای سیاستگذاری (به‌جای آزمون تجربیِ یک فرضیه محدود) ارائه کند؛ چارچوبی که بتواند در نهادهای حاکمیتی و تنظیم‌گر، به زبانِ برنامه، معیار و سازکار تبدیل شود. بدین منظور، سه مسیر تحلیلی به‌صورت هم‌افزا مورد توجه قرار گرفته است؛ مسیر نخست، مرور و بازسازی انتقادی ادبیات حکمرانی هوش مصنوعی است؛ به‌ویژه بحث‌هایی که نشان می‌دهد دولت‌ها باید همزمان نقش «تسهیل‌گر توسعه» و «تضمین‌کننده کاهش ریسک و مشارکت اجتماعی» را ایفا کنند و حکمرانی هوش مصنوعی را به‌مثابه حوزه‌ای کم‌توسعه‌یافته اما دارای فوریت در سیاست عمومی بنگرند (Chan, 2023). مسیر دوم، تمرکز بر ویژگی‌های خاصِ هوش مصنوعیِ مولد است؛ فناوری‌ای که علاوه بر سوگیری، مخاطراتی مانند توهم‌زایی، نشت اطلاعات حساس، دشواری کنترل و عدم شفافیت را به همراه دارد و بنابراین، تنظیم‌گریِ آن ناگزیر باید «پیش‌دستانه، مشارکتی و تطبیقی» طراحی شود (Walter, 2024: 1-29). مسیر سوم نیز، اتصال این مباحث به صورت‌بندی بومیِ هویت ملی در ایران است؛ یعنی تلاش برای ترجمه مفاهیم انتزاعی حکمرانی (داده، معیار، ممیزی، پاسخگویی) به زبانِ حوزه‌های هویتی (زبان فارسی، حافظه تاریخی، مرجعیت دینی، روایت انقلاب و انسجام اجتماعی).

محدودیت روشیِ مهم در پژوهش حاضر این است که بسیاری از بحث‌های هویتی در ایران، یا در سطح هنجاری باقی می‌ماند یا داده‌های عمومیِ کافی برای سنجش‌های کمّیِ دقیق در دسترس نیست؛ بنابراین، پژوهش، به‌جای ادعای «پیش‌بینی قطعی»، بر «طراحی گزینه‌های سیاستی و مسیر یادگیری» تکیه می‌کند. الگویی مبتنی بر پیشنهادِ پایلوت، ارزیابی، بازنگری و نهادینه‌سازی.

۲. پیشینه پژوهش

با توجه به شتاب فزاینده تحولات در حوزه هوش مصنوعی، گسترش کاربست آن در عرصه‌های مختلف زندگی فردی و اجتماعی و افزایش توجه دولت‌ها، از جمله جمهوری اسلامی ایران به ابعاد راهبردی این فناوری، تولیدات علمی مرتبط با هوش مصنوعی در سال‌های اخیر رشد

قابل توجهی یافته است. با این همه، مرور پژوهش‌های داخلی نشان می‌دهد که نسبت میان هوش مصنوعی و هویت ملی، به‌ویژه در صورت‌بندی ایرانی-اسلامی-انقلابی آن، به‌صورت مستقل و منسجم کمتر موضوع مطالعه قرار گرفته است.

در این میان، صادقیان (۱۴۰۴) با تفکیک دو سطح «حکمرانی هوش مصنوعی» و «حکمرانی بر هوش مصنوعی» کوشیده است چارچوبی برای مواجهه با چالش‌های این فناوری ارائه دهد. یافته‌های وی نشان می‌دهد که هوش مصنوعی می‌تواند به ارتقای دقت تصمیم‌گیری، بهینه‌سازی تخصیص منابع و افزایش شفافیت در حکمرانی یاری رساند، اما در همان حال، مسائل مهم اخلاقی، حقوقی و نظارتی را نیز پیش می‌کشد؛ ازاین‌رو، مدل سه‌بعدی حکمرانی هوش مصنوعی به‌عنوان پاسخ پیشنهادی این پژوهش مطرح شده است. کرمی و متقی‌دستنایی (۱۴۰۳) نیز با تمرکز بر تأثیر هوش مصنوعی بر دیپلماسی عمومی نشان داده‌اند که کار با کلان‌داده و بهره‌گیری از قابلیت‌های این فناوری، برای دیپلمات‌ها مزیت رقابتی ایجاد می‌کند و الگوهای ارتباطی دولت‌ها با افکار عمومی را دگرگون می‌سازد. از نظر آنان، هوش مصنوعی می‌تواند دامنه و اثربخشی ابتکارات دیپلماسی عمومی را افزایش داده و مناسبات سنتی این حوزه را به سوی سازوکارهای چابک‌تر و شبکه‌ای‌تر سوق دهد.

رحمانی و همکاران (۱۴۰۳) در بررسی نقش مدیریت استراتژیک اخلاق در عصر هوش مصنوعی، ضمن تأکید بر ظرفیت این فناوری در پیش‌بینی دقیق‌تر سناریوها و پشتیبانی از تصمیم‌گیری راهبردی، بر ضرورت مدیریت پیامدهای اخلاقی آن نیز انگشت نهاده‌اند. نتایج این پژوهش نشان می‌دهد که پیوند میان هوش مصنوعی و اصول اخلاقی، در صورت طراحی و هدایت مناسب، می‌تواند هم کیفیت تصمیم‌گیری را ارتقا دهد و هم از شدت مخاطرات احتمالی بکاهد. کرمپوری و اسلامی (۱۴۰۳) نیز با رویکردی تفسیری و مبتنی بر چارچوب نظری جان مارو، کوشیده‌اند حاکمیت هوش مصنوعی را از منظر پدیدارشناسی سیاسی تبیین کنند. یافته‌های آنان حاکی از آن است که این صورت از حاکمیت، با قاب‌بندی انسان و ریاضی‌سازی وجود و معنا همراه است؛ فرایندی که لزوما با تعالی فلسفی هستی و آزادی انسان همسو نیست. در امتداد همین ملاحظات، حامد اجاقی و همکاران (۱۴۰۳) در ارائه چارچوبی نظری برای روش‌شناسی هنجارگذاری هوش مصنوعی، چهار سطح تطوریِ ارشادی، توصیفی، ساختاری و ایجابی را مطرح کرده‌اند که هر یک به‌ترتیب کارکردهایی چون رونق شکوفایی، رشد ادراک، رفع موانع و تعمیم قواعد را دنبال می‌کند.

از سوی دیگر، محمدی‌منفرد و تقی‌پورجاوی (۱۴۰۳) تأثیر هوش مصنوعی بر ماهیت جنگ را از منظر اسناد رژیم صهیونیستی بررسی کرده‌اند. آنان با تکیه بر برنامه‌های ملی، راهبردها و ابتکارات مرتبط با هوش مصنوعی نشان داده‌اند که این رژیم در پی پیشبرد توسعه این حوزه در ابعاد نظامی، امنیتی و اقتصادی است؛ هرچند در اسناد رسمی، از تصریح مستقیم به برخی وجوه نظامی و امنیتی پرهیز شده است. همچنین رشیدی (۱۴۰۳) در تبیین چالش‌های دموکراسی در عصر دیجیتال به این نتیجه رسیده است که هوش مصنوعی، در کنار ظرفیت‌هایی برای بهبود عملکرد دموکراسی، می‌تواند زمینه‌ساز مخاطراتی جدی همچون دستکاری افکار عمومی، مخدوش شدن محیط اطلاعاتی، تعمیق شکاف دیجیتال، گسترش نابرابری و تبعیض، و در نهایت تقویت گرایش‌های تمامیت‌خواه شود.

در مجموع، مرور ادبیات نشان می‌دهد که هرچند پژوهش‌های داخلی درباره هوش مصنوعی در حال گسترش است، تمرکز غالب آنها بر ساحت‌هایی چون حکمرانی، اخلاق، دیپلماسی، جنگ و دموکراسی قرار دارد و نسبت این فناوری با هویت ملی هنوز به‌طور مستقیم و روشمند صورت‌بندی نشده است. این در حالی است که در ادبیات بین‌المللی، پیوند میان حکمرانی هوش مصنوعی، قدرت الگوریتمی، مخاطرات مولدها، و ضرورت سازوکارهای مشارکتی و پاسخگو، به‌صورت فزاینده‌ای در حال جلب توجه است. ازاین‌رو، پژوهش حاضر می‌کوشد این خلأ را در سطحی نظری و سیاستی، با تمرکز بر نسبت هوش مصنوعی و هویت ملی در ایران، مورد توجه قرار دهد.

۳. چشم‌انداز نظری: هوش مصنوعی به‌مثابه زیرساخت معنا و صورت‌بندی عدالت فرهنگی

برای فهم نسبت هوش مصنوعی و هویت ملی، می‌بایست از تلقی «هوش مصنوعی به‌عنوان ابزار خنثی» فاصله گرفت و آن‌را در جایگاه «زیرساخت معنا» بازشناخت (Bareis & Katzenbach, 2021). زیرساختی، شامل مجموعه‌ای از سازکارهای فنی و نهادی که نه تنها تعیین می‌کند چه صورت‌هایی از زبان، دانش و روایت در مقیاس وسیع بازتولید شود، بلکه اساسا قواعد «مشروعیتِ» گفتار و «امکانپذیریِ» بازنمایی را نیز دیکته می‌کند (Viberg et al., 2024; Cesur, 2024). اینجاست که پیوند میان ادبیات زیرساخت‌های دیجیتال و نظریه انتقادی قدرت

ضـرورت می‌یابـد. همانطـور کـه میجیـاس و کالـدری در کتـاب ربایـش داده[2] (۲۰۲۴) اسـتدلال می‌کننـد، آنچـه در اقتصـاد سیاسـی پلتفرم‌هـا رخ می‌دهـد، صرفـا اسـتخراج اطلاعـات نیسـت، بلکـه شـکلی نویـن از «اسـتعمار داده[3]» اسـت کـه در آن، منابـع فرهنگـی و معانـی بومـی به‌مثابـه مـواد خـام بـرای تغذیـه سیسـتم‌های مولـد اسـتخراج شـده و پـس از فـرآوری، به‌صـورت کلیشـه‌ها و انگاره‌هـای مسـلط بـه جوامـع بازگردانـده می‌شـود (Mejias & Couldry, 2024). ایـن فراینـد کـه در آن، الگوهـای داده‌کاوی را می‌تـوان میـراث‌دار سـاختارهای اسـتعماری گذشـته دانسـت، به‌طـور مسـتقیم بـا ایـده حاکمیـت فرهنگـی و هویتـی در تضـاد اسـت. چراکـه مدل‌هـای زبانـی بـزرگ، بـا تکیـه بـر داده‌هـای نامتـوازنِ عمدتـا انگلیسـی‌زبان و مبتنـی بـر معرفت‌شناسـی غربـی، بـه عاملانـی بـرای تثبیـت «هژمونـی الگوریتمـی» تبدیـل شـده اسـت کـه از طریـق نحـو زبـان و سـاختارهای زبانـیِ به‌ظاهـر خنثـی، سلسـله‌مراتب‌های فرهنگـی و اجتماعـی را بازتولیـد و تشـدید می‌کنـد. بـه بیـان دقیق‌تـر، مدل‌هـای زبانـی بـزرگ «ماشـین‌های رمزگـذاری، تکـرار و تقویـت سلسـله‌مراتب‌های اجتماعـی و فرهنگـی» اسـت کـه به‌عنـوان «عامـلان ضمنـی حکمرانـی نـرم» عمـل کـرده و از طریـق اولویت‌هـای نحـوی و سـبکی خـود، نـه فقـط آنچـه می‌گوییـم، بلکـه چگونگـی اندیشـیدن و ارتبـاط مـا بـا یکدیگـر را نیـز بازپیکربنـدی می‌کنـد (The syntax of power, 2025). ایـن نکتـه زمانـی روشـن‌تر خواهـد شـد کـه بـه نقـشِ هـوش مصنوعـی در تولیـد دانـش و هدایـت «مسـیرهای نیـل بـه دانـش» نگریسـته شـود؛ چراکـه مدل‌هـا، اگـر صرفـا بـر «محتـوا» تکیـه کـرده و زمینـه انسـانیِ تولیـد دانـش را نادیـده بگیـرد، پیش‌بینی‌هـا و پیشـنهادهای آنهـا محـدود و گاه گمراه‌کننـده خواهـد بـود. امـا هنگامـی کـه توزیـعِ تخصـص و رفتـار انسـانی وارد مدل‌سـازی شـود، تـوان پیش‌بینـی و کیفیـت پیشـنهادها بـه شـکل معنـاداری تغییـر می‌یابـد (Graziani et al., 2022: 3473-3504). ترجمـه سیاسـتیِ ایـن گـزاره بـرای مسـئله هویـت چنیـن اسـت: اگـر مدل‌هـای زبانـی و سـامانه‌های مولـد، بـدون درک بافـتِ اجتماعـی-فرهنگـیِ فارسـی و بـدون داده‌هـای نمایانگـر سـنت‌های فکـری، دینـی و تاریخـی ایـران توسـعه یابـد، خروجـیِ ظاهـرا «روان» آنهـا نـه تنهـا ممکـن اسـت در سـطح معناشـناختی، ساده‌سـاز، کلیشه‌سـاز یـا تحریف‌گـر باشـد، بلکـه در لایـه‌ای عمیق‌تـر، بـه تضعیـف ظرفیت‌هـای شـناختی و زبانـی جمعـی می‌انجامـد؛ یعنـی در همـان نقطـه‌ای کـه هویـت جمعـی بـه ظرافت‌هـای زبانـی، بینامتنیـت و زمینـه تاریخـی وابسـته اسـت. پژوهشـگران هشـدار داده‌انـد کـه توسـعه نامتـوازن هـوش مصنوعـی می‌توانـد بـه «همگن‌سـازی فرهنگـی» و «فرسـایش حاکمیـت فرهنگـی» منجـر شـود (BRAID Researchers,

2. Data Grab
3. Data Colonialism

.(2025; UN Expert Mechanism on the Right to Development, 2025

در ادبیات اخلاق و حکمرانی هوش مصنوعی، عدالت تنها به‌معنای «دسترسی برابر» نیست، بلکه به‌معنای توزیع منصفانه منافع و مضرات و نیز طراحی سازکارهایی است که از به حاشیه رفتن گروه‌ها و ارزش‌ها جلوگیری کند (Stanger et al., 2024). حال، اگر این ایده به سطح فرهنگی منتقل شود، «عدالت فرهنگی» یعنی تضمین اینکه زبان فارسی، گویش‌ها و سبک‌های متنوعِ بیان و نیز روایت‌های مرکزیِ تاریخ، دین و انقلاب، صرفا مصرف‌کننده محصولات فناورانه نبوده، بلکه در لایه داده، در معیارهای ارزیابی و در نهادهای تصمیم‌گیری، سهم مؤثر داشته باشند. چراکه بسیاری از بی‌عدالتی‌های فرهنگی در هوش مصنوعی، نه در «متن خروجی» بلکه در «نامرئیاتِ زیرساختی» رخ می‌دهد؛ یعنی در اینکه چه داده‌ای گردآوری و پاکسازی می‌شود، چه چیزی برچسب می‌خورد، چه چیزی مجاز به استفاده در آموزش است و در نهایت چه چیزی معیارِ «کیفیت» و «اعتبار» تلقی می‌شود. کیت کرافورد در کتاب «اطلس هوش مصنوعی[۴]» (۲۰۲۱) با تشریح زیرساخت‌های مادی و معرفتی هوش مصنوعی، به‌خوبی نشان می‌دهد که این سیستم‌ها به‌رغم ظاهر بی‌طرف و عینی خود، برای خدمت به منافع گروه‌های خاص و تشدید نظام‌های قدرتِ موجود طراحی شده است (Crawford, 2021). بنابراین، دفاع از عدالت فرهنگی در عصر هوش مصنوعی، مستلزم مقابله با «ربایش داده» و بازپس‌گیری حاکمیت بر روایت‌های هویتی است.

در این چارچوب، زبان فارسی صرفا یک متغیر فنی نیست، بلکه یک مسئله حاکمیتیِ فرهنگی است. وجودِ بنچمارک‌ها (معیارهای سنجش) و مجموعه چالش‌های استاندارد برای سنجش فهم زبان فارسی نشان می‌دهد که حتی در سطح ارزیابیِ مدل‌ها نیز، کمبودِ منابع و معیارهای پایدار می‌تواند مسیر پیشرفت را کند و ناهمگون سازد و از سوی دیگر، هرگاه معیارهای بومی ایجاد شده، امکان مقایسه، بهبود و سیاستگذاری مبتنی بر شواهد بیشتر شده است (Far-rell, 2025). بر همین مبنا، پیوند میان «عدالت فرهنگی» و «دارایی‌های داده‌ای/معیاری» قابل تشخیص است. اگر تولید و نگهداریِ پیکره‌داده‌ها و معیارهای فارسی، از سوی دولت و نهادهای عمومی و علمی، به‌عنوان کالای عمومی در نظر گرفته شود، آنگاه سیاست هویتی از سطح شعار به سطح زیرساخت منتقل خواهد شد.

4. Atlas of AI

اما عدالت فرهنگی فقط مسئله «فنی‌سازی زبان» هم نیست. هوش مصنوعی مولد، در دلِ مناقشات جهانیِ حکمرانی نیز قرار گرفته است. تحلیلِ قاب‌بندیِ اسناد سیاستی نشان می‌دهد که «حکمرانی»، اغلب به‌عنوان پاسخی برای حل مناقشه‌های اجتماعی پیرامون هوش مصنوعی عرضه می‌شود و بر نقش‌های فعال‌تر دولت و جامعه تأکید می‌گردد، هرچند در عمل با خطرِ کلی‌گویی یا گرفتارشدن در سازکارهای کم‌اثر مواجه است (Yu et al., 2022: 2646–2673). در موج جدیدِ بحث‌ها پیرامون هوش مصنوعی مولد نیز هشدار داده شده که بسیاری از ابتکارات بین‌المللی به «مدیریت ریسک» فروکاسته می‌شود و نقش عمومی و جهت‌دهی دموکراتیکِ فناوری، کماکان محدود است؛ وضعیتی که می‌تواند به تداوم عدم توازن قدرت میان چند کشور/شرکت و سایر جوامع بیانجامد (Sui & Shen & Zhou, 2024). برای ایران، این بدان معنا است که اگر «دیپلماسی داده» و حضور فعال در سازکارهای استانداردگذاری و گفت‌وگوهای منطقه‌ای/جهانی جدی گرفته نشود، حتی در صورت توسعه داخلی، قواعد بازیِ جهانی می‌تواند به‌شکلی ترسیم شود که زبان و هویت ایرانی را در حاشیه قرار دهد.

از این‌رو در چشم‌انداز نظریِ این پژوهش:

۱) هوش مصنوعی نه صرفا ابزار، بلکه زیرساخت تولید معنا و عرصه جدیدی برای اعمال «استعمار داده» و «حکمرانی نرم الگوریتمی» است و لذا سیاستگذاری آن، به‌طور جدایی‌ناپذیر، بخشی از سیاستگذاری فرهنگی و حاکمیتی است؛

۲) عدالت فرهنگی در هوش مصنوعی نیازمند مداخله فعال در لایه‌های زیرساختی داده، معیارها، ممیزی و حکمرانی مشارکتی (و نه صرفا تولید محتوا) است تا از بازتولید هژمونی‌های الگوریتمی و یکدست‌سازی فرهنگی جلوگیری کند؛

۳) زبان فارسی و مؤلفه‌های هویتی ایران می‌بایست به «استانداردهای قابل سنجش» و «نهادهای پاسخگو» ترجمه شود تا سیاستگذاری پیرامون آن، از سطح آرمان‌های کلی به سطح اقدامات اجراییِ مؤثر و مستقل منتقل شود.

٤. هویت ملی ایران در مواجهه با هوش مصنوعی

هویت ملی در ایران، در روایت رسمیِ پس از انقلاب، معمولا به‌صورت ترکیبی پویا از سه مؤلفه ایرانی، اسلامی و انقلابی صورت‌بندی می‌شود (شورای عالی انقلاب فرهنگی، ۱۳۹۱)

اهمیت این سه‌گانه برای بحث هوش مصنوعی در این است که هر مؤلفه، هم «حوزه‌های حساس» خاص خود را دارد (مثلا زبان و تاریخ، مرجعیت و مناسک، ارزش‌های سیاسی و نمادهای انقلاب) و هم «آسیب‌پذیری‌های» متفاوتی در برابر فناوری‌های مولد و توصیه‌گر دارد. برای نمونه، اگر هوش مصنوعی در حوزه خدمات دینی آنلاین یا پاسخگویی فقهی/کلامی به‌کار گرفته شود، مسئله فقط صحت اطلاعات نیست؛ بلکه موضوع «مرجعیت»، «اعتماد» و «شیوه صورت‌بندی پرسش و پاسخ» است. همچنین در حوزه هویت ایرانی، مسئله فقط تولید متن فارسی نیست؛ بلکه دقت در بینامتنیتِ ادبی، حساسیت‌های تاریخی و جلوگیری از کلیشه‌سازی درباره ایران و ایرانی است. در مؤلفه انقلابی نیز مسئله به روایت‌پردازی سیاسی، بازنمایی تاریخ معاصر و مدیریت قطبی‌سازی و اطلاعات نادرست گره می‌خورد؛ به‌ویژه وقتی هوش مصنوعی به تولیدکننده انبوهِ محتوا و «تقویت‌کننده» چرخه‌های رسانه‌ای تبدیل می‌شود.

٤-١. بُعد اسلامی هویت ملی: مرجعیت، مناسک و «تفسیر/فتوا» در محیط‌های مولد

اگر لایه اسلامیِ هویت ملی صرفا به «حضور نمادهای دینی» فرو کاسته نشود، با شبکه‌ای از مرجعیت، اعتماد، شیوه انتقال دانش دینی و تجربه‌زیسته دینداری روبه‌رو خواهیم بود که در آن، اعتبارِ گفتار دینی معمولا از مسیر زنجیره‌های شناخته‌شده (استاد-شاگرد، نهاد آموزشی، نظام اجتهاد و سازوکارهای پرسش و پاسخگویی) تثبیت می‌شود. در شرایطی‌که، ورود هوش مصنوعی مولد این الگو را دگرگون می‌کند؛ چراکه سامانه می‌تواند از «کانال انتقال» به «کنشگرِ تولیدکننده پاسخ» تبدیل شود و در جایگاه چت‌بات[5] دینی، مشاور اخلاقی، یا پاسخ‌گوی پرسش‌های اعتقادی ظاهر گردد؛ درست در همین نقطه، مسئله اصلی برای سیاستگذاری هویت‌محور این است که کاربر، پاسخ را به‌مثابه «سخن معتبر» تجربه می‌کند، حتی وقتی در پس‌زمینه آن، نه روش اجتهادی روشن وجود دارد و نه امکان پیگیری مسئولیتِ علمی و نهادی. چنان‌که پیشتر اشاره شد پژوهش‌های حکمرانی هوش مصنوعی مولد بر این تأکید می‌کنند که این سامانه‌ها در معرض خطاهای خاصی مانند توهم‌زایی، نشت اطلاعات حساس و دشواریِ کنترل و راستی‌آزمایی هستند و همین ویژگی‌ها، ریسک را در حوزه‌های ارزش‌محور (از جمله دین) تشدید می‌کند (Taeihagh, 2021).

در چنین زمینه‌ای، «ریسک هویتیِ اسلامی» را می‌توان در سه سازکار اصلی دید: (۱) تضعیف

5. Chatbot

مرجعیت از طریق برابرسازیِ صوریِ همه پاسخ‌ها (پاسخ عالم، پاسخ سامانه، پاسخ کانال ناشناس) در یک رابط کاربری یکسان؛ (۲) تقلیل‌گرایی معرفتی، به‌معنای تبدیل مسائل پیچیده فقهی/کلامی به نسخه‌های کوتاه و فوری که بافت تاریخی، اختلاف‌نظرها و شروط استنباط را حذف می‌کند و (۳) جابه‌جایی آرامِ معیار دینداری از «التزام آگاهانه» به «مصرفِ توصیه‌های فوری» که در بلندمدت می‌تواند صورت‌بندی هویت دینی را فردگرایانه‌تر کرده و پیوندهای آن را با نهادهای دانشی تضعیف کند (Latifi, 2024).

با تمام این‌ها، برخی فرصت‌های واقعی نیز در این زمینه قابل شناسایی است: اگر راهبرد ملی، به‌جای مواجهه سلبیِ کلی، به‌سمت «سامانه‌های بازیابی و آموزشِ تحت نظارت متخصص» حرکت کند، هوش مصنوعی می‌تواند در سازمان‌دهی منابع، تسهیل دسترسی به متون و آموزش لایه‌مندِ مباحث به طلاب و دانشجویان مفید باشد (Bartsch et al., 2025)، اما شرط تحقق این فرصت‌ها، روشن‌بودن «زنجیره پاسخگویی» و امکان ممیزی و تصحیح نهادی است، نه اتکای صرف به محبوبیت و مقبولیت عمومی.

۴-۲. بُعد ایرانی هویت ملی: زبان فارسی، حافظه تاریخی-ادبی و بازنمایی فرهنگی در داده و مدل

در بُعد ایرانی هویت ملی، زبان فارسی دارای نقشی زیرساختی است. چراکه هم حامل حافظه ادبی و تاریخی به شمار می‌رود و هم میدانِ بازتولید پیوندهای اجتماعی و احساسِ تعلق است. حال، نقش‌آفرینی هوش مصنوعی مولد نیز، دقیقا در همین نقطه برجسته می‌شود (Bommasani, 2022)؛ چراکه چنین سامانه‌ای، «تولید متن»، «بازنویسی» و «خلاصه‌سازی» را در مقیاس انبوه ممکن می‌سازد و به‌تدریج می‌تواند سبک نوشتار عمومی، معیارهای صحت و فصاحت و حتی دستورکارهای فرهنگی را شکل دهد مسئله کلیدی برای ایران این است که کیفیت و جهت‌گیری خروجی‌های فارسی، به‌طور مستقیم تابع سیاست داده (چه متونی گردآوری و وزن‌دهی می‌شود)، سیاست ارزیابی (چه چیزی «خروجی خوب» تلقی می‌شود) و سیاست استقرار (این سامانه‌ها در آموزش، رسانه، یا خدمات عمومی با چه محدودیت‌هایی به کار می‌رود) است. ادبیات سوگیری در سامانه‌های داده‌محور نشان می‌دهد که این اختلال، ضرورتا برآمده از «سوءنیت» نیست، بلکه می‌تواند در گردآوری، برچسب‌گذاری، پیش‌پردازش و حتی در انتخاب متغیرهای ظاهرا خنثی نیز بازتولید شود (Ntoutsi et al., 2020)؛ بنابراین،

اگـر «پیکره‌داده[۶]»هـای فارسـی نامتـوازن باشـد (مثـلا، صرفـا بـر چنـد منبـع محـدودِ رسـانه‌ای یـا سـبک معیـارِ رسـمی متمرکـز شـود)، خروجی‌هـا نیـز ناخواسـته به‌سـمت یکدست‌سـازی زبانـی و حاشـیه‌رانی گونه‌هـای بیانـیِ متنـوع میـل می‌کنـد.

در ایـن زمینـه، دو وجـه ریسـک هویتـی قابـل اشـاره اسـت: وجـه نخسـت، «کاهـش تنـوع مشـروع» درون زبـان و فرهنـگ اسـت؛ یعنـی گویش‌هـا، سـبک‌های ادبـی و صداهـای اجتماعـیِ کم‌نماینـده، در رقابـت الگوریتمـی بـه حاشـیه رانـده می‌شـود. وجـه دوم نیـز، «بازنمایـی بیرونـیِ کلیشـه‌ای» اسـت؛ چراکـه بسـیاری از سـامانه‌های مولـدِ عمومـی، مبتنـی بـر منابـع جهانـی آمـوزش می‌بیننـد و اگـر راهبـردی ملـی بـرای تغذیـه مدل‌هـا از منابـع داخلـی معتبـر، متنـوع و مسـتند فراهـم نباشـد، کلیشـه‌ها و خطاهـای بازنمایی‌شـده می‌توانـد بـا ظاهـری فنـی و قاطـع، در خروجی‌هـا تثبیـت شـود. در عیـن حـال کـه در برابـر چنیـن تهدیداتـی، فرصت‌هایـی شـامل «نـگاه بـه زبـان فارسـی به‌مثابـه کالای عمومـی دیجیتـال» و «سـرمایه‌گذاری روی پیکره‌داده‌هـای بـاز (کُرپوس‌هـای دسـترس‌پذیر)»، «معیارهـای سـنجش (بنچمارک‌هـای) بومـی» و «چرخـه بهبـود مسـتمر»، (Colace et al., 2025) می‌توانـد حضـور فارسـی را در زیسـت‌بوم هـوش مصنوعـی تقویـت کـرده و وابسـتگی بـه مدل‌هـای بیرونـی را کاهـش دهـد.

٤-٣. بُعد انقلابی هویت ملی: روایت سیاسی و عدالت‌خواهی در عصر محتوای مولد

بُعـد انقلابـی هویـت ملـی در ایـران، عـلاوه بـر لایه‌هـای ارزشـی (اسـتقلال، عدالت‌خواهـی، نفـی سـلطه)، بـا «روایـت» نیـز پیونـد خـورده اسـت. روایـت از گذشـته نزدیـک، روایـت از کنـش جمعـی و روایـت از نسـبت ایـران بـا جهـان. در عصـر محتـوای مولـد، ظرفیـت تولیـد انبـوه متـن، تصویـر و صـدا، میـدان روایـت را دگرگـون کـرده اسـت. جعـل عمیـق[۷]، بازترکیـب گزینشـی داده‌هـا و تولیـد روایت‌هـای ظاهـرا مسـتند می‌توانـد هزینـه دسـتکاری ادراک عمومـی را کاهـش دهـد و در عیـن حـال، کار تشـخیص حقیقـت را بـرای کاربـر عـادی دشـوارتر کنـد (Alrhmoun et al., 2020; Venugopal et al., 2024; Abrar et al., 2025). در ایـن وضعیـت، سیاسـتگذاری هویت‌محـور بایـد بپذیـرد کـه مسـئله فقـط «محتـوای نادرسـت» نیسـت، بلکـه تغییـر سـاختار اعتمـاد عمومـی اسـت. زیـرا، وقتـی مـرز میـان سـند، تفسـیر و جعـل مبهـم می‌شـود، جامعـه یـا در بدبینـی فراگیـر گرفتـار می‌شـود و یـا ممکـن اسـت بـه اتـکای بی‌رویـه بـه منابـع همسـو دچـار شـود.

۶. کُرپوس/ Corpus/ Corpora

7. Deepfake

۵. نقشه «ریسک هویتی» و حوزه‌های حساس در ایران: از محتوا تا زیرساخت حکمرانی

با کنار هم قرار دادنِ سه بُعد ایرانی، اسلامی و انقلابی، می‌توان نقشه‌ای از حوزه‌های حساس ترسیم کرد که در آن، ریسک هویتی صرفا پیامدِ یک خروجی نامناسب نیست، بلکه حاصلِ برهم‌کنشِ داده، معیارهای ارزیابی، طراحی رابط، مدل کسب‌وکار پلتفرم و سازکار پاسخگویی است. رویکردهای معاصر در حکمرانی هوش مصنوعی نشان می‌دهد که دولت‌ها معمولا میان دو الگو در نوسان‌اند: تأکید بر اصول اخلاقیِ کلی و انعطاف بالا، یا حرکت به‌سمت قواعد حوزه‌مند و ریسک‌محور؛ و در عمل، بسیاری از راهبردهای ملی، ترکیبی و «هیبریدی» است (Radu, 2021). برای ایران، ریسک‌محورکردنِ مسئله هویت این مزیت را دارد که حوزه‌ها را بر اساس «شدت پیامد» و «دشواری جبران»، سطح‌بندی می‌کند و مانع از سیاست‌های یکسان‌ساز می‌شود.

بر این مبنا، حوزه‌های «ریسک بسیار بالا» شامل کاربردهایی است که به‌طور مستقیم با مرجعیت دینی، آموزش رسمی ارزش‌ها و باورها و روایت تاریخ معاصر/انقلاب سروکار دارد؛ زیرا در این حوزه‌ها، خطا، فقط خطای اطلاعاتی نیست، بلکه خطای هویت‌ساز است و می‌تواند اعتماد نهادی و انسجام معنایی را مختل کند. حوزه‌های «ریسک بالا» شامل آموزش زبان و ادبیات فارسی، تعاملات فقهی-مذهبی، تولید محتوای رسانه‌ای عمومی و سامانه‌های توصیه‌گر محتوایی است؛ به‌ویژه توصیه‌گرها به این دلیل حساس است که نه تنها «چه چیزی دیده می‌شود»، بلکه «چه چیزی کمتر دیده می‌شود» را نیز تعیین می‌کند و از این راه، دستورکار فرهنگی را شکل می‌دهد (رحمانی و همکاران، ۱۴۰۳). حوزه‌های «ریسک متوسط/پایین» نیز شامل کاربردهای کم‌تماس (کمترمرتبط) با ارزش‌ها و هویت (مثلا بهینه‌سازی صنعتی) است که می‌توان در آنها بیشتر بر استانداردهای فنی و پاسخگویی سازمانی تکیه کرد، هرچند اصلِ امکان ممیزی و رسیدگی به پیامدهای پیش‌بینی‌نشده، حتی در چنین سطحی نیز می‌بایست محفوظ بماند.

بنابراین، سطح‌بندی ریسک، به‌طور مستقیم به نوع مداخله تنظیم‌گرانه جهت می‌دهد. در ریسک بسیار بالا، سیاستگذار ناگزیر است به‌سمت مجوزدهی حوزه‌مند، ضرورت حضور عامل انسان در حلقه تصمیم‌گیری (نه صرفا در سطح نمادین)، ممیزی پیشینی و پسینی و

سازکارهای بازخورد و جبران حرکت کند؛ زیرا در این حوزه‌ها، هزینه اجتماعیِ خطا بالا بوده و امکان اصلاح دیرهنگام محدود است. بنابراین، در ریسک بالا، می‌توان ترکیبی از قواعد الزام‌آور و خودتنظیمیِ تحت نظارت را به‌کار گرفت.

٦. ترجمه «ریسک هویتی» به الزامات فنی و نهادی

پس از شناسایی حوزه‌های حساس و تعیین سطوح مختلف ریسک، پرسش محوری این است که چگونه می‌توان مفهوم انتزاعی «ریسک هویتی» را به الزامات عملیاتی، قابل اندازه‌گیری و اجرایی تبدیل کرد که در تمام مراحل چرخه حیات سیستم‌های هوش مصنوعی (از فاز طراحی و توسعه گرفته تا آموزش، استقرار و پایش مستمر) مورد استفاده قرار گیرد. در بُعد فنی، اولین نقطه مداخله، لایه «داده» است؛ زیرا سوگیری‌ها و کاستی‌های داده‌محور، نه تنها در قالب تبعیض‌های آشکار، بلکه به شکل‌های پیچیده‌تری مانند پاکسازی ناخواسته ظرافت‌های فرهنگی، همگون‌سازی اجباری سبک‌ها و روایت‌ها و غلبه نامتعارف روایت‌های خارجی درباره ایران، در مدل‌ها بازتولید شده و در نهایت به‌صورت «دانش رسمی» در خروجی‌های سیستم‌ها تثبیت می‌شود. پژوهش‌های مرتبط با سوگیری در سامانه‌های داده‌ای نشان می‌دهد که مسیرهای ایجاد این اختلال، از مرحله گردآوری اولیه داده‌ها، انتخاب نمونه‌ها، برچسب‌گذاری و حتی در فرآیند پیش‌پردازش و تأثیرات اجتماعیِ موجود در داده‌ها، بسیار گسترده است و رفع آنها نیازمند رویکردی ترکیبی و چندبعدی است که همزمان از راهکارهای فنی تخصصی و سازکارهای حکمرانی مؤثر بهره ببرد (Ntoutsi et al., 2020: 5-8). در این جهت، برای حوزه‌های هویتی پرریسک، الزام اساسی می‌بایست «حاکمیت داده» باشد که شامل مستندسازی کامل منشأ داده‌ها، تعیین معیارهای شفاف برای ورود و حذف داده‌ها، ثبت شفاف تصمیم‌های پیش‌پردازش و تعریف شاخص‌های سنجش است که به وضوح نشان دهد مدل‌ها در برابر مفاهیم حساس هویتی (مانند دین، تاریخ، تنوع قومی و زبانی و سبک‌های زندگی) چه خطاهایی را نشان می‌دهند و این خطاها، در چه شرایط و زمینه‌هایی تشدید می‌شود.

سطح بعدی که نیازمند توجه ویژه است، «طراحی مدل و تعامل با کاربر» است؛ در این سطح می‌بایست به وضوح تعیین شود که سامانه تا چه حد اختیار ارائه پاسخ‌های قطعی را دارد، در چه زمان‌هایی باید عدم قطعیت و محدودیت‌های دانشی خود را آشکارا اعلام کند و در چه نقاط حساسی ضروری است تا فرآیند تصمیم‌گیری به متخصصان انسانی واگذار شود. این

همـان موقعیتـی اسـت کـه مفهـوم «انسـان در حلقـه تصمیـم» از یـک شـعار کلـی تبدیـل بـه یـک طـرح عملیاتـی دقیـق می‌شـود؛ به‌ویـژه در کاربردهـای حساسـی کـه به‌طـور مسـتقیم بـا مرجعیـت دینـی، آمـوزش رسـمی ارزش‌هـا و مفاهیـم، یـا روایت‌هـای تاریخـی مرتبـط اسـت و خطاهـای آن، پیامدهـای بلندمـدت هویت‌سـاز خواهـد داشـت.

در سـطح نهـادی نیـز، تبدیـل «ریسـک هویتـی» بـه سیاسـت‌های مؤثـر مسـتلزم ایجـاد ظرفیت‌هـای حکمرانـی هدفمنـد و شـفاف اسـت. در قالـب ایـن ظرفیت‌هـا، می‌بایسـت آشـکارا مشـخص شـود کـه «چـه نهـاد یـا شـخصی/جایگاهی» مسـئول صـدور مجوزهـای اولیـه اسـت، «چـه نهـاد مسـتقلی» وظیفـه ممیـزی مسـتمر سیسـتم‌ها را بـر عهـده دارد، «چـه سـازکاری» بـرای پاسـخگویی در قبـال خطاهـای هویتـی تعریـف شـده اسـت و «چـه فرآینـد تصمیم‌گیـری نهـادی» قـادر اسـت اختلافـات میـان نهادهـای فرهنگـی، دینـی و علمـی را در زمینـه اسـتانداردهای هـوش مصنوعـی حل‌وفصـل کنـد. مطالعـات حکمرانـی هـوش مصنوعـی نشـان می‌دهـد کـه دولت‌هـا در عصـر کنونـی هـم بـا چالـش سـرعت و گسـتردگی گذارهـای فناورانه-اجتماعـی و هـم بـا ریسـک‌هایی روبـه‌رو هسـتند کـه جوهـره آنهـا «پیش‌بینی‌ناپذیـری» اسـت؛ از ایـن‌رو، سـاختارهای حکمرانـی مؤثـر می‌بایسـت همزمـان بـه دو توانایـی کلیـدی مجهـز شـوند: «حمایـت از نـوآوری و خلاقیـت» و در عیـن حـال، «مدیریـت مؤثـر ریسـک‌های بالقـوه» (Taeihagh, 2021: 145-148). بـرای تطبیـق ایـن رویکـرد بـا شـرایط ایـران، یکـی از ضرورت‌هـا، فاصلـه گرفتـن از مدل‌هـای تصمیم‌گیـری سلسـله مراتبـی و تمرکـززده و نزدیـک شـدن بـه تعریـف زنجیـره‌ای شـفاف از نقش‌هـا و مسـئولیت‌ها اسـت. زنجیـره‌ای کـه می‌توانـد از «کمیته‌هـای تخصصـی» هـر حـوزه (متشـکل از متخصصـان در زمینه‌هـای دیـن، آمـوزش، رسـانه، تاریـخ و فرهنـگ) بـرای تعییـن خطـوط قرمـز و معیارهـای ارزیابـی گرفتـه تـا «نهـاد ارزیابـی مسـتقل» بـرای آزمون‌پذیـری و تأییـد ادعاهـای شـرکت‌های فنـاوری و در نهایـت «سـازکار پاسـخگویی» کـه خطاهـای هویتـی را بـا جدیتـی هماننـد خطاهـای امنیتـی و حریـم خصوصـی پیگیـری کنـد، تشـکیل شـود. چنیـن رویکـردی ضمـن جلـب اعتمـاد عمومـی، امـکان پاسـخگویی، اصـلاح و بهبـود مسـتمر را در برابـر چالش‌هـای پیچیـده هـوش مصنوعـی فراهـم مـی‌آورد.

۶-۱. از مفهوم تا ابزار: عدالت فرهنگی به‌مثابه چارچوب راهنمای سیاستگذاری و طراحی

بـرای آنکـه عدالـت فرهنگـی از سـطح هنجـار بـه سـطح ابـزار منتقـل شـود، توجـه بـه سـه مجموعـه مکمل الزامـی اسـت: ۱) مجموعـه داده‌ای، ۲) مجموعـه سـنجش/ممیزی و ۳) مجموعـه نهـادی. در مجموعـه

داده‌ای، اصل بر این است که ترکیب پیکره‌داده‌ها و تصمیم‌های مربوط به وزن‌دهی منابع، بخشی از «تصمیم هویتی» است و می‌بایست مستندسازی شده و قابل بازبینی باشد. در مجموعه سنجش/ممیزی، می‌بایست از سنجه‌های صرفا فنی (روانی متن، دقت دستوری) فراتر رفته و به سنجش تحریف مفاهیم کلیدی، کلیشه‌سازی، حذف تنوع زبانی و سوگیری در بازنمایی تاریخ و دین توجه شود؛ این کار نیز مستلزم تعریف دقیق ارتباط «لایه‌های شفافیت» (شفافیت در داده، در منطق مدل و در فرایند تصمیم‌گیری) با حوزه‌های ذی‌نفع از یک‌سو و دستیابی به «زبان مشترک» میان سیاستگذاران و متخصصان فنی، از سوی دیگر است (Graziani et al., 2022). مجموعه نهادی نیز معطوف به چگونگی طراحی «شفافیت» در تعامل انسان-سامانه است تا به‌واسطه آن، همزمان که زمینه ایجاد «اعتمادِ بی‌جا» را محدود می‌کند، امکان ارزیابی و تصحیح را افزایش دهد. در این چشم‌انداز، آنچه در زمینه ریسک هویتی در ایران آشکار به نظر می‌رسد، ضرورت توجه همزمان به سه لایه: ۱) معنا و روایت (دین، زبان، تاریخ/انقلاب)، ۲) زیرساخت (داده، مدل، پلتفرم) و ۳) حکمرانی (پاسخگویی، ممیزی، مشارکت) است.

۶-۲. شفافیت، تبیین‌پذیری و ممیزی فرهنگی: از «قابل توضیح بودن» تا «قابل داوری بودن»

در حوزه‌هایی که حساسیت هویتی بالا است، صرف ارجاع به «پیچیدگی مدل» یا «دامنه طبیعی خطاها» پذیرفتنی به نظر نمی‌رسد؛ بلکه لازمه آن، قائل شدن تمایز میان انواع شفافیت است. بخشی از شفافیت، فنی است (مانند دسترسی به مشخصات داده و پیکره‌داده)، بخشی از شفافیت، فرایندی است (اینکه چه کسی تصمیم گرفته چه داده‌ای وارد شود) و بخشی از آن، ارتباطی است (اینکه برای کاربر، حدود قابلیت و محدودیت سامانه چگونه بیان می‌شود). این درحالی است که واژگان و انتظارات این حوزه نیز با سطحی از آشفتگی همراه است و می‌بایست مفاهیمی مانند «قابل‌تفسیر بودن»، «قابل‌توضیح بودن» و «شفافیت»، با تعریف‌های دقیق و قابل اجرا از یکدیگر متمایز شود تا بتوان آنها را وارد مقررات و ارزیابی کرد (Graziani et al., 2022). برای آنکه مفهوم ممیزی فرهنگی از سطح یک توصیه هنجاری فراتر رود، می‌بایست به شاخص‌هایی عملیاتی و پروتکل‌هایی قابل اجرا ترجمه شود. در این زمینه، دست‌کم سه سطح سنجش قابل تفکیک است: نخست، سطح داده که در آن، تنوع و نمایندگی پیکره آموزشی از حیث جغرافیا، سبک زبانی، گونه‌های نوشتاری، حوزه دانشی و نسبت منابع داخلی معتبر سنجیده می‌شود؛ دوم، سطح خروجی که در آن، بسامد و شدت تحریف مفاهیم کلیدی هویتی،

کلیشه‌سازی در بازنمایی تاریخ و دین، حذف تنوع مشروع درون‌فرهنگی، نرخ خطای ارجاع و میزان آشکارسازی عدم‌قطعیت بررسی می‌گردد؛ و سوم، سطح فرایند تصمیم که در آن، حدود مداخله انسانِ متخصص، ثبت تغییرات مدل، مستندسازی به‌روزرسانی‌ها و سازوکار رسیدگی به خطاهای هویتی ارزیابی می‌شود. بر این اساس، ممیزی فرهنگی می‌تواند در سه مرحله طراحی گردد: ارزیابی پیشینی پیش از استقرار، پایش حین بهره‌برداری در کاربردهای حساس، و بازبینی پسینی بر مبنای شکایات، رخدادها و آزمون‌های تنش. نتیجه این فرایند آن است که «قابل توضیح بودن» صرف، جای خود را به «قابل داوری بودن» می‌دهد؛ یعنی سامانه نه فقط از حیث فنی، بلکه از حیث آثار هویتی و فرهنگی نیز در معرض ارزیابی منظم قرار می‌گیرد.

برای ایران، بازتاب عملی/اجرایی این تمایز آن است که در حوزه‌های حساس هویتی، هدف سیاستگذار می‌بایست بر «قابل داوری کردن خروجی‌ها و فرایندها» تمرکز یابد. به بیان دیگر، ممیزی دوره‌ای، آزمون‌های تنش[8] روی موضوعات حساس، سنجش «نرخ توهم» و خطای ارجاع و نیز الزام به گزارش‌دهی درباره به‌روزرسانی‌ها و تغییرات مدل، می‌تواند از جمله اهداف سیاستگذاری در نظر گرفته شود. این نگاه، ممیزی فرهنگی را از سطح سلیقه‌ای یا واکنشی خارج کرده و به یک سازکار نهادمند تبدیل می‌کند که می‌تواند با معیارهای صریح (مثلا در خصوص تحریف تاریخی، نسبت‌دادن نادرست گزاره‌های دینی، یا تولید کلیشه‌های قومی) فعالیت کند و در صورت نقض معیارها، مسیر اصلاح و جبران را فعال کند.

۶-۳. حکمرانی مشارکتی و «نمایندگی فرهنگی»

اگر مسئله هویت، «زیرساخت معنا» دانسته شود، آنگاه می‌توان این پرسش را مطرح کرد که چه گروه‌هایی صلاحیت/ضرورت نمایندگی در تصمیم‌های داده‌ای، معیارگذاری و ممیزی را خواهند داشت؛ زیرا بدون نمایندگی واقعی، شفاف و مؤثر، عدالت فرهنگی به‌سرعت به اخلاق‌نامه‌های کلی و بی‌اثر تبدیل خواهد شد. پژوهش‌هایی که بر حکمرانی مشارکتی برای هوش مصنوعی تأکید دارد، نشان می‌دهد که مشارکت عمومی و ذی‌نفعان می‌تواند زمینه «افزایش شمول» و «کاهش شکاف بهره‌مندی» را فراهم کند، به‌شرط آنکه مشارکت صوری نبوده و به سازکارهای واقعی اثرگذاری (دستورکار، حق دسترسی به اطلاعات و امکان پیگیری) متصل شود (Moon, 2023). در ایران، طراحی «میان‌رشته‌ای و میان‌نهادیِ» مشارکت، از ترجیحات اساسی

8. stress tests

است. نهادهای دینی و فرهنگی، نه در مقام مصرف‌کننده یا ناظر بیرونی، بلکه می‌بایست در تعیین دستورکار داده‌ای و معیارهای ارزیابی نقش داشته باشند؛ در عین حال که این مشارکت (با تصریح قواعدی روشن) نمی‌بایست به تصلّب یا یک‌صدایی منجر شود. چراکه هدف نهایی، تثبیت یک روایت واحد به هر قیمت نیست، بلکه ایجاد سازکاری است که تنوع درون‌فرهنگی را بازتاب داده و در عین حال، با دقت بیشتری از حوزه‌های پرریسک پاسداری کند.

با این همه، فرضِ مشارکت نهادی نباید ما را از تعارض‌های محتمل درون ساختار حکمرانی غافل کند. در ایران، بخشی از مناقشات نهادیِ قابل پیش‌بینی، از تفاوت در مأموریت و منطق عمل نهادها ناشی می‌شود: نهادهای امنیت‌محور معمولا بر پیشگیری، کنترل ریسک و محدودسازی سوءاستفاده تمرکز دارند؛ در حالی که نهادهای فرهنگی، آموزشی و هویتی بیشتر نگران صیانت از تنوع زبانی، مرجعیت‌های معرفتی، میراث معنایی و پرهیز از مداخله بیش‌ازحد یک‌سویه هستند. از این‌رو، حکمرانی مشارکتی تنها در صورتی از سطح یک مطلوب انتزاعی فراتر می‌رود که با تفکیک روشن نقش‌ها، تعیین رسمی حدود صلاحیت و پیش‌بینی سازوکار رفع اختلاف همراه شود. یک طراحی حداقلی می‌تواند بر چهار پایه استوار باشد: ۱) تعیین مکتوب نقش نهادهای داده‌پرداز، مجوزدهنده، ممیز و پاسخگو؛ ۲) توافق قبلی بر معیارهای ارزیابی در حوزه‌های حساس؛ ۳) تفکیک نهادی میان مرجع صدور مجوز و مرجع ممیزی؛ و ۴) ارجاع اختلافات به یک کمیته میان‌رشته‌ای با حضور نمایندگان حقوقی، فنی، فرهنگی، آموزشی و دینی. بدون چنین ترتیباتی، مشارکت نهادی ممکن است به‌جای افزایش مشروعیت و کارآمدی، به موازی‌کاری، تعارض صلاحیت و کندی در سیاستگذاری بینجامد. در نگاه سیاستگذاری، یک راه عملی، ایجاد «شورای نمایندگی فرهنگی در هوش مصنوعی» در کنار نهاد تنظیم‌گر است؛ شورایی که ترکیب آن شامل زبان‌شناسان، تاریخ‌پژوهان، متخصصان تعلیم‌وتربیت، فقیهان و متکلمان، جامعه‌شناسان فرهنگ و نمایندگان صنایع بوده و خروجی آن، نه توصیه‌های کلی، بلکه «پروتکل‌های ممیزی» و «راهنمای داده» برای حوزه‌های حساس باشد.

٦-٤. اقتصاد سیاسی حکمرانی و مسئله سلطه پلتفرم‌ها

حتی اگر ایران بهترین معیارها و شوراها را طراحی کند، بدون توجه به اقتصاد سیاسی هوش مصنوعی، سیاست‌ها یا اجرا نمی‌شود و یا به سطحی از «پوشش نمادین» فرو کاسته خواهد شد. در ادبیات سیاستی، تحلیل راهبردهای ملی نشان می‌دهد که بسیاری از دولت‌ها به‌جای

مقـررات سـخت، بـه اخلاق‌محـوری و ابهـام کارکـردی روی آورده‌انـد و ایـن شـرایط، رابطـه دولـت و بخـش خصوصـی را به‌سـمت ترتیبـات «هیبریـدی» سـوق داده اسـت (Radu, 2021: 178-193). هم‌زمـان، بازیگـران خصوصـی بـا پیشـبرد اسـتانداردهای اخلاقـی و خودتنظیمـی می‌تواننـد مسـیرهای وابسـتگی ایجـاد کننـد؛ یعنـی اسـتانداردهایی کـه زودتـر «جـا می‌افتـد»، بعدهـا به‌علـت هزینـه تغییـر، به‌صـورت مسـیر غالـب باقـی می‌مانـد، حتـی اگـر بـا منافـع عمومـی کامـلا همسـو نباشـد (Auld et al., 2022). در نتیجـه، اگـر ایـران «حاکمیـت هویتـی» را جـدی می‌گیـرد، می‌بایسـت هـم در سـطح زیرسـاخت (پیکره‌هـای بـاز فارسـی، بنچمارک‌هـا و خدمـات پایـه) سـرمایه‌گذاری کنـد و هـم در سـطح تنظیم‌گـری، از «وابسـتگی نهـادی» بـه پلتفرم‌هـای غیرشـفاف بکاهـد؛ زیـرا در غیـر این‌صـورت، قواعـد فرهنگـی و سیاسـی میـدانِ عمومـی عمـلا توسـط شـرکت‌هایی تعییـن می‌شـود کـه پاسـخگویی آنـان بـه جامعـه ایرانـی محـدود بـوده و منطـق فعالیـت آنـان نیـز عمدتـا تجـاری اسـت. آن‌چنان‌کـه برمی‌آیـد، مناسـبات قـدرت در حـوزه هـوش مصنوعـی، تنظیم‌گـری را به‌سـمت منافـع خصوصـی متمایـل می‌کنـد و اگـر «منافـع عمومـی» به‌عنـوان محـور مقـررات قـرار نگیـرد، نتیجـه می‌توانـد نوعـی حکمرانـی خودکامـهِ فناوری‌محـور را بـه بـار آورد (Dignam, 2020). بنابرایـن، روایـت بومـی ایـن دیـدگاه بـرای بحـث هویـت، ضـرورت توجـه همزمـان سیاسـتگذاری بـه «قـدرت بـازارِ مدل‌هـا» و «قـدرت زیرسـاخت‌های محتوایـی و توزیعـی» (پلتفرم‌هـا و توصیه‌گرهـا) اسـت.

۶-۵. مسیر یادگیری سیاستی: پایلوت، ارزیابی و بازنگری‌های دوره‌ای

از آنجـا کـه هـوش مصنوعـی (به‌ویـژه مولدهـا) بـا سـرعت بـالا تغییـر می‌کنـد و ریسـک‌های آن نیـز در عمـل آشـکار می‌شـود، سیاسـتگذاری می‌بایسـت «یادگیرنـده» باشـد. بـه عبارتـی، بـا پایلوت‌هـای محـدود آغـاز کنـد، شـواهد را گـرد آورد، از سـازکار اصـلاح سـریع برخـوردار باشـد و قواعـد را به‌صـورت دوره‌ای بازنگـری کنـد. پژوهش‌هـای مرتبـط بـا «یادگیـری حکمرانـی» در حوزه‌هـای دیگـر نشـان می‌دهـد کـه وقتـی مشـارکت ذی‌نفعـان بـه فراینـد اجـرا گـره می‌خـورد و شـواهد به‌صـورت منظـم جمـع‌آوری می‌شـود، امـکان اصـلاح طراحـی و افزایـش کارآمـدی سیاسـت فراهـم می‌شـود (Challies et al., 2017: 288-303). بـرای ایـران، یـک طراحـی واقع‌گرایانـه ایـن اسـت کـه در هـر حـوزه حسـاس (مثـلا آمـوزش زبـان فارسـی، پاسـخ‌گویی دینـی، روایـت تاریخـی یـا توصیه‌گرهـای رسـانه‌ای)، چنـد پایلـوتِ تحـت نظـارت راه‌انـدازی شـود کـه در آنهـا، هـم شـاخص‌های فنـی (نـرخ خطـا، توهـم، پایـداری خروجـی) و هـم شـاخص‌های فرهنگی-اجتماعـی (اعتمـاد، برداشـت هویتـی، میـزان یکدسـت‌سـازی روایـت) سـنجیده شـده و نتایـج به‌طـور منظـم بـه نهـاد تنظیم‌گـر و

شـورای نمایندگـی فرهنگـی گـزارش شـود. ایـن رویکـرد، به‌جـای برخوردهـای افراطـی یـا تفریطـی (ممنوعیـت کامـل یـا رهاسـازی کامـل)، امـکان می‌دهـد تـا توسـعه فناورانـه ادامـه یابـد، امـا در مسـیری کنترل‌شـده و پاسـخگو حرکـت کنـد. به‌ویـژه بـرای حوزه‌هایـی کـه خطاهـای آن «بازگشـت‌ناپذیر» یـا «تشدیدشـونده» اسـت و می‌توانـد موجـب فرسـایش اعتمـاد عمومـی شـود.

۷. بسته سیاستی برای حکمرانی هوش مصنوعی و هویت ملی در ایران

چنان‌کـه در بخش‌هـای پیشـین اشـاره شـد، «هویـت ملـی» در عصـر هـوش مصنوعـی، صرفـا موضـوعِ نمادهـا و بازنمایی‌هـای فرهنگـی نبـوده، بلکـه بـه سـطح «زیرسـاخت‌های تولیـد معنـا» منتقـل شـده اسـت (Bareis & Katzenbach, 2021)؛ عرصـه‌ای کـه داده‌هـا، مدل‌هـا، معیارهـای ارزیابـی و سـازکارهای توزیـع محتـوا تعییـن می‌کنـد چـه دانشـی معتبـر جلـوه کنـد، چـه روایتـی بیشـتر دیده شـود و چـه سـبک زبانـی و ارزشـی به‌تدریـج «معیـار» تلقی شـود. از ایـن منظـر، بایسـته سیاسـتگذاری در ایران، اگـر بـا هـدف پیشـبرد هم‌زمـان توسـعه فناورانـه و صیانت/بازآفرینـی هویـت در نظـر گرفتـه شـود، ایـن اسـت کـه بسـته‌ای از اقدامـات به‌هم‌پیوسـته را طراحـی کنـد کـه صرفـا اخلاق‌نامـه یـا توصیـه عمومـی نبـوده، بلکـه بـه ترتیبـات نهـادی، دارایی‌هـای داده‌ای، معیارهـای سـنجش و سـازکارهای پاسـخگویی متصـل شـود؛ نکتـه‌ای کـه در ادبیـات حکمرانـی هـوش مصنوعـی نیـز بـا تأکیـد بـر ضـرورت ظرفیـت دولـت بـرای مدیریـت ریسـک‌ها و پیامدهـای ناخواسـته و نیـز طراحـی ترکیـب ابزارهـای تنظیم‌گـری مطـرح شـده اسـت (Taeihagh, 2021: 137-157). بـر همیـن اسـاس، در ادامـه، بسـته‌ای پیشـنهادی در شـش محـور ارائـه می‌شـود کـه در آن، هرچنـد محورهـا به‌صـورت مفهومـی از هـم تفکیـک شـده اسـت، امـا در اجـرا می‌بایسـت ماننـد اجـزای یـک معمـاری واحـد دیـده شـده و بـا حلقه‌هـای بازخـورد و یادگیـری تدریجـی بـه‌روز شـود. همچنان‌کـه راهبردهـای ملـی در سـایر کشـورها نیـز عمدتـا «ترکیبـی و هیبریـدی» بـوده و بسـیاری از آنهـا، به‌صورتـی عامدانـه و آگاهانـه، درجـه‌ای از انعطـاف/ابهـامِ کارکـردی را حفـظ کـرده اسـت تـا بتوانـد با سـرعتِ تحـول فنـاوری سـازگار شـود (Radu, 2021).

- محـور اول: تعریـف و تثبیـت حوزه‌هـای حسـاس هویتـی و سـطح‌بندی ریسـک. تـا زمانی‌کـه در شـاکله سیاسـتگذاری، رسـما مشـخص نشـود کـه کـدام کاربردهـا دارای «ریسـک هویتـیِ بـالا» اسـت، مداخله‌هـا یـا بیش‌ازحـد کلـی خواهـد شـد (و در نتیجـه اجراشـدنی نخواهـد بـود) و یـا صرفـا در واکنـش بـه بحران‌هـای رسـانه‌ای شـکل خواهـد گرفـت (و در نتیجـه، ناپایـدار و سـلیقه‌ای خواهـد بـود). در ایـران، حساسـیت‌ها معمـولا در سـه سـاحت درهم‌تنیـده ایرانـی، اسـلامی و انقلابـی

ظهـور می‌کنـد؛ امـا بـرای تبدیـل ایـن سـاحت‌ها بـه ابـزار حکمرانـی، ترجمـه آنهـا بـه «حوزه‌هـای عملیاتـی» گریزناپذیـر اسـت. بـرای نمونـه، پاسـخ‌گویی دینـی و فقهـی، آمـوزش رسـمی زبـان و ادبیـات فارسـی، روایـت تاریـخ معاصـر، توصیه‌گرهـای محتـوا در رسـانه‌های جمعـی، خدمـات عمومـیِ مبتنـی بـر چت‌بات‌هـا و سـامانه‌های تولیـد تصویر/ویدئـو بـرای اهـداف اطلاع‌رسـانی یـا تبلیغاتـی، از جملـه ایـن حوزه‌هـا اسـت. همچنیـن، سـطح‌بندی ریسـک می‌بایسـت دارای پیامـدِ اجرایـی روشـن باشـد.

بـرای پرهیـز از ابهـام اجرایـی، هـر سـطح از ریسـک بایـد بـه مجموعـه‌ای از الزامـات متناظـر ترجمـه شـود. بـرای مثـال، در حوزه‌هـای پرریسـک (ماننـد پاسـخگویی دینـی، آمـوزش رسـمی زبـان و ارزش‌هـا، روایـت تاریـخ معاصـر و سـامانه‌های توصیه‌گـرِ اثرگـذار بـر افـکار عمومـی) حضـور انسـانِ متخصـص در حلقـه تصمیـم، ثبـت رویدادهـا، الـزام بـه مستندسـازی داده و تغییـرات مـدل، آزمون‌هـای تنـش دوره‌ای و سـازوکار گزارش‌دهـی اجبـاری ضـروری اسـت. در برابـر، در کاربردهـای کم‌ریسـک‌تر می‌تـوان از تنظیم‌گـری سـبک‌تر، خودتنظیمـیِ نظارت‌شـده و ارزیابی‌هـای دوره‌ای بهـره بـرد. چنیـن تفکیکـی، امـکان می‌دهـد کـه سیاسـتگذاری نـه بـه سـمت سـخت‌گیری فراگیـر و پرهزینـه بلغـزد و نـه در حـد توصیه‌هـای کلـی و غیرقابـل سـنجش باقـی بمانـد.

- محـور دوم: ایجـاد و تقویـت پیکره‌هـا و دارایی‌هـای داده‌ایِ فارسـی و بومـی به‌مثابـه کالای عمومـی. بـدون داده، نـه مـدلِ زبانـیِ باکیفیـت سـاخته می‌شـود و نـه ارزیابـیِ معتبـر ممکـن اسـت؛ امـا مسـئله فقـط «کمیـت داده» نیسـت، بلکـه کیفیـت، تنـوع، مستندسـازی و امـکان ممیـزیِ مسـیر تبدیـل متن/صوت/تصویـر بـه داده آموزشـی نیـز بسـیار مهـم اسـت. از یک‌سـو، تجربه‌هـا و مطالعـات نشـان می‌دهـد کـه بسـیاری از شکسـت‌ها و خطاهـای پروژه‌هـای هـوش مصنوعـی از ضعـف زیسـت‌بوم داده ناشـی می‌شـود و «ظرفیـت داده‌ای»، پیش‌شـرط ظرفیـت هـوش مصنوعـی اسـت (Gröger, 2021: 98-108). از سـوی دیگـر، ادبیـات سـوگیری نیـز یـادآور می‌شـود کـه هریـک از مراحـل گـردآوری، نمونه‌گیـری، برچسـب‌گذاری و پیش‌پـردازش داده می‌توانـد بـا میزانـی از سـوگیری همـراه گشـته و سـپس در مـدل تقویـت شـود (Ntoutsi et al., 2020: 5-8). بنابرایـن اگـر پیکره‌داده‌هـای فارسـی و بومـی به‌صـورت «دارایـی عمومـی بـا اسـتانداردهای روشـن» سـاخته نشـود، ظرفیـت ملـی، ناگزیـر یـا بـه داده‌هـای پراکنـده و غیرقابـل اطمینـان تکیـه خواهـد کـرد و یـا بـه مسـیرهای وابسـتگی بـه پیکره‌هـای بیرونی/غیربومـی خواهـد افتـاد کـه به‌طـور سـاختاری، سـبک‌ها و ترجیحـات معنایـی دیگـری را در خروجی‌هـا تثبیـت می‌کنـد. بـر ایـن اسـاس، از حیـث اجرایـی، چنـد لایـه بـرای ایـن محـور قابـل پیشـنهاد اسـت: الـف) تولیـد و نگهـداری پیکره‌هـای متنـی و گفتـاریِ معاصـر و کلاسـیک

(ادبی، حقوقی، علمی، رسانه‌ای، دینی) با تنوع سبکی و منطقه‌ای، ب) سازکارهای حقوقی و فنی برای رعایت حریم خصوصی و حقوق پدیدآورندگان، پ) استاندارد مستندسازی و «ثبت مشخصات داده» برای شفاف‌کردن منشأ، روش پالایش و محدودیت‌ها و ت) طراحی دسترسیِ طبقه‌بندی‌شده (باز/مشروط/محدود) تا هم مسیر نوآوری هموار شده و هم از حوزه‌های حساس در برابر سوءاستفاده‌های احتمالی پاسداری شود.

- محور سوم: طراحی معیارها و بنچمارک‌های بومی برای سنجش عدالت فرهنگی و هویتی. با فرض فراوانی پیکره‌داده‌های فارسی و دسترس‌پذیریِ دارایی‌های داده‌ایِ بومی، آنچه کماکان می‌تواند چالش‌آفرین باشد، فقدان یا کاستی «معیارهای سنجش» (بنچمارک‌ها) است. چراکه بدون این معیارها، ساحت سیاستگذاری، ابزاری برای پاسخ‌خواهی (مطالبه بازخورد) ندارد و توسعه‌دهنده نیز مقصد روشنی نخواهد داشت؛ در نتیجه، بحث هویت به سلیقه و برداشت فردی فروکاسته می‌شود. برای رفع این چالش، می‌بایست دو نوع معیار هم‌زمان طراحی شود: معیارهای فنیِ عملکرد (اعم از صحت، انسجام، پایداری در برابر ورودی‌های مخرب و میزان «هذیان‌گویی») و معیارهای فرهنگی/هویتی (مانند تحریف مفاهیم دینی، نسبت‌دهی نادرست نقل‌قول‌ها، کلیشه‌سازی درباره گروه‌های قومی و زبانی یا یک‌سویه‌سازی روایت تاریخ معاصر). در این میان، نکته قابل توجه در طراحی معیارهای سنجش، تفاوت «شفافیت» در زبان سیاستگذاری و زبان فنی است و اگر اصطلاحات «تبیین‌پذیری»/«شفافیت»/«قابلیتِ تفسیر» به‌صورت مبهم به‌کار برده شود، هم ارتباط میان‌رشته‌ای مختل شده و هم الزام‌آوری مقررات کاهش می‌یابد؛ ازاین‌رو، استفاده از یک طبقه‌بندی روشن برای مفاهیم تبیین‌پذیری و شفافیت، به خودِ طراحی مقررات نیز کمک می‌کند (Graziani et al., 2022). علاوه بر این، بنچمارک‌های بومی بهتر است عمومی و قابل بازتولید باشد تا جامعه پژوهشی و حتی نهادهای مدنی بتوانند نتایج را راستی‌آزمایی کنند. چراکه پاسخگویی بدون امکان آزمون مستقل، می‌تواند به گزارش‌دهی صوری تبدیل شده و گمراه‌کننده باشد.

- محور چهارم: نهادینه‌سازی حکمرانی مشارکتی و سازکار نمایندگی فرهنگی در کنار نهاد تنظیم‌گر. یکی دیگر از مسائل پیش‌روی سیاستگذاری در این حوزه، مواجهه توأمان با پیچیدگی‌های فنی و اختلاف نظر ارزشی/تأویلی در عرصه تنظیم‌گری است. بنابراین، راه‌اندازی سازکار نمایندگی واقعی و مشارکت سازمان‌یافته و مؤثر نیز، یکی از ضرورت‌ها است که در صورت فقدان یا ناکارآمدی آن، تصمیم‌ها ممکن است در تفریطِ تکنوکراتیک گرفتار شود (و

در نتیجه، مشروعیت اجتماعی آن شکننده شود) و یا به افراط سیاسی‌کاری/برخورد سلیقه‌ای دچار شود (و در نتیجه آن، پیش‌بینی‌پذیری برای نوآوری را مخدوش سازد). ادبیات سیاستگذاری در این عرصه نشان داده است که «چارچوب‌بندی حکمرانی» در اسناد هوش مصنوعی، غالبا با وعده مشارکت اجتماعی همراه است، اما همزمان با چالش‌های جدی مانند «مصادره» توسط ذی‌نفعان قدرتمند و یا دشواری رسیدن به اجماع روبه‌روست (Ulnicane, 2024: 158-177). از این‌رو، سازکار مشارکت می‌بایست از طراحیِ نهادی دقیقی برخوردار باشد (و صرفا به دعوت عمومی بسنده نشود). بدین منظور، پیشنهاد اجرایی این مقاله، تشکیل «شورای نمایندگی فرهنگی در هوش مصنوعی» است که در کنار نهاد تنظیم‌گر رسمی فعالیت کرده و ترکیب آن نیز چندرشته‌ای (اعم از زبان و ادبیات، تاریخ، علوم دینی، علوم اجتماعی، ارتباطات، حقوق فناوری، نمایندگان صنعت و جامعه نوآور و ...) است. کارکرد این شورا، مبتنی بر «تبدیل دغدغه‌های هویتی به تولیدات اجرایی» خواهد بود که می‌تواند طیف گسترده‌ای از موضوعات (مانند: تهیه فهرست حوزه‌های حساس، راهنمای داده و پالایش، پروتکل ممیزی، قواعد انتشار/برچسب‌گذاری و روند رسیدگی به شکایت‌های عمومی) را دربرگیرد. در این محور نیز، با هدف پرهیز از صوری‌شدن، می‌توان «اختیار اثرگذاری» بر فرایند مجوزدهی/رتبه‌بندی حوزه‌های پرریسک (و انتشار دوره‌ای نتایج فعالیت آن) را برای شورای مذکور در نظر گرفت.

- محور پنجم: ارتقای دانش و ظرفیت میان‌رشته‌ای در نهادهای آموزشی، رسانه‌ای و فرهنگی/دینی. پایداری حکمرانی هویتیِ هوش مصنوعی، صرفا با اتکا به ابزارهای سخت (مجوز، ممنوعیت، جریمه) بسیار دشوار بوده و بلکه ناممکن است؛ چراکه بخش بزرگی از پیامدهای هویتی، از مسیرهایی مانند مصارف روزمره، آموزش غیرمستقیم و گردش محتوا در رسانه‌ها رخ می‌دهد. بنابراین، در گامی مهم، بایسته است بکارگیری «ظرفیت انسانی» برای فهم فرصت‌ها و چالش‌ها و نیز برای نقد و ارزیابی خروجی‌های مدل‌ها، بیش‌ازپیش مورد توجه قرار گیرد. همچنین در سطح دانشگاه و آموزش عالی، تجربه سیاستگذاری آموزشی نشان می‌دهد که بسیاری از چارچوب‌های موفق، همزمان سه سطح را پوشش داده است: ۱) سطح آموزشی (کاربرد درست ابزارها)، ۲) سطح حکمرانی (حریم خصوصی، پاسخگویی و قواعد استفاده) و ۳) سطح عملیاتی (زیرساخت و آموزش کارکنان) (Chan, 2023). در ایران، به‌طور ویژه، «میان‌رشته‌ای‌سازیِ» این محور پیشنهاد می‌شود که به‌واسطه آن، تربیت افرادی که بتوانند بین زبان و فرهنگ از یک‌سو و بین داده/مدل/ارزیابی از سوی دیگر، اتصال و ارتباط مؤثر

برقرار کنند، محقق شود. همچنین می‌بایست برای حرفه‌های کلیدی (مانند معلم، روزنامه‌نگار، مدیر فرهنگی، مبلغ/پژوهشگر دینی) دوره‌های مستمری طراحی شود تا «مصرف منفعلانه» جای خود را به «کاربرد مسئولانه و قابل پرسش» بدهد.

- محور ششم: دیپلماسی داده و همکاری بین‌المللی هدفمند و مشروط. یکی از مهم‌ترین نکات سیاستگذاری هوش مصنوعی در ایران، توجه به پیوستگی اجتناب‌ناپذیر با زنجیره‌های جهانی ابزار، استاندارد و جریان ایده است و اگر سیاستگذاری ایرانی (به‌رغم توسعه داخلی) در صحنه استانداردگذاری و همکاری علمی بین‌المللی غایب باشد، معیارها و الگوهای فراگیر ممکن است علیه منافع و حوزه‌های حساس هویتی آن شکل گیرد. در عین‌حال که همکاری بین‌المللی این عرصه نیز می‌بایست گزینشی و مشروط طراحی شود. به بیان دیگر، انتقال دانش و ابزارهای متن‌باز، همکاری‌های علمی در ارزیابی و ایمنی و پروژه‌های مشترک برای زبان‌های کم‌منبع/کمتر پردازش‌شده (از جمله، عضویت فعال و مشارکت در پروژه‌های بین‌المللی نهادهایی مانند یونسکو) می‌تواند به تقویت ظرفیت داخلی کمک کند؛ اما در حوزه‌هایی که خطر وابستگی نهادی یا انتقال ناخواسته داده‌های حساس وجود دارد، اعمال چارچوب‌های حقوقی و فنی سخت‌گیرانه، غیرقابل چشم‌پوشی است. در چشم‌انداز نظری، حکمرانی جهانی هوش مصنوعی به‌شدت ناهمگون و متأثر از تفاوت‌های نهادی و قدرت دولت‌ها است و همین ناهمگونی، هم فرصتِ ائتلاف‌های بدیل را فراهم می‌سازد و هم خطر حاشیه‌نشینی را افزایش می‌دهد (Stanger et al., 2024). بنابراین، «دیپلماسی داده» بهینه برای ایران هم شامل بُعد دفاعی (حفظ دارایی‌های داده‌ای و استقلال) است و هم بُعد آفندی/پیش‌بَرنده (اثرگذاری در شبکه‌های علمی و فنی و ارائه روایت و استاندارد بدیل در حوزه‌های فرهنگی) را در بر دارد (کرمی و متقی دستنایی، ۱۴۰۳). در عین حال، هرگونه سیاستگذاری واقع‌بینانه در ایران باید محدودیت‌های بیرونی را نیز به‌صراحت در نظر گیرد. تحریم‌های مالی و فناورانه، محدودیت دسترسی به خدمات برخی ابرپلتفرم‌ها و دشواری تأمین سخت‌افزارهای پردازشی پیشرفته، امکان اتکای پایدار و بی‌هزینه به زنجیره‌های غالب جهانی را کاهش می‌دهد. از این‌رو، بومی‌سازی در اینجا نباید به معنای انزوای فناورانه فهم شود، بلکه باید به‌مثابه راهبردی برای کاهش آسیب‌پذیری ناشی از وابستگی، تنوع‌بخشی به منابع داده و همکاری، تقویت زیرساخت محاسباتی داخلی و استفاده هدفمند از مشارکت‌های منطقه‌ای و بین‌المللی مشروط صورت‌بندی گردد. در چنین شرایطی، دیپلماسی داده نه یک امر تزئینی، بلکه بخشی

از شرط امکان سیاست هوش مصنوعی هویت‌محور است؛ زیرا بدون توجه به قیود دسترسی، حتی طراحی دقیق معیارها و نهادها نیز ممکن است در سطح آرمانی باقی بماند.

۸. نتیجه‌گیری

«هوش مصنوعی و هویت ملی در ایران» مسئله‌ای در سطح حکمرانی است؛ چراکه با گسترش مدل‌های مولد و سامانه‌های توصیه‌گر، هوش مصنوعی به میانجی اصلی تولید و توزیع معنا تبدیل شده و از آموزش و پژوهش تا حافظه جمعی، دین‌ورزی و سرگرمی را متأثر می‌کند. در چنین وضعیتی، بی‌تفاوتی سیاستی به‌معنای پذیرش معیارها و داده‌هایی است که خارج از حوزه اختیار ملی شکل می‌گیرد. در چشم‌انداز نظری نیز، هوش مصنوعی صرفا ابزار حکمرانی نبوده، بلکه می‌تواند خود به یک صورت نوظهور از حکمرانی بدل شود. بر این اساس، هم‌سوسازی توسعه هوش مصنوعی با هویت ایرانی-اسلامی-انقلابی مستلزم مداخله هم‌زمان در سه لایه زیرساخت، سنجش و نهاد/جامعه است. بسته پیشنهادی مقاله نیز بر همین مبنا ارائه شده است: تعریف حوزه‌های حساس و سطح‌بندی ریسک، سرمایه‌گذاری در پیکره‌داده‌های فارسی، طراحی معیارهای بومی عدالت فرهنگی، حکمرانی مشارکتی با سازکارهای ممیزی، ارتقای دانش میان‌رشته‌ای و دیپلماسی داده. در این میان، «پاسخگویی» (به‌معنای روشن بودن منشأ داده‌ها، مفروضات مدل، نحوه ارزیابی خروجی‌ها و تعیین مسئولیت در رخدادهای پرریسک) محور پیونددهنده این اجزا است.

افق سیاستگذاری هویت ملی در عصر هوش مصنوعی، در عین‌حال که فرصتی کم‌نظیر برای تقویت فارسی و توسعه ابزارهای بومی به‌شمار می‌رود، دربردارنده ریسک‌هایی چون هذیان‌گویی، دستکاری‌پذیری، تولید محتوای حقیقت‌نما و اختلاف‌افکن و نیز چالش‌هایی نظیر سوگیری و مالکیت فکری است. از این‌رو، «صیانت هویتی» در این عرصه، اگر به‌صورت محدودیت‌های کلی دنبال شود، نوآوری را تضعیف و کاربران را به پلتفرم‌های غیرقابل کنترل سوق خواهد داد؛ همچنان‌که اگر در قالب ساختاری یادگیرنده با ابزارهای شفاف داده، معیار، ممیزی، مشارکت و ظرفیت‌سازی طراحی شود، می‌تواند همزمان توسعه فناورانه، عدالت فرهنگی و امکان بازآفرینی فعال هویت را در زیست‌بوم دیجیتال تقویت کند.

تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده است.

References

Abrar, A., et al. (2025), "Religious Bias Landscape in Language and Text-to-Image Models: Analysis, Detection, and Debiasing Strategies", available at: https://arxiv.org/pdf/2501.08441v1

Alrhmoun, A., et al. (2020), "Decoding Hate: Using Experimental Text Analysis to Classify Terrorist Content", *The Global Network on Extremism and Technology (GNET)*, available at: https://gnet-research.org/wp-content/uploads/2020/09/GNET-Report-Decoding-Hate-Using-Experimental-Text-Analysis-to-Classify-Terrorist-Content.pdf

Auld, G. et al. (2022), "Governing AI Through Ethical Standards: Learning from the Experiences of Other Private Governance Initiatives", *Journal of European Public Policy*, No. 29, pp. 1822-1844. https://doi.org/10.1080/13501763.2022.2099449

Bareis, J. & Katzenbach, C. (2021), "Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics", *Science, Technology & Human Values*, Vol. 47, No. 5, pp. 855-881.

Bommasani, R., et al. (2022), "On the Opportunities and Risks of Foundation Models", available at: https://arxiv.org/abs/2108.07258

BRAID Researchers (2025), "Response to the Call for input for EMRTD study: Artificial Intelligence, Cultural Rights, and the Right to Development", *Zenodo*. https://doi.org/10.5281/zenodo.17975180

Cesur, M. (2024), "Yapay Zeka Çağında Milli Tarih Anlatıları: Potansiyel Tehlikeler ve "Milli Yapay Zeka" Gerekliliği Üzerine Bir İnceleme", *Erciyes Akademi*, Vol. 38, no. 4, pp. 840-862. https://doi.org/10.48070/erciyesakademi.1544216.

Colace, F., et al. (2025), "New AI Challenges for Cultural Heritage Protection", *Journal of Cultural Heritage*, Vol. 71, pp. 112–125.

Chan, C. (2023), "A Comprehensive AI Policy Education Framework for Univer-

sity Teaching and Learning", *International Journal of Educational Technology in Higher Education*, No. 20. https://doi.org/10.1186/s41239-023-00408-3

Challies, E., et al. (2017), "Governance Change and Governance Learning in Europe: Stakeholder Participation in Environmental Policy Implementation", *Policy and Society*, No. 36, pp. 288-303. https://doi.org/10.1080/14494035.2017.1320854

Crawford, K. (2021), *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*, Yale University Press. https://katecrawford.net/atlas

Dignam, A. (2020), "Artificial Intelligence, Tech Corporate Governance and the Public Interest Regulatory Response", *Cambridge Journal of Regions Economy and Society*, No. 13, pp. 37-54. https://doi.org/10.1093/cjres/rsaa002

Dubois, L.E. & Walzak, L. (2025), "Blind Scouting: Using Artificial Intelligence to Alleviate Bias in Selection", *Personnel Review*. https://doi.org/10.1108/pr-02-2024-0130

Farrell, Henry (2025), "AI as Governance", *Annual Review of Political Science*. https://doi.org/10.1146/annurev-polisci-040723-013245

Graziani, M., et al. (2022), "A Global Taxonomy of Interpretable AI: Unifying the Terminology for the Technical and Social Sciences", *Artificial Intelligence Review*, No. 56, pp. 3473-3504. https://doi.org/10.1007/s10462-022-10256-8

Gröger, C. (2021), "There Is No AI Without Data", *Communications of the ACM*, No. 64, pp. 98-108. https://doi.org/10.1145/3448247

Karami, A. & Mottaghi Dastanaei, A. (2024), "Pathology of the Impact of Artificial Intelligence on Public Diplomacy", *Journal of Research on Global Relations*, Vol. 1, No. 3. pp. 81-146. doi: 10.22054/jrgr.2024.78585.1076. [in Persian]

Karampouri, M. & Eslami, R. (2024), "The Governance of Artificial Intellingence from the Perspective of Political Phenomenology", Journal of Philosophical Investigations, Vol. 18, No. 49, pp. 83-110. https://doi.org/10.22034/jpiut.2024.58516.3760. [in Persian]

Lane, J., et al. (2024), "Narrative AI and the Human–AI Oversight Paradox in Evaluating Early-Stage Innovations", *Harvard Business School Working Paper*, No. 25-001, available at: https://www.hbs.edu/ris/Publication%20Files/25-001_8ebbe0cb-2a19-453c-9014-1e301e8dd2fb.pdf

Latifi, H. (2024), "Challenges of Using Artificial Intelligence in the Process of Shi'i Ijtihad", *Religions (Special issue: Theology and Science: Loving Science, Discovering the Divine)*, Vol. 15, No. 5, available at: https://doi.org/10.3390/rel15050541

Mejias, U.A., & Couldry, N. (2024), *Data Grab: The New Colonialism of Big Tech and How to Fight Back*, University of Chicago Press.

Mohammadimonfared, H. & Taghipoor Javi, T. (2025), "The Role of Artificial Intelligence Technology in Transforming the Nature of War; Based on the Study of Strategic Documents of the Zionist Regime", *Journal of Fundamental and Applied Studies of the Islamic World*, Vol. 6, No. 4, pp. 285-307. https://doi.org/10.22034/fasiw.2025.483659.1394. [in Persian]

Moller, A.G., et al. (2025), "The Impact of Generative AI on Social Media: An Experimental Study", available at: https://arxiv.org/pdf/2506.14295

Moon, M. (2023), "Searching for Inclusive Artificial Intelligence for Social Good: Participatory Governance and Policy Recommendations for Making AI More Inclusive and Benign for Society", *Public Administration Review*. https://doi.org/10.1111/puar.13648

Ntoutsi, E., et al. (2020), "Bias in Data-driven Artificial Intelligence Systems; An Introductory Survey", *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery*, No. 10. https://doi.org/10.1002/widm.1356

Ojaghi, H.; Soleymani Roozbahani, F. & Zohoorian Nadali, I. (2024), "The Doctrine of Normativisation and Regulation of Artificial Intelligence", *Quarterly Journal of Government and Law*, Vol. 5, No. 1, pp. 123-150. available at: http://qgl.lri.ir/article_200143.html. [in Persian]

Pan, L., et al. (2024), "Complexity of Social Media in the Era of Generative AI", *National Science Review*, Vol. 12, No. 1, available at: https://doi.org/10.1093/nsr/nwae323

Radu, R. (2021), "Steering the Governance of Artificial Intelligence: National Strategies in Perspective", *Policy and Society*, No. 40, pp. 178-193. https://doi.org/10.1080/14494035.2021.1929728

Rahmani, S.; Zolfaghari, I. & Samavat, M. (2025), "Strategic Ethics Management in the Age of Artificial Intelligence", *Journal of Intelligent Strategic Management (JISM)*, Vol. 3, No. 4, pp. 105-120. doi: bumara.3.2.15564.35879841154. [in Persian]

Rashidi, A. (2025), "Artificial Intelligence and the Challenges of Democracy**",** *Bi-Quarterly Journal of Theoretical Politics Research,* Vol. 19, No. 36, pp. 351-387. available at:

https://political.ihss.ac.ir/ar/Article/48240. [in Persian]

Sadeghian, A. (2025), "Artificial Intelligence in Governance and the Governance of Artificial Intelligence", *Quarterly of New Approaches in Public Administration*, Vol. 2, No. 3, pp. 1-36. https://doi.org/10.71815/jnapa.2025.1200094. [in Persian]

Sourati, J. & Evans, J.A. (2023), "Accelerating Science with Human-aware Artificial Intelligence", *Nature Human Behaviour*, No. 7, pp. 1682-1696. https://doi.org/10.1038/s41562-023-01648-z

Stanger, A., et al. (2024), "Terra Incognita: The Governance of Artificial Intelligence in Global Perspective", *Annual Review of Political Science*. https://doi.org/10.1146/annurev-polisci-041322-042247

Sui, J.; Shen, H. & Zhou, X. (2024), "Impact of Cultural Tightness on Consumers' Preference for Anthropomorphic AI Services", *Psychology and Marketing*. https://doi.org/10.1002/mar.22086

Supreme Council of Cultural Revolution (2012), "Approval of the Cultural En-

gineering Map of the Country". Tehran: Secretariat of the Supreme Council of Cultural Revolution. Available at: https://qavanin.ir/Law/TreeText/?IDS=3468550008609607354. [in Persian]

Taeihagh, A. (2021), "Governance of Artificial Intelligence", *Policy and Society*, No. 40, pp. 137-157. https://doi.org/10.1080/14494035.2021.1928377

Ulnicane, I. (2024), "Governance Fix? Power and Politics in Controversies About Governing Generative AI", *Policy and Society*. https://doi.org/10.1093/polsoc/puae022

UN Expert Mechanism on the Right to Development (2025), "Artificial Intelligence, Cultural Rights and the Right to Development: APC and Derechos Digitales joint submission". https://www.apc.org/en/pubs/un-expert-mechanism-right-development-study-artificial-intelligence-cultural-rights-and-right

Venugopal, J.P., et al. (2024), "A Comprehensive Approach to Bias Mitigation for Sentiment Analysis of Social Media Data", *Applied Sciences*, Vol. 14, No. 23, available at: https://doi.org/10.3390/app142311471

Viberg, O., et al. (2024), "Cultural bias and cultural alignment of large language models", *PNAS Nexus*, Vol. 3, No. 9, available at: https://doi.org/10.1093/pnasnexus/pgae346

Walter, Y. (2024), "Embracing the Future of Artificial Intelligence in the Classroom: The Relevance of AI Literacy, Prompt Engineering, and Critical Thinking in Modern Education", *International Journal of Educational Technology in Higher Education*, No. 21, pp. 1-29. https://doi.org/10.1186/s41239-024-00448-3

Yu, T., et al. (2022), "Artificial Intelligence for Dunhuang Cultural Heritage Protection: The Project and the Dataset", *International Journal of Computer Vision*, No. 130, pp. 2646-2673. https://doi.org/10.1007/s11263-022-01665-x