



Unesco Chair
on E-Learning
and Teaching
Mehralborz University
IRAN



توصیه نامه درباره اخلاق هوش مصنوعی

مصوب ۲۳ نوامبر ۲۰۲۱

ترجمان اندیشه یونسکو

شماره هفتم
مرداد ۱۴۰۵



توصیه‌نامه درباره اخلاق هوش مصنوعی^۱

مصوب ۲۳ نوامبر ۲۰۲۱

1. Recommendation on the Ethics of Artificial Intelligence

فهرست مطالب

۵	دیباچه
۱۱	بخش یکم: دامنه کاربرد
۱۶	بخش دوم: اهداف و مقاصد
۱۹	بخش سوم: ارزش‌ها و اصول
۳۲	بخش چهارم: حوزه‌های اقدام سیاستی
۳۳	حوزه سیاستی ۱: ارزیابی اثرات اخلاقی
۳۵	حوزه سیاستی ۲: حکمرانی و راهبری اخلاقی
۳۹	حوزه سیاستی ۳: سیاست داده
۴۱	حوزه سیاستی ۴: توسعه و همکاری بین‌المللی
۴۲	حوزه سیاستی ۵: محیط زیست و اکوسیستم‌ها
۴۴	حوزه سیاستی ۶: جنسیت
۴۶	حوزه سیاستی ۷: فرهنگ
۴۸	حوزه سیاستی ۸: آموزش و پژوهش
۵۱	حوزه سیاستی ۹: ارتباطات و اطلاعات
۵۱	حوزه سیاستی ۱۰: اقتصاد و کار
۵۳	حوزه سیاستی ۱۱: سلامت و رفاه اجتماعی
۵۷	بخش پنجم: پایش و ارزیابی
۶۰	بخش ششم: بهره‌برداری و اجرای این توصیه‌نامه
۶۱	بخش هفتم: ترویج توصیه‌نامه‌ی حاضر
۶۲	بخش هشتم: مقررات نهایی

دیباچه

کنفرانس عمومی سازمان آموزشی، علمی و فرهنگی ملل متحد (یونسکو)، که در چهل‌ویکمین اجلاس خود از ۹ تا ۲۴ نوامبر ۲۰۲۱ در پاریس تشکیل جلسه داده است، با اذعان به آثار عمیق، پویا، مثبت و منفی هوش مصنوعی (AI) بر جوامع، محیط‌زیست، بوم‌سازگان‌ها^۱ و زندگی انسان، از جمله ذهن انسان؛ زیرا این فناوری از طریق شیوه‌های نوین استفاده، بر نحوه اندیشیدن، تعامل و تصمیم‌گیری انسان تأثیر می‌گذارد و حوزه‌های آموزش، علوم انسانی، علوم اجتماعی، علوم طبیعی، فرهنگ، ارتباطات و اطلاعات را متأثر می‌سازد؛

با یادآوری اینکه بر اساس اساسنامه یونسکو، این سازمان در پی آن است که از طریق گسترش همکاری میان ملت‌ها در زمینه‌های آموزش، علوم، فرهنگ، ارتباطات و اطلاعات، به صلح و امنیت یاری رساند تا احترام جهانی به عدالت، حاکمیت قانون، حقوق بشر و آزادی‌های بنیادینی که برای همه مردم جهان به رسمیت شناخته شده‌اند، تقویت شود؛ با اعتقاد راسخ به اینکه این توصیه‌نامه، به‌عنوان سندی هنجارساز که با رویکردی جهانی و بر پایه حقوق بین‌الملل تدوین شده و بر کرامت انسانی، حقوق بشر، برابری جنسیتی، عدالت اجتماعی و اقتصادی، توسعه، سلامت جسمی و روانی، تنوع، به‌هم‌پیوستگی، فراگیری، و حفاظت از محیط‌زیست و بوم‌سازگان‌ها تمرکز دارد، می‌تواند فناوری‌های هوش مصنوعی را در مسیری مسئولانه هدایت کند؛

با الهام از اهداف و اصول منشور ملل متحد؛

با توجه به اینکه فناوری‌های هوش مصنوعی می‌توانند خدمات بزرگی به بشریت ارائه دهند و همه کشورها قادرند از مزایای آن بهره‌مند شوند، اما در عین حال نگرانی‌های بنیادین اخلاقی نیز ایجاد می‌کنند؛ از جمله سوگیری‌هایی که ممکن است در این فناوری‌ها نهفته باشد یا تشدید شود و در نتیجه به تبعیض، نابرابری، شکاف دیجیتالی، محرومیت، و تهدیدی برای تنوع فرهنگی، اجتماعی و زیستی و نیز شکاف‌های اجتماعی یا اقتصادی بینجامد؛ همچنین ضرورت شفافیت و قابلیت فهم عملکرد الگوریتم‌ها و داده‌هایی که این الگوریتم‌ها بر اساس آن‌ها آموزش دیده‌اند، و نیز تأثیرات بالقوه آن‌ها بر کرامت انسانی، حقوق بشر و آزادی‌های بنیادین، برابری جنسیتی، دموکراسی، فرایندهای اجتماعی، اقتصادی، سیاسی و فرهنگی، رویه‌های علمی و مهندسی، رفاه حیوانات، محیط‌زیست و بوم‌سازگان‌ها؛

همچنین با اذعان به اینکه فناوری‌های هوش مصنوعی می‌توانند شکاف‌ها و نابرابری‌های موجود را، هم در درون کشورها و هم میان کشورها، عمیق‌تر کنند و اینکه باید عدالت، اعتماد و انصاف پاس داشته شود تا هیچ کشور و هیچ فردی، نه از طریق دسترسی عادلانه به فناوری‌های هوش مصنوعی و بهره‌مندی از مزایای آن‌ها،

۱. Ecosystems: بوم‌سازگان (اکوسیستم) به مجموعه موجودات زنده و محیط غیرزنده‌ای گفته می‌شود که در یک منطقه خاص با هم در ارتباط هستند و برهم اثر می‌گذارند.

و نه از حیث برخورداری از حمایت در برابر پیامدهای منفی آن‌ها، از قافله پیشرفت باز نماند؛ در عین حال با به رسمیت شناختن شرایط متفاوت کشورها و احترام به تمایل برخی افراد برای مشارکت نکردن در همه تحولات فناورانه؛

با آگاهی از اینکه همه کشورها با شتابگیری استفاده از فناوری‌های اطلاعات و ارتباطات و فناوری‌های هوش مصنوعی روبه‌رو هستند و در عین حال، نیاز روزافزونی به سواد رسانه‌ای و اطلاعاتی دارند؛ همچنین با توجه به اینکه اقتصاد دیجیتال، در کنار فرصت‌های مهم برای بهره‌مندی مشترک از منافع، به‌ویژه برای کشورهای با درآمد پایین و متوسط (LMICs)^۲، از جمله کشورهای کمتر توسعه‌یافته (LDCs)^۳، کشورهای در حال توسعه محصور در خشکی (LLDCs)^۴ و کشورهای کوچک جزیره‌ای در حال توسعه (SIDS)^۵، چالش‌های مهم اجتماعی، اقتصادی و زیست‌محیطی نیز به همراه دارد؛ چالش‌هایی که مستلزم به رسمیت شناختن، حفاظت و ترویج فرهنگ‌ها، ارزش‌ها و دانش بومی به‌منظور توسعه اقتصادهای دیجیتال پایدار است؛

با اذعان بیشتر به اینکه فناوری‌های هوش مصنوعی این ظرفیت را دارند که برای محیط‌زیست و بوم‌سازگان‌ها سودمند باشند و برای تحقق این مزایا، نباید آسیب‌ها و آثار منفی بالقوه بر محیط‌زیست و بوم‌سازگان‌ها نادیده گرفته شوند، بلکه باید به‌طور مؤثر مورد رسیدگی قرار گیرند؛

با تأکید بر اینکه پرداختن به مخاطرات و نگرانی‌های اخلاقی نباید مانعی بر سر راه نوآوری و توسعه باشد، بلکه باید فرصت‌های تازه‌ای فراهم آورد و پژوهش و نوآوری مبتنی بر اصول اخلاقی را تقویت کند تا فناوری‌های هوش مصنوعی بر پایه حقوق بشر، آزادی‌های بنیادین، ارزش‌ها، اصول، و تأملات اخلاقی و معنوی استوار شوند؛ همچنین با یادآوری اینکه کنفرانس عمومی یونسکو در چهلمین اجلاس خود در نوامبر ۲۰۱۹، با تصویب قطعنامه شماره ۳۷^۶، مدیرکل را مأمور کرد تا «یک سند هنجارساز بین‌المللی درباره اخلاق هوش مصنوعی، در قالب یک توصیه‌نامه» تهیه کند و آن را برای تصویب در چهل و یکمین اجلاس کنفرانس عمومی در سال ۲۰۲۱ ارائه دهد؛

با اذعان به اینکه توسعه فناوری‌های هوش مصنوعی مستلزم افزایش متناسب در سواد داده، سواد رسانه‌ای و سواد اطلاعاتی، و نیز دسترسی به منابع مستقل، کثرت‌گرا و قابل اعتماد اطلاعات است؛ از جمله در چارچوب تلاش‌ها برای کاهش مخاطرات ناشی از اطلاعات نادرست، اطلاعات گمراه‌کننده، گفتار نفرت‌پراکن و آسیب‌های ناشی از سوءاستفاده از داده‌های شخصی؛

با توجه به اینکه چارچوب هنجاری فناوری‌های هوش مصنوعی و پیامدهای اجتماعی آن، بر پایه چارچوب‌های حقوقی بین‌المللی و ملی، حقوق بشر و آزادی‌های بنیادین، اصول اخلاقی، ضرورت دسترسی به داده، اطلاعات و دانش، آزادی پژوهش و نوآوری، و

2. Low- and middle-income countries

3. Least developed countries

4. Landlocked developing countries

5. Small island developing States

6. 40 C/ Resolution 37

نیز رفاه انسان، محیط‌زیست و بوم‌سازگان‌ها استوار است و ارزش‌ها و اصول اخلاقی را با چالش‌ها و فرصت‌های مرتبط با فناوری‌های هوش مصنوعی، بر اساس درک مشترک و اهداف مشترک، پیوند می‌دهد؛

همچنین با اذعان به اینکه ارزش‌ها و اصول اخلاقی می‌توانند با ارائه رهنمودهایی متناسب با شتاب بالای توسعه فناوری، به تدوین و اجرای سیاست‌های مبتنی بر حقوق و مقررات قانونی یاری رسانند؛

همچنین با اعتقاد راسخ به اینکه استانداردهای اخلاقی مورد پذیرش جهانی برای فناوری‌های هوش مصنوعی، با رعایت کامل حقوق بین‌الملل، به‌ویژه حقوق بین‌الملل حقوق بشر، می‌توانند نقشی اساسی در شکل‌گیری هنجارهای مرتبط با هوش مصنوعی در سراسر جهان ایفا کنند؛

با در نظر داشتن اعلامیه جهانی حقوق بشر^۷ (۱۹۴۸)، اسناد چارچوب بین‌المللی حقوق بشر، از جمله کنوانسیون مربوط به وضعیت پناهندگان^۸ (۱۹۵۱)، کنوانسیون تبعیض (اشتغال و حرفه)^۹ (۱۹۵۸)، کنوانسیون بین‌المللی رفع همه اشکال تبعیض نژادی^{۱۰} (۱۹۶۵)، میثاق بین‌المللی حقوق مدنی و سیاسی^{۱۱} (۱۹۶۶)، میثاق بین‌المللی حقوق اقتصادی، اجتماعی و فرهنگی^{۱۲} (۱۹۶۶)، کنوانسیون رفع همه اشکال تبعیض علیه زنان^{۱۳} (۱۹۷۹)، کنوانسیون حقوق کودک^{۱۴} (۱۹۸۹)، کنوانسیون حقوق افراد دارای معلولیت^{۱۵} (۲۰۰۶)، کنوانسیون مبارزه با تبعیض در آموزش^{۱۶} (۱۹۶۰)، کنوانسیون حمایت و ترویج تنوع بیان‌های فرهنگی^{۱۷} (۲۰۰۵)، و نیز سایر اسناد، توصیه‌نامه‌ها و اعلامیه‌های بین‌المللی مرتبط؛

همچنین با توجه به اعلامیه سازمان ملل متحد درباره حق بر توسعه^{۱۸} (۱۹۸۶)، اعلامیه مسئولیت‌های نسل حاضر در قبال نسل‌های آینده^{۱۹} (۱۹۹۷)، اعلامیه جهانی اخلاق زیستی و حقوق بشر^{۲۰} (۲۰۰۵)، اعلامیه سازمان ملل متحد درباره حقوق مردمان بومی^{۲۱} (۲۰۰۷)، قطعنامه مجمع عمومی سازمان ملل متحد درباره بازنگری در اجلاس جهانی جامعه اطلاعاتی^{۲۲} (۱۲۵/۷۰/۷۰/A/RES) (۲۰۱۵)، قطعنامه مجمع عمومی سازمان ملل متحد با عنوان «دگرگون ساختن جهان ما: دستورکار ۲۰۳۰ برای توسعه پایدار»^{۲۳} (۱/۷۰/۷۰/A/RES)

7. Universal Declaration of Human Rights

8. Convention Relating to the Status of Refugees

9. Discrimination (Employment and Occupation) Convention

10. International Convention on the Elimination of All Forms of Racial Discrimination

11. International Covenant on Civil and Political Rights

12. International Covenant on Economic, Social and Cultural Rights

13. Convention on the Elimination of All Forms of Discrimination against Women

14. Convention on the Rights of the Child

15. Convention on the Rights of Persons with Disabilities

16. Convention against Discrimination in Education

17. Convention on the Protection and Promotion of the Diversity of Cultural Expressions

18. United Nations Declaration on the Right to Development

19. Declaration on the Responsibilities of the Present Generations Towards Future Generations

20. Universal Declaration on Bioethics and Human Rights

21. United Nations Declaration on the Rights of Indigenous Peoples

22. United Nations General Assembly resolution on the review of the World Summit on the Information Society

23. United Nations General Assembly Resolution on Transforming our world: the 2030 Agenda for Sustainable Development

(۲۰۱۵)، توصیه‌نامه درباره حفاظت و دسترسی به میراث مستند، از جمله در قالب دیجیتال^{۲۴} (۲۰۱۵)، اعلامیه اصول اخلاقی مرتبط با تغییرات اقلیمی^{۲۵} (۲۰۱۷)، توصیه‌نامه درباره علم و پژوهشگران علمی^{۲۶} (۲۰۱۷)، شاخص‌های جهان‌شمولی اینترنت^{۲۷} (که در سال ۲۰۱۸ توسط برنامه بین‌المللی توسعه ارتباطات یونسکو تأیید شد)، از جمله اصول ROAM (که در سال ۲۰۱۵ توسط کنفرانس عمومی یونسکو تأیید شد)، قطعنامه شورای حقوق بشر درباره «حق حریم خصوصی در عصر دیجیتال»^{۲۸} (A/HRC/RES/15/42/2009)، و نیز قطعنامه شورای حقوق بشر درباره «فناوری‌های نوظهور دیجیتال و حقوق بشر»^{۲۹} (A/HRC/RES/11/41/2009)؛

با تأکید بر اینکه باید توجه ویژه‌ای به کشورهای با درآمد پایین و متوسط (LMICs)، از جمله کشورهای کمتر توسعه‌یافته (LDCs)، کشورهای در حال توسعه محصور در خشکی (LLDCs) و کشورهای کوچک جزیره‌ای در حال توسعه (SIDS) معطوف شود؛ زیرا این کشورها، با وجود برخورداری از ظرفیت‌های خاص خود، در مباحث مربوط به اخلاق هوش مصنوعی کمتر نمایندگی شده‌اند و این امر نگرانی‌هایی را درباره نادیده گرفتن دانش بومی، تکرر فرهنگی، نظام‌های ارزشی و الزامات عدالت جهانی در مواجهه با آثار مثبت و منفی فناوری‌های هوش مصنوعی ایجاد می‌کند؛

همچنین با آگاهی از سیاست‌های ملی متعدد، سایر چارچوب‌ها و ابتکارهایی که از سوی نهادهای ذریع سازمان ملل متحد، سازمان‌های بین‌دولتی - از جمله سازمان‌های منطقه‌ای - و نیز بخش خصوصی، سازمان‌های حرفه‌ای، سازمان‌های غیردولتی و جامعه علمی در زمینه اخلاق و تنظیم‌گری فناوری‌های هوش مصنوعی تدوین شده‌اند؛

همچنین با اعتقاد راسخ به اینکه فناوری‌های هوش مصنوعی می‌توانند منافع مهمی به همراه داشته باشند، اما دستیابی به این منافع ممکن است هم‌زمان تنش‌هایی را پیرامون نوآوری، دسترسی نامتقارن به دانش و فناوری، از جمله کمبود سواد دیجیتال و سواد مدنی که توانایی عموم مردم را برای مشارکت در موضوعات مرتبط با هوش مصنوعی محدود می‌کند، و نیز موانع دسترسی به اطلاعات، کمبود ظرفیت‌های انسانی و نهادی، موانع دسترسی به نوآوری فناورانه، و نبود زیرساخت‌های فیزیکی و دیجیتالی و چارچوب‌های مقرراتی مناسب - از جمله در حوزه داده - تشدید کند؛ مسائلی که همگی باید مورد رسیدگی قرار گیرند؛

با تأکید بر اینکه تقویت همکاری و همبستگی جهانی، از جمله از طریق چندجانبه‌گرایی، برای تسهیل دسترسی عادلانه به فناوری‌های هوش مصنوعی و مقابله با چالش‌هایی که این فناوری‌ها برای تنوع و درهم‌تنیدگی فرهنگ‌ها و نظام‌های اخلاقی ایجاد می‌کنند، ضروری است؛ همچنین برای کاهش سوءاستفاده‌های احتمالی، بهره‌گیری

24. Recommendation Concerning the Preservation of, and Access to, Documentary Heritage Including in Digital Form

25. Declaration of Ethical Principles in relation to Climate Change

26. Recommendation on Science and Scientific Researchers

27. Internet Universality Indicators

28. The right to privacy in the digital age

29. Human Rights Council's resolution on "New and emerging digital technologies and human rights"

کامل از ظرفیت‌های هوش مصنوعی، به‌ویژه در حوزه توسعه، و تضمین اینکه راهبردهای ملی هوش مصنوعی بر پایه اصول اخلاقی هدایت شوند؛ با در نظر گرفتن کامل این واقعیت که توسعه شتابان فناوری‌های هوش مصنوعی، اجرای اخلاقی و حکمرانی آن‌ها، همچنین احترام به تنوع فرهنگی و حفاظت از آن را با چالش مواجه می‌کند و این ظرفیت را دارد که معیارها و ارزش‌های اخلاقی محلی و منطقه‌ای را دستخوش دگرگونی سازد؛

در این بیست‌وسومین روز از ماه نوامبر سال ۲۰۲۱، توصیه‌نامه حاضر درباره اخلاق هوش مصنوعی را تصویب می‌کند؛

۱. به کشورهای عضو توصیه می‌کند که مفاد این توصیه‌نامه را به‌صورت داوطلبانه اجرا کنند و برای تحقق اصول و هنجارهای مندرج در آن در قلمرو صلاحیت خود، مطابق با حقوق بین‌الملل، از جمله قوانین بین‌الملل حقوق بشر، اقدامات مقتضی را، از جمله هرگونه تدابیر قانونی یا سایر اقدامات لازم، متناسب با رویه‌های قانون اساسی و ساختارهای حکمرانی هر کشور، اتخاذ کنند؛

۲. همچنین به کشورهای عضو توصیه می‌کند که همه ذی‌نفعان، از جمله بنگاه‌های اقتصادی، را در اجرای این توصیه‌نامه مشارکت دهند تا هر یک نقش متناسب خود را ایفا کنند؛ و این توصیه‌نامه را به اطلاع مقامات، نهادها، سازمان‌های پژوهشی و دانشگاهی، و مؤسسات و سازمان‌های فعال در بخش‌های دولتی، خصوصی و جامعه مدنی که در حوزه فناوری‌های هوش مصنوعی فعالیت می‌کنند، برسانند تا توسعه و استفاده از فناوری‌های هوش مصنوعی، هم بر پایه پژوهش علمی معتبر و هم بر اساس تحلیل و ارزیابی اخلاقی، هدایت شود.



بخش یکم: دامنه کاربرد^۱

۱. این توصیه‌نامه به مسائل اخلاقی مرتبط با حوزه هوش مصنوعی، تا آنجا که در چارچوب مأموریت یونسکو قرار می‌گیرند، می‌پردازد. این سند، اخلاق هوش مصنوعی را به منزله تأملی هنجاری، نظام‌مند و مبتنی بر چارچوبی کل‌نگر، جامع، چندفرهنگی، پویا و متشکل از ارزش‌ها، اصول و اقدامات به‌هم‌پیوسته در نظر می‌گیرد؛ چارچوبی که می‌تواند جوامع را در مواجهه مسئولانه با آثار شناخته‌شده و ناشناخته فناوری‌های هوش مصنوعی بر انسان، جامعه، محیط‌زیست و بوم‌سازگان‌ها هدایت کند و مبنایی برای پذیرش یا رد این فناوری‌ها در اختیار آنان قرار دهد. این توصیه‌نامه، اخلاق را مبنایی پویا برای ارزیابی هنجاری و هدایت فناوری‌های هوش مصنوعی می‌داند که بر کرامت انسانی، رفاه انسان و پیشگیری از آسیب استوار بوده و ریشه در اخلاق علم و فناوری دارد.

۲. این توصیه‌نامه در پی ارائه یک تعریف واحد از هوش مصنوعی نیست؛ زیرا چنین تعریفی باید همگام با پیشرفت‌های فناوری، در طول زمان دستخوش تغییر شود. در عوض، هدف آن پرداختن به ویژگی‌های سامانه‌های هوش مصنوعی است که از منظر اخلاقی اهمیت بنیادین دارند. از این‌رو، این توصیه‌نامه سامانه‌های هوش مصنوعی را سامانه‌هایی می‌داند که توانایی پردازش داده‌ها و اطلاعات را به گونه‌ای دارند که رفتار هوشمندانه را بازنمایی می‌کند و معمولاً جنبه‌هایی مانند استدلال، یادگیری، ادراک، پیش‌بینی، برنامه‌ریزی یا کنترل را در بر می‌گیرد. در این رویکرد، سه عنصر جایگاهی اساسی دارند: **(الف)** سامانه‌های هوش مصنوعی، فناوری‌های پردازش اطلاعات هستند که با تلفیق مدل‌ها و الگوریتم‌ها، قابلیت یادگیری و انجام وظایف شناختی را ایجاد می‌کنند؛ وظایفی که به نتایجی همچون پیش‌بینی و تصمیم‌گیری در محیط‌های مادی و مجازی منجر می‌شوند. این سامانه‌ها به گونه‌ای طراحی می‌شوند که با درجات متفاوتی از خودمختاری، از طریق مدل‌سازی و بازنمایی دانش، بهره‌برداری از داده‌ها و محاسبه همبستگی‌ها عمل کنند. سامانه‌های هوش مصنوعی ممکن است از روش‌های گوناگونی استفاده کنند که از جمله، (اما نه محدود به) موارد زیر است:

۱. یادگیری ماشین، از جمله یادگیری عمیق و یادگیری تقویتی؛
۲. استدلال ماشینی، از جمله برنامه‌ریزی، زمان‌بندی، بازنمایی و استدلال دانش، جست‌وجو و بهینه‌سازی.

سامانه‌های هوش مصنوعی می‌توانند در سامانه‌های سایبری - فیزیکی، از جمله اینترنت اشیا، سامانه‌های رباتیک، رباتیک اجتماعی و رابط‌های انسان - رایانه به‌کار گرفته شوند؛ سامانه‌هایی که کنترل، ادراک، پردازش داده‌های گردآوری‌شده توسط حسگرها و عملکرد عملگرها را در محیطی که سامانه‌های هوش مصنوعی در آن فعالیت می‌کنند، در بر می‌گیرند.

(ب) پرسش‌های اخلاقی مربوط به سامانه‌های هوش مصنوعی، همه مراحل چرخه حیات این سامانه‌ها را در بر می‌گیرد. منظور از چرخه حیات، تمامی مراحل از پژوهش، طراحی و توسعه تا استقرار و بهره‌برداری، شامل نگهداری، بهره‌برداری، تجارت، تأمین مالی، پایش

و ارزیابی، اعتبارسنجی، پایان استفاده، برچیدن و خاتمه فعالیت سامانه است. افزون بر این، «کنشگران هوش مصنوعی» به هر شخص حقیقی یا حقوقی اطلاق می‌شود که در دست‌کم یکی از مراحل چرخه حیات سامانه‌های هوش مصنوعی نقش داشته باشد؛ از جمله پژوهشگران، برنامه‌نویسان، مهندسان، دانشمندان داده، کاربران نهایی، بنگاه‌های اقتصادی، دانشگاه‌ها و نهادهای عمومی و خصوصی.

(پ) سامانه‌های هوش مصنوعی مسائل اخلاقی جدیدی را مطرح می‌کنند که از جمله، (اما نه محدود به) تأثیر آن‌ها بر تصمیم‌گیری، اشتغال و کار، تعاملات اجتماعی، مراقبت‌های سلامت، آموزش، رسانه، دسترسی به اطلاعات، شکاف دیجیتال، حفاظت از داده‌های شخصی و حقوق مصرف‌کننده، محیط‌زیست، دموکراسی، حاکمیت قانون، امنیت و پلیس، کاربردهای دوگانه و نیز حقوق بشر و آزادی‌های بنیادین، از جمله آزادی بیان، حریم خصوصی و منع تبعیض است.

علاوه بر این، توانایی الگوریتم‌های هوش مصنوعی در بازتولید و تقویت سوگیری‌های موجود می‌تواند اشکال موجود تبعیض، پیش‌داوری و کلیشه‌سازی را تشدید کند. برخی از این مسائل به توانایی سامانه‌های هوش مصنوعی در انجام وظایفی مربوط می‌شود که پیش‌تر تنها موجودات زنده و در برخی موارد صرفاً انسان‌ها قادر به انجام آن بودند. این ویژگی‌ها موجب شده است که سامانه‌های هوش مصنوعی نقشی عمیق و نوین در فعالیت‌های انسانی، جامعه و نیز در رابطه انسان با محیط‌زیست و بوم‌سازگان‌ها ایفا کنند و زمینه‌ای تازه برای رشد کودکان و نوجوانان، شکل‌گیری درک آنان از جهان و از خود، فهم انتقادی رسانه‌ها و اطلاعات، و یادگیری تصمیم‌گیری فراهم آورند. در بلندمدت، سامانه‌های هوش مصنوعی ممکن است برداشت ویژه انسان از تجربه و عاملیت خود را به چالش بکشند و نگرانی‌های تازه‌ای درباره خودشناسی انسان، تعاملات اجتماعی، فرهنگی و زیست‌محیطی، خودمختاری، عاملیت، ارزش و کرامت انسانی پدید آورند.

۳. این توصیه‌نامه توجه ویژه‌ای به پیامدهای گسترده اخلاقی سامانه‌های هوش مصنوعی در حوزه‌های اصلی مأموریت یونسکو، یعنی آموزش، علم، فرهنگ، ارتباطات و اطلاعات دارد؛ همان‌گونه که در «مطالعه مقدماتی درباره اخلاق هوش مصنوعی» که در سال ۲۰۱۹ توسط کمیسیون جهانی اخلاق دانش و فناوری علمی یونسکو^۱ (COMEST) تهیه شد، تبیین شده است:

(الف) آموزش: زیرا زندگی در جوامعی که به‌سرعت در حال دیجیتالی شدن هستند، مستلزم شیوه‌های نوین آموزشی، تأمل اخلاقی، تفکر انتقادی، طراحی مسئولانه و مهارت‌های جدید است؛ به‌ویژه با توجه به پیامدهای هوش مصنوعی برای بازار کار، اشتغال‌پذیری و مشارکت مدنی.

(ب) علم: در گسترده‌ترین مفهوم آن، شامل همه حوزه‌های دانشگاهی از علوم طبیعی و علوم پزشکی تا علوم اجتماعی و علوم انسانی؛ زیرا فناوری‌های هوش مصنوعی ظرفیت‌ها و رویکردهای پژوهشی جدیدی ایجاد می‌کنند، بر برداشت ما از فهم و تبیین

1. UNESCO World Commission on Ethics of Scientific Knowledge and Technology (COMEST)

علمی اثر می‌گذارند و مبنای تازه‌ای برای تصمیم‌گیری فراهم می‌آورند.

(پ) هویت و تنوع فرهنگی: زیرا فناوری‌های هوش مصنوعی می‌توانند صنایع فرهنگی و خلاق را غنا بخشند، اما در عین حال ممکن است عرضه محتوای فرهنگی، داده‌ها، بازارها و درآمدها را در اختیار شمار اندکی از بازیگران متمرکز سازند و از این رهگذر پیامدهای منفی برای تنوع و تکثر زبان‌ها، رسانه‌ها، جلوه‌های فرهنگی، مشارکت و برابری به همراه داشته باشند.

(ت) ارتباطات و اطلاعات: زیرا فناوری‌های هوش مصنوعی نقش روزافزونی در پردازش، سازمان‌دهی و ارائه اطلاعات ایفا می‌کنند. روزنامه‌نگاری خودکار، تولید و عرضه الگوریتمی اخبار، و نیز پالایش و مدیریت محتوا در شبکه‌های اجتماعی و موتورهای جست‌وجو تنها نمونه‌هایی از موضوعاتی هستند که مسائلی همچون دسترسی به اطلاعات، اطلاعات نادرست، اطلاعات گمراه‌کننده، نفرت‌پراکنی، شکل‌گیری روایت‌های جدید اجتماعی، تبعیض، آزادی بیان، حریم خصوصی و سواد رسانه‌ای و اطلاعاتی را مطرح می‌کنند.



۴. این توصیه‌نامه خطاب به کشورهای عضو است؛ هم به‌عنوان کنشگران حوزه هوش مصنوعی و هم به‌عنوان مراجع مسئول تدوین چارچوب‌های قانونی و مقرراتی در سراسر چرخه حیات سامانه‌های هوش مصنوعی و ارتقای مسئولیت‌پذیری بنگاه‌های اقتصادی این سند همچنین با فراهم کردن مبنایی برای ارزیابی اخلاقی آثار سامانه‌های هوش مصنوعی در تمامی مراحل چرخه حیات آن‌ها، رهنمودهای اخلاقی لازم را برای همه کنشگران هوش مصنوعی، از جمله بخش‌های دولتی و خصوصی، ارائه می‌کند.



بخش دوم: اهداف و مقاصد^۱

۵. هدف این توصیه‌نامه آن است که مبنایی فراهم آورد تا سامانه‌های هوش مصنوعی در خدمت خیر و صلاح بشریت، افراد، جوامع، محیط‌زیست و بوم‌سازگان‌ها قرار گیرند و از بروز آسیب جلوگیری شود. همچنین این توصیه‌نامه در پی ترویج استفاده صلح‌آمیز از سامانه‌های هوش مصنوعی است.

۶. افزون بر چارچوب‌های اخلاقی موجود درباره هوش مصنوعی در سراسر جهان، این توصیه‌نامه در پی آن است که یک سند هنجاری مورد پذیرش جهانی ارائه کند که نه تنها بر تبیین ارزش‌ها و اصول، بلکه بر تحقق عملی آن‌ها از طریق توصیه‌های مشخص سیاستی تمرکز دارد و در این میان، بر فراگیری، برابری جنسیتی و حفاظت از محیط‌زیست و بوم‌سازگان‌ها تأکید ویژه می‌کند.

۷. از آنجا که پیچیدگی مسائل اخلاقی پیرامون هوش مصنوعی مستلزم همکاری ذی‌نفعان متعدد در سطوح و بخش‌های مختلف جوامع بین‌المللی، منطقه‌ای و ملی است، این توصیه‌نامه می‌کوشد زمینه پذیرش مسئولیت مشترک را بر پایه گفت‌وگویی جهانی و میان‌فرهنگی میان همه ذی‌نفعان فراهم آورد.

۸. اهداف این توصیه‌نامه عبارت‌اند از:

(الف) فراهم آوردن چارچوبی جهانی از ارزش‌ها، اصول و اقدامات برای هدایت دولت‌ها در تدوین قوانین، سیاست‌ها و سایر ابزارهای مرتبط با هوش مصنوعی، در انطباق با حقوق بین‌الملل؛

(ب) هدایت اقدامات افراد، گروه‌ها، جوامع، نهادها و شرکت‌های بخش خصوصی به منظور تضمین نهادینه شدن اصول اخلاقی در تمامی مراحل چرخه حیات سامانه‌های هوش مصنوعی؛

(پ) حمایت، ترویج و احترام به حقوق بشر و آزادی‌های بنیادین، کرامت انسانی و برابری، از جمله برابری جنسیتی؛ پاسداری از منافع نسل‌های حاضر و آینده؛ حفاظت از محیط‌زیست، تنوع زیستی و بوم‌سازگان‌ها؛ و احترام به تنوع فرهنگی در تمامی مراحل چرخه حیات سامانه‌های هوش مصنوعی؛

(ت) تقویت گفت‌وگوی چندذی‌نفعی، میان‌رشته‌ای و کثرت‌گرایانه و دستیابی به اجماع درباره مسائل اخلاقی مرتبط با سامانه‌های هوش مصنوعی؛

(ث) ترویج دسترسی عادلانه به پیشرفت‌ها و دانش در حوزه هوش مصنوعی و تسهیم منافع آن، با توجه ویژه به نیازها و مشارکت کشورهای با درآمد پایین و متوسط (LMICs)، از جمله کشورهای کمتر توسعه‌یافته (LDCs)، کشورهای در حال توسعه محصور در خشکی (LLDCs) و کشورهای کوچک جزیره‌ای در حال توسعه (SIDS).



بخش سوم: ارزش‌ها و اصول^۱

۹. ارزش‌ها و اصولی که در ادامه آمده‌اند باید در تمامی مراحل چرخه حیات سامانه‌های هوش مصنوعی رعایت شوند و در صورت لزوم و اقتضا، از طریق اصلاح قوانین موجود و تدوین قوانین، مقررات و دستورالعمل‌های کسب‌وکار جدید، ترویج و تقویت شوند. این اقدامات باید با حقوق بین‌الملل، از جمله منشور ملل متحد و تعهدات کشورهای عضو در زمینه حقوق بشر، مطابقت داشته باشند و همچنین با اهداف مورد توافق بین‌المللی در حوزه‌های اجتماعی، سیاسی، زیست‌محیطی، آموزشی، علمی و اقتصادی، از جمله اهداف توسعه پایدار سازمان ملل متحد^۱ (SDGs)، همسو باشند.

۱۰. ارزش‌ها نقش مهمی به‌عنوان آرمان‌های انگیزه‌بخش در شکل‌دهی به سیاست‌ها و اقدامات حقوقی ایفا می‌کنند. در حالی که مجموعه ارزش‌های بیان‌شده در ادامه، رفتار مطلوب را ترسیم کرده و بنیان اصول را تشکیل می‌دهد، این اصول ارزش‌های زیربنایی را به‌صورت ملموس‌تر تبیین می‌کنند تا امکان عملیاتی‌سازی آسان‌تر این ارزش‌ها در سیاست‌ها و اقدامات فراهم شود.

۱۱. هرچند تمامی ارزش‌ها و اصولی که در ادامه بیان می‌شوند، در هر زمینه عملی مطلوب هستند، ممکن است در برخی موقعیت‌ها میان این ارزش‌ها و اصول تنش‌هایی ایجاد شود. در چنین شرایطی، برای مدیریت این تنش‌های احتمالی، انجام ارزیابی متناسب با بافت و شرایط ضروری است؛ ارزیابی‌ای که اصل تناسب را رعایت کند و با حقوق بشر و آزادی‌های بنیادین سازگار باشد. در همه موارد، هرگونه محدودیت احتمالی بر حقوق بشر و آزادی‌های بنیادین باید دارای مبنای قانونی، معقول، ضروری و متناسب باشد و با تعهدات دولت‌ها بر اساس حقوق بین‌الملل سازگار باشد. مدیریت سنجیده چنین موقعیت‌هایی معمولاً مستلزم مشارکت طیف گسترده‌ای از ذی‌نفعان مناسب، بهره‌گیری از گفت‌وگوی اجتماعی و نیز بررسی‌های اخلاقی، ارزیابی دقیق و سنجش آثار است.

۱۲. قابلیت اعتماد و تمامیت چرخه حیات سامانه‌های هوش مصنوعی؛ برای تضمین آن ضروری است که فناوری‌های هوش مصنوعی در خدمت خیر و صلاح بشریت، افراد، جوامع، محیط‌زیست و بوم‌سازگان‌ها قرار گیرند و ارزش‌ها و اصول مندرج در این توصیه‌نامه را متجلی سازند. مردم باید دلایل موجهی برای اعتماد به سامانه‌های هوش مصنوعی داشته باشند؛ اینکه این سامانه‌ها می‌توانند منافع فردی و مشترک ایجاد کنند و در عین حال، تدابیر کافی برای کاهش مخاطرات آن‌ها اتخاذ شده است. یکی از الزامات اساسی اعتمادپذیری آن است که سامانه‌های هوش مصنوعی در سراسر چرخه حیات خود، حسب مورد، تحت پایش مستمر ذی‌نفعان مرتبط قرار گیرند. از آنجا که اعتمادپذیری حاصل عملیاتی شدن اصول این سند است، تمامی اقدامات سیاستی پیشنهادشده در این توصیه‌نامه با هدف ارتقای اعتمادپذیری در همه مراحل چرخه حیات سامانه‌های هوش مصنوعی تدوین شده‌اند.

1. United Nations Sustainable Development Goals (SDGs)

احترام، حمایت و ترویج حقوق بشر، آزادی‌های بنیادین و کرامت انسانی^۲

۱۳. کرامت ذاتی و خدشه‌ناپذیر هر انسان، بنیان نظام جهانی، تجزیه‌ناپذیر، سلب‌ناشدنی، به‌هم‌پیوسته و وابسته به هم حقوق بشر و آزادی‌های بنیادین را تشکیل می‌دهد. از این‌رو، احترام، حمایت و ترویج کرامت انسانی و حقوق بشر، همان‌گونه که در حقوق بین‌الملل، از جمله حقوق بین‌الملل حقوق بشر، مقرر شده است، باید در سراسر چرخه حیات سامانه‌های هوش مصنوعی تضمین شود. کرامت انسانی به شناسایی ارزش ذاتی و برابر هر انسان، صرف‌نظر از نژاد، رنگ پوست، تبار، جنسیت، سن، زبان، دین، عقیده سیاسی، منشأ ملی، منشأ قومی، منشأ اجتماعی، وضعیت اقتصادی یا اجتماعی هنگام تولد، معلولیت یا هر وضعیت دیگر مربوط می‌شود.

۱۴. هیچ انسان یا هیچ جامعه انسانی نباید در هیچ مرحله‌ای از چرخه حیات سامانه‌های هوش مصنوعی، از نظر جسمی، اقتصادی، اجتماعی، سیاسی، فرهنگی یا روانی آسیب ببیند یا به حاشیه رانده شود. در سراسر چرخه حیات سامانه‌های هوش مصنوعی، کیفیت زندگی انسان‌ها باید ارتقا یابد و تعریف «کیفیت زندگی» تا زمانی که ناقض یا سوءاستفاده‌کننده از حقوق بشر، آزادی‌های بنیادین یا کرامت انسان‌ها نباشد، باید به انتخاب افراد یا گروه‌ها واگذار شود.

۱۵. افراد ممکن است در سراسر چرخه حیات سامانه‌های هوش مصنوعی با این سامانه‌ها تعامل داشته باشند و از آن‌ها کمک دریافت کنند؛ از جمله در مراقبت از افراد آسیب‌پذیر یا افرادی که در موقعیت‌های آسیب‌پذیر قرار دارند، از جمله - اما نه محدود به - کودکان، سالمندان، افراد دارای معلولیت یا بیماران. در چنین تعاملاتی، هرگز نباید با افراد همچون شیء برخورد شود، کرامت آنان تضعیف گردد یا حقوق بشر و آزادی‌های بنیادینشان نقض یا مورد سوءاستفاده قرار گیرد.

۱۶. حقوق بشر و آزادی‌های بنیادین باید در سراسر چرخه حیات سامانه‌های هوش مصنوعی محترم شمرده شوند، مورد حمایت قرار گیرند و ترویج یابند. دولت‌ها، بخش خصوصی، جامعه مدنی، سازمان‌های بین‌المللی، جوامع فنی و دانشگاهی باید در تمامی مداخلات خود در فرایندهای مرتبط با چرخه حیات سامانه‌های هوش مصنوعی، به اسناد و چارچوب‌های حقوق بشر احترام بگذارند. فناوری‌های نوین باید ابزارهای تازه‌ای برای حمایت، دفاع و اعمال حقوق بشر فراهم آورند، نه آنکه موجب نقض یا تضعیف آن شوند.

1. VALUES

2. Respect, protection and promotion of human rights and fundamental freedoms and human dignity

شکوفایی محیط‌زیست و بوم‌سازگان‌ها^۳

۱۷. شکوفایی محیط‌زیست و بوم‌سازگان‌ها باید در سراسر چرخه حیات سامانه‌های هوش مصنوعی به رسمیت شناخته شود، مورد حفاظت قرار گیرد و ترویج یابد. افزون بر این، محیط‌زیست و بوم‌سازگان‌ها برای بشریت و سایر موجودات زنده ضرورتی اساسی دارند تا بتوانند از منافع پیشرفت‌های هوش مصنوعی بهره‌مند شوند.

۱۸. همه کنشگران دخیل در چرخه حیات سامانه‌های هوش مصنوعی باید از حقوق بین‌الملل و قوانین داخلی، استانداردها و رویه‌های مرتبط – از جمله تدابیر احتیاطی برای حفاظت و احیای محیط‌زیست و بوم‌سازگان‌ها و نیز توسعه پایدار – تبعیت کنند. آنان باید آثار زیست‌محیطی سامانه‌های هوش مصنوعی، از جمله، اما نه محدود به، ردپای کربنی آن‌ها را کاهش دهند؛ به کاهش تغییرات اقلیمی و مخاطرات زیست‌محیطی کمک کنند؛ و از بهره‌برداری، استفاده و دگرگونی ناپایدار منابع طبیعی که به تخریب محیط‌زیست و زوال بوم‌سازگان‌ها می‌انجامد، جلوگیری کنند.

تضمین تنوع و فراگیری^۴

۱۹. احترام، حمایت و ترویج تنوع و فراگیری باید در سراسر چرخه حیات سامانه‌های هوش مصنوعی، در چارچوب حقوق بین‌الملل، از جمله حقوق بشر، تضمین شود. این هدف می‌تواند از طریق مشارکت فعال همه افراد یا گروه‌ها، بدون توجه به نژاد، رنگ پوست، تبار، جنسیت، سن، زبان، دین، عقیده سیاسی، منشأ ملی، منشأ قومی، منشأ اجتماعی، وضعیت اقتصادی یا اجتماعی هنگام تولد، معلولیت یا هر وضعیت دیگر، تحقق یابد.

۲۰. دامنه انتخاب‌های مربوط به سبک زندگی، باورها، دیدگاه‌ها، شیوه‌های بیان یا تجربه‌های شخصی افراد، از جمله استفاده اختیاری از سامانه‌های هوش مصنوعی و مشارکت در طراحی این معماری‌ها، نباید در هیچ مرحله‌ای از چرخه حیات سامانه‌های هوش مصنوعی محدود شود.

۲۱. علاوه بر این، باید تلاش‌هایی، از جمله در قالب همکاری‌های بین‌المللی، برای رفع – و هرگز بهره‌برداری از – کمبود زیرساخت‌های فناورانه و دانش، آموزش و مهارت‌ها، و نیز چارچوب‌های حقوقی، به‌ویژه در کشورهای با درآمد پایین و متوسط، کشورهای کمتر توسعه‌یافته، کشورهای در حال توسعه محصور در خشکی و کشورهای کوچک جزیره‌ای در حال توسعه، که بر جوامع تأثیر می‌گذارند، صورت گیرد.

زیستن در جوامع صلح‌آمیز، عادلانه و به‌هم‌پیوسته^۵

3. Environment and ecosystem flourishing

4. Ensuring diversity and inclusiveness

5. Living in peaceful, just and interconnected societies

۲۲. همه کنشگران باید نقشی مشارکت‌جویانه و توانمندساز در تضمین شکل‌گیری جوامعی صلح‌آمیز و عادلانه ایفا کنند؛ جوامعی که بر آینده‌ای به‌هم‌پیوسته و در راستای خیر و صلاح همگان، و در هماهنگی با حقوق بشر و آزادی‌های بنیادین، استوار باشند. ارزش زیستن در جوامع صلح‌آمیز و عادلانه، بر ظرفیت سامانه‌های هوش مصنوعی برای مشارکت در سراسر چرخه حیات خود در تقویت پیوند میان همه موجودات زنده با یکدیگر و با محیط‌زیست طبیعی تأکید دارد.

۲۳. مفهوم به‌هم‌پیوستگی انسان‌ها بر این شناخت استوار است که هر انسان جزئی از یک کل بزرگ‌تر است؛ کلی که زمانی شکوفا می‌شود که همه اجزای تشکیل‌دهنده آن نیز امکان شکوفایی داشته باشند. زیستن در جوامع صلح‌آمیز، عادلانه و به‌هم‌پیوسته مستلزم پیوندی طبیعی، بی‌واسطه و بی‌چشمداشت از هم‌بستگی است؛ پیوندی که با تلاش مستمر برای برقراری روابطی صلح‌آمیز، همراه با گرایش به مراقبت از دیگران و از محیط‌زیست طبیعی، در گسترده‌ترین معنای این واژه، شناخته می‌شود.

۲۴. این ارزش ایجاب می‌کند که صلح، فراگیری، عدالت، انصاف و به‌هم‌پیوستگی در سراسر چرخه حیات سامانه‌های هوش مصنوعی ترویج شوند، به‌گونه‌ای که فرایندهای چرخه حیات این سامانه‌ها موجب جداسازی، شیء‌انگاری یا تضعیف آزادی و استقلال در تصمیم‌گیری، و نیز امنیت انسان‌ها و جوامع نشوند؛ افراد و گروه‌ها را در برابر یکدیگر قرار ندهند؛ یا همزیستی میان انسان‌ها، سایر موجودات زنده و محیط‌زیست طبیعی را تهدید نکنند.

سوم/دو: اصول^۱

تناسب و پرهیز از آسیب^۲

۲۵. باید پذیرفت که فناوری‌های هوش مصنوعی به‌خودی‌خود و لزوماً تضمین‌کننده شکوفایی انسان، محیط‌زیست و بوم‌سازگان‌ها نیستند. افزون بر این، هیچ‌یک از فرایندهای مرتبط با چرخه حیات سامانه‌های هوش مصنوعی نباید فراتر از آنچه برای دستیابی به اهداف مشروع یا مقاصد عینی ضرورت دارد، پیش برود و این فرایندها باید با شرایط زمینه مورد نظر متناسب باشند. در صورت احتمال بروز هرگونه آسیب به انسان‌ها، حقوق بشر و آزادی‌های بنیادین، جوامع، جامعه به‌طور کلی، محیط‌زیست یا بوم‌سازگان‌ها، باید اجرای رویه‌های ارزیابی خطر و اتخاذ تدابیری برای پیشگیری از وقوع چنین آسیبی تضمین شود.

۲۶. انتخاب استفاده از سامانه‌های هوش مصنوعی و نیز انتخاب روش هوش مصنوعی باید به شیوه‌های زیر توجه شود:

1. PRINCIPLES
2. Proportionality and Do No Harm

الف) روش انتخاب‌شده باید برای دستیابی به هدف مشروع مورد نظر، مناسب و متناسب باشد؛

ب) روش انتخاب‌شده نباید ارزش‌های بنیادینی را که در این سند بیان شده‌اند نقض کند؛ به‌ویژه نباید موجب نقض یا سوءاستفاده از حقوق بشر شود؛

ج) روش انتخاب‌شده باید با شرایط و زمینه استفاده متناسب باشد و بر مبنای علمی معتبر استوار باشد.

در موقعیت‌هایی که تصمیم‌ها آثار غیرقابل بازگشت یا بسیار دشوار برای جبران دارند، مانند تصمیم‌های مربوط به حیات و مرگ، تصمیم نهایی باید همواره توسط انسان اتخاذ شود. به‌ویژه، سامانه‌های هوش مصنوعی نباید برای امتیازدهی اجتماعی^۳ یا نظارت همگانی^۴ مورد استفاده قرار گیرند.

ایمنی و امنیت^۵

۲۷. باید از آسیب‌های ناخواسته (مخاطرات ایمنی) و نیز آسیب‌پذیری در برابر حملات (مخاطرات امنیتی) جلوگیری شود و این موارد در سراسر چرخه حیات سامانه‌های هوش مصنوعی مورد رسیدگی، پیشگیری و کاهش قرار گیرند تا ایمنی و امنیت انسان، محیط‌زیست و بوم‌سازگان‌ها تضمین شود. دستیابی به هوش مصنوعی ایمن و امن، مستلزم توسعه چارچوب‌های پایدار، حفظ حریم خصوصی و حفاظت از داده‌هاست که امکان آموزش و اعتبارسنجی بهتر مدل‌های هوش مصنوعی با استفاده از داده‌های باکیفیت را فراهم می‌آورند.

انصاف و عدم تبعیض^۶

۲۸. کنشگران هوش مصنوعی باید عدالت اجتماعی را ترویج کنند و از انصاف و عدم تبعیض، در هر شکل، مطابق با حقوق بین‌الملل حمایت نمایند. این امر مستلزم اتخاذ رویکردی فراگیر برای تضمین آن است که منافع فناوری‌های هوش مصنوعی برای همگان در دسترس و قابل بهره‌برداری باشد؛ با در نظر گرفتن نیازهای گروه‌های سنی مختلف، نظام‌های فرهنگی، گروه‌های زبانی گوناگون، افراد دارای معلولیت، دختران و زنان، و نیز افراد یا گروه‌های محروم، به‌حاشیه‌رانده‌شده، آسیب‌پذیر یا در موقعیت‌های آسیب‌پذیر. کشورهای عضو باید برای فراهم ساختن دسترسی فراگیر همگان، از جمله جوامع محلی، به سامانه‌های هوش مصنوعی با محتوای محلی و خدمات بومی، و با رعایت چندزبانگی و تنوع فرهنگی، تلاش کنند. همچنین باید برای کاهش شکاف‌های دیجیتال و تضمین دسترسی و مشارکت فراگیر در توسعه هوش مصنوعی اقدام نمایند. در سطح ملی، کشورهای عضو باید برابری میان مناطق روستایی و شهری و نیز میان همه افراد، صرف‌نظر

3. Social Scoring

4. Mass Surveillance

5. Safety and security

6. Fairness and non-discrimination

از نژاد، رنگ پوست، تبار، جنسیت، سن، زبان، دین، عقیده سیاسی، منشأ ملی، منشأ قومی، منشأ اجتماعی، وضعیت اقتصادی یا اجتماعی هنگام تولد، معلولیت یا هر وضعیت دیگر را از نظر دسترسی و مشارکت در چرخه حیات سامانه‌های هوش مصنوعی ترویج کنند. در سطح بین‌المللی، کشورهای دارای پیشرفت فناوری بیشتر، مسئولیت همبستگی با کشورهای کمتر توسعه‌یافته را بر عهده دارند تا از رهگذر دسترسی و مشارکت آن‌ها در چرخه حیات سامانه‌های هوش مصنوعی، منافع این فناوری‌ها به شکلی عادلانه‌تر در نظام جهانی اطلاعات، ارتباطات، فرهنگ، آموزش، پژوهش و نیز توسعه اجتماعی، اقتصادی و سیاسی توزیع شود.

۲۹. همه کنشگران باید تمامی تلاش‌های معقول را برای به حداقل رساندن و جلوگیری از تقویت یا تداوم کاربردها و پیامدهای تبعیض‌آمیز یا جانبدارانه در سراسر چرخه حیات سامانه‌های هوش مصنوعی به کار گیرند تا انصاف این سامانه‌ها تضمین شود. همچنین باید سازوکارهای مؤثر جبران خسارت در برابر تبعیض و تصمیم‌گیری‌های الگوریتمی جانبدارانه فراهم شود.

۳۰. علاوه بر این، شکاف‌های دیجیتالی و شکاف‌های دانشی، چه در داخل کشورها و چه میان کشورها، باید در سراسر چرخه حیات سامانه‌های هوش مصنوعی مورد توجه و رسیدگی قرار گیرند. این امر شامل برابری در کیفیت دسترسی به فناوری و داده‌ها، مطابق با چارچوب‌های حقوقی ملی، منطقه‌ای و بین‌المللی مرتبط، و نیز برابری در دسترسی به زیرساخت‌های ارتباطی، دانش، مهارت‌ها و مشارکت معنادار جوامع ذی‌نفع است، به‌گونه‌ای که با همه افراد به شکلی منصفانه رفتار شود.

پایداری^۷

۳۱. توسعه جوامع پایدار به تحقق مجموعه‌ای پیچیده از اهداف در ابعاد انسانی، اجتماعی، فرهنگی، اقتصادی و زیست‌محیطی وابسته است. ظهور فناوری‌های هوش مصنوعی، بسته به نحوه به‌کارگیری آن‌ها در کشورهای مختلف و در سطوح متفاوت توسعه، می‌تواند یا به تحقق اهداف پایداری کمک کند یا مانع دستیابی به آن‌ها شود. از این‌رو، ارزیابی مستمر آثار انسانی، اجتماعی، فرهنگی، اقتصادی و زیست‌محیطی فناوری‌های هوش مصنوعی باید با آگاهی کامل از پیامدهای این فناوری‌ها برای پایداری انجام گیرد؛ پایداری‌ای که مجموعه‌ای از اهداف همواره در حال تحول را در ابعاد گوناگون در بر می‌گیرد و در حال حاضر در اهداف توسعه پایدار سازمان ملل متحد تبلور یافته است.

حق بر حریم خصوصی و حفاظت از داده‌ها^۸

۳۲. حریم خصوصی، که حقی اساسی برای حمایت از کرامت انسانی، خودمختاری انسان

7. Sustainability

8. Right to Privacy, and Data Protection

و عاملیت انسانی است، باید در سراسر چرخه حیات سامانه‌های هوش مصنوعی محترم شمرده شود، مورد حمایت قرار گیرد و ترویج یابد. ضروری است داده‌های مورد استفاده در سامانه‌های هوش مصنوعی به شیوه‌ای گردآوری، استفاده، به اشتراک‌گذاری، بایگانی و حذف شوند که با حقوق بین‌الملل سازگار بوده و با ارزش‌ها و اصول مندرج در این توصیه‌نامه همخوانی داشته باشد، و در عین حال چارچوب‌های حقوقی ملی، منطقه‌ای و بین‌المللی مرتبط نیز رعایت شوند.

۳۳. باید چارچوب‌ها و سازوکارهای مناسب برای حفاظت و حکمرانی داده‌ها، با رویکردی چندذی‌نفعی، در سطح ملی یا بین‌المللی ایجاد شوند، از حمایت نظام‌های قضایی برخوردار باشند و در سراسر چرخه حیات سامانه‌های هوش مصنوعی تضمین گردند. این چارچوب‌ها و سازوکارها باید بر اصول و استانداردهای بین‌المللی حفاظت از داده‌ها در زمینه گردآوری، کنترل، پردازش، استفاده، افشا و حفاظت از داده‌های شخصی و نیز اعمال حقوق صاحبان داده استوار باشند و وجود هدفی مشروع و مبنای قانونی معتبر برای پردازش داده‌های شخصی، از جمله رضایت آگاهانه، را تضمین کنند.

۳۴. سامانه‌های الگوریتمی مستلزم انجام ارزیابی‌های مناسب اثرات بر حریم خصوصی هستند. این ارزیابی‌ها باید علاوه بر ملاحظات فنی، جنبه‌های اجتماعی و اخلاقی استفاده از این سامانه‌ها را نیز در بر گیرند و رویکرد «حفظ حریم خصوصی از بدو طراحی»^۹ را به صورت نوآورانه به کار گیرند. کنشگران هوش مصنوعی باید اطمینان حاصل کنند که در قبال طراحی و استقرار سامانه‌های هوش مصنوعی مسئولیت‌پذیر هستند، به گونه‌ای که اطلاعات شخصی در سراسر چرخه حیات این سامانه‌ها حفاظت شود.



نظارت و تصمیم‌گیری انسانی^{۱۰}

۳۵. کشورهای عضو باید اطمینان حاصل کنند که همواره امکان انتساب مسئولیت اخلاقی و حقوقی برای هر مرحله از چرخه حیات سامانه‌های هوش مصنوعی، و نیز هرگونه جبران خسارت مرتبط با این سامانه‌ها، به اشخاص حقیقی یا اشخاص حقوقی موجود وجود داشته باشد. نظارت انسانی نه تنها به نظارت مستقیم افراد، بلکه حسب مورد، به نظارت عمومی نیز اشاره دارد.

۳۶. ممکن است در برخی موارد، انسان‌ها به دلایل کارآمدی ترجیح دهند به سامانه‌های هوش مصنوعی اتکا کنند؛ با این حال، تصمیم‌گیری درباره واگذاری کنترل در شرایط محدود همچنان باید در اختیار انسان باقی بماند، زیرا انسان‌ها تنها به عنوان آخرین راه حل در فرایند تصمیم‌گیری و اقدام به سامانه‌های هوش مصنوعی متوسل می‌شوند. سامانه‌های هوش مصنوعی هرگز نمی‌توانند جایگزین مسئولیت‌پذیری و پاسخگویی نهایی انسان شوند. به عنوان یک قاعده، تصمیم‌های مربوط به حیات و مرگ نباید به سامانه‌های هوش مصنوعی واگذار شوند.

شفافیت و تبیین‌پذیری^{۱۱}

۳۷. شفافیت و تبیین‌پذیری سامانه‌های هوش مصنوعی غالباً از پیش شرط‌های اساسی برای تضمین احترام، حمایت و ترویج حقوق بشر، آزادی‌های بنیادین و اصول اخلاقی هستند. شفافیت برای آن ضروری است که نظام‌های مسئولیت‌پذیری ملی و بین‌المللی بتوانند به طور مؤثر عمل کنند. نبود شفافیت همچنین می‌تواند امکان به چالش کشیدن مؤثر تصمیم‌هایی را که بر پایه نتایج تولیدشده توسط سامانه‌های هوش مصنوعی اتخاذ شده‌اند، تضعیف کند، حق برخورداری از دادرسی عادلانه و جبران خسارت مؤثر را نقض نماید و دامنه کاربرد قانونی این سامانه‌ها را محدود سازد.

۳۸. باید در سراسر چرخه حیات سامانه‌های هوش مصنوعی، از جمله سامانه‌هایی که آثار فرامرزی دارند، برای افزایش شفافیت و تبیین‌پذیری آن‌ها تلاش شود تا از حکمرانی دموکراتیک پشتیبانی گردد. با این حال، میزان شفافیت و تبیین‌پذیری باید با توجه به زمینه و آثار هر سامانه متناسب باشد، زیرا ممکن است لازم باشد میان شفافیت و تبیین‌پذیری از یک سو، و اصول دیگری همچون حریم خصوصی، ایمنی و امنیت از سوی دیگر، توازن برقرار شود.

افراد باید به طور کامل آگاه شوند هرگاه تصمیمی بر اساس الگوریتم‌ها یا با اتکا به آن‌ها اتخاذ شده باشد، به ویژه هنگامی که چنین تصمیمی بر ایمنی یا حقوق بشر آنان تأثیر می‌گذارد. در این شرایط، آنان باید این حق را داشته باشند که از کنشگر مربوط به هوش مصنوعی یا نهاد عمومی ذی ربط درخواست توضیح کنند. همچنین افراد باید بتوانند به

10. Human oversight and determination

11. Transparency and explainability

دلایل تصمیمی که بر حقوق و آزادی‌هایشان اثر می‌گذارد دسترسی داشته باشند و این امکان را داشته باشند که اعتراض یا درخواست خود را نزد یکی از کارکنان تعیین شده در شرکت بخش خصوصی یا نهاد عمومی مسئول بررسی و اصلاح آن تصمیم ثبت کنند. همچنین کنشگران هوش مصنوعی باید در مواردی که یک محصول یا خدمت به‌طور مستقیم یا با کمک سامانه‌های هوش مصنوعی ارائه می‌شود، کاربران را به شیوه‌ای مناسب و به‌موقع آگاه سازند.

۳۹. از منظر اجتماعی - فنی، شفافیت بیشتر به شکل‌گیری جوامعی صلح‌آمیزتر، عادلانه‌تر، دموکراتیک‌تر و فراگیرتر کمک می‌کند. شفافیت امکان نظارت عمومی را فراهم می‌آورد؛ نظارتی که می‌تواند فساد و تبعیض را کاهش دهد و همچنین به شناسایی و پیشگیری از آثار منفی بر حقوق بشر کمک کند. هدف شفافیت، ارائه اطلاعات مناسب به مخاطبان مربوط است تا آنان بتوانند سامانه‌های هوش مصنوعی را درک کرده و به آن‌ها اعتماد کنند. در سطح هر سامانه هوش مصنوعی، شفافیت می‌تواند به افراد کمک کند تا دریابند هر مرحله از چرخه حیات سامانه چگونه طراحی و اجرا شده، چه تناسبی با زمینه و میزان حساسیت آن دارد و چه عواملی بر یک تصمیم مشخص اثر گذاشته‌اند. همچنین شفافیت می‌تواند مشخص کند که آیا تضمین‌های مناسبی، مانند تدابیر مربوط به ایمنی یا انصاف، در نظر گرفته شده‌اند یا خیر. در مواردی که آثار جدی بر حقوق بشر وجود دارد، شفافیت ممکن است مستلزم اشتراک‌گذاری کد یا مجموعه داده‌ها نیز باشد.

۴۰. تبیین‌پذیری به معنای قابل فهم ساختن و ارائه بینشی روشن درباره نتایج سامانه‌های هوش مصنوعی است. همچنین به قابل درک بودن داده‌های ورودی، خروجی و نحوه عملکرد هر یک از اجزای الگوریتمی و سهم آن‌ها در نتیجه نهایی سامانه اشاره دارد. از این‌رو، تبیین‌پذیری ارتباطی نزدیک با شفافیت دارد، زیرا نتایج و فرایندهای منتهی به آن‌ها باید با توجه به زمینه مربوط، قابل فهم و قابل ردگیری باشند. کنشگران هوش مصنوعی باید اطمینان حاصل کنند که الگوریتم‌های توسعه‌یافته از قابلیت تبیین برخوردارند. در کاربردهایی که آثار آن‌ها بر کاربران نهایی دائمی، غیرقابل بازگشت یا به هر نحو دیگر پرخطر است، باید تضمین شود که برای هر تصمیم منجر به اقدام، توضیحی معنادار ارائه می‌شود تا نتیجه حاصل شفاف تلقی شود.

۴۱. شفافیت و تبیین‌پذیری ارتباطی نزدیک با اتخاذ تدابیر مناسب برای مسئولیت‌پذیری و پاسخگویی دارند و از ارکان اساسی اعتمادپذیری سامانه‌های هوش مصنوعی به شمار می‌روند.

مسئولیت‌پذیری و پاسخگویی^{۱۲}

۴۲. کنشگران هوش مصنوعی و کشورهای عضو باید حقوق بشر و آزادی‌های بنیادین

را محترم بشمارند، از آن‌ها حمایت کنند و آن‌ها را ترویج دهند. آنان همچنین باید از محیط‌زیست و بوم‌سازگان‌ها حفاظت کنند و مسئولیت‌های اخلاقی و حقوقی متناسب با نقش خود را، مطابق با حقوق ملی و بین‌المللی - به‌ویژه تعهدات حقوق بشری کشورهای عضو - و رهنمودهای اخلاقی، در سراسر چرخه حیات سامانه‌های هوش مصنوعی بر عهده گیرند؛ از جمله در قبال کنشگران هوش مصنوعی که تحت قلمرو و کنترل مؤثر آن‌ها قرار دارند. مسئولیت اخلاقی و حقوقی مربوط به تصمیم‌ها و اقدامات مبتنی بر سامانه‌های هوش مصنوعی باید همواره، متناسب با نقش هر یک از کنشگران در چرخه حیات این سامانه‌ها، قابل انتساب و پیگیری باشد.

۴۳. باید سازوکارهای مناسب برای نظارت، ارزیابی اثرات، ممیزی و دقت لازم^{۱۳}، از جمله سازوکارهای حمایت از افشاگران، توسعه یابند تا پاسخگویی سامانه‌های هوش مصنوعی و آثار آن‌ها در سراسر چرخه حیاتشان تضمین شود. همچنین، طراحی‌های فنی و نهادی باید به گونه‌ای باشند که قابلیت حسابرسی و ردیابی^{۱۴} عملکرد سامانه‌های هوش مصنوعی را، به‌ویژه به‌منظور رسیدگی به تعارض با هنجارها و استانداردهای حقوق بشر و نیز تهدیدهای متوجه محیط‌زیست و سلامت بوم‌سازگان‌ها، تضمین کنند.

آگاهی و سواد^{۱۵}

۴۴. آگاهی عمومی و درک فناوری‌های هوش مصنوعی و ارزش داده‌ها باید از طریق آموزش باز و در دسترس، مشارکت مدنی، آموزش مهارت‌های دیجیتال و اخلاق هوش مصنوعی، سواد رسانه‌ای و سواد اطلاعاتی تقویت شود. این اقدامات باید به‌طور مشترک توسط دولت‌ها، سازمان‌های بین‌دولتی، جامعه مدنی، دانشگاه‌ها، رسانه‌ها، رهبران محلی و بخش خصوصی انجام گیرد و در عین حال تنوع زبانی، اجتماعی و فرهنگی موجود نیز مورد توجه قرار گیرد، تا مشارکت مؤثر عمومی تضمین شود و همه اعضای جامعه بتوانند درباره استفاده از سامانه‌های هوش مصنوعی تصمیم‌های آگاهانه بگیرند و از نفوذ و تأثیرگذاری ناروا مصون بمانند.

۴۵. یادگیری درباره آثار سامانه‌های هوش مصنوعی باید شامل یادگیری درباره، از طریق و برای حقوق بشر و آزادی‌های بنیادین باشد. این بدان معناست که رویکرد آموزش و فهم سامانه‌های هوش مصنوعی باید بر حقوق بشر و آزادی‌های بنیادین استوار باشد.

حکمرانی چندذینفعی، تطبیقی و همکاری^{۱۶}

۴۶. در استفاده از داده‌ها باید به حقوق بین‌الملل و حاکمیت ملی احترام گذاشته شود. این امر بدان معناست که دولت‌ها باید مطابق با حقوق بین‌الملل، داده‌هایی را که در

13. Due Diligence

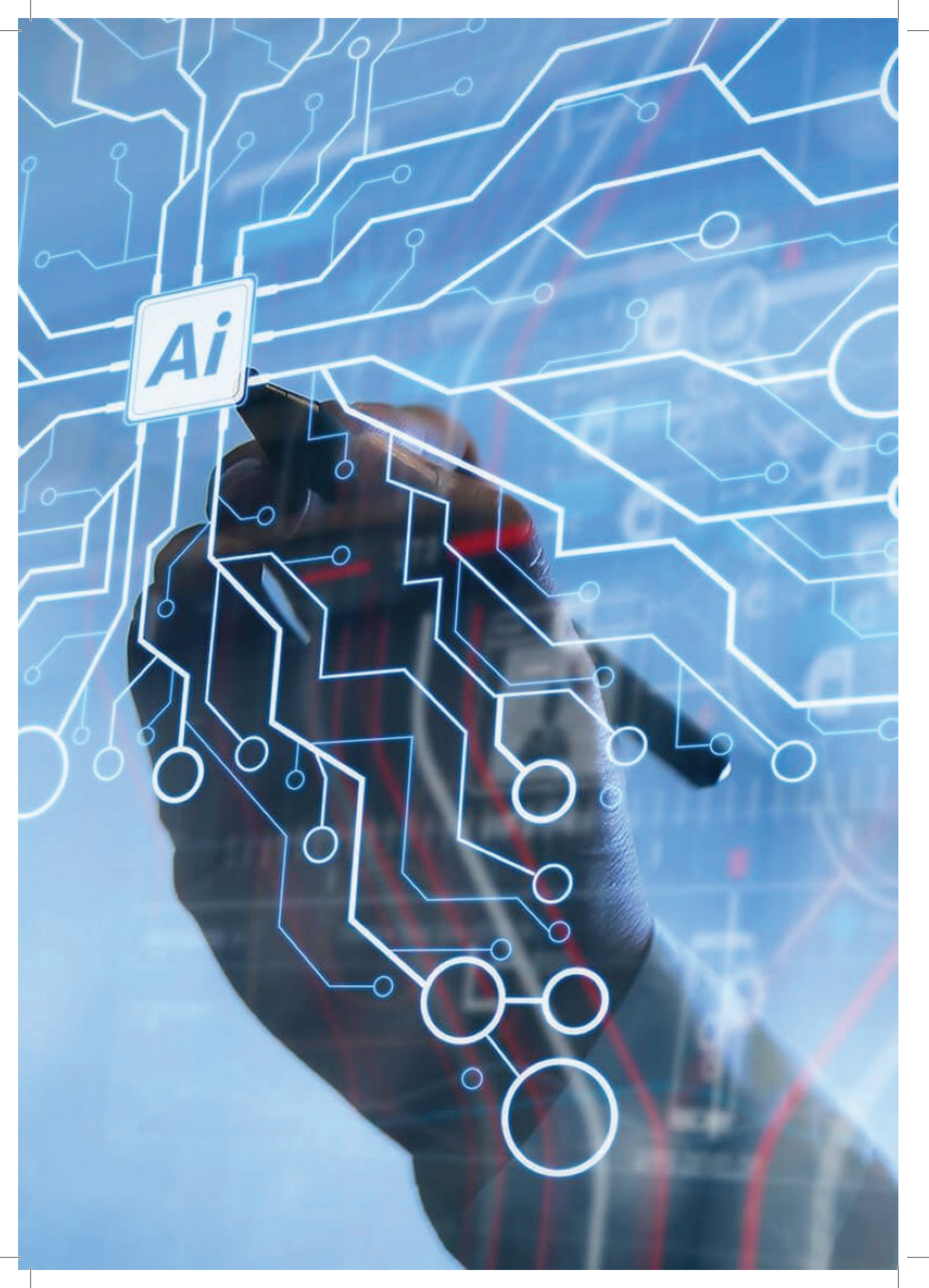
14. auditability and traceability

15. Awareness and literacy

16. Multi-stakeholder and adaptive governance and collaboration

قلمرو آن‌ها تولید می‌شود یا از قلمروشان عبور می‌کند، تنظیم و مدیریت کنند و تدابیری برای تنظیم مؤثر داده‌ها، از جمله حفاظت از داده‌ها، بر پایه احترام به حق حریم خصوصی و مطابق با حقوق بین‌الملل و سایر هنجارها و استانداردهای حقوق بشر اتخاذ نمایند.

۴۷. مشارکت ذی‌نفعان مختلف در سراسر چرخه حیات سامانه‌های هوش مصنوعی برای تحقق رویکردهای فراگیر در حکمرانی هوش مصنوعی ضروری است؛ رویکردهایی که امکان بهره‌مندی همگان از منافع این فناوری و مشارکت آن در توسعه پایدار را فراهم می‌کنند. ذی‌نفعان شامل، اما محدود به، دولت‌ها، سازمان‌های بین‌دولتی، جامعه فنی، جامعه مدنی، پژوهشگران و دانشگاهیان، رسانه‌ها، بخش آموزش، سیاست‌گذاران، شرکت‌های خصوصی، نهادهای حقوق بشر و برابری، نهادهای ناظر بر مقابله با تبعیض، گروه‌های جوانان و کودکان هستند. برای تسهیل همکاری، باید استانداردهای باز و قابلیت همکاری میان سامانه‌ها مورد استفاده قرار گیرد. همچنین باید تدابیری اتخاذ شود که تغییرات فناوری، ظهور گروه‌های جدید ذی‌نفع و امکان مشارکت معنادار گروه‌های به‌حاشیه‌رانده‌شده، جوامع و افراد را در نظر بگیرد و، در صورت اقتضا، درباره مردمان بومی، حق خودفرمانی بر داده‌هایشان^{۱۷} محترم شمرده شود.



Ai

The image features a hand holding a black pen, pointing towards a digital interface. The background is a light blue gradient with white circuit-like lines and nodes. A small white square with the text 'Ai' is positioned near the pen tip. The overall theme is technology and artificial intelligence.

بخش چهارم: حوزه‌های اقدام سیاستی^۱

۴۸. اقدام‌های سیاستی تشریح‌شده در حوزه‌های سیاستی زیر، ارزش‌ها و اصول مقرر در این توصیه‌نامه را به مرحله اجرا درمی‌آورند. اقدام اصلی بر عهده کشورهای عضو است تا تدابیر مؤثر، از جمله برای نمونه چارچوب‌ها یا سازوکارهای سیاستی، را ایجاد کنند و اطمینان یابند که سایر ذی‌نفعان، مانند شرکت‌های بخش خصوصی، مؤسسات دانشگاهی و پژوهشی و سازمان‌های جامعه مدنی نیز، از جمله از طریق تشویق همه ذی‌نفعان به توسعه ابزارهای حقوق بشر، حاکمیت قانون، دموکراسی، ارزیابی اثرات اخلاقی و دقت لازم، مطابق با رهنمودهایی از جمله اصول راهنمای سازمان ملل متحد درباره کسب‌وکار و حقوق بشر، به این تدابیر پایبند باشند. فرایند تدوین چنین سیاست‌ها یا سازوکارهایی باید با مشارکت همه ذی‌نفعان انجام شود و شرایط و اولویت‌های هر کشور عضو را مدنظر قرار دهد. یونسکو می‌تواند در تدوین، پایش و ارزیابی این سازوکارهای سیاستی، شریک و حامی کشورهای عضو باشد.

۴۹. یونسکو اذعان دارد که کشورهای عضو، از نظر آمادگی برای اجرای این توصیه‌نامه، در ابعاد علمی، فناورانه، اقتصادی، آموزشی، حقوقی، مقرراتی، زیرساختی، اجتماعی، فرهنگی و سایر ابعاد، در سطوح متفاوتی قرار دارند. همچنین تأکید می‌شود که «آمادگی» در اینجا مفهومی پویا و در حال تحول است. برای تسهیل اجرای مؤثر این توصیه‌نامه، یونسکو:

۱. روش‌شناسی ارزیابی آمادگی را تدوین خواهد کرد تا به کشورهای عضو علاقه‌مند در شناسایی وضعیت خود در مقاطع مختلف مسیر آمادگی، در طیفی از ابعاد گوناگون، کمک کند؛ و

۲. از کشورهای عضو علاقه‌مند در زمینه تدوین روش‌شناسی یونسکو برای ارزیابی اثرات اخلاقی^۱ فناوری‌های هوش مصنوعی، تبادل بهترین رویه‌ها، دستورالعمل‌های ارزیابی و سایر سازوکارها و مطالعات تحلیلی پشتیبانی خواهد کرد.

حوزه سیاستی ۱: ارزیابی اثرات اخلاقی

۵۰. کشورهای عضو باید چارچوب‌هایی برای انجام ارزیابی اثرات، از جمله ارزیابی اثرات اخلاقی، ایجاد کنند تا منافع، دغدغه‌ها و مخاطرات سامانه‌های هوش مصنوعی شناسایی و ارزیابی شوند و نیز تدابیر مناسب برای پیشگیری، کاهش و پایش خطرات، در کنار سایر سازوکارهای تضمین، پیش‌بینی شود. این ارزیابی‌ها باید آثار سامانه‌های هوش مصنوعی بر حقوق بشر و آزادی‌های بنیادین، به‌ویژه – اما نه محدود به – حقوق گروه‌های به‌حاشیه‌رانده‌شده و آسیب‌پذیر یا افراد در موقعیت‌های آسیب‌پذیر، حقوق کار، محیط‌زیست و بوم‌سازگان‌ها، و نیز پیامدهای اخلاقی و اجتماعی را شناسایی کنند و مشارکت شهروندان را در چارچوب ارزش‌ها و اصول مندرج در این توصیه‌نامه تسهیل نمایند.

1. Ethical Impact Assessment – EIA

۵۱. کشورهای عضو و شرکت‌های بخش خصوصی باید سازوکارهای دقت لازم و نظارت را برای شناسایی، پیشگیری، کاهش و پاسخگویی درباره نحوه رسیدگی به آثار سامانه‌های هوش مصنوعی بر احترام به حقوق بشر، حاکمیت قانون و جوامع فراگیر ایجاد کنند. کشورهای عضو همچنین باید بتوانند آثار اجتماعی - اقتصادی سامانه‌های هوش مصنوعی بر فقر را ارزیابی کرده و اطمینان حاصل کنند که شکاف میان افراد ثروتمند و فقیر، و نیز شکاف دیجیتال میان کشورها و درون آن‌ها، در نتیجه گسترش گسترده فناوری‌های هوش مصنوعی در حال حاضر و آینده افزایش نمی‌یابد.

برای تحقق این هدف، به‌ویژه باید پروتکل‌های الزام‌آور شفافیت، متناسب با حق دسترسی به اطلاعات، از جمله اطلاعات دارای منفعت عمومی که در اختیار نهادهای خصوصی قرار دارد، اجرا شوند. کشورهای عضو، شرکت‌های بخش خصوصی و جامعه مدنی باید آثار اجتماعی و روان‌شناختی توصیه‌های مبتنی بر هوش مصنوعی بر انسان‌ها و بر استقلال آنان در تصمیم‌گیری را بررسی کنند. سامانه‌های هوش مصنوعی که به‌عنوان مخاطره‌ای بالقوه برای حقوق بشر شناسایی می‌شوند، باید پیش از عرضه به بازار، در صورت لزوم و در شرایط واقعی، به‌طور گسترده توسط کنشگران هوش مصنوعی و در چارچوب ارزیابی اثرات اخلاقی آزمون شوند.

۵۲. کشورهای عضو و بنگاه‌های اقتصادی باید تدابیر مناسبی برای پایش همه مراحل چرخه حیات سامانه‌های هوش مصنوعی، از جمله عملکرد الگوریتم‌های مورد استفاده در تصمیم‌گیری، داده‌ها و نیز کنشگران هوش مصنوعی دخیل در این فرایند، به‌ویژه در خدمات عمومی و در مواردی که تعامل مستقیم با کاربر نهایی وجود دارد، به‌عنوان بخشی از ارزیابی اثرات اخلاقی، به اجرا بگذارند. تعهدات کشورهای عضو در زمینه حقوق بشر باید بخشی از ابعاد اخلاقی ارزیابی سامانه‌های هوش مصنوعی را تشکیل دهد.

۵۳. دولت‌ها باید چارچوبی مقرراتی اتخاذ کنند که به‌ویژه برای نهادهای عمومی، رویه‌ای مشخص جهت انجام ارزیابی اثرات اخلاقی سامانه‌های هوش مصنوعی فراهم سازد تا پیامدها پیش‌بینی، مخاطرات کاهش، از بروز آثار زیان‌بار جلوگیری، مشارکت شهروندان تسهیل و چالش‌های اجتماعی مدیریت شوند. این ارزیابی‌ها همچنین باید سازوکارهای نظارتی مناسب، از جمله قابلیت حسابرسی، قابلیت ردیابی و تبیین‌پذیری^۱ را ایجاد کنند تا ارزیابی الگوریتم‌ها، داده‌ها و فرایندهای طراحی امکان‌پذیر شود و علاوه بر آن، بازبینی مستقل سامانه‌های هوش مصنوعی را نیز در بر گیرد.

ارزیابی‌های اثرات اخلاقی، هر جا که مقتضی باشد، باید شفاف بوده و برای عموم در دسترس باشند. این ارزیابی‌ها همچنین باید چندرشته‌ای، چندذی‌نفعی، چندفرهنگی، کثرت‌گرا و فراگیر باشند. نهادهای عمومی باید موظف شوند سامانه‌های هوش مصنوعی را که توسط خود توسعه داده یا به‌کار گرفته‌اند، از طریق استقرار سازوکارها و ابزارهای مناسب، به‌طور مستمر پایش کنند.

1. Explainability

حوزه سیاستی ۲: حکمرانی و راهبری اخلاقی^۲

۵۴. کشورهای عضو باید اطمینان حاصل کنند که سازوکارهای حکمرانی هوش مصنوعی، فراگیر^۳، شفاف^۴، چندرشته‌ای^۵، چندجانبه^۶ (که امکان کاهش و جبران خسارت‌های فرامرزی را نیز در بر می‌گیرد) و چندذی‌نفعی باشند. به‌طور خاص، حکمرانی باید جنبه‌های آینده‌نگری، حمایت مؤثر، پایش آثار، اجرای مقررات و جبران خسارت را در بر گیرد.

۵۵. کشورهای عضو باید اطمینان حاصل کنند که خسارت‌های ناشی از سامانه‌های هوش مصنوعی، از طریق ایجاد سازوکارهای اجرایی قوی و اقدامات جبرانی، مورد بررسی قرار گرفته و جبران شوند تا اطمینان حاصل شود که حقوق بشر، آزادی‌های بنیادین و حاکمیت قانون، هم در جهان دیجیتال و هم در جهان فیزیکی، رعایت می‌شوند. چنین سازوکارها و اقداماتی باید شامل سازوکارهای جبران خسارت ارائه‌شده از سوی شرکت‌های بخش خصوصی و نهادهای عمومی باشند. برای تحقق این هدف، باید قابلیت حسابرسی و ردیابی سامانه‌های هوش مصنوعی تقویت شود.

علاوه بر این، کشورهای عضو باید ظرفیت‌های نهادی خود را برای اجرای این تعهد تقویت کنند و با پژوهشگران و سایر ذی‌نفعان همکاری داشته باشند تا هرگونه استفاده بالقوه مخرب از سامانه‌های هوش مصنوعی را شناسایی، پیشگیری و کاهش دهند.

۵۶. کشورهای عضو تشویق می‌شوند راهبردهای ملی و منطقه‌ای هوش مصنوعی را تدوین کنند و اشکال مختلف حکمرانی نرم^۷، مانند سازوکارهای صدور گواهی برای سامانه‌های هوش مصنوعی و شناسایی متقابل این گواهی‌ها را، متناسب با میزان حساسیت حوزه کاربرد، آثار مورد انتظار بر حقوق بشر، محیط‌زیست و بوم‌سازگان‌ها و سایر ملاحظات اخلاقی مندرج در این توصیه‌نامه، مدنظر قرار دهند.

چنین سازوکاری می‌تواند سطوح مختلف ممیزی سامانه‌ها، داده‌ها و میزان پایبندی به دستورالعمل‌های اخلاقی و الزامات رویه‌ای مرتبط با جنبه‌های اخلاقی را شامل شود. در عین حال، این سازوکار نباید به دلیل تحمیل بار اداری بیش از حد، مانع نوآوری شود یا شرکت‌های کوچک و متوسط، استارت‌آپ‌ها، سازمان‌های جامعه مدنی، و نیز نهادهای پژوهشی و علمی را در موقعیت نامطلوب قرار دهد.

این سازوکارها همچنین باید شامل مؤلفه‌ای برای پایش منظم باشند تا استحکام سامانه، حفظ یکپارچگی و استمرار رعایت دستورالعمل‌های اخلاقی در سراسر چرخه حیات سامانه هوش مصنوعی تضمین شود و در صورت لزوم، صدور مجدد گواهی انجام گیرد.

2. ETHICAL GOVERNANCE AND STEWARDSHIP

3. Inclusive

4. Transparent

5. Multidisciplinary

6. Multilateral

7. Soft Governance

۵۷. کشورهای عضو و نهادهای عمومی باید درباره سامانه‌های هوش مصنوعی موجود و سامانه‌های پیشنهادی، خودارزیابی شفاف^۸ انجام دهند. این ارزیابی باید، به‌ویژه، بررسی کند که آیا به‌کارگیری سامانه هوش مصنوعی مناسب است یا خیر؛ و در صورت مثبت بودن پاسخ، ارزیابی‌های تکمیلی برای تعیین مناسب‌ترین روش استفاده انجام شود. همچنین باید بررسی شود که آیا استفاده از آن سامانه منجر به نقض یا سوءاستفاده از تعهدات کشورهای عضو در زمینه حقوق بشر خواهد شد یا خیر؛ و اگر چنین باشد، استفاده از آن ممنوع شود.

۵۸. کشورهای عضو باید نهادهای عمومی، شرکت‌های بخش خصوصی و سازمان‌های جامعه مدنی را تشویق کنند که ذی‌نفعان مختلف را در حکمرانی هوش مصنوعی مشارکت دهند و ایجاد جایگاه مسئول مستقل اخلاق هوش مصنوعی^۹ یا سازوکار مشابهی را برای نظارت بر ارزیابی اثرات اخلاقی، ممیزی، پایش مستمر و تضمین رعایت اصول اخلاقی در سامانه‌های هوش مصنوعی مدنظر قرار دهند.

کشورهای عضو، شرکت‌های بخش خصوصی و سازمان‌های جامعه مدنی، با حمایت یونسکو، تشویق می‌شوند شبکه‌ای از مسئولان مستقل اخلاق هوش مصنوعی ایجاد کنند تا از این فرایند در سطوح ملی، منطقه‌ای و بین‌المللی پشتیبانی شود.

۵۹. کشورهای عضو باید توسعه و دسترسی به یک زیست‌بوم دیجیتال را برای توسعه اخلاق‌محور و فراگیر سامانه‌های هوش مصنوعی در سطح ملی تقویت کنند؛ از جمله با رفع شکاف‌های موجود در دسترسی به مراحل مختلف چرخه حیات سامانه‌های هوش مصنوعی، در عین حال که به همکاری‌های بین‌المللی نیز کمک می‌کنند. چنین زیست‌بومی، به‌ویژه، شامل فناوری‌ها و زیرساخت‌های دیجیتال و نیز سازوکارهای مناسب برای اشتراک‌گذاری دانش هوش مصنوعی است.

۶۰. کشورهای عضو باید، با همکاری سازمان‌های بین‌المللی، شرکت‌های فراملی، مؤسسات دانشگاهی و جامعه مدنی، سازوکارهایی ایجاد کنند تا مشارکت فعال همه کشورهای عضو، به‌ویژه کشورهای با درآمد کم و متوسط (LMICs) و به‌طور خاص کشورهای کمتر توسعه‌یافته (LDCs)، کشورهای در حال توسعه محصور در خشکی (LLDCs) و کشورهای جزیره‌ای کوچک در حال توسعه (SIDS)، در گفت‌وگوهای بین‌المللی مربوط به حکمرانی هوش مصنوعی تضمین شود.

این هدف می‌تواند از طریق تأمین منابع مالی، تضمین مشارکت برابر مناطق مختلف یا سایر سازوکارهای مناسب محقق شود. افزون بر این، برای تضمین فراگیر بودن این مجامع، کشورهای عضو باید رفت‌وآمد کنشگران هوش مصنوعی به قلمرو خود و از قلمرو خود را، به‌ویژه از کشورهای با درآمد کم و متوسط و به‌طور خاص کشورهای کمتر توسعه‌یافته، کشورهای محصور در خشکی و کشورهای جزیره‌ای کوچک در حال توسعه،

8. transparent self-assessment

9. Independent AI Ethics Officer

به‌منظور مشارکت در این مجامع تسهیل کنند.

۶۱. اصلاح قوانین موجود یا تدوین قوانین جدید ملی درباره سامانه‌های هوش مصنوعی باید با تعهدات کشورهای عضو در حوزه حقوق بشر سازگار باشد و احترام، حمایت و ترویج حقوق بشر و آزادی‌های بنیادین را در سراسر چرخه حیات سامانه‌های هوش مصنوعی تضمین کند.

ترویج این اهداف همچنین باید در قالب ابتکارهای حکمرانی، نمونه‌های موفق همکاری در زمینه سامانه‌های هوش مصنوعی، و دستورالعمل‌های فنی و روش‌شناختی ملی و بین‌المللی، همگام با پیشرفت فناوری‌های هوش مصنوعی، صورت گیرد. بخش‌های مختلف، از جمله بخش خصوصی، باید در رویه‌های خود در زمینه سامانه‌های هوش مصنوعی، با بهره‌گیری از ابزارهای موجود و ابزارهای جدید در کنار این توصیه‌نامه، حقوق بشر و آزادی‌های بنیادین را محترم بشمارند، از آن‌ها حمایت کنند و آن‌ها را ترویج دهند.

۶۲. دولت‌هایی که سامانه‌های هوش مصنوعی را برای حوزه‌های حساس از منظر حقوق بشر – مانند اجرای قانون، نظام رفاه اجتماعی، اشتغال، رسانه و اطلاع‌رسانی، خدمات سلامت و دستگاه قضایی – به‌کار می‌گیرند، باید سازوکارهای مناسبی برای پایش آثار اجتماعی و اقتصادی این سامانه‌ها ایجاد کنند.

این نظارت باید توسط نهادهای مستقل، از جمله مراجع حفاظت از داده‌ها، نهادهای نظارتی تخصصی و سایر مراجع مسئول انجام شود.

۶۳. دولت‌های عضو باید ظرفیت نظام قضایی را برای رسیدگی به پرونده‌های مرتبط با سامانه‌های هوش مصنوعی، مطابق حاکمیت قانون و استانداردهای بین‌المللی، تقویت کنند.

در صورت استفاده از هوش مصنوعی در فرایندهای قضایی، باید تضمین‌های کافی برای حفظ حقوق بنیادین، استقلال قضایی، نظارت انسانی و اعتماد عمومی به دستگاه عدالت فراهم شود.

۶۴. دولت‌های عضو باید اطمینان حاصل کنند که دولت‌ها و سازمان‌های چندجانبه نقش پیشرو در تضمین ایمنی و امنیت سامانه‌های هوش مصنوعی، با مشارکت همه ذی‌نفعان، ایفا می‌کنند.

همچنین لازم است استانداردهای بین‌المللی قابل اندازه‌گیری و قابل آزمون برای ایمنی، امنیت و شفافیت سامانه‌های هوش مصنوعی تدوین شود تا ارزیابی عینی میزان انطباق آنها امکان‌پذیر باشد.

علاوه بر این، دولت‌ها و بنگاه‌های اقتصادی باید از پژوهش‌های راهبردی درباره مخاطرات ایمنی و امنیت هوش مصنوعی حمایت کرده و تحقیقات در زمینه شفافیت، تبیین‌پذیری، فراگیری و سواد هوش مصنوعی را از طریق سرمایه‌گذاری بیشتر تقویت کنند.

۶۵. دولت‌های عضو باید سیاست‌هایی اتخاذ کنند تا اقدامات تمامی بازیگران هوش مصنوعی در سراسر چرخه عمر سامانه‌ها با حقوق بین‌الملل حقوق بشر، استانداردها و اصول مربوط سازگار باشد.

در این مسیر، تنوع فرهنگی و اجتماعی، آداب و رسوم محلی و سنت‌های دینی باید مورد توجه قرار گیرد، مشروط بر آنکه اصل جهان‌شمولی و برتری حقوق بشر حفظ شود.

۶۶. دولت‌های عضو باید سازوکارهایی ایجاد کنند که بازیگران هوش مصنوعی را ملزم کند هرگونه کلیشه‌سازی در داده‌ها یا خروجی‌های سامانه‌های هوش مصنوعی را - چه ناشی از طراحی و چه ناشی از سهل‌انگاری - شناسایی، افشا و اصلاح کنند.

همچنین باید اطمینان حاصل شود که داده‌های آموزشی سامانه‌ها موجب تقویت نابرابری‌های فرهنگی، اقتصادی یا اجتماعی، تعصب، اطلاعات نادرست یا محدود شدن آزادی بیان و دسترسی به اطلاعات نشوند. توجه ویژه‌ای نیز باید به مناطقی که با کمبود داده مواجه هستند، معطوف شود.

۶۷. دولت‌های عضو باید سیاست‌هایی برای افزایش تنوع و فراگیری در تیم‌های توسعه هوش مصنوعی و مجموعه داده‌های آموزشی تدوین کنند تا این تیم‌ها بازتاب‌دهنده تنوع جمعیتی کشورها باشند.

همچنین باید دسترسی برابر همه افراد، به‌ویژه گروه‌های حاشیه‌نشین در مناطق شهری و روستایی، به فناوری‌های هوش مصنوعی و منافع حاصل از آنها تضمین شود.

۶۸. دولت‌های عضو باید چارچوب‌های مقرراتی مناسب را تدوین، بازنگری و به‌روزرسانی کنند تا مسئولیت‌پذیری و پاسخگویی درباره محتوا و نتایج سامانه‌های هوش مصنوعی در تمام مراحل چرخه عمر آنها تضمین شود.

در صورت لزوم، باید نظام‌های مسئولیت مدنی ایجاد یا تفسیر قوانین موجود شفاف شود تا مسئولیت پیامدهای عملکرد سامانه‌های هوش مصنوعی به اشخاص حقیقی یا حقوقی قابل انتساب باشد.

مسئولیت نهایی همواره بر عهده انسان‌ها یا اشخاص حقوقی است و سامانه‌های هوش مصنوعی نباید شخصیت حقوقی مستقل داشته باشند.

۶۹. برای تدوین هنجارهای جدید یا اصلاح چارچوب‌های حقوقی موجود، دولت‌های عضو باید همه بازیگران مرتبط با هوش مصنوعی، از جمله پژوهشگران، جامعه مدنی، مجریان قانون، بیمه‌گران، سرمایه‌گذاران، تولیدکنندگان، مهندسان، وکلا و کاربران را در فرایند سیاست‌گذاری مشارکت دهند.

همچنین استفاده از ابزارهایی مانند سندباکس‌های مقرراتی^۱ و نمونه‌های آزمایشی سیاستی برای آزمون قوانین و مقررات پیش از تصویب رسمی آنها توصیه می‌شود.

۱. سندباکس نظارتی (Regulatory Sandbox) به‌عنوان ابزاری برای تنظیم نوآوری در حوزه‌های فناوری، به‌ویژه فین‌تک، در بسیاری از کشورها مورد استفاده قرار گرفته است.

۷۰. دولت‌های عضو باید الزامات روشنی برای شفافیت و تبیین‌پذیری سامانه‌های هوش مصنوعی تعیین کنند تا اعتمادپذیری کل چرخه عمر این سامانه‌ها تضمین شود. این الزامات باید طراحی و اجرای سازوکارهای ارزیابی آثار را نیز دربرگیرد و متناسب با نوع کاربرد، هدف استفاده، گروه مخاطبان و امکان‌پذیری اجرای هر سامانه تنظیم شود.

حوزه سیاستی ۳: سیاست داده^{۱۱}

۷۱. دولت‌های عضو باید راهبردهای حکمرانی داده را تدوین و اجرا کنند تا ارزیابی مستمر کیفیت داده‌های آموزشی مورد استفاده در سامانه‌های هوش مصنوعی تضمین شود. این ارزیابی باید شامل کفایت فرآیندهای گردآوری و انتخاب داده، اتخاذ تدابیر مناسب برای امنیت و حفاظت از داده‌ها، و ایجاد سازوکارهای بازخورد به‌منظور یادگیری از خطاها و به‌اشتراک‌گذاری تجربه‌ها و رویه‌های موفق میان تمامی بازیگران حوزه هوش مصنوعی باشد.

۷۲. دولت‌های عضو باید مطابق با حقوق بین‌الملل، تضمین‌های لازم را برای حمایت از حق حریم خصوصی برقرار کنند و نگرانی‌هایی مانند نظارت^{۱۲} و مراقبت گسترده را مورد توجه قرار دهند.

از جمله، دولت‌ها باید چارچوب‌های قانونی مناسب را تصویب یا تقویت کنند تا حفاظت کافی از داده‌های شخصی، مطابق با استانداردهای بین‌المللی، تضمین شود. همچنین دولت‌ها باید تمامی بازیگران حوزه هوش مصنوعی، به‌ویژه بنگاه‌های اقتصادی، را به رعایت استانداردهای بین‌المللی و انجام ارزیابی آثار حریم خصوصی^{۱۳} به‌عنوان بخشی از ارزیابی آثار اخلاقی تشویق کنند.

این ارزیابی‌ها باید آثار اجتماعی و اقتصادی گسترده‌تر پردازش داده‌های مورد نظر را نیز در نظر بگیرند و اصل حریم خصوصی در طراحی در توسعه سامانه‌ها رعایت شود. احترام، حفاظت و ترویج حق حریم خصوصی باید در سراسر چرخه عمر سامانه‌های هوش مصنوعی تضمین شود.

۷۳. دولت‌های عضو باید اطمینان حاصل کنند که افراد، حقوق خود نسبت به داده‌های شخصی‌شان را حفظ می‌کنند و این حقوق در چارچوبی حقوقی مورد حمایت قرار می‌گیرد. این چارچوب باید به‌ویژه موارد زیر را تضمین کند:

- شفافیت در پردازش داده‌ها؛
- تضمین‌های مناسب برای پردازش داده‌های حساس؛
- سطح مناسبی از حفاظت از داده‌ها؛

11. Data Policy

12. Surveillance

13. Privacy Impact Assessment

- نظام‌های مؤثر و معنادار پاسخگویی؛
- بهره‌مندی کامل صاحبان داده از حقوق خود؛
- امکان دسترسی، اصلاح و حذف داده‌های شخصی در سامانه‌های هوش مصنوعی، مگر در موارد استثنایی که مطابق قوانین بین‌المللی محدود شده باشد؛
- رعایت کامل قوانین حفاظت از داده‌ها، به‌ویژه هنگامی که داده‌ها برای مقاصد تجاری، مانند تبلیغات هدفمند، یا انتقال فرامرزی داده‌ها استفاده می‌شوند؛
- وجود نظام نظارت مستقل و مؤثر در چارچوب حکمرانی داده، به‌گونه‌ای که افراد کنترل داده‌های شخصی خود را حفظ کنند و در عین حال، جریان آزاد و مسئولانه اطلاعات در سطح بین‌المللی و دسترسی مشروع به داده‌ها نیز امکان‌پذیر باشد.

۷۴. دولت‌های عضو باید سیاست‌های ملی داده یا چارچوب‌های معادل آن را تدوین یا تقویت کنند تا امنیت کامل داده‌های شخصی و داده‌های حساس تضمین شود؛ زیرا افشای چنین داده‌هایی می‌تواند آسیب‌های جدی یا خسارات جبران‌ناپذیری به افراد وارد کند. این داده‌ها، از جمله، شامل موارد زیر هستند:

- اطلاعات مربوط به سوابق کیفری، دادرسی‌های قضایی، محکومیت‌ها و تدابیر امنیتی؛
- داده‌های زیستی^{۱۴}، ژنتیکی و سلامت؛
- و داده‌های شخصی مرتبط با نژاد، رنگ پوست، تبار، جنسیت، سن، زبان، دین، عقاید سیاسی، منشأ ملی، منشأ قومی، خاستگاه اجتماعی و سایر ویژگی‌های حساس افراد.

۷۵. دولت‌های عضو باید داده‌های باز^{۱۵} را ترویج کنند. در این راستا، دولت‌های عضو باید بازنگری در سیاست‌ها و چارچوب‌های مقرراتی خود، از جمله در زمینه دسترسی به اطلاعات و دولت باز، را مدنظر قرار دهند تا این سیاست‌ها با الزامات خاص هوش مصنوعی سازگار شوند. همچنین باید سازوکارهایی مانند مخازن باز برای داده‌ها و کدهای منبعی که از محل منابع عمومی تأمین مالی شده یا در اختیار نهادهای عمومی هستند، و نیز امانت‌های داده^{۱۶}، را به‌منظور حمایت از اشتراک‌گذاری ایمن، منصفانه، قانونی و اخلاقی داده‌ها، از جمله از طریق این سازوکارها، ترویج کنند.

۷۶. دولت‌های عضو باید استفاده از مجموعه داده‌های باکیفیت و قابل‌اعتماد را برای آموزش، توسعه و به‌کارگیری سامانه‌های هوش مصنوعی ترویج و تسهیل کنند و بر گردآوری و استفاده از این داده‌ها نظارت دقیق داشته باشند. این اقدام، در صورت امکان و قابلیت اجرا، می‌تواند شامل سرمایه‌گذاری برای ایجاد مجموعه داده‌های استاندارد طلایی، از جمله مجموعه داده‌های باز و قابل‌اعتماد، باشد که از تنوع کافی برخوردار بوده، بر مبنای معتبر قانونی، از جمله رضایت صاحبان داده‌ها در مواردی که قانون اقتضا می‌کند، گردآوری شده

14. Biometric

15. Open Data

16. Data Trusts

باشند. همچنین باید استانداردهایی برای حاشیه‌نویسی و برچسب‌گذاری مجموعه‌داده‌ها تشویق شود؛ از جمله تفکیک داده‌ها بر اساس جنسیت و سایر مبانی، به‌گونه‌ای که نحوه گردآوری مجموعه‌داده و ویژگی‌های آن به‌آسانی قابل تشخیص باشد.

۷۷. دولت‌های عضو، همان‌گونه که در گزارش هیئت عالی همکاری دیجیتال دبیرکل سازمان ملل متحد^{۱۷} نیز پیشنهاد شده است، با حمایت سازمان ملل متحد و یونسکو، باید رویکرد دارایی‌های مشترک دیجیتال^{۱۸} را در مورد داده‌ها، هر جا که مناسب باشد، اتخاذ کنند؛ قابلیت همکاری ابزارها، مجموعه‌داده‌ها و رابط‌های سامانه‌های میزبان داده را افزایش دهند و شرکت‌های بخش خصوصی را تشویق کنند داده‌هایی را که گردآوری می‌کنند، حسب مورد، برای پژوهش، نوآوری یا منافع عمومی، با همه ذی‌نفعان به اشتراک بگذارند. همچنین باید تلاش‌های بخش‌های عمومی و خصوصی را برای ایجاد سکوه‌های مشارکتی به‌منظور اشتراک‌گذاری داده‌های باکیفیت در فضاهای داده‌ای قابل اعتماد و امن ترویج کنند.



حوزه سیاستی ۴: توسعه و همکاری بین‌المللی^{۱۹}

۷۸. دولت‌های عضو و شرکت‌های فراملی باید با گنجاندن مباحث مربوط به مسائل اخلاقی هوش مصنوعی در مجامع بین‌المللی، بین‌دولتی و چندذی‌نفع مرتبط، اخلاق هوش مصنوعی را در اولویت قرار دهند.

17. United Nations Secretary-General's High-level Panel on Digital Cooperation

18. Digital Commons

19. DEVELOPMENT AND INTERNATIONAL COOPERATION

۷۹. دولت‌های عضو باید اطمینان حاصل کنند که استفاده از هوش مصنوعی در حوزه‌های توسعه، از جمله آموزش، علم، فرهنگ، ارتباطات و اطلاعات، مراقبت‌های سلامت، کشاورزی و تأمین غذا، محیط زیست، مدیریت منابع طبیعی و زیرساخت‌ها، برنامه‌ریزی و رشد اقتصادی و سایر حوزه‌ها، با ارزش‌ها و اصول مندرج در این توصیه‌نامه مطابقت داشته باشد.

۸۰. دولت‌های عضو باید از طریق سازمان‌های بین‌المللی برای فراهم کردن بسترهای همکاری بین‌المللی در زمینه هوش مصنوعی برای توسعه اقدام کنند؛ از جمله از طریق ارائه تخصص، تأمین مالی، داده، دانش تخصصی حوزه‌ای، زیرساخت و تسهیل همکاری چندذی‌نفعی به‌منظور رسیدگی به چالش‌های توسعه، به‌ویژه برای کشورهای با درآمد پایین و متوسط، به‌خصوص کشورهای کم‌توسعه‌یافته، کشورهای در حال توسعه محصور در خشکی و کشورهای جزیره‌ای کوچک در حال توسعه.

۸۱. دولت‌های عضو باید برای ترویج همکاری بین‌المللی در پژوهش و نوآوری هوش مصنوعی تلاش کنند؛ از جمله از طریق ایجاد مراکز و شبکه‌های پژوهش و نوآوری که مشارکت و نقش رهبری پژوهشگران کشورهای با درآمد پایین و متوسط و دیگر کشورها، از جمله کشورهای کم‌توسعه‌یافته، کشورهای در حال توسعه محصور در خشکی و کشورهای جزیره‌ای کوچک در حال توسعه را تقویت کنند.

۸۲. دولت‌های عضو باید پژوهش در زمینه اخلاق هوش مصنوعی را از طریق مشارکت سازمان‌های بین‌المللی، مؤسسات پژوهشی و نیز شرکت‌های فراملی ترویج کنند؛ به‌گونه‌ای که این پژوهش‌ها مبنایی برای استفاده اخلاقی از سامانه‌های هوش مصنوعی توسط نهادهای عمومی و خصوصی فراهم آورند. این امر شامل پژوهش درباره قابلیت اعمال چارچوب‌های اخلاقی خاص در فرهنگ‌ها و زمینه‌های مختلف و نیز امکان توسعه راهکارهای فناورانه عملی در چارچوب این اصول اخلاقی است.

۸۳. دولت‌های عضو باید همکاری و مشارکت بین‌المللی در حوزه هوش مصنوعی را برای کاهش شکاف‌های ژئوفناورانه تشویق کنند. تبادل فناوری و مشورت‌ها باید میان دولت‌های عضو و مردم آن‌ها، میان بخش‌های عمومی و خصوصی، و نیز میان کشورهای دارای بیشترین و کمترین سطح پیشرفت فناوری، با رعایت کامل حقوق بین‌الملل، انجام گیرد.

حوزه سیاستی ۵: محیط زیست و اکوسیستم‌ها^{۲۰}

۸۴. دولت‌های عضو و بنگاه‌های اقتصادی باید آثار زیست‌محیطی مستقیم و غیرمستقیم را در سراسر چرخه عمر سامانه‌های هوش مصنوعی ارزیابی کنند؛ از جمله، اما نه محدود به، ردپای کربنی، مصرف انرژی و آثار زیست‌محیطی استخراج مواد اولیه مورد نیاز برای پشتیبانی از تولید فناوری‌های هوش مصنوعی، و همچنین آثار زیست‌محیطی سامانه‌های

هوش مصنوعی و زیرساخت‌های داده را کاهش دهند. دولت‌های عضو باید اطمینان حاصل کنند که همه کنشگران حوزه هوش مصنوعی از قوانین، سیاست‌ها و رویه‌های زیست‌محیطی تبعیت می‌کنند.

۸۵. دولت‌های عضو باید، هرگاه لازم و مقتضی باشد، مشوق‌هایی را برای تضمین توسعه و به‌کارگیری راهکارهای مبتنی بر هوش مصنوعی که بر حقوق بشر و اصول اخلاقی استوار هستند، به‌منظور افزایش تاب‌آوری در برابر مخاطرات، پایش، حفاظت و احیای محیط زیست و اکوسیستم‌ها، و حفظ سیاره زمین، ایجاد کنند. این سامانه‌های هوش مصنوعی باید مشارکت جوامع محلی و بومی را در سراسر چرخه عمر سامانه‌های هوش مصنوعی در بر گیرند و از رویکردهای اقتصاد چرخشی و الگوهای مصرف و تولید پایدار حمایت کنند. برخی از نمونه‌های استفاده از سامانه‌های هوش مصنوعی، در صورت لزوم و اقتضا، عبارت‌اند از:

- (الف) حمایت از حفاظت، پایش و مدیریت منابع طبیعی.
- (ب) حمایت از پیش‌بینی، پیشگیری، کنترل و کاهش مشکلات مرتبط با تغییرات اقلیمی.
- (ج) حمایت از ایجاد نظام غذایی کارآمدتر و پایدارتر.
- (د) حمایت از تسریع دسترسی به انرژی پایدار و گسترش استفاده فراگیر از آن.
- (هـ) امکان‌پذیر ساختن و ترویج نهاده‌سازی زیرساخت‌های پایدار، الگوهای کسب‌وکار پایدار و تأمین مالی پایدار در راستای توسعه پایدار.
- (و) شناسایی آلاینده‌ها یا پیش‌بینی میزان آلودگی و در نتیجه کمک به ذی‌نفعان مربوط برای شناسایی، برنامه‌ریزی و اجرای مداخلات هدفمند به‌منظور پیشگیری و کاهش آلودگی و میزان مواجهه با آن.

۸۶. در انتخاب روش‌های هوش مصنوعی، با توجه به ماهیت بالقوه داده‌بر^{۲۱} یا منبع‌بر^{۲۲} برخی از این روش‌ها و آثار آن‌ها بر محیط زیست، دولت‌های عضو باید اطمینان حاصل کنند که کنشگران حوزه هوش مصنوعی، مطابق با اصل تناسب، روش‌های هوش مصنوعی کارآمد از نظر مصرف داده، انرژی و منابع را ترجیح می‌دهند. همچنین باید الزاماتی تدوین شود تا شواهد مناسب و کافی برای اثبات اینکه یک کاربرد هوش مصنوعی اثر موردنظر را خواهد داشت، یا اینکه تدابیر حفاظتی همراه با آن کاربرد می‌تواند استفاده از آن را توجیه کند، در دسترس باشد. چنانچه این امر امکان‌پذیر نباشد، باید اصل احتیاط^{۲۳} در اولویت قرار گیرد و در مواردی که آثار منفی نامتناسبی بر محیط زیست وجود داشته باشد، از به‌کارگیری هوش مصنوعی نباید استفاده شود.

21. data-intensive

22. resource-intensive

23. Precautionary Principle



حوزه سیاستی ۶: جنسیت^{۲۴}

۸۷. دولت‌های عضو باید اطمینان حاصل کنند که ظرفیت فناوری‌های دیجیتال و هوش مصنوعی برای کمک به تحقق برابری جنسیتی به حداکثر برسد و باید تضمین کنند که حقوق بشر، آزادی‌های بنیادین، ایمنی و تمامیت جسمانی دختران و زنان در هیچ‌یک از مراحل چرخه عمر سامانه‌های هوش مصنوعی نقض نشود. افزون بر این، ارزیابی تأثیرات اخلاقی^{۲۵} باید به‌طور فراگیر، دیدگاه جنسیتی را در بر گیرد.

۸۸. دولت‌های عضو باید از بودجه‌های عمومی خود، منابع مالی اختصاصی مرتبط با طرح‌های تأمین مالی پاسخگو به جنسیت اختصاص دهند، اطمینان حاصل کنند که سیاست‌های ملی دیجیتال شامل برنامه اقدام جنسیتی باشد و سیاست‌های مرتبط، برای مثال در حوزه آموزش نیروی کار، را با هدف حمایت از دختران و زنان تدوین کنند تا از حذف آنان از اقتصاد دیجیتال مبتنی بر هوش مصنوعی جلوگیری شود.

همچنین باید سرمایه‌گذاری ویژه‌ای برای اجرای برنامه‌های هدفمند و استفاده از ادبیات و رویکردهای ویژه جنسیتی، به‌منظور افزایش فرصت‌های مشارکت دختران و زنان در علوم، فناوری، مهندسی و ریاضیات^{۲۶} (STEM)، از جمله فناوری‌های اطلاعات و ارتباطات (ICT)، آمادگی شغلی، قابلیت اشتغال، توسعه برابر مسیر شغلی و رشد حرفه‌ای آنان، مورد توجه

24. Gender

25. Ethical Impact Assessment

26. Science, Technology, Engineering and Mathematics

قرار گرفته و اجرا شود.

۸۹. دولت‌های عضو باید اطمینان حاصل کنند که ظرفیت سامانه‌های هوش مصنوعی برای پیشبرد تحقق برابری جنسیتی بالفعل شود. آنان باید تضمین کنند که این فناوری‌ها شکاف‌های گسترده جنسیتی موجود در حوزه‌های مختلف دنیای غیردیجیتال را تشدید نکنند، بلکه در جهت رفع این شکاف‌ها عمل کنند.

این شکاف‌ها شامل موارد زیر است:

- شکاف دستمزد جنسیتی؛
- نابرابری در میزان حضور زنان و مردان در برخی حرفه‌ها و فعالیت‌ها؛
- کمبود نمایندگی در مناصب عالی مدیریتی، هیئت‌مدیره‌ها و گروه‌های پژوهشی حوزه هوش مصنوعی؛
- شکاف آموزشی؛
- شکاف در دسترسی، پذیرش، استفاده و استطاعت مالی فناوری‌های دیجیتال و هوش مصنوعی؛
- و توزیع نابرابر کار بدون دستمزد و مسئولیت‌های مراقبتی در جوامع.

۹۰. دولت‌های عضو باید اطمینان حاصل کنند که کلیشه‌های جنسیتی و سوگیری‌های تبعیض‌آمیز به سامانه‌های هوش مصنوعی منتقل نشوند و در عوض، این موارد شناسایی و به‌صورت پیش‌دستانه اصلاح شوند. همچنین باید برای جلوگیری از تشدید آثار منفی شکاف‌های فناورانه بر تحقق برابری جنسیتی و پیشگیری از خشونت‌هایی مانند آزار، زورگویی یا قاچاق دختران و زنان و سایر گروه‌های کم‌نماینده، از جمله در فضای برخط، اقدامات لازم صورت گیرد.

۹۱. دولت‌های عضو باید با ارائه و ترویج مشوق‌های اقتصادی، مقرراتی و سایر مشوق‌ها و برنامه‌های حمایتی، کارآفرینی، مشارکت و حضور زنان را در تمامی مراحل چرخه عمر سامانه‌های هوش مصنوعی تشویق کنند. همچنین باید سیاست‌هایی را اجرا کنند که مشارکت متوازن جنسیتی را در پژوهش‌های هوش مصنوعی در دانشگاه‌ها، حضور زنان در مدیریت ارشد شرکت‌های دیجیتال و هوش مصنوعی، هیئت‌مدیره‌ها و گروه‌های پژوهشی تضمین کند.

دولت‌های عضو باید اطمینان حاصل کنند که منابع مالی عمومی (برای نوآوری، پژوهش و فناوری) به برنامه‌ها و شرکت‌هایی با نمایندگی روشن جنسیتی اختصاص یابد و همچنین سرمایه‌گذاری‌های خصوصی نیز بر اساس اصول اقدام مثبت^{۲۷} به همین مسیر هدایت شوند.

علاوه بر این، باید سیاست‌هایی برای ایجاد محیط‌های کاری عاری از آزار تدوین و اجرا شود و هم‌زمان، انتقال بهترین رویه‌ها درباره ارتقای تنوع در سراسر چرخه عمر سامانه‌های

27. Affirmative action principles

هوش مصنوعی مورد تشویق قرار گیرد.

۹۲. دولت‌های عضو باید با ارائه مشوق‌هایی برای ورود دختران و زنان به این حوزه، ایجاد سازوکارهایی برای مقابله با کلیشه‌های جنسیتی و آزار در جامعه پژوهشی هوش مصنوعی، و تشویق نهادهای دانشگاهی و بخش خصوصی به اشتراک‌گذاری بهترین رویه‌ها برای ارتقای تنوع جنسیتی، تنوع جنسیتی را در پژوهش‌های هوش مصنوعی در دانشگاه و صنعت ترویج کنند.

۹۳. یونسکو می‌تواند در ایجاد یک مخزن^{۲۸} از بهترین رویه‌ها برای تشویق مشارکت دختران، زنان و گروه‌های کم‌نماینده در تمامی مراحل چرخه عمر سامانه‌های هوش مصنوعی نقش‌آفرینی کند.

حوزه سیاستی ۷: فرهنگ^{۲۹}

۹۴. دولت‌های عضو تشویق می‌شوند، هر جا که مناسب باشد، از سامانه‌های هوش مصنوعی در حفاظت، غنابخشی، شناخت، ترویج، مدیریت و افزایش دسترس‌پذیری میراث فرهنگی ملموس، مستند و ناملموس، از جمله زبان‌ها و دانش‌های در معرض خطر و نیز زبان‌ها و دانش‌های بومی، استفاده کنند. این امر می‌تواند از طریق ایجاد یا به‌روزرسانی برنامه‌های آموزشی مرتبط با کاربرد سامانه‌های هوش مصنوعی در این حوزه‌ها و نیز، در صورت اقتضا، با تضمین رویکردی مشارکتی که نهادها و عموم مردم را در بر گیرد، محقق شود.

۹۵. دولت‌های عضو تشویق می‌شوند آثار فرهنگی سامانه‌های هوش مصنوعی، به‌ویژه کاربردهای پردازش زبان طبیعی^{۳۰} مانند ترجمه خودکار و دستیارهای صوتی، را بر ظرافت‌ها و پیچیدگی‌های زبان و بیان انسانی بررسی و ارزیابی کنند. این ارزیابی‌ها باید مبنایی برای طراحی و اجرای راهبردهایی فراهم آورند که با کاهش شکاف‌های فرهنگی و افزایش درک متقابل میان انسان‌ها، منافع این سامانه‌ها را به حداکثر برسانند و در عین حال، پیامدهای منفی مانند کاهش میزان استفاده از برخی زبان‌ها را که ممکن است به نابودی زبان‌های در معرض خطر بینجامد، مورد توجه و رسیدگی قرار دهند.

۹۶. دولت‌های عضو باید آموزش هوش مصنوعی و آموزش‌های دیجیتال را برای هنرمندان و فعالان حرفه‌ای حوزه‌های خلاق ترویج کنند تا آنان بتوانند تناسب و قابلیت استفاده از فناوری‌های هوش مصنوعی را در حرفه خود ارزیابی کرده و در طراحی و پیاده‌سازی فناوری‌های مناسب هوش مصنوعی مشارکت داشته باشند. این امر از آن جهت اهمیت دارد که فناوری‌های هوش مصنوعی به‌طور فزاینده برای خلق، تولید، توزیع، انتشار و

28. Repository

29. Culture

30. Natural Language Processing – NLP

مصرف انواع کالاها و خدمات فرهنگی به کار گرفته می‌شوند؛ در عین حال باید اهمیت حفظ میراث فرهنگی، تنوع فرهنگی و آزادی هنری همواره مدنظر قرار گیرد.

۹۷. دولت‌های عضو باید آگاهی‌بخشی و ارزیابی ابزارهای هوش مصنوعی را در میان صنایع فرهنگی محلی و نیز بنگاه‌های کوچک و متوسط فعال در حوزه فرهنگ ترویج کنند تا از خطر تمرکزگرایی در بازار فرهنگی جلوگیری شود.

۹۸. دولت‌های عضو باید شرکت‌های فناوری و سایر ذی‌نفعان را برای ترویج عرضه متنوع و دسترسی چندگانه به جلوه‌های فرهنگی مشارکت دهند و به‌ویژه اطمینان حاصل کنند که نظام‌های توصیه‌گر الگوریتمی، دیده‌شدن و قابلیت کشف محتوای محلی را افزایش می‌دهند.



۹۹. دولت‌های عضو باید پژوهش‌های جدید را در حوزه تلاقی هوش مصنوعی و مالکیت فکری^{۳۱} تقویت کنند؛ برای مثال، به‌منظور تعیین اینکه آیا و چگونه آثار خلق‌شده با استفاده از فناوری‌های هوش مصنوعی می‌توانند تحت حمایت حقوق مالکیت فکری قرار گیرند. همچنین دولت‌های عضو باید ارزیابی کنند که فناوری‌های هوش مصنوعی چگونه بر حقوق یا منافع صاحبان حقوق مالکیت فکری، که آثارشان برای پژوهش، توسعه، آموزش یا اجرای کاربردهای هوش مصنوعی مورد استفاده قرار می‌گیرد، تأثیر می‌گذارند.

۱۰۰. دولت‌های عضو باید موزه‌ها، نگارخانه‌ها، کتابخانه‌ها و آرشیوهای ملی را تشویق کنند تا از سامانه‌های هوش مصنوعی برای برجسته‌سازی مجموعه‌های خود، ارتقای کتابخانه‌ها، پایگاه‌های داده و پایگاه‌های دانش خود، و نیز فراهم کردن دسترسی کاربران به این

31. Intellectual Property – IP

منابع استفاده کنند.

حوزه سیاستی ۸: آموزش و پژوهش^{۳۲}

۱.۱. دولت‌های عضو باید با سازمان‌های بین‌المللی، مؤسسات آموزشی و نهادهای خصوصی و غیردولتی همکاری کنند تا آموزش مناسب سواد هوش مصنوعی را در تمامی سطوح و برای عموم مردم در همه کشورها فراهم آورند. هدف از این اقدام توانمندسازی افراد و کاهش شکاف‌های دیجیتال و نابرابری‌های دسترسی دیجیتال ناشی از گسترش گسترده استفاده از سامانه‌های هوش مصنوعی است.

۱.۲. دولت‌های عضو باید کسب مهارت‌های پیش‌نیاز برای آموزش هوش مصنوعی را ترویج کنند؛ از جمله:

- سواد پایه؛
- مهارت‌های ریاضی؛
- برنامه‌نویسی و مهارت‌های دیجیتال؛
- سواد رسانه‌ای و اطلاعاتی؛
- تفکر انتقادی و خلاق؛
- کار تیمی؛
- مهارت‌های ارتباطی؛
- مهارت‌های اجتماعی - عاطفی؛
- و مهارت‌های اخلاق هوش مصنوعی.

این امر به‌ویژه در کشورها و مناطق یا بخش‌هایی از کشورها که در آموزش این مهارت‌ها با کمبودهای قابل توجهی مواجه هستند، اهمیت دارد.

۱.۳. دولت‌های عضو باید برنامه‌های عمومی آگاهی‌بخشی درباره تحولات هوش مصنوعی را ترویج کنند؛ از جمله در زمینه داده‌ها، فرصت‌ها و چالش‌های ناشی از فناوری‌های هوش مصنوعی، تأثیر سامانه‌های هوش مصنوعی بر حقوق بشر و پیامدهای آن، از جمله حقوق کودکان. این برنامه‌ها باید برای گروه‌های غیرفنی و نیز گروه‌های فنی به‌طور یکسان در دسترس و قابل استفاده باشند.

۱.۴. دولت‌های عضو باید ابتکارات پژوهشی درباره استفاده مسئولانه و اخلاقی از فناوری‌های هوش مصنوعی در آموزش، تربیت معلم و آموزش الکترونیکی و سایر حوزه‌های مرتبط را تشویق کنند تا فرصت‌های موجود در این حوزه تقویت و چالش‌ها و مخاطرات آن کاهش یابد. این ابتکارات باید با ارزیابی مناسب کیفیت آموزش و تأثیر استفاده از فناوری‌های هوش مصنوعی بر دانش‌آموزان و معلمان همراه باشد.

دولت‌های عضو همچنین باید اطمینان حاصل کنند که فناوری‌های هوش مصنوعی موجب توانمندسازی دانش‌آموزان و معلمان شده و تجربه آموزشی آنان را ارتقا می‌بخشند؛ با این توجه که جنبه‌های ارتباطی و اجتماعی و نیز ارزش شیوه‌های سنتی آموزش، در روابط معلم - دانش‌آموز و دانش‌آموز - دانش‌آموز، از اهمیت اساسی برخوردارند و هنگام بررسی به‌کارگیری فناوری‌های هوش مصنوعی در آموزش باید مورد توجه قرار گیرند.

سامانه‌های هوش مصنوعی مورد استفاده در فرایند یادگیری باید از نظر پایش، ارزیابی توانایی‌ها یا پیش‌بینی رفتار یادگیرندگان، مشمول الزامات سخت‌گیرانه باشند. هوش مصنوعی باید از فرایند یادگیری حمایت کند، بدون آنکه توانایی‌های شناختی را کاهش دهد یا اطلاعات حساس را استخراج کند، و این امر باید مطابق با استانداردهای مربوط به حفاظت از داده‌های شخصی انجام شود.

داده‌هایی که در جریان تعامل یادگیرنده با سامانه هوش مصنوعی و به‌منظور کسب دانش گردآوری می‌شوند، نباید در معرض سوءاستفاده، تصاحب غیرمجاز یا بهره‌برداری مجرمانه، از جمله برای مقاصد تجاری، قرار گیرند.

۱.۵. دولت‌های عضو باید مشارکت و نقش رهبری دختران و زنان، افراد متعلق به قومیت‌ها و فرهنگ‌های گوناگون، اشخاص دارای معلولیت، افراد به‌حاشیه‌رانده‌شده و آسیب‌پذیر یا افرادی که در شرایط آسیب‌پذیر قرار دارند، اقلیت‌ها و تمامی افرادی را که از مزایای کامل شمول دیجیتال برخوردار نیستند، در برنامه‌های آموزش هوش مصنوعی در تمامی سطوح ترویج کنند. همچنین باید پایش و تبادل بهترین رویه‌ها در این زمینه با سایر دولت‌های عضو را تشویق و تسهیل کنند.

۱.۶. کشورهای عضو باید، مطابق با برنامه‌ها و سنت‌های آموزشی ملی خود، برنامه‌های درسی اخلاق هوش مصنوعی را در تمامی سطوح توسعه دهند و همکاری متقابل میان آموزش مهارت‌های فنی و جنبه‌های انسانی، اخلاقی و اجتماعی آموزش هوش مصنوعی را ترویج کنند. دوره‌های آنلاین و منابع دیجیتال آموزش اخلاق هوش مصنوعی باید به زبان‌های محلی، از جمله زبان‌های بومی، توسعه یابند و تنوع محیط‌ها را در نظر بگیرند، به‌ویژه با تضمین دسترسی‌پذیری برای افراد دارای معلولیت.

۱.۷. کشورهای عضو باید پژوهش در حوزه هوش مصنوعی، به‌ویژه پژوهش اخلاق هوش مصنوعی، را ترویج و حمایت کنند، از جمله از طریق سرمایه‌گذاری یا ایجاد مشوق‌هایی برای بخش‌های عمومی و خصوصی، با اذعان به اینکه پژوهش سهم قابل‌توجهی در توسعه و بهبود فناوری‌های هوش مصنوعی دارد، با هدف ترویج حقوق بین‌الملل و ارزش‌ها و اصول تعیین‌شده در این توصیه‌نامه. کشورهای عضو همچنین باید بهترین رویه‌ها و همکاری با پژوهشگران و شرکت‌هایی که هوش مصنوعی را به‌صورت اخلاقی توسعه می‌دهند، به‌طور عمومی ترویج کنند.

۱۰۸. کشورهای عضو باید اطمینان حاصل کنند که پژوهشگران هوش مصنوعی در اخلاق پژوهش آموزش دیده‌اند و ملزم شوند ملاحظات اخلاقی را در طراحی‌ها، محصولات و انتشارات خود لحاظ کنند، به‌ویژه در تحلیل مجموعه داده‌هایی که استفاده می‌کنند، نحوه برچسب‌گذاری آن‌ها، و کیفیت و دامنه نتایج با کاربردهای احتمالی.

۱۰۹. کشورهای عضو باید شرکت‌های بخش خصوصی را تشویق کنند تا دسترسی جامعه علمی به داده‌های خود برای پژوهش را تسهیل کنند، به‌ویژه در کشورهای با درآمد پایین و متوسط، خصوصاً کشورهای کمتر توسعه‌یافته، کشورهای در حال توسعه محصور در خشکی و کشورهای جزیره‌ای کوچک در حال توسعه. این دسترسی باید مطابق با استانداردهای مرتبط حریم خصوصی و حفاظت از داده باشد.



۱۱۰. برای تضمین ارزیابی انتقادی پژوهش هوش مصنوعی و نظارت مناسب بر سوءاستفاده‌های بالقوه یا آثار نامطلوب، کشورهای عضو باید اطمینان حاصل کنند که هرگونه توسعه آینده در حوزه فناوری‌های هوش مصنوعی مبتنی بر پژوهش علمی دقیق و مستقل باشد، و پژوهش میان‌رشته‌ای را با مشارکت حوزه‌هایی فراتر از علوم، فناوری، مهندسی و ریاضیات (STEM)، مانند مطالعات فرهنگی، آموزش، اخلاق، روابط بین‌الملل، حقوق، زبان‌شناسی، فلسفه، علوم سیاسی، جامعه‌شناسی و روان‌شناسی ترویج کنند.

۱۱۱. با اذعان به اینکه فناوری‌های هوش مصنوعی فرصت‌های بزرگی برای پیشبرد دانش و عمل علمی فراهم می‌کنند، به‌ویژه در رشته‌هایی که به‌طور سنتی مبتنی بر مدل هستند، کشورهای عضو باید جوامع علمی را تشویق کنند تا از مزایا، محدودیت‌ها و خطرات

استفاده از آن‌ها آگاه باشند؛ از جمله تلاش برای اطمینان از اینکه نتایج حاصل از رویکردهای داده‌محور، مدل‌ها و درمان‌ها استحکام و اعتبار کافی دارند. همچنین کشورهای عضو باید از نقش جامعه علمی در مشارکت در سیاست‌گذاری و افزایش آگاهی درباره نقاط قوت و ضعف فناوری‌های هوش مصنوعی استقبال و حمایت کنند.

حوزه سیاستی ۹: ارتباطات و اطلاعات^{۳۳}

۱۱۲. کشورهای عضو باید از سامانه‌های هوش مصنوعی برای بهبود دسترسی به اطلاعات و دانش استفاده کنند. این می‌تواند شامل حمایت از پژوهشگران، دانشگاهیان، روزنامه‌نگاران، عموم مردم و توسعه‌دهندگان باشد تا آزادی بیان، آزادی‌های علمی و دانشگاهی، دسترسی به اطلاعات، و افشای فعالانه‌تر داده‌ها و اطلاعات رسمی تقویت شود.

۱۱۳. کشورهای عضو باید اطمینان حاصل کنند که کنشگران هوش مصنوعی به آزادی بیان و همچنین دسترسی به اطلاعات در زمینه تولید، تعدیل و گردآوری محتوای خودکار احترام گذاشته و آن را ترویج می‌کنند. چارچوب‌های مناسب، از جمله مقررات، باید شفافیت در ارتباطات و اپراتورهای اطلاعات آنلاین را ممکن سازند و تضمین کنند که کاربران به تنوع دیدگاه‌ها دسترسی دارند، و همچنین سازوکارهایی برای اطلاع‌رسانی سریع به کاربران درباره دلایل حذف یا سایر اقدامات روی محتوا، و سازوکارهای اعتراض برای جبران فراهم باشد.

۱۱۴. کشورهای عضو باید در مهارت‌های سواد دیجیتال و رسانه‌ای و اطلاعاتی سرمایه‌گذاری کرده و آن‌ها را ترویج دهند تا تفکر انتقادی و شایستگی‌های لازم برای درک استفاده و پیامدهای سامانه‌های هوش مصنوعی تقویت شود، به‌منظور کاهش و مقابله با اطلاعات نادرست، بداطلاعی و نفرت‌پراکنی، درک و ارزیابی بهتر از اثرات مثبت و بالقوه زیان‌بار سامانه‌های توصیه‌گر، باید بخشی از این تلاش‌ها باشد.

۱۱۵. کشورهای عضو باید محیط‌های توانمندسازی برای رسانه‌ها ایجاد کنند تا حقوق و منابع لازم برای گزارش مؤثر درباره منافع و مضار سامانه‌های هوش مصنوعی را داشته باشند، و همچنین رسانه‌ها را تشویق کنند تا در عملیات خود از هوش مصنوعی به‌صورت اخلاقی استفاده کنند.

حوزه سیاستی ۱۰: اقتصاد و کار^{۳۴}

33. COMMUNICATION AND INFORMATION

34. ECONOMY AND LABOUR

۱۱۶. کشورهای عضو باید تأثیر سامانه‌های هوش مصنوعی بر بازارهای کار و پیامدهای آن برای الزامات آموزشی را در تمامی کشورها، با تأکید ویژه بر کشورهایی که اقتصاد آن‌ها کاربرمحور است، ارزیابی و بررسی کنند. این امر می‌تواند شامل معرفی طیف گسترده‌تری از مهارت‌های «هسته‌ای» و میان‌رشته‌ای در تمامی سطوح آموزشی باشد تا به کارکنان فعلی و نسل‌های جدید فرصت عادلانه‌ای برای یافتن شغل در بازاری به‌سرعت در حال تغییر بدهد، و آگاهی آن‌ها را از جنبه‌های اخلاقی سامانه‌های هوش مصنوعی تضمین کند. مهارت‌هایی مانند «یادگیری نحوه یادگیری»، ارتباطات، تفکر انتقادی، کار تیمی، همدلی، و توانایی انتقال دانش میان حوزه‌ها، باید در کنار مهارت‌های تخصصی، فنی و حتی کارهای کم‌مهارت آموزش داده شوند. شفافیت درباره مهارت‌های مورد تقاضا و به‌روزرسانی برنامه‌های درسی بر این اساس، کلیدی است.

۱۱۷. کشورهای عضو باید توافق‌نامه‌های همکاری میان دولت‌ها، نهادهای دانشگاهی، مؤسسات آموزش فنی و حرفه‌ای، صنعت، سازمان‌های کارگری و جامعه مدنی را برای پر کردن شکاف میان مهارت‌ها و هم‌راستاسازی برنامه‌های آموزشی با پیامدهای آینده کار و نیازهای صنعت، از جمله بنگاه‌های کوچک و متوسط، حمایت کنند. روش‌های آموزش و یادگیری مبتنی بر پروژه برای هوش مصنوعی باید ترویج شوند تا امکان مشارکت میان نهادهای عمومی، شرکت‌های بخش خصوصی، دانشگاه‌ها و مراکز پژوهشی فراهم شود.



۱۱۸. کشورهای عضو باید با شرکت‌های بخش خصوصی، سازمان‌های جامعه مدنی و سایر ذی‌نفعان، از جمله کارگران و اتحادیه‌ها، همکاری کنند تا انتقال عادلانه برای کارکنان در معرض خطر تضمین شود. این شامل ایجاد برنامه‌های ارتقای مهارت و بازآموزی، ایجاد سازوکارهای مؤثر برای حفظ کارکنان در دوره‌های گذار، و بررسی برنامه‌های «شبکه ایمنی» برای کسانی است که امکان بازآموزی ندارند. کشورهای عضو باید برنامه‌هایی را برای پژوهش و رسیدگی به چالش‌های شناسایی‌شده، از جمله ارتقای مهارت و بازآموزی، تقویت

حمایت اجتماعی، سیاست‌ها و مداخلات صنعتی پیش‌دستانه، مزایای مالیاتی و اشکال جدید مالیات، تدوین و اجرا کنند. همچنین باید اطمینان حاصل کنند که بودجه عمومی کافی برای حمایت از این برنامه‌ها وجود دارد. مقررات مرتبط، مانند نظام‌های مالیاتی، باید به‌دقت بررسی و در صورت لزوم اصلاح شوند تا با پیامدهای بیکاری ناشی از خودکارسازی مبتنی بر هوش مصنوعی مقابله شود.

۱۱۹. کشورهای عضو باید پژوهشگران را تشویق و حمایت کنند تا تأثیر سامانه‌های هوش مصنوعی بر محیط‌های کاری محلی را برای پیش‌بینی روندها و چالش‌های آینده تحلیل کنند. این مطالعات باید رویکردی میان‌رشته‌ای داشته باشند و تأثیر سامانه‌های هوش مصنوعی را بر بخش‌های اقتصادی، اجتماعی و جغرافیایی، و نیز بر تعاملات انسان-ربات و روابط انسان-انسان بررسی کنند تا در زمینه بازآموزی و بهترین شیوه‌های استقرار مجدد توصیه ارائه دهند.

۱۲۰. کشورهای عضو باید اقدامات مناسبی را برای تضمین بازارهای رقابتی و حمایت از مصرف‌کننده اتخاذ کنند، با در نظر گرفتن تدابیر و سازوکارهای ممکن در سطوح ملی، منطقه‌ای و بین‌المللی، به‌منظور جلوگیری از سوءاستفاده از موقعیت‌های مسلط در بازار، از جمله توسط انحصارها، در ارتباط با سامانه‌های هوش مصنوعی در سراسر چرخه عمر آن‌ها، خواه داده، پژوهش، فناوری یا بازار باشد. کشورهای عضو باید از نابرابری‌های ناشی از این امر جلوگیری کرده، بازارهای مرتبط را ارزیابی و بازارهای رقابتی را ترویج کنند. باید توجه ویژه‌ای به کشورهای با درآمد پایین و متوسط، به‌ویژه کشورهای کمتر توسعه‌یافته، کشورهای در حال توسعه محصور در خشکی و کشورهای جزیره‌ای کوچک در حال توسعه، که به‌دلیل کمبود زیرساخت، ظرفیت انسانی و مقررات، در برابر سوءاستفاده از سلطه بازار آسیب‌پذیرتر هستند، معطوف شود. بازیگران هوش مصنوعی که در کشورهایی با استانداردهای اخلاقی تثبیت‌شده یا پذیرفته‌شده فعالیت می‌کنند، باید هنگام صادرات محصولات، توسعه یا به‌کارگیری سامانه‌های خود در کشورهایی که چنین استانداردهایی ممکن است وجود نداشته باشد، این استانداردها را رعایت کنند، در عین احترام به حقوق بین‌الملل و قوانین داخلی، استانداردها و رویه‌های آن کشورها.

حوزه سیاستی ۱۱: سلامت و رفاه اجتماعی^{۳۵}

۱۲۱. کشورهای عضو باید تلاش کنند از سامانه‌های مؤثر هوش مصنوعی برای بهبود سلامت انسان و حفاظت از حق حیات استفاده کنند، از جمله با کاهش شیوع بیماری‌ها، و در عین حال همبستگی بین‌المللی را برای مقابله با مخاطرات و عدم قطعیت‌های سلامت جهانی ایجاد و حفظ نمایند، و اطمینان حاصل کنند که به‌کارگیری سامانه‌های هوش مصنوعی در مراقبت‌های سلامت با حقوق بین‌الملل و تعهدات حقوق بشری آن‌ها سازگار است.

35. HEALTH AND SOCIAL WELL-BEING

کشورهای عضو باید تضمین کنند که فعالان در سامانه‌های هوش مصنوعی حوزه سلامت، اهمیت روابط بیمار با خانواده و کارکنان سلامت را در نظر بگیرند.

۱۲۲. کشورهای عضو باید اطمینان حاصل کنند که توسعه و به‌کارگیری سامانه‌های هوش مصنوعی مرتبط با سلامت به‌طور کلی و سلامت روان به‌طور خاص، با توجه کافی به کودکان و نوجوانان، به‌گونه‌ای تنظیم شود که ایمن، مؤثر، کارآمد، علمی و از نظر پزشکی اثبات‌شده باشند و نوآوری مبتنی بر شواهد و پیشرفت پزشکی را امکان‌پذیر سازند. همچنین در حوزه مرتبط با مداخلات سلامت دیجیتال، کشورهای عضو قویاً تشویق می‌شوند بیماران و نمایندگان آن‌ها را در تمامی مراحل مرتبط توسعه سامانه به‌طور فعال درگیر کنند.

۱۲۳. کشورهای عضو باید توجه ویژه‌ای در تنظیم راه‌حل‌های پیش‌بینی، تشخیص و درمان در مراقبت سلامت در کاربردهای هوش مصنوعی، از راه‌های زیر، داشته باشند:

- (الف) تضمین نظارت برای کاهش و تعدیل سوگیری؛
- (ب) تضمین اینکه متخصص، بیمار، مراقب یا کاربر خدمات به‌عنوان «کارشناس حوزه» در تیم، در تمامی مراحل مرتبط توسعه الگوریتم‌ها مشارکت داشته باشد؛
- (ج) توجه کافی به حریم خصوصی به دلیل نیاز بالقوه به پایش پزشکی و اطمینان از رعایت تمامی الزامات مرتبط ملی و بین‌المللی حفاظت از داده؛
- (د) تضمین سازوکارهای مؤثر تا افرادی که داده‌های شخصی آن‌ها تحلیل می‌شود، از استفاده و تحلیل داده‌های خود آگاه باشند و رضایت آگاهانه ارائه دهند، بدون اینکه دسترسی به خدمات سلامت محدود شود؛
- (ه) تضمین اینکه مراقبت انسانی و تصمیم‌نهایی در تشخیص و درمان همواره توسط انسان اتخاذ شود، در حالی که اذعان می‌شود سامانه‌های هوش مصنوعی می‌توانند در کار آن‌ها کمک کنند؛
- (و) تضمین، در صورت لزوم، بازبینی سامانه‌های هوش مصنوعی توسط کمیته اخلاق پژوهش پیش از استفاده بالینی.

۱۲۴. کشورهای عضو باید پژوهش درباره آثار و تنظیم آسیب‌های بالقوه سلامت روان مرتبط با سامانه‌های هوش مصنوعی، مانند افزایش سطوح افسردگی، اضطراب، انزوای اجتماعی، شکل‌گیری اعتیاد، قاچاق، افراط‌گرایی و اطلاعات نادرست، و موارد دیگر را برقرار کنند.

۱۲۵. کشورهای عضو باید دستورالعمل‌هایی برای تعاملات انسان-ربات و تأثیر آن‌ها بر روابط انسان-انسان، بر اساس پژوهش و با جهت‌گیری به توسعه آینده ربات‌ها، تدوین کنند، با توجه ویژه به سلامت روانی و جسمی انسان‌ها. توجه خاص باید به استفاده از ربات‌ها در مراقبت سلامت، مراقبت از سالمندان و افراد دارای معلولیت، در آموزش، و ربات‌های مورد استفاده برای کودکان، ربات‌های اسباب‌بازی، چت‌بات‌ها و ربات‌های همراه برای کودکان و بزرگسالان معطوف شود. افزون بر این، کمک فناوری‌های هوش مصنوعی باید برای افزایش ایمنی و ارگونومی استفاده از ربات‌ها، از جمله در محیط‌های کاری انسان-ربات،

به‌کار گرفته شود. همچنین باید به امکان استفاده از هوش مصنوعی برای دستکاری و سوءاستفاده از سوگیری‌های شناختی انسان توجه ویژه شود.



۱۲۶. کشورهای عضو باید اطمینان حاصل کنند که تعاملات انسان-ربات با همان ارزش‌ها و اصولی که برای هر سامانه هوش مصنوعی دیگر اعمال می‌شود، مطابقت دارند، از جمله حقوق بشر و آزادی اساسی، ترویج تنوع، و حفاظت از افراد آسیب‌پذیر یا در موقعیت‌های آسیب‌پذیر، مسائل اخلاقی مرتبط با سامانه‌های مبتنی بر هوش مصنوعی در حوزه‌های نورو تکنولوژی و رابط‌های مغز-رایانه باید مورد توجه قرار گیرد تا کرامت و خودمختاری انسانی حفظ شود.

۱۲۷. کشورهای عضو باید اطمینان حاصل کنند که کاربران بتوانند به راحتی تشخیص دهند آیا با یک موجود زنده تعامل دارند یا با یک سامانه هوش مصنوعی که ویژگی‌های انسانی یا حیوانی را تقلید می‌کند.

۱۲۸. کشورهای عضو باید سیاست‌هایی را برای افزایش آگاهی درباره انسان‌انگاری فناوری‌های هوش مصنوعی و فناوری‌هایی که احساسات انسانی را شناسایی و تقلید می‌کنند، از جمله در زبانی که برای اشاره به آن‌ها به کار می‌رود، اجرا کنند و جلوه‌ها، پیامدهای اخلاقی و محدودیت‌های احتمالی چنین انسان‌انگاری‌ای را ارزیابی نمایند، به‌ویژه در زمینه تعامل انسان-ربات و خصوصاً زمانی که کودکان درگیر هستند.

۱۲۹. کشورهای عضو باید پژوهش‌های مشترک درباره اثرات تعامل بلندمدت افراد با سامانه‌های هوش مصنوعی را تشویق و ترویج کنند و توجه ویژه‌ای به تأثیرات روان‌شناختی و شناختی این سامانه‌ها بر کودکان و نوجوانان داشته باشند. این امر باید با استفاده از هنجارها، اصول، پروتکل‌ها، رویکردهای میان‌رشته‌ای و ارزیابی تغییرات رفتارها و عادات، و نیز ارزیابی دقیق پیامدهای فرهنگی و اجتماعی پسینی انجام شود. افزون بر این، کشورهای عضو باید پژوهش درباره تأثیر فناوری‌های هوش مصنوعی بر عملکرد نظام سلامت و پیامدهای سلامت را تشویق کنند.

۱۳۰. کشورهای عضو، و همچنین همه ذی‌نفعان، باید سازوکارهایی را برای مشارکت معنادار کودکان و نوجوانان در گفت‌وگوها، مباحثات و تصمیم‌گیری‌ها در خصوص تأثیر سامانه‌های هوش مصنوعی بر زندگی و آینده آنان ایجاد کنند.

بخش پنجم: پایش و ارزیابی^۱

۱۳۱. دولت‌های عضو باید، متناسب با شرایط خاص، ساختارهای حکمرانی و ترتیبات قانون اساسی خود، سیاست‌ها، برنامه‌ها و سازوکارهای مرتبط با اخلاق هوش مصنوعی را با **شفافیت، اعتبار و به‌طور مستمر** پایش و ارزیابی کنند. این ارزیابی باید با بهره‌گیری از ترکیبی از **روش‌های کمی و کیفی** انجام شود.

برای حمایت از دولت‌های عضو، یونسکو می‌تواند اقدامات زیر را انجام دهد:

(الف) تدوین روش‌شناسی یونسکو برای ارزیابی اثرات اخلاقی^۲ فناوری‌های هوش مصنوعی، مبتنی بر پژوهش‌های علمی دقیق و منطبق با حقوق بین‌الملل بشر، همراه با ارائه راهنماهایی برای اجرای آن در تمامی مراحل چرخه عمر سامانه‌های هوش مصنوعی و نیز تهیه منابع آموزشی و برنامه‌های توانمندسازی به‌منظور کمک به دولت‌های عضو در آموزش مسئولان دولتی، سیاست‌گذاران و سایر بازیگران مرتبط با این روش.

(ب) تدوین روش‌شناسی ارزیابی آمادگی^۳ از سوی یونسکو برای کمک به دولت‌های عضو در شناسایی میزان آمادگی خود در مقاطع مختلف و در امتداد طیفی از ابعاد گوناگون.

(ج) تدوین روش‌شناسی یونسکو برای ارزیابی پیشینی (Ex Ante) و پسینی (Ex Post)^۴ اثربخشی و کارایی سیاست‌های اخلاق هوش مصنوعی و همچنین مشوق‌های مرتبط، بر اساس اهداف از پیش تعیین‌شده.

(د) تقویت تحلیل‌ها و گزارش‌های مبتنی بر پژوهش و شواهد^۵ درباره سیاست‌های اخلاق هوش مصنوعی.

(هـ) گردآوری و انتشار اطلاعات مربوط به پیشرفت‌ها، نوآوری‌ها، گزارش‌های پژوهشی، انتشارات علمی، داده‌ها و آمار مرتبط با سیاست‌های اخلاق هوش مصنوعی، از جمله از طریق بهره‌گیری از ابتکارات موجود، به‌اشتراک‌گذاری تجربیات

1. Monitoring and evaluation

2. Ethical Impact Assessment(EIA)

3. UNESCO readiness assessment methodology

4. UNESCO methodology to evaluate ex ante and ex post

5. research- and evidence-based analysis

موفق و یادگیری متقابل، با هدف پیشبرد اجرای این توصیه‌نامه.

۱۳۲. فرایندهای پایش و ارزیابی باید مشارکت گسترده همه ذی‌نفعان را تضمین کنند؛ از جمله، اما نه محدود به، افراد یا گروه‌های آسیب‌پذیر و افرادی که در شرایط آسیب‌پذیر قرار دارند. همچنین باید تنوع اجتماعی، فرهنگی و جنسیتی رعایت شود تا فرایندهای یادگیری بهبود یابد، ارتباط میان یافته‌ها، تصمیم‌گیری، شفافیت و پاسخگویی تقویت شود و نتایج از مشروعیت و کارآمدی بیشتری برخوردار باشند.

۱۳۳. به منظور ترویج بهترین سیاست‌ها و رویه‌های مرتبط با اخلاق هوش مصنوعی، لازم است ابزارها و شاخص‌های مناسبی برای ارزیابی اثربخشی و کارایی^۶ این سیاست‌ها بر اساس استانداردها، اولویت‌ها و اهداف تعیین‌شده تدوین شوند. این ارزیابی‌ها باید به‌طور ویژه وضعیت افراد متعلق به گروه‌های محروم، حاشیه‌نشین و آسیب‌پذیر یا افرادی که در شرایط آسیب‌پذیر قرار دارند، و نیز تأثیر سامانه‌های هوش مصنوعی را در سطوح فردی و اجتماعی مورد بررسی قرار دهند.

پایش و ارزیابی تأثیر سامانه‌های هوش مصنوعی و نیز سیاست‌ها و رویه‌های مرتبط با اخلاق هوش مصنوعی باید به‌صورت مستمر، نظام‌مند و متناسب با سطح مخاطرات مربوطه انجام شود. این فرایند باید بر چارچوب‌های مورد توافق در سطح بین‌المللی استوار باشد و ارزیابی نهادهای دولتی و خصوصی، ارائه‌دهندگان خدمات و برنامه‌ها، از جمله خودارزیابی‌ها^۷، مطالعات ردیابی^۸ و همچنین تدوین مجموعه‌ای از شاخص‌ها را دربر گیرد. گردآوری و پردازش داده‌ها باید مطابق با حقوق بین‌الملل، قوانین ملی مربوط به حفاظت از داده‌ها و حریم خصوصی و نیز ارزش‌ها و اصول مندرج در این توصیه‌نامه انجام شود.

۱۳۴. به‌طور خاص، دولت‌های عضو می‌توانند ایجاد یا به‌کارگیری سازوکارهای مختلف برای پایش و ارزیابی را مدنظر قرار دهند؛ از جمله:

- تشکیل کمیسیون اخلاق^۹؛
- ایجاد رصدخانه اخلاق هوش مصنوعی^{۱۰}؛
- راه‌اندازی مخزن یا پایگاه توسعه هوش مصنوعی منطبق با حقوق بشر و اصول اخلاقی؛
- مشارکت در ابتکارات موجود از طریق ترویج پایبندی به اصول اخلاقی در حوزه‌های مختلف صلاحیت یونسکو؛
- ایجاد سازوکارهایی برای تبادل تجربه و یادگیری متقابل؛
- بهره‌گیری از محیط‌های آزمون مقرراتی^{۱۱}؛

6. Effectiveness and Efficiency

7. Self-evaluations

8. Tracer Studies

9. Ethics Commission

10. AI ethics observatory

11. AI regulatory sandboxes

- و تدوین راهنمای ارزیابی برای همه بازیگران حوزه هوش مصنوعی، به‌منظور سنجش میزان انطباق فعالیت‌های آنان با توصیه‌های سیاستی مندرج در این سند.

بخش ششم: بهره‌برداری و اجرای این توصیه‌نامه^۱

۱۳۵. دولت‌های عضو و تمامی سایر ذی‌نفعانی که در این توصیه‌نامه شناسایی شده‌اند، باید ارزش‌ها، اصول و استانداردهای اخلاقی مرتبط با هوش مصنوعی را که در این توصیه‌نامه تبیین شده‌اند، رعایت، ترویج و حمایت کنند و تمامی اقدامات ممکن را برای تحقق و اجرای توصیه‌های سیاستی آن به عمل آورند.

۱۳۶. دولت‌های عضو باید تلاش کنند اقدامات خود را در راستای اجرای این توصیه‌نامه گسترش داده و تکمیل کنند. این امر از طریق همکاری با تمامی سازمان‌های دولتی و غیردولتی مرتبط در سطوح ملی و بین‌المللی، همچنین شرکت‌های فراملی و نهادهای علمی که فعالیت‌های آنان در چارچوب اهداف و حوزه شمول این توصیه‌نامه قرار می‌گیرد، امکان‌پذیر است.

در این راستا، تدوین روش‌شناسی یونسکو برای ارزیابی اثرات اخلاقی هوش مصنوعی^۲ و نیز ایجاد کمیسیون‌های ملی اخلاق هوش مصنوعی^۳ می‌تواند از ابزارهای مهم برای تحقق اهداف این توصیه‌نامه باشد.

1. Utilization and Exploitation of the Present Recommendation
2. UNESCO Ethical Impact Assessment methodology
3. National commissions for the ethics of AI

بخش هفتم:

ترویج توصیه‌نامه‌ی حاضر

۱۳۷. یونسکو رسالت نمایندگی اصلی سازمان ملل متحد برای ترویج و انتشار این توصیه‌نامه را عهده دار است و بر این اساس، ضمن احترام به مأموریت سایر نهادهای مرتبط سازمان ملل متحد و به منظور جلوگیری از تکرار کار، با آن‌ها همکاری خواهد کرد.

۱۳۸. یونسکو و نهادهایش از جمله، کمیسیون جهانی اخلاق، دانش، و فناوری علمی^۱(COMEST)، کمیته بین‌المللی اخلاق زیستی^۲(IBC) و کمیته بین دولتی اخلاق زیستی^۳(IGBC)، با سایر سازمان‌های دولتی و غیردولتی بین‌المللی، منطقه‌ای و زیرمنطقه‌ای نیز همکاری خواهند کرد.

۱۳۹. اگرچه در چارچوب یونسکو، مأموریت ترویج و حفاظت در حوزه اختیارات دولت‌ها و نهادهای بین دولتی قرار دارد، اما جامعه مدنی بازیگر مهمی برای دفاع از منافع بخش عمومی خواهد بود و بنابراین یونسکو باید مشروعیت آن را تضمین و ترویج کند.

1. World Commission on the Ethics of Scientific Knowledge and Technology (COMEST)
2. International Bioethics Committee (IBC)
3. Intergovernmental Bioethics Committee (IGBC)

بخش هشتم: مقررات نهایی

۱۴۰. این توصیه‌نامه باید به صورت یک سند یکپارچه و منسجم تفسیر شود. ارزش‌ها و اصول بنیادین مطرح‌شده در آن، مکمل یکدیگرند و باید در ارتباط متقابل با هم فهم و اجرا شوند؛ نه به صورت مجزا یا مستقل.

۱۴۱. هیچ‌یک از مفاد این توصیه‌نامه نباید به گونه‌ای تفسیر شود که موجب جایگزینی، تغییر یا تضعیف تعهدات یا حقوق دولت‌ها بر اساس حقوق بین‌الملل گردد، یا به هر دولت، بازیگر سیاسی، اقتصادی یا اجتماعی، گروه یا شخصی اجازه دهد فعالیت‌های انجام دهد یا اقدامی را دنبال کند که با حقوق بشر، آزادی‌های بنیادین، کرامت انسانی یا حفاظت از محیط زیست و زیست‌بوم‌ها - اعم از جاندار و بی‌جان - مغایرت داشته باشد.

